



UNIVERSITÀ DEGLI STUDI DI PADOVA
Dipartimento di Biomedicina Comparata ed Alimentazione
Dipartimento di Agronomia Animali Alimenti Risorse Naturali e Ambiente

Corso di Laurea Magistrale interfacoltà in
BIOTECNOLOGIE PER L'ALIMENTAZIONE

Tesi di Laurea in

**Selezione genomica di tori di razza Frisona
mediante un *panel* di marcatori SNP**

Relatore: ***Ch.mo Prof. Martino Cassandro***

Dip. di Agronomia Animali Alimenti Risorse Naturali e Ambiente

Correlatore: ***Dott. Fabio Maretto***

Dip. di Agronomia Animali Alimenti Risorse Naturali e Ambiente

Laureanda:

Anna Giulia Ferrarese

Matricola n. 1034614

ANNO ACCADEMICO 2012-2013

*Ai miei genitori.
Ad Alberto.*

Indice

RIASSUNTO	1
1. INTRODUZIONE.....	3
1.1 LA RAZZA FRISONA	3
1.2 IL MIGLIORAMENTO GENETICO ANIMALE	6
1.2.1 <i>Principali modelli utilizzati nella selezione animale tradizionale.....</i>	<i>7</i>
1.2.2 <i>Selezione genomica animale.....</i>	<i>14</i>
1.3 CARATTERI ECONOMICAMENTE IMPORTANTI NELLA SELEZIONE DEI BOVINI DA LATTE	22
2. SCOPO DELLA TESI.....	29
3. MATERIALI E METODI.....	31
3.1 ESTRAZIONE E PURIFICAZIONE DEL DNA GENOMICO	31
3.1.1 <i>Estrazione del DNA genomico dai campioni di seme mediante kit DNeasy® Blood & Tissue Kit (QIAGEN® Group).....</i>	<i>31</i>
3.1.2 <i>Estrazione manuale del DNA genomico dai campioni di seme.....</i>	<i>32</i>
3.2 QUANTIFICAZIONE DEL DNA GENOMICO ESTRATTO	33
3.2.1 <i>Quantificazione mediante saggio fluorimetrico Qubit®</i>	<i>33</i>
3.3 <i>GENOTYPING</i> DEI CAMPIONI MEDIANTE L'UTILIZZO DEL <i>GOLDENGATE GENOTYPING ASSAY FOR VERACODE</i> (ILLUMINA INC.)	34
3.3.1 <i>Preparazione dei campioni.....</i>	<i>35</i>
3.3.2 <i>Precipitazione</i>	<i>35</i>
3.3.3 <i>Allelic-Specific-Exstension (ASE).....</i>	<i>36</i>
3.3.4 <i>Estensione e ligazione.....</i>	<i>36</i>
3.3.5 <i>Polimerase Chain Reaction (PCR).....</i>	<i>37</i>
3.3.6 <i>Preparazione della piastra intermedia (INT Plate).....</i>	<i>38</i>
3.3.7 <i>Ibridazione e lettura della VeraCode Bead Plate.....</i>	<i>38</i>
3.4 RACCOLTA DEI DATI FENOTIPICI.....	39
3.5 ANALISI STATISTICA DEI DATI	40
4. RISULTATI	43

4.1 ELABORAZIONE DEI DATI MEDIANTE IL PROGRAMMA <i>GENOMESTUDIO DATA</i> <i>ANALYSIS</i> (ILLUMINA INC).....	43
4.2 ANALISI GENETICA	48
4.3 ELABORAZIONE DEI DATI FENOTIPICI.....	48
4.4 ANALISI DI ASSOCIAZIONE.....	54
5. DISCUSSIONE E CONCLUSIONI	59
6. BIBLIOGRAFIA	65
<i>RINGRAZIAMENTI</i>.....	73

Riassunto

I bovini di razza Frisona sono ad oggi considerati i maggiori produttori di latte al mondo. Il raggiungimento di questo importante traguardo è stato possibile grazie ai programmi di selezione che fin dal 1929 hanno interessato questa tipologia di bovini. Negli ultimi vent'anni i progressi compiuti dalla biologia molecolare hanno permesso di operare numerose scoperte nell'ambito della selezione animale: attraverso l'ausilio di marcatori molecolari come SNP e SSR, sono stati condotti numerosi studi finalizzati alla rilevazione e all'identificazione di QTL, allo scopo di intraprendere una selezione genomica assistita da marcatori molecolari (MAS, *Marker Assisted Selection*), in grado di portare ad una scelta più accurata, rapida e meno costosa dei migliori animali.

La finalità del presente lavoro di Tesi è stata quella di valutare l'associazione tra i dati genotipici ottenuti mediante *SNP genotyping*, condotto su campioni di seme bovino, e le informazioni fenotipiche registrate attraverso i convenzionali *progeny* e *performance test*, al fine di stabilire se vi siano dei presupposti per avviare programmi di selezione assistita da marcatori molecolari nella scelta e nel miglioramento dei tori da riproduzione di razza Frisona.

Dai risultati ottenuti è stato possibile constatare come l'utilizzo della selezione dei tori di razza Frisona mediante marcatori SNP sia un valido strumento per poter identificare, in breve tempo e con un minor dispendio economico, quelli che saranno i migliori torelli da avviare alle prove di progenie. I valori di sostituzione allelica ottenuti si sono dimostrati complessivamente in linea con quanto riportato in bibliografia e, nella gran parte dei casi, supportati da una buona significatività. Ovviamente questo nuovo approccio presenta ancora la necessità di confermare i dati genotipici con quelli fenotipici forniti dalle prove di progenie, ma permette in ogni caso di aumentare l'accuratezza nella selezione dei torelli e di ridurre in questo modo i costi legati al loro mantenimento.

1. Introduzione

1.1 La Razza Frisona

La razza Frisona, anche nota come Holstein-Fresian, rappresenta la tipologia di bovini da latte ad oggi riconosciuta come la più produttiva al mondo. La sua origine è da ricercarsi nel nord Europa ed in particolar modo nella zona geografica compresa tra la regione olandese Fresia e quella tedesca Holstein.

Per risalire alla nascita di questa razza, è necessario ritornare molto indietro nel tempo: si ritiene, infatti, che intorno al 100 A.C. i pastori nomadi dell'Asia centrale arrivarono con i loro bovini dal mantello nero fino alle regioni del Mare del Nord, occupando le zone limitrofe il Reno, dove le popolazioni locali erano solite allevare bovini con mantello bianco. Dall'incontro di questi due popoli si può ipotizzare la nascita della razza Holstein-Fresian caratterizzata da animali dal tipico mantello pezzato nero.

Per molti anni gli allevatori tedeschi operarono incroci mirati a sviluppare una razza costituita da animali in grado di utilizzare al meglio le vaste distese erbose tipiche dell'area geografica compresa tra la Germania e l'Olanda. Con il passare dei secoli, grazie a questo meccanismo di selezione, si è arrivati all'ottenimento di vacche da latte altamente produttive. Ebbe poi inizio una graduale e continua diffusione di questi bovini, che riuscirono in questo modo a soppiantare molte razze locali sia in modo diretto che attraverso incroci di sostituzione (Balasini, 1995), ottenendo in molti Paesi la formazione di "ceppi" con proprie caratteristiche morfo-funzionali.

Successivamente, con la scoperta e lo sviluppo commerciale del "nuovo mondo", gli allevatori e riproduttori americani iniziarono ad acquistare il loro bestiame in Olanda, in modo da poter importare la produzione di prodotti lattiero-caseari. La razza Frisona olandese infatti, grazie alle importazioni operate dai coloni fiamminghi, iniziò ad essere commercializzata negli Stati Uniti fin dal 1621 e, nella seconda metà del 1800, questi capi divennero la base genetica da cui ebbe origine la selezione del ceppo americano Holstein-Fresian, con capi specializzati nella produzione di latte e burro.

I primi animali ad essere importati nel nostro Paese furono, nel 1870, bovini olandesi che ben presto, intorno al 1908, raggiunsero i 60.000 capi (Balasini, 1995). La

distribuzione riguardò prevalentemente le regioni del Nord Italia in quanto zone a zootecnia più avanzata per via della fertilità del terreno e della presenza di una fitta rete di commerci. Come nel resto dell'Europa, anche in Italia si tentò inizialmente di mediare tra la produzione di latte e la resa in carne, cercando quindi di ottenere bovini a duplice attitudine. Tuttavia, col passare degli anni la Frisona venne sempre più selezionata per la produzione di latte, in quanto a quell'epoca per la produzione di carne le era preferita la Bruna Alpina. La spinta al miglioramento della Frisona al fine di farla diventare razza solo da latte si ebbe nel 1929 quando, nell'allevamento della Bonifica di Torre in Pietra, venne introdotto per la prima volta il toro *Carnation Producer* proveniente dalla *Carnation Milk Farm* di Seattle (U.S.A.) (Balasini, 1995). Questo avvenimento rappresentò una svolta nelle metodiche di selezione attuate fino a quel momento, le quali erano basate soprattutto su caratteri fenotipici come il colore del mantello e l'armonia della struttura fisica (Fusco, 1990). Durante la seconda guerra mondiale, la macellazione forzata del bestiame per necessità alimentari accentuò la pressione selettiva, in quanto i pochi animali risparmiati erano i più produttivi. Questo permise di evidenziare, a conflitto concluso, la netta superiorità economica e produttiva della razza Frisona, in particolare nei confronti della Bruna Alpina. Proprio per questo motivo, nel corso del dopoguerra ci fu una fortissima domanda di bovine Pezzate Nere e di bovine americane del tipo *Carnation*, che portò ad un conseguente incremento dell'importazione di capi olandesi e americani. Nello stesso periodo prese anche piede l'uso dell'inseminazione artificiale strumentale, che permise non solo di aumentare il numero di discendenti di un determinato toro, migliorandone i processi di selezione fino a quel momento utilizzati, ma anche di contrastare gravi malattie come la brucellosi.

Ad oggi, in Italia la razza Frisona Italiana ricopre un ruolo preponderante nella produzione di latte: dai dati A.N.A.F.I. (*Associazione Nazionale Allevatori Frisona Italiana*), emerge infatti come le vacche di razza Frisona allevate a livello Nazionale nel 2012 fossero 1.130.270, con una media di 90 vacche per allevamento (Tab. 1).

Tabella 1: Statistiche Nazionali relative alla razza Frisona Italiana fino all'anno 2012 (fonte: <http://www.anafi.it/>).

<i>Anno</i>	<i>Consistenza</i>		<i>Dati Nazionali</i>			
	<i>Allevamenti</i>	<i>N° vacche</i>	<i>Latte, Kg</i>	<i>Grasso, %</i>	<i>Proteina, %</i>	<i>Vacche per allevamento</i>
2002	15.106	1.088.178	8.346	3,62	3,32	72
2003	14.984	1.107.701	8.423	3,59	3,30	74
2004	14.823	1.100.543	8.592	3,71	3,32	74
2005	14.317	1.101.657	8.901	3,70	3,31	77
2006	14.069	1.102.655	8.945	3,70	3,33	78
2007	13.818	1.100.401	9.052	3,68	3,34	80
2008	13.510	1.101.868	8.979	3,68	3,37	82
2009	13.327	1.103.453	8.869	3,68	3,38	83
2010	13.164	1.113.859	9.125	3,70	3,36	85
2011	12.922	1.128.626	9.190	3,67	3,35	87
2012	12.578	1.130.270	9.320	3,72	3,38	90

Dal punto di vista morfologico (Fig 1), le vacche di razza Frisona Italiana presentano mantello pezzato nero e a volte rosso per la presenza del carattere recessivo; sono di statura alta, testa proporzionata e musello ampio; l'anteriore è armonico con il garrese ben pronunciato, le spalle fuse con il collo, gli arti anteriori e posteriori sono ben distanziati e i piedi forti con alta suola; la mammella è saldamente attaccata e presenta vene mammarie molto ramificate e non troppo grosse. Il legamento sospensorio mediano è forte e divide nettamente la mammella in due parti uguali. Il peso della femmina adulta è di 500 – 900 kg, l'altezza è di 130 – 150 cm e la durata della gravidanza è di circa 287 giorni. Per quanto riguarda i tori (Fig. 2), sono di mole maggiore: infatti, il peso del maschio adulto oscilla tra i 900 e i 1300 kg e l'altezza va da 138 a 155 cm. La valutazione morfologica viene espressa per tutte le femmine che abbiano partorito e per i maschi di oltre 18 mesi.



Figura 1: esempio di vacca Frisona Italiana



Figura 2: esempio di toro di razza Frisona Italiana

1.2 Il miglioramento genetico animale

Fin dagli inizi della storia e probabilmente a partire dai primi tentativi di addomesticazione avvenuti 12.000 – 14.000 anni fa, il miglioramento genetico animale si è basato sulla selezione degli individui fenotipicamente migliori e per questo adatti ad essere i progenitori delle future generazioni. A partire dalle origini dell'allevamento, l'uomo ha infatti progressivamente modificato ed accentuato le caratteristiche e le capacità del bestiame in modo da poter ottenere una maggior quantità di prodotti utili per il suo sostentamento, senza tuttavia essere realmente a conoscenza di tutti gli aspetti genetici che sono alla base di questo processo. Solo alla fine del ventesimo secolo, molti di questi concetti sono stati formalizzati in una nuova disciplina, chiamata genetica quantitativa. Da questo momento in poi, l'incremento dei caratteri quantitativi nel bestiame è stato associato all'accuratezza delle registrazioni fenotipiche degli individui e dei loro parenti, oltre alle conoscenze alla base delle variazioni genetiche per quei caratteri. Questo approccio, pur essendo di successo, si è dimostrato relativamente lento per caratteri misurati in tarda età, difficili e costosi da valutare o legati al sesso.

Più di 100 anni fa, attraverso la formazione di registri di bestiame e di società specializzate negli incroci, ebbe inizio la raccolta e l'organizzazione delle informazioni sul *pedigree* e sulle *performance* degli animali. Questi dati servirono per mettere a punto esperimenti di incrocio, allo scopo di migliorare la produttività, la salute e la morfologia dei capi allevati. Tale meccanismo si è protratto fino ai giorni nostri, in cui

molte specie bovine sono soggette alla selezione per molti caratteri utilizzando stime del merito genetico ottenuto attraverso complesse analisi statistiche e computazionali.

I progressi nella genetica molecolare avvenuti nella seconda metà dell'ultimo secolo hanno poi permesso di espandere, o addirittura di rimodellare, il panorama della selezione animale. In particolar modo, negli anni '90 l'associazione tra clonaggio molecolare e analisi del DNA condotta attraverso i marcatori microsatellite, ha portato allo sviluppo di una mappa completa di associazione genica per i bovini. Questa scoperta ha permesso poi di identificare la posizione cromosomica di *loci* ad azione quantitativa (QTL) responsabili di rilevanti variazioni nei caratteri di interesse economico, come la produzione di latte e carne. Tuttavia, nella gran parte dei casi, l'utilizzo diretto delle informazioni riguardanti il genotipo a partire da questi risultati, si è dimostrato molto più impegnativo di quanto detto in precedenza, soprattutto nella selezione genetica assistita da marcatori molecolari (Dekkers, 2004). La soluzione a questo inconveniente è stata possibile grazie ai grandi progressi compiuti negli ultimi decenni dalle tecniche molecolari, per mezzo delle quali, per esempio, si sono potuti utilizzare metodi per la determinazione simultanea e precisa del genotipo di centinaia di migliaia di posizioni di SNP (*Single Nucleotide Polymorphisms*) che, pur essendo meno informativi dei microsatelliti, a parità di *loci* analizzati permettono di monitorare in modo efficiente ed economico le variazioni genetiche tra molti campioni.

1.2.1 Principali modelli utilizzati nella selezione animale tradizionale

Come detto in precedenza, il percorso che ha portato alla formazione dei moderni metodi di selezione animale è stato lungo e laborioso. Utilizzando le conoscenze di statistica e genetica accumulate sin dagli anni '40 del secolo scorso, Lush ed Hazel (Lush, 1933; Hazel, 1943) svilupparono i principi dell'indice di selezione, che può essere considerato come il primo e vero metodo per ottimizzare la selezione artificiale. Tale metodo si basa sulla conoscenza della parentela tra individui e sulle informazioni fenotipiche rilevate sui caratteri di interesse. Esso consiste infatti nell'andare a sommare i diversi caratteri di interesse pesati in modo opportuno in base alla loro importanza economica, al fine di selezionare i migliori riproduttori senza necessariamente dover

ricorrere alla conoscenza del loro genoma e dell'effetto dei singoli geni che lo costituiscono. In questo modo è quindi possibile andare a misurare un carattere, come per esempio la produzione di latte, separando la componente genetica da quella ambientale (per esempio alimentazione, gestione dell'allevamento, ecc.), in modo da fornire una classifica dei migliori animali in base al loro valore genetico. Gli indici di selezione risultano essere pertanto molto vantaggiosi, in quanto permettono di massimizzare la risposta della selezione di uno o più caratteri usando molte sorgenti di informazione, come quelle relative alle caratteristiche proprie dell'individuo oppure quelle dei suoi parenti. Il postulato di Fisher, infatti, si basa sul fatto che un grande numero di geni controllano insieme caratteri fenotipici, e che ciascuno di questi fattori sia in grado di contribuire per una piccola quota alla varianza fenotipica complessiva. Una conseguenza importante del modello di Fisher è che l'ereditabilità di questi caratteri fenotipici può essere spiegata attraverso il lavoro di Mendel, comportando la possibilità di trovare correlazioni genetiche tra individui e permettendo di effettuare l'implementazione della selezione artificiale.

L'Indice di selezione (I) consiste perciò nel combinare diverse fonti di informazione per la stima del vero e sconosciuto valore genetico additivo (H) dell'animale X pari a:

$$H = a_1BV_1 + a_2BV_2 + \dots\dots\dots a_nBV_n$$

dove H è il genotipo aggregato dell'individuo X .

Attraverso questa formula il carattere i assume un peso economico misurato sulla base di a_i e BV_i , che rappresentano il valore di incrocio per ciascuno specifico carattere all'interno dell'aggregato BV (Maltecca et al., 2010).

L'indice di selezione (I) può essere quindi espresso come:

$$I = b_1x_1 + b_2x_2 + \dots + b_nx_n$$

dove:

I = indice di selezione dell'animale X , espresso come EBV (*Estimated Breeding Value*) o PTA (*Prediction Transmitting Ability*). Esso equivale a $0,5*EBV^1$;

b_i = coefficiente di ponderazione dell'indice di selezione dell'osservazione fenotipica dell' i -mo carattere;

x_i = osservazione fenotipica sull'animale oggetto di valutazione e dei suoi relativi parenti per l' i -mo carattere, espressa come deviazione dalla media di popolazione.

Le proprietà dell'indice di selezione sono quelle di minimizzare l'errore di predizione, ovvero la differenza tra lo sconosciuto valore di H ed il suo predittore I . Questo ci permette di massimizzare la correlazione tra H ed I , che corrisponde all'accuratezza dell'indice di selezione. Se queste proprietà sono rispettate allora esisterà una sola funzione lineare di X che soddisferà la seguente uguaglianza in termini matriciali:

$$I = b X$$

da cui

$$P b = G v$$

$$b = P^{-1} G v$$

dove :

P = matrice $n \times n$ di covarianze fenotipiche tra le osservazioni;

G = matrice $n \times m$ di covarianze genetiche tra le n osservazioni fenotipiche e gli m caratteri inclusi nel genotipo aggregato;

v = vettore colonna dei valori economici per gli m caratteri nel genotipo aggregato;

b = vettore colonna degli n coefficienti di ponderazione dell'indice di selezione.

In tal caso l'indice di selezione (I) può essere considerato come la miglior predizione lineare (*Best Linear Prediction*, BLP) del valore genetico additivo (H) dell'animale X , assumendo che gli effetti fissi (cioè tutti quei fattori ambientali di natura non genetica

¹ Gli EBV vengono calcolati utilizzando il metodo BLUP (*Best Linear Unbiased Prediction*) in cui si tengono in considerazione il *pedigree* e le informazioni di molti caratteri ricavate dalle *performance* dell'individuo e dei suoi parenti.

che possono influenzare la manifestazione del carattere in esame) siano noti, condizione che nella realtà spesso non viene mantenuta.

Nel 1975, C.R. Henderson definì e pubblicò l'uso del *Mixed Model Equation* (MME), anche conosciuto come metodo BLUP, ancora oggi utilizzato per le valutazioni genetiche nel campo animale. Il MME combina di fatto i metodi di stima dei minimi quadrati o GLS (*General Least Squares*) e l'indice di selezione, consentendo in questo modo di derivare stime non distorte dei valori genetici (BLUP) di individui campionati in differenti ambienti (per esempio allevamenti) o momenti temporali diversi (per esempio anni). Risulta così possibile ottenere i cosiddetti *Best Linear Unbiased Estimation* (BLUE). Tale metodologia permise di superare uno dei limiti principali dell'indice di selezione ovvero l'assunzione di conoscere a priori le vere medie degli effetti fissi non affette da errori sistematici. Il MME può essere così descritto in termini matriciali:

$$y = Xb + Zu + e$$

dove:

y = vettore delle osservazioni fenotipiche;

X = matrice di incidenza degli effetti fissi;

Z = matrice di incidenza delle osservazioni sull'animale;

b = vettore degli effetti fissi;

u = vettore degli effetti casuali;

e = vettore dei residui casuali con media zero e varianza σ_e^2 .

Le soluzioni per b e u possono essere derivate risolvendo il seguente sistema di equazioni del MME:

$$\begin{array}{|c|c|} \hline X'R^{-1}X & X'R^{-1}Z \\ \hline \end{array} \begin{array}{|c|} \hline b \\ \hline \end{array} = \begin{array}{|c|} \hline X'R^{-1}y \\ \hline \end{array}$$

$$\begin{array}{|c|c|} \hline Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \\ \hline \end{array} \begin{array}{|c|} \hline u \\ \hline \end{array} = \begin{array}{|c|} \hline Z'R^{-1}y \\ \hline \end{array}$$

Se da un lato il “metodo BLUP” illuminò il mondo delle valutazioni genetiche animali, dall'altro non venne di fatto applicato per oltre 20 anni per l'impossibilità di disporre di calcolatori in grado di gestire matrici di migliaia di righe e di colonne. Solo

all'inizio degli anni '70 si iniziarono a risolvere i primi modelli MME su archivi di grandi dimensioni e da allora il “modello BLUP” viene utilizzato come metodo standard nel settore del miglioramento genetico animale per la valutazione di vacche da latte, bovini da carne e suini, in moltissime Nazioni e centri di miglioramento genetico (Maltecca et al., 2010).

Diversi sviluppi sono stati effettuati nella storia del MME ed in particolar modo essi hanno riguardato l'analisi di modelli multivariati o a caratteri multipli (*Multiple Traits Model*) e l'uso di modelli a ripetibilità (*Repeatability Models*) nei quali si utilizzano informazioni ripetute sugli stessi animali, in modo da considerare gli effetti individuali comuni che permettono di stimare la ripetibilità di un carattere. La ripetibilità è un parametro molto importante per un carattere da migliorare geneticamente, perché consente di pianificare efficacemente a livello operativo quale sistema di raccolta dati sia più adeguato. In altri termini, con valori elevati di ripetibilità ci si può permettere di valutare accuratamente i riproduttori con sistemi di raccolta basati su rilevazioni uniche, mentre nel caso in cui i valori di ripetibilità siano ridotti si dovranno effettuare misurazioni ripetute dello stesso carattere per garantire un'accurata valutazione genetica. Altre evoluzioni del MME hanno riguardato la complessità dei modelli di valutazione genetica al fine di superare inadeguatezze di calcolo e di rilevamento dei dati in campo (Maltecca et al., 2010). Un esempio nel settore dei bovini da latte è quello riguardante il passaggio da modelli *sire model*, che considerano solo l'effetto del padre, a modelli più completi quali il *maternal grand-sire* o l'*Animal model*, definito come il modello ideale per la realtà della valutazione genetica dei riproduttori di bovine da latte (Quaas and Pollak, 1980). Altre evoluzioni metodologiche riguardano le analisi dei caratteri non distribuiti normalmente per i quali si sono definiti modelli a soglia o *threshold model* (Gianola e Foulley, 1983). Un esempio di tali caratteri può essere rappresentato dalle facilità al parto in bovini da latte e da carne o anche da caratteri legati alla riproduzione e fertilità per i quali non si assiste ad una variazione continua e teoricamente infinita, ma esistono poche classi o limiti nella curva di distribuzione dei dati. I più recenti sviluppi del MME hanno riguardato l'applicazione di *sire model* di tipo *multi-traits* per le valutazioni genetiche internazionali dei riproduttori bovini da latte e carne, il cosiddetto MACE (*Multiple Across Country Evaluation*). Questo metodo proposto da Schaeffer (2006) ha permesso di combinare le valutazioni genetiche dei riproduttori, elaborate con i diversi MME dai vari Paesi, in una singola valutazione internazionale. Tale MACE è attualmente in uso presso il centro di Interbull

(*International Bull Evaluation Service*), che ha sede ad Uppsala (Svezia) e di cui fanno parte 34 Paesi membri. Trimestralmente vengono inviate a questo centro le valutazioni genetiche di tutti i riproduttori disponibili su di un *ranking* mondiale (www.interbull.slu.se), rappresentando i dati di partenza delle valutazioni internazionali. Attraverso il modello MACE, Interbull procede alla valutazione di ogni carattere tenendo conto delle stime delle correlazioni genetiche tra tori nei differenti Paesi (in quanto specchio delle interazioni tra genotipo ed ambiente) e dell'ereditabilità che si presenterà in modo diverso per ogni Paese. Tali interazioni sono dovute principalmente alle differenze tra i sistemi di produzione, di raccolta dati e di valutazioni genetiche, nonché alle differenze climatiche esistenti tra i vari Stati. Quando la correlazione genetica tra due Paesi è significativamente diversa da 1, i risultati di MACE produrranno due differenti graduatorie in cui verranno assegnati i tori, con la possibilità di ricoprire una posizione diversa in base allo Stato di appartenenza. Questo significa, ad esempio, che la classifica Interbull della Francia può essere diversa da quella stilata per l'Olanda o per altri Paesi (Tartara, 1999). Il dato finale che si ottiene permette di valutare quanto simili siano le classifiche dei tori di uno Stato rispetto ad un altro (Fig. 3).

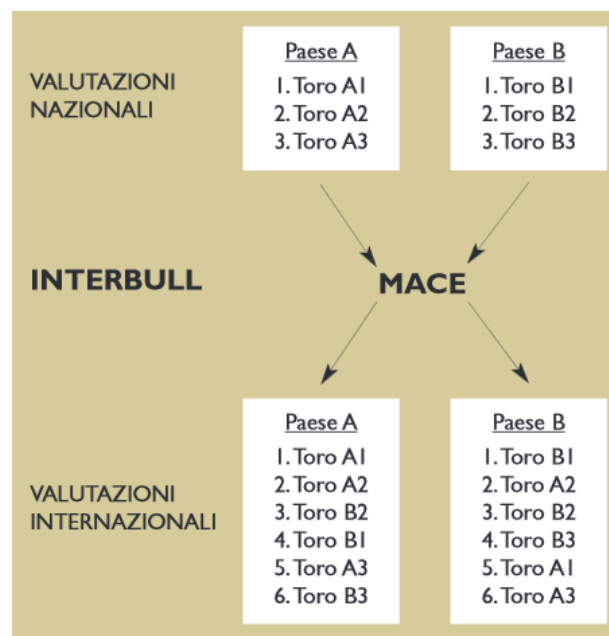


Figura 3: rappresentazione di come la procedura MACE, utilizzata da Interbull per le valutazioni internazionali, possa modificare cambiamenti nelle posizioni dei tori appartenenti a due Paesi distinti (Tartara, 1999).

Interbull, sulla base della propria scala, fornisce inoltre a ciascun Paese l'indice dei tori valutati in tutto il mondo (circa 70.000) (ANAFI, 2013) (Tab. 2).

Tabella 2: correlazioni genetiche per i caratteri per cui viene calcolato l'indice, tra Italia e alcuni Paesi partecipanti al servizio fornito da Interbull.

<i>Carattere</i>						
Paese	Proteina	Tipo	Cellule	Longevità	Fertilità	Facilità parto
USA	0,91	0,87	0,91	0,72	0,93	0,60
CAN	0,89	0,86	0,91	0,78	0,71	0,66
FRA	0,87	0,90	0,93	0,69	0,62	0,68
DFS ₁	0,89	0,78	0,93	0,61	0,85	0,75
NLD	0,85	0,85	0,88	0,47	0,83	0,71
DEU	0,86	0,88	0,94	0,76	0,82	0,65
AUS	0,75	0,52	0,81	0,43	0,85	0,59
NZL	0,75	0,58	0,76	0,30	0,65	-

DFS₁: Danimarca, Finlandia, Svezia (fonte: ANAFI, 2013).

Nella seconda metà degli anni '80, Gianola (1986) sviluppò un nuovo metodo Bayesiano che divenne un importantissimo strumento nella risoluzione dei problemi riscontrati fino a quel momento nella selezione animale. Secondo questa inferenza Bayesiana, non esiste una distinzione logica tra i parametri del modello (intesi come le variabili osservate) e i dati (cioè le variabili non osservate), perchè entrambi sono variabili casuali descritte da una distribuzione di probabilità congiunta specificata da un modello probabilistico. Questo metodo statistico permette inoltre di incorporare informazioni, dette informazioni a priori, di cui si dispone nel momento in cui si deve specificare il modello. Consideriamo per esempio un caso generico come:

$$\begin{aligned}
 p(\theta, y) &= p(\theta) p(y | \theta) \\
 &= p(y) p(\theta | y)
 \end{aligned}$$

dove $p(\theta | y)$ è la distribuzione congiunta dell'informazione sconosciuta (θ) e di quella nota (y). La distribuzione $p(\theta)$ è pesata verso quei valori di θ che sono considerati più probabili in accordo con l'informazione a priori utilizzata. In assenza di qualunque informazione, le *priors* (o informazioni a priori) hanno distribuzione rettangolare

nell'intervallo tra i valori di minimo e massimo del parametro (*flat priors*). Il nostro interesse è quindi quello di arrivare a conoscere il valore sconosciuto di θ e, alla luce della formula precedentemente riportata, possiamo ricavare che:

$$\begin{aligned} p(\theta / y) &= \frac{p(\theta) p(y / \theta)}{p(y)} \\ &\propto p(\theta) p(y / \theta) \end{aligned}$$

dove l'informazione a priori viene combinata con la *likelihood*², moltiplicandole tra loro ed ottenendo così una distribuzione a posteriori $p(\theta / y)$. Da questa equazione possiamo dedurre che il valore di $p(\theta / y)$ sarà tanto più elevato quanto più i valori di θ saranno supportati dai dati (quindi avranno una *likelihood* elevata) e saranno quindi coerenti con l'informazione assunta a priori ($p(\theta)$ elevata).

In generale il metodo statistico dell'inferenza Bayesiana risulta essere molto utile nel caso in cui si stiano utilizzando modelli con molte variabili interdipendenti tra loro e limitate da un intervallo particolare di valori. La presenza della *prior* permette infatti di imporre quali siano i valori massimi che le variabili possono assumere (Beaumont et al., 2004).

1.2.2 Selezione genomica animale

L'obiettivo principale del miglioramento dei bovini da latte è quello di incrementare i caratteri economicamente importanti. Tuttavia molti di questi caratteri sono di natura quantitativa, in quanto la loro espressione viene influenzata non solo da fattori ambientali, ma anche dall'azione di molti geni (Zhang et al., 1998). Le recenti innovazioni introdotte in ambito molecolare hanno permesso di ottenere una maggior efficienza ed accuratezza nella selezione animale ed in particolar modo, nella scelta dei torelli da avviare alle prove di progenie e delle madri di toro (Canavesi, 2010). Attraverso l'identificazione e il mappaggio di numerosi *loci* ad azione quantitativa

² Con il termine *likelihood* si intende una distribuzione condizionale che specifica la verosimiglianza dei dati osservati possedendo un qualsiasi valore dei parametri.

economicamente rilevanti nei bovini da latte e all'uso della selezione genomica assistita da marcatori molecolari, gli allevatori hanno avuto l'opportunità di ridurre i costi, di diminuire gli intervalli generazionali e di evitare la possibilità di *inbreeding* tra gli animali (Daetwyler et al., 2007).

Per capire meglio il significato del termine “selezione genomica” possiamo utilizzare la definizione fornita da Meuwissen (2007), con la quale egli afferma che la selezione genomica è una “*selezione simultanea di molte decine o centinaia di migliaia di marcatori, con i quali è possibile coprire l'intero genoma in modo denso, cosicché tutti i geni siano previsti essere in Linkage Disequilibrium con almeno alcuni dei marcatori usati*”. In particolar modo, la selezione genomica si basa sulla predizione dei valori di incrocio a partire da genotipi individuali (*Estimated Genomic Breeding Values*, GEBV) costituiti da un gran numero di marcatori di DNA sotto forma di polimorfismi di singoli nucleotidi (SNP), i quali sono in *Linkage Disequilibrium* con i loci ad azione quantitativa (QTL) codificanti caratteri con importante valore economico (Börner et al., 2012). Per quanto riguarda gli SNP, essi rappresentano una variazione, all'interno della sequenza di DNA, di un singolo nucleotide che differisce tra membri di una specie o anche solo all'interno di una coppia di cromosomi di un individuo. Tuttavia, proprio per la loro natura binaria, gli SNP risultano intrinsecamente meno informativi rispetto ad altri marcatori molecolari (come per esempio i microsatelliti). Nonostante questo, essi sono i *markers* più densamente distribuiti all'interno del genoma dei mammiferi, sono i più economici in termini di costo di analisi e, grazie ai recenti progressi compiuti in ambito tecnologico, è stato possibile sviluppare metodiche in grado di analizzare contemporaneamente un grandissimo numero di SNP. Per tutti questi motivi essi risultano essere i *markers* più adatti per la selezione genomica.

A partire dagli anni '90, vi è stato un progressivo interesse riguardo ai QTL che controllano caratteri economicamente importanti negli animali da reddito. Questo ha portato alla graduale identificazione di geni o regioni cromosomiche associati a differenze fenotipiche tra individui, famiglie, razze (Maltecca et al., 2010) e, in alcuni casi, si è riusciti ad individuare le varianti alleliche responsabili delle variazioni fenotipiche e ad incorporarle facilmente nelle stime dei valori di incrocio degli animali (Visscher, 1996). Sfortunatamente questi casi rappresentano delle eccezioni, in quanto l'utilizzo diretto delle informazioni genotipiche nei programmi di selezione attuati per mezzo di marcatori molecolari (*Marker Assisted Selection*, MAS) non sempre ha dato risultati positivi (Dekkers, 2004).

La prima metodologia sperimentata per attuare la selezione genomica è stata presentata da Meuwissen (Meuwissen et al., 2001). Egli, insieme ai suoi collaboratori, introdusse la selezione genomica come una strategia selettiva basata su valori riproduttivi predetti a partire da numerosi dati, ottenuti dall'utilizzo di marcatori molecolari. Questo metodo riconosce implicitamente il fatto che caratteri quantitativi, come quelli che influenzano il profitto delle produzioni animali, siano controllati dalla segregazione di un grande numero di *loci* multipli ad azione quantitativa (QTL) e assume, inoltre, che i marcatori siano in *Linkage Disequilibrium* (LD) con il QTL: quindi, più alta è la densità dei marcatori, maggiore sarà il livello di LD tra i marcatori e il QTL e, di conseguenza, la proporzione di varianza genetica che può essere spiegata attraverso i marcatori. In questo approccio vengono quindi utilizzati contemporaneamente centinaia di migliaia di SNP assumendo che tutti i QTL da saggiare siano in LD con almeno uno di questi marcatori. In seguito, viene effettuata una stima dell'effetto che ciascuno SNP può avere, mentre la stima del valore genetico di un individuo si ottiene sommando tutto gli effetti marcatori dell'individuo secondo la seguente formula:

$$GEBV = \sum_i^n x_i \hat{g}_i$$

dove:

GEBV: *Genomic Estimated Breeding Value* di un individuo ottenuto dalla contemporanea predizione degli *n* marcatori individuali;

x_i: genotipo dell'individuo rispetto ad uno specifico marcatore (per esempio se presenta 0, 1 o 2 copie di una specifica variante allelica);

ĝ_i: predizione dell'effetto dello SNP considerato (Maltecca et al., 2010).

Questo tipo di approccio permette, già dagli stadi più precoci della vita dell'animale, di poter calcolare in modo affidabile i GEBV per entrambi i sessi, con accuratezza superiore al 70%. Tutto ciò comporta uno spostamento dal modello animale BLUP e dalle informazioni sul *pedigree* finora utilizzati, verso un nuovo modello BLUP basato sulle informazioni fornite dagli SNP, rendendo meno rilevanti i dati ottenuti dall'osservazione della genealogia dell'animale (Köning et al., 2009).

I vantaggi apportati dalla selezione genomica nel miglioramento degli animali da reddito sono molteplici. Grazie all'analisi del DNA, è infatti possibile valutare i torelli

in prova ed avere un'indicazione delle varianti alleliche da loro possedute. Combinando queste informazioni con quelle ottenute dall'indice *pedigree* si può operare una scelta dei capi da avviare alla prova di progenie in modo più oggettivo, migliorando significativamente l'efficienza dell'analisi (Canavesi, 2010). La genomica, inoltre, consente di valutare il profilo genetico ereditato dai genitori in modo più diretto e quindi di determinare con più oggettività ed accuratezza le bovine (Fig. 4).

Carattere	Pedigree	Pedigree+DNA
Latte kg	35%	58%
Grasso kg	35%	68%
Proteina kg	35%	57%
Grasso %	35%	78%
Proteina %	35%	69%
Final Score	24%	42%
Longevità	27%	45%
SCS (Cellule)	30%	51%
Fertilità figlie	25%	41%

L'aumento di accuratezza corrisponde all'informazione portata da circa 11 figlie in lattazione (dati Usda, Van Raden, luglio 2008).

Figura 4: esempio di come la genomica possa portare ad un aumento dell'attendibilità dei dati registrati su giovani tori in prova (Canavesi, 2010).

Tradizionalmente, la selezione dei capi migliori viene fatta partendo dai dati raccolti nel tempo sulle *performance* dei parenti e sulle diverse relazioni tra le popolazioni. Questo significa che per quei caratteri che possono essere misurati solo in tarda età o che sono limitati dal sesso, l'intervallo generazionale aumenta notevolmente (Fig. 5).

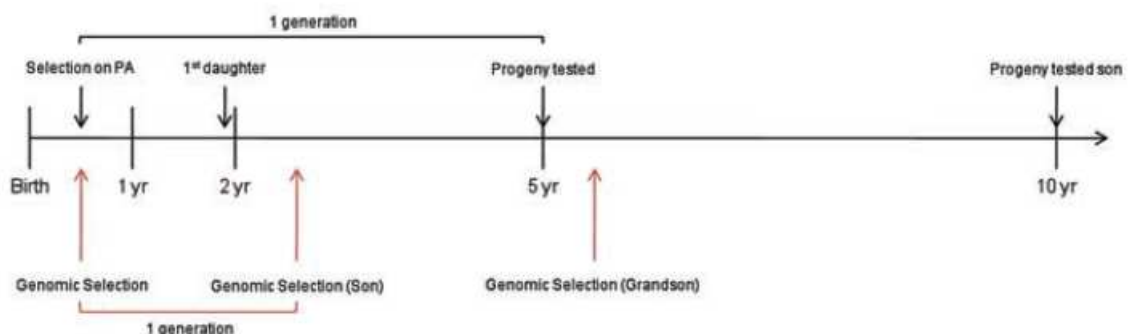


Figura 5: riduzione dell'intervallo generazionale attraverso l'uso della selezione genomica negli schemi di *progeny test* delle vacche da latte (Maltecca et al., 2010).

Come è possibile osservare nell'immagine presente in figura 5, negli schemi tradizionalmente utilizzati per la pre-selezione degli individui all'interno del *progeny test* (linea colorata in nero), l'analisi è interamente basata sul *Parental Average* (PA) e l'intervallo generazionale complessivo è influenzato dal tempo necessario per raccogliere le informazioni sulle figlie del toro in prova. Nella selezione genomica, invece (linea rossa), si trae beneficio dalle equazioni predittive sviluppate a partire da una popolazione di riferimento, riuscendo in questo modo a ridurre l'intervallo generazionale di circa 3 anni. Tutto questo è reso possibile dal fatto che la selezione genomica offre l'opportunità di fare affidamento su di una popolazione di riferimento in cui gli individui possiedono sia informazioni genotipiche che fenotipiche. Inoltre, con questo metodo, si è in grado di utilizzare le equazioni predittive sviluppate con la popolazione di riferimento per fare uno *screening* dei candidati prima che si siano registrati dati produttivi attraverso il *progeny test*. Tutto questo permette quindi di ridurre notevolmente l'intervallo generazionale e di ottenere un maggior progresso genetico (Schaeffer, 2006). Non meno importante è il beneficio ottenibile a livello economico. Con la sostituzione del *progeny test* con la selezione genomica nei programmi di selezione dei bovini da latte, si può infatti ottenere un guadagno genetico annuale e una maggior efficienza a livello economico. Una tradizionale prova su di un toro con 100 figlie, può infatti arrivare a costare all'allevatore fino a 20.000€, dovuti alle spese di nutrizione e avviamento dell'animale al *progeny test*. A questo costo iniziale vanno inoltre aggiunti 5.000€ del torello scelto sulla base del punteggio dei genitori, arrivando quindi ad una spesa di circa 25.000€. Nel caso in cui, invece, il torello sia sottoposto ad un programma misto in cui il tradizionale *progeny test* è affiancato dalla genomica, le informazioni relative al suo genotipo saranno disponibili precocemente e saranno già stati attribuiti i valori di incrocio. In questo modo il torello che verrà ritenuto il migliore potrà essere subito utilizzato per il *progeny test* ad un costo di 11.000€ (a cui verranno aggiunti 125€ per la genotipizzazione) (Tab. 3).

Tabella 3: fattori che influenzano i costi della selezione dei torelli (Täubert et al., 2011).

Source	per bull
Testing costs (from calf to 5 years)	20,000.00 €
purchase bull calf (CS)	5,000.00 €
genotyping per animal	125.00 €
purchase best genomic bull calves	11,000.00 €
purchase 2nd best genomic bull calves	6,000.00 €

La spesa si riduce ulteriormente nel caso in cui si decidesse di adottare solo la selezione genomica. Il costo finale sarebbe infatti di 282.500 € (500 genotipizzazioni x 125€ + 20 tori scelti x 11.000€) contro i 2.050.000€ della selezione tradizionale (100 tori testati x 20.000€ del costo del *test* + 10 tori scelti x 5.000€) e di 1.212.500€ del metodo misto (500 genotipizzazioni x 125€ + 10 tori selezionati come prima scelta x 11.000€ + 40 tori selezionati come seconda scelta x 6.000€ + 40 *progeny test* x 20.000€) (Täubert et al., 2011). Anche dal punto di vista del profitto si registrano enormi differenze. Con la selezione genomica è infatti possibile avere un ritorno economico tre volte superiore alle tradizionali metodiche di selezione. Vantaggi simili possono essere ottenuti con l'attuazione del modello misto (*progeny test* + genotipizzazione), che pur essendo meno vantaggioso della selezione genomica, apporta comunque maggiori guadagni rispetto al solo *progeny test* (Tab. 4).

Tabella 4: costi, ricavi e profitti in € dei tre diversi programmi relativi ad una sola bovina (Täubert et al., 2011).

Breeding program	CS	GS	MS
Return per generation	247.05	533.68	411.60
Return per year	49.13	155.25	88.52
Costs	8.42	1.13	4.32
Profit per generation	247.46	532.55	407.28
Profit per year	47.46	155.26	87.59

CS = conventional system, GS = genomic system, MS = mixed system

Altro aspetto da considerare al fine economico è l'incremento dell'accuratezza degli indici genomici e dell'intensità della selezione dei tori nei *Genomic Breeding Program* (GBP). Köning e collaboratori (2009) hanno infatti constatato come il profitto può aumentare fino al 50% nel caso in cui l'accuratezza degli indici genomici utilizzati salisse da 0,70 (valore basale) a 0,99 (Fig. 6).

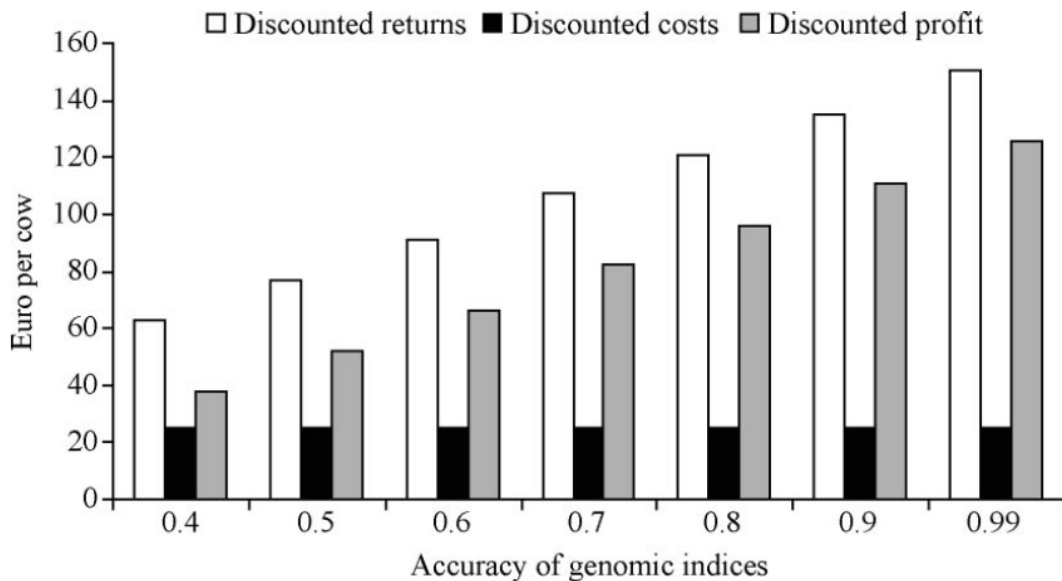


Figura 6: costi, ricavi e guadagni attuando un programma di selezione genomica per giovani tori variando l'accuratezza degli indici genomici (Köning et al., 2009).

Com'è possibile osservare dal grafico, un'accuratezza intorno allo 0,4 in un programma di selezione genomica (GBP) porta ad un ritorno economico molto simile a quello ottenibile con i tradizionali programmi di selezione. Al contrario, se l'accuratezza degli indici genomici utilizzati supera lo 0.7 il guadagno supera nettamente la spesa iniziale (Köning et al., 2009). Swalve e Köning (2007) hanno constatato come le variazioni di accuratezza dipendono moltissimo dalla qualità degli *Estimated Breeding Values* (EBV) convenzionalmente calcolati con il modello animale BLUP. Inoltre, da un recente studio di Habier e collaboratori (2007), è emerso come sia fondamentale tenere sempre in grande considerazione il contributo che il *Linkage Disequilibrium* (LD) fornisce all'accuratezza del *Genomic Estimated Breeding Value* (GEBV), in quanto si è visto che una riduzione dell'accuratezza del GEBV delle generazioni future è associato ad una diminuzione del LD.

Se da un lato l'avvento delle tecniche molecolari e la nascita della selezione genomica hanno contribuito a portare molte innovazioni e vantaggi al settore del miglioramento animale, dall'altro vi sono degli aspetti negativi. Innanzitutto per poter arrivare ad un'accuratezza del valore genomico tra il 50 e il 70% il numero di tori da provare è molto alto. È stato infatti stimato dai Paesi che implementano i loro programmi con la selezione genomica, che siano necessari almeno 16.000 tori per riuscire a minimizzare l'errore di stima. Purtroppo ad oggi nessun Paese da solo è in grado di costruire una popolazione di queste dimensioni ed è per questo che si stanno sviluppando accordi internazionali per sostenere questi progetti di ricerca (Canavesi, 2010). Altri limiti

riguardano l'identificazione dei QTL e la loro applicazione nella selezione assistita da marcatori. Nel caso di QTL associati a caratteri a bassa ereditabilità non vi sono ancora *panels* di marcatori abbastanza fitti da poter consentire l'identificazione di questi *loci*. Inoltre, sebbene siano stati condotti molti studi riguardo l'incorporazione di QTL che interagiscono nell'espressione dei caratteri di interesse, l'approccio ad oggi utilizzato risulta poco flessibile a causa della complessità dei modelli di interazione che considerano più di pochi geni alla volta. Infine, come in molti altri programmi di selezione assistita da marcatori, la porzione di genoma messa in evidenza dai *markers* è solo una minima parte della variazione totale presente. Questo dipende dagli effetti causati dal QTL e dal numero di marcatori utilizzati (Goddard e Hayes, 2009). Un'altra grande difficoltà sembra essere associata alla gestione del gran numero di risultati che la selezione genomica permette di ottenere. Infatti la possibilità di combinare diverse fonti di informazione in base al loro costo e alla loro correlazione con il genotipo, rende il lavoro di interpretazione dei dati molto complesso e comporta l'ottenimento di una grande varietà di schemi tra cui scegliere.

Le nuove sfide dei metodi di valutazione genetica quantitativa riguardano senza dubbio la possibilità di integrare le informazioni fornite dalla genetica molecolare (GM) con quelle sino ad ora disponibili nel campo della genetica quantitativa (GQ). L'identificazione di effetti QTL e più in generale la selezione assistita da marcatori, rappresenta solo una delle tantissime sinergie che potranno essere create. Inoltre, per quanto riguarda la selezione della Frisone Italiana, l'*Associazione Italiana Allevatori Frisone Italiana* (ANAFI) sta cercando di ampliare e diffondere tra gli allevatori l'uso della metodica molecolare a fini selettivi in modo che essi possano trarre benefici dalla scelta più mirata delle vacche su cui investire e sulla maggior accuratezza dei dati relativi ai tori in prova di progenie (Canavesi, 2009).

1.3 Caratteri economicamente importanti nella selezione dei bovini da latte

Come precedentemente descritto nel paragrafo 1.2, i grandi progressi compiuti negli ultimi decenni dalla genetica molecolare hanno portato all'identificazione di regioni genomiche in grado di influenzare l'espressione di caratteri quantitativi. Attraverso l'uso di marcatori molecolari, come per esempio gli SNP, è stato possibile identificare e mappare numerosi QTL associati a caratteri economicamente importanti nei bovini da latte.

I *loci* finora identificati possono essere suddivisi in tre diverse categorie: produzione del latte, qualità del latte e resistenza alle patologie.

Per quanto riguarda la scoperta ed il mappaggio dei QTL associati alla produzione del latte, gli studi iniziarono nel 1995 con il lavoro di Georges e collaboratori e da allora sono state ottenute numerose informazioni consultabili presso il database dell'*animal genome* (<http://www.animalgenome.org/QTLdb/cattle.html>). Grazie a questi studi è stato possibile constatare come all'interno del cromosoma 6 (BTA6, *Bos taurus* 6) vi siano due regioni distinte in grado di modulare l'espressione di caratteri coinvolti nella produzione del latte. All'interno di questo cromosoma è stato per esempio identificato il *cluster* delle caseine composto da quattro geni: *CSN1S1* (α_{s1} -caseina), *CSN2* (β -caseina), *CSN1S2* (α_{s2} -caseina) e *CSN3* (k-caseina) (Farrell et al., 2004). Queste molecole, secrete dalle cellule epiteliali della mammella durante la lattazione, costituiscono una grande fonte di sali minerali e amminoacidi (Larson, 1979) e rappresentano più del 95% delle proteine totali contenute nel latte bovino. Numerosi studi sono stati condotti al fine di analizzare come polimorfismi a carico dei geni codificanti le caseine possano influire sulla produzione e qualità del latte. Da essi è emerso come polimorfismi a carico del *locus CSN1S1* abbiano effetti significativi non solo sul contenuto di caseina, ma anche sulla quantità totale di proteine, grasso e sulle proprietà tecnologiche del latte. Inoltre sembra che anche polimorfismi a carico del gene per la k-caseina possano avere importanti risvolti su altri caratteri legati alla produzione del latte (Hayes et al., 2006). Sempre sul cromosoma 6, Cohen-Zinder e collaboratori (2005) scoprirono la presenza di uno SNP nella sequenza del gene *ABCG2* (*ATP-Binding cassette sub-family G*). Tale modifica, che comporta l'espressione di una tirosina al posto di una serina, si è visto essere associata ad una diminuzione della produzione di latte e ad un incremento della concentrazione di grasso e proteine. Inoltre, la proteina codificata dal gene *ABCG2* è in grado di trasportare xenobiotici e sostanze citotossiche attraverso la membrana plasmatica

(Litman et al., 2000). Per questo motivo è ragionevolmente plausibile che questo gene possa essere responsabile del trasporto del colesterolo nel latte. Un altro carattere essenziale per la produzione di latte è associato al gene per la Prolattina (*PRL*). Questo ormone riveste infatti un ruolo chiave nel mantenimento della lattazione e la sua soppressione può causare una drastica riduzione nella produzione di latte (Lacasse et al., 2012).

Altrettanto fondamentali per la produzione del latte sono i geni *LEP*, *LEPR*, *GHR* e *DGATI*. Il gene *LEP*, localizzato sul cromosoma 4 di *Bos taurus* (BTA4), è responsabile dell'espressione dell'ormone proteico Leptina. Tale molecola viene secreta dalle cellule del tessuto adiposo bianco ed è coinvolta nei processi metabolici di trasformazione del cibo in energia (Lagonigro et al., 2003). Inoltre Buchanan e collaboratori (2002) hanno rilevato che questo gene potrebbe incidere sulla manifestazione di altri caratteri come la produzione di latte, il bilancio energetico e la riproduzione. Il recettore per la Leptina è codificato dal gene *LEPR* situato sul cromosoma 3 (BTA3). Liefers e collaboratori (2004) costatarono in un loro studio come la presenza di polimorfismi nella sequenza di questo gene potessero portare variazioni nella concentrazione di Leptina nelle bovine gravide. Anche polimorfismi a carico del recettore dell'ormone della crescita (*GHR*) sono stati associati alla produzione del latte. Questo gene, situato sul cromosoma 20 (BTA20), controlla infatti le funzioni del recettore dell'ormone della crescita, una molecola che gioca un ruolo chiave nello sviluppo della ghiandola mammaria e nella produzione del latte (Bauman et al., 1985). Per quanto riguarda il gene *DGATI*, situato sul cromosoma 14 (BTA14), esso codifica l'enzima *Acetil-CoA diacilglicerolo aciltransferasi* che catalizza l'ultimo passaggio della sintesi dei trigliceridi (Grisart et al., 2002). Bovine prive di questo gene hanno manifestato malfunzionamenti nella sintesi dei trigliceridi, con conseguente riduzione o assenza della produzione di latte (Smith et al., 2000). Inoltre dallo studio condotto da Banos e collaboratori (2008) su bovine del Regno Unito, è emerso come polimorfismi nelle sequenze dei geni *GHR* e *DGATI* siano positivamente associati al bilancio energetico. In particolar modo gli animali dotati della variante allelica in cui è presente la mutazione, possiedono un via più efficiente di conversione energetica che consente loro di non intaccare le riserve presenti nel corpo, ma di utilizzare l'energia fornita dall'alimentazione per la sintesi del latte.

Un altro costituente fondamentale per una buona qualità del latte è il grasso. A questo proposito sono stati condotti numerosi studi volti all'identificazione di QTL associati alla sintesi di trigliceridi e acidi grassi del latte, ma ad oggi sono ancora pochi i *loci* di cui si conosce in modo certo l'associazione con la produzione di questi costituenti. Uno

di questi è il QTL identificato mediante analisi di associazione da Morris e collaboratori (2006). Essi riuscirono ad individuare sul cromosoma 19 del genoma bovino un *locus* in grado di influenzare la composizione di alcuni importanti acidi grassi del latte ed identificarono il *Fatty Acid Synthase (FASN)* come gene candidato del QTL. Da questo gene viene codificato un enzima multifunzionale che riveste un ruolo chiave nel metabolismo lipidico andando ad influenzare sia la composizione di acidi grassi nel tessuto adiposo dell'animale che nel contenuto di trigliceridi nel latte. Allo stesso modo anche i geni *PLIN (Perilipin)*, *LPAAT (Lysophosphatidic Acid Acyltransferase)* e *ACACA (Acetyl-CoA Carboxylase)* sono stati citati in diversi lavori per il loro ruolo fondamentale nella biosintesi di acidi grassi e controllo della composizione lipidica del latte (Matsumoto et al., 2012; Sadri et al., 2010; Mistry et al., 2002).

Al pari dei *loci* che influenzano la qualità e sintesi del latte, altrettanto importanti sono quelli che contengono geni in grado di influenzare la resistenza dell'animale alle patologie. In particolar modo la presenza di particolari varianti alleliche relative alle sequenze di geni come *CCL (Chemokine Ligand)*, *CCR (Chemokine Receptor)*, *FGF (Fibroblast Growth Factor)* e *STATs (Signal Transducers and Activators of Transcription)* sono state associate in diversi studi a funzioni di immunoregolazione e di maggior resistenza alle patologie. (Song et al., 2012; Widdison et al., 2010; Kuijk et al., 2012; Filsikowski et al., 2004). Da questi studi è infatti emerso come una grande espressione di questi geni a livello della ghiandola mammaria possa notevolmente ridurre l'insorgere di patologie quali mastiti e altre infezioni batteriche e di conseguenza, una considerevole riduzione del numero di cellule somatiche nel latte (Widdison et al., 2010) (Tab. 5).

Tabella 5: informazioni relative ai 65 SNP risultanti polimorfici tra quelli presi in esame durante questo lavoro di Tesi.

<i>SNP</i>	<i>Gene</i>	<i>Sequenza</i>	<i>Cromosoma</i>	<i>CP</i>	<i>TP</i>	<i>Pathway d'azione</i>
POU1F1_nors	<i>POU1F1</i>	GATGGCAGGGAGCTTAACCC[A/C]TTGTCTTTATAAGTTTCCTG	1	35013926	-	Controllo prolattina e ormone della crescita
rs109007595	<i>POU1F1</i>	TTCCACCTTCCTGGGAGCAG[T/C]GGAAACAAAGATAATGCAAA	1	35014129	NC	Controllo prolattina e ormone della crescita
rs41923484	<i>GHR</i>	CCGGGGGGTGCCATCTTCCA[C/G]CTCCTGCCAAGGGAGGGAGA	19	48768916	M	Ormone delle crescita
rs109231659	<i>GHR</i>	TTGAATCCACAGATACAAAG[T/G]AACCCAAAATACAGAAGGCC	20	32146186	S	Regolazione livelli GH e gestione del bilancio energetico
rs109136815	<i>GHR</i>	TACATTGCCCTGGCCCCCTCA[T/C]GTCGAGGCTGAATCACACGT	20	31891078	S	Regolazione livelli GH e gestione del bilancio energetico
rs110937773	<i>FGF2</i>	ATTTGGTTAGGGCCACTCT[A/G]TAACAGCCCCATTTGACTT	17	35247491	I	Formazione del tessuto mammario
rs29004170	<i>LEP</i>	GATGTGAACAAAACGGACCC[C/G]TGTGGGACTCGGCGGAGCAC	4	93249281	S	Utilizzo energia del tessuto adiposo
rs29004508	<i>LEP</i>	AATGCGCTGTGGACCCCTGT[A/T]TCGATTCTGTGGCTTTGGC	4	93263979	M	Utilizzo energia del tessuto adiposo
rs110559656	<i>LEP</i>	TAAACGTGGCCAAACTTAAC[A/G]CTAATTAATTTATTCACCTG	4	93257549	I	Utilizzo energia del tessuto adiposo
rs110684599	<i>PRL</i>	TTTCTCTGAAGTGATAAAC[A/C]TTTCGGTATCTAACTTGACT	23	35114464	U	Regolazione di funzioni endocrine e riproduttive
rs211032652	<i>PRL</i>	TATCACCTAGTCACCGAGGT[A/G]CGGGGTATGAAAGGAGCCCC	23	35106206	M	Regolazione di funzioni endocrine e riproduttive
rs109428015	<i>PRLR</i>	TCCCTAGACTCATTTTACTC[T/C]GAAGCAAACCTCAGCTTAAGA	20	39132325	S	Regolazione di funzioni endocrine e riproduttive
rs43349286	<i>LEPR</i>	TCTTACCTTTGCTGATAGCC[T/C]GTGAGAATGTTAAATTAATA	3	80092003	M	Variazioni del Livelli di Leptina
rs43706433	<i>TLR2</i>	TTCTACATCTGAAGCAGCCT[A/G]TTTTACTCGTGGACATTCTT	17	3952556	M	Risposta immunitaria
rs43706434	<i>TLR2</i>	TTCTCATCTCCCAAATCTGC[A/G]GACCCATAAAGTAGGAAATA	17	3952732	M	Risposta immunitaria
rs8193066	<i>TLR4</i>	TCCTGGTGTCTGTGGTAGGG[A/G]TCCTAGTCTACAAGTTCTAT	8	108838612	M	Risposta immunitaria
rs43710288	<i>CARD15</i>	AATTGAGAAACTCAGCCACC[A/T]GGATACCAGACTCTTGCTCT	18	19210671	M	Riconoscimento batterico attraverso l'immunità innata
rs41572288	<i>GRLF1</i>	ACGGATGAGCAGGCACGCGG[T/C]TCTCAGTCTGCAGTGCACATA	18	54450227	S	Risposta immunitaria
rs41255714	<i>CCL2</i>	TCCTCTGCAGTTTCAAGACC[A/G]TCCTGGGCAAGGAGTTATGT	19	16233476	U	Risposta immunitaria
rs41255713	<i>CCL2</i>	TCAGATTCCACTTCGTCAAT[T/C]ATTGAGGAGGATGCCTCCCG	19	16234934	M	Risposta immunitaria
rs41257077	<i>PI</i>	ACATCCCACCAGGAAGCAGC[A/G]TGCCACAAGATTGCCCCCAA	21	59582394	S	Regolazione risposta immunitaria
rs43765462	<i>LTF</i>	GTAGATTGCGCCTGTACCT[T/G]GGCTCCCGCTACTTGACCAC	22	53538186	S	Funzioni antimicrobiche e immunostimolatorie
rs43765461	<i>LTF</i>	GTGTGGTGTGCCGTGGGACC[T/C]GAGGAGCAGAAGAAGTGCCA	22	53538807	S	Funzioni antimicrobiche e immunostimolatorie
rs41257559	<i>CCR2</i>	CTCGAAGAACATTCAAGTTA[T/C]GCCTGTGCCCTTATTTTCC	22	53613730	S	Regolazione risposta immunitaria
rs41624917	<i>PLCE1</i>	CAAGGATTAATCCTCACTAA[T/C]GTCTCCCATTCAAACCCATT	26	15383866	I	Regolazione della risposta infiammatoria
rs42890834	<i>XDH</i>	CCTGACCCACTTCAACCAGA[A/G]GCTTGAGGGGTTTCAGCGTGC	11	14191183	M	Risposta antimicrobica
rs41704669	<i>LBP</i>	GGAATCCAGGATGGTGACCT[T/C]GACGGGCACCTTGCCCTCCC	13	67875446	U	Risposta immunitaria contro mastiti
rs109686238	<i>CCL3</i>	GGGACAGTGGCCCAAGCCTC[T/C]TTTGGATGGAAGGAACTGAA	19	14673538	U	Risposta immunitaria
rs135514413	<i>ETS2</i>	GTCGCTGCTGCAGTCTGGGA[A/C]CAGCCAGTCTGCTCTGCTGG	1	152886878	M	Sviluppo placenta e funzione immunoregolativa

<i>SNP</i>	<i>Gene</i>	<i>Sequenza</i>	<i>Cromosoma</i>	<i>CP</i>	<i>TP</i>	<i>Pathway d'azione</i>
rs43705173	<i>STAT1</i>	GGAAGTAAACTTTACAAATT [T/C]ATGAGTAGTATCTTCCATTT	19	79888611	M	Regolazione sintesi latte e risposte immunitarie
rs43706906	<i>STAT1</i>	GGACATCCTATAGGGAAAGA [C/G]AGCACGGATCATAAATGGAA	2	79923716	I	Regolazione sintesi latte e risposte immunitarie
rs137182814	<i>STAT5A</i>	ATCATCTCAGCCCTGGTGAC [C/G]AGGTGACTCCTGGCCACGCC	19	43045807	S	Regolazione sintesi latte e risposte immunitarie
rs109578101	<i>STAT5A</i>	ATTCCTGCCAGAGCCTTAGT [T/C]TGCTCAGCAGCTGTTTGAAT	19	43054393	NC	Regolazione sintesi latte e risposte immunitarie
rs133669403	<i>PPARGC1A</i>	TTGAACGTGACGCGCACAGG [A/G]GCGAATTTCCGGTGTGTGCGG	6	44875315	M	Aumento della percentuale di proteine nel latte
PPARGC1A_nors	<i>PPARGC1A</i>	ATCTCCCACCTGCAAATTTCA [A/C]AAGCCTTGTCAATGGTCAAG	6	44857081	-	Aumento della percentuale di proteine nel latte
rs109579682	<i>PPARGC1A</i>	AGGTAATGATGCACGTTTCGC [T/C]TCCCTCCGTGCCCTTATCC	6	44875251	NC	Aumento della percentuale di proteine nel latte
rs41577868	<i>ABCG2</i>	GTACGATCTCAAAAACCTTGG [T/G]GTTTTCAACTGGTAAACTCA	6	37983812	I	Sintesi del latte
rs110930452	<i>SPP1</i>	CAATTATGTATCTGTATTTA [T/C]TGGATCATTGAAATTTGGTA	20	30708512	M	Regolazione della lattazione
rs110981354	<i>CSN1S1</i>	CCTATGATAGGAGTGAATCA [C/G]GTAAGTGTGTGTCTGTCTGT	6	87155366	M	Proteine del latte
rs109299401	<i>CSN2</i>	CTTTCCTTCAGCCTGAAGTA [A/C]TGGGAGTCTCCAAAGTGAAG	6	87181542	M	Proteine del latte
rs43703011	<i>CSN2</i>	TCCCTTCCCTGGGCCATCC [A/C]TAACAGCCTCCCACAAAACA	6	87181619	M	Proteine del latte
rs43703013	<i>CSN2</i>	GTTGAGCCCTTACTGAAAG [C/G]CAGAGCCTGACTCTCACTGA	6	87181453	M	Proteine del latte
rs43703010	<i>CSN3</i>	CCAGCCAACCTACCATGGCAC [A/G]TCACCACACCCACATTTAT	6	87390576	M	Proteine del latte
rs43703015	<i>CSN3</i>	TGAGCCTACAAGTACACCTA [T/C]CATCGAAGCAGTAGAGAGCA	6	87390576	M	Proteine del latte
rs43703017	<i>CSN3</i>	CTTCTCCAGAAGTTATTGAG [A/G]GCCACCTGAGATCAACACA	6	87390632	M	Proteine del latte
rs134390757	<i>LXR-ALPHA</i>	TCTCTCTTTCGGGGTTCAGA [T/C]TCTGCAGTGGAGCTGTGGGA	15	78324597	S	Sintesi acidi grassi
rs137457402	<i>LPIN1</i>	AGGCTCTAATCCTATTTAAG [T/G]TTCTTGTTTTGCCTAATTTAG	11	86056573	I	Sintesi acidi grassi
rs136905033	<i>LPIN1</i>	TCACAAACCATGGGCCTTCT [T/C]ATCTTCAGAGAAATCCACA	11	86129986	U	Sintesi acidi grassi
rs43706495	<i>BTN1A1</i>	AGATGCAGAGCTGCCCTGTC [A/G]TCTGTCCCCAACGTGAGCG	23	31363023	NS	Regolazione formazione globuli di grasso nel latte
rs43349286	<i>LPAAT</i>	TCTTACCTTTGCTGATAGCC [A/G]GTGAGAATGTTAAATTAATA	3	80092003	M	Biosintesi dei trigliceridi
rs41608610	<i>DGKG</i>	TAAGTCACTTTGGCAACTTA [T/C]GGAGATAAAAAACAGAAAG	1	81589478	I	Metabolismo lipidico
rs110137537	<i>LIPE</i>	TGCACTTTCCAAGACGAAGA [A/C]ATCAACCTGGCCAGTGGCTT	18	51214707	NC	Controllo della lipolisi
rs109363579	<i>PLIN</i>	GGTAGGTAGGCGGTTTTCAGCT [C/G]TTCTTGCGCAGCTGGCTGTA	21	21504391	M	Sintesi acidi grassi
rs132839139	<i>ADRB2</i>	AGGATTAAGATGGATTTT [A/G]GAGTCAAACAGAACTGAGTT	7	62220606	NC	Controllo lipolisi
rs110562092	<i>ACACA</i>	GGAGCAAAACAGAAATTTCAA [A/G]GGAATATACAAAATCACCAG	19	13887927	I	Composizione acidi grassi nel latte
rs110757796	<i>FABP4</i>	CTCTCAGGACTCCTGAAGAC [A/T]CAGCTGAGGACACTTGCCAC	14	46835065	M	Controllo del metabolismo degli acidi grassi
rs110590698	<i>LPL</i>	GGAGTTGGCCCTCCCAGCC [T/C]GGTGTGACGGAGAGGACCA	8	67487606	M	Composizione acidi grassi nel latte
rs110454169	<i>AGPAT6</i>	GGAGTTGGCCCTCCCAGCC [T/C]GGTGTGACGGAGAGGACCA	27	36212557	NC	Regolazione sintesi acidi grassi del latte
rs109913786	<i>AGPAT6</i>	CCCAGTTTGGAGCTCTCGGA [T/C]ATTTTCTATTTCTGCCGAA	27	36220692	M	Regolazione sintesi acidi grassi del latte

rs133629324	<i>ORLI</i>	AGGTCTCTGATCTCATAAAG[A/C]AACAGCAAGCAAATATTACT	5	100247877	M	Produzione acidi grassi del latte
rs135588030	<i>ORLI</i>	CAGGAATTCATCCAGCAAAC[A/G]ATTGCCCATTCAGTTTCCC	5	100253752	M	Produzione acidi grassi del latte
rs41624761	<i>PLCBI</i>	AACTTGCTGAACTTTCTTAT[T/C]GAACTTGCTAGAGCGCAGAA	13	1655502	NC	Biosintesi acidi grassi
rs110270855	<i>PLCBI</i>	TAAAATGGCATTATTGCCCC[T/C]AAGACATCAGTATACAAAAT	13	1278678	NC	Biosintesi acidi grassi
rs41255693	<i>SCD-1</i>	TCTGGTTTCCCTGGGAGCTG[T/C]GGGTAAGTCAGCAGTCCACA	26	21144708	M	Sintesi acidi grassi del tessuto adiposo e steroli del latte
rs136334180	<i>SCD-1</i>	CAAGGCTAATTCTAGATAAC[A/G]GCAGGACAGGCTCAAAGTTG	26	21149234	NC	Sintesi acidi grassi del tessuto adiposo e steroli del latte

CP: Posizione Cromosomica secondo l'UMD 3.1 *Bos taurus* assembly

TP: Tipo di Polimorfismo, **S:** Sinonimo, **M:** Missenso, **I:** Introne, **U:** 5'-UTR, **NC:** Non Coding, **NS:** Non Senso

HEW: probabilità che lo SNP si trovi in equilibrio di Hardy & Weinberg

2. Scopo della Tesi

Negli ultimi vent'anni i progressi compiuti dalla biologia molecolare hanno permesso di operare numerose scoperte nell'ambito della selezione animale. A partire dai lavori di Meuwissen e collaboratori (2001) si iniziò a capire che l'espressione di caratteri economicamente importanti per il miglioramento dei bovini da latte non è influenzata da un singolo gene, ma da gruppi di geni situati su specifici *loci* cromosomici detti *loci* ad azione quantitativa o QTL.

Attraverso l'ausilio di marcatori molecolari come SNP e SSR, sono stati condotti numerosi studi finalizzati alla rilevazione ed identificazione di questi QTL, allo scopo di intraprendere una selezione genomica assistita da marcatori molecolari (MAS, *Marker Assisted Selection*) o di implementare le attuali metodiche di selezione genomica, che permetterebbero una riduzione di tempi e costi e una maggiore accuratezza nella selezione.

In questo ambito di ricerca si inserisce l'obiettivo del presente lavoro di Tesi. La principale finalità è stata infatti quella di valutare l'associazione di dati genotipici, ottenuti mediante *SNP genotyping* condotto su campioni di DNA prelevato da seme bovino, e le informazioni fenotipiche registrate attraverso i convenzionali *progeny* e *performance test*, al fine di stabilire se vi siano dei presupposti per avviare programmi di selezione assistita da marcatori molecolari nella scelta e nel miglioramento dei tori da riproduzione di razza Frisona.

3. Materiali e metodi

Un totale di 424 campioni di seme bovino sono stati consegnati dall'ente Intermizoo all'interno di *paillettes* di volume compreso tra 0,5 e 1 mL. I campioni sono stati conservati durante il trasporto in azoto liquido e, una volta giunti in laboratorio, sono stati trasferiti in provette *ependorf* da 1,5 mL e conservati a -20°C fino all'estrazione del DNA.

3.1 Estrazione e purificazione del DNA genomico

3.1.1 Estrazione del DNA genomico dai campioni di seme mediante *kit DNeasy® Blood & Tissue Kit* (QIAGEN® Group)

Dai campioni provenienti da Intermizoo e conservati a -20°C, sono stati prelevati 100µl di seme e posti in una *ependorf* da 1,5mL. A ciascuna provetta sono stati poi addizionati 100µl di *Buffer X2* (TrisHCl 20mM pH 8.0, EDTA 20mM pH 8.0, NaCl 200mM, SDS 4%; subito prima dell'uso sono stati aggiunti DTT (Ditiotritolo) 80mM e 12.5µl/mL di QIAGEN® proteinasi K). I campioni sono stati lasciati in incubazione a 56°C nel bagnetto termostato per circa un'ora per favorire la lisi cellulare. Successivamente ai campioni sono stati aggiunti 200µl di *Buffer AL* e, dopo aver agitato, sono stati addizionati 200µl di etanolo al 96%. In seguito, i campioni sono stati trasferiti nelle *DNeasy Mini spin column* (QIAGEN® Group) e centrifugati per 1 minuto a 6000xg. Dopo aver eliminato il liquido filtrato dalle colonnine ed aver addizionato 500µl di *Buffer AW1*, i campioni sono stati nuovamente centrifugati per 1 minuto a 8000rpm. Il liquido filtrato è stato nuovamente eliminato e alle colonnine sono stati addizionati 500µl di *Buffer AW2*. I campioni sono stati quindi posti in centrifuga per 3 minuti a 20000xg per permettere alle membrane delle colonnine di asciugarsi. Infine il liquido filtrato è stato eliminato e, dopo aver posto le colonnine in microtubi da 2ml, ad ogni campione sono stati addizionati 100µl di *Buffer AE* e lasciati incubare a

temperatura ambiente per 1 minuto. I campioni sono stati poi centrifugati per 1 minuto a 6000xg e, per ottenere una maggior quantità di DNA estratto, è stata ripetuta l'eluizione con 100µl di *Buffer AE* e successiva centrifugazione per 1 minuto a 8000rpm. Il DNA così estratto è stato conservato a -20°C.

3.1.2 Estrazione manuale del DNA genomico dai campioni di seme

Ai campioni di seme conservati a -20°C è stato aggiunto 1mL di PBS (*Phosphate Buffered Saline*, stock da 500mL contenente 4g di NaCl, 0,1g di KCl, 0,72g di Na₂HPO₄, 0,12g di KH₂PO₄). In seguito sono stati centrifugati per 5 minuti a 5000xg in modo da ottenere un *pellet* e, dopo aver eliminato il surnatante, i campioni sono stati risospesi in 1mL di PBS e centrifugati nuovamente per 5 minuti a 5000xg. Dopo aver eliminato il surnatante, ad ogni campione sono stati aggiunti 450µl di *Semen Extraction Buffer* (TrisHCl 10mM pH 8, EDTA 10mM portato a pH 8 con NaOH, NaCl 100mM, SDS 1%) precedentemente portato alla temperatura di 60°C. In seguito, dopo aver aggiunto 50µl di DTT (0,5M) e 10µl di proteinasi K (20mg/mL) i campioni sono stati lasciati in incubazione per 12 ore a 65°C nel bagnetto termostato per favorire la lisi cellulare. Ad ogni campione sono stati aggiunti 160µl di NaCl (5M) e, dopo averli agitati per circa 2 minuti, sono stati posti in centrifuga per 10 minuti a 15000xg. Il surnatante contenente il DNA estratto è stato trasferito in nuovi tubini e addizionato di 1mL di etanolo al 96% e 2µl di glicerolo. I campioni sono stati quindi centrifugati per 10 minuti a 15000xg per permettere la precipitazione del DNA. Il *pellet* è stato fatto asciugare mediante pompa a vuoto e risospeso in un volume di 50µl di ddH₂O.

3.2 Quantificazione del DNA genomico estratto

3.2.1 Quantificazione mediante saggio fluorimetrico Qubit®

Questo metodo di quantificazione prevede l'utilizzo di intercalanti fluorescenti in grado di legarsi in modo aspecifico al DNA, generando una fluorescenza rilevabile attraverso il fluorimetro Qubit® (life technologies™, Invitrogen). Il kit che è stato utilizzato è il *Qubit® dsDNA BR Assay* con cui è possibile rilevare la quantità di DNA genomico estratto. Inizialmente è stata preparata la soluzione contenente l'intercalante fluorescente (*Qubit® dsDNA BR reagent*, 1µl per ogni campione) e il *Qubit® dsDNA BR buffer* (199µl per ogni campione). Da ogni campione di DNA genomico precedentemente estratto sono stati prelevati 5µl e posti nelle *Qubit® assay tubes*, microtubi forniti con il kit in grado di far passare il fascio luminoso del fluorimetro e permettere in questo modo al fluoroforo legato al DNA di eccitarsi ed emettere la luminescenza nel caso di ibridazione. I campioni sono stati preparati nel modo seguente:

- 5µl di campione contenente il DNA genomico estratto + 195µl di soluzione contenente il buffer e l'intercalante;
- 10µl di *Qubit® dsDNA BR standard #1* (0ng/µl) + 190µl di soluzione contenente buffer e l'intercalante;
- 10µl di *Qubit® dsDNA BR standard #2* (100ng/µl) + 190µl di soluzione contenente buffer e l'intercalante.

Prima della lettura dei campioni, il fluorimetro è stato tarato mediante i due standard precedentemente preparati e successivamente si è passati ad analizzare i campioni misurando l'ipotetica quantità di DNA presente espressa in µg/mL.

3.3 Genotyping dei campioni mediante l'utilizzo del *GoldenGate Genotyping Assay for VeraCode* (Illumina Inc.)

Dopo aver effettuato la quantificazione del DNA genomico, ogni campione è stato normalizzato ad una concentrazione finale di 50 ng/μl. Successivamente i campioni sono stati fatti seccare a 45°C e risospesi in 6μl di ddH₂O.

Il primo passaggio del *GoldenGate Genotyping Assay* prevede l'attivazione del DNA attraverso biotinilazione, in modo che i campioni di DNA genomico si possano ibridare alle particelle paramagnetiche attraverso il legame Biotina-Streptavidina. In seguito avviene la fase di ibridazione in cui gli oligonucleotidi, il buffer di ibridazione e le particelle paramagnetiche vengono combinati con il DNA.

Per ciascun *locus* in cui è presente lo SNP, il kit fornisce tre oligonucleotidi: due sono allele specifici (ASO), mentre il terzo è *locus* specifico (LSO). Il *primer* LSO contiene inoltre una sequenza di riconoscimento che ha come bersaglio un particolare tipo di *VeraCode microbead*. Durante il processo di ibridazione dei *primer*, ASO e LSO si legano al DNA genomico ibridato alle particelle paramagnetiche. Vengono quindi effettuati numerosi lavaggi in modo da rimuovere le particelle in eccesso e gli oligonucleotidi non ibridati. I prodotti ottenuti dall'estensione dei *primer* ASO vengono quindi ligati a quelli generati dall'amplificazione di LSO, permettono di ottenere appropriate informazioni sia sul genotipo presente nel sito SNP che sulla sequenza di riconoscimento fornita da LSO. I prodotti della ligazione serviranno poi come templati per i *primer Forward* universali P1 (marcato con fluoroforo Cy3-) , P2 (marcato con fluoroforo Cy5-) e il *primer* universale *Reverse* P3 durante la successiva reazione di PCR. I filamenti marcati così ottenuti vengono quindi ibridati alla biglia complementare per mezzo del codice identificativo, permettendo di leggere in ogni campione il genotipo relativo al singolo SNP. Per identificare il codice delle *microbeads* e rilevare il segnale luminoso si utilizza il *BeadXpress*[®] *Reader* attraverso il quale, per mezzo di un laser, è possibile scansionare le biglie e generare l'immagine del codice di cui sono dotate, identificandole rapidamente e con grande specificità. Questo è possibile in quanto il codice viene ottenuto mediante un processo di incisione olografica. Quando la biglia viene colpita dal laser, l'immagine viene spezzettata in molte parti che costituiscono il codice identificativo della *microbead* che verrà letto dal *BeadXpress*[®] *Reader* (Fig. 7).

FIGURE 1: THE VERACODE BEAD INSCRIBING PROCESS

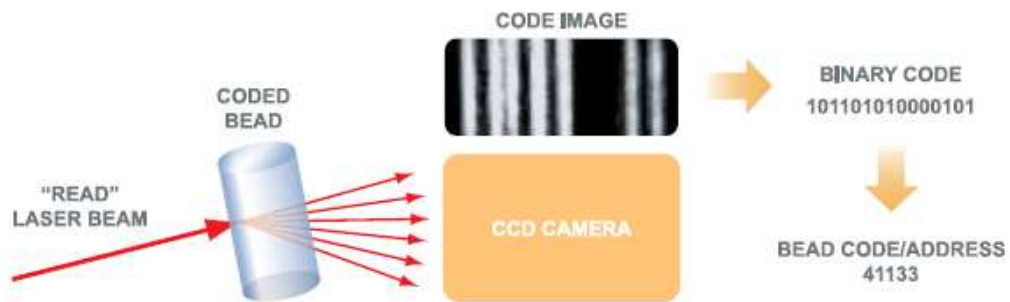


Figura 7: meccanismo di diffrazione dell'immagine olografica presente sulle bigliette e successiva lettura del codice identificativo.

3.3.1 Preparazione dei campioni

Sono stati posti 5 μ l del reagente MS1 e 5 μ l di DNA in ciascun dei 96 pozzetto di una piastra *Single-Use DNA* (SUD). Successivamente la piastra è stata posta in centrifuga a 250xg e agitata mediante vortex per 20 secondi a 2300 rpm. Dopo aver nuovamente centrifugato a 250xg, la piastra è stata posta ad incubare per 30 minuti a 95°C e nuovamente centrifugata a 250xg.

3.3.2 Precipitazione

A ciascuno dei 96 pozzetti della *SUD Plate* sono stati aggiunti 5 μ l del reagente PS1 (*Precipitation Solution*) e, dopo aver centrifugato la piastra a 250xg e agitato mediante vortex per 20 secondi a 2300 rpm, a ciascun pozzetto sono stato aggiunti 15 μ l di isopropanolo per poter permettere la precipitazione del DNA. La *SUD Plate* è stata nuovamente agitata mediante vortex per 20 secondi a 1600rpm e centrifugata per 20 minuti a 3000xg. Il surnatante che si è ottenuto è stato rimosso e, per garantire che tutto il liquido fosse stato eliminato, la piastra è stata capovolta su carta assorbente e posta in

centrifuga per 1 minuto a 8xg. Successivamente, la *SUD Plate* è stata posta ad asciugare a temperatura ambiente per 15 minuti e conservata a -20°C *over night*.

3.3.3 Allelic Specific Extension (ASE)

A ciascun pozzetto sono stati aggiunti 10µl di reagente RS1 (*Resuspension Solution*) e dopo aver centrifugato a 250xg e agitato la piastra mediante vortex per 1 minuto a 2300 rpm, è stata effettuata l'*Allelic Specific Extension*. 10µl di ciascun campione di DNA biotinilato sono stati trasferiti su di una nuova piastra (*ASE Plate*) e a ciascun pozzetto sono stati aggiunti 10µl di OPA (*Oligo Pool Assay*) e 30µl di OB1 (*Oligo Hybridization and Binding Buffer*). Dopo aver centrifugato a 250xg e agitato mediante vortex per 1 minuto a 1600rpm, la piastra è stata posta nell'incubatore impostato a 30°C e mantenuta a questa temperatura per 2 ore.

3.3.4 Estensione e ligazione

L'*ASE Plate* è stata centrifugata a 250xg e posta su di una piastra magnetica per circa 2 minuti, in modo da poter rendere più semplice la successiva rimozione del liquido contenuto nei pozzetti (circa 50µl) senza rischiare di aspirare anche le bigliette con il DNA ibridato. In seguito, sono stati aggiunti 50µl di AM1 (*Wash-Away Reagent*) al fine di rimuovere tutti gli oligonucleotidi aspecifici e rendere accessibile il DNA che invece si è ibridato. La piastra è stata quindi agitata mediante vortex per 20 secondi a 1600 rpm e posta nuovamente sul magnete per 2 minuti. Successivamente sono stati rimossi i 50µl di AM1 ed è stato ripetuto il lavaggio con questo reagente. Sono stati poi aggiunti 50µl di UB1 (*Universal Buffer*) e, dopo aver posto la piastra per 2 minuti sul magnete, è stato rimosso il liquido presente. Questo passaggio è stato ripetuto una seconda volta. A ciascun pozzetto sono stati in seguito aggiunti 37µl di MEL (*Reagent for Extension and Ligation*) e, dopo aver agitato per mezzo di un vortex per 1 minuto a 1600rpm, la piastra è stata posta in incubazione a 45°C per 15 minuti.

3.3.5 Polimerase Chain Reaction (PCR)

64µl di *Taq Polimerasi* e 50µl di *Uracil-DNA-Glicosilasi* sono stati posti all'interno di un *MMP tubes (Master Mix)*. 30µl della miscela così ottenuta sono stati addizionati a ciascun pozzetto. La piastra è stata in seguito centrifugata a 250xg e posta per 2 minuti sulla barra magnetica. Il surnatante presente nei pozzetti è stato eliminato, prestando attenzione a non rimuovere le biglie magnetiche. Sono stati poi aggiunti 50µl di UB1 e la piastra è stata nuovamente posta per 2 minuti sulla barra magnetica. Il surnatante è stato rimosso e a ciascun pozzetto sono stati addizionati 35µl di IP1. L'*ASE Plate* è stata posta sotto agitazione mediante l'uso di un vortex per 1 minuto a 1800rpm e successivamente incubata per 1 minuto a 95°C. La piastra è stata quindi posta nuovamente sul magnete per 2 minuti e da ciascun pozzetto sono stati prelevati 30µl di surnatante e posti in una nuova piastra. I campioni così ottenuti sono stati collocati nel termociclatore ed è stata avviata la PCR secondo il seguente schema:

	Temperatura	Tempo
	37°C	10 minuti
	95°C	3 minuti
x 34	95°C	35 secondi
	56°C	35 secondi
	72°C	2 minuti
	72°C	10 minuti
	4°C	5 minuti

I prodotti PCR sono stati centrifugati a 250xg e addizionati con 20µl di MPB (*Magnetic Particle Reagent*) in modo da poter legare i frammenti *double strand* ottenuti dalla reazione. I campioni sono stati risospesi e successivamente trasferiti su di una nuova piastra dotata di filtri. La procedura è stata ripetuta più volte in modo che il contenuto di ciascun pozzetto fosse stato trasferito completamente sulla nuova piastra. I campioni così ottenuti sono stati lasciati in incubazione per 60 minuti a temperatura ambiente.

3.3.6 Preparazione della piastra intermedia (*INT Plate*)

La piastra dotata di filtri, precedentemente utilizzata, è stata posta su di una piastra a 96 pozzetti *V-bottom plate* come riportato in figura (Fig. 8).



Figura 8: procedura di assemblaggio della *INT Plate*

La *INT Plate* così ottenuta è stata centrifugata a 1000xg per 5 minuti alla temperatura di 25°C. Successivamente sono stati addizionati 50µl di UB2 (*Universal Buffer*) a ciascun pozzetto della piastra e, dopo aver nuovamente centrifugato per 5 minuti a 1000xg e alla temperatura di 25°C, è stata rimossa la piastra dotata di filtri. Ai campioni filtrati sono stati addizionati 30µl di MH2 e la *filter plate* è stata riposizionata prestando grande attenzione affinché i pozzetti di entrambe coincidessero (A1 con A1). 30µl di NaOH 0,1M sono stati quindi addizionati a ciascun campione della piastra dotata di filtro al fine di normalizzare il pH della soluzione. In seguito, dopo aver centrifugato per 5 minuti a 1000xg e alla temperatura di 25°C, il filtro è stato rimosso.

3.3.7 Ibridazione e lettura della *VeraCode Bead Plate*

3mL del reagente MH2 sono stati posti in una *falcon* da 15mL insieme a 3mL di NaOH 0,1M (proporzione 1:1). 50µl della miscela così ottenuta sono stati addizionati a ciascun pozzetto della *INT Plate*. Sono stati trasferiti 100µl di ciascun campione in una *VeraCode Bead Plate* e posti nell'incubatore ad agitazione (*VeraCode Vortex Incubator*) per 3 ore a 850 rpm e alla temperatura di 45°C. Successivamente, dopo aver centrifugato la piastra a 250xg, sono stati aggiunti ad ogni pozzetto 200µl di VW1 (*Vera*

Wash Reagent). La piastra è stata quindi agitata e dopo aver atteso 2 minuti affinché le bigliette magnetiche si fossero depositate sul fondo del pozzetto, è stato aspirato il surnatante. Questi passaggi sono stati ripetuti per 7 volte. È stata quindi effettuata la lettura della *VeraCode Bead Plate* mediante il *BeadExpress® Reader*.

3.4 Raccolta dei dati fenotipici

Durante questo lavoro di tesi, per le valutazioni fenotipiche, sono stati presi in considerazione i dati ottenuti, a partire da ottobre 2011, da 238.475 campioni di latte individuale provenienti dalle 16.990 figlie dei tori testati. Per ciascuno di questi è stato valutato il contenuto di proteine, grasso, RCT (*Rennet Coagulation Time*) e a30 (consistenza del coagulo dopo 30 minuti dall'aggiunta del caglio) ricavati mediante l'utilizzo del Milko-Scan FT6000 (Foss Electric A/S). La predizione delle proprietà coagulative del latte (*Milk Coagulation Properties* o MCP) è stata effettuata, in accordo con quanto riportato da De Marchi e collaboratori (2012), attraverso la spettroscopia del medio infrarosso (*Mid-infrared spectroscopy*, MIR). Il numero di cellule somatiche presenti nel latte (SCC o *Somatic Cell Count*) è stato determinato per mezzo del Cell Fossomatic 250 (Foss Electric A/S) e trasformato in *Somatic Cell Score* (SCS) mediante la formula:

$$SCS = \frac{3 + \log_2 SCC}{100}$$

Sono stati inoltre registrati i dati relativi alla produzione giornaliera di latte, proteine e grasso in chilogrammi e il rapporto tra caseine:proteine (C : P) e grasso:proteine (G : P). In questo studio, i valori ricavati dal calcolo di C : P e G : P sono stati inseriti come possibili caratteri associati all'incremento dell'accuratezza delle stime genetiche per il MCP.

3.5 Analisi statistica dei dati

Per questa analisi, condotta dal Dott. Francesco Tiezzi, delle 16.990 vacche totali sono state considerate le 16.089 di cui erano noti il padre e la madre; sono stati inoltre tenuti in considerazione valori di RCT fino a 30 minuti. In seguito all'*editing*, sono stati presi in considerazione 63.470 dei 238.475 *records* fenotipici. Queste informazioni sono state utilizzate per il calcolo delle varianze secondo quanto riportato nel lavoro di Tiezzi e collaboratori (2013). Sono stati impiegati modelli univariati e bivariati, implementati in un contesto Bayesiano, mediante l'utilizzo del *software* GIBBS3F90 (consultabile *online* sul sito: <http://nce.ads.uga.edu/%7Eignacy/programs.html>) e la formula utilizzata per il calcolo è stata:

$$y = Xb + Z_h h + Z_p p + Z_a a + e$$

dove:

y: vettore dei *records* fenotipici per i caratteri analizzati (RCT, a30, quantità di latte, grasso, proteine, caseine, % di grasso, % di proteine, % di caseine, SCS, rapporto caseine:proteine e proteine:grasso);

b: vettore degli effetti fissi legati al parto suddiviso in 3 classi (1 parto, 2 parti, da 3 a 9 parti) e allo stadio di lattazione suddiviso in 12 classi mensili (1^a classe: da 6 a 35 giorni; 2^a classe: da 36 a 65 giorni; 3^a classe: da 66 a 95 giorni; 4^a classe: da 96 a 125 giorni; 5^a classe: da 126 a 155 giorni; 6^a classe: da 156 a 185 giorni; 7^a classe: da 186 a 215 giorni; 8^a classe: da 216 a 245 giorni; 9^a classe: da 246 a 275 giorni; 10^a classe: da 276 a 305 giorni; 11^a classe: da 306 a 335 giorni; 12^a classe: da 336 a 365 giorni);

h: vettore delle soluzioni per gli effetti casuali legati all'*Heard Test Date* (giorno di campionamento aziendale, HTD);

p: vettore delle soluzioni per l'effetto ambientale fisso a cui è soggetta la vacca;

a: vettore delle soluzioni per l'effetto genetico additivo dell'animale;

e: vettore dei residui casuali;

X, *Z_h*, *Z_p*, *Z_a*: matrici di incidenza di ordine appropriato.

Successivamente, dai valori così ottenuti, è stato possibile effettuare la stima degli EBV (*Estimated Breeding Value*). Per poter ottenere un dato indicativo del contributo fornito dal toro alla manifestazione dei caratteri fenotipici delle figlie, gli EBV

precedentemente ottenuti sono stati de-regressi secondo la formula riportata da Garrick e collaboratori (2009).

Infine, sono stati considerati gli EBV de-regressi relativi a 399 tori aventi una media di figlie pari a 41,86 (minimo 4 figlie massimo 642). Tali valori sono stati utilizzati per l'analisi di associazione condotta mediante il seguente modello:

$$dEBV = \mu + SNP + EGA + \varepsilon$$

dove:

dEBV: EBV de-regressi;

μ : media degli EBV de-regressi;

SNP: effetto fisso dello SNP;

EGA: Effetto Genetico Additivo, ossia il contributo genico che viene trasmesso alla progenie;

ε : errore residuo

4. Risultati

4.1 Elaborazione dei dati mediante il programma *GenomeStudio Data Analysis* (Illumina Inc)

I dati ottenuti dalla genotipizzazione effettuata mediante il *GoldenGate Genotyping Assay for VeraCode* sono stati analizzati per mezzo del software *GenomStudio Data Analysis* (Fig. 9). Tramite l'utilizzo di questo programma sono state normalizzate le intensità di fluorescenza relative ai singoli SNP ed è stato quindi effettuato il *genotyping* (Fig. 9).

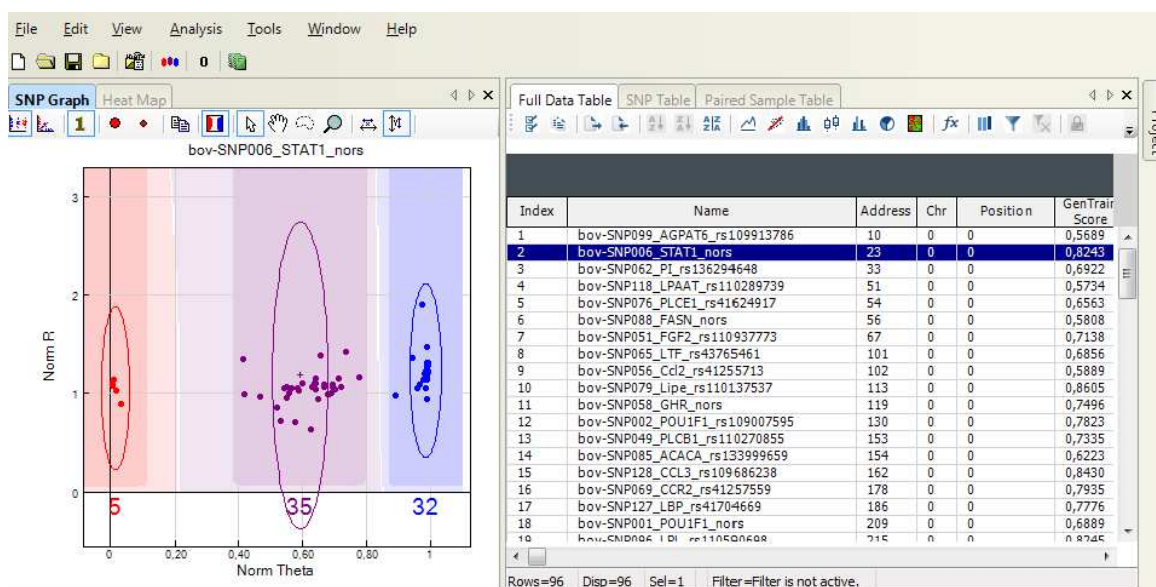


Figura 9: esempio di *cluster* relativo al gene STAT1 ottenuto mediante l'utilizzo del programma *GenomeStudio Analysis*

La designazione delle basi dell'allele relative al polimorfismo analizzato sono state eseguite seguendo il metodo TOP/BOT fornito da Illumina insieme al *GoldenGate Genotyping Assay for VeraCode*. Secondo questa metodica la situazione più semplice si verifica quando una delle possibili variazioni portate dallo SNP è rappresentata da un'adenina (A) e le restanti variazioni possono essere una citosina (C) o una guanina (G). In questo caso, la sequenza dello SNP viene designata come TOP e il nucleotide A viene nominato allele A. Allo stesso modo, i nucleotidi G o C rappresentano l'allele B. Nel caso in cui invece una delle possibili varianti dello SNP sia un timina (T) e le

restanti possibilità siano o una guanina (G) o una citosina (C), la sequenza dello SNP viene designata come BOT e l'allele A corrisponde al nucleotide T. L'allele B risulta invece essere associato alla presenza di G o di C (Fig. 10).

TABLE 1: STRAND AND ALLELE DESIGNATIONS FOR UNAMBIGUOUS SNPS

SNP Name ¹	Sequence ²	Strand Designation	Allele A	Allele B
rs363040	...AGGAGGCTAG[G/T]CTCGCAGAGC...	BOT	T	G
rs536477	...GAGATTTAGG[A/G]AAAGATGTGA...	TOP	A	G
rs684517	...ACCAGGTACT[C/T]TGAAC TTAC...	BOT	T	C
rs2034107	...CATCTCCCC[A/C]AAATCAGTTT...	TOP	A	C

¹SNP name and sequence are from dbSNP version 126.
²Sequence shortened to 10 base pairs on each side of SNP as noted by the ellipses (...) for illustrative purposes.
The unambiguous pairings used to determine Strand and Allele are noted in red.

Figura 10: esempio di designazione con metodo TOP/BOT.

Tuttavia, se lo SNP è rappresentato da [A/T] o da [C/G], è necessario applicare un differente metodo. Nel caso di uno SNP di tipo [A/T], la presenza di un'adenina viene indicata come TOP, mentre se vi è una timina BOT. In entrambi i casi, però, sia in presenza di una A che di una T, l'allele viene comunque designato come allele A generando in questo modo ambiguità. Lo stesso avverrà in caso di SNP di tipo [G/C]. Per ovviare a questo inconveniente, Illumina ha utilizzato la tecnica di “*sequence walking*” in modo da poter designare l'allele sia in caso di SNP nella forma [A/T] che [G/C]. Per poter fare ciò, lo SNP viene considerato in posizione n e le sequenze immediatamente prima e dopo vengono definite rispettivamente come $n-1$ e $n+1$. Allo stesso modo viene fatto per due basi a monte ($n-2$) e due basi a valle ($n+2$), ecc. (Fig. 11).

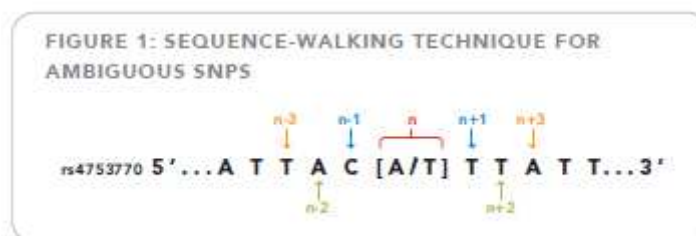


Figura 11: raffigurazione della tecnica di *sequence walking* utilizzata per SNP ambigui.

Utilizzando questo metodo, il *sequence walking* continua fino a quando non si rilevano due nucleotidi non ambigui, come per l'esempio riportato in figura 11 dove nelle posizioni $n-1$ ed $n+1$ vi è la coppia non ambigua costituita da una C da una T. A questo

punto, il filamento viene definito TOP o BOT andando a considerare se la A/T della coppia non ambigua si trova in 5' o 3' rispetto allo SNP. Nel caso dell'esempio in figura 11, il filamento è da considerarsi BOT in quanto la timina della coppia non ambigua si trova in 3' rispetto allo SNP. Per la designazione dell'allele, invece, in caso di SNP [A/T] con filamento TOP, l'allele viene designato con A in presenza di un'adenina nel sito polimorfico e B in presenza invece della timina; nel caso di filamento BOT l'allele A si ha in presenza di timina, mentre l'allele B in presenza di adenina. Lo stesso vale per SNP di tipo [G/C] in cui in caso di filamento TOP l'allele A equivale alla presenza di guanina e l'allele B di citosina, mentre nel caso di filamento BOT l'allele A equivale alla presenza di citosina e l'allele B di guanina (Fig. 12).

TABLE 2: STRAND AND ALLELE DESIGNATIONS FOR AMBIGUOUS SNPS

SNP Name ¹	Sequence ²	Strand Designation	Allele A	Allele B
rs1535632	...ACGGGGACAG[A/T]TATGTTAACT...	BOT	T	A
rs363334	...ATGAGTGAAT[C/G]AAGCACTATT...	TOP	C	G
rs7101540	...AAATTCAGAT[A/T]CAGAATCTTT...	TOP	A	T
rs7113791	...GATGGACAGG[A/T]TGACCTCTAG...	BOT	T	A
rs778833	...GGTTAAATGT[C/G]AAGGTGAGCT...	BOT	G	C
rs4933195	...ATGCTAATAA[A/T]ACATTAAAGT...	TOP	A	T
rs903997	...ATGAGAAAGT[C/G]TGAGAGTGCA...	TOP	C	G
rs1942968	...CTACATGACT[C/G]TTTATGTTAC...	BOT	G	C

¹SNP name and sequence are from dbSNP version 126
²Sequence shortened to 10 base pairs on each side of SNP as noted by the ellipses (...) for illustrative purposes.
The unambiguous pairings used to determine Strand and Allele are noted in red.

Figura 12: esempio di designazione con metodo TOP/BOT *sequence walking*.

I risultati relativi al genotipo, alle frequenze degli alleli A e B e alla probabilità che lo SNP si trovi in equilibrio Hardy & Weinberg (HWE) per ogni *locus* analizzato sono riportati in tabella 6.

Tabella 6: rappresentazione schematica dei genotipi associati ad ogni SNP, frequenze degli alleli A e B e calcolo della probabilità di equilibrio Hardy & Weinberg con relativa significatività.

SNP	GENE	GENOTIPO			FREQUENZE ALLELICHE		HWE ³
		AA	AB	BB	A	B	
POU1F1_nors	<i>POU1F1</i>	0,02	0,27	0,71	0,15	0,85	0,576
rs109007595	<i>POU1F1</i>	0,54	0,44	0,03	0,76	0,24	0,000***
rs41923484	<i>GHR</i>	0,03	0,25	0,72	0,16	0,84	0,194
rs109231659	<i>GHR</i>	0,38	0,13	0,49	0,44	0,56	0,000***
rs109136815	<i>GHR</i>	0,63	0,34	0,03	0,80	0,20	0,229
rs110937773	<i>FGF2</i>	0,09	0,54	0,37	0,36	0,64	0,000***
rs29004170	<i>LEP</i>	0,16	0,65	0,19	0,49	0,51	0,000***
rs29004508	<i>LEP</i>	0,09	0,47	0,44	0,32	0,68	0,095
rs110559656	<i>LEP</i>	0,54	0,42	0,04	0,75	0,25	0,028*
rs110684599	<i>PRL</i>	0,26	0,50	0,24	0,51	0,49	1,000
rs211032652	<i>PRL</i>	0,02	0,37	0,61	0,21	0,79	0,017*
rs109428015	<i>PRLR</i>	0,00	0,28	0,72	0,14	0,86	0,000***
rs43349286	<i>LEPR</i>	0,02	0,18	0,80	0,11	0,89	0,211
rs43706433	<i>TLR2</i>	0,43	0,51	0,05	0,69	0,31	0,000***
rs43706434	<i>TLR2</i>	0,03	0,32	0,65	0,19	0,81	0,265
rs8193066	<i>TLR4</i>	0,02	0,20	0,78	0,12	0,88	0,250
rs43710288	<i>CARD15</i>	0,20	0,57	0,24	0,48	0,52	0,009**
rs41572288	<i>GRLF1</i>	0,40	0,52	0,08	0,66	0,34	0,002**
rs41255714	<i>CCL2</i>	0,18	0,62	0,21	0,49	0,51	0,000***
rs41255713	<i>CCL2</i>	0,08	0,50	0,42	0,33	0,67	0,004**
rs41257077	<i>PI</i>	0,11	0,55	0,34	0,38	0,62	0,001**
rs43765462	<i>LTF</i>	0,07	0,38	0,55	0,26	0,74	0,704
rs43765461	<i>LTF</i>	0,62	0,30	0,08	0,77	0,23	0,003**
rs41257559	<i>CCR2</i>	0,17	0,56	0,27	0,45	0,55	0,008**
rs41624917	<i>PLCE1</i>	0,18	0,44	0,39	0,40	0,60	0,084
rs42890834	<i>XDH</i>	0,35	0,53	0,12	0,61	0,39	0,018*
rs41704669	<i>LBP</i>	0,54	0,19	0,26	0,64	0,36	0,000***
rs109686238	<i>CCL3</i>	0,67	0,30	0,02	0,82	0,18	0,401
rs135514413	<i>ETS2</i>	0,91	0,08	0,01	0,95	0,05	0,004**
rs43705173	<i>STAT1</i>	0,06	0,49	0,45	0,30	0,70	0,001**
rs43706906	<i>STAT1</i>	0,27	0,58	0,15	0,56	0,44	0,000***
rs137182814	<i>STAT5A</i>	0,02	0,69	0,28	0,37	0,63	0,000***
rs109578101	<i>STAT5A</i>	0,33	0,52	0,15	0,59	0,41	0,192
rs133669403	<i>PPARGCIA</i>	0,02	0,14	0,84	0,09	0,91	0,035*
PPARGC1A_nors	<i>PPARGCIA</i>	0,90	0,07	0,04	0,93	0,07	0,000***
rs109579682	<i>PPARGCIA</i>	0,04	0,36	0,60	0,22	0,78	0,317
rs41577868	<i>ABCG2</i>	0,30	0,58	0,12	0,59	0,41	0,000***
rs110930452	<i>SPP1</i>	0,20	0,50	0,30	0,45	0,55	0,922
rs110981354	<i>CSN1S1</i>	0,01	0,16	0,83	0,09	0,91	0,763
rs109299401	<i>CSN2</i>	0,88	0,11	0,01	0,93	0,07	0,024*
rs43703011	<i>CSN2</i>	0,13	0,49	0,38	0,37	0,63	0,351
rs43703013	<i>CSN2</i>	0,71	0,23	0,06	0,82	0,18	0,000***

³ L'equilibrio di Hardy & Weinberg è definito quando sono rispettati cinque requisiti: la popolazione in esame deve essere ideale, cioè di dimensioni infinite; al suo interno devono avvenire incroci casuali (Panmissia) e non deve essere soggetta all'azione di forze evolutive come selezione naturale, mutazioni e migrazioni.

rs43703010	<i>CSN3</i>	0,04	0,59	0,37	0,34	0,66	0,000***
rs43703015	<i>CSN3</i>	0,10	0,45	0,45	0,32	0,68	0,658
rs43703017	<i>CSN3</i>	0,58	0,33	0,09	0,75	0,25	0,007**
rs134390757	<i>LXR-ALPHA</i>	0,21	0,62	0,17	0,52	0,48	0,000***
rs137457402	<i>LPIN1</i>	0,03	0,35	0,62	0,21	0,79	0,239
rs136905033	<i>LPIN1</i>	0,02	0,16	0,83	0,09	0,91	0,081
rs43706495	<i>BTN1A1</i>	0,04	0,32	0,63	0,20	0,80	0,882
rs43349286	<i>LPAAT</i>	0,00	0,17	0,82	0,09	0,91	0,559
rs41608610	<i>DGKG</i>	0,54	0,45	0,01	0,76	0,24	0,000***
rs110137537	<i>LIPE</i>	0,11	0,57	0,32	0,39	0,61	0,000***
rs109363579	<i>PLIN</i>	0,15	0,48	0,37	0,39	0,61	1,000
rs132839139	<i>ADRB2</i>	0,80	0,19	0,01	0,90	0,10	0,601
rs110562092	<i>ACACA</i>	0,41	0,52	0,07	0,67	0,33	0,000***
rs110757796	<i>FABP4</i>	0,22	0,59	0,20	0,51	0,49	0,000***
rs110590698	<i>LPL</i>	0,88	0,11	0,01	0,94	0,06	0,218
rs110454169	<i>AGPAT6</i>	0,04	0,34	0,61	0,22	0,78	1,000
rs109913786	<i>AGPAT6</i>	0,03	0,32	0,65	0,19	0,81	0,642
rs133629324	<i>ORL1</i>	0,86	0,13	0,01	0,92	0,08	0,001**
rs135588030	<i>ORL1</i>	0,52	0,44	0,04	0,74	0,26	0,002**
rs41624761	<i>PLCB1</i>	0,37	0,53	0,11	0,63	0,37	0,007**
rs110270855	<i>PLCB1</i>	0,54	0,43	0,03	0,76	0,24	0,001***
rs41255693	<i>SCD-1</i>	0,07	0,34	0,59	0,24	0,76	0,230
rs136334180	<i>SCD-1</i>	0,08	0,51	0,41	0,33	0,67	0,004**

*: $P < a 0,05$

** : $P < a 0,01$

***: $P < a 0,001$

Com'è possibile osservare in tabella 6, 14 geni hanno presentato frequenza dell'allele A o B inferiore a 0,15. Questa soglia è indicativa soprattutto nel caso in cui la frequenza si riferisce all'allele favorevole. Infatti, per poter ottenere un incremento significativo di tale variante, al fine di aumentarne la presenza nella popolazione in esame, sarebbero necessari molto tempo e numerosi incroci mirati al miglioramento di solo quel carattere. È inoltre possibile osservare come 36 *loci* QTL dei 65 totali non si sono presentati in equilibrio HWE. L'equilibrio di Hardy e Weinberg si presenta quando vengono rispettati cinque requisiti:

- la popolazione in esame deve essere ideale, cioè di grandi dimensioni;
- gli incroci tra individui devono essere casuali (Panmissia);
- assenza di selezione naturale;
- assenza di mutazioni;
- assenza di migrazioni.

Nel caso in cui uno o più di questi requisiti non venga rispettato, la popolazione in esame non si trova in equilibrio HWE. Nella popolazione presa in esame nel seguente

lavoro di Tesi era attesa l'assenza di equilibrio HWE in quanto gli animali presenti sono sottoposti ad incroci mirati e a programmi di selezioni volti al miglioramento di specifici caratteri.

4.2 Analisi genetica

In seguito ad una ricerca bibliografica preliminare è stato possibile individuare 33 geni candidati responsabili della manifestazione dei caratteri di interesse. Per ciascuno di questi geni è stata condotta un'analisi sulle banche dati pubbliche (per esempio dbSNP) volta a rilevare la presenza di polimorfismi di tipo SNP localizzati nella sequenza codificante e/o nella regione regolatrice e/o intronica (Tab. 5 riportata nel paragrafo 1.3 dell'introduzione). 96 SNP appartenenti ai 33 geni candidati putativamente responsabili dei caratteri presi in considerazione (qualità, quantità e attitudine casearia del latte) sono stati inseriti nel *panel* di marcatori analizzati. Di questi, 3 SNP non hanno superato le soglie di qualità riportate dal *software GenomeStudio*. Dei restanti 93 SNP, 28 sono stati scartati in quanto aventi MAF (*Minor Allele Frequency*) inferiore al 5%⁴. I risultati riportati in tabella 5, presente nel paragrafo 1.3 dell'introduzione, sono quindi riferiti ai restanti 65 SNP più polimorfici.

4.3 Elaborazione dei dati fenotipici

Per quanto riguarda il quantitativo in chilogrammi di latte prodotto dalle bovine di razza Frisona in esame, il dato registrato si è dimostrato conforme alle aspettative. La media delle bovine provate si è infatti attestata intorno ai 29,81 kg/giorno, con un valore di mediana di 29,40 kg/giorno. Allo stesso modo anche il contenuto percentuale di gras-

⁴ Il *Minor Allele Frequency* o MAF rappresenta la frequenza con cui, nella popolazione in esame, si presenta l'allele più piccolo. Essa viene calcolata come frequenza dell'allele minore diviso il totale dei soggetti testati e rappresenta un parametro spesso utilizzato negli studi di *Genome Wide Association* (GWA) per caratteri complessi in quanto permette di discriminare se uno SNP è effettivamente polimorfico all'interno della popolazione saggiata. Per essere tale il valore MAF deve superare una certa soglia che può variare dal 5% all'1%.

so e proteina si è dimostrato conforme ai dati presenti in letteratura per la razza Frisona, con un valori medi di 3,82% per il grasso e di 3,45% per le proteine. Tale dato è stato inoltre confermato dal rapporto proteina/grasso (0,93%) indice di un buon funzionamento del metabolismo ruminale. Il quantitativo percentuale della caseina presente nei campioni esaminati (2,66%) indica una buona attitudine casearia, dato confermato anche dal rapporto caseina/proteine espresso in percentuale che, essendo pari al 0,78%, si è dimostrato leggermente superiore ai dati normalmente registrati nel latte delle bovine di razza Frisona (circa 0,75%).

La stima del contenuto di cellule somatiche nel latte (SCS), considerata un parametro fondamentale per la valutazione dello stato sanitario della mammella nonché della resa alla caseificazione del latte e della durata del periodo di conservazione dei prodotti lattiero-caseari, è stata calcolata per mezzo del parametro SCC, dotato di distribuzione normale ed analizzabile mediante un modello lineare. Per mezzo di questo valore è stato possibile evitare l'incidenza di errori dovuti all'eventuale presenza di dati anomali. Per poter interpretare le informazioni ottenute è stato necessario tenere in considerazione che:

- un SCS pari a 2 corrisponde ad un SCC equivalente a 50.000 cellule/mL;
- un SCS pari a 3 equivale ad un SCC pari a 100.000 cellule/mL;
- un SCS pari a 4 corrisponde ad un SCC pari a 200.000 cellule/mL.

Alla luce di queste considerazioni è possibile affermare come le bovine testate godessero di ottima salute in quanto il valore medio di SCS rilevato è stato di 3,11 corrispondente a circa 100.000 cellule/mL (Tab.6). Il valore infatti ritenuto convenzionalmente indicativo di un latte di ottima qualità è ≤ 100.000 cellule/mL (Reg. CE 853/2004).

L'attitudine casearia viene convenzionalmente valutata mediante l'uso del lattodinamografo, uno strumento in grado di valutare il cambiamento di consistenza del latte dopo l'aggiunta del caglio, producendo dei tracciati a forma di campana detti lattodinamogrammi. Grazie a questi tracciati è possibile ottenere importanti informazioni riguardo la qualità del latte quali:

- RCT o tempo di coagulazione: si misura in minuti ed è rappresentato dal tempo che intercorre dall'aggiunta del caglio e l'inizio del processo di coagulazione, in cui si rileva una variazione nella consistenza del campione;

- k20 o tempo di rassodamento: rappresenta il tempo necessario al coagulo per raggiungere una consistenza tale da determinare un'ampiezza del tracciato di 20 mm;
- a30 o consistenza del coagulo: si misura in millimetri e corrisponde all'ampiezza del tracciato raggiunta dopo 30 minuti dall'aggiunta del caglio (Fig. 13).



Figura 13: esempio di tracciato lattodinamografico

Secondo la classificazione formulata da Zannoni e Annibaldi (1981), un latte con attitudine casearia ottimale dovrebbe avere:

- RCT, minuti: 11'30'' - 18';
- k20, minuti: 5'30'' - 8'30'';
- a30, millimetri: 20 – 40.

Nel presente lavoro di Tesi, la valutazione delle proprietà coagulative del latte è stata condotta mediante l'utilizzo della spettroscopia al medio infrarosso (*Mid-Infrared Spectroscopy*, MIR), secondo quanto riportato da De Marchi e collaboratori (2012). Nel loro lavoro essi hanno infatti riscontrato come la metodica MIR possa essere utilizzata per questo tipo di valutazione, ottenendo una stima di accuratezza dei coefficienti di *cross-validation* pari al 70% per a30 e al 76% per RCT. Inoltre, l'utilizzo del MIR nella predizione di caratteri legati al latte, come la composizione di acidi grassi (De Marchi et al., 2011) e le proprietà coagulative (Dal Zotto et al., 2008; De Marchi et al., 2009), è stato proposto come valido metodo in grado di superare le limitazioni che si sono

presentate con l'uso del *computerized renneting meter* (Cassandro et al., 2008), del Formagraph (Cipolat-Gotet et al., 2012) e dell'Optigraph (Vallas et al., 2010). Tali metodiche non possono essere infatti applicate a livello di popolazione e per questo motivo non sono utili per scopi riproduttivi (Tiezzi et al., 2013).

I risultati ottenuti hanno dimostrato un *Rennet Coagulation Time* (RTC) intorno ai 22 minuti, mentre la consistenza media del coagulo a 30 minuti (a30) è stata di circa 24 millimetri. In tabella 7 sono indicate le statistiche descrittive dei caratteri oggetto di studio.

Tabella 7: informazioni relative ai caratteri qualitativi, quantitativi e all'attitudine casearia del latte registrati mediante rilevazione fenotipica (238.475 campionamenti) delle figlie (n = 16.990) dei tori presi in esame. **SD:** Deviazione Standard; **CV:** Coefficiente di variazione; **5° P e 95°P:** 5° e 95° percentile; **Skewness:** corrisponde all' indice di asimmetria con cui è possibile misurare di quanto la distribuzione di un valore non è simmetrica; **Kurtosis:** è un parametro statistico con cui è possibile descrivere l'andamento della curva della distribuzione di probabilità; **t:** produzione di latte totale; **SCS:** *Somatic Cell Score*; **RCT:** *Rennet Coagulation Time* o tempo di coagulazione; **a30:** consistenza del coagulo dopo 30 minuti dall'aggiunta del caglio.

<i>Carattere</i>	<i>Media</i>	<i>Mediana</i>	<i>SD</i>	<i>CV</i>	<i>5°P</i>	<i>95°P</i>	<i>Skewness</i>	<i>Kurtosis</i>
<i>Latte, t</i>	29,818	29,400	9,394	0,315	12,400	49,400	0,233	-0,010
<i>Grasso, kg</i>	1,113	1,083	0,360	0,324	0,484	1,934	0,627	1,028
<i>Proteine, kg</i>	0,998	0,994	0,277	0,278	0,458	1,559	0,100	0,077
<i>Caseine, kg</i>	0,778	0,776	0,217	0,279	0,353	1,215	0,080	0,071
<i>Grasso, %</i>	3,815	3,770	0,777	0,204	2,350	5,500	0,441	1,247
<i>Proteine, %</i>	3,414	3,380	0,419	0,123	2,710	4,350	0,564	0,527
<i>Caseine, %</i>	2,663	2,630	0,346	0,130	2,070	3,420	0,495	0,454
<i>SCS, n°/mL</i>	3,114	2,880	1,944	0,624	0,057	7,554	0,600	0,145
<i>RCT, min</i>	21,626	21,250	5,099	0,236	12,500	33,070	0,495	0,883
<i>a30, mm</i>	24,086	23,510	10,688	0,444	4,630	46,960	0,423	0,423
<i>Caseine/proteine, %</i>	0,783	77,914	1,980	0,025	74,050	82,337	0,116	0,907
<i>Proteine/grasso, %</i>	0,930	90,343	19,162	0,207	62,009	138,525	1,860	8,497

Per quanto riguarda la correlazione dei dati fenotipici (Tab. 8) le informazioni ottenute si sono presentate tutte significativamente associate tra loro. Il quantitativo di grasso e proteina si è dimostrato positivamente correlato alla produzione di latte (0,780 per il grasso e 0,930 per le proteine); allo stesso modo, il contenuto di caseine presente nei campioni è positivamente correlato al contenuto di proteine (0,996) e al quantitativo di latte (0,919). Al contrario, il contenuto di SCS ha presentato valori di associazione negativi con il quantitativo di latte (-0,251), grasso (-0,169) e proteine (-0,206). RCT e a30 hanno evidenziato una correlazione negativa pari a -0,694, mentre il carattere a30 ha evidenziato una limitata correlazione positiva con il contenuto percentuale di grasso (0,160), proteine (0,259) e caseine (0,270). Il carattere RCT, ha riportato una limitata correlazione positiva con la percentuale di proteine (0,113) e negativa con il quantitativo di latte (-0,161).

Tabella 8: correlazioni dei dati fenotipici. *t*: produzione di latte totale; *SCS*: Somatic Cell Score; *RCT*: Rennet Coagulation Time; *a30*: consistenza del coagulo dopo 30 minuti dall'aggiunta del caglio.

	<i>Latte, t</i>	<i>Grasso, kg</i>	<i>Proteine, kg</i>	<i>Caseina, kg</i>	<i>Grasso, %</i>	<i>Proteine, %</i>	<i>Caseine, %</i>	<i>SCS, n°/mL</i>	<i>RCT, min</i>	<i>a30, mm</i>	<i>Caseine/proteine, %</i>
<i>Grasso, kg</i>	0,780***										
<i>Proteine, kg</i>	0,930***	0,792***									
<i>Caseina, kg</i>	0,919***	0,776***	0,996***								
<i>Grasso, %</i>	-0,337***	0,290***	-0,211***	-0,217***							
<i>Proteine, %</i>	-0,503***	-0,246***	-0,176***	-0,158***	0,451***						
<i>Caseine, %</i>	-0,486***	-0,251***	-0,161***	-0,128***	0,414***	0,982***					
<i>SCS, n°/mL</i>	-0,251***	-0,169***	-0,206***	-0,217***	0,134***	0,208***	0,169***				
<i>RCT, min</i>	-0,161***	-0,151***	-0,135***	-0,139***	0,007***	0,113***	0,092***	0,234***			
<i>a30, mm</i>	-0,024***	0,065***	0,076***	0,086***	0,160***	0,259***	0,270***	-0,196***	-0,694***		
<i>Caseine/proteine, %</i>	-0,052***	-0,103***	0,027***	0,114***	-0,074***	0,193***	0,373***	-0,140***	-0,070***	0,127***	
<i>Grasso/proteine, %</i>	0,060***	-0,452***	0,122***	0,139***	-0,768***	0,117***	0,146***	-0,016***	0,047***	-0,011***	0,191***

*: $P < 0,05$ **: $P < 0,01$ ***: $P < 0,001$

4.4 Analisi di associazione

Attraverso l'utilizzo degli EBV de-regressi è stato possibile stimare l'associazione tra SNP e caratteri fenotipici presi in esame. I risultati, ottenuti dall'elaborazione statistica effettuata dal Dott. Francesco Tiezzi, sono riportati in tabella 9. I dati fenotipici e gli effetti di sostituzione allelica sono stati moltiplicati per le seguenti costanti per una più facile gestione dell'analisi e lettura dei risultati:

- i dati relativi al latte espresso in kg sono stati moltiplicati per 10;
- i dati relativi ai caratteri qualitativi del latte (grasso, proteine e caseine espressi in kg e in percentuale, SCS e rapporto caseine/proteine %) sono stati moltiplicati per 100;
- i dati ottenuti dal rapporto tra proteine e grasso % sono stati moltiplicati per 1;
- i dati relativi alle proprietà coagulative del latte (RCT e a30) sono stati moltiplicati per 10.

L'effetto di sostituzione allelica (allele A vs allele B) descritto per gli SNP analizzati è riferito all'incremento/decremento medio di unità di misura del carattere nella popolazione.

Com'è possibile osservare dai dati relativi agli effetti di sostituzione allelica riportati in tabella 9, i polimorfismi che si sono presentati più significativamente associati alla produzione di latte sono stati quelli a carico del gene *GHR* (*Growth Hormone Receptor*) con valori pari a 8,47 e 6,36. Per quanto riguarda gli aspetti qualitativi, espressi in termini di percentuale di grasso e proteine, le associazioni significativamente più positive sono state quelle riportate per gli SNP presenti nei geni *PRL* (*Prolactin Receptor*, 12,09 grasso % e 4,32 proteine %), *TLR2* (*Toll Like Receptor 2*, 14,03 grasso % e 5,11 proteine %), *CCR2* (*Chemokine Receptor 2*, 10,82 grasso % e 7,16 proteine %), *STAT5A* (*Signal Transducer and Activator of Transcription 5A*, 6,68 grasso % e 6,50 proteine %), *PPARGC1A* (*Peroxisome Proliferator Activated Receptor Gamma Coactivator 1 Alpha*, 6,93 grasso % e 2,73 proteine %), *ADRB2* (*Adrenergic Receptor Beta 2*, 6,36 grasso % e 4,40 proteine %), *AGPAT6* (*Acyl-CoA:Glycerol-3-Phosphate Acyltransferase 6*, 5,30 grasso % e 3,97 proteine %), *SCD-1* (*SteroylCoA Desaturase 1*, 8,90 grasso % e 4,34 proteine %). Per quanto riguarda il contenuto di cellule somatiche solo *POU1F1* (*Pituitary-Specific-Positive Transcription Factor 1*) e *CARD15* (*Caspase Recruitment Domain-Containing Protein 15*) hanno dimostrato associazione

significativa a questo carattere riportando valori pari a -13,05 *POU1F1* e -10,40 *CARD15*. Infine, per quanto riguarda le proprietà coagulative del latte, i valori più significativamente associati a RCT sono stati quelli rilevati per *POU1F1* (-5,94), *GRLF1* (*Glucocorticoid Receptor Repression Factor 1*, -7,85) e *CCL2* (6,23). Per quanto riguarda a30, invece, la maggiore significatività è stata registrata con *GRLF1* (13,96), *CCL2* (-18,31) e *STAT5A* (16,35).

Tabella 9: effetti degli SNP analizzati sui caratteri fenotipici. **RS:** Reference Sequence dello SNP secondo quanto riportato dal database dbSNP; **SCS:** Somatic Cell Score; **RCT:** Rennet Coagulation Time; **a30:** consistenza del coagulo dopo 30 minuti dall'aggiunta del caglio; **C:P:** rapporto tra caseine e proteine espresso in percentuale; **P:G:** rapporto tra proteine e grasso espresso in percentuale.

Gene	RS	Latte, kg	Grasso, kg	Proteine, kg	Caseine, kg	Grasso, %	Proteine, %	Caseine, %	SCS, n°/mL	RCT	a30	C:P	P:G
<i>POU1F1</i>	POU1F1_nors	0,450	-0,157	-0,450	-0,180	-1,361	-1,909	-1,121	-13,047 **	-5,495 **	7,839	10,533	-0,261
<i>POU1F1</i>	rs109007595	2,019	0,837	-0,131	0,079	-0,673	-2,297	-1,896	-1,303	-5,943 ***	9,586 **	-2,779	-0,861
<i>GHR</i>	rs41923484	0,646	-2,064	-1,029	-0,712	-10,010 **	-4,325	-3,541	-7,191	-1,098	-0,887	-3,710	1,553
<i>GHR</i>	rs109231659	8,466 †	-0,024	1,088	1,066	-13,128 †	-6,010 †	-4,355 ***	0,518	-3,797 *	-5,006	6,409	1,669 **
<i>GHR</i>	rs109136815	6,356 **	0,645	1,420 *	1,074	-5,549 *	-2,539	-2,011	1,722	-3,279	-0,037	-6,239	0,750
<i>FGF2</i>	rs110937773	2,194	-0,370	0,653	0,603	-5,301 *	-0,852	-0,140	1,587	-1,660	-1,110	14,808 *	1,098 *
<i>LEP</i>	rs29004170	-0,497	-0,845	-0,721	-0,519	-2,598	-1,688	-1,310	-1,531	-3,399 *	4,201	-4,089	0,435
<i>LEP</i>	rs29004508	0,745	1,105	0,540	0,559	3,289	2,062	1,713	0,213	1,113	-3,080	-2,994	-0,375
<i>LEP</i>	rs110559656	1,007	0,385	0,019	0,060	0,403	-0,415	-0,322	-0,162	-0,825	-0,051	-3,795	-0,125
<i>PRL</i>	rs110684599	-3,818	1,391	-0,602	-0,424	12,094 ***	4,324 **	3,673 **	-4,440	-1,283	2,193	-3,241	-2,005 ***
<i>PRL</i>	rs211032652	0,972	-1,301	-0,634	-0,370	-6,287	-3,840 **	-2,657 *	6,152	-2,489	0,601	12,109	0,470
<i>PRLR</i>	rs109428015	-1,718	0,382	0,364	0,258	3,673	3,128 **	2,356 *	-2,242	2,146	-0,224	-0,304	-0,157
<i>LEPR</i>	rs43349286	0,371	-0,134	-0,092	-0,037	-0,582	0,314	0,037	4,578	2,831	-4,200	-7,601	0,053
<i>TLR2</i>	rs43706433	-1,167	0,274	-0,136	-0,043	2,644	1,232	1,283	2,308	-3,755 **	7,758 **	7,898	-0,325
<i>TLR2</i>	rs43706434	-2,469	2,309 *	0,452	0,270	14,030 †	5,114 **	3,577 **	-4,539	0,395	6,744	-18,114	-2,285 ***
<i>TLR4</i>	rs8193066	0,584	0,818	2,060	1,652 *	0,831	5,089 **	4,514 **	2,200	-3,102	12,096	19,764	1,206
<i>CARD15</i>	rs43710288	-1,858	0,230	0,260	0,165	4,243	3,659 ***	2,924 **	-7,328 *	3,603 **	-0,832	3,685	-0,118
<i>GRLF1</i>	rs41572288	3,898 *	0,629	0,314	0,393	-4,127	-3,406 **	-2,580 **	-4,106	-7,849 †	13,963 †	1,516	-0,022
<i>CCL2</i>	rs41255714	-4,419 *	-0,726	-1,351	-1,094	4,985	1,463	1,034	10,056 *	1,794	-6,630	-3,364	-1,011
<i>CCL2</i>	rs41255713	-3,614 *	-1,515 *	-1,862 ***	-1,679 ***	1,130	-1,551	-1,501	7,188	6,232 †	-18,310 †	-14,112 *	-0,668
<i>PI</i>	rs41257077	1,431	-0,051	-0,359	-0,176	-1,899	-2,625 **	-1,612	6,310	-3,105 *	-0,584	9,005	-0,221
<i>LTF</i>	rs43765462	-1,533	-0,427	-0,521	-0,397	1,226	0,323	0,660	-3,067	-2,242	5,326	10,381	-0,346
<i>LTF</i>	rs43765461	-1,485	0,565	-0,630	-0,392	3,546	-0,321	-0,243	-2,645	-1,133	-1,912	1,547	-1,135 *
<i>CCR2</i>	rs41257559	-7,317 **	0,171	-0,397	-0,400	10,822 †	7,151 †	4,825 ***	0,389	1,359	9,180 *	-15,516	-0,830
<i>PLCE1</i>	rs41624917	-0,156	-0,241	0,575	0,318	-0,193	2,422	1,622	-4,852	-0,159	2,763	-8,186	0,686
<i>XDH</i>	rs42890834	-0,069	-0,849	-0,332	-0,018	-4,475	-0,738	-0,204	-7,239	-3,246 *	3,772	12,591	0,765
<i>LBP</i>	rs41704669	-5,440 **	-2,129 **	-1,036	-0,875	-0,056	2,530	2,008	-3,939	0,308	5,446	2,538	0,459
<i>CCL3</i>	rs109686238	-3,902 *	-1,013	-0,266	-0,313	1,061	2,732 *	2,133 *	-1,374	3,479 *	-1,104	5,694	0,291
<i>ETS2</i>	rs135514413	2,856	0,774	0,975	0,838	-2,759	-1,135	-0,862	0,893	-0,419	-1,058	3,395	0,322
<i>STAT1</i>	rs43705173	2,449	1,663 *	1,460 *	1,167 **	3,988	1,859	1,977	-1,915	-1,156	2,687	15,860 *	-0,357
<i>STAT1</i>	rs43706906	-1,649	-0,452	0,153	0,196	2,660	3,888 **	3,534 ***	3,343	-0,801	3,724	12,667	0,059

<i>Gene</i>	<i>RS</i>	<i>Latte, kg</i>	<i>Grasso, kg</i>	<i>Proteine, kg</i>	<i>Caseine, kg</i>	<i>Grasso, %</i>	<i>Proteine, %</i>	<i>Caseine, %</i>	<i>SCS, n*/mL</i>	<i>RCT</i>	<i>a30</i>	<i>C:P</i>	<i>P:G</i>
<i>STAT5A</i>	rs137182814	-0,056	1,777	1,643 *	1,387 *	6,676 **	6,504 †	5,114 ***	-4,345	-2,744	16,353 ***	0,742	-0,220
<i>STAT5A</i>	rs109578101	0,799	0,636	-0,115	0,020	0,910	-0,847	-0,341	-10,399 †	-1,967 *	1,741	8,326	-0,474
<i>PPARGCIA</i>	rs133669403	-0,860	-0,698	-1,107	-0,687	-3,047	-2,806 *	-1,745	-6,820	-4,362 **	3,589	14,850 *	-0,047
<i>PPARGCIA</i>	PPARGC1A_nors	2,480	1,299	0,656	0,555	0,855	-0,390	-0,583	3,208	-2,134	4,922	-11,615	-0,257
<i>PPARGCIA</i>	rs109579682	0,557	1,879 **	0,706	0,694	6,933 **	2,730 *	2,108 *	-0,508	-0,140	3,640	-0,326	-1,192 **
<i>ABCG2</i>	rs41577868	-0,174	-0,129	-0,193	-0,307	-0,160	-0,696	-0,610	4,464	-1,315	0,583	-5,185	-0,144
<i>SPP1</i>	rs110930452	1,580	0,127	-0,084	-0,040	-1,193	-1,462	-1,107	6,770	-2,955 *	3,742	-1,808	-0,038
<i>CSN1S1</i>	rs110981354	-0,503	0,872	0,485	0,530	0,919	1,590	1,496	-4,589	-0,448	3,373	10,446	0,198
<i>CSN2</i>	rs109299401	-1,853	0,130	0,168	0,045	2,106	2,004	1,249	1,567	1,965	-0,035	-1,538	0,012
<i>CSN2</i>	rs43703011	-0,987	0,743	-0,262	-0,073	4,436 *	1,408	1,538	-3,890	-3,481 **	8,453 **	13,126 *	-0,934 *
<i>CSN2</i>	rs43703013	-3,920	-0,107	-1,193	-1,021 *	4,460	0,331	-0,039	3,801	-0,878	2,119	-11,746	-1,063 *
<i>CSN3</i>	rs43703010	-1,921	0,058	-0,227	-0,299	2,939	1,468	0,728	-5,199	-1,502	6,943 *	-11,995 *	-0,335
<i>CSN3</i>	rs43703015	0,756	-0,012	-0,447	-0,033	-3,559	-2,962	-1,846	0,286	-4,354 *	2,996	15,776	-0,343
<i>CSN3</i>	rs43703017	0,609	0,394	0,244	0,233	0,083	0,699	0,804	-1,674	-2,231	6,016	9,254	-0,156
<i>LXR-ALPHA</i>	rs134390757	3,786 *	1,661 *	1,127	1,003 *	-0,378	-0,826	-0,467	-0,612	-2,152	1,356	5,851	0,010
<i>LPIN1</i>	rs137457402	0,849	-3,172 **	-1,517	-1,044	-13,946 ***	-6,428 **	-4,481 **	5,039	-0,693	-9,689	9,450	2,217 **
<i>LPIN1</i>	rs136905033	1,202	1,166	0,210	0,121	2,895	-0,875	-0,782	-3,459	-2,812	6,141	-2,775	-1,092 *
<i>BTN1A1</i>	rs43706495	1,862	2,867 ***	0,965	0,844	9,683 ***	2,155	1,515	0,763	-3,067	10,442 **	-4,220	-1,843 ***
<i>LPAAT</i>	rs43349286	-0,878	-0,137	-1,188	-1,205	1,583	-2,093	-2,757	-11,197	3,157	-3,383	-35,968 ***	-1,326
<i>DGKG</i>	rs41608610	-2,070	0,234	-0,890	-0,590	5,342 *	1,219	0,743	0,973	-3,672 *	9,680 **	-5,313	-1,194 *
<i>LIPE</i>	rs110137537	0,237	0,858	0,047	-0,023	3,060	-0,203	-0,209	1,572	-0,437	1,791	-1,597	-0,867 *
<i>PLIN</i>	rs109363579	0,754	-0,095	0,625	0,293	0,567	1,589	0,694	1,186	3,177	-2,947	-17,123 *	0,413
<i>ADRB2</i>	rs132839139	-1,839	-0,000	0,586	0,061	6,361 *	4,399 **	2,721 *	5,232	2,402	3,016	-22,392 **	-0,232
<i>ACACA</i>	rs110562092	-2,215	-0,434	-0,635	-0,633	1,397	0,692	0,097	2,401	3,288	-5,958	-17,269 *	-0,250
<i>FABP4</i>	rs110757796	-2,182	0,233	-0,799	-0,540	3,664	0,091	0,334	-1,296	-0,001	-1,305	6,033	-0,885 *
<i>LPL</i>	rs110590698	1,091	1,457 *	0,904	0,702	4,198	2,500 *	2,094 *	3,644	-1,248	4,537	8,508	-0,433
<i>AGPAT6</i>	rs110454169	-2,622	0,384	-0,080	0,122	5,303 *	3,967 **	3,406 **	-2,515	-1,059	9,126 *	13,833	-0,505
<i>AGPAT6</i>	rs109913786	3,902	1,705	1,335	1,301 *	1,754	1,655	1,794	-1,886	0,253	2,770	20,344 *	-0,066
<i>ORL1</i>	rs133629324	-3,366	0,107	-0,827	-0,708	5,023 *	1,462	0,833	1,511	1,900	-1,706	-11,380	-1,066 *
<i>ORL1</i>	rs135588030	-2,071	-0,289	-0,627	-0,354	0,397	-0,161	0,089	-1,705	-1,296	2,423	7,166	-0,065
<i>PLCB1</i>	rs41624761	0,248	0,124	-0,256	-0,260	0,132	-1,345	-1,226	-2,898	0,774	-1,720	-3,481	-0,209
<i>PLCB1</i>	rs110270855	2,065	3,021 ***	0,884	0,724	5,822 *	-0,006	-0,631	-5,973	0,872	2,822	-17,329 *	-1,592 **
<i>SCD-1</i>	rs41255693	-4,438	0,204	-0,204	-0,512	8,895 **	4,339 **	2,638	8,979	2,715	4,331	-23,044 **	-0,967
<i>SCD-1</i>	rs136334180	-5,076 *	-0,626	-0,930	-0,964	5,482 *	2,447	1,080	1,077	0,768	7,098	-25,281 **	-0,925

*: P < 0,1; **: P < 0,05; ***: P < 0,01; †: P < 0,001

5. Discussione e Conclusioni

La razza Frisona rappresenta ad oggi la tipologia di bovini da latte più produttivi al mondo. Questo è riscontrabile anche nel nostro Paese, in cui il numero di vacche Pezzate Nere allevate sul territorio italiano è di 1.130.270 capi, con una media di 90 bovine per allevamento (dati ANAFI, 2012). La peculiarità che contraddistingue la Frisona dalle altre razze è il grande quantitativo di latte che essa riesce a produrre durante la lattazione. Tale caratteristica produttiva è stata rilevata anche nel seguente lavoro di Tesi, nel quale i *records* fenotipici ottenuti dalle 16.990 vacche testate hanno evidenziato una produzione giornaliera media di latte pari a 29,8 kg, equivalenti a circa 9.089 kg/lattazione⁵. La percentuale di grasso e proteine, parametri fondamentali per una buona qualità del latte, hanno presentato discreti valori (3,81% grasso e 3,41% proteine), come i caratteri legati all'attitudine casearia del latte (RCT 21,62 minuti e a30 24,08 millimetri) (Tab. 7). Alla base di questi dati vi sono le numerose procedure di miglioramento che, fin dal 1929, sono state messe in atto al fine di far diventare la razza Frisona Italiana la principale produttrice di latte a livello Nazionale. Inoltre, è stato possibile riscontrare come queste informazioni, affiancate da quelle ottenute dalle correlazioni tra i caratteri fenotipici esaminati (Tab. 8), abbiano confermato quanto già riscontrato tra gli altri da Penasa e collaboratori (2012), ossia che il miglioramento delle proprietà tecnologiche del latte può essere condotto senza rischiare di ostacolare la selezione per caratteri produttivi e qualitativi.

L'obiettivo alla base della selezione dei bovini da latte è il miglioramento di tutti quei caratteri che risultano essere economicamente rilevanti. Tuttavia molti di essi sono di natura quantitativa, in quanto la loro espressione viene influenzata non solo da fattori ambientali, ma anche dall'azione di molti geni (Zhang et al., 1998). La soluzione a questo problema si è avuta grazie alle innovazioni introdotte nella genetica molecolare che hanno portato alla graduale identificazione di geni o regioni cromosomiche associati a differenze fenotipiche tra individui, famiglie, razze (Maltecca et al., 2010) e, in alcuni casi, all'individuazione delle varianti alleliche responsabili delle variazioni fenotipiche. Grazie a queste innovazioni è stato possibile ottenere una maggior efficienza ed accuratezza nella selezione animale ed in particolar modo, nella scelta dei torelli da

⁵ In questo caso si intende la lattazione standard della durata di 305 giorni.

avviare alle prove di progenie e delle madri di toro (Canavesi, 2010) oltre che un vantaggio per gli allevatori, i quali hanno avuto l'opportunità di ridurre i costi, diminuire gli intervalli generazionali e ridurre l'*inbreeding* tra gli animali (Daetwyler et al., 2007). Attraverso la selezione assistita da marcatori molecolari, come per esempio gli SNP, è stato possibile procedere all'identificazione dei QTL associati ai caratteri economicamente più rilevanti per la selezione delle bovine di razza Frisona.

Nel seguente lavoro di Tesi, sulla base delle informazioni bibliografiche riportate da diversi autori, riguardo il ruolo ricoperto da alcuni geni nella manifestazione di caratteri associati alla qualità e quantità di latte prodotto dalle bovine di razza Frisona, sono stati selezionati 33 geni candidati. Per ciascuno di essi, mediante un'analisi *in silico* condotta su banche dati pubbliche come dbSNP, sono stati identificati 96 SNP localizzati nella regione codificante e/o regolatrice e/o intronica ed utilizzati come *panel* di marcatori per la genotipizzazione dei 424 tori testati. Gli SNP monomorfici e/o con MAF inferiore al 5% sono stati scartati dalle successive analisi di associazione e i dati genotipici e le informazioni fenotipiche sono state opportunamente pesate per gli effetti fissi (casuali e non) e convertite in EBV de-regressi in modo da poter stimare il contributo genetico fornito dal toro alle figlie. Da questo studio di associazione è stato possibile ricavare gli effetti di sostituzione allelica relativi a ciascuno SNP (Tab. 9). Dalle informazioni così ottenute è possibile osservare come due dei tre polimorfismi a carico del gene *GHR* abbiano presentato valori di sostituzione allelica positivamente correlati alla produzione di latte. Questo dato si è presentato in accordo con quanto affermato da Bauman e collaboratori (1985), i quali riscontrarono come questo gene giochi un ruolo chiave nello sviluppo della ghiandola mammaria e nella produzione di latte. I valori che sono stati ottenuti sono stati di 0,85, con $P < 0,001$, e 0,64 con $P < 0,05$. Questi valori indicano come, nel caso in cui il toro trasmettesse alle figlie la variante allelica positiva per la manifestazione di questo carattere, la quantità di latte che esse andrebbero a produrre sarebbe superiore di 0,85 e 0,64 kg rispetto alla media delle bovine della popolazione. Inoltre, è interessante osservare come il polimorfismo che ha presentato la maggior significatività sia anche quello con frequenze alleliche tali da poter essere ancora inserito in programmi di implementazione genica (frequenza allele A pari a 0,44 e frequenza allele B pari a 0,56).

Per quanto riguarda gli aspetti qualitativi del latte, intesi come contenuto percentuale di grasso e proteine, i valori di sostituzione allelica che si sono presentati più significativamente associati sono stati quelli riportati per uno dei polimorfismi della

Prolattina (*PRL*). Tale SNP ha infatti presentato valori di associazione con la percentuale di grasso pari a 0,12 con $P < 0,01$ e di associazione con la percentuale di proteine pari a 0,04 con $P < 0,05$. Questi dati hanno confermato quanto riportato da Lacasse e collaboratori (2011) i quali hanno riscontrato come questo ormone non sia importante solo per il mantenimento della lattazione, ma anche per la regolazione della produzione delle componenti qualitative del latte quali appunto grasso e proteine. Anche in questo caso, come per il polimorfismo più significativo del gene *GHR*, la frequenza dell'allele più favorevole potrebbe essere ulteriormente implementata in quanto i valori che si sono presentati sono stati pari a 0,51 per l'allele A e 0,49 per l'allele B. Un altro SNP i cui coefficienti di sostituzione allelica hanno evidenziato una significativa associazione con la produzione di grasso è stato quello relativo al gene *BTN1A1*. I dati ottenuti relativi a questo gene si sono dimostrati in linea con quanto descritto da Bhattacharya e collaboratori (2004), i quali constatarono nel loro lavoro come questo gene fosse responsabile della sintesi e regolazione dei globuli di grasso all'interno del latte. I valori di correlazione con la percentuale di grasso sono stati infatti altamente significativi (0,09 con $P < 0,01$) indicando come le vacche che ereditino la variante allelica contenente tale polimorfismo siano in grado di produrre un latte contenente un quantitativo percentuale di grasso superiore dello 0,09% rispetto alla media delle bovine testate. Anche uno degli SNP presenti nella sequenza del gene codificante l'enzima *SteroylCoA Desaturase 1 (SCD-1)* responsabile della biosintesi degli acidi grassi, ha presentato valori di correlazione significativi rispetto al grasso % (0,08 con $P < 0,05$) e alle proteine % (0,04 con $P < 0,05$), in accordo con quanto riportato da Alim e collaboratori (2012). Per quanto riguarda il miglioramento del contenuto di proteine e grasso, l'eventuale implementazione dello SNP presente nel gene *SCD-1* risulta difficile a causa della già elevata frequenza dell'allele favorevole (B) all'interno della popolazione saggiata (frequenza allele B 0,76; frequenza allele A 0,24).

Tra i geni associati alla risposta immunitaria e alle funzioni antimicrobiche, solo *CARD15* e *STAT5A* hanno presentato valori di sostituzione allelica favorevoli alla riduzione del numero di cellule somatiche presenti nel latte. In particolar modo, è stato possibile constatare come la presenza del polimorfismo a carico di *CARD15* possa comportare una riduzione media del numero di cellule somatiche nel latte pari a -0,07 ($P < 0,1$) sulla media del latte proveniente dalle bovine testate. Inoltre, i valori delle frequenze alleliche di questo gene sono tali per cui sarebbe possibile inserirlo in un

programma di selezione assistita da marcatori molecolari (frequenza allele A pari a 0,48; frequenza allele B pari a 0,52). Per quanto riguarda *STAT5A*, il polimorfismo ad esso associato si è dimostrato migliorativo in quanto ha presentato un valore di sostituzione allelica pari a -0,10 con $P < 0,001$. A differenza di quanto appena detto per *CARD15*, le frequenze alleliche relative al polimorfismo di *STAT5A* si sono dimostrate tali per cui un programma di implementazione potrebbe risultare difficoltoso. La frequenza dell'allele B si è infatti dimostrata molto più presente nella popolazione testata rispetto a quella dell'allele A (frequenza allele B pari a 0,63; frequenza allele A pari a 0,37).

Le proprietà coagulative del latte (intese come RCT e a30) sono state ormai da molto tempo ritenute fondamentali per incrementare la qualità del latte. Tempi di coagulazione più veloci e una maggiore consistenza del coagulo dopo 30 minuti dall'aggiunta del caglio, dovrebbero infatti comportare una produzione di formaggio con caratteristiche più desiderabili e con una minor perdita di componenti rispetto al latte privo di proprietà coagulative favorevoli (Pretto et al., 2013). Alla luce di queste informazioni, nei risultati ottenuti sono stati rilevati quattro polimorfismi da considerarsi migliorativi per le proprietà coagulative del latte. Il primo è associato al gene *POU1F1* (*Pituitary Specific Positive Transcription Factor 1*), la cui funzione è legata al controllo dei livelli dell'ormone Prolattina e del bilancio energetico (Ogorevc et al., 2009). Tale polimorfismo si è dimostrato migliorativo sia per RCT (-0,59 con $P < 0,01$) che per a30 (0,96 con $P < 0,05$), indicando che la presenza nel genoma delle bovine della variante allelica contenente lo SNP porterebbe a tempi di coagulazione del latte più rapidi (-0,59 minuti rispetto alla media delle bovine testate) e una maggiore consistenza del coagulo dopo 30 minuti dall'aggiunta del caglio (0,96 millimetri in più rispetto alla media). Il secondo polimorfismo migliorativo per le proprietà coagulative del latte è stato a carico del gene *GRLF1*. Esso ha infatti presentato valori di sostituzione allelica pari a -0,79 ($P < 0,001$) per RCT e a 1,4 ($P < 0,001$) per a30. Anche il polimorfismo a carico del gene per la β caseina (*CSN2*) ha presentato valori migliorativi, in accordo con quanto rilevato da Larson e collaboratori (1979). Dai dati ottenuti è infatti emerso come la trasmissione della variante allelica associata a questo SNP comporti nelle bovine la produzione di un latte con RCT inferiore di 0,35 minuti ($P < 0,05$) rispetto alla media delle bovine testate e un incremento di 0,85 millimetri ($P < 0,05$) del coagulo dopo 30 minuti dall'aggiunta del caglio. L'ultimo SNP che si è presentato significativamente migliorativo per le proprietà coagulative del latte è associato al gene *DGKG* (*Diacylglycerol Kinase*

Gamma), coinvolto nel metabolismo lipidico. Dai dati ottenuti esso ha infatti presentato valori di RCT pari a -0,37 ($P < 0,1$) e di a30 pari a 0,97 ($P < 0,05$).

Dai risultati ottenuti in questo lavoro di Tesi è stato possibile constatare come l'utilizzo delle informazioni relative a marcatori SNP in geni candidati sia un valido strumento per poter selezionare precocemente i migliori torelli da avviare alle prove di progenie. I valori di sostituzione allelica ottenuti, si sono dimostrati complessivamente in linea con quanto riportato in bibliografia e, nella gran parte dei casi, supportati da una buona significatività. Ovviamente questo nuovo approccio presenta ancora la necessità di confermare i dati genotipici con quelli fenotipici forniti dalle prove di progenie e deve essere inoltre valutata nel contesto dell'attuale selezione genomica, ma permette in ogni caso di aumentare l'accuratezza nella selezione dei torelli e di ridurre in questo modo i costi legati al loro mantenimento. Per quanto riguarda le prospettive future, nella selezione della Frisone Italiana, ANAFI sta cercando di ampliare e diffondere tra gli allevatori l'uso della metodica molecolare a fini selettivi in modo che essi possano trarre benefici dalla scelta più mirata delle vacche su cui investire e sulla maggior accuratezza dei dati relativi ai tori in prova di progenie (Canavesi, 2009).

6. Bibliografia

- Alim M.A., Fan Y.P., Wu X.P., Xie Y., Zhang Y., Zhang S.L., Sun D.X., Zhang Y., Zhang Q., Liu L., Guo G., (2012). “Genetic effects of stearoyl-coenzyme A desaturase (SCD) polymorphism on milk production traits in the Chinese dairy population”, *Mol. Biol.Rep.*, 39(9): 8733 – 8740;
- Animal genome, <http://www.animalgenome.org/QTLdb/cattle.html>, 2010;
- Associazione Nazionale Allevatori Frisone Italiana (ANAFI), (1993). “Un nuovo indice di selezione dal luglio ‘93”, *Bianconero*;
- Balasini D., (1995), “Zootecnica speciale. Principali razze di animali domestici e tecniche di allevamento per le diverse produzioni”, *Edagricole*;
- Banos G., Wolliams J.A., Woodward B.W., Forbes A.B., Coffey M.P., (2008). “Impact of single nucleotide polymorphisms in Leptin, Leptin Receptor, Growth Hormone Receptor and Diacylglycerol Acyltransferase (DGAT1) gene loci on milk production, feed and body energy traits of UK dairy cows”, *J. Dairy Sci.*, 91: 3190 – 3200;
- Bauman D.E., Eppard P.J., De Geeter M.G., Lanza G.M., (1985). “Response of high producing dairy cows to long term treatment with pituitary somatotropin and recombinant somatotropin”, *J. Dairy Sci.*, 68: 1352 – 1358;
- Beaumont M.A., Rannala B., (2004). “The Bayesian devolution in genetics”, *Nat. Rev. Gen.*, 5: 251 – 261;
- Bhattacharya T.K., Misra S.S., Sheikh F.D., Dayal S., Vohra V., Kumar P., Sharma A., (2004). “Variability of milk fat globule membrane protein gene between cattle and riverine buffalo”, *DNA Seq*, 15 (5 – 6): 326 – 331;

- Boichard D., Bonaiti B., Barbat A., Mattalia S., (1995). “Three methods to validate the estimation of genetic trend for dairy cattle”, *J. Dairy Sci.*, 78: 431 – 437;
- Börner V., Reinsch N., (2012). “Optimising multi stage dairy cattle breeding schemes including genomic selection using decorrelated or optimum selection indices”, *Genetic Selection Evolution*, 44:1;
- Buchanan F.C., Fitzsimmons C.J., Van Kessel A.G., Thude T.D., Winkelmann-Sim D.C., Schmutz S.M., (2002). “Association of a missense mutation in the bovine leptin gene with carcass fat content and leptin mRNA levels”, *Genet. Sel. Evol.*, 7: 111 – 122;
- Canavesi F., (2009). “La selezione della Frisona nell’era della genomica”, *Bianconero*;
- Canavesi F., (2010). “Qual è la differenza fra i tori per la riproduzione?”, *Bianconero*;
- Cassandro M., Comin A., Ojala M., Dal Zotto R., De Marchi M., Gallo L., Carnier P., Bittante G., (2008). “Genetic parameters of milk coagulation properties and their relationship with milk yield and quality traits in Italian Holstein cows”, *J. Dairy Sci.*, 91: 371 – 376;
- Cipolat-Gotet C., Cecchinato A., De Marchi M., Penasa M., Bittante G., (2012). “Comparison between mechanical and near-infrared methods for assessing coagulation properties of bovine milk”, *J. Dairy Sci.*, 95: 6806 – 6819;
- Cohen-Zinder M., Seroussi E., Larkin D.M., Looor J.J., Everts-van der Wind A., Lee J.H., Drackley J.K., Band M.R., Hernandez A.G., Shani M., Lewin H.A., Weller J.I., Ron M., (2005). “Identification of a missense mutation in the bovine ABCG2 gene with major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle”, *Genome Research*, [5:9] 6 – 944;

- Daetwyler H.D., Villanueva B., Bijma P., Woolliams J.A., (2007). “Inbreeding in genome wide selection”, *J. Anim. Breed Genet.*, 124: 369 – 376;
- Dal Zotto R., De Marchi M., Cecchinato A., Penasa M., Cassandro M., Carnier P., Gallo L., Bittante G., (2008). “Reproducibility and repeatability of measures of milk coagulation properties and predictive ability of mid-infrared reflectance spectroscopy”, *J. Dairy Sci.*, 91: 4130 – 4112;
- De Marchi M., Fagan C.C., O’Donnell C.P., Cecchinato A., Dal Zotto R., Cassandro M., Penasa M., Bittante G., (2009). “Prediction of coagulation properties, titratable acidity, and pH of bovine milk using mid-infrared spectroscopy”, *J. Dairy Sci.*, 92: 423 – 432;
- De Marchi M., Penasa M., Cecchinato A., Mele M., Secchiari P., Bittante G., (2001). “Effectiveness of mid-infrared spectroscopy to predict fatty acid composition of Brown Swiss bovine milk”, *Animal*, 5: 1653 – 1658;
- De Marchi M., Penasa M., Tiezzi F., Toffanin V., Cassandro M., (2012). “Prediction of milk coagulation proprieties by Fourier transform mid-infrared spectroscopy”, *J. Dairy Sci.*, 96: 4707 – 4715;
- Dekkers J.C., (2004). Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons, *J. Anim. Sci.*, 82: E313 – E328;
- Farrell H.M. Jr., Jimenez-Flores R., Bleck G.T., Brown E.M., Butler J.E., Creamer L.K., Hicks C.L., Hollar C.M., Ng-Kwai-Hang K.F., Swaisgood H.E., (2004). “Nomenclature of the proteins of cows’ milk – sixth revision”, *J. Dairy Sci.*, 87: 1641 – 1674;
- Filsikowski K., Stralkowska N., Sloniewski K., Krzyżewki J., Zwierzchowski L., (2004). “Association of a sequence nucleotide polymorphism in exon 16 of the STAT5A gene with milk production traits in Polish Black-and-White (Polish Fresian) cows”, *Anim. Sci. and Reports*, 4: 515 – 522;

- Fusco R., (1990), “La Frisone italiana. Evoluzione, lotte e traguardi di cinque generazioni di allevatori”, *Edagricole*;
- Gianola D., Foulley J.L., (1983). “Sire evaluation for ordered categorical-data with a threshold-model”, *Genet. Sel. Evol.*, 15: 201 – 223;
- Gianola D., Fernando R.L., (1986). “Bayesian methods in animal breeding theory”, *J. Anim. Sci.*, 63: 217 – 244;
- Georges M., Nielsen D., Mackinnon M., Mishra A., Okimoto R., (1995). “Mapping quantitative trait loci controlling milk production in dairy cattle by employing progeny testing”, *Genetics*, 139: 907 – 920;
- Goddard M.E., Hayes B.J., (2009). “Mapping genes for complex traits in domestic animals and their use in breeding programs”, *Nat. Rev. Genet.*, 10: 381-391;
- Grisart B., Coppieters W., Farnir F., Karim L., Ford C., Berzi P., Cambisano N., Mini N., Reid S., Simon P., et al., (2002). “Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition”, *Genome Res.*, 12: 222 – 231;
- Habier D., Fernando R.L., Dekkers J.C.M., (2007). “The impact of genetic relationship information on genome-assisted breeding values”, *Genetics*, 177: 2389 – 2397;
- Hayes B., Hagesáther N., Ádnóy T., Pellerud G., Berg P.R., Lien S., (2006). “Effects on production traits of haplotypes among casein genes in Norwegian goats and evidence for a site of preferential recombination”, *Genetics*, 174: 455 – 456;
- Hazel L.N., (1943). “The genetic basis for constructing selection indexes”, *Genetics*, 28: 476 – 490;

- Henderson C.R., (1975). “Best linear unbiased estimation and prediction under a selection model”, *Biometrics*, 31: 423 – 447;

- Interbull, www.interbull.slu.se;

- Köning S., Simianer H., Willam A., (2009). “Economic evaluation of genomic breeding programs”, *J. Dairy Sci.*, 92: 382 – 391;

- Kujik E.W., van Tol L.T.A., Van de Velde H., Wubbolts R., Welling M., Geijsen N., Roelen B.A.J., (2012). “The role of FGF and MAP kinase in the segregation of the epiblast and hypoblast cell lineages in bovine and human embryos”, *Development and Stem Cells*, 139: 871 – 882;

- Lacasse P., Lollivier V., Bruckmaier R.M., Boisclair Y.R., Wagner G.F., Boutinaud M., (2011). “Effect of the Prolactin-release inhibitor quinagolide on lactating dairy cows”, *J. Dairy Sci.*, 94: 1302 – 1309;

- Lacasse P., Lollivier V., Dessauge F., Bruckmaier R.M., Ollier S., Boutinaud M., (2012). “New developments on the galactoproteic role of prolactin in dairy ruminants”, *Dom. Anim. Endo.*, 43: 154 – 160;

- Lagonigro R., Wiener P., Pilla F., Woolliams J.A., Williams J.L., (2003). “Short communication: a new mutation in the coding region of the bovine leptine gene associated with feed intake”, *Anim. Genet.*, 34: 371 – 374;

- Larson B.L.J., (1979). “Biosynthesis and secretion of milk proteins: a review”, *Dairy Res.*, 46: 161 – 174;

- Liefers S.C., Veerkamp R.F., te Pas M.F.W., Delavaud C., Chilliard Y., van der Lende T., (2004). “A missense mutation in the bovine leptin receptor gene is associated with leptin concentration during late pregnancy”, *Anim. Genet.*, 35: 138 – 141;

- Litman T., Brangi M., Hudson E., Fetch P., Abatti A., Ross D.D., Mityake K., Resau J.H., Bates S.E., (2000). “The multidrug-resistance phenotype associated with overexpression of the new ABC half-transporter, MXR (ABCG2)”, *J. Cell. Sci.*, 113: 2011 – 2021;
- Lush J.L., (1933). “The bull index problem in the light of modern genetics”, *J. Dairy Sci.*, 16: 501 – 522;
- Maltecca C., Parker K.L., Cassandro M., (2010). “Accomplishments and new challenges in dairy genetic evaluations”, *Ita. J. of Anim. Sci.*, 9: e68;
- Matsumoto H., Sasaki K., Bessho T., Kobayashi E., Abe T., Saszaki S., Oyama K., Mannen H., (2012). “The SNPs in the ACACA gene are effective on fatty acid composition in Holstein milk”, *Mol. Biol. Rep.*, 39: 8637 – 8644;
- Meuwissen T.H.E., Hayes B.J., Goddard M.E., (2001). “Prediction of total genetic value using genome wide dense marker maps”, *Genetics*, 157: 1819-1829;
- Meuwissen T.H.E., (2007). “Genomic selection: marker assisted selection on a genome wide scale”, *J. Anim. Breed. Genet.*, 124: 321-322;
- Mistry D.H., Medrano J. F., (2002). “Cloning and localization of the bovine and ovine Lysophosphatidic Acid Acyltransferase (LPAAT) genes that codes for an enzyme involved in Triglyceride biosynthesis”, *J. Dairy Sci.*, 85: 28 – 35;
- Morris C.A., Cullen N.G., Glass B.C., Hyndman D.L., Manley T.R., Hickey S.M., McEwan J.C., Pitchford W.S., Bottema C.D.K., Lee M.A.H., (2006). “Fatty acid synthase effects on bovine adipose fat and milk fat”, *Mammalian Genome*, 18: 64 – 74;
- Ogorevc J., Kunej T., Razpet A., Dovc P., (2009). “Database of cattle candidate genes and genetic markers for milk production and mastitis”, *Anim. Gen.*, 10: 1365 – 2052;

- Quaas R.L., Pollak E.J., (1980). “Mixed model methodology for farm and ranch beef cattle testing programs” , *J. Anim. Sci.*, 51: 1277 – 1287;
- Sadri H., Mielenz M., Morel I., Bruckmaier R.M., van Dorland H.A., (2010). “Plasma leptin and mRNA expression of lipogenesis and lipolisi-related factors in bovine adipose tissue around parturition”, *J. Anim. Physio. and Anim. Nutri.*, 95: 790 – 797;
- Schaeffer L.R., (2001). “Multiple trait international bull comparison”, *Livest. Prod. Sci.*, 69: 145 – 153;
- Schaeffer L.R., (2006). “Strategy for applying genome-wide selection in dairy cattle”, *J. Anim. Breed. Genet.*, 123: 369-376;
- Smith J.A., Lewis A.M., Wiener P., Williams J.L., (2000). “Genetic variation in the bovine myostatin gene in UK beef cattle: allele frequencies and haplotype analysis in the South Devon”, *Animal Genetics*, 31: 306 – 309;
- Song G., Riemer C., Dickins B., Kim H.L., Zhang L., Zhang Y., Hsu C.H., Hardison R.C., Green E.D., Miller W., (2012). “Revealing mammalian evolutionary relationships by comparative analysis of gene clusters”, *Genome Biol. Evol.*, 4: 586 – 601;
- Swalve H.H., Köning S., (2007). “Test herds in dairy cattle breeding programmes:1st communications: General considerations”, *Züchtungskunde*, 79: 449 – 462;
- Szyda J., Liu Z., Maschka R., Reinhardt F., Reents R., (2002). “Computer system for routine QTL detection and genetic evaluation under a mixed inheritance model in dairy cattle”, Proc. 7WCGALP, Communication #20 – 10;
- Tartara R., (1999). “La Frisone Italiana ed Interbull”, *Bianconero*;

- Täubert H., Rensing S., Reinhardt F., (2011). “Comparing conventional and genomic breeding programs with ZPLAN+”, *Interbull Bulletin* N°44;
- Tiezzi F., Pretto D., De Marchi M., Penasa M., Cassandro M., (2013). “Heritability and repeatability of milk coagulation proprieties predicted by mid-infrared spectroscopy during routine data recording, and their relationship with milk and quality traits”, *Animal*, 1 – 8;
- VanRaden P.M., Wiggans G. R., (1991). “Derivation, calculation and use of national animal model information”, *J. Dairy Sci.*, 74: 2737 – 2746;
- Visscher P.M., (1996). “Proportion of the variation in genetic composition in backcrossing programs explained by genetic markers”, *J. Hered.*, 87: 136 – 138;
- Widdison S., Siddiqui N., Easton V., Lawrence F., Ashley G., Werling D., Watson M., Coffey T. J., (2010). “The bovine chemokine receptors and their mRNA abundance in mononuclear phagocytes”, *BMC Genomics*, 11: 439;
- Zannoni M., Annibaldi S., (1981). “Standardization of the renneting ability of milk by Formagraph Pt. 1”, *Scienza e Tecnica Lattiero-Casearia*;
- Zhang Q., Boichard D., Hoeschele I., Ernst C., Eggen A., Murkve B., Pfister-Genskow M., Witte L. A., Grignola F. E., Uimari P., Thaller G., Bishop M. D., (1998). “Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree”, *Genetics*, 149: 1959 – 1973.

Ringraziamenti

Senza nemmeno accorgermene sono arrivata alla fine del mio percorso universitario e prima di concludere questa Tesi, mi sembra doveroso porgere alcuni ringraziamenti.

Un grazie speciale v`a sicuramente al Professor Martino Cassandro, che mi ha dato l'opportunità di realizzare questo progetto di Tesi, dimostrandomi sempre una grande disponibilità ed umanità. Ringrazio di cuore anche il Dottor Fabio Maretto, che mi ha supportato ed aiutato nelle analisi di laboratorio e nella stesura della Tesi, e il Dottor Francesco Tiezzi, che dall'altra parte del mondo ha perso ore di sonno per riuscire ad elaborare statisticamente i miei risultati.

Grazie anche a tutti i Dottorandi, Ricercatori, Borsisti e Tecnici di laboratorio che mi hanno sostenuto ed aiutato.

Una grazie di cuore v`a ovviamente ai miei genitori, Gianfranco e Rosetta, per l'appoggio e la pazienza che hanno sempre avuto nei miei confronti.

Ad Alberto, più che fare dei ringraziamenti, vorrei costruire un monumento per tutto l'amore che ha avuto e ha nei miei confronti e il grado di sopportazione che è riuscito a mantenere in questi mesi. Senza di lui non saprei davvero come sarei riuscita ad affrontare i momenti di ansia e frustrazione che ho vissuto durante la stesura di questa Tesi.

Infine ultimi, ma non certo per importanza, vorrei ringraziare tutti i membri della mia compagnia di amici per i bellissimi momenti trascorsi insieme, in cui con la vostra allegria mi avete permesso di staccare, anche solo per un attimo, dalla routine universitaria.

