Alma Mater Studiorum - Università di Bologna

Dipartimento di Chimica Industriale “Toso Montanari”

Corso di Laurea Magistrale in

Chimica Industriale

Classe LM-71 - Scienze e Tecnologie della Chimica Industriale

Dal DNA al profilo aromatico: digital PCR e fingerprinting dei composti volatili per l‘autenticità del riso basmati

Tesi di laurea sperimentale

CANDIDATO

Simone Barbera

RELATORE

Dott.ssa Alessandra Stefan

CORRELATORE

Dott.ssa Sonia Scaramagli

Dott. Pietro Morozzi

Dott.ssa Raffaella Bergami

Anno Accademico 2025-2026

Abstract

Le frodi alimentari costituiscono una delle sfide più significative da contrastare nel settore alimentare, soprattutto nell'ultimo periodo caratterizzato da una crescente globalizzazione e complessità delle filiere produttive agroalimentari. Questa sfida è resa ancora più complessa dal fatto che non esiste un'unica definizione di frode alimentare codificata dall'Unione Europea; generalmente si definisce una frode alimentare come un'azione intenzionale volta alla sostituzione, aggiunta o falsa rappresentazione di alimenti, ingredienti, imballaggi alimentari e dichiarazioni false a scopo di lucro. I prodotti maggiormente frodati sono quelli caratterizzati da un elevato valore economico, come la sostituzione di olio di oliva con altre tipologie di olio o la vendita di riso non basmati al posto di riso basmati; queste frodi non portano solamente un vantaggio economico a chi le commette, ma possono anche causare conseguenze sulla salute dei consumatori.

Il presente studio è volto allo sviluppo di strumenti analitici affidabili da affiancare ai metodi tradizionali di controllo della qualità, in modo da rafforzare la tutela del consumatore. Sono state condotte analisi biologiche e chimiche per verificare l'autenticità del riso basmati. In particolare, è stata effettuata un'analisi del DNA con la tecnica digital PCR (dPCR), utilizzando la chimica EvaGreen, per determinare la quantità di riso non basmati in campioni commercializzati come basmati sfruttando una delezione presente in un gene caratteristico della varietà aromatica. Parallelamente, sono stati analizzati i composti volatili (VOCs) mediante gascromatografia con rivelatore a ionizzazione di fiamma (GC-FID) e gascromatografia accoppiata a spettrometria a mobilità ionica (GC-IMS) per evidenziare le differenze tra i profili dei composti volatili nelle varietà basmati e non basmati, verificando un potenziale accumulo della 2-acetil-1-pirrolina (2-AP) nei profili dei risi basmati. I risultati dell'analisi dPCR hanno dimostrato l'efficacia del metodo nella quantificazione della adulterazione nel riso, mentre l'analisi dei VOCs ha mostrato la presenza degli stessi composti nelle diverse varietà, evidenziando una predominanza della 2-acetil-1-pirrolina nella varietà aromatica. È stato inoltre costruito un modello in grado di predire la varietà dei campioni di riso puri, ma non la percentuale di riso non basmati presente all'interno delle miscele.

In conclusione, questo studio permette di affermare la validità della tecnica dPCR per la quantificazione dell'adulterazione del riso basmati, mentre l'analisi dei VOCs è stata utile per sviluppare un modello predittivo delle diverse varietà di riso. Obiettivo futuro, grazie allo sviluppo di modelli basati su dataset più ampi, è quello di poter utilizzare l'analisi dei VOCs anche per quantificare l'adulterazione alimentare.

INDICE

INTRODUZIONE 1

Settore alimentare e frodi alimentari 1

Verifica dell'autenticità alimentare e identificazioni frodi 5

Riso 12

SCOPO DELLA TESI 22

MATERIALI E METODI 24

Campioni analizzati 24

Analisi di biologia molecolare (dPCR) 25

Procedura analitica 35

Elaborazione dei dati 37

Analisi chimica untargeted 38

Gascromatografia 38

FlavourSpec25® 39

Heracles II 46

Chemiometria 49

RISULTATI E DISCUSSIONE 54

Risultati biologia molecolare 54

Risultati analisi chimica untargeted 64

Risultati FlavourSpec25® 64

Risultati Heracles II 73

CONCLUSIONI 81

BIBLIOGRAFIA 83

INTRODUZIONE

Settore alimentare e frodi alimentari

Il settore alimentare è una parte fondamentale della vita e dell'economia di tutti i Paesi del mondo, non solo per la sopravvivenza, ma anche per il ruolo sociale che ricopre il cibo in molte situazioni quotidiane. Infatti, il cibo può avere una influenza sociale indiretta, che include le tradizioni culinarie e una influenza sociale diretta, che corrisponde alle interazioni con gli altri esseri viventi durante i pasti (Rozin, 1996).

Si stima che la popolazione mondiale possa raggiungere i 10 miliardi nel 2056, con conseguenze impattanti sul mondo e sui suoi abitanti. Anche se la produzione globale di cibo ha dimostrato durante la storia di riuscire sempre a superare la crescita dei consumi, si stima che entro il 2050 il mondo dovrà affrontare un divario alimentare significativo, causato dalla crescita della popolazione e dagli effetti dei cambiamenti climatici. Il riscaldamento globale odierno, che ha portato ad un aumento della temperatura media della Terra di 1.5°C, contribuirà alla diminuzione delle rese globali delle principali colture. Le produzioni agroalimentare e zootecnica sono tra i maggiori contribuenti alle emissioni dei gas serra in atmosfera. Al contempo, nonostante la crescita della popolazione, la crescente scarsità di terre, acqua ed energia, i conflitti e i cambiamenti climatici, stiamo attualmente vivendo una rivoluzione tecnologica, con scoperte e progressi in molti ambiti come la nutrizione, l'informatica, la genetica e la meteorologia (Webb, 2016).

Nell'annuario statistico 2024, pubblicato dall'Organizzazione delle Nazione Unite per l'Alimentazione e l'Agricoltura (FAO), si possono verificare le tendenze più significative che riguardano il settore agroalimentare nel mondo, evidenziando le pressioni ambientali a cui è sottoposta la produzione agricola, soprattutto legate all'aumento della temperatura terrestre. I dati principali che emergono da questo annuario mostrano come il valore agricolo globale sia aumentato del'89%, incidendo sulle economie di molte Nazioni, accompagnato però da una diminuzione della forza lavoro impiegata in questo settore, a causa del miglioramento delle tecnologie.

Accanto a zone in cui il tasso di obesità è in aumento si affiancano zone nel mondo in cui la denutrizione persiste: in Europa, nelle Americhe e in Oceania più del 25% della popolazione adulta è obesa, dimostrando la non accessibilità garantita a cibo sano e nutriente, mentre nel 2023 circa 733 milioni di persone sono risultate denutrite, residenti principalmente in Africa e in Asia.

Per quanto riguarda la produzione alimentare si è verificato un aumento del 56% dal 2000 al 2022 della produzione delle colture primarie, che includono canna da zucchero, mais, grano e riso. Queste colture insieme costituiscono quasi la metà della produzione agricola globale. Parallelamente la produzione mondiale di carne è aumenta del 55% e nel 2022 la carne di pollo ha superato la carne di maiale come carne più venduta e costituisce attualmente la tipologia maggiormente prodotta a livello globale.

Anche l'uso di pesticidi è aumentato, con un incremento del 70% tra il 2000 e il 2022, soprattutto nelle Americhe. Nello stesso periodo l'utilizzo dei fertilizzanti inorganici è aumentato del 37%, tra cui il più utilizzato è l'azoto, che costituisce il 58% dei fertilizzanti utilizzati. La produzione degli oli vegetali è cresciuta del 133%, dato correlato soprattutto all'aumento della produzione dell'olio di palma, mentre sussistono problemi legati al consumo del territorio e alle emissioni di gas serra, queste ultime aumentate del 10%, con il settore dell'allevamento che contribuisce per il 54%. Un problema enorme per le zone del Medio Oriente e dell'Africa del nord risiede nelle gravi carenze idriche che Stati come l'Arabia Saudita, il Kuwait e gli Emirati Arabi Uniti devono affrontare: si stima infatti che i Paesi ricchi del Medio Oriente prelevino quantità di acqua dolce fino a 9-40 volte superiori rispetto alle risorse idriche disponibili (L'Annuario Statistico 2024 della FAO rivela dati cruciali sulla sostenibilità dell'agricoltura a livello globale, sulla sicurezza alimentare e sull'occupazione nei sistemi agroalimentari, 2024).

Gli esseri umani hanno iniziato a coltivare la terra 10.000 anni fa utilizzando bastoni e pietre, iniziando a produrre utensili in metallo 5.000 anni fa, durante l'età del bronzo. Con i progressi tecnologici, soprattutto nel settore della meccanica e dell'elettronica, sono stati sostituiti i macchinari semplici utilizzati per secoli durante la storia dell'essere umano, come aratri trebbiatrici, seminatrici e le irroratrici, anche se versioni di questi strumenti sono tuttora utilizzate in alcune parti del mondo.

L'industria agroalimentare sta attraversando un periodo caratterizzato da cambiamenti tecnologici finalizzati all'aumento dell'automazione dei processi e all'incremento della produttività. L'automazione dei processi riguarda il campionamento del suolo, la mappatura dei campi e delle rese, il monitoraggio delle colture, la raccolta e la gestione dei dati di tracciabilità. Oggigiorno le macchine agricole sono in grado di elaborare colture in situ e si possono avvalere di collegamenti con computer e GPS per

connettersi ai satelliti e ricevere indicazioni estremamente precise su dove piantare le diverse colture e dove applicare pesticidi e fertilizzanti.

I miglioranti tecnologici relativi al settore alimentare non si limitano solo all'ambito agricolo, zootecnico e alla raccolta, ma vengono applicati anche nella lavorazione post-raccolta, nella trasformazione alimentare e nel confezionamento. In particolare, il confezionamento rappresenta una fase fondamentale nella filiera alimentare in quanto permette di mantenere la qualità degli alimenti, dalla loro produzione fino alla consumazione da parte del cliente (Suprem, 2013).

Parallelamente allo sviluppo e all'aumento dell'impatto economico del settore agroalimentare, sono aumentate le frodi alimentari in tutto il mondo. Secondo una stima effettata dalla Grocery Manufacturers Association, le frodi alimentari genererebbero un mercato globale compreso tra i 10 e i 15 miliardi di euro all'anno, interessando quasi il 10% di tutti gli alimenti commercializzati (Manning, 2016).

La frode alimentare è una pratica esercitata già nell'antichità: sono stati rinvenuti infatti sigilli contraffatti appartenenti all'epoca romana e indicazioni di vendite fraudolente di vino e olio nell'Antica Grecia. Con il passare del tempo si è ampliata sia l'area geografica che la numerosità dei prodotti soggetti alle frodi. Oggi le frodi alimentari sono sotto la competenza normativa della Food Drug Administration (FDA) e si differenziano dagli incidenti di sicurezza alimentare: le frodi sono atti intenzionali compiuti a scopo di lucro, mentre gli incidenti sono eventi non intenzionali che possono causare danni accidentali. Le frodi alimentari generano prodotti alimentari adulterati che costituiscono spesso un problema per la salute dei consumatori. Questo tema viene molto spesso sottovalutato anche a causa di una scarsa comprensione delle dichiarazioni riguardanti i rischi per la salute pubblica, soprattutto legati all'adulterazione o all'etichettatura ingannevole. La definizione di frode alimentare si differenzia dalla definizione di adulterazione economicamente motivata (EMA), termine definito dall'FDA che indica una modifica fraudolenta dell'alimento; il termine frode alimentare è generalmente utilizzato in senso più ampio per indicare la sostituzione o l'aggiunta di ingredienti, la falsa composizione di alimenti e ingredienti, la falsificazione degli imballaggi alimentari e delle etichette a scopo di lucro.

Solo negli ultimi anni è diventato possibile, grazie ai progressi tecnologici e scientifici, rilevare frodi e falsificazioni presenti nella filiera alimentare, dalla produzione alla distribuzione. Per definire un prodotto come autentico si devono rispettare 5 punti:

- Le normative sull'etichettatura
- La composizione degli ingredienti
- I protocolli di produzione
- Le pratiche e i procedimenti tecnologici
- L'identità genetica

Il problema dell'etichettatura errata è diventato rilevante in seguito allo scandalo del latte cinese nel 2008, durante il quale alcuni lotti di latte e prodotti per l'alimentazione infantile contenevano melamina, una sostanza tossica, per far apparire il latte più ricco di proteine e per mascherarne la cattiva qualità. Un altro scandalo è stato quello relativo alla carne di cavallo in Europa nel 2013, durante il quale in molti paesi europei è stata commercializzata, fino al 100%, carne di cavallo al posto di carne di manzo, dimostrando come il fenomeno delle frodi possa intaccare anche grandi scale di produzione (Huck, 2016).

Nel 2009 la FDA ha tenuto una riunione pubblica con lo scopo di migliorare la capacità di prevedere e prevenire l'EMA. Al fine di ridurre i rischi per la salute pubblica sono stati presi in considerazione tutti i prodotti regolamentati dalla FDA come alimenti, farmaci, mangimi, dispositivi medici e prodotti biologici. Questa riunione ha rappresentato un passo avanti notevole verso la riduzione delle frodi, riconoscendo ufficialmente il rischio emergente associato all'EMA in tutti i prodotti trattati dalla FDA. Sono stati identificati tre tipi di rischio connessi alle frodi alimentari per la salute pubblica: rischio diretto, indiretto e tecnico. Il rischio diretto si configura quando il consumatore è esposto a un rischio immediato come nel caso della presenza di un contaminante tossico nell'alimento. Il rischio indiretto incorre quando il consumatore è esposto al contaminante in maniera cronica per lungo tempo; questo rischio è legato soprattutto all'omissione dagli alimenti di vitamine o ingredienti benefici. Il rischio di frode alimentare tecnico è di natura immateriale e riguarda, per esempio, la falsificazione della documentazione relativa al contenuto di un prodotto o al suo luogo di origine.

Insieme al concetto di frode alimentare si è sviluppato il concetto di protezione alimentare, che include al suo interno la sicurezza alimentare, la qualità alimentare, la difesa alimentare e le frodi alimentari. I concetti di protezione alimentare sono stati incorporati nella Matrice del Rischio Alimentare (**Figura 1**), con lo scopo di comprendere meglio come la frode differisca dagli altri problemi alimentari, come ad

esempio la sicurezza. La Matrice del Rischio alimentare affronta principalmente la motivazione e la causa della truffa, ma non descrive l'effetto pratico (Spink, 2011).

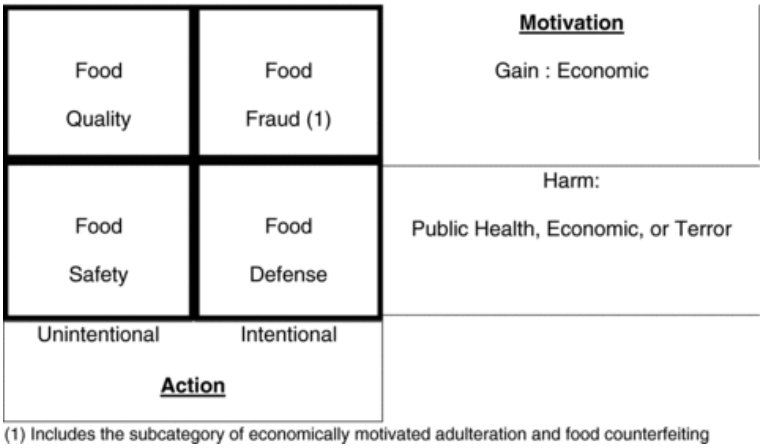


Figura 1: Matrice del Rischio Alimentare, nei 4 quadranti si trovano le 4 discipline, sull'asse x si trova l'intenzionalità e sull'asse y sono riportate le motivazioni e i rischi *(Spink, 2011)*.

La lotta contro le frodi alimentari è ostacolata dalla mancanza di una definizione univoca e dall'ambiguità del termine. Infatti, l'Unione Europea non ha ancora definito legalmente il concetto di frodi alimentari e all'interno dei documenti che contengono le strategie di prevenzione e mitigazione le definizioni di frode differiscono. Poiché l'industria alimentare è una delle poche industrie ad essere allo stesso tempo vittima e responsabile degli incidenti di frode alimentare, sarebbe necessario aumentare la chiarezza e la comprensione delle frodi. Nonostante la definizione di frode non sia assoluta, si definisce autenticità alimentare come la corrispondenza tra le caratteristiche del prodotto alimentare e le corrispondenti dichiarazioni riportate sul prodotto (CEN, 2019). L'integrità alimentare è stata invece definita dal Progetto UE Food Integrity del 2017 come la caratteristica di un alimento di essere integro, completo e in perfette condizioni (Robson, 2021).

Verifica dell'autenticità alimentare e identificazioni frodi

La globalizzazione, la crescita della popolazione e la sempre maggiore richiesta di alimenti stanno causando un aumento della complessità delle filiere alimentari, con conseguente aumento della produzione di cibo e relativo aumento delle frodi alimentari. Le frodi più comuni consistono nel contraffare o sostituire specie animali o vegetali con altre meno pregiate come, per esempio, la miscelazione di riso venduto

come basmati con riso non basmati o la vendita di una determinata specie di pesce al posto di un'altra più pregiata.

L'autenticazione alimentare è un settore in rapida crescita anche a causa della crescente consapevolezza dei consumatori riguardo alla qualità e alla sicurezza alimentare. I consumatori sono sempre più interessati a conoscere l'origine geografica e la qualità dei prodotti che bevono e mangiano come conseguenza della diversificazione e dello sviluppo dei mercati alimentari. L'autenticazione degli alimenti è il processo di verifica della conformità di un prodotto alimentare in base alle informazioni riportate sulla sua etichetta; queste informazioni possono includere l'origine, il metodo di produzione o le tecnologie di lavorazione. L'assicurazione della qualità e i metodi utilizzati per autenticare i prodotti ricoprono un ruolo di grande interesse dal punto di vista legale e commerciale. La strumentazione moderna permette di effettuare misurazioni precise e di determinare in alcuni casi l'origine degli alimenti. A partire dal XX secolo sono nate in tutto il mondo organizzazioni che stabiliscono degli standard e che sono addette al controllo dell'origine degli ingredienti e del processo di produzione, come il francese "Institut National des Appellations d'Origine" (INAO), la "Denominazione di Origine Controllata" in Italia, la "Denominación de Origen" in Spagna, "Wine of Origin" in Sud Africa o le "American Viticultural Areas" negli Stati Uniti. Seguendo le procedure indicate da queste organizzazioni si ottengono prodotti caratterizzati da elevata qualità che vengono commercializzati con prezzi più alti all'interno dei punti vendita, portando a maggiori benefici finanziari. Di contro, questi vantaggi stimolano la produzione di alimenti contraffatti e il commercio illegale. In Europa il principale problema legato all'autenticità è quello relativo all'origine degli ingredienti; per questo motivo la legislazione e le certificazioni sviluppate dall'Unione Europea per definire cibi e bevande di elevata qualità è estremamente abbondante. Un prodotto alimentare può essere etichettato come DOP (Denominazione di Origine Protetta), ovvero con la garanzia che i prodotti derivino da una determinata area geografica, oppure come IGP (Indicazione Geografica Protetta), che collega i prodotti ad un'area geografica in cui si è verificata almeno una delle fasi di produzione o di trasformazione. Esistono anche recenti normative che attribuiscono speciali qualità ai prodotti provenienti da aeree di montagna e insulari (Danezis, 2016).

Oltre che per i consumatori, l'autenticazione alimentare è motivo di interesse anche per i trasformatori alimentari che vogliono tutelarsi dalla concorrenza sleale. I sistemi europei per lo sviluppo e la protezione dei prodotti alimentari servono per favorire la

diversificazione della produzione agricola, per garantire maggiore protezione dall'uso improprio dei nomi dei prodotti, dalle imitazioni e per tutelare i consumatori fornendo le giuste informazioni (Reid, 2006).

Sono state sviluppate numerose tecniche analitiche utilizzabili per l'analisi dell'autenticità alimentare; al momento la maggior parte dei metodi utilizzati è di tipo mirato (*targeted*), ovvero metodi basati su analisi mirate alla ricerca di un composto adulterante, che può causare una modifica in una caratteristica del prodotto, limitando quindi l'analisi a pochi analiti. Negli ultimi anni sono stati sviluppati anche i metodi non mirati (*untargeted)*, caratterizzati da analisi non indirizzate a rilevare molecole note a priori, ma facendo un'impronta globale del campione, il più completa possibile. I metodi untargeted vengono utilizzati quando si vogliono trovare adulteranti non convenzionali o quando non sono note informazioni preliminari sui composti adulteranti. Questi metodi, a differenza dei metodi mirati, non richiedono inoltre procedure di estrazione complesse e specifiche per ogni analita, ma necessitano di una semplice preparazione dei campioni (Amaral, 2001).

Nel corso del tempo sono state sviluppate diverse tecniche analitiche per la verifica dell'origine degli alimenti. L'analisi molecolare, basata sullo studio dei geni e delle proteine, rappresenta una delle principali metodologie di autenticazione e, sebbene i metodi tradizionali siano ampiamente utilizzati per l'autenticazione degli alimenti, le tecniche genomiche e proteomiche stanno rapidamente affiancando o sostituendo tali approcci. I metodi basati sull'analisi degli acidi nucleici e delle proteine sono impiegati principalmente per l'identificazione delle specie, dato che il DNA è identico in tutte le cellule somatiche di un organismo e rimane invariato indipendentemente dal tessuto da cui viene estratto. L'elevata stabilità del DNA consente di analizzare prodotti alimentari sottoposti a trattamenti intensivi, nonché di rilevare contaminanti presenti in tracce. I metodi di autenticazione alimentare basati sul DNA (genomica) utilizzano tutto il genoma del campione e si basano sull'amplificazione specifica di frammenti di DNA tramite la reazione a catena della polimerasi (PCR), consentendo di amplificare tracce di materiale nucleotidico. In alternativa si può utilizzare l'analisi proteomica, che consiste nell'utilizzo di proteine come marcatori lungo l'intera filiera, consentendo di identificare prodotti specifici sulla base di proteine espresse e codificate dal DNA. Queste tecniche vengono impiegate per individuare informazioni errate sulle etichette di prodotti come gli OGM, i prodotti ittici e la selvaggina, per verificare l'autenticità di carni halal e kosher, per l'individuazione di carne equina o suina in prodotti etichettati

come carne bovina e per la determinazione dell'origine botanica degli alimenti (olio d'oliva, vino, cacao, riso).

Un altro tipo di tecnica che si può utilizzare per l'identificazione delle frodi è l'analisi cromatografica, che consente una separazione rapida e affidabile di diversi composti all'interno di matrici alimentari complesse che contengono lipidi, peptidi, carboidrati, amminoacidi, acidi grassi, organici e nucleici, coloranti, conservanti ecc. Tutti questi composti presentano dimensioni e caratteristiche chimiche (come la polarità e la carica) molto diverse tra loro. I *fingerprinting* chimici ottenuti con i metodi cromatografici possono essere poi utilizzati per differenziare e autenticare gli alimenti, rilevando la presenza di marker specifici o di differenze analitiche tra vari campioni, ottenendo profili caratteristici dei composti presenti in ciascun prodotto alimentare di cui costituiscono l'identità e l'origine. Vengono maggiormente usate tecniche cromatografiche ad alta risoluzione (gascromatografia o cromatografia liquida) accoppiate alla spettrometria di massa. La separazione mediante cromatografia liquida si basa su tre caratteristiche fondamentali dei composti chimici, polarità, carica elettrica e dimensione molecolare, e viene impiegata per l'analisi di proteine, amminoacidi, carboidrati, vitamine, composti fenolici. La gascromatografia invece risulta più indicata per l'analisi di molecole volatili o semivolatili. Le tecniche cromatografiche vengono soprattutto utilizzate per verificare l'autenticità di alimenti di alta qualità e valore commerciale come il miele, il vino, l'olio, lo zafferano, il latte, il riso, i funghi, ecc, che vengono spesso adulterati mediante aggiunta di ingredienti meno pregiati. L'autenticazione viene solitamente effettuata confrontando i profili cromatografici dei composti presenti nel campione in esame con i valori di riferimento predeterminati (standard).

Tutte le tecniche di *fingerprinting* generano un elevato volume di informazioni e gli strumenti di chemiometria sono fondamentali per la gestione di questi grandi dataset generati con le diverse tecniche analitiche. L'utilizzo di questi strumenti chemiometrici nel campo dell'autenticazione alimentare è ancora agli inizi, dimostrando mancanza di linee guida standardizzate per la descrizione della progettazione sperimentale e per la presentazione dei risultati nella letteratura scientifica; questo limita molto la possibilità di confrontare studi analoghi. Un altro limite consiste nel fatto che la maggior parte degli studi pubblicati non dispone di un numero di campioni sufficiente per condurre analisi significative e costruire modelli statistici robusti. Fortunatamente negli ultimi anni la situazione sta migliorando considerevolmente.

I risultati ottenuti con le analisi chemiometriche, nella maggior parte degli studi, si limitano a dimostrare la possibilità di discriminare tra diverse classi di dati distinguendo, ad esempio, campioni biologici da non biologici. Non esistendo ancora molte banche dati contenenti materiali di riferimento, vengono spesso condotti studi di carattere esplorativo analizzando dati preliminari per dimostrare che tali dati consentono di distinguere tra diverse classi. L'autenticazione accurata degli alimenti comporta aspetti tecnici, metodologici e normativi; i risultati devono pertanto essere: accurati, utilizzando materiali di riferimento standard; robusti, per gestire la variabilità della composizione del campione e i processi di lavorazione e devono essere confrontabili tra studi diversi. È fondamentale la creazione di database standardizzati che riportino informazioni dettagliate sull'origine e le caratteristiche dei prodotti, incluse informazioni sull'origine geografica, la specie o sottospecie di appartenenza e i metodi di produzione. Con il dataset ottenuto dall'analisi di un numero elevato di campioni si possono creare modelli predittivi che siano in grado di associare un campione incognito a una classe precedentemente definita. Con lo sviluppo rapido delle tecniche chemiometriche si avranno a disposizione strumenti molto potenti con cui identificare e contrastare le frodi alimentari. Tali strumenti dovranno essere integrati con infrastrutture di dataset standardizzati e strumenti matematici e informatici idonei all'autenticazione degli alimenti.

Oltre alle tecniche molecolari, alle tecniche cromatografiche e alle tecniche chemiometriche ci sono altri tipi di analisi che si possono eseguire nell'ambito della verifica dell'autenticità alimentare:

- L'analisi degli isotopi, che vengono suddivisi in base alla massa atomica in isotopi leggeri ($^{2}H/^{1}H$, $^{13}C/^{12}C$…) e isotopi pesanti ($^{87}Sr/^{86}Sr$). I rapporti isotopici degli elementi stabili variano in funzione delle condizioni climatiche, dell'origine geografica, delle caratteristiche geologiche e vengono utilizzati per l'analisi di vino, miele, succhi di frutta o sciroppo d'acero o come metodo per la distinzione tra vanillina naturale e sintetica. Queste analisi si possono anche combinare con altre tecniche analitiche come la analisi NMR e le tecniche gascromatografiche.
- La spettroscopia vibrazionale, ad esempio nel vicino infrarosso (NIR) e medio infrarosso (MIR), e la spettroscopia a fluorescenza rappresentano metodi rapidi ed economici e solitamente vengono accompagnate ai metodi chemiometrici e sottoposte ad analisi multivariata. Una metodologia emergente consiste nell'utilizzare la spettroscopia Raman, che coinvolge gruppi funzionali specifici di

alcune molecole, consentendo una identificazione qualitativa dei campioni. Queste tecniche consentono un'analisi rapida e non distruttiva del campione in esame e utilizzano le trasformate di Fourier per fornire i dati. La spettroscopia di fluorescenza, invece, presenta limiti di rilevabilità inferiori rispetto alle altre tecniche spettroscopiche e permette di rilevare idrocarburi poliaromatici e composti eterociclici dotati di scheletri molecolari rigidi. Nel complesso le diverse metodologie vengono utilizzate per la rilevazione di adulterazioni in latte e farina di soia (eventuale presenza di melammina), nel miele per identificare l'aggiunta di oli vegetali nell' olio extravergine d'oliva, per identificare l'adulterazione di spezie come il pepe nero macinato, per l'adulterazione della carne e in numerose altre applicazioni.

- l'analisi del profilo elementale di alimenti consente di determinare macroelementi come il sodio, il calcio e il potassio, oligoelementi come rame, zinco, e altri elementi presenti in tracce, tra cui metalli, iridio e oro, o terre rare (lantanio). Questi elementi sono presenti nel suolo, quindi sono fondamentali per il nutrimento delle piante e la loro concentrazione dipende dal tipo di terreno, dall'annata di produzione e dall'origine botanica della coltura. Gli elementi più caratterizzanti per determinare l'impronta caratteristica utile per definire l'origine geografica di un prodotto sono le terre rare, poiché la loro concentrazione è legata alla geologia del territorio e alle pratiche agricole effettuate durante la coltivazione. Anche la composizione elementale degli alimenti di origine animale riflette la presenza di minerali contenuti nel foraggio, nella vegetazione ingerita dagli animali, nell'acqua e in altri fattori legati all'allevamento. Le tecniche di analisi elementali vengono quindi utilizzate per determinare l'autenticità di numerosi prodotti come il vino, il miele, l'olio di oliva, le spezie, frutta e verdura.
- La risonanza magnetica nucleare (NMR) consente di quantificare un determinato metabolita anche con una preparazione del campione minima, ottenendo profili metabolici completi con una elevata sensibilità e altamente riproducibili. L'analisi NMR è stata utilizzata per individuare adulterazioni in numerosi prodotti alimentari come pesce, bevande spiritose, aceto e zafferano; un esempio di applicazione riguarda l'adulterazione del vino rosso dovuta all'aggiunta di antociani.
- le analisi sensoriali sono fondamentali per l'analisi delle proprietà organolettiche di un alimento e vengono effettuate da un gruppo (panel) di esaminatori appositamente selezionati e formati. Questi valutatori eseguono dei test

organolettici basati su caratteristiche come aspetto, aroma, sapore e consistenza del prodotto. Segue l'analisi statistica delle osservazioni, poiché il risultato delle analisi può variare sia tra persone diverse all'interno del panel, sia per lo stesso individuo in tempi diversi a causa della stanchezza, stress e altre variabili. Anche se i valutatori vengono addestrati adeguatamente, sono state sviluppate tecniche strumentali che riproducono gli stimoli sensoriali in modo da standardizzare l'analisi e fornire misure oggettive. I dati ottenuti dall'analisi con i panel sensoriali possono essere confrontati, ad esempio, con i dati ottenuti utilizzando strumenti come i nasi elettronici, caratterizzati da semplicità e rapidità di analisi. Le analisi sensoriali vengono applicate soprattutto per alimenti come vino, formaggio, birra e olio di oliva.

- la spettrometria di massa non accoppiata a tecniche cromatografiche comprende tecniche come la spettrometria di massa a reazione di trasferimento protonico (PTR-MS), che consente il monitoraggio quantitativo in tempo reale dei composti volatili organici (VOCs). Questa tecnica sfrutta una ionizzazione soft dei VOCs da parte di ioni idorossonio (H_3O^+), con la formazione di ioni protonati che successivamente vengono accelerati da un campo elettrico e separati. Il profilo ottenuto può essere utilizzato per costruire il *fingerprinting* utile per il riconoscimento dei campioni. Queste tecniche vengono utilizzate per prodotti come zafferano, caffè, mangimi per animali, succhi di frutta e tartufo e per verificare l'origine botanica delle specie vegetali come le spezie. Vengono utilizzate anche per rilevare la diluizione di olio di oliva con oli più economici e di latte ovino e caprino con quello vaccino.
- le tecniche immunologiche usano strumenti analitici chiamati "immunoassays", che sfruttano l'interazione specifica tra gli anticorpi e i loro antigeni. Questi strumenti consentono analisi rapide, sensibili e facili da eseguire. La tecnica immunologica più importante è l'ELISA (Enzyme-Linked Immunosorbent Assay), un test enzimatico a base di anticorpi coniugati con enzimi e viene impiegata per verificare l'autenticità di latticini, pesce, carne, per individuare la presenza di carne suina all'interno di carne bovina e per rilevare la presenza di organismi geneticamente modificati (OGM) (Danezis, 2016).

Riso

Da millenni uno degli alimenti più consumato dalla popolazione mondiale è un cereale, il riso (*Oryza sativa* L), grazie a caratteristiche uniche riguardanti la sua economicità e la sua facilità nell'essere coltivato in condizioni climatiche e geografiche diverse. Il riso è un alimento molto saziante ed è accessibile a persone appartenenti a diverse classi sociali, dalle più povere a quelle più ricche. Attualmente è l'alimento di base di metà popolazione mondiale e la sua coltivazione è molto antica. Il suo utilizzo, dagli inizi ad oggi, è sempre stato versatile; infatti, durante la storia ha svolto anche un ruolo culturale nelle cerimonie religiose, oltre essere stato utilizzato in ambiti non alimentari come nella produzione di cosmetici, medicinali e carta (World's Most Eaten Food, 2022).

A livello mondiale gli ettari di terreno dedicati alla coltivazione del riso sono circa 165 milioni, di cui l'89% si trova in Asia, dove rappresenta l'alimento base della dieta giornaliera. Complessivamente circa il 60% della popolazione mondiale si nutre regolarmente di riso, coltivato in tutti i continenti tranne che in Antartide, in ambienti con differenti regimi idrici e diverse tipologie di terreno. In Europa la superficie destinata alla coltivazione del riso ammonta a 650.000 ettari, meno dell'1% della superficie mondiale, soprattutto a causa del clima europeo che consente la crescita solo di alcune varietà di riso (molte prediligono un clima tropicale). L'Italia è il primo paese produttore di riso dell'UE, coprendo da solo il 46% della produzione europea con 240.000 ettari di terre coltivate (Riso nel mondo: una coltivazione globale di un chicco universale, s.d.).

Si stima che il riso costituisca l'alimento base per circa 3,5 miliardi di persone nel mondo, la maggior parte situata in Asia, regione caratterizzata dal fenomeno della povertà e della denutrizione cronica dove la coltivazione del riso è associata a tematiche di sicurezza alimentare e di instabilità politica. In questa regione il riso fornisce infatti fino al 50% dell'apporto calorico e una parte significativa dell'apporto proteico a circa 520 milioni di persone che vivono in condizioni di povertà. Anche nell'Africa subsahariana, nell'America Latina e nei Paesi caraibici la consumazione di riso è cresciuta costantemente nel tempo. Nei Paesi in via di sviluppo la filiera che riguarda la produzione di riso è la principale fonte di reddito e occupazione per oltre 200 milioni di persone.

A metà del '900 si è registrato un incremento significativo della produzione di riso, causando una vera e propria Rivoluzione Verde, che ha incrementato notevolmente la

produzione agraria nei Paesi in via di sviluppo. Questo fenomeno si è verificato soprattutto a causa dell'applicazione di fertilizzanti azotati, di fitofarmaci, di impianti di irrigazione e di nuovi investimenti di capitale alla produzione agricola, oltre all'impiego di varietà vegetali geneticamente selezionate. Il grafico in **Figura 2** mostra come la produzione di cereali e la resa di produzione abbiano avuto un aumento pari a quello della popolazione, mentre la superficie dei terreni utilizzata alla coltivazione sia rimasta pressoché costante. La popolazione dei Paesi in via di sviluppo è aumentata del 90% dal 1966 al 2000, mentre la produzione di riso grezzo è aumentata del 130%. In molti Paesi asiatici, la resa nella coltivazione e produzione del riso sono raddoppiati e in alcuni casi anche triplicati rispetto alla media precedente alla Rivoluzione Verde. Si prevede che la domanda di riso aumenterà del 70% nei prossimi 30 anni, soprattutto a causa dell'aumento demografico.

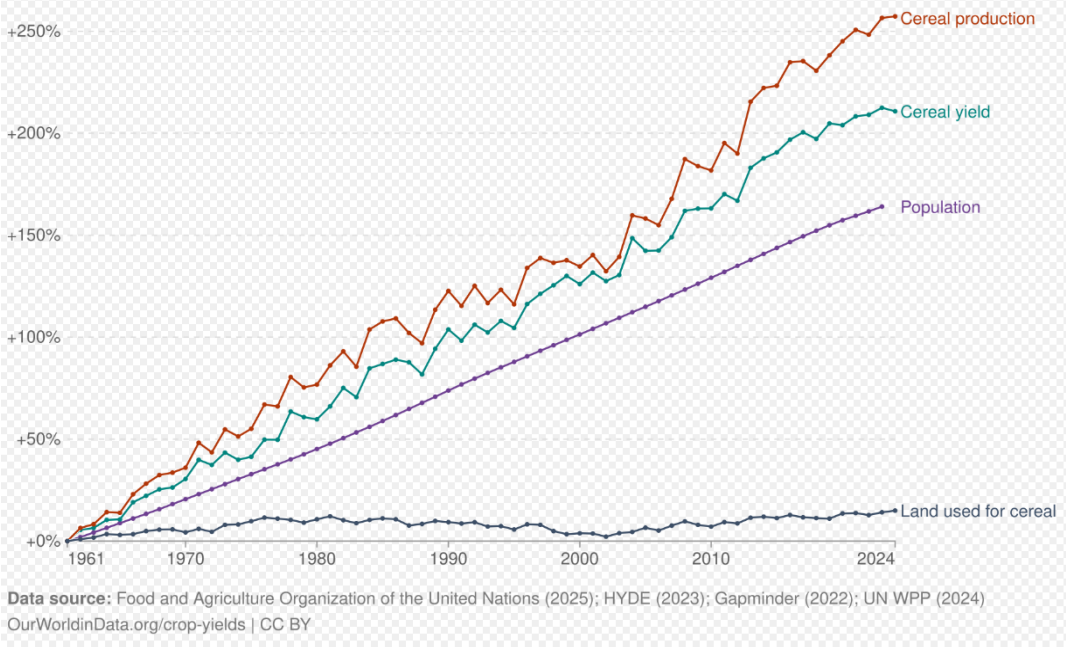


Figura 2: grafico che riporta sull'asse delle ordinate la crescita percentuale e sull'asse delle ascisse l'anno di riferimento. Si nota che in seguito alla Rivoluzione Verde la produzione e la resa di produzione dei cereali sia riuscita ad aumentare seguendo l'aumento demografico durante gli anni, senza aver interessato un aumento nell'uso dei terreni agricoli destinati ai cereali *(Rivoluzione verde, s.d.)*.

Attualmente il riso viene coltivato in oltre cento Paesi nel mondo, con una produzione annua superiore a 715 milioni di tonnellate di riso. Quindici paesi sono responsabili del 90% del raccolto mondiale di riso, solo la Cina e l'India rappresentano circa il 50% della produzione complessiva. Insieme a Indonesia, Bangladesh, Vietnam, Myanmar,

Thailandia, Filippine, Giappone, Pakistan, Cambogia, Repubblica di Corea, Nepal e Sri Lanka, i paesi asiatici coprono il 90% della produzione mondiale totale di riso. Tra i principali paesi produttori non asiatici figurano Brasile, Stati Uniti, Egitto, Madagascar e Nigeria, che nel loro complesso rappresentano il 5% della produzione globale. In Africa, il riso è l'alimento base che registra la crescita più rapida in quanto la produzione totale di cereali in Africa è aumentata costantemente, passando dal 9,3% del 1961 al 15,2% del 2007. Nonostante questo incremento, solo il 54% del consumo di riso è coperto dalla produzione locale.

Il riso (genere *Oryza*) è una pianta erbacea annuale semi acquatica caratterizzata da un fusto non duro e non legnoso che vive bene con le radici immerse nell'acqua, grazie alla presenza di vasi aeriferi che catturano l'ossigeno disciolto, evitando la marcitura delle radici. Il genere *Oryza* comprende circa 22 specie di cui 20 selvatiche, le 2 specie più importanti per il consumo umano sono *Oryza sativa* e *Oryza glaberrima*. La specie *Oryza sativa* venne coltivata inizialmente nel Sudest Asiatico circa 10.000 anni fa, mentre la specie *Oryza glaberrima* deriva dalle paludi alluvionali del fiume Niger e risale a circa 3000 anni fa. Attualmente la specie più coltivata a livello mondiale è l'*O. sativa*, che costituisce circa il 95% della produzione di riso globale, mentre la coltivazione di *O. glaberrima* è limitata solo ad alcune zone dell'Africa, dove però sta venendo lentamente sostituita dall'*O. sativa*.

La specie *O. sativa* è divisa in 3 sottospecie, coltivate in più di 100 Paesi:

- *Oryza sativa* subsp. Indica, a chicco (cariosside) lungo, coltivato soprattutto nell'Asia tropicale e subtropicale, zone caratterizzate da elevate temperature;
- *Oryza sativa* subsp. Japonica, a chicco corto o di lunghezza intermedia, coltivato in regioni con climi temperati come il Giappone o alcune zone della Cina settentrionale;
- *Oryza sativa* subsp. Javanica, coltivato nelle Filippine e nelle zone montuose del Madagascar e dell'Indonesia. È una sottospecie adattata ad ambienti umidi, con una diffusione molto limitata nel mercato globale. Esistono metodi principali per la coltivazione del riso: i sistemi di pianura irrigati (*irrigated)*, delimitati da argini, che possono coprire i 3/4 della produzione mondiale di riso; i sistemi di pianura alimentati ad acqua (*rain-fed)*, che sfruttano l'acqua piovana, generano il 20% del riso mondiale e vengono utilizzati nell'Asia meridionale, Sud-est Asiatico e in Africa; i sistemi di coltivazione del riso in altura (*upland*) che utilizzano terreni asciutti privi di irrigazione e che generano solo il 4% della produzione mondiale di riso.

(Muthayya, 2014). Essendo l'alimento più consumato nel mondo con un mercato internazionale pari a 372,30 miliardi di dollari, il riso è tra gli alimenti maggiormente a rischio di adulterazione e frode (Dimensioni del mercato del riso e fattori di crescita, 2026).

Il riso è costituito da un chicco (cariosside) racchiuso in un rivestimento esterno resistente (lolla) sotto al quale si trovano la crusca e il germe. Rimuovendo le due parti esterne si trova l'endosperma, che corrisponde al riso bianco, ovvero la tipologia di riso più consumata nel mondo, che viene generalmente arricchito di vitamine e minerali essenziali durante la lavorazione. La differenza tra il riso integrale e il riso bianco è che quello integrale è completo, ovvero possiede crusca, germe ed endosperma, di conseguenza possiede caratteristiche organolettiche e nutrizionali differenti rispetto al bianco, oltre ad avere una shelf-life più breve e tempi di cottura maggiori rispetto al riso bianco. In **Figura 3** è schematizzata la struttura di un chicco di riso con le sue componenti principali.

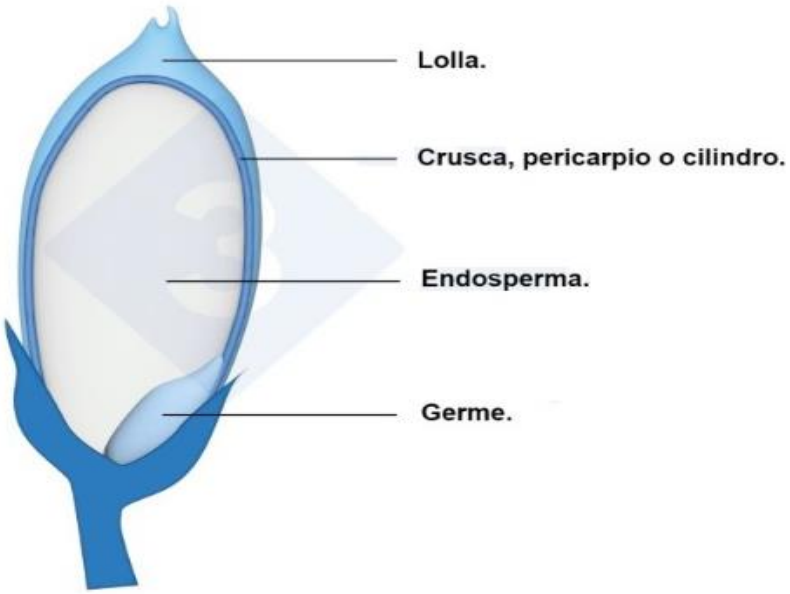


Figura 3: componenti principali di un chicco di riso: il germe costituisce la parte ventrale del chicco mentre la lolla, la crusca e l'endosperma la parte dorsale. La lolla costituisce il 20-22% del peso totale del chicco appena raccolto ed è costituita da lignina e da SiO_2, è povera di nutrienti ed ha funzione protettiva, la crusca costituisce il 2% del riso e l'endosperma è quella che rimane nel riso bianco dopo lavorazione e costituisce il 73-75% del chicco di riso *(Il riso come ingrediente nei mangimi per suini - Articoli, 2022)*.

Le diverse tipologie di riso possono essere distinte sulla base di quattro caratteristiche principali:

- grado di lavorazione: in base al livello di lavorazione di può distinguere tra riso integrale, semi-integrale e bianco. Il riso meno lavorato è quello integrale che conserva tutte le componenti ad eccezione della lolla; poi c'è il riso semi-integrale, in cui viene rimossa solo la lolla esterna e una parte della pula (rivestimento sotto la lolla che contiene fibre, vitamine B, minerali), e infine il riso bianco, che contiene solo l'endosperma (Lavorazione riso: le tipologie dall'integrale al raffinato, 2024)
- dimensione e forma del chicco: si distinguono il riso a cariosside corta, media e lunga, caratterizzati da diversa consistenza dopo cottura. Tra i risi a chicco lungo i più rinomati sono il riso basmati e il riso jasmine, mentre le varietà di riso a chicco medio e corto più consumate sono il riso carnaroli, il riso arborio e il vialone nano.
- contenuto e composizione di amido: le diverse varietà differiscono per la composizione dell'amido, costituito da diversi rapporti di amilosio e amilopectina. Tra le tipologie di riso maggiormente ricche di amido, aspetto che influenza la consistenza dopo la cottura, ci sono le varietà a chicco corto e medio.
- aroma/sapore: alcune varietà di riso presentano caratteristiche aromatiche marcate e tipiche. È noto, ad esempio, che le tipologie di riso basmati e jasmine abbiamo un sapore e un aroma diverso e più accentuato rispetto alle altre tipologie di riso.

La caratteristica che rende alcune tipologie di riso (basmati e jasmine) più aromatiche rispetto ad altre risiede nella biologia/genetica della pianta. La tipologia di riso aromatica più importante, coltivata in India e Pakistan, è il riso basmati, caratterizzato da chicchi a grano lungo, sottili e dall'aroma gradevole, di consistenza morbida e con la capacità di allungarsi senza gonfiarsi eccessivamente in cottura. L'aroma caratteristico del riso basmati è dovuto alla presenza di un composto volatile chiamato 2-acetil-1-pirrolina (2-AP). L'accumulo di tale molecola nelle varietà di riso aromatico è causato da una mutazione genetica nel gene BADH2, detto anche *fgr* (da fragranza), che provoca la perdita della funzione dell'enzima prodotto. Il gene è localizzato sul cromosoma 8 del genoma del riso e codifica per la betaina-aldeide deidrogenasi (BADH), enzima che serve a convertire la betaina aldeide a betaina, molecola coinvolta nel mantenere alta la tolleranza allo stress osmotico nelle piante:

$$\text{betaina aldeide} + NAD^+ + H_2O \rightleftarrows \text{betaina} + NADH + 2\ H^+$$

Nel riso comune il gene BADH2 è funzionante mentre nei genotipi aromatici possiede una mutazione che introduce un codone di stop a monte di una sequenza presente in altre BAD. Il risultato di tale mutazione comporta la perdita di funzione dell'enzima

BADH2, che viene prodotto come proteina tronca e non funzionante (enzima inattivo); l'assenza di attività enzimatica determina, nel riso basmati, un accumulo di 2-AP ottenuto mediante una via metabolica alternativa.

Per l'analisi dell'autenticità alimentare del riso basmati è possibile sfruttare la mutazione nel gene BADH2 in quanto nelle varietà aromatiche è presente una delezione di 8 bp che provoca lo shift della lettura e l'introduzione di un codone di stop nella sequenza genica. Per identificare la presenza della delezione si utilizza la PCR con due coppie differenti di primers: una coppia di primers (KMV) specifici per il riso che amplificano un tratto di 68 paia di basi (bp) del gene della fosfolipasi presente in tutte le varietà di riso (gene di riferimento o *housekeeping*), e una coppia di primers (NBR) specifici per il riso non basmati che amplificano un tratto di 80 bp del gene della betaina aldeide deidrogenasi (BADH2). Come descritto in precedenza è presente una delezione di 8 bp nel gene del riso aromatico (sequenza GATTATGG, **Figura 4**) che impedisce il corretto appaiamento dei primers. Questa delezione è responsabile della formazione della specifica fragranza nel riso aromatico *(flavoured*). Poiché il sito di legame del primer contiene sette delle otto basi eliminate nel riso basmati, si otterrà un'amplificazione solo se il riso analizzato è non basmati (Bucher, 2016). In tal modo la presenza o assenza del prodotto della PCR permette di discriminare le varietà basmati (aromatiche) da quelle non basmati (non aromatiche).

In **Tabella 1** vengono riportati i primers e le rispettive sequenze genetiche con gli ampliconi ottenuti (un amplicone è un frammento di DNA o RNA che viene prodotto in una reazione di amplificazione in PCR). La composizione chimica del riso varia enormemente in base alla varietà e alle caratteristiche dell'ambiente e del tipo di suolo dove viene coltivato; inoltre, il contenuto di amido, lipidi e proteine possono variare a seconda del tipo di tecnica analitica impiegata per la loro determinazione. In **Tabella 2** si riportano le sostanze maggiormente presenti nelle diverse tipologie di riso.

Tabella 1: sequenze e ampliconi dei primers utilizzati per il riconoscimento genetico delle varietà di riso *(Bucher, 2016)*.

Primers	Sequenza nucleotidica	Amplicone
NBR - specific primers (betaine aldehyde dehydrogenase gene)		80 bp
NoBa - forward	CAT GGT TTA TGT TTT CTG TTA GGT TG	NO
NoBa - reverse	TAG GAG CAG CTG AAG CCA TAA T	NO
Rice housekeeping primers (phospholipase gene)		68 bp
KMV - forward	TGG TGA GCG TTT TGC AGT CT	NO
KMV - reverse	CTG ATC CAC TAG CAG GAG GTC C	NO

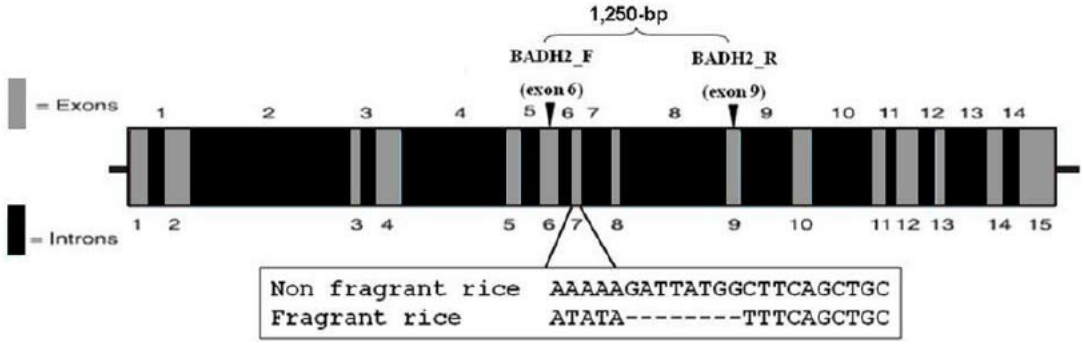


Figura 4: rappresentazione schematica della delezione delle 8 paia di basi all'interno del gene nel riso aromatico rispetto al riso non flavoured. Si nota la suddivisione del gene in introni e in esoni, gli intorni sono regioni non codificanti di un gene, mentre gli esoni possono essere codificanti. Entrambi vengono trascritti dall'RNA polimerasi, ma nell'mRNA maturo resteranno solamente gli esoni *(Srivong, 2007)*.

Tabella 2: quantità percentuale delle principali sostanze contenute all'interno di diverse tipologie di riso in base alla lavorazione che subisce il chicco *(la chimica del riso - Ute CiniselloB, s.d.)*.

Tipologia	Acqua	Proteine	Lipidi	Glucidi
Integrale	12	15,5	2,5	70
Sbiancato	12,5	7,5	1,3	78
Brillato	12,9	4,5	0,6	82,1

Poiché la composizione chimica cambia a seconda del grado di raffinazione, è possibile distinguere le seguenti tipologie di riso: integrale, ovvero il riso che ha subito sbramatura, quindi la rimozione della lolla; riso sbiancato, ovvero sottoposto a successive lavorazioni che rimuovono le parti più esterne come il pericarpo e il germe; riso brillato, cioè il riso che è stato trattato con talco e glucosio per conferire ai chicchi un aspetto più lucido. Le proteine maggiormente contenute nel riso presentano carenze di alcuni amminoacidi essenziali, come il triptofano e la lisina, mentre nei risi sottoposti a rimozione del germe si osserva una minore percentuale di lipidi al proprio interno (la chimica del riso - Ute CiniselloB, s.d.). I glucidi (chiamati anche carboidrati o zuccheri) sono composti organici formati da carbonio, ossigeno e idrogeno, svolgono numerose funzioni biologiche, rappresentano una delle principali fonti di energia e possiedono anche diverse funzioni strutturali.

Affinché un riso possa essere commercializzato come basmati deve possedere alcuni requisiti di purezza: in particolare deve contenere al suo interno meno del 7% di DNA di riso non basmati. Tale percentuale viene determinata mediante analisi dPCR, come indicato dalla Rice Association, British Rice Millers Association e il British Retail Consortium nel Basmati Rice Code of Practice (Basmati Rice Code of Practice Overview | PDF, 2017).

La maggior parte del riso aromatico prodotto a livello mondiale viene esportata da Pakistan e India, maggiori produttori di riso basmati, e dalla Thailandia, maggiore esportatrice di riso jasmine. Il prezzo del riso basmati può essere anche tre volte superiore al prezzo del riso non aromatico di alta qualità, a causa della minore resistenza alle malattie delle varietà aromatiche, che può portare ad una diminuzione nella resa produttiva.

Come descritto in precedenza, la 2-acetil-1-pirrolina (2-AP) è il composto più distintivo del riso aromatico rispetto a quello non aromatico, di conseguenza numerosi studi si sono focalizzati sullo sviluppo di metodi per la rilevazione e la quantificazione di 2-acetil-1-pirrolina nelle diverse cultivar di riso. Per l'analisi chimica della autenticità alimentare del riso basmati si sfrutta il fatto che la caratteristica fragranza del riso basmati è associata alla presenza di 2-acetil-1-pirrolina (**Figura 5**).

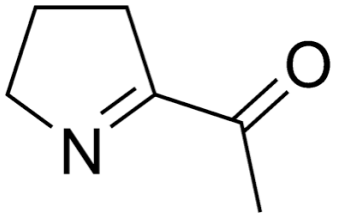


Figura 5: struttura molecolare della 2-acetil-1-pirrolina *(Fuso, 2023)*.

La 2-acetil-1-pirrolina (nome IUPAC: 1-(3,4-diidro-2H-pirrol-5-il)etanone; numero CAS: 85213-22-5 è stata identificata per la prima volta nel riso da Buttery, Ling e Juliano (1982) ed è considerata il composto aromatico più importante del riso, in particolare in quello aromatico. In tale studio, una concentrazione pari a 0,05 ppm di 2-acetil-1-pirrolina è stata descritta come caratterizzata da un odore simile a quello dei popcorn. In tale studio sono stati determinati i limiti di percezione olfattiva in acqua e in aria, corrispondenti rispettivamente a 0,1 nL/L e 0,02 ng/L. Questi valori, estremamente bassi, evidenziano l'importanza di questa molecola nel determinare l'aroma di un alimento. Oltre al riso, la 2-AP è un composto aromatico chiave in numerosi prodotti di origine vegetale e animale (Adams e De Kimpe, 2006; Wakte et al., 2017) (Wei. Handoko, 2017).

In seguito allo studio condotto da Yang, Shewfelt, Seong Lee e Kays, è stata analizzata la composizione aromatica di sei tipologie di riso caratterizzati da profili aromatici distinti: un campione di riso basmati, uno di jasmine, due cultivar coreane di tipo Japonica, un campione di riso nero e uno di riso non aromatico. Lo studio è stato compiuto analizzando lo spazio di testa attraverso gascromatografia-spettrometria di massa (GC-MS) e gascromatografia-olfattometria (GC-O) utilizzando un olfattometro, ovvero uno strumento impiegato per rilevare e misurare la diluizione degli odori, attraverso la determinazione della soglia di percezione olfattiva di determinate sostanze. Sono stati rilevati 36 composti odorosi presenti nei campioni sottoposti a riscaldamento, 25 dei quali caratterizzati da un'intensità dell'odore intermedia o superiore (intensità dell'odore ≥ 3 nella scala di intensità a 6 gradi). Questi 25 composti sono considerati i principali composti attivi dal punto di vista olfattivo: il composto che ha mostrato la soglia di percezione più bassa (0,02 ng/L) è la 2-acetil-1-pirrolina (2-AP), seguita da 11 aldeidi (con valori compresi tra 0,09 e 3,1 ng/L), dal guaiacolo (1,5 ng/L) e dall'1-otten-3-olo (2,7 ng/L). Oltre ai valori di soglia di percezione sono stati analizzati anche i valori di attività olfattiva (OAV, Odour Activity Value), definita come

il rapporto tra la concentrazione di ciascun analita e la sua soglia di percezione olfattiva. È emerso che i valori di OAV di 2-AP, esanale, del (E)-2-nonenale, ottanale, eptanale e nonanale rappresentano oltre il 97% del contributo degli OAV per ciascuna tipologia di riso analizzata. Tuttavia, il contributo dei singoli composti cambia in funzione alla tipologia di riso. In seguito alle analisi è emerso che sono 13 i composti attivi (odorosi) maggiormente responsabili delle differenze aromatiche nelle 6 diverse tipologie di riso: 2-AP, esanale, (E)-2-nonenale, ottanale, eptanale, nonanale, 1-otten-3-olo, (E)-2-ottenale, (E,E)-2,4-nonadienale, 2-eptanone, (E,E)-2,4-decadienale, decanale e guaiacolo (Yang, 2008).

SCOPO DELLA TESI

Lo scopo del presente studio è valutare, ed eventualmente ottimizzare, l'impiego di tecniche analitiche sensibili e selettive per verificare l'autenticità del riso commercializzato come basmati, in modo da rafforzare la tutela dei consumatori e potenziare gli strumenti per la tracciabilità e il controllo della qualità dei prodotti alimentari.

A tal fine, sono state condotte analisi presso i laboratori interni di Chimica e Biologia molecolare di Coop Italia, situati a Casalecchio di Reno (BO), per verificare l'autenticità di diversi campioni di riso commercializzati come basmati. Sono stati utilizzati due diversi approcci analitici: da un lato l'amplificazione e l'analisi del DNA tramite digital PCR (dPCR), dall'altro l'analisi e la caratterizzazione del profilo dei composti organici volatili ottenuti tramite gascromatografia con rivelatore a ionizzazione di fiamma (GC-FID) e gascromatografia accoppiata a spettrometria di mobilità ionica (GC-IMS).

Per quanto riguarda l'approccio biologico, sono state effettuate analisi del DNA di miscele sperimentali costituite da diverse percentuali di riso basmati e non basmati, per valutare la capacità del metodo di quantificare la percentuale di riso non basmati più simile al valore reale atteso. In seguito, il metodo è stato applicato a diversi campioni commerciali di riso basmati e non basmati, per verificarne l'autenticità, e controllare che nei campioni commercializzati come basmati non fosse superato il limite del 7% di riso non basmati al proprio interno.

L'approccio chimico ha previsto, invece, un'analisi dettagliata dei cromatogrammi dei composti volatili dei risi puri e delle miscele di riso basmati e non basmati, per verificare la presenza e l'accumulo di 2-acetil-1-pirrolina nelle varietà basmati. Sulla base dei profili ottenuti è stato creato un modello in grado di discriminare i campioni basmati dai non basmati e di prevedere le percentuali, espresse in peso su peso, dei componenti delle miscele analizzate.

L'obiettivo futuro consiste nell'analizzare e confrontare i cromatogrammi dei composti volatili di numerosi campioni di riso appartenenti a diverse tipologie, sfruttando il fatto che l'analisi è veloce e poco costosa. Successivamente, saranno sottoposti ad analisi di biologia molecolare solo i campioni di riso basmati che risultano avere una percentuale di riso non basmati superiore al limite previsto, riducendo così il tempo e il costo delle analisi.

Questi metodi non hanno ancora un valore legale, ma vengono utilizzati per il controllo qualità interno al laboratorio COOP Italia. In futuro, tuttavia, potrebbero essere validati e riconosciuti come metodi ufficiali. La tecnica gascromatografica, grazie ai tempi rapidi delle analisi e ai costi contenuti, è infatti già utilizzata anche in altri ambiti, per esempio durante i controlli effettuati presso i Laboratori Chimici dell'Agenzia delle Dogane e dei Monopoli (ADM) ed i Laboratori dell'Ispettorato centrale della tutela della qualità e della repressione frodi dei prodotti agroalimentari (ICQRF).

MATERIALI E METODI

Campioni analizzati

Per verificare la percentuale di riso non basmati all'interno dei campioni di riso commercializzati come puri, è stata eseguita l'analisi di 44 campioni di riso commerciali. Questi campioni sono venduti da 13 marchi anonimi differenti e hanno produttori e stabilimenti di produzione diversi. In **Tabella 3** si riporta l'elenco dei 44 campioni di riso analizzati, specificandone la varietà.

Tabella 3: specie e varietà dei campioni di riso puri commerciali analizzati.

SPECIE	**VARIETA'**	**Numero**	**TOT**
Oryza sativa, sottospecie Japonica	Carnaroli	9	
	Arborio	2	
	Ribe	2	
	Vialone nano	2	
	Baldo	3	
	Originario	1	
	Lungo A	1	
	Venere	1	
	Nerone 2	1	
	Volano	1	23
Oryza sativa, sottospecie Indica	Basmati	17	
	Lungo B	3	
	Jasmin	1	21

Sono state preparate miscele contenenti diverse percentuali, espresse in peso/peso, di riso non basmati (varietà Lungo B) e basmati (entrambi specie *Oryza sativa* L subsp. Indica). Le miscele finali erano così costituite: 1%, 3%, 5%, 7%, 10%, 25%, 50% e 75% di riso non basmati. In totale sono state preparate 19 miscele: 8 unendo il riso non basmati di marchio anonimo E (campione numero 233) con il riso basmati di marchio anonimo D (campione numero 232), 8 combinando il riso non basmati di marchio anonimo E con il riso basmati di marchio anonimo H (campione numero 234),

oltre ai risi puri corrispondenti alla miscela con 0% non basmati (analisi dei 2 risi basmati utilizzati per fare le miscele) e alla miscela con il 100% di riso non basmati.

Analisi di biologia molecolare (dPCR)

Estrazione DNA

Per l'estrazione del DNA è stato seguito il protocollo DNeasy Mericon Food Kit (Qiagen, Hilden, Germania). Il principio del kit si basa sull'utilizzo del CTAB (bromuro di cetil-trimetilammonio), della Proteinase K e di un Food lysis buffer.

Il bromuro di cetil-trimetilammonio è un detergente cationico che è utilizzato per l'estrazione degli acidi nucleici da una vasta gamma di matrici. In base alle condizioni saline utilizzate, il CTAB si può complessare con gli acidi nucleici (in presenza di una concentrazione bassa di sale) o con inibitori/contaminanti vari, come i metaboliti delle piante, polisaccaridi o proteine (in presenza di una concentrazione alta di sale).

La proteinasi K è un potente enzima utilizzato per la degradazione delle proteine nei campioni biologici; la sua azione facilita il rilascio degli acidi nucleici e garantisce la rimozione efficiente dei contaminanti in modo da ottenere campioni puliti e di alta qualità per il sequenziamento, la PCR ed altre applicazioni. Il suo utilizzo è particolarmente utile per lavorare con campioni complessi, come i tessuti, dove una degradazione proteica completa è fondamentale (Proteinase K for Efficient Protein Digestion, s.d.).

Il Food lysis buffer è una soluzione tampone che viene utilizzata per lisare le cellule presenti nei campioni alimentari e favorire il rilascio del materiale genetico. Gli inibitori e i componenti cellulari vengono precipitati in seguito a centrifugazione, mentre il DNA rimane in soluzione. Si utilizza successivamente cloroformio per rimuovere, tramite estrazione, i complessi CTAB-proteine, CTAB-polisaccaridi non precipitati e altri composti lipofilici che devono essere rimossi dai campioni.

Buffer utilizzati

Buffer PB: il tampone PB viene utilizzato nelle procedure di purificazione del DNA e consente un legame efficiente dei prodotti della PCR alla membrana della spin column. Contiene un'alta concentrazione di cloridrato di guanidina e isopropanolo (**Figura 6**) ed è utilizzato come tampone di legame. Il guanidinio denatura le proteine residue e destabilizza le strutture secondarie degli acidi nucleici, mentre l'isopropanolo crea le

condizioni di elevata forza ionica e polarità che permettono al DNA di adsorbirsi selettivamente alla membrana di silice della colonna (DNA Cleanup Buffers, s.d.).

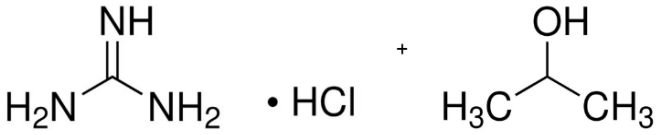


Figura 6: formula cloridrato di guanidina + isopropanolo (2-propanol, s.d.) (Guanidina - Aminoformamidina, Aminometanamidina, s.d.).

Buffer EB: è il tampone di eluizione usato a pH 8.5 per la purificazione del DNA ed è composto da 10 mM tris(idrossimetil)amminometano cloridrato (Tris-HCl, **Figura 7**). Il tampone serve per rilasciare il DNA purificato dalla membrana di silice delle colonne. Il DNA purificato può essere eluito anche in tampone TE (10 mM Tris-Cl, 1 mM EDTA, pH 8,0), ma l'EDTA (acido etilendiamminotetraacetico) potrebbe inibire le successive reazioni enzimatiche (Buffer EB, s.d.).

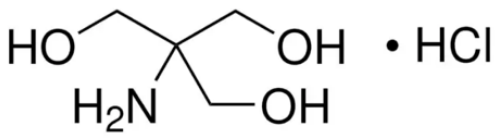


Figura 7: formula tris(idrossimetil)amminometano cloridrato (Tris-Cl) (Tris(idrossimetil)amminometano cloridrato EMPROVE® EXPERT, s.d.).

Buffer AW2: il tampone AW2 viene utilizzato come tampone di lavaggio, e richiede prima dell'uso l'aggiunta di una quantità appropriata di etanolo (96-100%) per ottenere una soluzione di lavoro. È composto da un glicole, NaCl 100 mM e Tris-HCl 10 mM, a pH 7.5. L'etanolo serve per mantenere il DNA legato alla silice mentre vengono rimossi gli inibitori della successiva reazione di PCR (Himmelreich, 2009).
Il kit viene utilizzato in una stazione semiautomatica di estrazione chiamata QIAcube (Qiagen, Hilden, Germany, **Figura 8**), in cui avviene un ciclo di estrazione che prevede la fase di lisi, adsorbimento, lavaggio ed eluizione.

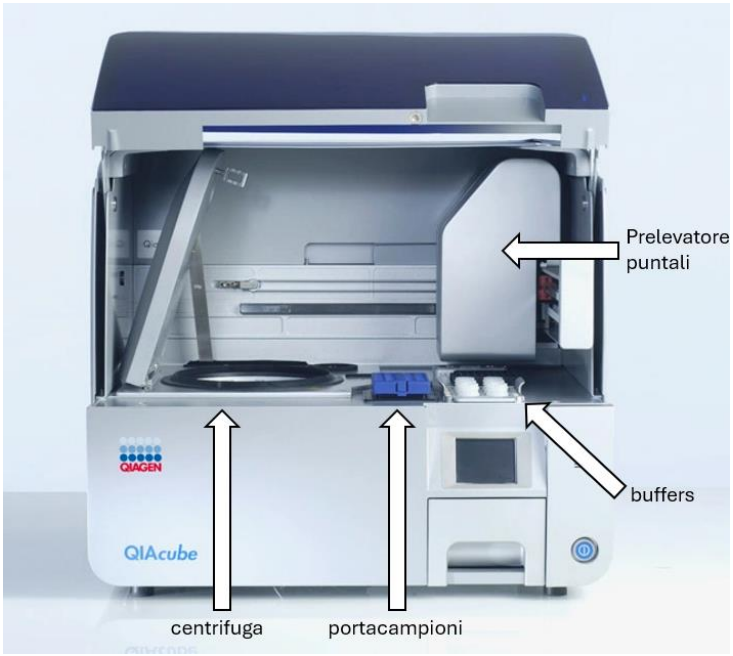


Figura 8: *rappresentazione dell'estrattore di DNA semiautomatico QIAcube (Qiagen QIAcube: Product Development – Helbling Projects, s.d.).*

Lettura della concentrazione del DNA

I fluorimetri QUBit™ (**Figura 9**) rilevano coloranti fluorescenti che sono altamente specifici per una molecola bersaglio di interesse presente nei campioni in esame (nel nostro caso il DNA). I coloranti si legano selettivamente al DNA, all'RNA o alle proteine ed emettono fluorescenza anche se i bersagli a cui si associano sono presenti in basse concentrazioni. L'intensità della fluorescenza è direttamente proporzionale alla quantità di DNA presente nel campione. I fluorimetri utilizzano una curva di calibrazione sviluppata da campioni standard a concentrazione nota. La concentrazione di un campione incognito di DNA, RNA o proteina viene calcolata confrontando le unità di fluorescenza relative (RFU) del campione con le RFU degli standard. La tecnologia di fluorescenza è così sensibile che i test richiedono appena 1 µL di campione, e la misura può essere effettuata anche se il campione è molto diluito (Qubit Thermo Fisher Scientific | Thermo Fisher Scientific, s.d.). Ogni QUBit™ dsDNA BR Assay Kit contiene: Qubit™ dsDNA BR Reagent (Component A) 200 x in DMSO, Qubit™ dsDNA BR Buffer (Component B), Qubit™ dsDNA BR Standard #1 (Component C) 0 ng/µL in TE buffer e Qubit™ dsDNA BR Standard #2 (Component D) 100 ng/µL in TE buffer.



Figura 9: raffigurazione del fluorimetro QUBit 3.0 *(Qubit® 3.0 Fluorometer - Unimed Scientifica srl, s.d.)*.

Digital PCR

La reazione a catena della polimerasi (PCR, Polymerase Chain Reaction) è una tecnica di laboratorio utilizzata per amplificare rapidamente milioni o miliardi di copie di uno specifico segmento di DNA, consentendone lo studio e l'analisi più nel dettaglio. La PCR prevede l'impiego di primers, ovvero brevi frammenti sintetici di DNA, che riconoscono e delimitano la regione di DNA da amplificare. L'amplificazione avviene attraverso cicli multipli di sintesi del DNA che portano ad amplificare la sequenza bersaglio (Polymerase Chain Reaction (PCR), s.d.). In **Figura 10** è riportato lo schema delle varie fasi di una reazione di Real-Time PCR, tecnica che consente di amplificare e quantificare uno specifico frammento, misurando l'andamento della fluorescenza emessa ad ogni ciclo di amplificazione.

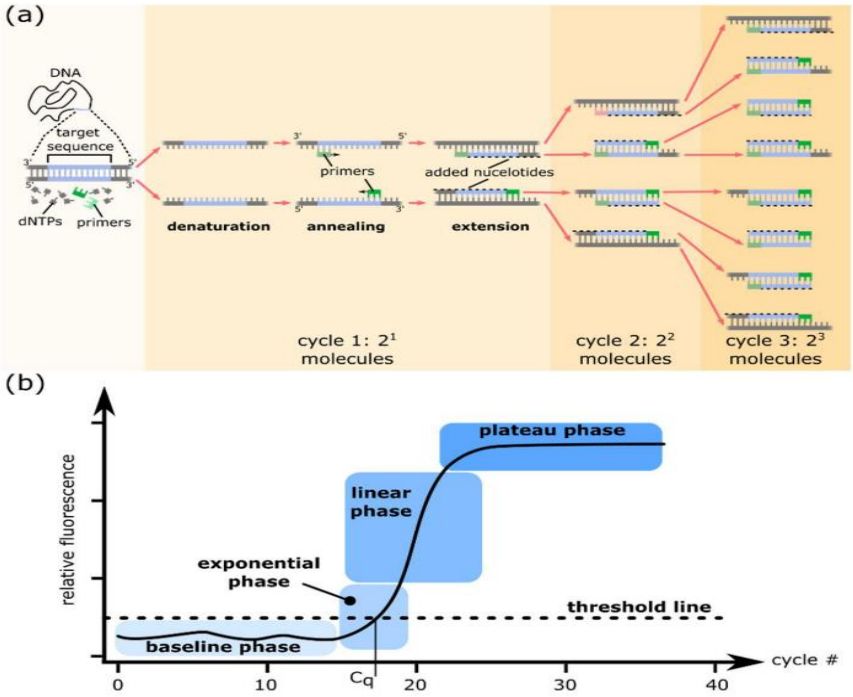


Figura 10: (a) principi della reazione a catena della polimerasi: ogni ciclo di PCR comprende tre fasi: (1) denaturazione del DNA a doppio filamento fornendo calore, (2) annealing dei primer alle loro sequenze di DNA bersaglio complementari, (3) estensione dei primer da parte di una DNA polimerasi termostabile. Una tipica reazione di PCR viene ripetuta 20-40 volte. Ogni ciclo può teoricamente portare al raddoppio del numero di molecole della sequenza bersaglio.

(b) diverse fasi di un grafico di amplificazione PCR in tempo reale, si riporta sull'asse delle ordinate la fluorescenza relativa e sull'asse delle ascisse il numero di cicli. Un tipico grafico di amplificazione PCR mostra una curva a forma sigmoide con 4 fasi distinte. (1) nei primi cicli di PCR, il segnale di fluorescenza dovuto al prodotto di amplificazione rimane a livello di fondo. La baseline è impostata per eliminare il segnale di fluorescenza di fondo. (2) durante la fase esponenziale, la quantità di prodotto PCR raddoppia ad ogni ciclo se l'efficienza di amplificazione è del 100%. La soglia è impostata al di sopra del segnale di background all'interno della fase esponenziale. Il ciclo di quantificazione (Cq) è il numero di cicli in cui il grafico di amplificazione interseca la linea di soglia, che è impostata significativamente al di sopra della baseline. L'amplificazione esponenziale si verifica perché tutti i reagenti sono freschi e disponibili, e la cinetica spinge la reazione. (3) la fase lineare indica che i reagenti stanno diventando limitati, il che si traduce in una riduzione dell'efficienza di amplificazione, infatti, man mano che la reazione procede, parte dei reagenti viene consumata a causa dell'amplificazione; quindi, le reazioni iniziano a rallentare e il prodotto della PCR non viene più raddoppiato a ogni ciclo. (4) la fase di plateau indica una saturazione del segnale siccome i reagenti sono esauriti e non viene generato o rilevato alcun prodotto PCR aggiuntivo. La reazione si è arrestata, non vengono più formati i prodotti e, se lasciata abbastanza a lungo, i prodotti della PCR inizieranno a degradarsi. Ogni reazione raggiungerà un plateau in un punto diverso, a causa della diversa cinetica di reazione per ogni campione. La fase di plateau è quella in cui la PCR tradizionale effettua la misurazione, nota anche come rilevamento del punto finale (Quan, 2018) (Real-Time vs. Digital PCR vs. Traditional PCR, s.d.).

Anche la Digital Polymerase Chain Reaction (dPCR) è una reazione di PCR avanzata utilizzata per amplificare e quantificare direttamente specifiche sequenze di acidi nucleici come DNA o RNA. Il principio di funzionamento è differente rispetto alla PCR quantitativa tradizionale (qPCR), consentendo risultati più precisi e sensibili, con una tolleranza maggiore alla presenza di inibitori biologici. La dPCR, inoltre, non dipende da curve standard per la quantificazione, né necessita di campioni di riferimento (Differenze tra qPCR e dPCR, s.d.).

La procedura viene schematizzata in **Figura 11**. A differenza della qPCR, nella Digital PCR il campione viene prima suddiviso in molte sub-reazioni PCR indipendenti, dette partizioni, in modo che ogni reazione contenga poche o nessuna sequenza target. Una

volta avvenuta l'amplificazione, se è presente la molecola bersaglio all'interno di una partizione si avrà un segnale positivo quantificabile tramite fluorescenza, altrimenti le partizioni prive di target avranno un segnale negativo. Il numero di partizioni positive e negative viene utilizzato per determinare la quantità assoluta di sequenza bersaglio. La frazione di partizioni positive all'amplificazione viene utilizzata per determinare la concentrazione della molecola target mediante analisi statistica basata sulla distribuzione di Poisson, che consente una quantificazione accurata.

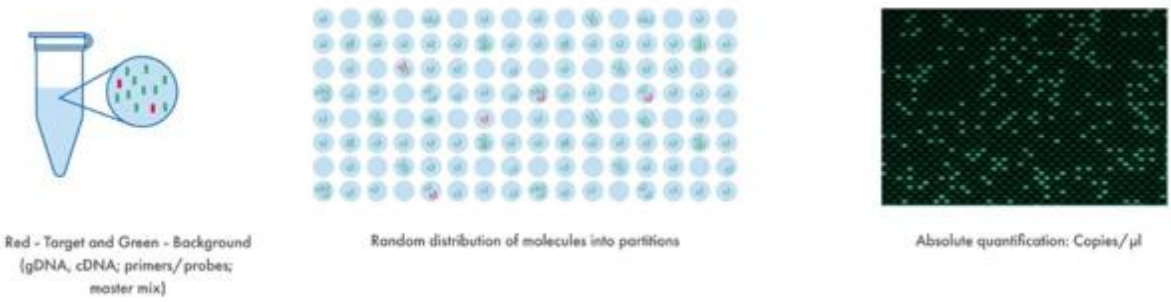


Figura 11: schematizzazione delle fasi della Digital PCR: preparazione del campione, partizionamento casuale delle molecole, amplificazione e analisi della fluorescenza (Fundamentals of digital PCR, s.d.).

Chimica EvaGreen:

La chimica utilizzata nella Digital PCR sfrutta la molecola fluorescente EvaGreen, un colorante di nuova generazione che si lega al DNA a doppio filamento e possiede diverse proprietà essenziali per la dPCR, come la stabilità, la sicurezza e un'intensa emissione di fluorescenza a seguito del legame al DNA.

Queste caratteristiche rendono il colorante EvaGreen molto apprezzato rispetto ad altri coloranti commerciali specifici per la PCR. Tale molecola utilizza un innovativo meccanismo di "rilascio su richiesta": il colorante è costituito da due unità monomeriche capaci di legare il DNA, collegate da un distanziatore flessibile. In assenza di DNA, EvaGreen assume una conformazione ad anello inattiva (non emette fluorescenza). In presenza di DNA, il colorante cambia conformazione una volta avvenuto il legame ed emette fluorescenza (**Figura 12**). Questo equilibrio fornisce continuamente colorante disponibile attivo man mano che il DNA viene amplificato durante la PCR (EvaGreen Dye for qPCR & Other Applications | Biotium., s.d.).

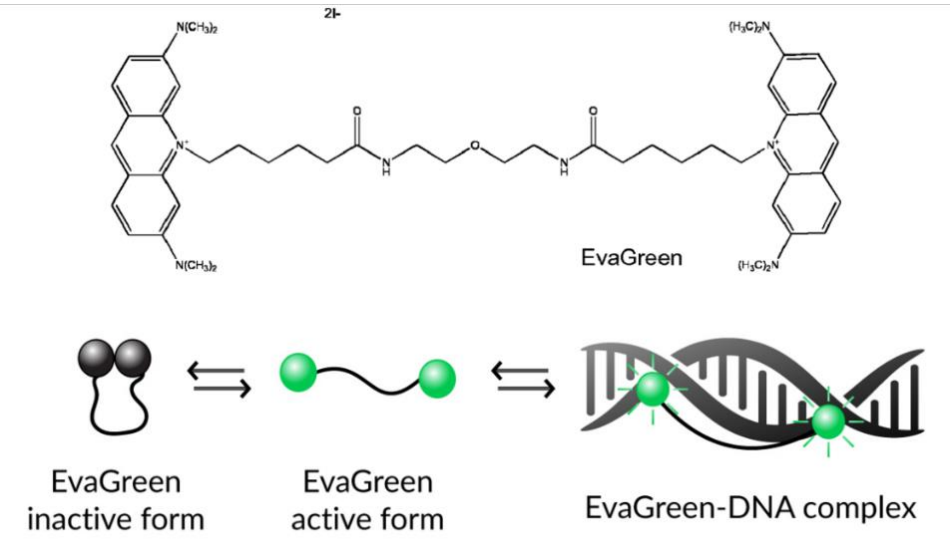


Figura 12: in alto si riporta la struttura chimica del colorante EvaGreen, in basso è mostrato il meccanismo di formazione della conformazione attiva del colorante in presenza della molecola di DNA (EP2565281A1 - Dye blends, s.d.) (EvaGreen® Dye For qPCR & Other Applications, s.d.).

Statistica di Poisson

Nella Digital PCR con il colorante EvaGreen, i campioni che risultano positivi possono contenere più di una copia della molecola target, di conseguenza una semplice somma delle partizioni positive non fornirà il numero corretto di molecole bersaglio presenti nel campione. Si utilizza quindi la distribuzione di Poisson per stimare correttamente il numero totale di copie della molecola bersaglio.

Questa strategia presuppone che ogni molecola abbia la stessa probabilità di trovarsi in una qualsiasi delle partizioni disponibili. La formula della distribuzione di Poisson è riportata in **Equazione 1** (Hu, 2008), dove con P si indica la probabilità di osservare k eventi in un dato intervallo. Nel caso della Digital PCR, k corrisponde alla probabilità che ci siano copie del DNA target all'interno delle partizioni. Il parametro λ è definito invece come il numero medio di eventi attesi nello stesso intervallo; quindi, nel nostro caso corrisponde al numero medio di copie di DNA target per partizione (**Equazione 2**) (Majumdar, 2015).

Equazione 1: $$P(k;\lambda) = -\ln\left(\frac{\lambda^k e^{-\lambda}}{k!}\right)$$

Equazione 2: $$\lambda = -\ln\left(\frac{\text{numero di partizioni valide - numero di partizioni positive}}{\text{numero di partizioni valide}}\right)$$

Utilizzando λ, la concentrazione target C (numero di copie di DNA) viene stimata tramite l'**Equazione 3**, dove V è il volume contenuto in ciascuna partizione (µL).

Equazione 3: $$C = \frac{\lambda}{V}$$

I risultati ottenuti in seguito alla distribuzione statistica di Poisson possono essere visualizzati attraverso grafici a dispersione 1D (**Figura 13**), nei quali ogni punto sul grafico corrisponde ad una singola partizione. Le partizioni positive, ovvero quelle contenenti una o più molecole bersaglio, vengono separate dalle partizioni negative sulla base di un valore soglia di fluorescenza.

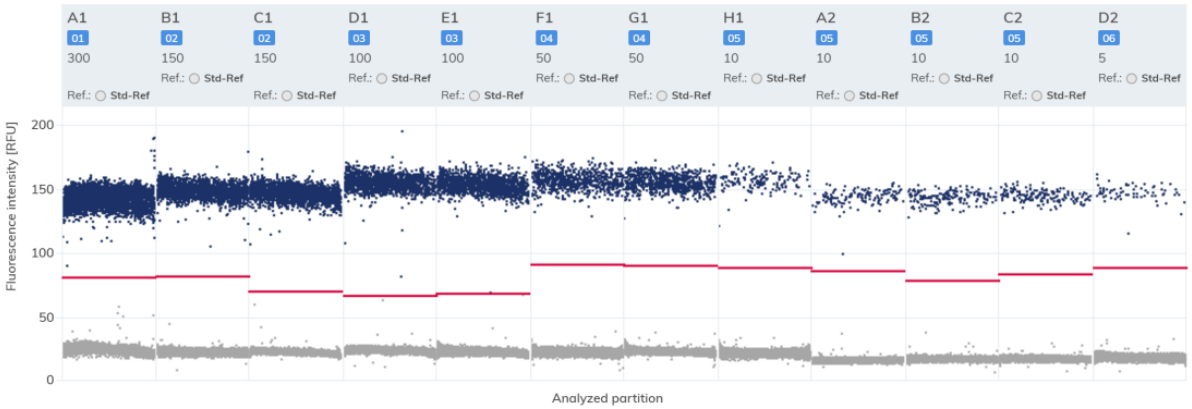


Figura 13: scatter 1D rappresentante le diluizioni seriali eseguite sul campione 92, sull'asse y si riporta l'intensità della fluorescenza in RFU e sull'asse delle x le partizioni analizzate. La linea di soglia è raffigurata in rosso, si nota come aumentando la diluizione del DNA le partizioni positive al primer KMV diminuiscano.

Di seguito sono riportati i componenti della Super Mastermix utilizzata per eseguire le analisi del DNA:

1. Due coppie di primers (forward e reverse) per misurare la quantità di DNA target amplificata rispetto alla quantità di DNA totale effettivamente amplificata.
2. Mastermix 3x (Qiagen, Hilden, Germania), composta da:

- Taq polimerasi termostabile hot-start: a basse temperature la DNA polimerasi viene mantenuta in uno stato inattivo dall'anticorpo QuantiNova e da un additivo, QuantiNova Guard, che stabilizza il complesso. Questo sistema impedisce l'estensione dei primers legati in modo non specifico; entro 2 minuti dall'aumento della temperatura a 95°C, QuantiNova e QuantiNova Guard si denaturano e la DNA polimerasi si attiva, consentendo l'avvio dell'amplificazione.
- dNTP: deossinucleotidi trifosfati utilizzati come elementi costitutivi per la sintesi del nuovo DNA.

- Sistema tampone: il tampone della PCR, contenente Tris-HCl e $MgCl_2$ come cofattore per l'attività della polimerasi, serve per mantenere il pH ottimale e la giusta forza ionica. La concentrazione di $MgCl_2$ è tipicamente 2,5-3 mM nella reazione finale.
- Colorante EvaGreen: colorante fluorescente intercalante che si lega al DNA a doppio filamento, consentendo il monitoraggio della fluorescenza in tempo reale (QIAcuity EvaGreen (EG) PCR Kit, s.d.), (Fast EvaGreen® qPCR Master Mix, s.d.).

Funzionamento del QIAcuity one:

Lo strumento QIAcuity one è basato sull'analisi di nanopiastre all'interno delle quali avviene il partizionamento della miscela di DNA, consentendo l'analisi simultanea di più campioni. La piastra utilizzata è costituita da 26 pozzetti contenenti ciascuno 24.000 nanopozzetti dove avviene la reazione di amplificazione. All'interno del QIAcuity one avvengono tre fasi:

1. *reaction partitioning* (o priming): la miscela di reazione viene distribuita sulla piastra in modo tale da riempire le migliaia di partizioni presenti. Successivamente vengono bloccati i microcanali che collegano le diverse partizioni, generando 24.000 compartimenti individuali.
2. *thermocycling*: avviene l'amplificazione nel termociclatore a effetto Peltier in base al profilo termico impostato che, nel nostro caso, prevedeva: 1 ciclo a 95°C per 5 minuti, 50 cicli a 95°C per 30 secondi e a 60°C per 1 minuto.
3. *imaging*: avviene la rilevazione della fluorescenza con una camera che acquisisce immagini delle singole partizioni. Va selezionato il canale di fluorescenza adatto che, nel nostro caso, è il canale Green.

La **Figura 14** riporta il procedimento schematizzato per la preparazione delle piastre, il lavoro automatizzato del QIAcuity One e l'analisi della fluorescenza ottenuta, mentre le componenti dello strumento QIAcuity One sono mostrate in **Figura 15**.

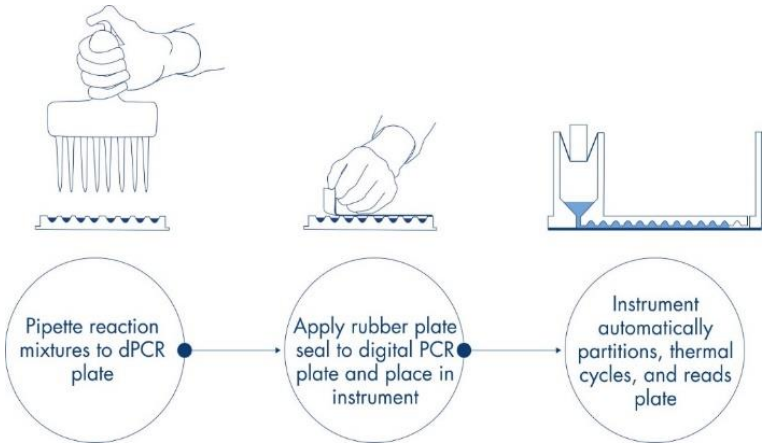


Figura 14: schema del procedimento per l'ottenimento dei risultati in seguito all'analisi svolta dal QIAcuity One: 1) inserire la soluzione composta da supermastermix + campione di DNA all'interno dei pozzetti della piastra, 2) applicare la gomma sopra alla piastra ed inserirla all'interno dello strumento, 3) lo strumento svolgerà automaticamente la partizione, i cicli termici e la lettura della fluorescenza e fornirà i risultati (QIAcuity Digital PCR System, s.d.).

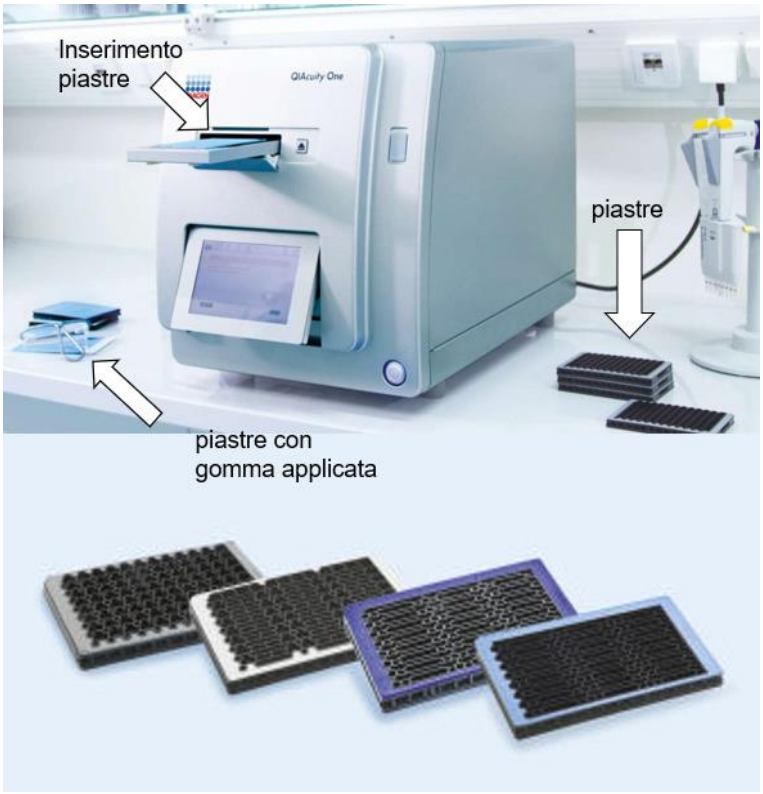


Figura 15: sopra viene mostrato lo strumento QIAcuity ONE con di fianco le piastre con e senza la gomma applicata. Sotto vengono mostrate da destra verso sinistra la piastra 26k x 96 pozzetti, 26k x 24 pozzetti, 8.5k x 24 pozzetti e 8.5k x 96 pozzetti. (QIAGEN QIAcuity dPCR, s.d.) (QIAcuity Nanoplates and Accessories, s.d.).

Procedura analitica

Estrazione DNA:

Materiali utilizzati:

- Macinatore Grindomix GM 200 Knife Mill (Retsch, Haan, Germania)
- Bilancia QUINTIX 1102 1CFR (Sartorius, Göttingen, Germania)
- Food lysis buffer (Qiagen, Hilden, Germany)
- Proteinase K, Qiagen
- Rotatore ROTATOR SB2 (Stuart, Vernon Hills, Illinois, Stati Uniti d'America)
- Stufa (Memmert, Schwabach, Germania)
- Centrifuga ALLEGRA X-22R (Beckman Coulter, Brea, California, Stati Uniti d'America)
- Cloroformio contenente 100-200 ppm di amileni come stabilizzanti, ≥ 99.5% (Sigma-Aldrich, Saint Louis, Missouri, Stati Uniti d'America)
- Cappa Chemfast (Carlo Erba, Milano, Italia)
- Centrifuga microfuge 22R centrifuge, Beckman Coulter
- Vortex Gen-2 digital (Scientific Industries, Bohemia, New York, Stati Uniti d'America)
- QIAcube, Qiagen.

Procedura:

1. Pesare 1 g di riso macinato dentro ad una eppendorf
2. Aggiungere 5 mL di Food Lysis Buffer e 25 µL di Proteinase K solution
3. Mescolare 2/3 volte, agitare con il vortex e inserire nel rotatore in stufa a 60°C per tutta la notte
4. Centrifugare a 6169 rpm per 5 min a 20°C
5. Prelevare 700 µL di surnatante e inserirli in una eppendorf da 2 mL
6. Aggiungere 500 µL di cloroformio sotto cappa
7. Agitare con il vortex le eppendorf per 1 min a 2240 rpm
8. Centrifugare per 15 minuti a 26°C a 14.000 g
9. Prelevare 350 µL di surnatante e inserirle in una eppendorf da 2 mL
10. Inserire le eppendorf nell'apposito spazio del QIAcube, e impostarlo per l'estrazione.

Al termine della procedura di estrazione si ottengono 70 µL di eluato contenenti il DNA.

Lettura della concentrazione del DNA:

Materiali utilizzati

- QUBit™ dsDNA BR Assay Kit (Thermo Fisher Scientific, Waltham, Massachusetts, Stati Uniti d'America)
- Qubit assay tubes, invitrogen, Thermo Fisher Scientific
- QUBIT 3.0 fluorometer, invitrogen, Thermo Fisher Scientific
- DNA estratto
- Vortex Gen-2 digital (Scientific Industries, Bohemia, New York, Stati Uniti d'America)

Procedura

1. Preparare una soluzione unendo in una eppendorf 199 µL di buffer e 1 µL di colorante per ogni campione da analizzare
2. Trasferire la miscela nel tubo di reazione e aggiungere 3 µL di DNA da quantificare
3. Agitare con il vortex e aspettare 2 minuti prima della lettura della concentrazione di DNA

Digital PCR:

Per eseguire la reazione di amplificazione è stata utilizzata una concentrazione finale di DNA pari a 2 ng/µL.

Materiali utilizzati:

- Mastermix 3x QIAcuity EG PCR (Qiagen, Hilden, Germany).
- Primers NoBa e KMV 100 µM EUROFINS GENOMIC, Qiagen
- Acqua RNase-free, Qiagen
- Centrifuga da laboratorio Mini-P25 (All-Sheng, Hangzhou, Cina)
- MicroAmp Clear Adhesive Film (Thermo Fisher Scientific, Waltham, Massachusetts, Stati Uniti d'America)
- Piastra 26x24000, Qiagen
- QIAcuity One, Qiagen
- Campioni DNA da analizzare
- Vortex Gen-2 digital (Scientific Industries, Bohemia, New York, Stati Uniti d'America)

Procedura:

1. Prima dell'utilizzo agitare con il vortex e centrifugare i primers, la Mastermix e i campioni di DNA
2. Preparare la miscela di reazione, in quantità superiore al necessario, miscelando acqua, primer forward e reverse alla concentrazione finale pari a 400 nM e l'opportuna quantità di Mastermix 3x.
3. Per preparare i campioni si miscelano 38 µL di questa soluzione e 2 µL di DNA nei 26 pozzetti di una piastra che viene sigillata con l'apposita pellicola trasparente e centrifugata.
4. I 40 µL presenti in ogni pozzetto della piastra vengono trasferiti nella piastra per dPCR (26 x 24.000 nanopozzetti) dove avviene il partizionamento. La piastra viene chiusa con l'apposita guarnizione.
5. Inserire la piastra all'interno del QIAcuity One e procedere con l'amplificazione.

<u>Elaborazione dei dati</u>

Per verificare l'ottimizzazione del metodo, è stato calcolato il limite di quantificazione (LOQ) della dPCR e il range lineare delle misure. Il limite di quantificazione è definito come la concentrazione minima di un analita che si può misurare con un livello di precisione e sicurezza accettabile. Il range lineare, invece, è definito come l'intervallo di concentrazioni all'interno del quale il segnale analitico varia in maniera direttamente proporzionale con la concentrazione degli analiti; in questo intervallo si possono determinare concentrazioni di campioni incogniti una volta determinata la retta di regressione lineare.

Per la determinazione del LOQ mediante dPCR, è stato analizzato un campione di riso non basmati (campione 92, riso baldo). Una volta determinata la concentrazione di DNA del campione, sono state preparate delle diluizioni seriali in modo da ottenere concentrazioni teoriche pari a 300, 150, 100, 50, 10, 5, 4, 3, 2 e 1 copie/µL per il primer target NoBa e 300, 150, 90, 30, 15, 9, 3, 2 e 1 copie/µL per il primer di controllo KMV. Per determinare il LOQ si calcola il coefficiente di variazione (CV%), definito come il rapporto tra la deviazione standard e la media delle misure per ciascuna diluizione seriale eseguita, espresso in percentuale. Al fine di avere un numero di copie/µL superiore al LOQ, il CV% non deve eccedere il 25% (Kralik, 2017). Per la verifica della linearità del metodo, è stato costruito un grafico riportando sull'asse delle ordinate la percentuale di riso non basmati calcolata tramite dPCR e sull'asse delle ascisse la

percentuale teorica di riso non basmati presente nelle miscele. I dati sono stati analizzati mediante regressione lineare ed è stato verificato l'ottenimento di un coefficiente di determinazione (R^2) il più prossimo possibile a 1.

Analisi chimica untargeted

Gascromatografia

La cromatografia gassosa (GC) è una tecnica analitica comunemente utilizzata per separare e analizzare i composti volatili e semivolatili presenti in una miscela complessa. La GC è una tecnica caratterizzata da un elevato potere risolutivo, rapidità di analisi e sensibilità. È ampiamente applicata in molti settori, tra cui quello ambientale, petrolifero, chimico, alimentare e farmaceutico (Gascromatografia (GC), s.d.).

La GC può essere applicata all'analisi di miscele contenenti da pochi a decine di grammi di composti con punti di ebollizione compresi da valori prossimi a zero fino a oltre 700 K, purché tali composti possano essere riscaldati a sufficienza senza decomporsi fino a raggiungere una pressione di vapore di pochi mmHg (Bartle, 2002). La GC condivide i principi fondamentali delle altre tecniche cromatografiche, distinguendosi principalmente per l'impiego di una fase mobile gassosa atta a trasportare l'analita attraverso la fase stazionaria (colonna cromatografica). La GC utilizza come fase mobile un gas inerte non reattivo (come l'elio), chiamato *carrier gas.* Se la fase stazionaria è un liquido viscoso (come un idrocarburo a catena lunga) si parla di cromatografia gas-liquido (GLC), se è solida (come gel di silice e allumina) si parla di cromatografia gas-solido (GSC). Il tubo capillare (colonna cromatografica) può avere uno spessore di pochi millimetri, ma può essere lungo fino a 10 metri. La separazione nella gascromatografia inizia con l'iniezione del campione in un iniettore riscaldato, dove avviene la vaporizzazione istantanea; successivamente, i componenti del campione vengono trasportati dalla fase mobile (gas di trasporto) attraverso la fase stazionaria in base alla loro volatilità e all'affinità relativa con la fase stazionaria: eluiscono più velocemente i composti che presentano minore affinità con la fase stazionaria. Ciò significa che i componenti lasciano la colonna in tempi diversi e possono essere rilevati singolarmente da un rilevatore che genera un segnale proporzionale alla quantità di analita presente. Tale segnale viene utilizzato per creare il cromatogramma, ovvero un grafico in cui sono riportati dei picchi di diversa intensità in funzione del tempo di ritenzione di ciascun componente, ovvero del tempo impiegato

per eluire dalla colonna. Nella gascromatografia, il cromatogramma fornisce anche informazioni quantitative sulla concentrazione dei componenti del campione, poiché l'area dei picchi è direttamente proporzionale alla concentrazione degli analiti. Nei gascromatografi accoppiati a rilevatori a ionizzazione di fiamma (FID), come Heracles II, si utilizza una fiamma per ionizzare i composti organici presenti nel campione, i cui analiti sono stati separati lungo la colonna cromatografica. La fiamma viene prodotta dalla combustione di idrogeno e aria pura secca (contenuto di acqua < 5 ppm nel caso di Heracles II) e produce atomi carichi positivamente ed elettroni. Lungo il percorso della fiamma sono disposti due elettrodi, quando i cationi si avvicinano all'elettrodo negativo avviene uno scambio di elettroni che genera una corrente elettrica che può essere registrata da un amperometro e convertita in un cromatogramma.

FlavourSpec25®

Teoria:

FlavourSpec25® è un gascromatografo accoppiato ad uno spettrometro a mobilità ionica ed effettua la ionizzazione dei composti tramite una sorgente radioattiva a trizio. Lo strumento è costituito da una colonna cromatografica multicapillare (MCC) lunga 20 cm con una fase stazionaria di tipo OV5 costituita dal 5% di difenil polisilossano e dal 95% di dimetilpolisilossano (**Figura 16**). Il gas carrier e il drift gas (gas di deriva) sono costituita da azoto N_2 con una purezza del 99,999%, prodotto a partire dall'aria. In **Figura 17** è raffigurato il gas cromatografo FlavourSpec25®, mentre in **Figura 18** si riporta la struttura interna del gascromatografo accoppiato allo spettrometro a mobilità ionica (FlavourSpec®25 with PAL3-RSI, 2024).

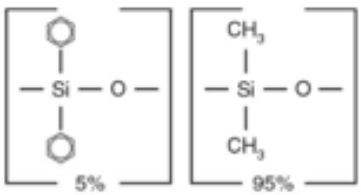


Figura 16: polimero difenil polisilossano/dimetil polisilossano (Colonna capillare GC Rtx-5, 30 m, ID 0,53 mm, 3,00 µm, s.d.).

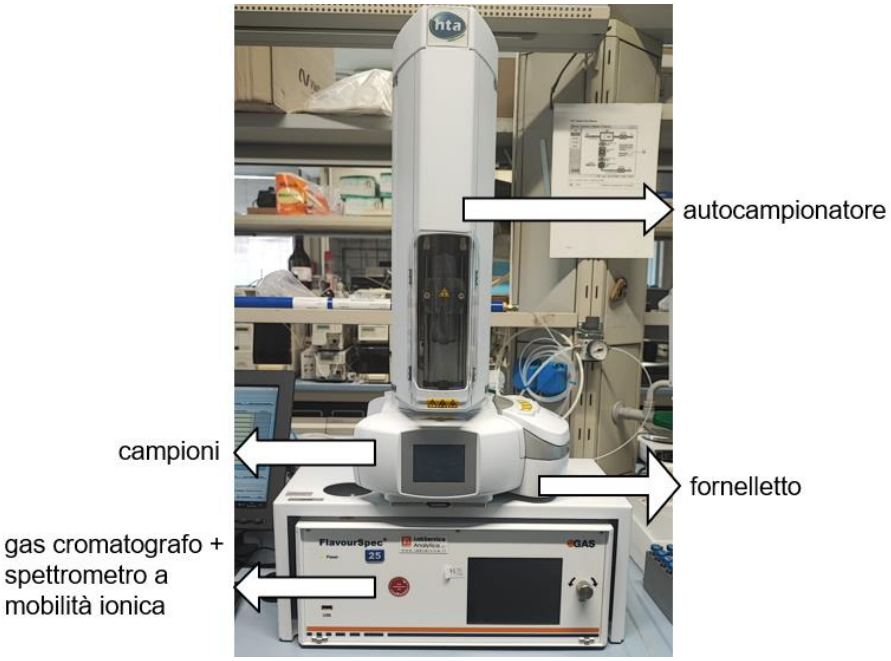


Figura 17: rappresentazione del sistema gas cromatografo FlavourSpec25® accoppiato con lo spettrometro a mobilità ionica.

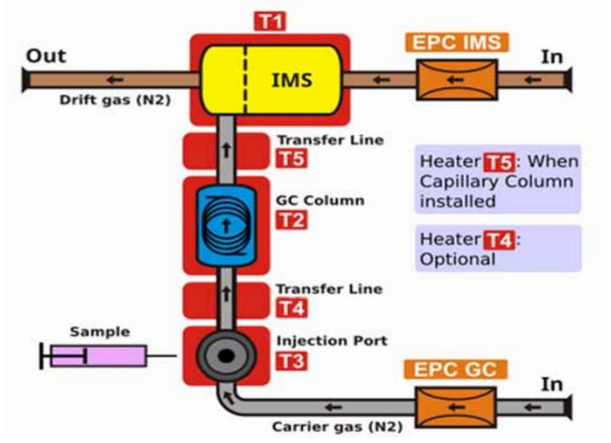


Figura 18: rappresentazione schematica di FlavourSpec25® raffigurante i sistemi di ingresso dei gas, il punto di iniezione, le linee di trasferimento e la colonna cromatografica. Tutte le componenti lavorano a temperature differenti.

Spettrometria a mobilità ionica:

La gascromatografia accoppiata alla spettrometria di mobilità ionica è una delle tecniche analitiche strumentali più comunemente utilizzate per identificare i composti aromatici volatili presenti negli alimenti.

La spettrometria di mobilità ionica (IMS) è una tecnica analitica per la separazione e l'identificazione dei composti contenuti in un campione sulla base della loro diversa mobilità in presenza di un carrier gas, di un drift gas e di un campo elettrico (Gabelica, 2018). Il nucleo di un sistema IMS è costituito dal tubo di deriva (drift tube), suddiviso in una regione di ionizzazione e una regione di separazione. Il principio generale di funzionamento dell'IMS è illustrato in **Figura 19**. I campioni vengono inizialmente riscaldati e vaporizzati, quindi trasferiti nella regione di ionizzazione mediante un carrier gas. Successivamente, le molecole neutre del campione e le molecole del gas carrier vengono ionizzate dalla sorgente ionica in modo da formare specie ioniche che, nella regione di separazione, presentano drift time (tempi di deriva) differenti. La ionizzazione del composto avviene tramite una sorgente radioattiva che cede protoni alle molecole di acqua in fase vapore presenti lungo la linea del gas, generando un complesso intermedio assimilabile a ioni H_3O^+. Questi intermedi successivamente cedono i protoni ai composti volatili nella camera di ionizzazione (**Reazione 2**, si indica con M un generico composto volatile): si tratta quindi di una ionizzazione soft poiché al composto volatile viene aggiunto solo uno ione H^+:

Reazione 2: $M + H^+(H_2O)_n \leftrightarrow MH^+(H_2O)_n^* \leftrightarrow MH^+(H2O)_{n-x} + xH_2O$

Gli ioni si muovono sotto l'influenza di un campo elettrico attraverso il drift gas e vengono rilevati in tempi diversi, fornendo informazioni utili sulla tipologia e sulla concentrazione degli analiti. Il drift time di uno ione che attraversa un debole campo elettrico omogeneo in presenza di un gas inerte è correlato alla mobilità dello ione (Wang, 2020). Al termine del percorso, gli ioni impattano su una piastra di Faraday, un rilevatore carico negativamente che trasforma la carica degli ioni in una corrente elettrica (misurata in V). Questo segnale viene quindi amplificato, convertito in voltaggio ed elaborato dal software dello strumento. L'intensità del segnale rilevato è direttamente proporzionale alla quantità di ioni che raggiungono la piastra, di conseguenza è direttamente proporzionale alla concentrazione del composto volatile analizzato. L'elaborazione dei dati ottenuti consente di generare un grafico bi- o tridimensionale del profilo dei composti volatili presenti che rappresenta l'impronta volatile univoca del campione in esame (*fingerprinting)*. Questa impronta consente di confrontare campioni differenti tra loro.

Mentre gli ioni vengono attirati dalla piastra di Faraday, un gas, chiamato drift gas (nel nostro caso azoto), viene alimentato in controcorrente rispetto al moto delle molecole, cosicché gli ioni vengono separati in base alla dimensione dato che quelli meno ingombranti effettuano meno urti con il drift gas e possano attraversare il drift tube in meno tempo rispetto agli ioni più pesanti. Le molecole vengono separate anche in base alla carica, in quanto quelle con maggiore carica positiva verranno attratte con più forza verso la piastra di Faraday (**Figura 19**).

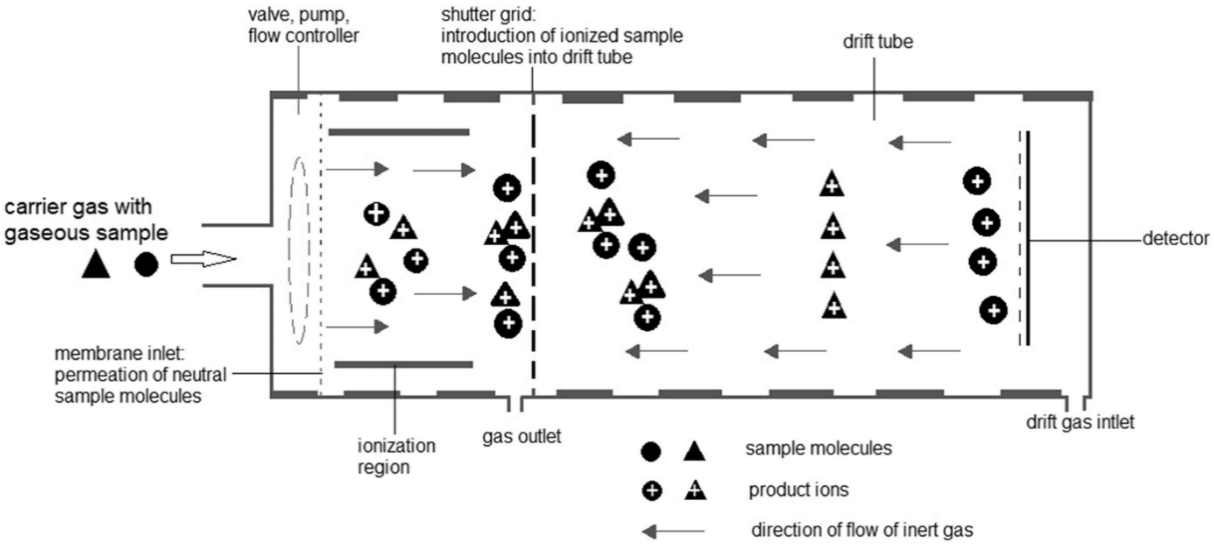


Figura 19: schema generale di uno spettrometro a mobilità ionica, a sinistra della griglia si trova la regione di ionizzazione, a destra si trova la regione di separazione (drift tube), in cui gli ioni vengono separati in base alla carica, alle dimensioni, alla geometria e vanno ad impattare un detector alla fine della regione di separazione (Borsdorf, 2006).

Procedura analitica

Sono stati esaminati i 44 campioni commerciali di riso precedentemente analizzati mediante analisi biologica, di cui 17 campioni di riso basmati e 27 campioni di riso non basmati di diverse tipologie e marche. Sono stati inoltre analizzate le 2 tipologie di miscele precedentemente esaminate con la Digital PCR, replicando le stesse percentuali di riso non basmati, espresse in peso/peso (0%, 1%, 3%, 5%, 7%, 10%, 25%, 50%, 75%, 100%), aggiunte al riso basmati, per un totale complessivo di 19 miscele analizzate. Tutti i campioni sono stati testati due volte mediante FlavourSpec25® in modo da avere un maggior numero di misure.

I gas cromatografi prelevano le vials dal portacampione e le inseriscono all'interno di un forno riscaldato, dove vengono scaldate in modo tale da favorire il rilascio dei

composti volatili nello spazio di testa. Successivamente, la siringa dell'autocampionatore preleva i composti volatili e li inietta nella colonna cromatografica, dove vengono separati in base alla loro affinità con la fase stazionaria. La separazione dipende anche dalla polarità della fase stazionaria, dalla temperatura di riscaldamento, dal flusso del gas di trasporto, dalla lunghezza e dimensione della colonna, dalla carica e geometria del composto volatile. In **Tabella 4** vengono riportate le condizioni operative utilizzate in questo studio.

Materiali utilizzati:

- bilancia PG5002 DELTARANGE (Mettler-Toledo, Columbus, Ohio Stati Uniti d'America)
- vials da 20 mL (Cole Palmer, Vernon Hills, Illinois, Stati Uniti d'America)
- tappi Kinesis, Cole Palmer
- crimper Wheaton (DVK Life Sciences, Wertheim, Germania)
- FlavourSpec25® (G.A.S. Gesellschaft für analytische Sensorsysteme mbH, Dortmund, Germania)

Procedura:

1. Preparare 2 vials da 20 mL per ciascun campione, pesando (4,00 ± 0,02) g di riso in ciascuna di esse, e 4 vials da 20 mL vuote che serviranno come bianco. Tutte le vials vanno sigillate ermeticamente
2. Posizionare le vials nel rack di FlavourSpec25®, collocando 2 bianchi all'inizio e 2 bianchi alla fine della sequenza delle analisi, quindi inserire il rack all'interno del portacampione

Tabella 4: condizioni operative del gas cromatografo FlavourSpec25® e dello spettrometro a mobilità ionica

FlavourSpec25®		
Parametri	Valori	Definizione
EPC_IMS	75.0 ml/min	Velocità del flusso del drift gas
EPC_GC	5.0 ml/min	Velocità del flusso del carrier gas
T1 IMS	45.0 °C	Temperatura di operatività dello spettrometro a mobilità ionica

T2 GC	40.0 °C	Temperatura di operatività del gas cromatografo
T3 injector	80.0 °C	Temperatura dell'iniettore
T4 transfer	80.0 °C	Temperatura della linea di trasferimento iniettore-colonna GC
T5 transfer	45.0 °C	Temperatura della linea di trasferimento colonna GC-IMS
Drift potential	5000 V	Voltaggio del campo elettrico
Injection volume	0.5 mL	Volume dello spazio di testa iniettato nel gas cromatografo
Time of analysis for a sample	34 min	Tempo totale di analisi del campione, comprende il tempo di riscaldamento preliminare, il tempo dell'iniezione e della corsa dei composti

Elaborazione dati:

In seguito alle analisi si ottengono dei cromatogrammi tridimensionali (**Figura 20**) nei quali**,** sugli assi x e y, vengono riportati rispettivamente il tempo di ritenzione del composto nella colonna cromatografica e il tempo di deriva nello spettrometro a mobilità ionica. Sull'asse z sono invece riportate le intensità dei segnali (altezze del picco), evidenziate da colori diversi.

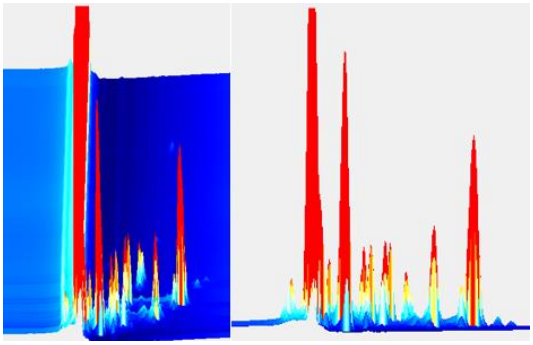


Figure 20: rappresentazione tridimensionale di un cromatogramma ottenuto a seguito dell'analisi GC-IMS con il gas cromatografo *FlavourSpec25®*.

Gli indici di ritenzione di Kovatz sono degli indici che esprimono i valori dei tempi di eluizioni dei composti organici in gascromatografia rispetto ai tempi di eluizioni di una miscela di chetoni. Vengono impiegati per standardizzare i dati ottenuti utilizzando

colonne cromatografiche a diversa polarità, così da avere dei tempi di ritenzioni uguali per gli stessi composti. Gli indici di ritenzione di Kovatz sono stati utilizzati in questo studio per associare ad ogni picco un possibile composto.

È stato utilizzato il software VOCal 0.4.33 per visualizzare i cromatogrammi tridimensionali ottenuti dalle analisi dei campioni di riso e per elaborare i dati utilizzando 2 moduli: DeNoise (G.A.S., Dortmund, Germania) e Percentile (G.A.S., Dortmund, Germania).

Il modulo DeNoise serve a ridurre il rumore di fondo e, quindi, ad aumentare il rapporto segnale/rumore rendendo più evidenti i picchi effettivamente presenti nel cromatogramma. Applicando questo modulo viene attenuato o eliminato anche il picco degli ioni reagenti (RIP= Reactant Ion Peak), permettendo di visualizzare i segnali deboli circostanti. In **Figura 21** si riporta un cromatogramma prima e dopo l'applicazione del modulo DeNoise.

Il modulo Percentile (**Figura 22**) invece permette di impostare, per ogni cromatogramma, un valore al di sotto del quale si trova la percentuale selezionata dei dati presenti in tutti i cromatogrammi ottenuti dalle analisi svolte. I percentili più utilizzati sono il 25°, il 50° (cinquantesimo percentile) e il 75°. È stato calcolato il cinquantesimo percentile per entrambe le varietà di riso (B, basmati e NB, non basmati) ottenendo quindi 2 cromatogrammi rappresentativi che sintetizzano le caratteristiche statistiche delle 2 classi. Infine, è stato generato il centesimo percentile unendo i due cromatogrammi corrispondenti ai cinquantesimi percentili delle 2 classi per ottenere un cromatogramma complessivo da cui selezionare le aree totali rilevanti.

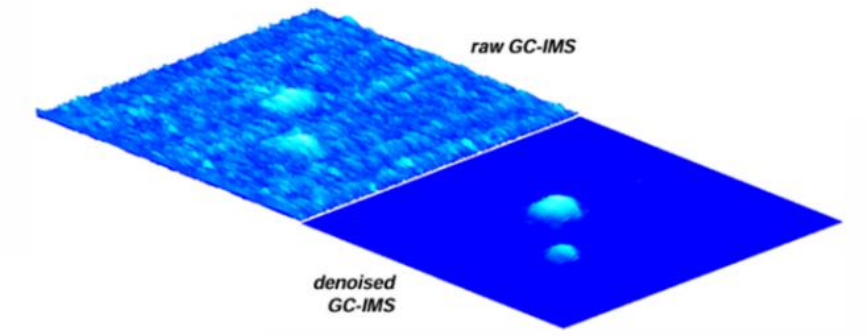


Figura 21: aumento del rapporto segnale rumore del cromatogramma tridimensionale in seguito all'applicazione del modulo DeNoise

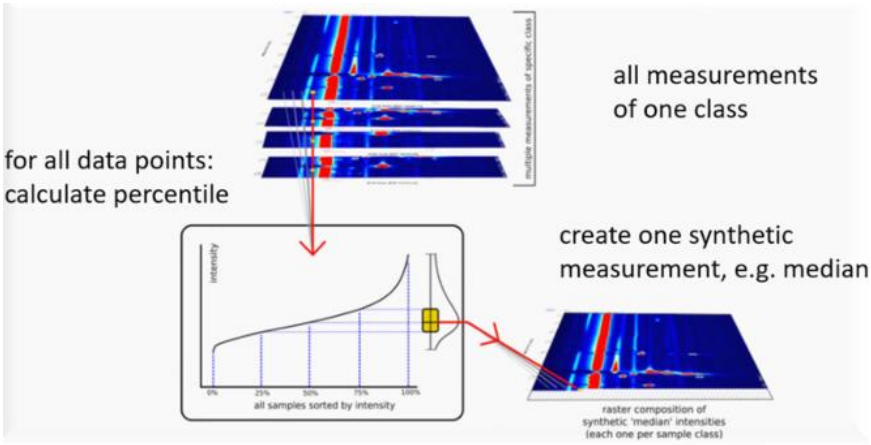


Figura 22: rappresentazione schematica del funzionamento del modulo Percentile.

Heracles II

Teoria

Heracles II è un gascromatografo costituito da 2 colonne cromatografiche metalliche capillari a polarità diverse: la prima è una colonna MXT-5 a bassa polarità, composta da 5% difenile e 95% dimetilpolisilossano, mentre la seconda è una colonna MXT-1701 a media polarità, composta da 14% cianopropilfenile e 86% dimetilpolisilossano (**Figura 23**) (Melucci, 2016). Le colonne sono collegate a due rilevatori a ionizzazione di fiamma alimentati a idrogeno (FID) e il carrier gas è idrogeno, che viene prodotto a partire da acqua milliQ. In **Figura 24** si riporta la rappresentazione del gas cromatografo Heracles II.

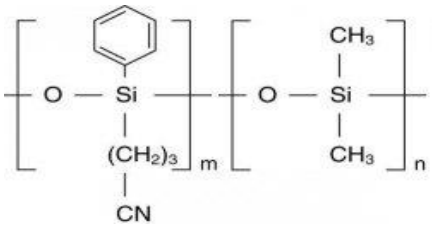


Figura 23: polimero cianopropilfenile/dimetil polisilossano (Optima® 624 colonne capillari per GC Ø int. 0,32 mm Ø est. 0,5 mm Spessore film 1,80 µm Lungh. 50 m, s.d.).

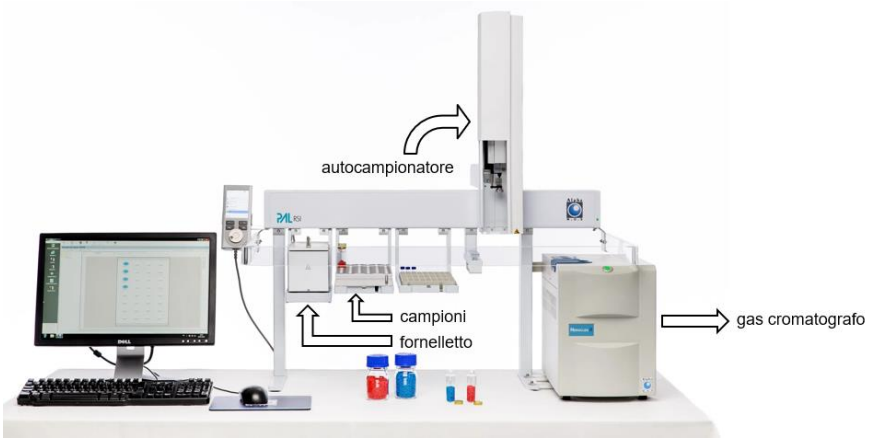


Figura 24: rappresentazione schematica del sistema con il gas cromatografo Heracles II con le sue componenti principali.

Procedura analitica

Materiali utilizzati:

- bilancia PG5002 DELTARANGE (Mettler-Toledo, Columbus, Ohio Stati Uniti d'America)
- vials da 20 mL (Cole Palmer, Vernon Hills, Illinois, Stati Uniti d'America)
- tappi Kinesis, Cole Palmer
- crimper Wheaton (DVK Life Sciences, Wertheim, Germania)
- Heracles II (Alpha MOS, Tolosa, Francia)

Procedura:

1. Preparare 4 vials da 20 mL per ciascun campione pesando (4,00 ± 0,02) g di riso in ciascuna di esse, e 4 vials da 20 mL vuote che serviranno come bianco. Tutte le vials vanno sigillate ermeticamente
2. Posizionare le vials nel rack di Heracles II, collocando i 4 bianchi all'inizio delle analisi, quindi inserire il rack all'interno del portacampione. Al termine delle analisi i 4 bianchi posti prima dei campioni verranno rianalizzati

Tabella 5: condizioni operative del gas cromatografo Hercales II.

Heracles II		
Parametri	Valori	Definizione
Injection volume	5 mL	Volume dello spazio di testa iniettato nel gas cromatografo

Injection speed	100 µL/s	Velocità di iniezione dei composti volatili nel gas cromatografo
Injection temperature	200 °C	Temperatura di iniezione dei composti volatili nel gas cromatografo
Incubation temperature	100 °C	Temperatura del fornelletto riscaldante
Trapping temperature	60 °C	Temperatura di intrappolamento dei composti volatili
Trapping duration	60 s	Tempo di intrappolamento dei composti volatili
Initial Oven Temperature	50 °C	Temperatura iniziale delle colonne cromatografiche
Final Oven Temperature	280°C	Temperatura finale delle colonne cromatografiche
Rate Oven Temperature	4.0 °C/s	Velocità di riscaldamento delle colonne cromatografiche
Time of analysis for a sample	21 min e 30 s	Tempo totale di analisi del campione, comprende il tempo di riscaldamento preliminare, il tempo dell'iniezione e della corsa dei composti

Elaborazione dati:

Il gas cromatografo Heracles II acquisisce un segnale analitico ogni 0,01 secondi; perciò, considerando il tempo dell'analisi (100 s) e che ogni analisi viene condotta utilizzando due colonne cromatografiche a polarità diversa, vengono generati due cromatogrammi grezzi per ogni campione per un totale di 20.002 punti. Per ridurre il costo computazionale e per facilitare il successivo trattamento statistico, i dati acquisiti vengono raggruppati mediando ogni 0,2 s, così da ottenere 2 cromatogrammi composti da 1001 punti totali. Questo consente di ridurre il numero dei dati e il rischio di *overfitting*, eliminando informazioni ridondati. L'overfitting è una condizione che può avvenire quando un modello di machine learning si adatta eccessivamente ai dati di

addestramento, non apprendendo regole generali e perdendo la capacità di predire le caratteristiche dei nuovi dati.

È stato utilizzato il software UnscramblerX per visualizzare i cromatogrammi ottenuti dalle analisi e per elaborare i dati utilizzando la tecnica della radice quadrata (**Figura 25**). Applicare la trasformazione con la radice quadrata ai dati cromatografici permette di equilibrare la differenza tra le intensità di ciascun picco, attenuando la predominanza dei picchi più elevati e valorizzando i picchi meno intensi, facendo in modo che anche le informazioni meno evidenti contribuiscano in modo non marginale all'analisi statistica. Infatti, non è detto che i picchi più intensi corrispondano necessariamente ai composti più determinanti per la discriminazione tra i diversi campioni.

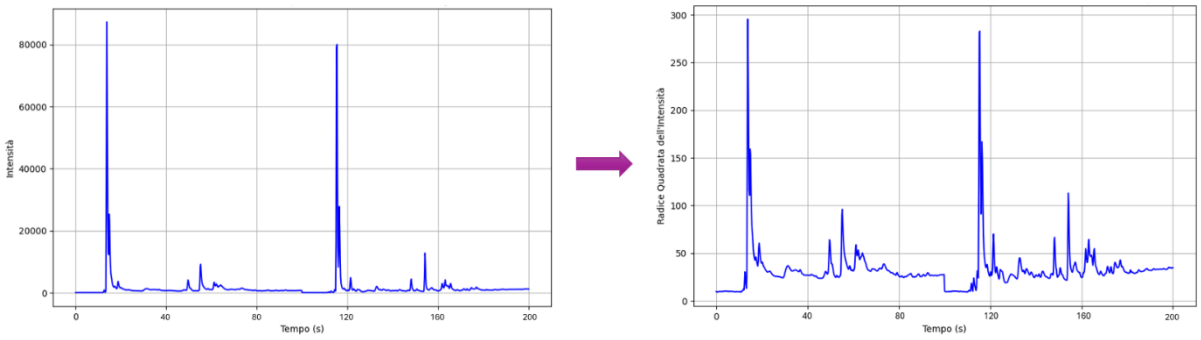


Figura 25: differenza delle intensità dei picchi cromatografici prima e doppo l'applicazione della radice quadrata. Sull'asse y si riportano le intensità dei picchi, sull'asse delle x i tempi di ritenzione dei composti.

Chemiometria

La chemometria è una branca della chimica che sfrutta metodi statistici, matematici, informatici e tecniche di statistica multivariata per la progettazione ottimale degli esperimenti. L'applicazione principale della chemiometria è rivolta ai problemi di natura chimica, con lo scopo di ottenere la maggior quantità possibile di informazioni rilevanti attraverso l'analisi dei dati ottenuti da determinate analisi chimiche (Heberger, 2008).

Analisi delle componenti principali:

I dati ottenuti dalle analisi mediante gascromatografia sono stati sottoposti ad analisi chemiometrica attraverso l'applicazione della statistica multivariata, ovvero la branca della statistica che tratta osservazioni composte da più di una componente. In particolare, è stata eseguita l'analisi delle componenti principali (PCA, *principal components analysis*) su tutte le misure effettuate sui campioni analizzati.

L'analisi delle componenti principali è un algoritmo matematico che riduce la dimensionalità dei dati preservando al contempo la maggior parte della variabilità del set di dati. Tale riduzione viene ottenuta individuando delle direzioni, denominate componenti principali, lungo le quali la variabilità dei dati è massima. A ciascuna componente principale è associata una varianza spiegata, definita come la quantità delle informazioni originali descritte da ciascuna componente principale. Le componenti principali sono nuove variabili che sono combinazioni lineari delle variabili originali e sono ortogonali tra loro. Utilizzando un numero ridotto di componenti principali, ciascun campione può essere rappresentato da pochi valori, anziché attraverso i valori di migliaia di variabili. È quindi possibile rappresentare graficamente i campioni, consentendo di valutare visivamente somiglianze e differenze tra di essi in modo più semplice e di stabilire se sia possibile raggrupparli secondo determinate caratteristiche.

La prima componente principale (PC1) è quella associata a varianza maggiore, seguita dalla seconda componente principale (PC2) e dalle altre in ordine decrescente di varianza. Per la rappresentazione grafica dei dati solitamente si considerano solo le prime due componenti principali, PC1 e PC2, omettendo le altre così da ridurre significativamente le variabili del dataset. In **Figura 26** si riporta l'esempio di un problema definito da due variabili originali che vengono trasformate nelle prime due componenti principali.

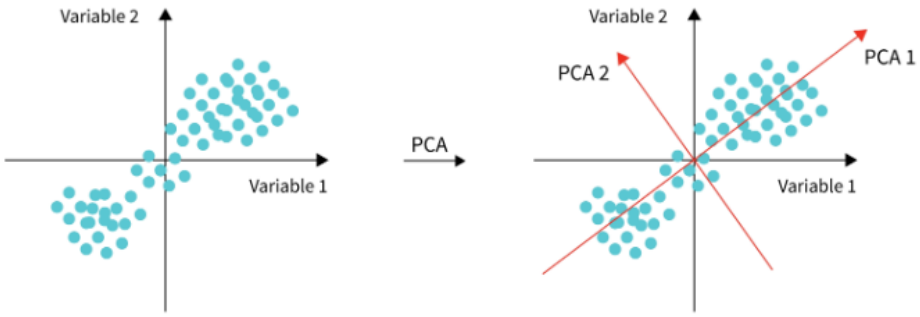


Figura 26: grafico variabile 2 vs variabile 1 a cui viene applicata una PCA, avviene una rotazione del sistema degli assi di riferimento, ottenendo un nuovo sistema di coordinate ortogonali (Principal Component Analysis (PCA), s.d.).

Calcolo delle componenti principali:

1. Standardizzazione delle variabili inziali: per garantire che ogni variabile abbia la stessa scala e contribuisca equamente all'analisi, vengono sottratti al dataset originale i valori medi delle variabili e si dividono questi valori per le deviazioni standard.
2. Calcolare la matrice di covarianza: la covarianza è un indice della correlazione di due o più variabili. Per capire come le combinazioni sono correlate bisogna guardare il segno delle variabili nella matrice: un valore positivo indica che le variabili aumentano/diminuiscono insieme, mentre un valore negativo vuol dire che, all'aumentare di una variabile, l'altra diminuisce. Un valore pari a zero indica che le variabili non sono correlate.
3. Calcolare gli autovettori e gli autovalori della matrice di covarianza: le componenti principali rappresentano, come autovettori, le direzioni della massima varianza nei dati, mentre gli autovalori rappresentano la quantità di varianza in ogni componente principale.
4. Selezionare le componenti principali: si scartano dall'analisi le componenti principali associate ad autovalori bassi
5. Trasformare i dati in un nuovo sistema di coordinate: i dati originali vengono proiettati dal vettore, creato dagli autovettori della matrice di covarianza, sui nuovi assi definiti dalle componenti principali (Cos'è l'analisi delle componenti principali (PCA)? - IBM, s.d.).

A seguito dell'analisi delle componenti principali si ottengono diversi grafici, che possono essere utilizzati per visualizzare i risultati ottenuti, tra cui lo *scores plot* e il *loading plot*.

Lo scores plot (**Figura 27**) è un grafico che riporta sugli assi le componenti principali con la relativa varianza percentuale in cui ogni punto rappresenta un campione. Tale grafico serve a raggruppare campioni con informazioni cromatografiche simili; infatti, punti più vicini tra loro suggeriscono campioni simili. Lo scores plot indica quindi come gli oggetti si distribuiscono rispetto alle nuove direzioni di massima varianza e permette di individuare possibili outliers o campioni anomali.

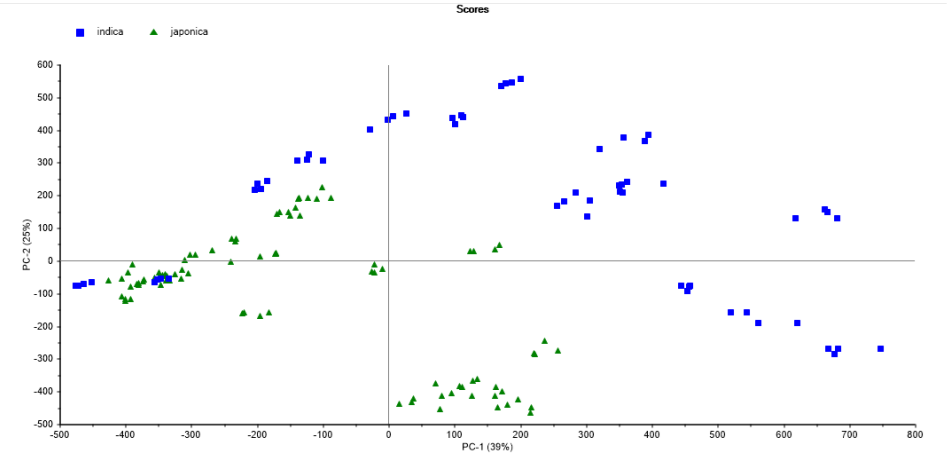


Figura 27: PCA eseguita su dei campioni di riso che divide i campioni Oryza sativa subsp. Indica (in blu) *dai* Oryza sativa subsp. Japonica (in verde) in base a determinate informazioni spettrali.

Il loadings plot (**Figura 28**) è un grafico che evidenzia il contributo di ogni variabile originale alla costruzione delle componenti principali. Ogni variabile è rappresentata da un punto la cui posizione mostra quanto è correlata con ciascuna componente principale e la relazione con le altre variabili nella stessa componente principale.

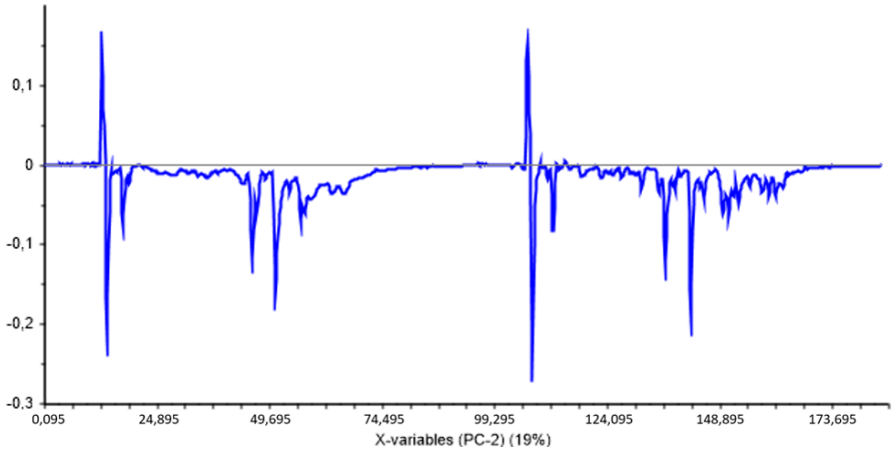


Figura 28: esempio di un grafico che riporta i loadings di ogni variabile in funzione della seconda componente principale

PLS-DA:

Il machine learning è il sottoinsieme dell'intelligenza artificiale che utilizza dati di addestramento per apprendere modelli tramite l'uso di algoritmi, con l'obiettivo di effettuare inferenze su nuovi set di dati, facendo previsioni senza avere istruzioni codificate. In questo studio è stato sviluppato un modello di classificazione dei campioni di riso, suddividendoli in basmati e non basmati, usando il metodo PLS-DA

(*Partial Least-Squares Discriminant Analysis*), che riduce la dimensionalità dell'oggetto dell'analisi mantenendo una buona capacità di classificazione. Sono stati scartati dall'analisi i campioni di riso basmati che eccedevano il limite di legge relativo alla quantità di riso non basmati al loro interno (7%) e i campioni di riso non basmati che presentavano una percentuale di riso non basmati al loro interno inferiore al 7%. La PLS-DA attua una separazione dei campioni basata sulle classi, le etichetta per massimizzare la separazione tra i vari gruppi attraverso l'estrazione di variabili latenti che colgono al meglio le varie differenze tra le classi (Classification with Multivariate Analysis: Focus on PLS-DA, s.d.). La variabile di categoria viene trasformata in una variabile numerica, che riesce a rappresentare il grado di appartenenza ad una classe su una scala compresa da 0 a 1 (0= totale diversità, 1= totale appartenenza alla classe in esame). La PLS-DA cerca quindi le direzioni lungo le componenti principali che portano alla migliore separazione tra le diverse classi.

In seguito alla creazione del modello, è stata eseguita la regressione dello stesso proiettando le miscele di riso, con lo scopo di ottenere una previsione quantitativa della percentuale di riso non basmati il più simile possibile a quella realmente presente nella miscela.

RISULTATI E DISCUSSIONE

Risultati biologia molecolare

L'analisi mediante dPCR permette di determinare il numero di copie per µl del DNA target, amplificato con il primer NoBa, per ogni campione analizzato, e il numero di copie per µl di DNA housekeeping, amplificato con il primer KMV. Dal rapporto tra questi due valori (**Equazione 4**) si ottiene la percentuale di riso non basmati presente all'interno del campione analizzato

Equazione 4: $$\%NB = \frac{\text{n° di copie/µl DNA target amplificato}}{\text{n° di copie/µl DNA housekeeping amplificato}}$$

Sono state preparate due tipologie di miscele a diverse percentuali, peso su peso, miscelando due varietà diverse di riso basmati con la stessa varietà di riso non basmati, come riportato nel paragrafo *campioni analizzati*. In **Tabella 6** vengono riportati i risultati del calcolo della percentuale di riso non basmati all'interno della miscela costituita dal campione di riso basmati 232 e il campione di riso non basmati 233, mentre in **Tabella 7** vengono riportati i risultati relativi alla miscela costituita dal campione di riso basmati 234 e il campione di riso non basmati 233.

In seguito al calcolo della percentuale di riso non basmati (% NB) in ciascuna miscela, è stata applicata la formula della propagazione dell'errore (**Equazione 5**), utilizzando gli intervalli di confidenza forniti dallo strumento. Sono stati così ottenuti i risultati finali, espressi come valore della percentuale di riso non basmati calcolata ± errore propagato.

Equazione 5: $$\sqrt{(CI)target^2+(CI)housekeeping^2}$$

In **Tabella 8** sono riportate le differenze assolute e le differenze percentuali tra i valori calcolati della percentuale di riso NB e i valori attesi determinati sulla base della composizione percentuale di riso non basmati presente in ciascuna la miscela.

Tabella 6: risultati ottenuti in seguito all'analisi dPCR della miscela composta dai campioni di riso 232/233. Si riportano il numero di copie/µl di DNA amplificato dai 2 primers, gli intervalli di confidenza e il risultato del calcolo della %NB

Campione	Target Cp/µl	Housekeeping Cp/µl	Target CI (95%)	Housekeeping CI (95%)	Intervallo di confidenza
NB00_232	0,847	72,07	49,00%	5,50%	1,2 ± 0,5
NB01_232	1,837	72,9	33,10 %	5,30%	2,5 ± 0,3
NB03_232	2,102	42,64	31,40 %	6,90 %	4,9 ± 0,3
NB05_232	3,03	54,02	26,40 %	6,20 %	5,6 ± 0,3
NB07_232	5,473	72,53	19,80 %	5,40 %	7,5 ± 0,2
NB10_232	7,366	49,19	17,00 %	6,50 %	15,0 ± 0,2
NB25_232	20,16	80,12	10,30 %	5,10 %	25,2 ± 0,1
NB50_232	30,13	57,18	8,40 %	6,20 %	52,7 ± 0,1
NB75_232	42,84	62,67	7,00 %	5,70 %	68,36 ± 0,09
NB100	117,9	142,8	4,30%	4,00%	82,56 ± 0,06

Tabella 7: risultati ottenuti in seguito all'analisi dPCR della miscela composta dai campioni di riso 234/233. Si riportano il numero di copie/µl di DNA amplificato dai 2 primers, gli intervalli di confidenza e il risultato del calcolo della %NB

Campione	Target Cp/µl	Housekeeping Cp/µl	Target CI (95%)	Housekeeping CI (95%)	Intervallo di confidenza
NB00_234	0,483	89,37	65,30%	5,00%	0,5 ± 0,7
NB01_234	1,042	79,25	43,80 %	5,10 %	1,3 ± 0,4
NB03_234	1,266	59,31	40,00 %	5,90 %	2,1 ± 0,4
NB05_234	4,415	82,17	21,50 %	5,10 %	5,4 ± 0,2
NB07_234	2,464	46,62	28,90 %	6,80 %	5,3 ± 0,3
NB10_234	7,032	102,2	17,30 %	4,60 %	6,9 ± 0,2
NB25_234	13,17	56,66	12,50 %	6,20 %	23,2 ± 0,1
NB50_234	33,81	78,67	7,80 %	5,30 %	43,0 ± 0,1
NB75_234	32,71	49,25	7,90 %	6,60 %	66,42 ± 0,10

Tabella 8: confronto dei risultati ottenuti per entrambe le miscele con la percentuale teorica di riso non basmati di cui è composta la miscela. Si riportano anche le differenze tra il rapporto ottenuto e il valore teorico e la differenza %, ottenuta dividendo la differenza per il valore atteso e moltiplicando per 100.

Campione	Rapporto numero di copie/µl T/H	Valore atteso	Differenza	Differenza %
NB00_232	1,2	0	-1,2	-
NB01_232	2,5	1	-1,5	-150
NB03_232	4,9	3	-1,9	-63
NB05_232	5,6	5	-0,6	-12
NB07_232	7,5	7	-0,5	-7
NB10_232	15,0	10	-5,0	-50
NB25_232	25,2	25	-0,2	-1
NB50_232	52,7	50	-2,7	-5
NB75_232	68,36	75	6,6	9
NB100	82,6	100	17,4	17
NB00_234	0,5	0	-0,5	-
NB01_234	1,3	1	-0,3	-30
NB03_234	2,1	3	0,9	30
NB05_234	5,4	5	-0,4	-8
NB07_234	5,3	7	1,7	24
NB10_234	6,9	10	3,1	31
NB25_234	23,2	25	1,8	7
NB50_234	43,0	50	7,0	14
NB75_234	66,42	75	8,6	11

È stata inoltre verificata la correlazione tra il valore percentuale di riso non basmati ottenuto e il corrispondente valore atteso, tracciando la retta di regressione lineare e calcolandone la pendenza e il coefficiente di determinazione. La regressione lineare relativa alla miscela 232/233 è riportata in **Figura 29**, mentre la regressione lineare della miscela 234/233 è riportata in **Figura 30**.

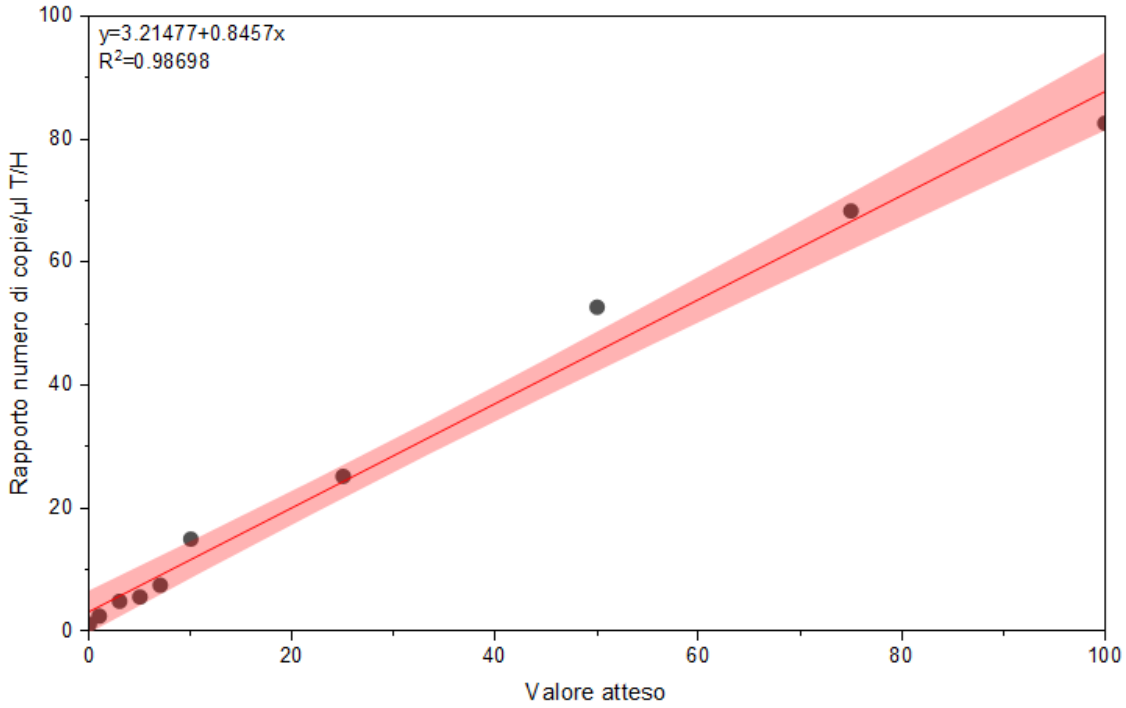


Figura 29: regressione lineare miscele riso 232/233, sull'asse delle ordinate è riportato il numero di copie/µl ottenuto in seguito ad analisi dPCR, sull'asse delle ascisse è riportato il valore atteso della percentuale di riso non basmati all'interno della miscela, si evidenziano anche la pendenza, il coefficiente di determinazione della retta e la banda di confidenza.

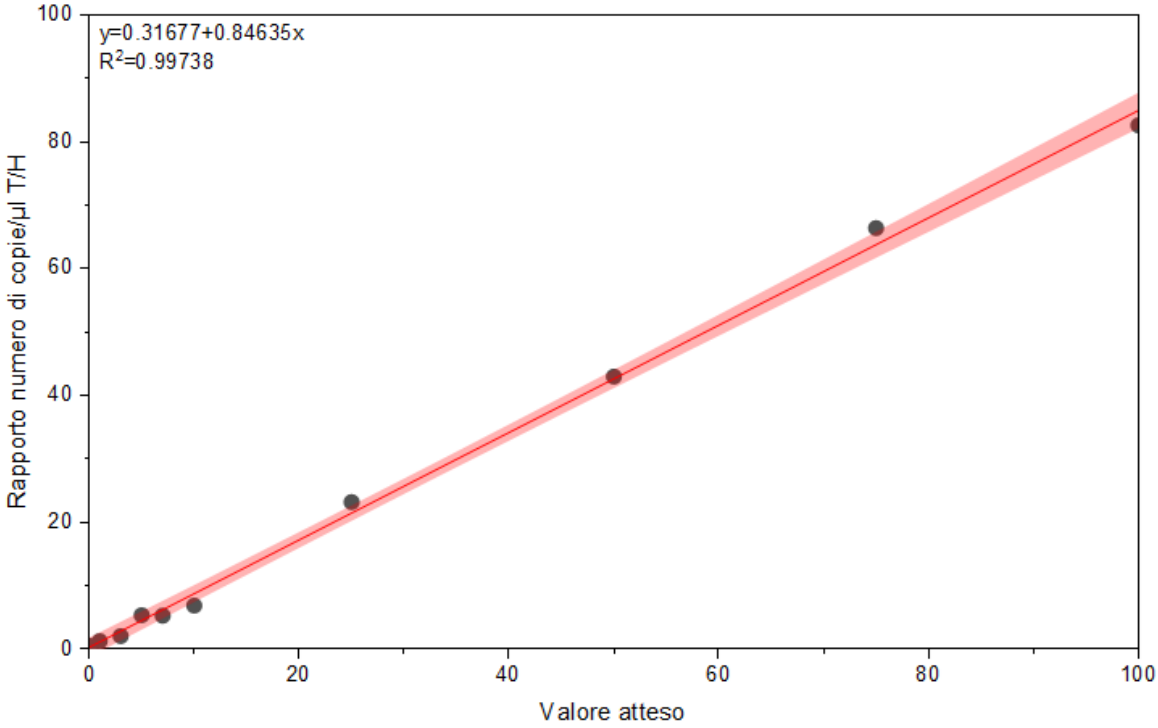


Figura 30: regressione lineare miscele riso 234/233, sull'asse delle ordinate è riportato il numero di copie/µl ottenuto in seguito ad analisi dPCR, sull'asse delle ascisse è riportato il valore atteso della percentuale di riso non basmati all'interno della miscela, si evidenziano anche la pendenza, il coefficiente di determinazione della retta e la banda di confidenza.

Dai dati ottenuti e dalle **Figura 29 e 30** si osserva una correlazione lineare tra la percentuale di riso non basmati determinata mediante le analisi genetiche e la composizione percentuale teorica presente all'interno delle miscele. Per entrambe le miscele, 232/233 e 234/233, si ottiene un coefficiente di determinazione che si discosta leggermente da 1; in particolare, la miscela 234/233 mostra una correlazione quasi perfettamente lineare.

La retta di regressione tracciata nel caso della miscela 232/233 tende a sovrastimare la composizione percentuale di riso non basmati presente nella miscela rispetto alla percentuale teorica di cui la miscela è composta fino alla miscela NB07. A partire da percentuali di riso non basmati più elevate, la correlazione tende invece a sottostimare la percentuale rispetto al valore teorico.

La retta di regressione relativa alla miscela 234/233 sovrastima la percentuale di riso non basmati nelle miscele NB00 e NB01 rispetto alla percentuale teorica, sottostimando invece i valori delle miscele contenenti più dell'1% di riso non basmati.

In **Figura 31** e in **Figura 32** si riportano, rispettivamente, le regressioni lineari delle miscele di riso 232/233 e 234/233, considerando le miscele composte da una percentuale di riso non basmati compresa tra 0% e 25%. Questa analisi è stata eseguita per verificare come cambiano le caratteristiche della retta di regressione considerando esclusivamente le miscele composte da basse percentuali di riso non basmati.

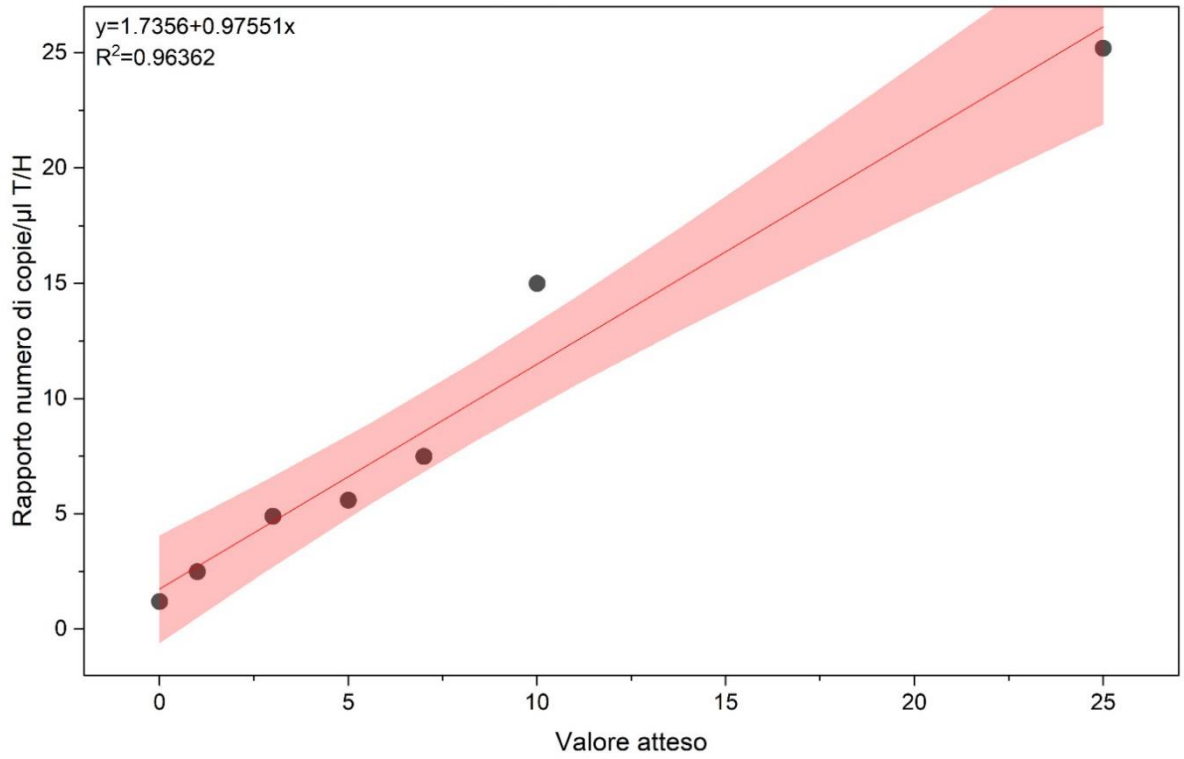


Figura 31: regressione lineare miscele riso 232/233 fino alla miscela NB25, sull'asse delle ordinate è riportato il numero di copie/µl ottenuto in seguito ad analisi dPCR, sull'asse delle ascisse è riportato il valore atteso della percentuale di riso non basmati all'interno della miscela, si evidenziano anche la pendenza, il coefficiente di determinazione della retta e la banda di confidenza.

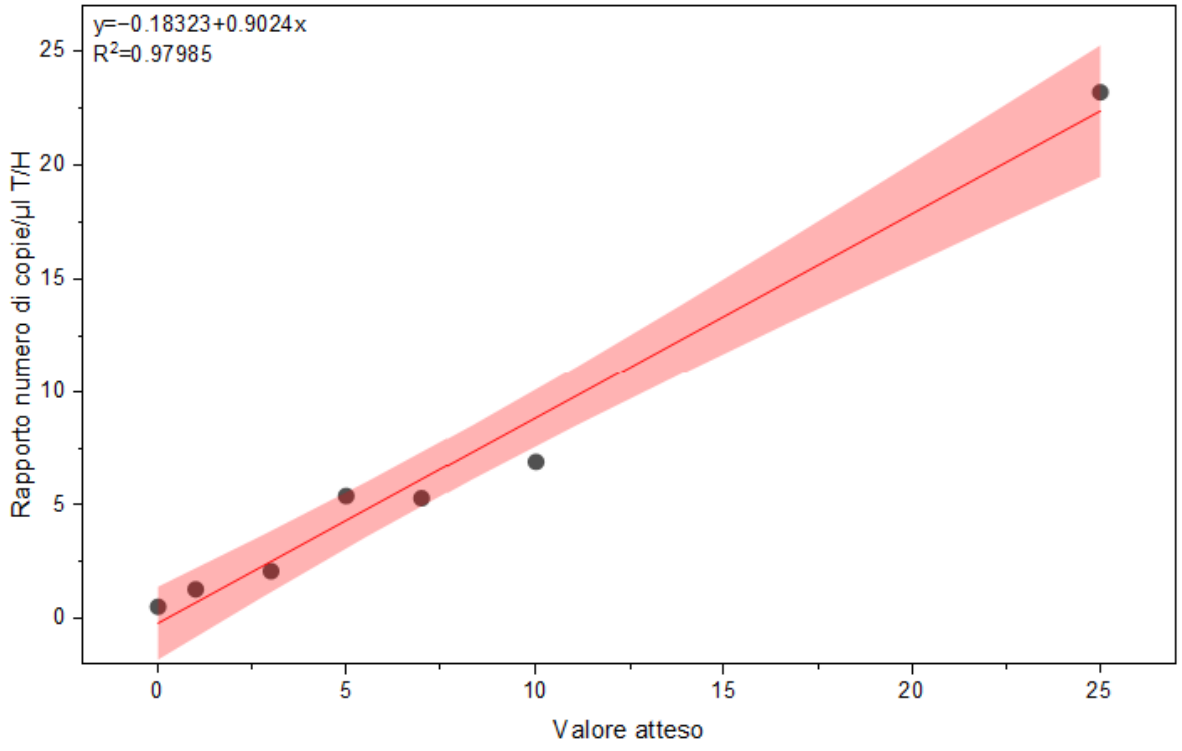


Figura 32: regressione lineare miscele riso 234/233 fino alla miscela NB25, sull'asse delle ordinate è riportato il numero di copie/µl ottenuto in seguito ad analisi dPCR, sull'asse delle ascisse è riportato il valore atteso della percentuale di riso non basmati all'interno della miscela, si evidenziano anche la pendenza, il coefficiente di determinazione della retta e la banda di confidenza.

Dalle **Figure 31** e **32** si nota come, alle percentuali più basse di riso non basmati nella miscela, il discostamento tra la percentuale teorica di NB e quella ottenuta dall'analisi sia maggiore rispetto a quello osservato con le percentuali più elevate di riso non basmati nella miscela. Si osserva, inoltre, come il coefficiente di determinazione diminuisca per entrambe le miscele quando la retta di regressione viene calcolata fino alla miscela NB25.

Per ottimizzare il metodo viene calcolato il limite di quantificazione (LOQ) eseguendo diluizioni seriali del DNA del campione di riso 92. Per determinare il LOQ si è stata calcolata la media (μ) e la deviazione standard (σ) delle repliche di ogni diluzione eseguita, dalle quali è stato ricavato il coefficiente di variazione (CV%) secondo l'**Equazione 6**. In **Tabella 9** si riportano il numero di copie/μL del DNA target ottenuto dall'analisi, il numero di copie/μL teoricamente atteso, la media, la deviazione standard e il CV% per ciascuna diluizione seriale.

Equazione 6: $$CV\% = \frac{\sigma}{\mu} * 100$$

Tabella 9: si riportano il numero di copie/ μL desiderato e quello ottenuto in seguito alle diluizioni seriali del campione di DNA del riso 92 con i primers NoBa, in seguito si sono calcolati la media, la deviazione standard e il CV%.

	numero repliche									
copie attese	1	2	3	4	5	6	7	media	SD	CV%
300	264.2	-	-	-	-	-	-	264,2	-	-
150	136,4	137,9	-	-	-	-	-	137,2	1,1	0,8
100	91,2	94,03	-	-	-	-	-	92,6	2,0	2,2
50	43,25	42,96	-	-	-	-	-	43,1	0,2	0,5
10	8,425	7,615	9,844	9,662	-	-	-	8,9	1,1	11,9
5	4,068	3,871	4,696	3,928	3,893	4,315	3,705	4,1	0,3	8,3
4	3,374	3,201	-	-	-	-	-	3,3	0,1	3,7
3	2,482	3,015	3,145	2,696	-	-	-	2,8	0,3	10,6
2	1,826	2,097	1,62	2,085	2,534	-	-	2,0	0,3	16,9
1	0,331	0,483	0,654	1,188	1,043	0,857	0,903	0,8	0,3	39,3

In **Tabella 10** si riportano il numero di copie/µL del DNA housekeeping amplificato ottenute dall'analisi, il numero di copie/µL teoricamente atteso, la media, la deviazione standard e il CV% per ogni diluizione seriale.

Tabella 10: si riportano il numero di copie/µL desiderato e quello ottenuto in seguito alle diluizioni seriali del campione di DNA del riso 92 con i primers KMV, in seguito si sono calcolati la media, la deviazione standard e il CV%.

	Numero repliche										
Copie attese	1	2	3	4	5	6	7	8	Media	SD	CV%
300	315,7	-	-	-	-	-	-	-	315,7	-	-
150	174,3	146,1	-	-	-	-	-	-	160,2	19,9	12,4
90	108	99,37	-	-	-	-	-	-	103,7	6,1	5,9
30	33,75	32,59	-	-	-	-	-	-	33,2	0,8	2,5
15	16,11	17,27	15,6	17,88	-	-	-	-	16,7	1,0	6,3
9	9,946	8,307	10,69	9,244	8,694	8,805	-	-	9,3	0,9	9,6
3	2,875	3,864	3,866	3,671	2,631	2,597	2,662	3,057	3,2	0,6	17,7
2	2,149	1,905	2,21	1,705	1,419	-	-	-	1,9	0,3	17,4
1	1,025	1,078	0,928	1,029	1,386	0,607	-	-	1,0	0,3	24,9

Dalle **Tabelle 9 e 10** si evince che il LOQ per il DNA target NoBa è pari a 2 copie/µL, mentre il LOQ del DNA housekeeping KMV è pari a 1 copie/µL.

Dopo aver ottimizzato il metodo, sono stati analizzati i 44 campioni di riso commerciali, 17 dei quali appartenenti alla varietà basmati. I risultati sono riportati in **Tabella 11**. In **Figura 33** è visibile il grafico a barre che illustra la percentuale di riso non basmati presente all'interno di tutti i campioni analizzati, evidenziando la differenza tra i campioni basmati e i non basmati ed il limite di legge, corrispondente al 7% di NB. In seguito alla determinazione della concentrazione di DNA, sono state eseguite diluizioni dei campioni puri al fine di ottenere una concentrazione finale di DNA pari a 2 ng/µL.

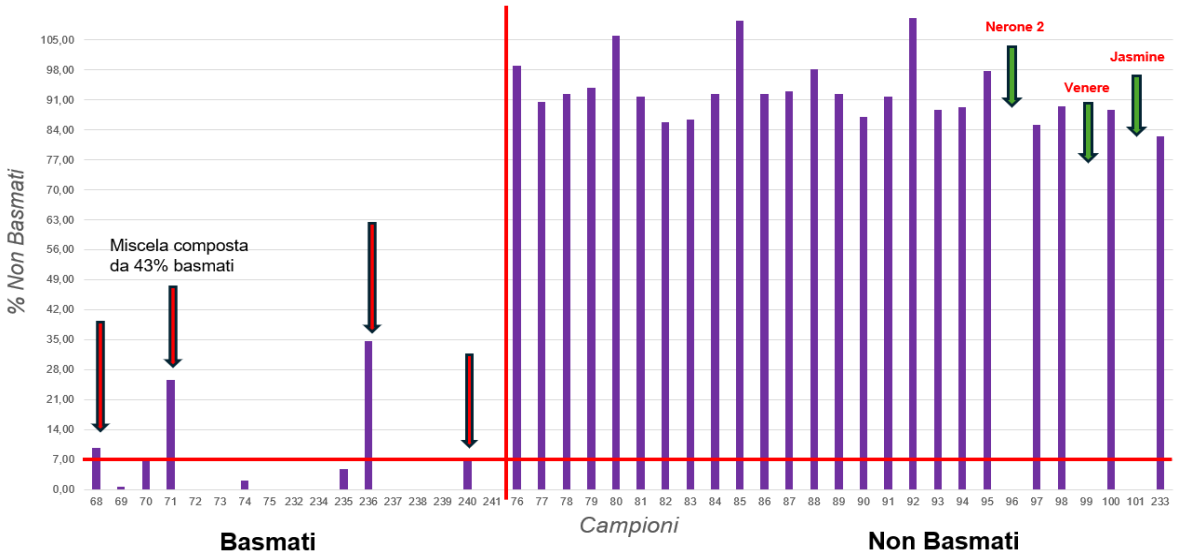


Figura 33: grafico delle percentuali di riso non basmati all'interno di tutti i campioni. A sinistra della barra verticale rossa si trovano i risultati dei risi basmati, a destra quelli dei risi non basmati; la barra orizzontale rossa evidenzia il limite di legge del 7%. Si evidenziano con frecce rosse i campioni basmati che hanno una percentuale di NB inferiore al LOQ, con le frecce verdi si evidenziano i campioni non basmati che hanno una percentuale di non basmati inferiore al 7%.

Tabella 11: si riportano il numero di copie/µL, l'intervallo di confidenza e i risultati del calcolo della percentuale di non basmati contenuto all'interno di tutti i campioni di riso commerciali analizzati

Tipologia	Campione	Target cp/µL	Housekeeping cp/µL	Target CI (95%)	Housekeeping CI (95%)	Intervallo di confidenza
Basmati	68	34,22	355,1	7,7%	2,6%	9,64 ± 0,08
Basmati	68	45,4	448,1	6,7%	2,3%	10,12 ± 0,07
Basmati	69	2,02	315,6	31,8%	2,7%	0,6 ± 0,3
Basmati	70	29,74	431,5	8,4%	2,4%	6,89 ± 0,09
Basmati	71	121,50	475,2	4,2%	2,3%	25,57 ± 0,05
Basmati	72	0,59	378,4	59,1%	2,5%	ASSENTE
Basmati	73	0,71	228,5	54,4%	3,1%	ASSENTE
Basmati	74	6,81	320,9	17,4%	2,7%	2,1 ± 0,2
Basmati	75	1,11	244,9	42,8%	3,0%	ASSENTE
Basmati	232	0,847	72,07	49,0%	5,50%	ASSENTE
Basmati	234	0,483	89,37	65,3%	5,00%	ASSENTE
Basmati	235	3,434	71,21	24,7%	5,60%	4,8 ± 0,3
Basmati	236	24,49	70,59	9,2%	5,40%	34,7 ± 0,1
Basmati	237	1,645	87,7	35,8%	4,90%	ASSENTE
Basmati	238	0,321	79,69	80,0%	5,10%	ASSENTE

Basmati	239	1,593	51,74	36,4%	6,30%	ASSENTE
Basmati	240	6,498	88,1	18,0%	4,90%	7,4 ± 0,2
Basmati	241	0,159	34,85	106%	7,70%	ASSENTE
Carnaroli	76	304,90	307,7	2,8%	2,7%	99,09 ± 0,04
Carnaroli	77	281,70	311,2	2,9%	2,7%	90,52 ± 0,04
Carnaroli	78	306,40	331,5	2,8%	2,7%	92,43 ± 0,04
Carnaroli	79	337,30	359,2	2,6%	2,6%	93,90 ± 0,04
Carnaroli	80	333,3	314,3	2,6%	2,7%	106,05 ± 0,04
Carnaroli	81	320,7	376,7	2,7%	2,5%	85,13 ± 0,04
Carnaroli	81	373,8	407,4	2,5%	2,4%	91,75 ± 0,03
Carnaroli	82	308,3	359,4	2,8%	2,6%	85,78 ± 0,04
Carnaroli	83	225,2	260,9	3,4%	2,9%	86,32 ± 0,04
Carnaroli	84	283,4	306,9	2,9%	2,8%	92,33 ± 0,04
Arborio	85	339,5	310,1	2,8%	2,7%	109,50 ± 0,04
Arborio	86	330,0	357,6	2,7%	2,6%	92,29 ± 0,04
Ribe	87	414,3	445,1	2,4%	2,3&	93,08 ± 0,03
Ribe	88	326,6	332,5	2,7%	2,6%	98,24 ± 0,04
Vialone nano	89	295,5	320,0	2,8%	2,8%	92,35 ± 0,04
Vialone nano	90	265,1	304,5	3,0%	2,9%	87,08 ± 0,04
Volano	91	290,0	315,9	2,8%	2,8%	91,81 ± 0,04
Baldo	92	344,0	312,6	2,6%	2,7%	110,04 ± 0,04
Baldo	93	340,0	383,0	2,7%	2,5%	88,77 ± 0,04
Baldo	94	361,0	403,8	2,6%	2,4%	89,40 ± 0,04
Originario	95	411,8	420,9	2,5%	2,4%	97,84 ± 0,03
Nerone 2	96	0,5	289,7	62,0%	2,8%	ASSENTE
Lungo B	97	392,5	461,3	2,5%	2,3%	85,09 ± 0,03
Lungo B	98	348,5	389,7	2,6%	2,5%	89,43 ± 0,04
Venere	99	1,9	134,4	33,6%	1,7%	ASSENTE
Lungo A	100	324,7	365,8	2,7%	2,6%	88,76 ± 0,04
Jamin	101	1,3	319,6	40,0%	2,7%	ASSENTE
Lungo B	233	117,9	142,8	4,30%	4,00%	82,56 ± 0,06

Dalla **Figura 33** si nota che, delle 17 varietà di riso basmati analizzate, 13 presentano una percentuale di non basmati inferiore al limite di legge, mentre 4 eccedono la soglia del 7%. Tra questi, il campione di riso 71 è dichiarato in etichetta essere una miscela contenente il 43% di riso basmati. Per i campioni che sono risultati contenere una percentuale di riso NB inferiore al LOQ, l'analisi si può considerare solo di tipo qualitativo. I campioni non basmati mostrano percentuali di riso non basmati al proprio interno molto superiori rispetto ai campioni basmati, ad eccezione di tre campioni, che corrispondono al riso Nerone 2, al riso Venere e al riso Jasmine, evidenziati con le frecce rosse. Dalla letteratura (Nader, 2026) è noto che il riso Venere e Jasmin possiedono la stessa delezione genica associata al riso basmati; quindi, sono entrambe varietà profumate. Dal registro varietale per l'annata agraria 2025/2026 rilasciato dall'ente Ente Nazionale Risi (ENTERISI) (Risi, 2025), risulta che il riso Nerone 2 possiede il pericarpo nero aromatico o profumato naturalmente. Questo potrebbe spiegare la bassa percentuale di riso non basmati rilevata al suo interno, anche se in letteratura non sono presenti articoli che confermino la presenza della delezione nel gene BADH2 che causa il profumo e il sapore tipici del riso basmati.

Risultati analisi chimica untargeted

Risultati FlavourSpec25®

Identificazione VOCs:

Sono stati analizzati i 44 campioni precedentemente analizzati con la tecnica dPCR con il gas cromatografo FlavourSpec25® per verificare la possibilità di riconoscere un riso basmati da uno non basmati tramite la correlazione tra il profilo dei picchi dei composti volatili con la quantificazione molecolare della tecnica biologica.

Inizialmente è stata effettuata una parte esplorativa dei picchi cromatografici attraverso gli indici di Kovatz: nel databank Arochembase è stato trovato che la molecola 2-acetil-1-pirrolina ha un tempo di ritenzione pari a 400 secondi (indice di Kovatz= 922). In **Figura 34** si riportano i cinquantesimi percentili dei cromatogrammi dei risi basmati e non basmati.

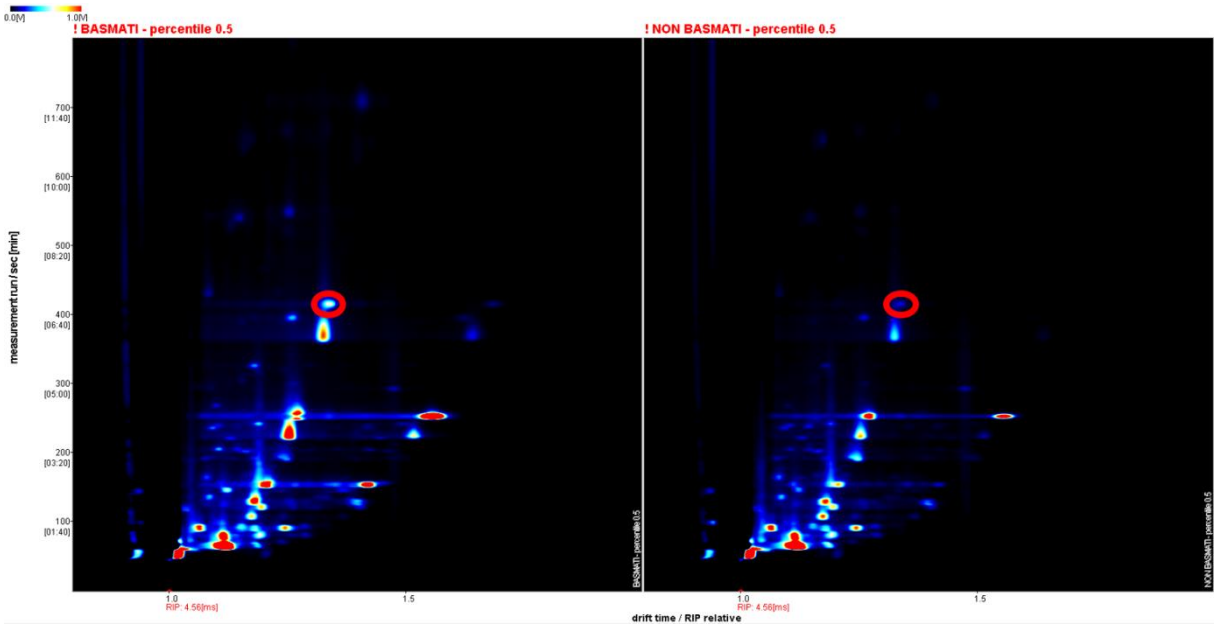


Figure 34: a sinistra si riporta il 50° percentile dei campioni di riso basmati, a destra il 50° percentile dei campioni di riso non basmati, in entrambi i cromatogrammi sintetici viene evidenziato il probabile picco corrispondente alla 2-acetil-1-pirrolina, identificato attraverso gli indici di Kovatz.

Dai cromatogrammi si nota la presenza di un picco cromatografico a tempo di ritenzione pari a 400 s, che risulta più intenso nei campioni basmati rispetto ai campioni non basmati. Anche se non si può affermare con certezza, il picco viene identificato come corrispondente alla 2-acetil-1-pirrolina, che da letteratura risulta sovraespressa nelle varianti basmati del riso. Si nota come i picchi dei 2 cromatogrammi siano simili, questo suggerisce che i marker delle 2 tipologie di riso siano gli stessi, ma abbiano intensità diverse.

In seguito all'identificazione della 2-acetil-1-pirrolina, gli indici di Kovatz sono stati utilizzati per l'identificazione degli altri picchi più intensi del cromatogramma: in **Figura 35** si riportano i nomi dei composti identificati nel cromatogramma in cui è stato eseguito il percentile 1. Si nota che ogni composto viene rilevato due volte agli stessi tempi di ritenzione, ma a drift time differenti, poiché a drift time maggiori si trova il dimero del composto analizzato. Se la concentrazione del composto voltile è elevata, si formano più ioni e quindi è più probabile che si formino cineticamente dei cluster protonati.

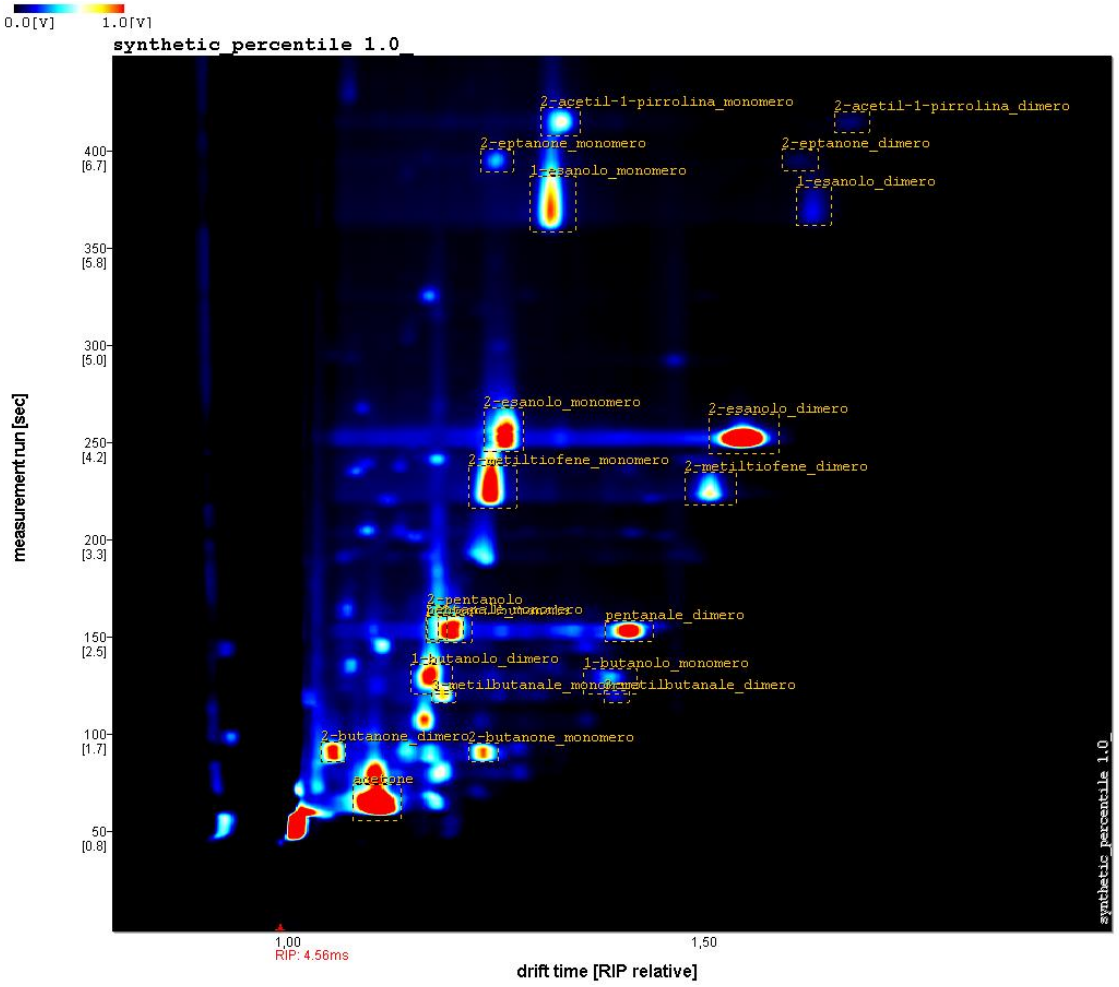


Figura 35: picchi cromatografici più rappresentativi del fingerprinting dei composti volatili di un campione di riso, identificati mediante gli indici di Kovats.

Da letteratura emerge che gli alcoli siano probabilmente prodotti da una reazione secondaria dell'ossidazione degli acidi grassi insaturi, formandosi in seguito alla rottura delle aldeidi. Questo spiegherebbe la presenza marcata in **Figura 35** dei picchi relativi all'1-esanolo, al 2-esanolo, all'1-butanolo, al 2-pentanolo, al 3-metil butanale e al penatale. Gli alcoli sono i secondi maggiori composti presenti nel riso cucinato, in particolare risulta che l'esanolo sia l'alcol maggiormente presente nel riso, soprattutto nelle varietà basmati, contribuendo al sapore. Risulta anche che i tiofeni ed altri composti eterociclici siano i maggiori composti volatili che vengono prodotti dalla reazione di Maillard, questo spiega la presenza del 2-metil tiofene. La reazione di Maillard consiste in una serie di reazioni a cascata che iniziano con la reazione di un gruppo amminico e di uno zucchero riducente e formano melanoidine, ovvero macromolecole che conferiscono al cibo cotto il caratteristico colore, odore e sapore.

(Hu L. G., 2020). Il 2-eptanone è un composto che si genera dalla decomposizione dell'acido oleico, che costituisce il 36% degli acidi grassi nei chicchi di riso. (Zaplin, 2013). Da letteratura è noto come l'acetone sia uno dei composti volatili presenti nel riso (Sun, 2022).

PCA:

È stata applicata la statistica multivariata eseguendo l'analisi delle componenti principali al dataset ottenuto dalle analisi gascromatografiche (**Figura 36**), rimuovendo i dati cromatografici dei campioni basmati che sono risultati da analisi biologica contenere più del 7% di riso non basmati al proprio interno e i campioni non basmati che risultano "profumati", quindi, aventi meno del 7% di riso non basmati al proprio interno. È stato così ottenuto un nuovo dataset composto da 37 campioni, di cui 13 di riso basmati e 24 di riso non basmati. Dal grafico degli scores plot è possibile osservare una separazione abbastanza netta lungo la PC-1 e la PC-2 tra i campioni di riso basmati, evidenziati in viola, e i campioni di riso non basmati, evidenziati in verde, dimostrando una marcata separazione tra risi basmati e non basmati lungo le prime due componenti principali, confermando la presenza di differenze sostanziali nel profilo dei composti volatili.

In **Figura 37** si riporta il grafico dei loadings plot, che mostra l'intensità dei loadings di ogni composto principale rilevato nei cromatogrammi; per semplificare il grafico si riportano i composti rilevati sotto forma di dimeri e non i monomeri. L'esame del *loadings plot* mostra che la maggior parte dei composti identificati contribuisce positivamente alla caratterizzazione dei campioni basmati, mentre acetone e 2-butanone risultano associati prevalentemente ai campioni non basmati.

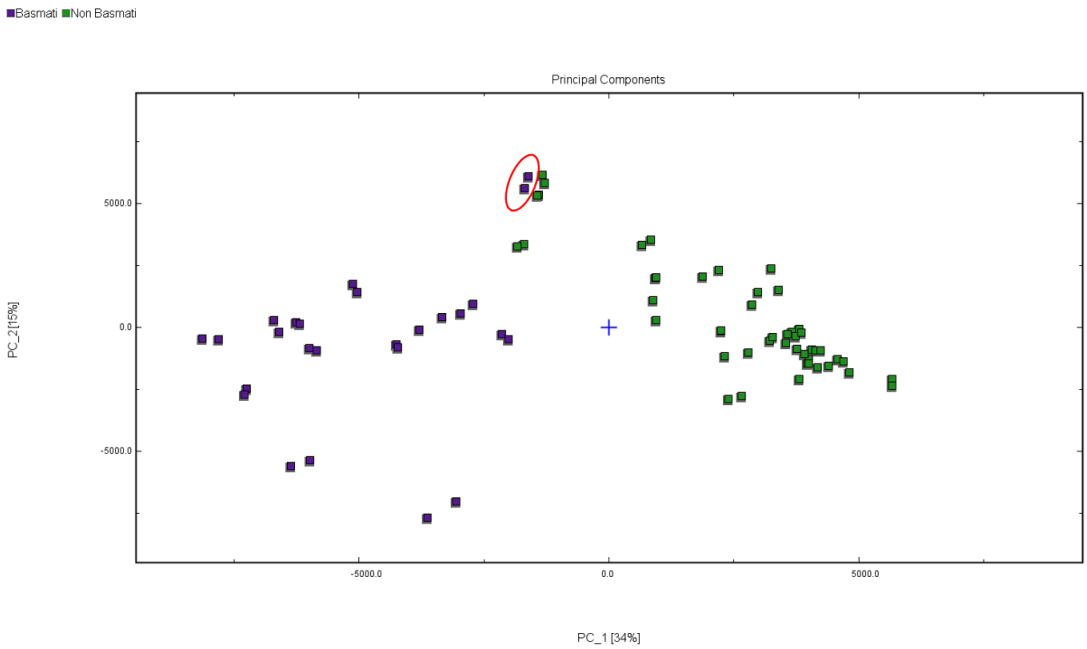


Figura 36: scores plot relativo ai dati cromatografici ottenuto in seguito ad analisi GC-IMS. Ogni punto indica un campione di riso, in viola sono riportati i campioni basmati e in verde i campioni non basmati. Si mette in evidenza il riso 70 cerchiato in rosso, che dall'analisi biologica risulta contenere una percentuale di riso non basmati leggermente sotto il limite di soglia 7%.

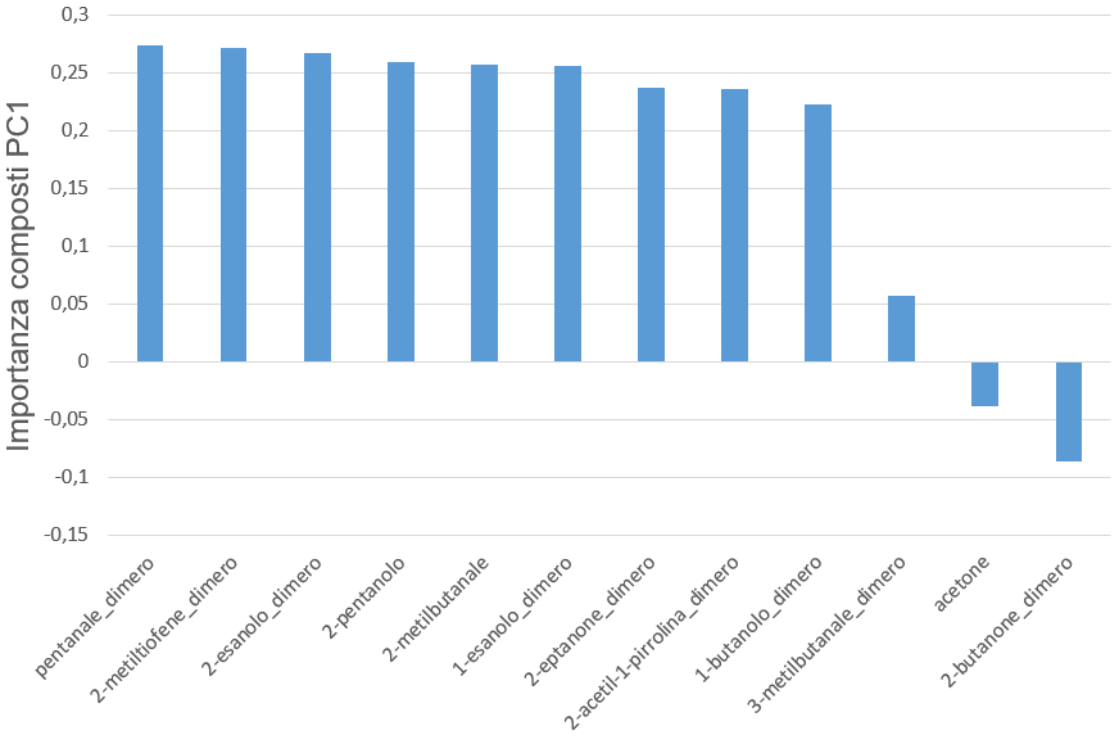


Figura 37: grafico dei loadings riferito ai composti più intensi del cromatogrammi, si omettono i monomeri per semplificare la trattazione.

PLS-DA:

Il modello è stato creato utilizzando il metodo PLS-DA per la classificazione del riso in basmati e in non basmati, rimuovendo i campioni anomali basmati che hanno % NB > 7% e i non basmati che hanno % NB assente. Si riportano i risultati del modello del training set in **Tabella 12** e la matrice di confusione in calibrazione per i risi basmati in **Tabella 13**. La matrice di confusione è una matrice che aiuta la valutazione delle prestazioni di un modello di classificazione nel machine learning, riporta nelle colonne i valori effettivi di un set di dati (nel nostro caso i risi effettivamente basmati e non basmati) e nelle righe i valori predetti di un set di dati. Per il training set sono state usate l'80% delle misure effettuate (60 misure), si riporta solo la matrice di confusione in calibrazione dei campioni basmati siccome è equivalente alla matrice di confusione in validazione dei campioni basmati e alle matrici di confusione in validazione e calibrazione dei campioni non basmati. Per la creazione del modello è stata eseguita una full cross validation-1, ovvero una tecnica statistica usata per la validazione del modello in cui il dataset viene diviso in una parte che serve per l'addestramento del modello (training set) e una parte che serve per testarlo e validarlo (test set).

Tabella 12: risultati del modello del training set in calibrazione e validazione

	True positive	False positive	True negative	False negative	Specificità	Sensibilità	Accuratezza
Cal	21	0	39	0	1	1	1
CV	21	0	39	0	1	1	1

Tabella 13: matrice di confusione in calibrazione dei campioni di riso basmati

	Basmati	Non basmati
Basmati	21	0
Non basmati	0	39

Dalla **Tabella 12** e dalla matrice di confusione si nota come tutti i campioni vengano descritti correttamente sia in validazione che in calibrazione, sia per i campioni basmati che non basmati; infatti, si hanno valori di specificità (**Equazione 7**), sensibilità (**Equazione 8**) e accuratezza (**Equazione 9**) uguali ad 1.

Equazione 7: $$\text{Specificità} = \frac{TN}{TN+FP}$$

Equazione 8: $$\text{Sensibilità} = \frac{TP}{TP+FN}$$

Equazione 9: $$\text{Accuratezza} = \frac{TN+TP}{TP+\ TN+FP+FN}$$

Il modello è stato creato scegliendo 14 componenti, ovvero le nuove variabili sintetiche create per riassumere l'informazione contenuta nei dati di partenza. La scelta è stata effettuata grazie al grafico riportato in **Figura 38**, che riporta il *root mean square error*, ovvero la differenza media tra i valori predetti dal modello e i valori reali in funzione del numero delle nuove componenti. Si nota che scegliendo 14 componenti si ha un punto di minimo locale, ovvero non si incorre in una situazione in cui l'errore sia troppo alto né nella situazione in cui il numero delle componenti sia troppo elevato aumentando molto la complessità del modello e andando incontro al rischio di overfitting. In **Figura 39** si riporta il grafico riguardante i Variable Importance in Projection (VIP), in cui ogni picco rappresenta un composto volatile, le altezze dei picchi fanno riferimento al contributo alla discriminazione tra riso basmati e non basmati; i composti con VIP > 1 sono quelli maggiormente rilevanti per la classificazione.

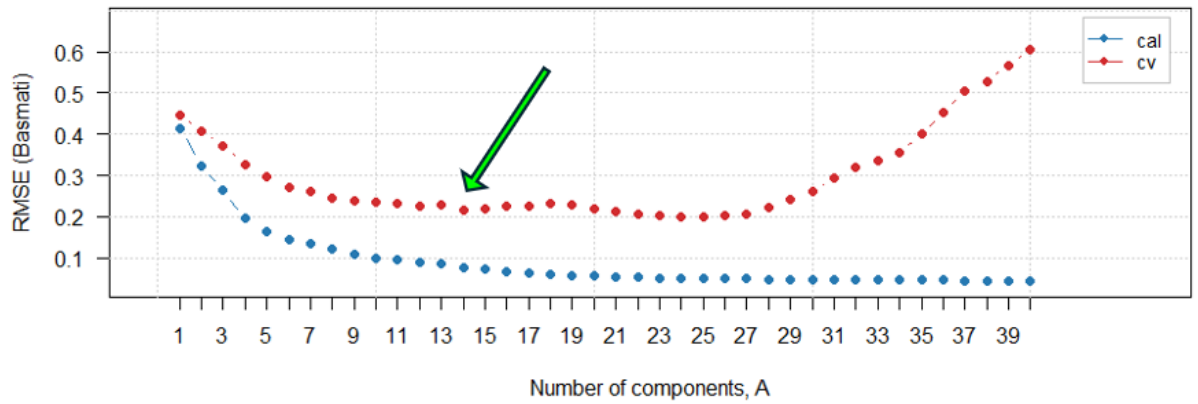


Figura 38: grafico root mean square error vs numero di componenti, si evidenzia con la freccia verde il minimo locale utilizzato per la scelta del numero di componenti, scegliendo più componenti il modello diventa sempre più complesso con una diminuzione dell'errore minima.

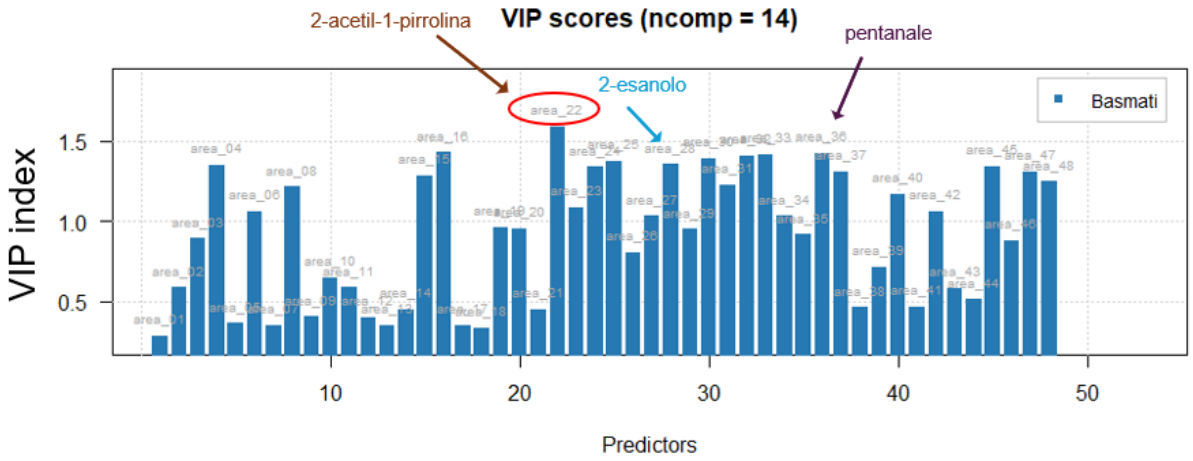


Figura 39: valori VIP (Variable Importance in Projection) del modello PLS-DA sviluppato sui dati GC-IMS (FlavourSpec25®). Per ciascun composto volatile è riportato il contributo alla discriminazione tra riso basmati e non basmati.

Dal grafico in **Figura 39** si nota come l'area più significativa per la discriminazione tra riso basmati e non basmati sia quella relativa alla 2-acetil-1-pirrolina, come da letteratura scientifica, oltre a quella del 2-esanolo e del pentanale.

Il restante 20% delle misure (14 misure) è stato utilizzato per la predizione dei campioni basmati e non basmati con il modello appena creato: si riporta la matrice di confusione (**Tabella 14**) e il grafico delle predizioni dei campioni basmati e non basmati (**Figura 40**).

Tabella 14: matrice di confusione in predizione dei campioni di riso basmati e non basmati

	Basmati	Non basmati
Basmati	5	0
Non basmati	0	9

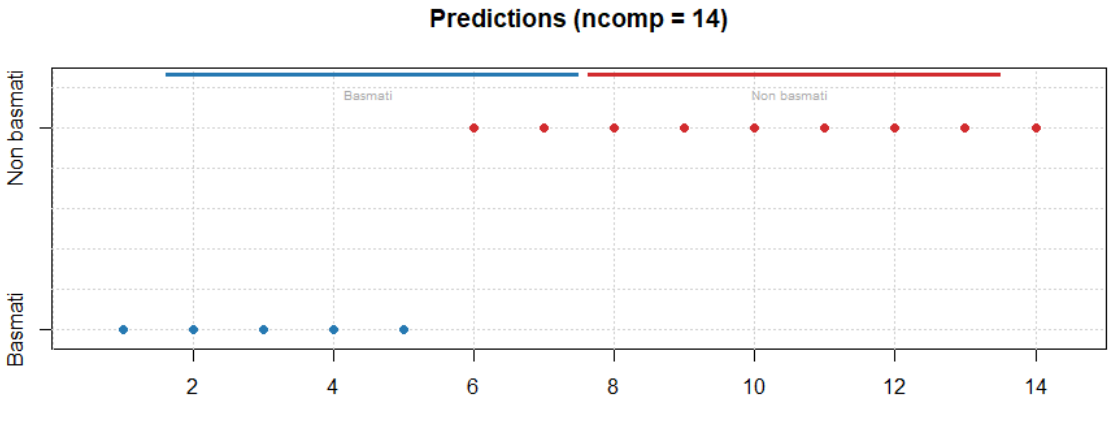


Figura 40: grafico predizioni dei campioni di riso, riporta sull'asse y la tipologia predetta dal modello e con colori diversi la tipologia reale dei campioni di riso, ogni punto corrisponde ad un campione di riso.

Dalla matrice di confusione in predizione e dal grafico in **Figura 40** si evince come tutti i campioni di riso vengano predetti correttamente, confermando la robustezza e la bontà del modello di classificazione creato.

Regressione del modello: vengono proiettate le miscele di riso comprese le miscele NB00 e NB100 applicando il modello creato, si riporta in **Figura 41** il risultato della proiezione. Si nota come vengano predette come non basmati le miscele NB75 e NB 100, quindi quelle contenenti riso non basmati in maggioranza, mentre le miscele a maggioranza basmati e 50% basmati e 50% non basmati vengano predette come basmati. In seguito, si è provato a predire la percentuale precisa di riso non basmati all'interno delle miscele usando 4 componenti, si riporta la correlazione lineare **Figura 42**.

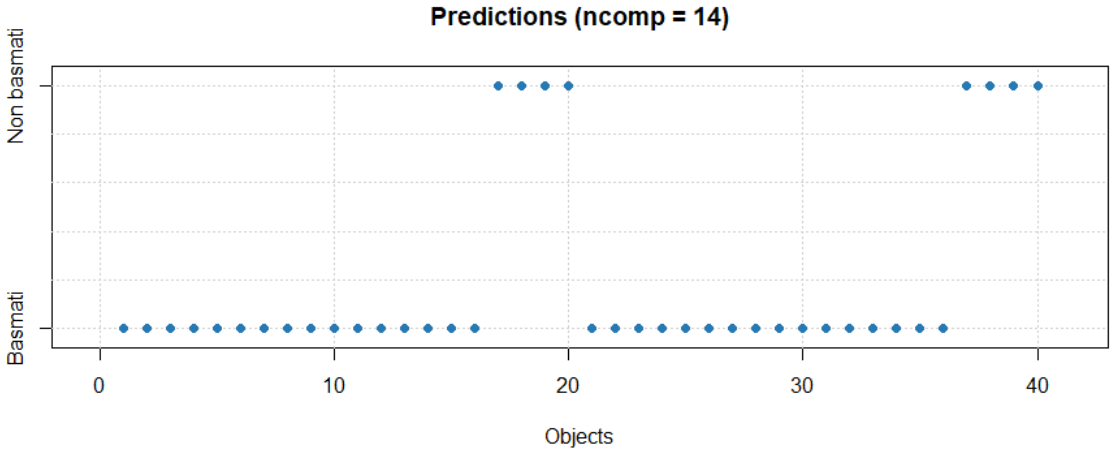


Figura 41: grafico predizioni delle miscele di riso, sull'asse y è riportata la tipologia predetta dal modello, ogni punto corrisponde ad una miscela di riso, vengono predette come non basmati solo le miscele NB75 e NB100.

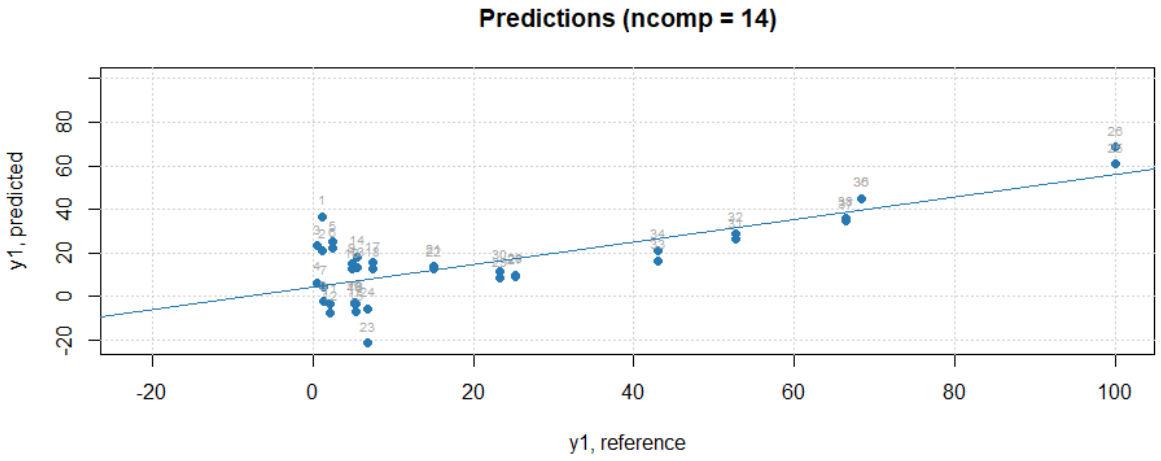


Figura 42: percentuale di riso non basmati predetta dal modello PLS-DA sviluppato sui dati GC-IMS (FlavourSpec25 ®) in funzione della percentuale reale nelle miscele sperimentali. In ascissa è riportata la percentuale reale di riso non basmati, in ordinata quella predetta.

Si nota dalla **Figura 42** come i valori predetti delle percentuali di non basmati all'interno delle miscele discostino parecchio dai valori reali, soprattutto per le miscele contenenti percentuali alte e basse di riso non basmati, dimostrando una scarsa capacità predittiva del modello. Si ottiene una retta con pendenza 0.519 e coefficiente di determinazione= 0.545.

Risultati Heracles II

Identificazione VOCs:

Si ripete il procedimento eseguito con FlavourSpec25® con il gascromatografo Heracles II, analizzando i 44 campioni precedentemente analizzati con la tecnica dPCR.

Per elaborare gli output dei composti volatili è stato utilizzato il software UnscramblerX e sui cromatogrammi ottenuti è stata applicata la radice quadrata, ottenendo dei cromatogrammi bidimensionali. Si effettua la parte esplorativa dei picchi utilizzando gli indici di ritenzione di Kovatz, nel databank Arochembase si trova che la molecola 2-acetil-1-pirrolina ha un tempo di ritenzione pari a 34-36 s nella colonna cromatografica a bassa polarità (indice di Kovatz= 922) e 137-138 s nella colonna cromatografica ad alta polarità (indice di Kovatz= 1015). In **Figura 43** si riporta il cromatogramma raw ottenuto dall'analisi dei campioni di riso basmati e non basmati.

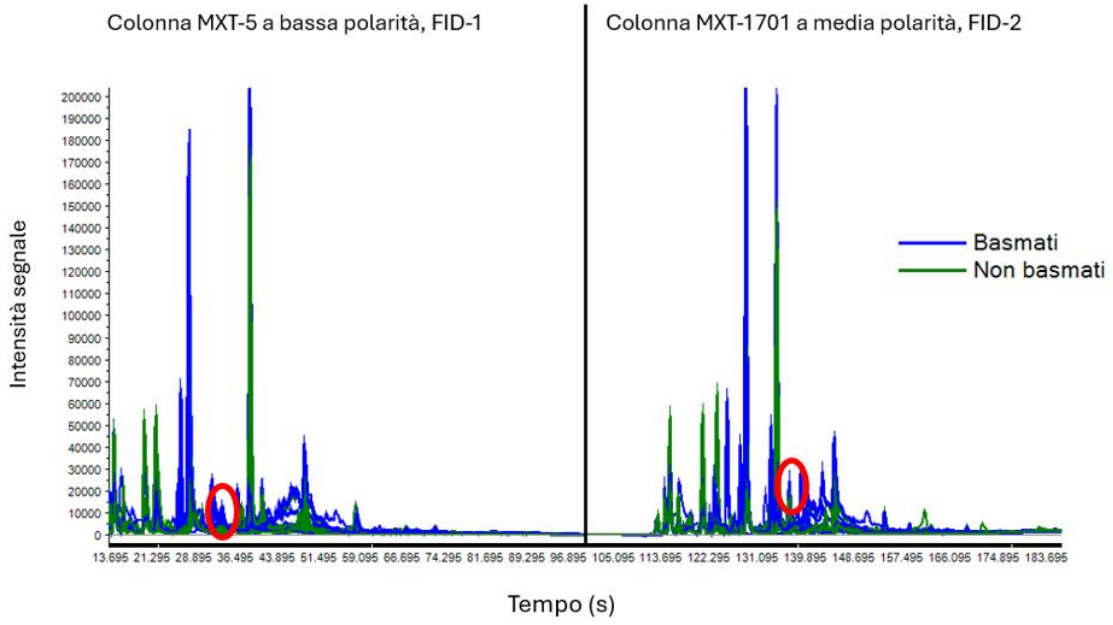


Figura 43: cromatogramma ottenuto dall'analisi di tutti i campioni di riso commerciali, in rosso viene evidenziato il picco più probabilmente associabile alla 2-acetil-1-pirrolina. Il cromatogramma è diviso in 2 parti dalla linea nera, corrispondenti ai 2 spettri ottenuti dalle 2 colonne cromatografiche a polarità diversa.

Dal cromatogramma in **Figura 43** si notano 2 picchi poco intensi, uno che cade a t= 34.20 s presente solo nei risi basmati e l'altro che cade a t= 137.3 s, più intenso nei basmati ma presente anche nei non basmati. Questi picchi possono essere associati con una buona probabilità alla presenza di 2-acetil-1-pirrolina.

In seguito all'identificazione dei picchi probabilmente corrispondenti alla 2-acetil-1-pirrolina è stato utilizzato il databank Arochembase per associare ai picchi più intensi dei probabili composti, utilizzando gli indici di ritenzione di Kovatz. In **Figura 44** si riporta il cromatogramma dei campioni di riso puri a cui è stata applicata la radice

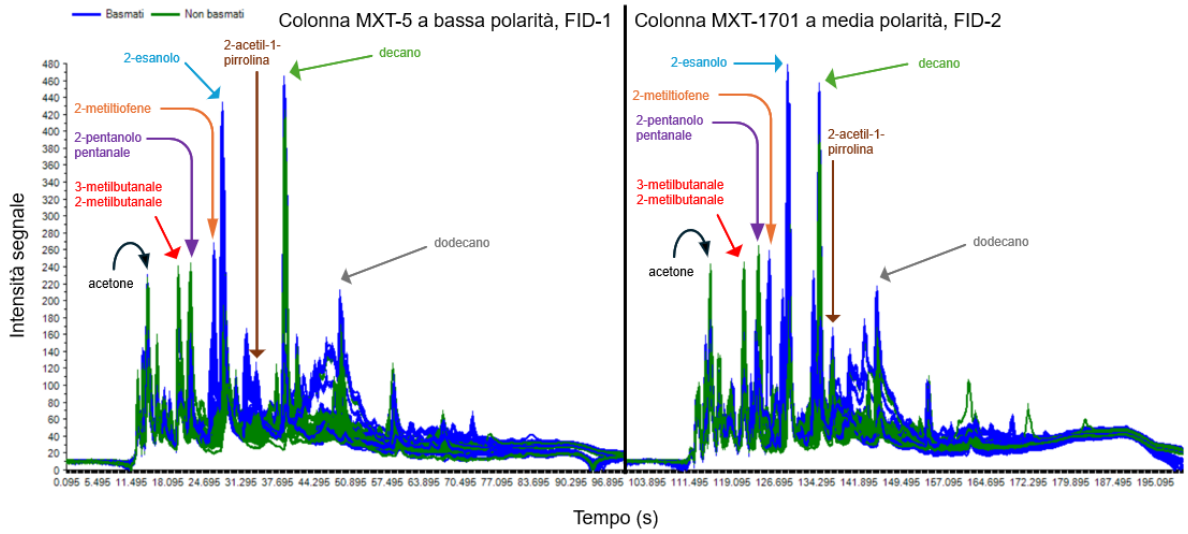


Figura 44: cromatogrammi ottenuti sulle due colonne a diversa polarità del gascromatografo. Sono evidenziati con colori diversi i composti a cui sono più probabilmente attribuibili i picchi più intensi, identificati mediante gli indici di Kovats.

quadrata.

Come mostrato anche dall'analisi svolta con FlavourSpec25®, il 2-esanolo risulta essere una delle molecole più presenti tra i composti volatili emessi dal riso riscaldato, altri picchi intensi corrispondono al 2-metil-tiofene e al 3-metilbutanale, al 2-penatnolo/pentanale e all'acetone, tutti composti rilevati anche dal gascromatografo FlavourSpec25®. Composti che non sono stati individuati dalla precedente analisi sono il decano e il dodecano che, essendo alcani, non vengono protonati dalla sorgente radioattiva all'interno del gascromatografo FlavourSpec25® e quindi non vengono rilevati dal detector, è noto da letteratura che alcune paraffine, specialmente il decano, sono state identificate all'interno di campioni di riso profumato cotto (Lee, 2022).

PCA:

Si esegue una analisi delle componenti principali al dataset trattato con la radice quadrata, rimuovendo i dati cromatografici relativi ai campioni basmati risultanti avere una percentuale di NB superiore al 7% e i dati cromatografici relativi ai campioni privi di riso non basmati (%NB assente); si ottiene così un nuovo dataset composto da 37 campioni, di cui 13 di riso basmati e 24 di riso non basmati. Dallo scores plot in **Figura 45** si nota una distinzione abbastanza netta lungo la PC-1 e la PC-2 tra i campioni di riso basmati e i campioni di riso non basmati.

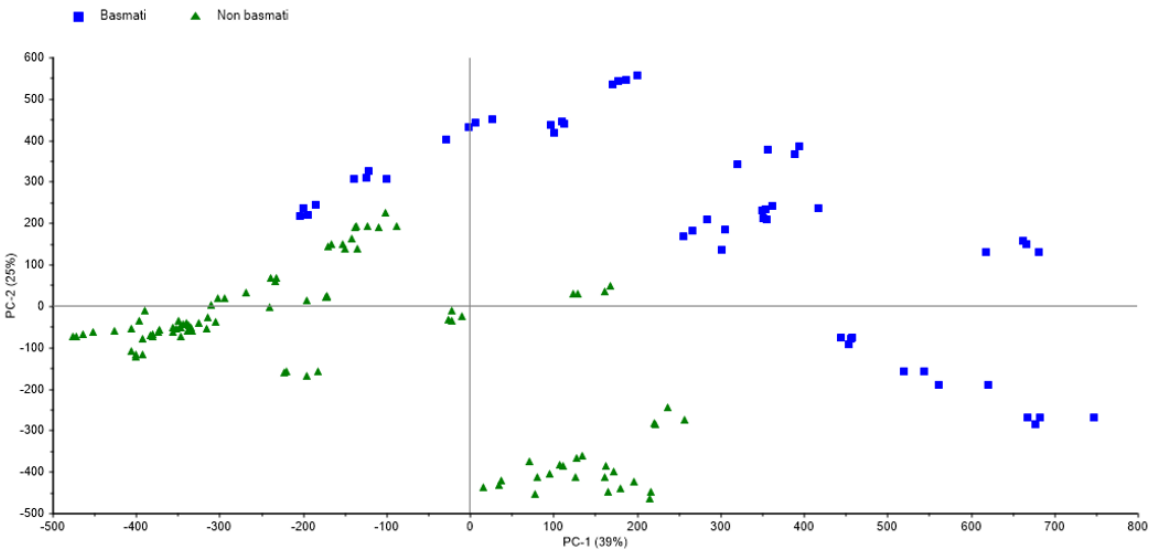


Figura 45: scores plot relativo ai dati cromatografici ottenuto in seguito ad analisi GC-FID. Ogni punto indica un campione di riso, in verde sono rappresentati i campioni di riso non basmati, in blu i campioni di riso basmati. Vengono rimossi i dati cromatografici relativi ai campioni di riso basmati che eccedono la soglia di legge riguardante la percentuale di NB al proprio interno e i campioni di riso non basmati avente percentuale di NB assente.

In **Figura 46** si riporta il loadings plot relativo alla prima componente principale mentre, in **Figura 47,** si riporta il loadings plot relativo alla seconda componente principale. Ad ogni picco più intenso nei loadings plot è stato associato un probabile composto tramite gli indici di Kovatz.

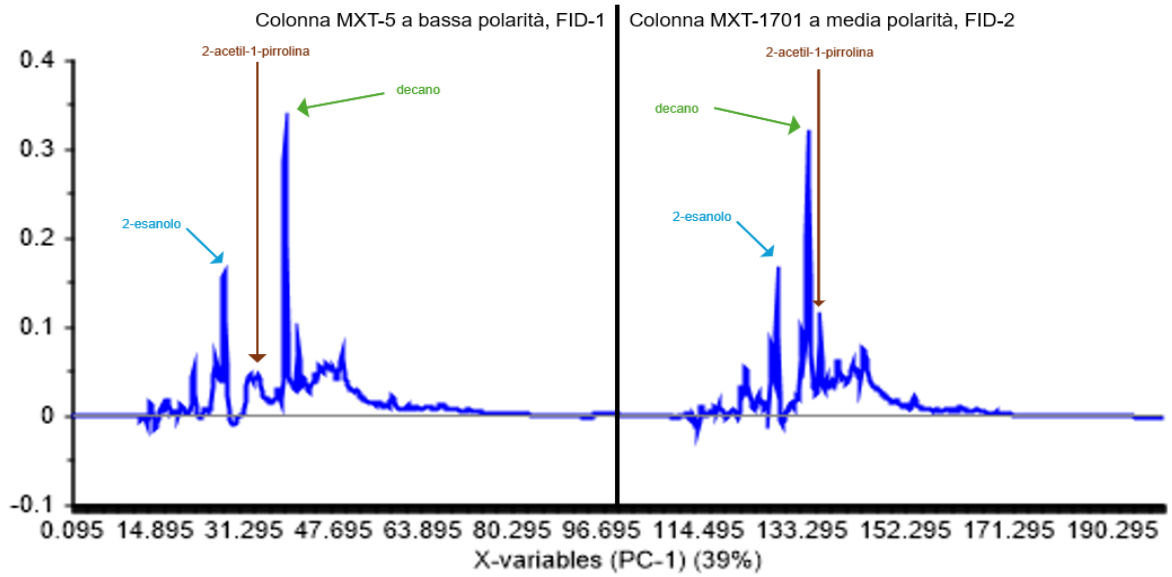


Figura 46: loadings plot relativo alla PC-1, suddiviso in due sezioni corrispondenti alla colonna a bassa polarità e a quella a media polarità. In ascissa è riportato il tempo di ritenzione, in ordinata il valore di loading, indicativo del contributo di ciascun picco alla separazione tra le due classi. Ai picchi più intensi è stato associato un probabile composto mediante gli indici di Kovats; in marrone è evidenziato il picco attribuito alla 2-acetil-1-pirrolina.

Dai loadings plot si nota come sia per la prima componente principale che per la seconda componente principale i picchi più intensi corrispondano al 2-esanolo e al decano; quindi, questi sono i 2 composti che maggiormente discriminano un riso basmati da uno non basmati. È presente un picco meno intenso rispetto a quello di altri composti al tempo di ritenzione corrispondente a quello della 2-acetil-1-pirrolina.

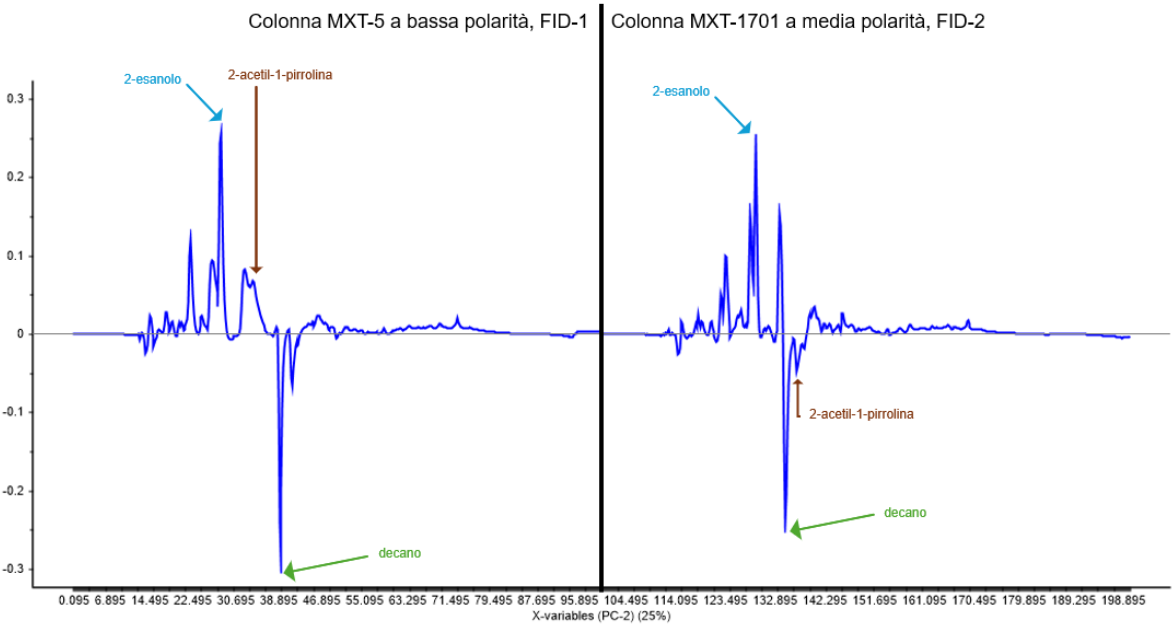


Figura 47: loadings plot relativo alla PC-2, suddiviso in due sezioni corrispondenti alla colonna a bassa polarità e a quella a media polarità. In ascissa è riportato il tempo di ritenzione, in ordinata il valore di loading, indicativo del contributo di ciascun picco alla separazione tra le due classi. Ai picchi più intensi è stato associato un probabile composto mediante gli indici di Kovats; in marrone è evidenziato il picco attribuito alla 2-acetil-1-pirrolina.

PLS-DA:

È stato utilizzato il metodo PLS-DA per la creazione del modello di classificazione del riso in basmati e non basmati. A tale scopo, vengono rimossi i campioni anomali di riso basmati che presentano una percentuale di non basmati maggiore al 7% e i non basmati che hanno una percentuale di NB pari a zero. Si riportano i risultati del modello del training set in **Tabella 15** e la matrice di confusione in calibrazione per i risi basmati in **Tabella 16**. Per il training set si sono utilizzate l'80% delle misure (117 misure) e si è eseguita una full cross validation-1 per la creazione del modello.

Tabella 15: risultati del modello del training set in calibrazione e validazione

	True positive	False positive	True negative	False negative	Specificità	Sensibilità	Accuratezza
Cal	41	0	76	0	1	1	1
CV	41	0	76	0	1	1	1

Tabella 16: matrice di confusione in calibrazione dei campioni di riso basmati

	Basmati	Non basmati
Basmati	41	0
Non basmati	0	76

Dalla **Tabella 15** e dalla matrice di confusione si nota come tutti i campioni vengano descritti correttamente sia in validazione che in calibrazione, sia per i campioni basmati che non basmati; infatti, si hanno valori di specificità, sensibilità e accuratezza uguali ad 1. Per la creazione del modello si sono scelte 17 componenti consultando il grafico riportato in **Figura 48**, in cui si sceglie un punto in cui il RMSE è in un minimo locale. In **Figura 49** si riporta il grafico raffigurante i *Variable Importance in Projection* (VIP), assegnando ad ogni picco più intenso un probabile composto.

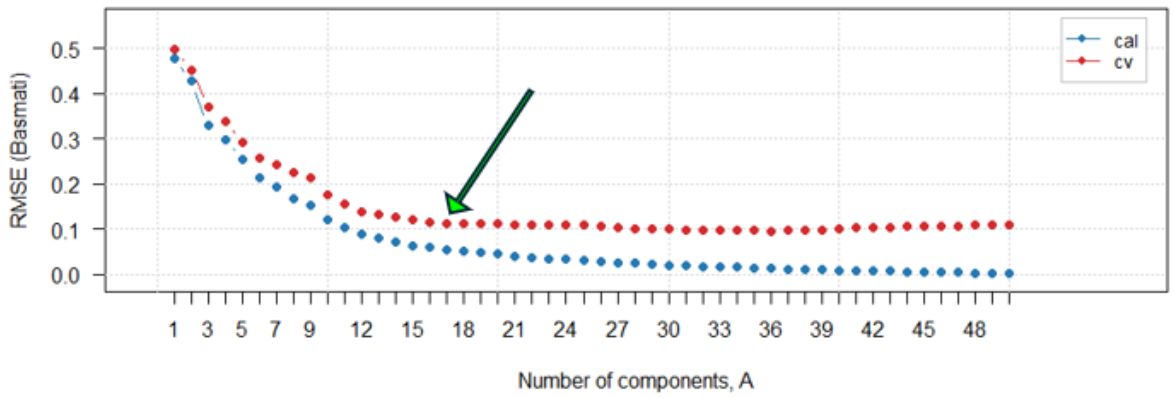


Figura 48: grafico root mean square error vs numero di componenti, si evidenzia con la freccia verde il minimo locale utilizzato per la scelta del numero di componenti.

Dal grafico in **Figura 49** si nota come i picchi più significativi per la discriminazione tra riso basmati e non basmati corrispondano agli stessi composti che risultavano più significativi per i loadings della PCA, in particolare 2-esanolo e decano.

Una volta creato il modello, si utilizzano il restante 20% delle misure (29 misure) per predire la varietà dei campioni, si riportano la matrice di confusione (**Tabella 17**) e il grafico delle predizioni dei campioni in basmati e non basmati (**Figura 50**).

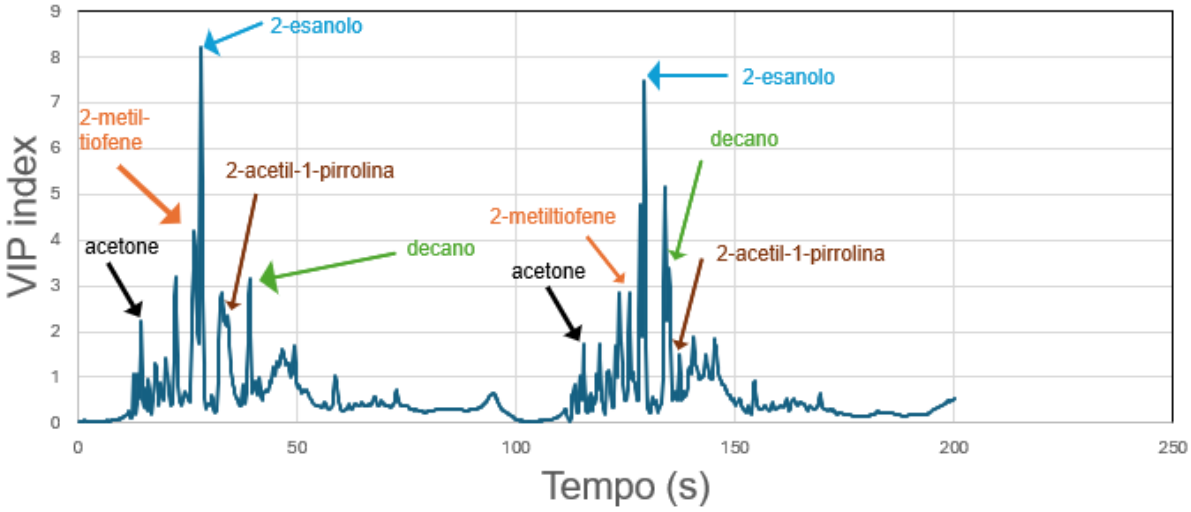


Figura 49: valori VIP (Variable Importance in Projection) del modello PLS-DA sviluppato sui dati GC-FID (Heracles II). Per ciascun composto volatile è riportato il contributo alla discriminazione tra riso basmati e non basmati; a sinistra i segnali della colonna a bassa polarità, a destra quelli della colonna a media polarità. I composti con VIP > 1 sono quelli maggiormente rilevanti per la classificazione.

Tabella 17: matrice di confusione in predizione dei campioni di riso basmati e non basmati

	Basmati	Non basmati
Basmati	10	0
Non basmati	0	19

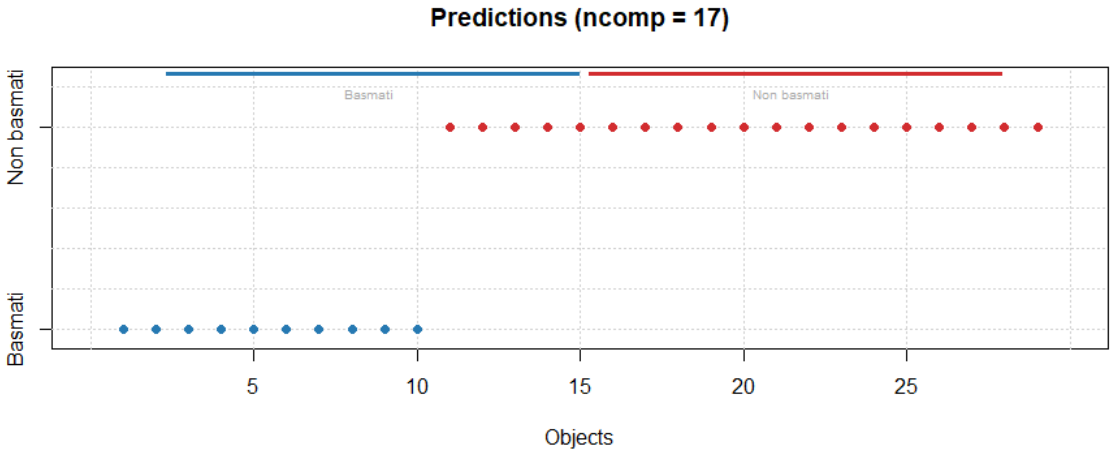


Figura 50: grafico predizioni dei campioni di riso, riporta sull'asse y la tipologia predetta dal modello e con colori diversi la tipologia reale dei campioni di riso. Ogni punto corrisponde ad un campione di riso, in rosso sono rappresentati i campioni non basmati e in blu i campioni basmati.

La matrice di confusione in predizione e il grafico delle predizioni mostrano come tutti i campioni di riso vengano predetti correttamente, confermando la affidabilità del modello di classificazione creato.

Regressione del modello: vengono proiettate tutte le miscele di riso (anche NB00 e NB100) sul modello creato, il risultato è mostrato in **Figura 51:** si nota come vengano predette come basmati tutte le miscele di riso, tranne la miscela NB100. Infine, si è provato a predire la percentuale di riso non basmati all'interno delle miscele usando 4 componenti, il risultato della predizione del test set è mostrato in **Figura 52**.

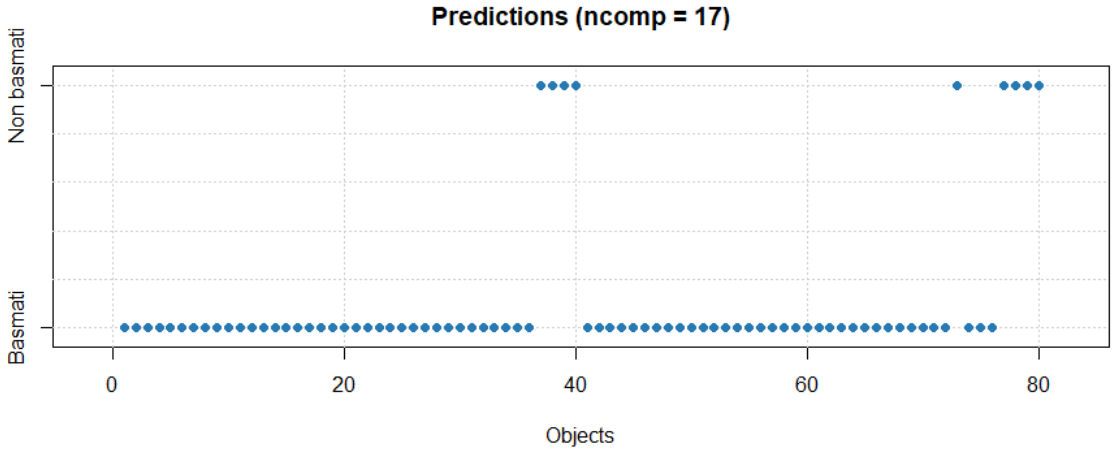


Figura 51: grafico predizioni delle miscele di riso, sull'asse y è riportata la tipologia predetta dal modello, ogni punto corrisponde ad una miscela di riso, viene predetta come non basmati solo la miscela NB100.

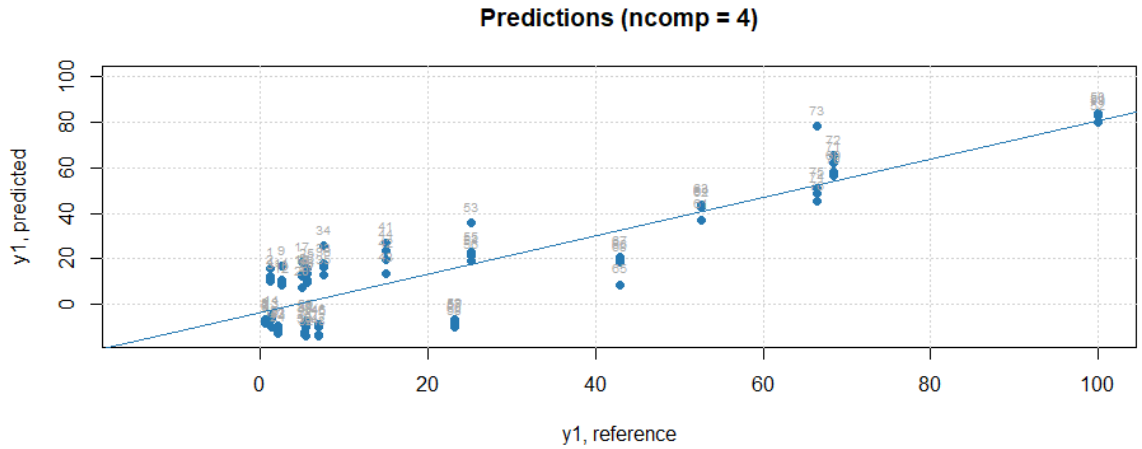


Figura 52: grafico predizioni della percentuale di riso non basmati presente all'interno di ogni miscela, sull'asse y sono riportate le percentuali predette di non basmati all'interno delle miscele, sull'asse x la percentuale di non basmati reale all'interno delle miscele.

Analizzando la correlazione lineare si nota che i valori predetti delle percentuali di riso non basmati all'interno delle miscele discostino dai valori reali, soprattutto a basse percentuali, dimostrando una scarsa capacità predittiva del modello. Si ottiene infatti una retta con pendenza 0.843 e coefficiente di determinazione pari a 0.72, valori che risultano comunque migliori rispetto a quelli ottenuti con la regressione effettuata sui dati cromatografici ottenuti con FlavourSpec25®.

CONCLUSIONI

L'obiettivo principale di questo progetto era sviluppare e integrare tecniche chimiche e biologiche per verificare l'autenticità del riso basmati, al fine di evitare frodi alimentari e tutelare i produttori, importatori e consumatori.

La digital PCR si è dimostrata una tecnica efficace e precisa per determinare la percentuale di riso non basmati all'interno dei campioni commerciali analizzati, con un limite di quantificazione, LOQ, pari a 1 copia/µL per il primer housekeeping KMV e di 2 copie/µL per il primer target NoBa. L'analisi delle miscele ha evidenziato una correlazione lineare tra la percentuale di riso non basmati misurata e quella teorica, ottenendo misure quantitative affidabili specialmente per i valori inferiori al 15%. I risultati dimostrano che la tecnica dPCR, basata sulla chimica EvaGreen, può essere impiegata per le analisi di controllo della qualità e dell'autenticità del riso basmati. Tale tecnica consente di individuare frodi alimentari, rispettando il criterio del 7% come limite di cut off per la percentuale di riso non basmati all'interno di varietà basmati, come indicato dalle Rice Association, British Rice Millers Association e il British Retail Consortium.

L'analisi dei composti volatili (VOCs) nei campioni di riso puro ha dimostrato come i profili cromatografici delle due tipologie di riso presentino gli stessi picchi, ma si differenzino per la diversa intensità dei composti caratteristici. A seguito dell'analisi untargeted, è stato individuato il picco più probabilmente associabile alla 2-acetil-1-pirrolina, caratterizzato da un indice di ritenzione compatibile con quelli presenti nei database e da una maggiore intensità nella tipologia di riso basmati. L'analisi dei VOCs ha evidenziato inoltre la presenza di altri composti maggiormente presenti nelle varietà basmati, come il 2-esanolo, che potrebbero contribuire al profumo e al sapore tipico del riso basmati.

L'analisi delle componenti principali (PCA) ha permesso di separare le tipologie di riso in basmati e in non basmati. Inoltre, una volta escluse le informazioni cromatografiche dei campioni anomali, è stato sviluppato un modello chemiometrico capace di distinguere correttamente le due varietà nei campioni puri. L'analisi di regressione delle miscele ha tuttavia evidenziato che questo modello, pur essendo in grado di separare i campioni puri basmati dai non basmati, non possiede una sensibilità sufficiente per rilevare la percentuale esatta di riso non basmati all'interno delle miscele. In particolare, il modello non riesce a discriminare con precisione le miscele contenenti

percentuali superiori o inferiori al 7% in peso di riso non basmati; pertanto, il modello non può essere utilizzato come unico metodo verificare l'autenticità del riso basmati. L'analisi del *fingerprinting* dei campioni volatili non risulta quindi sufficientemente affidabile per la quantificazione della percentuale di riso non basmati sopra ai limiti di legge, forse a causa del numero limitato di miscele analizzate. Il *fingerprinting* può tuttavia essere utilizzato come metodo complementare alla dPCR, che invece risulta essere una tecnica affidabile e precisa, sebbene richieda tempi e costi maggiori (richiede estrazione, preparazione delle piastre, amplificazione, reagenti e kit specifici).

In futuro, con l'aumento del numero di campioni analizzati e addestrando il modello su miscele con percentuali di riso non basmati prossime al 7% (peso/peso), si potranno sviluppare modelli chemiometrici più sensibili capaci di distinguere se un campione di riso è fuori dal limite di riferimento. In tal modo, eseguita l'analisi dei composti volatili di un elevato numero di campioni, si potrebbe poi ridurre tale numero per le analisi genetiche, sottoponendo a digital PCR solo i campioni con un contenuto di riso non basmati vicino al valore soglia.

In conclusione, questo studio ha dimostrato come sia necessario integrare discipline diverse e metodi analitici differenti per confermare l'autenticità di un prodotto alimentare e individuare possibili frodi alimentari. Infatti, a causa del miglioramento delle filiere produttive e della globalizzazione del mercato delle materie prime, in futuro si presenteranno sempre maggiori sfide per garantire la sicurezza e l'autenticità degli alimenti.

BIBLIOGRAFIA

(3213 - How much ethanol should be added to Buffer AP3/E (renamed to Buffer AW1) and Buffer AW (renamed to Buffer AW2) in the DNeasy Plant kit?, s.d.). (s.d.). Tratto da QIAGEN: https://www.qiagen.com/ko-us/resources/faq/3213

2-propanol. (s.d.). Tratto da Sigma-Aldrich: https://www.sigmaaldrich.com/IT/it/product/sigald/34863

Amaral. (2001). Target and Non-Target Approaches for Food Authenticity and Traceability. *Foods*.

Bartle, M. (2002). History of gas chromatography. *TrAC Trends in Analytical Chemistry*, 547-557.

Basmati Rice Code of Practice Overview | PDF. (2017, Giugno 30). Tratto da Scribd: https://it.scribd.com/document/808293566/BRITISH-CODE-OF-PRACTICE-BASMATI-RICE-2017

Borsdorf, E. (2006). Ion Mobility Spectrometry: Principles and Applications. *Applied Spectroscopy Reviews* , 323–375.

Bucher, K. (2016). Duplex digital droplet PCR for the determination of non-Basmati rice in Basmati rice (Oryza sativa) on the base of a deletion in the fragrant gene. *Eur Food Res Technol* , 927–934.

Buffer EB. (s.d.). Tratto da QIAGEN: https://www.qiagen.com/us/products/discovery-and-translational-research/lab-essentials/buffers-reagents/buffer-eb

Classification with Multivariate Analysis: Focus on PLS-DA. (s.d.). Tratto da elearningunite: https://elearning.unite.it/pluginfile.php/365055/mod_resource/content/1/PLS-DA%2025-26.pdf

Colonna capillare GC Rtx-5, 30 m, ID 0,53 mm, 3,00 µm. (s.d.). Tratto da https://it.restek.com/p/10285: https://it.restek.com/p/10285

Cos'è l'analisi delle componenti principali (PCA)? - IBM. (s.d.). Tratto da IBM: https://www.ibm.com/it-it/think/topics/principal-component-analysis

Danezis, T. C. (2016). Food authentication: Techniques, trends & emerging approaches. *Food authentication: Techniques, trends & emerging approaches*, 123-132.

Differenze tra qPCR e dPCR. (s.d.). Tratto da Gentaur: https://gentaur.it/enciclopedia/differenze-tra-qpcr-e-dpcr

Dimensioni del mercato del riso e fattori di crescita. (2026, febbraio 25). Tratto da mordorintelligence.it: https://www.mordorintelligence.it/industry-reports/rice-market

DNA Cleanup Buffers. (s.d.). Tratto da QIAGEN: https://www.qiagen.com/ko-us/products/discovery-and-translational-research/lab-essentials/buffers-reagents/dna-cleanup-buffers

EP2565281A1 - Dye blends. (s.d.). Tratto da Google Patents: https://www.google.com/search?sca_esv=65f9ea52ca19315a&rlz=1C1VDKB_itIT1044IT1044&sxsrf=APpeQnvHGyiGXqPSruZ6oBmwd94xeOHfVA:1783517169420&udm=2&fbs=ABfTbFWgwlc2MNw7uknLQP9cFo05YtUSpNgn2klEDQSYENClc1REDvMK2B5VqcHL1AzgYSb8XqNEQ7kenNbDD5dNYRCQ0jmk5wcmLJYPRESC

EvaGreen Dye for qPCR & Other Applications | Biotium. (s.d.). Tratto da biotium: https://biotium.com/technology/pcr-dna-amplification/evagreen-dye-for-qpcr/

EvaGreen® Dye For qPCR & Other Applications. (s.d.). Tratto da biotium: https://www.google.com/search?sca_esv=65f9ea52ca19315a&rlz=1C1VDKB_itIT1044IT1044&sxsrf=APpeQnseyCKbjYiqqrPukYY0vPAW1Hkbpw:1783517313800&udm=2&fbs=ABfTbFWgwlc2MNw7uknLQP9cFo05YtUSpNgn2klEDQSYENClc1REDvMK2B5VqcHL1AzgYSb8XqNEQ7kenNbDD5dNYRCQ0jmk5wcmLJYPRESC

Fast EvaGreen® qPCR Master Mix. (s.d.). Tratto da Cambridge Bioscience: https://www.bioscience.co.uk/userfiles/pdf/Fast%20EvaGreen.pdf

FlavourSpec®25 with PAL3-RSI. (2024). Tratto da GAS Dortmund : https://www.gas-dortmund.de/data-file-gas/downloads/Manuals/2024-10-10%20FlavourSpec25%20User%20Manual%204.61%20EN.pdf

Fundamentals of digital PCR. (s.d.). Tratto da QIAGEN: https://www.qiagen.com/us/knowledge-and-support/knowledge-hub/bench-guide/pcr/digital-pcr/what-is-digital-pcr

Fuso. (2023). Sensi chimici. La scienza degli odori e dei sapori. *Annales. Proceedings of the Academy of Sciences of Bologna. Class of Physical Sciences.*

Gabelica, M. (2018). Fundamentals of ion mobility spectrometry. *Current Opinion in Chemical Biology*, 51-59.

Gas Chromatography: Principle, Parts, Steps, Procedure, Uses. (s.d.). Tratto da Microbe Notes: https://microbenotes.com/gas-chromatography/

Gascromatografia (GC). (s.d.). Tratto da Sigma-Aldrich: https://www.sigmaaldrich.com/IT/it/applications/analytical-chemistry/gas-

chromatography?srsltid=AfmBOooRPftNK--sRuercaUsfpx5fw0o1g8_GMTriuS9b73wtTetJVcY

Guanidina - Aminoformamidina, Aminometanamidina. (s.d.). Tratto da Sigma-Aldrich: https://www.sigmaaldrich.com/IT/it/substance/guanidinehydrochloride955350011?srsltid=AfmBOopTE9sfNMZE7Ov8GaBaN0-H4QVQnJDitAsQenAmt88fJvN3YNiy

Heberger. (2008). Chemoinformatics—multivariate mathematical–statistical methods for data evaluation. *Medical Applications of Mass Spectrometry*, 141-169.

Himmelreich. (2009). Method for isolating and purifying nucleic acids .

Hu. (2008). Poisson distribution and application.

Hu, L. G. (2020). Volatile compounds, affecting factors and evaluation methods for rice aroma: A review. *Trends in Food Science & Technology*, 136-146.

Huck, P. H.-P. (2016). An industry perspective of food fraud. *Current Opinion in Food Science*, 32-37.

Il riso come ingrediente nei mangimi per suini - Articoli. (2022, Giugno 20). Tratto da 3tres3: https://www.3tres3.com/it/articoli/il-riso-come-ingrediente-nei-mangimi-per-suini_12610/

Kralik, R. (2017). A Basic Guide to Real Time PCR in Microbial Diagnostics: Definitions, Parameters, and Everything. *Frontiers in Microbiology* .

L'Annuario Statistico 2024 della FAO rivela dati cruciali sulla sostenibilità dell'agricoltura a livello globale, sulla sicurezza alimentare e sull'occupazione nei sistemi agroalimentari. (2024). Tratto da Food and Agriculture Organization: https://www.fao.org/newsroom/detail/fao-statistical-yearbook-2024-reveals-critical-insights-on-the-sustainability-of-agriculture-food-security-and-the-importance-of-agrifood-in-employment/it

la chimica del riso - Ute CiniselloB. (s.d.). Tratto da uteciniselloBalsamo.it: https://www.uteciniselloBalsamo.it/Dispense/scientifica/2024_2025_la_chimica_della_cucina/dispensa_13.pdf

Lavorazione riso: le tipologie dall'integrale al raffinato. (2024, Maggio 2). Tratto da Mundi Riso: https://www.mundiriso.it/2024/05/02/lavorazione-riso-tipologie-e-caratteristiche/

Lee, Y. J. (2022). Identification of volatile organic compounds related to the eating quality of cooked japonica rice. *Scientific Reports*.

Majumdar. (2015). Digital PCR Modeling for Maximal Sensitivity, Dynamic Range and Measurement Precision.

Manning. (2016). Food fraud: policy and food chain. *Current Opinion in Food Science*, 16-21.

Melucci, B. T. (2016). Rapid direct analysis to discriminate geographic origin of extra virgin olive oils by flash gas chromatography electronic nose and chemometrics. *Food Chemistry*, 263-273.

Muthayya, S. M. (2014). An overview of global rice production, supply, trade, and consumption. *Annals of the New York Academy of Sciences*, 7-14.

Nader, S. K. (2026). DNA markers for Basmati authentication have the power to distinguish among other global economically relevant rice varieties and to evaluate their genetic background. *Food Control*.

Optima® 624 colonne capillari per GC Ø int. 0,32 mm Ø est. 0,5 mm Spessore film 1,80 µm Lungh. 50 m. (s.d.). Tratto da Unimed Scientifica: https://www.unimedscientifica.it/prodotto/optima-624-colonne-capillari-per-gc-int-025-mm-est-05-mm-spessore-film-180-m-lungh-30-m/

Polymerase Chain Reaction (PCR). (s.d.). Tratto da National Human Genome Research Institute : https://www.genome.gov/genetics-glossary/Polymerase-Chain-Reaction-PCR

Principal Component Analysis (PCA). (s.d.). Tratto da Scaler Topics: https://www.scaler.com/topics/nlp/what-is-pca/

Proteinase K for Efficient Protein Digestion. (s.d.). Tratto da QIAGEN: https://www.qiagen.com/ko-dm/product-categories/discovery-and-translational-research/enzymes-for-molecular-biology/proteinase-k

QIAcuity Digital PCR System. (s.d.). Tratto da QIAGEN: https://www.qiagen.com/us/products/instruments-and-automation/pcr-instruments/qiacuity-digital-pcr-system?cmpid=CM_PCR_dPCR_Traffic_0123_SEA_GA_ROW&s_kwcid=AL!16953!3!701116944449!p!!g!!qiagen%20dpcr&gad_source=1&gad_campaignid=9644320942&gbraid=0AAAAAD-R

QIAcuity EvaGreen (EG) PCR Kit. (s.d.). Tratto da QIAGEN: https://www.qiagen.com/us/products/instruments-and-automation/accessories/qiacuity-eg-pcr-kit

QIAcuity Nanoplates and Accessories. (s.d.). Tratto da QIAGEN: https://www.qiagen.com/us/products/instruments-and-automation/accessories/qiacuity-nanoplates-and-accessories?catno=250001

Qiagen QIAcube: Product Development – Helbling Projects. (s.d.). Tratto da helbing-inc.com: https://helbling-inc.com/en/cases/qiacube

QIAGEN QIAcuity dPCR. (s.d.). Tratto da Dipylon medical: https://dipylonmedical.com/product/qiagen-qiacuity/

Quan, S. B. (2018). dPCR: A Technology Review. *Bio-MEMS for Precision Medicine*.

Qubit Thermo Fisher Scientific | Thermo Fisher Scientific. (s.d.). Tratto da Thermo Fisher SCientific: https://www.thermofisher.com/it/en/home/industrial/spectroscopy-elemental-isotope-analysis/molecular-spectroscopy/fluorometers/qubit.html?ef_id=EAIaIQobChMIzYau9o_DlQMVxqSDBx0qfgc-EAAYASAAEgJQhPD_BwE:G:s&s_kwcid=AL!3652!3!679462091897!e!!g!!qubit%20fluoro

Qubit® 3.0 Fluorometer - Unimed Scientifica srl. (s.d.). Tratto da Unimed Scientifica : https://www.unimedscientifica.it/prodotto/qubit-3-0-fluorometer/

Real-Time vs. Digital PCR vs. Traditional PCR. (s.d.). Tratto da Thermo Fisher Scientific: https://www.thermofisher.com/it/en/home/life-science/pcr/real-time-pcr/real-time-pcr-learning-center/real-time-pcr-basics/real-time-vs-digital-vs-traditional-pcr.html

Reid, O. D. (2006). Recent technological advances for the determination of food authenticity. *Trends in Food Science & Technology*, 344-353.

Risi, E. N. (2025). *REGISTRO VARIETALE per l'annata agraria 2025/2026 aggiornato al 28 agosto 2025 (Articolo 6 del decreto legislativo 4 agosto 2017, n.131).* Tratto da Ente Nazionale RIsi : https://www.enterisi.it/upload/enterisi/gestionedocumentale/REGISTROVARIETALEcampagna2025-2026persito-def_784_114852.pdf

Riso nel mondo: una coltivazione globale di un chicco universale. (s.d.). Tratto da LaBalocchina: https://labalocchina.com/blogs/notizie/riso-nel-mondo-una-coltivazione-globale-di-un-chicco-universale?srsltid=AfmBOooagpv0UmGvI-rvjQ7UmWL1QuUdWaE_4c9yQn8ybHuCbyzC63dG

Rivoluzione verde. (s.d.). Tratto da Wikipedia: https://it.wikipedia.org/wiki/Rivoluzione_verde

Robson, D. H. (2021). A comprehensive review of food fraud terminologies and food fraud mitigation guides. *Food Control*.

Rozin. (1996). Socioiocultural influences on human food selection. *APA PsycNet*, 233-263.

Spink, M. (2011). Defining the Public Health Threat of Food Fraud. *Food Science*.

Srivong, W. P. (2007, Gennaio 01). Characterization of a Fragrant Gene and Enzymatic Activity of Betaine Aldehyde Dehydrogenase in Aromatic and Nonaromatic Thai Rice Cultivars.

Sun, L. C. (2022). An HS-GC-IMS analysis of volatile flavor compounds in brown rice flour and brown rice noodles produced using different methods. *LWT*.

Suprem, M. K. (2013). A review on application of technology systems, standards and interfaces for agriculture and food sector. *Computer Standards & Interfaces*, 355-364.

Tris(idrossimetil)amminometano cloridrato EMPROVE® EXPERT. (s.d.). Tratto da Sigma-Aldrich: https://www.sigmaaldrich.com/IT/it/product/mm/108219?srsltid=AfmBOooNN1iPiCq8eobG7ZC5Lc-h8fSZf73vL1zA3M5mzfwH5FUWC59w

Wang, C. S. (2020). Recent progress in food flavor analysis using gas chromatography–ion mobility spectrometry (GC–IMS). *Food Chemistry*.

Webb, B. (2016). Global Challenges for the 21st Century: the Role and Strategy of the Agri-Food Sector. *Animal reproduction*, 133-142.

Wei. Handoko, P. M. (2017). Evaluation of 2-acetyl-1-pyrroline in foods, with an emphasis on rice flavour. *Food Chemistry*, 531-544.

World's Most Eaten Food. (2022, Dicembre 18). Tratto da AGC Catering Equipment : https://www.agcequipment.com.au/blog/worlds-most-eaten-food/

Yang, S. S. (2008). Comparison of Odor-Active Compounds from Six Distinctly Different Rice Flavor Types. *Journal of Agricultural and Food Chemistry*, 2780–2787.

Zaplin, L. L. (2013). Production of high oleic rice grains by suppressing the expression of the OsFAD2-1 gene. *Funct Plant Biol.*, 996-1004.