

RELAZIONE TECNICO-SCIENTIFICA RELATIVA ALLE ATTIVITÀ SVOLTE DA AGRIS SARDEGNA NELL'AMBITO DEL PROGETTO "CHEESR" – Sottomisura 10.2 PSRN 2014-2020.

Le attività di supporto tecnico –scientifico che AGRIS Sardegna ha realizzato per ASSO.NA.PA. si sono concretizzate nella messa a punto di una proposta di schema di selezione genomico per la razza ovina Sarda e più in generale per le razze ovine da latte.

PREMESSA

Le attività di sperimentazione si sono da un lato concentrate sull'individuazione di mutazioni causative o marcatori *in linkage disequilibrium* (LD) per una eventuale applicazione della selezione assistita da marcatori. Attualmente i marcatori più utilizzati sono i Single Nucleotide Polymorphisms (SNP) per i quali sono stati messi a punto dei sistemi di analisi altamente efficienti e anche relativamente economici (SNP chip) che variano per numero di SNP analizzati con un'unica analisi (da 5K a 800K). L'avvento delle nuove tecnologie di re-sequenziamento (*Next Generation Sequencing*, NGS) a costi sostenibili consente oggi di abbinare l'uso degli SNP chip alla conoscenza dell'intera sequenza del genoma di un animale con una conseguente maggiore probabilità di identificazione delle mutazioni causative. L'impiego combinato di SNP chip e NGS è dunque al momento lo strumento più performante per le applicazioni di selezione assistita da marcatori.

L'altro aspetto su cui si è rivolta la sperimentazione è l'implementazione della selezione genomica (GS) che si basa sull'utilizzo di una popolazione di riferimento costituita da individui con dati fenotipici e genotipici attraverso i quali è possibile stimare equazioni di predizione dell'effetto additivo dei segmenti genomici. L'impiego di queste equazioni consente in seguito la determinazione del valore genomico dei candidati alla selezione dei quali si conosce esclusivamente il genotipo. L'applicazione della GS nei bovini da latte si fonda sull'utilizzo di una popolazione di riferimento costituita da tori con un gran numero di figlie per i quali è possibile stimare un accurato valore genetico. Anche negli ovini da latte si sta tentando di applicare la GS ma con difficoltà maggiori legate alla mancanza di un numero sufficiente di arieti con un numero adeguato di figlie, al ridotto numero di caratteri misurati nelle aziende in selezione e al valore economico del singolo capo che limitano la sostenibilità del costo della genotipizzazione.

La popolazione femminile di riferimento

L'applicazione degli strumenti della genomica richiede generalmente una popolazione di riferimento numerosa sulla quale effettuare precise misure di un gran numero di fenotipi per un lungo periodo di tempo. Le limitazioni riportate in precedenza per l'impiego della genomica negli ovini da latte hanno portato alla valutazione di un approccio per la razza Sarda di una popolazione di riferimento femminile (PRF). Tale approccio consente infatti di poter effettuare misure accurate per un gran numero di caratteri su una popolazione ridotta con un'incidenza inferiore dei costi per le misurazioni e genotipizzazioni. La popolazione è stata ottenuta dall' "allevamento genomico" presente nell'azienda di Monastir di AGRIS. Essa è costituita annualmente da circa 1000 pecore in lattazione con una rimonta del 25%, generata dall'accoppiamento delle pecore adulte con arieti provenienti dal Centro Arieti di razza Sarda e più in generale da allevamenti del LG. Fino ad oggi sono state generate 4850 pecore da circa 180 arieti. Su tutte le pecore sono stati misurati numerosi fenotipi relativi a: caratteri produttivi (produzione di latte, contenuto in grasso, proteina e lattosio, peso e nota di stato corporeo), caratteri legati alla facilità di mungitura e alla conformazione della mammella, caratteri legati alla sanità dell'animale (contenuto in cellule somatiche, mastiti cliniche, conta delle uova fecali, test ELISA per la paratubercolosi e la *visna maedi*, esami istopatologici per la paratubercolosi), caratteri legati al contenuto in acidi grassi del latte e caratteri riproduttivi (fertilità e prolificità).

La selezione assistita da marcatori

Nel corso del progetto sono state effettuate diverse analisi per l'identificazione di QTL che influenzano i caratteri misurati nelle diverse generazioni della PRF. Grazie alla particolare struttura in famiglie della PRF si è potuto utilizzare un metodo di identificazione che combina le informazioni di linkage entro famiglia e quelle di LD a livello di popolazione (LDLA analysis) e prevede l'impiego dell'analisi delle componenti principali della matrice IBD (*Identity By Descent*) al fine di includere tutte le informazioni di somiglianza genetica degli aplotipi base della popolazione di riferimento. Sono state individuate un elevato numero di regioni del DNA significativamente associate a molti dei caratteri misurati nella PRF (Tabella 1).

Tabella 1 - QTL significativi alla soglia 5% genome-wide per i caratteri misurati

Carattere misurato	Cromosoma (numero di regioni nel cromosoma)
<i>Caratteri produttivi¹</i>	
Quantità di Latte	2(2), 3(1), 5(1), 10(1), 11(1), 12(1),13(1), 16(1), 20(1), 22(1), 25(1)
Quantità di Grasso	2(1), 10(1), 11(1), 15(1), 20(1), 22(1), 25(1)
Quantità di Proteina	2(1), 3(1), 10(1), 11(1), 12(1), 15(1), 20(1), 25(1)
Tenore in Grasso	1(2), 2(2), 3(1), 4(1), 5(1), 6(1), 9(1), 10(1), 12(2), 13(1), 14(1), 16(1), 17(1), 18(1), 19(1), 20(1), 25(1)
Tenore in Proteina	1(2), 2(2), 3(2), 4(1), 5(1), 6(1), 8(1), 9(1), 10(1), 11(1), 12(1), 13(1), 15(1), 16(1), 17(1), 18(1), 26(1)
<i>Morfologia mammaria²</i>	
Altezza della Mammella	9(1), 24(1)
Posizione dei Capezzoli	3(3), 6(1), 7(1), 8(1), 9(1), 10(1), 16(2), 17(1), 20(1)
Solco Mediano	2(1), 5(1)
Grado di Sospensione della Mammella	1(1), 2(1), 4(1), 5(2), 7(1), 8(1), 9(3), 10(1), 14(1), 16(1), 20(1), 25(1)
<i>Caratteri qualità del Latte³</i>	
C4:0	17(1)
C10:0	11(1)
C14:0	11(1)
C14:1	7(1), 22(1)
C14:1/C14:0	7(1), 22(1)
C15:0	16(1)
C16:0	26
C16:1	10(1), 22(1)
C16:1/C16:0	22(1)
C18:0	8(1)
C18:1/C18:0	19(1), 22(1)
CLA/acido vaccenico	22(1)
<i>Caratteri Sanitari⁴</i>	
Contenuto Cellule Somatiche	2(1), 3(2), 4(1)
Conta Uova Fecali	1(1), 3(1), 4(1), 8(1), 12(1), 15(2), 17(1), 19(1), 20(1), 21(1)
Test ELISA per la paratuberculosis	20(1)
Esame Istopatologico per paratuberculosis	20(1)

¹Usai *et al.*, 2014; ²Casu *et al.*, 2018; ³Casu *et al.*, 2014; ⁴Carta *et al.*, 2016.

In quelle più significative sono state effettuate ulteriori analisi genomiche per identificare le mutazioni causali o in LD per i caratteri misurati. La definizione più precisa dei polimorfismi in queste regioni consentirà l'impiego delle mutazioni causali o in LD a fini selettivi, impiegando chip a bassa densità che includono SNP con elevata capacità predittiva delle performance per i caratteri di interesse. Attualmente già si dispone di polimorfismi utili da inserire nei chip che riguardano sia caratteri produttivi (produzione di latte e contenuto in grasso e proteine) che funzionali (resistenza alla Scrapie, Visna Maedi, Nematodi, morfologia mammaria e resistenza alle mastiti). L'impiego di tali polimorfismi può anche essere utilizzato per l'attribuzione delle parentele nel caso di gruppi di monta con più arieti.

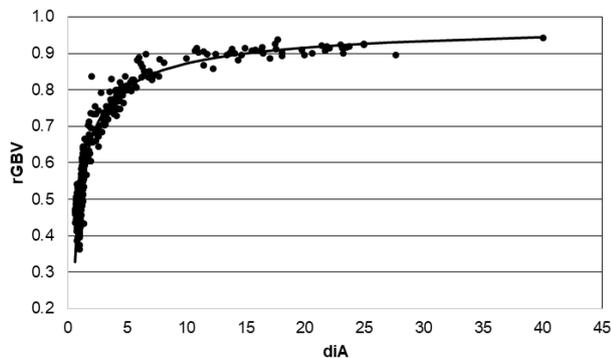
La selezione genomica

Gli obiettivi principali della sperimentazione per l'applicazione della selezione genomica nello schema di selezione sono stati la verifica della rappresentatività della variabilità genetica della razza nella PRF e la valutazione dell'accuratezza della predizione del valore genomico attraverso la PRF. A tal fine tutte le analisi sono state effettuate impiegando la produzione di latte e il tenore in proteina come caratteri di riferimento in quanto esistono numerose stime pregresse dei parametri genetici che possono consentire la messa a punto di modelli di analisi adeguati anche per i caratteri innovativi che sono specificamente oggetto di studio in questo progetto.

Per quantificare l'informazione molecolare di un singolo ariete rappresentata nella PRF, Usai *et al.* (2018) hanno utilizzato il valore della diagonale (diA) della parte dell'inversa della matrice di parentela additiva (A) che include le pecore della PRF e l'ariete considerato. Per valutare la relazione tra accuratezza dell'indice genomico e la porzione di genoma rappresentato nella PRF è stato considerato un campione di validazione costituito da diverse categorie di arieti del LG: 144 arieti di FA con figlie in PRF, 105 arieti di FA senza figlie

in PRF, 288 arieti impiegati in monta naturale senza figlie in PRF. Tutti gli arieti disponevano della valutazione genetica convenzionale per la produzione di latte (EBV). Per tutti gli arieti è stato stimato il valore genomico (GEBV) per la produzione di latte utilizzando i dati produttivi della PRF e la matrice di parentela genomica (**G**) tra gli arieti e la PRF. La matrice genomica consente di stimare una parentela genetica fra due individui in funzione della condivisione di tratti identici di DNA, anche quando questi non risultano parenti in base alle informazioni anagrafiche. Comunemente per il calcolo dell'accuratezza del GEBV si impiega la correlazione fra GEBV e EBV di un campione di validazione di riproduttori che abbiano un EBV sufficientemente accurato da poter essere considerato un valido stimatore del valore genetico reale. Nello schema di selezione della razza Sarda, a causa del ridotto impatto della FA e dell'impiego di sistemi semplificati di misurazione delle produzioni, solo per pochi arieti si ottengono accuratèzze del loro EBV così elevate. Infatti, poco meno del 50% degli arieti genotipati del LG hanno un numero di figlie superiore a 50. Per tale motivo, in questo caso, si è preferito calcolare l'accuratezza teorica della stima del GEBV (rGBV) degli arieti del campione di validazione come $\{1 - [SE_i^2 / (G_{ii} * \sigma_g^2)]\}^{1/2}$ dove i rappresenta l' i -esimo ariete e SE l'errore standard della stima. Il valore di rGBV così calcolato è risultato compreso tra 0,36 e 0,94 con una media di 0,58, valore superiore al valore medio dell'accuratezza teorica dell'indice pedigree di arieti avviati alla prova di progenie nel LG (0,52). Come si può osservare nella figura 1, diA mostra una buona capacità predittiva dell'accuratezza della stima del valore genomico.

Figura 1 – Relazione tra l'accuratezza della stima del valore genomico (rGBV) e il parametro diA che quantifica l'informazione molecolare di un singolo ariete rappresentata nella PRF.



Pertanto, attraverso diA è possibile scegliere gli arieti da genotipare e quelli da utilizzare come padri nella PRF con l'obiettivo di massimizzare il numero di arieti del LG che mostrano una rGBV attesa oltre una soglia di riferimento. In pratica si è individuata la soglia di parentela tra arieti e PRF che consente di ottenere accuratèzze attese superiori o uguali a quelle di un tradizionale indice pedigree. In questa analisi il 49% dei 537 arieti superava questa soglia.

Per confermare la capacità predittiva della valutazione genomica basata sulla PRF si è confrontata in un campione di 119 giovani arieti, l'accuratezza teorica della stima del valore genomico ottenuta utilizzando PRF come popolazione di riferimento con quella ottenuta utilizzando una popolazione di riferimento maschile costituita da 414 arieti adulti con figlie nel LG attraverso un modello GBLUP che utilizzava le loro DYD. I risultati hanno dimostrato che la media di rGBV era superiore utilizzando PRF (0,49 vs 0,42). È stata inoltre studiata la relazione tra la porzione di genoma di un ariete rappresentata nella popolazione maschile, stimata come nel caso precedente dal diA corrispondente alla porzione di matrice di parentela che include l'ariete e i maschi della popolazione di riferimento, e l'accuratezza dei GEBV. Le relazioni tra i diA e le relative accuratèzze stimate con la PRF o con la popolazione di riferimento maschile sono state poi utilizzate per predire l'accuratezza teorica dei GEBV di 310 giovani arieti del LG nati nel 2013 utilizzando come popolazione di riferimento la PRF o alternativamente 4.189 arieti nati tra il 2003 e il 2010, che simulano la popolazione di riferimento maschile che sarebbe stata disponibile nel caso dell'applicazione di un approccio "bovino" alla GS. I risultati mostrano che il valore medio di rGBV attesa tra le due popolazioni è simile. Tale risultato conferma che la capacità predittiva della valutazione genomica basata sulla PRF è, per la produzione di latte, analoga a quella che si sarebbe ottenuta genotipizzando tutti i maschi del LG nati dal 2003 al 2010. Questo risultato conferma che per i caratteri misurati solo all'interno di PRF, si ottengono GEBV dei giovani arieti comparabili a quelli ottenibili con un approccio "bovino" ma con un enorme risparmio in termini di misure fenotipiche. Per i caratteri misurati sia in PRF che nel LG, la PRF è essenziale per innescare la selezione genomica in attesa che, con il crescere del numero di arieti con figlie genotipate nel LG, si possano combinare

le informazioni genomiche derivanti dalla PRF con quelle provenienti dalla popolazione di riferimento maschile attraverso la metodologia Single-Step GBLUP.

Infine, con l'obiettivo di fornire criteri pratici per gestire la scelta dei riproduttori da impiegare nella PRF e di identificare una soglia di parentela con PRF che consenta agli allevatori di prevedere l'accuratezza del GEBV dei loro arieti è stato stimato il valore atteso di rGBV in funzione del numero di parenti presenti nella PRF. A tal fine è stata utilizzata la regressione della rGBV di un ariete con il numero di parenti in PRF con coefficiente di parentela di 0,50 e il numero di parenti con coefficiente di parentela di 0,25. I risultati hanno dimostrato che arieti con almeno 5 figlie in PRF raggiungono un valore di accuratezza della stima del valore genomico di 0,58 mentre i figli di arieti con almeno 5 figlie in PRF (vale a dire con 5 mezze sorelle) raggiungono un valore di 0,54.

Dal punto di vista applicativo, una volta posto a regime lo schema, solo gli arieti avviati alla prova di progenie verranno genotipati e di questi alcuni saranno impiegati come padri della PRF. Considerando un tasso di rimonta di 250 pecore nella PRF sarà possibile ottenere un GEBV con accuratezza oltre la soglia del pedigree tradizionale per 50 nuovi arieti all'anno e per i loro figli.

Una prima valutazione genomica è stata eseguita per i caratteri produzione di latte, contenuto in proteina e resistenza ai nematodi. Gli indici latte e proteina sono stati stimati per ottenere valutazioni e distribuzioni di riferimento anche per i caratteri innovativi. I dati produttivi sono stati registrati nell'allevamento genomico di Monastir dal 2000 al 2020. Si tratta di 15001 record della produzione di latte in lattazione espressa in Equivalente Pecora Matura (LEPM) di 4487 pecore, 14990 record della percentuale di proteina in lattazione (TP) di 4483 pecore e 16964 record del Contenuto individuale di Uova Fecali (FEC) come variabile di riferimento per la resistenza ai nematodi espressa come valore logaritmico ($\text{LOGFEC} = \ln(\text{FAC}+14)$) di 4302 pecore.

Di seguito le statistiche descrittive delle variabili considerate:

Carattere	N	Media	Dev. St.	Min	Max
LEPM	15001	217	55	20	419
TP	14990	5.16	0.41	2.71	7.51
LOGFEC	16964	4.82	1.43	2.64	9.96

Per la stima delle componenti delle varianze e dei valori genomici si è utilizzato il single-step Genomic BLUP con modelli a misure ripetute. Per la produzione di latte e il tenore in proteina in lattazione il modello è:

$$Y = \text{GACS} + \text{ACNM} + A + \text{APA} + E$$

dove Y è uguale a LEPM o TP, GACS è l'effetto fisso dato dall'interazione fra annata e classe di età, ACNM è l'effetto fisso dato dall'interazione fra l'annata, classe di età e epoca di parto, A è l'effetto casuale genetico additivo, APA è l'effetto casuale ambiente permanente animale e E è l'effetto casuale errore.

Per il Contenuto individuale di Uova Fecali il modello è:

$$\text{LOGFEC} = \text{DC} + \text{ETA} + \text{NA*GP} + A + \text{APA} + E$$

dove DC è l'effetto fisso data di campionamento, ETA è l'effetto fisso età dell'animale, NA*GP è l'effetto fisso dato dall'interazione tra numero di agnelli portati e la distanza in giorni dal parto, A è l'effetto casuale genetico additivo, APA è l'effetto casuale ambiente permanente animale e E è l'effetto casuale errore.

Il file delle parentele è stato ricavato dall'archivio anagrafico di ASSO.NA.PA. risalendo fino ai bisnonni degli animali genotipati per un totale di 15853 animali.

Il file dei genotipi comprendeva 7031 animali di cui 5383 pecore e 1648 arieti.

Le analisi genomiche sono state realizzate utilizzando la suite di programmi BLUPf90 (Miszta et al., 2002).

Nella tabella seguente si riportano le stime delle componenti della varianza, l'ereditabilità e la ripetibilità per i caratteri presi in considerazione.

Tabella 2 – Varianza Genetica additiva (Va), ambiente permanente animale (Vapa), residua (Ve) e totale (Vtot), ereditabilità (h²) e ripetibilità (r) per quantità di latte in lattazione (LEPM), tenore proteico in lattazione (TP) e contenuto in uova fecali (FEC, resistenza ai nematodi)

Carattere	Va	Vapa	Ve	Vtot	h ²	r
LEPM	693	535	906	2134	0.32	0.58
TP	0.065	0.011	0.030	0.107	0.61	0.72
FEC	0.38	0.09	1.19	1.66	0.23	0.28

I valori genomici (GEBV) sono stati calcolati per 7031 capi di cui 5040 pecore del gregge genomico, 896 pecore allevate in 3 allevamenti iscritti al LG e 1648 arieti. Di quest'ultimi, 299 sono padri o antenati delle pecore del gregge genomico mentre 1349 sono arieti di allevamenti iscritti al LG.

Di questi animali 1489 fanno parte del campione analizzato nell'ambito del progetto CHEESR di cui 408 pecore del gregge genomico e 1081 arieti del LG. Altri 9 campioni di arieti facevano parte di quest'ultimo campione ma non sono stati inseriti nella valutazione genomica per problemi legati alla qualità dell'analisi del DNA. Al totale dei 1489 capi genotipizzati nell'ambito del progetto vanno aggiunte le analisi genomiche che sono state ripetute in doppio su 6 soggetti.

Tabella 3 – Media e deviazione standard dei valori genomici stimati per quantità di latte in lattazione (LEPM), tenore proteico in lattazione (TP) e contenuto in uova fecali (FEC, resistenza ai nematodi)

	N	GEBV LEPM	GEBV TP	GEBV FEC
Femmine	5383	0.23 ± 20.62	0.01 ± 0.22	-0.07 ± 0.46
Allev. Bonassai	127	13.74 ± 12.25	-0.17 ± 0.14	0.11 ± 0.25
Allev esterno 1	121	0.98 ± 9.22	-0.01 ± 0.11	0.42 ± 0.20
Allev. esterno 2	95	4.83 ± 7.69	-0.03 ± 0.12	0.30 ± 0.20
Gregge genomico	5040	-0.22 ± 21.02	0.02 ± 0.23	-0.10 ± 0.46
Maschi	1648	1.42 ± 12.69	-0.07 ± 0.15	0.22 ± 0.35
Giovani Arieti LG (=>2017)	688	1.14 ± 10.58	-0.08 ± 0.13	0.30 ± 0.26
Vecchi Arieti LG (<2017)	661	0.51 ± 10.68	-0.06 ± 0.13	0.25 ± 0.33
Arieti Gregge genomico	299	4.09 ± 19.24	-0.05 ± 0.22	-0.02 ± 0.46
Totale	7031	0.51 ± 19.07	-0.01 ± 0.21	0.00 ± 0.45

Nel file allegato alla presente relazione si riportano i dati individuali dei valori genomici e la *reliability* di tutti gli animali genotipizzati (analisi con chip).

Prospettive

La sperimentazione ha dimostrato che l'applicazione della genomica è una concreta opportunità per incrementare l'efficienza e la sostenibilità economica dello schema di selezione della razza Sarda. L'approccio proposto consente la stima del valore genomico con sufficiente accuratezza non solo per la produzione del latte ma per tutti i caratteri misurati nella PRF con precisione pari o superiore a quella di un pedigree tradizionale per la produzione del latte. Per i caratteri misurati anche nelle aziende iscritte al LG sono attese accuratezze superiori grazie alla possibilità di cumulare le informazioni derivanti da PRF con quelle delle popolazioni di riferimento maschili. È cruciale inoltre mantenere una accurata registrazione delle parentele per regolare il flusso dei riproduttori da e verso PRF e selezionare i giovani maschi per i quali è conveniente realizzare l'analisi genomica. In quest'ottica sarebbe auspicabile avviare il programma di attribuzione delle parentele utilizzando le informazioni del DNA che consentirebbe agli allevatori di ridurre l'impegno della gestione dei gruppi di monta, incrementando nel contempo la percentuale di genealogie conosciute.

Sassari, 2 giugno 2021

Il Responsabile Scientifico AGRIS Sardegna
Dottor Antonello Carta

