



UNIVERSITÀ
DEGLI STUDI
FIRENZE

**Scuola di
scienze matematiche
fisiche e naturali**

Corso di laurea in chimica
Curriculum scienze chimiche

Studio del fenotipo metabolico individuale: stabilità nel tempo e fattori di variabilità

Study of the individual metabolic phenotype: stability over time and factors of variability

Relatore

Claudio Luchinat

Correlatore

Leonardo Tenori

Candidato

Tommaso di Francescantonio

INDICE

- 1 INTRODUZIONE
 - 1.1 Metabolomica e le altre “omiche”
 - 1.2 Metodi di analisi in metabolomica
 - 1.3 Analisi statistica in metabolomica
 - 1.4 Analisi metabolica di urine
 - 1.5 Fingerprint metabolico
 - 1.6 Applicazioni della metabolomica
- 2 PROGETTO METREF
 - 2.1 MetRef 1
 - 2.2 MetRef 2-3
- 3 ORGANIZZAZIONE E SCOPO DEL LAVORO
- 4 MATERIALI E METODI
 - 4.1 Collezione dei campioni
 - 4.2 Preparazione dei campioni
 - 4.3 Acquisizione degli spettri
 - 4.4 Processing degli spettri
 - 4.5 Analisi statistica
- 5 RISULTATI E DISCUSSIONE
 - 5.1 Spettri NMR di urina umana
 - 5.2 Analisi MetRef 4
 - 5.3 Stabilità del fenotipo
 - 5.4 Analisi delle variabilità individuali
- 6 CONCLUSIONI
- 7 ABBREVIAZIONI
- 8 BILIOGRAFIA
- 9 RINGRAZIAMENTI

1 INTRODUZIONE

1.1 METABOLOMICA E LE ALTRE “OMICHE”

Negli ultimi venti anni si è osservato un intenso sviluppo delle scienze omiche nell'ambito della biologia dei sistemi¹.

Questo sviluppo è da attribuire al sequenziamento dell'intero genoma e all'utilizzo di tecniche “high-throughput” le quali hanno garantito un maggior numero di informazioni sulla biologia degli esseri viventi².

Lo scopo delle scienze “omiche” è quello di studiare ed interpretare l'insieme di dati generato per dare una visione globale dello stato fisiopatologico del sistema biologico preso in esame.

Le prime tecniche “omiche” sviluppate sono state la genomica, la proteomica e la trascrittomica.

Più precisamente la genomica studia il sequenziamento dell'intero genoma ed è la più sviluppata e meno recente delle scienze dette omiche.

La proteomica invece studia il prodotto dei geni ovvero le proteine e quindi mira ad individuare e quantificare l'espressione proteica da parte del genoma.

La trascrittomica si concentra sul processo di trascrizione cellulare ovvero studia l'identificazione e la quantificazione dei trascritti di RNA all'interno della cellula.

Lo studio e lo sviluppo delle scienze omiche ha evidenziato, a differenza di quanto ipotizzato dal dogma centrale della biologia molecolare³, che i processi cellulari non sono unidirezionali ma sono interconnessi tra loro attraverso meccanismi retroattivi.

In pratica, attraverso la sola conoscenza della genomica non si può prevedere l'evoluzione del metabolismo di un sistema biologico.

I processi cellulari di un organismo più o meno complesso non possono essere regolati da un sistema rigido a blocchi ma da una struttura organizzata ed interconnessa che regola i flussi di energia e di materia all'organismo.

Alle scienze omiche tradizionali, recentemente si è affiancata una nuova disciplina: La metabolomica.

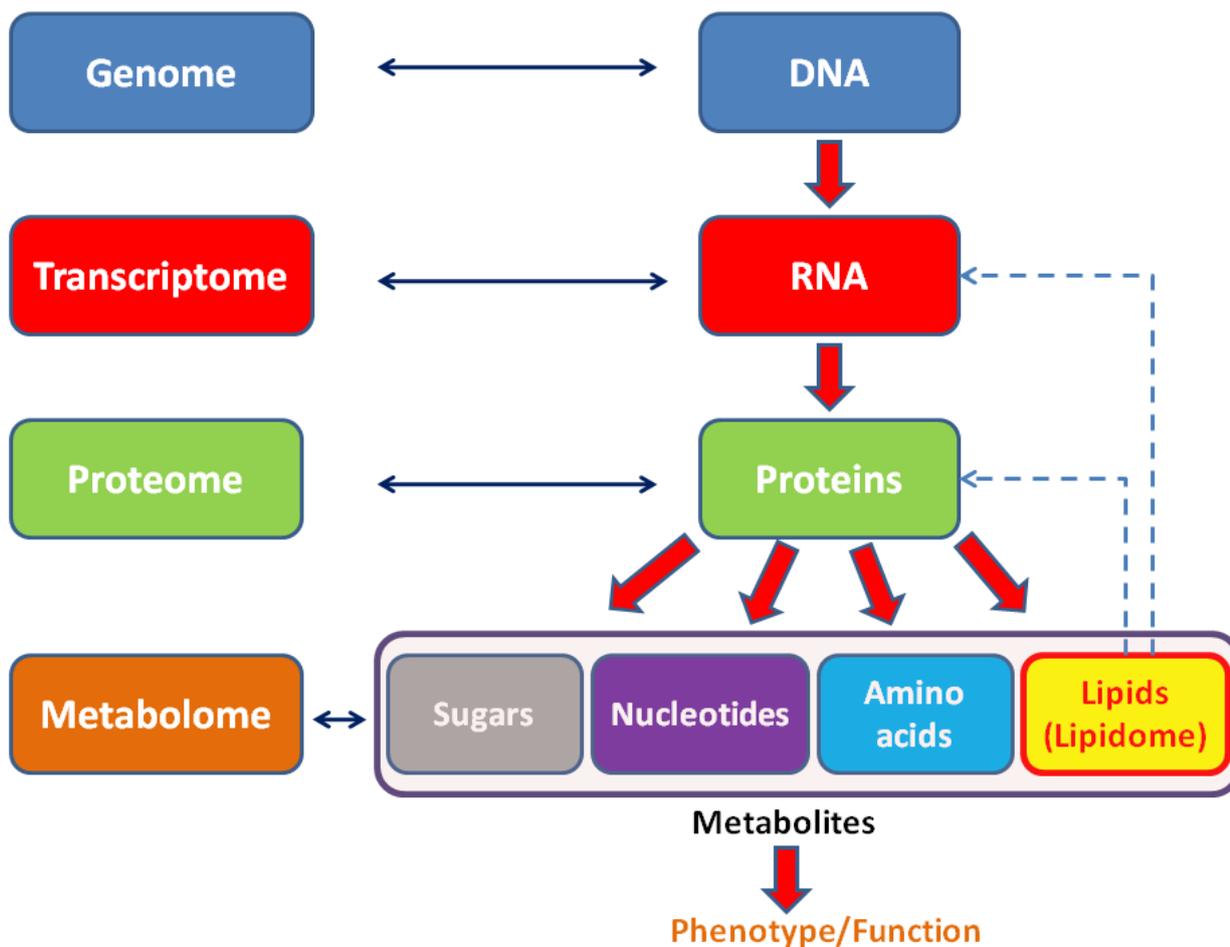


Figura 1 "schema delle scienze omiche"

La metabolomica viene definita come lo studio della descrizione quantitativa di tutte le molecole dal peso molecolare inferiore ai 1000Da presenti in un campione biologico⁴

Questa scienza va a complementare le più tradizionali scienze omiche attraverso l'investigazione delle proprietà che non possono essere direttamente valutate dalla proteomica e dalla genomica.

Le modificazioni epigenetiche⁶, o anche i cambiamenti dell'espressione del Dna che non possono essere valutati dalla conoscenza della sequenza del Dna stesso, possono altresì influenzare indirettamente il metaboloma⁴.

Appare ovvio che un sistema biologico complesso risenta e si modifichi continuamente in relazione allo stato fisiologico e alle interazioni con l'ambiente esterno generando nuove vie metaboliche.

La metabolomica studiando e quantificando i metaboliti presenti nei liquidi biologici offre una visione del sistema istantanea, fornendo informazioni utili alla comprensione dei processi in atto nell'organismo analizzato.

I metaboliti, ovvero le piccole molecole che vanno a costituire il metaboloma, possono essere viste come il prodotto finale dell'espressione genica e dell'attività proteica, definendo così il fenotipo biochimico di un organismo biologico.

Il metaboloma sarà quindi costituito dall'insieme di tutte le molecole di piccole dimensioni (metaboliti) presenti in una cellula, in un tessuto, in un organismo in un determinato stato fisiologico.

Gli studi metabolomici recentemente stanno riscuotendo un enorme successo nell'ambito biochimico tanto che negli ultimi anni vi è stato un notevole incremento delle pubblicazioni in questo settore.

Ciò è dovuto al fatto che il metaboloma fornisce la misura quantitativa dei metaboliti presenti nell'organismo e inoltre, essendo il gradino finale della cascata biologica che connette geni, preteine e metaboliti, è il più utile a descrivere l'attività della cellula a livello funzionale.

A differenza del proteoma o del genoma il profilo metabolico è più sensibile e rapido ai cambiamenti dovuti agli stimoli esterni, inclusi i cambiamenti a livello fisiopatologico.

La causa di questa caratteristica è da cercare nei meccanismi post-trascrizionali e post-traduzionali che fanno del metaboloma l'immagine più rappresentativa del fenotipo individuale⁵.

Quindi la metabolomica assume un ruolo fondamentale nel determinare istante per istante la risposta metabolica conseguente a stimoli esterni, a differenza della genomica che fornisce solo una previsione statica di come può evolvere un sistema.

Gli studi metabolomici si basano per la maggior parte sull'analisi di biofluidi facilmente ottenibili dai mammiferi (principalmente esseri umani) quali sangue (siero e/o plasma) e urine.

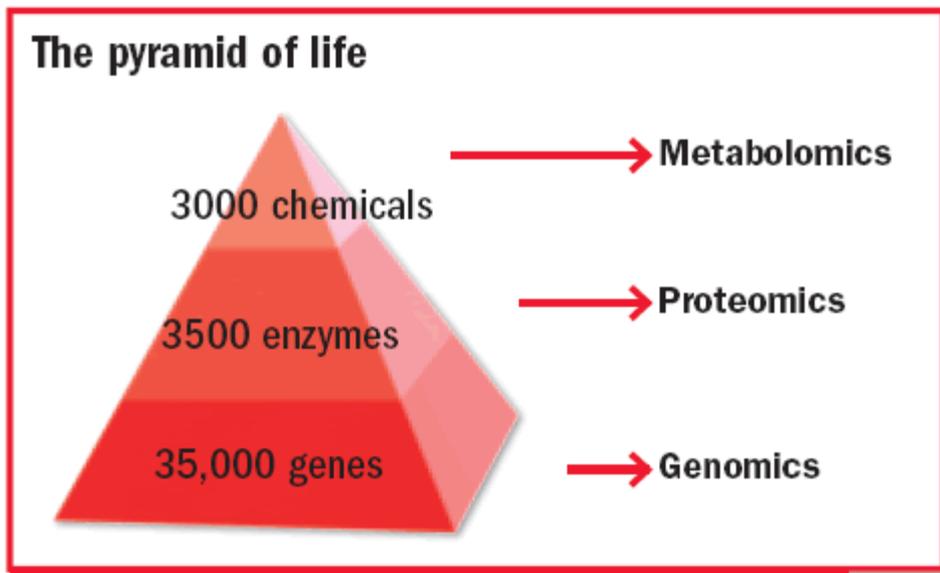
Un altro aspetto non meno importante è anche la quantità di elementi da analizzare. Infatti a livello metabolomico si studiano al massimo qualche migliaio di metaboliti decisamente in quantità minore rispetto alle decine di migliaia di geni (35000 geni nel genoma umano) e alle centinaia di migliaia di proteine.

La composizione di un biofluido è direttamente legata alla funzione delle cellule riservate alla sua produzione. Esso fornisce importanti informazioni sullo stato funzionale e biochimico non solo delle cellule che lo hanno prodotto, ma anche sulle possibili alterazioni sistemiche legate ad una malattia o a fattori esterni quali dieta, farmaci o stati emotivi.

L'utilizzo principale di sangue e urine è determinato non solo dalla quantità di informazioni contenute ma anche alla maggiore reperibilità e a tecniche di estrazione meno invasive.

Sono state condotte analisi anche su campioni di saliva, sudore, condensato di respiro e liquidi cerebrospinali⁷.

Un campione di biofluido umano contiene numerosi metaboliti riconosciuti stimati intorno alle 3000 unità, che rispetto agli oltre 35000 geni e 3500 enzimi di un essere umano forniscono un ulteriore vantaggio alla metabolomica riducendo il numero di elementi da determinare.



Al livello di ricerca l'integrazione della metabolomica con la genomica e la proteomica può aiutare a comprendere a più livelli di informazione quello che è l'intero processo cellulare.

1.2 Metodi di Analisi in metabolomica

Al fine di ottenere un'esauriente analisi metabolomica è necessario ottenere la determinazione e la quantificazione dei metaboliti che forniscono l'immagine dello stato biochimico del sistema biologico preso in esame.

L'analisi metabolomica può essere condotta attraverso due strategie complementari tra loro per livello di informazione.

La prima tiene conto delle variazioni metaboliche complessive senza andare ad analizzare le concentrazioni dei metaboliti presenti.

Lo scopo è di ottenere una visione globale dello stato fisiologico dell'organismo attraverso l'analisi spettrale accoppiata alla statistica multivariata.

Questo tipo di analisi fornisce una rappresentazione numerica dei dati ottenuti che costituisce un'impronta metabolica digitale ("fingerprint metabolico")⁸.

La seconda strategia di indagine (“metabolic profiling”^{6,7}) mira all’analisi qualitativa e quantitativa dei singoli metaboliti.

L’identificazione e la quantificazione delle molecole che costituiscono il metaboloma è necessaria e fondamentale per determinare biomarkers legati a specifiche vie metaboliche.

Le tecniche più utilizzate per entrambi gli approcci sono indubbiamente la spettroscopia di risonanza magnetica(NMR) e la spettrometria di massa(MS) che garantiscono efficienza, velocità e bassi costi.

Queste tecniche sono in grado di quantificare e determinare strutturalmente centinaia di metaboliti.

L’analisi NMR è una tecnica non distruttiva che fornisce una risposta analitica lineare e riproducibile che ne permette l’utilizzo sia per il fingerprint che per il profiling.

Nonostante non presenti una grande sensibilità (limite di rilevazione attorno al micromolare), l’approccio basato sull’NMR è vantaggioso in termini di tempo e costi operativi.

Una scansione infatti richiede pochi minuti e la preparazione dei campioni è facile e veloce.

La spettrometria di massa è una tecnica distruttiva molto più sensibile (limite di rilevabilità intorno al nano/picomolare, a seconda degli strumenti).

Per questo motivo la MS viene impiegata nel processo di profiling, garantendo non solo la quantificazione dei metaboliti ma fornendone anche la struttura.

Questa tecnica viene spesso accoppiata a tecniche cromatografiche quali HPLC (high performance liquid chromatography) e GC (gas-chromatography) per la separazione degli analiti^{9,8}.

Recentemente, nonostante la forte affermazione dell’NMR e della MS, alcune tecniche quali Raman e Ft-IR sono state rivalutate per la capacità di determinare efficientemente i gruppi funzionali in pochi secondi³.

1.3 Analisi statistica in metabolomica

Tramite le analisi metabolomiche si ottengono notevoli quantità di dati che devono essere interpretati. I metodi chemiometrici aiutano a gestire questa enormità di informazione.

Le misurazioni ottenute vengono trasformate in matrici numeriche che riportano, per ciascun campione misurato, l'insieme dei chemical shifts dello spettro NMR. Generalmente questo è realizzato attraverso il processo definito "bucketing", che consiste nel suddividere lo spettro in intervalli di ampiezza predefinita (es. 0.02 ppm) e quindi nel calcolare l'integrale dello spettro in tali zone (chiamate appunto buckets). I buckets divengono così le variabili utilizzate nelle successive analisi statistiche. Questo metodo permette di: 1) diminuire il numero totale di variabili (che da circa 64000 chemical shifts divengono circa 500 buckets) e 2) compensare le piccole variazioni di posizione dei segnali dovute al pH.

La chemiometria viene definita da Malinowski¹¹ come l'utilizzo di metodi matematici e statistici al fine di utilizzare, interpretare e predire dati "chimici".

La chemiometria quindi si avvalora della statistica multivariata per:

- Avere una visione globale delle variazioni, determinare le differenze globali e le relazioni tra variabili e campioni
- Poter realizzare predizioni sui nuovi campioni analizzati
- Determinare la presenza di clusters, ovvero raggruppamenti di campioni secondo una variabile o solo globalmente, ma anche di outsiders cioè campioni che si discostano particolarmente dal sistema preso in esame
- Analizzare i singoli metaboliti e osservare i responsabili delle differenze negli spettri

I metodi chemiometrici si dividono in due principali gruppi ovvero i metodi unsupervised e supervised.

I metodi unsupervised sono utilizzati prevalentemente per una scansione preliminare ed hanno come scopo quello di evidenziare le differenze globali,

riducendo le variabili e massimizzando la varianza, fornendoci informazioni sulle clusterizzazioni ma anche sulla presenza di outliers.

Una delle tecniche principalmente utilizzate nei metodi unsupervised è la *Principal Components Analysis* o PCA¹².

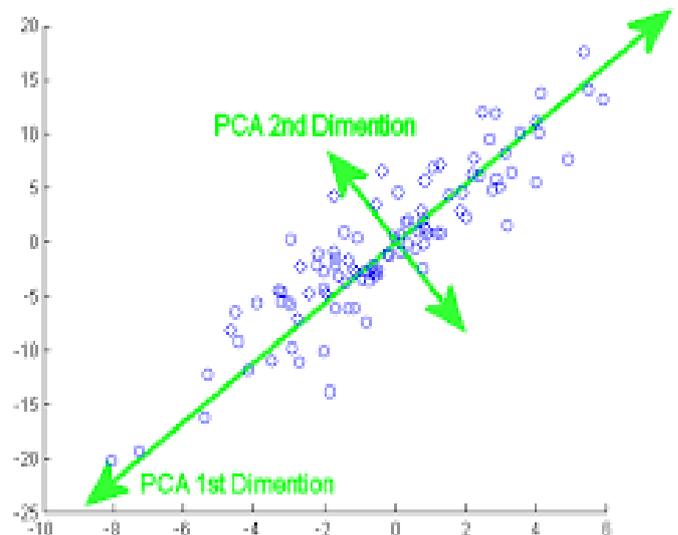
Lo scopo della PCA è quello di trovare le differenze e le similarità tra i campioni e quindi mettere in relazione le variabili tra di loro.

Nella PCA i dati sono definiti tramite una matrice ($n \times p$) in cui il primo asse fornisce la massima varianza dei dati mentre il secondo asse, perpendicolare al primo, fornisce la successiva massima varianza dei dati. La matrice utilizzata può essere oltretutto estesa con assi aggiuntivi.

La PCA può essere utilizzata per definire quali variabili sono più importanti in un processo.

Ogni campione potrà quindi essere rappresentato anche da sole poche componenti in quanto le prime sono quelle che ci danno la massima varianza.

Per le analisi statistiche è necessario conoscere sia i factors (le colonne della matrice) che gli objects (le righe della matrice).



Se lo scopo dell'esperimento è quello di determinare quale variabile è più significativa si può osservare la matrice dei loadings composta nelle colonne dagli autovettori della matrice di covarianza mentre le righe rappresentano le variabili originali.

Per le analisi e la comprensione delle relazioni tra i campioni determinando la presenza di clusters o outliers è necessario studiare la matrice degli scores che definisce come gli objects si trovano rispetto al sistema di riferimento nuovo.

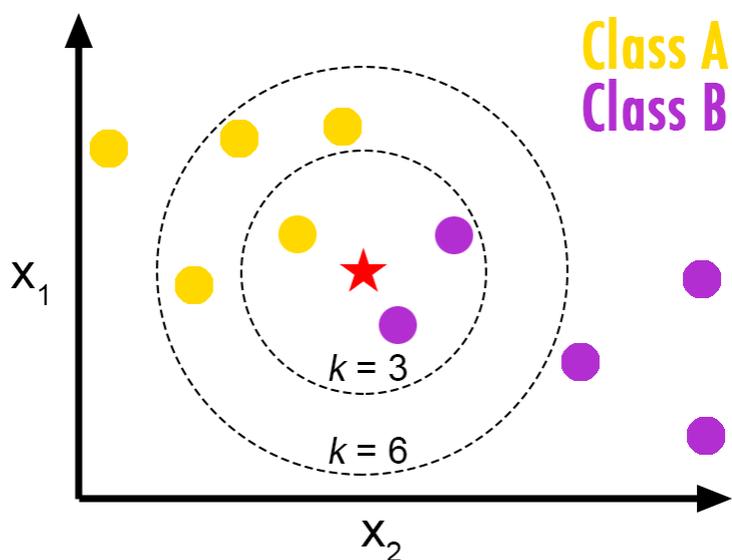
Nei metodi supervised il sistema viene istruito con informazioni aggiuntive rispetto ai metodi unsupervised quali ad esempio il numero delle classi in

modo tale da poter determinare la presenza di raggruppamenti o/e semplicemente similarità tra i diversi gruppi presenti nei dati.

Una delle tecniche utilizzate è la PLS (Partial least squares regression).

Questa tecnica pone in relazione la matrice X delle variabili dipendenti con la matrice Y delle variabili indipendenti quindi cerca la direzione multidimensionale dello spazio X che spieghi la massima direzione della varianza multidimensionale dello spazio Y (ossia massimizza la covarianza tra X e Y).

Un'altra tecnica Supervised, fondamentale per le analisi metabolomiche, è la K-NN (K- Nearest Neighbors).



Questo metodo statistico ci permette di fare delle predizioni su campioni futuri ovvero stabilita la base dei campioni selezionata, assegna i nuovi campioni all'elemento più significativo.

Crea un intorno del campione, tramite una funzione di distanza prestabilita, al fine di assegnare il campione prescelto alla categoria maggioritaria nel suo intorno.

Distance functions

Euclidean	$\sqrt{\sum_{i=1}^k (x_i - y_i)^2}$
Manhattan	$\sum_{i=1}^k x_i - y_i $
Minkowski	$\left(\sum_{i=1}^k (x_i - y_i ^q) \right)^{1/q}$

1.4 Analisi metabolomica di urine

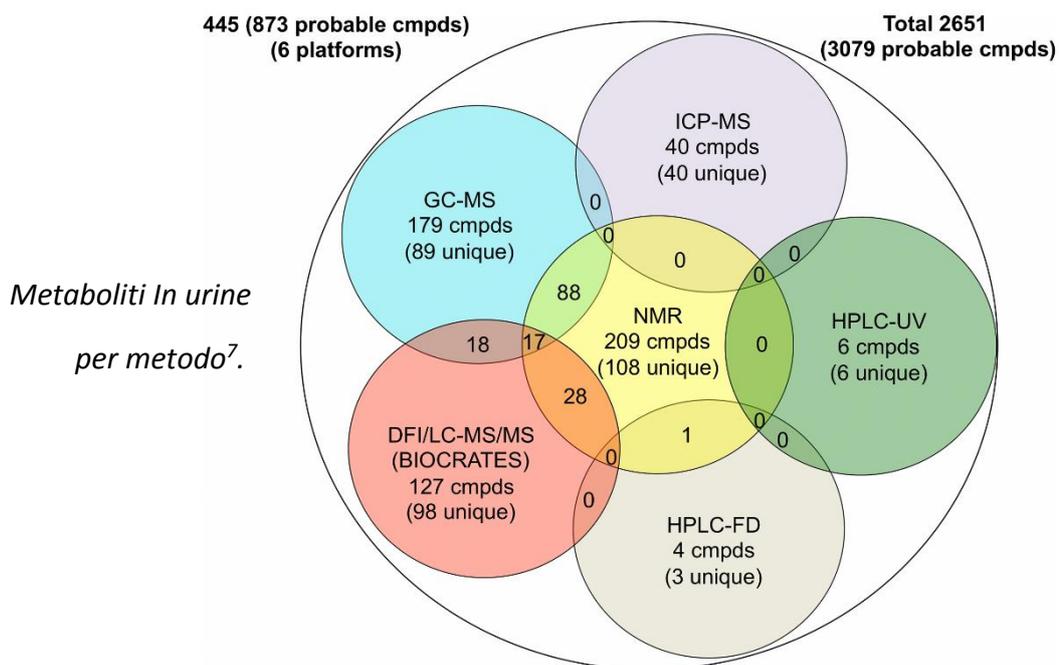
Le urine sono il biofluido più largamente utilizzato nelle ricerche metabolomiche. La raccolta di urine non è invasiva, non richiede una preparazione del paziente e sostanzialmente non presenta procedure complicate da eseguire.

I metaboliti nelle urine, come prodotto di normali e patologici processi cellulari, sono strettamente legate ai fenotipi individuali.

Per questa ragione sono stati fatti numerosi progressi verso l'utilizzo degli studi metabolici urinari come strumenti diagnostici.

Come materiale di scarto biologico, le urine contengono numerosi prodotti della decomposizione metabolica di cibi e liquidi ingeriti, farmaci, sostanze inquinanti e scarti endogeni. Non tutti questi composti sono stati caratterizzati o identificati e costituiscono quindi una sfida per l'analisi metabolomica.

La numerosità di prodotti presenti nelle urine (circa 2560 metaboliti confermati¹³) sono una fonte di ricchezza di dati e di informazione biochimica che, associata alla facilità nel reperimento e preparazione dei campioni fanno delle urine il substrato ideale per gli studi di fingerprinting.



1.5 Fingerprint metabolico

Un ambito di ricerca interessante della metabolomica consiste nella determinazione di un fenotipo metabolico individuale ricavato dagli insieme di metaboliti presenti in un biofluido che definiscono il fingerprint individuale.

Questa impronta metabolica è caratteristica per ogni organismo e differisce tra un individuo e l'altro e lo scopo della ricerca metabolica è quello di caratterizzare tale fingerprint, studiarne le proprietà e determinare l'influenza su di esso di fattori genetici e ambientali.

Il profilo metabolico individuale (metabotipo, metabotype in inglese) viene definito come "una descrizione multiparametrica dello stato fisiologico di un individuo basata su dati metabolomici"¹⁴.

Questi "metabotypes" possono essere rappresentati come l'immagine speculare dei processi cellulari e i cicli metabolici di un individuo in un determinato stato fisiopatologico.

Le difficoltà nella determinazione dei fingerprint metabolici consiste nelle numerose fluttuazioni delle concentrazioni dei metaboliti dovute a svariati fattori legati alle condizioni di vita dell'individuo.

Solo recentemente infatti si sono raccolte evidenze sperimentali che confermano le ipotesi di una presenza di un'impronta genica individuale^{15,16,17}.

La conferma di un profilo metabolico identificabile, soprattutto nei campioni di urina, ha permesso lo sviluppo di numerose applicazioni in campo medico-biologico.

Il fingerprint metabolico in quanto fotografia di uno stato fisiopatologico può essere utilizzato come strumento per studiare i meccanismi biochimici delle patologie, individuare l'insorgenza di una malattia, fare diagnosi, prognosi e stratificazione dei pazienti.

1.6 Applicazioni della metabolomica

La metabolomica ha numerose applicazioni differenti in quanto potendo usufruire sia di strumenti statistici sia bioinformatici può rivelarsi uno strumento efficiente nella determinazione dei processi biochimici presenti in un organismo.

Fornendo l'immagine dei cicli metabolici di un individuo svolge un ruolo fondamentale nella comprensione dell'integrità strutturale di un organismo e del suo stato fisiologico, per questa ragione può essere utilizzata al fine di determinare e caratterizzare gli effetti di processi patologici ma anche la tossicità di un farmaco¹⁸.

Anche attraverso l'utilizzo di metodi statistici trova spazio nella diagnostica in campo medico soprattutto nel riconoscimento sano-malato e nella successiva comprensione del meccanismo di sviluppo di una malattia¹⁹.

Fino a questo punto la metabolomica è stata in grado di fornire numerose informazioni in campo medico sullo studio di malattie quali cancro, celiachia, diabete e disturbi sia respiratori che intestinali²⁰.

Grazia alla ricchezza di informazioni fornita, non solo è possibile differenziare il sano dal malato, ma anche prevedere l'insorgenza di una malattia.

Recentemente, data la possibilità di riconoscere un'impronta molecolare, la metabolomica ha trovato sviluppo anche nel campo agro-alimentare. La varietà dell'analisi metabolomica fornisce non solo gli strumenti necessari alla determinazione della qualità di un prodotto ma anche della sua provenienza.

Ad esempio in uno studio condotto nell'istituto dove è stato svolto questo tirocinio sono stati raccolti dati interessanti che permettono di identificare perfettamente la stalla di provenienza a partire da un campione di latte bovino.

2 Progetto MetRef

Il progetto MetRef¹⁶ (metabolomic reference) è un lavoro di ricerca dei professori Bertini e Luchinat cominciato nel 2005 allo scopo di determinare la presenza di un fingerprint metabolico in campioni di urine umane e studiarne la stabilità.

2.1 MetRef 1

Il primo studio sul fenotipo metabolico individuale iniziò nel 2005 con la raccolta di 40 campioni di urine (raccolti giornalmente e di mattina) per 22 partecipanti al progetto (880 campioni).

La scelta di 40 campioni per ciascun individuo si fondava sulla creazione di una base statistica sufficientemente ampia necessaria al riconoscimento del fenotipo individuale.

Gli studi condotti tramite analisi NMR rivelarono, come si osserva dall'immagine sottostante (figura 1), che nonostante le fluttuazioni giornaliere del metaboloma i campioni di ogni individuo si raggruppavano secondo il rispettivo donatore.

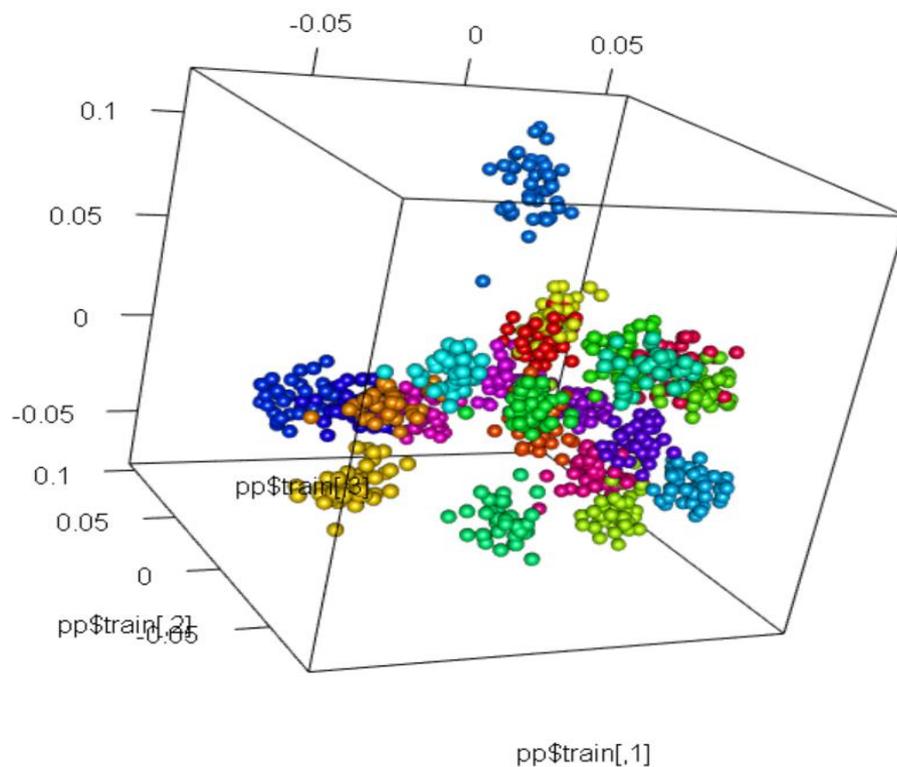


Figura 1. PCA/CA Individui MetRef 1

Lo studio ha dato informazioni importanti sul riconoscimento individuale fornendo materiale sufficiente ad ipotizzare la presenza di una parte stabile del profilo metabolico, identificabile come fingerprint individuale.

La tabella in figura 2 rappresenta l'accuratezza con la quale sono stati identificati gli individui.

L'elevate percentuali di riconoscimento dell'individuo forniscono non solo informazioni sulla stabilità del profilo metabolico individuale ma anche sulla precisione della sua individuazione

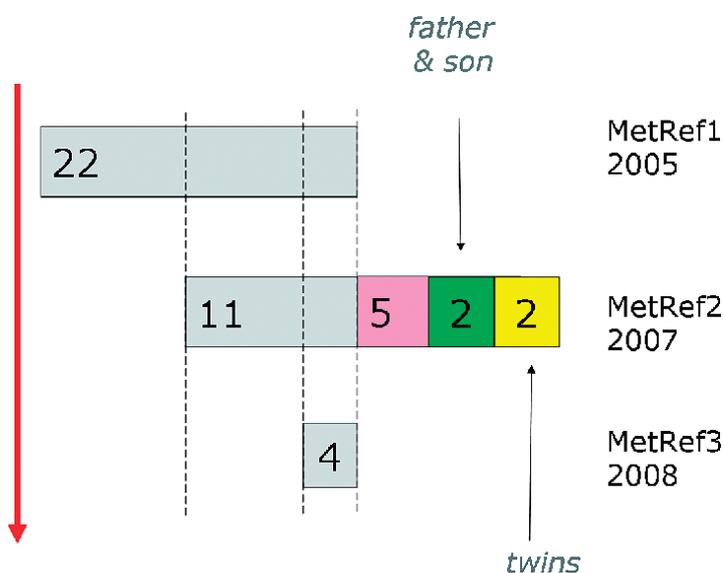
Lo studio ha inequivocabilmente dimostrato l'esistenza di un profilo metabolico individuale¹⁶ che, nonostante la variabilità giornaliera dovuta a dieta e stile di vita, permette di identificare ciascun individuo con elevata accuratezza. Tuttavia il periodo di raccolta di 40 giorni non è stato abbastanza prolungato per stabilire la stabilità nel tempo del fenotipo individuale.

AD	99.3
AF	100.0
AG	99.7
AH	95.2
AI	100.0
AO	100.0
AP	100.0
AR	100.0
AS	100.0
AT	100.0
AU	100.0
AW	100.0
AX	100.0
AZ	98.1
BC	100.0
BD	99.6
BE	100.0
BF	100.0
BG	98.5
BH	100.0
BI	100.0
BK	99.3

a

Figura 2. Percentuali accuratezza Individui MetRef 1

2.2 Metref 2-3



Lo studio successivo iniziò nel 2007 e si concluse nel 2008 e si è basato sulla raccolta di 40 campioni di urina donati da 20 individui sani in un'età compresa tra 25 e i 55 anni.

Di questi individui solo 15 avevano partecipato alla precedente raccolta.

Lo scopo dello studio era quello di determinare se il fenotipo metabolico fosse stabile in un breve periodo di circa 3 anni.

Il primo passaggio è stato quello di utilizzare i metodi adoperati nell'analisi precedente al fine di determinare se fossero ancora riconoscibili gli individui nella loro unicità.

Le analisi confermarono i precedenti risultati sulla presenza di un fenotipo metabolico individuale.

Qui sotto sono riportate le percentuali di accuratezza per il riconoscimento individuale.

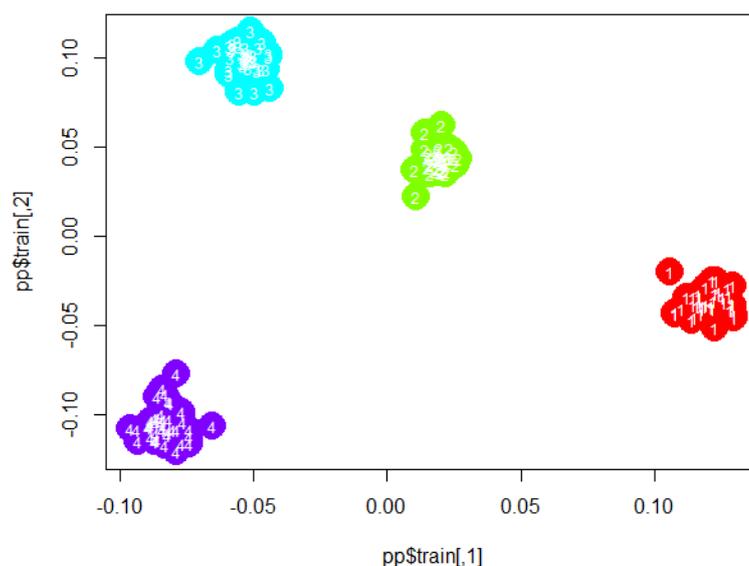
Accuratezza MetRef 2007

AI	100.0
AO	100.0
AR	99.5
AS	100.0
AU	100.0
AW	98.8
BC	95.9
BF	99.5
BG	99.3
BH	100.0
BI	99.2
BQ	99.3
BS	97.7
BT	100.0
BU	99.3
BV	100.0
BX	99.0
BZ	100.0
TA	97.1
TB	97.1

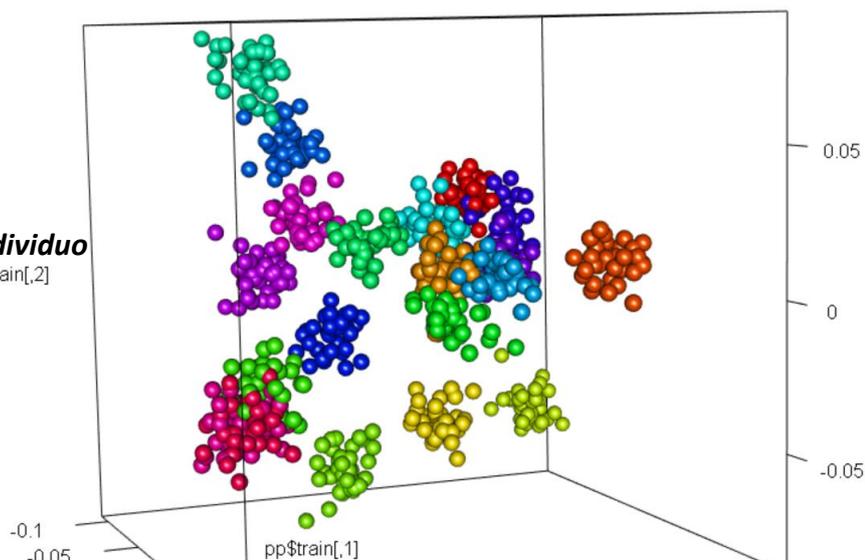
Accuratezza MetRef 2008

AR	100.0
AS	99.9
AU	100.0
AW	100.0

PCA/CA MET3



PLS/CA MetRef 2007 per individuo



Successivamente sono state condotte analisi per determinare se l'individuo fosse riconoscibile nel tempo, ovvero stabilire se il fingerprint metabolico si conservasse inalterato negli anni.

A tale scopo è stata utilizzata una predizione statistica per verificare che i nuovi campioni acquisiti fossero riconosciuti sulla base di un confronto con i campioni precedentemente raccolti.

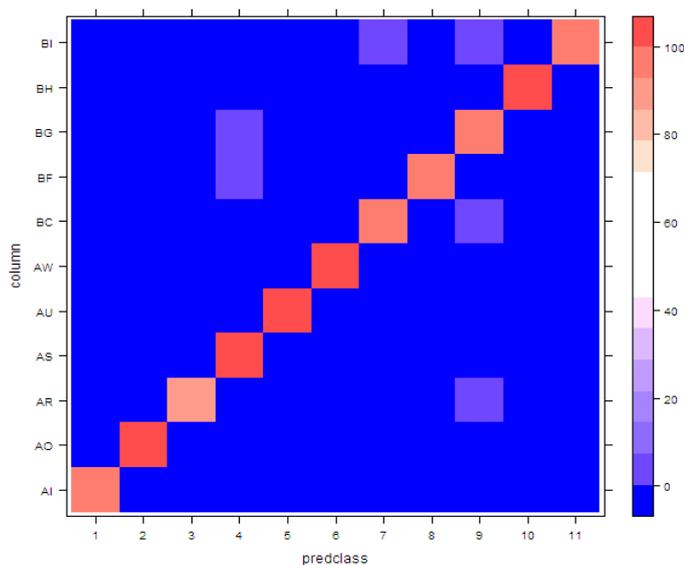
L'immagine sottostante rappresenta il levelplot sulla predizione del gruppo di individui di MetRef 1 contro quelli di MetRef 2 (dal grafico sono stati esclusi quelli non a comune).

Dalle percentuali di predizione venne confermato¹⁶ che esiste un fenotipo metabolico individuale e che questi non varia sensibilmente nel breve periodo.

AI	97.3
AO	100.0
AR	89.2
AS	100.0
AU	100.0
AW	100.0
BC	97.3
BF	97.3
BG	97.3
BH	100.0
BI	97.3

Tabella percentuali predizioni metref1

Contro metref 2

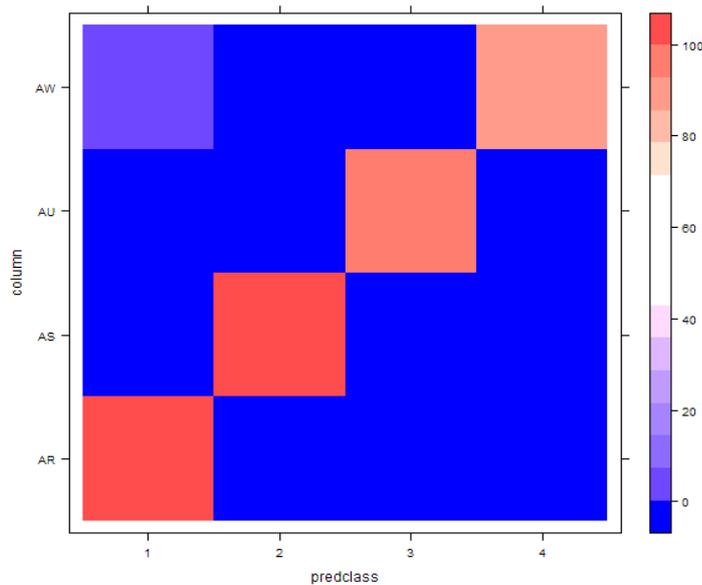


Levelplot predizione individui metref1 contro metref2

Allo stesso modo anche le predizioni²² sugli individui di MetRef1 contro MetRef3 portano agli stessi risultati.

	AR	AS	AU	AW	
AR	100	0	0.0	2.5	media:96.975
AS	0	100	0.0	0.0	
AU	0	0	97.5	0.0	
AW	0	0	0.0	90.0	

Tabella percentuali MetRef3 contro MetRef4



Levelplot MetRef 3 contro MetRef 4

Oltre a fornire la conferma della presenza di un fingerprint metabolico stabile questo lavoro di ricerca ha dato ulteriori informazioni interessanti.

I risultati ricavati dai dendogrammi (grafici che mettono in relazione gli individui per similitudine del fingerprint individuale) mostrano come la coppia di gemelli omozigoti TA e TB siano gli individui più vicini tra loro

(altezza della forcina che li divide). Questo ha portato all'ipotesi che il corredo genetico contribuisca in maniera determinante al fingerprint, tuttavia non è la sola variabile importante, altrimenti sarebbero indistinguibili. Ne consegue che i fattori esterni sono altrettanto importanti.

Questo dimostra che la conoscenza del solo genoma, come riportato nel dogma centrale della biologia molecolare, non basti ad identificare un individuo o spiegare i processi cellulari in atto e che fattori esterni (stile di vita, dieta o altro) e interni (espressione genica) possono influire simultaneamente sullo sviluppo di un organismo.

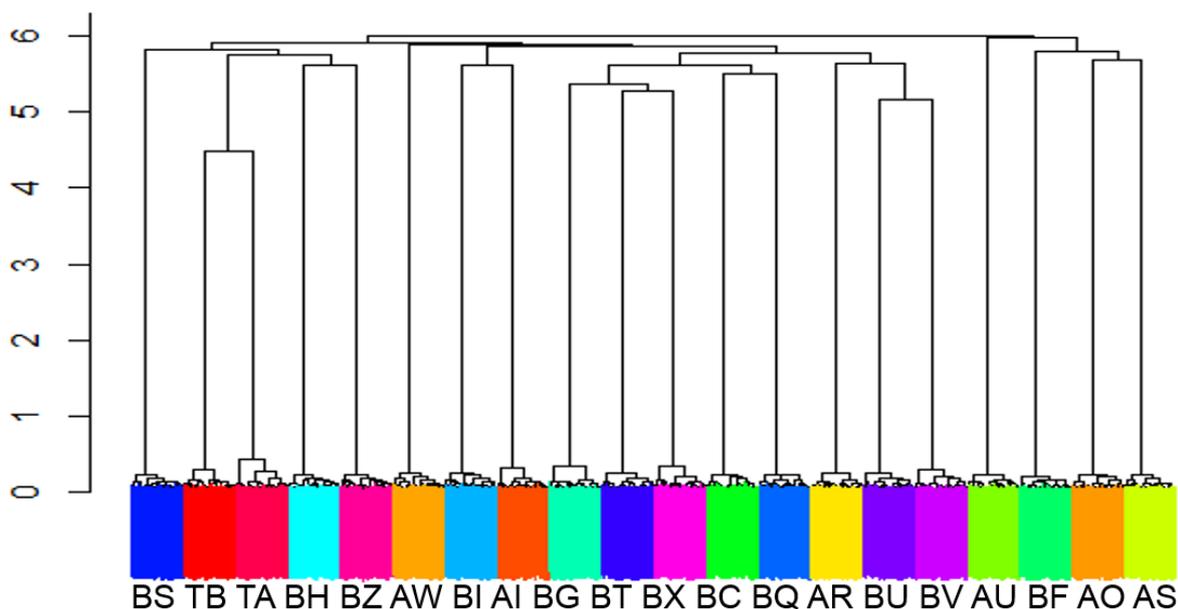


Fig.10 Dendrogramma individui metref 2007

Si osserva in ugual modo come anche BU e BV, legati dalla relazione padre-figlio, presentano spettri più vicini tra loro rispetto agli altri individui mantenendo però una certa distanza ad immagine della loro unicità individuale.

3 Organizzazione e scopo del lavoro

Il presente lavoro si inquadra nel più ampio progetto chiamato MetRef condotto dal gruppo di ricerca di metabolomica afferente al Professor

Claudio Luchinat presso il centro di risonanze magnetiche (CERM) dell'università degli studi di Firenze.

Tale progetto prevede lo studio di campioni di urina umani con tecnica NMR al fine di determinare un profilo metabolico individuale e studiarne la stabilità e i fattori di variabilità.

In questo lavoro di tesi sono presentati i risultati ottenuti al fine di determinare se il fingerprint metabolico sia stabile nel tempo.

4 Materiali e Metodi

4.1 Collezione dei campioni

Per questo progetto sono stati collezionati 20 campioni di urina da 12 volontari sani partecipanti al progetto in un'età compresa tra i 25 e i 70 anni.

Ogni campione è stato raccolto al primo mattino di ogni giorno in contenitori sterili e poi conservato ad una temperatura compresa tra i 2 e gli 8 gradi centigradi per un massimo di due ore.

Successivamente prima di essere depositati alla temperatura di -80°C nel congelatore della Da Vinci european Biobank i campioni sono stati sottoposti ad una centrifugazione (2500g per 5 minuti alla temperatura di 4°C) e filtrazione attraverso una membrana con pori di 0,20 nanometri, allo scopo di eliminare i prodotti della degradazione cellulare²¹ che possono costituire un elemento di interferenza nelle analisi successive.

4.2 Preparazione dei campioni

Prima del trattamento ogni campione estratto dal congelatore è stato fatto scongelare a temperatura ambiente e successivamente agitato per ottenere la massima uniformità.

Dopo la fase di scongelamento un'aliquota di 630 μL veniva prelevata dal campione e centrifugata a 14000g per 5 minuti.

540 dei 630 microlitri prelevati sono stati miscelati con 60 μL di un buffer di fosfato di potassio (1.5 M KH_2PO_4 in D_2O con 10mM di trimetilsilil

propionato di sodio(TMSP) ad un pH di 7.4) in modo tale da ottenere un rapporto di 9:1 di urina-solvente.

Della miscela ottenuta vengono estratti 450 μ L e riposti in tubicini da 4,25mm(BRucker biospin srl) necessari per il proseguo delle analisi con NMR²¹.

4.3 Acquisizione degli spettri

Per ogni campione di urina sono stati acquisiti spettri ¹H NMR utilizzando uno spettrometro Bruker 600Mhz operante alla frequenza di Larmor per il protone di 600.13 MHz.

Lo spettrometro è dotato di un'unità di tuning and matching automatica(ATM), di un *cryoprobe* CPTCI ed di un campionatore automatico.

Una termocoppia BTO 2000 garantisce la stabilità della temperatura di acquisizione con una sensibilità di 0.1K che permette di lavorare alla temperatura prestabilita.

Ogni campione viene conservato per tre minuti, prima dell'acquisizione, nel prebehead dello strumento ovvero un'anticamera, per equilibrare il campione alla temperatura di esercizio dello strumento che nel caso di questo lavoro di tesi è di 300K.

Per ogni campione di urina è stato acquisito uno spettro ¹H NMR con sequenza NOESY-presaturation. Sono state effettuate un totale di 64 scansioni con 64000 data points con una larghezza spettrale di 12019Hz, tempi di acquisizione di 2.7 s, tempi di rilassamento di 4s e mixing time di 0.1 s.



4.4 Processing degli spettri

Le FIDs (free induction decays) sono state moltiplicate per una funzione esponenziale equivalente ad un allargamento di 1hz prima dell'applicazione della trasformata di fourier.

Per le distorsioni della baseline e della fase gli spettri sono stati corretti automaticamente e successivamente calibrati rispetto al picco a 0.00ppm del TMSP usando *Topspin 3.2*(*Brucker Biospin srl*).

Escludendo le zone dello spettro comprese tra $\delta=4.5-6.0$ ppm ovvero quelle contenenti i picchi dell'urea e dell'acqua e gli estremi con $\delta<0.5$ e $\delta>9.5$, le restanti parti sono state suddivise in bins (intervalli di chemical shift) di 0.02ppm. Le corrispondenti aree spettrali sono state poi integrate utilizzando il software AMIX (*Brucker Biospin srl*).

Le variazioni di pH sono la maggior fonte di spostamento dei picchi perciò ogni campione è stato sottoposto a procedure standardizzate per evitare interferenze legate al pH.

Nell'analisi multivariata sono stati utilizzati bin di 0.02ppm allo scopo di limitare la maggior parte questi *chemical shift*.

4.5 Analisi statistica

La matrice di dati ottenuta è stata analizzata con metodi chemiometrici applicati utilizzando R, un *software open source per analisi statistica*, costituito da un insieme di applicazioni che possono essere utilizzate per la gestione e l'analisi dei dati e la produzione di grafici.

Allo scopo di ottenere una maggiore comprensione dei dati in nostro possesso fornendo allo stesso tempo una visione generale è stata utilizzata l'analisi delle componenti principali ovvero la PCA.

Sono state utilizzate anche tecniche *supervised* come la PLS per massimizzare la discriminazione tra individui combinandole talvolta con la CA usando la funzione standard di R "cancor".

Per ogni set di dati per individuo è stata applicata una classificazione utilizzando la tecnica KNN ($k=7$).

Le percentuali di accuratezza invece sono state ricavate attraverso il metodo chiamato cross-validazione.

Infine sono state ricavate le concentrazioni dei metaboliti integrando nelle regione spettrali relative i picchi delle molecole prese in esame.

Le intensità dei picchi sono proporzionali alle concentrazioni dei metaboliti e sono state analizzate in funzione delle diverse proprietà intrinseche che caratterizzano i campioni, così da ottenere una visione complessiva dell'andamento delle concentrazioni metaboliche in base ai parametri analizzati e avendo la possibilità di determinare se queste avessero subito cambiamenti significativi.

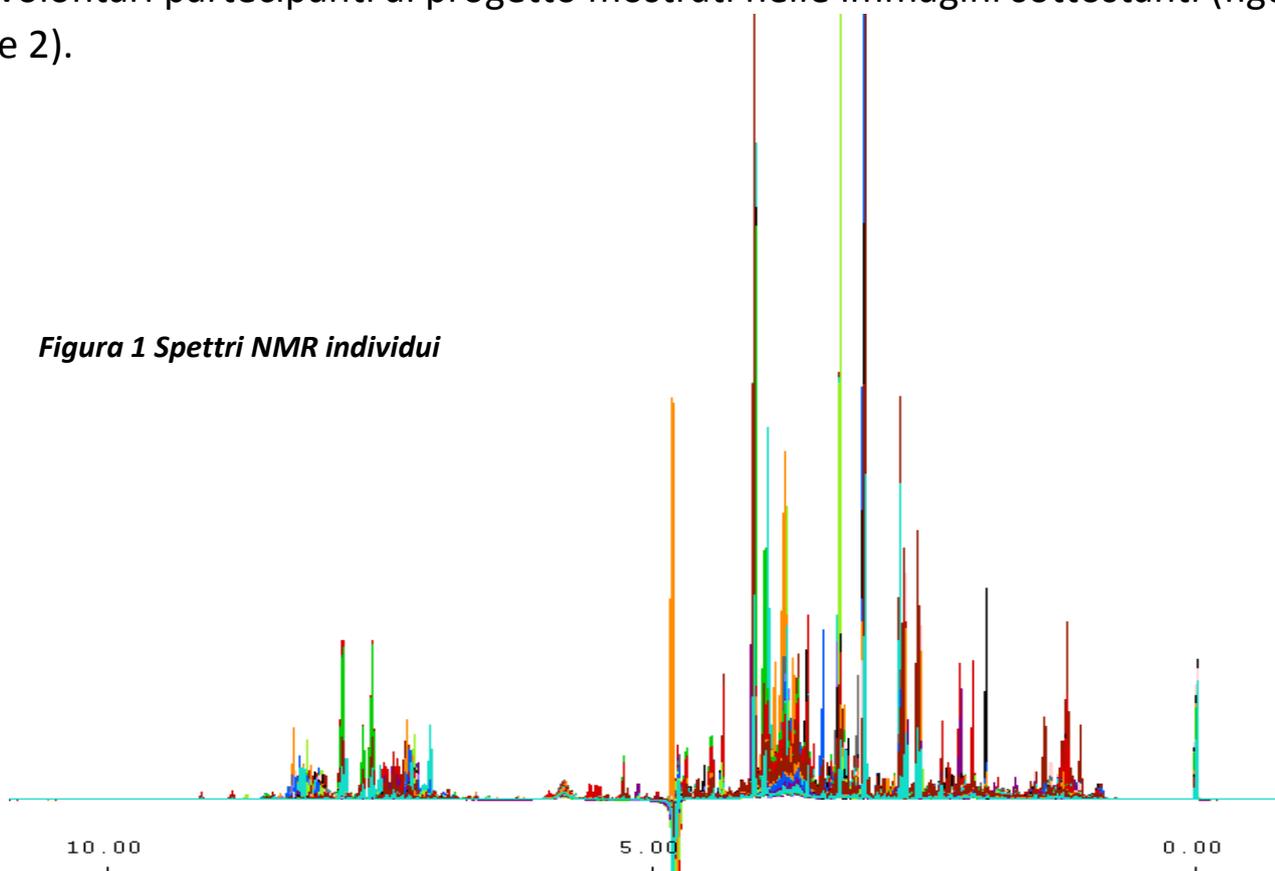
Un p-valore inferiore agli 0.05 è considerato statisticamente significativo.

5 Risultati e Discussione

5.1 Spettri NMR di Urina Umana

Nel corso di questo lavoro di tesi sono stati ottenuti 440 spettri NMR dei volontari partecipanti al progetto mostrati nelle immagini sottostanti (figura 1 e 2).

Figura 1 Spettri NMR individui



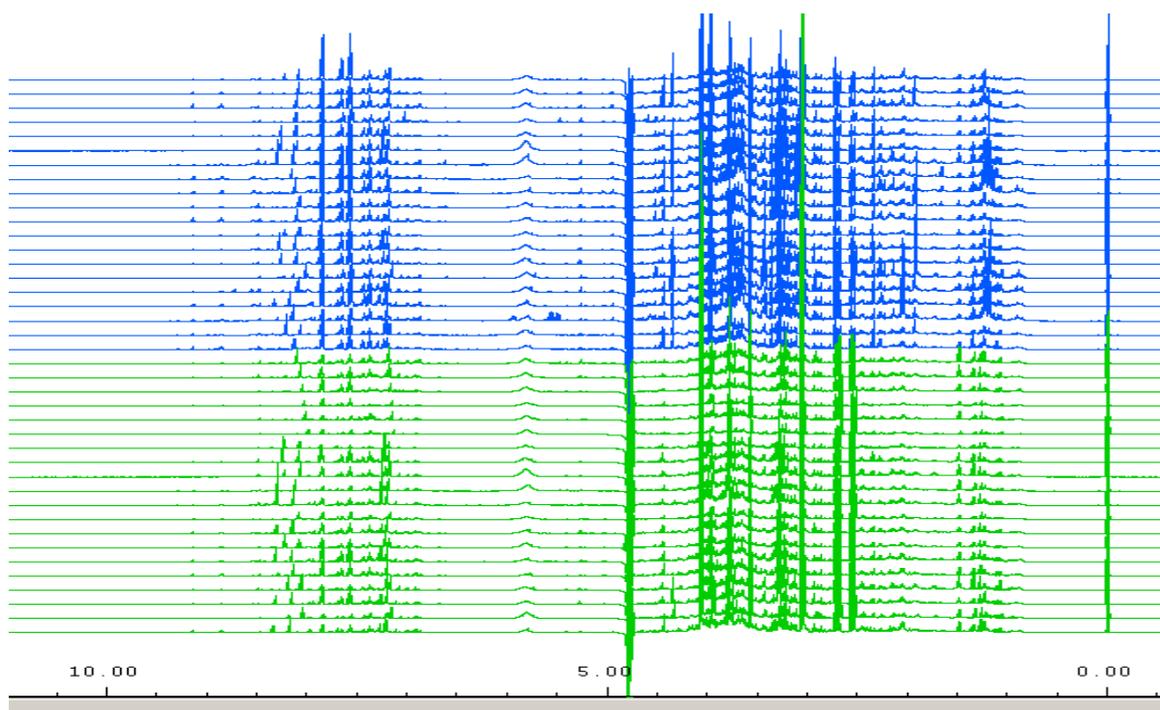


Figura 2. Spettri NMR individui

Questi spettri sono stati ottenuti da 12 volontari partecipanti al progetto i quali hanno partecipato ad almeno uno dei progetti precedenti. Il riassunto dello schema di campionamento usato nel corso degli anni è raffigurato nell'immagine sottostante (Figura 3)

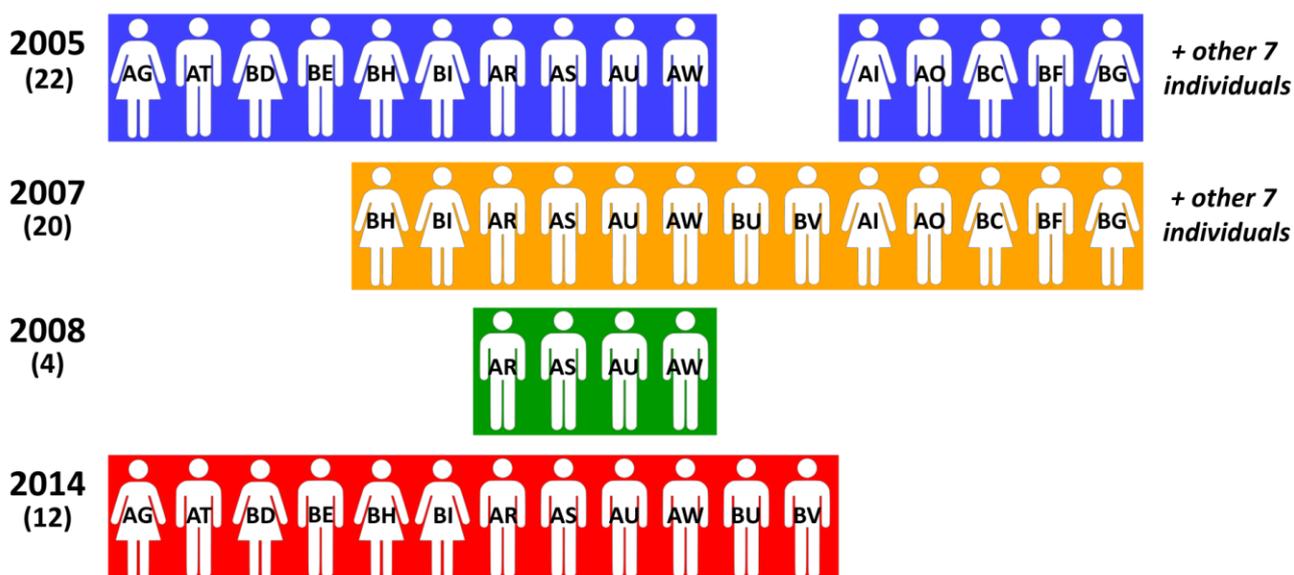


Figura 3. Volontari partecipanti ai 4 progetti

5.2 Analisi Metref 4

Il primo passaggio in questo lavoro di tesi è stato confermare la presenza di un fenotipo metabolico individuale negli spettri ricavati dalla raccolta del 2015.

In questo studio sono state utilizzate le stesse tecniche statistiche utilizzate nelle precedenti raccolte ovvero sia i metodi di analisi unsupervised come la PCA sia quelli supervised come la KNN.

Le prime tecniche di analisi utilizzate sono la PCA e la PLS, le quali forniscono una visione complessiva delle posizioni (presenza di cluster o outliers) dei nostri campioni. Le Figure 4 e 5 raffigurano le analisi effettuate ed evidenziano la presenza di alcuni outliers.

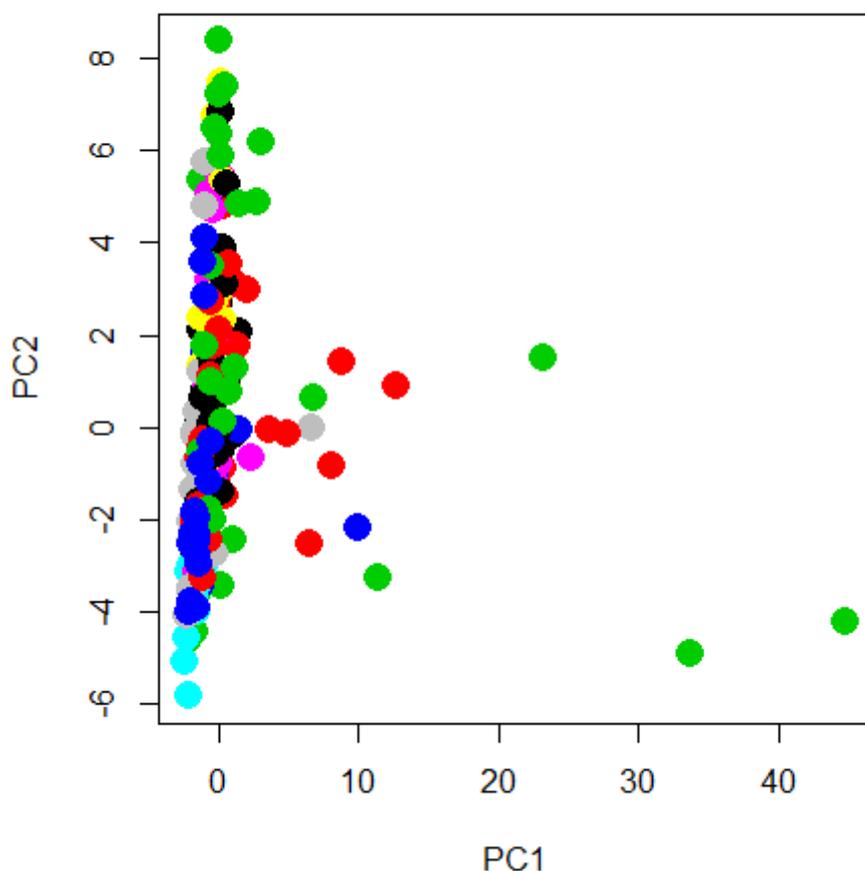


Figura 4. Grafico analisi PCA individui MetRef4

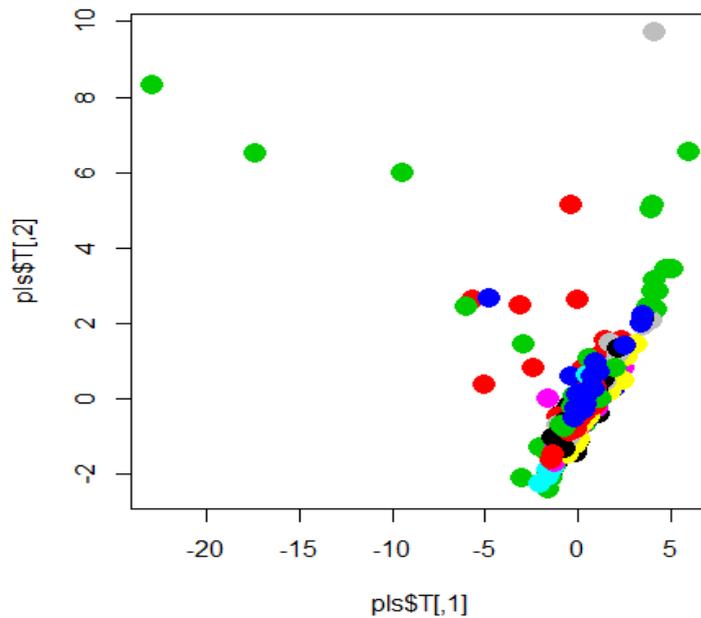


Figura 5. PLS Individui raccolta 2015

Da queste analisi si osserva che nonostante la presenza della maggior parte dei campioni in una zona ben delimitata, esistono degli outliers che si allontanano dalla ragione principale. Da questo tipo di analisi preliminare risulta evidente che le differenze intraindividuali sono rilevabili mentre quelle interindividuali non sono ben definite.

Passando a metodi supervised quali la PCA/CA si nota una clusterizzazione più definita che individua soprattutto le differenze tra campioni dello stesso individuo (figura 6).

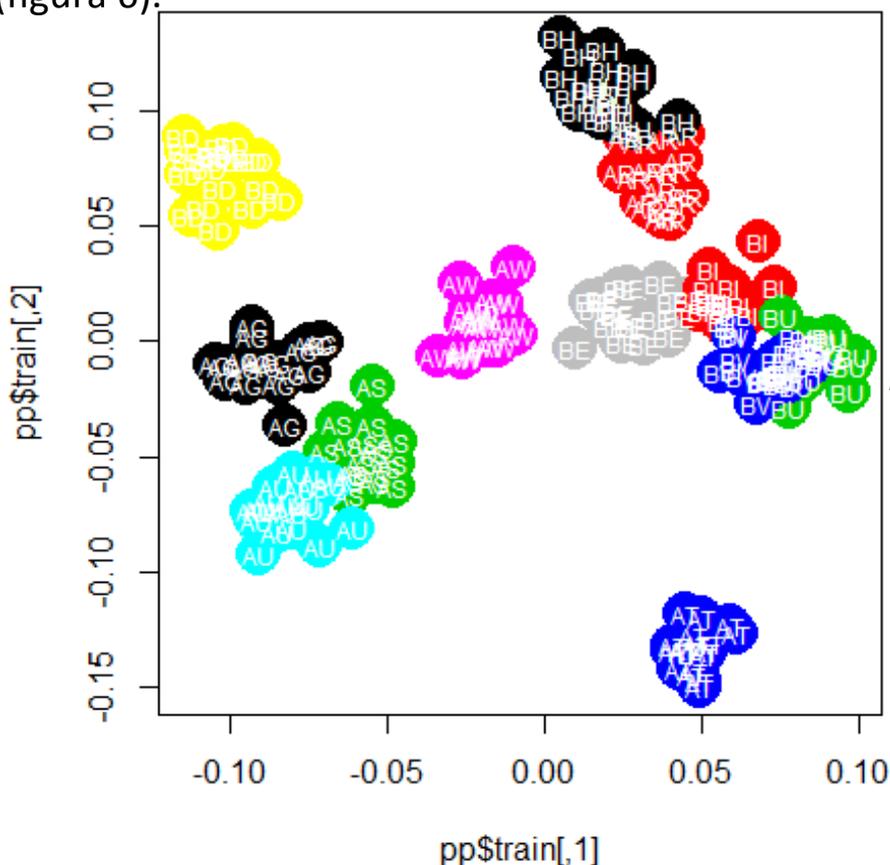


Figura 6. PCA/CA

Il grafico in figura 6 evidenzia la presenza di un fingerprint metabolico individuale che nonostante le possibili variazioni dello spettro metabolomico giornaliero, rimane ben definito.

Utilizzando una visualizzazione 3D del grafico precedentemente mostrato (figura 6) vengono evidenziate maggiormente le differenze interindividuali.

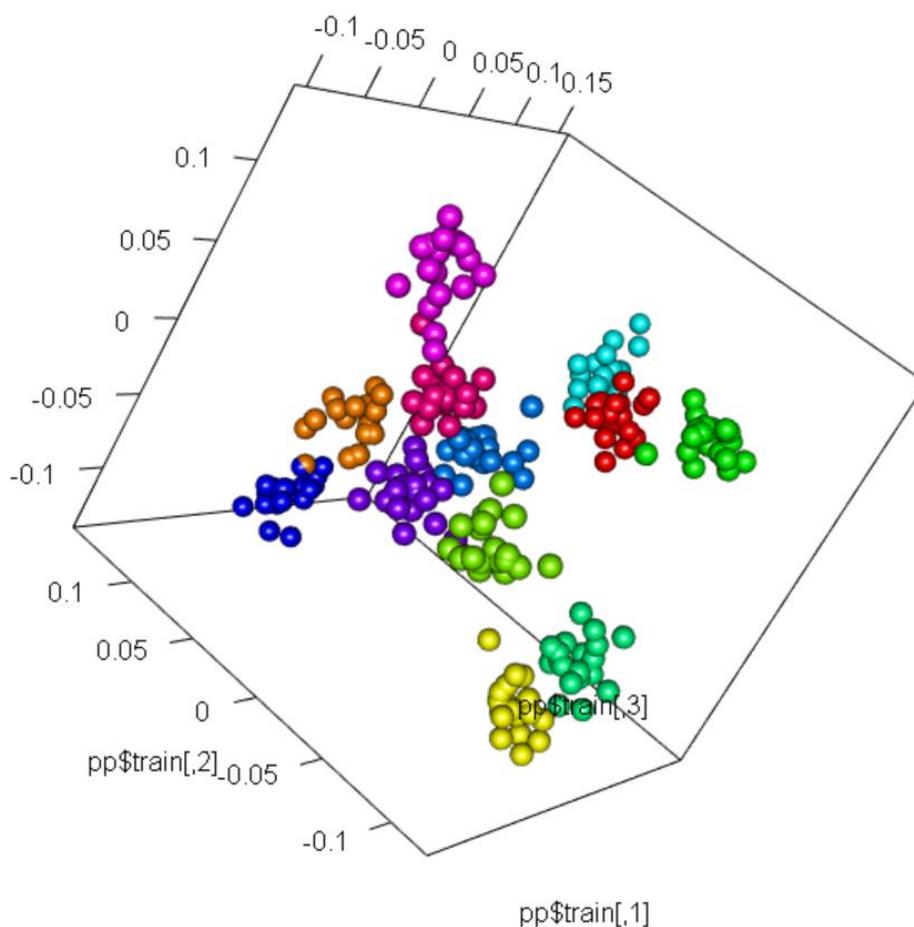


Figura 6. Grafico 3D PCA/CA individui MetRef4 per colore

Anche negli stessi raggruppamenti di campioni colorati per individuo si osservano degli outliers ovvero elementi che si discostano di più rispetto al resto del gruppo, ciò è dato dalle fluttuazioni che avvengono nel nostro

metaboloma ed esprimono quindi la mutabilità del profilo metabolico giornalmente.

Quindi si può affermare che lo spazio metabolico di ogni individuo rappresenta il concetto di un fenotipo metabolico flessibile come conseguenza metabolica dell'adattamento, a stimoli esterni.

Il nostro metabolismo ogni giorno è capace di adattarsi agli stimoli esterni al fine di mantenere le condizioni di equilibrio e questo si osserva nei cambiamenti delle concentrazioni dei metaboliti nello spettro NMR.

Dopo aver confermato visivamente le previsioni sulla presenza di un fenotipo metabolico già dimostrato nei precedenti progetti è stato necessario un calcolo sull'accuratezza del riconoscimento del singolo individuo.

Il metodo di analisi utilizzato è chiamato cross-validazione e fornisce le percentuali di assegnamento di ogni campione nell'intorno dello spazio metabolico di un individuo.

Il test statistico di cross-validazione sugli individui del MetRef4 è stato ripetuto 1000 volte con uno schema casuale (Montecarlo) in modo da ottenere la massima precisione statistica. I risultati sono riportati in Figura 7. Si osserva che le percentuali di riconoscimento sono elevate per ogni

AG	94.8
AR	94.7
AS	100.0
AT	100.0
AU	100.0
AW	100.0
BD	100.0
BE	100.0
BH	95.7
BI	100.0
BU	99.4
BV	100.0

individuo confermando l'ipotesi di un fingerprint metabolico stabile e individuabile con precisione.

Rispetto alle precedenti raccolte la media del riconoscimento è calata lievemente e si attesta intorno al 98.7%.

Quindi si può affermare che le variazioni giornaliere legate a stimoli esterni non influiscono in modo rilevante sulla determinazione dell'individuo che presenta quindi un fenotipo metabolico ben definito

5.3 Stabilità del fenotipo a lungo e breve termine

Lo scopo principale di questo lavoro di tesi è stabilire se il fenotipo metabolico individuale sia stabile nel tempo.

Per adempiere a questo obiettivo è necessario ricorrere ad analisi statistiche in grado di determinare le percentuali di riconoscimento dei campioni durante le raccolte.

Queste analisi chiamate "predizioni" confrontano i campioni degli individui di due o più raccolte assegnandoli all'intorno più vicino al campione stesso.

Mettendo a confronto due o più gruppi diversi di raccolta si includono anche gli individui non a comune fornendo quindi una maggiore probabilità di errore.

Il primo confronto effettuato è stato con gli individui della terza raccolta MetRef. Si nota dalla tabella in figura 8 come le percentuali di riconoscimento a breve termine rimangono soddisfacenti come nei precedenti risultati ottenuti confrontando MetRef2 e MetRef3 contro la prima raccolta del 2005.

	AR	AS	AU	AW
AR	90	5	0	0
AS	5	95	0	0
AU	5	0	100	0
AW	0	0	0	100

Figura 8. Percentuali predizioni MetRef 2008 contro MetRef 2014

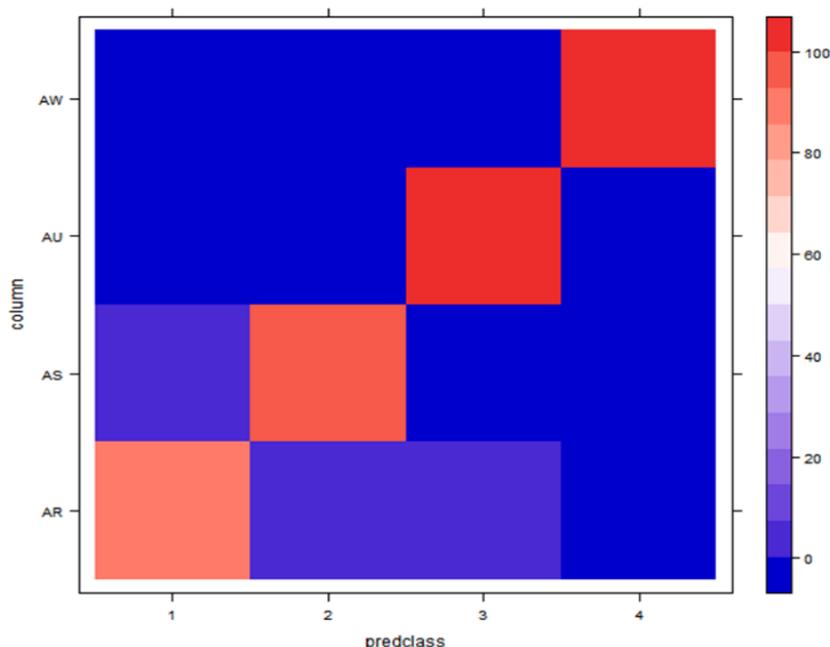


Figura 9. Levelplot predizioni (figura 8)

Con una media intorno al 96.25 percentuale si può affermare che il fenotipo metabolico è stabile nel breve periodo ed è riconoscibile con una buona approssimazione. Il grafico che si osserva in figura 9 rappresenta il levelplot degli individui di MetRef3 contro quelli di MetRef4 ed esprime visivamente le percentuali di riconoscimento.

Procedendo nello studio della stabilità a breve termine sono state elaborate le percentuali e quindi il grado di riconoscimento degli individui a comune tra il progetto svoltosi nel 2007 e quello attuale del 2015.

Dalla figura 10 (riportata sotto) si osserva che il riconoscimento individuale peggiora sensibilmente soprattutto per gli individui AW, BU e BV.

	AR	AS	AU	AW	BH	BI	BU	BV
AR	90	0	0	0	0	0	0	0
AS	0	90	5	0	0	0	10	0
AU	0	0	95	0	0	0	5	0
AW	0	0	0	10	0	0	0	0
BH	5	0	0	0	95	0	0	0
BI	0	0	0	0	0	100	5	0
BU	0	0	0	0	0	0	70	0
BV	0	0	0	0	0	0	0	85

Figura 10. Tabella riconoscimento MetRef2 contro MetRef4

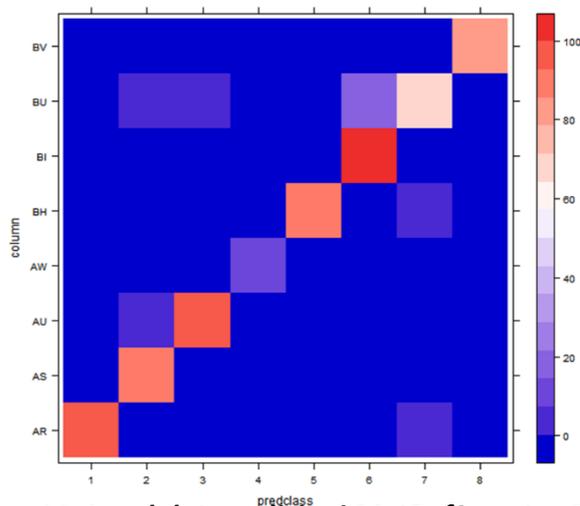


Figura 11. Levelplot predizioni MetRef2 contro MetRef4

Le cause dello scarso riconoscimento individuale sono da ricercare nelle loro condizioni fisiopatologiche.

Per l'individuo BV la percentuale di assegnazione è ancora accettabile ai fini statistici e potrebbe derivare sia da un errore umano o strumentale ma anche da una lieve patologia che altera moderatamente il profilo metabolico. Per il volontario AW lo scenario è completamente diverso e verrà quindi analizzato successivamente (cap. 5.3).

Dopo aver svolto lo studio del fingerprint per il breve periodo è stato elaborato anche quello sul lungo periodo ovvero il confronto tra i campioni raccolti nel 2005 con quelli raccolti nel 2014-2015.

I dati raccolti espressi in Figura 11 evidenziano come per alcuni individui il riconoscimento metabolomico è elevato (anche 100%) mentre per altri, rispettivamente BD, AG e AW, è estremamente basso.

	AG	AR	AS	AT	AU	AW	BD	BE	BH	BI
AG	45	0	0	0	0	0	0	0	0	0
AR	0	95	0	0	0	0	0	0	0	0
AS	0	0	90	0	0	10	0	0	0	0
AT	0	0	0	100	0	0	0	0	0	0
AU	5	0	0	0	100	0	0	0	0	0
AW	0	0	0	0	0	55	0	0	0	0
BD	0	0	0	0	0	15	35	0	0	0
BE	0	0	0	0	0	0	0	100	0	0
BH	0	5	0	0	0	0	0	0	95	0
BI	0	0	0	0	0	0	0	0	0	100

Figura 12. Tabella percentuali riconoscimento MetRef4 contro MetRef1

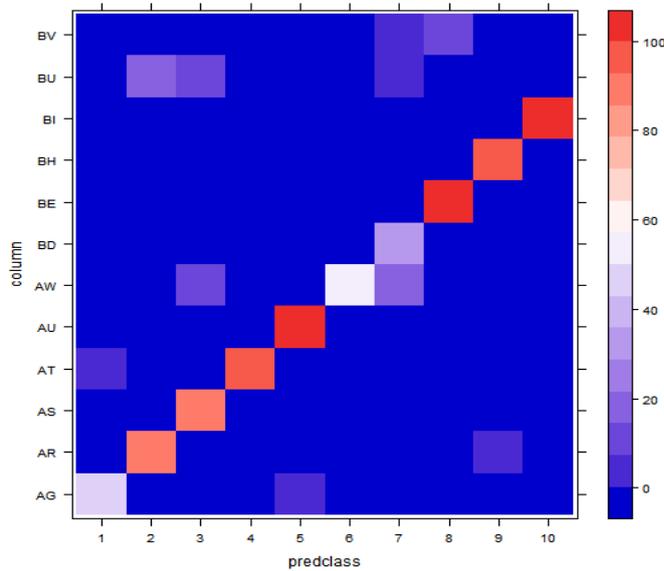


Figura 13. Levelplot predizioni MetRef4 contro MetRef1

Da questi risultati si può affermare che il profilo metabolico è generalmente stabile e quindi identificabile nel lungo periodo.

Le cause della scarsa riconoscibilità di determinati individui sarà da ricercare nelle variazioni delle concentrazioni dei metaboliti e potranno essere spiegate solo conoscendo lo stato fisiopatologico dell'individuo al momento della raccolta. Per questo motivo ad ogni volontario partecipante al progetto è stato chiesto di compilare un questionario informativo al fine di conoscere le condizioni dell'individuo al momento della raccolta e se nel corso degli anni trascorsi tra una raccolta e l'altra fossero avvenuti eventi (malattie, cambio di residenza ecc..) che avessero condizionato il loro stato fisiologico. Per ottenere una maggiore visuale dell'accuratezza sul riconoscimento è stata effettuata una predizione degli individui di MetRef4 contro tutti i campioni delle raccolte precedenti ottenendo i risultati riportati in figura 14 e 15.

	AG	AR	AS	AT	AU	AW	BD	BE	BH	BI	BU	BV
AG	35	0	0	0	0	0	0	0	0	0	0	0
AR	0	95	5	5	0	0	0	0	0	0	5	0
AS	0	0	95	0	0	0	0	0	0	0	10	0
AT	0	0	0	95	0	0	0	0	0	0	0	0
AU	5	0	0	0	100	0	0	0	0	0	0	0
AW	5	0	0	0	0	40	0	0	0	0	0	0
BD	0	0	0	0	0	0	25	0	0	0	0	0
BE	0	0	0	0	0	0	0	100	0	0	0	0
BH	0	5	0	0	0	0	0	0	100	0	0	0
BI	0	0	0	0	0	0	0	0	0	100	0	0
BU	0	0	0	0	0	0	0	0	0	0	75	0
BV	0	0	0	0	0	0	0	0	0	0	0	70

Figura 14. Tabella percentuali predizioni MetRef4 contro Totale

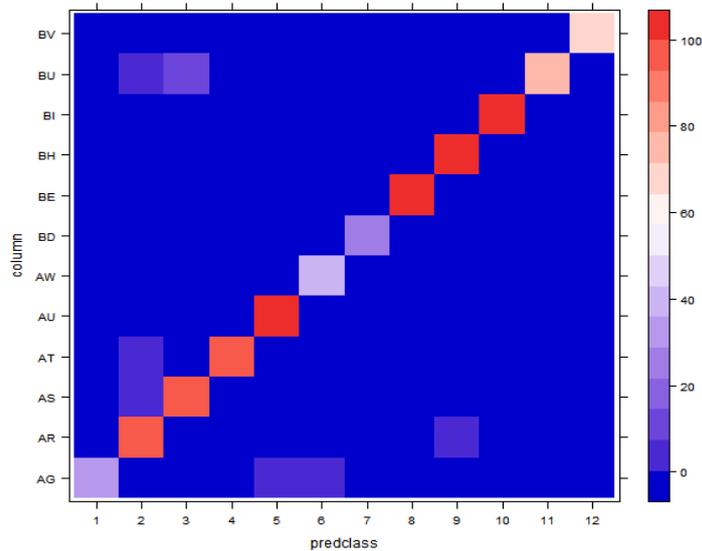


Figura 15. Levelplot predizioni MetRef4 contro Totale

Dal grafico in 3D della PCA/CA (figura 16) effettuata sugli individui che hanno partecipato a tutte le raccolte viene evidenziato, attraverso i raggruppamenti di campioni (colorati in modo differente), come per gli individui stabili nel tempo tutti i cluster sono vicini tra loro mentre per quelli non stabili (AW) si vengono a formare gruppi più distanti (AW 4).

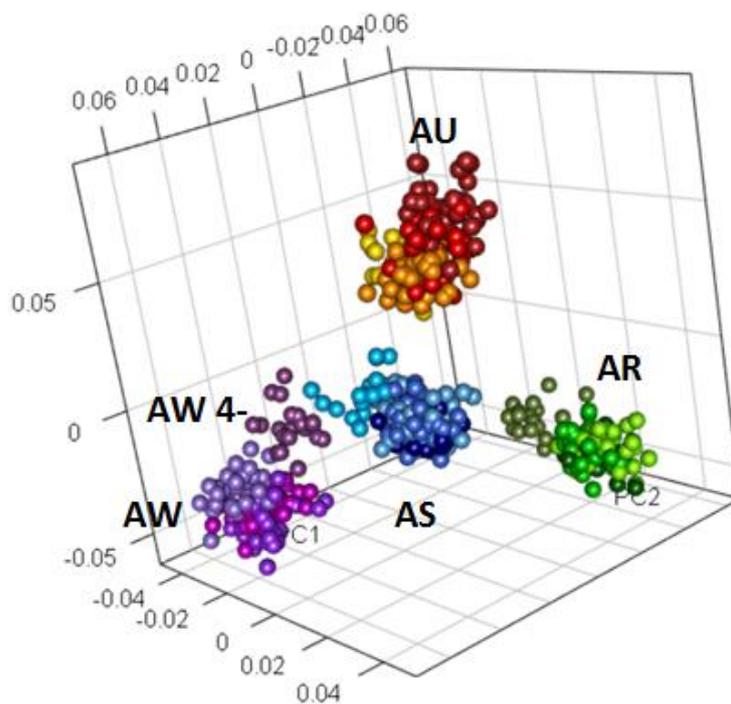


Figura 16. PCA/CA individui a comune

L'individuo AW presenta una clusterizzazione dei campioni raccolti nei MetRef del 2005 e del 2008 mentre quelli esaminati nel 2015 si distanziano, facendo presumere che negli anni trascorsi dal 2008 al 2014 il soggetto AW abbia subito un'alterazione del suo stato fisiologico.

5.4 Analisi delle stabilità e variabilità individuali

Le analisi statistiche utilizzate per la determinazione delle percentuali di riconoscimento individuale sul lungo e breve periodo hanno evidenziato che non tutti i partecipanti al progetto MetRef hanno mantenuto il loro profilo metabolico invariato nel tempo.

Attraverso il questionario fornito ai partecipanti abbiamo appreso maggiori informazioni sulle loro vite durante gli anni trascorsi tra le varie raccolte.

Nel corso dei 10 anni non sono avvenuti, per la maggior parte, cambiamenti significativi.

La maggior parte infatti ha mantenuto lo stesso posto di lavoro (eccetto BE), nessuno si è trasferito (a parte AS e BV che hanno cambiato casa nella stessa città e BE che si è trasferito ma apparentemente senza nessuna variazione nel suo profilo metabolico) o ha cambiato la sua dieta (solo BV e AG hanno modificato l'apporto di carboidrati nella loro alimentazione).

Quindi nessuno dei partecipanti ha riportato cambiamenti significativi alla propria vita durante l'arco di 10 anni a parte AG, AW e BD che sono quelli che presentano le più basse percentuali di riconoscimento.

Negli studi precedenti AW era sempre ben riconoscibile e il suo profilo metabolico era stabile nonostante nel periodo tra il 2005 e il 2008 si fosse trasferito dal sud al nord Europa. Quindi si può stabilire che i cambiamenti di residenza, dieta o stile di vita influiscono in maniera ridotta sul profilo metabolico di un individuo^{23,24}. Questo stesso individuo AW non ha riscontrato nessun cambiamento fisiopatologico rilevante sino alla primavera del 2014, quando dopo un trattamento antibiotico consistente è diventato intollerante al lattosio. Non è ben chiaro se i due eventi siano correlati tra loro ma sicuramente gli antibiotici possono influire sulla capacità dell'intestino di produrre l'enzima lattasi.

È oltretutto appurato che l'assunzione di antibiotici porta alla modifica della microflora intestinale^{25,26} che è direttamente correlata alle concentrazioni dei metaboliti delle urine e quindi al metaboloma^{27,26}.

Per comprendere al meglio le variazioni del metaboloma di AG è stato necessario proseguire con l'analisi dei metaboliti.

In questo studio di profiling sono stati individuati 28 metaboliti riportati nella tabella sottostante (figura 17).

Figura 17. Tabella Metaboliti

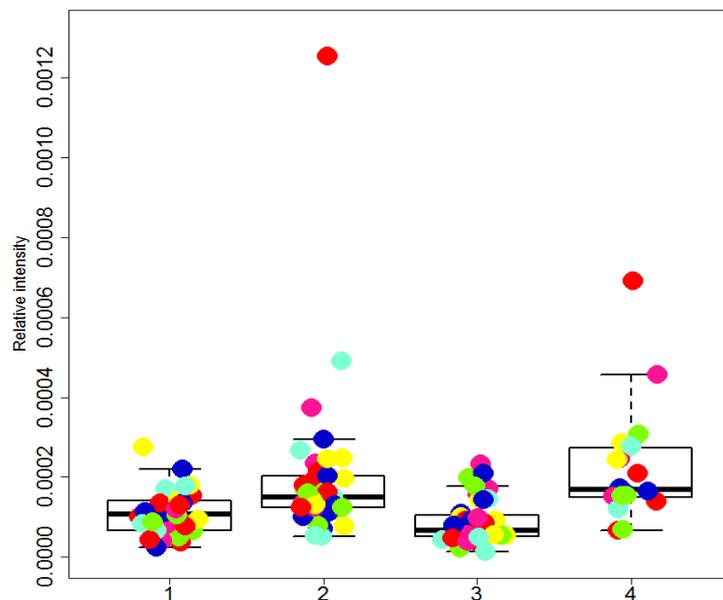
acetoacetato	glicina
inoxylsolfato	colina
Tirosina	creatina
Fucosio	dimetilglicina
mHppa	citrato
fenilacetilglicina	piruvato
ascorbato	glutammato
idrossibutirrato	arginina
lattato	alanina
trigonellina	idrossivalerato
metilnicotinammide	valerato
formiato	valina
ippurato	isoleucina
treonina	leucina

Le variazioni delle concentrazioni dei metaboliti sono classificate attraverso un valore chiamato "*P-Value*" il quale se inferiore agli 0,05 diventa statisticamente significativo ovvero il metabolita presenta delle differenze rilevanti.

L'individuo AW mostra variazioni più consistenti in determinati metaboliti quali la Trigonellina, l'Ippurato e l'idrossibutirrato riportati nelle immagini successive.

**Figura 18. Cambiamenti
Trigonellina durante le 4 raccolte
metref.**

pValue= 1.81×10^{-8}



**Figura 19. Cambiamenti
ippurato.**

PValue= 0.037

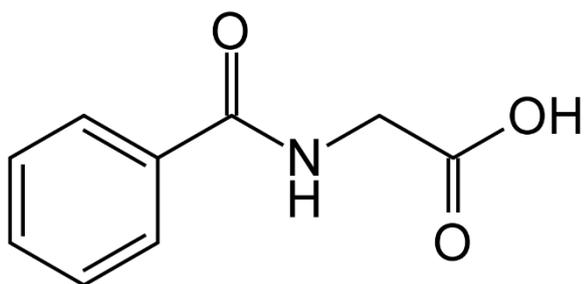
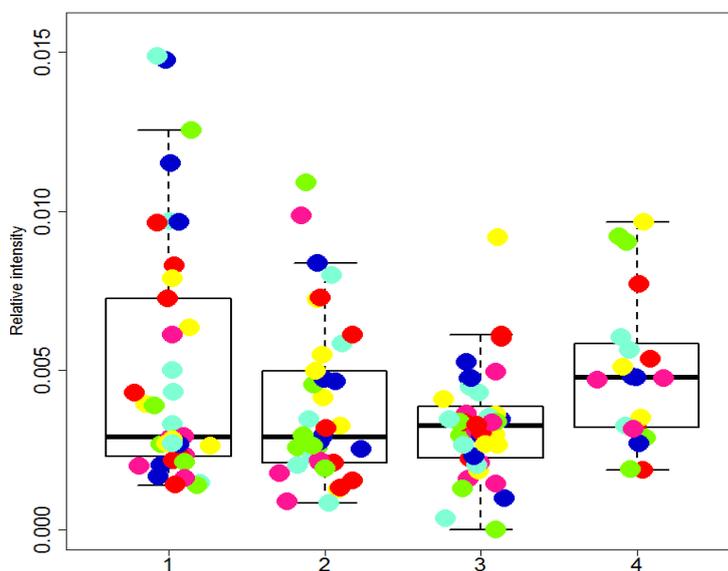


Figura 20. Formula Acido Ippurico

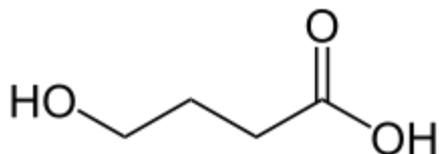


Figura 21. Formula acido idrossibutirrico

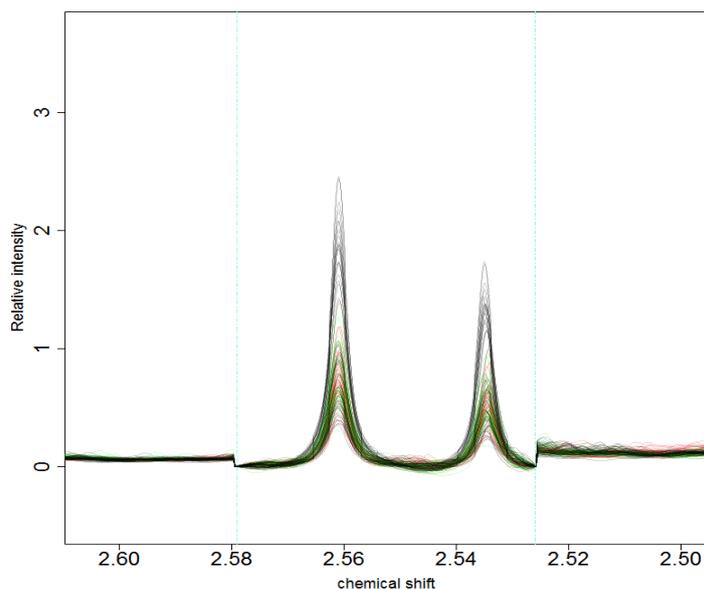


Figura 22. Picchi del Citrato

Questi 3 metaboliti sono strettamente legati all'attività della flora intestinale. Si osserva anche un aumento significativo del citrato che può essere associato ad un'inflammatione sistemica e ad un'anomala attività metabolica.

Nel periodo tra il 2005 e il 2014 il soggetto BD ha avuto due gravidanze ma al momento della raccolta non stava allattando. Nello stesso periodo però ha subito un'asportazione della tiroide a causa di una formazione di una massa tumorale. Non ci sono studi che analizzano questo tipo di intervento da un punto di vista metabolomico ma è possibile affermare che ci sia una correlazione tra questo evento e la bassa percentuale di riconoscimento dell'individuo. Infatti la tiroide è quella ghiandola che regola il meccanismo energetico del metabolismo e per questo motivo i pazienti sottoposti ad un intervento di asportazione della tiroide sono costretti ad assumere ormoni tiroidali per tutto il proseguo della vita.

Nel soggetto BD molti metaboliti hanno subito una variazione della loro concentrazione che altera quindi il profilo metabolomico abbassando le percentuali di riconoscimento.

I metaboliti maggiormente responsabili di questa alterazione sono la colina, la tirosina e l'isoleucina. La tirosina ad esempio è il precursore, nella tiroide, della formazione degli ormoni triiodotironina e tiroxina e quindi il cambiamento nella sua concentrazione può essere correlato all'asportazione della ghiandola tiroidale. L'aumento e poi la diminuzione della concentrazione dei metaboliti che contengono colina nelle urine è spesso associata alla formazione di una massa tumorale come dimostrato in numerosi studi sulle cellule cancerogene^{28,29}. Queste informazioni avvalorano la teoria della correlazione l'intervento alle tiroide e il cambiamento del profilo metabolico. Altri studi andranno condotti per accertare la correlazione causa-effetto di questo tipo di patologia.

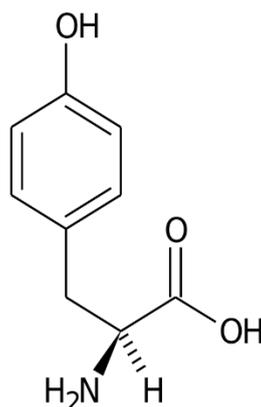


Figura 23 Formula delle tirosina

Nell'immagine sottostante sono riportati in una tabella riassuntiva le variazioni delle concentrazioni dei metaboliti degli individui con bassa percentuale di riconoscimento (figura 24).

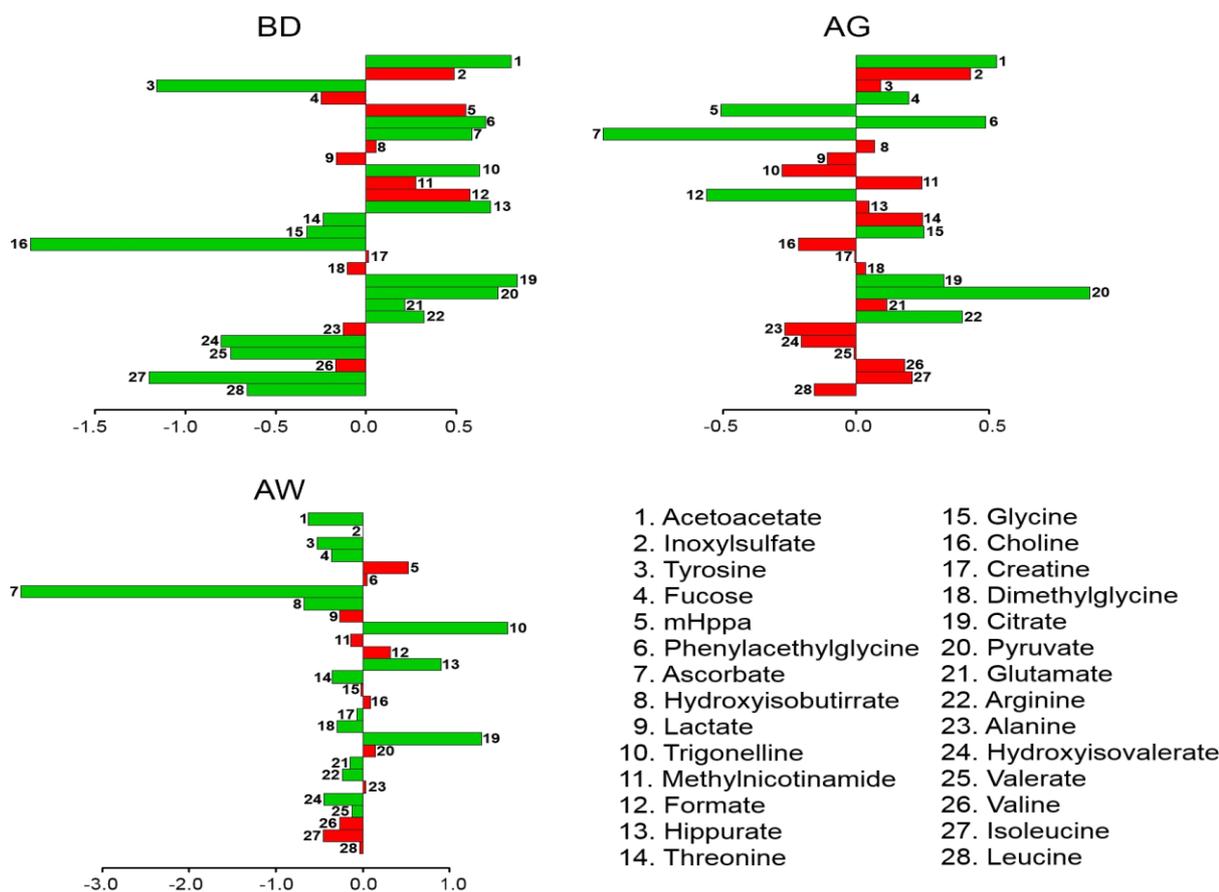


Figura 24. Tabella delle variazioni dei metaboliti nei 3 individui. In verde i metaboliti che hanno subito una fluttuazione significativa mentre in rosso quelli che sono rimasti costanti. Al lato la tabella con i numeri associati ad ogni metabolita.

Per il soggetto AG sono meno i metaboliti che subiscono delle variazioni durante il corso degli anni.

AG tra il 2005 e il 2014 ha avuto 2 gravidanze e nell'ultima raccolta ancora allattava. Questa informazione fornisce dati importanti al fine di stabilire le variabilità incorse durante il periodo proposto.

L'aumento di alcuni metaboliti come l'ascorbato e il piruvato (riportato in figura 25 e 26) non possono essere gli unici responsabili della scarsa accuratezza sulla predizione.

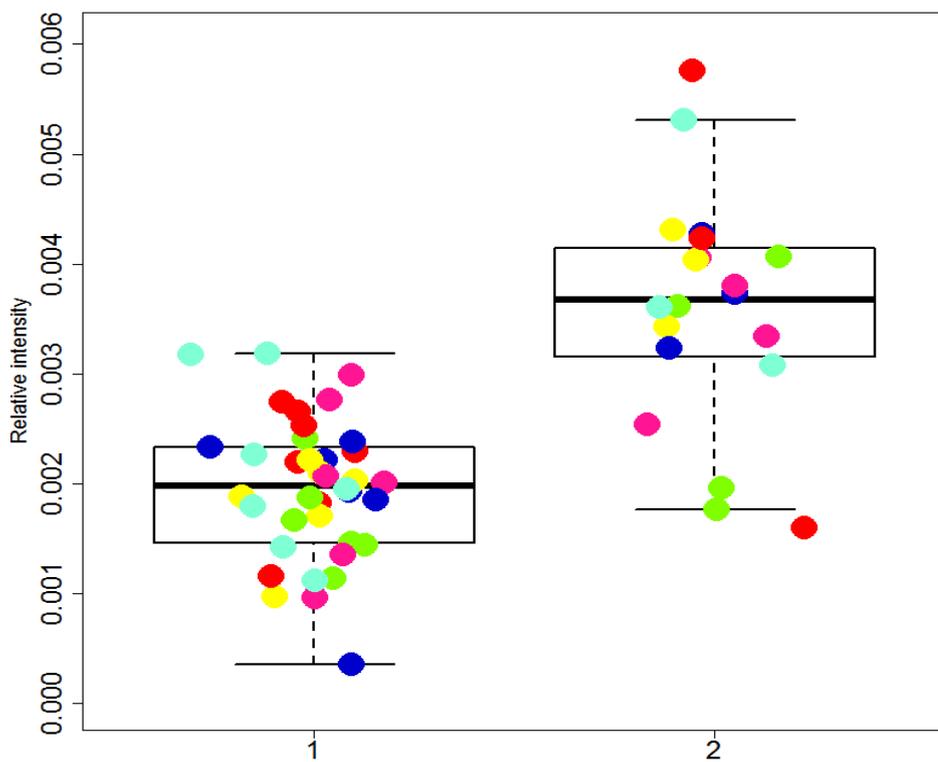


Figura 25. Grafico rappresentante la variazione di Piruvato nel soggetto AG

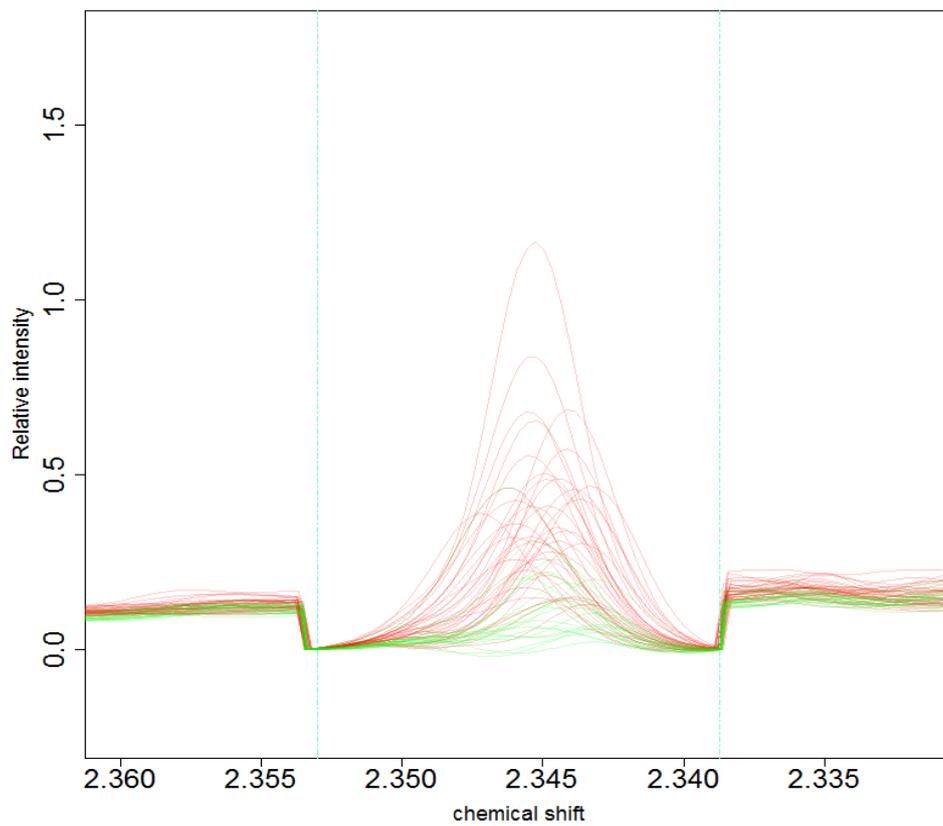


Figura 26. Picco del piruvato

Si può ipotizzare che la causa principale dello scarso riconoscimento di AG sia legata all'allattamento e quindi alla presenza di lattosio nelle urine. Infatti un individuo normale non assorbe dalla dieta elevate quantità di lattosio quindi generalmente i picchi relativi a questo metabolita saranno presenti ma non intensi. Dagli spettri analizzati si ricava che nei campioni di AG raccolti nel 2014 non solo è presente il lattosio ma anche in quantità superiori alla norma. In figura 27 viene riportato lo spettro relativo ai picchi del lattosio.

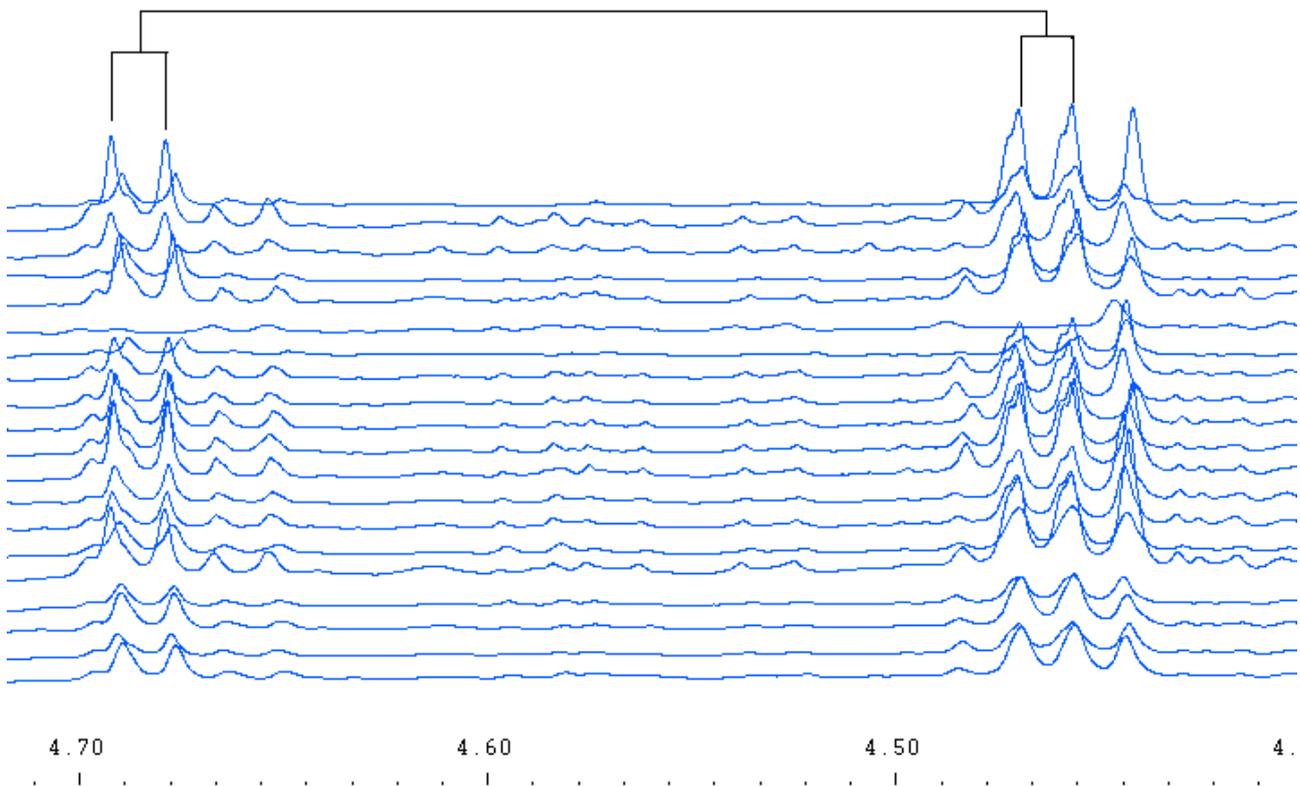


Figura 27. Spettro dei picchi del lattosio dell'individuo AG

Non è sorprendente che il fenotipo metabolico nelle urine venga condizionato durante la fase di allattamento. L'aumento della concentrazione del lattosio e della glicina durante la fase di allattamento è stata già da tempo confermata^{30,31} e il lattosio è il metabolita caratteristico che sancisce l'allattamento.

Nelle figure 28 e 29 sono riportati i grafici relativi alla presenza di glicina nell'individuo AG.

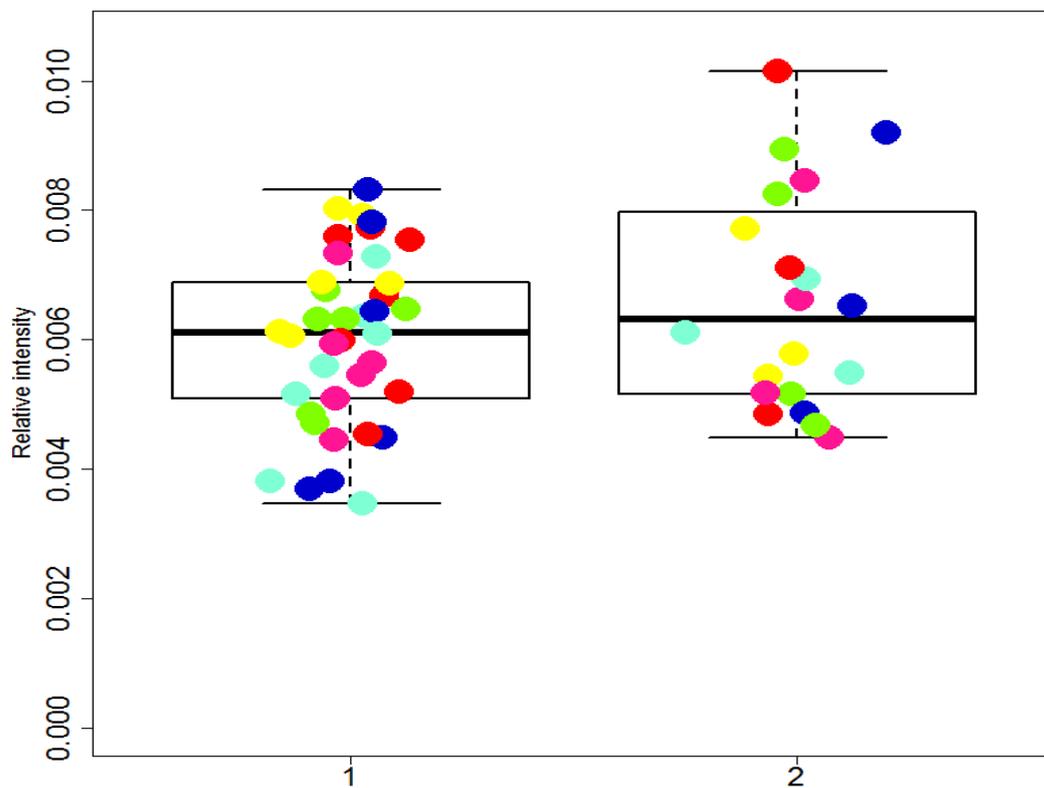


Figura 28. Schema delle variazioni della glicina nell'individuo AG nel tempo.

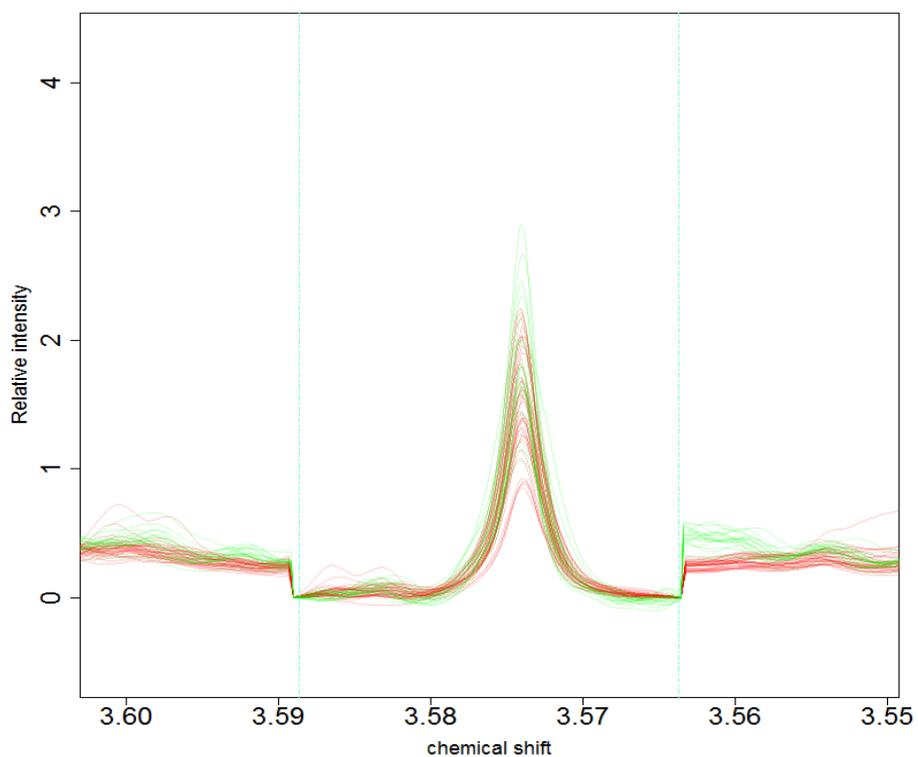


Figura 29. Picco della glicina dell'individuo AG

Durante l'analisi dei dati, avendo constatato che alcuni individui erano scarsamente riconosciuti nel tempo, ad alcuni di questi è stato chiesto di portare ulteriori campioni al fine di valutare se il profilo metabolico non solo variasse nel tempo ma anche ritornasse alla sua condizione iniziale.

L'individuo che ha fornito i campioni è stato AG che è anche il più semplice da analizzare in quanto non presenta una patologia ma solo una condizione fisiologica (l'allattamento).

Nel periodo di raccolta AG naturalmente non allattava dando così la possibilità di studiare se effettivamente ci fosse un ritorno alle condizioni iniziali.

I risultati ottenuti, come si osserva nella comparazione tra gli spettri della primavera del 2014 e l'inverno del 2015 riportata in figura 30, confermano le ipotesi precedentemente formulate.

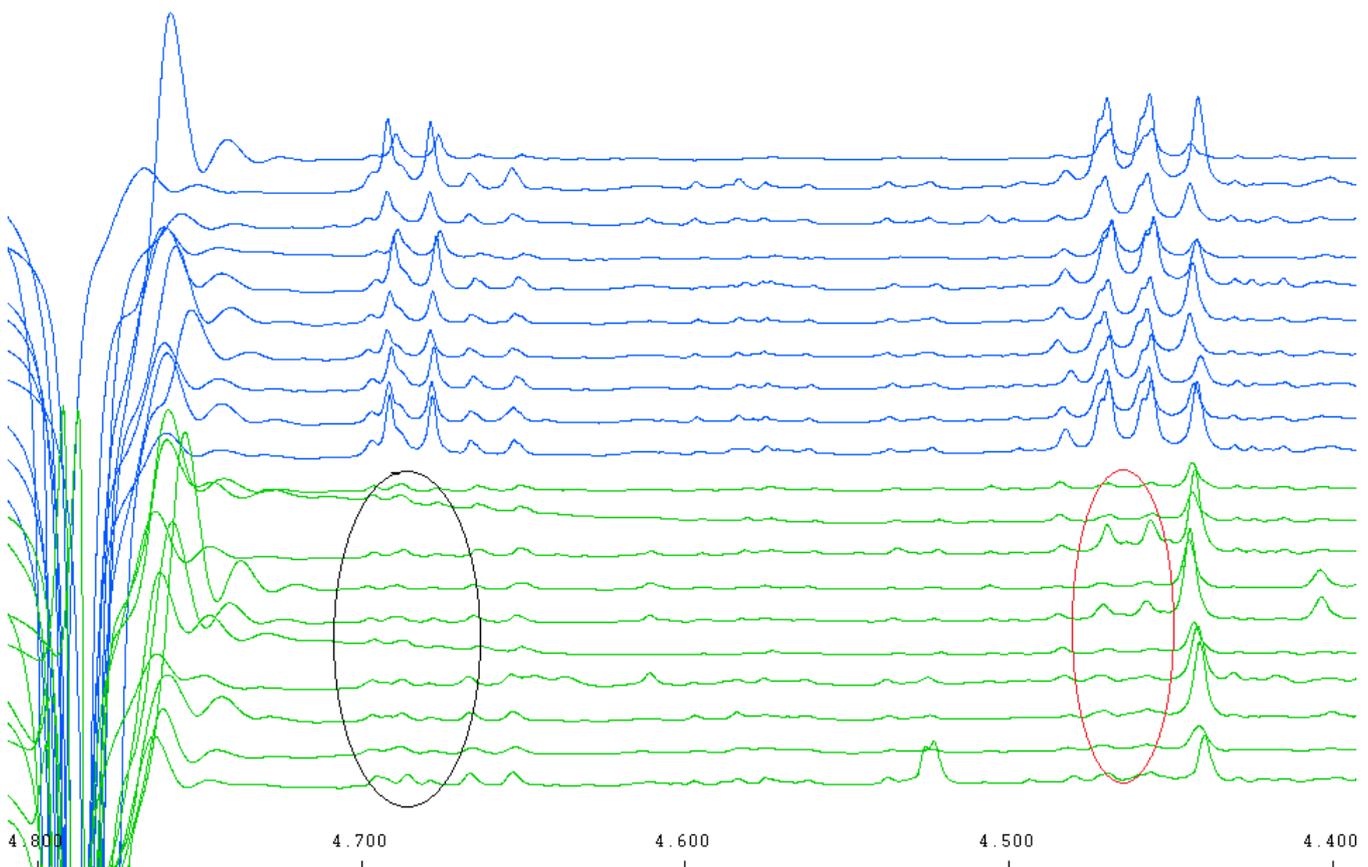


Figura 30. Comparazione degli spettri dei campioni della prima raccolta di metref4 (blu) e quelli della seconda raccolta (verde) dell'individuo AG

I picchi a 4.68 ppm sono quelli relativi agli zuccheri presenti nelle urine e forniscono la riprova che la concentrazione della parte zuccherina è diminuita drasticamente nel corso dei pochi mesi che separano le due raccolte.

Più significativa è la regione del doppietto (intorno a 4.46 ppm) rappresentante il lattosio. Si osserva quasi una scomparsa completa rispetto ai picchi della prima raccolta confermando l'ipotesi precedentemente stipulata e aprendo la strada a successive analisi per dimostrare che l'individuo è nuovamente riconoscibile ottimamente.

Si può concludere che i tre individui erano scarsamente riconoscibili perché durante la raccolta della primavera del 2014 sono stati soggetti ad una condizione fisiologica "stressante" e prolungata che ha portato uno spostamento del loro *spazio* metabolico. Si può oltretutto ipotizzare che sia per AG che per AW la situazione sia temporanea mentre per BD, soggetta ad una più intensa patologia, il processo di riequilibrio sia più lungo e non completamente reversibile.

6 Conclusioni

In questo progetto sono stati studiati i fenotipi metabolici di 12 individui nell'arco di 8-10 anni, realizzando il più esteso studio dei profili metabolici individuali fino ad oggi realizzati.

È stato dimostrato che gli individui sani mantengono il loro profilo metabolico inalterato sia nel breve che nel lungo periodo. Quindi gli stimoli esterni giornalieri non sono così influenti da determinare un'alterazione dello spazio metabolico individuale.

In assenza di un evento patologico rilevante si è potuto osservare che il fenotipo individuale è stabile anche nel corso di 10 anni, al contrario l'esposizione prolungata ad una patologia conduce ad uno spostamento dello spazio metabolico proporzionale alla gravità dell'evento.

In ugual modo si può stabilire che alla scomparsa della patologia il profilo metabolico individuale tende a ristabilire le condizioni di equilibrio precedenti all'inizio della condizione di stress.

Facendo un'analogia possiamo considerare il profilo metabolico individuale come una curva di potenziale: la condizione di massima stabilità può essere espressa come una *bucca* di potenziale dalla quale è difficile uscire, ma in presenza di un cambiamento energetico consistente (la patologia) si esce dalla condizione di stabilità fino ad un avvallamento a più alta energia che sarà una condizione di stabilità temporanea fino al ritorno alla condizione iniziale (fine della patologia).

In conclusione la metabolomica con lo studio del profilo metabolico individuale potrebbe avere, nell'ambito medico, un'importanza rilevante nel monitoraggio di un paziente durante l'avanzamento di una malattia, una terapia o anche una dieta, grazie alla rapidità e la ricchezza di informazioni che caratterizza questa tecnica.

7 Abbreviazioni

MetRef= metabolomic reference,

Met=MetRef,

NMR= nuclear magnetic resonance

MS= Mass spectrometry

HPLC= High pressure liquid chromatography

IR= infra-red

PCA= principal component analysis

PLS= partial least square

CA= canonical analysis

KNN= k-nearest neighbours

FID= free induction decay

8 Bibliografia

- 1) Nicholson, J. K. & Wilson, I. D. Opinion: understanding 'global' systems biology: metabonomics and the continuum of metabolism. *NatRevDrug Discov* **2**, 668-676 (2003)
- 2) Joyce, A. R. & Palsson, B. Ø. The model organism as system: integrating 'omics' data sets. *Nat. Rev. Mol. Cell Biol.* **7**, 198-210 (2006).
- 3) Nicholson, J. K., Lindon, J. C. & Holmes, E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica Fate Foreign Compd. Biol. Syst.* **29**, 1181-1189 (1999).
- 4) Hollywood, K., Brison, D. R. & Goodacre, R. Metabolomics: Current technologies and future trends. *PROTEOMICS* **6**, 4716-4723 (2006).
- 5) The emerging science of epigenomics. Callinan PA, Feinberg AP *Hum Mol Genet.* 2006 Apr 15; 15 Spec No 1():R95-101
- 6) Ellis, D. I., Dunn, W. B., Griffin, J. L., Allwood, J. W. & Goodacre, R. Metabolic fingerprinting as a diagnostic tool. *Pharmacogenomics* **8**, 1243-1266 (2007).
- 7) Lindon, C. J., Nicholson, J. K. & Everett, J. R. NMR Spectroscopy of Biofluids. *Annu Rep NMR Spectro* **38**, 1-88 (1999).
- 8) Robertson, D. G., Reily, M. D. & Baker, J. D. Metabonomics in Pharmaceutical Discovery and Development. *J. Proteome Res.* **6**, 526-539 (2007).
- 9) Veselkov K. A., Vingara L. K., Masson P., Robinette S. L., Want E., Li J. V., Barton R. H., Boursier-Neyret C., Walther B., Ebbels T. M., Pelczar I., Holmes E., Lindon J. C., Nicholson J. K. (2011) Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery. *Anal. Chem.* **83**, 5864–587

- 10) Metabolic phenotyping in health and disease. *Holmes E, Wilson ID, Nicholson JK*
Cell. 2008 Sep 5; 134(5):714-7.
- 11) *Malinowski [E.R. (1991) Factor Analysis in Chemistry, Second Edition]*
- 12) Implementing a Principal Component Analysis (PCA) in Python step by step. Sebastian Raschka on April 13, 2014
- 13) The human urine metabolome. Bouatra, Aziat, Mandal, Chi Guo et al. September 04, 2013
- 14) Gavaghan, C. L., Holmes, E., Lenz, E., Wilson, I. D. & Nicholson, J. K. An NMR- based metabonomic approach to investigate the biochemical consequences of genetic strain differences: application to the C57BL10J and Alpk:ApfCD mouse. *FEBS Lett. 484, 169-174 (2000).*
- 15) Assfalg, M. *et al.* Evidence of different metabolic phenotypes in humans. *Proc.Natl.Acad.Sci.U.S.A* **105**, 1420-1424 (2008).
- 16) Bernini, P. *et al.* Individual human phenotypes in metabolic space and time. *J ProteomeRes* 8, 4264-4271 (2009).
- 17) Holmes, E. *et al.* Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* 453, 396-400 (2008).
- 18) Nicholson, J. K., Connelly, J., Lindon, J. C. & Holmes, E. Metabonomics: a platform for studying drug toxicity and gene function. *Nat. Rev. Drug Discov.* 1, 153-161 (2002).
- 19) Madsen, R., Lundstedt, T. & Trygg, J. Chemometrics in metabolomics—A review in human disease diagnosis. *Anal. Chim. Acta* 659, 23-33 (2010).
- 20) Calabrò *et al.* A Metabolomic Perspective on Coeliac Disease. *Autoimmune Dis.* 2014, (2014).
- 21) Bernini, P.; Bertini, I.; Luchinat, C.; Nincheri, P.; Staderini, S.; Turano, P. Standard operating

- procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. *J. Biomol. NMR* 2011, 49 (3-4), 231–243.
- 22) Assfalg, M.; Bertini, I.; Colangiuli, D.; Luchinat, C.; Schäfer, H.; Schütz, B.; Spraul, M. Evidence of different metabolic phenotypes in humans. *Proc. Natl. Acad. Sci.* **2008**, 105 (5), 1420–1424.
- 23) Saccenti, E.; Tenori, L.; Verbruggen, P.; Timmerman, M. E.; Bouwman, J.; van der Greef, J.; Luchinat, C.; Smilde, A. K. Of monkeys and men: a metabolomic analysis of static and dynamic urinary metabolic phenotypes in two species. *PLoS One* 2014, 9 (9), e106077.
- 24) Wallner-Liebmann, S.; Gralka, E.; Tenori, L.; Konrad, M.; Hofmann, P.; Dieber-Rotheneder, M.; Turano, P.; Luchinat, C.; Zatloukal, K. The impact of free or standardized lifestyle and urine sampling protocol on metabolome recognition accuracy. *Genes Nutr.* 2015, 10 (1), 441.
- 25) Dethlefsen, L.; Huse, S.; Sogin, M. L.; Relman, D. A. The Pervasive Effects of an Antibiotic on the Human Gut Microbiota, as Revealed by Deep 16S rRNA Sequencing. *PLoS Biol* 2008, 6 (11), e280.
- 26) Dethlefsen, L.; Relman, D. A. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl. Acad. Sci.* 2011, 108 (Supplement_1), 4554–4561.
- 27) Nicholson, J. K.; Holmes, E.; Kinross, J.; Burcelin, R.; Gibson, G.; Jia, W.; Pettersson, S. Host-Gut Microbiota Metabolic Interactions. *Science* 2012, 336 (6086), 1262–1267.
- 28) Glunde, K.; Jacobs, M. A.; Bhujwala, Z. M. Choline metabolism in cancer: implications for diagnosis and

- therapy. *Expert Rev. Mol. Diagn.* 2006, 6 (6), 821–829.
- 29) Miccoli, P.; Torregrossa, L.; Shintu, L.; Magalhaes, A.; Chandran, J.; Tintaru, A.; Ugolini, C.; Minuto, M. N.; Miccoli, M.; Basolo, F.; et al. Metabolomics approach to thyroid nodules: A high-resolution magic-angle spinning nuclear magnetic resonance–based study. *Surgery* 2012, 152 (6), 1118–1124.
- 30) Sachse, D.; Bærug, A.; Sletner, L.; I. Birkeland, K.; Nakstad, B.; Jenum, A. K.; Berg, J. P. Urine NMR metabolomics analysis of breastfeeding biomarkers during and after pregnancy in a large prospective cohort study. *Scand. J. Clin. Lab. Invest.* 2014, 74 (3), 264–272.
- 31) achse, D.; Sletner, L.; Mørkrid, K.; Jenum, A. K.; Birkeland, K. I.; Rise, F.; Piehler, A. P.; Berg, J. P. Metabolic Changes in Urine during and after Pregnancy in a Large, Multiethnic Population-Based Cohort Study of Gestational Diabetes. *PLoS ONE* 2012, 7 (12).

9 Ringraziamenti

Ringrazio innanzitutto il professor Luchinat che mi ha permesso di scoprire un campo di ricerca così ampio e affascinante come quello della metabolomica.

Di tutte le persone del CERM un ringraziamento particolare va a Leonardo che mi ha convinto ad entrare in questo affascinante mondo e soprattutto per la sua disponibilità durante tutta questa lunga esperienza.

Non posso che ringraziare la mia mentore nonché compagna di avventure, Veronica che senza la sua conoscenza, la sua “energia” prorompente e la sua infinita (anche se lei dice il contrario) pazienza questo periodo sarebbe stato

sicuramente più lungo e meno radioso.

Ringrazio anche i due metabolomici Claudio e Alessia per l'aiuto che mi hanno sempre fornito e per avermi fatto apprezzare con la loro gentilezza e simpatia la stanza numero 8.

Un ringraziamento sentito va anche al mitico Roberto che con la sua simpatia e conoscenza del rugby mi ha fatto sentire sempre a casa.

Ma il ringraziamento più assoluto va ai miei genitori che mi hanno cresciuto e mi hanno aiutato a diventare quello che sono, sostenendomi ogni giorno e spronandomi sempre a dare il massimo e sopportandomi anche nei momenti più bui.

Voglio ringraziare anche tutti i parenti che chi più e chi meno non mi hanno mai fatto mancare il loro supporto fin da piccino.

Un ringraziamento speciale va alla persona che mi è stata più vicino in questo periodo, Elisa che è ormai da due anni il mio mondo e che nonostante le sue fughe universitarie verso lidi stranieri mi è sempre vicina non facendomi mancare mai il suo sostegno, ascoltandomi nei miei momenti di follia e aiutandomi ogni giorno ad andare avanti.

Non posso non dimenticare i miei migliori amici di sempre Matte, Giova e Pagno che fin da piccoli sono stati non solo amici ma anche elementi fondamentali del mosaico chiamato vita.

Scrivendo questo paragrafo non posso non pensare ai miei amici chimici che mi hanno accompagnato in questa avventura "chimica" iniziata col mitico gruppo "Zwaifen" e momentaneamente conclusati con il più ampio "bandatcha".

Ringrazio anche tutto il gruppo "il bandino" che con le loro stranezze e risate hanno allietato le serate e le vacanze fiorentine.

Grazie anche a Dami e Cocco, compagni di pesca insuperabili, sempre pronti a passare una giornata in riva a un lago chiacchierando e filosofeggiando.

Grazie davvero a tutte quelle persone che con la loro presenza hanno reso il tempo trascorso meno faticoso e che mi hanno aiutato nel conseguimento di questo primo traguardo che segna la fine di un'esperienza davvero indimenticabile.

