



UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Biomedicina comparata ed alimentazione

Corso di laurea magistrale in Biotecnologie per l'alimentazione

**ANALISI DEL PROFILO DI METILAZIONE IN
MUSCOLO SCHELETRICO DI BOVINI TRATTATI
CON PROMOTORI DELLA CRESCITA**

Relatore

Prof. Luca Bargelloni

Correlatore

Dott.ssa Roberta Carraro

Laureanda

Michela Salmaso

Matricola n. 1058738

ANNO ACCADEMICO 2013-2014

Alla mia “grande famiglia” che
mi ha dato ali per volare e
radici sulle quali contare

ABSTRACT

Despite the European ban (Council Directive 2008/97/EC), the illegal use of growth-promoters in meat production is still present and confirmed by the numbers of non-compliant results reported in the National Monitoring Plan of 2013. There is evidence, however, that the real incidence of growth-promoter abuse might be underestimated. The major concern is the use of new anabolic compounds, unknown and not subjected to routine controls, the administration at low doses and the combination of different substances as sexual steroids, β_2 -agonists and corticosteroids in cocktails in order to exploit synergic growth effects and elude official controls. To overcome these illegal practices, constant innovation in screening and validation methods is necessary. In this contest the innovative methods include the so-called *omics* techniques (genomics, transcriptomics, proteomics and metabolomics) based on the simultaneous analysis of thousands of markers might help finding signatures of growth-promoter abuse. The present study implemented an epigenomic-based strategy to obtain DNA methylation patterns for the identification of possible gene markers for anabolic treatments in beef cattle. For this purpose, two different groups of animals were studied: a group was experimentally treated with dexamethasone and estradiol benzoate and the other group was not treated to represent the control.

The adopted epigenomic approach coupled a target enrichment technique using MBD (*methyl-binding protein*) with bisulfite sequencing conversion of DNA. Sequence data were then analysed using specific bioinformatic tools. As a result, three differentially methylated genomic regions were identified, which might represent putative biomarkers of treatment. However, this requires further analysis to validate these regions and increase the number of specimens tested to increase statistical significance.

RIASSUNTO

Nonostante il divieto imposto dalle Direttiva europea 2008/97/CE, l'utilizzo illecito di promotori della crescita nell'allevamento del bovino da carne è ancora presente e confermato dalle non conformità riportate nel Piano nazionale per la ricerca dei Residui (PNR 2013). Tuttavia ci sono prove che la reale incidenza dell'uso illegale di promotori della crescita sia sottostimata. La principale preoccupazione riguarda lo sviluppo di nuovi composti anabolizzanti, sconosciuti e non soggetti ai controlli di routine, la somministrazione a basse dosi e la combinazione di sostanze diverse quali steroidi sessuali, β_2 -agonisti, e corticosteroidi in *cocktails* per sfruttare l'effetto sinergico sulla crescita ed eludere i controlli ufficiali. Per far fronte a queste pratiche illegali è necessaria una costante innovazione nei metodi di screening e validazione. In questo contesto le metodiche innovative includono le tecnologie omiche (genomica, trascrittomica, proteomica e metabolomica) basate sull'analisi simultanea di migliaia di marcatori che potrebbero essere d'aiuto nell'identificazione della somministrazione illecita di promotori della crescita. Nello specifico questo studio ha implementato una strategia epigenomica per ottenere i profili di metilazione del DNA di muscolo scheletrico al fine di identificare possibili marcatori genomici indicativi del trattamento anabolizzante nei bovini da carne. A tale scopo sono stati studiati due diversi gruppi di animali, il primo trattato con desametasone ed estradiolo benzoato, il secondo non trattato (controllo).

Questo studio sperimentale ha adottato un approccio epigenomico abbinando una tecnica di arricchimento con MBD (*methyl-binding protein*) al trattamento con bisulfite seguito da sequenziamento del DNA. L'analisi è stata resa possibile utilizzando strumenti bioinformatici. Sono state identificate tre regioni genomiche differenzialmente metilate che potrebbero rappresentare possibili biomarcatori di trattamento. Tuttavia ulteriori analisi per validare le regioni individuate e l'incremento del numero di campioni testati si rendono necessari al fine di giungere a risultati affidabili e statisticamente significativi.

INDICE	1
1. INTRODUZIONE	1
1.1 Normativa vigente in materia di trattamenti illeciti	1
1.2 Principali categorie di promotori della crescita	2
1.2.1 Tireostatici	2
1.2.2 Glicocorticoidi	2
1.2.3 Estrogeni	4
1.2.4 β -agonisti	6
1.3 Risultati del PNR 2013	6
1.4 Metodi di indagine diagnostica	8
1.5 Modificazioni epigenetiche	11
1.5.1 La metilazione del DNA	12
1.5.2 Il meccanismo di metilazione del DNA: le proteine MCP	13
1.5.3 La metilazione nel mondo animale	15
1.5.4 Implicazione dell'epigenesi nei processi fisiologici	16
1.5.5 Modificazioni epigenetiche associate ad interazione con l'ambiente	19
1.5.5.1 Contaminanti ambientali	20
1.6 Metodiche epigenetiche	21
2. OBIETTIVI	25
3. MATERIALI E METODI	27
3.1 Campionamento	27
3.2 Creazione di librerie genomiche attraverso arricchimento con proteine specifiche e trattamento con bisulfito	28
3.2.1 Estrazione e quantificazione del DNA genomico	28
3.2.2 Frammentazione meccanica del dsDNA	29
3.2.3 Arricchimento selettivo del DNA frammentato in base al grado di metilazione	30
3.2.4 <i>End-repair</i> , <i>A-tailing</i> e ligazione di adattatori	31
3.2.5 Trattamento con bisulfito	32
3.2.6 Amplificazione e sequenziamento delle librerie	32
3.2.7 Controllo qualità delle sequenze e <i>trimming</i>	34
3.2.8 <i>Mapping</i> delle sequenze sul genoma di riferimento	34
3.2.9 <i>Methylation extractor</i>	36
3.2.10 <i>Methylkit</i>	37
3.3 Metodica <i>2b-RAD</i>	38
3.3.1 Estrazione e quantificazione del DNA genomico	38
3.3.2 Creazione delle librerie <i>WGA</i>	40
3.3.3 Creazione delle librerie <i>2b-RAD</i>	41
3.3.4 Purificazione e sequenziamento delle librerie	43
4. RISULTATI E DISCUSSIONE	45
4.1 Librerie genomiche arricchite: risultati preliminari	45
4.2 Analisi bioinformatiche	46

4.2.1 Controllo qualità e <i>trimming</i>	46
4.2.2 <i>Mapping</i>	47
4.2.3 <i>Methylation extractor</i>	49
4.2.4 Statistiche descrittive di <i>Methylkit</i>	49
4.2.5 Correlazione tra campioni	52
4.2.6 <i>Clustering</i>	52
4.2.7 Analisi per componenti principali (PCA)	53
4.2.8 Analisi per posizione	54
4.2.9 Analisi per regione	54
4.2.10 Analisi di annotazione per posizione e per regione	55
4.3 Librerie <i>2b-RAD</i> : risultati preliminari	58
4.3.1 Analisi bioinformatiche	59
5. CONCLUSIONI	63
6. BIBLIOGRAFIA	65
7. ALLEGATO 1	69

1. INTRODUZIONE

1.1 Normativa vigente in materia di trattamenti illeciti

A partire dagli anni '60 del secolo scorso sono stati dedicati numerosi studi all'alimentazione e alla selezione genetica nell'allevamento dei bovini da carne per migliorare gli indici di conversione, ridurre il ciclo di allevamento e ottimizzare le rese alla produzione ed alla macellazione (Nebbia, 2009). Anche la somministrazione di alcuni composti farmacologicamente attivi (tireostatici, glicocorticoidi di sintesi, steroidi sessuali e β -agonisti) si è dimostrata utile a questi scopi attraverso vari meccanismi fisiologici (Carraro et al., 2009). Tuttavia l'utilizzo di questi farmaci come promotori della crescita in animali in produzione zootecnica, parzialmente accettato in altri Paesi, è vietato nell'Unione Europea dal 1988 (Decisione della Commissione 88/199/CEE) a causa dei potenziali effetti negativi sugli animali stessi e della conseguente presenza, rischiosa per la salute dei consumatori, di residui nel prodotto finito. In particolare la somministrazione di sostanze ad azione ormonale, tireostatica e quella di β -agonisti nelle produzioni animali è vietata dalla Direttiva 2008/97/CE, che modifica la Direttiva 96/22/CE, ed è stata recepita dal governo italiano con il D.Lgs. 158/06; quest'ultimo definisce il trattamento illecito come "utilizzazione di sostanze o prodotti non autorizzati, ovvero di sostanze o prodotti autorizzati, a fini o condizioni diversi da quelli previsti dalle disposizioni vigenti" vietando l'utilizzo di qualsiasi sostanza, anche autorizzata, con lo scopo di migliorare le performance produttive. La Direttiva 96/23/CE "Concernente le misure di controllo su talune sostanze e sui loro residui negli animali vivi e nei loro prodotti" regola il monitoraggio di sostanze farmacologicamente attive e, a questo scopo, richiede agli Stati membri di adottare Piani Nazionali di Controllo per verificare la conformità, in allevamento e nella prima fase di trasformazione, di residui di medicinali veterinari, di antiparassitari e contaminanti ambientali con i limiti massimi residuali (limiti massimi tollerati per i contaminanti) fissati dalle normative vigenti (Regolamento UE 37/2010). Il limite massimo residuale di sostanze farmacologicamente attive è stato definito come "la concentrazione massima di residui risultante dall'uso di un medicinale veterinario (espressa in mg/kg o μ g/kg sulla base del peso vivo) che la Comunità può ammettere che sia consentita legalmente o riconosciuta accettabile negli alimenti" (Regolamento 2377/90/CE). Per tutti i farmaci con un limite massimo residuale, destinati ad animali

in produzione, devono essere inoltre definiti i tempi di sospensione, ovvero una finestra di tempo adeguata per prevenire l'accumulo di residui nel prodotto finale (Cannizzo et al., 2013) e garantire così la sicurezza del consumatore. Particolare attenzione è stata riservata a sostanze utilizzate illecitamente perché vietate o perché impiegate con una posologia diversa per scopi diversi da quelli autorizzati. È questo il caso dei promotori della crescita, infatti, nonostante il divieto e la realizzazione di piani di monitoraggio persiste la somministrazione di queste sostanze e, dai sequestri effettuati dagli Organi di vigilanza e controllo, emerge l'esistenza di un fiorente mercato nero di sostanze e preparazioni ad hoc (Nebbia, 2009; Pegolo et al., 2012). Lo sviluppo di nuovi composti e la somministrazione a bassi dosaggi di promotori della crescita, o la combinazione di questi in "cocktail" formati anche da decine di molecole diverse (Nebbia, 2009), richiedono lo sviluppo di metodi innovativi di indagine diagnostica (Pegolo et al., 2012).

1.2 Principali categorie di promotori della crescita

1.2.1 Tireostatici

Ad oggi i derivati tiouracilici (metil-, propil- e feniltiouracile) e imidazolici (metimazolo) riconducibili al gruppo dei tireostatici non sembrano più utilizzati come promotori della crescita e questo è dimostrato dall'assenza di non conformità rilevata dagli ultimi Piani di monitoraggio in ambito europeo (Nebbia, 2009). I tireostatici inibendo la sintesi di ormoni tiroidei provocano un aumento della ritenzione idrica nelle masse muscolari e del peso corporeo dovuto principalmente al rallentamento della motilità dei prestomaci. I tireostatici sono stati tra i primi composti ad essere utilizzati a scopo auxinico nell'allevamento dei bovini da carne, tuttavia la disponibilità di altre molecole più performanti e meno rischiose in fase di screening ne ha ridotto drasticamente l'utilizzo come singoli (Nebbia, 2009).

1.2.2 Glicocorticoidi

Da modificazioni strutturali del composto endogeno cortisolo o idrocortisone derivano i glicocorticoidi di sintesi, frequentemente utilizzati in medicina veterinaria a scopo terapeutico come potenti antiinfiammatori nel trattamento di patologie dell'apparato locomotore e muscolo-scheletrico, malattie autoimmuni, allergie e forme di shock di diversa origine e malattie batteriche acute e croniche in associazione a chemio antibiotici (Nebbia, 2009). Oltre all'attività antiinfiammatoria e immunosoppressiva, i

glicocorticoidi di sintesi a dosi farmacologiche hanno effetti importanti nella gluconeogenesi, nell'accumulo di glicogeno in ambito epatico e nel catabolismo proteico e lipidico. I glicocorticoidi a dosaggi terapeutici hanno inoltre una serie di effetti negativi sul muscolo scheletrico inibendone lo sviluppo e causandone atrofia (Carraro et al., 2009). D'altro canto il loro utilizzo come promotori della crescita, da soli o in associazione ad altre sostanze come steroidi sessuali e β -agonisti, si spiega a dosi giornaliere molto inferiori rispetto a quelle farmacologiche per un periodo limitato di tempo (30-50 giorni) (Nebbia, 2009). Questo si traduce in un aumento dell'ingestione, stimolato dal metabolismo glucidico (Carraro et al., 2009), in un miglioramento sostanziale dell'incremento medio giornaliero, in un moderato aumento delle masse muscolari e in un colore meno intenso delle carcasse (Gottardo et al., 2008). Tra i principali corticosteroidi di sintesi (desametasone, betametasone, fumetasone, beclometasone, prednisolone, flormetolone), il desametasone, insieme al prednisone, è quello maggiormente studiato e più frequentemente responsabile di positività (Nebbia, 2009). La somministrazione del desametasone può avvenire per via intramuscolare, endovenosa e orale, aggiungendo il farmaco alla razione, nel caso di trattamenti illeciti. In generale è rapidamente assorbito e se somministrato per os il picco ematico è raggiunto dopo circa 2 ore, mentre è ancora più rapido nel caso di trattamento per via intramuscolare (Nebbia, 2009).

L'organo target, in particolare in bovini da carne trattati illecitamente con desametasone, è il timo che presenta diminuzione del peso, atrofia e severe modificazioni istologiche quali deplezione di cellule linfocitarie e infiltrazione di tessuto adiposo nel parenchima (Cannizzo et al., 2010). Per quanto riguarda l'analisi ufficiale svolta nell'ambito del Piano Nazionale per la ricerca dei Residui (PNR), il fluido biologico d'elezione in allevamento è rappresentato dall'urina nella quale i cortisonici sono reperibili sia durante il trattamento sia dopo la sua sospensione per circa 3-10 giorni in relazione alla dose, al composto e alla via di somministrazione. Al macello l'indagine analitica riguarda invece il fegato che presenta concentrazioni rilevabili fino a 6-8 giorni circa dall'ultimo trattamento (Nebbia, 2009). Tuttavia lo studio di Cannizzo et al. (2010) evidenzia che il desametasone (0.7 mg/giorno per os per 40 giorni), ma soprattutto il prednisolone (15 mg/giorno per os per 30 giorni; 30 mg/giorno per os per 35 giorni), tende a diminuire durante il periodo di somministrazione e a scomparire

sia nelle urine che nel fegato dei bovini dopo 6 giorni dall'ultimo trattamento, rendendo difficile svelare casi di illeciti al macello su queste matrici. Per l'analisi di entrambe le matrici, la metodica di screening prevista è l'ELISA (*Enzyme-Linked Immunoassorbent Assay*), mentre le tecniche di conferma sono rappresentate da GC-MS/MS e LC-MS/MS (limite di rilevabilità nelle urine pari a 2.0 µg/kg, mentre nel fegato 2-10 µg/kg), ovvero gas cromatografia o cromatografia liquida abbinate a *Tandem Mass Spectrometry* (PNR 2014).

1.2.3 Estrogeni

L'estradiolo benzoato è un estere sintetico dell'estradiolo che dopo la somministrazione è idrolizzato a 17β-estradiolo, estrogeno naturale tipico del mondo animale. Il 17β-estradiolo è prodotto principalmente nell'ovaio, ma anche in altri tessuti come la corteccia surrenale e, nei soggetti gravidi, il corpo luteo e la placenta; anche le gonadi maschili sintetizzano quantità apprezzabili di questo estrogeno naturale (Nebbia, 2009). Oltre alla funzione centrale nel ciclo riproduttivo, gli estrogeni hanno effetti importanti in numerosi processi fisiologici quali l'anabolismo proteico e la ritenzione di minerali (Pezzolato et al., 2011) promuovendo l'accrescimento corporeo e lo sviluppo delle masse muscolari. L'azione auxinica degli estrogeni è legata innanzitutto alla presenza in organi e tessuti di recettori intracellulari specifici; questi ultimi vengono attivati e dissociati dalle *Heat Shock Proteins (HSPs)* grazie al legame con il ligando (estrogeno); il complesso ormone-recettore interagisce con determinate sequenze geniche modulandone il profilo di espressione e promuovendo la sintesi di proteine responsive. In particolare l'effetto del composto estrogenico è direttamente proporzionale alla concentrazione di recettori degli estrogeni (ER) presenti in quel dato organo o tessuto. Nel caso dei bovini è stato infatti calcolato che il tenore medio di ER nei muscoli degli arti posteriori è circa il doppio rispetto a quello dei muscoli dell'addome (Nebbia, 2009). In particolare le differenze nell'azione anabolizzante delle diverse molecole dipende principalmente dalla loro affinità di legame con i recettori, ma anche dall'intensità dell'effetto di primo passaggio e dalla velocità di eliminazione. Gli estrogeni agiscono anche indirettamente stimolando il rilascio dell'ormone della crescita da parte dell'adenoipofisi e di altri fattori di crescita dal fegato (Nebbia, 2009).

A fini anabolizzanti, le diverse vie di somministrazione dipendono dalla natura del composto estrogenico scelto e dal veicolo utilizzato. La via principale è quella intramuscolare nei gruppi muscolari meno pregiati come quelli dei quarti anteriori o del collo; spesso le preparazioni sono composte da esteri di vari acidi in veicoli oleosi in grado di garantire il lento rilascio del principio attivo. Un'altra modalità accertata è quella dermica attraverso preparazioni *pour on* da localizzare nella regione del dorso sfruttando così la liposolubilità caratteristica dei composti steroidei e la difficoltà nell'individuare il sito d'inoculo in fase ispettiva. Sono utilizzati illecitamente anche impianti sottocutanei (autorizzati in vari Paesi extraeuropei), costituiti da una matrice inerte sulla quale vengono adsorbite le molecole d'interesse, inseriti in aree non destinate al consumo umano come la cartilagine auricolare o la plica caudale. La via orale è invece quella meno praticata per motivi di carattere cinetico riconducibili alle reazioni biotrasformative di fase I che determinano nel caso di molti composti estrogenici la formazione di metaboliti generalmente meno attivi. L'eccezione è rappresentata da estrogeni di tipo sintetico quali il dietilstilbestrolo e, più recentemente il nandrolone che subiscono meno intensamente l'effetto di primo passaggio, soprattutto dal punto di vista ossidativo, e danno origine a metaboliti con un effetto ormonale sostanzialmente invariato rispetto alla molecola parente (Nebbia, 2009). Per quanto riguarda la distribuzione e l'accumulo nei tessuti edibili, generalmente le concentrazioni più elevate di estrogeni sono presenti nell'ordine, in fegato, rene, grasso e muscolo. Grazie all'elevata liposolubilità i composti estrogenici sono dotati di un volume di distribuzione elevato anche per i tessuti a lenta perfusione garantendo una lunga durata d'azione nell'organismo animale che si traduce nella somministrazione di dosi più basse e ad una minore frequenza di trattamento.

La ricerca chimica dei residui e la valutazione delle eventuali positività deve necessariamente considerare la sintesi endogena degli estrogeni naturali modulata in rapporto al sesso e alla fase del ciclo estrale delle femmine. I mezzi di indagine diagnostica si basano su modificazioni cliniche, anatomo-istopatologiche e su analisi di laboratorio. In particolare sull'animale in vita è possibile evidenziare alcune modificazioni comportamentali e morfo-funzionali ben caratterizzate, di vario grado in relazione alla natura e alla dose del composto, e rilevabili ad un attento esame clinico: aumento della libido, aumento di volume delle masse muscolari, riduzione delle

dimensioni del testicolo, ipertrofia del clitoride, edema vulvare accompagnato da un aumento delle secrezioni vaginali e ipertrofia dell'apparato mammario nella femmina. Le alterazioni anatomo-istopatologiche riguardano in particolare le ghiandole sessuali accessorie (prostata e ghiandole bulbo-uretrali nel maschio, ghiandole del Bartolino nella femmina) dove si osservano iperplasia e metaplasia squamosa dell'epitelio tubulo-acinoso. Pezzolato et al. (2011) propongono un nuovo metodo istologico altamente sensibile basato sull'identificazione di una sola caratteristica (la metaplasia delle ghiandole bulbo-uretrali) in grado di svelare il trattamento illecito di 17 β -estradiolo anche dopo due settimane dalla sospensione del trattamento in vitelli da carne. A partire dal 2008 infatti il test istologico è stato introdotto nel PNR come test di screening per la valutazione delle modificazioni isto-anatomo-patologiche, in organi target di bovini regolarmente macellati, riconducibili ad un possibile trattamento con promotori della crescita (Cannizzo et al., 2013). Per la ricerca del 17 β -estradiolo le metodiche chimico-analitiche stabilite dal PNR sono RIA (*Radio Immuno Assay*) e ELISA come metodi di screening, e GC-MS/MS e LC-MS/MS come conferma (limite di rilevabilità pari a 0.04 μ g/kg) nella matrice siero in bovini allevati; mentre per gli estrogeni di sintesi la ricerca viene eseguita sulla matrice muscolo con le medesime tecniche di screening e conferma con un limite di rilevabilità pari a 2.0 μ g/kg.

1.2.4 B-agonisti

I β -agonisti intensificano il tasso di crescita stimolando i recettori β -adrenergici sulla superficie cellulare. Nel tessuto muscolare la loro stimolazione rallenta il catabolismo proteico promuovendo l'ipertrofia, mentre nel tessuto adiposo promuovono la lipolisi agendo di fatto come ripartitori dell'energia. Il principale β -agonista utilizzato come promotore della crescita è il clenbuterolo, un β_2 -agonista dotato di un'intensa azione terapeutica sulla muscolatura liscia, ma impiegato a scopi illeciti come auxinico sfruttando la sua efficacia a dosi molto basse (Nebbia, 2009).

1.3 Risultati del PNR 2013

Dal punto di vista statistico, le non conformità riguardanti le molecole sopra descritte sono riportate annualmente nel PNR elaborato dal Ministero della Salute in collaborazione con le Regioni, le Province Autonome, i Laboratori Nazionali di Riferimento per la ricerca dei residui e gli Istituti Zooprofilattici Sperimentali. Il PNR stabilisce la frequenza e la modalità del campionamento indicando il gruppo di

sostanze che devono essere monitorate per ciascuna categoria di animali, mangimi e prodotti animali nella prima fase di trasformazione. In particolare i cortisonici dal 2012 non sono più compresi nella categoria A -Sostanze ad effetto anabolizzante e sostanze non autorizzate- gruppo A3 (steroidi), ma nella categoria B -Medicinali veterinari e sostanze contaminanti- sottogruppo B2f -Altre sostanze esercitanti un'attività farmacologica- accogliendo la richiesta della Commissione europea. I cortisonici sono farmaci veterinari autorizzati, hanno un LMR stabilito sulla base della valutazione del rischio e di conseguenza non possono essere considerati solamente come sostanze proibite; infatti "l'utilizzo di cortisonici a fini o condizioni diverse da quelli previsti dalle vigenti disposizioni, compresa l'assenza di prescrizione, l'assenza di registrazione nel registro dei trattamenti, la mancanza dell'annotazione del veterinario responsabile sul registro di scorta in caso di trattamento, configura un trattamento illecito". Nella relazione sui risultati dei controlli ufficiali effettuati nell'anno 2013, i campioni analizzati e considerati non conformi per la presenza dei residui sono stati complessivamente 46, pari allo 0.12% del totale. Mentre 6 sono risultati irregolari per la presenza di residui di sostanze appartenenti alla categoria A (13%), 40 sono i campioni irregolari per la presenza di residui di sostanze presenti in categoria B, pari all'87%. Per quanto riguarda quest'ultima categoria, il gruppo B3 -Altre sostanze e agenti contaminanti- rappresenta la principale causa di non conformità (33%), seguito dal gruppo B1 -Sostanze antibatteriche, compresi sulfamidici e chinolonici) con il 28% e dal gruppo B2 con il 26% (Figura 1). Come per gli altri anni, anche per il 2013 il settore maggiormente coinvolto è quello dei Bovini (41.3%), seguito dai Suini (26.1%) e dal settore Latte (19.6%). Generalmente rispetto agli ultimi quattro anni è stata riscontrata una flessione nella rilevazione delle non conformità pari al 45%. Inoltre a seguito dello spostamento dei cortisonici dal gruppo A3 al gruppo B2f, i campioni irregolari per la categoria A sono drasticamente diminuiti, ma le non conformità sono diminuite anche per il gruppo B2f rispetto al 2012 e in particolare non è mai stata riscontrata la presenza di prednisone, molto spesso associata a desametasone e prednisolone (PNR 2013).

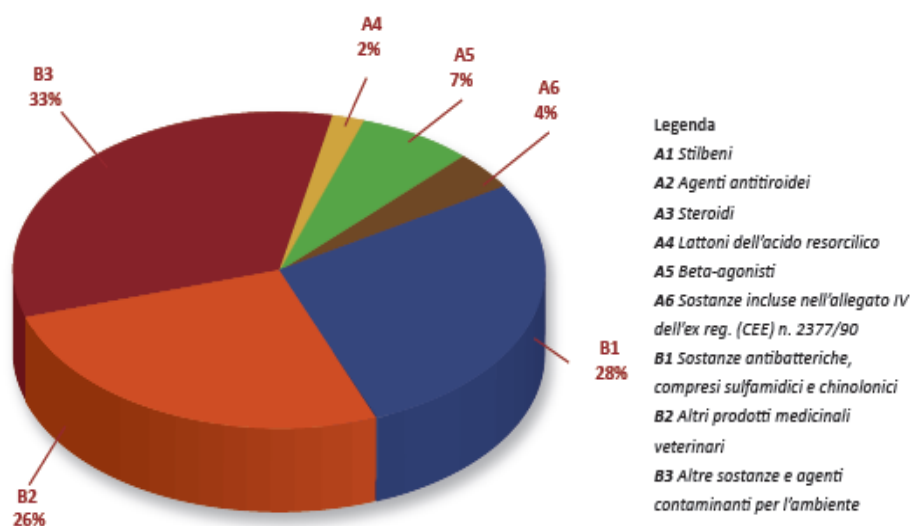


Figura 1 Distribuzione percentuale delle non conformità nei singoli gruppi di sostanze (PNR 2013)

1.4 Metodi di indagine diagnostica

Dal punto di vista normativo, la Decisione della Commissione 2002/657/CE attua la Direttiva 96/23/CE del Consiglio relativamente al rendimento dei metodi analitici e all'interpretazione dei risultati fissando i requisiti necessari per la validazione degli stessi nell'ambito dell'identificazione di sostanze vietate o della quantificazione delle concentrazioni per quelle autorizzate; garantendo così la comparabilità degli esiti tra Laboratori di controllo ufficiali. Le tecniche maggiormente utilizzate per la determinazione del trattamento illecito con promotori della crescita si basano sulla messa in evidenza diretta delle sostanze farmacologicamente attive o dei loro metaboliti negli organi target o nelle urine per mezzo di metodi chimico-fisici (Nebbia, 2009). Questi ultimi sono ad oggi gli unici metodi analitici accettati dalla legislazione Europea per l'identificazione dell'utilizzo illegale di farmaci a scopo anabolizzante negli animali in produzione zootecnica (Stella, 2014). In letteratura sono riportate diverse tecniche, talvolta con varianti nuove e più performanti, utilizzabili dai Laboratori ufficiali; tuttavia è possibile razionalizzarle e raggrupparle in metodi immunometrici impiegati in fase di screening e metodi riconducibili alla spettrometria di massa per la conferma nella rilevazione del residuo. I metodi immunometrici comprendono il saggio immuno-enzimatico ELISA e la tecnica del dosaggio radioimmunologico RIA che si differenzia dall'ELISA perché per rilevare la positività, al posto dell'enzima, viene utilizzato un radio nuclide coniugato all'anticorpo. Queste metodiche sono

ampiamente utilizzate in fase di screening per la rapidità, la facilità di esecuzione e la discreta sensibilità nel rilevare basse concentrazioni di principi attivi. Sono però caratterizzati da bassa specificità perché, a causa della cross-reattività dovuta a interferenti endogeni (ad esempio nel caso degli steroidi sessuali), solo raramente sono in grado di distinguere analiti diversi, anche se di struttura chimica affine. Infine, a causa dell'effetto matrice, permettono il dosaggio di una sola molecola per test (Nebbia, 2009). La spettrometria di massa invece risente meno dell'effetto matrice e permette l'individuazione contemporanea di più sostanze grazie ad un'elevata specificità e sensibilità di analisi. Il principio di base risiede nella capacità di separare una miscela di ioni in funzione del loro rapporto massa/carica tramite campi magnetici statici o oscillanti. In particolare le molecole del campione in esame attraversano un fascio di elettroni ad energia nota e così ionizzate non sono più stabili e si frammentano in ioni più leggeri in funzione della loro struttura chimica. Le metodiche di conferma ufficiali utilizzano il metodo di analisi *Tandem Mass Spectrometry – MS/MS-* che consente di ripetere più volte la frammentazione su ioni-frammenti già analizzati, in combinazione con tecniche di separazione iniziale quali la gas cromatografia (GC) e la cromatografia liquida (LC –*Liquid Chromatography-*). La gas cromatografia è una tecnica che può essere attuata velocemente per i composti volatili, ma per tutti gli altri è necessaria una laboriosa preparazione del campione per la derivatizzazione, allungando così i tempi di esecuzione. La cromatografia liquida invece si adatta ad analisi di routine perché ha sensibilità e specificità più elevate abbinate a tempi di analisi più contenuti, e affinità più marcata con le diverse matrici grazie all'eluizione in fase mobile caratterizzata da un liquido o una miscela di liquidi. Tuttavia queste metodiche di conferma richiedono tempi lunghi, costi elevati e personale altamente qualificato per l'esecuzione dell'analisi. Inoltre l'introduzione nel mercato illegale di composti chimicamente sconosciuti e la somministrazione di diverse molecole a basse dosi mettono in crescente difficoltà l'adeguatezza dei metodi analitici ufficiali nella determinazione del trattamento (Cannizzo et al., 2013). Per questo motivo sono in continuo sviluppo metodi indiretti basati sulla valutazione (esame istologico) e sulla quantificazione (biomarcatori di trattamento –*omics technologies-*) degli effetti fisiologici derivanti dal trattamento illecito con promotori della crescita (Pegolo et al., 2012).

Nello specifico l'esame istologico è una metodica economica e relativamente veloce che mette in evidenza l'illecito in seguito alla valutazione macro e microscopica degli effetti in organi bersaglio. Con questo esame indiretto non si individua il composto alla base del trattamento, ma la recente decisione di affiancarlo ai metodi chimico-analitici è da ricercare nella rapida metabolizzazione a cui vanno incontro molti promotori della crescita e nella difficoltà di ricercare con le metodiche ufficiali tutte le singole molecole impiegate (Nebbia, 2009).

Le tecnologie omiche (genomica, trascrittomica, proteomica e metabolomica) rappresentano un gruppo di metodiche innovative indirette che forniscono strumenti utili per l'identificazione simultanea, nelle diverse matrici biologiche, di potenziali biomarcatori molecolari, ovvero indicatori specifici del trattamento illecito e della risposta biologica che ne deriva. Nello specifico i biomarcatori possono fornire risposte di tipo quantitativo sull'esposizione alla sostanza farmacologicamente attiva, sulla rilevanza della dose in grado di suscitare un effetto fisiologico e sull'eventuale comparsa di un effetto a posteriori. Negli ultimi anni sono stati pubblicati numerosi studi che utilizzano le metodologie legate alla trascrittomica per l'analisi del profilo di espressione genico e la ricerca di biomarcatori di espressione indicativi dell'utilizzo di sostanze anabolizzanti nei bovini da carne. La crescente applicazione della trascrittomica in ambito tossicologico ha portato alla definizione di una nuova disciplina scientifica quale la tossicogenomica (Pegolo et al., 2012) che costituisce l'insieme delle tecniche molecolari utilizzate per la caratterizzazione degli effetti dati dagli xenobiotici sugli organismi animali. Dal punto di vista operativo è possibile distinguere tra due approcci: a *gene candidato* o *globale e simultaneo*. Nel primo caso si analizza l'espressione di un numero ristretto di geni, selezionati sulla base di studi in merito presenti in letteratura, attraverso la metodica di *RT-qPCR* (*PCR Real Time* quantitativa); nel secondo approccio si analizza invece l'espressione trascrittomica globale in un singolo esperimento, utilizzando la tecnologia *cDNA microarray* (Riedmaier et al., 2012) che permette di identificare geni differenzialmente espressi tra controllo e trattato. Questa tecnologia è ad oggi il *gold standard* per lo screening dei cambiamenti di espressione genica causati dal trattamento con promotori della crescita (Riedmaier et al., 2012). Con le tecnologie di sequenziamento *high throughput* è stato possibile sviluppare una nuova metodica per l'analisi dell'espressione genica,

chiamata *RNA-sequencing*, che permette di analizzare ciascun trascritto genico con una sensibilità molto simile alla *RT-qPCR* (Riedmaier et al., 2012). Anche lo studio degli effetti dei promotori della crescita sul proteoma è stato recentemente utilizzato come strumento per la scoperta di potenziali biomarcatori di trattamento. Con le metodiche relative alla proteomica è possibile identificare alterazioni quali-quantitative del profilo proteico in esame; generalmente le proteine sono prima separate utilizzando l'elettroforesi bidimensionale (2DE -*Bi-Dimensional Electrophoresis*) e poi caratterizzate e identificate attraverso la misurazione della massa dei peptidi (MALDI-MS, -*Peptide-mass fingerprinting*-) (Stella et al., 2014). La metabolomica rappresenta una tecnologia emergente nell'ambito degli illeciti. Quest'ultima si focalizza sulla misurazione su larga scala di metaboliti potenzialmente in grado di essere biomarcatori di risposta rispetto al trattamento (Pinel et al., 2010). L'obiettivo principale di queste tecnologie molecolari è di identificare validi biomarcatori di trattamento che possano essere utilizzati per sviluppare test di screening applicabili su vasta scala (Nebbia et al., 2010).

1.5 Modificazioni epigenetiche

Con il termine epigenetica si fa riferimento a cambiamenti sul profilo trascrizionale che avvengono indipendentemente da modificazioni dell'informazione contenuta, o meglio della struttura di base delle sequenze. Le alterazioni epigenetiche mediano infatti le influenze dell'ambiente sull'espressione genica (Foley et al., 2009) e così sul fenotipo, nel normale sviluppo dell'organismo, nella differenziazione cellulare e nella manifestazione di alcune patologie, anche tumorali. È stato stimato che le alterazioni di tipo epigenetico sono 100 volte più frequenti delle mutazioni genetiche e questo può accadere in maniera casuale o in risposta a stimoli ambientali (Foley et al., 2009). Tra i fattori ambientali sono stati annoverati deficit nutrizionali riguardanti in particolare i folati, le vitamine B6 e B12 e gli aminoacidi colina e metionina coinvolti nel *pathway* metabolico che determina la disponibilità di S-adenosilmetionina (SAM), principale donatore di metili per la metilazione delle proteine istoniche e del DNA (Feil and Fraga, 2012); ma anche stress esterni, cure parentali ed esposizione a contaminanti. Mentre il genoma in un organismo pluricellulare è lo stesso in ogni sua cellula, l'epigenoma, ovvero l'insieme delle informazioni epigenetiche, è diverso per ogni tipo cellulare, è dinamico, soggetto a cambiamenti durante il corso della vita, ma

anche sufficientemente stabile nella mitosi e nella meiosi per essere ereditabile (Salozhin et al., 2005). In modelli animali è stato dimostrato che, se modificazioni di tipo epigenetico avvengono nella linea germinale, l'interazione tra geni e ambiente può avere conseguenze nelle generazioni future con importanti effetti ecotossicologici (Head et al., 2014).

I fattori trascrizionali che riconoscono e si legano a sequenze specifiche del DNA (promotori, *enhancers* e *silencers*) giocano un ruolo cruciale nella regolazione dell'espressione genica, tuttavia sono in grado di spiegare solo in parte la complessità del fenomeno all'interno delle cellule (Bird, 2002). Anche la struttura della cromatina, in particolare a livello di istoni, e la sequenza stessa del DNA sono soggette a modificazioni di tipo chimico che incidono sul profilo di espressione dei geni. Gli istoni sono proteine strutturali responsabili dell'impacchettamento del DNA in cromatina alle quali sono associate una serie di modificazioni (metilazione, acetilazione, fosforilazione e ubiquitinazione) a livello di residui aminoacidici specifici. Ad oggi le modificazioni istoniche più studiate sono la metilazione e l'acetilazione dei residui di lisina delle estremità N-terminali degli istoni H3 e H4. I risultati che emergono mostrano che la metilazione in alcune posizioni è associata ad una repressione della trascrizione, mentre in altre ad uno stato attivo della cromatina; nonostante il duplice effetto, la metilazione dei residui aminoacidici risulta essere principalmente coinvolta nel fenomeno del silenziamento della trascrizione (Salozhin et al., 2005). La metilazione del DNA è uno tra i più importanti meccanismi epigenetici coinvolti nel controllo dell'espressione genica.

1.5.1 La metilazione del DNA

La metilazione del DNA è una modificazione enzimatica che negli eucarioti è localizzata generalmente nelle citosine in corrispondenza della sequenza dinucleotidica simmetrica 5'-CG-3' (siti CpG). La maggior parte dei siti CpG non metilati si trova in cluster all'interno di corte sequenze di DNA (200-2000 bp) denominate *CpG islands* (Feil and Fraga, 2012) localizzate soprattutto nelle regioni promotoriali, ma anche associate a sequenze ripetute (Foley et al., 2009), ai primi due esoni e al primo introne del trascritto primario sottolineandone l'importanza nella regolazione dell'espressione genica (Strömqvist et al., 2010). Solitamente queste regioni non vengono metilate durante i primi stadi di sviluppo e lo stabilirsi dell'espressione genica tessuto-specifica

(Salozhin et al., 2005). Bassi livelli di metilazione, ovvero ipometilazione delle *CpG islands* in regioni intra o extra geniche, hanno evidenziato un aumento nell'attività di trascrizione del gene specifico; d'altro canto vi è una forte correlazione tra ipermetilazione di questi siti e *gene silencing* (Strömqvist et al., 2010). Il livello di espressione genica è generalmente più elevato per geni con livelli di metilazione più bassi (Choi et al., 2010). Nelle piante si verifica infatti anche in altri contesti: CHG e CHH dove H può rappresentare adenina, timina o citosina (Feil and Fraga, 2012).

In particolare l'enzima DNA metiltransferasi (DNMT) catalizza l'aggiunta del gruppo metile (CH₃), trasferito dal donatore di metili S-adenosilmetionina (SAM), al carbonio in posizione 5 della citosina coinvolta e trasformata così in 5-metilcitosina (Strömqvist et al., 2010). Nel genoma dei mammiferi è stata identificata una famiglia di DNA metiltransferasi: DNMT1 che sembra avere un ruolo fondamentale nel mantenere i pattern di metilazione durante la replicazione del DNA perché in grado di metilare i filamenti di DNA emimetilati (Bird, 2002); DNMT3A, DNMT3B e DNMT3L capaci di stabilire pattern di metilazione de novo nei primi stadi di sviluppo (Head, 2014). Nel regno animale i vertebrati hanno il numero più elevato di citosine metilate distribuite su gran parte del genoma (Bird, 2002); circa l'80% dei siti CpG sono soggetti a metilazione (Salozhin et al., 2005). In particolare nei mammiferi la percentuale di genoma metilato varia tra 60 e 90% (Head, 2014).

1.5.2 Il meccanismo di metilazione del DNA: le proteine MCP

L'espressione genica può essere influenzata dalla metilazione del DNA direttamente o indirettamente. Molti fattori di trascrizione possono legarsi in modo specifico a sequenze promotoriali solo se gli eventuali siti CpG presenti sono privi di metilazione; in caso contrario non riescono ad interagire e così ad attivare la trascrizione del gene corrispondente (Bird, 2002). Nel secondo meccanismo invece sono coinvolti complessi proteici attratti dalla metilazione dei dinucleotidi CpG. In particolare è stata caratterizzata una famiglia di 5 proteine nucleari denominate *methyl-CpG-binding protein*: MeCP2, MBD1, MBD2, MBD3 e MBD4. Per l'interazione di MeCP2 è sufficiente anche un solo sito CpG metilato nella regione d'interesse e il dominio funzionale principalmente responsabile del legame con il DNA è localizzato nella parte N-terminale della stessa ed è chiamato MBD (*methyl-DNA-binding domain*). MeCP2 è la proteina più studiata all'interno della famiglia perché regola l'espressione di BDNF

(*brain-developed neurotrophic factor*), una neurotropina determinante nel corretto sviluppo del sistema nervoso; inoltre la sua mancanza o la perdita di funzionalità portano alla manifestazione di sintomi analoghi a quelli manifestati da pazienti affetti dalla sindrome di Rett (grave ritardo nell'acquisizione del linguaggio e della coordinazione motoria, disfunzioni della memoria, movimenti stereotipati e irregolarità nella respirazione) dove in molti casi sono state identificate mutazioni nella sequenza genica codificante per MeCP2. Anche MBD1 ha un dominio MBD che le permette di interagire con sequenze di DNA metilate, ma ne possiede anche altri capaci di legare fattori associati all'inattivazione della cromatina (*CAF- chromatin associated factor*). A differenza di MBD2 e MBD4, MBD1 è in grado di interagire anche quando il pattern di metilazione non è correttamente mantenuto (Salozhin et al., 2005). MBD4, oltre ad essere coinvolta nella repressione della trascrizione, appartiene al sistema di riparazione del DNA che interviene sulla correzione delle citosine mutate. Le citosine metilate rappresentano infatti degli *hot spot* di mutazione perché in seguito a deaminazione spontanea sono soggette alla conversione per transizione in timine (Salozhin et al., 2005). MBD2 ha un'identità di sequenza del 70% nei confronti di MBD3, ma mentre dal punto di vista funzionale MBD2 è semplicemente una subunità di un'altra *methyl-CpG-binding protein* (MeCP1), MBD3 fa parte di uno dei principali complessi di rimodellamento ATP-dipendente della cromatina: NuRD (*nucleosome remodeling and histone deacetylase complex*) che, attraverso la deacetilazione dei residui istonici N-terminali, permette la formazione di cromatina condensata (eterocromatina) così l'inattivazione del processo di trascrizione (Salozhin et al., 2005). L'importanza della presenza di MBD3 è stata sottolineata da un modello animale di topo *Knockout*, per il gene di MBD3, per il quale questa mancanza è risultata letale già nella prima fase embrionale (Salozhin et al., 2005). Di recente è stata identificata in vitro anche *Kaiso*, il cui legame ad un preciso motivo nucleotidico (5' m^5CGm^5CG) sembra essere altamente condizionato dalla presenza di citosine metilate (Bird, 2002). L'interazione con il DNA non è garantita dalla presenza di un dominio MBD, ma da un dominio caratterizzato da un motivo strutturale denominato *zinc fingers* (Salozhin et al., 2005).

1.5.3 La metilazione nel mondo animale

Anche se durante i primi stadi di sviluppo avvengono diverse modificazioni epigenetiche associate alla differenziazione cellulare tessuto-specifica, la maggior parte dei cambiamenti si conclude con la fase iniziale dell'embriogenesi (Salozhin et al., 2005) durante la quale avviene una sorta di rimodellamento a livello epigenetico (Foley et al., 2009). Le variazioni dello stato di metilazione sono state studiate principalmente in *Mus musculus* (Figura 2) e *Danio rerio*, specie modello tipiche della ricerca biomedica.

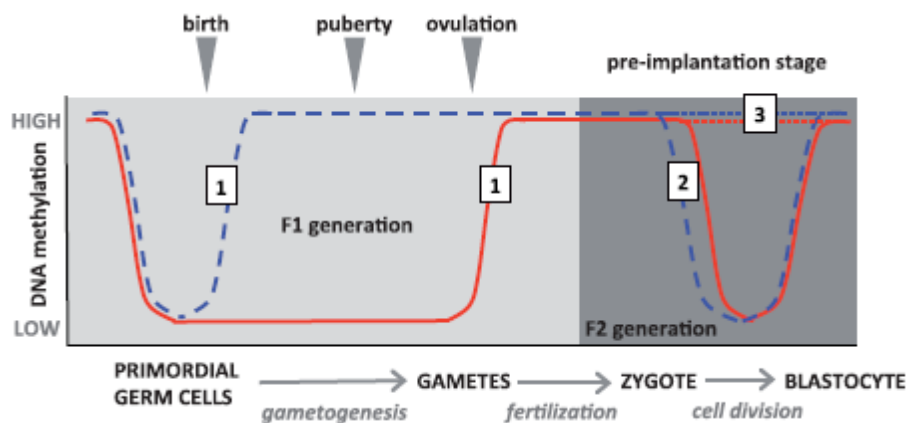


Figura 2 Schematizzazione del *remodeling* epigenetico in *Mus musculus* (Head, 2014)

La prima riduzione nel livello globale di metilazione (demetilazione) si verifica nelle cellule germinali primordiali durante la fase di proliferazione di spermatogoni e oogoni (Bird, 2002). Nel topo è stato visto che la metilazione degli spermatozoi si ristabilisce quasi subito, mentre per gli oociti solo nel momento dell'ovulazione (Head, 2014). La seconda demetilazione dell'intero genoma, esclusi i geni *imprinted*, avviene brevemente dopo la fecondazione, ma prima dell'impianto dell'embrione. A differenza del topo, nella specie modello *Xenopus laevis* l'embrione mantiene elevati livelli di metilazione durante tutto il primo periodo dell'embriogenesi (Head, 2014). Successivamente alla fase di impianto la metilazione è progressivamente ripristinata de novo secondo il *pattern* somatico ereditato (Head, 2014). Anche se la metilazione non può essere mantenuta a livello di singola citosina nei siti CpG (si stima che la metilazione venga persa con una frequenza del 5% per sito ad ogni divisione cellulare) (Bird, 2002), lo stato globale delle isole CpG sembra mantenersi stabile durante lo sviluppo (Bird, 2002). Inoltre quando il *pattern* di metilazione supera indenne la fase di

remodeling durante la formazione dei gameti può essere ereditato dalla generazione successiva (Foley et al., 2009).

A differenza dei vertebrati, gli invertebrati hanno un *pattern* di metilazione del DNA estremamente variabile intraspecie: il genoma di alcuni è privo di basi metilate (*Drosophila melanogaster*, il nematode *Caenorhabditis elegans*), mentre quello di altri presenta livelli di metilazione simili a quelli dei vertebrati (la falena *Mamestra brassicae*). Anche altri insetti come formiche, afidi, api, vespe, mosche, scarabei presentano sostanziali e interessanti livelli di metilazione (Head, 2014). Nel *phylum* dei molluschi, riferimento frequente per il monitoraggio della salubrità dell'ambiente marino, sono state individuate citosine metilate nelle specie *Daphnia pulex*, *Daphnia magna*, *Crassostrea gigas*, *Chlamys farreri* e *Donax trunculus*. La differenza principale riguarda i siti target per la metilazione animale, ovvero i CpG, che invece negli invertebrati sono concentrati in estesi domini di DNA metilato intervallati da equivalenti regioni non metilate. Inoltre le regioni codificanti tendono a presentare una maggiore metilazione rispetto a quelle intergeniche (Head, 2014). La diversità riscontrata nel regno animale potrebbe riflettere differenze nel ruolo e nella funzionalità della metilazione (Bird, 2002) che negli invertebrati sembra regolare lo *splicing* alternativo, ovvero un processo post- trascrizionale attraverso il quale, mediante un diverso arrangiamento degli esoni (regioni di mRNA codificanti), da uno stesso gene possono derivare diverse isoforme proteiche.

1.5.4 Implicazione dell'epigenesi nei processi fisiologici

I principali processi biologici nei quali è coinvolta la metilazione del DNA sono: l'*imprinting* genomico, l'inattivazione del cromosoma X, il silenziamento dei trasposoni e l'oncogenesi.

L'*imprinting* genomico è un fenomeno che si verifica nella maggior parte dei mammiferi, ma ad oggi non è chiaro se sia presente anche in altri organismi vertebrati (Head, 2014). In particolare nel genoma dei mammiferi vi è una sottoclasse di loci omologhi, definiti geni *imprinted*, la cui espressione dipende dall'origine, paterna o materna (Salozhin et al., 2005); di conseguenza alcuni di questi geni sono espressi se ereditati dal padre, altri se ereditati dalla madre. In particolare i due alleli di questi geni sono influenzati in modo differenziale dalla metilazione del DNA e dalle modificazioni

istoniche alle quali i loci *imprinted* vanno in contro nella linea germinale, quando i cromosomi parentali segregano nei gameti. Uno dei due alleli parentali viene infatti mantenuto in uno stato in parte represso dalla metilazione della sequenza a livello di isole CpG promotoriali di solito non metilate. In genere i geni sottoposti a *imprinting* mostrano caratteristiche comuni e spesso sembrano essere raggruppati in grandi *cluster* all'interno dei cromosomi. Inoltre l'espressione di molti geni *imprinted* all'interno di un determinato dominio è spesso controllata da un comune elemento regolativo (Salozhin et al., 2005). In generale, i geni espressi dall'allele paterno, al contrario di quelli espressi dall'allele di origine materna, implementano la crescita e la differenziazione cellulare. La normale regolazione ed espressione dei geni *imprinted* è determinante per l'ontogenesi: alterazioni nell'espressione possono portare allo sviluppo di tumori o a malattie genetiche congenite come la sindrome di *Prader-Willy* e la sindrome di *Angelman* causate dalla perdita di uno o più geni, presenti nel braccio lungo del cromosoma 15, espressi dall'allele paterno nel primo caso e materno nel secondo.

Come l'*imprinting* genomico, anche l'inattivazione del cromosoma X è coinvolta nel silenziamento di un solo allele senza influenzare l'altro con la differenza che tutti i geni sono coinvolti (Bird, 2002). Questo normale processo biologico coinvolge le femmine di mammifero che, a differenza dei maschi, possiedono due cromosomi X e di conseguenza due alleli per ogni locus *X-linked*. Con l'inattivazione casuale di uno dei due cromosomi X femminili (cellule diverse di uno stesso organismo possono avere un diverso cromosoma X attivo), allo stadio di blastocisti, l'espressione di questi geni è compensata e resa equivalente al genere maschile. In generale, considerando l'intero cromosoma, il pattern di metilazione dell'X inattivo è invertito rispetto a quello dell'X attivo: le isole CpG delle regioni promotoriali sono metilate, mentre si presentano non metilate nel cromosoma X attivo. Tuttavia un recente studio ha dimostrato che il livello di metilazione delle sequenze geniche del cromosoma X trascrizionalmente attivo è più del doppio rispetto a quello delle sequenze geniche dell'X inattivo (Figura 3).

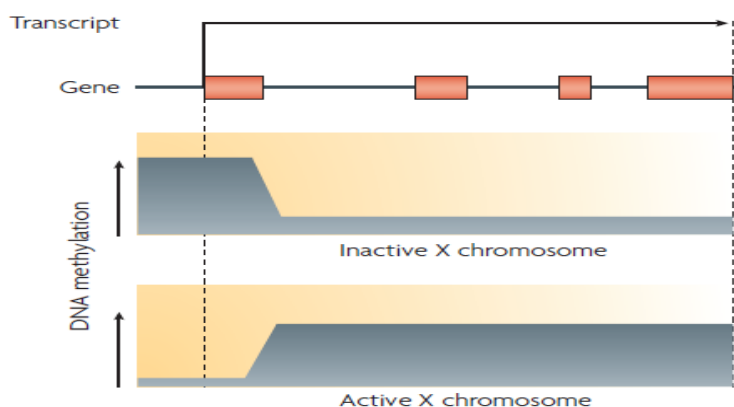


Figura 3 Stato di metilazione dei cromosomi X inattivo e attivo (Suzuki and Bird, 2008)

Una possibile spiegazione può essere che la maggior parte del genoma dei mammiferi è altamente metilata e di conseguenza anche esoni e introni sono coinvolti. Tuttavia è difficile determinare se la metilazione sia indirizzata nei confronti dei geni o sia uno stato di *default* che interessa regioni geniche come del resto regioni non codificanti. Inoltre dato che lo stato di metilazione dei due cromosomi X di partenza è lo stesso, l'ipometilazione del cromosoma X inattivo potrebbe essere il risultato di una demetilazione, durante l'inattivazione, rispetto allo stato normale (Suzuki and Bird, 2008).

Un'altra importante funzione della metilazione del DNA, riscontrata in particolare nei genomi di piante (Suzuki and Bird, 2008) è la repressione trascrizionale di elementi trasponibili. I trasposoni sono sequenze di DNA in grado di spostarsi da un sito ad un altro del genoma inserendosi in regioni codificanti alterandone la funzionalità o modificandone il livello di espressione se l'inserimento avviene nelle regioni promotoriali. L'inattivazione di questi elementi è necessaria per prevenire danni al DNA dovuti a riarrangiamenti (mutazioni geniche, malattie e in alcuni casi tumori) e inoltre per garantire il corretto funzionamento del meccanismo di espressione genica eliminando le possibili interferenze (Bird, 2002). In questo caso il coinvolgimento della metilazione potrebbe essere richiesto per la capacità di stabilizzare un'inattivazione trascrizionale già organizzata (Bird, 2002). Infatti il trascritto di regioni genomiche interessate dall'inserzione di elementi trasponibili può essere degradato da brevi sequenze di RNA d'interferenza (siRNA) a singolo filamento in grado di legarsi in modo complementare all'mRNA in oggetto (Salozhin et al., 2005).

1.5.5 Modificazioni epigenetiche associate ad interazione con l'ambiente

La maggior parte degli eventi associati a cambiamenti nel livello di metilazione globale del genoma (cicli di demetilazione e metilazione de novo) si concludono durante le prime fasi di sviluppo embrionale o comunque prima della nascita. Anche nella fase di sviluppo successiva e nel corso della vita il profilo epigenetico è dinamico e soggetto a variazioni, di minore entità, che sono il risultato dell'interazione dell'organismo con l'ambiente in cui vive, le possibili sostanze tossiche presenti e le abitudini alimentari (Salozhin et al., 2005). Tuttavia la situazione cambia drammaticamente nel caso di insorgenza di malattie complesse come asma, allergie, obesità, diabete di tipo II, malattie cardiovascolari e neurologiche, ma soprattutto di neoplasie (Foley et al., 2009). In particolare il *pattern* di metilazione legato ai tumori presenta delle caratteristiche anomale che lo contraddistinguono: demetilazione generale del genoma, singolare ipermetilazione dei siti CpG di determinati geni, di geni di controllo ad espressione costitutiva (geni *housekeeping*) e di oncosoppressori (Salozhin et al., 2005). Le patologie tumorali si distinguono sulla base della localizzazione differenziale dell'ipermetilazione o dell'ipometilazione, ma soprattutto per quanto riguarda lo stato di metilazione di oncogeni e oncosoppressori. La diminuzione del livello trascrizionale di geni oncosoppressori in tessuti neoplastici è dovuta ad un aumento della metilazione delle *CpG islands* presenti di circa il 100% rispetto al tessuto normale. La differenza osservata nei tumori è più rilevante di quella riscontrata nelle altre patologie influenzate da modificazioni epigenetiche, dove il livello di metilazione delle isole CpG tra tessuti affetti e non si differenzia del 10% (Foley et al., 2009). Data l'importanza della metilazione del DNA nella patogenesi tumorale, di recente sono stati effettuati numerosi studi volti ad individuare possibili marcatori epigenetici di malattia in tessuti target e linee cellulari integrando così i dati derivanti dalle indagini genomiche e trascrittomiche. Questo approccio integrato permette di avere un quadro più chiaro sulla malattia per quanto riguarda i meccanismi, gli effetti dell'ambiente e dell'ereditarietà (Birney, 2011). Lo studio di Ma et al. (2014) è andato oltre alla caratterizzazione dell'epigenoma e all'individuazione di possibili biomarcatori dell'alterazione tissutale perché ha esaminato le correlazioni nel profilo di metilazione tra tessuti target e tessuti surrogato dimostrando che si può predire lo stato di metilazione sulla base di questi ultimi. Studi precedenti hanno infatti evidenziato che il *pattern* di metilazione è altamente conservato tra i diversi tessuti e inoltre che la

correlazione, per quanto riguarda i tessuti, tra individui è elevata (0.852) (Byun et al., 2009). Questo ha permesso di sviluppare lo studio di metilazione, riguardante la fibrillazione atriale, sul tessuto di atri e arterie utilizzando il sangue, o meglio i leucociti come matrice surrogato attraverso la messa a punto di calibrazioni statistiche. Di conseguenza i biomarcatori epigenetici identificati nella matrice surrogato possono essere utilizzati in fase di *screening* e diagnosi precoce (Ma et al., 2014).

1.5.5.1 Contaminanti ambientali

Lo studio degli effetti dei contaminanti ambientali sulla metilazione del DNA rappresenta un ambito di ricerca attuale e in rapida espansione. Diversi studi su specie di mammiferi (*Mus musculus*), pesci e invertebrati infatti dimostrano che numerosi contaminanti influenzano il pattern di metilazione a livello globale, generalizzato o di determinati loci genomici. In particolare è stato studiato l'effetto epigenomico delle seguenti classi di sostanze: metalli pesanti, composti che alterano il sistema endocrino, pesticidi e inquinanti dell'aria (Head et al., 2014). In molti casi le variazioni epigenetiche sono state correlate all'espressione fenotipica dell'organismo. L'esposizione al dietilstilbestrolo, estrogeno di sintesi, ha mostrato in *Mus musculus* una diminuzione nel livello di metilazione di promotori genici nei tessuti degli organi riproduttivi e in particolare dell'oncogene *c-fos*. Il dietilstilbestrolo è infatti associato ad un aumento di adenocarcinomi della vagina e della cervice uterina nella progenie di sesso femminile, mentre ad ipotrofia testicolare nell'apparato genitale maschile (Foley et al., 2009). Per quanto riguarda il metilmercurio, mentre in *Mus musculus* si osserva un'ipermetilazione del promotore di *BDNF* (*brain-derived neurotrophic factor*) che ne reprime l'espressione genica e sfocia in sintomi depressivi; nell'orso polare è associato ad una ipometilazione del tessuto nervoso. In topi esposti invece all'inquinamento atmosferico è stata osservata un'ipermetilazione a livello di spermatozoi che sembra rimanere anche successivamente senza una nuova esposizione. Quando le alterazioni epigenetiche si verificano nella linea germinale gli effetti legati ai contaminanti possono essere ereditati dalle generazioni successive non esposte. Tuttavia l'unico esempio che lo descrive riguarda l'esposizione di ratti al fungicida *vinclozolin* o al pesticida *methoxychlor*, associata, oltre che a cambiamenti globali nella metilazione del DNA, ad una diminuita fertilità della progenie maschile per diverse generazioni senza nuove esposizioni ai contaminanti (Head et al., 2012). In conclusione l'effetto dei

contaminanti sullo stato di metilazione del DNA dipende sostanzialmente dalla sostanza considerata e dalla regione genomica interessata.

1.6 Metodiche epigenetiche

La comprensione dei molteplici aspetti della metilazione richiede poi di conoscere quali sono le regioni genomiche interessate. In quest'ottica la disponibilità crescente di genomi sequenziati parallelamente all'avanzamento del *next-generation sequencing* ha permesso alle metodiche molecolari ad alta risoluzione l'analisi puntuale del profilo di metilazione dell'intero genoma (Harris et al., 2010). Negli anni sono stati pubblicati numerosi studi riguardanti analisi del *pattern* di metilazione e con esse sono stati illustrati molteplici approcci metodologici riconducibili principalmente alle seguenti tecniche di pre-trattamento del genoma: trattamento con bisulfito, digestione con enzimi di restrizione metilazione-sensibile e arricchimento per affinità con i siti CpG metilati (Zilberman and Henikoff, 2007). Ad oggi il sequenziamento di librerie trattate con bisulfito è considerato il gold standard perché offre un'informazione di tipo quantitativo circa lo stato di metilazione di ciascuna citosina (Ivanov et al., 2013). Il trattamento con sodio bisulfito provoca la deaminazione delle citosine non metilate in uracile proteggendo invece dalla conversione le citosine metilate. Con la successiva amplificazione tramite PCR (*Polymerase chain reaction*) l'uracile è sostituito dalla base timina. I genomi convertiti con bisulfito possono essere analizzati attraverso approcci basati su array o su tecnologie di sequenziamento classico (*Sanger*) o *high throughput* (*454 sequencing*, *Illumina HiSeq*, *Solexa bisulphite sequencing*) (Susuki and Bird, 2008). Anche se il trattamento con sodio bisulfito non è facilmente adattabile all'analisi su *array* sono state sviluppate diverse piattaforme per lo studio dei polimorfismi di metilazione (Zilberman and Henikoff, 2007). La tecnologia *BeadArray* sviluppata da Illumina, a differenza delle altre, permette un'analisi quantitativa ad alta risoluzione (*single-base resolution*) saggiando fino a 1536 loci CpG simultaneamente in 96 campioni. Per ogni locus da saggiare sono disegnati due *primers*, ciascuno marcato con un fluoroforo diverso a seconda del diverso stato di metilazione (citosina metilata non convertita o citosina non metilata) della sequenza in oggetto. Nella fase di amplificazione questi *primers* sono utilizzati in combinazione con *primers* sempre locus specifici, ma insensibili alla metilazione. Infine i prodotti di amplificazione si ibridano al supporto e viene rilevato il segnale di fluorescenza che permette di definire lo stato di

metilazione locus specifico e di discriminare i diversi campioni secondo la variabile biologica oggetto di indagine (Zilberman and Henikoff, 2007). Tuttavia questo approccio necessita di un ampio set di *primers* per lo screening di loci d'interesse. Proprio per la difficoltà nella progettazione e nella preparazione degli *array*, la metodica d'elezione per l'analisi di genomi trattati con bisulfito è il sequenziamento (*WGBS*, *MethylC-seq*, *BS-seq*), in particolare il *next-generation sequencing* in grado di generare un'elevata quantità di dati in un'unica corsa (180-200 milioni di *reads*) offrendo un risultato quantitativo sicuro e robusto riferito all'intero genoma con risoluzione a singola base (Zilberman and Henikoff, 2007). L'efficienza del trattamento con bisulfito nella completa conversione delle citosine non metilate è essenziale per ottenere un output altamente informativo e accurato. Inoltre un altro problema è la diminuzione del grado di complessità dei genomi trattati con bisulfito perché, nell'analisi bioinformatica successiva al sequenziamento, può influenzare l'allineamento univoco delle *reads* al genoma di riferimento. Per questo motivo sono state sviluppate strategie di sequenziamento basate su sequenze più lunghe, generalmente più complesse, e *paired-end* per avere, oltre all'informazione della sequenza, un vincolo in più per l'allineamento, ovvero la distanza tra *reads* accoppiate (Laird, 2010). Tuttavia attualmente il *whole-genome bisulfite sequencing* non è utilizzabile come metodo di screening negli studi epigenomici a causa dell'elevato costo del sequenziamento e dell'elevato numero di campioni richiesti per avere risultati statisticamente significativi (van den Oord, 2013).

Una valida alternativa al sequenziamento dell'intero genoma è rappresentata dall'analisi della metilazione mediante arricchimento selettivo del DNA frammentato grazie al legame con anticorpi monoclonali (*MeDIP –Methylation analysis by DNA immunoprecipitation*) o con *methyl-binding protein* (MBD) utilizzando la metodica *MBD affinity purification* (MAP) (Suzuki and Bird, 2008). Tuttavia mentre gli anticorpi monoclonali, commercialmente disponibili, sono in grado precipitare citosine metilate presenti in tutti i contesti (CpG, CHG o CHH), il dominio proteico delle MBD ha un'elevata affinità nei confronti solo delle citosine metilate all'interno delle *CpG islands* (Suzuki and Bird, 2008). Entrambe le metodiche richiedono un'elevata densità di metilazione. Di conseguenza potrebbero avere difficoltà nel discriminare tra frammenti di DNA con siti target distanti e frammenti non metilati (Zilberman and Henikoff, 2007), come nel caso dei genomi dei mammiferi altamente metilati, ma con

bassa densità di siti CpG e isole CpG generalmente non metilate. Per l'analisi *whole-genome* di questo tipo di genomi è riportata anche la possibilità di arricchimento per frammenti non metilati, isolandoli dalla frazione legata ad anticorpi o MBD (Zilberman and Henikoff, 2007). Infine le regioni genomiche d'interesse, metilate o non metilate, possono essere rilevate attraverso tecnologia *microarray* o sequenziamento. Diversamente dal sequenziamento dove viene sottoposta ad analisi solo la frazione arricchita d'interesse, la metodologia *microarray* richiede l'ibridazione di entrambe, marcate con fluorofori diversi (Zilberman and Henikoff, 2007) per identificare regioni genomiche differenzialmente metilate in studi caso-controllo.

Una metodica classica per lo studio dei profili di metilazione è la digestione del genoma con enzimi di restrizione arricchendo sia per DNA metilato (*HpaII*) che per DNA non metilato (McrBC). L'utilizzo degli enzimi di restrizione è particolarmente utile nel confronto tra campioni con pochi polimorfismi (Zilberman and Henikoff, 2008). Anche la metodica RRBS (*Reduced Representation of Bisulfite-Seq*) si basa sull'utilizzo di un'endonucleasi di restrizione (*MspI* che taglia in corrispondenza di siti CpG metilati), ma è abbinata anche a trattamento con bisulfito riducendo la porzione di DNA da sequenziare e offrendo la possibilità di quantificare la metilazione su regioni ricche in CpG ad elevata risoluzione (Stevens et al., 2013).

Per la presente sperimentazione è stata messa a punto una strategia integrata che prevede l'arricchimento del DNA genomico frammentato con MBD seguito dal trattamento con bisulfito della frazione metilata purificata. La selettività dell'arricchimento rimane la stessa del protocollo originale (Ivanov et al., 2013), mentre la precisa quantificazione dei livelli di metilazione a livello di singola citosina è garantita dal trattamento con bisulfito. In particolare per la prima volta viene applicata una metodica *whole-genome sequencing* per individuare marcatori epigenetici di trattamento mediante confronto tra campioni di muscolo scheletrico di bovini trattati con promotori della crescita e controlli.

2. OBIETTIVI

Oggetto del presente studio è lo sviluppo di marcatori genomici innovativi per rilevare l'utilizzo illecito di promotori della crescita nell'allevamento del bovino da carne. Attualmente l'identificazione e la caratterizzazione di possibili biomarcatori di trattamento in muscolo scheletrico è stata condotta attraverso l'analisi dei profili trascrizionali (trascrittomica) o dei pattern di espressione proteici (proteomica). Lo scopo del lavoro è quello di mettere in luce regioni genomiche differenzialmente metilate in seguito alla somministrazione di promotori della crescita, da proporre come potenziali biomarcatori epigenetici. Per lo studio dei pattern di metilazione (trattato-controllo) è stato determinante l'utilizzo di strumenti bioinformatici e statistici che, grazie alla loro spiccata capacità di supportare e mediare elevate quantità di dati, hanno reso possibile l'analisi del genoma sequenziato. L'intento è quello di esplorare le potenzialità di nuovi biomarcatori genomici nel rilevare la risposta biologica legata alla somministrazione di sostanze anabolizzanti e contribuire quindi alla sicurezza sanitaria del prodotto carne.

3. Materiali e metodi

3.1 Campionamento

La presente sperimentazione si inserisce in un ampio programma di campionamento che ha coinvolto 40 bovini maschi di razza Frisona allevati nella provincia di Mantova (SaZoo-UNIALLEVA, Gonzaga) e macellati a 6-7 mesi d'età con un peso vivo compreso tra i 200 e i 290 kg. Gli animali, suddivisi in 4 gruppi distinti, sono stati trattati sperimentalmente con diverse combinazioni di promotori della crescita. Un quinto gruppo, costituito da 10 capi, non è stato sottoposto a trattamento per rappresentare il controllo. Per tutti i gruppi di allevamento è stato seguito lo stesso programma di razionamento caratterizzato da sostitutivo del latte in polvere (proteina 22.5%, grasso 22.5%, ceneri 6%, fibra 0.3%, calcio 0.65%, fosforo 0.65%, sodio 0.5%; vitamina A 25.000 IU/kg, vitamina C 50 mg/kg, vitamina D3 3.700 IU/kg, vitamina E 75 mg/kg e Cu 5 mg/kg), per due volte al giorno, fino a raggiungere gradualmente i 16 L/capo/giorno, e da 0.5 kg/capo/giorno di paglia d'orzo (sostanza secca 90%, fibra 41%, ceneri 6%, proteina 2% e grasso 1%) addizionata ad un mese d'età. La dieta è stata formulata seguendo le linee guida indicate dalla Decisione n. 182/97 della Commissione europea. La tabella 1 riassume i principali dati relativi alla sperimentazione.

Tabella 1 Suddivisione dei capi e relativi trattamenti

Gruppo	Numero capi	Trattamento
A	8	Estradiolo benzoato + Brotizolam
B	8	Prednisolone
C	6	Estradiolo benzoato + Desametasone
D	8	Ractopamina
E	10	Controlli

In particolare in questo studio sono stati analizzati 16 campioni di muscolo *longissimus dorsi* appartenenti ai 10 animali del gruppo di controllo (gruppo E) e ai 6 a cui sono stati somministrati estradiolo benzoato per via intramuscolare (5 mg/capo/settimana) per 6 settimane e desametasone per os (0.4 mg/capo/giorno) per 31 giorni (gruppo C). Il campionamento è stato eseguito il giorno stesso della macellazione (luglio 2011). Il

tessuto prelevato è stato stoccato in *RNAlater* e conservato a -20 °C. Successivamente le carcasse sono state distrutte secondo la Direttiva europea n. 74/2003 recepita con il Decreto legislativo n. 158 del 16 marzo 2006. La sperimentazione animale è stata condotta secondo quanto stabilito dalla Direttiva europea n. 609/86, recepita dal Governo italiano con il Decreto legislativo n. 116/92, e approvata dal Ministero della Salute.

3.2. Creazione di librerie genomiche attraverso arricchimento con proteine specifiche e trattamento con bisulfito

3.2.1 Estrazione e quantificazione del DNA genomico

Per l'estrazione di DNA dai campioni di tessuto muscolare è stato utilizzato il kit *ZR Genomic DNATM-Tissue MidiPrep* che abbina digestione e lisi a purificazione del DNA totale attraverso colonnina (*Zymo-SpinTM column/filter assembly*), seguendo il protocollo riportato nel manuale.

Prime prove di estrazione di DNA da tessuto muscolare hanno dimostrato come questo tipo di tessuto, strutturato e ad elevato contenuto in proteine, non consenta di ottenere rese elevate in termini di quantità e qualità di DNA. Per questo motivo sono stati introdotti alcuni accorgimenti: in fase di digestione è stata addizionata una maggiore quantità di proteinasi K (55 µL ad una concentrazione di 20 mg/mL) e l'incubazione a 55 °C è stata prolungata a 4-5 ore, mentre in fase finale l'eluizione è stata eseguita due volte con lo scopo di aumentare la quantità di estratto.

La quantità e la qualità del DNA genomico purificato sono state valutate tramite un fluorimetro QUBIT (*Life technologies, Invitrogen*) e uno spettrofotometro NanoDrop (*Thermo scientific*).

Il saggio di quantificazione al QUBIT è basato sulla rilevazione della fluorescenza emessa da uno specifico fluorocromo che si lega in modo aspecifico al dsDNA: la concentrazione viene quindi determinata attraverso una retta di calibrazione data dall'impiego di due standard di DNA a concentrazione nota. Il saggio utilizzato è il *dsDNA BR Assay*, rapido, altamente selettivo e accurato nella misurazione di concentrazioni di DNA comprese tra i 100 pg/µL e i 1000 ng/µL.

A differenza del QUBIT, il NanoDrop sfrutta la capacità degli acidi nucleici di assorbire la luce ultravioletta, con un picco massimo di assorbimento alla lunghezza d'onda di

260 nm, registrandone l'intensità luminosa che ne deriva. L'assorbanza del campione, direttamente proporzionale alla sua quantità, viene calcolata utilizzando il valore di intensità del campione stesso insieme a quello di riferimento dato dal controllo negativo. Il grado di purezza del DNA è dato invece dal calcolo dei rapporti A_{260}/A_{280} e A_{260}/A_{230} : il primo è indicativo per la presenza di sostanze proteiche che assorbono a 280 nm, mentre il secondo riflette la contaminazione da parte di solventi chimici utilizzati in fase di estrazione o carboidrati. Il DNA estratto è considerato puro se tali rapporti presentano valori compresi tra 1.8 e 2.2. Il NanoDrop ha dei vantaggi tecnici notevoli rispetto ad un classico spettrofotometro perché permette di porre una minima quantità dell'estratto in soluzione (1 μ L) a diretto contatto con il sistema ottico, grazie alla sola tensione superficiale, eliminando l'utilizzo di altre provette e riducendo così il rischio di contaminazione. Inoltre è capace di quantificare campioni altamente concentrati senza necessità di diluirli fino ad una concentrazione massima di 3700 ng/ μ L.

Infine con l'elettroforesi su gel di agarosio concentrato allo 0.8% è stata valutata la qualità del DNA estratto in termini di peso molecolare indicativo, integrità ed eventuale contaminazione.

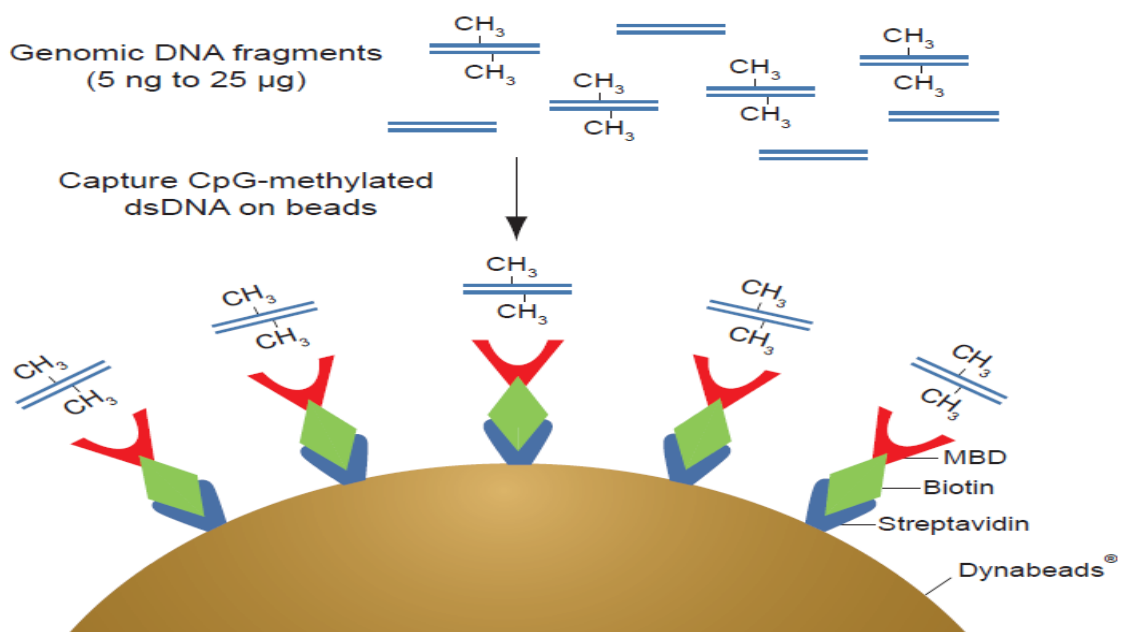
3.2.2 Frammentazione meccanica del dsDNA

La frammentazione meccanica del DNA genomico è stata effettuata con il sonicatore *Covaris S220* per ottenere frammenti uniformi, di 50-100 pb, con lo scopo di ottimizzare l'output del *Next-generation sequencing*. Il funzionamento si basa sulla tecnologia AFA (*Adaptive Focused Acoustics*) che produce ed emette energia acustica ad una frequenza di 500 kHz, attraverso un trasduttore collegato ad un computer, nei confronti del campione in oggetto. In particolare per la presente applicazione sono stati utilizzati i seguenti parametri: *Duty cycle* 10%, *Cycles per burst* 200 e *Intensity* 5 ad una temperatura controllata di 5.5-6°C. La quantità di campione analizzata all'interno dei *microTUBE AFA™* è pari a 50 μ L ad una concentrazione di 40 ng/ μ L, portati a 130 μ L con il *buffer* specifico.

Per verificare la lunghezza dei frammenti ottenuti è stata eseguita un'elettroforesi su gel di agarosio 1,8% con un marcatore di peso molecolare *100 pb DNA ladder* come *reference*.

3.2.3 Arricchimento selettivo del DNA frammentato in base al grado di metilazione

Nel presente studio l'utilizzo del kit commerciale *MethylMiner™ Methylated DNA Enrichment kit (Invitrogen™)* ha permesso di procedere, a partire dal dsDNA frammentato, all'arricchimento ad alta risoluzione di singoli siti metilati CpG (*Cytosine bases located 5' to guanine*). Le proteine MBD (*Methyl-DNA-binding domain*), agganciate a biglie magnetiche *Dynabeads®* associate a streptavidina attraverso il legame con biotina, sono in grado di riconoscere e legare in modo specifico i frammenti che contengono citosine metilate. L'elevata affinità di legame delle proteine *MethylMiner™ MBD-Biotin* offre un saggio più sensibile rispetto al classico arricchimento con anticorpi, mentre le *Dynabeads®* semplificano il recupero dei frammenti dsDNA metilati assicurando una elevata ripetibilità sperimentale. Dopo aver rimosso la frazione di DNA non metilato, si procede all'eluizione dei frammenti d'interesse dalle *Dynabeads®*. L'eluizione può essere effettuata con due diverse modalità: se si vogliono ottenere più frazioni distinte a seconda del grado di metilazione dei frammenti si utilizza una serie di buffer a crescente concentrazione di NaCl (da 200 a 2000 mM), mentre, come nella presente sperimentazione, per ottenere e una singola frazione si utilizza un buffer altamente concentrato (2000 mM di NaCl) capace di recuperare dal 70 al 90% dei frammenti catturati. Il dsDNA eluito viene poi precipitato con etanolo 70% e risospeso in 60 µL di acqua *DNase-free*. La seguente figura illustra come vengono selezionati i frammenti metilati.



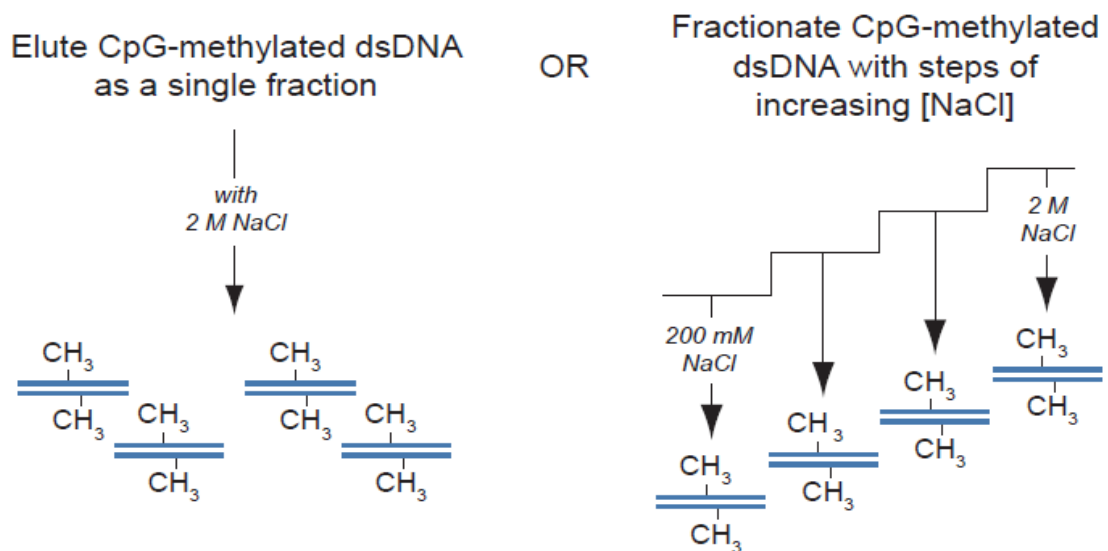


Figura 4 Schema di funzionamento del kit *MethylMiner™ Methylated DNA Enrichment*

Dal punto di vista applicativo, la quantità di partenza è stata di 3 µg di DNA frammentato ad una concentrazione di 25 ng/µL, raggiunta eseguendo opportune diluizioni con volumi di acqua compresi tra 40 e 400 µL, portata poi ad un volume finale di 500µL. In questo studio i frammenti di dsDNA metilati rappresentano circa il 3% del DNA di partenza, in linea con quanto riportato nel manuale. Infine, per avere una misurazione accurata del dsDNA, è stata eseguita una quantificazione al QUBIT e un'elettroforesi su gel di agarosio.

3.2.4 End-repair, A-tailing e ligazione di adattatori

La fase di ligazione di sequenze di adattatori specifiche è stata condotta utilizzando il kit commerciale *Illumina® TruSeq® Nano DNA Sample Preparation Kit LT (low-throughput)*, allo scopo di ottenere frammenti *double strand* idonei alle successive analisi. 100 ng di 350 pb-dsDNA sono stati sottoposti ad una prima fase di *end-repair* mediante conversione delle estremità *protruding*, derivanti dalla frammentazione meccanica, in estremità tronche *blunt-end*, grazie all'attività di esonucleasi 3'-5' e polimerasi 5'-3' contenute nella *End-Repair Mix*. I frammenti così riparati sono stati purificati e selezionati in maniera stringente in base alla lunghezza scelta con l'utilizzo di biglie magnetiche *Sample purification beads*. Successivamente, per evitare la formazione di templati concatenati durante la ligazione, è stata aggiunta una *A-tail* alle estremità 3' *blunt-end* dei frammenti. Infine specifici adattatori, che possiedono una timina in più per la complementarietà di legame con la A-tail, sono stati incorporati ai

frammenti. In particolare questi adattatori a doppio filamento con estremità *protruding* sono costituiti da un sito che permette ai *primer* di sequenziamento di legarsi, da un *index* univoco per i frammenti di ogni campione e da una sequenza all'estremità più esterna, compatibile con la *flow cell* del sequenziatore. Inoltre questi adattatori sono stati progettati con tutte le citosine metilate in modo da non interferire con la successiva fase di trattamento con bisulfito.

3.2.5 Trattamento con bisulfito

Per la conversione del DNA metilato è stato scelto il kit commerciale *MethylCode™ Bisulfite Conversion Kit (Life Technologies)* che consente in un unico passaggio la denaturazione e la conversione dei templati di DNA sottoponendoli al seguente ciclo impostato in un classico termociclatore (*PTC-200 Peltier Thermal Cycler*):

98°C per 10 minuti (denaturazione)

64°C per 2.5 ore (conversione con bisulfito)

La quantità ottimale di DNA di partenza è compresa tra i 200 e i 500 ng in un volume di 20 µL a cui vengono aggiunti 130 µL di *CT Conversion Reagent* che converte le citosine non metilate in uracile (e in timina nella seguente amplificazione) senza alterare quelle metilate. Infine grazie all'utilizzo di colonnine il DNA viene purificato e recuperato in 10 µL di *elution Buffer*.

3.2.6 Amplificazione e sequenziamento delle librerie

Le librerie sono state prodotte con una classica PCR (*Polymerase Chain Reaction*), ottimizzata *ad hoc* per l'inserimento del passaggio di trattamento con bisulfito. Per questo motivo è stata utilizzata la *PfuTurbo Cx HotStart DNA polymerase (Agilent technologies)*, una polimerasi capace di riconoscere la base azotata uracile derivante dal trattamento con bisulfito, e di intensificare l'attività di *proofreading* per evitare di introdurre errori durante il processo. Dal punto di vista operativo sono stati aggiunti 15 µL di DNA ad una mix PCR (*buffer 10X, dNTPs, primer Illumina, polimerasi PfuTurbo Cx e acqua DNase-free*) per un volume finale di 50 µL. In particolare sono stati effettuati 14 cicli di amplificazione per evitare di introdurre errori e per avere una rappresentazione il più omogenea possibile dei frammenti all'interno di ogni singola libreria.

La purificazione finale degli amplificati è stata condotta utilizzando il kit commerciale *AGENCOURT® AMPURE® XP* che possiede una tecnologia (*Agencourt's solid-phase*

paramagnetic bead technology) basata sulla capacità di biglie magnetiche di legare, grazie ad un *buffer* specifico, frammenti a partire dalle 100 bp, separandoli con un lavaggio in etanolo dai possibili contaminanti. Infine i prodotti di PCR purificati sono stati eluiti dalle biglie in *elution buffer*.

Successivamente le librerie sono state valutate quantitativamente al QUBIT e al *Bioanalyzer (2100 Agilent)* per avere un'informazione precisa sulla distribuzione in lunghezza dei frammenti presenti. Il *Bioanalyzer* è uno strumento che permette di eseguire un'elettroforesi su gel miniaturizzata grazie all'allestimento del corrispondente chip *Agilent High Sensitivity DNA* con reagenti specifici (*High Sensitivity DNA dye concentrate, gel matrix, marker e ladder*). Questo chip è caratterizzato da una serie di microcanali interconnessi attraverso i quali i frammenti ds DNA dei campioni in oggetto e del *marker* si separano elettroforeticamente su gel, vengono rilevati mediante fluorescenza ed elaborati dal software (*Agilent 2100 Expert*) in immagini gel simili e elettroferogrammi. L'analisi è rapida, precisa, accurata e richiede un consumo minimo del campione (1 μ L).

Le librerie genomiche sono state sequenziate con tecnologia *Illumina High throughput* (che consente di ottenere un output di circa 180 milioni di *reads*) splittando il *pool* dei campioni in oggetto in 2 diverse *lane* della *flow cell* per ridurre eventuali *bias* legati a questa variabile. In una prima fase i templati, resi a singolo filamento, vengono immobilizzati alla superficie della *flow cell* ricca di *primers*, per amplificare ogni frammento più di un migliaio di volte producendo così diversi milioni di *clusters* di ssDNA per *lane* (*solid-phase bridge amplification*). Successivamente tali *clusters* vengono sequenziati grazie alla loro amplificazione (*sequencing by synthesis*) per mezzo di dNTPs marcati che ad ogni ciclo vengono incorporati e, eccitati da un laser, emettono un segnale di fluorescenza caratteristico identificando di volta in volta la base. I dati sulle sequenze vengono registrati in un *software* che permette di seguire l'analisi in tempo reale. È stato scelto un approccio di sequenziamento "100 paired-end".

3.2.7 Controllo qualità delle sequenze e *trimming*

Le informazioni di sequenza di ogni campione, salvate in file in formato FastQ (sequenziamento *paired-end*), sono state analizzate mediante l'utilizzo di un server multicore con 256 Gb di memoria RAM.

Un primo controllo qualità ha previsto il conteggio delle sequenze e la rimozione di tutte quelle con uno *Phred quality score* inferiore a 20. Il *Phred quality score* è un punteggio ideale per valutare la qualità delle sequenze in oggetto perché è calcolato in relazione all'accuratezza con la quale le basi vengono identificate dal sequenziatore: un valore pari a 20 indica che la probabilità che quella base sia stata chiamata in modo scorretto è di 1 su 100 corrispondente ad un'accuratezza del 99%. È stato quindi condotto un *trimming* dei due file per eliminare gli adattatori e rimuovere le *reads* con lunghezza inferiore alle 20 bp, utilizzando i seguenti parametri :

- *paired*: indica il metodo di sequenziamento utilizzato
- *trim 1*: toglie una base al 3' di ciascuno dei due filamenti delle *reads*
- *stringency 10*: elimina gli adattatori riconosciuti da almeno 10 basi
- *length 80*: elimina sequenze inferiori a 80 bp

3.2.8 Mapping delle sequenze sul genoma di riferimento

I file FastQ sono stati analizzati mediante il programma *Bismark* (version 0.12.1, www.bioinformatics.babraham.ac.uk/projects/bismark/), il quale esegue un allineamento delle diverse *reads* contro un genoma di riferimento e, allo stesso tempo, definisce lo stato di metilazione di tutte le citosine (*cytosine methylation call*).

Affinchè il *mapping* avvenga correttamente il genoma di riferimento (*Bos taurus* UMD3.1) deve essere precedentemente convertito in silico allo stesso modo delle librerie genomiche trattate con bisulfite (citosine non metilate in timine e di conseguenza adenine in guanine). Successivamente grazie a *Bowtie 1*, strumento che esegue il *mapping* di brevi sequenze, vengono corsi 4 allineamenti in parallelo per determinare il migliore allineamento possibile di ciascuna *reads* (Figura 5).

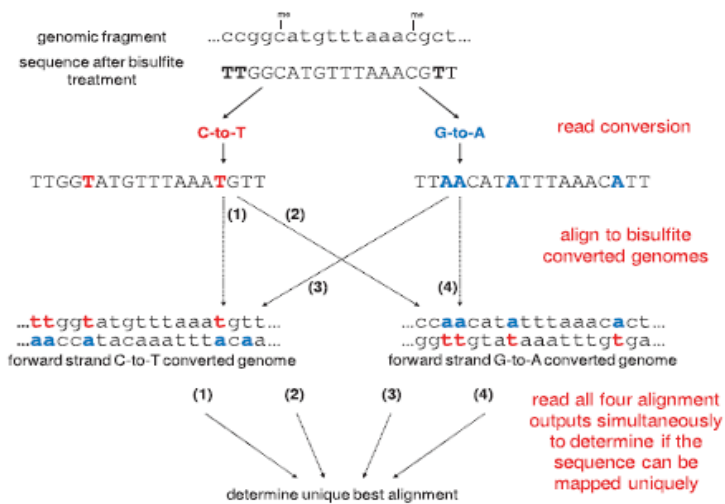


Figura 5 Algoritmo utilizzato da *Bowtie* per il *mapping*

Il *mapping* richiede inoltre di specificare se le librerie da analizzare sono state costruite in modo direzionale o non direzionale. Nel presente studio le librerie non sono direzionali in quanto sono stati sequenziati tutti i 4 possibili filamenti di una data *read*; *Bismark* esegue quindi l'allineamento non solo nei confronti dei 2 filamenti *forward* e *reverse* originali, ma anche dei 2 filamenti complementari a questi. *Bismark* restituisce i dati relativi al *mapping* dei due filamenti di ciascuna *read*, identificata da un codice specifico, all'equivalente sul genoma di riferimento e sulla chiamata delle citosine distinguendole in metilate e non. Oltre a questo *Bismark* offre per ogni campione un report contenente statistiche descrittive sull'efficienza del *mapping* (percentuale di *reads* allineate in modo univoco, con allineamento multiplo e non allineate) e una prima indicazione sulla percentuale di citosine metilate separata per contesto: CpG, CHG o CHH (H = A, T o C). Le informazioni ottenute da *Bismark* sono salvate di default in formato SAM.

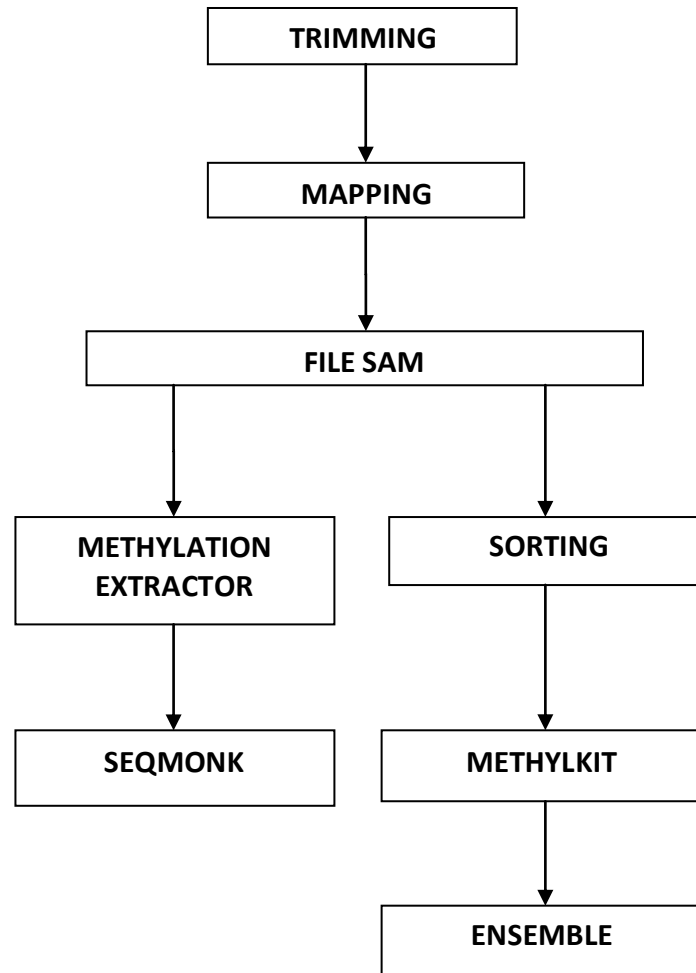


Figura 6 Flow chart dell'analisi bioinformatica

3.2.9 Methylation extractor (Bismark)

L'estrazione dei dati di metilazione dei campioni è stata eseguita grazie al tool *Methylation extractor* di *Bismark*. Questa prima analisi di metilazione è stata condotta inserendo due parametri utili per l'analisi di librerie *paired-end*:

-*no overlap*: elimina le informazioni ridondanti sulle citosine quando si sovrappongono più *reads* a livello di una data regione genomica

-*comprehensive*: unisce le informazioni che derivano dai quattro filamenti considerati nell'allineamento e le suddivide in tre *file* a seconda del contesto in cui si trovano (CpG, CHG e CHH)

Inoltre *Methylation extractor* produce, nel caso di librerie *paired-end*, due grafici denominati *M-bias plot* che mostrano le citosine, in percentuale e in numero assoluto,

per ciascuna posizione all'interno delle *reads* considerando anche i tre diversi contesti in cui si trovano.

L'output di *Methylation extractor* può essere direttamente visualizzato con il *genome viewer SeqMonk*. Quest'ultimo è stato sviluppato per supportare l'analisi di dati derivanti dal sequenziamento *high-throughput* e in particolare per visualizzare graficamente le *reads* sequenziate contro il genoma di riferimento annotato. Nel presente studio *SeqMonk* ha permesso di effettuare una prima valutazione qualitativa sull'omogeneità dei dati tra i campioni in oggetto e di analizzare, in fase conclusiva, regioni genomiche di interesse grazie ad informazioni sull'annotazione.

3.2.10 Methykit

Le informazioni contenute nei file sono state quindi "ordinate" per cromosoma e per posizione delle *reads* secondo un procedimento denominato *sorting*, con l'obiettivo di razionalizzare l'analisi dei profili epigenetici grazie al programma *Methykit* che gira nell'ambiente di analisi R (www.r-project.org). In particolare *Methykit* processa l'allineamento per ottenere la percentuale di metilazione all'interno delle isole CpG con il relativo *coverage* che indica quante *reads* mappano in una data posizione. I singoli campioni filtrati in base al *coverage* scelto sono stati distinti a seconda della variabile biologica d'interesse (trattamento) in controlli e trattati per eseguire una prima analisi per singola posizione, seguita da un'analisi per regioni genomiche d'interesse. *Methykit* calcola inoltre il coefficiente di correlazione tra i diversi campioni ed esegue un'analisi per componenti principali (*PCA-Principal Component Analysis*) che riduce il numero di variabili che incidono sui singoli in gruppi (componenti) per eliminare la ridondanza.

I risultati sono stati estratti impostando idonei valori di *q-value* e definendo la percentuale soglia affinché due posizioni/regioni siano considerate differenzialmente metilate.

Nell'indagine per regioni si considerano anche i parametri riguardanti la lunghezza della regione da analizzare (*win size*) e l'ampiezza dell'intervallo per interrogare la successiva (*step size*). L'annotazione è stata eseguita su *Ensembl Genome Browser* (www.ensembl.org) per conoscere la localizzazione e la funzione di regioni genetiche specifiche.

3.3 Metodica 2b-RAD

Per la produzione di librerie 2b-RAD sono stati nuovamente estratti e processati gli stessi otto campioni utilizzati per la costruzione, attraverso arricchimento con *MBD proteins* e trattamento con bisulfito, delle librerie genomiche precedenti. In particolare per ogni estratto sono state prelevate due aliquote: 2 μL sono stati destinati alla *Whole Genome Amplification (WGA)* e 4 μL alla metodica 2b-RAD. Le librerie WGA sono state costruite per ottenere il *reference* interno nella procedura 2b-RAD. In figura 7 è riportato in maniera sintetica il procedimento seguito.

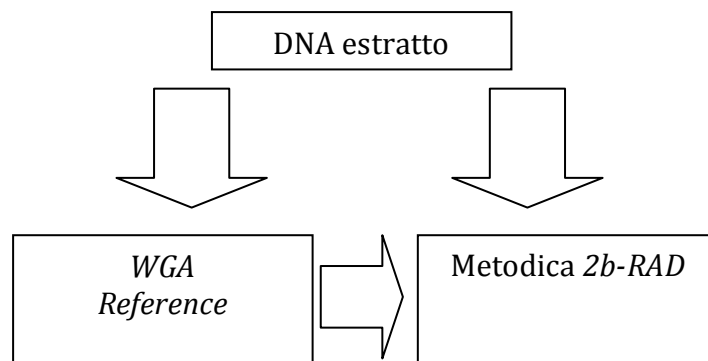


Figura 7 *Flow chart* dell'analisi di metilazione

3.3.1 Estrazione e quantificazione del DNA genomico

In sette dei campioni in oggetto era presente una quantità di tessuto muscolare di 20-25 mg, sufficiente per partire con la tecnica di estrazione scelta. In particolare è stato utilizzato un protocollo messo a punto dal gruppo di ricerca e basato sulla precipitazione del DNA genomico indotta da alte concentrazioni di sali. Prima di procedere con l'estrazione, il tessuto è stato pesato e omogeneizzato meccanicamente con delle forbicine all'interno della provetta. Per la lisi cellulare sono stati aggiunti 300 μL di SSTNE (SSTNE, 1 litro: 17.532 g NaCl, 6.055 g Tris, 1 mL EDTA 0.2 M, 76.08 mg EGTA, 1 mL spermidina 363 mg/5 mL acqua, 1 mL spermina 261 mg/5 mL acqua), 30 μL di SDS (sodio dodecilsolfato) disciolto a 40 °C e 10 μL di proteinasi K concentrata 10 mg/mL. Successivamente i campioni sono stati incubati a 55 °C per 3 ore e a 70°C per 10 minuti prima di aggiungere 20 μL dell'enzima *RNasi A* (10 mg/mL) e incubare nuovamente il campione per 30 minuti a 37 °C. La precipitazione del lisato è stata effettuata con l'aggiunta di 240 μL di cloruro di sodio 5M e seguita, dopo aver posto i campioni in ghiaccio per 10 minuti, da una centrifugazione a 4°C di 15 minuti a 13000 rpm. Sono stati quindi conservati e trasferiti in nuove provette 500 μL del surnatante,

dov'è presente il DNA in soluzione, ai quali sono stati addizionati 500 μ L di isopropanolo per concentrare il DNA incubando i campioni a -20 °C overnight. La precipitazione è avvenuta mediante una centrifugazione di 15 minuti a 13000 rpm. Il pellet di DNA formatosi è stato lavato con 800 μ L di etanolo concentrato al 70% che ha agito per 2 ore a temperatura ambiente per rimuovere i sali co-precipitati. I campioni sono stati nuovamente centrifugati ed il surnatante è stato eliminato. Dopo un periodo di essiccazione di 5-6 minuti a 50 °C, il pellet è stato risospeso in 50 μ L di *elution buffer*. La quantificazione del DNA genomico estratto è stata effettuata mediante fluorimetro QUBIT, utilizzando il saggio *dsDNA BR Assay*, per determinare la concentrazione ng/ μ L, mentre per il calcolo dei rapporti di assorbanza A_{260}/A_{280} e A_{260}/A_{230} è stato utilizzato lo spettrofotometro *μ Drop™ (Thermo scientific)*. Quest'ultimo ha lo stesso principio di funzionamento di un classico spettrofotometro, ma ha la peculiarità di analizzare in modo semplice e preciso piccole quantità di DNA in soluzione ($< 2\mu$ L), a diretto contatto con il sistema ottico, di più campioni contemporaneamente. In particolare la *μ Drop™ Plate* possiede un'area dedicata costituita da 16 posizioni, ordinate in due file da otto, che permette di porre direttamente piccoli volumi di 16 estratti. La *μ Drop™ Plate* così allestita è stata inserita nel lettore *Multiskan®GO* che ha restituito il risultato dell'analisi attraverso lo *Skant® Software*.

Prima della misurazione dei campioni, la *μ Drop™ Plate* è stata preparata con 16 replicati di *elution buffer* per determinare il background, ovvero la media dei valori di "bianco" da sottrarre alle assorbanze dei campioni, determinate dal software, per poi calcolare i rapporti.

I campioni sono stati valutati anche qualitativamente grazie ad un'elettroforesi su gel di agarosio concentrato allo 0,8% con un marcatore molecolare *100 pb DNA ladder*.

Dopo l'analisi dei dati relativi alla quantificazione, per i campioni con rapporti di assorbanza non ottimali è stata eseguita una purificazione dell'estratto con colonnine *Amicon® Ultra-0.5 mL Centrifugal Filters Devices (Millipore)* adattando il protocollo al caso specifico. In particolare l'estratto in oggetto è stato portato ad un volume di 500 μ L con acqua *DNase-free*, trasferito nella colonnina inserita in un tubino, e centrifugato per 5 minuti a 12200 rpm. Successivamente la colonnina con il DNA legato è stata separata dal tubino contenete il filtrato e inserita capovolta in un nuovo tubino per ottenere, mediante una breve centrifugazione a 3250 rpm, circa 28 μ L di estratto concentrato 18 volte.

3.3.2 Creazione delle librerie WGA

Per ottenere il *reference* interno ad ogni campione è stata prelevata una quantità minima di estratto (2 µL) grazie alla quale, in seguito ad opportune diluizioni in *elution buffer*, sono stati raggiunti i 20 ng in 2.5 µL richiesti dal kit commerciale *REPLI-g® Mini Kit (Qiagen)* per la *Whole Genome Amplification (WGA)* basata sulla tecnologia innovativa *Multiple Displacement Amplification (MDA)*. La resa costante per 50 µL di reazione è di 10 µg, mentre la lunghezza media dei prodotti di amplificazione è superiore a 10 Kb, compresa tra 2 e 100 Kb offrendo la possibilità di procedere con molteplici analisi genomiche come *SNP (Single Nucleotide Polymorphism) genotyping*, analisi mediante *SSR (Short Simple Repeat)*, impiego di enzimi di restrizione come nel presente studio (metodica *2b-RAD*) e *NGS*.

La *WGA* permette un'amplificazione uniforme dell'intero genoma, minimamente soggetta a errori grazie all'elevata processività della DNA polimerasi fagica *Phi29* in grado di sintetizzare frammenti di 70 Kb senza dissociarsi dal template (Dean et al., 2002) garantendo allo stesso tempo massima fedeltà di replicazione con l'attività di *proofreading* esonucleasica 3'-5'. Per impedire che quest'ultima possa degradare anche i *primers* esamerici presenti nel kit, le loro estremità 3' sono opportunamente modificate in 3' tiofosfato.

L'amplificazione basata su tecnologia *MDA* inizia con una fase di denaturazione alcalina (2.5 µL di buffer D1) che permette di ottenere un template a singolo filamento uniforme e poco frammentato rispetto alla classica denaturazione ad elevata temperatura (Dean et al., 2002). La reazione si realizza a temperatura ambiente in 3 minuti e viene interrotta dall'aggiunta di 2.5 µL di buffer di neutralizzazione N1. Per la fase di amplificazione vengono aggiunti 40 µL di Master Mix (acqua *DNase-free*, *REPLI-g Mini DNA Polymerase* e *REPLI-g Mini Reaction Buffer* che le permette di dissolvere eventuali strutture secondarie prevenendo così *slippage*, interruzioni e dissociazioni) ai 10 µL di DNA denaturato, il quale si replica durante un'incubazione a 30 °C overnight (10-16 ore). In particolare questa amplificazione isotermica inizia con il legame dei *primers* esamerici in siti multipli del DNA *single strand* e prosegue con la sintesi del filamento complementare operata dalla DNA polimerasi *Phi29* che procede lungo il template sollevando in più punti il filamento neo sintetizzato (tecnologia *MDA*). I filamenti spostati rappresentano lo stampo per nuovi eventi di replicazione (*primers*

esamerici e aggancio della DNA polimerasi *Phi29*) creando una struttura ramificata per l'amplificazione esponenziale di DNA ad elevato peso molecolare. Infine la replicazione viene interrotta inattivando la DNA polimerasi a 65 °C per 3 minuti.

Il controllo qualità del prodotto di amplificazione è stato eseguito con un'elettroforesi su gel di agarosio 0.8% con un marcatore di peso molecolare *100 pb DNA ladder*, mentre la quantità è stata misurata al QUBIT.

3.3.3 Creazione delle librerie *2b-RAD*

Il *2b-RAD* (*Restriction site-associated DNA*) è una metodica molecolare basata sul sequenziamento di frammenti genomici uniformi prodotti da enzimi di restrizione di tipo IIB in grado di tagliare il dsDNA in corrispondenza di sequenze specifiche. La ridotta complessità del genoma che ne deriva e la semplicità di esecuzione lo rendono particolarmente adatto per studi di *genotyping*, mappaggio di *QTL* (*Quantitative Trait Loci*) e filogenesi, specialmente in organismi con genoma non sequenziato o di grande dimensione (Wang et al., 2012).

In particolare nel presente studio il *2b-RAD* è stato utilizzato per la caratterizzazione dei profili di metilazione grazie all'utilizzo dell'endonucleasi metilazione sensibile *BcgI* (*NEB*), capace di tagliare il DNA a monte e a valle del suo sito di riconoscimento generando frammenti di 34 bp con 2 estremità *protruding* al 3' di ogni filamento. Tuttavia essendo metilazione sensibile non taglia se sono presenti citosine metilate. In figura 8 è riportato il sito di taglio di *BcgI*.

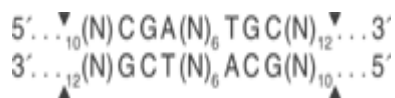


Figura 8 Frammenti generati da *BcgI*

Il protocollo *2b-RAD* consiste in una serie di fasi che portano alla costruzione di librerie genomiche normalizzate di frammenti ristretti, ligati e amplificati. La quantità di partenza di dsDNA è di 4 µL ad una concentrazione compresa tra i 200 e i 1000 ng. A questi, per la fase di digestione, sono stati aggiunti 2 µL di mix costituita da *buffer CutSmart* (*NEB*) che prepara le condizioni di lavoro per l'enzima di restrizione *BcgI*, *SAM* (*S-adenosil-methionine*) che facilita l'incontro tra il DNA e l'enzima, e acqua

DNase-free. La reazione è stata condotta in un termociclatore (*PTC-200 Peltier Thermal Cycler*) in cui è stato impostato il seguente ciclo:

37 °C per un'ora

65 °C per 20 minuti

I frammenti prodotti sono stati quindi ligati ad adattatori standard aggiungendo 20 µL di mix ai 6 µL di digerito. Nella mix sono presenti, oltre ai due adattatori, acqua *DNase-free*, *T4 ligase buffer*, *T4 ligase* (Cabru) e ATP (NEB) che permette all'enzima ligasi di catalizzare la reazione. La ligazione è stata eseguita inserendo nel termociclatore il seguente ciclo:

3 ore a 16 °C (attivazione enzima)

10 minuti a 65 °C (disattivazione enzima)

A questo segue l'amplificazione dei frammenti ristretti e ligati in modo corretto, resa possibile grazie a due coppie di *primers*: la prima è costituita da un *primer 2b-RAD forward* e un *primer 2b-RAD reverse barcode* con una parte della sequenza che riconosce l'adattatore e una parte che identifica univocamente la libreria; mentre la seconda è caratterizzata da due *primer 2b-RAD* che amplificano il costrutto prodotto in precedenza. Dal punto di vista operativo prima vengono dispensati 2.5 µL di *primer 2b-RAD reverse barcode* per ogni libreria, poi si aggiungono 12 µL di dsDNA ligato e infine 35.5 µL di mix costituita dai rimanenti *primers*, acqua *DNase-free*, *Phusion high-fidelity reaction buffer* (NEB), *dNTPs* e *Phusion high-fidelity DNA polymerase* (NEB). La reazione si svolge in 12 cicli dove ogni ciclo è caratterizzato come segue:

98 °C per 5 secondi

60 °C per 20 secondi

72 °C per 5 secondi

Sia gli adattatori utilizzati in fase di ligazione che i *primers* per l'amplificazione sono sintetizzati appositamente per il sequenziamento Illumina HiSeq (Wang et al., 2012). Infine l'amplificato è stato valutato qualitativamente con un'elettroforesi su gel di agarosio 1.8% utilizzando un marcatore di peso molecolare *100 pb DNA ladder*.

3.3.4 Purificazione e sequenziamento delle librerie dsDNA

Il prodotto di amplificazione è stato purificato eseguendo le operazioni di un protocollo messo a punto dal gruppo di ricerca e basato sull'impiego di biglie magnetiche *SPRIselect*[®] *Beckman Coulter*. In particolare la metodica sfrutta la capacità delle biglie (presenti 1.7 volte in più rispetto al volume dell'amplificato) di legare selettivamente i frammenti dsDNA di interesse. Successivamente le strip vengono poste in una piastra magnetica che attira in parete le biglie in emulsione lasciando al centro il surnatante costituito dai reagenti utilizzati in fase di amplificazione (buffer, dNTPs e *primer* non legati). Inoltre il lavaggio con etanolo 85% (180 µL/pozzetto) ha permesso di purificare ulteriormente i frammenti dsDNA eliminando i possibili contaminanti rimasti. Tolte le strip dal magnete, sono stati addizionati 25 µL di *acqua DNase-free* per pozzetto in modo che le biglie rilasciassero in sospensione i frammenti grazie all'utilizzo di un classico agitatore. Infine le strip sono state poste nuovamente nella piastra magnetica per prelevare soltanto il prodotto di amplificazione purificato.

Le singole librerie *2b-RAD* purificate sono state quantificate al QUBIT per calcolare la quantità di campione da inserire all'interno del pool, e successivamente al *Bioanalyzer* con il corrispondente chip *Agilent High Sensitivity DNA* per visualizzare la distribuzione in pb dei frammenti presenti. Anche il pool finale di 15 µL è stato valutato quantitativamente al QUBIT prima di essere sequenziato con tecnologia Illumina *high throughput*.

4. RISULTATI E DISCUSSIONE

4.1 Librerie genomiche arricchite: risultati preliminari

Il DNA estratto dal muscolo *longissimus dorsi* di bovini appartenenti al gruppo di trattamento scelto e al gruppo di controllo è stato quantificato e valutato qualitativamente. In tabella 2 sono riportate le concentrazioni ng/ μ L e i rapporti A_{260}/A_{280} e A_{260}/A_{230} , mentre la figura 9 mostra la qualità rivelata con elettroforesi su gel di agarosio. Al fine di ottimizzare la costruzione delle librerie, sono stati selezionati, tra gli estratti iniziali, gli 8 estratti quantitativamente e qualitativamente migliori.

Tabella 2 Concentrazioni e rapporti

Campione LD	Concentrazione ng/ μ L	A_{260}/A_{280}	A_{260}/A_{230}
Controlli			
BT03	58.6	1.85	1.60
BT04	70.6	1.85	1.95
BT05	44.8	1.86	1.88
BT06	42	1.81	1.16
Trattati			
BT27	62.4	1.90	1.98
BT28	46.8	2.00	1.89
BT29	36.7	1.70	1.93
BT32	41.4	1.97	1.60

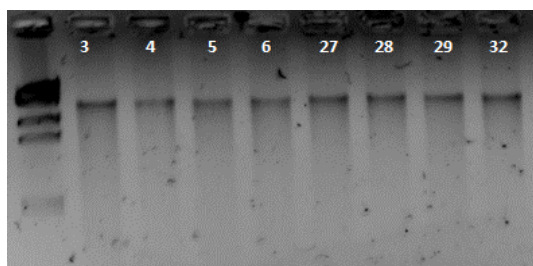


Figura 9 Elettroforesi su gel d'agarosio 1.8 %

4.2 Analisi bioinformatiche

4.2.1 Controllo qualità e *trimming*

Con il sequenziamento Illumina *High throughput* per ogni libreria genomica sono state ottenute circa 45 milioni di *reads*. I dati che derivano dal sequenziamento spesso sono soggetti ad errori nella chiamata delle basi che possono influenzare negativamente l'analisi bioinformatica successiva. In particolare il tasso di errore osservato per Illumina *HiSeq* 2000 è dello 0,26 % (Quail et al., 2012). La scorretta assegnazione delle basi verso l'estremità 3' delle *reads* in corrispondenza degli adattatori può favorire selettivamente alcuni nucleotidi rispetto ad altri (Hansen et al., 2012). Per questo motivo prima di procedere con l'allineamento i data set di ogni campione sono stati filtrati secondo i parametri qualitativi del *trimming*.

L'effetto del *trimming* in base al *Phred quality score*, il cui valore soglia è stato impostato a 20, risulta simile in tutti i dataset e un esempio grafico è riportato in figura 10. Le sequenze che si trovano nella "zona verde" dopo il *trimming* sono quelle idealmente sequenziate con un'accuratezza superiore al 99 %, che vengono scelte per le successive analisi.

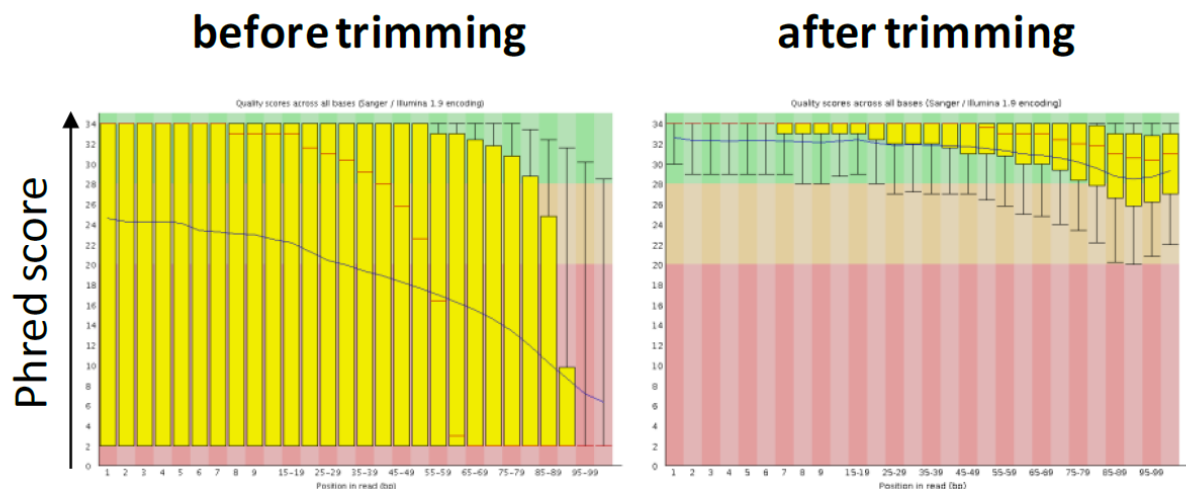


Figura 10 Effetto del *Phred quality score*

La figura 11 riporta un grafico descrittivo della distribuzione delle sequenze in *Phred quality score*. Il dataset "grezzo" contiene numerose sequenze con *Phred* molto basso (1-3), ma dopo la loro rimozione la maggior parte delle sequenze ha un *Phred* elevato (30-34).

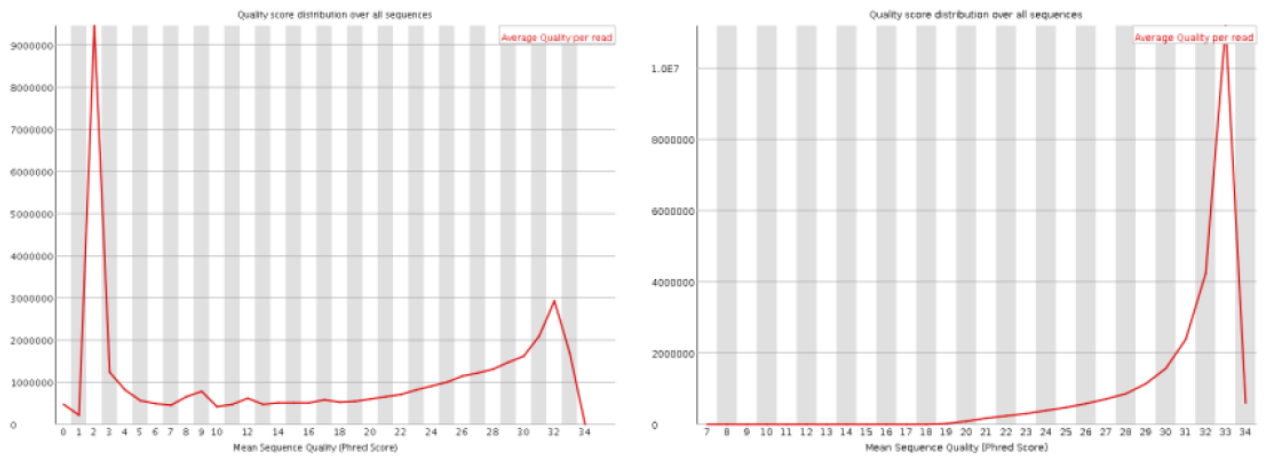


Figura 11 Dati quantitativi di *trimming* sulla base del *Phred quality score*

La figura 12 mostra come la presenza degli adattatori sbilanci il contenuto nucleotidico in citosine verso la terminazione 5' delle *reads*. La rimozione degli adattatori stabilizza invece il contenuto di citosine all'1 % per tutta la lunghezza della sequenza.

before trimming

after trimming

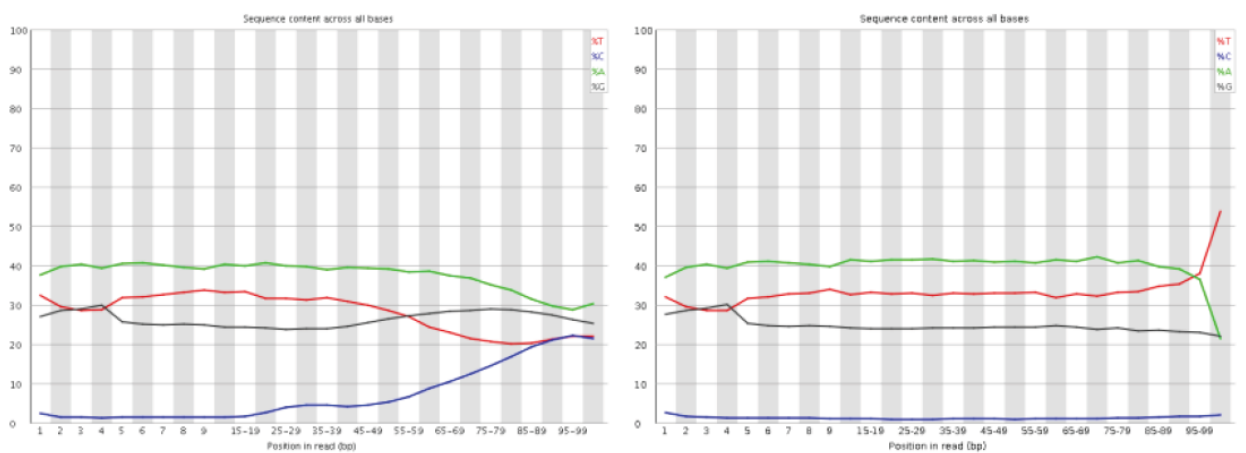


Figura 12 Effetto della rimozione degli adattatori

4.2.2 Mapping

Le sequenze filtrate sono state allineate contro il genoma di riferimento (*Bos taurus* UMD3.1) convertito in silico con bisulfito con l'utilizzo del programma *Bismark*. L'efficienza del *mapping* è del 40-50% poiché le sequenze che mappano in modo univoco sono comprese tra i 6 e i 15 milioni. Generalmente l'efficienza del *mapping* è più bassa per le specie di interesse zootecnico rispetto all'uomo o al topo, come rilevato da Doherty and Couldrey (2014): gli autori riportano infatti valori pari al 42.2%

di efficienza di *mapping* per campioni di longissimus dorsi di ovino in un esperimento di *whole-genome bisulfite sequencing*. Tali valori di efficienza di mapping potrebbero essere dovuti a un incompleto o non preciso assemblaggio del *reference genome* di queste specie rispetto a specie ampiamente più studiate come l'uomo e il topo. Inoltre la presenza di ampie regioni ripetute all'interno del genoma potrebbe spiegare la bassa percentuale di *reads* che mappano in modo univoco. Le *reads* non allineate o che presentano allineamento multiplo non sono state incluse nell'analisi riducendo ulteriormente di 30-40 milioni di unità il numero di sequenze iniziali. Bismark offre anche statistiche relative al numero e allo stato di metilazione delle citosine. Anche il numero di citosine presente sul totale di *reads* analizzate varia ampiamente tra i diversi campioni: esso è compreso tra 300 e 800 milioni circa, in linea con la diversa abbondanza di *reads* presenti per campione. La percentuale di citosine metilate all'interno dei siti maggiormente coinvolti dalla metilazione, ovvero i siti CpG, è compresa invece in un breve intervallo tra 83 e 85% mostrando omogeneità tra i campioni. Quest'ultimo dato è indice di validità e solidità della metodica sperimentale utilizzata, perché la percentuale di siti CpG metilati corrisponde a quella mediamente riscontrata nei vertebrati (Salozhin et al., 2005). I dati relativi all'output di *Bismark* sono riportati in Tabella 3.

Tabella 3 Statistiche descrittive di *mapping* per campione

	Numero di sequenze con unico allineamento	Efficienza di mapping (% di reads mappano univocamente)	Numero di citosine analizzate	% citosine metilate in CpG
BT03	13.050.662	48.9 %	633.709.793	84.1 %
BT04	7.091.927	43.1 %	341.096.149	84.6 %
BT05	10.595.633	50.4 %	533.888.353	84.2 %
BT06	8.155.419	47.6 %	399.810.619	84.6 %
BT27	8.135.604	42.3 %	396.055.171	85.1 %
BT28	9.583.572	41.8 %	487.240.122	82.8 %
BT29	6.238.461	40.2 %	302.599.768	85.2 %
BT32	15.490.218	46.0 %	776.726.061	83.7 %

4.2.3 Methylation extractor

Grazie a *Methylation extractor*, strumento analitico di *Bismark*, sono state semplificate e razionalizzate le informazioni, derivanti dall'allineamento, relative alla numerosità e allo stato di metilazione di tutte le citosine per quanto riguarda il contesto CpG. In particolare, rispetto alle statistiche fornite da *Bismark*, i valori sono leggermente più bassi, ma sostanzialmente sovrapponibili perché l'analisi di questo software ha un metodo di rilevazione più mirato e preciso nei confronti delle molteplici informazioni che provengono dalle diverse *reads*; nelle *reads* viene osservata la presenza di timina o di citosina e contando quelle contenenti quest'ultima si può stabilire se ogni singola citosina presente sul genoma sia metilata (*reads* con citosina presenti in numero maggiore) o meno. Complessivamente le citosine analizzate variano tra i 266 e i 740 milioni, mentre la percentuale di metilazione relativa alle citosine nei siti CpG è compresa tra 82.5 e 84.8%, come si può rilevare dalla tabella 4.

Tabella 4 Statistiche descrittive di *Methylation extractor* per campione

STATISTICHE DESCRITTIVE PER CAMPIONE - METHYLATION EXTRACTOR			
	Numero di citosine analizzate	Totale citosine metilate in Cpg	% citosine metilate in Cpg
BT03	608.249.272	97.888.972	83.9 %
BT04	308.101.160	53.237.278	84.2 %
BT05	519.170.424	80.718.751	84.1 %
BT06	366.120.024	62.735.774	84.3 %
BT27	352.876.022	64.047.397	84.7 %
BT28	460.947.716	79.622.732	82.5 %
BT29	266.461.603	49.807.122	84.8 %
BT32	740.868.349	125.456.183	83.5 %

4.2.4 Statistiche descrittive di *Methylkit*

I dati di ogni campione provenienti dall'allineamento sono stati strutturati da *Methylkit* in due diversi istogrammi per visualizzare la percentuale di singole citosine metilate nei contesti CpG e il *coverage* per singoli siti CpG. In particolare in Figura 13 sono riportati gli istogrammi delle sequenze *forward* e *reverse* di uno dei campioni analizzati. Il grafico mostra le frequenze, tradotte poi nelle corrispondenti percentuali, delle isole CpG che presentano un dato livello di metilazione. La percentuale di regioni CpG che

ha il 100% di metilazione, ovvero tutti i siti CpG che mappano in quel punto sono metilati, è del 65% nei filamenti *forward* e del 65.1% considerando i *reverse*. Considerando tutti i campioni la presenza di regioni CpG altamente metilate è compresa tra il 63 e il 67%, sempre con una bassa, talvolta irrisoria, proporzione di emimetilate o non metilate.

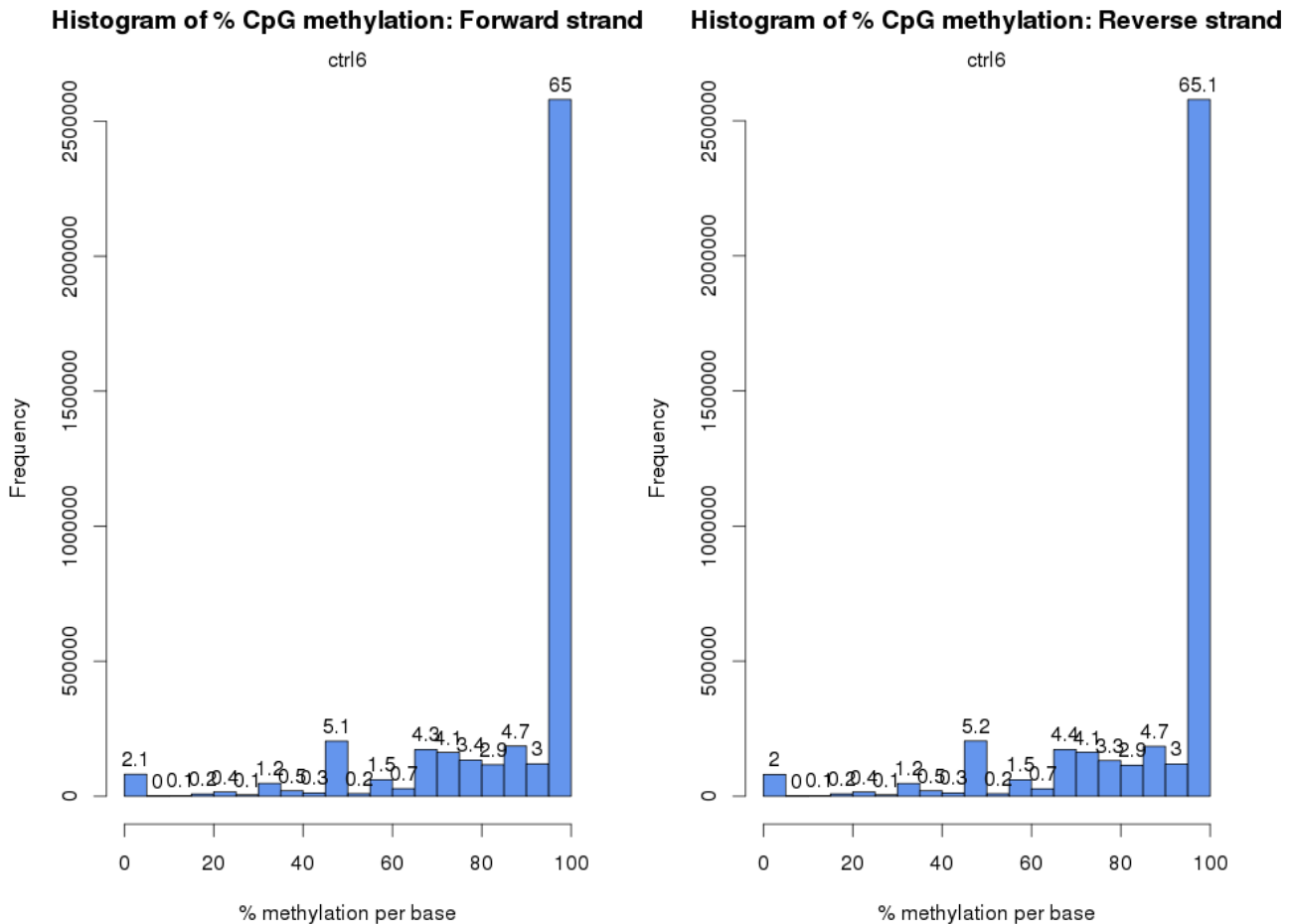


Figura 13 Percentuale di metilazione all'interno delle isole CpG

Generalmente questo istogramma dovrebbe presentare una distribuzione bimodale con un secondo picco in corrispondenza di basse percentuali di metilazione (Akalin et al., 2012) suggerendo la presenza nel genoma di due stati alternativi presenti in ugual misura: silenziamento (elevate percentuali di metilazione) e potenziale attività trascrizionale relativa a regioni scarsamente metilate (Suzuki and Bird, 2008). Tuttavia, la tecnica di arricchimento selettivo con MBD, che recupera in maggior misura la frazione metilata del DNA, giustifica la distribuzione unimodale del grafico ottenuto.

Anche l'informazione sul *coverage* è riportata in un istogramma (Figura 14). La distribuzione del *coverage* rappresenta un altro indicatore di qualità importante

perché permette di capire se l'esperimento è influenzato o meno da errori in fase di amplificazione che potrebbero avvantaggiare alcune *reads* rispetto ad altre (Akalin et al., 2012) indebolendo così la valenza del risultato finale. Se questo si fosse verificato, ci sarebbe stato un secondo picco sulla destra del grafico in corrispondenza di un valore di *coverage* elevato. Per correggere eventuali *bias* legati alla PCR è possibile filtrare il campione in base al *coverage*, opportunamente scelto, eliminando le posizioni con un numero elevato di *reads* che mappano in quel punto. Inoltre è utile eliminare, come nel presente studio, anche le posizioni con un basso *coverage* per incrementare il potere statistico dell'analisi. Ad esempio scegliendo un *coverage* 10X come cut-off (ovvero scegliendo tutte le singole posizioni sul genoma con almeno 10 *reads* allineate in loro corrispondenza), equivalente a 1 nella scala logaritmica del grafico, si eliminano tutte le *reads* a sinistra, o meglio 4 colonne con *coverage* molto bassi che non permettono di stabilire con accuratezza lo stato di metilazione della citosina considerata.

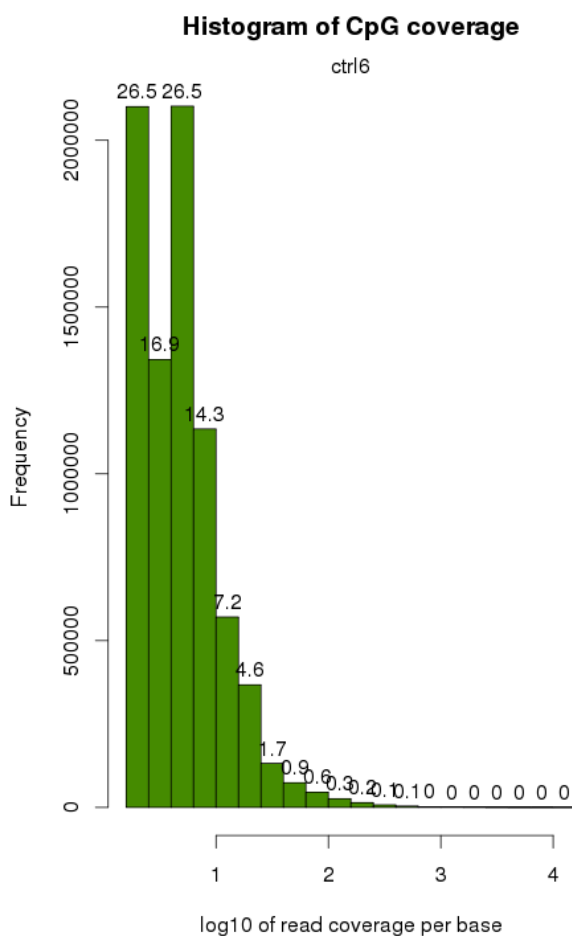


Figura 14 Distribuzione del *coverage* per singoli siti CpG

4.2.5 Correlazione tra i campioni

Il coefficiente di correlazione è stato calcolato a coppie tra le percentuali di metilazione di tutti i campioni in esame (Figura 15). I valori stimati sono piuttosto elevati (77-85%) suggerendo che il profilo di metilazione a livello di singola citosina è molto simile tra trattati e controlli.

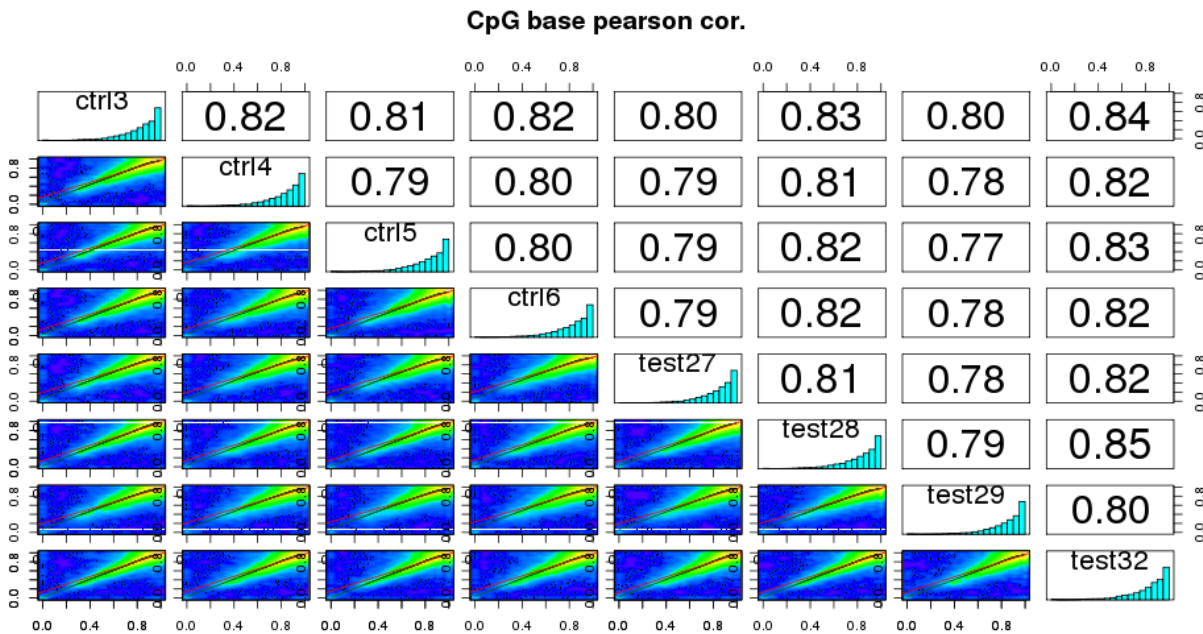


Figura 15 Correlazioni tra i profili di metilazione

4.2.6 Clustering

I campioni sono stati raggruppati in base alla similarità tra profili di metilazione per evidenziare le relazioni epigenetiche esistenti (Figura 16). Nel dendrogramma i campioni non sono suddivisi in due gruppi distinti sulla base del trattamento, ma il campione BT29 include tutti gli altri dimostrando che non vi è una netta differenza tra i campioni di controllo e i trattati e probabilmente, a livello di intero genoma, il trattamento non incide in modo tale da discriminare su questa base i campioni.

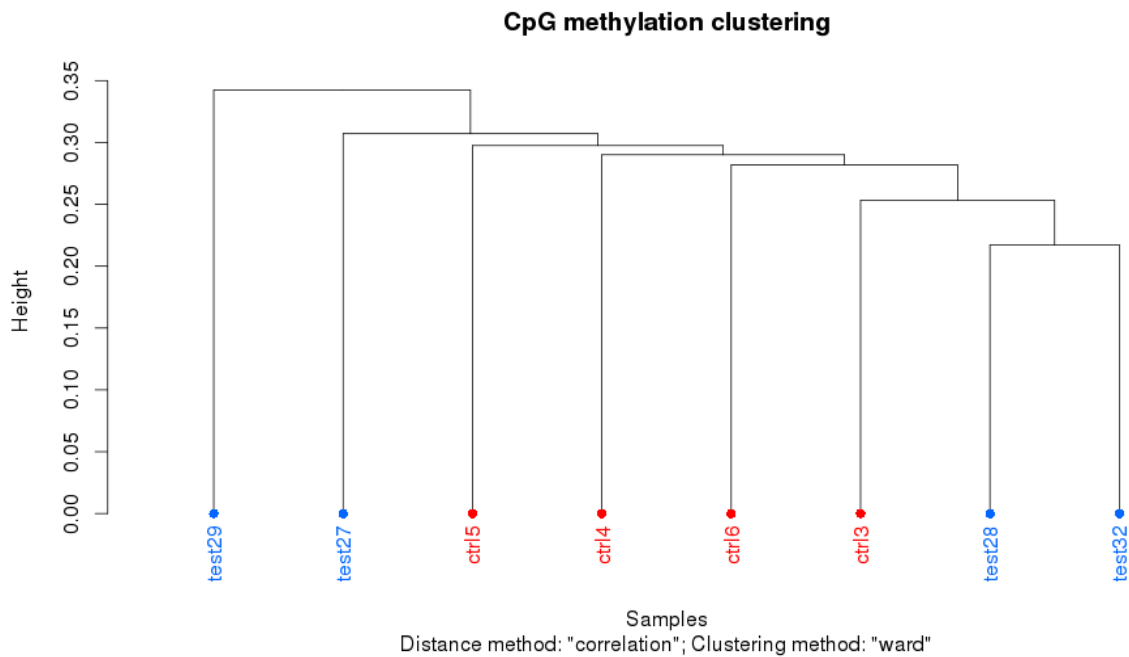


Figura 16 Dendrogramma di similarità

4.2.7 Analisi per componenti principali (PCA)

Per analizzare il profilo di metilazione dei campioni in oggetto è stata condotta anche un'analisi PCA. In Figura 17 è riportato il grafico a dispersione nel quale i campioni sono distinti in base alle due componenti principali (PC1 e PC2). La componente che più incide sulla varianza è in ascissa e in base a questa il campione BT29 risulta distante rispetto a tutti gli altri campioni; mentre per quanto riguarda la componente in ordinata i campioni maggiormente differenziati rispetto agli altri sono BT05 e BT06. Idealmente nel grafico si dovrebbe osservare una separazione netta tra il gruppo dei controlli e il gruppo dei trattati, ma come per le precedenti analisi di correlazione e *clustering* questa distinzione non è apprezzabile. Questo rivela che la variabile biologica in oggetto, il trattamento, non incide in modo più marcato rispetto alle altre variabili nel distinguere il profilo di metilazione tra controlli e trattati.

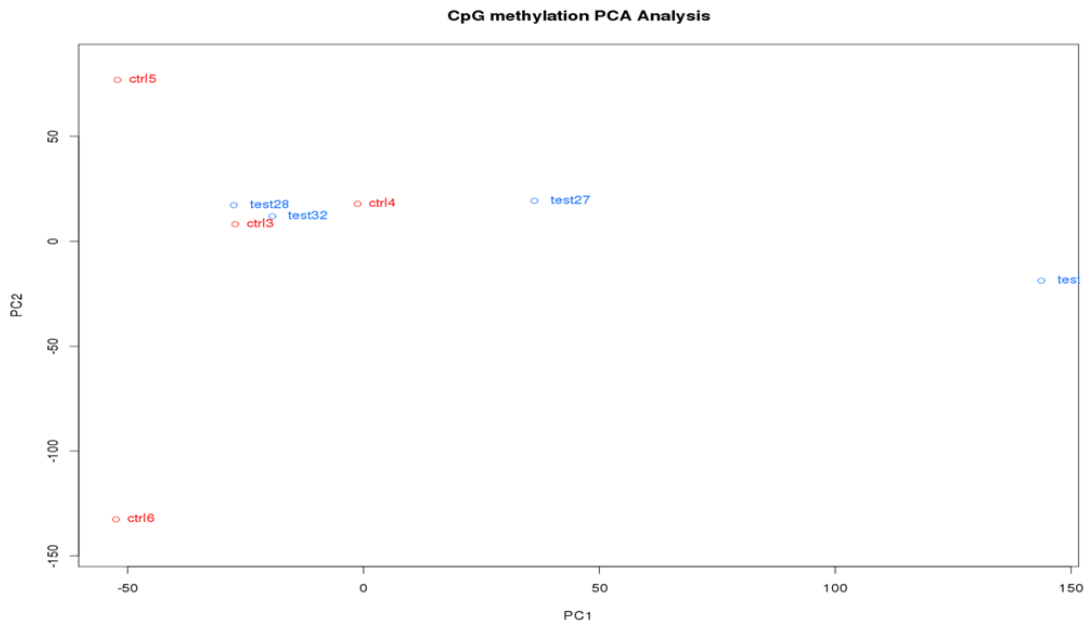


Figura 17 Analisi PCA dei profili di metilazione

4.2.8 Analisi per posizione

La prima tipologia di analisi è stata condotta al fine di rilevare differenze di metilazione a livello di singole posizioni. È stato utilizzato dapprima un approccio molto stringente (*coverage* 10X, differenza di metilazione minima pari al 25% e *q-value* di 0.01) per poi diminuire via via la stringenza nelle analisi successive (ad esempio tramite la diminuzione del *coverage* e l'aumento del *q-value*). Con l'analisi più stringente si ottengono 102 posizioni differenzialmente metilate in 4 cromosomi con una significatività elevata. Le posizioni in oggetto sono ordinate in base al cromosoma in cui si trovano e la differenza di metilazione, ipometilazione (-) o ipermetilazione (+) rispetto al controllo, è riportata nell'ultima colonna (Allegato 1). Le analisi meno stringenti confermano le analisi più conservative anche se, come è logico attendersi, vengono incluse posizioni aggiuntive.

4.2.9 Analisi per regione

Anche in questo caso sono state effettuate analisi più o meno stringenti mediante modificazione dei parametri di *coverage*, *q-value* e *win size/tile window*, questi ultimi caratteristici dell'analisi per regione. L'analisi per regione più stringente ha permesso di individuare 3 regioni cromosomiche differenzialmente metilate in base ad un cut-off del 25% (*q-value*=0.01) interrogando regioni di 100 bp con un intervallo di 100 bp.

4.2.10 Analisi di annotazione per posizione e per regione

Methylkit permette di annotare sul genoma le posizioni e le regioni differenzialmente metilate per conoscere la loro localizzazione su regioni promotoriali, esoni, introni e regioni intergeniche (Tabella 5). Dai risultati emerge che la maggior parte delle posizioni (61.2%) si trova in regioni intergeniche, mentre solo l'8.4% è localizzata in regioni codificanti. Per quanto riguarda le regioni individuate, la metà si trova su sequenze introniche, l'altra metà su regioni intergeniche. Inoltre l'annotazione è stata indirizzata in modo mirato nei confronti dei siti CpG (Tabella 6). Le posizioni differenzialmente metilate si trovano per circa l'80% nelle vicinanze delle isole CpG o in altre regioni genomiche, diverse dal contesto CpG, mentre solo il 16% si localizza in esse. Nessuna delle regioni invece è annotata nei confronti di isole CpG, ma o è nelle vicinanze o in altri contesti del genoma.

Tabella 5 Analisi per posizione e per regione sull'intero genoma

	Promotori (%)	Esoni (%)	Introni (%)	Regioni intergeniche (%)
posizioni	0	8.42	30.33	61.23
regioni	0	0	50	50

Tabella 6 Analisi per posizione e per regione in isole CpG

	Isole CpG (%)	Regioni vicine a CpG (%)	Altro (%)
posizioni	16.29	12.92	70.78
regioni	0	50	50

La prima regione identificata, di lunghezza 100 pb, è localizzata sul cromosoma 1 tra 145375500 e 145375600 bp (Figura 18, area verde).

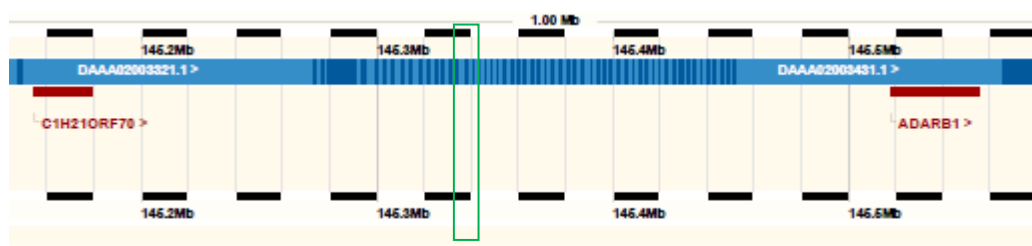


Figura 18 Parte di sequenza del cromosoma 1

Con *Ensembl Genome Browser* è possibile visualizzare che la regione è presente in un'area cromosomica intergenica, tra le regioni codificanti C1H21ORF70 implicata nella biogenesi della proteina ribosomale SLX9 e ADARB1 che codifica per una adenosin deaminasi.

La seconda regione differenzialmente metilata è lunga 100 pb ed è localizzata sul cromosoma 10 tra 104228800 e 104228900 bp (Figura 19, area verde).

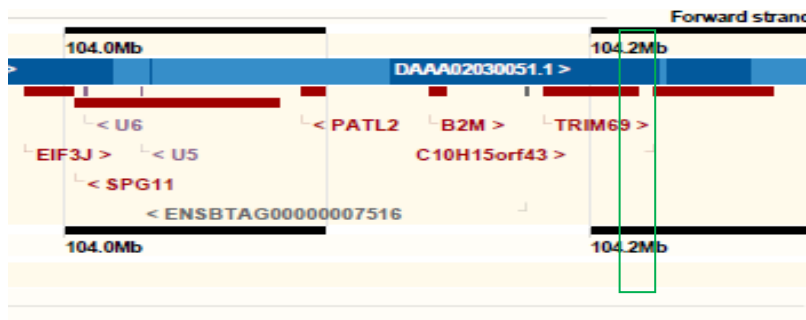


Figura 19 Parte di sequenza del cromosoma 10

La regione è presente all'interno di una regione codificante, tra i geni TRIM 69 gene presente nelle cellule della linea germinale che se overespresso induce apoptosi e C10H15orf43 (non caratterizzato).

Infine il cromosoma 11 contiene la terza regione 100 bp differenzialmente metilata compresa tra 74294900 e 74295000 bp (Figura 20, area verde).

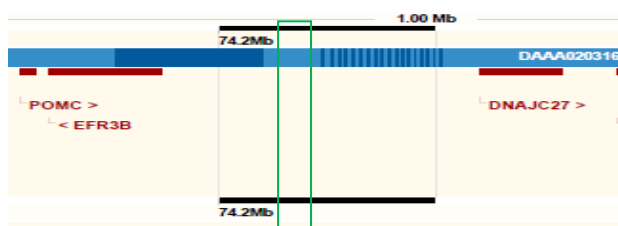


Figura 20 Parte di sequenza del cromosoma 11

Anche questa regione è presente in una regione intergenica, compresa tra le sequenze geniche codificanti di POMC e DNAJC27. In particolare POMC codifica per la pro-opiomelanocortina (POMC), un proormone dal quale, attraverso specifiche modificazioni e maturazioni proteolitiche attuate da peptidasi all'interno di vescicole esocitiche, si originano diversi tipi di ormoni ipofisari: ACTH (Adreno Cortico Tropic Hormone- ormone adrenocorticotropo) e β -lipotropina secreti dall'adenoipofisi e γ -lipotropina, α -MSH (*Melanocyte Stimulating Hormone*), β -MSH e β -endorfina sviluppati nel lobo intermedio. Nello specifico l'ACTH stimola la secrezione di glucocorticoidi, tra i

quali il più importante è il cortisolo, mineralcorticoidi e ormoni sessuali da parte della corticale del surrene. Tra gli steroidi sessuali l'unico ad avere importanza biologica è l'androgeno deidroepiandrosterone (DHEA) secreto da entrambi i sessi. Il gene DNAJC27 è invece implicato nella sindrome dell'iperstimolazione ovarica.

La quasi totalità delle posizioni e delle regioni differenzialmente metilate individuate dal confronto dei profili di metilazione tra controlli e trattati è localizzata in regioni intergeniche e sequenze introniche; solo una piccola parte si trova in corrispondenza di isole CpG. Tuttavia, anche se queste regioni non comprendono siti promotoriali o non si trovano in corrispondenza di sequenze geniche codificanti, esse potrebbero comunque essere implicate nella regolazione dell'espressione genica, perché possono rappresentare fattori di regolazione cis-agenti le cui regioni target possono trovarsi anche molto distanti (Suzuki and Bird, 2008).

La fase successiva all'identificazione di regioni differenzialmente metilate, sulla base dei dati derivanti dal sequenziamento, è rappresentata dalla validazione con metodiche PCR mirate come la *Real Time* PCR che permette simultaneamente di amplificare e quantificare la sequenza genomica in oggetto. Per il presente studio ad oggi è stata proposta la messa a punto di un saggio *Real Time* PCR ad hoc, ovvero metilazione sensibile. In particolare i primers sito-specifici utilizzati devono essere sintetizzati e trattati con bisulfito convertendo i siti non metilati in uracile lasciando inalterate le citosine metilate. Grazie alla presenza nella mix di reazione di dNTPs marcati con fluorofori diversi è possibile rilevare il segnale di fluorescenza per singola base e discriminare in modo preciso tra siti metilati e non metilati.

Negli ultimi anni sono stati pubblicati diversi studi con l'ambizione di caratterizzare l'epigenoma umano o di specie modello (*Mus musculus*, *Danio rerio*) in relazione a specifiche sequenze geniche (Ivanov et al., 2013) o a particolari malattie (Ma et al., 2014) e patologie tumorali (Choi et al., 2010). Tuttavia solo di recente l'analisi del pattern di metilazione del DNA sta acquisendo importanza in specie animali non modello per comprendere l'effetto dell'esposizione a contaminanti ambientali (Head, 2014) e, in animali in produzione zootecnica, per ottenere un ulteriore livello di informazione per il miglioramento genetico di caratteristiche d'interesse (Couldrey et al., 2014). In particolare Couldrey et al. (2014) e Doherty and Couldrey (2014) hanno analizzato il muscolo *longissimus dorsi* di una razza ovina da carne per determinare il

metiloma e capire in che misura il potenziale epigenetico può essere coinvolto nell'ipertrofia muscolare. Per l'analisi bioinformatica in parte sono stati implementati gli stessi software (*Bismark* e *Seqmonk*) con risultati paragonabili soprattutto in termini di efficienza di *mapping* (42%).

Il *whole genome bisulfite sequencing* può tuttavia presentare una serie di limitazioni a seconda dell'ipotesi da testare e del disegno sperimentale (Doherty and Couldrey, 2014) perché per ottenere risultati statisticamente significativi è necessario un ampio numero di replicati biologici. Nella presente sperimentazione per determinare regioni differenzialmente metilate in grado di discriminare significativamente in base al trattamento è quindi necessario, oltre alla validazione delle regioni già individuate, aumentare il numero di campioni per testare la ripetibilità del risultato su scala più ampia includendo anche trattamenti con altri potenziali promotori della crescita per capire se l'analisi può essere anche ripetibile.

4.3 Librerie *2b-RAD*: risultati preliminari

Le concentrazioni ottenute e i rapporti A_{260}/A_{280} e A_{260}/A_{230} degli estratti sono riportati in tabella 7, mentre la valutazione qualitativa è stata eseguita sulla base dell'elettroforesi su gel. Generalmente i valori quantitativi sono migliori rispetto alla precedente sperimentazione grazie alla messa a punto di una metodica classica basata sulla precipitazione del DNA indotta da elevata concentrazione salina.

Tabella 7 Concentrazioni e rapporti

Campione LD	Concentrazione ng/ μ L	A_{260}/A_{280}	A_{260}/A_{230}
Controlli			
3	192	1.89	1.84
4	308	1.76	1.50
5	244	1.89	1.97
6	175	1.71	1.33
Trattati			
27	82.1	1.63	1.03
28	22.33	1.82	1.50
29	49	1.92	0.77
32	111	1.74	1.26

4.3.1 Analisi bioinformatiche

Attualmente è stato eseguito il controllo qualità delle librerie sequenziate grazie all'utilizzo del software FastQC (*Babraham bioinformatics*, UK). Quest'ultimo fornisce un report per ogni libreria (*WGA* e *2b-RAD*) riferita al singolo campione. Il primo dato è relativo al numero di *reads* distinte per ogni campione in base alla sequenza identificativa che le caratterizza (Tabella 8). La differenza nel numero di *reads* tra i diversi campioni è anche di 1-1.5 volte, mentre le *reads* non assegnate corrispondono al 3% del totale. Dai grafici relativi alla qualità per posizione all'interno della sequenza 50 base pair (Figura 21) e per sequenze all'interno della libreria (Figura 22) emerge che il sequenziamento è avvenuto in modo accurato e le diverse *reads* presentano un elevato *Phred quality Score*. L'enzima metilazione sensibile *BcgI* produce frammenti di 36 bp, di conseguenza la parte finale della sequenza corrisponde all'adattatore. Le basi non assegnate (2%) sono presenti all'interno della sequenza solamente nel punto di legame dell'adattatore con il frammento e inoltre nel sito di riconoscimento dell'enzima tra 21 e 23 bp (Figura 23). Altro parametro importante è il contenuto medio di siti CpG all'interno dei frammenti, maggiore rispetto alla situazione teorica proprio per la natura del frammento stesso prodotto dall'enzima *BcgI* (Figura 24).

Tabella 8 Conteggio *reads*, librerie *WGA* (W) e librerie *2b-RAD*

CAMPIONI	READS
Bt3	13884016
Bt3W	9298032
Bt4	11460755
Bt4W	9026569
Bt5	9050724
Bt5W	12352396
Bt6	9656348
Bt6W	11826634
Bt27	7762945
Bt27W	10195887
Bt28	6564808
Bt28W	7679092
Bt29	5399244
Bt29W	12104922
Bt32	15975453
Bt32W	11672101
Unassigned	6090319

Per la prima volta la metodica *2b-RAD* metilazione-sensibile viene applicata per comprendere come l'epigenoma può variare in base al trattamento illecito con promotori della crescita in bovini da carne. Rispetto al *whole-genome bisulfite sequencing* e all'approccio integrato con l'arricchimento con MBD, il *2b-RAD* è in grado di indirizzare l'analisi nei confronti di una frazione genomica ancora più piccola, ma altamente informativa per la capacità dell'enzima Bcgl di produrre frammenti uniformi con siti CpG al loro interno. La relativa semplicità del protocollo e il sequenziamento così mirato rendono il *2b-RAD* una metodica molto promettente nella caratterizzazione e nello studio di alterazioni tra profili di metilazione a confronto. Tuttavia la normalizzazione dei dati per il confronto con il *reference* interno (*WGA*) rappresenta ad oggi il primo obiettivo per l'immediato futuro. Inoltre può essere interessante mappare i frammenti provenienti dal *2b-RAD* sulle precedenti regioni per capire se sono all'interno o meno di esse.

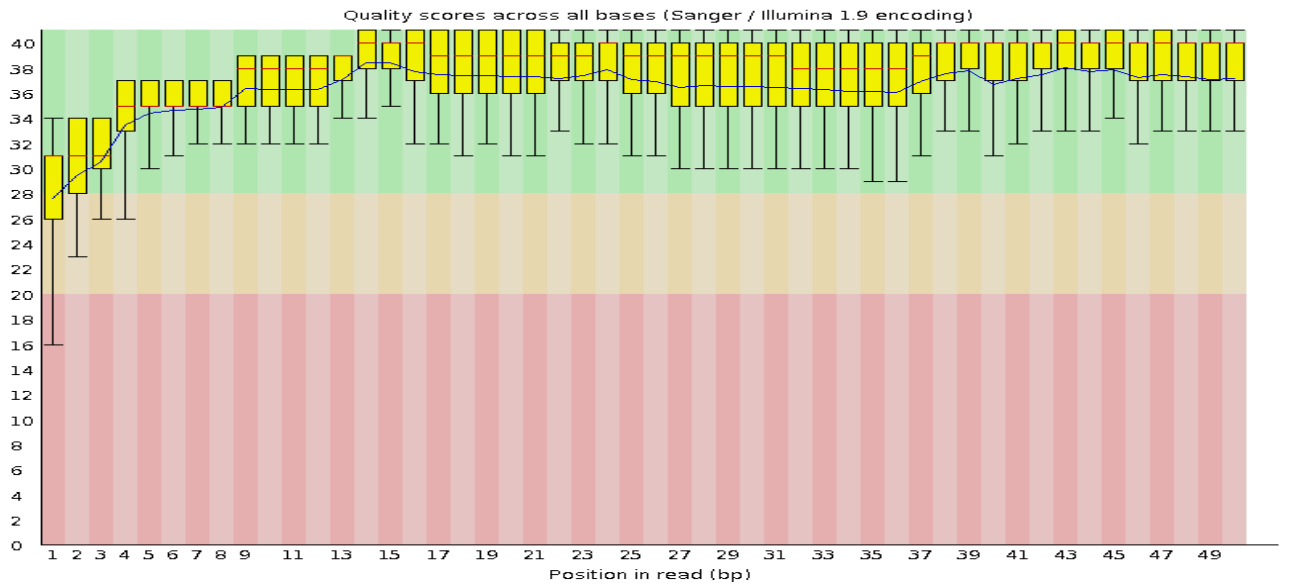


Figura 21 Qualità per posizione

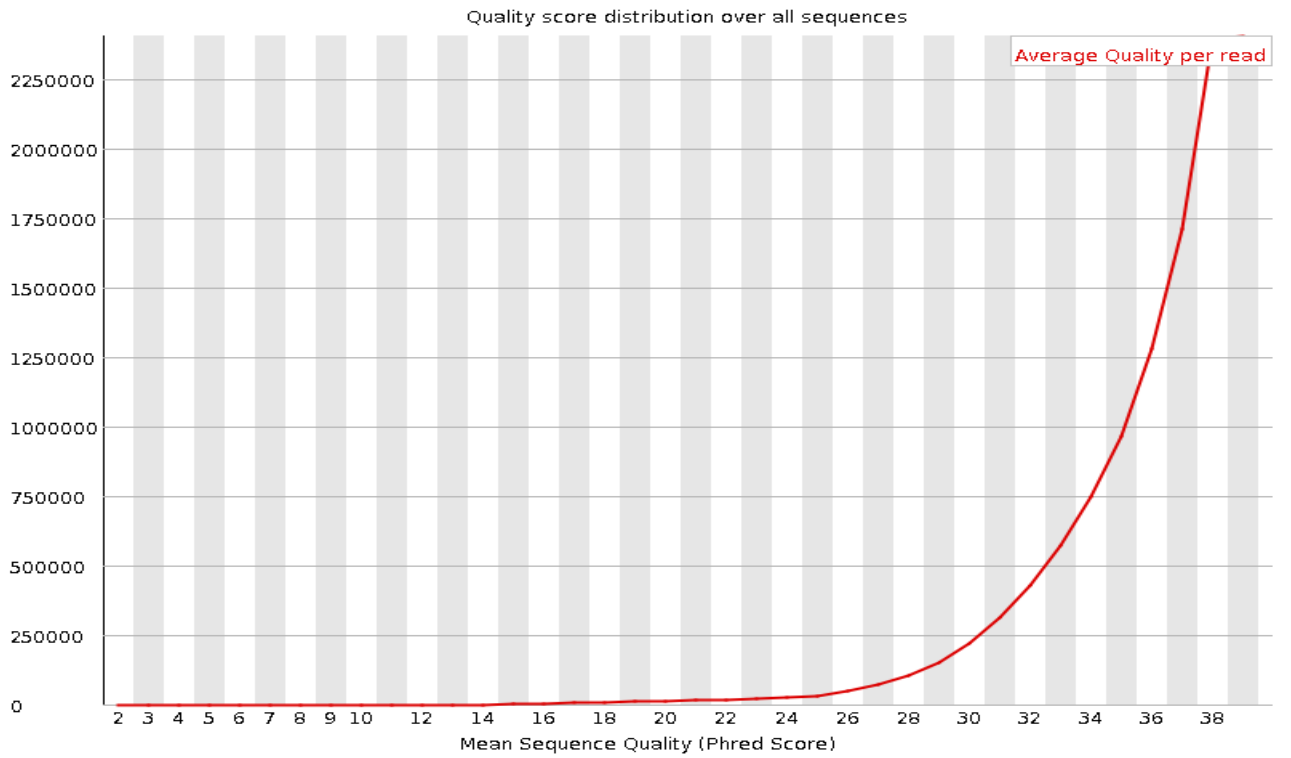


Figura 22 Qualità per sequenza

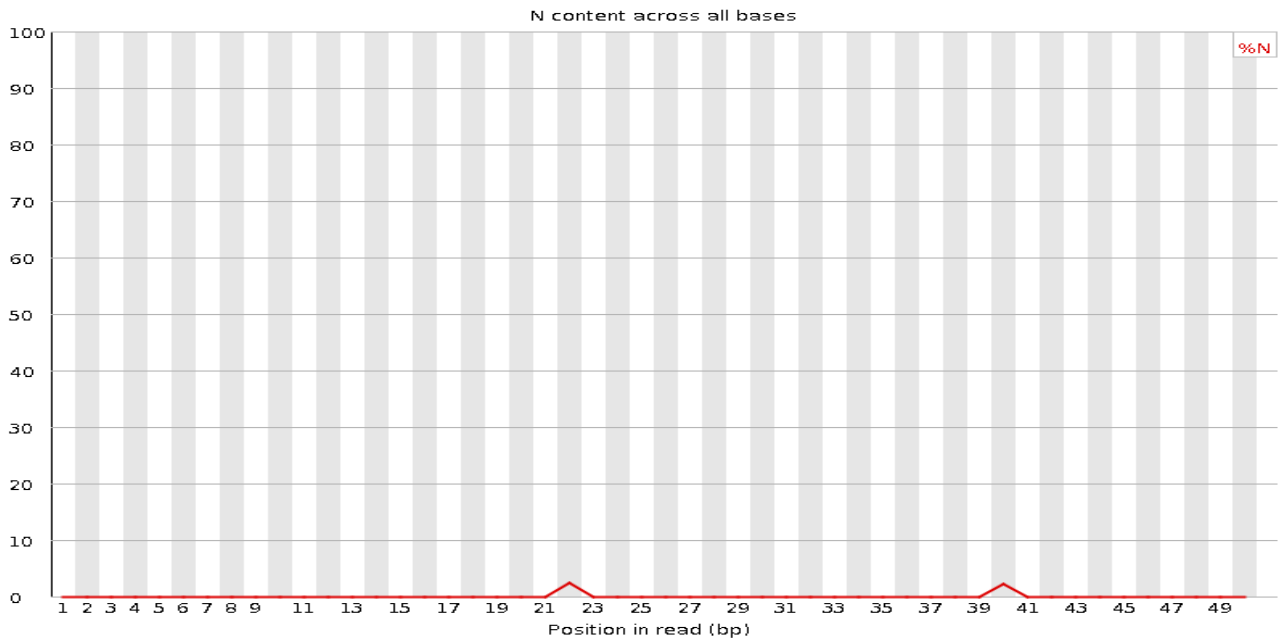


Figura 23 Basi non assegnate

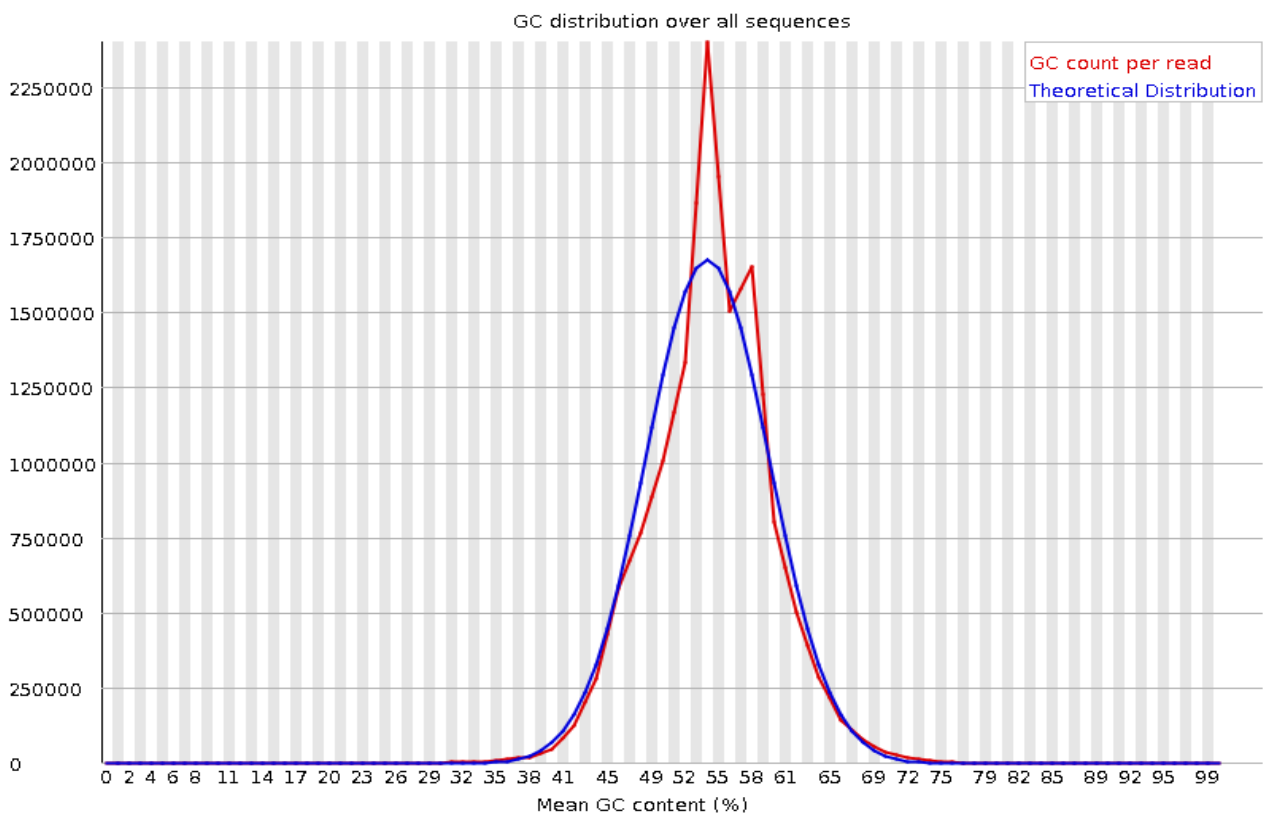


Figura 24 Contenuto medio in siti CpG

5. CONCLUSIONI

Nel presente lavoro è stato utilizzato un approccio epigenetico per l'identificazione di possibili marcatori molecolari in muscolo scheletrico, indicatori di trattamento illecito con promotori della crescita nell'allevamento del bovino da carne. In particolare è stato messo a punto un metodo di metilazione del DNA capace di arricchire selettivamente per la frazione metilata del genoma (arricchimento con MBD) e mettere in evidenza i target della metilazione, ovvero le citosine, convertendo unicamente quelle non metilate in uracile (trattamento con bisulfito). Questo tipo di approccio si è dimostrato idoneo per l'analisi in oggetto perché capace di determinare lo stato di metilazione a livello di singola citosina. La fase finale di elaborazione dei dati è stata condotta grazie all'ausilio di software bioinformatici specifici per studi di epigenomica. Il confronto dei profili di metilazione tra animali trattati e controlli ha rilevato tre regioni differenzialmente metilate, che potrebbero rappresentare potenziali biomarcatori di trattamento illecito con sostanze anabolizzanti. Tuttavia l'analisi non ha rilevato una netta distinzione tra le due situazioni a confronto, di conseguenza è necessario incrementare il numero di replicati biologici per verificare tali sequenze genomiche ed inferire positivamente sulla significatività statistica del dato finale. Inoltre per proporre biomarcatori di provata utilità in fase di screening potrebbe essere vantaggioso inserire altri trattamenti per comprendere in maniera globale l'effetto del trattamento stesso a livello epigenetico al fine di ottenere risultati ripetibili.

6. BIBLIOGRAFIA

- Akalin A, Kormaksson M, Li S, Garret-Bakelman FE, Figueroa ME, Melnick A and Mason CE. 2012. MethylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, 13:R87.
- Bird A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev.* 16: 6-21.
- Birney E. 2011. Chromatin and heritability: how epigenetic studies can complement genetic approaches. *Trends in Genetics*, Vol. 27, No. 5.
- Byun HM, Siegmund KD, Pan F, Weisenberger DJ, Kanel G, Laird PW and Yang AS. 2009. Epigenetic profiling of somatic tissues from human autopsy specimens indentifies tissue- and individual-specific DNA methylation patterns. *Hum Mol Genet*, 18, 4808-4817.
- Cannizzo FT, Capra P, Divari S, Ciccottelli V, Biolatti B and Vincenti M. 2011. Effects of low-dose dexamethasone and prednisolone long term administration in beef calf: chemical and morphological investigation. *Analytica Chimica Acta* 700 95-104.
- Cannizzo FT, Pegolo S, Starvaggi Cucuzza L, Bargelloni L, Divari S, Franch R, Castagnaro M and Biolatti B. 2013. Gene expression profiling of thymus in beef cattle treated with prednisolone. *Research in Veterinary Science* 95 540-547.
- Carraro L, Ferraresso S, Cardazzo B, Romualdi C, Montesissa C, Gottardo F, Patarnello T, Castagnaro M and Bargelloni. 2009. Expression profiling of skeletal muscle in young bulls treated with steroidal growth promoters. *Physiological Genomics* 38:138-148.
- Choi J-H, Li Y, Guo J, Pei L, Rauch TA, Kramer RS, Macmil SL, Wiley GB, Bennet LB, Schnabel JL, Taylor KH, Kim S, Xu D, Sreekumar A, Pfeifer GP, Roe BA, Caldwell CW, Bhalla KN, Shi H. 2010. Genome-wide DNA methylation maps in follicular lymphoma cells determined by methylation-enriched bisulfite sequencing. *PLoS ONE* 5(9): e13020. Doi:10.1371/journal.pone.0013020.
- Couldrey C, Brauning R, Bracegirdle J, Maclean P, Henderson HV and McEvan JC. 2014. Genome-Wide DNA methylation patterns and transcription analysis in sheep muscle. *Plos One*.

- Dean F, Hosono S, Fang L, Wu X, Farugi F, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J, Driscoll M, Song W, Kingsmore SF, Egholm M, and Lasken R. 2002. Comprehensive human genome amplification using multiple displacement amplification. *PNAS* Vol. 99 no. 8 5261-5266.
- Doherty R and Couldrey C. 2014. Exploring genome wide bisulfite sequencing for DNA methylation analysis in livestock: a technical assessment. *Frontiers in Genetics* doi: 10.3389/fgene.2014.00126.
- Feil R and Fraga MF. 2012. Epigenetics and the environment: emerging patterns and implications. *Nature reviews Genetics* 97-109.
- Foley DL, Craig JM, Morley R, Olsson CJ, Dwyer T, Smith K and Saffery R. 2009. Prospects for epigenetic epidemiology. *American Journal of Epidemiology* 2009; 169:389-400.
- Gottardo F, Brscic M, Pozza G, Ossensi C, Contiero B, Marin A and Cozzi G. 2008. Administration of dexamethasone per os in finishing bulls. I. Effects on productive traits, meat quality and cattle behavior as indicator of welfare. *Animal*, 2:7, pp 1073-1079.
- Hansen KD, Langmead B and Irizarry RA. 2012. BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biology* 2012, 13:R83.
- Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Douney SL, Johnson BE, Fouse SD, Delaney A, Zhao Y, Olshen A, Ballinger T, Zhou X, Forsberg KJ, Gu J, Echipare L, O'Geen H, Lister R, Pelizzola M, Xi Y, Epstein CB, Bernstein BE, Hawkin RD, Ren B, Chung WY, Gu H, Bock C, Gnirke A, Zhang MQ, Haussler D, Ecker JR, Li W, Farnham PJ, Waterland RA, Meissner A, Marra MA, Hirst M, Milosavljevic A and Costello JF. 2010. Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nature Biotechnology* Vol 28 Num 10.
- Head JA, Dolinoy DC and Basu N. 2012. Epigenetics for Ecotoxicologists. *Environmental Toxicology and Chemistry*, Vol. 31, No. 2.
- Head JA. 2014. Patterns of DNA methylation in animals: an ecotoxicological perspective. *Integrative and Comparative Biology*, pp. 1-10 doi:10.1093/icb/icu025.

- Ivanov M, Kals M, Kacevsca M, Metspalu A, Ingelman-Sundberg A and Milani L. 2013. In-solution hybrid capture of bisulfite-converted DNA for targeted bisulfite sequencing of 174 ADME genes. *Nucleic Acids Research*, 2013, Vol. 41, No. 6 e72.
- Krueger F and Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatic Applications Note*, vol. 27 no. 11 2011, pages 1571-1572.
- Laird PW. 2010. Principles and challenges of genomewide DNA methylation analysis. *Nature reviews Genetics*.
- Ma B, Wilker EH, Willis-Owen SAG, Byun H, Wong KCC, Motta V, Baccarelli AA, Schwartz J, Cookson WOCM, Khabbaz K, Mittleman MA, Moffatt MF and Liang L. 2014. Prediction DNA methylation level across human tissues. *Nucleic Acids Research*, 2014, Vol. 42, No. 6 3515-3528.
- Ministero della Salute. Piano Nazionale Residui (Relazione finale anno 2013).
- Ministero della Salute. Piano Nazionale per la ricerca dei Residui (PNR 2014).
- Nebbia C. Residui di farmaci e contaminanti ambientali nelle produzioni animali. 2009.
- Nebbia C, Urbani A, Carletti M, Gardini G, Balbo A, Bertarelli D and Girolami F. 2010. Novel strategies for tracing the exposure of meat cattle to illegal growth-promoters. *The Veterinary Journal* 189 34-42.
- Pegolo S, Gallina G, Montesissa C, Capolongo F, Ferraresso S, Pellizzari C, Poppi L, Castagnaro M and Bargelloni L. 2012. Transcriptomic markers meet the real world: finding diagnostic signatures of corticosteroid treatment in commercial beef samples. *BMC Veterinary Research*, 8:205.
- Pezzolato M, Maurella C, Varello K, Meloni D, Bellino C, Borlatto L, Di Corcia D, Capra P, Caramelli M and Bozzetta E. 2011. High sensitivity of a histological method in the detection of low-dosage illicit treatment with 17 β -estradiol in male calves. *Food Control* 1668-1673.
- Pinel G, Weigel S, Antignac JP, Mooney MH, Elliott C, Nielen MWF and Le Bizec B. 2010. Targeted and untargeted profiling of biological fluids to screen for anabolic practices in cattle. *Trends in Analytical Chemistry*, Vol. 29, No. 11.

- Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP and Gu Y. 2012. A tale of three next generation sequencing platforms comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13:341.
- Riedmaier I, Benes V, Blake J, Bretschneider N, Zinser C, Becker C, Meyer HHD and Pfaff MW. 2012. RNA-Sequencing as useful screening tool in the combat against the misuse of anabolic agents. *Analytical Chemistry*, 84, 6863-6868.
- Salozhin SV, Prokhorchuk EB and Georgiev GP. 2005. Methylation of DNA-One of the major epigenetic markers. *Biochemistry (Moscow)*, Vol. 70, No. 5, pp. 525-532.
- Stella R, Arrigoni G, Biancotto G, Krogh M, Vascellari M, Lega F, Pozza G, Angeletti R, Andrighetto I and James P. 2014. Confirmation of protein biomarkers of corticosteroids treatment in veal calves sampled under field conditions. *J. Proteome Res.* 13, 1794-1799.
- Stevens M, Cheng JB, Li D, Xie M, Hong C, Maire CL, Ligon KL, Hirst M, Marra MA, Costello JF and Wang T. 2013. Estimating absolute methylation levels at single-CpG resolution from methylation enrichment and restriction enzyme sequencing methods. *Genome Research*.
- Strömquist M, Tooke N and Brunström B. 2010. DNA methylation levels in the 5' flanking region of the vitellogenin I gene in liver and brain of adult zebrafish (*Danio rerio*)-Sex and tissue differences and effects of 17 α -ethinylestradiol exposure. *Elsevier Aquatic Toxicology* 98 (2010) 275-281.
- Suzuki MM and Bird A. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nature reviews Genetics*.
- Van den Oord E JCG, Bukszar J, Rudolf G, Nerella S, McClay JL and Xie LY. 2013. Estimation of CpG coverage in whole methylome next-generation sequencing studies. *BMC Bioinformatics* 2013, 14:50.
- Wang S, Meyer E, McKay J and Matz M. 2012. 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nature methods*, DOI:10.1038/NMETH.2023.
- Zilberman D and Henikoff S. 2007. Genome-wide analysis of DNA methylation patterns. *Development* 134, 3959-3965.

7. ALLEGATO 1

In rosso le differenze di metilazione significative tra trattati e controlli

chr1	9523247	9523248	31.4902705268154
chr1	11274358	11274359	-28.8911426639621
chr1	20159129	20159130	30.5916305916306
chr1	56557478	56557479	39.2921632358252
chr1	59433556	59433557	-37.3983739837398
chr1	64827125	64827126	30.8627581612259
chr1	90932665	90932666	35.6009563138448
chr1	123063065	123063066	34.3785151856018
chr1	131455959	131455960	-43.341567482596
chr1	138031211	138031212	30.4444444444444
chr1	138504790	138504791	-29.578313253012
chr1	143569579	143569580	61.7647058823529
chr1	146848872	146848873	38.7174833635814
chr1	147312723	147312724	47.718965750643
chr1	147495894	147495895	44.2360379346681
chr1	151402828	151402829	-28.3603238866397
chr1	152686953	152686954	32.8172817281728
chr1	153554621	153554622	-32.3308270676692
chr1	153834860	153834861	26.2295081967213
chr1	153834959	153834960	-29.002284843869
chr1	155639476	155639477	-32.1382007822686
chr1	155809888	155809889	-44.3563728598605
chr1	155809916	155809917	-44.2538307794804
chr1	157280007	157280008	-30.9401709401709
chr1	157733164	157733165	26.8758738542799
chr10	13912333	13912334	37.5
chr10	15026319	15026320	44.2847625392595
chr10	36218473	36218474	-46.900406504065
chr10	39031249	39031250	-28.5255882836926
chr10	62302882	62302883	-27.2384034519957
chr10	62302914	62302915	-28.9953488372093
chr10	75727362	75727363	30.1050013815971
chr10	77222484	77222485	-45.7073104570731
chr10	77389976	77389977	25.0962067069819
chr10	79900711	79900712	28.3008849557522
chr10	81027476	81027477	-50.453339230429
chr10	83380210	83380211	46.8574438202247
chr10	101373152	101373153	-33.3077316948285
chr11	2346104	2346105	-70.9601291764998
chr11	2348953	2348954	30.9039548022599
chr11	2362583	2362584	-68.5283687943262
chr11	2428084	2428085	52.1249359959037
chr11	2989316	2989317	-39.4883980989656
chr11	3846126	3846127	25.9862068965517
chr11	3846179	3846180	49.2143306096794
chr11	5352507	5352508	-52.6603554445248
chr11	10547640	10547641	49.2467532467532

chr11	12347950	12347951	41.7660656495608
chr11	12554953	12554954	32.9980842911877
chr11	12968337	12968338	-32.1703720008805
chr11	19725601	19725602	-42.2222222222222
chr11	22707682	22707683	60.51459673429
chr11	27073583	27073584	-32.9545454545455
chr11	29295054	29295055	-65.737003058104
chr11	45324778	45324779	-36.0576006103376
chr11	53729877	53729878	43.9644481190575
chr11	59745071	59745072	-31.0062893081761
chr11	81432625	81432626	28.9336450278727
chr11	81725238	81725239	-40.7608695652174
chr11	81790863	81790864	30.894457655021
chr11	86334930	86334931	26.3247163410925
chr11	97196114	97196115	-45.2648851383029
chr11	98215441	98215442	-35.7278330962541
chr11	100808821	100808822	33.2972972972973
chr11	101130184	101130185	-28.0921206453121
chr11	101829999	101830000	29.404839764552
chr11	102327873	102327874	29.5542046605876
chr11	103456790	103456791	73.815987933635
chr11	103617258	103617259	29.1079812206573
chr11	103670298	103670299	31.021897810219
chr11	104570597	104570598	-28.4411636995224
chr11	104570968	104570969	-37.7575757575758
chr11	104647988	104647989	26.836611802489
chr11	104654912	104654913	-31.9140292816093
chr11	104680067	104680068	33.0965909090909
chr11	104930567	104930568	33.6263736263736
chr11	105936113	105936114	-33.6300309597523
chr11	105938971	105938972	-29.7299922299922
chr11	106349763	106349764	30.1984126984127
chr11	106358469	106358470	-40.8991228070175
chr11	106463959	106463960	-25.518341307815
chr11	106557838	106557839	45.7763975155279
chr11	106811713	106811714	27.710843373494
chr11	107302100	107302101	26.2184468943912
chr12	2249528	2249529	-26.9757033248082
chr12	2257160	2257161	-29.2798110979929
chr12	2257283	2257284	26.5443609022556
chr12	11961014	11961015	26.1099365750528
chr12	12421292	12421293	30
chr12	16628180	16628181	-25.4524886877828
chr12	23066771	23066772	35.7142857142857
chr12	24218995	24218996	-25.2164502164502
chr12	32620602	32620603	74.2268041237113
chr12	33114531	33114532	32.0932697403286
chr12	34531848	34531849	26.2100456621005
chr12	34573809	34573810	-34.4603672189879
chr12	34895192	34895193	-33.4671705243269
chr12	35646233	35646234	-38.0281690140845
chr12	40255983	40255984	-32.5094260578132

chr12	41827532	41827533	33.6666666666667
chr12	49641970	49641971	-34.6899224806202
chr12	49689356	49689357	26.4959723820483

Ringraziamenti

Al professore Luca Bargelloni per i preziosi consigli.

A Roberta per la pazienza e il supporto.

Alle mie amiche Erica, Roberta e Serena per i momenti spensierati e le lunghe chiacchierate “terapeutiche”.

Ai miei compagni di laboratorio e frisbee Andrea, Fabio e Giulia.

Al nonno che mi difende sempre.

A mia sorella Gioia che si arrabbia alla velocità della luce, ma per me c'è sempre.

A Nicolò che è il mio piccolo sostenitore.

A Riccardo che mi ha “sostituita” egregiamente.