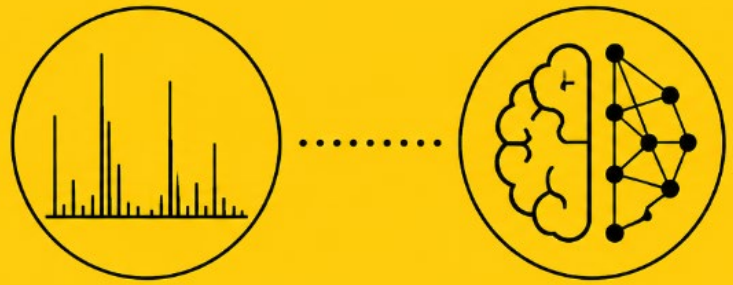
IZS TERAMO

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

INTEGRAZIONE TRA SPETTROMETRIA DI MASSA MALDI-TOF E MODELLI DI INTELLIGENZA ARTIFICIALE



Applicazioni su *Brucella* e *Campylobacter*

RICERCA CORRENTE: IZS AM 01/22 RC - MALD-IA

TERESA ROMUALDI

IZS TERAMO

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

MALDI-TOF MS

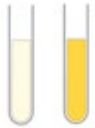
Una rivoluzione nell'identificazione dei microrganismi

IMPORTANZA DELL'IDENTIFICAZIONE RAPIDA

- ✓ Ridurre i tempi diagnostici
- ✓ Usare terapie mirate e tempestive
- ✓ Combattere l'antibiotico-resistenza
- ✓ Garantire sicurezza alimentare, salute animale, protezione ambientale e benessere umano

LIMITI DEGLI ALTRI METODI

TEST FENOTIPICI
Lenti

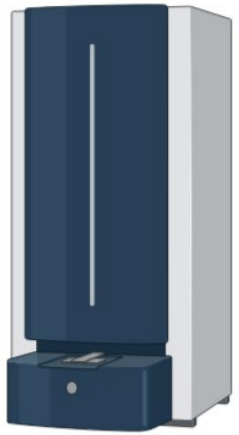


TEST BIOCHIMICI
Scarsa risoluzione

TEST MOLECOLARI
(PCR/Sequenziamento)
Costosi e analisi complessa

INTRODUZIONE DI MALDI-TOF MS

- ✓ Rapido
- ✓ Accurato
- ✓ Economico
- ✓ Facile esecuzione
- ✓ Facile interpretazione del risultato



Da giorni a pochi minuti

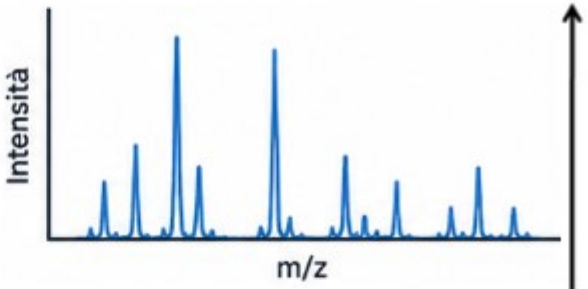
IZS TERAMO

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

MALDI-TOF MS

Principio di funzionamento

Database matching

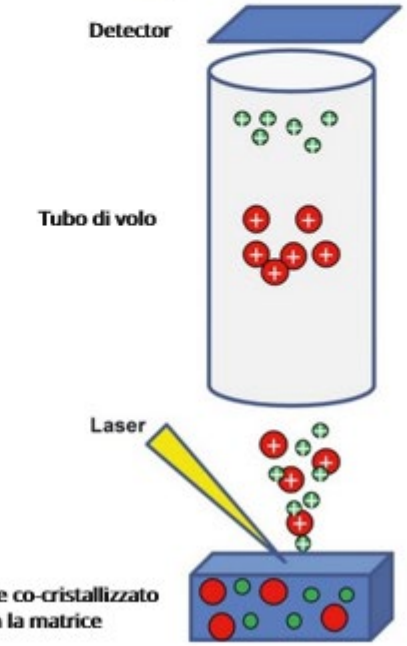


Generazione dello spettro di massa

Report d'identificazione

Sample Name	Sample ID	Organism (best match)	Score Value	Organism (second-best match)	Score Value
C7 (+++) (A)	2025.TE.13564.1.1 (Standard)	Campylobacter jejuni	2.30	Campylobacter jejuni	2.26
C8 (+++) (A)	2025.TE.13564.1.1 (Standard)	Campylobacter jejuni	2.24	Campylobacter jejuni	2.20
C9 (+) (B)	2025.TE.13568.1.1 (Standard)	Campylobacter lanienae	1.82	No Organism Identification Possible	1.37
C10 (-) (B)	2025.TE.13568.1.1 (Standard)	No Organism Identification Possible	1.54	No Organism Identification Possible	1.39

Score range	Interpretazione	Classe di consistenza
2.00 - 3.00	Identificazione attendibile a livello di specie	A
1.70 - 1.99	Identificazione attendibile a livello di genere	B
0.00 - 1.69	Mancata identificazione	C



Detection degli ioni

Separazione TOF in base al rapporto m/z

Accelerazione degli ioni

Ionizzazione MALDI mediante impulso laser

IZS TERAMO

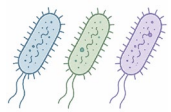
ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

PROGETTO MALD-IA

IA per l'analisi avanzata degli spettri MALDI-TOF

LIMITI DEL MALDI-TOF

Dipendenza dalle banche dati
Identificazione limitata a specie presenti nei database commerciali



Difficoltà con specie strettamente correlate
Limitata capacità discriminante tra ceppi filogeneticamente affini

Limitazioni nella sottotipizzazione
Non è possibile distinguere ceppi diversi della stessa specie

PROGETTO MALD-IA

Scopo
Sfruttare tutte le informazioni presenti nello spettro proteico acquisito al MALDI in modo ottimale

Approccio
Combinare i dati acquisiti tramite MALDI-TOF con strumenti di intelligenza artificiale/machine learning

Outcome attesi

- Migliorare discriminazione tra specie strettamente correlate
- Possibile sottotipizzazione di ceppi della stessa specie

IZS TERAMO

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

TASK *Brucella*

Potenziamento dell'identificazione di specie

Problema

- La brucellosi continua a rappresentare una zoonosi di rilevanza veterinaria e sanitaria in Italia.
- L'identificazione rapida a livello di specie è fondamentale per la sorveglianza epidemiologica.
- Le specie del genere *Brucella* sono filogeneticamente molto affini e scarsamente rappresentate nelle librerie commerciali MALDI-TOF.

Strategia adottata

- → Raccolta di ceppi rappresentativi delle principali specie di *Brucella* (collezione CNR Brucellosi).
- → Acquisizione di spettri MALDI-TOF.
- → Creazione di MSPs e addestramento di algoritmi di machine learning a partire dai dati di spettri proteici.

Obiettivo finale

- ✓ Realizzare una libreria MALDI-TOF e un modello predittivo in grado di identificare rapidamente e accuratamente le specie di *Brucella*, supportando le attività diagnostiche e di sorveglianza epidemiologica.

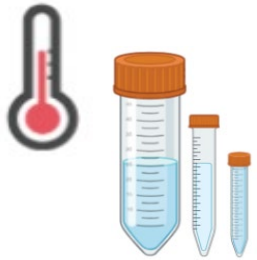
TASK *Brucella*

Workflow di acquisizione degli spettri

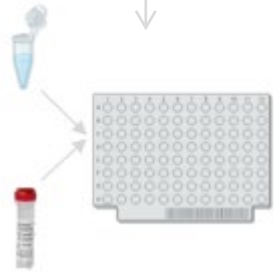
SPECIE	N° ceppi	N° spettri
B. abortus bv1	555	4218
B. abortus bv3		
B. melitensis		
B. suis bv2		
B. ceti		
Others		

Specie selezionate

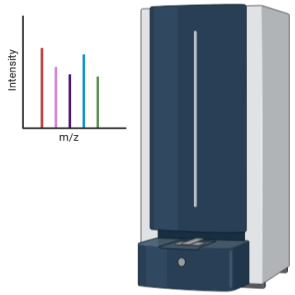
Coltura su *Brucella* Agar a 37°C, in aerobia con 5-10% CO_2 per 48-72 ore



Inattivazione per bollitura (100°C per 15 minuti) ed estrazione proteica (protocollo etanolo/acido formico/acetonitrile)



Allestimento piastra MALDI
(1 µL di estratto proteico + 1 µL di matrice HCCA)



Acquisizione spettri MALDI-TOF
- 24 spettri (8x3) per training set + MSPs
- 3 spettri (3x1) per test set

Creazione MSPs per espandere le librerie commerciali

Creazione dataset per addestramento di algoritmi di ML (cartella con raw data + file Excel con metadati)

Esempio di dato grezzo in formato .txt dove ad ogni valore m/z corrisponde un valore di intensità

IZS
TERAMO
ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

TASK *Brucella*

Workflow di sviluppo del modello RF

Costruzione del dataset | Preprocessing dei dati | Training del modello RF | Valutazione del modello

Classe	N campioni	Test set	Training set
B. abortus bv1	149	35	114
B. abortus bv3	173	42	131
B. ceti	32	4	28
B. melitensis	104	25	79
B. suis bv2	82	16	66
Others	15	4	11

IZS TERAMO

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

TASK *Brucella*

Risultati e discussioni

Accuratezza per spettri 91,9%		Classe reale					
		B. abortusbv1	*B. abortusbv3*	*B. ceti*	*B. melitensis*	*B. suisbv2*	Others
Classe predetta	*B. abortusbv1*	105	11	0	0	0	0
	B. abortusbv3	0	167	0	2	1	0
	B. ceti	0	0	36	0	0	0
	B. melitensis	0	0	0	200	12	24
	B. suisbv2	0	1	0	1	160	13
	Others	0	0	0	0	0	59

Accuratezza per NRG 95,2%		Classe reale					
		B. abortusbv1	*B. abortusbv3*	*B. ceti*	*B. melitensis*	*B. suisbv2*	Others
Classe predetta	*B. abortusbv1*	35	3	0	0	0	0
	B. abortusbv3	0	38	0	1	0	0
	B. ceti	0	0	4	0	0	0
	B. melitensis	0	1	0	25	0	1
	B. suisbv2	0	0	0	0	16	1
	Others	0	0	0	0	0	2

CLASSE	PARAMETRO	VALORE
B. abortusbv1	sensibilità	100,0%
	specificità	98,4%
B. abortusbv3	sensibilità	92,8%
	specificità	99,8%
B. ceti	sensibilità	100,0%
	specificità	100,0%
B. melitensis	sensibilità	99,5%
	specificità	93,6%
B. suisbv2	sensibilità	92,5%
	specificità	97,7%
Others	sensibilità	61,5%
	specificità	100,0%

- Il modello RF mostra alte prestazioni complessive, con sensibilità e specificità elevate per la maggior parte delle classi.
- Le migliori performance si osservano per *B. ceti* e *B. abortusbv1,* mentre la classe Others ha una sensibilità più bassa (0,61), probabilmente dovuta all'eterogeneità del gruppo.
- L'analisi a livello di NRG (ID unico campione) riduce la variabilità intra-ceppo e conferma la robustezza del modello.

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

TASK *Brucella*

Sviluppo della web application

Upload degli spettri

Analisi mediante algoritmo ML addestrato

Predizione della specie con probabilità associata

Distribuzione delle probabilità per ciascuna classe

Classificazione spettro MALDI

Seleziona il file dello spettro (.txt)

Browse... 2026.TE.6248.1.1_G7.txt

Upload complete

Esegui previsione

Classe predetta

B_suisbv2

Probabilità: 93.82%

Probabilità per classe

classe	probabilita
B_suisbv2	93.82%
Others	3.25%
B_melitensis	1.29%
B_abortusbv3	0.91%
B_ceti	0.49%
B_abortusbv1	0.23%

IZS TERAMO

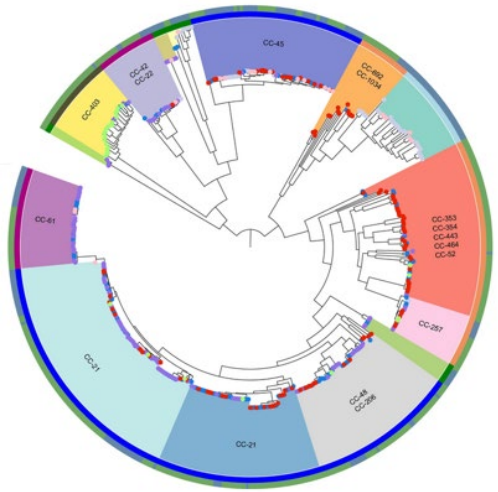
ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

Task *Campylobacter*

Tipizzazione rapida di *C. jejuni*

Problema

- Alcuni CC di *C. jejuni* mostrano un forte associazione con specifici serbatoi animali, ma la loro caratterizzazione mediante MLST richiede tempi di risposta lunghi.
- Il MALDI-TOF consente l'identificazione di specie, ma non la sottotipizzazione di ceppi di *Campylobacter jejuni.*
- Necessità di uno strumento rapido e accurato per supportare sorveglianza epidemiologica e outbreak investigation.



Strategia adottata

- → Costruire un database di spettri MALDI-TOF rappresentativi dei principali Clonal Complex (CC) e Sequence Types (ST) di *Campylobacter jejuni* isolati nel panorama italiano ed europeo a partire da diverse matrici.
- → Addestramento di algoritmi di Machine Learning sugli spettri proteici acquisiti.

Obiettivo finale

- ✓ Predire rapidamente i Clonal Complex di *C. jejuni* da dati MALDI-TOF, fornendo un supporto immediato alle indagini epidemiologiche e alla source attribution.

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

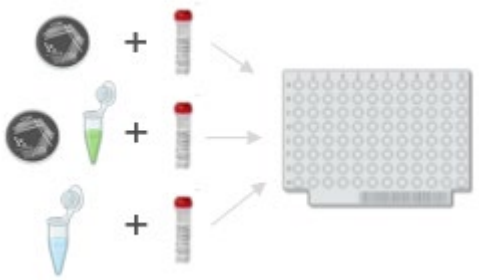
Task *Campylobacter*

Workflow di acquisizione spettri MALDI-TOF

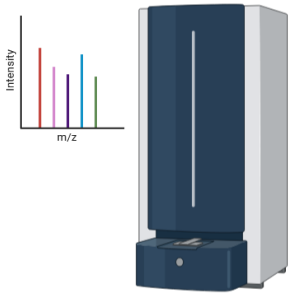
CC	N° ceppi	N° spettri
21	671	14215
353		
354		
45		
257		
Others		

Clonal complex selezionati

Coltura su *Campylobacter* CCDA Agar, a 42°C, in microaerofilia, per 48 ore



Allestimento piastra MALDI
- Spot diretto (colonia + 1 µL HCCA)
- Spot diretto con AF (colonia + 1 µL AF + 1 µL HCCA)
- Estratto proteico (1 µL estratto proteico + 1 µL HCCA)



Acquisizione spettri MALDI-TOF
- 27 spettri (3 spot x 3 modalità, acquisiti in triplicato)

Creazione dataset per addestramento di algoritmi di ML (cartella con raw data + file Excel con metadati)

PARTNER SCIENTIFICI COINVOLTI NELL'ACQUISIZIONE DEGLI SPETTRI E COSTRUZIONE DEL DATABASE

IZS TERAMO

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

IZS TERAMO

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

Task *Campylobacter*

Workflow di sviluppo dei modelli Random Forest e Neural Network

Costruzione del dataset | Preprocessing dei dati | Training dei modelli (RF/NN) | Valutazione del modello

	TOTAL		TRAINING		VALIDATION			
CC	#Spectra	#Strains	#Spectra	#Strains	#Spectra	#Strains	%Tr	%Val
CC21	4109	222	3206	180	903	42	81.08%	18.92%
CC257	692	37	530	29	162	8	78.38%	21.62%
CC353	2830	131	2340	106	490	25	80.92%	19.08%
CC354	948	42	798	36	150	6	85.71%	14.29%
CC45	1377	55	1038	42	339	13	76.36%	23.64%
CCOth	4259	184	3618	156	641	28	84.78%	15.22%
SUM	**14215**	**671**	**11530**	**549**	**2685**	**122**	81.82%	18.18%

IZS TERAMO

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

Task *Campylobacter*

Risultati preliminari e discussioni modelli RF/NN

Random Forest

		Classe reale					
		CC21	CC257	CC353	CC354	CC45	OtherCC
Classe predetta	CC21	824	19	14	0	6	0
	CC257	8	128	0	0	0	0
	CC353	6	0	429	0	0	0
	CC354	0	0	0	118	0	0
	CC45	0	0	4	0	252	4
	OtherCC	47	15	43	32	81	641

CLASSE	Accuratezza	Sensibilità	Specificità	Precisione	MCC
CC21	0.9635	0.9214	0.9848	0.9686	0.9179
CC257	0.9911	0.9074	0.9964	0.9423	0.9200
CC353	0.9825	0.9265	0.9950	0.9763	0.9406
CC354	0.9896	0.8667	0.9968	0.9420	0.8981
CC45	0.9758	0.8319	0.9966	0.9724	0.8865

Neural Network

		Classe reale					
		CC21	CC257	CC353	CC354	CC45	OtherCC
Classe predetta	CC21	832	10	4	0	2	11
	CC257	8	147	0	0	1	0
	CC353	11	0	454	0	0	0
	CC354	5	3	0	130	0	0
	CC45	0	1	3	0	282	4
	OtherCC	47	1	29	20	54	626

CLASSE	Accuratezza	Sensibilità	Specificità	Precisione	MCC
CC21	0.9628	0.9324	0.9781	0.9557	0.9162
CC257	0.9844	0.7901	0.9968	0.9412	0.8545
CC353	0.9750	0.8755	0.9973	0.9862	0.9149
CC354	0.9881	0.7867	1.0000	1.0000	0.8814
CC45	0.9661	0.7434	0.9983	0.9844	0.8388

- Entrambi i modelli mostrano elevate capacità discriminative, con accuratezza >96% e specificità prossima al 99% per tutti i principali CC.
- Random Forest risulta complessivamente più performante e bilanciato, con sensibilità e MCC superiori rispetto a Neural Network.
- Neural Network presenta maggiori falsi negativi per CC45 e CC257, mentre Random Forest ne migliora significativamente il riconoscimento.
- I CC di maggiore interesse epidemiologico non vengono quasi mai classificati come "Other CC« da entrambi i modelli, garantendo un basso rischio di falsi negativi nelle classi prioritarie.

IZS TERAMO

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

MALD-IA
Prospettive future

Ampliamento della library spettrale per il genere *Brucella*

Ottimizzazione del modulo di tipizzazione per *Campylobacter jejuni*

Sviluppo di piattaforme web interattive

MALD-IA e SPECTRA-VET: sinergia bidirezionale

IZS TERAMO

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELL'ABRUZZO E DEL MOLISE "G. CAPORALE"

Ministero della Salute

GRAZIE PER L'ATTENZIONE

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

IZS TERAMO SCIENZE STATISTICHE E GIS

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA