



## Az.9 Elaborazione e gestione delle informazioni raccolte.

**IOVI:** Analisi genetiche nuovi caratteri. **Parametro:** pubblicazione risultati analisi genetiche e genomiche

**Anno 2°: 1 pubblicazione.**

*Pubblicazione di dati genomici che derivano dall'attività svolta nel progetto Dual Breeding.*

### Coefficienti di consanguineità genomica

La consanguineità è la probabilità che un animale abbia allo stesso locus entrambi gli alleli identici per origine (identical by descent).

La consanguineità può essere riferita anche come la frazione di geni omozigoti per discendenza complessivamente presenti nel genoma di un animale. E' possibile che un animale sia consanguineo solo se i suoi due genitori sono parenti tra di loro. La consanguineità è espressa con il coefficiente di consanguineità o di inbreeding ( $F_{PED}$ ) che viene generalmente calcolato usando i dati di pedigree.

Il coefficiente di consanguineità calcolato in questo modo ha diversi limiti:

- Non tiene conto delle relazioni di parentela tra gli animali della popolazione fondatrice (assume infatti che tutti gli animali della popolazione fondatrice non sono parenti);
- Necessita di registrazioni complete dei dati di pedigree per le due linee parentali, in modo da considerare le relazioni di parentela entro linea e tra linee;
- Assume che tutte le registrazioni di parentela sono corrette;
- Non tiene conto della stocasticità degli eventi di ricombinazione durante le meiosi nelle generazioni successive;
- Non tiene conto della selezione verso alcune regioni genomiche (*selection signature*).

Quindi il coefficiente di consanguineità così calcolato è intrinsecamente non preciso.

Per riuscire a superare i problemi legati al calcolo del coefficiente di consanguineità basato sui dati di pedigree è necessario calcolare direttamente a livello genomico questo parametro.

Analizzando il genoma di ciascun individuo quindi non è più necessario disporre dei dati di pedigree. Quindi il coefficiente di consanguineità genomico non risente più dei limiti che queste informazioni possono avere.

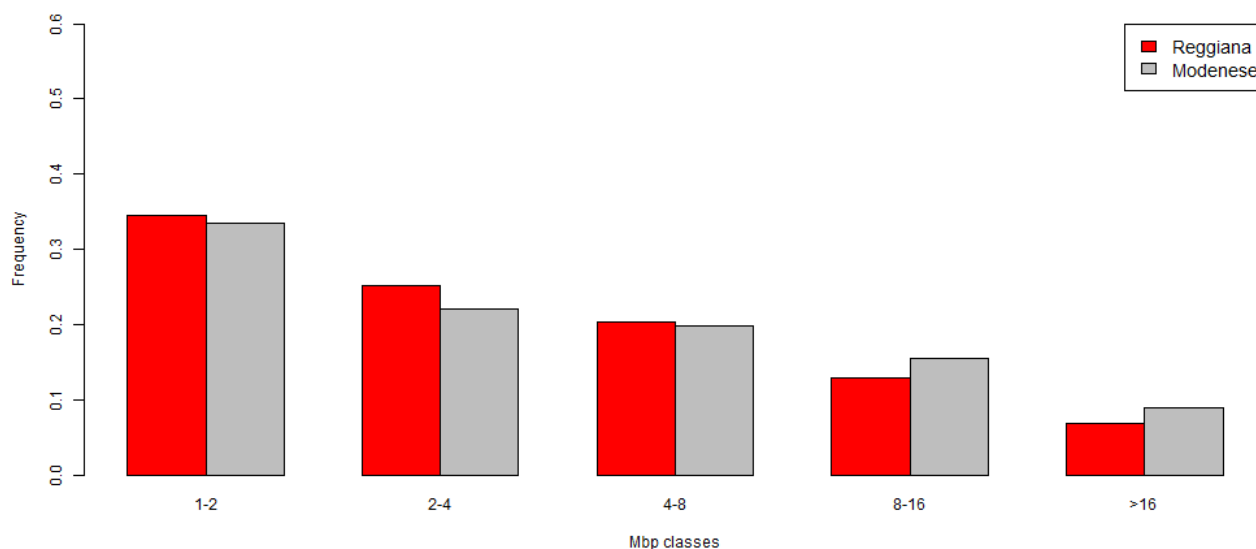


Grazie ai pannelli SNPchip, che permettono di valutare il genotipo di centinaia di migliaia di polimorfismi del DNA (marcatori del DNA o SNP) che si posizionano lungo tutto il genoma, è possibile calcolare un coefficiente di consanguineità genomica molto preciso. La presenza di regioni di SNP omozigoti (dall'inglese Runs of Homozygosity, abbreviato in ROH) lungo il genoma può infatti indicare che i genitori di un individuo gli abbiano trasmesso informazioni identiche che avevano a loro volta ricevuto da parenti comuni.

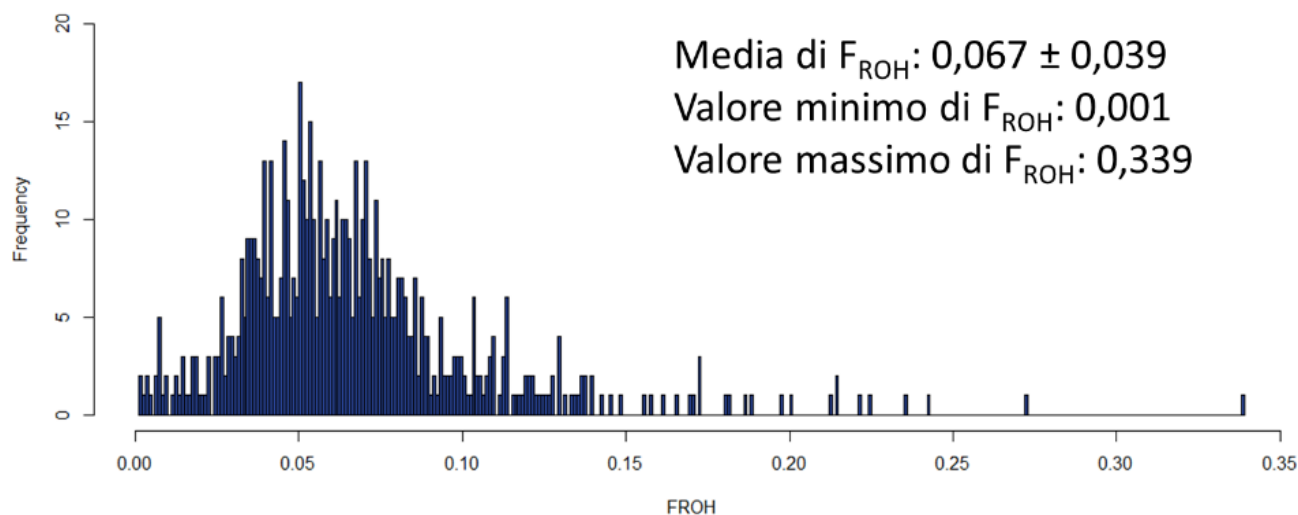
Il coefficiente di consanguineità genomica, che riflette il livello di consanguineità dovuto ad incroci tra animali imparentati, è calcolato come la porzione del genoma di un individuo che presenta ROH. Questa porzione può variare da 0 ad 1 e il coefficiente così ottenuto è indicato come FROH.

Se le regioni del genoma con ROH sono molto lunghe (più di 8 Mbp o di 16M bp) significa che gli eventi di accoppiamento tra consanguinei sono avvenuti di recente. Per questo, dopo aver identificato la presenza di ROH e la porzione di genoma coperto da esse, è interessante classificare le ROH in base alla loro lunghezza. Un'alta proporzione di ROH corte (1-4 Mbp) indica che la gestione degli accoppiamenti nella popolazione è riuscita a limitare la consanguineità.

Come mostrato in **Figura 1**, in queste popolazioni, le due classi di ROH corte (1-2 Mbp e 2-4 Mbp) sono preponderanti e rispecchiano una limitata consanguineità. Anche globalmente il valore di FROH degli individui è limitato (valore medio 0,067 +- 0.039) salvo per pochissimi individui in cui comunque non supera lo 0.35 (**Figura 2**).



**Figura 1.** Rappresentazione della distribuzione dei parametri di lunghezza dei Runs of Homozygosity (ROH) nelle razze Reggiana e Modenese su cui sono calcolati i coefficienti di consanguineità genomica.



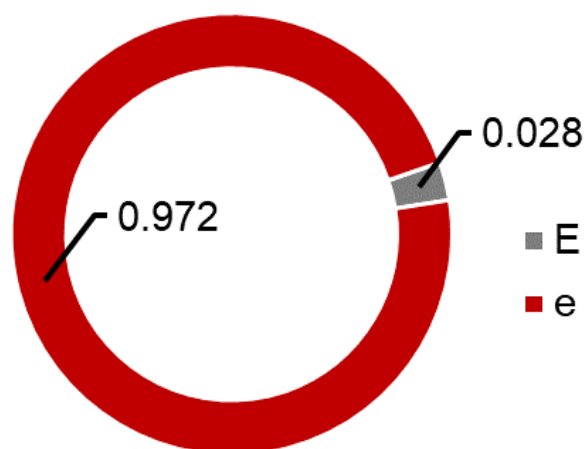
**Figura 2.** Distribuzione della frequenza dei valori di consanguineità genomica  $F_{ROH}$  calcolata utilizzando i Runs of Homozygosity nella razza Reggiana.

I risultati ottenuti da queste analisi permetteranno di utilizzare FROH ad integrazione di FPED per la pianificazione degli accoppiamenti con lo scopo di ridurre il livello di consanguineità reale nella razza Reggiana.

### Frequenze alleliche al gene MC1R nella razza Reggiana

La razza Reggiana si caratterizza morfologicamente per il mantello rosso fromentino. Diversi alleli ( $E^+$ ,  $E^D$ ,  $e$ ) del gene *MC1R* sono associati a particolari effetti sul colore del mantello. Nella Reggiana, il colore rosso è dato dalla presenza in omozigosi (due copie) dell'allele  $e$ . Tale caratteristica risulta utile per tracciare e autenticare geneticamente il formaggio Parmigiano Reggiano prodotto con il latte di solo vacche di razza Reggiana.

L'analisi del gene *MC1R* di 400 bovini di razza Reggiana ha evidenziato che non tutti gli animali sono omozigoti  $e/e$ . Una piccola frazione di questi animali è infatti portatrice dell'allele  $E$  considerando tutti gli altri alleli diversi da “ $e$ ” (**Figura 1**).



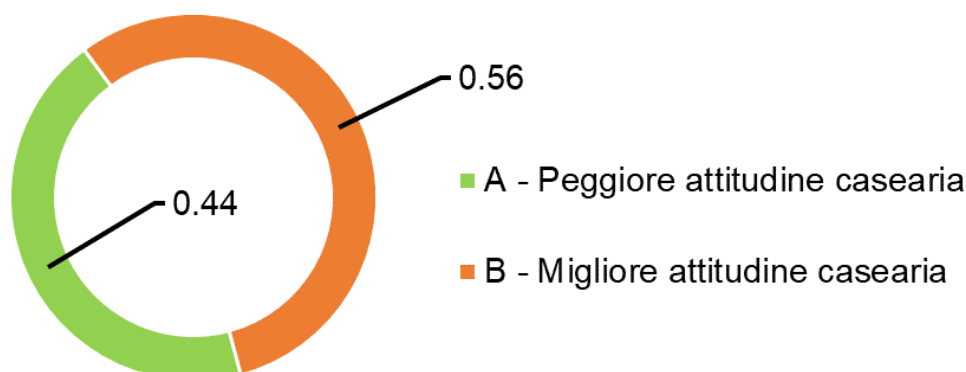
**Figura 1.** Frequenze alleliche al gene *MC1R* nella razza Reggiana. La distribuzione delle frequenze è riportata graficamente come frazione del cerchio. L'allele "e" è quello che determina il colore rosso fromentino, caratteristico della razza Reggiana.

I risultati ottenuti da queste analisi permettono di utilizzare le informazioni genomiche del gene *MC1R* per l'iscrizione degli animali al Libro Genealogico della razza.

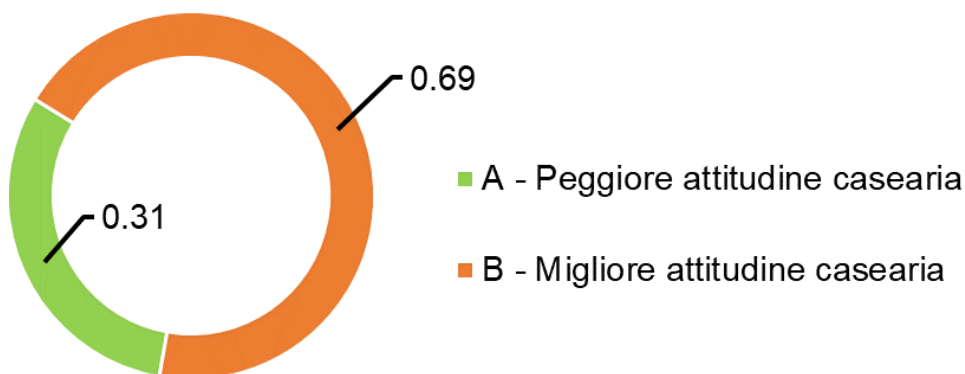
## Frequenze alleliche del gene k-caseina (*CSN3*) nelle razze Reggiana e Modenese

La K-caseina è una proteina del latte molto importante per la caseificazione. La proteina è codificata dal gene *CSN3*. Le diverse forme alleliche (A e B) di tale gene sono state associate ad una diversa attitudine casearia. In particolare, l'allele B migliora la reazione caglio/caseina, la velocità di coagulazione e la dimensione delle micelle mentre l'allele A mostra una minore propensione alla caseificazione.

L'analisi del gene *CSN3* di 400 bovini di razza Reggiana e di 120 bovini di razza Modenese ha evidenziato che l'allele B, associato ad una migliore attitudine casearia, è il più frequente sia nella popolazione di Reggiana (**Figura 1**) che nella popolazione di Modenese (**Figura 2**).



**Figura 1.** Frequenze alleliche del gene *k-caseina* (*CSN3*) nella razza Reggiana. La distribuzione delle frequenze è riportata graficamente come frazione del cerchio.



**Figura 2.** Frequenze alleliche del gene *k-caseina* (*CSN3*) nella razza Modenese. La distribuzione delle frequenze è riportata graficamente come frazione del cerchio.

I risultati ottenuti da queste analisi permettono di utilizzare le informazioni genomiche relative alla variabilità del gene *CSN3* per selezionare i portatori dell'allele favorevole ed aumentare ulteriormente la frequenza dell'allele B nelle popolazioni delle razze Reggiana e Modenese.



## Studio di associazione genome wide nella razza Reggiana per la percentuale di grasso nel latte.

Uno studio di associazione genome wide (GWAS) è un'indagine dell'intero genoma al fine di identificare regioni del DNA associate ad un determinato carattere e quindi potenzialmente responsabili dell'espressione del fenotipo. In questa tipologia di studi viene quindi valutata una sorta di "correlazione" tra un fenotipo (esempio: colore del mantello) e uno delle centinaia di migliaia di polimorfismi del DNA (marcatori del DNA) che si posizionano lungo tutto il genoma. Come risultato dell'analisi si ottiene un livello di significatività per ciascun polimorfismo testato.

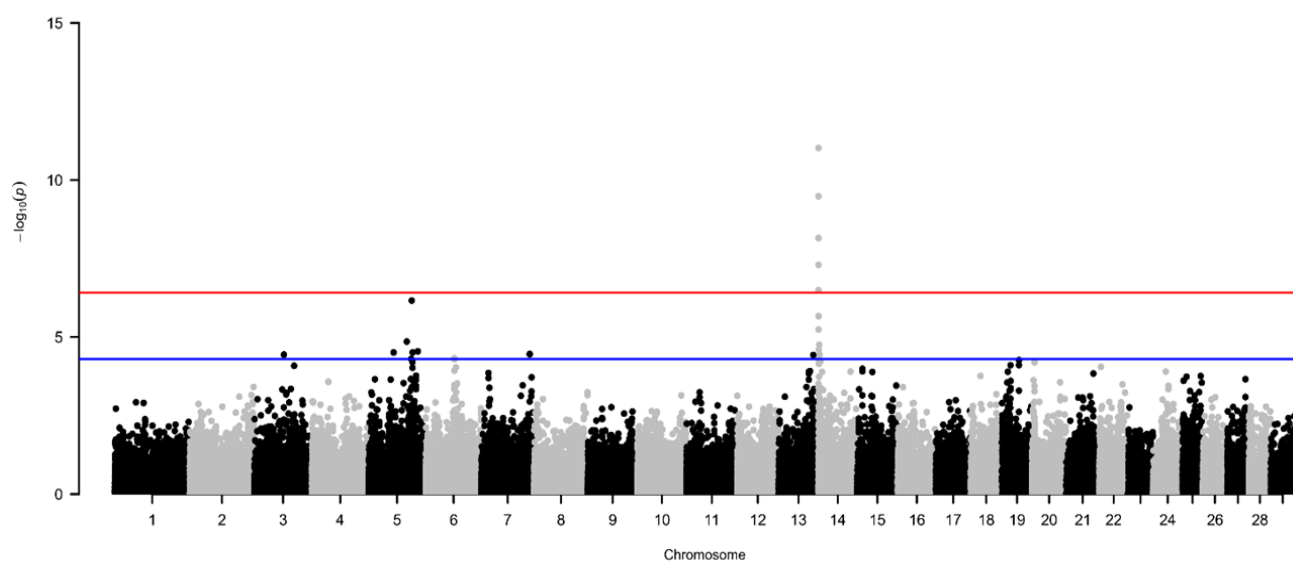
Utilizzando i dati genomici derivati dai pannelli di SNP genotipizzati nei bovini di razza Reggiana è stato possibile effettuare uno studio di associazione genome-wide per alcuni caratteri produttivi. Nel caso specifico sono stati uniti i dati genomici di 1234 bovini di razza Reggiana e gli indici genetici per la percentuale di grasso nel latte.

Il grafico in **Figura 1** riporta i risultati dell'analisi di associazione tra la percentuale di grasso nel latte e i 139459 marcatori del DNA che si sono analizzati utilizzando il chip di SNP della GeneSeek (HD GGP) per 1234 bovini. Ogni punto rappresenta un marcatore del DNA a posizione nota sul genoma bovino (asse x). L'asse delle x è poi anche suddiviso per cromosomi. Il genoma bovino è infatti diviso in 29 cromosomi autosomici che sono indicati, appunto, su quest'asse. Sull'asse delle y è riportato il  $p$ -value di associazione, ovvero la significatività dell'associazione genotipo-fenotipo.

La rappresentazione grafica di tutto il genoma riportata in Figura 1 non permette di distinguere tutti i punti (cioè gli SNP) in quanto sono troppo densi. Tuttavia, ciò che interessa sono i diversi punti che si elevano al di sopra delle linee blu e rosse che indicano i diversi livelli di significatività che gli SNP raggiungono nello studio di associazione. Più il valore è alto e più il risultato è significativo. Questo significa che gli SNP in quelle regioni del genoma sono significativi e che mutazioni in alcuni geni in quelle regioni determinano un effetto rilevante sulla percentuale di grasso del latte.

I risultati dell'analisi così strutturata mostrano un picco di associazione (cioè diversi puntini, che sono gli SNP nelle corrispondenti posizioni cromosomiche, al di sopra delle linee rosse e blu) sul cromosoma 14. In questa regione del cromosoma 14 sono presenti geni che influenzano la percentuale di grasso del latte. Sappiamo già, da altri studi, che in quella regione è presente il gene DGAT1 la cui variabilità in molte razze bovine influenza direttamente il contenuto di grasso nel latte.

I risultati ottenuti permetteranno di selezionare i bovini in base al genotipo in questa regione per aumentare la qualità del latte.



**Figura 1.** Risultati dello studio di associazione genome wide nella razza Reggiana per la percentuale di grasso nel latte. La figura rappresenta un Manhattan plot che unisce le informazioni di significatività degli SNP (asse delle Y) con la loro posizione nel genoma (asse delle X).