

AZ. 7 – IOV 7.1 - Produzione report con studio su GAS-MAS

**STUDIO DI ASSOCIAZIONE GENOMICA PER LA FERTILITÀ ESPRESSA COME INTERVALLO PARTO CONCEPIMENTO DELLA RAZZA RENDENA**

**A cura di: Enrico Mancin e Roberto Mantovani**

Dipartimento di Agronomia, Animali, Alimenti, Risorse Naturali e Ambiente - Dafnae

**Premessa**

La selezione genomica è ormai ampiamente utilizzata nel miglioramento genetico di numerose specie e razze di interesse zootecnico grazie alla possibilità di eseguire su larga scala la genotipizzazione degli animali (Miształ et al., 2020). Il punto cruciale di questo processo risiede nell'uso di un elevato numero di marcatori presenti nel DNA di ogni individuo, noti come SNP (polimorfismi a singolo nucleotide), che sono uniformemente distribuiti sul genoma, facilmente identificabili con analisi a basso costo e completamente automatizzabili. Grazie alla loro elevata numerosità, questi marcatori sono facilmente associabili a geni con azione quantitativa o QTL, ossia geni che influenzano caratteri di rilevanza economica e zootecnica (latte, carne, fertilità, salute, longevità, ecc.; Miształ et al., 2020). Con decine di migliaia di SNP, scelti per rappresentare l'intero genoma, si prevede sempre una vicinanza tra un marcatore SNP e un particolare gene o segmento di DNA di interesse per la selezione. La selezione genomica basata sugli SNP (GAS/MAS) consente quindi di identificare gli animali potenzialmente miglioratori per uno o più caratteri, misurando il potenziale genetico individuale basato sulla quantificazione delle associazioni tra geni QTL e marcatori SNP (Miształ et al., 2020). Per determinare rapidamente ed efficacemente il genotipo di migliaia di SNP in molti individui, si utilizza la tecnologia del "DNA-chip", costituita da un supporto solido come vetro, plastica o chip di silicio, che forma un array (vettore) sulla cui superficie sono presenti normalmente centinaia di migliaia di sonde a DNA, ciascuna delle quali corrisponde a uno specifico marcatore SNP a cui si lega il frammento di DNA dell'individuo. Nel corso del PSRN (Dualbreeding e Dualbreeding-fase2) sono stati raccolti numerosi campioni biologici e ottenuti profili genomici di numerosi maschi e femmine della razza Rendena. Sulla base di questi profili genomici, è stato condotto uno studio per valutare e identificare possibili associazioni tra marcatori SNP e il carattere relativo alla fertilità nelle bovine di razza Rendena, espresso come intervallo parto-concepimento, e oggetto di analisi genetiche nell'ambito del PSRN. L'obiettivo era identificare possibili QTL coinvolti nell'espressione di questo importante carattere di fertilità, con l'intento finale di utilizzare questi QTL in processi di selezione di tipo GAS/MAS per la razza Rendena per questo importante carattere di benessere animale: è noto infatti come animali fertili, che riescono,

cioè, a concepire nel minor tempo possibile dopo il parto, sono indubbiamente indice di buono stato di benessere dell'animale.

## **Dati e Metodi**

### Rilievi Genomici

I profili genomici iniziali comprendevano 3.445 animali genotipizzati con 33K (33.000 SNP circa; ottenuti con il chip a bassa densità GGP Bovine LD Array 33K di Illumina) e con chip a più alta densità 150K (150.000 SNP circa; chip GGP Bovine HD Array 150K di Illumina). Inizialmente sono stati rimossi dal dataset gli animali con genotipizzazioni aventi un call-rate inferiore al 90% (inquinamento del campione biologico) e sono stati eliminati dai profili genomici tutti i marcatori SNP con call rate marcatore inferiore al 10% e con una Minor Allele Frequency (MAF)  $<0.05$ . In entrambi i casi, infatti, la variante SNP identificata risulta poco presente o presente con bassa frequenza nel gruppo di animali analizzato, mettendo in luce, in definitiva, poche possibilità di poter avere stretti legami con geni ad azione quantitativa (QTL) legati alla fertilità. Quindi si è proceduto all'armonizzazione dei profili genomici ottenuti con differenti chip mediante imputazione dei marcatori SNPs per tutti i soggetti 33K utilizzando come riferimento il panel a 150K; questa operazione che è stata resa possibile dall'utilizzo delle relazioni di parentela tra tori e vacche sottoposti a genotipizzazione utilizzando allo scopo uno specifico programma (FImpute3). Al termine del processo di imputazione sono stati infine rimossi tutti gli animali con bassa percentuale di imputazione, cioè con una soglia inferiore al 90% di call rate. Il dataset finale conteneva quindi 1.416 animali genotipizzati con una densità media di 113.400 SNP (113K).

### Rilievi fenotipici

Come carattere fertilità è stato utilizzato il parto concepimento. Il parto concepimento è l'intervallo tra la data del parto e l'inseminazione in cui si è ottenuta la gravidanza. Dati appartenenti all'ultima lattazione di animali ancora vivi sono stati inizialmente considerati informazioni censurate. Tuttavia, poiché i record *censored* (incompleti) erano solo il 2% dei fenotipi, sono stati rimossi dal dataset finale per semplicità. La procedura di data editing è stata fatta come descritto in Mancin et al., 2020. Dopodiché solo animali aventi genotipo e fenotipo sono stati considerati (1.341 animali totali).

## **Modello di associazione**

Sono stati individuati tra tutti i marcatori SNP, quelli che presentavano un valore significativo di associazione fenotipica con la fertilità (ssGBLUP; Aguilar et al., 2019).

Studi precedenti hanno dimostrato che ssGBLUP fornisce potenzialmente valori genomici (gEBV) più accurati e meno distorti rispetto agli alternativi metodi multi-step, specialmente in presenza di piccole popolazioni e caratteri limitati dal sesso (Christensen e Lund, 2010).

Il modello matematico utilizzato in questo studio è stato il seguente:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{W}\mathbf{p} + \mathbf{Z}\mathbf{a} + \mathbf{e}$$

dove  $\mathbf{y}$  è il vettore dei record parto-concepimento di una determinata vacca e  $\mathbf{X}$ ,  $\mathbf{W}$  e  $\mathbf{Z}$  sono le matrici di incidenza degli effetti fissi, dell'effetto dell'ambiente permanente e gli effetti genetici, rispettivamente. Il vettore  $\mathbf{b}$  include gli effetti fissi del mese allevamento anno, numero lattazione. I vettori  $\mathbf{p}$ ,  $\mathbf{a}$ ,  $\mathbf{e}$  rappresentano i seguenti effetti casuali: l'ambiente permanente, la componente genetica additiva e l'errore residuo.

In ssGBLUP, gli effetti genetici additivi vengono campionati dalla distribuzione  $\mathbf{a} \sim \mathbf{N}(\mathbf{0}, \mathbf{H}\sigma_a^2)$ . La matrice  $\mathbf{H}$  contiene informazioni di parentela sia anagrafiche, sia genomiche. In ssGBLUP, la matrice inversa della struttura di (co)varianza dell'effetto additivo è rappresentato da  $\mathbf{H}^{-1}$ , descritta come:

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

dove  $\mathbf{A}^{-1}$  e  $\mathbf{A}_{22}^{-1}$  sono gli inversi della matrice delle relazioni genealogiche rispettivamente per tutti gli animali e solo per gli animali genotipizzati.  $\mathbf{G}^{-1}$  è l'inversa della matrice di relazione genomica.

In questo caso si è scelto di rappresentare l'associazione genotipo fenotipo tramite percentuale di varianza spiegata da una finestra di 20 SNPs. Finestre con una varianza genetica sopra l'1% vengono considerate associate al fenotipo di fertilità.

## Risultati e conclusioni

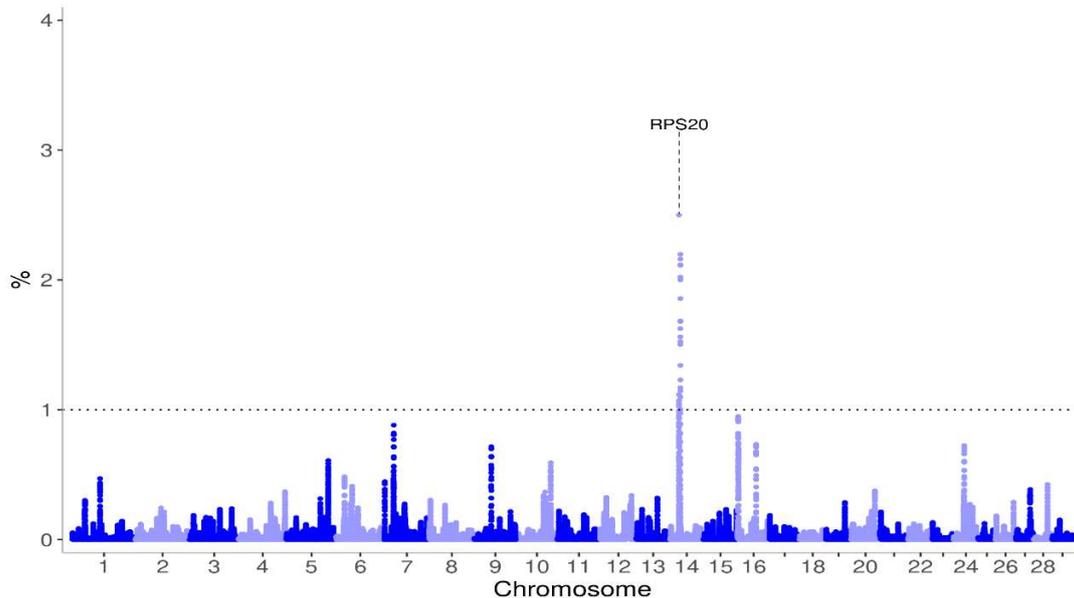
Attraverso il metodo impiegato è stato possibile identificare associazioni statisticamente significative tra la fertilità, espressa come intervallo parto-concepimento, e alcuni SNP presenti nel cromosoma 14 (UCSC Genome Browser, 2002; Figura 1).

Un grande picco è stato trovato nel cromosoma 14, situato nella regione del gene RPS20, che codifica per la proteina "ribosomal protein S20". Questo gene è stato associato alla fertilità bovina attraverso il nostro studio di associazione genome-wide (GWAS).

Il gene RPS20 è noto per il suo ruolo nella regolazione della trasduzione del segnale e nell'omeostasi cellulare, che sono processi cruciali per la riproduzione e la fertilità. Questa evidenza è supportata anche da altri studi che hanno evidenziato l'importanza di RPS20 nella fertilità bovina. Ad esempio, un GWAS

ha identificato due loci principali che influenzano la facilità di parto e i tratti di crescita nei bovini, suggerendo un legame tra questi loci e la fertilità (Pausch et al., 2011).

Figura 1: Manhattan plot rappresentante l'associazione tra SNP e con fertilità nella razza Rendena



Un altro studio ha rivelato inoltre un'ampia variazione nell'espressione genica di 13 geni legati alla fertilità utilizzando tecniche avanzate come RNA-Seq, ISO-Seq e CAGE-Seq nei bovini Brahman (Ross et al. 2022). Questo sottolinea ulteriormente l'importanza del gene RPS20 nella regolazione della fertilità.

Un'altra ricerca ha esplorato le *signature of selection* o firme di selezione positiva nelle razze bovine selezionate per diversi caratteri, trovando evidenze che suggeriscono la selezione naturale di varianti genetiche che influenzano la fertilità (Taye et al. 2017).

Infine, uno studio ha mostrato che l'assenza di cellule del cumulo durante la maturazione in vitro influisce sul metabolismo lipidico negli ovociti bovini, evidenziando come il metabolismo cellulare e l'ambiente cellulare siano fondamentali per la fertilità (Auclair et al., 2013).

Questi risultati indicano che il gene RPS20 potrebbe giocare un ruolo chiave nella fertilità bovina e suggeriscono la necessità di ulteriori ricerche per confermare queste scoperte. La connessione tra RPS20 e la fertilità potrebbe aprire nuove strade per migliorare i programmi di selezione genetica e le pratiche di gestione riproduttiva nei bovini.

**Bibliografia:**

- Aguilar, I., Legarra, A., Cardoso, F., Masuda, Y., Lourenco, D., & Misztal I. Frequentist p - values for large - scale - single step genome - wide association, with an application to birth weight in American Angus cattle. *Genetics Selection Evolution* 51 (2019).
- Christensen, O.F., Lund, M.S. Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution* 42 (2010).
- Misztal, I., Lourenco, D., Legarra A. Current status if genomic evaluation. *Journal of Animal Science* 98 (2020): 1-14.
- Mancin, Enrico et al. Non-genetic effects affecting fertility traits in local Reggiana cattle. *Acta Fytotechnica et Zootechnica*, 23 (2020): 338-346.
- Pausch, Hubert, Krzysztof Flisikowski, Simone Jung, Reiner Emmerling, Christian Edel, K. U. Götz and Ruedi Fries. Genome-Wide Association Study Identifies Two Major Loci Affecting Calving Ease and Growth-Related Traits in Cattle. *Genetics* 187 (2011): 289 - 297.
- Ross, Elizabeth M., H. B. Sanjana, Loan T. Nguyen, Yuanyuan Cheng, Stephen Moore and Ben J. Hayes. Extensive Variation in Gene Expression is Revealed in 13 Fertility-Related Genes Using RNA-Seq, ISO-Seq, and CAGE-Seq From Brahman Cattle. *Frontiers in Genetics* 13 (2022)
- Taye, Mengistie, Wonseok Lee, Soomin Jeon, Joon Yoon, Tadelle Dessie, Olivier Hanotte, Okeyo Ally Mwai, Stephen J. Kemp, Seoae Cho, Sung-Jong Oh, Hak Koy Lee and Heebal Kim. Exploring evidence of positive selection signatures in cattle breeds selected for different traits. *Mammalian Genome* 28 (2017): 528-541.
- Auclair, Sylvain, Rustem E. Uzbekov, Sébastien Elis, Laura Sanchez, Igor I. Kireev, Lionel Lardic, Rozenn Dalbies-Tran and Svetlana Uzbekova. Absence of cumulus cells during in vitro maturation affects lipid metabolism in bovine oocytes. *American journal of physiology. Endocrinology and metabolism* 304 6 (2013): E599-613.