Università degli studi di Padova

Dipartimento di Medicina Animale, Produzioni e Salute

Corso di Laurea Magistrale a ciclo unico in

Medicina Veterinaria

South Tyrolean milk authentication by direct analysis in real time (DART)

Laureanda

Claudia Simeoni

Matricola 1120448

Relatore

Prof. Severino Segato

Correlatrice

Dott.ssa Ilaria Lanza

Anno Accademico 2020-2021

Università degli studi di Padova

Dipartimento di Medicina Animale, Produzioni e Salute

Corso di Laurea Magistrale a ciclo unico in

Medicina Veterinaria

Autenticazione di latte dell'Alto Adige mediante *direct analysis in real time* (DART)

Laureanda

Claudia Simeoni

Matricola 1120448

Relatore

Prof. Severino Segato

Correlatrice

Dott.ssa Ilaria Lanza

Anno Accademico 2020-2021

A Irene e Clara

Table of contents

RIASSUNTO

La globalizzazione, il mercato libero e l'aumento del benessere umano hanno caratterizzato gli ultimi decenni e hanno portato ad una crescente attenzione verso ciò che mangiamo e verso il background produttivo delle risorse alimentari. In particolare, viene data priorità al benessere animale e alla sostenibilità ambientale. A reazione di questo contesto, il mercato mondiale ha visto nascere numerosi prodotti a "valore aggiunto", molto apprezzati dai consumatori del primo mondo per le loro particolari caratteristiche organolettiche, nutrizionali o etiche.

Poiché il valore commerciale dei prodotti certificati si basa proprio sui "valori aggiunti", le rispettive materie prime di alta qualità risultano molto vulnerabili al fenomeno dell'adulterazione quindi, a tutela della fiducia dei consumatori, sono sorte numerose etichette e certificazioni che garantiscano protocolli produttivi specifici.

L'industria lattiero-casearia, grazie al grande assortimento dei propri prodotti, si è dimostrata particolarmente adatta ad offrire articoli dal "valore aggiunto".

La determinazione dell'autenticità alimentare e il rilevamento delle contraffazioni, è divenuto un problema di rilievo all'interno di questo settore ed è di notevole preoccupazione per consumatori, riveditori e per le stesse autorità di vigilanza.

Di conseguenza, consumatori e produttori, necessitano dello sviluppo di tecniche analitiche affidabili ed omnicomprensive che possano autenticare le informazioni riportate in etichetta.

In questo contesto, lo studio metabolomico può dare informazioni dettagliate della composizione alimentare, fornendo un ampio screening dei metaboliti anche in matrici alimentari complesse quali il latte.

La *direct analysis in real time*, associata alla spettrometria di massa ad alta risoluzione (DART-HRMS) è tra le tecniche che si utilizzano per ottenere un'impronta digitale metabolica del latte.

Questo studio, quindi, si pone come scopo l'identificazione di particolari biomarcatori del latte, riconducibili a specifici sistemi alimentari delle vacche.

Per ottenere questo, sono stati eseguiti 88 campionamenti di latte crudo di massa (cisterna), prodotto in allevamenti provenienti da due aree geografiche differenti, le quali

adottavano cinque sistemi alimentari (dai quali derivano i cinque gruppi presi in esame). I primi quattro gruppi appartenevano ad un'area di pianura e i loro sistemi alimentari erano principalmente composti da: i) HMS elevato uso di insilato di mais ii) MMS misto di insilato di mais e altri insilati ii) FCG erba fresca e insilato diverso dal mais e iv) HAY fieno di pascolo permanente, senza uso di insilati. L'ultimo gruppo (APS), invece, è stato campionato nella provincia dell'Alto Adige e la sua dieta era principalmente composta di v) pascolo alpino e fieno alpino. I campioni di latte sono stati diluiti in solventi polari e non polari e successivamente analizzati con la DART sia in modalità positiva che negativa. Gli spettri ottenuti, sono stati concatenati mediante *low-level* data fusion e una prima discriminazione è stata provata mediante *partial least square discriminant analysis* (PLS-DA).

A causa della scarsa capacità discriminatoria, i cinque gruppi sono stati riassortiti in tre nuovi gruppi: LLS, dalla fusione di HMS e MMS, rappresentava l'uso di insilati di pianura; LLF, fondendo FCG e HAY, rappresentava l'uso dei foraggi di pianura; mentre APS è rimasto a rappresentare il gruppo di montagna. Per discriminare i tre gruppi, è stata eseguita nuovamente la PLS-DA. Successivamente, sui 25 segnali ionici m/z più discriminanti, è stata effettuata una *hierarchical cluster analysis* (HCA) dalla quale si è ottenuta una *heatmap* delle correlazioni tra ioni discriminanti e razioni alimentari a confronto.

Dal tentativo di assegnazione dei valori m/z discriminanti, è risultato che il gruppo LLS era rappresentato maggiormente da derivati chetoacidi, creatinina, metil 2-furoato, 3-idrossi-2-metilglutarato o 2-idrossi-2-etilsuccinato, dimetilfumarato, glucosio, glucosamina, N-acetilglucosamina e acido oleico. Il gruppo LLF è stato principalmente discriminato da acetolattato, norgramina e MAG 20:2; mentre il gruppo APS da acido lattico e i MAG 16:0 e 18:0.

Una validazione esterna finale del sistema analitico è stata raggiunta utilizzando un dataset di validazione su cui è stata effettuata la *linear discriminant analysis* (LDA).

È quindi possibile affermare che la DART-HRMS, associata ad una PLS-DA, può essere un efficace strumento di autenticazione alimentare e che questo strumento può essere impiegato nella prevenzione delle frodi alimentari.

ABSTRACT

The last decades have been characterized by increasing globalization, free trade, human welfare, and consequent attention to what we eat and the backgrounds involved behind food production, with particular regard to animal welfare and environmental sustainability. In this contest, a lot of added value products have reached the global market, appreciated by "first world" consumers for their organoleptic, nutritional, and/or ethical characteristics. The dairy industry is a very thriving trade in this field, thanks to its great differentiation of products. To protect consumers' trust, certifications and labels have been designed *ad hoc* for every added-value category.

Certified products based their commercial value on regulatory specifications protocols. Foods and ingredients presenting high-value are the most vulnerable for adulteration. Determination of food authenticity and detection of adulteration in dairy products has become an important issue within the food sector and a major concern for consumers, retailers, food processors, and regulatory authorities. Thus, consumers and producers parties ask for the development of robust and comprehensively analytical techniques that could allow authentication of the label's information (e.g. feeding system).

In this context, metabolomic approach can be a useful authentication technique that provides a detailed picture of food composition by screening and profile metabolites in complex matrices. Direct analysis in real time coupled to high resolution mass spectrometry (DART-HRMS) is one of the most applied techniques for metabolic fingerprinting of milk.

This study aims to identify particular biomarkers of milk according to the specific forage feeding systems. To do so, 88 raw bulk milk samples produced by cows fed with five different diets and coming from two areas, were sampled. Four of the groups belonged to a lowland area and their diets were mainly composed of i) HMS, high maize silage; ii) MMS, mixed maize and crop silages; iii) FCG, fresh grass/crop silages, and iv) HAY lowland permanent meadow hay. The last group was APS, alpine pasture system and it was sampled on the South Tyrol area, and its diet was mainly composed of v) alpine pasture and/or alpine hay. Milk samples were diluted in both polar and non-polar solvents and analyzed in both positive and negative DART ion-mode. The obtained spectra were

pre-processed and then concatenated with a low-level data fusion approach. On the train set (70% of data), a partial least square discriminant analysis (PLS-DA) was performed to attempt a discrimination among the feeding groups. Due to poor discrimination capability, samples of the five diet groups were re-attributed to only three experimental groups: LLS (lowland silages, gathered HMS and MMS), LLF (lowland forages, gathered FCG and HAY), and APS which remained the same. The PLS-DA was performed again to discriminate the three groups. Furthermore, a hierarchical cluster analysis (HCA) was applied to the 25 most discriminative m/z signals (ions) and a heatmap (matrix of correlations among ions and dietary groups) was obtained.

LLS group was discriminated by m/z values whose tentative assignment was ketoacid derivate, creatinine, methyl 2-furoate, 3-hydroxy-2-methylglutarate or 2-hydroxy-2-ethylsuccinate, dimethyl fumarate, glucose, glucosamine, N-acetyl-glucosamine, and oleic acid. LLF feeding system was discriminated mainly by acetolactate, norgramine and MAG (20:2) while APS group by lactic acid, MAGs (16:0) and (18:0).

A validation of the statistical modeling approach was carried out by performing a linear discriminant analysis (LDA) on the test set (30% of the samples). The resulting confusion matrix showed reliable predictive performance only for the comparison between alpine and lowland milk samples, meanwhile, a relative high misclassification rate (around 0.55) was observed among samples of LLS and LLF theses.

It is possible to affirm that DART-HRMS coupled to a PLS-DA approach could be a powerful tool for food authentication even though further analyses and modeling steps need to be performed.

PREFACE

The thesis was carried out with the analytical support of the *Chimica Sperimentale - Istituto Zooprilattico Sperimentale delle Venezie (IZSVe sede di Vicenza*) with regard to the direct analysis in real time coupled to high resolution mass spectrometry (DART-HRMS). Thanks to Dr. Roberto Piro, Alessandra Tata, Andrea Massaro, Marco Bragolusi, Brunella Miano, and Alessandro Negro for the analytical support.

An analytical support was made also by the following labs: *Analisi e controlli funzionali latte qualità - A.R.A.V.* (*Associazione Regionale Allevatori del Veneto, sede di Padova e Vicenza*) and *Chimico-NIRS-XRF - Dept. of Animal Medicine, Production and Health* of Padova University (*Campus di Agripolis*). Thanks to Dr. Sandro Tenti, Massimo Mirisola and Lorenzo Serva for the analytical support and Giorgia Riuzzi and Barbara Contiero for technical assistance in milk sampling and data processing.

The data sampling was performed with the availability of the Federazione Latterie Alto Adige (Sennereiverband) and its farmers and A.R.A.V. and 14 specialized dairy farms located in the Vicenza and Treviso province.

This research was funded by *Fondazione Cariverona* (SAFIL project – call 2016) and by *MIPAAF and SusAn, an ERA-Net co-funded under the European Union's Horizon 2020* research and innovation program (www.era-susan.eu), under *Grant Agreement no. 69623*.

1. INTRODUCTION

1.1 Dairy farming and industry

Milk and dairy products represent an important group of foodstuffs that provides high-value nutrients to a large segment of population. As a consequence of this spread, dairy products are subject to fraud. The main reason for adulteration is the commercial profit by this frauds are commonly based on partial or total substitution of the declared material with cheaper and more easily available components. As a consequence, consumers pay more for a lower quality commodity and may also encounter "ethical" issues, such as commercial milk labeled as organic, or real health problems, such as the presence of allergens or toxic contaminants like melamine as occurred in China in 2008 (Hrbek et al. 2014).

Another field of adulteration is the geographical origin declaration, on which is based the commercial value of many "origin-protected" dairy products like PDO, PGI, or mountain high-quality products (Brescia et al. 2002) (Tenori et al. 2018).

For the latter, the preservation of milk production in mountain areas is a successful strategy to sustain the growth of local communities and supports many other local economic sectors like tourism and craftsmanship: dairy industry goes beyond simple cost-benefit analysis. However, highland production is generally complicated because of lower productivity, higher labor costs, and the limited degree of mechanization. For this reason, mountain products have to be recognized, valued, and paid for their complementary services. However, a single label cannot assure against counterfeits. It can only be trusted if analytical methods for the authentication are developed (Asaduzzaman et al. 2020) (Scampicchio et al. 2016).

1.2 Milk Quality

From a biological point of view, milk is the secretion produced by the mammary gland of female mammals, to feed the newborn. It is a white, opalescent liquid, with a sweetish taste and a characteristic scent (Treccani n.d.). However, from a juridical point of view, only the milking product of a healthy cow can be defined as simply "milk" (art. 15 R.D. 9.5.1929, n. 994 n.d.).

Milk is an important human diet constituent. The main nutrients are lipids, proteins, and lactose but it is also well appreciated for the presence of bioactive compounds such as immunoglobulins and other immune proteins. The beneficial activities of milk, besides it being a source of energy, include anti-cancer, anti-microbial, anti-inflammatory, and immunosuppression properties (Boudonck et al. 2009). Two of the major consumers' concerns are the authentication of geographical origin and the nutritional value of milk linked to its composition (Tenori et al. 2018). Milk's chemical composition depends on several factors such as breed, metabolism, season, health status, nutrition, and milking protocols (Lamanna et al. 2011) (Tian et al. 2016). Each of these factors influences the metabolic pathways, resulting in milk metabolites variability.

Milk nutritional quality depends on the concentration of protein (like casein), fat, lactose, and somatic cell count. The composition also affects cheese-making attitude and can be helpful in herd health monitoring (e.g. the SCC in subclinical mastitis monitoring) (Lindmark-Månsson, Fondén, and Pettersson 2003) (Alhussien and Dang 2018).

As stated by the European Commission (The commission of the European Communities 2001), milk quality determination shall be officially performed through reference methods certified by internationally recognized authorities such as the International Organization for Standardization (ISO n.d.), International Dairy Federation (IDF n.d.), and Association of Official Agricultural Chemists (AOAC n.d.). Reference methods are presented in *Table 1* and can be used to calibrate routine methods.

Table 1 International reference methods for milk quality analysis (ICAR 2017)

Fat	
Gravimetry (Rose-Gottlieb)	ISO 1211 | IDF 1
Gravimetry (modified Mojonnier)	AOAC 989.05
Crude (or total) Protein	
Tritrimetry (Kjeldahl)	ISO 8968 | IDF 20
	AOAC 991:20
	AOAC 991:21
	AOAC 991:22
	AOAC 991:23
Casein	
Tritrimetry (Kjeldahl)	ISO 17997 | IDF 29
	AOAC 998:05
	AOAC 998:06
	AOAC 998:07
Lactose	
HPLC	ISO 22662 | IDF 198
Urea	
Differential pH-method	ISO 14637 | IDF 195
Somatic Cell Count (SCC)	
Direct microscopic somatic cell count	ISO 13366-1 | IDF 148-1

Reference methods are expensive, complex, and specific for each compound, therefore a lot of rapid chemical methods have been developed and authorized in last years: instruments based on multi-analytical approach calibrated on reference methods, but faster and easier to handle (Grelet et al. 2015). One of the most successful is the Foss Milkoscan™ FT 6000 (Foss Electric, Hillerød, Denmark) which is based on Fourier Transform Mid-Infrared Spectroscopy.

Fourier Transform Mid-Infrared Spectroscopy FT-MIR, firstly marked in 1993, is the worldwide recognized method for routine composition and quality milk testing and provides fast, non-destructive quantification of milk chemical components (Grelet et al. 2015). MIR spectroscopy allows the detection of fundamental vibrational transitions in the spectral range from 2,500 to 50,000 nm (rather than NIR's spectral range from 780 to 2,500 nm) which is the region where is possible to detect C-O, C-C, C-O-H, and C-H

stretching (Ferreira et al. 2014). Milkoscan FT6000 allows recording milk fat, protein, casein, lactose, urea composition, and chemical traits, pH, b-hydroxybutyrate, freezing point, chloride, calcium.

1.3 Milk authentication

During the last years, the interest in authentication tools has increased a lot and many technologies and methods have been developed. One of the most validated approaches is the use of mass spectrometry techniques to detect the sample's metabolites which are then analyzed through chemometric studies (Dettmer, Aronov, and Hammock 2007). This approach also allows to distinguish, among known and unknown metabolites, specifical biomarkers that can serve as a signature for milk or create a fingerprint of the samples which, in most cases, is able to authenticate these samples according to how they are produced (Boudonck et al. 2009).

Mass Spectroscopy at High Resolution (HRMS) is even more frequently coupled with ambient ionization methods, such as Direct Analysis in Real Time (DART), for food quality analysis and food authentication. DART-HRMS is a rapid, easy but excellent technique to investigate food composition with a high sample throughput (Hrbek et al. 2014) (Cody, Laramée, and Durst 2005) (Dettmer, Aronov, and Hammock 2007) (Hajslova, Cajka, and Vaclavik 2011) (Dal Martello 2020).

Based on these assumptions, the following work has been carried out to create a fingerprint of milk samples in order to identify specific biomarkers able to discriminate their main roughage dietary source by using DART-HRMS technique and low-level data fusion chemometric technique.

As shown by Scampicchio (Scampicchio et al. 2016), the limited discriminatory capacity of these techniques considered individually is overcome by combining different multi-variables techniques. In this case, the discriminative potential has been used to discern Alps-originated milk samples from the lowland dairy systems ones.

1.3.1 Metabolomics

Metabolomics is the study of cells by measuring profiles of their metabolites (Shulaev 2006). Metabolites are the substrates, intermediates, and products of cellular processes and are usually molecules less than 1500 Da in size (Wishart et al. 2007). The metabolome embodies the complete set of metabolites in a biological cell, tissue, organ, or organism (Jordan et al. 2009).

Metabolomics is a functional genomic system, which includes genomic decoding but also the analysis of genomic expression. Along with genomics, transcriptomics, and proteomics, metabolomics allows the study of cellular physiology and also its correlation with different environmental and nutritional perturbations. Despite the other analysis's building block, metabolomics is the endpoint of this "omics cascade" and therefore the closest to phenotype (Dettmer, Aronov, and Hammock 2007). Metabolomics produces large amounts of data, which should be processed and analyzed by specialized mathematical, statistical, and bioinformatics tools. There are three major approaches used in metabolomics studies: (i) targeted analysis, (ii) metabolite profiling, and (iii) metabolic fingerprinting (Shulaev 2006). Another school of thought describes the first two categories as targeted metabolomics analyses and the metabolic fingerprint as untargeted metabolomic analyses (Cevallos-Cevallos et al. 2009).

- Targeted analysis is a quantitative approach used to measure the concentration of a limited group of known metabolites. Therefore, the structure of the target metabolite should be known a priori and the purified form must be available. This is a standard that is not fulfilled for a lot of metabolites. (Shulaev 2006)
- Metabolite profiling is the research and analysis of a known metabolite set in the studied sample. Mainly used in human health diagnosis as an extension of functional genomic, it allows the study of gene mutation through the changes of the cellular metabolome (Shulaev 2006).
- Metabolic fingerprinting does not measure a specific metabolite or a group of them but describes the unique metabolites pattern that characterizes a specific cellular line, setting out the peculiar fingerprint of the biological material at issue. Some of the found metabolites may remain unknown. (Shulaev 2006) (Dettmer, Aronov, and Hammock 2007).

Metabolomic studies can also be classified as discriminative, informative, or predictive, depending on the analysis and data manipulation (Cevallos-Cevallos et al. 2009).
The discriminative analysis is used to find differences between sample groups by the only use of multivariate data analysis techniques such as principal component analysis, without the necessity of statistical models.
Informative analyses have been aimed to obtain sample intrinsic information. Identification and quantification of known and unknown metabolite, in order to describe a metabolite database.
With predictive studies, statistical models based on metabolite profiles are developed and used to predict a variable that cannot be quantified by other means. Prediction is the analysis method employed in this work. (Cevallos-Cevallos et al. 2009)

1.3.2 Mass Spectrometry

Mass spectrometry (MS) is a technique which aim is the identification of a sample through the ratio mass/charge of each molecule that composes the sample, where *m* is the atomic mass number (sum of neutrons and protons, also called nucleon) and *z* is the valency (number of hydrogen atoms that can bind to a generic element forming its binary compound). (Shulaev 2006) (Audi, Wapstra, and Thibault 2003) (Scheer et al. 1998)
Following the mass spectrometry principle, the molecules of the sample are ionized using an electron beam at known energy. The ions will separate and sort according to their characteristic mass and charge; the unique ionic fingerprint is called “mass spectrum” (Sano et al. 2005) (Reusch 2013) (Spettrometria di massa 2019). In general, the MS technique is used in combination with separating techniques such as chromatography or non-chromatographic approach; in particular, matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF-MS) or ambient mass spectrometry techniques as direct analysis in real time (DART-MS) (Medina et al. 2019). There are different spectroscopic techniques available to achieve the metabolites panel: high-resolution mass spectrometry (HRMS), coupled with DART, is the one chosen for this research (Shulaev 2006) (Dal Martello 2020). Conventional MS calculates the m/z ratio, basing on nominal masses of compounds while High Resolution Mass Spectrometry (HRMS) measures the exact

masses of each compound so that the detection capability of the study is extended (ALEX HYDE n.d.).

Due to its high sensitivity and wide range of covered metabolites, MS has become one of the most chosen analytical platforms for many metabolomics studies to obtain and analyze metabolites (Shulaev 2006) (Dettmer, Aronov, and Hammock 2007).

1.3.3 DART- HRMS

Direct Analysis in Real Time High Resolution Mass Spectrometry is a non-contact, new ambient ion source for mass spectrometry at atmospheric pressure (Medina et al. 2019) (Cody, Laramée, and Durst 2005). DART technique presents several advantages: it allows direct analysis of several kinds of samples, both solid and liquid, in the open air at atmospheric conditions with low molecular mass as well, without the need for sample preparation by chromatographic techniques. The response is instantaneous, providing real-time information. Moreover, the sampling at atmospheric conditions rather than exposed to high electrical potentials preserves the integrity of the sample. (Cody, Laramée, and Durst 2005) (Gross 2014) (Weston 2010) (Jorabchi, Hanold, and Syage 2013) (Chernetsova, Morlock, and Revelsky 2011).

DART analytical technique has been extensively used in the field of food authentication, during the last years. It was largely employed in the authentication of alcoholic beverages like wine or beer, but also the assessment of vegetables' farming practice. Nevertheless, DART has proved to be particularly capable in animal fat detection and rapid profiling of triacylglycerols (TAGs) (Vaclavik et al. 2011) (Medina et al. 2019) (Cubero-Leon, Peñalver, and Maquet 2014).

In particular, DART-HRMS has already been used by Hrbek for dairy authentication, with highly effective in detecting the adulteration of dairy products with vegetal oils (Hrbek et al. 2014) and by Tata to predict the health status of dairy cows (Tata et al. 2021).

DART ionization consists of a tube within which a heated gas stream allows the formation of a distal plasma discharge of ions, electrons, and metastable species (excited-state atoms and/or molecules of gas). These charged particles are immitted in the open air, where the

sample is instantly ionized and submitted to thermal desorption before entering the mass spectrometer (Cody, Laramée, and Durst 2005) (Song et al. 2009) (Gross 2014).

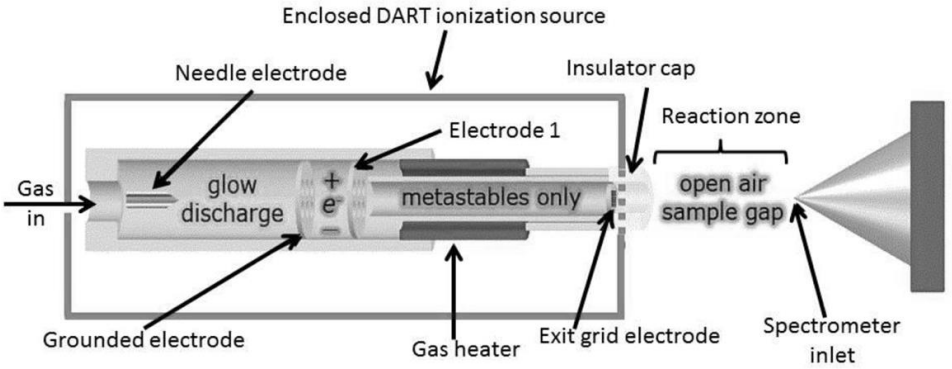


Figure 1. Detailed cutaway view of DART ionization source (JEOL n.d.)

DART-HRMS zones:

- DART ionization source zone is a pipe divided into three chambers. The flow of helium, or sometimes nitrogen, is introduced into the first chamber. A corona discharge, which is an electrical discharge resulting from the ionization of air, is generated by an electric potential of several kV realized between a needle electrode and a perforated disk electrode at a ground electric potential at the end of the chamber (Chang, Lawless, and Yamamoto 1991) (Cody, Laramée, and Durst 2005) (Gross 2014).

  The contiguous second chamber ends with another perforated disk electrode with an electric potential which serves as an eliminator of the cationic species of the plasma. The remaining plasma enters into the third chamber, which can be optionally heated, and the gas flows throughout an exit grid electrode which mediates the purification from anions and electrons. The grid electrode serves as an ion repellent, which prevents ion-ion recombination, a process that might result in a signal loss

- Reaction zone is the 5-25mm long open-air zone between the DART ionization source and the HRMS, where sample ionization and thermal desorption occurs (Hajslova, Cajka, and Vaclavik 2011). The sample can be solid, liquid, or gaseous (Cody, Laramée, and Durst 2005). The gas metastable species react with ambient

atmosphere components to form reactive species, ionizing the neutral analytes of samples (Hajslova, Cajka, and Vaclavik 2011).

Two different types of ionization processes can be performed, to comply with different analyte properties so that different spectra are achievable from the same sample (Jorabchi, Hanold, and Syage 2013).

- Positive-ion formation: the helium metastable species react with atmospheric water, returning protonated clusters. Protons transfer from these clusters to the analyte molecule yielding in the formation of $[M+H]^+$ ion with the release of electrons.

- Negative-ion formation: electrons formed during the DART ionization source formation allow the formation of negatively-charged oxygen clusters. In this way, molecules of analytes are deprotonated $[M-H]^-$ (Hajslova, Cajka, and Vaclavik 2011).

The spectra obtained from the DART-HRMS modalities can be merged with a data fusion method, into a single dataset representative of the sample. Data fusion method consists in combining data from different analytical, multimodal sources to provide a more accurate characterization of a sample, reducing interferences or error rates. Three different types of data fusion can be performed: low-, mid-, high- level data fusion.

- Low-level data fusion: raw data from all sources are concatenated into a common data matrix after suitable preprocessing and weighting. It is the one chosen for this work.
- Mid-level data fusion, also known as *feature level data fusion*, is based on the extraction of relevant characteristics from each dataset separately. These relevant scores are then merged into a single combination dataset to be analyzed according to a multivariate approach.
- High-level data fusion, also called “decision level”, consists of independent models calculated from each dataset and they are, only subsequently, merged together. (Schwolow et al. 2019)

2. AIM OF THE RESEARCH

The research aimed to investigate the effect of the main dietary roughage source on milk composition and chemical traits.

Moreover, the research aimed to detect the differences between the main quality traits of the South Tyrolean alpine milk, compared to that produced in the lowland intensive dairy systems throughout a one-year experimental period.

The main goal of the research was to test the capability of the direct analysis in real time coupled to high resolution mass spectrometry (DART-HRMS) to discriminate among the milk samples of the experimental dietary thesis. In addition to the detection of the DART biomarkers, we also tried their identification and to assess if their presence in the milk has an appreciable or depreciative role.

3. MATHERIALS AND METHODS

3.1. Experimental design

The study involved 14 dairy farms in the middle of the lowland area of Po Valley (North East of Italy, 45°19′49″N 9°47′56″E) and six farms in the Alpine area of South-Tyrol (46°30'0" N 11°19'59" E). The farms were selected to represent average herd size, breeds, and milk production characterizing both the intensive (lowland) and extensive (Alpine) dairy systems of the Italian dairy production chain during every season. In the intensive system, the main breeds represented were Italian Friesian and Italian Brown, with at least 95% of the lactating cows in each herd belonged to one or the other (*Table 2*). In the Alpine system, the main breeds were Brown Swiss, Alpine Grey, Simmental, Pinzgauer. Great variability was appreciable between mountain realities: stables' and pastures' altitudes were spread within 1000 m asl and 2050 m asl, therefore pastures' composition was very heterogeneous with tens of different plant species. During winter, alpine cows are fed with locally-produced dry forages or silages, with great variability of raw materials among the farms.

All lowland farms were associated with the Regional Breeders' Association (Veneto Region), ensuring herd performances were recorded monthly over the experimental period (*Table 2*). In the case of the Tyrolean farms, they were associated with Sennereiverband Südtirol (South Tyrol Dairy Federation). In the intensive system (lowland), herd dry matter intake (DMI) was recorded at each sampling visit (5 recordings *per* farm across one-year experimental period) by calculating the difference between the total amount of TMR distributed to the lactating cows and refusals after 24 h or before the subsequent distribution. In the extensive system (Alpine Tyrol), herd dry matter intake (DMI) referred only to the summer season (2^{nd} and 3^{rd} milk sampling in July and September, respectively). It was estimated according to daily theoretical consumption of 10 kg of DM of alpine grazing pasture for each lactating dairy cow, which receive also a supplement of dried and fresh forages and concentrates as described in *Table 3a*.

Table 2. Herd breed incidence (%) and descriptive statistics (average + standard deviation) according to the 5 dietary feeding groups based on the main roughage source.

	Lowland				Alpine
	HMS	**MMS**	**FCG**	**HAY**	**APS**
Breed percentage (%)					
Italian H. Friesian	50	100	79	52	1
Italian Brown	50	-	21	48	4
Tyrolean Grey	-	-	-	-	21
Simmental	-	-	-	-	70
Pinzgauer	-	-	-	-	4
Herd descriptive statistics					
Lactating cows (n)	109 (±32)	122 (±27)	62 (±12)	71 (±16)	21 (±15)
[1] **DIM** (n)	198 (±26)	177 (±17)	178 (±21)	165 (±22)	n.e.
Calving interval (d)	427 (±28)	410 (±24)	412 (±20)	401 (±17)	385 (±23)
[2] **DMI** (kg/d/cow)	23.2 (±1.0)	24.0 (±1.1)	22.6 (±1.5)	22.1 (±1.4)	n.e.
Milk (kg/d/cow)	30.5 (±4.2)	32.7 (±3.4)	29.5 (±3.4)	28.8 (±2.4)	20.8 (±3.4)
[3] **FPCM** (kg/d/cow)	30.2 (±4.0)	32.1 (±2.9)	28.5 (±3.1)	28.0 (±2.1)	20.8 (±3.1)

Main roughage source: HMS, high maize silage; MMS, mixed maize/crop silages; FCG, fresh grass/crop silages; HAY, lowland permanent meadow hay; APS, alpine pasture/alpine hay. [1]Days in milk. [2]Dry matter intake (kg of DM per cow and day). [3] Fat Protein Corrected Milk; 4.0% fat, 3.3% true protein; true protein estimated as 93% of the crude protein. n.e., not estimated

The experimental protocol was designed to allocate each farm to one of five feeding groups that represent the main dietary roughage source. The five dietary feeding groups were formulated according to the percentage (%) of the roughage source: i. high maize silage (HMS, maize silage > 28; 4 farms and 20 milk samples); ii. mixed maize/crop silages (MMS, maize silage < 20 and crop silages > 20; 3 farms and 18 milk samples); iii. fresh grass/crop silages (FCG, crop silage < 20 and fresh-cut grass > 10; 4 farms and

19 milk samples); iv. hay (HAY, lowland permanent meadow hay > 40; 3 farms and 13 milk samples); v. alpine pasture system (APS, alpine pasture and dried forages > 65, 6 farms and 18 milk samples). On all lowland farms, cows were fed total mixed rations (TMR) meanwhile, in the Alpine area, they graze during the summer on natural pastures and receive a daily amount of 4 kg (on dry matter basis) of concentrates per lactating cow, which defines all the extensive farms' samples as high-input rearing (Kühl, Flach, and Gauly 2020). During winter alpine cows are fed with local-produced hay or silage. Lowland-TMR and Alpine-supplements were formulated to cover the herd's nutritional requirements (available energy and protein) based on the NRC standard (Nutr. Requir. Dairy Cattle 2001). All forages were produced on the farms although some concentrate feeds were purchased. Average rations for the four experimental groups (% on DM) and their diet proximate compositions (% on DM) are reported in *Table 3a* and *Table 3b* respectively.

Table 3a Ingredients (%) on dry matter (DM) basis of the 5 dietary feeding groups based on the main roughage source.

	Lowland				Alpine
	HMS	**MMS**	**FCG**	**HAY**	**APS [1]**
Ingredients (% on DM)					
Maize Silage	33	16	0	0	0
Hay [2]	11	10	23	44	12
Crop silages [3]	8	24	18	8	3
Fresh-cut green grass	0	0	12	3	5
Alpine pasture	0	0	0	0	60
Amylaceous concentrates [4]	26	27	32	31	12
Protein concentrates [5]	17	19	11	10	8
Residual[6]	5	4	4	4	5

Main roughage source: HMS, high maize silage; MMS, mixed maize/crop silages; FCG, fresh grass/crop silages; HAY, lowland permanent meadow hay; APS, alpine pasture/alpine hay.

[1]Data of APS group referred to the grazing period (12 samplings out 18) and considered a theoretical DMI of 20 kg of DM as following: 12 kg of pasture, 4 kg of amylaceous and protein concentrates, 3 kg of a mix of hay/fresh cut grass/ensiled grass and 1 kg of residual. [2]Permanent meadow and alfalfa. [3]Sorghum, wheat, alfalfa, Italian ryegrass. [4]Manly maize and barley grain derivate (meal, extruded, rolled, flaked). [5]Mainly soybean and sunflower products. [6]Straw, bran, beet pulps, min-vitamin premix.

Table 3b Proximate composition (average + standard deviation) on dry matter (DM) basis of the 5 dietary feeding groups based on the main roughage source.

	Lowland				Alpine
	HMS	**MMS**	**FCG**	**HAY**	**APS [1]**
Proximate composition (% on DM)					
DM (%)	55.0 (±5.0)	55.8 (±4.8)	59.8 (±6.8)	68.0 (±5.2)	51.0 (±7.2)
Crude protein	14.0 (±0.5)	13.9 (±0.6)	13.9 (±0.9)	13.8 (±1.0)	14.8 (±1.5)
Crude fat	2.7 (±0.4)	2.9 (±0.5)	2.7 (±0.5)	2.5 (±0.5)	2.1 (±0.4)
Crude ash	7.9 (±0.7)	7.6 (±0.4)	8.0 (±0.4)	7.9 (±0.6)	8.8 (±0.6)
aNDF	37.0 (±1.9)	37.4 (±2.4)	37.8 (±3.4)	40.1 (±3.8)	43.1 (±4.4)
ADF	21.9 (±1.4)	22.3 (±1.5)	21.8 (±2.3)	23.1 (±1.9)	25.9 (±2.2)
Non-fiber carbohydrates[7]	38.4 (±1.8)	38.2 (±1.9)	37.6 (±2.6)	35.7 (±3.3)	31.2 (±3.5)
Starch	22.4 (±1.8)	21.8 (±2.7)	21.1 (±3.7)	19.9 (±1.5)	17.1 (±1.9)

Main roughage source: HMS, high maize silage; MMS, mixed maize/crop silages; FCG, fresh grass/crop silages; HAY, lowland permanent meadow hay; APS, alpine pasture/alpine hay.

[1]Data of APS group referred to the grazing period (12 samplings out 18) and considered a theoretical DMI of 20 kg of DM as following: 12 kg of pasture, 4 kg of amylaceous and protein concentrates, 3 kg of a mix of hay/fresh cut grass/ensiled grass and 1 kg of residual. [7]NFC was calculated as 100 minus (CP + CF +CA +aNDF).

3.2 Milk and ration sampling and analysis

During 2019, five and three raw bulk milk samples were collected on each lowland and Tyrolean farm over 5 (March, May, July, September, and December) and 3 (June, July, and September), respectively. The first mountain sampling has to be considered as late spring sampling because grazing pasture by lactating dairy cows had not yet begun. Thus, a total of 88 (70 lowland and 18 Alpine) raw bulk milk samples were analyzed for proximate composition, chemical traits, and DART metabolites. Two lowland farms (a FCG and a HAY farm) altered diet ingredients depending on the seasonal supply of feeds and changed TMR formulation over the experimental period, essentially changing groups. The original FCG farm changed once into MMS ration. The HAY farm changed twice into MMS ration. However, according to Rego et al. (Rego et al. 2016), we ensured at least three weeks between the TMR change and milk sampling. At each sampling, the current lowland TMR were collected and formulations recorded. In the Tyrolean farms, grazing pasture, mix of fresh and dried forage, and concentrate supplements were sampled as the main ingredients of the ration of the lactating dairy cows. The milk and ration (TMR and ingredients) samples were refrigerated and carried to the laboratory immediately after the sampling and milk sub-samples for wet chemistry, near infrared (NIR) spectroscopy, and DART analysis, kept at 2± °C in dark conditions.

TMR and raw ingredients samples were analyzed for dry matter (DM), crude protein (CP), crude fat (CF), crude ash (CA), neutral detergent fiber (aNDF), acid detergent fiber (ADF), and starch by means of a FOSS 5000 scanning monochromator bench-top near-infrared (NIR) instruments (Foss NIRSyste m, Hillrød, DK), using the calibration curve as described by (Andrighetto et al. 2018). Non-fiber carbohydrates (NFC) was calculated as complement to 100.

The milk proximate composition (crude protein, casein, lipids, lactose, ash) and chemical traits (urea, pH, β-hydroxybutyrate) were recorded by a Fourier transform mid-infrared (FT-MIR) spectroscopy technique using a MilkoScan FT6000 (Foss Electric A/S, Hillerød, Denmark). Additionally, the somatic cell count (SCC) was performed by a Fossomatic 5000 (Foss Electric A/S, Hillerød, Denmark).

3.3 DART-HRMS

3.3.1 Sample extraction

Two different extraction procedures were applied to the milk samples. In the first one, 50µL of milk were suspended in 1mL of water and methanol (H_2O:MeOH; 20:80 v/v) solution (MilliQ water and Methanol HPLC-grade with 99.9% purity, from VWR International, Radnor, USA), vortexed for 30 s, sonicated for 15 minutes and centrifuged for 5 min at 12000 g to extract the polar metabolites (Dettmer, Aronov, and Hammock 2007). In the second protocol, 50µL of milk were diluted in 10 mL of pure ethyl acetate (EtAc) (99.9% purity, Carlo Erba Reagents, Cornaredo, Italy), vortexed for 30 s, then sonicated for 15 minutes to extract the more lipophilic, non-polar metabolites (Dettmer, Aronov, and Hammock 2007). A volume (1mL) of the extract was pipetted into a small tube and centrifuged for 5 min at 12000 g.

Subsequently, the two methanol diluted samples would be analyzed one in negative-ion mode and the other in positive-ion mode, the same goes for the ethanol diluted samples, to obtain four analytical variables. This metabolites fractionation allows differentiated analysis and the expansion of the achievable dataset. (Riuzzi et al. 2021).

3.3.2 DART-HRMS Analysis

The instrumental analysis was carried out using a DART SVP 100 ion source (IonSense, Saugus, USA) coupled with an Exactive Orbitrap (Thermo Fisher Scientific, Waltham, USA). The DART source was coupled with a Dip-it$^{(R)}$ sampler (IonSense, Saugus, MA, USA). To facilitate the ions to pass from the DART source to the mass spectrometer, a vapor interface was installed. The distance between the DART gun and the ceramic transfer tube of the vapour interface was 12 mm. The parameters of the DART and the Orbitrap analyzer were set as described by Riuzzi et al. (Riuzzi et al. 2021). The resolution was set to 70.000 FWHM and the mass range was 75–1125 Da in both positive and negative ion modes.

All DART-MS analyses were run with an automated gain control target setting of 3×10^6. Melting point tubes were inserted into the autosampler holder, and then 5 □L of each extract were spotted individually onto them. Subsequently, the spotted melting point tubes were automatically moved at a constant speed of 0.3 mm/s through the DART gun exit and ceramic tube of the Vapur interface. The time of desorption from the surface of each tip was about 20 s.

The samples were analyzed in triplicate, and XCalibur QualBrowser software (Thermo Fisher Scientific, Waltham, USA) was used to visualize the entire spectra in a .raw format. These were converted to mzML files using Proteowizard (Holman, Tabb, and Mallick 2014) and then opened with mMass software (http://www.mmass.org/) to interpret the mass spectrometry data. The *m/z* values were tentatively assigned by consulting the online METLIN (https://metlin.scripps.edu) and HUMAN METABOLOME DATABASE (www.hmdb.ca) libraries. Prior to statistical analysis, the spectra of the four datasets (two extraction solvents and two ion modes) were converted into .csv files with Rstudio 3.6.1 software (RStudio Team, 2016; RStudio Integrated Development for R; RStudio, Inc., Boston, USA).

3.4 Data processing and statistical analysis

3.4.1 Statistical analysis of milk proximate composition and chemical traits

Milk proximate composition and chemical traits were analyzed using a linear mixed model that included the fixed effects of dietary group and the random effect of the farm (SAS PROC MIXED). Pairwise comparisons among levels of all the factors were performed using Bonferroni correction. The hypotheses of the linear model on the residuals were graphically assessed. This first statistical model was performed using SAS 9.4 software (SAS Institute Inc., Cary, NC, USA).

3.4.2 DART statistical modelling according five dietary groups

The triplicate spectral data were averaged and statistically analyzed using Rstudio 3.6.1 software and the MetaboAnalyst 5.0 web portal (www.metaboanalyst.ca) for comprehensive and integrative metabolomic data analysis (Pang et al., 2020).

The isotopes were removed from the signals recorded in the four datasets, and the *m/z* values aligned with a tolerance of 0.008 Da. All ion signals with more than 75% of missing values (no detected ion intensity) were removed. For ions with less than 75% of missing values, those missing values were replaced with half of the value of the lowest recorded *m/z* intensity. The signals of each spectrum were also normalized by sum, whereas each feature was normalized by Pareto scaling. As reported in *Figure 2*, the four dataset blocks were merged (concatenated) by performing a low-level data fusion approach (Borràs et al. 2015). The merged dataset was split into train (70% of the data, n = 63) and validation set (30% of the data, n = 25). The merged train set was submitted to a partial least squared discriminant analysis (PLS-DA) with the aim of distinguishing between the five dietary groups.

Subsequently, only ions with coefficients >30 were retained. The 25 selected ions were submitted to hierarchical cluster analysis (HCA) with Pearson distance and Ward linkage to show the correlation between groups and the selected ions.

The twenty-five *m/z* values extrapolated by PLS-DA were used to construct a linear discriminant analysis (LDA) model on the training set using Rstudio 3.6.1. Its capability to correctly classify the samples according to the dietary groups was verified on the training set by 10 fold cross-validation. Furthermore, the LDA model was performed against the independent validation set withheld previously. The predictions of this blind verification were arranged in a confusion matrix and a set of statistical measurements (accuracy, sensitivity, specificity, precision, and Matthews correlation coefficient) were calculated to assess the predictive discriminating capacity of the supervised classifier LDA model based on the 25 most informative ions sorted by the PLS-DA (Bisutti et al. 2019).

3.4.3 DART statistical modelling according three dietary groups

To increase the discriminative capacity to separate milk samples, the multivariate statistical approach was carried out again according to only three experimental dietary groups. The three dietary theses were: i. lowland silages (LLS for brevity) gathered the samples of HMS and MMS (thus the main roughage source was maize silage and other cereal silage); lowland dried/ensiled forages (LLF for brevity) gathered the samples of FCG and HAY (thus the main roughage source was grass and permanent meadow hay and forage crop silage but without maize silage); iii. APS group remained unchanged. After that, the DART modelling was repeated as described in the previous paragraph 3.4.2 and as illustrated in *Figure 2*.

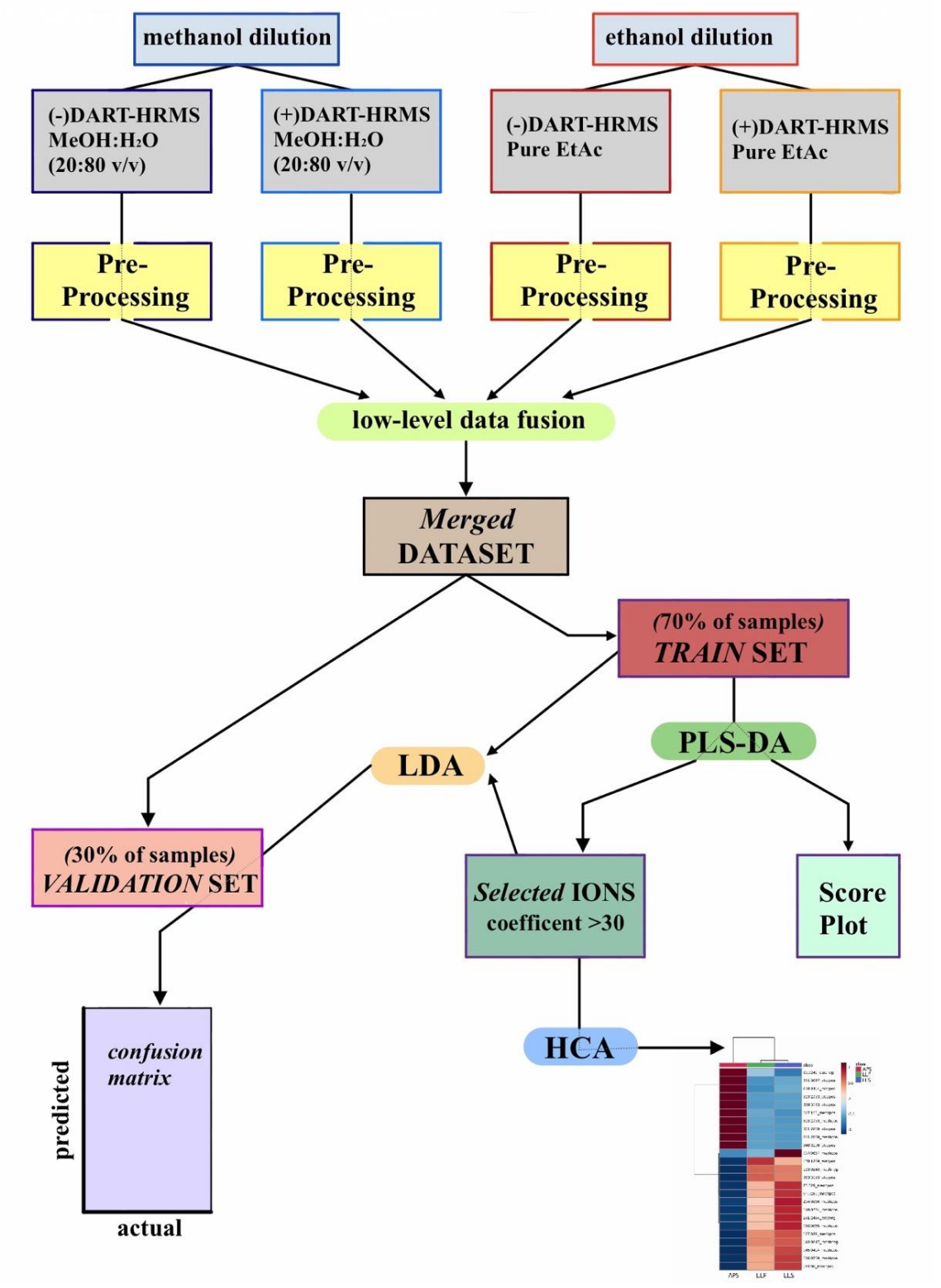


Figure 2. Flow chart of (+/−) direct analysis in real time high resolution mass spectrometry (DART-HRMS) signatures to discriminate milk samples according to dietary forage groups. After DART-HRMS data pre-processing (yellow boxes, pre-

processing refers to the removing of isotopes and ions with > 75% of missing values as well as normalization of signals by sum and Pareto scaling), the four pre-processed datasets were submitted to a low-level data fusion (light green box). The low-level fused dataset was randomly separated into train (n = 63, bordeaux box) and validation (n = 25, red box) set.

A partial least squared discriminant analysis (PLS-DA) was carried out (green box) on the merged train dataset. Subsequently, only ions with coefficient > 30 were retained and used to perform a hierarchical cluster analysis (HCA, blue box) to visualize the correlation between groups and selected variables.

The 25 selected ions were used to build a LDA (orange box) that was cross-validated on the fused train set and then validated on the fused validation set (blind validation). A confusion matrix (lilac box) was built with the predictions of the cross-validation to facilitate the calculation of the accuracy, sensitivity, specificity and Matthews correlation coefficient (MCC).

4. RESULTS AND DISCUSSION

A total of 88 raw bulk milk were sampled in 18 farms during 2019. They were divided into five groups according to the main feeding system and the geographical rearing area. 70 came from raw bulk milk sampled in intensive, lowland areas of which: 20 from a high maize silage-based diet (HMS), 18 from a mixed maize/crop silages diet (MMS), 19 from fresh-cut grass diet (FCG), 13 from a permanent meadow hay diet, with the exclusion of silages (HAY). The last 18 came from cows reared extensively on the Alps with a diet mainly composed of mountain dry matter (hay or silages) and pasture (APS).

4.1 Milk proximate composition and chemical traits

Milk composition is strictly connected with the feeding system and it is well known the influences that a lot of diet-parameters have on milk traits (Lindmark-Månsson, Fondén, and Pettersson 2003).

Over the last few decades, maize silage has become, with its high percentage of starch (27-35%), the major forage component in many TMR formulations. Maize silage towards grass silage allows an increase in DMI (Dry Matter Intake) up to 2kg d^{-1}, in milk yield and in protein content (Khan et al. 2015).

As can be found in several studies, a hay/grass-based feeding system (HAY+APS) provides a higher NDF (neutral detergent fiber) intake. Higher NDF percentage means higher chewing activity and rumination time, rising of rumen pH, and rumen health condition. The consequences of high NDF on milk should be an increase in fat content but a lower protein content and a linear decrease in milk production (Beauchemin 1991). The rise of fat concentration can be explained by noting that a higher NDF in the feed decreases both propionate and valerate milk levels, with a consequent increase of acetate, butyrate, isobutyrate, and isovalerate concentrations. These variations lead to an increase of acetate/propionate ratio (C2:C3) > 2.2 which is known to be an important parameter for milk fat concentration as well as a sign of decreased ruminal acidosis risk in cows (Sejrsen, Hvelplund, and Nielsen 2009).

The decrease in production can be also simply explained: a higher NDF content implies a lower intake potential and a consequent reduction of available energy, which results in lower milk production.

How can be seen in *Table 3b*, HMS and MMS feeding systems provide a greater starch intake, while HAY and ASP diets are higher in dry matter and NDF. The higher the NDF is, the lower the yielded milk is: APS milk production (kg/d/cow) is averagely 10 kg fewer than in lowland breeding.

On the other side, the high quantity of concentrates, typical of intensive high-inputs rearing, leads to an increase of DMI with a consequent increase of available energy for milk secretion and milk proteins synthesis (Nielsen et al. 2006) (Asaduzzaman et al. 2020).

Milk samples were analyzed for proximate composition and the results are reported in *Table 4.* SCC was normalized with a log-transformation which allowed us to obtain SC score [d].

Table 4. Effect of dietary forage on milk proximate composition and chemical traits

	Lowland				Alpine	SEM[1]	p-value
	HMS (n = 20)	**MMS** (n = 18)	**FCG** (n = 19)	**HAY** (n = 13)	**APS** (n=18)		
Crude protein (g/100 g)	3.52[a]	3.48[ab]	3.38[b]	3.45[ab]	3.53[a]	0.04	0.009
Casein (g/100 g)	2.70[a]	2.66[ab]	2.59[b]	2.61[b]	2.72[a]	0.03	0.014
Crude fat (g/100 g)	4.15[ab]	3.92[b]	3.90[b]	3.92[b]	4.25[a]	0.09	0.043
Lactose (g/100 g)	4.80	4.82	4.76	4.78	4.82	0.03	0.084
SCC score (units)[2]	3.92	3.77	4.08	4.01	4.07	0.17	0.152
Urea (mg/dL)	24.3	24.9	23.9	24.7	25.4	1.4	0.420
BHB[3]	0.056	0.056	0.062	0.055	0.068	0.010	0.300
Milk native pH	6.65	6.68	6.65	6.66	6.65	0.01	0.197

Main roughage source: HMS, high maize silage; MMS, mixed maize/crop silages; FCG, fresh grass/crop silage; HAY, lowland permanent meadow hay; APS, alpine pasture/ alpine hay. Within the dietary groups are reported the number of sampling in brackets.
[1]SEM, standard error of the mean; [2]SCC score, $\log_2$ (SCC/100,000) + 3; [3]BHB, β-Hydroxybutyrate (mmol/L).
[a-c]Least squares mean in a row without a common superscript differ ($p < 0.05$)

Considering these milk productions and NDF content we expected, as explained above, low-quality milk composition for HMS and MMS the most productive breeding methods, and high fat – poor protein milk content for APS, which is the least productive one.
Moreover, in APS milk we expected high content of urea and somatic cells, due to a poorly balanced ration formulation and poor hygiene conditions, especially by using a hand-portable milking machine compare to the mechanized milking parlor.

Contrary to expectations, the intensive rearing and the extensive farm had very similar chemical compositions: HMS and MMS milk were significant (p-value < 0.05) high in crude protein, casein, and crude fat, even higher than FCG and HAY. Lactose was almost the same for every thesis even if, according to Asaduzzaman et al. (Asaduzzaman et al. 2020), it was expected to be higher in high-yielding realities.

APS had significant (p-value < 0.05) high crude fat levels, as expected, but also high crude protein and casein levels, overcoming FCG and HAY which had both a higher DMI. Moreover milk urea content and SCC were low, similarly to highly specialized rearing.

Even if not expected, these results were comparable with other studies' results thus, the quality results of this work should not be considered an exception (Sturaro et al. 2013) (Scampicchio et al. 2016).

The high-quality content of HMS rearing might be explained by the significant presence of Brown Swiss cows (almost one-half of the herds, how shown in *Table 2*) which are well known for a great milk content of crude protein, casein, and crude fat (Zanon et al. 2020). Brown Swiss were probably included in the HMS herds to improve milk cheese-making properties. Moreover, it can be assumed that the dairy systems that invest more in increasing milk yield, pay also more attention to rations formulation and feeding practices to have a more valuable product (Sturaro et al. 2013).

The low urea and SCs content in APS milk, suggests that even in the traditional and less mechanized systems, the farmers are able to adopt the necessary good management practices to maintain the first quality classification (Sturaro et al. 2013). Also on mountain small farms, hygiene and udder health status are factors to care about. Moreover, considering the diet proximate composition shown in *Table 3b*, nutrients and energy balance seem to be similar to the proximate composition of intensive rearing. With at least 3 kg of concentrates daily administrated, APS can be classified as high-input management which explains, how previously outlined, the high milk protein content in *Table 4.*

As could be expected, no discriminative marker between the feeding theses was identifiable by the only milk chemical analysis. That is why metabolomic analysis associated with chemometric study should be performed on milk samples. (Scampicchio et al. 2016)

4.2 DART - HRMS modelling and milk authentication

4.2.1 Authentication based on five milk-feeding groups

For subsequent analysis, milk samples were diluted in both water:methanol and pure ethyl acetate (80:20 v/v) and then vortexed as preparation for DART analysis. DART-HRMS spectra were acquired in both positive and negative ion-mode. A mass spectrum is composed of m/z values in the X-axis and its relative intensity (or abundance %) in Y-axis (McLafferty and Turecek 1993). Before the statistical analysis, a pre-processing approach was carried out toward the (+/-) DART-HRMS dataset by deisotoping, alignment, missing values removing, normalization by sum, and Pareto scaling. The preprocessed data were then concatenated with a low-level data fusion, which integrates the multiple data sources into a single, more useful information set. A 70% of the samples have constituted the train set which has undergone two different discriminant analysis models.

A first statistical and graphical separation was attempted by performing a partial least square-discriminant analysis (PLS-DA) between the 5 dietary groups.

PLS-DA is a supervised, classification method, which means that it aims to assess a particular sample – or group of them- and its confidence interval, to the appropriate belonging group. Being a supervised model, the classification groups are settled by the operator. (Schwolow et al. 2019) (Medina et al. 2019).

The results are graphically shown in the score-plot of *Figure 3*. The axes of the graphical space are built with the first two principal *components* (that are the theoretical statistical representation of the raw variability of DART signals) resulted from the PLS-DA, where "principal" stands for the components which the higher percentage of the variance explained by the model.

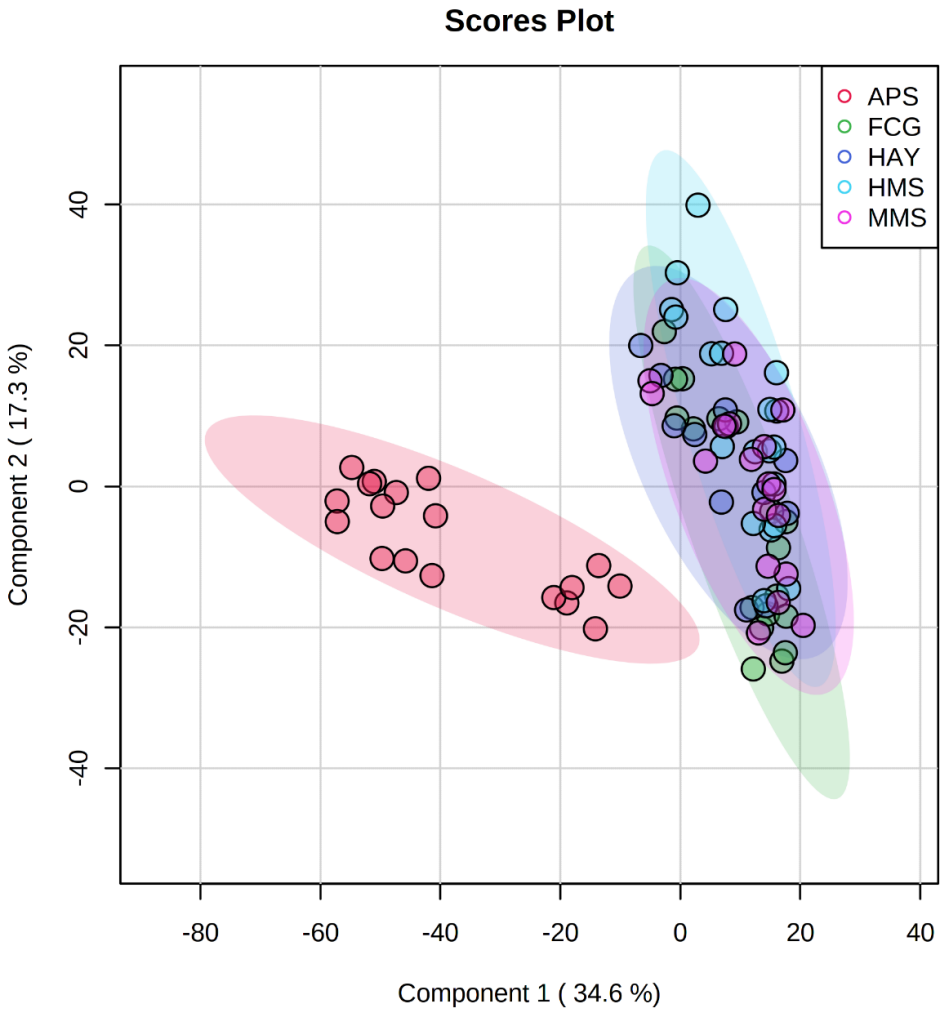


Figure 3. PLS-DA score plot of feeding groups based on milk (+/-) DART-HRMS signatures. Ninety-five percent ellipses confidence interval s (0.95-CI) are drawn around each centroid of grouping.

The roughage dietary groups are: alpine pasture/alpine hay (APS) in red, fresh grass/crop silages (FCG) in green, lowland permanent meadow hay (HAY) in blue, high maize silage (HMS) in light blue, and mixed maize/crop silages (MMS) in pink. FCG, HAY, HMS and MMS are experimental groups from lowland farms meanwhile APS is an experimental group from South Tyrolean farms.

On the resulting PLS-DA scores plotting, it is possible to distinguish APS (reported in red) from all the other groups which, however, are not discriminable ones from each other because they overlapped. It can be assessed that the 5-group PLS-DA model was able to discriminate between alpine and lowland milk but it was unable to separate milk samples

within the lowland dairy systems. According to this reduced discriminant capacity, the results related to the heatmap and the LDA model are not further discussed.

4.2.2 Authentication based on three milk-feeding groups

After joining the milk samples of the lowland theses (as described in paragraph 3.4.3), a similar statistical approach was carried out again with the aim to improve the discriminative capacity of the model. Therefore, for this second approach, the three dietary theses were classified as lowland silages (LLS for brevity, a thesis that gathered the samples of HMS and MMS), lowland dried/ensiled forages (LLF for brevity, a thesis that gathered the samples of FCG and HAY) and the original APS group. It has been noticed that LLS is an intensive feeding system based on maize silage as the main fibrous source (maize silage is a forage but with a high nutritive value) plus a mix of other cereal silage and a very low inclusion of hays. The LLF is also an intensive feeding system based on grass and permanent meadow hay as a fibrous dietary component and a limited amount of forage crop (alfalfa, Italian ryegrass) ensiled to enhance the nutritive value of the ration (the total mixed ration, TMR). The APS is an extensive feeding system based on seasonality: permanent, polyptych pastures during summer and hay or silages locally produced, during winter. Fresh, dried, or fermented, the diet is in any case representative of the local vegetation, which is very variable depending on area and altitude. Despite the great variability of pasture types, they are always rich in botanical essences usually not available in the lowland. The most abundant forage species are *Agrostis tenuis, Anthoxanthum alpinum, Festuca rubra, Nardus stricta, Phleum alpinum, Poa alpina, Trifolium repens,* and *Achillea millefolium.* (Ziliotto, Scotton, and Da Ronch 2004) (Orlandi, Clementel, Bovolenta 2005).

The score plot of the PLS-DA based on 3 experimental dietary groups is reported in *Figure 4.*

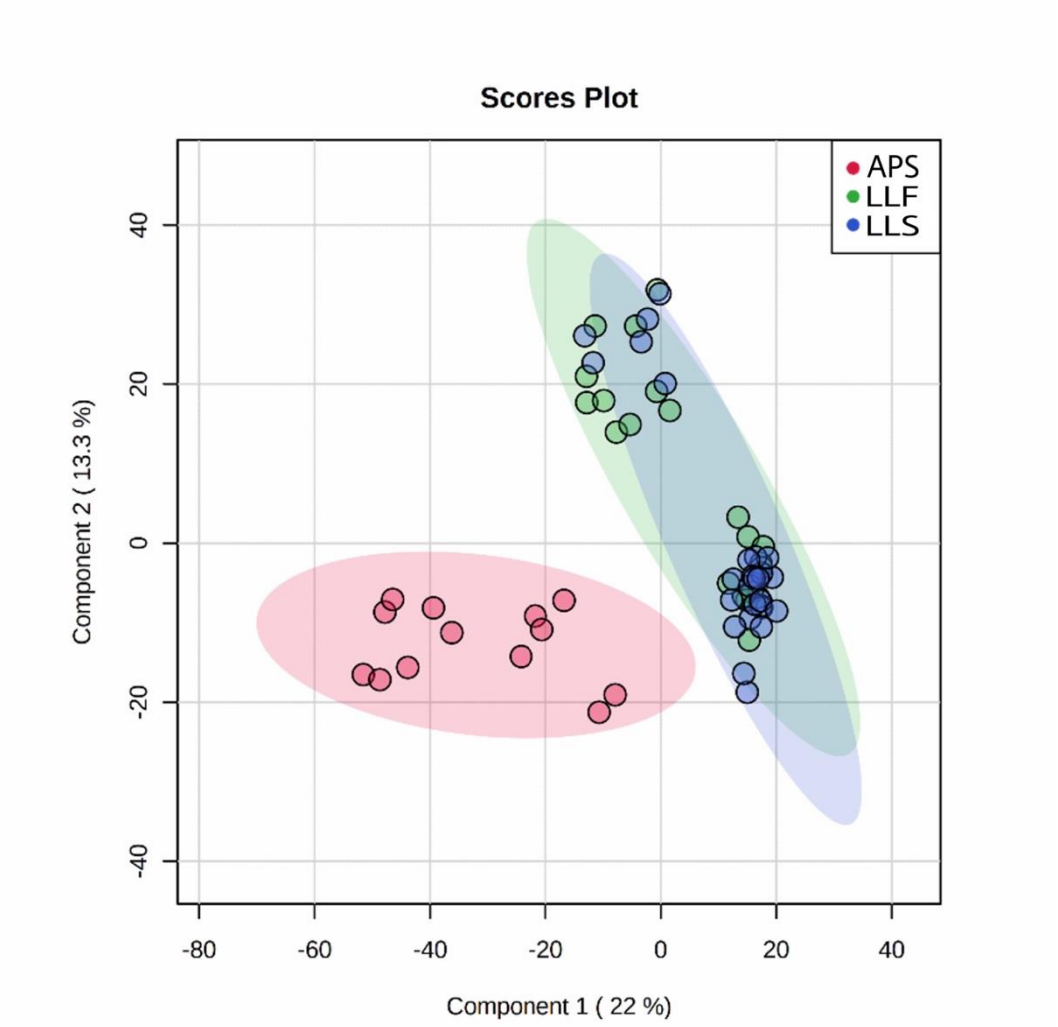


Figure 4. PLS-DA score plot of feeding groups based on milk (+/-) DART-HRMS signatures. Ninety-five percent ellipses confidence interval s (0.95-CI) are drawn around each centroid of grouping.

The roughage dietary groups are: APS, alpine pasture system (South Tyrol); LLS, lowland silages, LLF lowland forages (lowland of Veneto region). See the paragraph 3.4.3 and page 22 for more details about the 3 dietary theses.

With this new model, the APS group remains clearly discriminable from the other groups and a slight improvement is achieved in discriminating lowland forage group (reported in green), which is maize silage-free, from lowland silage one (reported in blue) that is mainly based on whole plant and grain maize silages as well as others maize-similar (i.e., sorghum and wheat silages). It can also be noted how lowland's score plots, despite their overlapping, are mainly distributed into two groups, containing both silage (LLS) and

forage (LLF) data. This graphical result suggests that a sort of discrimination not only based on roughage dietary source has been achieved. According to the bibliography, the discrimination hypothesis might be based on intra-group breed-effect or season-effect (Scampicchio et al. 2016) (Zanon et al. 2020).

HCA and heatmap

Of the ions (m/z values) arising from the PLS-DA, those having a coefficient higher than 30 were selected and submitted to hierarchical cluster analysis (HCA) with the application of Pearson correlation distance to reveal the 25 most discriminative DART-HMRS signatures. HCA is an unsupervised, pattern recognition method, which means that it clusters the data into groups considering their nearness in the multidimensional space and visualizes the data to highlight their differences and similarities (Jiménez-Carvelo et al. 2019) (Medina et al. 2019). Being unsupervised, sample classification and prior information about the sample identity are unknown while performing this analysis (Dettmer, Aronov, and Hammock 2007). The result of HCA is presented in a dendrogram that detects groups of similar individuals not necessarily equivalent to the research groups (Brescia et al. 2002). HCA allows also the construction of a heatmap using a correlation matrix between the selected ions and the dietary roughage group. The positivity or negativity of the correlation between every selected ion to each of the three dietary groups is graphically represented by red or blue color, respectively. The strength of the correlation is represented by the intensity of the color itself. The results are graphically shown in *Figure 5.*

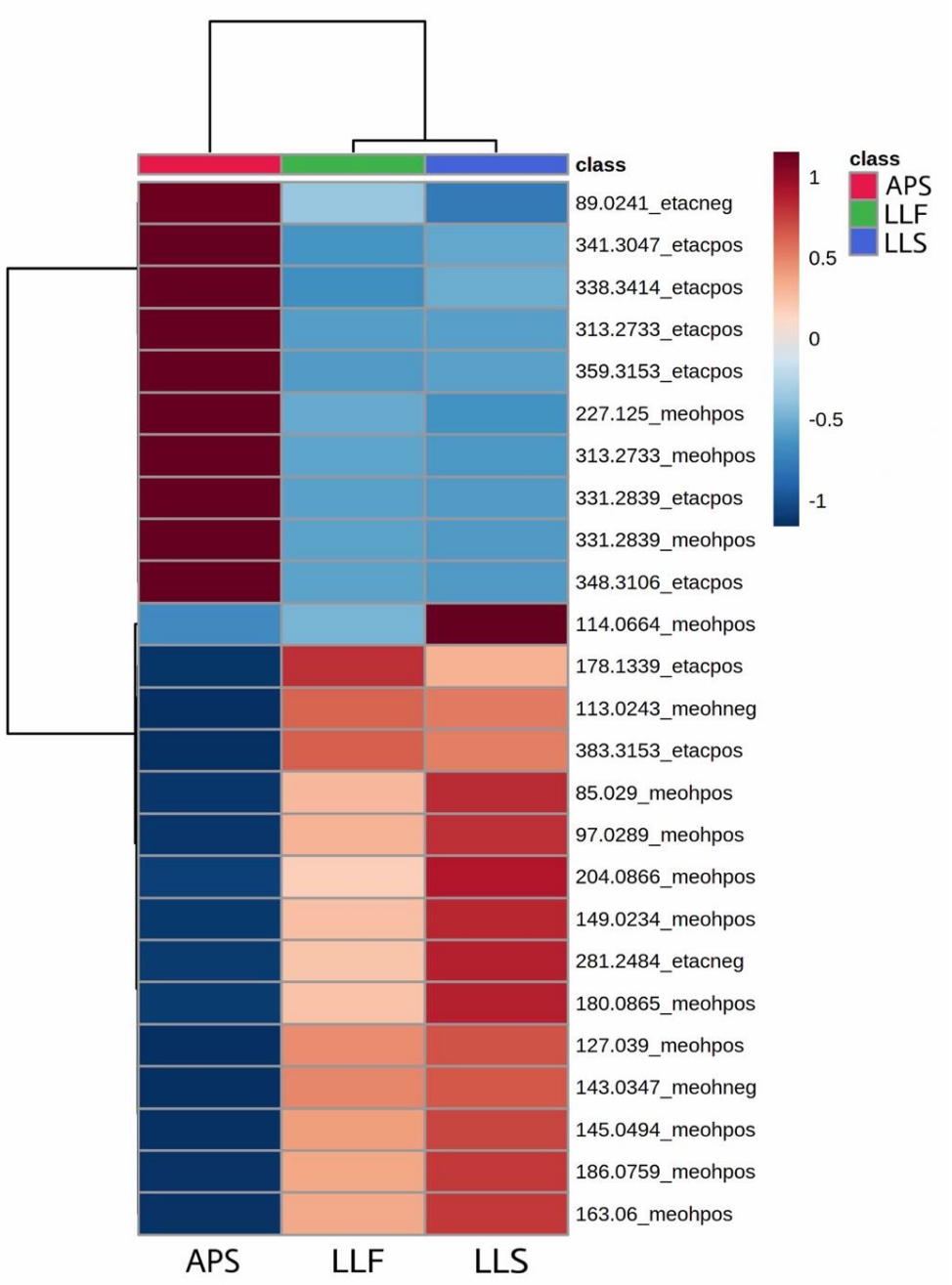


Figure 5. Heatmap obtained by hierarchical clustering analysis (HCA) of the informative molecular features (Pearson distance, Ward clustering algorithm). It shows the correlation between extrapolated features (metabolic ions) and the three dietary groups of the study. The red-brown (positive) and blue (negative) color scales indicate the degree of correlation between metabolic ions and feeding regimen; the two shorter Pearson's distance-tree clusters among the forage types (columns) and metabolites (rows) are represented by the branch height (the lower a node is vertically, the more similar its subtree is).

Abbreviations are: APS, Alpine pasture system; LLS, lowland silages, LLF lowland forages. APS is an Alpine (South Tyrol) dietary group meanwhile LLS and LLF are lowland (Veneto region) dietary groups. See the paragraph 3.4.3 and page 22 for more details about the 3 dietary theses.

In the first instance, the unsupervised HCA chemometric algorithm resulted in two main classes which can be classified as mountain-milk and lowland-milk clusters. However, the lowland-milk cluster was slightly divided into the two silage (LLS) and forage (LLF) sub-clusters. This HCA behavior suggests a better discrimination capacity between Alp and lowland milk-based samples, while only a poorly detectable difference within the lowland theses. Similar performances can be noticed by focusing on the heatmap reported in *Figure 5*. APS group is strongly positive correlated to ten putative ions related to the actual m/z values: (-) EtAc 89.0241, (+) EtAc 341.3047, 338.3414, 313.2733, 359.3153, 331.283 348.3106, and (+) MeOH 227.125, 313.2733, 331.2839. On the contrary, the same m/z values are negatively correlated to both lowland theses.
Lowland theses are correlated to the ions' block going from (+) EtAc 178.1339 to (+) MeOH 163.06, which is actually strongly negative correlated with APS group, sanctioning a clear division between the two dairy systems, which differ according to the environmental (mountain vs. lowland) and botanical origin of the roughage (forage) source as well as the herd characteristics.
In particular, lowland forage group (LLF) is more positively correlated with (+) EtAc 178.1339, 383.3153, and (-) MeOH 113.0243 ions, while lowland silage group (LLS) is stronger positively correlated with (+) MeOH 114.0664, 85.029, 97.0289, 204.0866, 149.0234, 180.0865, 127.039, 145.0494, 186.0759, 163.06, (-) EtAc 281.2484, and (-) MeOH 143.0347. The only strong detectable difference between LLS and LLF seems to be (+) MeOH 114.0664, which is strongly positively correlated with LLS and strongly negatively correlated with LLF. These results suggest that, within the lowland cluster composed of LLS and LLF, ions correlations might be interchangeable.

DART-HRMS fingerprinting of milk

Despite the study is characterized by a relatively small-sized and slightly unbalanced experimental feeding regimen, the main challenge is related to model building aiming to extrapolate the m/z values with the higher discriminative capability. For each of the most informative ions, a tentative molecule assignment of the *m/z* values was carried out with the use of the online METLIN (https://metlin.scripps.edu) and HUMAN METABOLOME DATABASE (www.hmdb.ca) libraries. The results of the tentative identification of metabolites corresponding to the APS's ions and lowland groups' ions are shown in *Table 5* and *Table 6*, respectively.

Table 5. List of discriminant (+/-) DART-HRMS metabolites detected in milk samples (discriminant model based on 3 dietary groups). The experimental *m/z*, theoretical *m/z*, error (ppm), elemental formula, type of ion, ion mode and extraction procedure, tentative assignment, and literature references are reported (***continues…***)

Dietary thesis	DART-HRMS *m/z*	Theoretical *m/z*	Error (ppm)	Elemental formula	Type of ion	Instrument ion mode and extraction solvent	Tentative assignment	Reference
Alpine Pasture (APS)	89.0241	89.0244	-3.37	$C_3H_6O_3$	$[M-H]^-$	(-) Pure EtAc	lactic acid	(Mordenti, Brogna, and Formigoni 2017) (Melzer et al. 2013) (Riuzzi et al. 2021) (Hrbek et al. 2014)
	227.1250	-		-	-	(+) MeOH:H_2O (80:20 v/v)	-	
	313.2733	313.2743	-3.19	$C_{19}H_{38}O_4$	$[M-H_2O+H]^+$	(+) Pure EtAc (+) MeOH:H_2O	MAG (16:0)	(Roda et al. 2015) (Corazzin et al. 2019)
	331.2839	331.2843	-1.2	$C_{19}H_{38}O_4$	$[M-H]^+$	(+) Pure EtAc (+) MeOH:H_2O	MAG (16:0)	(Roda et al. 2015) (Corazzin et al. 2019)
	341.3047	341.3056	-2.05	$C_{21}H_{42}O_4$	$[M-H_2O+H]^+$	(+) Pure EtAc	MAG (18:0)	(Corazzin et al. 2019) (Segato et al. 2017)
	348.3106	348.3108	-0.57	$C_{19}H_{38}O_4$	$[M+NH_4]^+$	(+) Pure EtAc	MAG (16:0)	(Roda et al. 2015) (Corazzin et al. 2019)
	359.3153	395.0958	-1.39	$C_{21}H_{42}O_4$	$[M-H]^+$	(+) Pure EtAc	MAG (18:0)	(Corazzin et al. 2019) (Segato et al. 2017)

Abbreviations are: APS, Alpine pasture system; LLS, lowland silages; LLF lowland forages. APS is an Alpine (South Tyrol) dietary group meanwhile LLS and LLF are lowland (Veneto region) dietary groups. See the paragraph 3.4.3 and page 22 for more details about the 3 dietary theses.

MAG, monoacylglycerol.

For the APS group, six out of seven discriminant m/z values were identified and assigned: lactic acid, MAG (16:0) three times, MAG (18:0) two times. 227.1250 m/z value is a discriminant for the milk of APS feeding system but the tentative assignment to the molecule failed.

The identification revealed redundancies of m/z values associated with the same metabolite MAG (16:0) and MAG (18:0). In the case of data fusion, a detection of similar molecular features from the original dataset [(+) or (-) both pure EtAc or MeOH: H2O, that it means two ion modes for two extraction solvents] may occur as a result of the chemometric modeling (Calderón-Santiago et al. 2016). Attempts to refine redundancies and, thus filtering off redundant m/z from the datasets in which their intensity is lower, have been carried out.

Lactic acid (89.0241 m/z) presence in milk is largely reported by the bibliography (Melzer et al. 2013), (Mordenti, Brogna, and Formigoni 2017), (Riuzzi et al. 2021) but it is usually related to a high starch intake, which is not a condition attributable to APS feeding system (*Table 3b).* However, it could be related to a ruminal unbalance between rapidly fermentable (i.e., starch and sugars from barley and/or early-stage plant rich in leaves) and structured (NDF)-carbohydrates. Additionally, it may be the consequence of a feeding condition based on the use of starch-rich supplements administered before and after (typically during the milking in the morning and evening) the daily grazing of high-mountain swards. Moreover, lactic acid has already been proposed as a biomarker candidate for mastitis since a positive correlation exists between lactic acid and SCC (57.80%) (Melzer et al. 2013). However, it cannot fully explain the lactic acid presence in APS milk since it has been already shown how low SCC in alpine milk is (*Table 4).*

MAG (16:0) (313.2733 m/z 331.2839 m/z and 348.3106 m/z) is the monoacylglycerol composed of glycerol and the saturated fatty acid (SFA) palmitic acid. Palmitic acid discriminant capacity is in line with Borreani et al. (Borreani et al. 2013) results, which reports that palmitic acid concentration is higher in hay-based milk production instead of silage-milk one.

MAG (18:0) (341.3047 m/z and 359.3153 m/z) is the monoacylglycerol composed of glycerol and the SFA stearic acid. According to Corazzin et al. and Segato et al. (Corazzin et al. 2019) (Segato et al. 2017), it is correlated to high-input and hay-based feeding systems.

Milk and milk derivatives are important sources of fatty acids (FA) and there is still an open debate about the effects of FA on human health. Saturated fatty acids (SFA) have been associated with human cardiovascular health problems, while monounsaturated fatty acids (MUFAs) and polyunsaturated fatty acids (PUFAs) are accountable for a lot of beneficial effect on health. Particularly appreciated are the PUFAs, to which belong omega 3, omega 6, and the conjugated linoleic acid (CLA) as well. (Roda et al. 2015) (Borreani et al. 2013) (Nielsen et al. 2006) (Corazzin et al. 2019).

Mountain milk is bibliographically associated with a higher content of beneficial MUFA, PUFA, and particularly CLA so that the detection of saturated monoacylglycerols in APS group might seem a contradiction (Sozooalp 2004) (Dhiman et al. 1999) (Roda et al. 2015) (Segato et al. 2017). However, it must be noticed that DART-HRMS coupled with the proposed chemometric model, lead to the detection of discriminating biomarkers and not to an evaluation of the milk composition. Therefore, the presence of saturated MAGs, doesn't mean that APS milk had a higher content of saturated fatty acids (SFAs).

The C16:0, cis-9 C18:1 and C18:2 n-6 were the most abundant SA, MUFA and PUFA, respectively. Generally, the *sn*-1 and *sn*-2 position are mainly esterified by palmitic acid, whereas short FA such as butyric, caproic, and caprylic acids prefer the sn-3 position of the glycerol backbone. The PUFAs that are present in milk in low amounts prefer the primary positions (*sn*-1) of the TAG. It is well known that milk contains a potent indigenous lipoprotein lipase (LPL) which has an optimum of activity at the *sn*-1 and *sn*-3 position of triglycerides (Collins, McSweeney, and Wilkinson 2003) (Cossignani, Pollini, and Blasi 2019). Therefore, as a result of the early enzymatic activity of LPL in milk, trace of MAG rich in C18:0 and C16:0 may occur and they may be play a role as metabolites of specific dairy systems. The different FA distribution over the 3 *sn* positions and the environmental-specific activity of LPL (mountain vs. lowland) maybe a key-factor of the milk authentication chemometric approach, even if this lipidomic analysis has been going to enhance by further studies. Indeed, large botanical diversity and environmental eating conditions/forage preservation methods (e.g., outdoor/pasture grazing vs. indoor/hay) may influence the milk microbiota originating from teat skin and the following specific enzymatic activities of the microflora conveyed in the milk (Segato et al. 2019) (Moreira et al. 2018) (Rocchetti et al. 2020).

Whether they derive from in-source DART fragmentations or from hydrolysis phenomena in milk, there is no clear explanation for the detection of MAG as milk biomarkers for the APS samples ($r = 0.7$ on average).

Table 6. ***(…continued)*** List of discriminant (+/-) DART-HRMS metabolites detected in milk samples (discriminant model based on 3 dietary groups). The experimental *m/z*, theoretical *m/z*, error (ppm), elemental formula, type of ion, ion mode and extraction procedure, tentative assignment, and literature references are reported.

Dietary thesis	DART-HRMS *m/z*	Theoretical *m/z*	Error (ppm)	Elemental formula	Type of ion	Instrument ion mode and extraction solvent	Tentative assignment	Reference
Lowland forage (LLF)	113.0243	113.0239	3.54	$C_5H_8O_4$	$[M-H-H_2O]^-$	(-) MeOH:H_2O (80:20 v/v)	acetolactate	(Mohr et al. 1997)
	178.1339	178.1339	0	$C_{10}H_{12}N_2$	$[M+NH_4]^+$	(+) Pure EtAc	norgramine	(Riuzzi et al. 2021)
	383.3153	383.3156	-0.78	$C_{23}H_{42}O_4$	$[M + H]^+$	(+) Pure EtAc	MAG (20:2)	(Hrbek et al. 2014) (Borreani et al. 2013) (Riuzzi et al. 2021)
Lowland silage (LLS)	85.0290	85.0290	0	$C_4H_6O_3$	$[M-H_2O+H]^+$	(+) MeOH:H_2O (80:20 v/v)	ketoacid derivate	
	97.0289	97.0290	1.05	$C_5H_6O_3$	$[M-H_2O+H]^+$	(+) MeOH:H_2O (80:20 v/v)	ketoacid derivate	
	114.0664	114.0662	1.75	$C_4H_7N_3O$	$[M + H]^+$	(+) MeOH:H_2O (80:20 v/v)	creatinine	(Scano et al. 2014) (Foroutan and et al. 2019) (Tenori et al. 2018) (Riuzzi et al. 2021) (Sun et al. 2017)
	127.0390	127.0390	0	$C_6H_6O_3$	$[M + H]^+$	(+) MeOH:H_2O (80:20 v/v)	methyl 2-furoate	(Hrbek et al. 2014) (Riuzzi et al. 2021)

Lowland silage (LLS)	143.0347	143.0344	2.1	$C_{30}H_{48}O_3$	$[M-H]^-$	(-) MeOH:H_2O (80:20 v/v)	3-hydroxy-2-methylglutarate or 2-hydroxy-2-ethylsuccinate	(Hrbek et al. 2014) (Riuzzi et al. 2021)
	145.0494	145.0495	0.7	$C_6H_8O_4$	$[M + H]^+$	(+) MeOH:H_2O (80:20 v/v)	dimethyl fumarate	(Melzer et al. 2013) (Riuzzi et al. 2021)
	149.0234	-		-	-	(+) MeOH:H_2O (80:20 v/v)	-	
	163.0600	163.0607	-4.3	$C_6H_{12}O_6$	$[M-H_2O+H]^+$	(+) MeOH:H_2O (80:20 v/v)	glucose	(Mordenti, Brogna, and Formigoni 2017) (Melzer et al. 2013) (Riuzzi et al. 2021)
	180.0865	180.0861	2.22	$C_6H_{12}O_6$	$[M+NH_4-H_2O]^+$	(+) MeOH:H_2O (80:20 v/v)	glucosamine	
	204.0866	204.0872	-2.9	$C_8H_{15}NO_6$	$[M-H2O+H]^+$	(+) MeOH:H_2O (80:20 v/v)	N-acetyl-glucosamine	
	281.2484	281.2486	-0.7	$C_{18}H_{34}O_2$	$[M-H]^-$	(-) Pure EtAc	oleic acid	(Capuano et al. 2014) (Yang et al. 2019)

Abbreviations are: APS, Alpine pasture system; LLS, lowland silages; LLF lowland forages. APS is an Alpine (South Tyrol) dietary group meanwhile LLS and LLF are lowland (Veneto region) dietary groups. See the paragraph 3.4.3 and page 22 for more details about the 3 dietary theses.

MAG, monoacylglycerol.

For all the LLF m/z values, a tentative assignment was given: acetolactate, norgramine, and MAG (20:2).

Acetolactate (113.0243 m/z) presence is already bibliographically reported (Mohr et al. 1997).

Norgramine (178.1339 m/z) molecule is a ubiquitous presence in milk coming from intensive farms based on barley, cereals, and silage feeding systems (Riuzzi et al. 2021).

MAG (20:2) (383.3153 m/z) is the monoacylglycerol of an omega 6: the polyunsaturated eicosadienoic acid. According to Corazzin (Corazzin et al. 2019), herbage intake increased the level both of PUFA and MUFA.

Ten out of eleven discriminant m/z values were assigned for the LLS group: two ketoacid derivates, creatinine, methyl 2-furoate, 3-hydroxy-2-methylglutarate or 2-hydroxy-2-ethylsuccinate, dimethyl fumarate, glucose, glucosamine, N-acetyl-glucosamine, and oleic acid. For 149.0234 m/z value, discriminating value for LLS milk, no tentative assignment was achieved.

Creatinine (114.0664 m/z value) was found as a potential marker of the health status of cows and in the diagnosing of the heat stress status in dairy cows, which may be derived from the phosphocreatine in the muscle tissue that has been mobilized for energy supply (Sun et al. 2017).

Methyl 2-furoate (127.0390 m/z) was bibliographically reported by Riuzzi and Hrbek (Riuzzi et al. 2021) (Hrbek et al. 2014).

Dimethyl fumarate (145.0494 m/z) is a fumarate derivate also found by Melzer (Melzer et al. 2013).

For 143.0347 m/z value, a double identity has been assigned: **3-hydroxy-2-methylglutarate** or **2-hydroxy-2-ethylsuccinate**. The ion identification is not certain information but a tentative performed with the help of ions databases. In case of great uncertainty, two different assignments can be given to the same m/z value. 2-hydroxy-2-ethylsuccinate was also found in milk by Riuzzi (Riuzzi et al. 2021).

Glucose (163.0600 m/z) finding in the LLS group could be connected with the high starch intake, as shown in *Table 3b*, and with the rapidly fermentable sources of energy (Sun et al. 2017).

Oleic acid (281.2484 m/z) is considered a marker of corporal fat mobilization (Yang et al. 2019). Despite for this MUFA (C18:1 cis-9) was detected a strong seasonal effect (higher in summer than in winter), it is more frequently present in milk from ration based on ensiled or fresh grass compare to hay-based ones (Capuano et al. 2014)
No information was found in the bibliography for ketoacid derivate, glucosamine, and N-acetyl-glucosamine, as milk biomarkers.

Validation of multivariate discriminant model

The twenty-five *m/z* values extrapolated by PLS-DA were used to construct a linear discriminant analysis (LDA) model on the training set. Its capability to correctly classify the samples according to the dietary groups was verified on the training set by 10 fold cross-validation. Furthermore, the LDA model was performed against the independent validation set. The predictions of this blind verification were reported in the confusion matrix reported below (*Table 7*), and a set of statistical measurements (accuracy, sensitivity, specificity, precision, and Matthews correlation coefficient) were calculated to assess the predictive discriminating capacity of the supervised classifier LDA model.
In this blind verification, that was carried out in the independent validation set, the classification model correctly predicted the APS samples (as indicated by the MCC value, the overall correct classification was 1.00). However, pooling together the two maize silage- and the maize silage free-based (LLS and LLF), milk collections did not permit a correct prediction of the two lowland feeding groups (the overall correct classification was lower than 0.20). This set of results demonstrates that the DART-HRMS chemometric approach, coupled with robust multivariate modelling, is able to authenticate the "all seasons" raw bulk mountain milk compare to the lowland dairy production. On the contrary, the relatively high misclassification rate between the two lowland theses confirmed that the intensive dairy systems would be characterized by a relatively low variability in terms of feeding system. Moreover, the seasonal feeding variations linked to cows' lactation stage and health status, and type of farming management, might reduce chemical differences among milk samples from dietary regimens with or without maize silage leading to a limited reliability capacity of metabolomics in tracing the dairy products.

Table 7. Confusion matrix reporting the predicted class probabilities (number of true positives, true negatives, false positives, and false negatives) and the descriptive statistic (sensitivity, specificity, accuracy, precision, and Matthews correlation coefficient) of the milk samples based on the predictions of the LDA model on the (+ / -) DART-HRMS validation set (n = 25)

		Actual class		
		APS	**LLF**	**LLS**
Predicted class	**APS**	**5**	0	0
	LLF	0	**6**	8
	LLS	0	3	**3**
Predictive statistics				
Sensitivity		1.00	0.67	0.27
Specificity		1.00	0.50	0.79
Accuracy		1.00	0.56	0.56
Precision		1.00	0.43	0.50
Matthews correlation coefficient		1.00	0.16	0.07

Abbreviations are: APS, Alpine pasture system; LLS, lowland silages, LLF lowland forages. APS is an Alpine (South Tyrol) dietary group meanwhile LLF and LLS are lowland (Veneto region) dietary groups. Bold values represent the samples that were classified correctly.

5. CONCLUSIONS

During 2019, 88 raw bulk milk samples were collected from intensive lowland dairy farms and extensive South Tyrolean farms. Milk samples were divided into five groups according to geographical origin and to the main dietary roughage: HMS (prevalence of maize silage, lowland), MMS (prevalence of cereal and forage silages, lowland), FCG (prevalence of ensiled crop and fresh forage, lowland), HAY (prevalence of dried forage, lowland), and APS (mountain forage, South Tyrol). The milk chemical composition (n=88) showed high-quality traits in the most productive groups (HMS and MMS) and in the alpine group as well; particularly, APS had quality values comparable to highly specialized rearing. The samples were analyzed by a (+/-) DART-HRMS, then the spectra sub-datasets were merged through a low-level data fusion. 75% of the samples have undergone a partial least squared discriminant analysis (PLS-DA), according to the five experimental groups. After a first attempt of modelling, the five dietary theses were redistributed into three groups: LLS (HMS and MMS), LLF (FCG and HAY), and APS, and the PLS-DA was performed again. Both multivariate statistical analyses have revealed a great ability in discriminating the Alp group (APS) from the lowland ones, while only a slight increase in discriminatory capacity among the lowland groups was achieved passing from the five-groups to the three-groups approach.

The statistical modelling approach allowed to identify the metabolites which contributed the most to discriminating the groups. By eliminating the m/z values with a coefficient lower than 30, the remaining 25 most informative ions were submitted to a hierarchical cluster analysis (HCA), which results in an unsupervised clustering of the data and the construction of a heatmap. Again, the clustering revealed a reliable capacity of the chemometric model into dividing South Tyrolean data (APS) from the lowland ones, which were only secondly separated into two different categories representative of LLS and LLF data. The heatmap provided a graphical representation of the m/z signals which discriminated the groups and a tentative assignment of these DART signatures was carried out. APS group was discriminated most by lactic acid and MAGs, while LLF group was mainly discriminated by MAG 20:2, acetolactate, and norgramine. Among the molecules that discriminated against LLS feeding system were found ketoacid derivates,

creatinine, methyl 2-furoate, 3-hydroxy-2-methylglutarate or 2-hydroxy-2-ethylsuccinate, dimethyl fumarate, glucose, glucosamine, N-acetyl-glucosamine, and oleic acid.

Finally, the discriminative capacity of DART analysis was performed by a linear discriminant analysis (LDA) based on the 25 most informative ions and validate on the independent validation set, composed of the samples withheld previously, throughout blind cross-validation. The classification performances of the LDA model were summarized in a confusion matrix, which underlined a correct prediction of the APS (misclassification rate of 0.0) meanwhile an overlapping between lowland-milk samples was still detected (misclassification rate of 0.55).

This study showed that a DART-HMRS coupled with a PLS-DA analysis was a successful approach to perform fast and accurate discrimination of milk samples only in the comparison between mountain (APS group) and lowland (LLS and LLF) feeding systems. APS milk production is connected with the use of mountain produced feedstock and the practice of summer alpine pasture in the Italian province of South Tyrol.

Milk produced with this feeding system seemed to have a higher market value due to a more sustainable production background perceived by consumers (deepening in the following appendix). Nevertheless, even if certain m/z values were founded for APS group and a tentative assignment was achieved, none of them was attributable to molecules with a marketable organoleptic value. It can be assessed that this modeling approach could be applied to certify milk South Tyrolean origin, basing on the specifical feeding system which reflects on specifical retrievable ions. However, these ions cannot yet be employed as biomarkers of a higher quality of milk.

APPENDIX

SOUTH TYROLEAN MILK

South Tyrol

Trentino-South Tyrol is one of the five Italian autonomous regions, is situated in the northeast part of the country and, particularly, is divided into two autonomous provinces: Trentino and South Tyrol which, bordering with Austria, is the northern one (46°30'0" N 11°19'59" E) (*Figure I*). Since 1972 it enjoys a special self-government form in order to safeguard the peculiar territory and culture.

The province has 531.178 inhabitants (situation at 31.12.18) of which 69,41% belong to the German language group and only 26,06% to the Italian group. People are spread among 7.400 m^2 of land, of which two-thirds are situated at an altitude greater than 1.500 m asl and only 14% below 1.000 m asl, so large areas are alpine pastures and 40% of the territory is protected landscape (Giunta provinciale di Bolzano 2019).

In addition to these particular geographical conditions, the rare plain valleys are mainly occupied by viticulture, orchards, and the few cities. As a result of this background, the dairy production realities are mainly dislocated on the mountains, forming a high quote capillary net.

Figure I. Physical map of Trentino-South Tyrol region, delimitation of the two provinces and their borders (Image from ontheworldmap.com).

The alpine meadow scene is mainly constituted by permanent, polyptych, natural, or spontaneous pastures. Spontaneous meadow originates from tree-felling and its maintenance strictly depends on the regular, annual presence of grazing cows (Ziliotto, Scotton, and Da Ronch 2004). Considering the great variability of pasture types, the most abundant forage species are *Agrostis tenuis, Anthoxanthum alpinum, Festuca rubra, Nardus stricta, Phleum alpinum, Poa alpina, Trifolium repens*, and *Achillea millefolium* (Orlandi, Clementel, Bovolenta 2005).

Traditional Dairy Farms

Market conditions, technological progress, and economic pressure have led to an intensification of dairy production in the last years. Consequently, milk production has averagely shifted to more profitable regions while in mountain areas, a high percentage of agricultural land has been abandoned, with severe consequences on landscapes, biodiversity, migration of people, and cultural loss because mountain dairy farms are of high importance for the local economy and the maintenance of traditional landscapes

(Kühl, Flach, and Gauly 2020). Counter-trend, South Tyrol has recognized the value and preserved its small-size extensive farms in the last decades.

Of the almost 4.500 farms, 70% of them are carried out as side jobs because this activity rarely supports a family on its own (Sennereiverband 2019). The average rearing density is 15 cows and 8 calves, the average milk production ranges from 25 to 30 L/d depending on the employed breeds (Federazione latterie Alto Adige 2020).

The traditional approach consists of an extensive grassland-based system. In wintertime the animals are raised in closed stalls, mainly tie, usually fed with hay silage or dried forages obtained from local meadow, and concentrates. During summer cows are usually led to the highland pasture, approximately from May to September, depending on the altitude and the climate (Sturaro et al. 2013) (Asaduzzaman et al. 2020). In fact, 34% of the South Tyrolean surface is classified as pasture. This annual cattle drive on and off the mountains is connected to folkloric events and popular traditions which attract many tourists every year (Altoadige-Tirolo 2020).

Two-thirds of South Tyrolean farms are single-breed (Zanon et al. 2020). The breeds mainly chosen, range from high-yielding types like Brown Swiss (31.7%), Simmental (29.0%), and Holstein Friesian (19.9%), to the more rural breeds like Alpine Grey (13.2%), Pinzgauer (1.8%), Jersey (1%) and other breeds (3.4%) like Pustertaler Sprinzen and to find a balance between productive ambitions and resilience necessity (Rinderzuchtverband 2019). Brown is considered the most interesting for alpine dairy farming to achieve optimal milk quality. Simmental is well appreciated for its good milk performance, coupled with its robustness, high carcass value, high market value of calves, and adaptability to the mountain farming system (Zanon et al. 2020). For Alpine Grey rearing subsidies are provided since it is considered an endangered breed (Kühl, Flach, and Gauly 2020).

As already seen, farms are distributed all over the valleys and mountains, making the productive scenario really heterogeneous. Some farms are located in almost inaccessible areas, reachable only by forest roads or cableways, like the one shown in *Figure II*. For these and a lot of other reasons, production costs are clearly higher compared to lowland areas or other countries: specifically, the cost per kg ECM (energy corrected milk) is between 58.1 c/kg and 70.6 c/ costs, averagely more than countries like Germany (36.5 c/kg), France (36.8 c/kg) or rest of Italy (30.5 c/kg) (Kühl, Flach, and Gauly 2020).

In order not to disadvantage them, some measures have been studied: for example, insemination costs are the same for big accessible farms as for distant small rearing (Vereinigung der südtiroler Tierzuchtverbände. 2019).

Figure II. A dairy farm in Varna, South Tyrol. Here the production keeps going during the whole year, not only in summer. The nearest paved street is distant 6 kilometers from here.

Chain

Except for the rare self-productive and selling realities, which usually concentrate their activity on the summer huts, milk is collected every day by refrigerated tank trucks. They pick up milk directly from the farm or from a collecting center and they transfer it, in less than 24 h, to one of the nine dairy cooperatives dislocated on the territory (*Figure IV*). Considering the geographical location of the farms, milk collection is often tricky and not always paved (*Figure III*): some situations are even solved with the use of cableways (Federazione latterie Alto Adige 2020).

Figure III. A refrigerated tank truck that every day deals with gravel paths to collect milk on pastures

Figure IV. The nine dairy cooperatives of the South Tyrolean province (Federazione latterie Alto Adige 2020).

Every cooperative has its own production chains, its brand, and flagship products (i.e. Psairer Bergkäserei produces only organic products) (Federazione latterie Alto Adige 2020).

Above these small industrial realities is the oversight of the "Sennereiverband Südtirol, Federazione latteria Alto Adige" (South Tyrol Dairy Federation) which coordinates South Tyrolean milk production, its quality standards, laboratory controls, and farmers education.

The federation deals with raw milk chemical test, with the test for milk-detectable cow-infections such as IBR, brucellosis, enzootic

bovine leucosis (4.588 analyzed samples during 2019), with truck tank checks, and with instruments' calibration.

In order to ensure the measures' reliability, the federation's lab performs crossed analysis with other certified labs in Europe.

On the final product, traceability proves are performed as well as residue tests like dioxin, aflatoxin, PBC, organophosphorus compounds, benzimidazoles, avermectin.

Throughout 2019, 3.829 products have been checked, 57.141 microbiological analysis have been performed, 19.631 chemical analysis, 8.380 physical analysis, 3.191 sensory analysis, taking into consideration 1.334 different parameters (Federazione latterie Alto Adige 2020)

Aim

70% of Italian annual milk production (12 mio.t) is yielded in three regions: Veneto, Lombardy, Emilia-Romagna. Considering the higher cost of production (Kühl, Flach, and Gauly 2020) due to peculiar difficulties, it is clear that South Tyrolean dairy production cannot compete on quantity with other regions and states (Sennereiverband 2019). Farms cannot evolve in big, high populated rearing because of the morphological conformation of the territory and the impossibility to use big motorized tools. Also, they cannot employ high-yielding breeds because high resiliency is requested on the alpine pasture and those breeds are often not appropriate for grazing (Kühl, Flach, and Gauly 2020).

Quality

In this perspective, South Tyrolean farmers have to maximize efficiency and profitability through the enhanced nutritional and organoleptic quality of their milk and dairy-products (Zanon et al. 2020).

In the dairy field, to be considered high quality by the Ministry of Health, milk has to fulfill:

- Composition requirements:
  - Fat content not lower than 3,50%
  - Protein content not lower than 32,0 g/L
- Hygienic-sanitary requirements:
  - Bacterial charge at +30 °C not greater than 100.000 /ml
  - Somatic cell content not greater than 300.000 /ml
  - Lactic acid content not greater than 30 p.p.m.

(Ministero della Sanità 1991) [D.M n 185/1991]

Actually, South Tyrolean milk greatly fulfills such parameters, as shown in *Table I*

Table I. Average South Tyrolean milk content in 2019.

Components	**Value**
Somatic Cell Count SCC/ml	204.000
Fat %	4.17
Protein %	3.55
Lactose %	4.74
Urea mg/dl	22.1
Casein %	2.82

Source: (Vereinigung der südtiroler Tierzuchtverbände. 2019)

Figure V. The logos of the "Qualità Alto Adige-Qualität Südtirol" brand, in both South Tyrolean languages. From www.provincia.bz.it

In order to protect and support high-quality industry, South Tyrol has established with the provincial law n.12 of 22 December 2005, an umbrella brand for its products. The brand covers all the DOP, IGP, DOC products like cheese and other dairies, speck, wines, honey, eggs, apple juice, liquors, groceries, but can also be used for productive enterprises and service offices. The logo (*Figure V*) is a guarantor of quality and geographical origin, certifying for a 100% South Tyrolean production, processing, and optimizing marketing efforts (Regolamento marchio ombrello Alto Adige 2018).
Regarding dairy, the umbrella logo equals to:

- Only South Tyrolean farm milk
- Total absence of GMO
- Quality controls from the milking to the processing
- Animal welfare care

(Federazione latterie Alto Adige 2020)

Quality has also been pursued with the development of some specific products which better embodied the South Tyrolean bond with mountain and agricultural production, like "Heumilch" (Hay-Milk). It is an Austrian traditional dairy product, which disciplinary forbids silage and GMO use. Europe recognizes "Heumilch" as "Traditional Specialty Guaranteed" (TSG) which allows its production in European states different from Austria, as long as the production specification is being respected (Latte fieno 2020). The diet mainly consists of pasture or hay, depending on the season, with the integration of bran

and protein crops (Federazione latterie Alto Adige 2020). In 2019, 18,8% of the South Tyrolean milk belonged to "Heumilch" category, and 4,0% was organic "Heumilch", both productions were greater than the simple organic production (0,17%) and both had a positive trend (*Table II*). (Sennereiverband 2019).

Table II. South Tyrolean milk production in 2019 and its distribution into categories.

	2019	**Compare to 2018**
Delivered bovine milk	399,108 t	- 1.7 %
….of which		
Organic	711 t	- 33.0%
"Heumilch"	75,156 t	+ 19.8%
Organic "Heumilch"	15,905 t	+ 24.0%

Source: (Sennereiverband 2019).

To uphold quality and cover higher production costs, the revenue per kg to farmers is averagely greater than in other productive areas. The majority of the small-size realities would be loss-making if there weren't subsidies connected with feeding strategy and farm structure, premium payments for value-added milk products to support extensive farms. Data in *Table III* refer to 2019 production (Sennereiverband 2019) (Kühl, Flach, and Gauly 2020).

Table III. Average revenue (in €) per kg of milk in South Tyrol and other productive areas in the world.

Area	Revenue	Area	Revenue
South Tyrolean high quality milk	50.4 cent/kg	Germany	33.1 cent/kg
South Tyrolean organic milk	70.4 cent/kg	French	35.0 cent/kg
Average South Tyrolean milk	51.2 cent/ kg	Switzerland	56.5 cent/kg
EU	34.3 cent/kg	New Zealand	30.4 cent/kg
Lombardy	40.6 cent/kg	USA	36.6 cent/kg
Austria	32.7 cent/kg		

Source: (Sennereiverband 2019)

Added Value

Summer pasture is not a direct profitable practice, it has been demonstrated the statistically significant link between production costs and the percentage of the herd on alpine pastures (*Figure VI*) (De Ros G., Baldessari E. 2005).

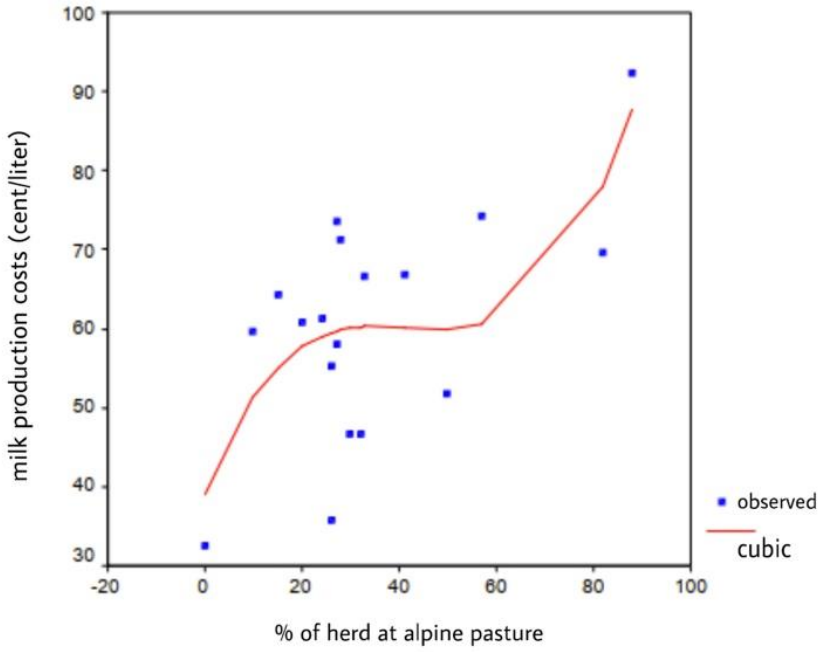
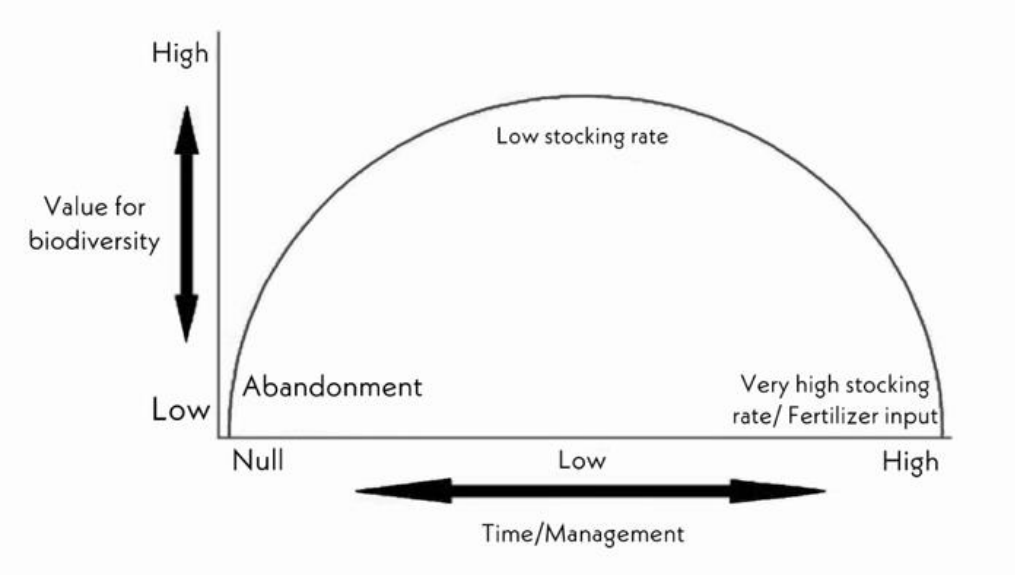


Figure VI. Milk production costs and utilization level of alpine pasture in 18 dairy farms province of Trento analyzed between 2000 and 2002. Edited from (De Ros G., Baldessari E. 2005)

Nevertheless, there are a lot of side plus-values in maintaining the cattle-drive in the mountains.

Biodiversity. It is long already known that the presence of grazing animals and the limitation to the reforestation of pasturing areas, enriches the landscape variability and above all biodiversity. The main reason is due to frequent cows' cut of the grass, which allows small vegetal species to survive despite the presence of strong, bigger species that would prevail if left free to grow. Consequently, a greater vegetation variability offers habitat opportunities for more wildlife species (Ziliotto, Scotton, and Da Ronch 2004) (*Figure VII*). Furthermore, mountain dairy farms are mostly populated by different local breeds, thereby preserving their genetic heritage, like in the case of Tyrolean Grey cattle that are listed as endangered (Berton et al. 2016) (Kühl, Flach, and Gauly 2020).



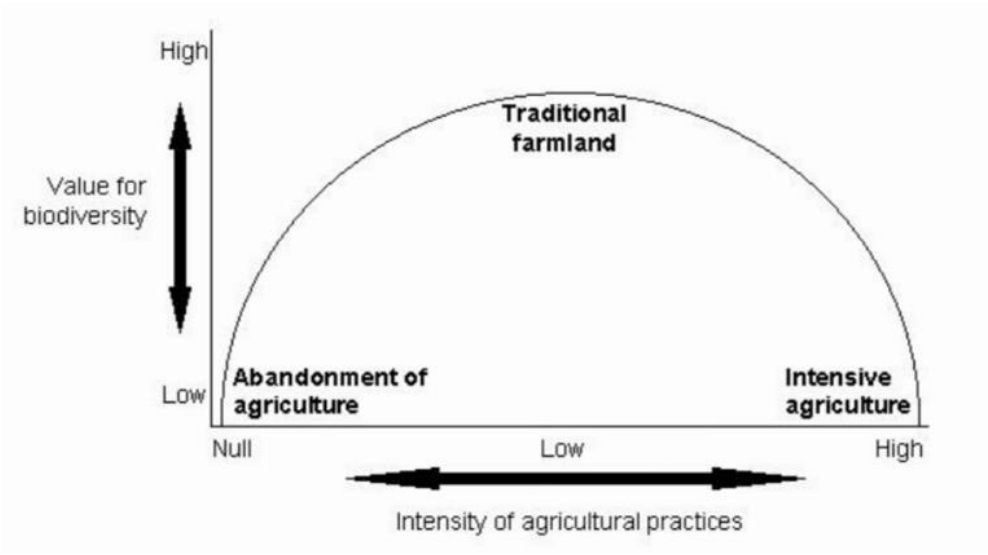


Figure VII. Impact of land use and breeding intensity on biodiversity. Edited from (Russo 2006).

- Ecosystem Service. Pastures' presence and the correlated human work, along with the construction of dry stone walls, play an important role to prevent natural hazard events like wildfires, avalanches, and erosions (Gusmeroli 2005).

Moreover, they reduce valley meadows' eutrophication (Ziliotto, Scotton, and Da Ronch 2004) and play a role in carbon sequestration (Corazzin et al. 2019).

- Sustainability. We can see there are a lot of strong points in favor of dairy farms in mountainous areas when using a multi-indicators approach to assess its sustainability without only considering the direct economical reward. Aside from the maintenance of local communities, the most interesting element is the low competition with human-edible resources because of the energy taken from otherwise unusable grasslands:

  He (human-edible) Feed Conversion Ration = 0.72 ± 0.46 MJ feed/MJ milk

  Considering that conventional dairy cows usually get more energy in human-edible feedstuffs than what they produce in milk, this ratio lower than 1, sets out grazing dairy cows like efficient food producers (Berton et al. 2016). In other words, traditional rearing consumes less arable land than the conventional one (Sturaro et al. 2013).

- Cultural heritage. Most of the mountain community life revolves around cattle rearing and driving to the pasture. The ascent and descent from the Alps are often celebrated with popular festivals, which are very attractive for tourists. Food, tales, and above all languages, are influenced by farm life. Stable communities are keepers of spoken knowledge and traditions (Corti 2005).

- Touristic appeal. A great interest in alpine pasture traditions has developed in the last years and even more, people are fascinated by what this reality includes. Farm accommodations, guided tours, trekking along the pasture paths and the alpine huts, rural food experiences, and popular festivals turn mountain rural communities into profitable holiday destinations (Corti 2005) (Sozooalp 2004). This is the major element to counterbalance lower-yielding quantity and higher production costs from an economical point of view.

The South Tyrolean dairy organizations and the province government aim to the development and rewarding of these values by providing funds, incrementing advertising campaigns, and certifying added values (Vereinigung der südtiroler Tierzuchtverbände. 2019).

Also Europe recognizes the relevance of promoting added values to protect the natural landscape, as declared in the European Landscape Convention signed by the Council of Europe in 2000 (Council of Europe 2000). In 2008 the European Environment Agency charted the map of the HNV, namely High Nature Value farmland, to identify farming systems and areas with a high biodiversity value (*Figure VIII*) (European Environment Agency 2012).

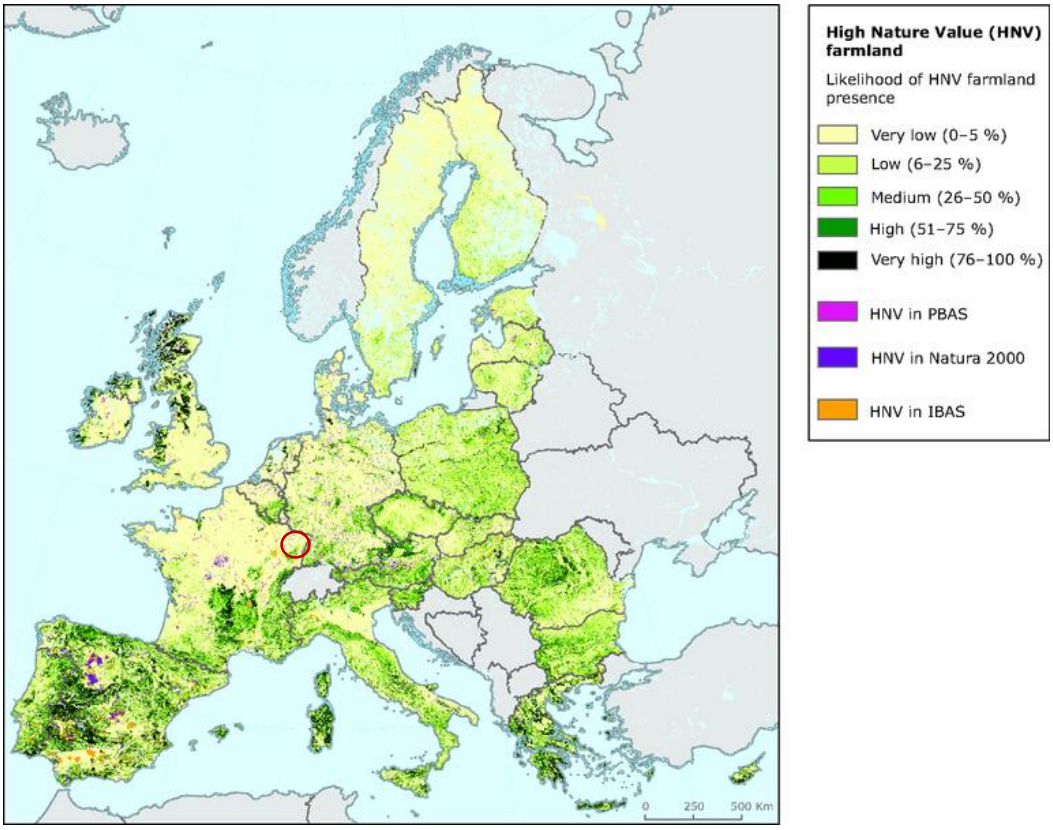


Figure VIII. European distribution of high nature value (HVN) farmlands with particular attention on the South Tyrolean area (red circle). Edited from (European Environment Agency 2012)

Improvement

According to the 2019 activity report of the breeder association "Vereinigung der südtiroler Tierzuchtverbände" (Union of South Tyrolean Animal Breeding Associations), 80% of the farms join voluntarily to performance tests, in order to improve the rearing performances and monitor the land advancement. The test is repeated in nine different periods of the year and consists of a comprehensive analysis of the animals and the rear, separating them according to age and breed to obtain a breed study as well, as shown in *Table IV* and *Table V*. Milk analysis are performed by the "Sennereiverband" in certified labs: protein, fat, urea, lactose, casein, somatic cells, and the derivable health state of cows. Furthermore, reproductive data are collected, like calving interval (CI), days open, fertility rates to artificial insemination or natural service (a practice still performed, even if decreasing), and infectious disease state like IBR, TBC, brucellosis, enzootic leucosis are monitored. These results are also shared with " Associazione Italiana Allevatori" (AIA) (Italian Breeders Association) in the project LEO, for the study and protection of bovine breeds and disease monitoring. (Vereinigung der südtiroler Tierzuchtverbände. 2019). During the last years, due to these constant monitoring and upgrades, a progressive and regular improvement in milk production (kg) was achieved, as shown in *Figure IX*.

Table IV. Herd breed incidence, breed's age, and reproductive performances (average data)

Breed	**age**		**Days Open**	**age at first calving**		**ASPP[1]**
	Years	Months		Years	Months	
Brown Swiss	4	7	159	2	7	2.2
Simmental	4	7	121	2	6	1.9
Alpine Grey	5	2	123	2	8	1.9
Holstein Friesian	4	1	160	2	5	2.1
Pinzgauer	4	9	133	2	7	2.1
Average	4	7	141	2	6	2.1

Source (Vereinigung der südtiroler Tierzuchtverbände. 2019). [1]Average service per pregnancy.

Table V. Average data of South Tyrolean dairy cows population in 2019.

Age	4 years and 7 months
Age at first calving	2 years and 6 months
Days open	141 days
ASPP[1]	2,1

Source (Vereinigung der südtiroler Tierzuchtverbände. 2019). [1]Average service per pregnancy.

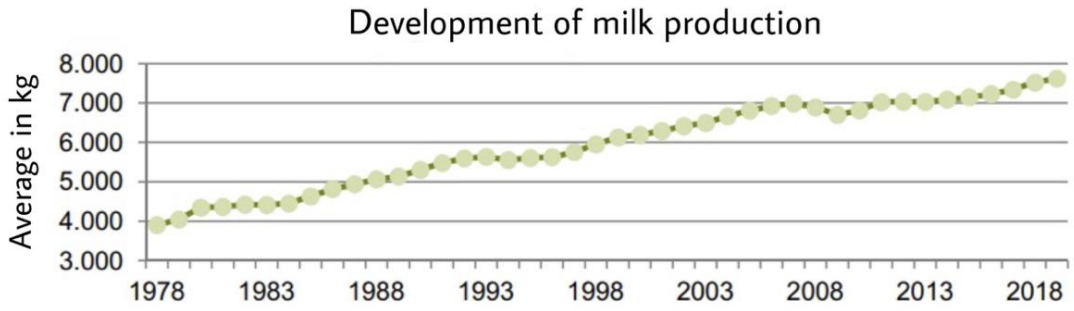


***Figure IX*.** Development of average milk production in South Tyrol from 1978 to 2018. Edited from (Vereinigung der südtiroler Tierzuchtverbände. 2019)

Communication

The fourth point that the South Tyrolean dairy organization aims to, to empower the local productive reality, is the communicative efficiency throughout every part of the capillary chain: from the most remote farm to the central federation and the government.

The SESAM project aims to develop user-friendly computer programs and to teach farmers how to use them (Vereinigung der südtiroler Tierzuchtverbände. 2019). Nevertheless, the internet is not available everywhere, so milk analysis results are communicated via SMS or Fax. A lot of workshops are organized by the breeders association but an important transversal capillary work is done by five farm advisers. The advisers go from farm to farm to perform random tests on milk, cows, work tools, verify

feed certifications but also to help and update farmers according to the “Sennereiverband” guidelines. (Sennereiverband 2019)

Conclusions

In conclusion, the conversion to high input, intensive, high-yielding rearing would seem to be economically advantageous for South Tyrolean small-scale dairy farmers. However, the geographical conformation would not allow such conversion on mountains and, moreover, this could result in problems regarding animal welfare, impact on environment and landscape, the regions attractiveness for tourists, loss of cultural heritage, and public acceptance. Furthermore, a low percentage of roughage used for milk production could lead to less fatty acids in the milk (Borreani et al. 2013). To support extensive mountain farms, the European community provides financial help (EC 2008) and has recently developed mountain product labels to attract the growing number of consumers
who perceive milk produced at higher altitudes as an ethical, natural, and sustainable choice (Asaduzzaman et al. 2020).
Also the Province of South Tyrol promotes the development and valorization of the traditional dairy industry with economical subsidies and scientific studies (Kühl, Flach, and Gauly 2020).

REFERENCES

ALEX HYDE. “High-Resolution Mass Spectrometry: More than Exact Mass.” https://www.bioanalysis-zone.com/high-resolution-mass-spectrometry-bioanalysis/#1.

Alhussien, Mohanned Naif, and Ajay Kumar Dang. 2018. “Milk Somatic Cells, Factors Influencing Their Release, Future Prospects, and Practical Utility in Dairy Animals: An Overview.” *Veterinary World* 11(5): 562–77.

“Altoadige-Tirolo.” 2020. https://www.altoadige-tirolo.com/il-rientro-dalle-malghe/.

Andrighetto, Igino et al. 2018. “Proposal and Validation of New Indexes to Evaluate Maize Silage Fermentative Quality in Lab-Scale Ensiling Conditions through the Use of a Receiver Operating Characteristic Analysis.” *Animal Feed Science and Technology* 242(May): 31–40.

AOAC. “No Title.” https://www.aoac.org.

“Art. 15 R.D. 9.5.1929, n. 994.”

Asaduzzaman, Mohammad et al. 2020. “Short-Wave near Infrared Spectroscopy for the Quality Control of Milk.” *Journal of Near Infrared Spectroscopy* 28(1): 3–9.

Audi, G., A.H. Wapstra, and C. Thibault. 2003. “The Ame2003 Atomic Mass Evaluation.” *Nuclear Physics A*.

Beauchemin, K. A. 1991. “Effects of Dietary Neutral Detergent Fiber Concentration and Alfalfa Hay Quality on Chewing, Rumen Function, and Milk Production of Dairy Cows.” *Journal of Dairy Science* 74(9): 3140–51.

Berton, M., E. Sturaro, M. Ramanzin, and G. Bittante. 2016. “Multi-Indicators Approach for the Evaluation of Efficiency of Mountain Dairy Farms.”

Bisutti, Vittoria et al. 2019. “Multivariate and Machine Learning Approaches for Honey Botanical Origin Authentication Using near Infrared Spectroscopy.” *Journal of Near Infrared Spectroscopy* 27(1): 65–74.

Borràs, Eva et al. 2015. “Data Fusion Methodologies for Food and Beverage Authentication and Quality Assessment - A Review.” *Analytica Chimica Acta* 891: 1–14.

Borreani, G. et al. 2013. “Effect of Different Feeding Strategies in Intensive Dairy Farming Systems on Milk Fatty Acid Profiles, and Implications on Feeding Costs in Italy.” *Journal of Dairy Science* 96(11): 6840–55.

Boudonck, Kurt J., Matthew W. Mitchell, Jacob. Wulff, and John A. Ryals. 2009.

"Characterization of the Biochemical Variability of Bovine Milk Using Metabolomics." *Metabolomics* 5(4): 375–86.

Brescia, M. A. et al. 2002. "Characterization of the Geographical Origin of Italian Red Wines Based on Traditional and Nuclear Magnetic Resonance Spectrometric Determinations." *Analytica Chimica Acta* 458(1): 177–86.

Calderón-Santiago, Mónica, María A. Fernández-Peralbo, Feliciano Priego-Capote, and María D. Luque De Castro. 2016. "MSCombine: A Tool for Merging Untargeted Metabolomic Data from High-Resolution Mass Spectrometry in the Positive and Negative Ionization Modes." *Metabolomics* 12(3).

Capuano, Edoardo et al. 2014. "Verification of Fresh Grass Feeding, Pasture Grazing and Organic Farming by Cows Farm Milk Fatty Acid Profile." *Food Chemistry* 164: 234–41.

Cevallos-Cevallos, Juan M. et al. 2009. "Metabolomic Analysis in Food Science: A Review." *Trends in Food Science and Technology* 20(11–12): 557–66.

Chang, Jen Shih, Phil A. Lawless, and Toshiaki Yamamoto. 1991. "Corona Discharge Processes." *IEEE Transactions on Plasma Science*.

Chernetsova, Elena S, G E Morlock, and Igor A Revelsky. 2011. "DART Mass Spectrometry and Its Applications in Chemical Analysis." *Russian Chemical Reviews*.

Cody, Robert B., James A. Laramée, and H. Dupont Durst. 2005. "Versatile New Ion Source for the Analysis of Materials in Open Air under Ambient Conditions." *Analytical Chemistry*.

Collins, Yvonne F., Paul L.H. McSweeney, and Martin G. Wilkinson. 2003. "Evidence of a Relationship between Autolysis of Starter Bacteria and Lipolysis in Cheddar Cheese during Ripening." *Journal of Dairy Research* 70(1): 105–13.

Corazzin, Mirco et al. 2019. "Fatty Acid Profiles of Cow's Milk and Cheese as Affected by Mountain Pasture Type and Concentrate Supplementation." *Animals* 9(2): 1–13.

Corti, M. 2005. "Le Valenze Turistiche Ed Educative Del Sistema Delle Alpi Ascolive: Indagine Sugli Eventi Turistici Sul Tema Dell'alpeggio." : 53–89.

Cossignani, L., L. Pollini, and F. Blasi. 2019. "Invited Review: Authentication of Milk by Direct and Indirect Analysis of Triacylglycerol Molecular Species." *Journal of Dairy Science* 102(7): 5871–82.

Council of Europe. 2000. "European Landscape Convention, Florence, 20 October 2000." *ETC No. 176*.

Cubero-Leon, Elena, Rosa Peñalver, and Alain Maquet. 2014. "Review on Metabolomics for Food Authentication." *Food Research International* 60: 95–107.

Dal Martello, Alessandro. 2020. “Metabolic Fingerprinting of Milk by Mid-Level Data Fusion of (+/-) DART-HRMS Spectra Relating to the Main Forage Source.”

Dettmer, Katja, Pavel A. Aronov, and Bruce D. Hammock. 2007. “Mass Spectrometry-Based Metabolomics.” *Mass Spectrometry Reviews*.

Dhiman, T. R., G. R. Anand, L. D. Satter, and M. W. Pariza. 1999. “Conjugated Linoleic Acid Content of Milk from Cows Fed Different Diets.” *Journal of Dairy Science* 82(10): 2146–56.

“European Environment Agency.” 2012. https://www.eea.europa.eu/data-and-maps/data/high-nature-value-farmland.

Federazione latterie Alto Adige. 2020. “AltoAdigelatte.” https://www.altoadigelatte.com/chi-siamo/settore-lattiero/.

Ferreira, D S, O F Galão, J A L Pallone, and R J Poppi. 2014. “Comparison and Application of Near-Infrared (NIR) and Mid-Infrared (MIR) Spectroscopy for Determination of Quality Parameters in Soybean Samples.” *Food Control* 35(1): 227–32.

Foroutan, Aidin, and et al. 2019. “Chemical Composition of Commercial Cow’s Milk.” *Journal of Agricultural and Food Chemistry*.

Giunta provinciale di Bolzano. 2019. *Manuale Dell’Alto Adige*.

Grelet, C. et al. 2015. “Standardization of Milk Mid-Infrared Spectra from a European Dairy Network.” *Journal of Dairy Science*.

Gross, Jürgen H. 2014. “Direct Analysis in Real Time-a Critical Review on DART-MS.” *Analytical and Bioanalytical Chemistry* 406(1): 63–80.

Gusmeroli, F. 2005. “Il Piano Di Pascolamento: Strumento Fondamentale per Una Corretta Gestione Del Pascolo.” : 27–41.

Hajslova, Jana, Tomas Cajka, and Lukas Vaclavik. 2011. “Challenging Applications Offered by Direct Analysis in Real Time (DART) in Food-Quality and Safety Analysis.” *TrAC - Trends in Analytical Chemistry* 30(2): 204–18.

Holman, Jerry D., David L. Tabb, and Parag Mallick. 2014. “Employing ProteoWizard to Convert Raw Mass Spectrometry Data.” *Current Protocols in Bioinformatics* (SUPPL.46): 1–9.

Hrbek, Vojtech, Lukas Vaclavik, Ondrej Elich, and Jana Hajslova. 2014. “Authentication of Milk and Milk-Based Foods by Direct Analysis in Real Time Ionization-High Resolution Mass Spectrometry (DART-HRMS) Technique: A Critical Assessment.” *Food Control* 36(1): 138–45.

ICAR. 2017. “Section 12 - Guidelines for Milk Analysis.” (January): 14.

IDF. “No Title.” https://www.fil-idf.org.

ISO. “No Title.” https://www.iso.org/home.html.

JEOL. “JEOL Ltd.” https://www.jeol.co.jp/en/.

Jiménez-Carvelo, Ana M., Antonio González-Casado, M. Gracia Bagur-González, and Luis Cuadros-Rodríguez. 2019. “Alternative Data Mining/Machine Learning Methods for the Analytical Evaluation of Food Quality and Authenticity – A Review.” *Food Research International*.

Jorabchi, Kaveh, Karl Hanold, and Jack Syage. 2013. “Ambient Analysis by Thermal Desorption Atmospheric Pressure Photoionization.” *Analytical and Bioanalytical Chemistry*.

Jordan, Kate W. et al. 2009. “Metabolomic Characterization of Human Rectal Adenocarcinoma with Intact Tissue Magnetic Resonance Spectroscopy.” *Diseases of the Colon and Rectum*.

Khan, Nazir A. et al. 2015. “Nutritive Value of Maize Silage in Relation to Dairy Cow Performance and Milk Quality.” *Journal of the Science of Food and Agriculture*.

Kühl, Sarah, Laura Flach, and Matthias Gauly. 2020. “Economic Assessment of Small-Scale Mountain Dairy Farms in South Tyrol Depending on Feed Intake and Breed.” *Italian Journal of Animal Science* 19(1): 41–50.

Lamanna, Raffaele, Angela Braca, Elvio Di Paolo, and Giovanna Imparato. 2011. “Identification of Milk Mixtures by 1H NMR Profiling.” *Magnetic Resonance in Chemistry*.

“Latte Fieno.” 2020. https://it.wikipedia.org/wiki/Latte_fieno.

Lindmark-Månsson, Helena, Rangne Fondén, and Hans Erik Pettersson. 2003. “Composition of Swedish Dairy Milk.” *International Dairy Journal* 13(6): 409–25.

McLafferty, F. W., and F. Turecek. 1993. “Interpretation of Mass Spectra.” In *Vol. 4*, ed. U. S Book. , 1993.

Medina, Sonia et al. 2019. “Current Trends and Recent Advances on Food Authenticity Technologies and Chemometric Approaches.” *Trends in Food Science and Technology*.

Melzer, N. et al. 2013. “Investigating Associations between Milk Metabolite Profiles and Milk Traits of Holstein Cows.” *Journal of Dairy Science* 96(3): 1521–34.

Ministero della Sanità. 1991. “Gazzetta Ufficiale D.M. 185.” http://www.granlatte.it/media/documents/granlatte_www/DL185_91.pdf.

Mohr, Britta et al. 1997. “A New Method for the Determination of 2-Acetolactate in Dairy Products.” *International Dairy Journal*.

Mordenti, Attilio Luigi, Nico Brogna, and Andrea Formigoni. 2017. “REVIEW: The Link between Feeding Dairy Cows and

Parmigiano-Reggiano Cheese Production Area." *Professional Animal Scientist* 33(5): 520–29.

Moreira, Gisela M.M. et al. 2018. "Effect of Ripening Time on Proteolysis, Free Amino Acids, Bioactive Amines and Texture Profile of Gorgonzola-Type Cheese." *Lwt* 98(September): 583–90.

Nielsen, Tina S., Ellen M. Straarup, Mogens Vestergaard, and Kris Sejrsen. 2006. "Effect of Silage Type and Concentrate Level on Conjugated Linoleic Acids, Trans-C18:1 Isomers and Fat Content in Milk from Dairy Cows." *Reproduction Nutrition Development* 46(6): 699–712.

Nutrient Requirements of Dairy Cattle. 2001. Nutrient Requirements of Dairy Cattle.

Orlandi, Clementel, Bovolenta, Dovier S. 2005. "Caratterizzazione Chimica e Nutrizionale Delle Principali Specie Pascolive Alpine." *Sozooalp*: 171–80.

Rego, Oldemiro A. et al. 2016. "Changes in Milk Production and Milk Fatty Acid Composition of Cows Switched from Pasture to a Total Mixed Ration Diet and Back to Pasture." *Italian Journal of Animal Science* 15(1): 76–86.

Regolamento Marchio Ombrello Alto Adige. 2018.

Reusch, William. 2013. "Mass Spectrometry." https://www2.chemistry.msu.edu/faculty/reusch/virttxtjml/Spectrpy/MassSpec/masspec1.htm.

Rinderzuchtverband. 2019. "Rinderzuchtverband." https://www.rinderzuchtverband.it/it/.

Riuzzi, Giorgia et al. 2021. "Authentication of Forage-Based Milk by Mid-Level Data Fusion of (+/−) DART-HRMS Signatures." *International Dairy Journal* 112: 104859.

Rocchetti, Gabriele et al. 2020. "Milk Metabolomics Based on Ultra-High-Performance Liquid Chromatography Coupled with Quadrupole Time-of-Flight Mass Spectrometry to Discriminate Different Cows Feeding Regimens." *Food Research International* 134(January): 109279. https://doi.org/10.1016/j.foodres.2020.109279.

Roda, Gabriella, Stefano Fialà, Michela Vittorini, and Francesco Secundo. 2015. "Fatty Acid Composition and Fat Content in Milk from Cows Grazing in the Alpine Region." *European Food Research and Technology* 241(3): 413–18.

De Ros G., Baldessari E., Ventura W. 2005. *I Costi Dell'alpeggio Sono Sostenibili?*

Russo, D. 2006. "Effects of Land Abandonment on Animal Species in Europe: Conservation and Management Implications." *Ecologia*.

Sano et al., Akihiro. 2005. 2 *Mass Spectrometer System*.

Scampicchio, Matteo et al. 2016. "Multi-Method Approach to Trace the

Geographical Origin of Alpine Milk: A Case Study of Tyrol Region." *Food Analytical Methods* 9(5): 1262–73.

Scano, Paola, Antonio Murgia, Filippo M. Pirisi, and Pierluigi Caboni. 2014. "A Gas Chromatography-Mass Spectrometry-Based Metabolomic Approach for the Characterization of Goat Milk Compared with Cow Milk." *Journal of Dairy Science* 97(10): 6057–66.

Scheer, Elke et al. 1998. "The Signature of Chemical Valence in the Electrical Conduction through a Single-Atom Contact." *Nature*.

Schwolow, Sebastian, Natalie Gerhardt, Sascha Rohn, and Philipp Weller. 2019. "Data Fusion of GC-IMS Data and FT-MIR Spectra for the Authentication of Olive Oils and Honeys—Is It Worth to Go the Extra Mile?" *Analytical and Bioanalytical Chemistry*.

Segato, Severino et al. 2017. "Identification of Lipid Biomarkers to Discriminate between the Different Production Systems for Asiago PDO Cheese." *Journal of Agricultural and Food Chemistry* 65(45): 9887–92.

———. 2019. "1H NMR Metabolic Profile to Discriminate Pasture Based Alpine Asiago PDO Cheeses." *Animals* 9(10): 1–11.

Sejrsen, K, T Hvelplund, and M O Nielsen. 2009. Ruminant physiology *Ruminant Physiology*.

Sennereiverband. 2019. *Tätigkeitsbericht*.

Shulaev, Vladimir. 2006. "Metabolomics Technology and Bioinformatics." *Briefings in Bioinformatics* 7(2): 128–39.

Song, Liguo, Andrew B. Dykstra, Huifang Yao, and John E. Bartmess. 2009. "Ionization Mechanism of Negative Ion-Direct Analysis in Real Time: A Comparative Study with Negative Ion-Atmospheric Pressure Photoionization." *Journal of the American Society for Mass Spectrometry*.

Sozooalp, Quaderni. 2004. "Il Sistema Delle Malghe Alpine."

"Spettrometria Di Massa." 2019. *12 luglio*. https://it.wikipedia.org/wiki/Spettrometria_di_massa.

Sturaro, E. et al. 2013. "Dairy Systems in Mountainous Areas: Farm Animal Biodiversity, Milk Production and Destination, and Land Use." *Livestock Science* 158(1–3): 157–68.

Sun, Hui Zeng et al. 2017. "Lactation-Related Metabolic Mechanism Investigated Based on Mammary Gland Metabolomics and 4 Biofluids' Metabolomics Relationships in Dairy Cows." *BMC Genomics*.

Tata, Alessandra et al. 2021. "Serum Metabolomic Profiles of Paratuberculosis Infected and Infectious Dairy Cattle by Ambient Mass Spectrometry." *Frontiers in Veterinary Science* 7(January): 1–10.

Tenori, Leonardo et al. 2018. "NMR Metabolomic Fingerprinting Distinguishes

Milk from Different Farms." *Food Research International*.

The commission of the European Communities. 2001. "Commission Regulation (EC) No 213/2001 of 9 January 2001 Laying down Detailed Rules for the Application of Council Regulation (EC) No 1255/1999 as Regards Methods for the Analysis and Quality Evaluation of Milk and Milk Products and Amending Regulations (." *Official Journal of the European Communities* 72(371): 63–67. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32001R0213&from=EN.

Tian, He et al. 2016. "Integrated Metabolomics Study of the Milk of Heat-Stressed Lactating Dairy Cows." *Scientific Reports*.

Treccani. "Latte." https://www.treccani.it/enciclopedia/latte/.

Vaclavik, Lukas et al. 2011. "Authentication of Animal Fats Using Direct Analysis in Real Time (DART) Ionization-Mass Spectrometry and Chemometric Tools." *Journal of Agricultural and Food Chemistry* 59(11): 5919–26.

Vereinigung der südtiroler Tierzuchtverbände. 2019. *Tätigkeitsbericht 2019_klein.Pdf*.

Weston, Daniel J. 2010. "Ambient Ionization Mass Spectrometry: Current Understanding of Mechanistic Theory; Analytical Performance and Application Areas." *Analyst* 135(4): 661–68.

Wishart, David S. et al. 2007. "HMDB: The Human Metabolome Database." *Nucleic Acids Research*.

Yang, Y., G. Ferreira, B. A. Corl, and B. T. Campbell. 2019. "Production Performance, Nutrient Digestibility, and Milk Fatty Acid Profile of Lactating Dairy Cows Fed Corn Silage- or Sorghum Silage-Based Diets with and without Xylanase Supplementation." *Journal of Dairy Science* 102(3): 2266–74.

Zanon, Thomas et al. 2020. "Bulk Milk Quality as Affected by Cattle Breed Composition of the Herd in Mountain Area." *Annals of Animal Science* (June).

Ziliotto, U., M. Scotton, and F. Da Ronch. 2004. "I Pascoli Alpini: Aspetti Ecologici e Vegetazionali." *Convegno SoZooAlp Il sistema delle malghe alpine: aspetti agro-zootecnici, paesaggistici e turistici*.

Acknowledgments

Desidero ringraziare:
Il professor Severino Segato per avermi coinvolta in questo entusiasmante progetto che mi ha permesso di portare un pezzo della mia Terra nel mio percorso di studi.
La Sennereiverband per aver risposto con entusiasmo al progetto e in particolare Annemarie Kaser e Michael Pichler per la loro grandissima disponibilità.
Gli allevatori altoatesini che mi hanno accolta nelle loro stalle rispondendo pazientemente alle domande, nonostante le mie gaffe linguistiche.
Thomas Facchini, Ilaria Lanza, Alessandro Dal Martello e Samantha Endrizzi per l'aiuto concreto che mi hanno dato nel compimento di questo lavoro.

A conclusione di questo percorso, vorrei rivolgere un grande grazie:
Ai miei genitori, Midars e Bongo, ai quali va il merito principale di questo mio traguardo e di molti altri nella mia vita, grazie al loro sostegno e al loro esempio.
A mia sorella Gloria, che ha sempre creduto in me più di un'intera squadra di cheerleaders.
Ai miei immancabili compagni di squadra Mirko e Luca, sempre pronti a sbattermi in faccia la verità ma pronti anche ad aiutarmi, incoraggiarmi e (goffamente) sostenermi.
Ai miei coinquilini e amici: Giulia, Anna, Clelia e Gianbarni, che hanno condiviso con me gioie e dolori della coinquilinanza, dell'università, e qualche bottiglia di vino.
Alle altre persone che, chi prima e chi dopo, hanno riempito questo percorso di avventure, risate, dilemmi ed esperienze indimenticabili: Lucia, Alex, Rachel, Checco, Cris, Mattia e Gianluca.
Alla mia mentore Stefania.
Alla famiglia Barin, che in questi anni mi ha accolta in casa loro come una figlia.
Infine vorrei ringraziare Daniel, in cui la vita mi ha fatto incappare dopo immensi assurdi giri. Grazie per la sua incredibile capacità di rendere ogni cosa spettacolare e per farmi credere che non ci sia nulla nella vita a cui io non possa arrivare.