



UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI

Population structure and signatures of selection in Italian Holstein through genome-wide analysis of imputed SNP data

Christian Persichilli^{1*}, Maurizio Marusi², Jan-Thijs van Kaam², Raffaella Finocchiaro², Gabriele Senczuk¹, Marika Di Civita¹, Fabio Pilla¹, Martino Cassandro^{2,3}

¹Department of Agriculture Environment and Food Science, University of Molise, Via F. De sanctis, 86100 CB Campobasso, Italy; ²Italian Holstein, Brown and Jersey Breeders Association, Via Bergamo, 292, 26100 CR Cremona, Italy; ³Department of Agronomy, Food, Natural resources, Animals and Environment (DAFNAE), University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy

*c.persichilli@studenti.unimol.it

Introduction



UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI



Among Italian dairy cattle, the Holstein is the most reared breed to produce Parmigiano Reggiano PDO cheese, which represents one of the most renowned products in the entire Italian dairy industry.

For this reason, ANAFIBJ defined the Cheesemaking and Sustainability Index – Parmigiano Reggiano (ICS-PR)

Aims



UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI



1. Explore the genetic structure of the Italian Holstein:
 - Assess its distinctiveness from the North American population
 - Identify possible distinctiveness of Italian Holstein selected for Parmigiano Reggiano production.
2. Perform selection signature analyses to identify possible genomic regions carrying important production traits.

Materials & Methods

Data source & Quality control



UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI



280 Genotypes from ANAFIBJ database:

- 101 **North-American Holstein**
- 74 **Italian Holstein**
- 105 **Italian Holstein from Parmigiano Reggiano area**

Genotypes from
multiple sources
(8k-200k)



Imputation (PedImpute)
86,841 SNPs



Data filtering (maf 0.05 &
marker missingness 0.01)
79,464 SNPs

Materials & Methods

Genetic diversity & Population structure



UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI



- Genomic diversity indices (F_{ROH} ; H_e ; H_o ; N_e)
- MDS analyses
- Admixture analyses (k2 – k5)



Materials & Methods

Signatures of selection



UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI



F_{ST} analysis with two approaches:

- Genome-wide F_{ST} analysis
- window-based F_{ST} analysis

Threshold chosen is the **top 0.05% F_{ST}** value (99.95th percentile) for both the analyses

In the **window-based** method only **windows containing more than 4 SNP** have been considered

For both methods we looked for **annotated genes** and we investigated **QTLs** using the cattle database (ARS-UCD1.2; QTLdb)

Diversity indices



UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI



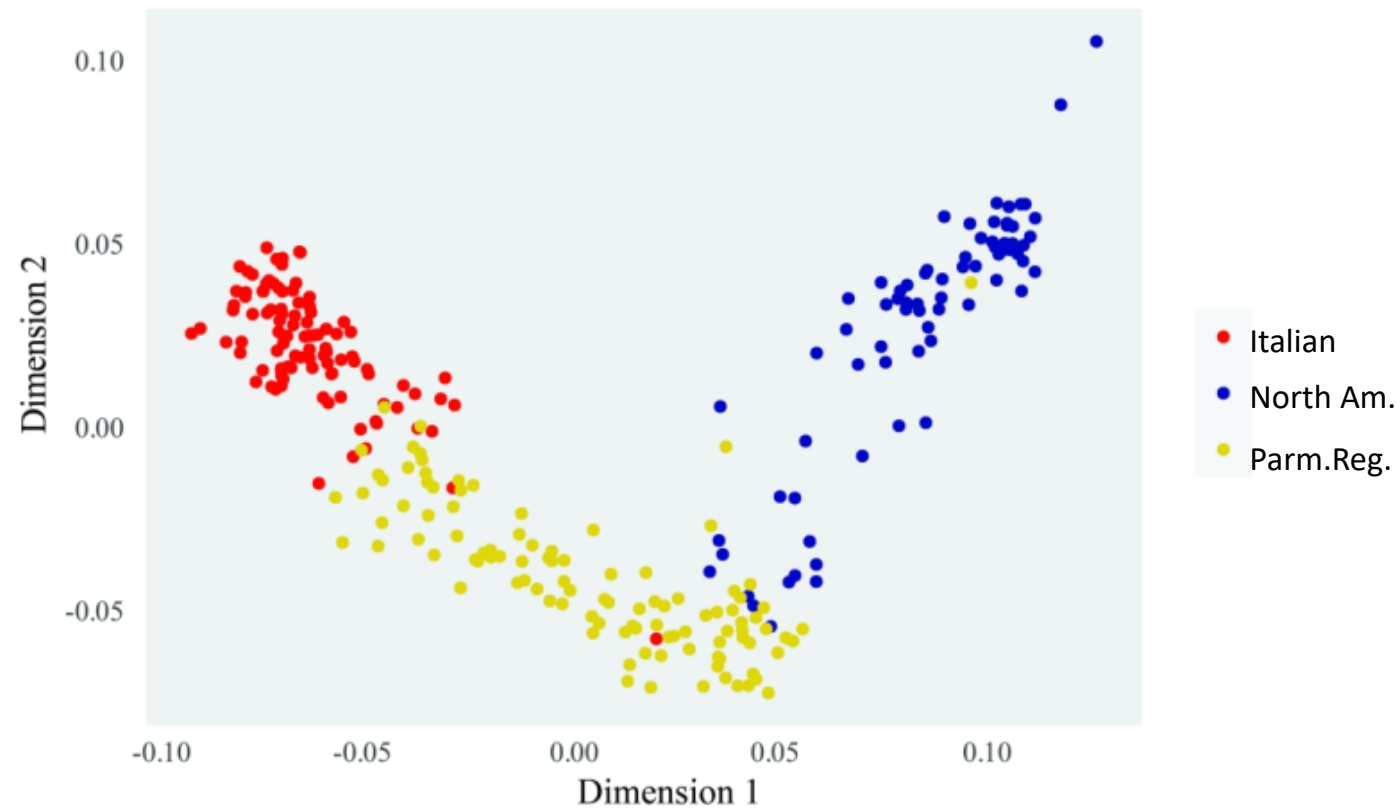
Population	Code	N°	Ho ± SD	He ± SD	F_{ROH}	N_e
Italian Holstein	itHOL	74	0.373 ± 0.160	0.354 ± 0.144	0.102	43.39
Parmigiano Reggiano Holstein	prHOL	105	0.363 ± 0.138	0.368 ± 0.135	0.125	89.05
North American Holstein	acHOL	101	0.342 ± 0.171	0.332 ± 0.160	0.168	39.97

Multi-Dimensional Scaling Analysis



UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI



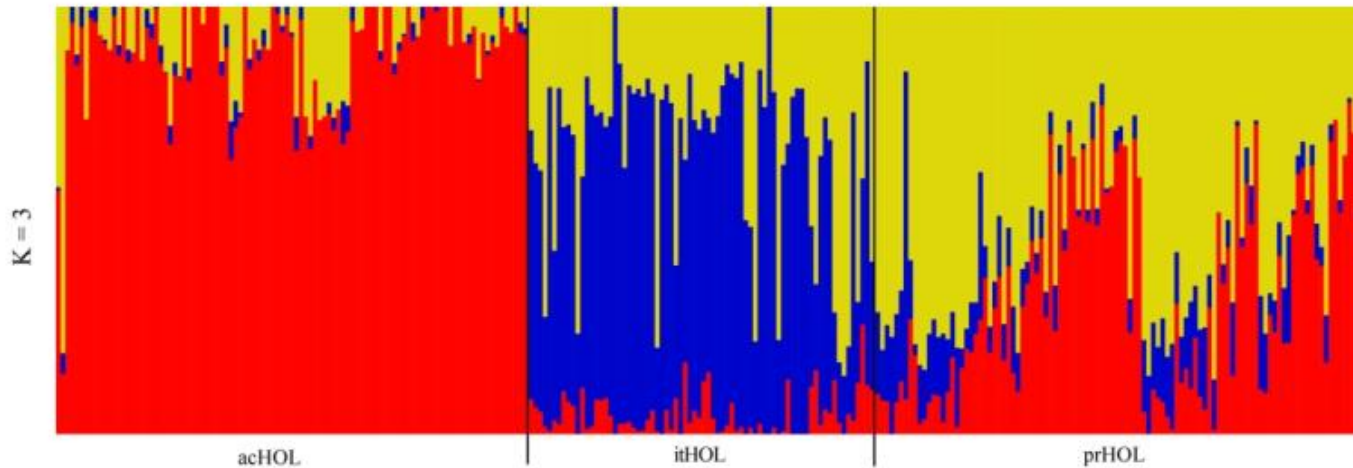
The Multi-dimensional scaling analysis shows a **clear differentiation** among the three sub-population.

Admixture



UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI



ADMIXTURE analysis at the most reliable number of K value ($K = 3$).

North-American and *Italian* samples result to be **highly differentiated**

The *Parmigiano* area sub-population results to **have its own structure** even if it is not as clearly defined as the others.

Signatures of selection

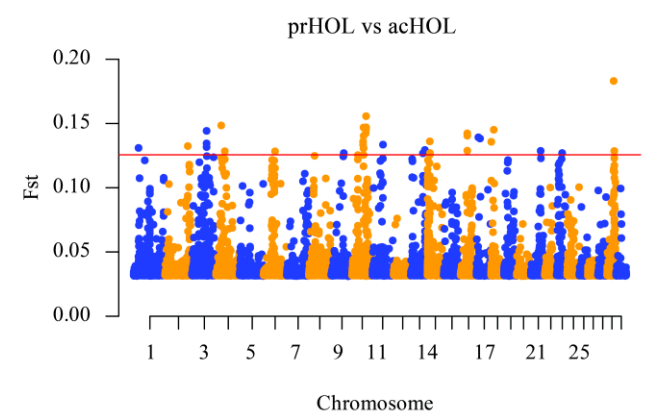
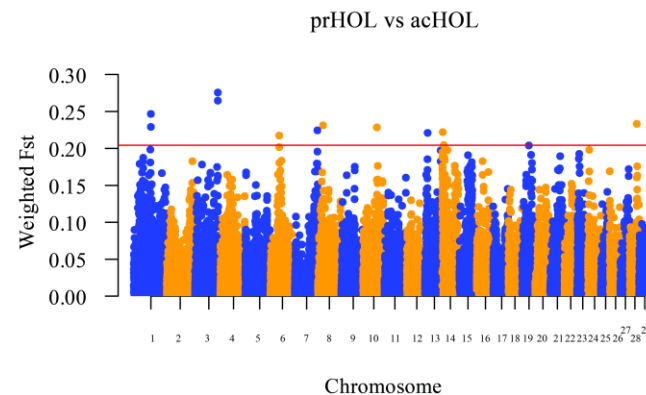
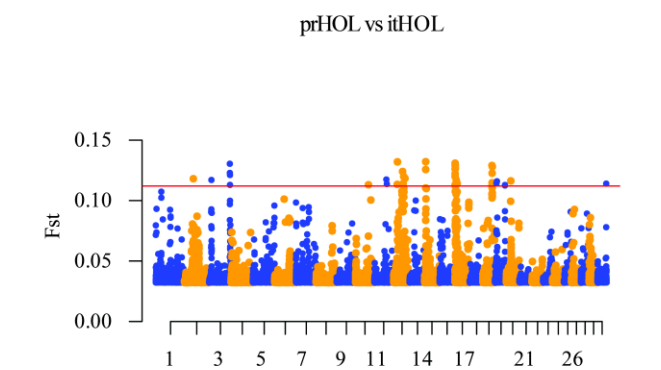
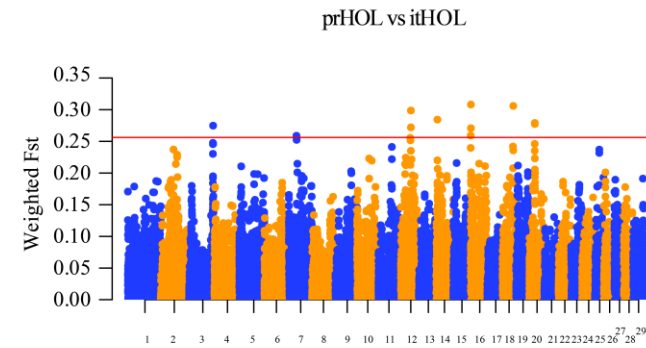
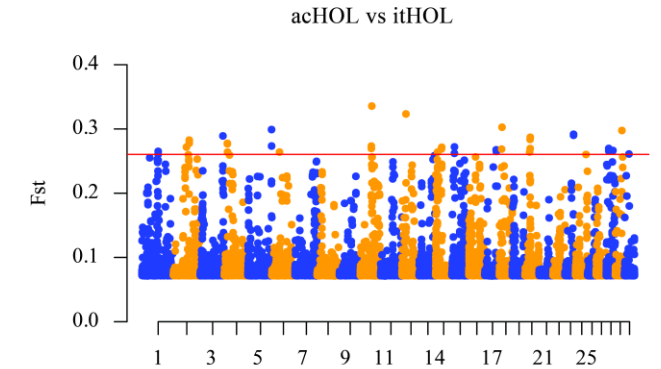
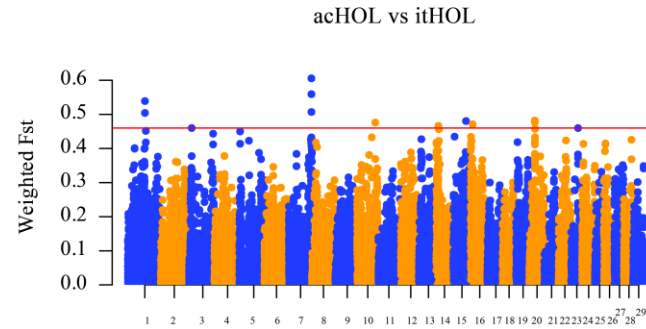
BayeScan analysis **40 top markers** for each comparison involving:

- **itHOL vs acHOL 25 genes** in 12 chromosomes
- **prHOL vs acHOL 27 genes** in 9 chromosomes
- **prHOL vs itHOL 43 genes** in 8 chromosomes.

In the window-based analysis:

12 significant windows for each comparison

- **itHOL vs acHOL 10 windows**, 9 genes in 7 chromosomes,
- **prHOL vs acHOL 8 windows** 8 genes in 5 chromosomes
- **prHOL vs itHOL 6 windows** 15 genes located in 4 chromosomes.



Conclusions

Genetic diversity & Population structure

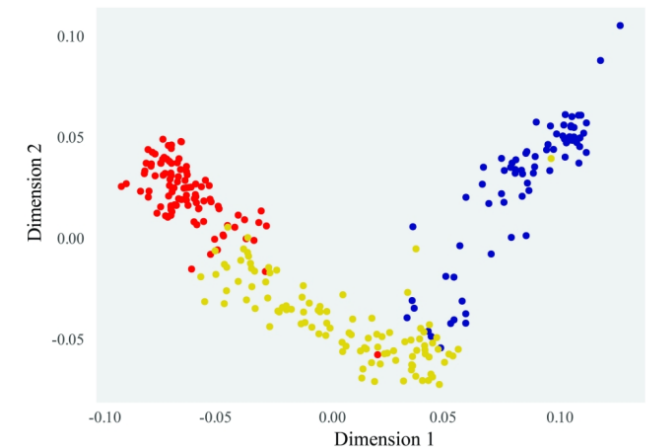


UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI



- **Italian Holstein** results **differentiated** by the North-American population.
- Inside the Italian population a **further differentiation** is evident **based on** the geographical area and **production purpose** (Parmigiano Reggiano)
- **Italian Holstein** in general results to be **more diverse** and **less inbred** than the North-American population, but the sub-population from the **Parmigiano Reggiano** area has even **lower inbreeding** values and **higher diversity** indices (Ne; He)



Population	Code	N°	Ho ± SD	He ± SD	F _{ROU}	Ne
Italian Holstein	itHOL	74	0.373 ± 0.160	0.354 ± 0.144	0.102	43.39
Parmigiano Reggiano Holstein	prHOL	105	0.363 ± 0.138	0.368 ± 0.135	0.125	89.05
North American Holstein	acHOL	101	0.342 ± 0.171	0.332 ± 0.160	0.168	39.97

Conclusions

Signatures of selection



UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

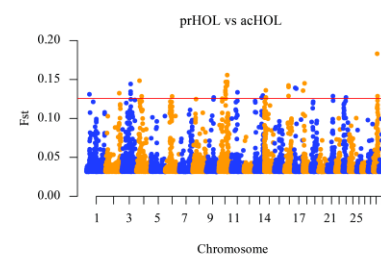
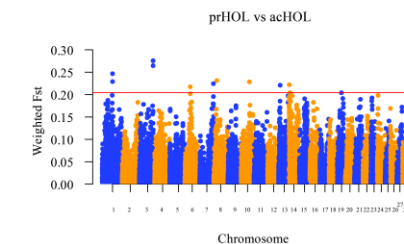
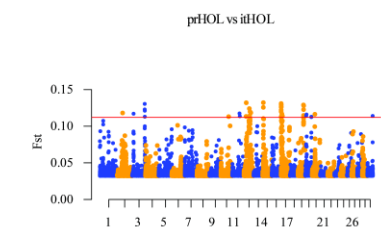
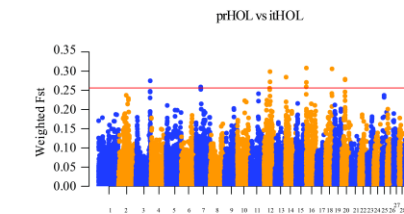
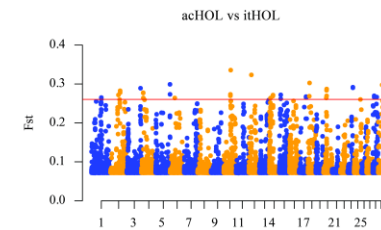
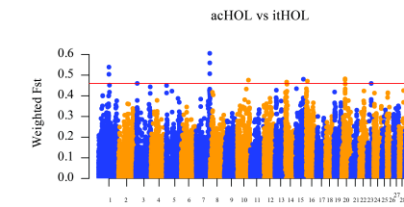
DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI



The Bayescan analysis, identified 55 genes having a role in productive traits, including 17 genes related to milk production (*GSK3A*, *SUPT4H1*, *RNF43*, *VPS8*, *KSR2*, *ZFPM1*, ***KCNMA1***, *TMEFF2*, *CYP7B1*, *RELN*, *ZFPM1*, *HDAC9*, *LIPE*, *FBXO33*, *C4A*, *IRS1* and *KIAA0586*).

The windows-based F_{ST} analysis highlighted 29 genes, 15 of which have been previously detected to be associated with productive traits including milk yield and quality (*DCDC1*, ***VPS8***, *PTP4A3*, *ST3GAL1*, *GPR20* and *SLC24A2*).

Nine genes showed concordant results between the two approaches (*VPS8*, *CHIT1*, *MYBPH*, *CHI3L1*, *PPP1R15B*, *PIK3C2B*, *ZNF404*, *ZNF45* and *ZNF226*). Most of them linked to dairy traits.



Conclusions

Signatures of selection



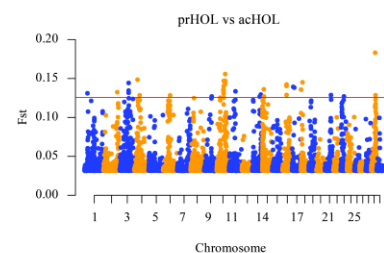
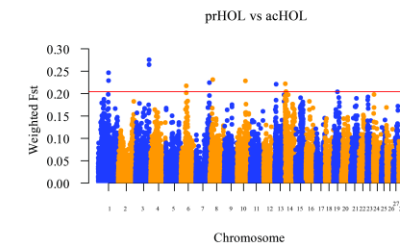
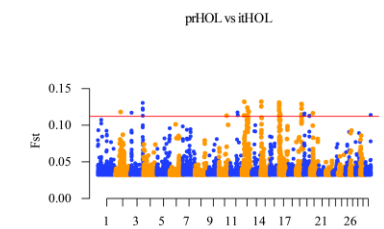
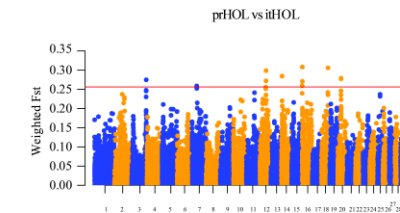
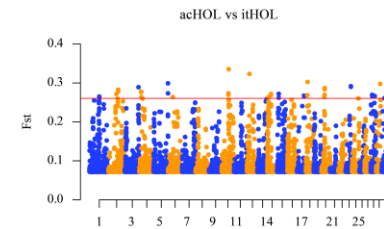
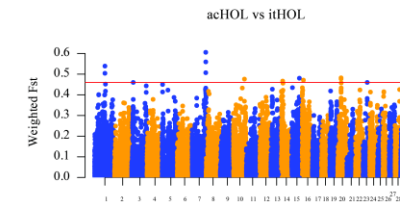
UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI



Most of the genes involved milk traits that overlapped between the two approaches, were between the two Italian sub-populations and the North-American Holstein.

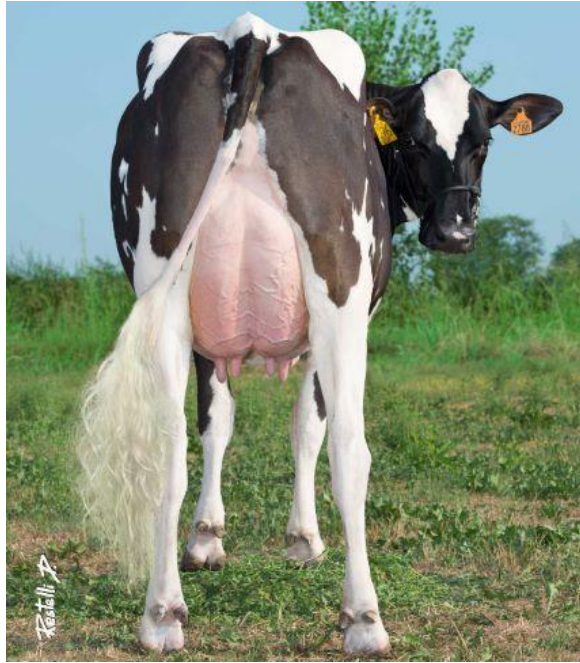
Future directions are to replicate the analyses using an haplotypic approach



Credits

Special thanks should be given to ANAFIBJ for making the dataset available and for the imputation of the genotypes





Thank you for your
kind attention



UNIVERSITÀ
DEGLI STUDI
DEL MOLISE

DIPARTIMENTO DI AGRICOLTURA,
AMBIENTE E ALIMENTI

