

Utilizzo di informazioni genomiche ai fini del monitoraggio della diversità genetica

Conservation, Health and Efficiency Empowerment of Small Ruminant (CHEESR)

Nell'ambito dell'azione 6.1 – Step 3, le informazioni genomiche ottenute da 3343 soggetti di specie caprina (appartenenti alle razze Camosciata delle Alpi, Garganica, Nicastrese e Saanen) e 2820 di specie ovina (appartenenti alle razze Comisana, Massese, Delle Langhe, Fabrianese, Istriana – Carsolina, Gentile di Puglia e Sarda) sono state utilizzate per il calcolo dei parametri genomici, quali l'inbreeding genomico e la distanza genetica media tra le razze, ed è stata calcolata la correlazione tra inbreeding stimato da pedigree e inbreeding genomico.

Nello specifico, nella specie caprina i soggetti genotipizzati erano ripartiti nelle diverse razze come mostrato nella Tabella 1. Il dataset delle razze caprine è stato sottoposto a Quality Control (QC) e le analisi sono state svolte su un dataset conforme ai seguenti requisiti qualitativi: percentuale di SNP mancanti per genotipo inferiore al 5% e percentuale di SNP mancanti per soggetto inferiore al 5%.

Razza	Soggetti pre-QC	Soggetti post-QC
CAMOSCIATA DELLE ALPI	2178	2161
GARGANICA	114	114
NICASTRESE	101	100
SAANEN	950	945
TOTALE	3343	3320

Tabella 1. Numero di soggetti di specie caprina genotipizzati, prima e dopo l'applicazione del Quality Control (QC)

Per ogni individuo, sono state calcolate le Runs of Homozygosity (ROH) utilizzando i seguenti parametri del software PLINK 1.9:

- homozyg-window-het pari a 0
- homozyg-snp pari a 20
- homozyg-kb pari a 1000
- homozyg-gap pari a 500
- homozyg-density pari a 100
- homozyg-window-snp pari a 20
- homozyg-window-missing pari a 1

Per quanto riguarda, invece, l'inbreeding genomico (F_{ROH}) è stato calcolato a partire dalla formula di MacQuillan et al. 2008 riportata di seguito:

$$F_{ROH} = \frac{L_{ROH}}{L_{AUTO}}$$

dove L_{ROH} è la lunghezza delle regioni di ROH del singolo individuo mentre L_{AUTO} è la lunghezza totale degli autosomi coperti da SNP.

I valori di F_{ROH} medi per razza, insieme ai valori medi di numero di ROH e lunghezza media e totale delle ROH, sono riportati nella Tabella 3. Invece, in Figura 1 viene riportata la distribuzione dei valori di F_{ROH} nei soggetti analizzati, suddivisi per razza.

L'inbreeding genomico è stato valutato anche in funzione delle classi di ROH che lo hanno generato ed in questo modo è stato possibile datare, in termini di numero di generazioni, eventi selettivi passati. Infatti, la porzione di F_{ROH} relativa alle ROH più lunghe è dovuta ad eventi di inbreeding più recenti; al contrario, a ROH più corte sono associati eventi di inbreeding più antichi. Come riferimento è stato preso il lavoro di Onzima et al. 2018 dove gli autori hanno proposto una valutazione delle ROH per classi di lunghezza, da 2 a 16 Mega basi (Mb), riuscendo a datare l'inbreeding da 3 fino a 25 generazioni fa. Le classi valutate nel progetto CHEESR e la datazione in termini di generazioni è riportata in Tabella 2.

Classe di lunghezza delle ROH	Datazione inbreeding
$ROH \leq 2 \text{ Mb}$	~ 25 generazioni fa
$2 \text{ Mb} < ROH \leq 4 \text{ Mb}$	$\sim 25 - 12$ generazioni fa
$4 \text{ Mb} < ROH \leq 8 \text{ Mb}$	$\sim 12 - 6$ generazioni fa
$8 \text{ Mb} < ROH \leq 16 \text{ Mb}$	$\sim 6 - 3$ generazioni fa
$ROH > 16 \text{ Mb}$	< 3 generazioni fa

Tabella 2. Classi di ROH valutate e relativa datazione in termini di generazioni.

Nella Figura 2, viene rappresentata la distribuzione dei valori medi di F_{ROH} per classe di lunghezza delle ROH.

Da questi risultati si osserva che la razza Garganica presenta, complessivamente, il valore medio di F_{ROH} più elevato (0.077) e dovuto ad eventi più recenti in quanto prevale la componente di F_{ROH}

relativa a ROH di lunghezza superiore a 16 Mb; mentre la Nicastrese presenta il valore medio di F_{ROH} più basso (0.036) ed equamente ripartito nelle diverse classi di ROH. La razza Saanen e la Camosciata delle Alpi presentano un valore di F_{ROH} medio rispettivamente pari a 0.071 e 0.065 con una porzione di inbreeding uniformemente distribuita tra le varie classi di ROH considerate ad eccezione della classe di ROH più lunga (>16 Mb), rappresentata dal colore giallo, in cui la razza Saanen presenta un inbreeding superiore alla razza Camosciata, indice di un aumento dell'inbreeding in tempi recenti. In tutte le razze caprine analizzate, il 75% dei soggetti (intervallo interquartile) non presenta un valore di F_{ROH} superiore a 0.1; solo un numero ridotto di soggetti, appartenenti alla razza Camosciata delle Alpi, ha valori superiori a 0.3.

Razza	Numero ROH	Lunghezza media	Lunghezza totale	F_{ROH}
CAMOSCIATA DELLE ALPI	36.137	160.587	4.275	0.065
GARGANICA	29.289	189.246	5.295	0.077
NICASTRESE	20.080	87.718	3.505	0.036
SAANEN	35.795	173.550	4.723	0.071

Tabella 3. Numero, lunghezza media e totale delle ROH e inbreeding genomico (F_{ROH}) medio nelle razze caprine genotipizzate.

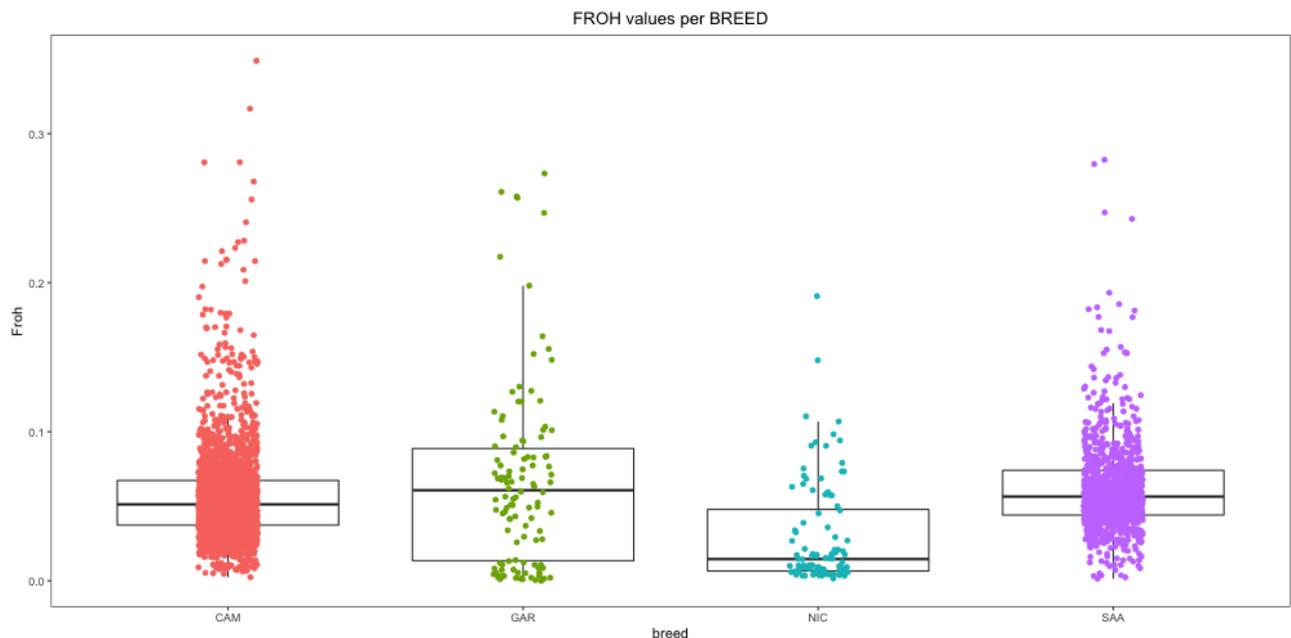


Figura 1. Distribuzione dei valori di inbreeding genomico (F_{ROH}) nei soggetti caprini genotipizzati. CAM = Camosciata delle Alpi, GAR = Garganica, NIC = Nicastrese, SAA = Saanen.

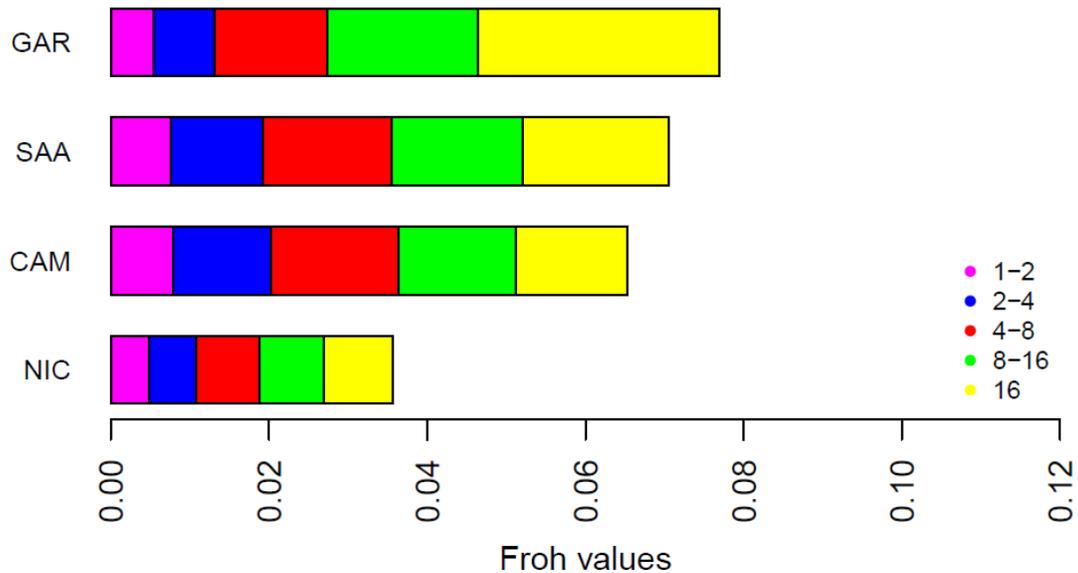


Figura 2. Distribuzione dei valori di inbreeding genomico (F_{ROH}) medi delle razze caprine genotipizzate; ogni colore rappresenta una diversa classe di lunghezza di ROH (in giallo la classe $ROH > 16 Mb$, in verde la classe $8 Mb < ROH \leq 16 Mb$, in rosso la classe $4 Mb < ROH \leq 8 Mb$, in blu la classe $2 Mb < ROH \leq 4 Mb$ e in fucsia la classe $ROH \leq 2 Mb$). CAM = Camosciata delle Alpi, GAR = Garganica, NIC = Nicastrese, SAA = Saanen.

Successivamente è stata calcolata la correlazione tra inbreeding genomico F_{ROH} e inbreeding da pedigree F_{PED} (Tabella 4). I risultati relativi alla razza Nicastrese non sono mostrati in quanto i soggetti campionati ad essa appartenenti presentavano valori di F_{PED} pari a 0; mentre, i risultati relativi alla razza Garganica non sono stati riportati in quanto non significativi. La correlazione è risultata altamente significativa per le razze Camosciata delle Alpi e Saanen, con un coefficiente di correlazione pari a circa 0.30.

Razza	Correlazione $F_{ROH} - F_{PED}$	P-value
CAMOSCIATA DELLE ALPI	0.26	< 0.0001
SAANEN	0.31	<0.0001

Tabella 4. Correlazione tra inbreeding genomico (F_{ROH}) e da pedigree (F_{PED}) nelle razze caprine genotipizzate.

Al fine di stimare le distanze genetiche tra le diverse razze genotipizzate, sono state calcolate le distanze di Reynolds (Reynolds et al., 1983), che misurano la divergenza genetica sulla base dei coefficienti di *coancestry*. La matrice così ottenuta è riportata nella Tabella 5. Come atteso, si osserva una maggior vicinanza genetica delle razze autoctone del Sud Italia, Garganica e Nicastrese (0.098); e delle razze alpine, Saanen e Camosciata (0.122). Le due razze più lontane sono risultate la razza Saanen e la razza Garganica (0.139).

	CAMOSCIATA DELLE ALPI	GARGANICA	NICASTRESE	SAANEN
CAMOSCIATA DELLE ALPI	0	0.137	0.136	0.122
GARGANICA	0.137	0	0.098	0.139
NICASTRESE	0.136	0.098	0	0.137
SAANEN	0.122	0.139	0.137	0

Tabella 5. Matrice delle distanze di Reynolds tra le razze caprine genotipizzate.

Per quanto riguarda la specie ovina, viene riportato in Tabella 6 il numero di soggetti genotipizzati e analizzati per razza. Come per la specie caprina, il dataset delle razze ovine è stato sottoposto a Quality Control (QC) e le analisi sono state svolte su un dataset conforme ai seguenti requisiti qualitativi: percentuale di SNP mancanti per genotipo inferiore al 5% e percentuale di SNP mancanti per soggetto inferiore al 5%.

Razza	Soggetti pre-QC	Soggetti post-QC
COMISANA	541	536
FABRIANESE	108	107
GENTILE DI PUGLIA	107	105
ISTRIANA-CARSOLINA	81	81
MASSESE	375	371
OVINO DELLE LANGHE	104	104
SARDA	1504	1486
TOTALE	2820	2790

Tabella 6. Numero di soggetti delle razze ovine genotipizzati, prima e dopo l'applicazione del Quality Control (QC).

Per tutte le razze ovine analizzate è stato calcolato l'inbreeding genomico medio e per classe di lunghezza di ROH, la correlazione tra l'inbreeding genomico e l'inbreeding stimato da pedigree e le distanze di Reynolds. Nella Tabella 7 vengono riportati la descrizione del numero e della lunghezza delle ROH e i coefficienti di inbreeding genomico (F_{ROH}) medi per razza. Nelle Figure 3 e 4 si possono osservare, rispettivamente, la distribuzione dei valori di F_{ROH} nei soggetti analizzati e dei valori di F_{ROH} medi per razza, suddivisi a seconda delle classi di lunghezza delle ROH. I risultati indicano che la razza Istriana – Carsolina presenta, mediamente, un valore di F_{ROH} più alto (0.141) seguita dalle razze Sarda (0.134), Fabrianese (0.128), Delle Langhe (0.118), Massese (0.078), Comisana (0.51) e Gentile di Puglia (0.045), che è risultata avere il valore medio di F_{ROH} più basso. Analizzando l'inbreeding genomico per classe di lunghezza di ROH maggiore di 16 Mb (imputabile all'attività selettiva più recente), rappresentato dal colore giallo, è possibile notare che le razze Gentile di Puglia (0.010), Comisana (0.004) e Massese (0.008) presentano un inbreeding ridotto rispetto a tutte le altre razze. Invece le razze Delle Langhe e Sarda presentano un inbreeding recente rispettivamente pari a 0.021 e pari a 0.026, il quale risulta leggermente più contenuto se confrontato con l'inbreeding delle

due classi precedenti. L'inbreeding recente più elevato è risultato per la Fabrianese e la razza Istriana – Carsolina che presentano dei valori rispettivamente pari a 0.061 e 0.051. Per il 75% dei soggetti di tutte le razze è stato calcolato un F_{ROH} inferiore a 0.2 e solamente un soggetto di razza Fabrianese presenta un valore superiore a 0.4.

Razza	Numero ROH	Lunghezza media	Lunghezza totale	F_{ROH}
COMISANA	40.159	125.312	3.071	0.051
FABRIANESE	40.411	311.737	6.600	0.128
GENTILE DI PUGLIA	27.267	109.966	3.621	0.045
ISTRIANA-CARSOLINA	52.827	343.599	6.025	0.141
MASSESE	56.097	189.754	3.371	0.078
OVINO DELLE LANGHE	66.558	288.171	4.263	0.118
SARDA	71.863	327.903	4.340	0.134

Tabella 7. Numero, lunghezza media e totale delle ROH e inbreeding genomico (F_{ROH}) medio nelle razze ovine genotipizzate.

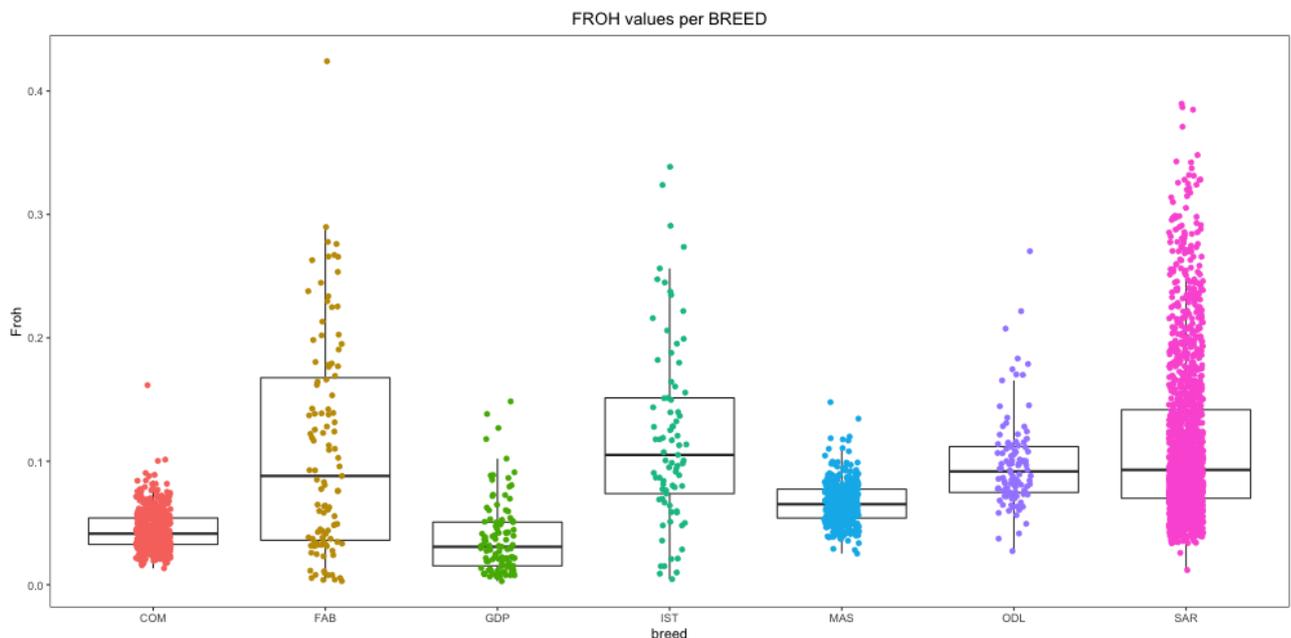


Figura 3. Distribuzione dei valori di inbreeding genomico (F_{ROH}) nei soggetti ovini genotipizzati. COM = Comisana, FAB = Fabrianese, GDP = Gentile di Puglia, IST = Istriana-Carsolina, MAS = Massese, ODL = Ovino delle Langhe, SAR = Sarda.

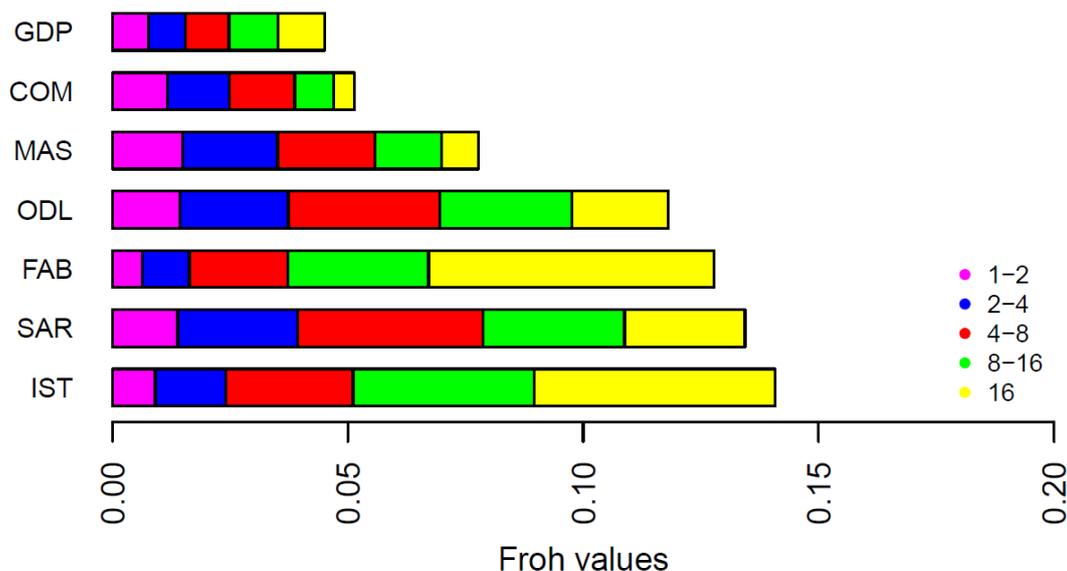


Figura 4. Distribuzione dei valori di inbreeding genomico (F_{ROH}) medi delle razze ovine genotipizzate; ogni colore rappresenta una diversa classe di lunghezza di ROH (in giallo la classe $ROH > 16 Mb$, in verde la classe $8 Mb < ROH \leq 16 Mb$, in rosso la classe $4 Mb < ROH \leq 8 Mb$, in blu la classe $2 Mb < ROH \leq 4 Mb$ e in fucsia la classe $ROH \leq 2 Mb$). COM = Comisana, FAB = Fabrianese, GDP = Gentile di Puglia, IST = Istriana-Carsolina, MAS = Massese, ODL = Ovino delle Langhe, SAR = Sarda.

La Tabella 8 mostra i risultati dell'analisi della correlazione tra i valori di F_{PED} e F_{ROH} . La correlazione è risultata altamente significativa per le razze Comisana, Massese, Delle Langhe e Sarda, con un coefficiente che variava da 0.32 nella razza Massese fino a 0.80 nella razza Sarda. I due coefficienti di inbreeding non sono risultati significativamente correlati nelle razze Fabrianese, Gentile di Puglia e Istriana e pertanto non vengono riportati i risultati in tabella. Le diverse correlazioni riscontrate tra razze e tra specie dipendono dalla variabilità nella profondità e completezza dei pedigree analizzati. Infatti, quando nessuno o pochi ascendenti sono noti F_{PED} risulta essere sottostimato; al contrario, con pedigree completi e profondi, come quello della razza Sarda, i valori di F_{PED} risultano essere molto più simili a quelli di F_{ROH} .

Razza	Correlazione $F_{ROH} - F_{PED}$	P-value
COMISANA	0.38	< 0.0001
MASSESE	0.32	<0.0001
OVINO DELLE LANGHE	0.44	<0.0001
SARDA	0.80	<0.0001

Tabella 8. Correlazione tra inbreeding genomico (F_{ROH}) e da pedigree (F_{PED}) nelle razze caprine genotipizzate.

Infine, la matrice delle distanze di Reynolds misurate tra le diverse razze ovine genotipizzate è riportata nella Tabella 9. Le razze risultate tra loro più vicine sono la razza Comisana e la razza Gentile di Puglia (0.122); mentre le più lontane sono risultate la razza Istriana – Carsolina e la razza Delle Langhe (0.190). Anche in questo caso emerge una maggior vicinanza genetica legata alla localizzazione geografica delle diverse razze.

	COMISANA	FABRIANESE	GENTILE DI PUGLIA	ISTRIANA-CARSOLINA	MASSESE	OVINO DELLE LANGHE	SARDA
COMISANA	0.000	0.139	0.122	0.162	0.147	0.172	0.148
FABRIANESE	0.139	0.000	0.126	0.161	0.146	0.173	0.156
GENTILE DI PUGLIA	0.122	0.126	0.000	0.153	0.136	0.162	0.142
ISTRIANA-CARSOLINA	0.162	0.161	0.153	0.000	0.170	0.191	0.178
MASSESE	0.147	0.146	0.136	0.170	0.000	0.180	0.163
OVINO DELLE LANGHE	0.172	0.173	0.162	0.191	0.180	0.000	0.185
SARDA	0.148	0.156	0.142	0.178	0.163	0.185	0.000

Tabella 9. Matrice delle distanze di Reynolds tra le razze ovine genotipizzate.

Con il progetto CHEESR, Asso.Na.Pa. per la prima volta ha avuto l'opportunità di raccogliere e analizzare dati genomici appartenenti a diverse razze ovicaprine, allevate in diversi allevamenti localizzati in tutta Italia. I risultati presentati in questa relazione rappresentano un'importante svolta per il settore ovicaprino in quanto, attraverso l'analisi e il confronto dei dati di pedigree e dei dati genomici, è possibile valutare e trarre utili indicazioni riguardo la gestione delle popolazioni.