

UNIVERSITA' DEGLI STUDI DI MILANO– BICOCCA

Facoltà di Medicina e Chirurgia

Dipartimento di Medicina Sperimentale

Dottorato di Ricerca in Tecnologie Biomediche XXV Ciclo

Coordinatore: Chiar.ma Prof.ssa Marina Del Puppo



**Alterazioni del peptidoma serico di pazienti con
tumore renale valutate tramite “label free” (nLC-ESI-
MS/MS) e “peptide profiling” (MALDI-MS/MS)**

Candidato:
Dott.ssa Squeo Valeria
Matricola 734504

Tutor: Chiar.mo Prof. Fulvio MAGNI

*A te che sei la felicità perché sei sei tu nei miei sogni
a te che sei il sole che scalda il mio cuore...*

INDICE

1. INTRODUZIONE	1
1.1 Il rene: anatomia e fisiologia	2
1.1.1 Anatomia renale	2
1.1.2 Fisiologia renale	3
1.2. Il carcinoma renale	4
1.2.1 Cenni di epidemiologia	4
1.2.2 Classificazione.....	5
1.2.3 Carcinoma a cellule chiare	7
1.2.4 Carcinoma renale papillare.....	8
1.2.5 Carcinoma renale cromofobo	9
1.2.6 Eziologia.....	10
1.2.7 Fattori prognostici molecolari	12
1.2.8 Strumenti di diagnosi.....	14
1.2.9 Fattori prognostici anatomici e istologici	18
1.2.10 Strategie Terapeutiche	22
1.3 Proteomica	24
1.3.1 Introduzione alla proteomica	24
1.3.2 Tecniche di proteomica.....	27
1.3.4 Analisi di miscele complesse: elettroforesi e cromatografia.....	28
1.3.4.1 Cromatografia liquida HPLC.....	30
1.3.3 Analisi dei fluidi biologici: la tecnologia SELDI e Clinprot	32
1.4 La spettrometria di massa	36
1.4.1 Introduzione e principi fondamentali	36
1.4.2 Spettro di massa.....	38
1.4.3 Spettrometro di massa: elementi fondamentali.....	41
1.4.3.1 Sistema di introduzione	41
1.4.3.2 Sorgente ionica	42
1.4.3.3 Analizzatore di ioni.....	44
1.4.3.4 Rilevatore.....	47
1.4.4 Tecniche di desorbimento.....	48
1.4.4.1 Elettrospray ionization.....	49

1.4.4.2 MALDI- TOF	52
1.4.5 Spettrometria di massa Tandem	60
1.4.6 Frammentazione	62
1.4.7 Proteomica quantitativa mediante spettrometria di massa	63
1.4.8 Quantificazione proteica con tecnica Label Free	65
2. SCOPO DELLA TESI	67
3. MATERIALI E METODI	71
3.1 Raccolta dei campioni	72
3.2 Prefrazionamento dei campioni attraverso tecnica ClinProt	72
3.3 Studio di profiling con tecnica MALDI-TOF	74
3.4 Identificazione peptidica con tecnica MALDI-TOF	75
3.5 Analisi bioinformatica e biostatistica	76
3.6 Identificazione peptidica attraverso tecnica nLC-ESI MS/MS	87
3.7 Tecnica Label-Free: quantificazione peptidica/proteica	93
4. RISULTATI	94
4.1 Popolazione e dati clinici dello studio	95
4.2 Analisi proteomica ed efficacia diagnostica dei pattern	96
4.3 Identificazione dei peptidi	99
4.4 Validazione	122
5. CONCLUSIONI	125
5.1 Validazione	128
5.2 Identificazione peptidica	129
5.3 Espressione proteica	133
6. BIBLIOGRAFIA	139

1.INTRODUZIONE

1.1 Il rene: anatomia e fisiologia

1.1.1 Anatomia renale

Il rene è un organo simmetrico, bilateralmente situato presso la parete posteriore dell'addome, dietro al peritoneo, ai lati della colonna vertebrale all'altezza delle ultime vertebre dorsali (Fig 1).

Figura 1: struttura del rene (Immagine tratta da Wheater. Istologia e anatomia microscopica, a cura di Young B, Heath JW. CE Ambrosiana).

I reni hanno forma caratteristica simile a quella di un fagiolo, con l'asse maggiore disposto longitudinalmente; il margine concavo, rivolto medialmente, presenta un'apertura, *ilo renale*, in corrispondenza del quale penetrano i vasi sanguigni renali e si diparte il corrispettivo uretere.

In sezione frontale, il rene presenta una parte periferica più chiara e una parte interna più scura: la parte periferica è detta *zona corticale* mentre quella più interna è la *zona midollare*.

La zona midollare è suddivisa da in diverse porzioni triangolari ciascuna avente al vertice le cosiddette *piramidi di Malpighi*. Il parenchima tra le piramidi è costituito dalla stessa corticale che in questa sede forma *le colonne del Bertin*.

Il parenchima di un rene è costituito da circa un milione di unità funzionali, i nefroni, a loro volta costituiti da una parte filtrante, il corpuscolo renale, cui fa seguito una parte riassorbente e secernente, il tubulo renale.

I corpuscoli renali sono formazioni sferoidali che si trovano esclusivamente nella corticale del rene e operano l'ultrafiltrazione del plasma. Nei reni sono circa due milioni, con una superficie filtrante totale di 1,5-2 m². Ciascun corpuscolo è costituito da due unità funzionali: la capsula di Bowman e il glomerulo renale. La capsula di Bowman è una formazione sferica formata da una membrana ialina in cui si possono distinguere il polo arterioso attraverso il quale entra l'arteriola afferente ed esce l'arteriola glomerulare efferente, e il polo urinifero che è il punto in cui inizia il tubulo renale.

Il glomerulo renale rappresenta l'insieme dei capillari che collegano l'arteriola afferente a quella efferente.

L'altra parte unità funzionale del nefrone è il *tubulo prossimale* che è organizzato con una successione di tratti con proprie caratteristiche fisiologiche e morfologiche. Il tubulo contorto prossimale rappresenta la prima porzione in prossimità del corpuscolo di Malpighi da cui origina.

Si dirama all'interno della corticale e, a sua volta, da origine all'ansa di Henle, che si addentra nella piramide midollare, compie un'ansa e ritorna nella corticale dove prosegue nel tubulo

contorto distale. Questo tratto prosegue con un dotto collettore che convoglierà l'urina all'interno del calice renale.

Il tubulo renale si contorce e si ripiega su se stesso in modo che il tubulo distale passa tra le arteriole afferenti ed efferenti formando l'apparato iuxtaglomerulare costituito dalle cellule iuxtaglomerulari (miociti della parete dell'arteriola afferente i quali si sono differenziati in cellule endocrine) e dalla macula densa in contatto con la porzione del tubulo distale^{1,2,3}.

1.1.2 Fisiologia renale

Il sistema urinario è deputato alla filtrazione di quantità elevate di plasma, al riassorbimento di quelle sostanze che l'organismo deve conservare e all'eliminazione e/o escrezione di quelle che devono essere espulse. Ogni minuto si formano circa 120 ml di ultra filtrato, anche se la quantità di urina prodotta è pari al 1 ml/min, quindi più del 99% dell'ultrafiltrato glomerulare viene riassorbito.

Il sangue che, arrivando dall'arteria renale, passa attraverso i capillari glomerulari viene ultrafiltrato dando origine ad un soluto simile a plasma ma molto povero in proteine. Il filtro glomerulare permette il passaggio di acqua e soluti a basso peso molecolare questo fa sì che vengano trattenute, in base al peso molecolare, molecole come le proteine mentre c'è un libero scambio di acqua ed elettroliti. Il riassorbimento di elettroliti ed acqua è anche regolato rispettivamente dall'ormone aldosterone e dalla vasopressina di cui il rene è l'organo bersaglio.

Inoltre il rene è anche sede di produzione di composti a funzione endocrina infatti nell'apparato iuxtaglomerulare viene prodotta la renina un enzima proteolitico che, nel circolo sanguigno, induce la produzione dell'angiotensina I che, per successiva idrolisi genera angiotensina II. Quest'ultimo ormone è un potente vasocostrittore ad alto effetto ipertensivo.

Un altro ormone prodotto dal sistema renale è l'eritropoietina che stimola la sintesi di emoglobina negli eritroblasti (elemento immaturo della serie rossa del sangue) e induce la differenziazione delle cellule staminali eritroidi (emocitoblasti) in eritroblasti e quindi in eritrociti.

Infine il rene è responsabile anche della produzione calcitriolo o 1,25-diidrossicolecalciferolo. La sua attività è stimolata con effetto induttorio dalla carenza alimentare di calcio e fosfato, dall'ipocalcemia ed ipofosfatemia e dagli ormoni paratiroidei, della crescita, prolattina ed estrogeni^{4,5,6}.

La Società Italiana di Nefrologia (SIN), sulla base dei dati aggiornati a tutto il 2009, conferma che in Italia, le persone che devono abitualmente fare ricorso alla dialisi sono oltre 45mila;

coloro che hanno subito un trapianto di rene sono quasi 16mila (15.793 a tutto il 31 dicembre 2010); e sono tra i 5 e i 6 milioni gli italiani che soffrono di un danno renale.

L'insufficienza renale può essere causata da differenti affezioni e quelle più frequenti sono le lesioni renali secondarie all'ipertensione arteriosa, all'arteriosclerosi e al diabete. Seguono per frequenza le glomerulo nefriti; le cosiddette nefropatie interstiziali, talora infettive o secondarie ad un'ostruzione delle vie urinarie, o più spesso legate ad un'inappropriato uso di farmaci; le malattie ereditarie, tra i quali i reni policistici. L'insufficienza renale, anche lieve accresce sensibilmente il tasso di morbilità e mortalità: il tasso di sopravvivenza dei dializzati italiani è del 55 per cento e tale percentuale si abbassa notevolmente se si considerano i tumori renali.

Il carcinoma renale è il terzo tumore urologico più frequente dopo quello alla prostata e alla vescica e appare in crescita negli ultimi decenni^{7,8}.

1.2. Il carcinoma renale

1.2.1 Cenni di epidemiologia

Nel 2011 il carcinoma renale (RCC) ha ricevuto una grande attenzione presso i maggiori congressi urologici e oncologici tenutisi in tutto il mondo anche se si attesta solo al terzo posto tra le più comuni neoplasie urologiche dopo il cancro prostatico e uroteliale.

L'RCC conta più di 50.000 nuovi casi negli USA e circa 15.000 ogni anno in Germania⁹. Dati aggiornati al 2002 dell'Istituto Superiore della Sanità confermano che sono stati diagnosticati 8.200 (5.600 uomini e 2.600 donne) e rappresentano il 3% di tutte le neoplasie nella popolazione italiana.

In Italia l'incidenza di questa neoplasia è in crescita, mentre la mortalità è in riduzione come si può vedere dal grafico. Questo trend favorevole è probabilmente dovuto alla maggior diffusione della diagnostica per immagini, che tuttavia ha migliorato la percentuale di diagnosi precoce, con un effetto positivo sulle possibilità terapeutiche.¹⁰

Elevato rimane comunque il numero delle vittime affette da RCC a causa della sua crescita aggressiva, dell'alto tasso di metastasi precoci e della mancanza di una terapia in fase avanzata.

L'incidenza dell'RCC è aumentata in modo significativo negli ultimi 30 anni. In particolare un'analisi retrospettiva (1975 - 2006) condotta dal SEER (*Surveillance Epidemiology and End Results*) su un database di 63.843 pazienti con carcinoma renale ha rivelato che negli Stati Uniti tra il 1975 e il 2006 l'incidenza è passata dal 7,4 al 17,6 per 100.000 abitanti.

L' incremento medio annuo è stato del 3,6% nel periodo compreso tra il 1976-1990 mentre questa tendenza è leggermente diminuita negli anni che seguirono. Tra il 1991-2006, infatti, la percentuale di incidenza medi annuale ha raggiunto il 2,9%.¹¹

Un altro interessante risultato è emerso da questa analisi: l'aumento dell'incidenza ha colpito soprattutto i pazienti più anziani fino al 1990 e più giovani dopo il 1991. Questo significa anche che, al momento della diagnosi, la maggior parte dei pazienti con carcinoma renale aveva più di 65 anni fino al 1991, mentre il 55,3% dei soggetti colpiti da RCC aveva un'età inferiore ai 65 anni nel 2006. Il carcinoma a cellule renali ha, quindi, un andamento crescente in base all'età, è infatti raro prima dei vent'anni, rappresentando solo il 2% dei tumori pediatrici, senza differenza di genere, e fino alla quarta decade si annoverano pochi casi; l'incidenza aumenta poi progressivamente raggiungendo un picco tra i sessanta ed i settanta anni.

Il RCC è più frequente nel sesso maschile rispetto a quello femminile, con un rapporto maschio-femmina di 3-2 ed in particolare ha un tasso di incidenza europea di 16.7 tra gli uomini e 7.6 nelle donne per 100000 persone/anno¹².

1.2.2 Classificazione

I carcinomi renali sono una classi eterogenea di tumori, che possono derivare da cellule specifiche all'interno del nefrone o nei dotti collettori.

Le neoplasie del rene sono classificate secondo il sistema redatto dall'Organizzazione Mondiale della Sanità (OMS) nel 2004 che è stato sviluppato a partire dalla classificazione formulata da Heidelberg e Rochester nel 1997. La classificazione OMS 2004 è molto importante ai fini dell'inquadramento clinico e terapeutico perché definisce numerosi sottotipi di neoplasie, che presentano aspetti istopatologici, alterazioni genetiche e comportamenti clinici distinti.

In questa classificazione sono inclusi sia tumori benigni sia tumori maligni ed una categoria indicata come "carcinoma a cellule renali inclassificabile", a cui dovrebbero essere assegnate quelle neoplasie che non rientrano in nessuna delle categorie ben stabilite.¹³

Istopatologia: classificazione delle neoplasie del rene secondo WHO 2004

Neoplasie a cellule renali

Maligne

- Carcinoma a cellule renali a cellule chiare
- Carcinoma a cellule renali a cellule chiare multiloculare
- Carcinoma a cellule renali papillare
- Carcinoma a cellule renali cromofobo
- Carcinoma dei dotti collettori del Bellini
- Carcinoma a cellule renali midollare
- Carcinoma con traslocazione Xp11

Carcinoma associato a neuroblastoma
Carcinoma mucinoso a cellule tubulari e fusate
Carcinoma a cellule renali inclassificabile

Benigne

Adenoma papillare
Oncocitoma

Neoplasie metanefriche

Adenoma metanefrico
Adenofibroma metanefrico
Tumore stromale metanefrico

Neoplasie nefroblastiche

Residui nefrogenici
Nefroblastoma
Nefroblastoma cistico

Neoplasie mesenchimali

Sarcoma a cellule chiare
Tumore rabdoide
Nefroma mesoblastico congenito
Tumore renale ossificante dell'infanzia
Leiomioma
Angiosarcoma
Rabdomiosarcoma
Istiocitoma fibroso maligno
Emangiopericitoma
Osteosarcoma
Angiomiolipoma
Angiomiolipoma epiteloide
Leiomioma
Emangioma
Linfangioma
Tumore a cellule juxtaglomerulari
Tumore a cellule interstiziali della midollare
Schwannoma
Tumore fibroso solitario

Neoplasie miste mesenchimali ed epiteliali

Nefroma cistico
Tumore misto epiteliale e stromale
Sarcoma sinoviale

Neoplasie neuroendocrine

Neoplasie ematopoietiche e linfoidi

Neoplasie a cellule germinali

Neoplasie metastatiche

I tumori benigni (adenoma, angiomiolipoma, fibroma, lipoma, leiomoma, emangioma, tumore delle cellule iuxtaglomerulari) non sono frequenti e la loro incidenza è compresa tra il 3-7% dei casi. Possono insorgere da qualsiasi cellula renale presente nel rene, possono raggiungere grandi dimensioni e sono difficilmente differenziabili dalle forme maligne anche con le moderne

tecniche radiologiche. Nelle maggior parte dei casi, la diagnosi delle forme benigne è data dall'esame istologico del pezzo operatorio, quindi solo dopo intervento chirurgico. Solo l'angiomiolipoma renale può avere una diagnosi pre-operatoria pressoché sicura, dal momento che le sue immagini all'ecografia e alla TC sono caratteristiche.

Tra le neoplasie maligne, quelle con una più alta percentuale di incidenza sono le neoplasie a cellule chiare, il tumore renale papillare e cromofobo che insieme costituiscono oltre il 95 % dei casi¹⁴.

1.2.3 Carcinoma a cellule chiare

Il carcinoma renali a cellule chiare (ccRCC) rappresenta circa il 75% dei tumori renali ed è quindi la neoplasia più frequente. Nel 28-37% dei casi di ccRCC si verifica necrosi coagulativa soprattutto in pazienti classificati con un alto grado nucleare G4 secondo la scala di Fuhrman .

La trasformazione sarcomatoide avviene nel 5% circa di tale neoplasia¹⁵ ed è evento più frequente nei tumori a cellule chiare, ma non esclusivo di questo istotipo. Un cenno particolare meritano le forme a presentazione cistica, tra le quali ve ne sono alcune assai poco aggressive¹⁶. Nel caso in cui vi siano cisti ma non una componente tumorale solida si parla di carcinoma a cellule renali cistico multiloculare. Questo è considerato un sottotipo del RCC a cellule chiare ed ha un'eccellente prognosi.

La neoplasia si presenta macroscopicamente di colorito giallastro conferito dall'abbondante presenza di colesterolo, lipidi neutri e fosfolipidi, talora con aree emorragiche, necrotiche o parzialmente cistiche.

Le cellule chiare rotondeggianti o poligonali, che caratterizzano il ccRCC, derivano da un artefatto, causato dalla routinaria procedura istologica che, rimuovendo il glicogeno ed i lipidi all'interno dei citoplasmici, li fa apparire otticamente vuoti al microscopio ottico dopo colorazione con ematossilina ed eosina.

Il ccRCC può anche contenere una percentuale variabile di cellule con citoplasma granulare eosinofilo: raramente, queste cellule granulari sono predominanti o l'unico tipo di cellula.

Questo tipo di neoplasia può crescere in vari modelli strutturali e tra questi i tipi più comuni sono quello alveolare compatto, tubolare, e microcistico. Nel modello alveolare compatto (Fig 4), le cellule epiteliali si aggregano in modo solido con pattern microacinare e sono separate da una delicata ramificazione di tessuto connettivo altamente vascolarizzato.

Figura 2: struttura alveolare compatta di ccRCC (Immagine tratta da e- medicine Medscape)

Nel modello tubolare (Fig 5), le cellule epiteliali sono disposte linearmente lungo i setti di tessuto connettivo vascolarizzato a formare il lume centrale. Nelle regioni di degenerazione cistica, queste strutture si dilatano a formare un pattern microcistico e / o macrocistico con spazi luminali contenenti materiale necrotico, chiaro fluido eosinofilo e globuli rossi^{17,18}.

Figura 3: struttura tubulare di ccRCC (Immagine tratta da e- medicine Medscape)

CCRCCs spesso contengono più di un modello architettonico e possono essere osservati anche regioni di architettura papillare o pseudo papillare (Fig 6).

Figura 4: struttura papillare di ccRCC (Immagine tratta da e- medicine Medscape)

La maggior parte dei carcinomi a cellule chiare generano una debole risposta infiammatoria, tuttavia vi può essere la presenza di infiltrati linfocitari o neutrofilici con cellule natural killer e vi è un'associazione tra la forte infiltrazione linfocitaria e l'outcome sfavorevole.

CCRCCs hanno una maggiore propensione alla metastatizzazione attraverso il sistema vascolare che non per il sistema linfatico, con la presenza di cellule maligne all'interno di piccole vene intrarenali. Così l' invasione del seno renale di solito avviene dall'interno della vena renale, portando ad una maggiore propensione per le metastasi a distanza che per la diffusione loco-regionale e i bersagli più comuni sono il polmone (33-72%), i linfonodi intra-addominali (3-35%), il sistema scheletrico (21-25%), il cervello (7-13%), il sistema epatico (5-10%) Il carcinoma a cellule renali a cellule chiare è conosciuto per la metastatizzazione in tempi lunghi, anche dopo 10 anni dalla diagnosi/trattamento^{19,20}.

1.2.4 Carcinoma renale papillare

Il carcinoma a cellule renali di tipo papillare (pRCC) rappresenta circa il 10-15% delle neoplasie del rene. Sebbene la maggior parte pRCCs sono unilaterali, pRCC è il più comune tumore renale bilaterale o multifocale. Macroscopicamente si può presentare in forma solida od avere una degenerazione necrotica centrale simil-cistica. Microscopicamente la caratteristica principale è l'aggregazione papillare con assi fibrovascolari con macrofagi e/o con necrosi colesterinica²¹. La differenza tra il pRCC e il ccRCC è individuabile dalla minore vascolarizzazione intratumorale, misurata in termini di densità microvascolare (MVD).

Esistono due sottotipi di carcinoma papillare: il tipo 1 in cui la papilla presenta un epitelio monostratificato, a citoplasma basofilo e con nuclei privi di nucleoli evidenti ed il tipo 2, che è invece costituito da papille rivestite da epitelio pseudo-stratificato, a citoplasma più eosinofilo e con grado nucleare più elevato ²² (Fig 7).

a b

Figura 7 : a) tumore renale papillare di tipo I; b) tumore renale papillare di tipo II (Immagine tratta da Radiographics :Papillary Renal Cell Carcinoma: Radiologic-Pathologic Correlation and Spectrum of Disease

pRCC tende a metastatizzare meno frequentemente rispetto a ccRCC e colpisce il sistema scheletrico, il cervello e il polmone che è la sede di metastatizzazione più frequente. ²³ La prevalenza di metastasi viscerali è del 5,7% -11% e i pazienti colpiti da questo tipo di metastasi hanno una prognosi infausta con una sopravvivenza media significativamente inferiore (9,1 mesi) rispetto ai pazienti con carcinoma metastatico ccRCC (22 mesi) ^{24,25}.

1.2.5 Carcinoma renale cromofobo

Il carcinoma a cellule renali di tipo cromofobo (ChRCC) rappresenta circa il 5% delle neoplasie del rene. Macroscopicamente, si presenta come una massa solitaria, circoscritta, e non incapsulata con un' omogenea superficie chiara di colore castano. La dimensione media del tumore è 6,0 centimetri, ed è più grande di quella di altri sottotipi ²⁶. Microscopicamente contiene delle grandi cellule poligonali con membrana cellulare prominente ²⁷. Il citoplasma è chiaro e resistente alle colorazione con ematossilina ed eosina. I vasi sanguigni tumorali hanno pareti spesse ²⁸ (Fig 8)

Figura 8: carcinoma renale cromofobo (Immagine tratta da Journal of Experimental & Clinical Cancer Research Chromophobe renal cell cancer - review of the literature and potential methods of treating metastatic disease)

Il ChRCC comprende un gruppo eterogeneo di carcinomi, compresi il tipo classico, eosinofilo e il tipo misto. La variante eosinofila (contenenti più dell' 80% di cellule eosinofile) è spesso bilaterale (11%) e multifocale (22%) e presenta aree simili a quelle dell'oncocitoma (architettura nidificata alveolare o laminare con granularità eosinofila, accentuazione periferica del citoplasma). Il tipo classico (contenenti più di 80% di cellule chiare) è associata a necrosi o cambiamento sarcomatoide con un' architettura alveolare o laminare.

L'RCC cromofobo con istologia mista ha un'architettura variabile che può contenere una miscela di cellule chiare e cellule eosinofile ²⁹.

1.2.6 Eziologia

Molti sono gli studi che hanno valutato i numerosi fattori che sembrano influire sull'insorgenza del carcinoma a cellule renali ma, a tutt'oggi, non è stato identificato alcun agente eziologico specifico.

Un ruolo importante nello sviluppo della malattia sembra essere rivestito dal fumo di sigaretta: esiste infatti una chiara correlazione dose-risposta in quanto nei fumatori accaniti il rischio relativo aumenta di 2,3 volte rispetto ai fumatori moderati. E' stato osservato, inoltre una riduzione del rischio con la cessazione del fumo: l'eliminazione del consumo di tabacco potrebbe prevenire dal 16 al 28% dei casi di carcinoma del rene considerando una percentuale di fumatori tra il 20 e il 40% nella popolazione adulta. Si osserverebbe inoltre una riduzione del rischio negli ex-fumatori da più di 15 anni sino al 25%, rispetto ai fumatori^{30,31}.

E' stato calcolato l'obesità ha contribuito allo sviluppo dell' RCC in circa il 30% dei casi in Europa e 40% negli Stati Uniti e in Canada tuttavia non è chiaro se l'obesità sia anche associata con l'aggressività e la sopravvivenza alla patologia. Un recente studio conferma che il rischi di sviluppo della malattia aumenta di 1,48 per ogni unità di BMI per pazienti obesi ($BMI > 30 \text{ kg/m}^2$) rispetto a soggetti normopeso. Diversi possibili meccanismi sono stati proposti per spiegare l'associazione tra obesità e un aumentato rischio di alcuni tipi di cancro tra cui quello del rene. Le cellule adipose producono adipochine che possono stimolare o inibire la crescita delle cellule: la leptina, che è più abbondante in persone obese, sembra promuovere la proliferazione cellulare, mentre l' adiponectina, che è meno abbondante in persone obese, può esercitare effetti antiproliferativi. Inoltre le cellule adipose sono coinvolte nella produzione di citochine e ormoni ad azione pro-infiammatoria e molti studi hanno confermato le reciproche influenze fra stati flogistici e fase di iniziazione e di progressione di molti tumori. Il grasso corporeo, inoltre, influenza direttamente alcuni ormoni circolanti, come insulina, IGF (insulin-like growth factor) ed estrogeni, creando un ambiente che favorisce la carcinogenesi e che limita l'apoptosi³².

E' ormai noto che l'ipertensione è un fattore di rischio per lo sviluppo di cancro e il carcinoma renale sembra essere strettamente associato a questa patologia. I pazienti ipertesi hanno un rischio di morire per carcinoma renale maggiore di 2,43 rispetto ai soggetti normotesi; pazienti con ipertensione di stadio 3 hanno un rischio di quasi 4 volte superiore. Molti studi, tuttavia

associano il rischio di carcinoma a cellule renali all' esposizione a lungo termine a terapie diuretiche³³.

Anche se il carcinoma renale non può essere considerato una malattia professionale, diversi studi hanno approfondito la relazione tra esposizioni professionali e la sua insorgenza e si è visto che agenti cancerogeni industriali nel campo della lavorazione della gomma, dei coloranti organici e delle raffinerie metallurgiche, oltre all'esposizione specifica ad amine aromatiche, è correlata ad un alto rischio di sviluppare un carcinoma al rene³⁴.

Esiste una predisposizione familiare e genetica allo sviluppo del carcinoma renale: alcune sindromi ereditarie sono, infatti, associate all'aumentato rischio di sviluppare particolari istotipi di carcinoma renale (10). Tali sindromi includono la sindrome del carcinoma renale a cellule chiare di tipo familiare, il carcinoma renale di tipo papillare ereditario, la sindrome della leiomiomatosi associata a carcinoma renale, la sindrome di Birt-Hogg-Dube e la malattia di von Hippel-Lindau³⁵.

Gli individui affetti da oncocitoma renale familiare sono in grado di sviluppare oncocitoma multifocale bilaterale, o neoplasie oncocitiche nel rene mentre coloro che sono colpiti da carcinoma renale papillare ereditario sviluppano un carcinoma renale papillare bilaterale multifocale.

La leiomiomatosi associata a carcinoma renale è una malattia ereditaria a trasmissione autosomica dominante, caratterizzata dallo sviluppo di leiomiomi cutanei ed uterini multipli e da carcinomi renali papillari di tipo 2 o dei dotti collettori, con una elevata capacità metastatica ed una prognosi infausta. Nella sindrome di Birt-Hogg-Dube, si manifestano tumori benigni del follicolo pilifero ma con alto rischio di sviluppo di neoplasie renali. La malattia di von Hippel-Lindau (VHLD) rappresenta, invece, la sindrome ereditaria più frequente, con una incidenza annuale negli Stati Uniti di 1 caso ogni 35000 persone. Si tratta di una sindrome ereditaria a trasmissione autosomica dominante caratterizzata da una predisposizione a neoplasie primitive del cervello e del cervelletto, che si verificano nel 65% dei casi. Nel 20% dei pazienti insorgono invece neoplasie delle ghiandole surrenaliche, tumori endocrini del pancreas o neoformazioni cistiche pancreatiche multiple. Cisti renali e carcinomi renali si sviluppano in circa il 70% dei pazienti con tale sindrome. Le neoplasie renali nella VHLD sono sempre carcinomi a cellule chiare, si sviluppano precocemente e tendono ad essere bilaterali e multifocali^{36, 37, 38}.

1.2.7 Fattori prognostici molecolari

Il carcinoma renale comprende un gruppo eterogeneo di diversi sottotipi di neoplasie e le analisi genomiche hanno dimostrato che le differenze non si basano solo su criteri istopatologici, ma anche su specifiche mutazioni genetiche³⁹. Una più accurata comprensione delle caratteristiche biologiche e genomiche della malattia costituisce uno strumento efficace non solo per classificare più adeguatamente neoplasie già esistenti ma anche per individuare la neoplasia stessa in una fase pre-invasiva ed inoltre, per sviluppare una terapia mirata⁴⁰.

Le alterazioni a carico del gene VHL appaiono essere responsabili di circa il 60% dei carcinomi a RCC. Il gene VHL codifica per una proteina biologicamente attiva in grado di inibire la proliferazione cellulare, bloccando la risposta cellulare a stimoli ipossici. Il prodotto del gene VHL (Pvhl) è infatti in grado di formare un complesso che funge da ligasi per l'ubiquitina, attivandola. In condizioni di normale ossigenazione cellulare ed in seguito al legame con l'ubiquitina, il complesso VHL sarebbe in grado di riconoscere e degradare la subunità alfa dei fattori di trascrizione noti come Hypoxia inducible factors (HIF-1 alfa, HIF-2 alfa e HIF-3 alfa) Nell'RCC, l'alterazione altamente specifica è la delezione del cromosoma 3p in corrispondenza di distinte regioni cromosomiche: 3p21.3 e 3p25. In condizioni di ipossia o di mutazione/alterata funzione di VHL, HIF-1 α , espresso costitutivamente, dimerizza con la subunità HIF-1 β traslocata nel nucleo e, legandosi alle sequenze consenso, determina la trascrizione di geni proangiogenici indotti dall'ipossia, quali Vascular Endothelial Growth Factor VEGF, Platelet Derived Growth Factor PDGF, Transforming Growth Factor TGF α , Anidrasi Carbonica IX ed Eritropoietina (EPO)⁴¹.

Il ruolo del HIF nello sviluppo del RCC a cellule chiare è stato dimostrato in diversi studi immunohistochimici nei quali una più intensa colorazione si verifica in campioni di pazienti affetti da ccRCC rispetto a campioni di soggetti sani o provenienti da altri rari istotipi di tumore renale. In vitro, linee cellulari prive di pVHL sono in grado di generare tumori e d'altra parte il ripristino dell'espressione di pVHL ne impedisce lo sviluppo solo quando HIF-1 α non è costitutivamente espresso, suggerendo che l'espressione aberrante di HIF1 α ha un ruolo cruciale nella formazione di tumori⁴².

Diversi studi dimostrano l'importanza di VEGF, un gene indotto da HIF che riveste un ruolo fondamentale nell'angiogenesi. VEGF si lega ai recettori tirosin chinasi VEGFR1 (Flt1) e VEGFR2 (Flk1), che innescano diversi network biologici che inducono la proliferazione cellulare, la migrazione e la degradazione della matrice extracellulare⁴³. Rivet et al. hanno

dimostrato un' aumentata espressione di VEGF in tessuti di pazienti con CCRcc rispetto a campioni non tumorali di tessuto⁴⁴. Schips et al hanno mostrato livelli elevati di VEGF nel siero di pazienti con carcinoma renale⁴⁵. In accordo con questi risultati, Sato et al hanno dimostrato un aumento dell' mRNA di VEGF in campioni RCC rispetto a campioni sani⁴⁶. Questi risultati giustificano l'uso di inibitori VEGFR nella terapia avanzata per CCRcc.

Recenti studi si sono concentrati sul ruolo della Rapamycine (mTOR) nella generazione tumorale anche nei tumori renale a cellule chiare. mTOR è una serina / treonina chinasi intracellulare che regola la crescita e la proliferazione cellulare. Viene attivata da segnali provenienti da pathway innescato da diversi fattori di crescita e in particolare mTOR è un effettore a valle di PI3Kinase/Akt e Ras / MAP chinasi. L'up- regulation di molti fattori di crescita come VEGF, PDGFbeta e TGFalfa che si verifica in pazienti affetti da ccRCC, ha spinto i ricercatori a considerare mTOR come un potenziale mediatore di sviluppo del tumore e, quindi, un possibile bersaglio terapeutico. Inoltre, è stato dimostrato che HIF1 α è regolata anche da mTOR, rafforzando il razionale per l'utilizzo di inibitori di mTOR nella terapia CCRcc.⁴⁷

Nell'ccRCC la predisposizione alla tumorigenisi è quindi dipendente da diversi network biologici :

- l'aumento dell'espressione di fattori di crescita, che promuovono la proliferazione e l'angiogenesi

delle cellule epiteliali

- la down-regulation delle integrine, come l' e-caderina, che favorisce la migrazione cellulare,

- l'aumento dell' espressione delle anidrasi carbonica 9 e 12, che, acidificando il microambiente, favoriscono la crescita cellulare e il comportamento invasivo.

Tuttavia, studi in vivo suggeriscono che per generare questo fenotipo maligno di tumore renale l'inattivazione VHL non è sufficiente, anzi VHL è solo un fattore iniziatore di tumori genesi e, quindi, ulteriori mutazioni saranno responsabili dello sviluppo del cancro^{48,49}.

Il *carcinoma a cellule renali di tipo papillare* si caratterizza geneticamente per le trisomie, che riguardano i cromosomi 3q, 7, 8, 12, 16, 17, 20 e per la perdita del cromosoma Y⁵⁰. Queste alterazioni sono presenti tanto nelle forme solitarie che in quelle multifocali ed è stato dimostrato che avvengono precocemente nell'evoluzione della neoplasia. Sono state anche evidenziate differenze tra le alterazioni che coinvolgono il carcinoma papillare di tipo 1 e quello di tipo 2. Nel primo, si presentano con maggior frequenza alterazioni a carico del cromosoma 17q; nel secondo, quelle a carico del cromosoma 9p 51. Il proto-oncogene c-Met, localizzato sul cromosoma 7, codifica per il recettore dell'Hepatocyte Growth Factor/Scatter Factor un fattore di crescita epatocita rio che media la crescita cellulare, la mobilità, l'invasione e la

differenziazione morfogenetica. Sono state trovate delle mutazioni tanto nella forma ereditaria del tumore quanto in carcinomi che si osservano sporadicamente. Infine, i pazienti affetti dalla sindrome ereditaria "leiomiomatosi e carcinoma renale", che si caratterizza per un aumentato rischio di insorgenza di leiomiomi sia cutanei che uterini e di carcinomi a cellule renali di tipo papillare di tipo 2 in forma solitaria, presentano caratteristicamente un'alterazione a livello del gene della fumarato-idratasi, trasmessa con meccanismo autosomico dominante⁵².

Le alterazioni cromosomiche caratteristiche osservate nel *carcinoma a cellule renali di tipo cromofobo* sono le monosomie. Esse riguardano diversi cromosomi quali l'1, 2, 6, 10 e 17⁵³. Raramente, si possono identificare polisomie attribuibili a fenomeni di disgiunzione mitotica durante l'anafase. Sia la variante classica sia quella eosinofila presentano le medesime monosomie, seppur in percentuale diversa. La sindrome di Birt-Hogg-Dubè è una genodermatosi che si caratterizza per la presenza di fibrofollicolomi sulla cute del capo e del collo, meno frequentemente di fibromi molli e tricodiscomi e di cisti polmonari, che causano pneumotorace e tumori renali multipli⁵⁴. Tra questi ultimi, il più comune è il cosiddetto "tumore ibrido", definito come una neoplasia che mostra aree di commistione tra il carcinoma cromofobo e l'oncocitoma. Anche gli altri istotipi tumorali possono insorgere in tale sindrome; in particolare il carcinoma a cellule renali di tipo cromofobo con una frequenza del 35%, il carcinoma a cellule renali chiare nel 9% dei casi, l'oncocitoma nel 5% e il carcinoma a cellule renali di tipo papillare nel 2%. Il gene responsabile di questa sindrome è stato localizzato sul cromosoma 17p11.2 e codifica per una proteina chiamata follicolina, il cui ruolo non è completamente chiarito, ma sembra essere quello di un oncosoppressore. Sebbene la progressione oncocitoma/carcinoma a cellule renali di tipo cromofobo sia stata suggerita sulla base dei dati ottenuti dallo studio dei tumori renali di pazienti con la sindrome di Birt-Hogg-Dubè, altre informazioni, ottenute dallo studio delle forme sporadiche di tali neoplasie nonché delle lesioni osservate nelle oncocitosi renali in pazienti senza tale sindrome, non sono in accordo con tale ipotesi^{53, 55}.

Altri autori hanno valutato il numero di copie di DNA e anomalie del signaling osservando che l'inattivazione del pathway del gene Myc è associata con un andamento più aggressivo nei carcinomi papillari del rene di tipo 2. Analisi computazionali di dati di espressione genica hanno rivelato che una firma trascrizionale indicativa di attivazione del pathway Myc è correlata alla prognosi, in quanto è presente solo nei carcinomi papillari di tipo 2 a comportamento clinico aggressivo^{56, 57}.

1.2.8 Strumenti di diagnosi

Molti masse renali sono asintomatiche e non palpabili fino agli stadi più avanzati della malattia e la principale responsabile del ritardo con cui compare la sintomatologia e la posizione anatomica del rene ⁵⁸. Il tessuto adiposo entro la Fascia di Gerota e nel retroperitoneo conferisce protezione ai visceri e alle strutture somatiche circostanti il rene, rendendo tardivo il loro interessamento e quindi, la comparsa di dolore persistente o di una tumefazione palpabile. La classica triade di sintomi costituita da dolore al fianco, ematuria macroscopica e massa addominale palpabile è rara e si verifica solo nel 6-10% dei casi ^{59,60}. Il 20% dei carcinomi renali si presenta con i segni sistemici di malattia avanzata e fra questi, i più frequenti sono: calo ponderale fino alla cachessia, febbre cronica possibilmente accompagnata da sudorazioni notturne, anemia, dolore osseo metastatico, tosse cronica da coinvolgimento polmonare, cefalea secondaria a ripetizioni intracraniche o comparsa di linfadenopatia cervicale. Durante il decorso della neoplasia, nel 10-40% dei casi può comparire una sindrome paraneoplastica che regredisce dopo l'asportazione del tumore. Si distinguono due tipi di sindromi paraneoplastiche: endocrine e non endocrine la cui insorgenza è probabilmente dovuta ad un'anomala produzione di sostanze ormonali o citotossiche, abitualmente o non abitualmente secrete dal rene o per una risposta abnorme del sistema immunitario alla neoplasia (Tab.1).

ENDOCRINE	NON ENDOCRINE
<i>Ipercalcemia</i>	<i>Disfunzione epatica non metastatica</i>
<i>Ipertensione</i>	<i>Amiloidosi</i>
<i>Policitemia</i>	<i>Anemia</i>
<i>Galattorea</i>	<i>Neuromiopia</i>
<i>Sindrome di Cushing</i>	<i>Vaculopatia</i>
<i>Alterazione metabolismo glucidico</i>	<i>Nefropatia</i>
	<i>Coagulopatia</i>
	<i>Tiroide subacuta</i>
	<i>Tosse ostinata</i>
	<i>Sindrome Penfigoide</i>

Tabella 1: principali sindromi endocrine e non endocrine

L' ipercalcemia paraneoplastica è dovuta alla disidratazione secondaria alla poliuria, per incapacità del rene a concentrare le urine.

Livelli elevati di renina sierica documentati in pazienti affetti da tumori renali di alto grado inducono ipertensione nel 37% dei pazienti. La policitemia è dovuta all'iperincretazione di

eritropoietina prodotta sia dalle cellule tumorali sia dalle cellule peri-neoplastiche, ischemiche per la compressione indotta dal tumore. Tuttavia l'eritropoietina prodotta dalle cellule tumorali sarebbe, in forma inattiva, per cui il quadro ematico presente nei pazienti con carcinoma renale è caratterizzato più spesso da anemia, piuttosto che da policitemia ^{61,62}

Altre sindromi paraneoplastiche di tipo endocrino sono dovute a produzione di sostanze ormonali quali gonadotropina corionica, ACTH, glucagone ed insulina, da parte delle cellule neoplastiche.

Tra le sindromi paraneoplastiche non endocrine, alcune sono estremamente rare come la tiroidite subacuta e la Sindrome Penfigoide ⁶³: la loro associazione con il carcinoma renale non è ben definita.

Di più frequente riscontro è la disfunzione epatica non metastatica e l'amiloidosi che si ipotizza sia dovuta ad un'eccessiva produzione di amiloide dovuta alla prolungata stimolazione del sistema immunitario da parte di sostanze necrotiche liberate dal tumore.

Tra le altre sindromi paraneoplastiche non endocrine elencate nella tabella 1 vi sono vasculiti ⁶⁴, coagulopatie ⁶⁵, nefropatie da deposizione di catene leggere, tosse persistente ⁶⁶ ma la rarità con la quale si manifestano non ha permesso di studiarne né i meccanismi patogenetici né di definire il loro significato prognostico.

L'esame fisico ha solo un ruolo limitato nella diagnosi di RCC, tuttavia una massa addominale palpabile, una palpabile linfadenopatia cervicale e un'edema bilaterale degli arti inferiori, che suggerisce un coinvolgimento venoso, dovrebbero avviare esami radiologici.

Gli esami di laboratorio rappresentano uno strumento molto utile nella diagnosi. I parametri di laboratorio più comunemente valutati sono la creatinina sierica, la GFR, l'emoglobina, gli eritrociti, il velocità di eritrosedimentazione, la fosfatasi alcalina e i livelli ematici di Calcio, di emoglobina e di Lattico-deidrogenasi (LDH).

Un'aumentata concentrazione della creatinina sierica è un indice della compromissione della funzione renale.

Gli elevati livelli di calcio nei pazienti sono dovuti alla sintesi di paratormone e peptidi ormonali attivi da parte di cellule tumorali paratiroideo-simili, evidenziabili mediante tecniche radioimmunologiche ⁶⁷. Studi recenti hanno dimostrato che i peptidi tumorali paratormono-simili, sotto lo stimolo di fattori immunologici come TGF (Trasforming growth factor), TNF (Tumor necrosis factor), IL-1 (Inteleuchina-1) e OAF (Osteoclast activating factor), legano i

recettori del paratormone, inducendo riassorbimento del calcio a livello osseo e ridotta eliminazione a livello renale ^{68,69}.

Valori di emoglobina inferiori ai parametri di riferimento e dosaggi di LDH 1,5 volte superiori alla norma sono stati osservati soprattutto nei pazienti sottoposti a terapia di prima o seconda linea per RCC in fase metastatica.

Attualmente, più del 50% di RCCs vengono rilevati incidentalmente utilizzando tecniche di imaging eseguite in genere per motivi clinici diversi e non legati ad una specifica sintomatologia urologica ^{70,71}. La maggior parte dei tumori renali sono diagnosticati con l'ecografia addominale o TC mentre le tecniche di imaging come ecografia, TAC o risonanza magnetica possono essere utilizzate per distinguere le masse solide dalle cisti renali.

La tomografia computerizzata (TC) o la risonanza magnetica (RM) vengono eseguite sia prima e sia dopo somministrazione di un mezzo di contrasto per via endovenosa e forniscono importanti informazioni quali:

- funzione e la morfologia del rene;
- estensione del tumore primario con l'eventuale diffusione extrarenale;
- coinvolgimento venoso;
- ingrossamento dei linfonodi locoregionali;
- condizione delle ghiandole surrenali e del fegato

Dato che la maggior parte dei carcinomi renali presentano una ricca vascolarizzazione, è apprezzabile un significativo potenziamento della densità della massa dopo l'infusione di mezzo di contrasto.

La TC con contrasto iodato iniettato rapidamente in vena e con l'assunzione di scansioni in fasi diverse (precoce o cortico-midollare, parenchimografica ed escretoria), permette di acquisire tutte le informazioni necessarie per verificare e/o confermare la diagnosi, fornire informazioni sul rene controlaterale ed evidenziare eventuale patologia di altra natura, sia a livello urinario che extraurinario.

La TC con contrasto presenta elevata sensibilità e specificità nella scoperta di un tumore renale di dimensioni > 1.5-2 cm, in base al diverso comportamento della modalità di transito ed impregnazione di contrasto nel tempo, rispetto al tessuto renale sano. La TC è particolarmente importante per la stadiazione del tumore perché consente di documentare la vascolarizzazione arteriosa e venosa del rene e quindi, un'anatomia vascolare utile per la pianificazione terapeutica. Anche le metastasi da carcinoma renale sono in genere ben delineabili per la loro ricca vascolarizzazione quando sono localizzate in sede epatica, splenica e pancreatica ^{72,73,74}.

La RM è ampiamente utilizzata nello studio dei tumori renali perché permette di individuare l'extracapsulare del tumore, dell'infiltrazione vascolare e per la caratterizzazione delle piccole masse solide (< 2 cm), utilizzando sequenze specifiche⁷⁵. Un altro vantaggio non trascurabile è la assenza di radiazioni ionizzanti e l'impiego di un mezzo di contrasto meno nefrotossico rispetto ai composti organo iodati che possono provocare problemi di tossicità soprattutto in pazienti con funzionalità renale compromessa

Una radiografia addominale, così come la scintigrafia scheletrica e la PET, vengono impiegate per la stadiazione ed il follow-up, ma non presentano vantaggi diagnostici ma sono per lo più impiegate per la ricerca di lesioni metastatiche secondarie a carico dell'apparato scheletrico e di altri organi.

Le biopsie sono sempre più utilizzate nella diagnosi e nel follow-up di tumori renali^{76,77} e hanno dimostrato un'elevata specificità e sensibilità per la presenza di neoplasie maligne^{76,78} al fine di determinarne l'istotipo e il grado della massa valutata renale ma nel 10% dei casi il risultato che si ottiene può dirsi conclusivo.

In ogni caso, nessuna di queste metodiche può essere considerata la migliore in assoluto per tutti i pazienti con carcinoma renale. Per questo motivo, le tecniche vengono spesso combinate per ottenere informazioni più complete, soprattutto quando si intende procedere con l'asportazione di tumori di grandi dimensioni.

1.2.9 Fattori prognostici anatomici e istologici

Il fattore più importante per la prognosi è la valutazione dello stadio. Il sistema di classificazione tumorale del Joint Committee on Cancer (AJCC) è conosciuto anche come sistema TNM (*Tumor, Node, Metastasis*) viene generalmente utilizzato per la stadiazione sia per scopi clinici che scientifici e valuta diversi fattori anatomici (Tab. 2)⁷⁹

La lettera T seguita da un numero da 1 a 3 descrive dimensioni del tumore e la diffusione ai tessuti circostanti. Valori di T più alti indicano un tumore più grande e / o la diffusione più estesa ai tessuti in prossimità del rene. La lettera N seguita da un numero da 0 a 2 indica se il tumore si è diffuso ai linfonodi vicino al rene e, in caso affermativo, quanti sono interessati. La lettera M seguita da a 0 o 1 indica se il cancro si è diffuso in organi distanti o ai linfonodi che non sono vicino ai reni⁸⁰

Tumore Primitivo (T)

<i>Tx</i>	Tumore primitivo non definibile
<i>T0</i>	Tumore primitivo non evidenziabile
<i>T1</i>	Tumore della dimensione massima di 7 cm, confinato al rene

T1a	Tumore ≤ 4 cm
T1b	Tumore > 4 cm e < 7 cm
T2	Tumore della dimensione massima oltre i 7 cm, confinato al rene
T2a	Tumore > 7 cm e ≤ 10 cm, confinato al rene
T2b	Tumore > 10 cm, confinato al rene
T3	Tumore che invade le maggiori vene o i tessuti perinefrici, ma non la ghiandola surrenale ipsilaterale, e che non va oltre la fascia di Gerota
T3a	Tumore che invade la vena renale e le sue ramificazioni o il tessuto perineale, ma che non va oltre la fascia di Gerota
T3b	Tumore che invade macroscopicamente la vena cava al di sotto del diaframma
T3c	Tumore che invade macroscopicamente la vena cava al di sopra del diaframma o la parete della vena cava
T4	Tumore che invade oltre la fascia di Gerota (includendo adiacente invasione della ghiandola surrenale ipsilaterale)

Linfonodi Regionali (N)

Nx	Linfonodi regionali non valutabili
N0	Linfonodi regionali liberi da metastasi
N1	Metastasi in un singolo linfonodo regionale
N2	Metastasi in più linfonodi regionali
N3	Metastasi in un singolo linfonodo >5 cm (dimensione massima)

Metastasi a Distanza (M)

M0	Metastasi a distanza assenti
M1	Metastasi a distanza presenti

Tabella 2: classificazione TNM (*Tumor, Node, Metastasis*) per la stadiazione dei tumori del Joint Committee on Cancer (AJCC)

La sopravvivenza causa-specifica a 5 anni varia dall' 88 al 99% negli RCC pT1, dal 70.5 all'82% nei pT2, dal 10 al 60% nei pT3 e fino al 20% nei pT4. Inoltre, nei pazienti con malattia metastatica, la sopravvivenza causa-specifica a 5 anni varia dal 10 al 30% dei casi.

I pazienti con coinvolgimento dei linfonodi regionali presentano percentuali di sopravvivenza significativamente peggiori rispetto ai pazienti con malattia confinata, ma migliori rispetto a quelli con metastasi a distanza ⁸⁰. I pazienti con linfonodi positivi vengono attualmente classificati in due sottogruppi, sulla base dell'interessamento di un singolo linfonodo coinvolto (N1) o di un numero maggiore (N2). Tuttavia, molti dati della letteratura recente hanno evidenziato come a questa suddivisione non corrisponda una differente prognosi. La positività di minimo 4 linfonodi o del 60% dei linfonodi asportati o la presenza di un'invasione extranodale

della metastasi sono i parametri proposti per migliorare la stratificazione prognostica dei pazienti con metastasi linfonodali^{81,82}.

Un miglior quadro della classificazione anatomica del carcinoma renale si può avere dalla tabella 3. nella quale sono elencate le principali caratteristiche dei diversi stadi tumorali:

Raggruppamento in stadi

<i>Stage</i>	T	N	M
<i>I</i>	T1	N0	M0
<i>II</i>	T2	N0	M0
<i>III</i>	T1 or T2	N1	M0
	T3	N0 or N1	M0
<i>IV</i>	T4	Ogni N	M0
	Ogni T	Ogni N	M1

Tabella 3: principali caratteristiche dei diversi stadi tumorali secondo la classificazione TNM

Stadio I: il tumore ha una dimensione di circa 7 cm o minore, è limitato al rene. Non vi è diffusione ai linfonodi o ad organi distanti.

Stadio II: il tumore è più grande di 7,0 centimetri, ma ancora limitato al rene. Non vi è diffusione ai linfonodi o ad organi distanti.

Stadio III: include diverse combinazioni di categorie T e N. Questo stadio comprende tumori di qualsiasi dimensione, con o senza diffusione di tessuto adiposo intorno al rene, con o senza diffusione nelle grandi vene che portano dal rene al cuore, con la diffusione ad un linfonodo vicino, ma senza diffusione a linfonodi distanti o ad altri organi. Lo stadio III comprende anche i tumori con diffusione al tessuto adiposo intorno ai reni, alle grandi vene, tumori che non si sono diffusi a tutti i linfonodi o ad altri organi.

Stadio IV: include diverse combinazioni di T, N, M. Questo stadio comprende tutti i tumori che si sono diffuse direttamente attraverso il tessuto adiposo e le fasce di tessuto che circondano il rene. Vengono inclusi anche qualsiasi tipo di cancro che si è diffuso a più di un linfonodo in prossimità del rene, a qualsiasi linfonodo non vicino al rene, o ad eventuali altri organi, come i polmoni, le ossa, o il cervello.

Lo Stadio I-II ha una prognosi favorevole con una sopravvivenza compresa tra 80-100% dei casi a 10 anni, mentre per lo stadio III-IV la prognosi sfavorevole con una sopravvivenza del 40-5% a 10 anni.

Tuttavia gli esiti del carcinoma risultano significativamente peggiori con l'aumentare dell'età nei pazienti con tumore di media grandezza tra i 4 e i 7 cm; la sopravvivenza relativa a 5 anni è

inferiore nei soggetti malati di età più avanzata (tra 50 e 79 anni) rispetto ai pazienti più giovani (tra 30 e 39 anni)^{83,84}.

La diagnosi istologica nel carcinoma renale viene stabilita dopo la rimozione chirurgica dei tumori del rene o dopo biopsia^{85,86}. Il sistema di classificazione per grado nucleare di Fuhrman (Tab. 4) è la classificazione più generalmente accettata, ed è un importante fattore prognostico nell' RCC e permette di stratificare la morfologia nucleare e la presenza di nucleoli da G1 (nucleo linfocita-simile) a G4 (voluminoso, atipico, polimorfo e con nucleoli evidenti).

Grado 1 -G1-	cellule tumorali di nucleo piccolo(circa 10 µm), uniforme e rotondo privo di nucleoli
Grado 2 -G2-	cellule tumorali con nucleo più voluminoso (circa 15 µm) ed irregolare nel contorno e nucleoli evidenti ad in gradimento 400x.
Grado 3 -G3-	cellule tumorali con nuclei ancora più grandi (circa 20 µm) con evidenti irregolarità dei contorni e nucleoli prominenti anche a piccolo ingrandimento (100x)
Grado 4 -G4-	cellule tumorali con nucleo bizzarro, multi lobulato e deposizioni grossolane di cromatina

Tabella 4: sistema di classificazione per grado nucleare di Fuhrman

La stratificazione dell'informazione prognostica relativa al grading nucleare sulla base del differente istotipo tumorale delle neoplasie analizzate ha portato alla conclusione che questo sistema fornisce una valida informazione prognostica solo se applicato ai tumori a cellule chiare. Al contrario, esistono dati recenti della letteratura che tendono a riconsiderare la correttezza dell'applicazione del grading nucleare di Fuhrman nell'ambito degli RCC ad istotipo papillare e a negarla in quelli a istotipo cromofobo⁸⁷.

Gli RCC papillari e cromofobi presentano uno stadio patologico ed un grading istologico più favorevole rispetto agli RCC a cellule chiare ed hanno una minore probabilità di dare metastasi a distanza^{88,89}. Inoltre, gli RCC a cellule chiare hanno una sopravvivenza causa-specifica significativamente peggiore rispetto agli RCC papillari e cromofobi. Al contrario, non sono state riportate in letteratura differenze significative nella sopravvivenza tra i carcinomi papillari ed i cromofobi. La percentuale di sopravvivenza causa-specifica a 5 anni varia tra il 43 e l'83% per gli RCC a cellule chiare; tra il 61 ed il 90% per gli RCC papillari e tra l'80 ed il 100% per i cromofobi⁸⁸.

La necrosi coagulativa rappresenta un'altra variabile istologica che può presentarsi nei carcinomi renali: si osserva con maggiore frequenza nei pazienti con RCC ad istotipo papillare, in confronto con quelli ad istotipo a cellule chiare o cromofobi. Tuttavia, il valore predittivo

indipendente della necrosi coagulativa è stato dimostrato solo nei pazienti con RCC a cellule chiare ^{90,91}. Anche nei pazienti con RCC ad istotipo cromofobo, il riscontro all'esame istologico di un quadro di necrosi coagulativa potrebbe predire una maggiore aggressività di queste neoplasie ⁹². Tuttavia, altri studi sono arrivati a conclusioni diverse, non confermando il valore prognostico della necrosi coagulativa ^{93,94}. Le differenze principali nei risultati degli studi sono probabilmente imputabili alla mancata standardizzazione della definizione di necrosi coagulativa.

Un altro parametro istologico in grado di influenzare negativamente la prognosi dei pazienti con RCC è il riscontro istologico di una differenziazione sarcomatoide. La sopravvivenza causa-specifica dopo 2 anni dal trattamento chirurgico nei pazienti con differenziazione sarcomatoide è risultata pari al 30%, al 40% ed al 25% in relazione all'istotipo a cellule chiare, papillare o cromofobo. Nei pazienti senza differenziazione sarcomatoide le corrispondenti percentuali sono risultate rispettivamente pari all'84%, al 96%, ed al 96% ⁹¹

1.2.10 Strategie Terapeutiche

Nel 1969, Robson pubblicò quello che è stato considerato il più influente e più citato articolo sui tumori del rene nella letteratura urologica ⁹⁵. Questo studio retrospettivo condotto su 88 pazienti affetti da tumore a cellule renali e sottoposti a nefrectomia radicale evidenziò l'efficacia, in termini di sopravvivenza, della chirurgia radicale rispetto alla nefrectomia semplice. Le ragioni di un intervento così esteso erano giustificate anche dal fatto che, all'epoca, i tumori renali si presentavano prevalentemente come grosse masse e la stadiazione, basata su urografia venosa ed angiografia, era difficile ed approssimativa. Ancora oggi, tale approccio è considerato il gold standard di riferimento nel trattamento dei tumori del rene, in assenza di terapie mediche e radioterapiche efficaci. Tuttavia, con l'introduzione ed il diffuso utilizzo dell'ecografia e della TC, agli inizi degli anni '80 si è assistito ad un cambiamento della storia naturale del tumore, con una diagnosi sempre più precoce e questo ha portato alla progressiva affermazione della chirurgia conservativa, inizialmente riservata alle situazioni di necessità, anche per tumori renali di piccole dimensioni ed in presenza di un rene controlaterale normale. Attualmente, per i tumori renali < 4-5 cm, localizzati e singoli, la nefrectomia radicale è considerata un overtreatment, mentre risulta indicata la chirurgia conservativa ⁹⁶. Serie numericamente importanti di nefrectomia radicale riportano sopravvivenze cancro specifiche comprese tra il 90% ed il 100%

per neoplasie in stadio patologico T1⁹⁷⁻⁹⁹. La sopravvivenza cancro specifica dopo nefrectomia parziale è del tutto sovrapponibile a quella della nefrectomia radicale. I pazienti con malattia chirurgicamente resecabile, tuttavia, al momento della diagnosi, presentano per il 30% una metastasi a distanza, un'ulteriore 30-40%, con patologia localizzata clinicamente al momento della nefrectomia, svilupperà successivamente metastasi⁹⁸. Questi dati rendono evidente la necessità di un'efficace terapia adiuvante ma, relativamente pochi sono gli studi di terapia adiuvante nel RCC per diversi fattori: scarsa efficacia della chemioterapia, tossicità di alcuni schemi di immunoterapia e difficoltà di arruolamento negli studi clinici cooperativi multicentrici.

La radioterapia può giocare un ruolo palliativo nei pazienti metastatici da carcinoma renale se viene pianificata con frazionamenti intensificati, tecniche più evolute ed impiegata con selettività di scelta; i pazienti con discrete condizioni generali e affetti da oligometastasi appaiono beneficiare maggiormente dal trattamento radiante, che può offrire loro un controllo sintomatico duraturo e potenzialmente, favorirne una più lunga sopravvivenza libera da progressione.¹⁰⁰

Le manipolazioni ormonali hanno oramai un valore storico e scarsa validità terapeutica.¹⁰¹

Per quello che riguarda i farmaci chemioterapici, attualmente, non esiste un regime standard di chemioterapia per il carcinoma renale metastatico. I regimi citotossici tradizionali falliscono nel tentativo di impattare significativamente con la storia naturale di questa malattia, anche se talvolta possono essere utilizzati a scopo palliativo, in seconda battuta, soprattutto nei pazienti poor risk, che comunque sono poco responsivi sia alle terapie immunologiche sia a quelle bersaglio-mirate. I farmaci chemioterapici meritano comunque di essere oggetto di sperimentazioni cliniche, anche nelle diverse situazioni prognostiche, ad esempio in combinazione alle nuove terapie biologiche. La chemioterapia, infine, pur con risultati aleatori, rimane il trattamento palliativo di scelta nei pazienti con carcinoma di tipo sarcomatoide ed in quelli con carcinoma di Bellini^{102,103}.

Nello studio condotto da Cochrane Collaboration Group sono stati riassunti i risultati degli studi clinici randomizzati dell'immunoterapia con citochine in pazienti affetti da RCC. L'insieme dei dati di diverse strategie immunoterapiche ha fornito una percentuale del 12.9% di probabilità complessiva di risposta obiettiva (vs 2.5% per gli approcci non immunoterapici e 4.3% per il placebo), con una sopravvivenza media tuttavia di soli 13.3 mesi con un modesto vantaggio in termini di sopravvivenza rispetto ad altre terapie normalmente utilizzate^{104,105}.

L'inibizione dei processi di neoangiogenesi è una strategia terapeutica promettente nella maggior parte dei tumori solidi, ma sembra esserlo particolarmente in questa neoplasia, per la quale, al momento attuale, sono state ampiamente studiate tre classi di farmaci ad attività antiangiogenica:

gli inibitori diretti di VEGF (Bevacizumab), gli inibitori dell'attività tirosin-chinasica dei recettori di VEGF (Sunitinib e Sorafenib) e gli inibitori di mTOR (mammalian Target of Rapamycin), target legato alla sintesi proteica e, ancora una volta, all'angiogenesi (Temsirrolimus ed Everolimus). I risultati positivi ottenuti con questi nuovi farmaci stanno completamente ridisegnando la strategia terapeutica del carcinoma renale avanzato^{106,107,108}.

Markers genetici e molecolari possono, quindi, aumentare la capacità di predire la prognosi dei pazienti in quanto i trattamenti tradizionali sono inefficaci nel trattamento adiuvante del RCC.

1.3 Proteomica

1.3.1 Introduzione alla proteomica

Negli anni '90, il maggiore impulso alla medicina e alla ricerca è stato senza dubbio dato dalla genetica: nacque infatti la genomica, una branca della biologia molecolare che si occupa dello studio del genoma degli organismi viventi. È una scienza che si basa sulla bioinformatica per l'elaborazione e la visualizzazione dell'enorme quantità di dati che produce. Un progetto pubblico di sequenziamento iniziò nel 1986 sotto il nome di "Progetto genoma umano". Ma un'azienda privata riuscì a completare (nel 2000) il sequenziamento del genoma umano in anticipo sul progetto pubblico. La sequenza fu pubblicata nel 2001 in due articoli paralleli sulle riviste Science e Nature, che riunivano i risultati di entrambi i progetti; si trattava di una bozza pari al 90% della sequenza e ancora con notevoli probabilità di errori ma una sequenza accurata pari al 99% del genoma umano è stata pubblicata nel 2003.

Negli ultimi anni la genomica ha fornito una enorme quantità di dati ed informazioni su molte patologie attraverso l'individuazione di alterazioni sulla struttura e sull'espressione poligenica^{109,110}. Con la conclusione del Progetto Genoma è diventato definitivamente chiaro come la conoscenza delle sequenze genomiche non sia di per sé sufficiente a spiegare tutti gli eventi biologici complessi e tutte le patologie rendendo così necessario lo studio dei prodotti genici: indispensabile, quindi, lo studio del proteoma. Il termine proteoma viene introdotto da Wilkins et al nel 1994 e con esso si fa riferimento al complemento proteico del genoma¹¹¹. Il suo studio si rende necessario innanzitutto perché la correlazione tra quantità di mRNA e quantità di proteina intracellulari è scarsa e variabile tra le diverse proteine: in seguito alla trascrizione da DNA ad RNA messaggero quest'ultimo può subire processi di maturazione differenti, detti di *splicing* alterativo, che portano alla formazione di differenti prodotti genici prima di essere tradotti in una sequenza aminoacidica.

Il fatto ancor più rilevante è che molte proteine, dopo la traduzione, subiscono modificazioni post-traduzionali, come ad esempio glicosilazioni e fosforilazioni, che giocano un ruolo fondamentale nella modulazione dell'attività delle proteine e rendono il numero di isoforme delle proteine di fatto superiore a quello prevedibile sulla sola base dell'analisi degli acidi nucleici.

Per questi motivi all'interno dell'organismo umano è stimata la presenza di 30.000-40.000 geni, mentre il numero di proteine codificate è pari a circa 2 milioni. Questo significa che, statisticamente, l'informazione di un singolo gene può portare alla codifica di circa 50 specie proteiche differenti.

Inoltre le modificazioni post-traduzionali sono il risultato di più vie metaboliche intersecantesi, è un processo altamente regolato, controllato e suscettibile di cambiamenti in seguito a processi morbosi o esposizione a farmaci e tossine^{112,113}

Poiché sono le proteine ad essere direttamente coinvolte nei diversi processi biochimici che si svolgono all'interno di una cellula e che possono essere alla base di una patologia, è parsa chiara la necessità di effettuare studi approfonditi sulle proteine presenti all'interno della cellula o in un tessuto patologico al fine di ottenere informazioni utili a comprendere gli eventi biochimici che avvengono all'interno delle cellule in risposta a determinati stimoli di varia natura oppure che portano alla loro "trasformazione" in cellule affette da una determinata patologia¹¹⁴.

La proteomica costituisce l'insieme delle tecnologie e degli approcci utilizzati per lo studio delle proteine e rappresenta il principale strumento per colmare il divario tra le informazioni che derivano dalla sequenza del genoma e l'effettiva fisiologia della cellula tramite lo studio dei prodotti genici e le loro interazioni. La proteomica si sviluppa su diversi livelli (Tab. 5) ed in particolare analizza e caratterizza la struttura, la funzione e l'espressione di tutte le proteine nell'ambito di una cellula, un tessuto, un liquido biologico o un organismo in un dato momento e soprattutto di inserirla esattamente all'interno dei differenti *pathway* cellulari in modo da poter analizzare in maniera completa situazioni fisiologiche o patologiche complesse, effetti tossici di farmaci ed agenti inquinanti e le interazioni fra stimoli di varia natura tramite l'analisi di fluidi biologici come liquor, siero ed urine, e materiali biologici più complessi come organi, tessuti e cellule^{115,116}

TIPOLOGIA

Protomica sistemica

OBIETTIVI

Identificazione e caratterizzazione del proteoma, oltre che costruzione di database, al fine di creare mappature di riferimento delle proteine espresse in vari tessuti e fluidi biologici

<i>Proteomica differenziale o di espressione</i>	Studio dell'espressione differenziale delle proteine in cellule diverse di uno stesso organismo ed in momenti di vita differenti di una stessa cellula. Confronto dei profili di espressione proteica in varie situazioni biologiche e dell'evoluzione del proteoma a seguito di trattamenti farmacologici
<i>Proteomica funzionale</i>	Studia le interazioni fisiche e funzionali fra le proteine ed i pathway nei quali sono coinvolte
<i>Proteomica strutturale</i>	Si occupa di determinare la struttura tridimensionale delle proteine.

Tabella 5: diverse tipologie di proteomica

Gli esperimenti di proteomica spesso implicano l'analisi simultanea di alcune migliaia di specie proteiche derivanti da campioni biologici complessi. La separazione dei peptidi e delle proteine è dunque l'elemento chiave dell'analisi proteomica, in quanto permette di semplificare miscele proteiche complesse e portare le molecole alla sorgente di ionizzazione.

La scelta della metodologia di separazione dipende dal fine sperimentale e dal livello a cui si desidera indagare le proteine. Due sono gli approcci proteomici usati: l'approccio bottom-up e quello top-down ¹¹⁷.

Nell'approccio bottom-up, le miscele proteiche complesse sono digerite enzimaticamente in miscele peptidiche ancora più complesse, che verranno in seguito frazionate mediante vari step cromatografici prima di essere sottoposte ad indagini in spettrometria di massa tandem (MS/MS). Gli spettri MS/MS sono registrati in modo da rilevare il maggior numero di peptidi possibili ed i risultati sottoposti alla ricerca in banca dati al fine di identificare le proteine presenti nelle miscele originali. Questo approccio è altamente informativo e adatto ad essere automatizzato ¹¹⁸. Queste procedure tipicamente identificano un numero molto limitato di peptidi per proteina ma sufficienti per identificare il gene da cui la proteina viene trascritta. Questo approccio funziona bene per quegli organismi in cui l'assunzione che un gene codifichi per una proteina è valida mentre è meno efficace per gli eucarioti superiori, dove processi come lo splicing alternativo, l'RNA editing e le modificazioni post-traduzionali fanno corrispondere ad un singolo gene diverse specie proteiche. Dunque le strategie basate sui peptidi permettono l'identificazione dei geni da cui derivano le proteine ma non identificano adeguatamente le differenti isoforme funzionali delle proteine ¹¹⁹.

Nella strategia top-down, invece, lo studio si focalizza sulla proteina intatta e le differenti isoforme possono essere isolate prima della loro identificazione e caratterizzazione mediante spettrometria di massa. La frammentazione delle proteine non avviene in soluzione ma in fase

gassosa, direttamente nello spettrometro di massa, in modo da creare una serie di ioni indicativi della sequenza. La differenza nei valori m/z degli ioni frammento definiscono la posizione e la sequenza degli amminoacidi nella proteina ¹²⁰. Molti sono i metodi e gli spettrometri usati per questo tipo di analisi. Un esempio è costituito dall'ESI-FTMS capace di analizzare proteine intatte grazie alla sua elevata risoluzione. Le proteine sono frammentate attraverso la tecnica ECD (Electron Capture Dissociation). In effetti la spettrometria FTMS in seguito all'alta risoluzione permette di usare miscele semplici di proteine riducendo la dipendenza dell'approccio proteomico dai metodi separativi e restituisce dati che consentono di rilevare errori nella sequenza amminoacidica e modificazioni post-traduzionali ¹²¹

1.3.2 Tecniche di proteomica

Il campo della proteomica si è sviluppato fortemente in quest'ultimi anni, anche grazie alla tecnologia ed alla ricerca applicata nello studio di tecniche o strumentazioni, che hanno permesso di rilevare quantità di proteine dell'ordine di picogrammi così come riuscire a separare le proteine stesse nell'ordine di microgrammi/nanogrammi.

Inizialmente, la proteomica si è focalizzata sulla creazione di mappe proteiche di estratti cellulari mediante l'utilizzo della tecnica di *elettroforesi bidimensionale (2DE-PAGE)*, che separa le proteine prima per punto isoelettrico e poi in base al peso molecolare ^{122,123}. La successiva identificazione avviene tramite spettrometria di massa con una tecnica denominata peptide mass fingerprinting (PMF). L'analisi del proteoma può essere quindi suddivisa in una prima estrazione di tutti i costituenti proteici dal materiale di studio (tessuto, linee cellulari o fluidi biologici) seguita dalla separazione della miscela di proteine con un efficace sistema elettroforetico. Dopo idrolisi enzimatica in gel delle proteine di interesse, vengono acquisite delle di informazioni strutturali tramite spettrometria di massa tipo matrix-assisted laser desorption ionization – time of flight (MALDI-TOF) ed infine, l'identificazione viene ottenuta grazie a strumenti bioinformatici che permettono il confronto delle masse sperimentali dei peptidi proteolitici ottenuti con quelle teoriche attese dalla sequenza amminoacidica delle proteine presenti nei database. Uno degli svantaggi di questa tecnica è che non tutte le proteine possono essere separate come, ad esempio, proteine molto idrofobiche o proteine di membrana, inoltre possono essere analizzate solo proteine con un punto isoelettrico comprese tra 4 -10 e con un peso molecolare compreso tra 20-100 KDa. Altri problemi relativi a questa tecnica sono la scarsa riproducibilità e l'elevata laboriosità che non può essere sopperita dall'automazione.

Le recenti innovazioni tecnologiche nel campo della proteomica hanno tuttavia permesso di aumentare la capacità separativa attraverso la cromatografia liquida con la strategia MudPit (Multidimensional Protein Identification Technology). Tale tecnica consiste in una separazione cromatografia in due-dimensioni, prima dell'analisi in spettrometria di massa. Sfruttando le proprietà fisiche uniche di un peptide come il punto isoelettrico e l'idrofobicità, possono essere separate miscele complesse di proteine e peptidi. La prima dimensione è normalmente costituita da una colonna cromatografica a forte scambio cationico (SCX), poiché queste hanno capacità di carico elevate. La seconda dimensione è invece una colonna cromatografica a fase inversa (RP), che integra la precedente colonna in quanto è molto efficace nel rimuovere sali e ha il vantaggio di essere compatibile con la spettrometria di massa elettrospray (ESI-MS). Dopo aver ottenuto lo spettro di massa, lo strumento può automaticamente selezionare uno o più ioni e promuovere la loro frammentazione mediante collision induced decomposition (CID) con gas quali azoto o argon. Vengono così prodotte numerose serie di ioni (MS/MS), dalle quali si possono ottenere informazioni complete o parziali sulla sequenza amminoacidica. Inoltre, mediante questa metodica, è possibile caratterizzare numerose modificazioni post-traduzionali come ad esempio la fosforilazione^{124,125}.

Tuttavia il principale scopo delle indagini proteomiche è quello di identificare proteine che presentano differenze quali/quantitative e tali modificazioni possono presentarsi non solo negli organi, nei tessuti e nelle cellule, ma possono riflettersi anche nei liquidi biologici. Due sono le tecniche automatizzabili che sono state sviluppate per individuare potenziali marcatori biologici all'interno dei fluidi biologici in grado di caratterizzare e predire una patologia e alla creazione di profili proteici che permettano una diagnosi precoce e un chiarimento dei meccanismi di insorgenza della malattia¹²⁶: SELDI e CLINPROT.

1.3.4 Analisi di miscele complesse: elettroforesi e cromatografia

Prima di determinare in dettaglio le proprietà, la composizione in amminoacidi e la sequenza di una proteina è necessario separarla da tutte le altre. L'elettroforesi è una tecnica che si basa sulla migrazione delle proteine cariche in un campo elettrico e non viene solitamente usata per purificare grandi quantità di proteine ma è una tecnica vantaggiosa per valutarne alcune proprietà come ad esempio il punto isoelettrico¹²⁵.

L'*elettroforesi* delle proteine viene di solito effettuata in gel costituiti da polimeri caratterizzati da numerosi legami crociati che agiscono come setacci molecolari e rallentano la migrazione delle proteine in base al loro rapporto carica/massa molecolare.

La *cromatografia* è una tecnica di separazione che consiste si basa sulla diversa attitudine che ogni molecola o ione possiede nel distribuirsi fra due differenti fasi.

I primi tentativi di tecniche cromatografiche sono attribuibili al botanico russo Michail Tsweett e risalgono agli inizi del'900. Egli riempì una colonna con CaSO₄ (gesso) in polvere, vi depositò in testa un estratto di foglie verdi ed eluì con etere di petrolio. I vari pigmenti fogliari si separarono in bande colorate, e per questo che fu coniato il termine cromatografia¹²⁶.

La cromatografia si avvale di due fasi: la fase stazionaria, generalmente un solido, e di una fase mobile detta anche eluente che può essere liquida o gassosa in base alla tecnica utilizzata.

Nella cromatografia su colonna la fase mobile fluisce verso il basso per gravità (nei dispositivi più semplici) o per l'azione della pressione (nel caso di mezzi più sofisticati come per l'HPLC o nella GC). La cromatografia trova fondamento nella diversa affinità di un analita per le due fasi, mobile e stazionaria: maggiore è la propensione di un componente per la F.S. e più tempo questi impiegherà ad attraversare la colonna; diversamente se l'analita ha maggiore affinità verso la F.M., questi ultimerà il percorso della colonna con maggior rapidità¹²⁷.

Oggi con il termine cromatografia si indicano in genere tutte le varie tecniche separative, applicabili a miscele di sostanze come si può vedere dalla figura 9.

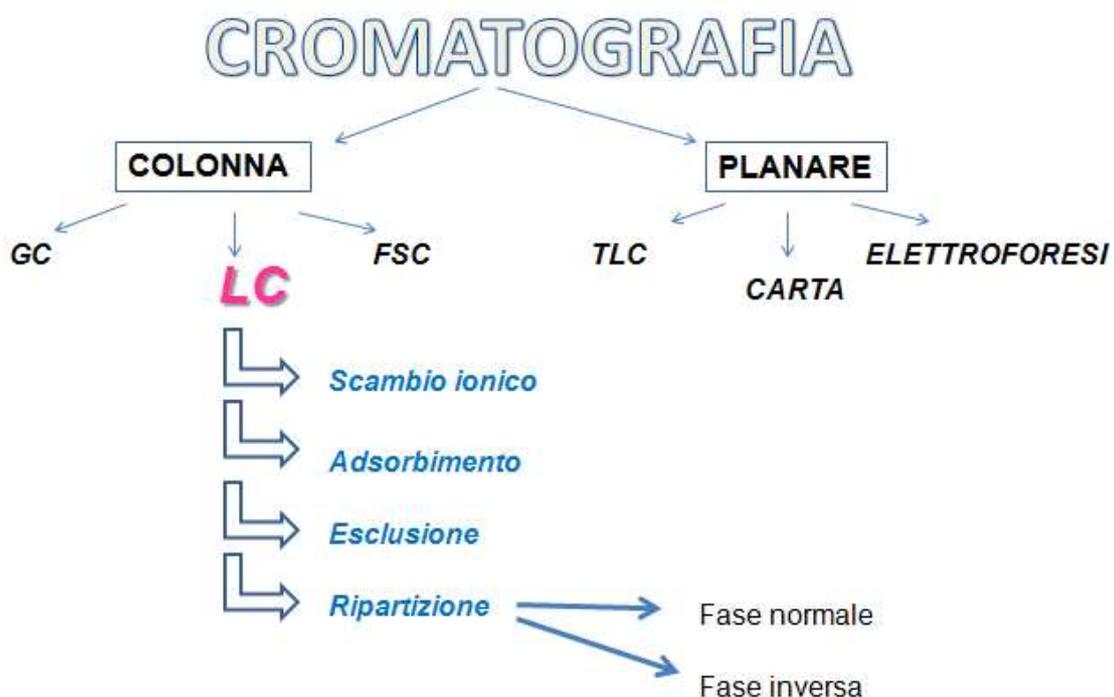


Figura 9: tecniche di separazione cromatografica

Nella *cromatografia per scambio ionico* la ritenzione si basa sull'attrazione tra gli ioni del soluto e i siti carichi legati alla fase stazionaria. La fase stazionaria è costituita da una resina a scambio ionico, ossia un supporto organico al quale sono legati dei gruppi ionici o ionizzabili I

gruppi ionici della resina trattengono per mezzo di interazioni elettrostatiche, controioni di carica opposta che possono essere scambiati con gli ioni presenti nella fase mobile. Il meccanismo di separazione è basato perciò sulla diversa affinità che i diversi soluti (ionici) presentano nei confronti dei gruppi attivi della resina. Gli scambiatori sono *cationici* quando contengono delle cariche negative e i cationi sono attratti elettrostaticamente da questi siti di carica opposta mentre *anionici* quando contengono dei gruppi carichi positivamente in grado cioè di legare degli ioni di carica negativa¹²⁸.

Nella *cromatografia di adsorbimento*, la fase stazionaria è un materiale solido costituito da particelle con grande area superficiale: in questo caso il solido non è un semplice supporto inerte, ma è in grado di adsorbire superficialmente i soluti da separare. Il meccanismo di separazione è quindi determinato dalla diversa affinità dei soluti per la fase solida costituita in genere da gel di silice, allumina, carbone attivo.

La *cromatografia per esclusione molecolare*, detta anche *gel filtrazione*, separa le molecole in base alle loro dimensioni. La matrice presente nella colonna è un polimero con molti legami crociati che generano pori con un diametro specifico. Molecole più grandi potranno migrare nella colonna più velocemente in quanto non possono entrare nei pori della matrice al contrario, quelle più piccole arriveranno con un percorso più lungo perché potranno entrare nei pori.

Nella *cromatografia di ripartizione*, la fase stazionaria è un liquido, immiscibile con la fase mobile, che impregna un supporto solido inerte; questo tipo di cromatografia è basato sulla diversa ripartizione di due o più soluti fra le due fasi liquide. Si possono distinguere due differenti metodiche:

-*cromatografia liquida in fase normale*: la fase stazionaria è più polare della fase mobile (che è un solvente organico); gli analiti vengono eluiti in ordine di polarità crescente.

-*cromatografia liquida in fase inversa*: la fase stazionaria è meno polare della fase mobile. La fase stazionaria è costituita da gruppi alchilsilanici legati chimicamente alla silice (butile, ottile, ottadecile). Come fase mobile si possono utilizzare solventi organici (metanolo, acetonitrile, tetraidrofurano), ma anche tamponi acquosi e l'eluizione degli analiti avviene in ordine di polarità decrescente¹²⁷.

1.3.4.1 Cromatografia liquida HPLC

Il metodo strumentale della cromatografia liquida ad alta pressione HPLC è frutto dell'evoluzione tecnologica delle tecniche di cromatografia su colonna e, ad oggi, è considerata la tecnica cromatografia più diffusa, efficace e versatile perché è adatta sia per analizzare composti con bassa tensione di vapore come macromolecole, proteine, polisaccaridi, composti

ionici, che per analizzare prodotti termolabili come aminoacidi, coloranti, acidi nucleici. Questi composti vengono iniettati in colonna senza subire modificazioni chimiche a differenza della Gas Cromatografia (GC) il cui passaggio limitante è rappresentato dalla derivatizzazione ovvero una reazione che permette di aumentare la volatilità degli analiti, di ottenere una migliore risoluzione e di aumentare la sensibilità analitica abbassando i limiti di determinazione strumentali. Il problema è che la derivatizzazione risulta critica in quanto deve garantire una reazione possibilmente completa, sicuramente irreversibile e riproducibile, in tempi brevi. L'HPLC presenta quindi il vantaggio di poter analizzare composti non facilmente vaporizzabili o che non sopportano alte temperature perchè si decompongono (termolabili).

I principi dell' HPLC sono quelli della cromatografia per adsorbimento e di ripartizione, ma le fasi stazionarie sono impaccate in colonne chiuse, con materiali di granulometria molto fine (5-10 µm): in tal modo viene aumentata la superficie di contatto fra fase mobile e fase stazionaria e l'impaccamento diviene più omogeneo.

Utilizzando queste colonne è necessario che la fase mobile venga fatta fluire ad alta pressione perché, attraverso colonne con impaccamento a granulometria così fine, il flusso dell'eluente diventa molto lento, quindi con l'impiego di pompe particolari, capaci di applicare pressioni di 50-150 atm, diventa possibile ottenere flussi di alcuni ml/min, sufficienti ad ottenere l'eluizione in tempi ragionevolmente brevi.

Nei sistemi a gradiente di eluizione, i due solventi impiegati (uno a polarità bassa o nulla e l'altro a polarità elevata) vengono prelevati dai due serbatoi in modo programmato per mezzo di due pompe e quindi mandati ad alta pressione in una camera di miscelazione; da qui la miscela passa in colonna.

I campioni possono essere iniettati in colonna attraverso apposite microsiringhe per HPLC perforando un setto e inserendo il campione in una camera di iniezione, con il flusso di eluente bloccato, quindi riattivando il flusso con valvole di iniezione che rendono possibile l'introduzione del campione senza interrompere il flusso dell'eluente (anche a pressioni elevate) Il campione iniettato va a finire in un serpentino mentre l'eluente continua a fluire nella colonna, spinto dalla pompa. Ruotando la valvola, si attiva un altro percorso dell'eluente che, prima di entrare in colonna, passa attraverso il serpentino in cui era stato introdotto il campione trasportandolo all'interno della colonna^{125, 126}.

Le colonne per HPLC sono generalmente d'acciaio inossidabile e sono adatte a resistere a pressioni elevate. Sono chiuse ai due estremi con setti porosi di acciaio sintetizzato, resistenti alle pressioni di esercizio; servono a bloccare la fase stazionaria all'interno della colonna.

In commercio si trovano colonne a lunghezza di 5 cm a 25 cm in quanto maggiore è la lunghezza migliore sarà l'efficienza separativa. L'esperienza dimostra che l'efficienza di una colonna è direttamente collegata al numero di piatti teorici.

Il numero di piatti teorici (N) caratteristici di una colonna è dato dalla relazione:

$$N = 16 (t_R/w_b)^2$$

dove W è l'ampiezza del picco e T_v è il tempo di ritenzione, in altre parole il tempo che intercorre fra il momento dell'iniezione e la cresta del picco.

Invece facendo riferimento al tempo di ritenzione corretto, si definisce il numero dei piatti effettivi

$$N_{\text{eff}} = 16 (t'_R/w_b)^2$$

Per migliorare l'efficienza di una colonna, bisogna quindi fare in modo che il numero di piatti N sia più grande possibile. Ciò può essere conseguito o aumentando la lunghezza della colonna od ottimizzando tutti quei parametri che determinano il valore di N quindi aumentando la velocità lineare (in mm/sec) della fase mobile

Un altro componente fondamentale di uno strumento per HPLC è il rivelatore, dispositivo sensibile al passaggio di un composto e le cui prestazioni si valutano in base ai parametri di sensibilità, intervallo di linearità, limite di rivelabilità^{127, 128}.

Tra i principali sistemi di rivelazione vi sono:

- rivelatori conduttometrici
- rivelatori rifrattometrici
- rivelatori spettrofotometrici UV
- rivelatore a fluorescenza
- spettrometri di massa

1.3.3 Analisi dei fluidi biologici: la tecnologia SELDI e Clinprot

La tecnica SELDI (*surface-enhanced laser desorption/ionization*) sfrutta la capacità delle cromatografia di separare le proteine in base alle loro proprietà. Il sistema di prefrazionamento del campione è costituito da superfici, ProteinChip, che possono essere funzionalizzate con gruppi chimici ad esempio gruppi cationici, anionici, catene idrofobiche, e gruppi polari o modificate chimicamente con l'uso di anticorpi, recettori, ligandi, enzimi e acidi nucleici.

Figura 10: immagine di Protein Chip (Immagine tratta dal sito <http://biophy.uchicago.edu/seldi.php>)

Solo un gruppo di proteine presenti nel campione biologico possono essere trattenute dalla superfici attivate e quindi rilevate attraverso analisi MALDI-TOF (*Matrix Assisted Laser Desorption Ionization-Time of Flight*). I profili spettrali ottenuti vengono poi processati con opportuni software di indagine statistica per l'identificazione di segnali differentemente espressi nelle popolazioni in studio¹²⁹.

La tecnica SELDI può essere applicata ad un numero elevato di campioni permettendo di conferire all'indagine una forte valenza statistica, inoltre, possono essere analizzate quantità molto piccole di campione (0.5-400 µl) e i risultati vengono ottenuti in breve tempo.

Una valida alternativa al SELDI è rappresentata dalla tecnologia ClinProtTM che utilizza biglie magnetiche ricoperte da superfici funzionalizzate che permettono la pre-purificazione del campione prima dell'analisi in MALDI-TOF^{130, 131}.

Nella metà degli anni 90, Girault S. et al. scrivono il primo lavoro riferito a questo approccio descrivendo l'uso di biglie magnetiche funzionalizzate con streptavidina per legare peptidi biotinilati. I peptidi sono stati eluiti, miscelati con matrice ed analizzati tramite MALDI-TOF. Gli autori descrivono questa tecnologia come un promettente strumento in grado di studiare le interazioni peptide/proteina. Successivamente, l'applicazione di questa tecnologia è stata applicata alla scoperta dei biomarcatori¹³². Un Anti-PSA-IgG-biotina è stato legato alla streptavidina che rivestiva le biglie magnetiche e usate per arricchire e purificare dal PSA il siero con eccellenti risultati.

Gli studi si sono quindi interrotti fino all'inizio del XX secolo, quando, guidato dal successo della tecnologia SELDI, l'approccio legato alle biglie magnetiche ha iniziato ad essere utilizzato più frequentemente negli studi clinici come testimoniato dall'incremento del numero di pubblicazioni volte, soprattutto, alla ricerca di biomarker in fluidi biologici.

Villanueva e i suoi collaboratori sono stati i pionieri nel profiling di campioni di siero con l'uso di biglie magnetiche funzionalizzate con catene di carbonio con diversa lunghezza (C1, C2, C3, C8 e C18) e con superficie porosa o non porosa. La diversa capacità di legame esercitata dalle biglie magnetiche espressa come numero dei picchi osservati in uno spettro MALDI-TOF, ha mostrato che biglie a catena C8 avevano una capacità estrattiva maggiore rispetto alle altre. Recentemente, uno studio comparativo tra pazienti affetti da cancro al seno e soggetti sani ha valutato la migliore strategia di analisi con biglie magnetiche funzionalizzate con vari gruppi chimici: a interazione idrofobica (C8), a debole scambio cationico (WCX), immobilized metal-

ion affinity (IMAC-Cu) 9. Il risultato emerso è che con le biglie magnetiche WCX si ottengono il maggior numero di segnali e quindi, costituiscono un ottimo punto di partenza per questo tipo di analisi¹³³.

Con un altro lavoro, Callesen e coll.10 hanno valutato non solo le biglie magnetiche commercialmente disponibili (C8 e IMAC-Cu), ma anche microcolonne personalizzate (C8, C18, IMAC-Cu e anionexchange) utilizzando sia i protocolli indicati dal produttore sia protocolli modificati ad hoc¹³⁴. Dopo un' accurata ottimizzazione delle procedure analitiche, il più alto numero di picchi riproducibili (più di 400) è stato ottenuto con la purificazione ottenuta con biglie IMAC-Cu e analisi MALDI-TOF utilizzando 2,6-DHAP come matrice.

Uno studio dettagliato sulle condizioni ottimali per il pre-frazionamento del siero è stato descritto con le biglie magnetiche C18 che ha confermato anche l'elevata riproducibilità di questa tecnica¹³⁵.

Oltre alle biglie magnetiche impiegate in uno studio di proteomica, vi sono altri fattori che possono influenzare la qualità degli spettri di massa: l'origine e la raccolta di materiale biologico, l'ottimizzazione e la tecnica MS utilizzata sono le variabili indagate in diversi lavori di ricerca. Nel 2005 Baumann et al hanno dimostrato che cicli di gelo-scogelamento possono notevolmente influenzare il profilo proteico del siero ottenuto utilizzando biglie magnetiche (C3, C8, C18). Tuttavia, con un protocollo altamente standardizzato per la raccolta del sangue, per la preparazione del siero, del tempo di scongelamento, la variabilità intra e inter-saggio, misurata con i segnali compresi tra 1-10 KDa e con intensità bassa, media e alta, era compresa rispettivamente fra il 3% -23% e del 6% -33% (% CV). Due studi sono stati effettuati per indagare gli effetti della manipolazione del siero sul profilo spettrale: il primo ha mostrato che la temperatura di conservazione è uno dei fattori principali tra i maggiori responsabili delle modifiche degli spettri di massa, Il secondo studio ha dimostrato invece che campioni di siero e plasma possono essere conservati a 4 ° C fino a 24 ore senza alcuna modifica significativa del profilo proteico¹³⁶.

In un altro studio è stato dimostrato che alcuni fattori quali temperatura e tempo della preparazione del siero, l'umidità e la temperatura del processo di cristallizzazione influenzano i profili ottenuti con la tecnologia MALDI-TOF. In particolare la percentuale di umidità durante il processo di cristallizzazione tra campione e matrice influisce sulla formazione di cristalli indispensabili per l'analisi in spettrometria MALDI-TOF. Recentemente, è stata valutato anche l'utilizzo di diverse matrici negli studi di profiling sierico¹³⁷.

Esistono diverse matrici, ognuna adatta ad una particolare esigenza analitica Ad esempio l'acido α -Ciano-4-Idrossicinnammico (HCCA) è adatto per analisi di peptidi o proteine con massa

inferiore a 10 KDa e quindi per la maggior parte degli studi di proteomica, mentre l'acido Sinapinico (SA) è adatto per proteine intere o peptidi con massa superiore ai 10 KDa. Gli autori hanno dimostrato che la scelta di una matrice adeguata ha una maggiore influenza rispetto al processo utilizzato per la preparazione dei campioni ¹³⁸.

Tutte le procedure del processo di pre-purificazione dei campioni possono essere automatizzate mediante l'utilizzo di una piattaforma robotica, come ad esempio il *clinprot robot*.

L'automazione, oltre a rendere più veloce ed economico il processo di purificazione, permette di ottenere una maggiore riproducibilità rispetto ad una preparazione dei campioni manuale con semplici passaggi.

Nella fase di legame (**binding**) il fluido biologico viene miscelato con una sospensione di biglie magnetiche e una soluzione di legame, al fine di ottenere l'estrazione specifica di peptidi dalla matrice biologica.

Successivamente grazie all'ausilio del separatore magnetico, le biglie rimangono adese allo stesso e, in tal modo, viene favorita l'eliminazione del surnatante e, quindi, delle sostanze legate in modo aspecifico alle biglie (fase di eliminazione del surnatante)

Durante la fase di lavaggio (**washing**) il robot effettua una serie di 3 lavaggi consecutivi con un tampone di lavaggio. Le proteine legate in maniera aspecifica, i sali e i contaminanti, che possono interferire con la successiva analisi in spettrometria di massa vengono, in questa fase, eliminate.

Viene, dunque, aggiunta e miscelata una soluzione per favorire la rottura del legame tra proteine/peptidi e biglie per permettere l'eluizione degli analiti. (**elution**).

Nell'ultima fase, quella di deposito, (**spotting**), l'eluato viene miscelato nella matrixplate con la soluzione di matrice e, dopo il deposito di alcuni microlitri di questa miscela sul target (AnchorChip 600/384) e l'evaporazione del solvente, il campione co-cristallizza con la matrice.

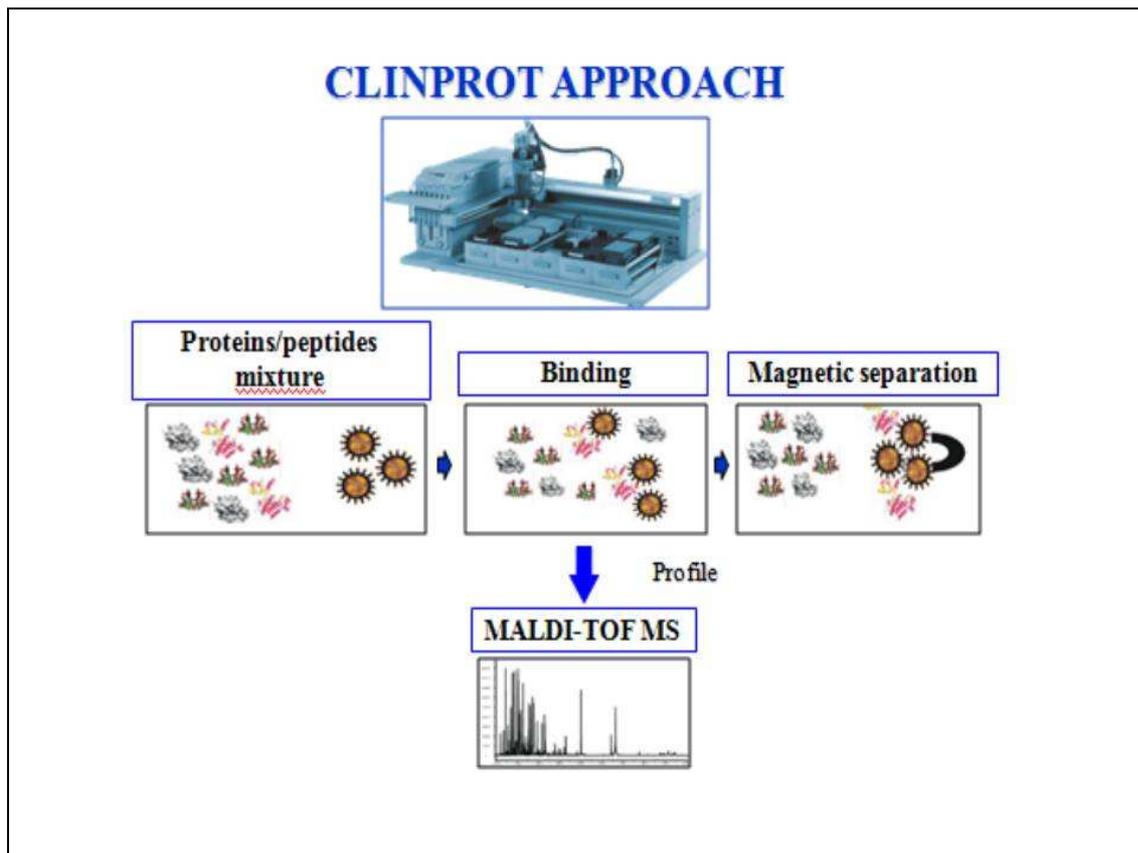


Figura 11: strumentazione e passaggi relativi alla tecnica Clinprot

1.4 La spettrometria di massa

1.4.1 Introduzione e principi fondamentali

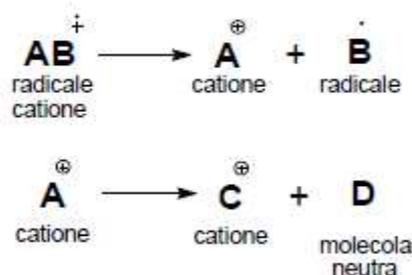
La spettrometria di massa trae le sue origini, nella prima parte di questo secolo, dalle esperienze del fisico J.J.Thompson, il quale mise in evidenza la formazione di elettroni e “radiazioni positive” in un tubo posto sotto vuoto, nel quale veniva applicata una differenza di potenziale elettrico. Thompson osservò che tale nuova tecnica poteva essere utile al chimico per analizzare sostanze, come descritto nel suo libro *“Rays of Positive Electricity and Their Application to Chemical Analysis”*.¹ Nonostante questa lungimirante osservazione, l’applicazione primaria della spettrometria di massa rimase confinata nel campo della fisica per circa trent’anni. Essa è stata usata in particolare per individuare uno svariato numero di isotopi, per determinare la loro abbondanza relativa e per misurare la loro massa esatta, ossia la massa atomica con una precisione di 1 parte per milione o anche superiore. Queste fondamentali e importanti misure hanno permesso di creare le basi per sviluppi successivi in diversi campi che vanno dalla geocronologia alla ricerca biochimica¹³⁹.

In proteomica, la spettrometria di massa non rappresenta soltanto la procedura di elezione per ottenere misure precise della massa dei polipeptidi o dei loro prodotti di frammentazione proteolitica. Le informazioni più rilevanti che si possono avere dalle analisi effettuate con le configurazioni più evolute dello spettrometro di massa (tandem MS) consentono infatti di descrivere degli stessi polipeptidi anche la struttura primaria e, quindi, di ottenerne la completa identificazione. Peraltro, l'elevata affidabilità delle identificazioni ottenibili utilizzando i dati forniti dall'analisi spettrometrica si basa proprio sulla altissima precisione con la quale si possono misurare le masse molecolari. La misura sperimentale della massa di un polipeptide comporta infatti un errore non superiore allo 0,001%. Per fare un esempio la massa relativa di un polipeptide di PM 35.000 si determina con un errore inferiore a 4 unità di massa atomica. Con questa precisione delle misure è quindi possibile rilevare con accuratezza anche modificazioni post-traduzionali delle proteine allo studio. Diminuendo la Mr delle molecole, la precisione delle misure è ancora più elevata. Per esempio, la misura della massa di un oligopeptide di Mr = 1000 dà un errore minore di 10 ppm¹⁴⁰.

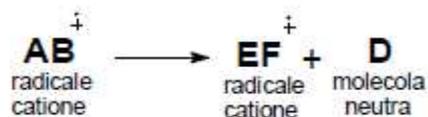
La spettrometria di massa e' basata sulla ionizzazione di una molecola, il catione radicalico che si forma (*ione molecolare*) in parte si frammenta dando molecole e/o radicali neutri (che lo strumento non rileva), in parte generando cationi e/o radicali cationi (*ioni frammento*). Lo ione molecolare e i vari ioni che si originano per frammentazione (cationi e radicali cationi) vengono discriminati sulla base del loro rapporto massa/carica e rivelati da un detector.



1. Processi di frammentazione:



2. Processi di riassetamento:



Molti di questi ioni sono molto instabili e si frammentano e riassetano in 10⁻¹⁰-10⁻⁸ secondi. Quando questi ioni posseggono una struttura che permette loro di sopravvivere per almeno 10⁻⁶ secondi e possono essere registrati.

La formazione di ioni di campione in fase gassosa è un pre-requisito essenziale per i processi di separazione e di rivelazione tipici in uno spettrometro di massa. Fino a non molto tempo fa gli spettrometri di massa richiedevano il campione in fase gassosa, ma grazie agli sviluppi più recenti, descritti più avanti, l'applicabilità della spettrometria di massa è stata estesa fino a includere anche campioni in fase liquida o inglobati in una matrice solida. Il campione, che può essere solido, liquido o gassoso, viene introdotto in una camera ad alto vuoto mediante un opportuno *sistema di introduzione*.

All'interno della *sorgente ionica* il campione organico, trasformato in forma gassosa viene bombardato da una corrente di elettroni emessa. Il fascio di elettroni viene fatto incidere perpendicolarmente alla direzione di propagazione del gas.

Gli ioni prodotti, che si trovano in fase gassosa, vengono separati nell'*analizzatore* sulla base del loro rapporto massa/carica (m/z). L'analizzatore è un tubo ad U (piegato a 180° che viene attraversato dagli ioni accelerati fino a giungere al collettore e quindi al rivelatore. Esso è posto tra i due poli di un campo magnetico orientato e variabile. Il campo magnetico sottopone la particella positiva ad una forza centripeta che va ad opporsi a quella centrifuga derivante dall'accelerazione.

Solo quando la forza centripeta bilancia perfettamente quella centrifuga la particella percorrerà una traiettoria circolare attraversando interamente il tubo ad U e giungendo al *rilevatore*. Nel rivelatore gli ioni generano un segnale elettrico proporzionale al numero di ioni presenti.

Figura 12: schema a blocchi di uno spettrometro di massa (Immagine tratta da "What is MASS SPECTROMETRY", ASMS (American Society for Mass Spectrometry)).

Il sistema di elaborazione dati registra questi segnali elettrici in funzione del rapporto m/z e li converte in uno *spettro di massa*, una intensità che dipenderà dalla abbondanza naturale di tali isotopi ¹⁴¹.

1.4.2 Spettro di massa

Lo spettro di massa è unico per ciascun composto e può essere utilizzato come "finger print" chimico per caratterizzare il campione o per identificarlo in miscele molto complesse.

La presentazione grafica è costituita da una serie di istogrammi inseriti in un diagramma su cui in ascissa sono riportati i valori m/z dei picchi relativi alle specie ioniche presenti ed in ordinata l'intensità o abbondanza in unità arbitrarie (Fig. 11). Il rapporto massa/carica (m/z) è il rapporto

tra la massa m dello ione, misurata rispetto alla massa del carbonio che per convenienza è 12,000000, ed il numero z di cariche elettrostatiche (misurate rispetto a quello dell'elettrone, cui viene assegnata una carica elettrostatica negativa unitaria) presenti sullo stesso ione.

In generale quando una molecola è ionizzata, per rimozione di un elettrone si origina uno ione molecolare il quale genera nello spettro di massa un picco denominato picco ionico molecolare, che compare a un valore m/z numericamente uguale al peso molecolare nominale dell'analita e costituisce il primo precursore di tutti gli ioni frammento nello spettro di massa. Questo ione molecolare, infatti, può contenere sufficiente energia interna dando luogo a delle frammentazioni successive che portano alla formazione di specie neutre, cationiche o radicaliche.



Lo spettrometro di massa sarà perciò in grado di separare tutti i frammenti carichi (A^+ e $A^{+\bullet}$), che, a loro volta, possono contenere energia interna sufficiente per generare successivi frammenti ionici anch'essi analizzabili attraverso analisi spettrometrica¹⁴².

In un'analisi spettrometrica con una sorgente ionica ESI spesso le molecole generano ioni multi carica e questo accade, ad esempio, nel processo di ionizzazione elettronica quando una molecola perde due elettroni: si ottiene una ione molecolare doppiamente carico che produce nello spettro di massa un picco ad un valore m/z numericamente uguale alla metà del peso molecolare del composto. Analogamente se una molecola acquista nH^+ durante il processo di ionizzazione electrospray si ottiene uno ione molecolare protonato che produce nello spettro di massa un picco ad un valore di m/z numericamente uguale ad $\frac{1}{n}$ del peso molecolare del composto aumentato di n unità $(M+n/n)$ ¹⁴³.

Il picco più intenso di uno spettro di massa è il picco base e viene utilizzato come base unitaria rispetto alla quale normalizzare le abbondanze relative di tutti gli altri ioni. Per definizione, quindi, l'intensità relativa del picco base è 100% mentre le intensità degli altri picchi sono espresse in % Relative Intensity (R. I.).

La capacità dello spettrometro di massa di separare due picchi adiacenti è detta risoluzione (R) ed è data dalla relazione:

$$R = M_n / (M_n - M_m)$$

dove M_n è il valore di massa superiore i due picchi adiacenti e M_m è il valore di massa inferiore¹⁴⁴.

La risoluzione di uno strumento è un parametro fondamentale per la misura delle masse esatte da cui si può derivare la composizione elementare del composto. Nei casi in cui la risoluzione dello

spettrometro di massa è sufficientemente elevata, infatti, ad ogni picco si accompagna uno o più picchi ad (m+1), (m+2), etc. Tali segnali sono dovuti alla presenza di isotopi stabili degli elementi che entrano nella composizione elementare dell'analita. L'intensità di questi picchi isotopici segue la distribuzione isotopica naturale degli elementi (Tab.6)

ABBONDANZA ISOTOPICA NATURALE						
<i>Elemento</i>	<i>X</i>	<i>Massa %</i>	<i>X+1</i>	<i>Massa %</i>	<i>X+2</i>	<i>Massa %</i>
Idrogeno	¹ H	99,99	² H	0,01		
Carbonio	¹² C	98,9	¹³ C	1,1		
Azoto	¹⁴ N	99,6	¹⁵ N	0,4		
Ossigeno	¹⁶ O	99,76	¹⁷ O	0,04	¹⁸ O	0,20
Fluoro	¹⁹ F	100				
Fosforo	³¹ P	100				
Zolfo	³² S	95,02	³³ S	0,76	³⁴ S	4,22
Cloro	³⁵ Cl	75,77			³⁷ Cl	24,23
Bromo	⁷⁹ Br	50,5			⁸¹ Br	49,5
Iodio	¹²⁷ I	100				

Tabella 6: abbondanza isotopica naturale di alcuni costituenti organici

Si potrebbe continuare con i picchi isotopici (m+3), (m+4), etc, il cui contributo, trascurabile per le molecole con pochi atomi di carbonio diventa significativo per molecole ad alto peso molecolare le quali presentano "cluster isotopici" complessi anche per il contributo additivo degli isotopici degli altri elementi ¹⁴⁵.

Le qualità che uno spettro di massa acquisito deve avere per poter essere utilizzato sono:

- un buon rapporto segnale/rumore di fondo
- l'intensità dei picchi non deve superare la soglia di acquisizione, poiché in questo caso risulterebbe difficile attribuire loro un valore esatto
- picchi ben risolti, ovvero alti e stretti

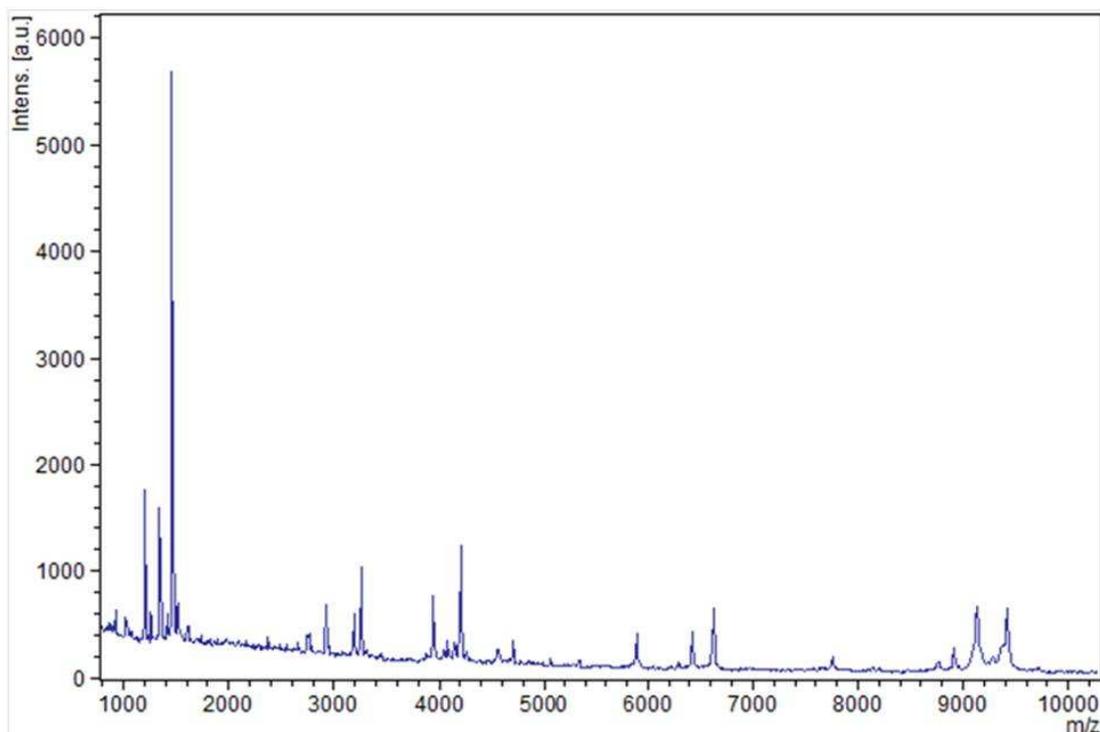


Figura 11: esempio di uno spettro di massa

1.4.3 Spettrometro di massa: elementi fondamentali

1.4.3.1 Sistema di introduzione

A causa della sua grande sensibilità lo spettrometro di massa è spesso usato come rilevatore in GC e in HPLC. L'accoppiamento delle tecniche cromatografiche con la spettrometria di massa ha consentito di superare il problema analitico della carenza di informazioni identificative proprie dei comuni rilevatori associati a GC e HPLC. I modi d'introdurre il campione all'interno dello strumento sono vari, normalmente dipendono dalla volatilità e dalla natura del campione, ed anche dal metodo di ionizzazione utilizzato.

Per miscele complesse si può ricorrere alla preventiva separazione mediante gascromatografia (GC) o cromatografia liquida a alte prestazioni (HPLC); in questi casi il sistema d'introduzione è una vera interfaccia tra le due strumentazioni la cui funzione è quella di rendere compatibile l'abbinamento di sistema cromatografici, che funzionano sotto pressione, con lo spettrometro di massa, che funziona sotto vuoto spinto.

Due sono le tipiche interfacce per collegamento GC – MS: il separatore a getto (o di Rihaghe) e l'interfaccia open-split, ma nel GC-MS si può eliminare l'impiego dell'interfaccia quando si utilizzano colonne capillari. Mentre tipica interfaccia per il collegamento HPLC-MS sono la "particle beam", la "thermospray" e la "electrospray". La termospray e l'electrospray

rappresentano però anche due specifici metodi di ionizzazione a pressione atmosferica (API = atmospheric pressure ionization) ¹²⁷.

1.4.3.2 Sorgente ionica

La sorgente ionica è una parte dello spettrometro di massa nella quale le molecole gassose, liquide o solide vengono trasformate in ioni desolvatati in fase gassosa.

La sorgente ionica è costituita da camera di ionizzazione all'interno del quale avvengono anche i processi di frammentazione che portano alla formazione di un certo numero di specie ioniche, con definiti valori di m/z , i cui segnali daranno luogo allo spettro di massa.

il fascio di ioni che viene così prodotto viene focalizzato e accelerato attraverso una serie di elettrodi che sono presenti in sorgente.

Generalmente, prima di entrare nell'analizzatore, il fascio ionico viene in parte intercettato da un elettrodo che genera una corrente chiamata "corrente ionica totale" o TIC (total ion current). Questa corrente è proporzionale alla quantità di ioni raccolti, i quali sono a loro volta proporzionali alla quantità di sostanze che si trova in quel momento nella sorgente e soggetta al processo di ionizzazione. Nel caso di scansioni continue, la TIC può essere anche generata dal sistema di acquisizione ed

elaborazione dati sommando le intensità di tutti i picchi registrati negli spettri di massa. Nel caso di analisi GC-MS ed HPLC-MS il profilo della TIC è solitamente sovrapponibile al cromatogramma ottenuto con i rilevatori cromatografici classici.

Esistono diversi tipi di sorgenti ioniche, queste vengono scelte in funzione delle caratteristiche chimico-fisiche dell'analita e che perciò sfruttano principi fisici diversi per generare ioni gassose da molecole neutre ¹²⁶.

La **sorgente a ionizzazione elettronica** (EI), detta anche ad impatto elettronico, è il metodo più diffuso e per lo più viene usato per specie volatili e termicamente stabili.

Le molecole neutre vengono introdotte già in fase gassosa e grazie ad fascio di elettroni ad alta energia emessi da un filamento, avviene la ionizzazione. Questo fascio di elettroni di bombardamento si muove su un percorso a spirale per effetto del campo magnetico locale generato da due piccoli magneti, aumentando così la probabilità di interazione degli elettroni con le molecole neutre. L'elettrone in movimento può essere considerato come un pacchetto di energia presentando un movimento ondulatorio analogo a quello di un fotone; durante l'approccio elettrone/molecola il campo elettrico dell'onda e il campo magnetico della molecola si modificano mutuamente. In seguito al processo di ionizzazione si innescano una serie di processi di frammentazioni unimolecolari.

L'energia degli elettroni di bombardamento è molto più grande dell'energia di un normale legame chimico ed inoltre l'interazione elettronica genera uno ione molecolare radicalico in uno stato fortemente eccitato, per cui esso può, a volte, decomporsi prima di essere rivelato per ciò viene considerato un metodo di ionizzazione hard.

La **ionizzazione chimica** (CI) è concettualmente simile a quella EI, ma, a differenza di quest'ultima, la pressione nella camera di ionizzazione della sorgente viene mantenuta a un valore compreso tra 0,1 Torr e 2,0 Torr per introduzione di un gas reagente che può essere metano, isobutano, ammoniaca oppure idrogeno che portano rispettivamente alla formazione di ioni secondari del tipo CH_5^+ , C_4H_9^+ , NH_4^+ e H_3^+ . Tali cationi sono acidi forti che possono quindi protonare con una reazione acido-base di Brønsted praticamente qualsiasi molecola organica ma se la molecola è sprovvista di un sito basico, piuttosto che protonarsi questa specie tenderà a perdere un protone e questo ione pseudo-molecolare presenta una scarsa tendenza a frammentare.

La sorgente **Atmospheric Pressure Chemical Ionization** (APCI) è formata da un nebulizzatore in ceramica riscaldato contenente un capillare nel quale l'eluato della colonna viene spinto per essere vaporizzato; la temperatura (da 200 a 500 °C) e una corrente di azoto permettono l'evaporazione di solvente e analiti.

Un potenziale elettrico, positivo o negativo ionizza le molecole: l'atmosfera nella sorgente, composta principalmente dai vapori del solvente oltre che da azoto, viene ionizzata da un ago che produce una "corona discharge" (Fig 12).

Figura 12: schema di funzionamento della sorgente APCI in modalità positiva: la scarica elettrica dell'ago ionizza le molecole gassose del solvente. Queste trasferiscono poi la carica acquisita agli analiti (S solvente, M analita). (Immagine tratta dal sito The University of Bristol, School of Chemistry)

I vapori di solvente e azoto vengono istantaneamente ionizzati dall'ago e reagiscono chimicamente con gli analiti in fase gassosa, trasferendo loro la carica elettrica. Per lavorare con una sorgente APCI occorre che l'analita sia sufficientemente volatile e termicamente stabile e che la fase mobile possa accettare e donare protoni quando essa si trovi in fase gassosa. Dall'analisi condotta in APCI, di norma, si forma esclusivamente lo ione pseudomolecolare $[\text{M}+\text{H}]^+$ e talvolta si possono osservare perdite di molecole di H_2O ma nessuna frammentazione; di conseguenza, la corrente ionica totale dell'analita, di norma, è associata ad un solo ione. La

sorgente APCI è generalmente preferita per l'analisi di molecole poco polari e con PM fino a 4000 D.

La spettrometria di massa APCI consente l'approccio a una notevole serie di applicazioni nei campi più svariati e apre effettivamente nuove prospettive alla ionizzazione chimica, offrendo una efficienza

di ionizzazione assai più elevata della ionizzazione chimica tradizionale. La APCI è adatta per composti mediamente polari, ma può fornire buoni risultati con composti molto polari, specialmente quando la fase mobile preveda l'impiego di tamponi, sia volatili che non volatili. La APCI può essere direttamente interfacciata a colonne analitiche HPLC standard con diametro interno di 4,6 mm, dato che essa è in grado di accettare flussi di 1-2 ml/min.

per supplire alla scarsa frammentazione osservata in condizioni normali di ionizzazione è possibile eseguire un'analisi MS-MS che permette di incrementare ulteriormente la sensibilità con il segnale in assoluto meno intenso. Occorre tenere presente che in questo modo si genera la frammentazione di tutto quello che si trova in sorgente in quel momento quindi è necessario un efficiente stadio di separazione per essere sicuri di ionizzare un componente puro. Il segnale di corrente ionica ottenuto in queste condizioni sarà più elevato, ma il rumore chimico non viene eliminato come nella MS-MS^{125,126}.

1.4.3.3 Analizzatore di ioni

L'analizzatore consente di differenziare gli ioni generati in base al loro rapporto massa/carica (m/z).

Gli analizzatori possono essere suddivisi in due gruppi fondamentali:

1. gli analizzatori della quantità di moto come gli analizzatori magnetici, che separano gli ioni in funzione del prodotto massa per velocità;
2. gli analizzatori dinamici, come gli analizzatori a tempo di volo (TOF) o gli analizzatori a trappola ionica (IT) oppure gli analizzatore a risonanza ionica ciclotronica (FT-ICR).

Gli analizzatori più comunemente usati sono i settori magnetici ed elettrici, i quadrupoli, le trappole ioniche (ion traps), gli spettrometri basati sulla risonanza ionica ciclotronica in trasformata di Fourier (FT-ICR) e gli analizzatori a tempo di volo (Time of Flight, TOF)¹³⁹.

Gli analizzatori a *settori magnetici* deviano le traiettorie degli ioni in percorsi circolari il cui raggio dipende dal rapporto momento/carica dello ione. Ioni con m/z più elevato seguono traiettorie di raggio più ampio rispetto a quelli con m/z più basso, così che ioni con differenti

valori di m/z vengono dispersi nello spazio. Cambiando le traiettorie degli ioni mediante variazioni del campo magnetico, ioni di diverso rapporto m/z possono essere focalizzati sul rivelatore. Gli spettrometri di massa a doppia focalizzazione utilizzano una combinazione di campi magnetici ed elettrici per separare gli ioni. Il campo elettrico accelera gli ioni inviandoli in un condotto ricurvo, dove il campo magnetico induce una deflessione del percorso degli ioni accelerati in base al loro rapporto m/z . La fenditura agisce da filtro per selezionare uno specifico rapporto m/z .

Figura 13: schema di funzionamento di un analizzatore a settore magnetico (Immagine tratta dal sito <http://www.mabra.it/dida/massa/fm.htm>)

L'equazione mostra la dipendenza della traiettoria di uno ione, con un determinato rapporto m/z , dal valore del campo magnetico H , da quello elettrico V e dal raggio di curvatura r .

$$\frac{m}{z} = \frac{H^2 r^2}{2V}$$

La scansione temporale del campo magnetico/elettrico consentirà a tutti gli ioni di giungere al rivelatore ¹⁴⁶.

L'*analizzatore quadrupolare* (Fig. 14) è costituito da quattro barre di metallo disposte parallelamente a cui viene applicata una differenza di potenziale, generata da una corrente continua alternata. Gli ioni a causa di questa differenza, subiranno nel loro transito delle oscillazioni che potranno essere stabili, permettendo così agli ioni di uscire dal quadrupolo, oppure instabili inducendo così la collisione dello ione con le barre del quadrupolo. A determinati valori della tensione applicata, solo ioni aventi un certo rapporto m/z usciranno dal quadrupolo e saranno convogliati al rivelatore. Variando nel tempo la tensione applicata, tutti gli ioni saranno messi in condizione di percorrere l'analizzatore ¹⁴⁷.

Figura 14: schema di funzionamento di un analizzatore quadrupolare (Immagine tratta dal sito <http://www.dia.unisa.it/~ads/BIOINFORMATICA/Proteomica/glossario.html>)

La *trappola ionica* opera con un principio simile a quello del quadrupolo, ma anziché permettere agli ioni di attraversare il campo quadrupolare la trappola ionica può trattenere tutti gli ioni al

suo interno che vengono rilasciati progressivamente verso il rivelatore, facendo variare il campo elettrico.

E' considerato un sistema a tre elettrodi in quanto vi sono due elettrodi uno sopra e uno sotto chiamati caps ed end, mentre gli altri due sono rappresentati da un elettrodo ad anello posto nel centro.

Il campo elettrico che si genera intrappola gli ioni all'interno della cavità: gli ioni che possiedono una traiettoria stabile vengono momentaneamente trattenuti nella trappola, mentre gli ioni che presentano una traiettoria instabile saranno espulsi (*Fig 15*)¹²⁸.

Figura 15: schema di funzionamento di uno spettrometro a trappola lineare (Immagine tratta dal sito Thermo Scientific)

Altri due tipi di analizzatori, sempre più diffusi nei laboratori di ricerca, sono lo spettrometro di massa a risonanza ionica ciclotronica in trasformata di Fourier (*Fourier Transform Ion Cyclotron Resonance*, FT-ICR) e lo spettrometro di massa a tempo di volo (*Time Of Flight*, TOF). Le caratteristiche particolari di questi due analizzatori li rendono specialmente utili per lo sviluppo della spettrometria di massa in nuovi campi di applicazione.

In uno spettrometro FT-ICR (*Fig. 16*) gli ioni vengono intrappolati elettrostaticamente in una cella cubica posta all'interno di un campo magnetico costante. L'applicazione di un impulso a radiofrequenza tra due piastre della cella che fungono da trasmettitore (piastre di eccitazione) induce un moto orbitale coerente. Gli ioni in moto orbitale generano un debole segnale "immagine" che viene raccolto dalle piastre della cella che agiscono come ricevitore (piastre di rivelazione). La frequenza del segnale generato da un particolare ione è uguale alla sua frequenza orbitale che a sua volta è inversamente correlata al suo rapporto m/z . L'intensità del segnale per una data frequenza è proporzionale al numero di ioni aventi il corrispondente valore di m/z . Il segnale totale viene amplificato e quindi scomposto nelle varie componenti di frequenza generando così lo spettro di massa. Se la pressione all'interno della cella è molto bassa, ossia se viene mantenuto un vuoto molto spinto, il moto orbitale degli ioni può essere mantenuto per tempi relativamente lunghi e la frequenza misurata ripetutamente numerose volte per ottenere una accuratezza molto elevata. Gli strumenti FT-ICR sono in grado quindi di produrre spettri di massa con una altissima risoluzione¹⁴⁸.

Figura 16: schema di funzionamento di uno spettrometro di massa FT-ICR. (Immagine tratta dal sito [The University of Bristol, School of Chemistry](#))

L' analizzatore a tempo di volo (TOF, Time of Flight) è uno strumento abbastanza semplice ed è generalmente combinato con sorgenti MALDI. E' costituito da un tubo di deriva collegato da una parte alla sorgente e dall'altra al rivelatore. Gli ioni prodotti nella sorgente vengono accelerati con la stessa energia potenziale (V) per cui la loro energia cinetica ($1/2mv^2$) è la stessa. La velocità, v, di uno ione dipende dalla seguente formula:

$$v = \sqrt{\frac{2zeV}{m}}$$

la velocità dello ione dipende dalla massa, così che se un gruppo di ioni è accelerato, ed attraversa una regione dello spettrometro priva di campi, gli ioni arriveranno al detector con tempi differenziati, in dipendenza delle relative velocità, questo significa che ioni con massa più alta sono caratterizzati da una velocità più bassa rispetto a quelli a massa minore quindi al rivelatore arriveranno prima gli ioni più veloci ad m/z più bassi e poi gli ioni più lenti ad m/z più alti ¹⁴⁹.

1.4.3.4 Rilevatore

Il rilevatore è un sistema che permette di convertire e registrare l'informazione portata da sorgenti ed analizzatori. Tutti i rilevatori possiedono una superficie in grado di raccogliere gli ioni e neutralizzare la carica attraverso meccanismi di raccolta o donazione di elettroni. Si realizza quindi un trasferimento di elettroni ed un flusso di corrente che può essere amplificato ed infine convertito in un segnale processabile da un computer. La maggioranza dei rivelatori funziona ad impatto ionico o per cattura ionica.

Gli ioni prodotti nella sorgente e separati dall'analizzatore, in funzione del rapporto m/z, raggiungono il rivelatore con correnti ioniche che vanno da 10^{-9} A a 10^{-17} A. Il segnale che verrà generato sarà direttamente proporzionale al numero di ioni che raggiungono il detector e verrà trasformato in uno spettro di massa.

Quattro sono le tecniche che vengono utilizzate per la rivelazione degli ioni: il moltiplicatore elettrico, il Daly detector, il Faraday-cup e i vari focal plane detector.

Il detector più comune è il moltiplicatore elettrico che è costituito da una serie di dieci o venti dinodi (elettrodi) composti da una lega rame-berillio. Quando gli ioni impattano sul primo dinodo di conversione si ha l'emissione di un certo numero di elettroni per ione. Gli elettrodi emessi dal primo dinodo colpiranno il secondo producendo un effetto a cascata. L'ultimo dinodo è collegato con un pre-amplificatore che converte la corrente in voltaggio in modo da amplificare e registrare il segnale.

Una variante del moltiplicatore elettrico è rappresentato dal CEM (Channel Electron Multiplier) cioè un moltiplicatore elettrico a tubo costituito da un dinodo continuo dove gli elettroni collidono continuamente contro le pareti interne del moltiplicatore producendo il tipico effetto a cascata.

L'altro tipo di detector è detto Daly ed è un rivelatore elettro-ottico che utilizza un plate su cui impatta un fascio di ioni, provocando, anche in questo caso, un rilascio di elettroni. Gli elettroni così prodotti impattano su di uno schermo al fosforo che emette fotoni che vengono poi trasmessi ad una schiera di fotodiodi (photodiode array) su cui sono collocati composti dei sensori

Il Faraday cup è un detector costituito da un plate, ma più spesso una tazza (cup), di metallo, su cui impatta un fascio di ioni, connesso a terra per mezzo di un resistore. Su tale plate avviene la neutralizzazione della carica degli ioni quando questi lo colpiscono generando, così un flusso di corrente attraverso il resistore.

Il detector focal plane è caratterizzato da una lastra fotografica, quando gli ioni impattano sulla lastra, le bande che ne risultano sono rivelate mediante uno sviluppo fotografico e successivamente convertite nei rispettivi segnali elettrici mediante l'utilizzo di micro densitometri ¹⁵⁰.

1.4.4 Tecniche di desorbimento.

Le tecniche di desorbimento hanno portato una rivoluzione nell'applicazione della spettrometria di massa all'analisi di superfici e a molecole grandi, non volatili o termicamente poco stabili. Con il termine ionizzazione per desorbimento si indica il processo per cui una molecola viene sia evaporata da una superficie, sia ionizzata: le principali tecniche sono elencate in tabella 7 ¹³⁹.

<i>Tecnica di ionizzazione</i>	<i>Mezzo ionizzante</i>
Bombardamento con atomi veloci (<i>Fast Atom Bombardment</i> , FAB)	Impatto di atomi a velocità elevata su un campione disciolto in una matrice liquida
Spettrometria di massa di ioni secondari (<i>Secondary Ion Mass Spectrometry</i> , SIMS)	Impatto di ioni a velocità elevata su un film sottile di campione depositato su un substrato metallico (o disciolto in una matrice liquida, <i>Liquid SIMS</i>)
Desorbimento in plasma (<i>Plasma Desorption</i> , PD)	Impatto di frammenti da fissione nucleare, ad esempio dal ²⁵² Cf, su un campione solido depositato su una lamina metallica
Ionizzazione laser assistita da matrice (<i>Matrix Assisted Laser Desorption Ionization</i> , MALDI)	Impatto di fotoni ad alta energia su un campione inglobato in una matrice organica solida

Desorbimento di campo (<i>Field Desorption</i> , FD)	Imposizione di un forte campo elettrico su un campione depositato su una speciale sonda metallica
Ionizzazione per elettro-nebulizzazione (<i>Electrospray Ionization</i> , ESI)	Formazione di particelle liquide cariche dalle quali gli ioni vengono emessi per desorbimento e/o desolvatazione

Tabella 7:tecniche di ionizzazione per desorbimento

Per le prime quattro tecniche elencate nella tabella, il campione viene desorbito e ionizzato mediante un processo che sfrutta il bombardamento della superficie del campione con atomi ad alta energia, frammenti da fissione nucleare, ioni o fotoni. L'impatto deposita energia sul campione, direttamente o tramite la matrice, e porta al trasferimento delle molecole campione in fase gassosa e alla loro ionizzazione.

Nel desorbimento di campo il campione viene depositato in forma di sottile pellicola su uno speciale filamento che viene posto in un campo elettrico di potenziale molto elevato. In questa situazione gli ioni, formati in seguito all'estrazione di un elettrone dalla molecola indotta dal forte campo elettrico, vengono facilmente e rapidamente estratti nello spettrometro di massa. Nella ionizzazione per elettro-nebulizzazione gli ioni si formano da soluzioni che vengono finemente nebulizzate attraverso un ago tenuto ad alto potenziale elettrico. Gli ioni dell'analita nelle goccioline cariche vengono volatilizzati in seguito alla rimozione del solvente per evaporazione. Questo avviene a pressione atmosferica entro un campo elettrico generato tra l'ago e una lente di estrazione degli ioni.

Questi metodi di ionizzazione possono produrre diversi tipi di ioni, a seconda dei parametri utilizzati e del tipo di campione. La formazione di ioni molecolari positivi $M^{+\bullet}$ e pseudo-molecolari $[M+H]^+$ è accompagnata dalla formazione di addotti con metalli alcalini, come $[M+Na]^+$ e, particolarmente nel caso della ionizzazione per elettro-nebulizzazione, da quella di ioni multiprotonati (e quindi multicarica) come $[M+nH]^{n+}$. L'estensione della dissociazione delle specie molecolari (frammentazione) dipende a sua volta dal metodo di ionizzazione^{151,152}.

1.4.4.1 Elettrospray ionization

L'*elettrospray ionization*, ESI, è un semplice metodo per analizzare piccole e grandi molecole ed opera a pressione atmosferica ed a temperature moderate. Particolarmente adatta a molecole non volatili e termicamente poco stabili è una delle metodologie principali per ionizzare peptidi e proteine. Una soluzione contenente l'analita viene fatta fluire attraverso un capillare alla fine del quale, per la contemporanea azione di un getto di aria compressa e dell'applicazione di un elevato voltaggio (5000/6000 V per l'analisi in positivo oppure 5000-6000 V nel caso di analisi

in negativo), la soluzione iniettata viene nebulizzata. Se il potenziale è positivo, gli ioni positivi si accumulano sulla superficie delle gocce, le quali seguiranno il percorso stabilito dal “cono di Taylor” (Fig 17).

Figura 17: schema di una sorgente di ionizzazione ESI (Tratta dal sito di Bruker Daltonik GmbH)

Le gocce attraversano un gradiente di potenziale e pressione tale da indurre una riduzione del diametro per desolvatazione sino al fenomeno d’esplosione di Coulomb, quando praticamente il diametro delle gocce è talmente piccolo (limite di Rayleigh), che la forza repulsiva tra le cariche, vicine tra loro, supera la tensione superficiale delle gocce.

L’equazione che lega la carica della goccia Q con il suo raggio R è:

$$Q^2 = 64\pi^2\epsilon_0\gamma R^3$$

dove:

Q = carica della goccia

γ = tensione superficiale del liquido

ϵ^0 = permittività nel vuoto

R = raggio della goccia

Il passaggio che avviene in seguito, ovvero il passaggio dalla fase liquida a quella gassosa mediante una “esplosione” delle gocce cariche rimane ancora controversa e per la sua spiegazione sono state proposte due differenti teorie o modelli: il modello della carica residua o modello di Dole ed il modello dell’evaporazione degli ioni o modello di Iribarne e Thomson¹⁵³.

Secondo il modello di Dole la goccia madre, in seguito alle esplosioni coulombiane, genera gocce figlie che a loro volta vanno incontro allo stesso processo fino a quando non si arriva ad avere una singola molecola carica per goccia.

Secondo il modello di Iribarne e Thomson invece, il passaggio degli ioni in fase gassosa avviene a partire dalle gocce figlie che hanno una densità di carica Q sufficientemente elevata, che corrisponde ad un determinato raggio R :

$$Q = 10^{-17} \text{ C}$$

$$R < 10 \text{ nm}$$

Gli ioni vengono emessi dalle gocce figlie in seguito all’evaporazione del solvente; durante il processo la densità di carica rimane costante e questo permette al fenomeno di ripetersi. Attualmente nessuno dei due modelli può essere considerato non corretto, anzi la teoria predominante è che il processo di ionizzazione electrospray avvenga secondo il modello di Dole

per molecole di grandi dimensioni e secondo il modello di Iridarne e Thomson per le molecole di piccole dimensioni .il campione, in questo modo, viene ionizzato mediante l'esplosione coulombica delle goccioline di nebulizzato e quindi posto nelle condizioni ottimali per essere analizzato.

Gli ioni sono accelerati dal campo elettrico prodotto dal voltaggio di circa 5000V applicato. Questi ioni accelerati entrano nello strumento attraverso un orifizio e transitano in una camera nella quale fluisce N₂ (gas tampone). Dall'altro lato di questa camera si trova il micro orifizio che dà accesso alla parte dello strumento sotto vuoto spinto. Tra i due orifizi viene applicato un voltaggio di 40-150 mV, maggiore è il voltaggio, maggiore sarà sia l'accelerazione subita dagli ioni del campione sia la forza cinetica delle collisioni fra questi e le molecole di gas tampone quindi migliore sarà lo stripping di contro-ioni e molecole di solvente da ioni di campioni transitanti. Se la collisione è eccessivamente energetica può causare frammentazione. Nella camera ad alto vuoto gli ioni vengono separati in base al loro rapporto m/z. A voltaggi OR bassi (40-60 mV) si osservano preferenzialmente ioni ad alta carica, quindi bassi valori di m/z, in quanto gli ioni molto carichi risentono maggiormente della scarsa accelerazione fornita dall'OR, ma sono meno puliti perché lo stripping di solvente, sali, etc. è meno efficiente. A voltaggi più alti (maggiori di 90 mV) aumenta il segnale dovuto agli ioni con bassa carica e quindi alti valori di massa / carica, in quanto il voltaggio OR è sufficientemente alto per attrarli all'orifizio interno, il segnale risulta essere anche più pulito (stripping più efficiente) ma gli ioni più carichi potrebbero risentire di frammentazione. Gli ioni che vengono prodotti in fase gassosa possiedono diverse importanti proprietà: lo stato di carica degli ioni in fase gassosa riflette lo stato di carica della fase condensata, anche se qualche volta può essere modificato in seguito a collisione; il trasferimento di ioni in fase gassosa non è un processo energetico, infatti gli ioni si "raffreddano" nello step di desolvatazione; coinvolge la scissione di legami non covalenti comportando la perdita delle molecole di solvatazione perciò è considerata una tecnica non spinta energeticamente e non distruttiva nei confronti del campione. Inoltre una delle principali caratteristiche della ionizzazione mediante ESI è la formazione degli ioni multi carica (Fig.18).

Figura 18: schema di funzionamento della sorgente ESI (Immagine tratta dal sito di Thermo Scientific)

In presenza di basse concentrazioni di campione diventa cruciale una diminuzione del flusso per aumentare la sensibilità dell'analisi.

Per rispondere a questa esigenza esistono due tipologie di sorgenti *electrospray*: la sorgente micro- *electrospray* e la sorgente nano- *electrospray*¹⁵⁴. Di particolare interesse è la sorgente nano- *electrospray* mediante la quale è possibile analizzare un campione utilizzando un flusso

all'interno della sorgente di 50-500 nl/min. In questo tipo di sorgente il capillare metallico della sorgente *electrospray* è sostituito da un capillare in silice fusa avente un diametro interno di 10-15 µm; il capillare è rivestito esternamente da un sottile strato metallico in modo che possa condurre corrente. La peculiarità di questo tipo di sorgente è rappresentata dalle ridotte dimensioni delle gocce di analita che si vengono a formare. Tali gocce hanno una densità di carica superiore ed un rapporto superficie/volume maggiore, i quali permettono un desorbimento degli ioni più veloce se confrontato con la sorgente *electrospray*. Un'altra caratteristica della sorgente, anch'essa dovuta alle ridotte dimensioni delle gocce, è che non necessita dell'utilizzo dei gas di solvatazione, inoltre la distanza fra il capillare ed il controlettrodo è minore e variabile e la differenza di potenziale applicata è minore, circa 1000 V, in modo da evitare scariche a corona all'interno della sorgente ¹⁵³.

1.4.4.2 MALDI- TOF

La tecnica di ionizzazione in seguito a desorbimento laser assistito da matrice MALDI è stata introdotta alla fine degli anni 80; il gruppo di Karas e Hillenkamp ed il gruppo di Tanaka, hanno proposto nello stesso anno l'aggiunta al campione di un largo eccesso di matrice (generalmente in rapporto 100:1 o 10.000:1) in grado di assorbire la lunghezza d'onda del laser preservando l'analita ed aiutandolo nel processo di ionizzazione. In seguito sono stati ottenuti notevoli miglioramenti soprattutto dovuti all'osservazione ed agli esperimenti empirici piuttosto che alle conoscenze chimico-fisiche riguardo a questa tipologia di ionizzazione. Infatti il reale meccanismo della ionizzazione MALDI è ancora oggi un importante argomento di ricerca e discussione ¹⁵⁵.

La spettrometria di massa MALDI-TOF viene spesso utilizzata nelle fasi iniziali dei progetti di proteomica durante le quali viene eseguito uno screening per l'identificazione di un gran numero di candidati proteici per successive e più approfondite analisi; in questa fase risultano essere molto importanti la produttività ed il numero di proteine identificate positivamente. Per questa prima fase si ricorre in genere al mappaggio dei peptidi ottenuti in seguito a digestione enzimatica mediante spettrometria di massa MALDI-TOF poiché, oltre ai vantaggi precedentemente evidenziati, questi strumenti possiedono un'elevata risoluzione, cruciale per questo tipo di approccio, ed inoltre offrono la possibilità di automatizzare l'analisi fino all'identificazione della proteina. Le due caratteristiche che rendono automatizzabile le analisi per l'identificazione delle proteine mediante questa tipologia di strumenti sono il loro

relativamente facile utilizzo, e la facile interpretazione degli spettri ottenuti in quanto con la tecnica di ionizzazione MALDI si ottengono quasi esclusivamente ioni monocarica $[M+H]^+$ (Tab 8) ¹⁵⁶.

Tipologia di Campione	Ioni Positivi	Ioni Negativi
Bassa Polarità	$M^+ // [M+H]^+$	$M^- // [M-H]^-$
Media Polarità	$M^+ [M+H]^+ // [M+Cat]^+$ $[2M+H]^+$ MW > 3000: $[M+nH]^{n+}$, n = 2, 3	$M^- [M-H]^- // [2M-H]^-$ MW > 3000: $[M-nH]^{n-}$, n = 2, 3
Alta Polarità	$[M+H]^+ [M+Cat]^+ // [2M+H]^+$ $[2M+Cat]^+$ MW > 3000: $[M+nH]^{n+}$, n = 2, 3	$[M-H]^- // M^- [2M-H]^-$ MW > 3000: $[M-nH]^{n-}$, n = 2, 3
Ionico K^+A^-	$K^+ // [K_n+A_{n-1}]^+ [KA]^{n+}$	$A^- // [A_n+K_{n-1}]^- [KA]^{n-}$
Cat: Cationizzazione mediante Li^+ , Na^+ , K^+ ed altri ioni metallici An: Anionizzazione mediante Cl^- , Br^- , I^- , HSO_4^- ed altri anioni Addotti con la matrice si possono formare con tutti i tipi di analiti, es. $[M+H+matrice]^+$.		

Prima di essere analizzato, il campione deve essere depositato, su di un target metallico, insieme ad una apposita matrice; l'evaporazione del solvente in cui sono disciolti campione e matrice, provoca la co-cristallizzazione del campione con la matrice stessa, che è generalmente costituita da un acido organico debole. Esistono diversi tipi di matrice ognuna delle quali viene scelta in base al tipo di analisi da effettuare, quindi in base al campione in analisi ed al tipo di laser che si utilizza per la ionizzazione. La matrice ha il compito di assorbire la lunghezza d'onda del laser, proteggere l'analita, evitare l'aggregazione delle molecole di analita per ottenere ioni molecolari e desorbire l'analita in maniera indipendente dalle sue caratteristiche intrinseche.

Esistono diverse matrici, ognuna adatta ad una particolare esigenza analitica (Tab. 9). Ad esempio l'acido α -Ciano-4-Idrossicinammico (HCCA) è adatto per analisi di peptidi o proteine con massa inferiore a 10 KDa e quindi per la maggior parte degli studi di proteomica, mentre l'acido Sinapinico (SA) è utilizzato per rilevare proteine intere o peptidi con massa superiore ai 10 KDa (Fig.19) ¹⁵⁷.

MATRICE	APPLICAZIONI
Acido α -ciano-4-idrossicinnamico (CHCA)	Peptidi (<10kDa), lipidi, carboidrati
Acido Sinapinico o acido Trans-3,5-dimetossi-4-idrossi cinnamico	Peptidi e proteine di grosse dimensioni (10-150 kDa), glicoproteine, proteine di membrana
Acido 2,5-diidrossi benzoico (DHB)	Peptidi, proteine, carboidrati, glicoproteine, glicolipidi, polimeri, lipidi, molecole organiche
Acido 3-idrossipicolinico (HPA)	Oligonucleotidi > 3.5 kDa
2,4,6-triidrossiacetofenone (THAP)	Oligonucleotidi < 3.5 kDa
Acido picolinico (PA)	Oligonucleotidi
Miscele di PA e HPA	Oligonucleotidi > 10 kDa
6-azo2-tiotimina (ATT)	DNA

Tabella 8: matrici più utilizzate per le analisi MALDI-TOF e loro applicazioni

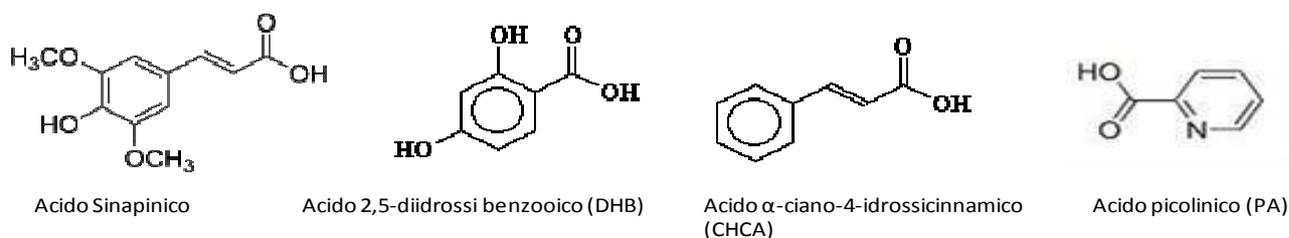


Figura 19: formula di struttura di alcune fra le matrici più utilizzate per le analisi MALDI-TOF

Una volta selezionata la matrice opportuna, si procede con la scelta del metodo di cristallizzazione da adottare in quanto anch'esso varia a seconda dell'analita e della tipologia di analisi da eseguire e può influenzare la riuscita dell'analisi. Esistono alcune indicazioni riguardo al metodo da adottare, ma in generale è sempre l'esperienza dell'operatore a determinare il tipo di deposito da utilizzare.

I due metodi maggiormente utilizzati per le analisi di peptidi preparati in soluzione sono il *double layer* ed il *dried droplet*. Il metodo *double layer* consiste nel formare un primo strato di matrice sul quale viene in seguito depositato il campione ed infine depositare sopra al campione un ulteriore strato di matrice. Questo tipo di deposito presenta il vantaggio di fornire cristalli di piccole dimensioni ed omogenei, per questo motivo viene spesso utilizzato per analisi che richiedono una grande precisione in quanto fornisce misure molto riproducibili da uno sparo di laser all'altro. Presenta tuttavia il difetto di una preparazione laboriosa.

Il metodo *dried droplet* invece prevede che la matrice venga miscelata con l'analita ed in seguito questa miscela venga depositata sul *target*. Questo metodo genera cristalli di dimensioni maggiori ed eterogenei. Lo svantaggio di questo tipo di deposito risiede proprio nell'eterogeneità

dei cristalli e nella loro maggiore dimensione in quanto queste caratteristiche influenzano la ionizzazione dell'analita. Maggiori sono le dimensioni dei cristalli, maggiore sarà la loro resistenza agli impulsi laser, per questo come risultato si otterranno misure lievemente meno precise. Il vantaggio nell'utilizzo di questo metodo di deposito risiede nella sua facilità di esecuzione e quindi nella possibilità di eseguirlo in maniera automatizzata.

Una volta eseguito il deposito è possibile inoltre, quando necessario, eseguire una desalificazione del campione direttamente sul *target* mediante l'aggiunta di acqua acidificata sul deposito. L'acqua viene in seguito allontanata, insieme ai sali, ed il campione viene fatto ricristallizzare per aggiunta di matrice.

Un altro parametro importante per il desorbimento e la ionizzazione dell'analita è la lunghezza d'onda del laser (*Tab. 10*); la sua scelta dipende soprattutto dalla matrice utilizzata e dalla natura dell'analita¹⁵⁸.

LASER	LUNGHEZZA D'ONDA	ENERGIA (kcal/mol)	ENERGIA (eV)	DURATA DELL'IMPULSO
Azoto	337 nm	85	3.68	< 1 ns- pochi ns
Nd: YAG μ 3	355 nm	80	3.49	5 ns
Nd: YAG μ 4	266 nm	107	4.66	5 ns
XeCl	308 nm	93	4.02	25 ns
KrF	248 nm	115	5.00	25 ns
ArF	193 nm	148	6.42	15 ns
Er:YAG	2.94 μ m	9.7	0.42	85 ns
CO ₂	10.6 μ m	2.7	0.12	100 ns + 1 μ s coda

Tabella 9: laser maggiormente utilizzati per le analisi MALDI-TOF e relative caratteristiche

Una volta effettuato il deposito, il *target* viene inserito all'interno dello strumento mediante un sistema di introduzione del campione che permette il mantenimento del vuoto sia all'interno della sorgente di ionizzazione che del tubo di volo, impedendo così l'ingresso nello spettrometro a molecole di ossigeno, azoto ed altri gas presenti nell'aria. La presenza di queste molecole, infatti, provocherebbe collisioni con le molecole dell'analita con una conseguente variazione del loro tempo di volo. Inoltre determinerebbe la frammentazione delle stesse e la possibilità di reazioni chimiche indesiderate, come ad esempio ossidazioni.

Il processo di desorbimento/ionizzazione, può essere scomposto in diverse fasi: la prima fase è costituita dall'assorbimento dell'energia del fascio di fotoni della luce laser da parte della matrice. Solitamente il laser emette ad una lunghezza d'onda nel campo dell'IR o dell'UV; in seguito all'assorbimento di energia, la matrice ionizza e subisce un cambiamento di fase

passando dallo stato solido a quello di gas supercompressso.

Durante il cambiamento di fase la matrice cede la carica ad alcune molecole di analita mediante il trasferimento di un protone. In seguito al cambiamento di fase la matrice si espande a velocità supersonica ed in fase gassosa avviene l'ulteriore ionizzazione delle molecole di analita .

Nonostante rimangano ancora da chiarire alcuni aspetti sul processo di ionizzazione MALDI, questo è conosciuto soprattutto nelle sue fasi fondamentali: la protonazione, la deprotonazione e la cationizzazione avvengono in fase gassosa. Inoltre è noto che la ionizzazione avviene in due fasi distinte chiamate ionizzazione primaria e ionizzazione secondaria; i processi della ionizzazione primaria sono dovuti ad una associazione fra un meccanismo di foto-ionizzazione, imputabile ai protoni adsorbiti dalla matrice, ed un meccanismo termico dovuto all'elevata temperatura presente nel plasma d'espansione che è di circa 500°K.

Il processo di ionizzazione primaria dà come risultato la formazione di molecole di matrice allo stato eccitato



Il processo di ionizzazione secondaria ha luogo nel plasma di espansione per un processo di trasferimento protonico in fase gassosa; durante questo tipo di trasferimento protonico possono avvenire due tipi di reazione:

reazione Matrice-Matrice:



e reazione Matrice-Analita:



Durante entrambi i processi di ionizzazione il parametro determinante per la ionizzazione dell'analita è l'affinità protonica. Le proteine ed i peptidi hanno un' affinità protonica pari a circa 240 kcal/mol ⁶, mentre la maggior parte delle matrici comunemente utilizzata ha valori di affinità protonica inferiori, che al massimo raggiungono le 223 kcal/mol. Essendo la reazione di protonazione di peptidi e proteine favorita dal punto di vista termodinamico, questa reazione risulta essere preponderante. Il processo è esotermico e l'energia liberata può causare la frammentazione dell'analita ¹⁵⁹. Questo permette inoltre di spiegare il processo di frammentazione degli ioni metastabili che viene descritto come una conseguenza dell'aumento di temperatura durante il processo di desorbimento. Se al contrario viene utilizzata una matrice con un' affinità protonica uguale o superiore a quella dell'analita, si avrà una diminuzione della frammentazione.

Una volta formati, gli ioni dell'analita subiscono un' accelerazione verso l'analizzatore a tempo di volo (TOF, Time of Flight) che è generalmente combinato con sorgenti MALDI . Tale

accelerazione viene impressa mediante l'applicazione di un campo elettrico, di potenziale (V), il quale fornisce ad ogni ione la medesima energia cinetica (E_C) pari a zeV dove z rappresenta la carica dello ione, mentre e rappresenta la carica di un elettrone ($1,602 \times 10^{-19}$ C).

Essendo l'energia cinetica :

$$E_C = \frac{1}{2}mv^2$$

si ricava che:

$$E_C = zeV$$

Da queste equazioni è possibile ricavare la velocità (v) ed il tempo di percorrenza (t) di uno ione all'interno del tubo di volo di lunghezza L :

$$zeV = \frac{1}{2}mv^2$$

da cui:

$$t = L\sqrt{\frac{m}{2zeV}}$$

e :

$$v = \sqrt{\frac{2zeV}{m}}$$

Risulta quindi evidente che la velocità è inversamente proporzionale alla massa dello ione. Considerando ioni monocarica, quelli con massa maggiore impiegheranno un tempo maggiore a raggiungere il rivelatore rispetto a quelli con massa minore¹⁶⁰.

In conclusione gli ioni vengono separati in base al tempo impiegato a percorrere il tubo di volo. Infatti il tempo di volo è il parametro che viene realmente registrato dallo strumento che, mediante un'apposita calibrazione, fornisce come risultato il valore di m/z degli ioni. Tale parametro è influenzato dalla lunghezza stessa del tubo di volo la quale, a sua volta, determina il potere risolutivo dello strumento. Infatti, più è lungo il percorso che gli ioni devono percorrere, migliore sarà la separazione di ioni con massa simile. La nascita concettuale dell'analizzatore a tempo di volo risale al 1946, ma l'elettronica di acquisizione e rivelazione non erano ancora sufficientemente sviluppate per poter rendere utilizzabile questo tipo di analizzatore. Solamente una decina di anni dopo questo poté essere utilizzato efficacemente.

Nonostante fosse noto da tempo, questo tipo di analizzatore iniziò ad essere maggiormente utilizzato in seguito allo sviluppo della sorgente MALDI.

In particolare, esistono due tipologie di analizzatori a tempo di volo: lineare e *reflectron* (Fig. 1,7). L'analizzatore lineare è più semplice dal punto di vista costruttivo, ma ha minor potere risolutivo, mentre il *reflectron* è più complesso, anche dal punto di vista teorico, ma fornisce una

risoluzione maggiore anche se non è utilizzabile per tutti gli analiti (Fig. 20).

In un analizzatore *reflectron*, viene accoppiato al normale tubo di volo uno specchio elettrostatico che ha il compito di rallentare, riaccelerare e rifocalizzare gli ioni mediante un gradiente crescente di potenziali generati da una serie di anelli o griglie. Il *reflectron* svolge una duplice funzione: aumenta lo spazio che gli ioni devono percorrere per raggiungere il rivelatore ed annulla le piccole differenze di energia cinetica che ioni aventi la stessa massa possono acquisire a causa di piccole disomogeneità del campo elettrico applicato. Tutto questo contribuisce ad aumentare il potere risolutivo dello strumento; questo tipo di analizzatore può così arrivare ad avere potere risolutivo fino a 25.000, però è adatto unicamente per molecole con massa inferiore a 5.000 Da, come ad esempio peptidi, mentre l'analizzatore lineare virtualmente non possiede limiti per quanto riguarda il range di massa analizzabile. Un'altra differenza che si riscontra fra un analizzatore a tempo di volo lineare ed uno *reflectron* è che il primo può essere utilizzato solamente per analisi MS mentre con il secondo sono possibili anche delle analisi MS/MS, che consentono di ottenere informazioni parziali sulla sequenza di peptidi e proteine.

Nel tempo sono state apportate delle modifiche all'analizzatore lineare, queste hanno portato ad un aumento della risoluzione di questo analizzatore ed allo sviluppo del *reflectron*. Come già visto, in seguito alla formazione degli ioni in sorgente, questi vengono accelerati verso il tubo di volo; questo processo non avviene in continuo, ma gli ioni vengono prodotti ed accelerati a pacchetti discreti. Questi pacchetti vengono accelerati verso il tubo di volo mediante un potenziale decrescente (nel caso di ioni positivi) il quale conferisce a tutti gli ioni la medesima energia cinetica. Gli ioni però all'interno della sorgente non si formano tutti alla stessa distanza dall'inizio del tubo di volo a causa del ritardo che si ha fra la ionizzazione e l'estrazione, questo comporta una piccola differenza nei loro tempi di volo dell'ordine del centinaio di nanosecondi. A questa differenza inoltre va sommata la differenza nel tempo di volo dovuta alla non perfetta omogeneità del campo elettrico applicato. Questi due fenomeni, sommati alle collisioni che possono avvenire fra gli ioni durante il processo di ionizzazione/desorbimento e che causano una dispersione dell'energia cinetica, fanno sì che l'analizzatore lineare abbia una risoluzione bassa e generi dei picchi allargati.

Figura 20: schema di analizzatori a tempo di volo lineare (a) e *reflectron* (b). (Immagine tratta dal manuale d'uso dell'ultraflex Extreme, Bruker Daltonics).

La risoluzione al problema della differente distanza dall'ingresso del tubo di volo è stata ottenuta mediante l'estrazione ritardata. Tale metodica si basa sul principio secondo il quale ioni con velocità iniziali differenti possono essere focalizzati mediante l'uso di potenziali di accelerazione appropriati. L'introduzione di un ritardo comune a tutti gli ioni che si formano in sorgente ne

permette la focalizzazione nello spazio correggendo la perdita di risoluzione ⁶⁰. L'introduzione di lenti per l'estrazione ritardata ha portato così ad un aumento della risoluzione da 1.000 a 5.000 per gli analizzatori a tempo di volo lineari e ad un considerevole aumento della risoluzione, anche per gli analizzatori *reflectron*.

L'analizzatore a tempo di volo *reflectron* inoltre consente di effettuare analisi MS/MS. Le analisi MS/MS eseguibili su peptidi e proteine con uno spettrometro di massa MALDI-TOF con *reflectron*, sono di due tipi: ISD (*In Source Decay*) e PSD (*Post Source Decay*). Nell'analisi ISD è di fondamentale importanza il ruolo svolto dall'estrazione ritardata la quale, trattenendo gli ioni in sorgente, aumenta la probabilità di collisione, e quindi di frammentazione, degli ioni dell'analita all'interno della sorgente stessa. Gli ioni frammento così generati vengono rifocalizzati dalle lenti di estrazione ed in seguito all'accelerazione acquisiscono tutti la medesima energia cinetica iniziale e vengono analizzati in modalità *reflectron* ⁶¹. Con questo tipo di analisi è possibile ottenere informazioni sulla sequenza all'N-terminale di proteine pure ed intatte oltre che di peptidi.

Per ottenere invece informazioni sulla sequenza di peptidi, si può utilizzare la tecnica del PSD. Negli spettrometri di massa MALDI-TOF con *reflectron* è possibile ottenere informazioni strutturali dalla frammentazioni degli ioni metastabili all'interno del tubo di volo. Tali ioni accumulano, durante il processo di ionizzazione, una quantità di energia interna tale da portare alla loro frammentazione all'interno del tubo di volo ¹⁶¹. Questi ioni prodotto non possono essere separati da un analizzatore lineare poiché vengono generati all'interno del tubo di volo in assenza di campi elettrici e quindi avranno la stessa velocità dello ione *parent*. Questi frammenti possiedono energie cinetiche differenti che variano a seconda delle loro dimensioni, quindi possono essere separati mediante l'utilizzo del *reflectron*, che svolge il ruolo di rifocalizzare gli ioni in base alla loro energia. Ioni con rapporto m/z maggiore penetrano più a fondo nel *reflectron* e raggiungono il rivelatore più tardi rispetto a ioni con m/z minore; pertanto la separazione in base al rapporto massa/carica dei frammenti viene effettuata solo dal *reflectron*.

Tale tecnica è anche utile per lo studio di peptidi recanti modificazioni post-traduzionali come la fosforilazione ⁶³. Fino a poco tempo fa non esistevano *reflectron* in grado di rifocalizzare tutti i frammenti in un singolo step, quindi per rifocalizzare i frammenti in tutto il range di massa era necessario ridurre progressivamente il potenziale applicato al *reflectron*, acquisire più spettri corrispondenti ai diversi intervalli di massa ed in sede di analisi, con l'utilizzo di opportuni programmi, incollare i segmenti per ottenere lo spettro totale; questo richiede l'utilizzo di una grande quantità di campione. Attualmente, con l'utilizzo di *reflectron* a gradiente quadratico, è possibile rifocalizzare tutti i frammenti in un unico passaggio minimizzando la quantità di

campione necessaria.

Per consentire analisi di questo tipo, alcuni spettrometri MALDI-TOF possiedono un *ion selector* posto subito dopo la sorgente in grado di selezionare solo lo ione *parent* di interesse.

La tecnica MALDI-TOF MS presenta tuttavia alcuni svantaggi tra cui la soppressione ionica, a causa della quale la copertura di sequenza delle proteine risulta minore rispetto a quella ottenuta con altri spettrometri. Tale fenomeno è dovuto alla differente capacità di ionizzare dei diversi peptidi in una miscela. Alcuni di essi, ad esempio quelli particolarmente ricchi in residui acidi, ionizzano meglio di altri provocandone la soppressione nello spettro di massa finale. Questo è anche il motivo per cui l'intensità del segnale nello spettro ottenuto con ionizzazione MALDI non può essere utilizzata come misura quantitativa dell'abbondanza di un peptide nel campione.

Un ulteriore svantaggio è la scarsa frammentazione ottenibile con la metodica PSD, che attualmente non può competere con quella ottenibile mediante analizzatori a quadrupolo o a trappola ionica^{162,163}.

1.4.5 Spettrometria di massa Tandem

La spettrometria di massa tandem è un utile strumento nella determinazione strutturale delle molecole perché permette di ovviare al problema della scarsa frammentazione o dell'assenza della frammentazione con una relativa mancanza di informazioni strutturali.

Uno spettrometro tandem è costituito da due analizzatori disposti in serie, dove il primo analizzatore (MS-1) ha la funzione di selezionare tra i vari ioni presenti in uno spettro lo ione desiderato chiamato generalmente ione precursore. Tale ione viene successivamente fatto collidere con un opportuno gas di collisione (He o Ar) all'interno di una cella. I frammenti da esso derivanti detti "ioni figli",

generati dalla dissociazione dello ione molecolare a causa degli urti col gas, vengono separati dal secondo analizzatore (MS-2) in base a loro rapporto m/z . La selezione di una singola specie ionica monoisotopica del precursore, genera uno spettro di massa formato solo da specie monoisotopiche con una ulteriore semplificazione dello spettro nell'interpretazione. Questa risulta una tecnica molto utile, ad esempio per la caratterizzazione della struttura primaria di peptidi perché consente di avere informazioni sulla composizione amminoacidica.

Gli spettrometri Tandem si dicono a geometria ibrida o non ibrida, a secondo che accoppiano due analizzatori diversi o meno; gli accoppiamenti classici sono: doppio quadrupolo con quadrupolo per la collisione (Q1-Qcoll-Q2), quadrupolo accoppiato con TOF (Q-TOF), magnetico/TOF e TOF/TOF¹⁶⁴.

Anche la trappola ionica (ion trap) può essere inclusa tra i sistemi Tandem sebbene sia costituita da un unico analizzatore in quanto l'analisi MS/MS può essere eseguita nel tempo. Nell'ion trap, infatti, è possibile inizialmente intrappolare tutti gli ioni presenti nello spettro primario, successivamente isolare lo ione desiderato (espellendo gli altri dalla trappola), poi indurre la dissociazione dello ione isolato, e infine analizzare i frammenti generati all'interno della stessa trappola. Un'analisi della progressione nello spazio significa che la selezione dello ione, la sua dissociazione indotta e l'analisi dei frammenti generati avvengono in spazi diversi (cioè in diversi settori dello spettrometro ¹⁶⁵).

La maggior parte delle identificazioni proteiche viene effettuata con analisi nanoelectrospray-MS/MS ma grazie agli sviluppi tecnologici degli ultimi anni, anche la tecnologia MALDI TOF/TOF permette di ottenere gli spettri MS/MS di peptidi e proteine. In particolare lo spettrometro di massa LIFT TOF/TOF è costituito da una sorgente MALDI con estrazione ritardata, un selettore ad alta risoluzione, un apparecchio LIFT, che consente di aumentare il potenziale ionico, e un secondo stage di messa a fuoco con un con la successiva fase di post-accelerazione (sorgente 2).

Il sistema LIFT è costituito da quattro elettrodi disposti in modo da formare tre camere, la camera LIFT, la camera di accelerazione a impulsi e la fase di accelerazione (Figura 3-9) e sono localizzate dietro il PCIS (precursor ion selector). Il selettore dello ione precursore (PCIS) come mostrato in Figura 3-6 funziona come un filtro di massa per separare un particolare ione genitore e relativi frammenti da tutti gli altri ioni derivanti da un'analisi MS / MS.

Nel sistema LIFT (Fig 21), quando il precursore e frammenti relativi entrano nella prima camera, i quattro elettrodi hanno diversi potenziali:

1. Elettrodo 1 e 2 sono collegati a terra.
2. Elettrodo 3 è collegato a 19 kV.
3. Elettrodo 4 è sempre collegato a terra.

Figura 21: schema di funzionamento tecnologia LIFT. (Immagine tratta dal manuale d'uso dell'ultraflex Extreme, Bruker Daltonics).

In genere uno quando lascia la sorgente possiede un'energia pari a 8 keV.

Mentre gli ioni si muovono all'interno della prima cella, il potenziale elettrico dell'elettrodo No.: 1 e No.: 2 arriva a di 19 kV. Poiché non vi è alcuna differenza di tensione e nessun campo tra i due elettrodi il movimento di ioni non è influenzato in alcun modo quindi le particelle continuano a muoversi come prima ma con un potenziale elettrico maggiore.

All'elettrodo 3 viene, successivamente, applicato un potenziale elettrico (2 - 3kV) mentre frammenti di precursori e affini sono ancora all'interno delle celle 2 quindi quando entrano nella cella 3 le particelle subiscono un' ulteriore accelerazione.

Le due fasi di accelerazione a differenza di potenziale di 19 kV aumentano in misura sufficiente le energie cinetiche del precursore e dei frammenti. Ad esempio dopo l'accelerazione con il sistema LIFT l'energia dello ione genitore raggiunge 27 keV (8KeV + 19 keV). Un frammento con un peso molecolare di 1/20 dello ione genitore ha l'energia di 19,4 keV (8keV/20 + 19 keV). Questa energia corrisponde a circa il 28% dello ione genitore e permette allo ione figlia di essere focalizzata ed arrivare al rivelatore con lo ione genitore.

1.4.6 Frammentazione

Gli esperimenti di massa Tandem sono gli strumenti fondamentali per l'identificazione di una proteina in quanto permettono di ottenere informazioni sulla sequenza dei peptidi perchè, all'interno dello spettrometro, questi frammentano secondo processi noti. Una prima nomenclatura degli ioni frammento generati dai peptidi è stata proposta all'inizio degli anni '80 da Roepstorff e Fohlmann 71. E' stato osservato che i peptidi frammentano principalmente a livello del legame peptidico formando ioni chiamati di tipo b ed y: la frammentazione b si otterrà quando la carica sarà trattenuta dal residuo amminico mentre la frammentazione y si otterrà quando la carica sarà trattenuta dal gruppo carbossi-terminale ¹⁶⁶.

Altri frammenti possono generarsi mediante rottura in altre zone del legame oppure su altri legami generando ioni chiamati di tipo a, c, x, z (*Fig. 22*). Brevemente, se la carica è trattenuta nella zona N-terminale del peptide, i frammenti ottenuti per frammentazione a bassa energia vengono indicati con le lettere a, b, e c a seconda della zona del legame peptidico che si rompe, mentre se la carica viene trattenuta nella zona C-terminale del peptide i frammenti vengono indicati con le lettere x, y e z a seconda della zona del legame peptidico che si rompe.

Figura 22: meccanismo di frammentazione di peptidi all'interno di uno spettrometro di massa (Immagine tratta da V.H. Wysocki et al. / *Methods* 35 (2005) 211-222)

E' inoltre possibile la formazione di altri frammenti, d, v e w, i quali sono invece il risultato di una frammentazione ad elevata energia delle catene laterali degli aminoacidi (*Fig 23*).

Figura 23: meccanismo di frammentazione di peptidi all'interno di uno spettrometro di massa (Immagine tratta da V.H. Wysocki et al. / *Methods* 35 (2005) 211-222)

Per poter risalire alla sequenza aminoacidica di un peptide si devono conoscere le regole di frammentazione. Principalmente, lavorando con analizzatori a trappola ionica, si ottengono ioni di tipo b ed y. La differenza di massa fra due frammenti peptidici consecutivi e dello stesso tipo, fornisce la massa corrispondente ad uno dei 20 amminoacidi ^{167,168}. Inizialmente il sequenziamento dei peptidi veniva effettuato mediante il cosiddetto sequenziamento de novo che consiste nello studio degli spettri MS/MS al fine di determinare l'esatta sequenza dei peptidi triptici isolati e frammentati. L'interpretazione manuale degli spettri MS/MS è lunga e laboriosa ma attualmente sono a disposizione degli operatori programmi che aiutano ad elaborare dei dati e a velocizzare l'operazione di sequenziamento dei peptidi. Inoltre sono attualmente disponibili anche algoritmi per la ricerca in banca dati a partire dai dati sperimentali ottenuti da esperimenti MS/MS. Tutto questo porta alla possibilità di identificare una proteina con elevata confidenza statistica anche se si possiedono dati MS/MS relativi ad uno solo dei peptidi che la compongono ¹⁶⁹.

1.4.7 Proteomica quantitativa mediante spettrometria di massa

Oltre all'individuazione di componenti proteiche, la proteomica si occupa anche della misurazione dell'abbondanza delle proteine in sistemi biologici. Mentre la maggior parte degli sforzi iniziali in proteomica si sono concentrati sulla identificazione delle proteine, recenti sviluppi della tecnologia legata alla spettrometria di massa hanno fornito utili strumenti per lo studio dei cambiamenti quantitativi delle proteine. La proteomica quantitativa è importante per la comprensione della funzione molecolare di ciascun componente della proteina e potrebbe permettere di approfondire le conoscenze sui meccanismi molecolari di molteplici processi biologici ¹⁷⁰.

La tecnologia più conosciuta e più utilizzata per gli studi proteomici è sicuramente l'elettroforesi bidimensionale o 2-DE (Two Dimensional Polyacrylamide Gel Electrophoresis), in cui le proteine vengono separate sulla base della sola carica elettrica nella prima dimensione (Isoelectrofocusing) e quindi nella seconda sulla base della sola massa molecolare (Elettroforesi su gel di poliacrilamide/SDS). Con questo tipo di tecnica separativa, partendo da miscele complesse di proteine quali liquidi biologici o lisati cellulari, si ottiene la separazione su un singolo gel di parecchie centinaia di proteine. Con sistemi computerizzati di analisi di immagine si possono quindi, confrontando ad esempio campioni di tessuto sano e frammenti biotici di tumore, individuare le proteine che presentano un diverso grado di espressione.

L'identificazione delle proteine così individuate viene effettuata con notevole velocità e accuratezza tramite spettrometria di massa e con l'ausilio di banche dati internazionali. Inoltre usando tecniche di immuno-blotting le isoforme di una particolare proteina possono essere evidenziate e analizzate, tramite l'uso di anticorpi specifici. Le tecniche basate sulla 2D-E, a causa della loro complessità, scarsa riproducibilità e resa, sono difficilmente proponibili per applicazioni cliniche che prevedano studi multicentrici su un adeguato numero di campioni e sono di fatto scarsamente utilizzabili per campioni biologici quali il plasma in cui proteine come l'albumina e le immunoglobuline, presenti in elevate concentrazioni, mascherano le proteine di origine tissutale potenzialmente interessanti per la caratterizzazione delle condizioni patologiche¹⁷¹.

Negli ultimi anni lo sviluppo di tecniche proteomiche "non-gel based" come il MudPIT hanno apportato potenti strumenti per lo studio e la caratterizzazione dell'espressione proteica nei sistemi biologici¹⁷². La spettrometria di massa è stata a lungo utilizzata a scopo quantitativo solo per piccole molecole ad esempio per ricerche nel campo farmacologico con l'utilizzo di analoghi del composto di interesse marcati con isotopi stabili. All'analisi in massa, i composti marcati presentano picchi analoghi a quelli del composto di riferimento in quanto non ci sono differenze chimiche tra le specie e possono essere analizzati nello stesso esperimento. Queste caratteristiche hanno condotto allo sviluppo di diversi approcci di quantificazione con isotopi stabili sia mediante vie biosintetiche che chimiche¹⁷³.

Esistono principalmente tre diversi approcci quantitativi:

- a) marcatura isotopica metabolica: prevede l'incorporazione di marcatori isotopici durante i processi di metabolismo cellulare e sintesi proteica; uno delle più usate è la SILAC (Stable Isotope Labelling by Amino acids in Cell culture)¹⁷⁴.
- b) marcatura chimica: le più utilizzate sono la tecnologia ICAT (Isotope-Coded Affinity Tag), ICPL (Isotope Coded Protein Labelling), ITRAQ (Isobaric Tags for Relative and Absolute Quantification) e l'incorporazione di O¹⁸ durante taglio enzimatico della proteina;
- c) quantificazione assoluta: la misura della quantità esatta di una proteina o peptide richiede a differenza di quella relativa un valutazione qualitativa prima di quella quantitativa in modo che l'entità da misurare sia definita. Esempi sono l'utilizzo di uno standard interno e la tecnica VICAT (Visible Isotope-Code Affinity Tags)¹⁷⁵.

Il metodo di marcatura chimica delle proteine più utilizzato è l'ICAT. Tale metodo prevede l'introduzione di un marcante costituito da isotopi stabili nelle proteine ottenute da una condizione normale ed una patologica. Le due popolazioni di proteine, marcate ciascuna con un

differente isotopo, vengono miscelate, estratte con una colonna di affinità ed infine analizzate attraverso LC-ESI-MS/MS. Successivamente le proteine vengono identificate e la loro differenza di espressione viene valutata sulla base della diversa intensità degli ioni prodotti dalla spettrometro. Una metodologia di marcatura innovativa nel campo della proteomica è poi rappresentata dalla tecnologia ICPL, basata sulla marcatura isotopica di tutti i gruppi aminoacidici liberi nelle proteine. Questo approccio fornisce una quantificazione altamente riproducibile ed accurata, estese informazioni sulla sequenza proteica, tra cui modificazioni post-traduzionali e isoforme, e la compatibilità con tutte le tecniche di separazione attualmente disponibili¹⁷⁶.

Tuttavia anche queste tecniche presentano delle limitazioni come ad esempio la complessità della preparazione del campione, elevato costo dei reagenti, problemi relativi alla marcatura incompleta e la necessità di utilizzare elevate concentrazioni di campione¹⁷⁷.

1.4.8 Quantificazione proteica con tecnica Label Free

Nel campo della proteomica quantitativa, la tecnica label-free viene sempre più utilizzata per ovviare ai problemi legati all'uso di marcatori e ottenere quantificazioni semplici e veloci¹⁷⁸.

Il metodo label-free prevede diversi passaggi: in primo luogo la preparazione del campione che può includere l'estrazione, riduzione, alchilazione e digestione delle proteine. Il campione viene, quindi, separato con cromatografia liquida e analizzato tramite spettrometria MS/MS per ottenere l'identificazione e la quantificazione peptidica¹⁷⁹.

Due sono le misurazioni che possono essere eseguite per la quantificazione proteica: la prima misurazione è basata sull'intensità del segnale dello ione molecolare nello spettro di massa o su un profilo cromatografico. Questo tipo di quantificazione si basa sulla rilevazione dell'intensità dello ione molecolare, utilizzando uno spettrometro di massa ad alta risoluzione munito di analizzatori quali TOF, FT-ICR o *Orbitrap*¹⁸⁰. L'intensità di questo ione viene standardizzata e applicata a tutti gli altri campioni presi in esame. A seguire, specifici *tools* bioinformatici permettono il *match* dei peptidi corrispondenti, valutando il valore di *m/z*, la carica ionica, il tempo di ritenzione, la selezione di peptidi specifici e la loro identificazione. In particolare, il *software* IDEAL-Q è uno strumento efficiente e robusto per le analisi di quantificazione *label-free*. Esso accetta dati grezzi, come quelli derivanti dalle maggiori piattaforme di ricerca, e utilizza un algoritmo per predire il tempo di eluizione di un peptide non identificato in una specifica corsa LC-MS/MS ma identificato in altre analisi in modo da quantificare il numero più elevato possibile di peptidi. Successivamente il tempo di eluizione viene utilizzato per

individuare i *clusters* di picchi relativi al peptide di interesse, che vengono quindi processati tramite metodi statistici e computazionali esuccessivamente validati attraverso la soglia segnale/rumore, lo stato di carica e la distribuzione isotopica in modo da eliminare il rumore di fondo ¹⁸¹.

Il secondo metodo è basato sul calcolo degli spettri di proteine identificati con analisi MS/MS (spectral count) con le stesse proteine analizzate con tecnologia LC-MS/MS oppure LC/LC-MS/MS in una serie di datasets. Questo tipo di analisi è possibile perché un incremento dell'abbondanza di una proteina contribuisce ad un incremento dei peptidi proteolitici e viceversa. Un incremento del numero di peptidi corrisponde solitamente ad una maggiore copertura della sequenza proteica, del numero di peptidi identificati e del numero totale di spettri MS/MS identificati per ciascuna proteina (spectral count). Liu et al. ha dimostrato che la tecnica dello spectral count mostra una forte correlazione lineare con l'abbondanza relativa delle proteine ($r=0.9997$) con un range dinamico di due ordini di grandezza ¹⁸².

2.SCOPO DELLA TESI

Il carcinoma del rene (RCC) rappresenta il 3% dei tumori solidi ed in media più del 90% di tutte le neoplasie maligne renali che insorgono in soggetti adulti ¹⁸³. Il carcinoma del rene non rappresenta un' unica entità patologica, ma uno spettro di malattia al quale concorrono differenti neoplasie primitive del rene, ognuna con peculiari caratteristiche istologiche, differente decorso clinico, differente risposta alla terapia e differente modificazione genetica alla base della sua insorgenza ¹⁸⁴. Dettagliati studi di morfologia e di citogenetica molecolare hanno permesso di identificare differenti sottotipi di neoplasie epiteliali del rene adulto e i principali includono: carcinoma a cellule chiare, papillare, cromofobo e l'oncocitoma.

Si tratta di una neoplasia due volte più comune tra gli uomini che non tra le donne, con una età media di diagnosi intorno ai 60 anni ¹⁸⁵. La sua incidenza è aumentata in modo significativo negli ultimi trenta anni, sia in Europa che negli Stati Uniti ^{184, 186}, dove si calcolano circa 36000 nuovi casi ogni anno, un terzo dei quali ad esito infausto ¹⁸⁷. Sebbene infatti i pazienti con neoplasie renali localizzate abbiano attualmente una sopravvivenza a 5 anni di circa il 90%, con un notevole incremento rispetto al 40% dei primi anni '60, solo il 18% di quelli con malattia avanzata risulta vivo a due anni dalla diagnosi ¹⁸⁸.

La diagnosi di RCC avviene raramente nelle fasi iniziali e gli unici strumenti di diagnosi sono l'ecotomografia, la TAC e RMN. Non esiste un regime standard di chemioterapia per il carcinoma renale metastatico: i regimi citotossici tradizionali falliscono nel tentativo di impattare significativamente con il decorso della malattia. L'unica terapia radicale attuabile nel paziente con tumore localizzato è la nefrocapsulectomia allargata ¹⁸⁹.

A tutt'oggi non sono stati evidenziati marcatori per una diagnosi precoce di RCC nei pazienti asintomatici o per il monitoraggio durante il trattamento post chirurgico della malattia. La ricerca di molecole da utilizzare come biomarcatori è attualmente perseguita con diversi approcci sia genomici che proteomici.

In particolare sta aumentando l'interesse per l'applicazione della proteomica alla diagnostica clinica ed alla medicina preventiva. L'approccio proteomico permette l'individuazione di livelli di espressione di proteine e di loro modificazioni post-traduzionali determinate da reazioni di fosforilazione, glicosilazione, metilazione, deamidazione, ipotizzando la possibilità di identificazione di un cluster di biomarcatori coinvolti nel processo tumorale. Poiché le proteine sono deputate alla maggior parte dei processi biologici degli organismi, la spettrometria di massa (MS), grazie alla sua elevata sensibilità e specificità, è stata riconosciuta come uno strumento indispensabile per gli studi di proteomica ¹⁹⁰.

I recenti sviluppi della spettrometria di massa hanno permesso di raggiungere elevati livelli di selettività, al fine di riconoscere una determinata sostanza anche se presente in miscela con molte altre, ed alti standard di sensibilità, nell'ordine delle femto (10^{-15}) o atto (10^{-18}) moli.

Un aspetto fondamentale dell'analisi proteomica è costituito dallo studio del *processing* che il prodotto genico primario subisce, fino al raggiungimento della sua completa e definitiva funzione. Varianti proteiche strutturali e funzionali sono prodotte grazie all'azione di specifici enzimi, coinvolti in un numero elevato di processi cellulari e sono molto importanti per la determinazione funzionale di una proteina e possono influenzare anche la sua stabilità conformazionale, la localizzazione e l'emivita cellulare.

La spettrometria di massa è in grado di fornire informazioni strutturali qualitative e quantitative, per la determinazione di macromolecole nel campo della ricerca biomedica e tutti questi vantaggi analitici hanno portato a considerare la spettrometria di massa la tecnica d'elezione per l'analisi di sostanze presenti in basse concentrazioni in matrici complesse, quali sono quelle di origine biologica^{191,192}.

Data l'elevata complessità dei campioni analizzati e la necessità di eseguire un'analisi proteomica affidabile e di qualità, è essenziale un' iniziale procedura di purificazione dei fluidi biologici. Spesso una fase di pre-frazionamento del campione viene applicata allo scopo di rimuovere interferenti come sali e contaminanti ed aumentare così il numero di proteine/peptidi rilevabili. In questo modo è possibile ottenere profili proteici da fluidi periferici di pazienti e selezionare da essi un cluster di segnali in grado di distinguere lo stato di malattia.

Un approccio innovativo completamente automatizzato nella fase di prefrazionamento prima dell'analisi in MS di siero, plasma, urina, lisati e prodotti cellulari per la ricerca di nuovi marcatori biologici, è rappresentato dalla tecnologia Clin Prot. Tale tecnica rappresenta un utile strumento di profiling grazie alla purificazione dei campioni con microsferi magnetiche funzionalizzate, che sono in grado di selezionare gruppi molecolari limitati per la successiva indagine in spettrometria di Massa MALDI-TOF (Matrix Assisted Laser Desorption Ionization Time of Flight).

Il progetto di questa ricerca è stato dunque rivolto allo studio e alla caratterizzazione del peptidoma serico di soggetti sani e pazienti con carcinoma renale, così come alla diagnosi differenziale delle forme di tumore renale maligno da lesioni benigne. A questo proposito sono stati purificati, mediante tecnica ClinProt che utilizza microparticelle magnetiche funzionalizzate, il siero di un ampio numero di pazienti affetti da ccRCC e soggetti controllo, estendendo lo studio anche a pazienti affetti da non-ccRCC. Le analisi di profiling peptidico sono state condotte mediante SM MALDI-TOF seguita da elaborazione dei profili

spettrali per all'individuazione di clusters con potenziale diagnostico. Inoltre si è cercato di identificare proteine/peptidi con alterata espressione serica. A questo scopo sono state effettuate analisi mediante tecnologia nLC-ESI MS/MS e MALDI-TOF MS/MS. I risultati del "profiling" tramite MALDI-TOF sono stati verificati utilizzando un approccio proteomico diverso basato su una quantificazione relativa label-free in nLC-ESI MS/MS.

3.MATERIALI E METODI

3.1 Raccolta dei campioni

I campioni di siero sono stati raccolti da pazienti e soggetti volontari sani presso diversi istituti ospedalieri: Ospedale di Desio (Desio, Milano, Italia), Fondazione Policlinico dell'Ospedale Maggiore di Milano (Milano, Italia) nel periodo compreso tra Marzo 2003 e Giugno 2010.

La raccolta del siero dei pazienti affetti da carcinoma renale è avvenuto il giorno prima dell'intervento chirurgico. Tutti i volontari sani e i pazienti hanno preventivamente firmato e dato il consenso alla raccolta dei campioni. I protocolli e le procedure attuati durante lo studio sono stati

approvati dal comitato etico locale e le analisi sono state effettuate in accordo con la Dichiarazione di Helsinki.

Metodo

Ciascuna aliquota è stata preparata raccogliendo 10 mL di sangue in vacutainer. La provetta è stata lasciata in posizione verticale per 30 minuti a temperatura ambiente in modo da favorire il processo della coagulazione. Il siero è stato poi separato dal coagulo tramite centrifugazione a 2000 rpm per 10 minuti a 4°C e suddiviso in piccole aliquote, le quali sono state subito congelate a -80 °C fino al momento dell'analisi. Le aliquote di siero così ottenute sono state scongelate una sola volta prima di essere sottoposte al processo di purificazione.

3.2 Prefrazionamento dei campioni attraverso tecnica ClinProt

Materiali e solventi

L'estrazione specifica di proteine/peptidi di tutti i campioni di siero in studio è stata effettuata con biglie magnetiche a fase inversa di tipo C18 (RPC18 Dynabeads).

Inoltre sono stati utilizzati i seguenti solventi:

- Metanolo, CH₃OH
- Acetonitrile, CH₃CN
- Acetone, CH₃COCH₃
- Etanolo, C₂H₆O
- Acido trifluoroacetico (TFA), CF₃COOH

- Soluzione 0.1% di TFA, acido trifluoroacetico CF_3COOH .

Tutti i solventi utilizzati sono di grado analitico.

Per l'analisi dei campioni in MALDI-TOF MS è stata impiegata come matrice l'acido α -ciano-4-idrossicinnamico (CHCA) 0,3 g/L solubilizzato in Acetone/Etano 1:2.

Strumenti

Tutti i passaggi di purificazione sono stati eseguiti in maniera automatizzata con l'ausilio di una piattaforma robotizzata a 8 canali (*ClinProt Robot*- Bruker Daltonics, Germany). L'automazione, oltre a rendere più veloce ed economico il processo di purificazione, permette di ottenere una maggiore riproducibilità rispetto ad una preparazione dei campioni manuale.

Metodo

Il protocollo di purificazione si basa su diversi passaggi fondamentali.

- Fase di legame (**binding**): 10 μL di biglie magnetiche C18 vengono miscelati con 40 μL di TFA 0.1 % e 5 μL di siero; dopo 1 minuto di incubazione il surnatante viene rimosso.
- Fase di lavaggio (**washing**) : 3 lavaggi delle biglie con 25 μL di TFA 0.1 % ; il surnatante viene eliminato ad ogni passaggio.
- Fase di eluizione (**elution**): : gli analiti vengono eluiti con 15 μL di una soluzione al 50% di acetonitrile/ H_2O dopo 1 minuto di incubazione.
- Fase di deposito (**spotting**): 2 μL di eluato vengono miscelati con 10 μL di matrice in soluzione (0,3 g/L di 4-Hydroxy- α -cyanocinnamic acid HCCA in Acetone/Etano 1:2) e 0.8 μL di questa miscela vengono depositati su target MTP AnchorChipTM 600/384.

La componente di proteine/peptidi eluita da ogni campione e non depositata sul target è stata riunita con gli altri eluati della popolazione di appartenenza (controlli e pazienti in base alla patologia); un'aliquota è stata quindi congelata a -80°C fino al momento dell'analisi in nano-LC-ESI MS/MS.

Altre 2 aliquote sono stati preparate miscelando 50 μL di siero di 80 pazienti (ccRCC n=62 e non-ccRCC n=18) e 80 controlli.

3.3 Studio di profiling con tecnica MALDI-TOF

Materiali, standard e solventi

Per la calibrazione dello strumento, in modalità lineare è stata utilizzata una miscela di calibrazione adeguata per proteine a basso peso molecolare (range 1000 a 10000 Da): Protein Mix I in 0.1% TFA.

In modalità reflector (range 1-3 KDa), invece, per ottenere la massa molecolare accurata delle proteine/peptidi contenuti nel campione da analizzare è stata utilizzata la miscela di calibrazione PepMix in 0.1% TFA.

Strumenti e software

Lo spettrometro di massa UltrafleXtreme™ MALDI-TOF/TOF MS (Bruker Daltonics, Germany) è stato utilizzato per ottenere gli spettri di massa. Le condizioni strumentali per l'analisi in lineare di tutti i campioni di siero sono state:

- ion source 1 = 25.13 kV
- ion source 2 = 23.77 kV
- lens = 6.46 kV
- pulsed ion extraction = 300 ns
- high gating factor with matrix suppression up to 900 Da

Il processo di ionizzazione è stato condotto utilizzando un laser Smartbeam™ ad una frequenza di 1000 Hz con una potenza del 60% del valore massimo.

Le analisi sono state eseguite con il metodo automatico AutoXcute (versione 3.3.108.0) del software FlexControl versione 3.3 (Bruker Daltonics, Germany).

Gli spettri in lineare sono stati acquisiti in un range di massa compreso tra 1000-12000 Da.

In modalità reflectron, invece, i parametri impostati sullo strumento sono i seguenti:

- ion source 1 = 25.13 kV
- ion source 2 = 22.43 kV
- lens = 7.45 kV
- deflection mode con matrix suppression = 400 Da.

Metodo

Prima di iniziare l'analisi in massa dei campioni, lo strumento viene calibrato con miscele esterne di calibrazione e a questo proposito sul target sono state depositate due miscele di calibrazione standard che contengono polipeptidi a valori di massa noti. Tali valori consentono la costruzione di una curva di calibrazione al fine di convertire i tempi di volo in valori di m/z.

Dopo la calibrazione con la miscela ProtmixI sono stati acquisiti i campioni depositati sul target. Lo spettro di ciascun campione è stato ottenuto sommando gli spettri ottenuti da “spari” (bombardamenti con fotoni) successivi in posizioni diverse del pozzetto in cui è stato depositato il campione; in questo modo si riesce a migliorare il rapporto S/N dello spettro. Poiché la matrice non cristallizza in modo uniforme è stato analizzato tutto il pozzetto contenente il campione per poter ottenere il maggior numero di informazioni ed il più possibile riproducibili. Lo spettro di ciascun campione, quindi, è stato ottenuto sommando i migliori spettri ottenuti da sei differenti posizioni all’interno di ciascun deposito mediante 200 shots per un totale di 1200 shots. Con il programma di acquisizione AutoXcute, i profili spettrali sono stati ottenuti attraverso la sommatoria dei migliori profili scartando quelli con scarsa risoluzione e basso rapporto segnale/rumore di fondo (S/N). Un altro vantaggio del software è quello di impostare i parametri necessari per acquisire gli spettri di massa dei campioni attraverso una più rapida acquisizione rispetto a quella “manuale”.

Una volta conclusa l’analisi in lineare è stata effettuata una misurazione in reflectron per ottenere le masse accurate dei segnali individuati attraverso l’analisi di profiling.

Innanzitutto è stata eseguita una calibrazione esterna attraverso la miscela Pepmix e, successivamente, è stata effettuata l’acquisizione dei profili spettrali in modalità manuale in modo da verificare che i segnali fossero ben risolti e che gli spettri fossero tutti ben calibrati.

3.4 Identificazione peptidica con tecnica MALDI-TOF

Strumenti e software

Lo spettrometro di massa UltrafleXtreme™ MALDI-TOF/TOF MS (Bruker Daltonics, Germany) in modalità LIFT TOF/TOF è stato utilizzato per l’identificazione dei peptidi.

Le analisi sono state eseguite utilizzando i seguenti parametri:

- Ion Source 1 = 7,5 kV;
- Ion Source 2 = 6,7 kV;
- Lens = 3,6 kV;
- Reflector 1 = 29,5 kV;
- Reflector 2 = 13,95 kV;
- Lift 1 = 19 kV;
- Lift 2 = 3,15 kV;
- Pulsed ion extraction 80 ns
- High-energy collision-induced dissociation (CID): N/A

Il motore di ricerca Mascot (versione 2.3.02) è stato utilizzato per l'analisi in banca dati impostando i seguenti parametri:

- Tassonomia: ristretta alla sezione umana
- Database Swissprot (Febbraio 2012– 20317 sequenze);
- Enzimi o modificazioni fisse o variabili: N/A;
- Tolleranze MS e MS/MS : 1 Da.

Le identificazioni accettate sono state quelle con uno score più elevato delle soglie di identità di Mascot.

Metodo

Le masse dei frammenti sono state analizzate dopo aver acquisito la massa accurata in modalità reflectron. Lo strumento viene nuovamente calibrato con la miscela di calibrazione Pepmix per procedere quindi con l'acquisizione automatica degli spettri grazie al software Flex Control™. Ciascuno spettro è stato acquisito dopo 500 shots e sono state così ottenute le masse esatte dello ione parent e dei relativi frammenti per eseguire l'identificazione dei peptidi su Mascot.

3.5 Analisi bioinformatica e biostatistica

Software

- FlexAnalysis™ 3.3 per l'elaborazione degli spettri acquisiti al MALDI-TOF.
- ClinProTools™ 2.2 per l'indagine statistica degli spettri e la visualizzazione delle differenze di espressione.

Metodo

Gli spettri ottenuti sono stati inizialmente rielaborati con il software FlexAnalysis™ 3.3 per migliorarne la qualità in modo da poter procedere con l'analisi attraverso il software statistico ClinProTools™ 2.2 che permette il confronto tra spettri multipli, la rilevazione e la visualizzazione di pattern proteici in grado di discriminare le popolazioni di studio.

FlexAnalysis™

La qualità degli spettri è stata ottimizzata con il software FlexAnalysis™ attraverso diversi passaggi fondamentali. Innanzitutto è stato eseguito un primo confronto esclusivamente visivo delle eventuali differenze di segnale e della presenza di spettri non soddisfacenti o non acquisiti

in automatico dallo strumento (segnali poco intensi e in scarso numero, segnale del rumore di fondo elevato oppure *shots* insufficienti) attraverso la modalità di visualizzazione di tutti gli spettri acquisiti su una o più finestre (Fig.24). Gli spettri qualitativamente insufficienti sono stati facilmente riacquisiti reinserendo il target nello strumento.

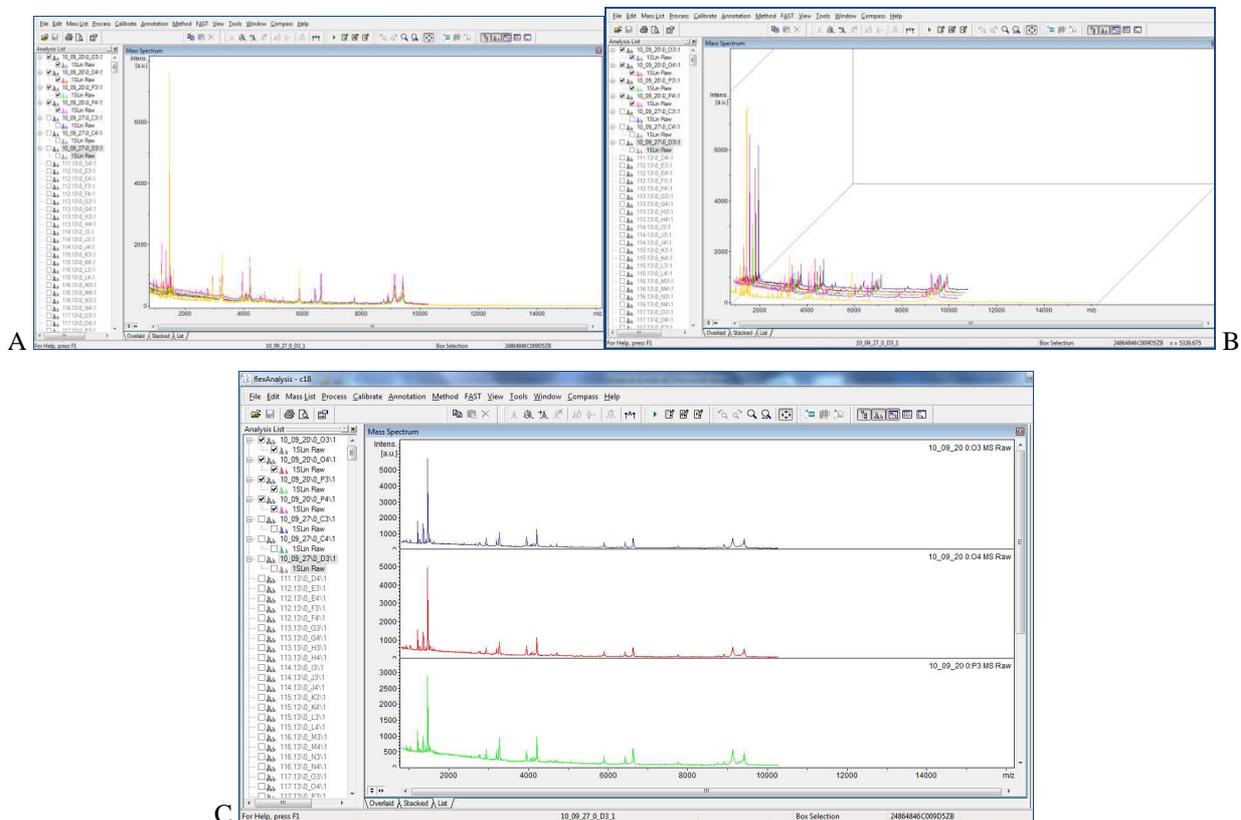


Figura 24 Finestra principale del software FlexAnalysis. In questa schermata gli spettri possono essere visualizzati sovrapposti (A) in 3D (B) oppure in più sotto finestre (C)

Dopo questa analisi preliminare, gli spettri acquisiti in MALDI-TOF sono stati, quindi, allineati attraverso uno spettro già acquisito e calibrato in una precedente analisi. Attraverso questa rielaborazione i segnali ottenuti da diversi campioni in differenti analisi possono essere fra loro sovrapponibili e, di conseguenza, comparabili.

Per ottenere i valori di m/z dei picchi e parametri indispensabili quali l'intensità e risoluzione di ciascun picco il software è dotato dei seguenti algoritmi:

- *Apex*: consente di visualizzare tutti i valori di massa dei picchi in corrispondenza di ogni apice e per rilevare i picchi utilizza le derivate prima e seconda. È necessario impostare alcuni parametri: l'intensità relativa, l'ampiezza minima di un segnale (peak width) affinché sia identificato come picco (es. 0,1 m/z ; se $< 0,1 m/z$ è rumore di fondo) e la soglia segnale/rumore (mai < 3). Il parametro peak width è un indice che permette di

accettare positivamente solo i segnali che presentano un valore uguale o maggiore a quello impostato ed è molto importante poiché se tale parametro è impostato ad un valore troppo elevato, c'è il rischio che i segnali bassi siano considerati come rumore.

- *Centroide*: calcola il valore di massa nella parte media del segnale ad una determinata altezza; è possibile quindi impostare la percentuale di altezza che il software deve considerare rispetto alla linea di base (valore generalmente raccomandato è 80%). Questo algoritmo utilizza lo stesso meccanismo di calcolo dell'algoritmo Apex, tuttavia risulta più accurato, poiché quando i picchi si presentano frastagliati un valore m/z calcolato in corrispondenza degli apici risulta poco affidabile.
- *Sum*: calcola le cosiddette “pseudo derivate”, invece delle reali derivate calcolate dagli algoritmi Apex e Centroide. Se il valore della pseudo derivata supera una certa soglia, il picco è considerato come tale. Inoltre, Sum esegue una correzione interna della linea di base e un calcolo del rumore di fondo, al fine di selezionare correttamente i picchi e quantificare i rapporti segnale/rumore.
- *Snap*: considera il cluster isotopico di ogni peptide e calcola le masse monoisotopiche utili per un successivo chiarimento sulla struttura e sequenza di un particolare peptide; inoltre esegue automaticamente una correzione interna della linea di base e una determinazione del rumore. È opportuno, quindi, utilizzare Snap solo con dati grezzi non processati precedentemente, su quali non è mai stata eseguita una sottrazione della linea di base o uno smoothing. Anche in questo caso è necessario impostare alcuni parametri, tra cui l'intensità relativa, la soglia segnale/rumore, il numero massimo di picchi considerati ed, infine, il fattore di qualità che indica quanto il picco reale si discosta da quello isotopico. Questo algoritmo rappresenta l'algoritmo migliore per un'analisi di spettrometria di massa su peptidi ed è quello che è stato utilizzato per l'analisi degli spettri dei campioni in esame.

Grazie all'interfaccia *BioTools* presente su questo software è stato, inoltre, possibile collegarsi al motore di ricerca *Mascot* al fine di identificare le proteine o i peptidi di interesse.

Grazie al software statistico Clinprotools™ è stato possibile non solo creare modelli predittivi per la selezione dei picchi, ma anche validare i modelli diagnostici, utilizzando campioni di prova di cui già si conosce già la popolazione di appartenenza, al fine di determinare la sensibilità e la specificità degli stessi.

Nella schermata principale del software ClinProTools™ sono visibili quattro finestre riportanti le principali informazioni relative ai profili spettrali in esame.

La finestra *Spectra View* (Fig.25) permette di visualizzare lo spettro singolo, lo spettro medio totale e lo spettro medio di ciascuna classe. Sull'asse x vengono riportati i valori di m/z, mentre l'asse y i valori di intensità dei picchi secondo unità arbitrarie.

Le operatività connesse a questa funzione sono molteplici. Innanzitutto è possibile ottenere dell'area dei picchi di ogni spettro come si può vedere dalle bande azzurre riportate nella figura sottostante selezionando "peak calculation", inoltre si può includere ed escludere manualmente gli spettri non processati dal software o determinati picchi, nonché forzare i picchi statisticamente più significativi evidenziandoli al fine di utilizzarli nella costruzione di un modello di classificazione. Infine, è possibile visualizzare in questa finestra anche alcune informazioni statistiche relative a tutti i picchi in esame o ad un numero ristretto di essi (es. media e deviazione standard, distribuzione dei picchi).

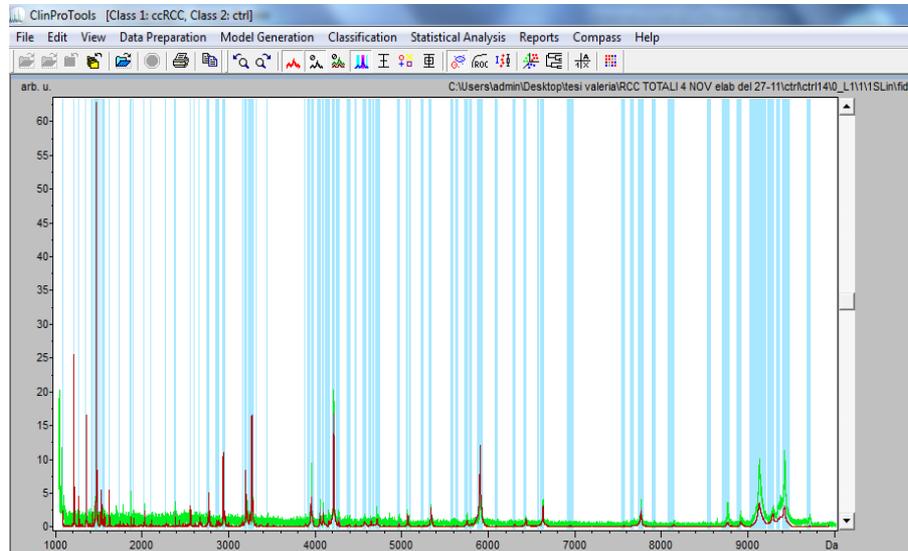


Figura 25. Finestra di visualizzazione *Spectra View*

La finestra *Model List View* (Fig.26): consente di creare modelli di classificazione per le popolazioni in esame. Per ogni modello sono riportati il nome, il tipo di algoritmo utilizzato, i valori di "Cross Validation" (indice predittivo di bontà del modello per future classificazioni) e "Recognition Capability" (capacità del cluster di classificare correttamente i campioni con cui è stato elaborato).

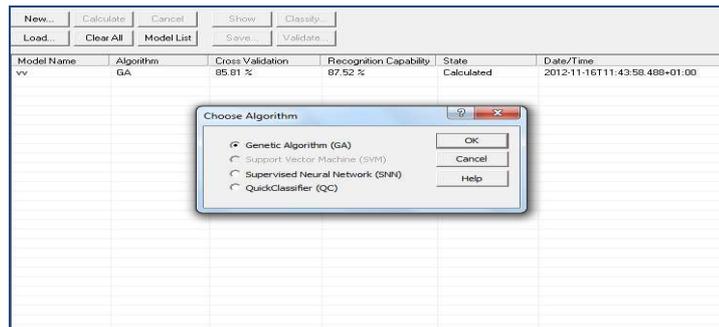


Figura 26. Finestra di visualizzazione *Model List View*

Per creare un modello per le popolazioni in esame il software permette di scegliere tra i quattro algoritmi di seguito brevemente illustrati.

Il *Genetic Algorithm (GA)* è utilizzato per selezionare le migliori combinazioni di picchi statisticamente significativi in grado di separare le popolazioni in esame (Fig.27).

Da ricordare che gli Algoritmi Genetici sono una famiglia di tecniche di ottimizzazione che si ispirano alla teoria della selezione naturale di Charles Darwin che regola l'evoluzione biologica. I sistemi naturali esibiscono un alto grado di robustezza ed efficacia nel risolvere una vasta serie di problemi per la loro sopravvivenza grazie al processo evolutivo basato sulla riproduzione selettiva degli individui migliori, sulla ricombinazione genetica dei loro cromosomi e su alcune mutazioni casuali.

Così l'algoritmo seleziona i modelli più idonei, mentre quelli meno predittivi vengono scartati. L'idoneità è definita dal gruppo di tecniche e dai calcoli per la selezione. I modelli selezionati sono poi sottoposti a ricombinazione (crossing over), dove coppie di modelli vengono scomposte, modificate e poi riunite. Tutto questo può essere seguito da una mutazione, dove un singolo picco del modello viene sostituito con un altro. Con i modelli più idonei (migliori dei loro progenitori) viene eseguita nuovamente la ricombinazione in modo da ottenere una nuova generazione e il processo si ripete fino a raggiungere il modello migliore. In ogni modello che si desidera creare mediante algoritmo genetico, è possibile selezionare il numero massimo di picchi utilizzati per la costruzione, il numero massimo di generazioni e il numero di spettri nelle vicinanze che il modello deve considerare per la classificazione dei profili di interesse. Inoltre, altri due parametri importanti per la definizione del modello migliore sono il valore di "Mutation Rate", che indica la probabilità che si verifichi una mutazione, e il valore di "Crossover Rate", che è la probabilità che si verifichi crossing over.

Il risultato finale di un modello di questo tipo è rappresentato dalla combinazione di picchi che separa nel miglior modo possibile le popolazioni in esame.

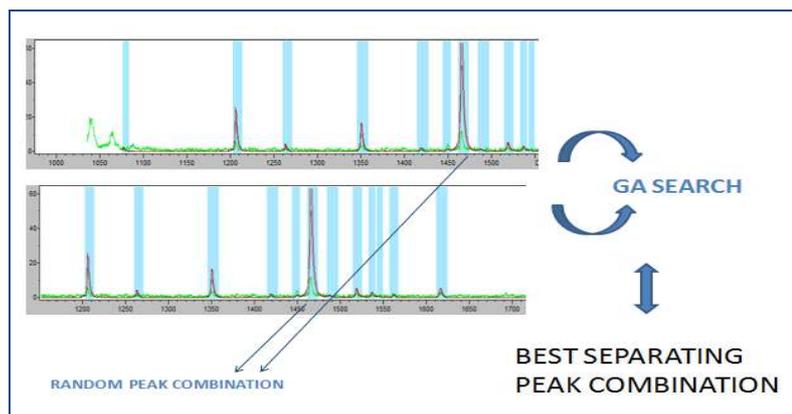


Figura 27. Meccanismo di funzionamento utilizzato per algoritmo GA

L'algoritmo *Support Vector Machine (SVM)* seleziona, invece, una combinazione di picchi che permettano di separare le popolazioni in esame attraverso l'utilizzo di un vettore (Fig. 28). L'idea principale di questo classificatore è quella di rappresentare i picchi appartenenti a classi differenti in punti nello spazio. SVM costruisce una separazione tra questi punti tale che siano separati dal più ampio gap possibile.

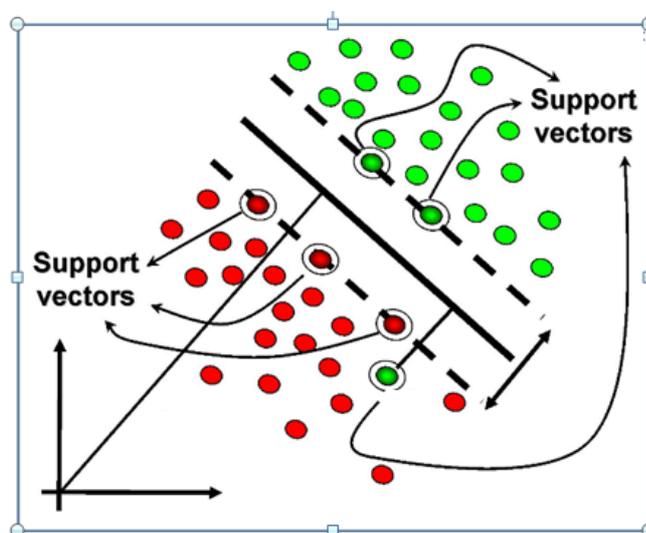


Figura 28. Meccanismo di funzionamento utilizzato per l'algoritmo Support vector machine

Un altro algoritmo che può essere selezionato è il *QuickClassifier (QC)* seleziona tutte le aree/intensità dei picchi con relativi valori di p-value e calcola la media pesata di questi. Opera, quindi, un confronto dei valori delle medie pesate ed indica sia l'insieme di picchi in grado di fornire una migliore separazione sia la probabilità di corretta classificazione per ogni popolazione in esame. La classe con il più alto indice di probabilità rappresenta la classe predittiva

L'ultimo dei quattro algoritmi è il *Supervised Neural Network (SNN)*. Il meccanismo di funzionamento implica un'identificazione iniziale di un certo numero di "prototipi", ovvero spettri caratteristici per ogni classe. L'algoritmo lavora, poi, in modo da distribuire i campioni nel piano in relazione all'area/intensità dei diversi segnali e seleziona i set di dati in cui le classi sono meglio separate, ovvero in cui campioni appartenenti allo stesso gruppo vengono correttamente classificati in relazione ai prototipi.

Un'altra finestra disponibile sulla schermata iniziale del software ClinProtools è *Gel/Stack View*. Attraverso questa visualizzazione vengono mostrati tutti i segnali dei campioni che costituiscono le due classi in analisi attraverso un'immagine virtuale di gel (pseudo-gel). Due finestre possono essere visualizzate attraverso questa modalità: *Gel View* e *Stack View*.

La finestra *Gel View* (Fig 29) visualizza i picchi degli spettri sottoforma di spots di diversa densità sottoforma di gel virtuale. Sull'asse x sono riportati i valori di m/z, mentre sull'asse y a sinistra il numero degli spettri in successione.

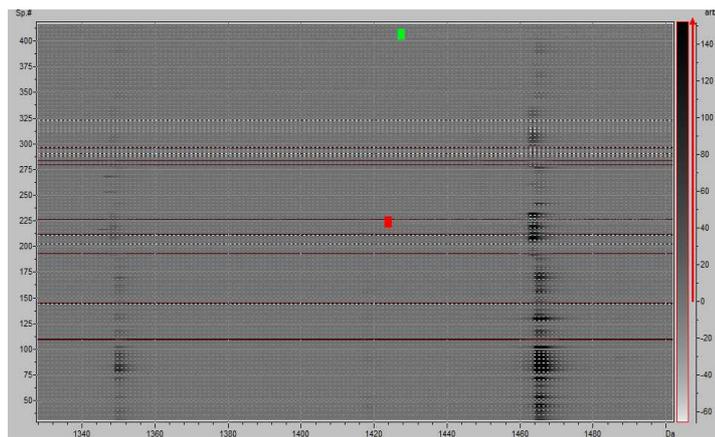


Figura 29. Visualizzazione *Gel view*.

La finestra *Stack View* (Fig.30), invece, mostra tutti gli spettri delle classi in un'immagine 3D in modo da poter valutare il corretto allineamento dei picchi di ciascuno spettro. Sull'asse x sono riportati i valori di m/z, sull'asse y le intensità dei picchi in unità arbitrarie e sull'asse z l'ordine di caricamento delle classi.

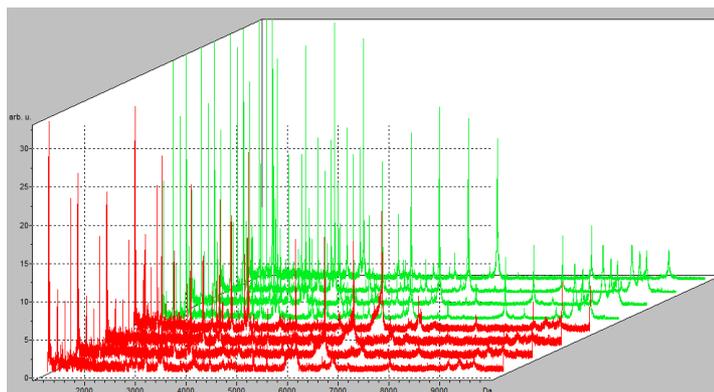


Figura 30. Visualizzazione *Stack View*

L'ultima finestra è la *Peak Statistics View* nella quale sono disponibili tre modalità per la visualizzazione dei segnali più discriminanti: *2D Peak Distribution View*, *ROC Curve View* e *Single Peak Variance View*.

La finestra *2D Peak Distribution View* (Fig.31) mostra la distribuzione dei primi due picchi statisticamente significativi, rilevati dal software, in tutti gli spettri delle classi in esame. Il numero e il valore m/z di questi picchi è indicato sugli assi x e y. Solitamente l'asse x riporta i valori di area/intensità del primo picco, che presenta infatti minore p-value, mentre l'asse y i valori di area/intensità del secondo picco. È comunque possibile sostituire questi due picchi scelti automaticamente dal software con altri segnali a piacere e valutarne la distribuzione. Le ellissi rappresentano la deviazione standard delle aree/intensità dei picchi o l'intervallo di confidenza del 95%.

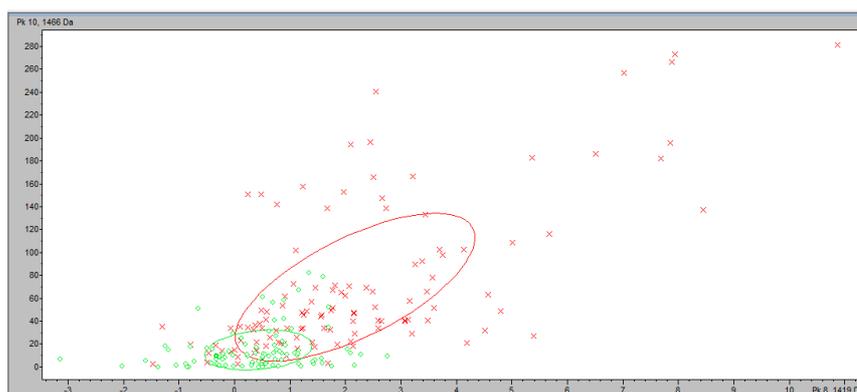


Figura 31. Finestra *2D Peak Distribution View*

La finestra *ROC Curve View* mostra la curva ROC (Receiver Operating Characteristic curve) relativa ad un determinato picco precedentemente selezionato (Fig 32).

La curva ROC offre una visione grafica circa la specificità e sensibilità di un test e, nel caso di questo software, fornisce una stima della capacità discriminante del picco. L'asse x riporta la specificità in termini di falsi positivi, mentre l'asse y la sensibilità.

La capacità del picco di classificare correttamente i soggetti controllo e i pazienti è valutata attraverso il valore AUC, che rappresenta l'area sotto la curva. Un valore AUC uguale a 0 indica una capacità discriminante del picco nulla, mentre un valore AUC uguale a 1 un'ottima capacità discriminante.

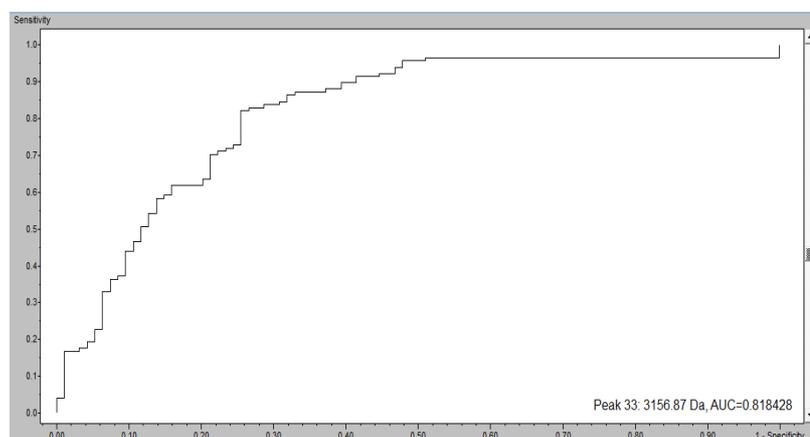


Figura 32. Finestra di visualizzazione ROC curve

La finestra *Single Peak Variance View* (Fig 33) mostra tre tipologie di dati statistici relativi ad un determinato picco precedentemente selezionato: “box and whiskers”, media e deviazione standard e distribuzione del picco. Tutte queste informazioni sono ottenute dai valori di area/intensità del picco selezionato e sono riportate separatamente per ogni classe in studio.

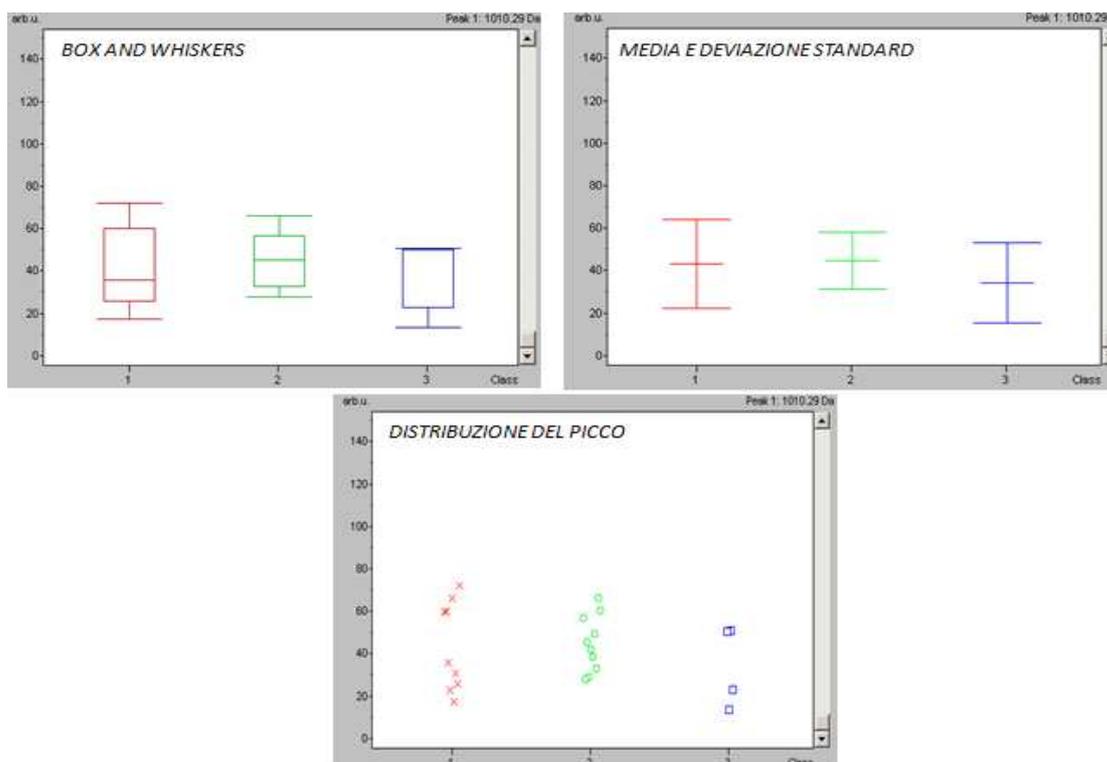


Figura 33. Tre diverse modalità di visualizzazione della finestra *Single Peak Variance View*

Per una corretta elaborazione dei dati e un ulteriore riallineamento degli spettri utilizzando i picchi di maggiore intensità è stata effettuata, innanzitutto, una normalizzazione sulla corrente ionica totale (TIC) di tutti gli spettri in elaborazione. Successivamente per ogni picco è stata calcolata l'area e l'intensità, che sono correlate all'abbondanza relativa della sostanza nei diversi campioni di studio.

Con il software è stato possibile calcolare lo spettro medio di ogni campione per ottenere una corretta analisi statistica. Infatti le funzioni "Support Spectra Grouping" ed "Enable Similarity Selection" hanno permesso di selezionare lo spettro il cui profilo è più simile alla media complessiva in questo modo solo uno spettro, dei quattro acquisiti per ogni campione mediante analisi MALDI-TOF, è stato utilizzato nell'elaborazione.

L'indagine di possibili pattern proteici in grado di discriminare le popolazioni in esame è stata condotta tramite utilizzo degli algoritmi matematici elencati precedentemente.

I modelli ottenuti sono stati, quindi, validati attraverso una cross-validation. Il software permette due tipi di validazione. Nell' *internal validation* possono essere classificati gli stessi spettri che sono stati utilizzati per creare il modello. Grazie all' *external-validation* nuovi campioni non impiegati nella costruzione del modello classificativo possono essere classificati utilizzando il modello di analisi, ottenendo una previsione della classe di appartenenza.

La validazione determina le dimensioni della classe, la specificità e la sensibilità del modello ottenuto con l'insieme di dati prova e determina la capacità di previsione del modello in base alla percentuale dei dati di prova correttamente classificati.

Due importanti funzioni del software sono state impiegate per rilevare le differenze di espressione tra le classi in esame: Peak Statistic Report e Correlation Matrix Report.

Peak Statistic Report sono elencati tutti i picchi rilevati dal software durante la "Peak Statistic Calculation" (ricalibrazione degli spettri, calcolo degli spettri medi e calcolo dei picchi) nell'ordine dal più discriminante a quello meno differente a livello di espressione. Sulla base di un t-test/Anova per insiemi di dati parametrici ed un Wilcoxon/Kruskal-Wallis per test non parametrici, viene calcolato per ogni picco il valore *p-value* (*PTTA* e *PWKW*) indicativo della probabilità che la differenza di espressione trovata nelle popolazioni in esame non sia fondata su un evento casuale. Dunque questo indice rappresenta il livello di significatività effettiva ed è indipendente dall'altezza e dall'area del picco. Se il *p-value* è $\leq 0,05$ il segnale è statisticamente significativo ai fini di una buona separazione, mentre se il *p-value* è $> 0,05$ il segnale non permette una discriminazione delle classi. Il valore *PAD* rappresenta il valore *p-value* ottenuto da un test Anderson-Darling ed indica se il segnale è parametrico ($PAD > 0,05$) o non parametrico ($PAD \leq 0,05$) e quindi suggerisce se è necessario rifarsi al *PTTA* o al *PWKW*. Nella tabella sono riportati anche il numero del picco, il valore di massa, la media delle aree di ogni classe e la differenza tra queste (*DAve*) ed, infine, le deviazioni standard di ogni classe in esame (*StdDev*).

Correlation Matrix Report: è una matrice ottenuta mediante analisi di correlazione tra tutti i picchi identificati dal software e riportati nel Peak Statistic Report. Questa analisi può essere eseguita su più classi a confronto oppure su una singola classe. Logicamente, questa funzione deve essere preceduta dalla ricalibrazione degli spettri e dal calcolo dei picchi. Ogni riquadro della matrice presenta un determinato colore a seconda del tipo di correlazione tra i segnali dell'incrocio: rosso se è presente correlazione positiva, blu per una correlazione negativa e grigio quando non vi è correlazione. Quando il valore di correlazione si avvicina a +1, significa che è presente correlazione positiva e quindi ad un aumento di intensità del picco corrisponde un aumento del segnale correlato all'interno della stessa classe in esame. Quando, invece, il valore si avvicina a -1, significa che è presente correlazione negativa e quindi ad un aumento di intensità del picco corrisponde una riduzione del segnale correlato all'interno della stessa classe in esame. Infine, quando il valore si avvicina a ± 0 , non è presente correlazione tra i due segnali.

I parametri utilizzati dal software per le analisi sono stati ottimizzati in base alla tipologia di campione studiato.

3.6 Identificazione peptidica attraverso tecnica nLC-ESI MS/MS

Materiali

- Puntali da pipetta *ZiptipTM μ -C18* (Millipore, Badford, USA)
- Centrifuga da vuoto *Heto Speed-Vac* (Heto Lab. Equipment A/S, Copenhagen, Denmark)

Strumenti e software

- Spettrometro di Massa *MaXis UHR-QToF* (Bruker Daltonics, Germany)
- HPLC *EasynLCTM* (Proxeon Biosystems, Odense, Denmark)
- Software *Compass DataAnalysisTM v.4.0 Sp4* (Bruker Daltonics, Germany)

Il motore di ricerca Mascot (versione 2.3.02) è stato utilizzato per l'analisi in banca dati impostando i seguenti parametri:

- Tassonomia: ristretta alla sezione umana
- Database Swissprot (Febbraio 2012– 20317 sequenze);
- Enzimi o modificazioni fisse: N/A;
- Modificazioni variabili: Acetyl (N-term); Acetyl (K), Amidated (Cterm), Oxidation (M); Phospho (Y) / Phospho (ST), Sulfo (S) / Sulfo (Y) / Sulfo (T)
- Tolleranze MS e MS/MS : 5-10 ppm 0,1-0,5 Da.

Sono state considerate le soglie di Mascot per la definizione di omologia e di identità e il decoy database come filtri per la determinazione della significatività dei peptidi.

Metodo

Le frazioni di eluato ottenute grazie al prefrazionamento con l'utilizzo di biglie magnetiche RP-C18 del siero sono state riunite ottenendo in questo modo due pool di sieri: il primo derivante dagli 80 pazienti analizzati mentre l'altro ottenuto dagli 80 soggetti sani.

Dopo essere stati concentrati in centrifuga da vuoto, i due pool di siero sono stati desalificati e privati delle biglie rimanenti utilizzando puntali da pipetta *ZiptipTM μ -C18* preventivamente equilibrate con 0.1% TFA, lavate con 0.1% TFA ed eluite con 50% Me OH/0.1% TFA come da protocollo della ditta produttrice.

I pool sono stati diluiti prima dell'iniezione per evitare la saturazione dei segnali ed analizzati a diverse concentrazioni per aumentare il numero delle identificazioni.

Le frazioni così ottenute sono state iniettate ad un flusso di 2 $\mu\text{L}/\text{min}$ attraverso il sistema EasynLCTM accoppiato al sistema ibrido MaXis UHR-QToF secondo il protocollo descritto da Mainini et al.

I dati grezzi MS/MS vengono poi rielaborati col software Compass DataAnalysisTM grazie al quale è stato possibile correggere i valori per la *lock-mass* (m/z 1221.9906), effettuare una deconvoluzione ed infine ottenere una lista di picchi. Per escludere ioni con rapporto segnale/rumore minore di 1 e intensità inferiore a 20 *counts* è stata attivata la funzione *Peakfinder*.

I frammenti ottenuti dall'analisi nano liquid chromatography (nLC)-ESI MS/MS sono stati identificati con il motore di ricerca Mascot secondo i parametri di cui sopra.

DataAnalysis 4.0

Il software permette di innanzitutto di visualizzare i cromatogrammi relativi all'acquisizione in massa, gli spettri MS e MS/MS attraverso diverse modalità:

- *TIC (Total Ion Current)*: visualizzazione della corrente ionica totale rilevata nel tempo.
- *Base Peak Chromatogram*: visualizzazione della corrente ionica del del picco di maggiore intensità (picco base)
- *Extracted Ion Chromatogram*, visualizzazione del cromatogramma di un determinato ione di cui si conosce il valore m/z monocarica ottenuto con l'analisi MALDI-TOF. È necessario trasformare questo valore in massa bicarica o tricarica, in modo da inserire un intervallo di massa (solitamente un dalton in più e uno in meno rispetto al valore m/z di interesse) utile al software per l'estrazione i frammenti di questo ione. Questa operazione è indispensabile, poiché con la tecnica ESI si ottengono generalmente ioni a cariche multiple..

I cromatogrammi dei dati MS/MS ottenuti vengono rielaborati, come già illustrato precedentemente. Tale operazione comprende anche la deconvoluzione che viene effettuata al fine di passare dallo spettro di massa sperimentale multicarica ad uno spettro di massa monocarica in modo da ottenere il peso molecolare vero della molecola.

Successivamente viene effettuata una rilevazione automatica dei compounds presenti in tutti gli spettri MS e MS/MS che permette di eseguire un'identificazione accurata dei frammenti collegandosi al programma informatico Biotools da cui aprire il motore di ricerca MS/MS Ion Search di Mascot (Fig.34)

Database Mascot

Mascot è un potente motore di ricerca che utilizza dati provenienti da analisi in spettrometria di massa per l'identificazione di proteine. La ricerca all'interno di Mascot può essere effettuata con differenti approcci in relazione ai dati analitici in possesso. In particolare il *Peptide Mass Fingerprint (PMF)* consente di caricare, all'interno del database, i valori di massa di peptidi provenienti dalla digestione, enzimatica o chimica, di una o più proteine. Con *MS/MS Ion Search* è possibile effettuare la ricerca con dati ottenuti con un'analisi MS/MS, ma in questo caso occorre fornire un'informazione aggiuntiva ovvero la metodica usata per l'analisi MS/MS.

Un altro approccio solitamente usato per la ricerca è *Sequence Query* in cui ad uno o più peptidi, utilizzati per il PMF, possono essere associate informazioni parziali sulla sequenza, sulla composizione amminoacidica, sui valori di massa/carica di ioni provenienti da analisi MS/MS o altri ancora.

L'identificazione delle proteine, per questo lavoro, è avvenuta attraverso uno dei più accurati database: *SwissProt*. La sua caratteristica fondamentale è che in una singola voce sono contenute la sequenza di ciascuna proteina e le modificazioni note in questo modo nel database non sono contenute informazioni ridondanti e le ricerche sono veloci.

Per una corretta identificazione, sono state inseriti alcuni parametri fondamentali per la ricerca con *MS/MS Ion Search*: i valori di massa dello ione *parent* e dei frammenti generati a partire da quest'ultimo durante l'analisi, il tipo di spettrometro di massa con il quale si è eseguita l'analisi, il tipo di frammentazione avvenuta e la carica degli ioni ed infine è stato specificato che le masse sperimentali fanno riferimento al valore di massa monoisotopica dei peptidi.

Un altro importante parametro da indicare è la *tolleranza* che indica la misura dello scarto massimo che ci può essere tra i dati sperimentali e quelli teorici e riveste un ruolo fondamentale sulla significatività statistica dell'analisi.

Minore è il valore di tolleranza utilizzato per la ricerca, maggiore è l'affidabilità del risultato ottenuto, al contrario, inserendo un margine di errore troppo ampio, si potrebbe ottenere l'identificazione di un falso positivo e non della proteina corretta.

Tale parametro può essere espresso in *Dalton*, in percentuale o in parti per milione (ppm) e per l'identificazione finale (*MS/MS Tol*) sono stati inseriti due valori: quello per il ione *parent* (*Mass Tol MS*) e quello per i singoli ioni frammento.

Per la presente ricerca non sono stati usati enzimi per la digestione della proteina ma il database permette di indicare l'enzima utilizzato e gli eventuali siti di taglio persi per ottenere la massima discriminazione possibile.

E' noto, inoltre, che proteine e peptidi possono essere modificate in centinaia di modi: alcune modifiche riguardano la funzione biologica, altri sono elementi di manipolazione del campione. Nel database possono essere inserite due tipi di modifiche. Una *modifica quantitativa*, ad esempio, potrebbe essere l'alchilazione delle cisteine. Poiché tutte le cisteine sono state modificate, questo comporta un cambiamento nella massa della cisteina, al contrario, la maggior parte delle *modificazioni post-traduzionali* non si applicano a tutti i residui. Ad esempio, una fosforilazione potrebbe influenzare solo una sola delle serine contenute in un peptide. Le modifiche non quantitative possono aumentare il tempo richiesto della ricerca e ridurre la sua specificità perché il software deve analizzare tutte le possibili combinazioni di residui modificati e non modificati che si adattano alla massa molecolare del peptide.

Una terza classe di modifiche che possono verificarsi in un peptide sono quelle dovute ad un esperimento di quantificazione con isotopi stabili. Un peptide potrà legare un isotopo pesante o leggero ma mai una miscela di isotopi leggeri e pesanti quindi all'interno del database possono essere effettuate due ricerche separate uno per isotopi leggeri e uno per isotopi pesanti

Una variabile che può essere inserita al fine di ridurre i tempi di ricerca e semplificare il risultato ottenuto è la *tassonomia* che permette che proteine omologhe appartenenti ad altre specie vengano eliminate.

Set Mascot search form defaults

Database	SwissProt NCBIInr contaminants cRAP
Taxonomy	All entries
Enzyme	Trypsin
Allow up to	1 missed cleavages
Fixed modifications	Acetyl (K) Acetyl (N-term) Acetyl (Protein N-term) Amidated (C-term) Amidated (Protein C-term)
Variable modifications	Acetyl (K) Acetyl (N-term) Acetyl (Protein N-term) Amidated (C-term) Amidated (Protein C-term)
Show all mods.	<input type="checkbox"/>
Quantitation	None
Peptide tol. ±	1.2 Da # ¹³ C 0
MS/MS tol. ±	0.6 Da
Peptide charge	1+
Monoisotopic	<input checked="" type="radio"/> Average <input type="radio"/>
Data format	Mascot generic (MS/MS only)
Instrument	Default (MS/MS only)
Decoy	<input type="checkbox"/>
Error tolerant	<input type="checkbox"/>
Report top	AUTO hits

Figura 34. Schermata di Mascot

Le identificazioni basate su dati provenienti da analisi con spettrometria di massa sono sempre valutate con uno score probabilistico quindi l'informazione che si ottiene rappresenta la probabilità che il match ottenuto sia un evento casuale. Lo score in Mascot viene assegnato con l'algoritmo *MOWSE (Molecular Weight Search) scoring algorithm* che classifica le proteine risultanti dalla ricerca in banca dati in base ad uno score che rappresenta una misura della probabilità assoluta (P) che la proteina identificata sia un falso positivo e non la proteina effettivamente analizzata. Lo score è quindi definito con la seguente formula:

$$\text{score} = -10 \text{Log}_{10} (P)$$

Nel caso di analisi MS/MS, viene assegnato uno *score* anche ad ogni singolo ione riconosciuto (*Ion Score*), per ognuno dei quali ci sarà un valore soglia variabile in base a diversi parametri tra cui la lunghezza della sequenza e la tolleranza scelta per lo ione *parent*.

La soglia di significatività statistica riportata in Mascot corrisponde ad un p-value < 0,5 ovvero una possibilità inferiore al 5% che l'identificazione possa essere attribuibile al caso. Questo valore è predefinito e viene segnalato nella pagina di riepilogo dei risultati attraverso degli istogrammi dove ogni barra rappresenta una determinata proteina, mentre l'area non evidenziata con colore verde indica le proteine statisticamente significative.

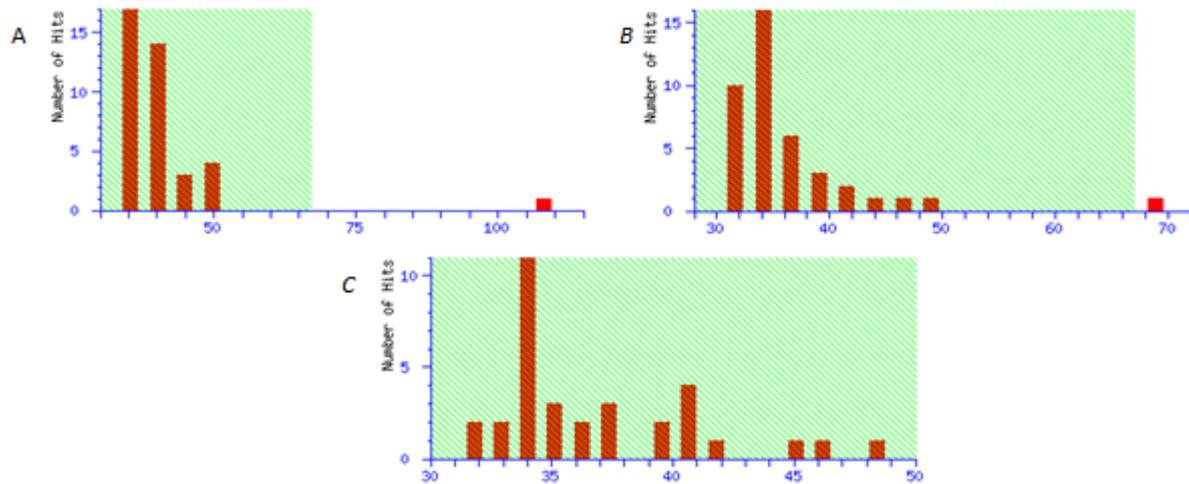


Figura 35. Significatività statistica di una proteina identificata in Mascot

Nella figura di cui sopra sono riportati tre possibili risultati ottenuti per l'identificazione di una proteina. Nel caso A il risultato ottenuto ha un alto score e un'alta significatività, nel caso B, invece, lo score cade proprio vicino alla finestra di significatività. Nell'ultimo caso, figura C, l'identificazione non è statisticamente significativa.

Nella pagina dei risultati, al di sotto dell'istogramma, sono riportate un tutte le possibili molecole riconosciute dal software che presentano un profilo simile. Nella prima posizione di questo elenco è riportata la proteina che più probabilmente (score più alto) costituisce la molecola di interesse; di quest'ultima è possibile visualizzare alcune informazioni generali tra cui il peso molecolare, il punto isoelettrico, lo score, la funzione e la copertura della sequenza amminoacidica (in rosso sono evidenziati i tratti di sequenza riconosciuta o meglio i peptidi utilizzati per l'identificazione, mentre in nero quelli non allineati). In particolare, il parametro "Intensity Coverage" indica la percentuale di copertura della sequenza della proteina identificata, cioè i picchi o peptidi che sono stati utilizzati da Mascot e che hanno contribuito al riconoscimento.

Grazie a Mascot è possibile collegarsi direttamente alla banca dati SwissProt per avere una descrizione più dettagliata della proteina.

3.7 Tecnica Label-Free: quantificazione peptidica/proteica

Strumenti e software

- Software *IDEAL-Q* (versione 1.6.02)
- Software *CompassXport* versione 3.0.4 (Bruker Daltonics, Germany).

Metodo

Le frazioni ottenute dalla purificazione dei due pool costituiti dagli 80 sieri e dagli 80 controlli sono stati analizzati con 3 run ciascuno con lo spettrometro MaXis UHR-QToF e i dati ottenuti sono stati convertiti in formato mzXML utilizzando il software CompassXport.

Invece nei file in formato mgf ottenuti dall'identificazione con Mascot e con l'output di elaborazione DataAnalysisTM, il numero di *compounds* (cmpd) è stato sostituito con il numero di *scan*. In questo modo è stato possibile esportare i file in formato XML al fine di allineare i dati degli spettri (mzXML) con i risultati della ricerca. L'allineamento dei picchi è stato eseguito con i parametri impostati con predefiniti degli strumenti Q-TOF.

La normalizzazione è stata effettuata utilizzando la mediana dell'abbondanza peptidica e il quantile del *ratio* del livello peptidico. Le abbondanze peptidiche sono state determinate utilizzando gli stessi parametri e una soglia di *score* peptidico impostata a 37 (*score* per identità o estesa omologia) come calcolati dall'algoritmo di *Mascot*.

I *ratios* peptidici ottenuti con la tecnica ESI MS/MS sono stati confrontati con quelli calcolati a partire dagli spettri MALDI-TOF. Per raggiungere questo scopo, gli spettri di massa MALDI-TOF ottenuti in modalità lineare (MALDI-LM) per il siero di pazienti e di controlli, sono stati utilizzati per calcolare il *ratio* m/z tra pazienti e controlli.

L'allineamento dei *ratios* m/z ottenuti in MALDI-LM con quelli dell'analisi LC-ESI è stato effettuato utilizzando come riferimento gli spettri di massa acquisiti in MALDI-TOF modalità *reflectron* (MALDI-RM), mediante l'accertamento dell'identità di alcuni di essi con un'analisi MALDI-TOF/TOF e la validazione aggiuntiva data dalle misurazioni delle masse accurate nei profili ottenuti dalle analisi MALDI FT-ICR.

Inoltre, per confrontare i *ratios* peptidici MALDI e i corrispondenti *ratios* ESI normalizzati, quantificati mediante approccio *label-free* è stato utilizzato un *t-test* per dati appaiati. È stato impostato un p -value < 0.05 per verificare che la media delle differenze tra le due analisi, non sia statisticamente significativa.

4.RISULTATI

4.1 Popolazione e dati clinici dello studio

Lo studio è stato condotto su tre gruppi di soggetti: 92 soggetti controllo (ctrl) (68 uomini e 24 donne), 85 soggetti affetti da tumore renale a cellule chiare (ccRCC) (54 uomini e 31 donne) e 29 soggetti affetti da tumori con differente sottotipo istologico (non-ccRCC) (15 uomini e 14 donne). L'età media dei soggetti controllo era di 46 anni con un range compreso tra 30 e 64 anni mentre per i pazienti l'età media alla diagnosi era di 64 anni con un range compreso tra 33 e 87 anni.

Come visibile dalla tabella 11, i pazienti sono stati valutati secondo la stadiazione TNM (tumor, node, metastasis) del 2009, e, prima dell'exeresi chirurgica della lesione renale, sono stati sottoposti ad un'analisi istologica nella quale sono stati valutati diversi parametri: il sistema di *grading* di Fuhrman, il riconoscimento del sottotipo istologico, il differenziamento sarcomatoide e cistico, la necrosi del tumore, l'infiltrazione tumorale microvascolare e della pelvi urinaria.

	N° di PAZIENTI
NUMERO TOTALE	114
Età media ± SD alla diagnosi	64 ± 12.04
Età media alla diagnosi (range)	33-87
SESSO	
Maschi	69
Femmine	45
STADIAZIONE	
Tumore primario (T)	
pT1	74
pT2	23
pT3	1
pT4	0
sconosciuto	4
Linfonodi (N)	
NX	68
N0	28
N1	1
sconosciuto	5
GRADO	
G1	4
G2	68
G3	21
G4	1
sconosciuto	8
ISTOLOGIA	
RCC a cellule chiare	85
RCC papillare	9
RCC a cellule chiare + papillare	3
RCC Cromofobo	2
Oncocitoma	6
Angiomiolipoma	5
Altri sottotipi	4
TIPO TUMORALE	
Maligno	102
Benigno	12

Tabella 10. Caratteristiche e classificazione secondo la scala TNM del 2009 dei pazienti inclusi nello studio

I pazienti sono stati quindi divisi in due gruppi: ccRCC (n=85) e non-ccRCC (n=29)

4.2 Analisi proteomica ed efficacia diagnostica dei pattern

I campioni di siero dei tre gruppi di soggetti analizzati in questo studio, soggetti controllo, pazienti affetti da ccRCC e pazienti non-ccRCC, sono stati sottoposti a pre-frazionamento con biglie magnetiche a fase inversa C18 prima dell'analisi in spettrometria di massa MALDI-TOF. Gli spettri dopo essere stati acquisiti in un range di massa compreso tra 1 e 12 kDa, sono allineati ed elaborati e, successivamente, valutati a livello statistico mediante il software *ClinProTools*TM.

L'analisi statistica condotta sui profili serici ottenuti, ha permesso l'elaborazione di uno spettro medio dei profili serici per ognuna delle classi in studio (fig. PRRO 36) e dal loro confronto sono stati individuati più di 100 segnali in comune e circa 40 segnali con una AUC che differisce in modo statisticamente significativo tra i tre gruppi ($p < 0,05$ con analisi ANOVA ad una via).

Data l'elevata complessità dei profili spettrali, l'analisi statistica e l'individuazione dei *patterns* discriminanti, sono state effettuate mediante l'utilizzo di due algoritmi forniti dal software: l'Algoritmo Genetico (GA) e il *Support Vector Machine* (SVM) in modo da fornire la più accurata classificazione possibile delle tre popolazioni in studio.

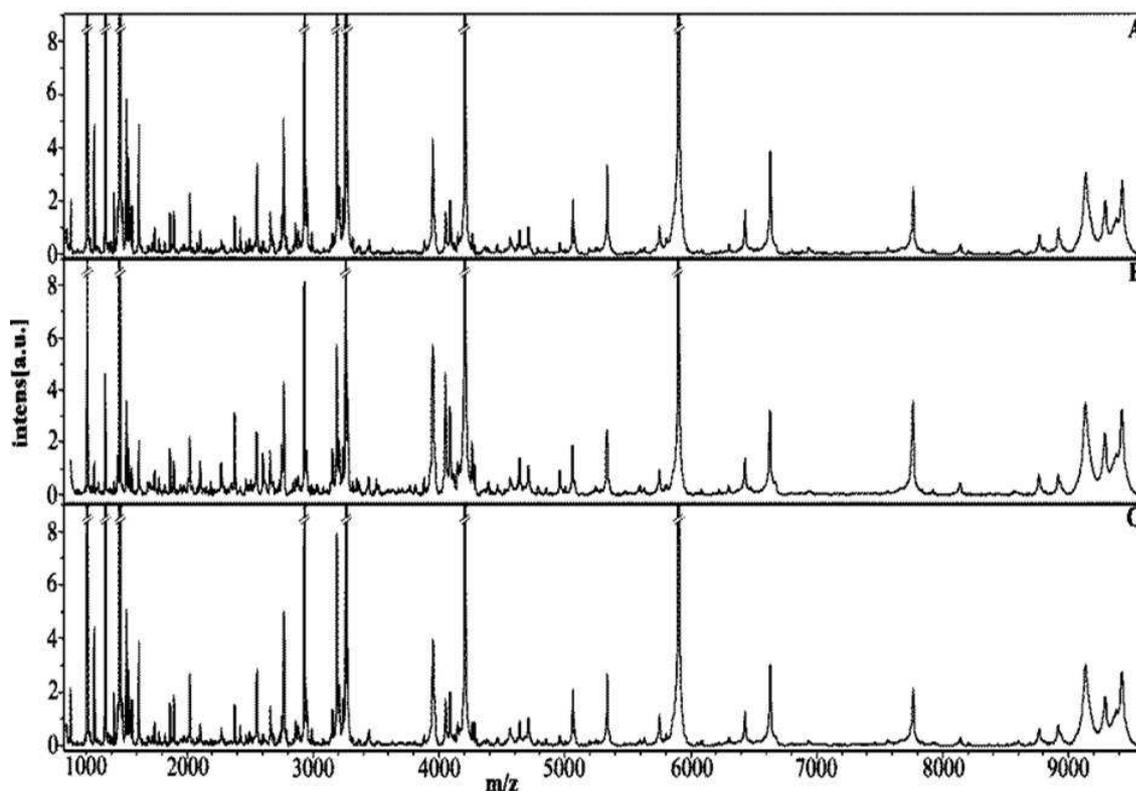


Figura 36. Profili spettrali serici medi delle tre classi in studio. A. Profilo dei pazienti cc-RCC B. Profilo dei controlli C. Profilo pazienti non cc- RCC

L'intera casistica è stata inizialmente suddivisa in due gruppi: uno composto da soggetti con tumore maligno (n=102) mentre l'altro formato da soggetti con masse renali benigne (n=12) e soggetti controllo (n=92). Grazie all'analisi statistica, è stato individuato un cluster discriminante formato da cinque segnali con m/z 1264, 1419, 2423, 2789 e 4962 Da in modalità lineare. La combinazione di questi picchi ha mostrato un miglioramento della *performance* rispetto a quella di ogni singolo ione e la validazione interna del cluster, nella fase di training, ha ottenuto valori di specificità e sensibilità rispettivamente di 90% e 94% (Tab. 12). L'analisi statistica effettuata con le curve ROC, ha inoltre valutato il valore di AUC per i cinque ioni che è risultato compreso in un *range* di 0,84 e 0,75. Secondo i criteri suggeriti da Swets¹⁹³, valori compresi tra 0,7 e 0,9 suggerisco un test moderatamente accurato mentre AUC superiori a 0,9 indicano un test altamente accurato, pertanto i segnali individuati hanno un buon potere diagnostico. A conferma, una *cross-validation* di tipo *k-fold* ($k=10$) è stata applicata per valutare la capacità predittiva di questo modello e i risultati ottenuti dalla classificazione mostravano elevati valori di specificità 84% e sensibilità 89%.

Tutti i segnali individuati sono risultati sovra-espressi ($p<0,05$) nei pazienti con tumori maligni rispetto ai soggetti controllo più soggetti con masse renali benigne a parte lo ione ione a m/z 4962 che risultava sotto-espresso ($p<0,05$).

m/z	MALIGNANT vs BENIGN			MALIGNANT vs BENIGN			
		SPECIFICITY	SENSITIVITY		true false	true true	classe precision
1264							
1419	TRAINING PHASE	90.4%	94.1%	Pred false	85	5	94.4%
2422	TEST PHASE	83.7%	89.2%	Pred true	9	87	90.6%
4961				spec, sens	90.4%	94.1%	
2789							

Tabella 11. Performances classificative del confronto tra tumori maligni *versus* lesioni renali benigne e controlli

Successivamente sono state effettuate tre ulteriori e differenti classificazioni caso/controllo: controlli *versus* ccRCC, controlli *versus* non-ccRCC e ccRCC *versus* non-ccRCC. Per ognuna delle comparazioni caso/controllo sono stati ottenuti i valori di specificità e sensibilità riportati in Tabella 4B e 4C. Mediante analisi statistica degli spettri MALDI-TOF in modalità lineare, è stato individuato un *cluster* di due peptidi a m/z 1466 (AUC=0,89) e a m/z 4052 (AUC=0,85) aventi un'accuratezza elevata nella discriminazione di soggetti controllo e pazienti affetti da ccRCC. La loro combinazione ha mostrato una specificità e una sensibilità entrambe del 92% nella fase di *training*, mentre nella fase di validazione la specificità e la sensibilità sono risultate rispettivamente dell'88% e dell'85%. Allo scopo di migliorare l'efficienza discriminativa di questo modello predittivo basato su due ioni, ad esso sono stati aggiunti altri ioni i cui valori di

AUC erano compresi tra 0,89 e 0,70 che corrispondono ad un test moderatamente accurato²⁰⁸. Un *pattern* di quattro (m/z in LM 1466, 2378, 2789 e 4052) ed uno di sei segnali (m/z in modalità lineare 1466, 2378, 2789, 4052, 4089 e 4151) hanno permesso di migliorare la capacità classificativa come si può vedere dai risultati riportati in Tabella 13.

La combinazione dei sei ioni nell'ultimo *cluster* ha mostrato indici di specificità (96%) e sensibilità (94%) più elevati se confrontati con i valori ottenuti con gli altri *patterns* aventi un numero minore di segnali. L'analisi di *cross-validation* ha confermato la buona capacità diagnostica ottenuta nella fase di *training* (93,5% di specificità e 91% di sensibilità). Due di questi ioni, a m/z 1466 e 2789, erano presenti in maggiore concentrazione nei pazienti affetti da ccRCC, mentre gli altri mostravano una concentrazione maggiore nei soggetti controllo. Dopo l'aggiunta di nuovi segnali nel *pattern*, non sono stati riscontrati miglioramenti nella *performance* dello stesso, inoltre al fine di evitare il fenomeno dell'*overfitting* dei dati, è stato deciso di non aumentare il numero degli ioni inclusi nel *cluster* discriminante.

m/z	CTRL vs ccRCC			CTRL vs ccRCC			
1466		SPECIFICITY	SENSITIVITY		true false	true true	classe precision
4052	TRAINING PHASE	92.4%	91.8%	Pred false	77	6	92.8%
	TEST PHASE	88.1%	84.7%	Pred true	6	71	92.2%
				spec, sens	92.4%	91.8%	
m/z	CTRL vs ccRCC			CTRL vs ccRCC			
1466		SPECIFICITY	SENSITIVITY		true false	true true	classe precision
2378	TRAINING PHASE	96.7%	90.6%	Pred false	80	7	92.0%
2789	TEST PHASE	89.1%	89.4%	Pred true	3	70	95.9%
4052				spec, sens	96.7%	90.6%	
m/z	CTRL vs ccRCC			CTRL vs ccRCC			
1466		SPECIFICITY	SENSITIVITY		true false	true true	classe precision
2378	TRAINING PHASE	95.7%	94.1%	Pred false	79	5	94.0%
2789	TEST PHASE	93.5%	90.6%	Pred true	4	72	94.8%
4052				spec, sens	95.7%	94.1%	
4088							
4150							

Tabella 12. Performances classificative del confronto tra controlli versus pazienti cc-RCC

Un'ulteriore confronto statistico eseguito tra soggetti sani dai pazienti affetti da non-ccRCC ha messo in evidenza un *pattern* di tre ioni (m/z in LM 1419, 2603, 4264) con una forte capacità diagnostica (rispettivamente 96% e 90% di specificità e sensibilità nella fase di *training* e 95% e 90% nella fase di validazione) (Tab.14). Lo ione a m/z 1419 (AUC=0,85) è stato osservato essere più espresso ($p<0,05$) nei pazienti affetti da non-ccRCC rispetto ai soggetti controllo, mentre gli altri segnali erano sottorappresentati.

m/z	CTRL vs non-ccRCC		CTRL vs non-ccRCC			
		SPECIFICITY	SENSITIVITY	true false	true true	classe precision
1419						
2603	TRAINING PHASE	95.7%	89.7%	79	3	96.3%
4264	TEST PHASE	94.6%	89.7%	4	23	85.2%
				spec, sens	95.7%	89.7%

Tabella 13. Performances classificative del confronto tra controlli versus pazienti non cc-RCC

E' stata inoltre eseguita una comparazione tra pazienti affetti da ccRCC e quelli non ccRCC tramite un'analisi statistica eseguita con lo stesso approccio utilizzato in precedenza per la selezione dei segnali basato, su due algoritmi e con l'inclusione di sei o più picchi ma non è stato possibile individuare un cluster che permettesse di discriminare le due popolazioni.

4.3 Identificazione dei peptidi

L'analisi in spettrometria di massa mediante nLC-ESI MS/MS e la ricerca in database con il server Mascot hanno condotto all'identificazione dei peptidi delle frazioni ottenute dopo pre-frazionamento con biglie magnetiche RP-C18 del siero di controlli e pazienti affetti da tumore renale.

Sono stati identificati circa 200 peptidi unici aventi uno *score* maggiore della soglia di identità peptidica e con un *false discovery rate* <1,5% appartenenti a 32 diverse proteine e 262 peptidi ottenuti utilizzando come *cut-off* la soglia di omologia con un *false discovery rate* <5% con i quali è stata costruita una *library*. Più del 25% di tutti i peptidi identificati presenti nella *library* sono stati individuati inserendo nei parametri di ricerca di *Mascot* alcune possibili modificazioni post-traduzionali. Tra queste, le più frequenti risultavano essere, come già riportato per i peptidi endogeni da Tinoco *et al.*209, la solfatazione, la fosforilazione, l'acetilazione N-terminale e l'amidazione C-terminale.

Il confronto dei segnali differentemente espressi nelle classi in studio ha permesso di individuare 30 peptidi identificati solo nei pazienti e 53 solo nei soggetti sani.

Circa il 60% delle 201 isoforme peptidiche identificate, con uno *score* maggiore della soglia di identità, è stato individuato sia nei campioni dei pazienti che in quelli dei controlli.

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
725,250	21	27	A	DSGEGDF	L	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
738,475	12	18	W	PIVRIAA	H	ADCYA_HUMAN	Adenylate cyclase type 10 OS=Homo sapiens GN=ADCY10 PE=1 SV=3	
738,480	399	405	K	PLVRVGV	F	YK031_HUMAN		
741,300	587	592	D	DKETCF	A	ALBU_HUMAN	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	
748,376	27	34	D	FLAEGGGV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
806,352	618	625	G	SSRGGENT	E	CC114_HUMAN	Coiled-coil domain-containing protein 114 OS=Homo sapiens GN=CCDC114 PE=2 SV=3	
824,279	419	425	H	CQLAHSS	S	P12LL_HUMAN	Putative POM121-like protein 1-like OS=Homo sapiens PE=5 SV=3	Sulfo (S)
856,324	35	41	N	DNEEGFF	S	FIBB_HUMAN	Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2	
863,403	26	34	G	DFLAEGGGV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
904,465	1093	1100	S	FINNPSLT	V	SCN3A_HUMAN	Sodium channel protein type 3 subunit alpha OS=Homo sapiens GN=SCN3A PE=1 SV=2	
904,477	27	35	D	FLAEGGGVR	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
908,450	321	327	Y	QTFNPR	T	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
920,424	25	34	E	GDFLAEGGGV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
920,468	160	167	Y	YIAMDIVP	V	TBX15_HUMAN	T-box transcription factor TBX15 OS=Homo sapiens GN=TBX15 PE=2 SV=2	
920,470	186	193	Y	YIAMDIVP	V	TBX18_HUMAN	T-box transcription factor TBX18 OS=Homo sapiens GN=TBX18 PE=2 SV=3	
941,461	1312	1319	I	HWESASLL	R	CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	
957,448	1097	1105	L	SSVTYASMI	P	CTTB2_HUMAN	Cortactin-binding protein 2 OS=Homo sapiens GN=CTTNBP2 PE=1 SV=1	
958,367	316	323	P	NNSYADFE	R	DAPK3_HUMAN	Death-associated protein kinase 3 OS=Homo sapiens GN=DAPK3 PE=1 SV=1	
970,367	34	41	V	NDNEEGFF	S	FIBB_HUMAN	Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2	
979,410	1060	1067	L	KKPNEDMC	L	BAZ2B_HUMAN	Bromodomain adjacent to zinc finger domain protein 2B OS=Homo sapiens GN=BAZ2B PE=1 SV=3	Oxidation (M)
1014,393	35	43	N	DNEEGFFSA	R	FIBB_HUMAN	Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2	
1019,504	26	35	G	DFLAEGGGVR	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1031,458	136	144	L	VSLYEAGIT	V	GTR10_HUMAN	Solute carrier family 2, facilitated glucose transporter member 10 OS=Homo sapiens GN=SLC2A10 PE=1 SV=2	Phospho (ST)
1035,487			L	DTDLEALFN	S	WWTR1_HUMAN	WW domain-containing transcription regulator protein 1 OS=Homo sapiens GN=WWTR1 PE=1 SV=1	Amidated (C-term)
1049,467	24	34	G	EGDFLAEGGGV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1054,545	1311	1319	R	IHWESASLL	R	CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
1071,440			G	ETAVCDNGAGH	C	CO2_HUMAN	Complement C2 OS=Homo sapiens GN=C2 PE=1 SV=2	
1071,443	86	94	P	VSDVEMEYI	M	GBRA4_HUMAN	Gamma-aminobutyric acid receptor subunit alpha-4 OS=Homo sapiens GN=GABRA4 PE=1 SV=2	Amidated (C-term)
1076,525	25	35	E	GDFLAEGGGVR	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1087,417	221	229	D	PADYFAPEV	Q	NALD2_HUMAN	N-acetylated-alpha-linked acidic dipeptidase 2 OS=Homo sapiens GN=NAALAD2 PE=1 SV=1	Sulfo (Y)
1096,588	89	99	D	PEVRPTSAAVA	-	APOC3_HUMAN	Apolipoprotein C-III OS=Homo sapiens GN=APOC3 PE=1 SV=1	
1106,488	23	34	S	GEGDFLAEGGGV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1112,397	295	303	S	QEYDPTEEC	L	ZN287_HUMAN	Zinc finger protein 287 OS=Homo sapiens GN=ZNF287 PE=2 SV=1	
1128,436	34	43	V	NDNEEGFFSA	R	FIBB_HUMAN	Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2	
1144,438	8899	8909	K	VSVGDSASLQC	Q	TITIN_HUMAN	Titin OS=Homo sapiens GN=TTN PE=1 SV=3	Sulfo (S)
1170,567	671	682	L	GLPGPPDVPDHA	A	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
1187,560	357	366	I	QHTPISIKNA	R	PMGT1_HUMAN	Protein O-linked-mannose beta-1,2-N-acetylglucosaminyltransferase 1 OS=Homo sapiens GN=POMGNT1 PE=1 SV=2	Sulfo (T)
1188,671	2	12	M	PVTVTRTTITT	T	MYADM_HUMAN	Myeloid-associated differentiation marker OS=Homo sapiens GN=MYADM PE=1 SV=2	
1193,520	315	324	R	TATSEYQTFI	N	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
1193,520	22	34	D	SSEGDFLAEGGGV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1193,593	354	363	R	ELLESYIDGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
1200,556	319	327	S	EYQTFNPR	T	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
1205,568	24	35	G	EGDFLAEGGGVR	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1210,646	1310	1319	H	RIHWESASLL	R	CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	
1215,502	163	172	A	EQCWNGGFIY	L	PDZD2_HUMAN	PDZ domain-containing protein 2 OS=Homo sapiens GN=PDZD2 PE=1 SV=4	
1221,555	423	434	I	PVATLSQGSSPV	T	GMEB1_HUMAN	Glucocorticoid modulatory element-binding protein 1 OS=Homo sapiens GN=GMEB1 PE=1 SV=2	Sulfo (T)
1228,488	104	114	V	TDEQEGFAEGF	V	JUN_HUMAN	Transcription factor AP-1 OS=Homo sapiens GN=JUN PE=1 SV=2	
1231,463	44	55	P	WEDEGAEDGAPG	R	CNO_HUMAN	Protein cappuccino homolog OS=Homo sapiens GN=CNO PE=1 SV=1	
1243,511	32	45	E	GAGGDGGPDAGVW	S	TEAD2_HUMAN	Transcriptional enhancer factor TEF-4 OS=Homo sapiens GN=TEAD2 PE=1 SV=2	
1255,569	290	300	G	TYMCIAENRVG	K	ROBO2_HUMAN	Roundabout homolog 2 OS=Homo sapiens GN=ROBO2 PE=1 SV=2	
1258,565	705	715	R	VGFYESDVMGR	G	A2MG_HUMAN	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=3	
1262,589	23	35	S	GEGDFLAEGGGVR	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1266,440	461	472	G	EDDSSEADSPSS	S	THA_HUMAN	Thyroid hormone receptor alpha OS=Homo sapiens GN=THRA PE=1 SV=1	Acetyl (N-term)
1269,397	413	424	T	VSSGASTATNSE	S	MUC21_HUMAN	Mucin-21 OS=Homo sapiens GN=MUC21 PE=1 SV=2	2 Phospho (ST)
1269,400			H	SSPSKSGMSAD	A	NRK_HUMAN	Nik-related protein kinase OS=Homo	2 Sulfo (T)

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
							sapiens GN=NRK PE=2 SV=2	
1269,420	413	424	T	VSSGASTATNSE	S	MUC21_HUMAN	Mucin-21 OS=Homo sapiens GN=MUC21 PE=1 SV=2	2 Phospho (ST)
1287,588	318	327	T	SEYQTFNPR	T	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
1291,407			S	PTTSTPEPDTS	T	SCMH1_HUMAN	Polycomb protein SCMH1 OS=Homo sapiens GN=SCMH1 PE=1 SV=1	2 Sulfo (T)
1297,415			E	GMDLPDSSGLF	Q	ZN304_HUMAN	Zinc finger protein 304 OS=Homo sapiens GN=ZNF304 PE=2 SV=2	2 Sulfo (S)
1297,462	1129	1141	C	GETPEDSCSEGST	A	SCN5A_HUMAN	Sodium channel protein type 5 subunit alpha OS=Homo sapiens GN=SCN5A PE=1 SV=2	
1303,597	2	12	M	SDKPDMAEIEK	F	TYB4_HUMAN	Thymosin beta-4 OS=Homo sapiens GN=TMSB4X PE=1 SV=2	Acetyl (K)
1308,547	21	34	A	DSGEGDFLAEGGGV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1324,521	77	87	L	HDSFVENEDYA	G	SNX6_HUMAN	Sorting nexin-6 OS=Homo sapiens GN=SNX6 PE=1 SV=1	
1330,527	509	524	D	DSGGGSDSGTPGPPS G	G	KANK3_HUMAN	KN motif and ankyrin repeat domain-containing protein 3 OS=Homo sapiens GN=KANK3 PE=2 SV=1	
1330,528	894	905	T	ALGMMESPMVSM	V	MD13L_HUMAN	Mediator of RNA polymerase II transcription subunit 13-like OS=Homo sapiens GN=MED13L PE=1 SV=1	3 Oxidation (M)
1346,487	328	339	C	NCSGRSEECTFD	R	LAMC3_HUMAN	Laminin subunit gamma-3 OS=Homo sapiens GN=LAMC3 PE=1 SV=3	
1349,621	22	35	D	SPEGDFLAEGGGV	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1368,499			D	SRESASSRRSS	Q	SPKAP_HUMAN	A-kinase anchor protein SPHKAP OS=Homo sapiens GN=SPHKAP PE=1 SV=1	2 Sulfo (S)
1368,520	1206	1216	D	SRESASSRRSS	Q	SPKAP_HUMAN	A-kinase anchor protein SPHKAP OS=Homo sapiens GN=SPHKAP PE=1 SV=1	2 Phospho (ST)
1371,590	219	233	T	SSTYSSSVQGGGQGA	H	RAI1_HUMAN	Retinoic acid-induced protein 1 OS=Homo sapiens GN=RAI1 PE=1 SV=2	
1379,584	20	34	T	ADSGEGDFLAEGGG V	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1379,741	2	15	M	AAPRPSAIVSVS	A	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	Acetyl (N-term)
1384,615	578	589	S	SYSKQFTSSTSY	N	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1386,835	2086	2097	Y	PTRILGNFLTKK	Y	PIEZ1_HUMAN	Piezo-type mechanosensitive ion channel component 1 OS=Homo sapiens GN=PIEZO1 PE=1 SV=4	
1387,553	803	816	V	WGEGAGAGEPEAEE	E	S12A9_HUMAN	Solute carrier family 12 member 9 OS=Homo sapiens GN=SLC12A9 PE=1 SV=1	
1387,567			Q	QASEADAAAATFR	A	GALT_HUMAN	Galactose-1-phosphate uridylyltransferase OS=Homo sapiens GN=GALT PE=1 SV=3	Sulfo (T)
1387,571			Y	GAPYAAAAAAQSYP	G	IRX6_HUMAN	Iroquois-class homeodomain protein IRX-6 OS=Homo sapiens GN=IRX6 PE=2 SV=3	Sulfo (Y)
1387,580			Q	QASEADAAAATFR	A	GALT_HUMAN	Galactose-1-phosphate uridylyltransferase OS=Homo sapiens GN=GALT PE=1 SV=3	Phospho (ST)
1388,636	317	327	A	TSEYQTFNPR	T	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
1389,616	590	601	Y	NRGDSTFESKSY	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1395,615	576	588	K	SSSYSKQFTSSTS	Y	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
1396,538	270	280	Q	QQREDEDDYEA	A	5NT1B_HUMAN	Cytosolic 5'-nucleotidase 1B OS=Homo sapiens GN=NT5C1B PE=2 SV=2	
1401,564	75	87	S	NEESGDSQQPSQP	S	PO2F1_HUMAN	POU domain, class 2, transcription factor 1 OS=Homo sapiens GN=POU2F1 PE=1 SV=2	
1410,600			C	TSAGRSDGKMWCA	T	MMP2_HUMAN	72 kDa type IV collagenase OS=Homo sapiens GN=MMP2 PE=1 SV=2	Acetyl (N-term)
1432,797	467	478	R	TLLVFEVQQPFL	F	IC1_HUMAN	Plasma protease C1 inhibitor OS=Homo sapiens GN=SERPING1 PE=1 SV=2	
1434,647	237	251	D	TTVSGDTETGQSPGV	F	EMIL2_HUMAN	EMILIN-2 OS=Homo sapiens GN=EMILIN2 PE=1 SV=3	
1434,892			L	AIQLLKRLATVAP	M	NU188_HUMAN	Nucleoporin NUP188 homolog OS=Homo sapiens GN=NUP188 PE=1 SV=1	Acetyl (N-term)
1445,689	347	360	L	VEGEIAEAAEKAT	S	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
1446,597			D	SAAVASGGAQTLAL A	G	OGFR_HUMAN	Opioid growth factor receptor OS=Homo sapiens GN=OGFR PE=1 SV=3	2 Sulfo (S)
1446,626	10	21	W	KLVSSNFDDYM	K	MYP2_HUMAN	Myelin P2 protein OS=Homo sapiens GN=PMP2 PE=1 SV=3	
1446,645	836	847	K	RDILFLFDGSAN	L	CO6A3_HUMAN	Collagen alpha-3(VI) chain OS=Homo sapiens GN=COL6A3 PE=1 SV=5	Sulfo (S)
1448,752	1308	1309	I	THRIHWESASLL	R	CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=3	
1457,725	346	359	A	LVEGEIAEAAEKA	T	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
1459,673	316	327	T	ATSEYQTFNPR	T	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
1464,648	21	35	A	DSGEGDFLAEGGGV R	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1464,658	23	37	T	TTGVSNSEEGIDVGT	L	PKHA9_HUMAN	Putative protein PLEKHA9 OS=Homo sapiens GN=PLEKHA8P1 PE=5 SV=1	
1464,680	112	125	H	GSGPSKDSFGELSR	A	APC2_HUMAN	Adenomatous polyposis coli protein 2 OS=Homo sapiens GN=APC2 PE=1 SV=1	Acetyl (N-term)
1471,647	577	589	S	SSYSKQFTSSTSY	N	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1475,668			M	WYVSMPTISSI	P	NU1M_HUMAN	NADH-ubiquinone oxidoreductase chain 1 OS=Homo sapiens GN=MT-ND1 PE=1 SV=1	Sulfo (S)
1486,617	219	232	Y	SSELNAPSQESDSH	P	RBM25_HUMAN	RNA-binding protein 25 OS=Homo sapiens GN=RBM25 PE=1 SV=3	
1486,621	1517	1530	H	SQTAGRTLTSSSPD	S	ZN831_HUMAN	Zinc finger protein 831 OS=Homo sapiens GN=ZNF831 PE=2 SV=4	Sulfo (S) / Sulfo (T)
1486,666	634	646	I	FSLGHCPNQDWLA	V	TLE2_HUMAN	Transducin-like enhancer protein 2 OS=Homo sapiens GN=TLE2 PE=1 SV=2	
1502,598	700	721		GDGGGSAGGGAGG GSGGGAGSP	G	CAD22_HUMAN	Cadherin-22 OS=Homo sapiens GN=CDH22 PE=2 SV=2	
1503,812	23	35	G	SPMYSIITPNILR	L	CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	
1514,642	448	463	P	EPAALGSSSAHCTNN G	V	CLC14_HUMAN	C-type lectin domain family 14 member A OS=Homo sapiens GN=CLEC14A PE=1 SV=1	
1517,558	105	116	G	RYYCKAENGVGV	P	MDGA1_HUMAN	MAM domain-containing glycosylphosphatidylinositol anchor protein 1 OS=Homo sapiens GN=MDGA1 PE=1 SV=1	Sulfo (Y)

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
1517,579	1049	1061	M	EEEGEEFTEGEHP	E	CC165_HUMAN	Coiled-coil domain-containing protein 165 OS=Homo sapiens GN=CCDC165 PE=1 SV=4	
1517,690			F	TDGQWGLFPGEGQQ	A	TET3_HUMAN	Methylcytosine dioxygenase TET3 OS=Homo sapiens GN=TET3 PE=2 SV=3	Amidated (C-term)
1525,680	2	15	M	GEDAAQAEKFQHPG	S	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	Acetyl (N-term)
1535,685	20	35	T	ADSGEGDFLAEGGG VR	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1539,647	201	214	D	SDQDSCRLSIDSQS	S	PAX8_HUMAN	Paired box protein Pax-8 OS=Homo sapiens GN=PAX8 PE=1 SV=2	
1539,659	1079	1092	G	ETDPTTIFPCKEKG	K	TREF1_HUMAN	Transcriptional-regulating factor 1 OS=Homo sapiens GN=TRERF1 PE=1 SV=1	
1542,710			T	YRVAVSRTQDTPA	F	SIPA1_HUMAN	Signal-induced proliferation-associated protein 1 OS=Homo sapiens GN=SIPA1 PE=1 SV=1	Sulfo (Y)
1544,605	21	35	A	DSGEGDFLAEGGGV R	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Sulfo (S)
1544,612	495	508	I	LQENGCLSDSDMFS	Q	LPPRC_HUMAN	Leucine-rich PPR motif-containing protein, mitochondrial OS=Homo sapiens GN=LRPPRC PE=1 SV=3	
1544,614	21	35	A	DSGEGDFLAEGGGV R	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Phospho (ST)
1549,602	200	215	S	SSGSASDSEESPEGQP	V	ZN653_HUMAN	Zinc finger protein 653 OS=Homo sapiens GN=ZNF653 PE=1 SV=1	
1557,666	1703	1723	A	EGGGSAGGGGAGA GVASGPE	L	SHAN1_HUMAN	SH3 and multiple ankyrin repeat domains protein 1 OS=Homo sapiens GN=SHANK1 PE=1 SV=2	
1558,679	576	589	K	SSSYSKQFTSSTSY	N	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1558,773	346	360	A	LVEGEIAEAAEKAT	S	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
1560,721	315	327	R	TATSEYQTFNPR	T	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
1582,549	5	26	S	QGGGGNSGGGGCGG GGSSGGCG	T	LRCH2_HUMAN	Leucine-rich repeat and calponin homology domain-containing protein 2 OS=Homo sapiens GN=LRCH2 PE=2 SV=2	
1592,711	513	527	R	HPDEAAFFDTASTGK	T	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1611,873	630	644	L	QGAKIPKPEASFSPR	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
1615,630	394	408	R	TAMLDGNYSDGEVND	L	GBRR1_HUMAN	Gamma-aminobutyric acid receptor subunit rho-1 OS=Homo sapiens GN=GABRR1 PE=2 SV=2	Oxidation (M)
1615,642	20	35	T	ADSGEGDFLAEGGG VR	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Sulfo (S)
1615,652	20	35	T	ADSGEGDFLAEGGG VR	G	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Phospho (ST)
1615,794	344	359	A	SALVEGEIAEAAEK A	T	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
1616,896	23	36	G	SPMYSIITPNILRL	E	CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	
1624,880	1338	1351	N	GFKSHALQLNRRQI	R	CO4A_HUMAN	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1	
1625,805	85	99	F	WDLDPVVRPTSAVA A	-	APOC3_HUMAN	Apolipoprotein C-III OS=Homo sapiens GN=APOC3 PE=1 SV=1	
1625,823	499	513	P	QLSSARPSAQRNSQ	P	PKHG1_HUMAN	Pleckstrin homology domain-containing family G member 1 OS=Homo sapiens GN=PLEKHG1 PE=1 SV=2	
1625,839	1337	1350	R	NGFKSHALQLNRRQ	I	CO4A_HUMAN	Complement C4-A OS=Homo	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
							sapiens GN=C4A PE=1 SV=1	
1625,840	1337	1350	R	NGFKSHALQLNNRQ	I	CO4B_HUMAN	Complement C4-B OS=Homo sapiens GN=C4B PE=1 SV=1	
1629,810	345	360	S	ALVEGEIAEEAAEKAT	S	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
1630,818	404	417	R	MSLSSFSVNRPFLLF	F	A2AP_HUMAN	Alpha-2-antiplasmin OS=Homo sapiens GN=SERPINF2 PE=1 SV=3	
1636,867	254	267	F	LSALEEYTKKLNTQ	-	APOA1_HUMAN	Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	
1648,751	590	603	Y	NRGDSTFESKSYKM	A	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1685,899	99	113	P	EEHPVLLTEAPLNPK	A	ACTB_HUMAN	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1	
1691,542			G	GEYDDQTYETYDN	S	KHDR2_HUMAN	KH domain-containing, RNA-binding, signal transduction-associated protein 2 OS=Homo sapiens GN=KHDRBS2	Sulfo (T)
1710,978	344	360	A	SALVEGEIAEEAAEKAT	S	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
1716,842	344	360	A	SALVEGEIAEEAAEKAT	S	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
1719,789	590	604	Y	NRGDSTFESKSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1721,890			N	LIIVGAGVIMACFYNN	I	S38A9_HUMAN	Plasma protease C1 inhibitor OS=Homo sapiens GN=SERPING1 PE=1 SV=2	Acetyl (N-term)
1724,957	629	644	Y	LQGAKIPKPEASFSPR	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
1725,880	667	683	S	SRQLGLPGPPDVPDHA	Y	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
1735,775	2695	2713	V	TAEVAAPSTSSSATSPEG	P	SRCAP_HUMAN	Helicase SRCAP OS=Homo sapiens GN=SRCAP PE=1 SV=3	
1735,784	590	604	Y	NRGDSTFESKSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Oxidation (M)
1736,894	23	38	A	AEDDLPTVELQGVVPR	G	F13A_HUMAN	Coagulation factor XIII A chain OS=Homo sapiens GN=F13A1 PE=1 SV=4	
1738,920	1337	1351	R	NGFKSHALQLNNRQI	R	CO4B_HUMAN	Complement C4-B OS=Homo sapiens GN=C4B PE=1 SV=1	
1738,923	1337	1351	R	NGFKSHALQLNNRQI	R	CO4A_HUMAN	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1	
1757,776	511	526	R	HRHPDEAAFFDTASTG	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1776,669	377	390	K	PFSCSVCSKSFNRR	E	ZN865_HUMAN	Zinc finger protein 865 OS=Homo sapiens GN=ZNF865 PE=3 SV=1	Sulfo (S)
1776,963	1305	1319	S	SKITHRIHWESASLL	R	CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	
1780,819	2	16	M	SDKPDMAEIEKFDKS	K	TYB4_HUMAN	Thymosin beta-4 OS=Homo sapiens GN=TMSB4X PE=1 SV=2	Acetyl (K)
1784,749	583	598	Q	FTSSTSYNRGDSTFES	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1785,847	671	687	L	GLPGPPDVPDHAAYHPF	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
1787,879	343	360	L	ASALVEGEIAEEAAEKAT	S	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
1790,764	127	142	Y	PGMQNERHGQSFVDS	S	CHD7_HUMAN	Chromodomain-helicase-DNA-binding protein 7 OS=Homo sapiens GN=CHD7 PE=1 SV=3	Oxidation (M)
1793,895	54	71	R	EEAPSLRPAPPISGGY	R	FIBB_HUMAN	Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
1803,874	344	361	A	SALVEGEIAEEAAEK ATS	R	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
1814,884	2	16	M	AQQAADKYLYVDK NF	I	MYH9_HUMAN	Myosin-9 OS=Homo sapiens GN=MYH9 PE=1 SV=4	Acetyl (N-term)
1815,925	168	184	F	SVVPSPKVSDTVVEP YN	A	TBB4B_HUMAN	Tubulin beta-4B chain OS=Homo sapiens GN=TUBB4B PE=1 SV=1	
1815,930	168	184	F	SVVPSPKVSDTVVEP YN	A	TBB3_HUMAN	Tubulin beta-3 chain OS=Homo sapiens GN=TUBB3 PE=1 SV=2	
1818,938	249	266	A	NTQPRGPPASSPAPA PKF	S	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
1822,906	349	363	E	DKTERELLESYIDGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
1827,791	586	601	S	STSYNRGDSTFESKS Y	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1852,808	584	600	F	TSSTSYNRGDSTFES KS	Y	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1854,848	1321	1337	R	SEETKENEGFTVTAE GK	G	CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	
1858,779	2	18	M	GEDAAQAEKFQHPG SDM	R	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	Acetyl (N-term)
1863,995	1304	1319	R	SSKITHRIHWESASLL	R	CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	
1874,890	221	237	Q	EILESEEKGDPNKPS GF	R	PDLI1_HUMAN	PDZ and LIM domain protein 1 OS=Homo sapiens GN=PDLIM1 PE=1 SV=4	
1885,871	511	527	R	HRHPDEAAFFDTAST GK	T	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
1886,897	186	204	F	ESRAQLGGPEAAKS DETA	K	HSPB1_HUMAN	Heat shock protein beta-1 OS=Homo sapiens GN=HSPB1 PE=1 SV=2	
1888,021	628	644	Y	YLQGAKIPKPEASFS PR	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
1895,020	585	600	S	CSEKSHRVDPPQVKV KR	H	AFIL1_HUMAN	Actin filament-associated protein 1-like 1 OS=Homo sapiens GN=AFAPIL1 PE=1 SV=2	
1895,020	1337	1352	R	NGFKSHALQLNNRQI R	G	CO4B_HUMAN	Complement C4-B OS=Homo sapiens GN=C4B PE=1 SV=1	
1895,024	1337	1352	R	NGFKSHALQLNNRQI R	G	CO4A_HUMAN	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1	
1921,016	343	360	V	RSPGAPGPLTLKEVE ELE	Q	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
1921,975	21	38	S	NAAEDDLPTVELQG VVPR	G	F13A_HUMAN	Coagulation factor XIII A chain OS=Homo sapiens GN=F13A1 PE=1 SV=4	
1923,917	2	20	M	AGSVADSDAVVKLD DGHLN	N	LEGL_HUMAN	Galectin-related protein OS=Homo sapiens GN=LGALS1 PE=1 SV=2	Acetyl (N-term)
1929,640			L	STLSGGIFGGQGNSL LC	S	WDR87_HUMAN	WD repeat-containing protein 87 OS=Homo sapiens GN=WDR87 PE=2 SV=3	4 Phospho (ST)
1929,651			E	GETEPTSTKDPRTG SA	G	MRP6_HUMAN	Multidrug resistance-associated protein 6 OS=Homo sapiens GN=ABCC6 PE=1 SV=2	3 Sulfo (T)
1949,996	53	71	K	REEAPSLRPAPPISG GGY	R	FIBB_HUMAN	Fibrinogen beta chain OS=Homo sapiens GN=FBG PE=1 SV=2	
1950,920	407	424	A	GNAQTSACLIVPKPA IPS	S	DCC_HUMAN	Netrin receptor DCC OS=Homo sapiens GN=DCC PE=1 SV=2	2 Phospho (ST)
1951,949	348	363	L	EDKTERELLESYIDG R	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
1953,060	96	113	R	VAPEEHPVLLTEAPL NPK	A	ACTB_HUMAN	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1	
1958,050	390	406	Y	VKVTSIQDWVQKTIA EN	-	HPT_HUMAN	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1	
1983,970	27	44	K	NPLPSKETIEQEKQA GES	-	TMSL3_HUMAN	Thymosin beta-4-like protein 3 OS=Homo sapiens GN=TMSL3 PE=2 SV=1	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
1999,876	583	600	Q	FTSSTSYNRGDSTFESKS	Y	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2000,870	576	593	K	SSSYSKQFTSSTSYNRGD	S	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2014,890	578	595	S	SYSKQFTSSTSYNRGDST	F	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2015,870	584	601	F	TSSSTSYNRGDSTFESKSY	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2020,097	1304	1320	R	SSKITHRIHWESASLLR	S	CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	
2022,000	2	18	M	SDKPDMAEIEKFDKSKL	K	TYB4_HUMAN	Thymosin beta-4 OS=Homo sapiens GN=TMSB4X PE=1 SV=2	Acetyl (N-term)
2026,990	669	687	R	QLGLPGPPDVPDHAAYHPF	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
2030,991	221	238	Q	EILESEEKGDPNKPSGFR	S	PDLI1_HUMAN	PDZ and LIM domain protein 1 OS=Homo sapiens GN=PDLIM1 PE=1 SV=4	
2040,040	129	147	A	ALLSPYSYSTTAVVTNPKE	-	TTHY_HUMAN	Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1	
2051,080	627	644	K	YYLQGAKIPKPEASFSR	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
2059,701			S	DFGYLTCETYLPGSYGD	L	COLQ_HUMAN	Acetylcholinesterase collagenic tail peptide OS=Homo sapiens GN=COLQ PE=1 SV=2	2 Sulfo (Y)
2062,070	183	200	A	QPPSSSLVIDKESEVYKML	Q	PDLI1_HUMAN	PDZ and LIM domain protein 1 OS=Homo sapiens GN=PDLIM1 PE=1 SV=4	
2065,030	347	363	S	LEDKTERELLESYIDGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
2078,090	52	71	K	KREEAPSLRPAPPPIGGGY	R	FIBB_HUMAN	Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2	
2082,086	663	683	P	GVLSSRQLGLPGPPDVPDHA	Y	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
2101,919	577	595	S	SSYSKQFTSSTSYNRGDST	F	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2107,060	1429	1447	K	DDPDAPLQPVTPLQLFEGR	R	CO4A_HUMAN	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1	
2107,060	1429	1447	K	DDPDAPLQPVTPLQLFEGR	R	CO4B_HUMAN	Complement C4-B OS=Homo sapiens GN=C4B PE=1 SV=1	
2112,050			L	VKQEGRLRFLEQELETIT	I	PLTP_HUMAN	Phospholipid transfer protein OS=Homo sapiens GN=PLTP PE=1 SV=1	Phospho (ST)
2112,070	26	44	E	KNPLPSKETIEQEKQAGES	-	TMSL3_HUMAN	Thymosin beta-4-like protein 3 OS=Homo sapiens GN=TMSL3 PE=2 SV=1	
2113,955	204	225	T	STLSGGGGGATADSLGGKEEK	E	HXC6_HUMAN	Homeobox protein Hox-C6 OS=Homo sapiens GN=HOXC6 PE=2 SV=3	Sulfo (S)
2121,855	600	619	K	SYKMADEAGSEADHEGTHST	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2127,930	582	600	K	QFTSSTSYNRGDSTFESKS	Y	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2140,051	511	531	F	SRPGGRASVDTKEAEGAPQVE	A	CALD1_HUMAN	Caldesmon OS=Homo sapiens GN=CALD1 PE=1 SV=2	
2147,150	1431	1449	D	PDAPLQPVTPLQLFEGR	R	CO4A_HUMAN	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1	
2149,990			V	AFLLIQWYIYGFSLSAV	Y	CXA1_HUMAN	Gap junction alpha-1 protein OS=Homo sapiens GN=GJA1 PE=1 SV=2	Phospho (ST)
2152,065	346	363	K	SLEDKTERELLESYIDGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
2157,960	586	604	S	STSSTYNRGDSTFESKSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2162,940	583	601	Q	FTSSTSYNRGDSTFESKSY	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
2186,310			W	LKKARVVRLGEVPY KMVKG	F	XPC_HUMAN	DNA repair protein complementing XP-C cells OS=Homo sapiens GN=XPC PE=1 SV=4	Oxidation (M)
2187,082	36	56	K	VNPFPRGDSEPPAP GAQRAQ	M	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
2188,951	576	595	K	SSSYSKQFTSSTSYN RGDST	F	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2214,990	21	40	A	SEAEDASLLSFMQGY MKHAT	K	APOC3_HUMAN	Apolipoprotein C-III OS=Homo sapiens GN=APOC3 PE=1 SV=1	
2241,112	25	44	Q	EKNPLPSKETIEQEK QAGES	-	TMSL3_HUMAN	Thymosin beta-4-like protein 3 OS=Homo sapiens GN=TMSL3 PE=2 SV=1	
2243,110	233	250	Q	EQARFWNVRTFFVL SQES	S	NB5R5_HUMAN	NADH-cytochrome b5 reductase-like OS=Homo sapiens GN=CYB5RL PE=1 SV=3	
2245,200	249	267	S	FKVSFLSALEEYTKK LNTQ	-	APOA1_HUMAN	Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	
2253,012	192	209	N	SKIAFKIIRQEPSDSP MF	I	DSG1_HUMAN	Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2	Sulfo (S)
2270,123	667	687	S	SRQLGLPGPPDVPDH AAYHPF	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
2280,160	345	363	K	KSLEDKTERELLESYI DGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
2280,220	256	278	P	PASSPAPAPKFSPVTP KFTPVAS	K	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
2283,190	50	69	C	LEEKGPMEFELLPGES NKIPR	L	SRGN_HUMAN	Serglycin OS=Homo sapiens GN=SRGN PE=1 SV=3	
2288,060	511	531	R	HRHPDEAAFFDTAST GKTFPG	F	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2291,001	584	603	F	TSSTSYNRGDSTFES KSYKM	A	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Oxidation (M)
2308,080			G	DPEDAPPPGSPQA LAFPEHGS	V	OPLA_HUMAN	5-oxoprolinase OS=Homo sapiens GN=OPLAH PE=1 SV=3	Amidated (C-term)
2317,240	746	764	L	FQVLPWLKEKLQDE DLGFL	-	CFAB_HUMAN	Complement factor B OS=Homo sapiens GN=CFB PE=1 SV=2	
2318,130	1861	1882	S	RCVACMVDSSLGVW VTLKGSAAH	V	ARHGH_HUMAN	Rho guanine nucleotide exchange factor 17 OS=Homo sapiens GN=ARHGEF17 PE=1 SV=1	
2323,120	178	198	S	GAVEGKEELPDENK SLEETLH	T	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
2323,120			R	SCQKSPAQQEPPQRQ TVGTK	Q	AFF1_HUMAN	AF4/FMR2 family member 1 OS=Homo sapiens GN=AFF1 PE=1 SV=1	2 Acetyl (K); Acetyl (N-term)
2324,050	135	155	R	SLDRNLPSDSQDLGQ HGLEED	F	SRGN_HUMAN	Serglycin OS=Homo sapiens GN=SRGN PE=1 SV=3	
2345,750	810	830	S	TEGSEMTTVSTTGAE TTTDST	E	MUC22_HUMAN	Mucin-22 OS=Homo sapiens GN=MUC22 PE=1 SV=1	Sulfo (T)
2346,043	584	604	F	TSSTSYNRGDSTFES KSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2349,815			G	YGHGYGGLGCGYGR GYGGYGYG	C	KR202_HUMAN	Keratin-associated protein 20-2 OS=Homo sapiens GN=KRTAP20-2 PE=2 SV=1	2 Sulfo (Y)
2352,119	282	302	G	EKTPKDESANQEEPE ARVPAQ	S	VASP_HUMAN	Vasodilator-stimulated phosphoprotein OS=Homo sapiens GN=VASP PE=1 SV=3	
2357,160	666	687	L	SSRQLGLPGPPDVPD HAAYHPF	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
2362,050	605	626	S	DGQGGYRVGVDGLL VTDAQPEH	S	SEM4G_HUMAN	Semaphorin-4G OS=Homo sapiens GN=SEMA4G PE=2 SV=1	Sulfo (Y)
2369,171	24	44	T	QEKPLPSKETIEQE KQAGES	-	TMSL3_HUMAN	Thymosin beta-4-like protein 3 OS=Homo sapiens GN=TMSL3 PE=2 SV=1	
2374,130	656	675	N	SKPLGPLQTLMENLS SNRFY	S	NRK_HUMAN	Nik-related protein kinase OS=Homo sapiens GN=NRK PE=2 SV=2	Sulfo (Y)
2377,200	1429	1449	K	DDPDAPLQPVTPLQL	R	CO4B_HUMAN	Complement C4-B OS=Homo	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
				FEGRRN			sapiens GN=C4B PE=1 SV=1	
2377,203	1429	1449	K	DDPDAPLQPVTPPLQL FEGRRN	R	CO4A_HUMAN	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1	
2389,250	398	418	L	MIEQNTKSPLFMGK VVNPTQK	-	A1AT_HUMAN	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	
2393,116			V	PEPEVTSIVATVVVDN ENTVTIS	T	PR40A_HUMAN	Pre-mRNA-processing factor 40 homolog A OS=Homo sapiens GN=PRPF40A PE=1 SV=2	Sulfo (T)
2408,250	344	363	E	KKSLEDKTERELLES YIDGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
2420,130	448	470	C	ESQGEVQPNFSTSS PLLSSALN	T	ASXL2_HUMAN	Putative Polycomb group protein ASXL2 OS=Homo sapiens GN=ASXL2 PE=1 SV=1	
2420,139			K	DSHSLTTNIMEILRG DFSSANN	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Amidated (C-term)
2421,070	602	624	Y	KMADEAGSEADHEG THSTKRGHA	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2421,123	101	122	K	DSHSLTTNIMEILRG DFSSANN	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2422,070	583	603	Q	FTSSTSYNRGDSTFES KSYKM	A	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2424,250	168	190	F	SVVPSPKVSDTVVEP YNATLSVH	Q	TBB4B_HUMAN	Tubulin beta-4B chain OS=Homo sapiens GN=TUBB4B PE=1 SV=1	
2437,118	101	122	K	DSHSLTTNIMEILRG DFSSANN	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Oxidation (M)
2438,070			Q	FTSSTSYNRGDSTFES KSYKM	A	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Oxidation (M)
2439,120	218	240	K	QVQDLLSEDHGCSLP ATPQISD	L	CC149_HUMAN	Coiled-coil domain-containing protein 149 OS=Homo sapiens GN=CCDC149 PE=1 SV=2	
2461,270	247	267	L	ESFKVSFLSALEEYT KKLNTQ	-	APOA1_HUMAN	Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	
2465,060	577	598	S	SSYSKQFTSSTSYNR GDSTFES	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2470,219	23	44	E	TQEKNPSPKETIEQE KQAGES	-	TMSL3_HUMAN	Thymosin beta-4-like protein 3 OS=Homo sapiens GN=TMSL3 PE=2 SV=1	
2470,230	271	291	T	GQAFLYHQVRCMM AILFLIGQ	G	PUS3_HUMAN	tRNA pseudouridine(38/39) synthase OS=Homo sapiens GN=PUS3 PE=1 SV=3	2 Oxidation (M)
2471,120	135	156	R	SLDRNLPSDSQDLGQ HGLEEDF	M	SRGN_HUMAN	Serglycin OS=Homo sapiens GN=SRGN PE=1 SV=3	
2476,022	530	553	I	TSPTPTKEQPKSTAS GSSGESMDS	V	RGL1_HUMAN	Ral guanine nucleotide dissociation stimulator-like 1 OS=Homo sapiens GN=RGL1 PE=1 SV=1	Sulfo (T)
2476,981	258	281	I	NADAYDSIEGPCNE EAAAPEAPE	N	CI082_HUMAN	Uncharacterized protein C9orf82 OS=Homo sapiens GN=C9orf82 PE=1 SV=2	
2482,310	660	683	N	FRPGVLSSRQLGLPG PPDVPDHAA	Y	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
2491,310	402	423	M	IIVPTDTQNIFFMSKV TNPQQA	-	AACT_HUMAN	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2	
2493,110	583	604	Q	FTSSTSYNRGDSTFES KSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2495,990			S	DENCFEDAVYWEMP EGNSTVY	I	RHCG_HUMAN	Ammonium transporter Rh type C OS=Homo sapiens GN=RHCG PE=1 SV=1	Amidated (C-term)
2496,949			S	DQETGEQVSSPSSHP ALHTTTE	D	NIBL1_HUMAN	Niban-like protein 1 OS=Homo sapiens GN=FAM129B PE=1 SV=3	2 Sulfo (T)
2496,972			G	MQYYLSSQDFDLSL QRQES	S	MUL1_HUMAN	Mitochondrial ubiquitin ligase activator of NFKB 1 OS=Homo sapiens GN=MUL1 PE=1 SV=1	2 Sulfo (Y)
2502,334	397	418	F	LMIEQNTKSPLFMGK VVNPTQK	-	A1AT_HUMAN	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	
2506,125	579	600	S	YSKQFTSSTSYNRGD STFESKS	Y	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
2508,180	611	633	E	MVVEAAQKANAHG FIMELQDGYS	T	ABCB9_HUMAN	ATP-binding cassette sub-family B member 9 OS=Homo sapiens GN=ABCB9 PE=1 SV=1	
2508,280	2	22	M	SDKPDMAEIEKFDKS KLKKTTE	T	TYB4_HUMAN	Thymosin beta-4 OS=Homo sapiens GN=TMSB4X PE=1 SV=2	Acetyl (K)
2509,107	583	604	Q	FTSSTSYNRGDSTFES KSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Oxidation (M)
2519,260	89	111	K	AIQLTYNPDESSKPN MIDAATLK	S	FIBG_HUMAN	Fibrinogen gamma chain OS=Homo sapiens GN=FGG PE=1 SV=3	
2528,260	68	90	R	NLDFLKAVDTNRAS VGQDSPEPR	S	PXDC2_HUMAN	Plexin domain-containing protein 2 OS=Homo sapiens GN=PLXDC2 PE=1 SV=1	
2530,270	15	38	A	VPPNNSNAEDDLPT VELQGVVPR	G	F13A_HUMAN	Coagulation factor XIII A chain OS=Homo sapiens GN=F13A1 PE=1 SV=4	
2537,300	343	363	F	EKKSLIEDKTERELLE SYIDGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
2552,094	576	598	K	SSSYSKQFTSSTSYN RGDSTFES	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2558,200	590	612	C	PESAMEMNEPQAPA PSLFRQCLL	W	SC5A2_HUMAN	Sodium/glucose cotransporter 2 OS=Homo sapiens GN=SLC5A2 PE=1 SV=1	
2577,220	101	123	K	DSHSLTTNIMEILRG DFSSANNR	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2581,333	617	639	R	NVHSGSTFFKYYLQ GAKIPKPEA	S	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
2590,056	36	57	Y	LESCFSAKNGTPRQP SICSYSR	A	NAP1_HUMAN	Nef-associated protein 1 OS=Homo sapiens GN=C9orf156 PE=1 SV=2	Sulfo (S)
2590,390	253	278	P	RGPPASSPAPAPKFSP VTPKFTPVAS	K	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
2593,160	578	600	S	SYSKQFTSSTSYNRG DSTFESKS	Y	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2601,303	14	38	R	AVPPNNSNAEDDL PTVELQGVVPR	G	F13A_HUMAN	Coagulation factor XIII A chain OS=Homo sapiens GN=F13A1 PE=1 SV=4	
2602,161	135	157	R	SLDRNLPSDSQDLGQ HGLEEDFM	L	SRGN_HUMAN	Serglycin OS=Homo sapiens GN=SRGN PE=1 SV=3	
2609,330	2	23	M	SDKPDMAEIEKFDKS KLKKTET	Q	TYB4_HUMAN	Thymosin beta-4 OS=Homo sapiens GN=TMSB4X PE=1 SV=2	Acetyl (N-term)
2611,080	1213	1241	S	SQGSSVDAGCGTGSS GSACATPVESPLCP	S	FRPD4_HUMAN	FERM and PDZ domain-containing protein 4 OS=Homo sapiens GN=FRMPD4 PE=1 SV=1	
2615,350	10	34	V	LGTGLTECILSGIMSV NGKKVLHMD	R	GDIA_HUMAN	Rab GDP dissociation inhibitor alpha OS=Homo sapiens GN=GDI1 PE=1 SV=2	
2618,151	1486	1510	C	NSEPKELAGDPEAVP ESDEPPPGGE	Q	MYST4_HUMAN	Histone acetyltransferase MYST4 OS=Homo sapiens GN=MYST4 PE=1 SV=3	
2618,160			R	SLDRNLPSDSQDLGQ HGLEEDFM	L	SRGN_HUMAN	Serglycin OS=Homo sapiens GN=SRGN PE=1 SV=3	Oxidation (M)
2626,329	663	687	P	GVLSSRQLGLPGPPD VPDHAAYHPF	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
2642,060	21	45	D	PALVSSTPESGVFFPS GPEGLDAAA	S	GATA1_HUMAN	Erythroid transcription factor OS=Homo sapiens GN=GATA1 PE=1 SV=1	Phospho (ST)
2646,330	65	91	S	AAPVPTAPAAGAPL MDFGNDFVPPAPR	G	RTN4_HUMAN	Reticulon-4 OS=Homo sapiens GN=RTN4 PE=1 SV=2	
2646,330	65	91	S	AAPVPTAPAAGAPL MDFGNDFVPPAPR	G	RTN4_HUMAN	Reticulon-4 OS=Homo sapiens GN=RTN4 PE=1 SV=2	
2658,250	605	629	A	DEAGSEADHEGTHS TKRGHAKSRPV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2665,400	168	192	F	SVVPSPKVSDTVVEP YNATLSVHQTL	V	TBB4B_HUMAN	Tubulin beta-4B chain OS=Homo sapiens GN=TUBB4B PE=1 SV=1	
2668,360	617	640	R	NVHSGSTFFKYYLQ GAKIPKPEAS	F	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
2671,168	600	624	K	SYKMADEAGSEADH EGTHSTKRKGA	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2679,200			K	SSSYKQFTSSTS RGDSTFESK	S	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Amidated (C-term)
2680,189	577	600	S	SSYSKQFTSSTS RGDSTFESKS	Y	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2681,170	740	762	F	DDSDVQQLSEDEV TQTAYILFY	Q	UBP31_HUMAN	Ubiquitin carboxyl-terminal hydrolase 31 OS=Homo sapiens GN=USP31 PE=1 SV=2	
2684,370	342	363	L	FEKKSLEDKTERE ESYIDGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
2687,340	36	61	K	VNPFPRPGDSEPP GAQRAQMGRVG	E	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
2693,358	146	169	D	LEIDSLSSLLDDM NDPFKARVS	S	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
2715,240	135	158	R	SLDRNLPSDSQDL HGLEDFML	-	SRGN_HUMAN	Serglycin OS=Homo sapiens GN=SRGN PE=1 SV=3	
2723,382	662	687	R	PGVLSRQLGLPG DVPDHAAYHPF	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
2725,300	409	433	H	FGLEGDEESTMLE VSPKKSTVLQ	Q	TLN1_HUMAN	Talin-1 OS=Homo sapiens GN=TLN1 PE=1 SV=3	
2727,390	658	683	R	MNFRPGVLSSRQ LGLPGPPDVPDHA	Y	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
2729,180	2	24	M	TEKAPEPHVEEDD DELDSKLNLY	K	GDIR2_HUMAN	Rho GDP-dissociation inhibitor 2 OS=Homo sapiens GN=ARHGDI2 PE=1 SV=3	Acetyl (K)
2737,210	574	598	R	GKSSYSKQFTSST YNRGDSTFES	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2752,430	25	48	R	DAHKSEVAHRFK DLGEENFKALVL	I	ALBU_HUMAN	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	
2756,220	578	601	S	SYSKQFTSSTS DSTFESKSY	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2757,405	13	38	R	RAVPPNNSNAEED LPTVELQGVVPR	G	F13A_HUMAN	Coagulation factor XIII A chain OS=Homo sapiens GN=F13A1 PE=1 SV=4	
2758,290	21	45	A	SEAEDASLLSFMQ GYMKHATKTAKD	A	APOC3_HUMAN	Apolipoprotein C-III OS=Homo sapiens GN=APOC3 PE=1 SV=1	
2763,331	400	425	Y	ALTSEEAERSDGP VQPAVLQVHQTS	-	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
2767,221	576	600	K	SSSYKQFTSSTS RGDSTFESKS	Y	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2767,242	213	239	K	GLNEEVITGSGK NSWKSEAEGESSDT	Q	DDX4_HUMAN	Probable ATP-dependent RNA helicase DDX4 OS=Homo sapiens GN=DDX4 PE=2 SV=2	
2768,220			D	EAVKFDMGAYKSS KDDAKITVIL	R	IL1A_HUMAN	Interleukin-1 alpha OS=Homo sapiens GN=IL1A PE=1 SV=1	2 Phospho (ST); Phospho (Y)
2770,480	244	267	L	PVLESFKVSFLS ALEEYTKKLNTQ	-	APOA1_HUMAN	Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	
2783,460	196	219	K	KVYVGGLESRVL KYTAQNMELQNK	V	CREB3_HUMAN	Cyclic AMP-responsive element-binding protein 3 OS=Homo sapiens GN=CREB3 PE=1 SV=1	Oxidation (M)
2805,137	297	320	E	SSPLPKEVNDSP RAADKSPESQNL	I	ZN532_HUMAN	Zinc finger protein 532 OS=Homo sapiens GN=ZNF532 PE=1 SV=2	Sulfo (S)
2805,170	983	1007	I	SEHVANYEQCIFG PSFDHNSNGNPEQ	N	ANR31_HUMAN	Ankyrin repeat domain-containing protein 31 OS=Homo sapiens GN=ANKRD31 PE=5 SV=2	
2813,089	1061	1083	S	AYERSMMSAYERS MSPMADRSM	M	SON_HUMAN	Protein SON OS=Homo sapiens GN=SON PE=1 SV=4	6 Oxidation (M)
2815,320	548	574	R	GSESGIFTNTKES SSHHPGIAEFPSRG	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2828,390			Q	QYTQKLFIPNIT TKNSGSYACHTTN	S	CEAM8_HUMAN	Carcinoembryonic antigen-related cell adhesion molecule 8 OS=Homo sapiens GN=CEACAM8 PE=1 SV=2	Amidated (C-term)
2828,400	20	44	K	KTETQEKNPSPK ETI	-	TMSL3_HUMAN	Thymosin beta-4-like protein 3	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
				EQEKQAGES			OS=Homo sapiens GN=TMSL3 PE=2 SV=1	
2828,410			A	GELSRQLEEKESIVS QLSRKQAF	T	MYH3_HUMAN	Myosin-3 OS=Homo sapiens GN=MYH3 PE=1 SV=3	Phospho (ST)
2829,330	21	46	A	SEAEDASLLSFMQGY MKHATKTAKDA	L	APOC3_HUMAN	Apolipoprotein C-III OS=Homo sapiens GN=APOC3 PE=1 SV=1	
2834,033	1072	1095	S	MSQPSQMSQPGLSQP ELSQDSYLG	D	RENT1_HUMAN	Regulator of nonsense transcripts 1 OS=Homo sapiens GN=UPF1 PE=1 SV=2	Sulfo (S)
2843,250	577	601	S	SSYSKQFTSSTSYNR GDSTFESKSY	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2860,330	603	629	K	MADEAGSEADHEGT HSTKRGHAKSRPV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2866,430	2	25	M	SDKPDMAEIEKFDKS KLKKTETQE	K	TYB4_HUMAN	Thymosin beta-4 OS=Homo sapiens GN=TMSB4X PE=1 SV=2	Acetyl (K)
2868,260	893	919	R	HLSQDTGSPSGMRP WEDLPSQDTGSPS	R	FA5_HUMAN	Coagulation factor V OS=Homo sapiens GN=F5 PE=1 SV=4	
2879,480	661	687	F	RPGVLSRRLGLPGP PDVDPHAAAYHPF	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
2904,330	296	320	K	SVPGLNVDMEEEEE EDIDLVLKHL	R	SYTL4_HUMAN	Synaptotagmin-like protein 4 OS=Homo sapiens GN=SYTL4 PE=1 SV=2	
2916,550	250	278	N	TQPRGPPASSPAPAP KFSPVTPKFTPVAS	K	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
2917,480			I	QTVSGILQGPFDLGN QLLGLKGVMEMM	C	UN45B_HUMAN	Protein unc-45 homolog B OS=Homo sapiens GN=UNC45B PE=2 SV=1	Acetyl (K)
2922,440	237	262	Q	VEKETQPGGQNVVF VDEGEMIFKKNT	R	SYTL4_HUMAN	Synaptotagmin-like protein 4 OS=Homo sapiens GN=SYTL4 PE=1 SV=2	
2928,320	579	603	S	YSKQFTSSTSYNRGD STFESKSYKM	A	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2930,284	576	601	K	SSYSKQFTSSTSYNR RGDSTFESKSY	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2945,340	574	600	R	GKSSYSKQFTSSTSYNR GDSTFESKSY	Y	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2945,410	297	322	T	FGRKKVREGESHAE NETKSEDLPSE	Q	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
2973,415	141	167	G	GGTGSGMGTLMLNK IREEYPDRIMNSF	S	TBB1_HUMAN	Tubulin beta-1 chain OS=Homo sapiens GN=TUBB1 PE=1 SV=1	
2983,398	466	491	R	GDKLFGPDLKLVPP MEEDYPQFGSPK	-	A2AP_HUMAN	Alpha-2-antiplasmin OS=Homo sapiens GN=SERPINF2 PE=1 SV=3	Sulfo (Y)/ Sulfo (S)
2983,410	466	491	R	GDKLFGPDLKLVPP MEEDYPQFGSPK	-	A2AP_HUMAN	Alpha-2-antiplasmin OS=Homo sapiens GN=SERPINF2 PE=1 SV=3	Phospho (ST)
2988,420	602	629	Y	KMADEAGSEADHEG THSTKRGHAKSRPV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
2999,360	579	604	S	YSKQFTSSTSYNRGD STFESKSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3008,501	143	169	S	SIDLEIDSLSSLLDDM TKNDPFKARVS	S	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
3013,430	398	425	G	SYALTSEEAEERSDGD PVQPAVLQVHQT	-	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
3015,356	579	604	S	YSKQFTSSTSYNRGD STFESKSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Oxidation (M)
3015,360	78	107	P	NLTESIPSPPPQPPQ GPSQSPLSSCCSSF	S	MAGC2_HUMAN	Melanoma-associated antigen C2 OS=Homo sapiens GN=MAGEC2 PE=1 SV=1	
3026,550	660	687	N	FRPGVLSRRLGLPG PPDVPDHAAYHPF	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
3030,590	249	278	A	NTQPRGPPASSPAPAP KFSPVTPKFTPVAS	K	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
3044,470	936	960	Q	AEKKKMQQNIQELE EQLEEEESARQ	K	MYH9_HUMAN	Myosin-9 OS=Homo sapiens GN=MYH9 PE=1 SV=4	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
3058,380	576	602	K	SSSYSKQFTSSTSYN RGDSTFESKSYK	M	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3083,620	25	51	R	DAHKSEVAHRFKDL GEENFKALVLI AF	A	ALBU_HUMAN	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	
3086,390	578	604	S	SYSKQFTSSTSYNRG DSTFESKSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3104,650	467	491	R	TLLVFEVQQPFLFVL WDQQHKFPVF	M	IC1_HUMAN	Plasma protease C1 inhibitor OS=Homo sapiens GN=SERPING1 PE=1 SV=2	
3115,400	574	601	R	GKSSSYSKQFTSSTSYN YNRGDSTFESKSY	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3127,680	247	277	S	LANTQPRGPPASSPA PAPKFSVPVTPKFTPV A	S	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
3140,590	659	687	M	NFRPGVLSSRQLGLP GPPDVPDHAAYHPF	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
3151,490	601	629	S	YKMADEAGSEADHE GTHSTKRGHAKSRP V	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3155,619	617	644	R	NVHSGSTFFKYLLQ GAKIPKPEASFSR	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
3156,609	119	147	N	DSGPRRYTIAALLSP YSYSTAVVTNPKE	-	TTHY_HUMAN	Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1	
3173,420	577	604	S	SSYSKQFTSSTSYNRG GDSTFESKSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3189,382	848	874	V	FHLSCHVPTLTNFP GEWICTFCRDL	K	TIF1A_HUMAN	Transcription intermediary factor 1-alpha OS=Homo sapiens GN=TRIM24 PE=1 SV=3	Sulfo (S)
3189,420	576	603	K	SSSYSKQFTSSTSYN RGDSTFESKSYKM	A	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3205,410	576	603	K	SSSYSKQFTSSTSYN RGDSTFESKSYKM	A	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Oxidation (M)
3214,710	247	278	S	LANTQPRGPPASSPA PAPKFSVPVTPKFTPV AS	K	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
3220,710	337	363	C	GLRPLFEKKSLEDKT ERELLESYIDGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
3238,517	600	629	K	SYKMADEAGSEADH EGTHSTKRGHAKSRP V	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3243,360	590	619	Y	NRGDSTFESKSYKM ADEAGSEADHEGHT ST	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3250,970			T	MTVSTTTSVPTTTSIP TTTSVPVTTT	V	HAVR1_HUMAN	Hepatitis A virus cellular receptor 1 OS=Homo sapiens GN=HAVCR1 PE=1 SV=2	8 Sulfo (T)
3254,510	600	629	K	SYKMADEAGSEADH EGTHSTKRGHAKSRP V	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Oxidation (M)
3257,490	528	557	K	TFPGFFSPMLGEFVS ETESRGSSEGI FTNT	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3260,457	576	604	K	SSSYSKQFTSSTSYN RGDSTFESKSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3271,635	658	687	R	MNFRPGVLSSRQLGL GPPDVPDHAAYHPF	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	
3276,445	133	160	K	RGEEIASSGQSPHIKS NLLSGLQSLWR	A	CI021_HUMAN	UPF0308 protein C9orf21 OS=Homo sapiens GN=C9orf21 PE=2 SV=1	3 Acetyl (K)
3276,450	215	247	S	SSSPGGSRRGGSPHHS DCKNGGGVGGGELD KKDQ	E	ZN703_HUMAN	Zinc finger protein 703 OS=Homo sapiens GN=ZNF703 PE=1 SV=1	Sulfo (S)
3276,450			K	SSSYSKQFTSSTSYN RGDSTFESKSYKMA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Oxidation (M)
3276,470	4188	4216	A	EYYRLHLEGYHGTA GDSMSYHSGSVFSA R	D	TENX_HUMAN	Tenascin-X OS=Homo sapiens GN=TNXB PE=1 SV=3	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
3276,480	171	200	L	SRGDSTYTFHEGPVL STSWRPGDSALS YTC	R	SLAF9_HUMAN	SLAM family member 9 OS=Homo sapiens GN=SLAMF9 PE=2 SV=2	
3278,989			S	WEEGSSGSSSDEEHG GGGMRVGPQYQAV	V	RCOR1_HUMAN	REST corepressor 1 OS=Homo sapiens GN=RCOR1 PE=1 SV=1	5 Sulfo (S)
3287,630			R	MNFRPGVLSSRQLGL PGPPDVPDHAAYHPF	R	ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	Oxidation (M)
3298,440			K	PQDMAEKDITFEAPP LGTQPCNPTPENGE A	P	TNR8_HUMAN	Tumor necrosis factor receptor superfamily member 8 OS=Homo sapiens GN=TNFRSF8 PE=1 SV=1	Amidated (C-term)
3306,560	173	204	D	PPGPPLLSKPGDEDD APPKMGEPAGGRYE HPG	L	ARI1B_HUMAN	AT-rich interactive domain-containing protein 1B OS=Homo sapiens GN=ARID1B PE=1 SV=2	Acetyl (K)
3314,750	35	62	Q	QASHYLYVRAHGVV QGKSTWPQKRTLF	V	RRP7A_HUMAN	Ribosomal RNA-processing protein 7 homolog A OS=Homo sapiens GN=RRP7A PE=1 SV=2	
3343,540	511	540	R	HRHPDEAAFFDTAST GKTFPGFFSPMLGEF	V	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3346,702	7	38	R	TAFGGRRAVPPNNS NAAEDDLPTVELQG VVPR	G	F13A_HUMAN	Coagulation factor XIII A chain OS=Homo sapiens GN=F13A1 PE=1 SV=4	
3362,550	395	425	R	YEGSYALTSEEAEERS DGDPVQPAVLQVHQ TS	-	SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	
3368,480	15	43	E	EMGSGDYDSMKEPC FREANANFNKIFLPT	I	CXCR4_HUMAN	C-X-C chemokine receptor type 4 OS=Homo sapiens GN=CXCR4 PE=1 SV=1	
3426,700	111	142	E	EEIFSPPPPPPEEGGP EAPIPPPPQPREKVS	S	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
3432,860	45	76	R	GHRPLDKKREEAPSL RPAPPISGGGYRAR PA	K	FIBB_HUMAN	Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2	
3440,642	462	491	K	GFPRGDKLFGPDLKL VPPMEEDYPQFGSPK	-	A2AP_HUMAN	Alpha-2-antiplasmin OS=Homo sapiens GN=SERPINF2 PE=1 SV=3	Sulfo (Y)/ Sulfo (S)
3440,650	462	491	K	GFPRGDKLFGPDLKL VPPMEEDYPQFGSPK	-	A2AP_HUMAN	Alpha-2-antiplasmin OS=Homo sapiens GN=SERPINF2 PE=1 SV=3	
3440,650	462	491	K	GFPRGDKLFGPDLKL VPPMEEDYPQFGSPK	-	A2AP_HUMAN	Alpha-2-antiplasmin OS=Homo sapiens GN=SERPINF2 PE=1 SV=3	Phospho (ST)
3440,650			S	CCYFCGTVCSLIHL LALRIPFYRSNVINH	F	OR5L1_HUMAN	Olfactory receptor 5L1 OS=Homo sapiens GN=OR5L1 PE=2 SV=1	Phospho (Y)
3442,608	511	541	R	HRHPDEAAFFDTAST GKTFPGFFSPMLGEF V	S	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3445,570	574	604	R	GKSSYSKQFTSSTS YNRGDSTFESKSYK MA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3461,568	574	604	R	GKSSYSKQFTSSTS YNRGDSTFESKSYK MA	D	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Oxidation (M)
3522,530	592	624	R	GDSTFESKSYKMAD EAGSEADHEGTHST KRGHA	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3737,810	279	316	S	KFSPGAPGGSGSQPN QKLGHPALSAGTGS PQPPSFTY	A	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
3738,600	420	451	V	SEYQQFQDAKAVLE EDEVTEEAEMEPED KGH	-	TBB1_HUMAN	Tubulin beta-1 chain OS=Homo sapiens GN=TUBB1 PE=1 SV=1	
3740,030	108	142	L	VTLAAHLPAEFTPAV HASLDKFLASVSTVL TSKYR	-	HBA_HUMAN	Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=2	
3741,840	111	145	E	EEIFSPPPPPPEEGGP EAPIPPPPQPREKVSSI D	L	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
3792,680	590	624	Y	NRGDSTFESKSYKM ADEAGSEADHEGTH STKRGHA	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
3825,970	467	497	R	TLLVFEVQPPFLFVL	P	IC1_HUMAN	Plasma protease C1 inhibitor	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
				WDQQHKFPVFMGR VYD			OS=Homo sapiens GN=SERPING1 PE=1 SV=2	
3827,950	185	222	K	SSTKPAAGGTAPLPP WKSPSSSQPLQVPA PAQSQTQF	H	ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	
3948,970	2	38	M	SETSRTAFGGRRRAVP PNNNSAAEDDLPTVE LQGVVPR	G	F13A_HUMAN	Coagulation factor XIII A chain OS=Homo sapiens GN=F13A1 PE=1 SV=4	Acetyl (N-term)
4050,656			G	VHFPPQDQRPQGQY PSMMMATKNHGTP MPYEQY	P	ZC12C_HUMAN	Probable ribonuclease ZC3H12C OS=Homo sapiens GN=ZC3H12C PE=1 SV=2	2 Sulfo (Y)
4050,920	429	467	C	DAPDSPRPVKASRED SGLFSPIRSSAFSPLG GCTPAECF	C	AKA11_HUMAN	A-kinase anchor protein 11 OS=Homo sapiens GN=AKAP11 PE=1 SV=1	
4050,920			F	VGQLKSSLTCTDCG YCSTVFDPFWDLSLP IAKRQYP	E	UBP2_HUMAN	Ubiquitin carboxyl-terminal hydrolase 2 OS=Homo sapiens GN=USP2 PE=1 SV=2	Acetyl (N-term) + Acetyl (K)
4052,690			N	HQSGSSSHQGGATTT PLASGASVNTPFINLP ALWRSV	A	SBSN_HUMAN	Suprabasin OS=Homo sapiens GN=SBSN PE=2 SV=1	4 Sulfo (T)
4072,890			F	MKTCFRGTVRLYNR QNCCAQFQWQPVNT GRGIAF	S	CX022_HUMAN	Uncharacterized protein CXorf22 OS=Homo sapiens GN=CXorf22 PE=2 SV=3	Sulfo (T) / Sulfo (Y) /Phospho (ST)
4082,536			M	CTGDKDQVNEGRKS FPSGHSSFAFAGLAF ASFYLA	G	PPC1B_HUMAN	Phosphatidate phosphatase PPAPDC1B OS=Homo sapiens GN=PPAPDC1B PE=1 SV=2	5 Sulfo (S)
4088,000	229	267	Q	ADRHSLMAKYIMDL ERLDPAGAAETFHV GLPGALGGHDG	L	JAK3_HUMAN	Tyrosine-protein kinase JAK3 OS=Homo sapiens GN=JAK3 PE=1 SV=2	
4088,020	328	363	R	TFGSGEADCGLRPLF EKKSLEDKTERELLE SYIDGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	
4089,883	592	629	R	GDSTFESKSYKMAD EAGSEADHEGTHST KRGHAKSRPV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
4111,575	90	125	M	DHTFSTTSKDGECS YTSLSIDICYPQEDS TYFTG	I	RTN1_HUMAN	Reticulon-1 OS=Homo sapiens GN=RTN1 PE=1 SV=1	Sulfo (Y)
4150,160	921	958	V	TLKVFLDDVNDLSP AFIPSSYSVKVLEDLP VGTVIAWL	E	FAT3_HUMAN	Protocadherin Fat 3 OS=Homo sapiens GN=FAT3 PE=2 SV=2	
4150,166	467	500	R	TLLVFEVQQPFLFVL WDQQHKFPVFMGR VYDPR	-	IC1_HUMAN	Plasma protease C1 inhibitor OS=Homo sapiens GN=SERPING1 PE=1 SV=2	
4166,170	827	862	M	HCYLDLVYQQWTES RKPTPIPPMVGQT NKSLTI	H	PAPP2_HUMAN	Pappalysin-2 OS=Homo sapiens GN=PAPPA2 PE=1 SV=4	
4167,973	328	363	R	TFGSGEADCGLRPLF EKKSLEDKTERELLE SYIDGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	Sulfo (T)
4167,980	328	363	R	TFGSGEADCGLRPLF EKKSLEDKTERELLE SYIDGR	I	THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	Phospho (ST)
4178,031			G	FEIPINGLSEFVDYED LVELAPGKFQLVAE NRRYQ	R	ITIH2_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H2 OS=Homo sapiens GN=ITIH2 PE=1 SV=2	Sulfo (S)
4180,900	1630	1665	N	VVQLNFLKLLSATAR QNFTYSCQNAAAAL DEATGDY	S	CO5A3_HUMAN	Collagen alpha-3(V) chain OS=Homo sapiens GN=COL5A3 PE=1 SV=3	Phospho (ST) + Phospho (Y)
4180,940	536	574	P	MLGEFVSETESRGE SGIFTNTKESSSHHPG IAEFPSRG	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
4182,161	788	829	P	AAMVPSQAPYLVPA FPLPAATSPGREYAA PGTAPEGLHGLPL	S	PER3_HUMAN	Period circadian protein homolog 3 OS=Homo sapiens GN=PER3 PE=2 SV=4	

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
4206,990	556	597	K	LMMEQENGLRGSP GPKGDMGSPGPKGD RGFPPTGPIGPLG	H	COHA1_HUMAN	Collagen alpha-1(XVII) chain OS=Homo sapiens GN=COL17A1 PE=1 SV=3	3 Oxidation (M)
4207,050	634	669	A	QSFNFRGLCLHKDY VAGREFVAWKDTHP DAFPNQL	T	PA24F_HUMAN	Cytosolic phospholipase A2 zeta OS=Homo sapiens GN=PLA2G4F PE=2 SV=3	
4360,030	590	629	Y	NRGDSTFESKSYKM ADEAGSEADHEGTH STKRGHAKSRPV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
4387,248			I	ALVVEMVLTFLFSA ALLATRSVITINKNG PFDFAAQPVDE	V	CRAS1_HUMAN	Putative uncharacterized protein CRYM-AS1 OS=Homo sapiens GN=CRYM-AS1 PE=5 SV=1	Sulfo (S)
4398,120	21	61	A	SEAEDASLLSFMQGY MKHATKTAKDALSS VQESQVAQQARG	W	APOC3_HUMAN	Apolipoprotein C-III OS=Homo sapiens GN=APOC3 PE=1 SV=1	
4518,960	27	64	Q	FVREPEDEEEEEEEE EDEDLELELVLE RKPAAGLS	A	RTN4_HUMAN	Reticulon-4 OS=Homo sapiens GN=RTN4 PE=1 SV=2	
4523,090	589	629	S	YNRGDSTFESKSYK MADEAGSEADHEGT HSTKRGHAKSRPV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
4524,082			L	FLEKMSEPLIRSSFS DRKFSVTSRGSIDDV FNCNLSP	R	CM033_HUMAN	Uncharacterized protein C13orf33 OS=Homo sapiens GN=C13orf33 PE=2 SV=1	2 Sulfo (S)
4610,120	588	629	T	SYNRGDSTFESKSYK MADEAGSEADHEGT HSTKRGHAKSRPV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
4622,490	384	423	L	SALVETRTIVRFNRPF LMIIVPTDTQNIFFMS KVTNPKQA	-	AACT_HUMAN	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2	
4625,220	1285	1334	E	KGDPGDVGGPGPPG ASGEPGAPGPPGKRG PSGHMGREGREGEK GAKGEPG	P	CO5A3_HUMAN	Collagen alpha-3(V) chain OS=Homo sapiens GN=COL5A3 PE=1 SV=3	
4784,030	576	619	K	SSSYSKQFTSSTSYN RGDSTFESKSYKMA DEAGSEADHEGTHS T	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
4786,570	377	418	F	LEAIPMSIPPEVKFNK PFVFLMIEQNTKSPLF MGKVVNPTQK	-	A1AT_HUMAN	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	
4899,340	21	65	A	SEAEDASLLSFMQGY MKHATKTAKDALSS VQESQVAQQARGW VTD	G	APOC3_HUMAN	Apolipoprotein C-III OS=Homo sapiens GN=APOC3 PE=1 SV=1	
4956,370	21	66	A	SEAEDASLLSFMQGY MKHATKTAKDALSS VQESQVAQQARGW VTDG	F	APOC3_HUMAN	Apolipoprotein C-III OS=Homo sapiens GN=APOC3 PE=1 SV=1	
5061,352	528	574	K	TFPGFFSPMLGEFVS ETESRGESEGIFTNTK ESSSHHPGIAEFPSRG	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
5099,300			N	IYPGDAGKAGDQLV PDNLKETDKEKGEV GKSWDSWSDSEEEF FE	C	RB3GP_HUMAN	Rab3 GTPase-activating protein catalytic subunit OS=Homo sapiens GN=RAB3GAP1 PE=1 SV=3	Acetyl (N-term) + 3Acetyl (K)
5189,450	528	575	K	TFPGFFSPMLGEFVS ETESRGESEGIFTNTK ESSSHHPGIAEFPSRG K	S	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
5333,346	576	624	K	SSSYSKQFTSSTSYN RGDSTFESKSYKMA DEAGSEADHEGTHS TKRGHA	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
5349,290	389	434	H	AASGKDSRLAVTD PTRPGATTKITTTSTY ISASTLKVNKTKGL	I	KAT6B_HUMAN	Histone acetyltransferase KAT6B OS=Homo sapiens GN=KAT6B PE=1 SV=3	Sulfo (S)

Calc. mass (Da) in all pre-fractionated serum samples	start	stop		sequence		Accession (UNIPROT)	Description	PTM
5349,341	576	624	K	SSSYKQFTSSTSYN RGDSTFESKSYKMA DEAGSEADHEGTHS TKRGHA	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	Oxidation (M)
5543,535	373	419	L	SSRGRERRVFLFEQIII FSEALGGGVRRGGTQ PGYVYKNSIKVSLCLG L	E	ARHGP_HUMAN	Rho guanine nucleotide exchange factor 25 OS=Homo sapiens GN=ARHGEF25 PE=1 SV=2	7 Phospho (ST)+Phospho (Y)
5593,600	260	310	G	QDIAPGSNWRRLYS HLATDSVDPADSGN YSCMAGNKSGDVKY VTYMVNVLE	H	CDON_HUMAN	Cell adhesion molecule-related/down-regulated by oncogenes OS=Homo sapiens GN=CDON PE=1 SV=2	
5801,630	576	628	K	SSSYKQFTSSTSYN RGDSTFESKSYKMA DEAGSEADHEGTHS TKRGHAKSRP	V	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
5900,700	576	629	K	SSSYKQFTSSTSYN RGDSTFESKSYKMA DEAGSEADHEGTHS TKRGHAKSRPV	R	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
6929,210	511	574	R	HRHPDEAAFFDTAST GKTFFGFFSPMLGEF VSETESRGSESGIFTN TKESSHHPGIAEFPS RG	K	FIBA_HUMAN	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	
7763,210			K	DAKISFLQKAIDVVV MVSGEPLLAKPARIV AGHEPERTNELLQIIG KCCLNKLSSDDAVR RVLAGEKGEVKG	R	MIPT3_HUMAN	TRAF3-interacting protein 1 OS=Homo sapiens GN=TRAF3IP1 PE=1 SV=1	Acetyl (N-term)
8909,390	23	101	G	TQQPQQDEMPSPFTL TQVKESLSSYWESA KTAAQNLYEKTYLP AVDEKLRDLYSKST AAMSTYTGIFTDQVL SVLKGEE	-	APOC2_HUMAN	Apolipoprotein C-II OS=Homo sapiens GN=APOC2 PE=1 SV=1	

Dei 101 segnali presenti negli spettri MALDI- Linear Mode (LM)), 43 sono stati identificati nella library sulla base di un confronto delle Masse molecolari relative (Mr) individuate in LC-ESI con quelle determinate in MALDI-TOF LM e e Reflector Mode (RM) e anche mediante MALDI FT-ICR. Gli errori medi associati alle misurazioni delle masse molecolari (MME *mass measurement error*) rilevate in MALDI ed ESI relativamente ai peptidi identificati sopra la soglia di identità erano $12,0 \pm 26,1$ ppm per RM e 101 ± 192 ppm per LM L'errore medio nella misurazione della massa dai profili FT-ICR rispetto alle masse esatte identificate in LC-ESI MS è stato determinato essere $0,11 \pm 0,88$ ppm.

Reflectron mode m/z	Linear mode m/z	FT-ICR m/z	MME in ppm (from precision profile)	Calc. m/z of [M+H] ⁺	Calc. mass (Da) in all pre-fractionated serum samples	Mascot significance	Accession (UNIPROT)	Sequence	PTM
---------------------	-----------------	------------	-------------------------------------	---------------------------------	---	---------------------	---------------------	----------	-----

Reflectron mode m/z	Linear mode m/z	FT-ICR m/z	MME in ppm (from precision profile)	Calc. m/z of [M+H] ⁺	Calc. mass (Da) in all pre-fractionated serum samples	Mascot significance	Accession (UNIPROT)	Sequence	PTM
1077,565	1077,720	1077,533	0,48	1077,532	1076,525	+	FIBA_HUMAN	GDFLAEGGGVR	
1206,609	1206,940	1206,575	0,29	1206,575	1205,568	+	FIBA_HUMAN	EGDFLAEGGGVR	
1263,632	1263,930	1263,596	-0,12	1263,596	1262,589	+	FIBA_HUMAN	GEGDFLAEGGGVR	
1350,662	1351,060	1350,628	-0,11	1350,628	1349,621	+	FIBA_HUMAN	SGEGDFLAEGGGVR	
1418,581	1419,090								
1449,767	1450,130	1449,760	0,10	1449,760	1448,752	+	CO3_HUMAN	THRIHWESASLL	
1465,690	1466,130	1465,655	-0,14	1465,655	1464,648	+	FIBA_HUMAN	DSGEGDFLAEGGGVR	
				1465,687	1464,680	+	APC2_HUMAN	GSGPSKDSFGELSR	Acetyl (N-term)
				1465,665	1464,658	+	PKHA9_HUMAN	TTGVSNSEEGIDVGT	
1487,679	1487,930			1487,624	1486,617	+	RBM25_HUMAN	SSELNAPSQESDSH	
				1487,628	1486,621	+	ZN831_HUMAN	SQTAGRTLTS SSPD	Sulfo (S)/ Sulfo (T)
				1487,673	1486,666	+	TLE2_HUMAN	FSLGHCPNQDWLA	
1518,711	1519,160								
1536,717	1537,150	1536,692	-0,50	1536,692	1535,685	+	FIBA_HUMAN	ADSGEGDFLAEGGGVR	
1545,750	1545,150			1545,612	1544,605	+	FIBA_HUMAN	DSGEGDFLAEGGGVR	Sulfo (S)
				1545,621	1544,614	+	FIBA_HUMAN	DSGEGDFLAEGGGVR	Phospho (ST)
				1545,619	1544,612	+	LPPRC_HUMAN	LQENGCLSDSDMFS	
1561,767	1562,250	1561,729	0,64	1561,728	1560,721	+	THRB_HUMAN	TATSEYQTFNPR	
1616,697	1616,770			1616,801	1615,794	+	SDPR_HUMAN	SALVEGEIAEAAEKA	
				1616,649	1615,642	+	FIBA_HUMAN	ADSGEGDFLAEGGGVR	Sulfo (S)
				1616,659	1615,652	+	FIBA_HUMAN	ADSGEGDFLAEGGGVR	Phospho (ST)
				1616,637	1615,630	+	GBRR1_HUMAN	TAMLDGNYSDGEVND	
1739,948	1740,440	1739,929	-0,45	1739,930	1738,923	+	CO4A_HUMAN / CO4B_HUMAN	NGFKSHALQLNNRQI	
1865,019	1865,380	1865,002	-0,24	1865,003	1863,995	+	CO3_HUMAN	SSKITHRIHWESASLL	
1896,037	1896,640	1896,031	-0,26	1896,031	1895,020	+	CO4A_HUMAN / CO4B_HUMAN	NGFKSHALQLNNRQIR	
				1896,027	1895,020	+	AF1L1_HUMAN	CSEKSHRVPQVKVQR	
2021,117	2021,650	2021,104	0,15	2021,104	2020,097	+	CO3_HUMAN	SSKITHRIHWESASLLR	
2104,251	2104,870								
2271,108	2271,800	2271,130	-0,35	2271,131	2270,123	+	ITIH4_HUMAN	SRQLGLPGPPDVPDHAAYH PF	

Reflectron mode m/z	Linear mode m/z	FT-ICR m/z	MME in ppm (from precision profile)	Calc. m/z of [M+H] ⁺	Calc. mass (Da) in all pre-fractionated serum samples	Masco t significance	Accession (UNIPROT)	Sequence	PTM
2378,200	2378,330	2378,209	-0,19	2378,210	2377,200	+	CO4A_HUMAN / CO4B_HUMAN	DDPDAPLQPVTPLQLFEGR RN	
2422,107	2422,870	2422,129	0,46	2422,130	2421,123	+	FIBA_HUMAN	DSHSLTTNIMEILRGDFSSA NN	
				2422,077	2421,070	+	FIBA_HUMAN	KMADEAGSEADHEGTHST KRGHA	
2423,091	2422,870	2423,081	0,31	2423,082	2422,070	+	FIBA_HUMAN	FTSSTSYNRGDSTFESKSYK M	
2553,083	2553,510	2553,101	-0,23	2553,101	2552,094	+	FIBA_HUMAN	SSSYSKQFTSSTSYNRGDST FES	
2602,284	2603,290	2602,311	-0,07	2602,311	2601,303	+	F13A_HUMAN	AVPPNNSNAAEDDLPTVEL QGVVPR	
		2603,168	0,04	2603,168	2602,161	+	SRGN_HUMAN	SLDRNLPSDSQDLGQHGLE EDFM	
2659,242	2659,730			2659,257	2658,250	+	FIBA_HUMAN	DEAGSEADHEGTHSTKRGH AKSRPV	
2753,405	2753,790	2753,436	-0,16	2753,437	2752,430	+	ALBU_HUMAN	DAHKSEVAHRFKDLGEENF KALVL	
2768,223	2768,580	2768,228	-0,05	2768,228	2767,221	+	FIBA_HUMAN	SSSYSKQFTSSTSYNRGDST FESKS	
2788,987	2789,180								
2861,344	2861,870	2861,334	0,03	2861,334	2860,330	+	FIBA_HUMAN	MADEAGSEADHEGTHSTK RGHAKSRPV	
2882,531	2883,060								
2931,291	2931,710	2931,292	0,18	2931,292	2930,284	+	FIBA_HUMAN	SSSYSKQFTSSTSYNRGDST FESKSY	
2951,874	2951,950								
	2989,850	2989,428	-0,54	2989,429	2988,420	+	FIBA_HUMAN	KMADEAGSEADHEGTHST KRGHAKSRPV	
3156,637	3157,340	3156,630	1,03	3156,627	3155,619	+	ITI4_HUMAN	NVHSGSTFFKYLLQGAKIP KPEASFSPR	
		n.d.		3157,616	3156,609	+	TTHY_HUMAN	DSGPRRYTIAALLSPYSYST TAVVTNPKE	
3190,448	3190,870	3190,428	0,38	3190,427	3189,420	+	FIBA_HUMAN	SSSYSKQFTSSTSYNRGDST FESKSYKM	
				3190,389	3189,382	+	TIF1A_HUMAN	FHLSCHVPTLTNFPSEWIC TFCRDLS	Sulfo (S)
3206,458	3206,410			3206,417	3205,410	+	FIBA_HUMAN	SSSYSKQFTSSTSYNRGDST FESKSYKM	Oxidation (M)
3239,571	3240,020	3239,526	0,57	3239,524	3238,517	+	FIBA_HUMAN	SYKMADEAGSEADHEGTH STKRGHAKSRPV	
3261,497	3261,850	3261,465	0,30	3261,464	3260,457	+	FIBA_HUMAN	SSSYSKQFTSSTSYNRGDST FESKSYKMA	
3277,496	3277,000			3277,457	3276,450	+	FIBA_HUMAN	SSSYSKQFTSSTSYNRGDST FESKSYKMA	Oxidation (M)

Reflectron mode m/z	Linear mode m/z	FT-ICR m/z	MME in ppm (from precision profile)	Calc. m/z of [M+H] ⁺	Calc. mass (Da) in all pre-fractionated serum samples	Masco t significance	Accession (UNIPROT)	Sequence	PTM
				3277,452	3276,445	+	CI021_HUMAN	RGEEIASSGQSPHIKSNLLSGSLQSLWR	3 Acetyl (K)
				3277,457	3276,450	+	ZN703_HUMAN	SSSPGGSRGGSPHHSCKN GGGVGGGELDKKDQ	Sulfo (S)
				3277,477	3276,470	+	TENX_HUMAN	EYYRLHLEGYHGTAGDSM SYHSGSVFSAR	
				3277,487	3276,480	+	SLAF9_HUMAN	SRGDSTYTFHEGPVLSTSW RPDGDSALSYTC	
	3314,800								
	3446,080			3446,577	3445,570	+	FIBA_HUMAN	GKSSSYSKQFTSSTSYNRGD STFESKSYKMA	
	3881,440								
	3934,070								
	3951,020								
				4051,927	4050,920	+	AKA11_HUMAN	DAPDSPRPVKASREDSGLFS PIRSSAFSPGGCTPAECF	
	4052,050			4051,663	4050,656	+	ZC12C_HUMAN	VHFPPQDRPQGQYPSMM MATKNHGTPMPYEQY	2 Sulfo (Y)
				4051,927	4050,920	+	UBP2_HUMAN	VGQLKSSLCTDCGYCSTV FDPFWDLSLPIAKRGYP	Acetyl (N-term) + Acetyl (K)
	4089,110	4089,021	-0,45	4089,023	4088,020	+	THRB_HUMAN	TFGSGEADCGLRPLFEKKSLEDKTERELLESYIDGR	
	4108,100								
	4120,400								
		4151,162	-2,51	4151,173	4150,166	+	IC1_HUMAN	TLLVFEVQQPFLFVLWDQQ HKFPVFMGRVYDPA	
	4151,450			4151,167	4150,160	+	FAT3_HUMAN	TLKVFLDDVNDSPAFIPSS YSVKVLEDLPVGTVIAWL	
	4165,920								
				4207,997	4206,990	+	COHA1_HUMAN	LMMEQENGLRGSPPGKGD MGSPGPKGDRGFPPTGIP GPLG	3 Oxidation (M)
4208,447	4207,870			4208,057	4207,050	+	PA24F_HUMAN	QSFNFTRGLCLHKDYVAGR EFVAWKDTHPDAFPNQL	
	4245,600								
	4264,490								
	4280,440								
	4380,710								
	4393,470								
	4465,940								
	4565,490								
	4624,980								

Reflectron mode m/z	Linear mode m/z	FT-ICR m/z	MME in ppm (from precision profile)	Calc. m/z of [M+H] ⁺	Calc. mass (Da) in all pre-fractionated serum samples	Mascoat significance	Accession (UNIPROT)	Sequence	PTM
	4641,930								
	4677,910								
	4687,440								
	4709,030								
	4786,260	4785,045	0,99	4785,040	4784,030	+	FIBA_HUMAN	SSSYSKQFTSSTSYPNRGDST FESKSYKMADEAGSEADHE GTHST	
		n.d.		4787,574	4786,570	+	A1AT_HUMAN	LEAIPMSIPPEVKFNKPFVFL MIEQNTKSPFLFMGKVVNPT QK	
	4852,680								
	4961,940								
	5002,200								
	5062,990	5062,378	3,69	5062,359	5061,352	+	FIBA_HUMAN	TFPGFFSPMLGEFVSETESR GSESGIFTNTKESSHHPGIA EFPSRG	
	5077,900								
	5246,270								
	5261,850								
	5291,090								
	5334,920	5334,359	0,97	5334,354	5333,346	+	FIBA_HUMAN	SSSYSKQFTSSTSYPNRGDST FESKSYKMADEAGSEADHE GTHSTKRGHA	
	5595,160			5594,607	5593,600	+	CDON_HUMAN	QDIAPGSNWRRLYSHLATD SVDPADSGNYSCMAGNKS GDVKYVTYMVNVLE	
	5632,200								
	5751,390								
	5804,820								
	5902,810			5901,707	5900,700	+	FIBA_HUMAN	SSSYSKQFTSSTSYPNRGDST FESKSYKMADEAGSEADHE GTHSTKRGHAKSRPV	
	6087,820								
	6224,640								
	6302,590								
	6431,320								
	6475,940								
	6526,420								
	6588,020								
	6629,700								
	6934,840								
	7565,400								

Reflectron mode m/z	Linear mode m/z	FT-ICR m/z	MME in ppm (from precision profile)	Calc. m/z of [M+H] ⁺	Calc. mass (Da) in all pre-fractionated serum samples	Mascoat significance	Accession (UNIPROT)	Sequence	PTM
	7668,840								
	7764,630			7764,217	7763,210	+	MIPT3_HUMAN	DAKISFLQK AIDVVVMVSG EPLLAKPARIVAGHEPERTN ELLQIIGKCCLNKLSSDDAV RRVLAGEKGEVKG	Acetyl (N-term)
	7923,230								
	8141,420								
	8205,360								
	8567,140								
	8602,430								
	8766,140								
	8918,270								
	9134,030								
	9289,830								
	9380,880								
	9424,360								
	9716,640								

4.4 Validazione

Al fine di validare i risultati ottenuti dall'analisi MALDI-TOF, i rapporti peptidici ottenuti da questa tecnica sono stati confrontati con quelli risultanti dalla strategia di quantificazione relativa Label-Free (LF) mediante nLC-ESI MS/MS applicata agli stessi campioni.

Per l'analisi LF sono stati preparati due pools di siero con i campioni analizzati con metodo MALDI: il primo costituito da 80 controlli mentre l'altro composto da 62 pazienti affetti da ccRCC e 18 non-ccRCC. Per calcolare il ratio m/z paziente/controllo, sono stati utilizzati gli spettri di massa ottenuti con MALDI-TOF LM per ogni paziente e controllo inserito nei pools.

Analogamente a quanto riportato per l'identificazione dei peptidi, è stato eseguito l'allineamento delle masse. Sono state determinate le masse molecolari relative (Mr) mediante MALDI-RM ed ESI per i peptidi quantificati con uno *score* maggiore della soglia di identità. Le differenze medie tra le due Mr sono state di $8,4 \pm 13$ ppm. Una validazione aggiuntiva dell'allineamento delle masse è stata effettuata anche grazie alle misurazioni accurate effettuate nei profili MALDI FT-ICR. Anche in questo caso per i peptidi quantificati con uno *score* maggiore della soglia di

identità è stato valutato l'errore medio associato alle misurazione delle masse molecolari MME nei profili FT-ICR con le masse calcolate in ESI ed è stato trovato un valore di $0,14 \pm 0,92$ ppm. I bassi MMEs relativi al confronto tra le masse calcolate per i peptidi identificati e quantificati in ESI e le relative Mr protonate presente nei profili MALDI LM, RM, FT-ICR hanno avvalorato l'accuratezza dell'allineamento. Tre dei 28 segnali peptidici allineati e quantificati tramite LF sono stati anche confermati tramite identificazione MALDI TOF MS/MS: tutti e tre sono stati riconosciuti come frammenti della catena alfa del Fibrinogeno (*accession number* di SwissProt: P02671) con uno *score* in *Mascot* maggiore della soglia di identità ($p < 0,05$). In particolare, il segnale a m/z 1263,63, determinato in MALDI-TOF in modalità *reflector* (Mr calcolata 1262,589), è stato identificato (*score*=93) come un peptide avente sequenza GEGDFLAEGGGVR. Gli altri due ioni a m/z 1350,66 (Mr calcolata 1349,621) e m/z 1465,69 (Mr calcolata 1464,648) sono stati identificati come peptidi aventi rispettivamente sequenza SGEGDFLAEGGGVR (*score*=112) e DSGEGDFLAEGGGVR (*score*=130).

Come si può vedere dalla tabella i segnali individuati in MALDI-RM e 6 in MALDI-LM sono stati presumibilmente identificati mediante approccio LF con uno *score* di *Mascot* maggiore della soglia di identità. I segnali a m/z 1449,77 e 2753,40 osservati in MALDI-RM sono stati identificati con uno *score* maggiore della soglia di omologia.

Gli ioni aventi Mr 2601,30 e 2602,17 identificati nell'analisi LC-ESI MS/MS e aventi tempi di ritenzione (r.t.) diversi, potrebbero corrispondere ai segnali a m/z 2603 e 2602,28 osservati negli spettri MALDI-TOF acquisiti rispettivamente in modalità lineare e reflector. Il primo segnale è stato identificato come un frammento di F13A (*Coagulation factor XIII A chain*) (r.t. 82 min) mentre il secondo come un frammento di SRGN (*Serglycin*) (r.t. 77 min). Entrambi i peptidi sono stati osservati nei profili MALDI FT-ICR rispettivamente a valori di m/z di 2602,311 e 2603,168.

Inoltre, sono stati individuati tramite approccio LF due peptidi: uno a Mr 3155,62 (r.t. 74 min) identificato come un frammento di ITIH4 (*Inter-alpha-trypsin inhibitor heavy chain H4*) e l'altro a Mr 3156,61 (r.t. 93 min) come un frammento di TTHY (*Transthyretin*), che potrebbero corrispondere al segnale presente a m/z 3156,64 osservato nello spettro MALDI-RM. Tuttavia, nei profili MALDI FT-ICR, è stato possibile collegare ai dati LC-ESI solo lo ione a m/z 3156,630.

Lo ione a m/z 4786 presente nello spettro MALDI-LM non è stato osservato nello spettro reflector corrispondente. L'analisi MALDI FT-ICR lega questo peptide a quello identificato come un frammento di FIBA (*Fibrinogen alpha chain*) corrispondente ad una Mr di 4784,03 Da.

I *ratios* peptidici ottenuti dall'analisi LC-ESI MS/MS sono stati confrontati con quelli calcolati dagli spettri MALDI-TOF: il *t-test di Student* per dati appaiati non ha messo in evidenza alcuna differenza statisticamente significativa tra i *ratios* ($p=0.4547$).

QUALITY of PEPTIDE IDENTIFICATION in SAMPLES			N° of PEPTIDE IDENTIFICATIONS					
ALL pre-fractionated serum samples	Kidney tumor	Controls	With PTM	Without PTM	TOTAL	With PTM	Without PTM	TOTAL
IDENTITY	IDENTITY	IDENTITY	7	79	86	17	184	201
	IDENTITY	HOMOLOGY	1	13	14			
	IDENTITY	not identified	1	29	30			
	HOMOLOGY	IDENTITY	0	18	18			
	not identified	IDENTITY	8	45	53			
HOMOLOGY	HOMOLOGY	HOMOLOGY	6	19	25	111	151	262
	HOMOLOGY	not identified	76	63	139			
	not identified	HOMOLOGY	29	69	98			
Total (IDENTITY+HOMOLOGY)						128	335	463

Tabella 14: Peptidi identificati mediante nLC-ESI MS/MS nel siero purificato con biglie magnetiche RP-C18 di pazienti e controlli. L'identificazione si basava sullo *score* di *Mascot* maggiore della soglia di identità (rosso) o della soglia di omologia (blu). PTM= modificazioni post-traduzionali

5. CONCLUSIONI

Negli ultimi anni ha assunto sempre più importanza, ai fini diagnostici, lo studio dei profili proteomici, che presenta numerosi vantaggi rispetto all'analisi del trascrittoma poiché l'espressione genetica non sempre è correlata con l'assetto proteico funzionale e quantitativo in quanto è regolato sia da eventi post-traduzionali sia da fattori ambientali.

Il proteoma dei fluidi biologici rifletterebbe lo stato di salute dell'organismo, dato che essi trasportano proteine secrete dai tessuti. Per questo motivo, l'investigazione diretta nei fluidi periferici di proteine correlate ad un alterato stato fisiopatologico rappresenta un valido approccio, perché i campioni possono essere ottenuti facilmente con metodi di raccolta per nulla o poco invasivi. L'espressione delle proteine, in termini qualitativi, quantitativi ma anche di struttura, modificazioni, interazioni proteina-proteina e distribuzione spaziale, rappresenta un utile strumento al fine di comprendere i meccanismi fisiopatologici della malattia, identificare biomarcatori per una diagnosi precoce della malattia oppure eventuali bersagli di nuove molecole ad attività farmacologica.^{194,195} È importante inoltre sottolineare che l'uso di singoli biomarcatori per la diagnosi di una malattia non è risolutivo, in quanto molte patologie sono di origine poligenica, quindi la predittività di metodi diagnostici basati sul dosaggio di marcatori multipli risulta più affidabile soprattutto quando si voglia utilizzarli per la diagnosi precoce.

Numerosi sforzi sono stati compiuti nel campo della ricerca al fine di sviluppare nuove metodiche adeguate per una diagnosi precoce del RCC o per un suo monitoraggio continuo. Tuttavia ad oggi non sono stati individuati biomarcatori specifici, che permettano attraverso l'analisi di fluidi biologici il riconoscimento di questa patologia né in stadio precoce né già avanzato. Per molti anni a riguardo sono state impiegate tecniche di proteomica classica, che consistono nell'estrazione proteica da tessuti renali normali e tumorali, seguita da separazione in elettroforesi bidimensionale ed identificazione con spettrometria di massa delle proteine differenzialmente espresse¹⁹⁶. L'elettroforesi bidimensionale risulta però una tecnica lunga e laboriosa, quindi difficilmente applicabile per uno screening diagnostico della popolazione; inoltre consente facilmente di rilevare proteine intatte e loro frammenti oltre 10 kDa, ma presenta forti limitazioni per proteine/peptidi al di sotto di questa soglia¹⁹⁷. Negli ultimi anni, tuttavia, le tecniche utilizzate per lo studio del proteoma hanno subito un notevole impulso, che ha condotto ad un significativo miglioramento della capacità risolutiva dei metodi analitici e ad uno sviluppo di sofisticati programmi bioinformatici per un aumento della specificità e selettività di analisi.

Oggi l'interesse scientifico è rivolto all'utilizzo di tecniche di profiling altamente sensibili e versatili, basate su processi cromatografici di purificazione prima dell'analisi in spettrometria di

massa¹⁹⁸. Il frazionamento del campione può essere utilizzato al fine di aumentare il numero di proteine individuabili in MS (arricchimento di proteine poco abbondanti) mediante ad esempio cromatografia liquida bidimensionale (MudPit) o di ridurre la complessità del campione mediante utilizzo di superfici attivate. La tecnica di estrazione su fase solida basata sul frazionamento *off-line* del proteoma mediante l'utilizzo di biglie magnetiche o un target MALDI solido con superficie attivata, prima dell'analisi in spettrometria di massa, come ad esempio le tecniche ClinProt e SELDI, permette una ricerca *high-throughput* nei fluidi periferici di proteine/peptidi alterati dalla patologia^{199,200}.

L'obiettivo di questo lavoro di ricerca è stato quello di costruire i profili proteomici differenziali di sieri di controlli e di pazienti con RCC, estendendo l'analisi anche a pazienti affetti da non-ccRCC, mediante tecnica di estrazione in fase solida associata alla spettrometria di massa. L'applicazione di questa tecnologia avanzata, che impiega biglie magnetiche attivate con opportune fasi per l'estrazione da fluidi biologici di proteine/peptidi, consente di valutare il proteoma al fine di identificare marcatori predittivi della patologia. In questo studio la ricerca è stata focalizzata maggiormente su piccoli peptidi sierici ($M_r < 5000$ Da) ottenuti dal pre-frazionamento dei campioni con biglie magnetiche a fase inversa (RP-C18).

In primo luogo l'analisi era volta ad individuare un pattern di segnali in grado di differenziare tumori maligni da masse renali benigne e soggetti sani. Da un punto di vista clinico, infatti, è più interessante focalizzarsi sulla malignità del tumore in modo supportare la diagnosi di casi dubbi ed avere ricadute diverse sull'intervento chirurgico ed il follow up del paziente. L'analisi statistica effettuata sui profili medi delle popolazioni in studio, ha permesso di individuare un pattern di cinque segnali con elevate percentuali di specificità e sensibilità rispettivamente del 90% e del 94% (Tab.12) ottenute nella fase di *training*, le quali sono state successivamente confermate nel test di validazione che hanno riportato valori di specificità e sensibilità rispettivamente di 83,7% e 89,2%.

Successivi test analitici hanno permesso di studiare anche la natura delle lesioni renali effettuando differenti comparazioni tra le popolazioni in esame: 2 classificazioni caso/controllo (ccRCC versus controlli e non-ccRCC versus controlli) ed un'ulteriore classificazione ccRCC versus non-ccRCC. Sono stati quindi testati diversi clusters aventi numero crescente di ioni, da 2 a 6 e grazie a questi test è stato individuato un pattern di sei segnali con valori di m/z in modalità lineare 1466, 2378, 2789, 4052, 4089 e 4151 Da in grado di discriminare i pazienti affetti da ccRCC dai controlli. I valori di AUC di tali segnali compresi tra 0,89 e 0,70, mostravano elevata accuratezza e hanno mostrato elevati indici di specificità (96%) e sensibilità (94%) nella fase di

training confermati anche dalle alte percentuali di specificità e sensibilità ottenute nella fase di validazione.

Un ulteriore cluster di tre segnali (m/z 1419, 2603 e 4264 Da) ha permesso di discriminare i pazienti affetti da carcinoma renale non-RCC dai soggetti controllo anche in questo caso con elevati indici di specificità e sensibilità rispettivamente 95,7% e 89,7% nella fase di training e 94,6% e 89,7% nella fase di validazione.

Il confronto statistico effettuato sulle due popolazioni di pazienti, RCC versus non-RCC non ha purtroppo condotto alla costruzione di un cluster affidabile probabilmente a causa dell'elevata eterogeneità dei pazienti stessi e dell'esiguo numero di campioni.

5.1 Validazione

I risultati ottenuti dalla pre-purificazione mediante biglie magnetiche RP-C18 combinata con l'analisi MALDI-TOF richiedono generalmente una conferma con un metodo di validazione indipendente. Solitamente questi metodi prevedono analisi immunologiche basate sull'utilizzo di western-blot con anticorpi specifici e non cross-reattivi. Tuttavia questo tipo di validazione non può essere eseguito a livello peptidico a causa della piccola dimensione dei peptidi, della possibilità che esse possano non essere di *ratios* di quantificazione.²⁰¹

Recentemente è stato pubblicato uno studio sull'iperplasia prostatica benigna in cui digeriti triptici di proteine estratte da urine con tecnica *high-throughput* MALDI-TOF sono state identificate e calcolati i *ratios* di quantificazione controllo /paziente attraverso la tecnologia LC-ESI MS/MS in combinazione con la marcatura con isotopi stabili.

In questo lavoro di tesi si è pensato pertanto di applicare un approccio simile per la validazione dei risultati ottenuti con l'analisi MALDI: è stata, quindi, effettuata una quantificazione relativa label free LF basata su un metodo di ionizzazione alternativo.

La metodologia proteomica LF è stata applicata per la valutazione quantitativa a livello peptidomico e non mirata al livello proteico come solitamente riportato per le miscele triptiche complesse²⁰².

I *ratios* peptidici eseguiti sui pools di siero degli stessi pazienti e controlli ottenuti dall'approccio LF sono stati confrontati con quelli ottenuti con la tecnologia MALDI. Il confronto dei *ratios* dei segnali peptidici quantificati sia in MALDI che in ESI è stato reso possibile attraverso un accurato allineamento delle masse. È stato quindi possibile sovrapporre molte delle identificazioni ottenute mediante LC-ESI MS/MS con i segnali presenti nei profili MALDI-LM dei campioni di siero analizzati. L'allineamento dei *ratios* m/z dell'analisi MALDI-LM con quelli ottenuti in LC-ESI è stato eseguito attraverso l'utilizzo degli spettri di massa acquisiti in

MALDI-TOF con modalità *reflector* e successivamente confermato dai profili MALDI FT-ICR e dalle identificazioni MALDI MS/MS. Come riportato nei risultati, gli MME sono risultati essere molto bassi intermini di ppm. L'analisi statistica non ha messo in evidenza alcuna differenza statisticamente significativa ($p < 0,05$) tra i *ratios* peptidici determinati in MALDI ed in ESI, rinforzando quindi la loro validità.

Solo in due casi sono stati identificati in LC-ESI MS/MS due peptidi che potrebbero corrispondere ad un singolo ione presente negli spettri MALDI-LM (m/z 2603 e 3157). Ciononostante, i *ratios* per entrambi i possibili peptidi identificati tramite nLC-ESI MS/MS, che possono essere assegnati agli ioni osservati in MALDI-TOF, erano simili. Il *ratio* del segnale a m/z 2603 è stato osservato essere sottoespresso sia per i due peptidi identificati in ESI che per il segnale osservato in MALDI. Per quanto riguarda l'altro segnale a m/z 3157 Da, che probabilmente è presente in maggiore quantità nei controlli in base ai risultati MALDI, nei dati ottenuti in ESI non sono state riscontrate alterazioni per il peptide a Mr 3156,61 mentre quello a Mr 3155,62 risultava essere presente in maggiore concentrazione nei soggetti sani. Ventotto dei 101 ioni presenti negli spettri MALDI-LM sono stati quantificati anche mediante approccio LF.

5.2 Identificazione peptidica

L'identificazione di potenziali marcatori correlati ad uno stato di salute alterato rappresenta un passaggio fondamentale nello sviluppo di saggi specifici, inclusi saggi immunologici al fine di ottenere la validazione di un efficace strumento diagnostico della patologia²⁰³.

Va inoltre sottolineata l'importanza che l'identificazione di ciascuna proteina potrebbe assumere nel miglioramento della conoscenza biologica, della comprensione della funzione e della regolazione di molecole bioattive e della regolazione di molecole bioattive e di prodotti di degradazione nei fluidi biologici in stati patologici²⁰⁴.

In questo studio le identificazioni sono state ottenute grazie alla tecnologia nLC-ESI MS/MS che permette di ottenere identificazioni più accurate e affidabili rispetto all'approccio MALDI.

Tale tecnologia, infatti, risulta essere veloce e permette di ridurre il tempo necessario alla preparazione del campione specialmente con la combinazione di un robot com'è stato realizzato in questo progetto. Tuttavia l'identificazione di peptidi con la tecnica MALDI è invece più complessa, solitamente meno sensibile e informativa, soprattutto a causa della singola carica del precursore, a differenza di quanto avviene nella tecnologia LC-ESI MS/MS dove le specie ottenute sono tutte multicarica.

Pochi anche in letteratura gli studi in cui l'identificazione è stata condotta con la tecnica MALDI e la più ampia lista di peptidi identificati in siero umano è stata pubblicata da Villanueva et al.²⁰⁵

Anche Tiss et al. sono stati in grado di identificare più di 90 picchi nei profili MALDI di siero umano, derivanti da 13 proteine e con una piccola sovrapposizione con i dataset già pubblicati in letteratura²⁰⁶.

Per questi motivi la scelta della tecnica di identificazione si è basata su un'analisi in LC ESI MS/MS e grazie all'aumento delle energie di frammentazione CID e l'ottimizzazione dei tempi di accumulo degli spettri MS/MS, insieme con un adeguato setting dei parametri di acquisizione è stato possibile ottenere una migliore frammentazione dei peptidi endogeni.

Pertanto dall'analisi delle frazioni purificate con biglie magnetiche è stato possibile costruire una library di circa 460 peptidi endogeni di cui di cui circa 200 peptidi unici con uno score di Mascot superiore alla soglia di identità e 262 peptidi aventi uno score sopra la soglia di omologia. La costruzione di questo database ha permesso di identificare 43 dei 101 segnali presenti negli spettri maldi acquisiti in modalità lineare grazie all'allineamento delle masse molecolari relative individuate in LC-ESI con quelle determinate in MALDI-TOF e in MALDI FT-ICR.

Il match tra le identificazioni ottenute in nLC-ESI MS/MS e i segnali osservati negli spettri MALDI è stato reso ancora più valido dal basso valore in ppm del MME dell'allineamento delle masse, come precedentemente descritto.

Se si considera l'intera library, 7 dei 12 peptidi identificati e presenti nei modelli diagnostici appartengono a proteine implicate nella coagulazione del sangue o nell'attivazione del sistema del complemento.

MALDI-TOF				nLC-ESI MS/MS				
Modalità Reflectron m/z	Modalità Lineare m/z	Massa calc. (Da) nei campioni di siero dopo pre-frazionamento	AUC	Popolazione con aumentato valore di concentrazione	Significato in Mascot	Accession (UNIPROT)	Sequenza	PTM
1263.632	1263.930	1262.589	0.77	Maligni	+	FIBA_HUMAN	GEGDFLAEGGGVR	
1418.581	1419.090		0.84/0.85	Maligni e non-ccRCC				
1465.690	1466.130	1464.648	0.89	ccRCC	+	FIBA_HUMAN	DSGEGFLAEGGGVR	
2378.200	2378.330	2377.200	0.77	Controlli	+	CO4A_HUMAN / CO4B_HUMAN	DDPDAPLQPVTVPLQLFEGRRN	
2422.107	2422.870	2421.123	0.79	Maligni	+	FIBA_HUMAN	DSHSLTINMIEILRGDFSSANN	
		2421.070			+	FIBA_HUMAN	KMADEAGSEADHEGTHSTKRGA	
2423.091	2422.870	2422.070			+	FIBA_HUMAN	FTSSTSYNRGDSFESKSYKM	
2602.284	2603.290	2601.303	0.84	Controlli	+	F13A_HUMAN	AVPPNNSNAEEDLPTVELQGVVPR	
		2602.161			+	SRGN_HUMAN	SLDRNLPSDSQDLGQHGLEEDFM	
2788.987	2789.180		0.75/0.88	Maligni e ccRCC				
	4052.050	4050.920	0.85	Controlli	+	AKA11_HUMAN	DAPDSRPVKASREDSGLFSPIRSSAFSLGGCTPAECF	
		4050.656			+	ZC12C_HUMAN	VHFPQDQRQPGQYPSMMMATKNHGTMPYEQY	2 Sialfo (Y)
		4050.920			+	UBP2_HUMAN	VGQLKSSLTCTDCGYCSTVDFPFDLPLPIAKRGYP	Acetyl (N-term) + Acetyl (K)
	4089.110	4088.020	0.69	Controlli	+	THRB_HUMAN	TFGSGEADCGLRPLFEKKSLEDKTERELLESYIDGR	
	4151.450	4150.166			+	IC1_HUMAN	TLLVFEVQQPFLFVLDQDQHKFPVFMGRVYDPA	
	4264.490				0.83			
	4961.940		0.83	Benigni e controlli				

Tabella 15. Elenco dei segnali utilizzati per la costruzione dei modelli diagnostici e di screening. L'identificazione dei peptidi è stata condotta sulla base degli score di Mascot maggiore della soglia di identità ("+") o di omologia ("-"). La tabella riporta inoltre i valori di AUC (Area Under the Curve) di tutti gli ioni sulla base delle loro curve ROC (Receiver Operating Characteristics) e la popolazione in cui la concentrazione del relativo segnale risulta essere aumentata in maniera significativa. PTM= modificazioni post-traduzionali

Tre ioni a m/z 1264,1466 e 2423 osservati negli spettri MALDI-LM, identificati come frammenti della catena alfa del Fibrinogeno (FIBA), sono presenti in concentrazioni maggiori nei campioni dei pazienti affetti da RCC. I primi due peptidi sono frammenti del Fibrinopeptide A (FPA) (Mr 1535,69) che deriva dalla degradazione del fibrinogeno a fibrina con il rilascio dei Fibrinopeptidi A e B. Invece, lo ione a m/z 2423 corrisponde alla sequenza aminoacidica tra Asp101 e Asn122 della catena alfa del Fibrinogeno (FIBA).

Tali risultati sono in accordo con quanto già osservato in uno studio sul tumore della mammella, della vescica e del fegato condotto da Villanueva et al²⁰⁷. Gli autori hanno dimostrato che alcuni biomarcatori non sono direttamente espressi dai tessuti patologici, ma sono prodotti di degradazione derivanti dall'attività di esopeptidasi che contribuiscono alla generazione non solo di peptidi serici cancro specifici, ma anche specifici per il tipo di carcinoma. E' stato ipotizzato che differenti tipi di tumore secernono proteasi distinte che, attraverso la loro attività catalitica, possono generare profili peptidici sierici unici.

Inoltre 21 dei 43 peptidi identificati e presenti nei profili MALDI sono stati riportati nello studio di Villanueva sebbene siano state utilizzate biglie magnetiche con fasi stazionarie differenti (RP-C18 e C8). La tabella 17 mostra i ratios peptidici delle tre differenti neoplasie studiate da Villanueva e quelli derivanti dal tumore renale indagati in questo studio.

MALDI-TOF			nLC-ESI MS/MS	PEPTIDE IDENTIFICATION		Villanueva et al [11] Ratio of medians (patients/controls)		
Reflectron mode monoisotopic m/z	Linear mode monoisotopic m/z	Linear mode Ratio (kidney tumor/CTRL)	Calc. mass (Da) in all pre-fractionated serum samples by nLC-ESI MS/MS	Sequence	Accession (UNIPROT)	Prostate cancer	Bladder cancer	Breast cancer
1077.565	1077.720	0.750	1076.525	GDFLAEGGGVR	FIBA_HUMAN	0.54	0.50	0.95
1206.609	1206.940	2.540	1203.588	EGDFLAEGGGVR	FIBA_HUMAN	0.50	0.44	0.69
1263.632	1263.930	5.760	1262.589	GGDFLAEGGGVR	FIBA_HUMAN	0.20	0.24	0.23
1350.662	1351.060	3.634	1349.621	SGGDFLAEGGGVR	FIBA_HUMAN	0.47	0.46	0.35
1449.767	1450.130	1.019	1448.732	THRHVESASLL	CO3_HUMAN	1.885	2.646	4.37
1465.690	1466.130	4.734	1464.648	DSGGDFLAEGGGVR	FIBA_HUMAN	0.66	0.55	0.30
1536.717	1537.150	1.924	1535.685	ADSGDFLAEGGGVR	FIBA_HUMAN	0.58	0.60	0.54
1739.948	1740.440	1.063	1738.923	NGKSHALQLNRRQI	CO4A_HUMAN / CO4B_HUMAN	0.75	0.66	2.76
1865.019	1865.380	0.976	1863.995	SKKTRHWHVESASLL	CO3_HUMAN	2.18	3.33	0.30
1896.037	1896.640	1.147	1895.020	NGKSHALQLNRRQI	CO4A_HUMAN / CO4B_HUMAN	1.27	3.33	2.95
2021.117	2021.650	1.130	2020.097	SKKTRHWHVESASLLK	CO3_HUMAN	1.05	0.96	0.93
2271.108	2271.800	0.396	2270.125	SRQLGLGPPDPVPOHAAVHPF	ITIH_HUMAN	4.58	2.64	1.46
2553.083	2553.510	1.430	2552.094	SSSYSRQFTSSSYNRGDSIPES	FIBA_HUMAN	1.35	1.31	1.50
2602.284	2603.290	0.314	2601.303	AVPNSNNAEDELPTVLOGVVPFR	F13A_HUMAN	2.84	2.31	4.73
2679.242	2679.730	0.658	2678.250	DEAGSEADHEGTHSTUNGHAQSPV	FIBA_HUMAN	1.37	1.18	2.45
2768.223	2768.580	1.195	2767.221	SSSYSRQFTSSSYNRGDSIPESIS	FIBA_HUMAN	0.77	0.63	0.98
2931.291	2931.710	1.357	2930.284	SSSYSRQFTSSSYNRGDSIPESISY	FIBA_HUMAN	1.24	0.13	0.97
3156.637	3157.340	0.470	3155.619	NVHSGSTPKIYYLQGANPPEASFPFR	ITIH_HUMAN	3.43	10.88	1.33
3190.448	3190.870	1.485	3189.420	SSSYSRQFTSSSYNRGDSIPESISYTM	FIBA_HUMAN	0.52	0.60	0.94
3339.571	3340.020	0.882	3338.517	SYVMADIEAGSEADHEGTHSTUNGHAQSPV	FIBA_HUMAN	1.00	1.00	1.00
3361.497	3361.850	1.501	3360.457	SSSYSRQFTSSSYNRGDSIPESISYQMA	FIBA_HUMAN	0.33	0.08	0.74

Tabella 16. Confronto tra i *ratio* dei peptide sierici determinati per i pazienti affetti da tumore renale e quelli riportati da Villanueva et al. per il tumore della prostata, della vescica e della mammella in stadio avanzato rispetto ai soggetti sani. L'identificazione dei peptidi è stata condotta sulla base degli *score* di Mascot maggiore della soglia di identità (rosso) o di omologia (blu)

La sovrapposizione di molti di essi mette ancor più in evidenza l'importanza di proteasi specifiche nel sangue. Utilizzando criteri robusti per la valutazione della sotto- (*ratio fold change* $\leq 0,5$) o sovra-concentrazione (*ratio fold change* $\geq 1,5$) come già suggerito in letteratura è possibile notare come il *trend* di espressione peptidica tra il tumore renale e le altre tre patologie

mostrano un basso livello di coerenza: il 38% per il carcinoma prostatico, il 19% per il cancro della vescica e il 24% per il tumore della mammella (Tab.17). Questo dato, pertanto, rafforza il fatto che ogni tipologia tumorale potrebbe essere caratterizzata da un pattern degradativo tipico.

In accordo con quanto riportato dalla letteratura per altre patologie anche per il tumore renale sono state identificate alterazioni di frammenti della catena alfa del Fibrinogeno. Oltre a quelle già illustrate precedentemente, lo ione con Mr 2422,07, che potrebbe corrispondere allo stesso segnale a m/z 2423 nello spettro MALDI-TOF in modalità lineare, e a m/z 2423,09 osservato in MALDI-TOF modalità *reflectron* è stato identificato come un frammento della catena alfa del Fibrinogeno con sequenza compresa tra Phe583 e Met603.

Il segnale era sovra espresso nel siero di pazienti affetti da RCC e questo potrebbe riflettere l'aumentata espressione ed attivazione di proteasi specifiche per questo tipo di tumore. Tale risultato è però in contrasto con lo studio di Villanueva in cui è stata riscontrata una diminuzione del 60% di questi peptidi.

Gli altri quattro peptidi identificati (Mr 2377,20, 2601,30, 4088,02 e 4150,17) erano tutti sottoespressi nei pazienti con tumore e derivano dal Complemento C4-A o B, dalla catena alfa del fattore XIII della coagulazione (F13A), dalla Protrombina (THRB) e dall'inibitore della proteasi plasmatica C1 (IC1).

La ridotta espressione di entrambi i fattori C4-A e B non risulta inaspettata in stati patologici in quanto entrambi rivestono ruoli rilevanti nella clearance immunitaria, la memoria e le funzioni effettrici della risposta immunitaria umorale²⁰⁸. La carenza di C4-A e B è stata già studiata in letteratura ed è spesso associata non solo a tumori ma anche a malattie infettive e autoimmuni²⁰⁹.

Sono stati inoltre osservati, nell'analisi LC-ESI MS/MS, due ioni aventi Mr 2601,30 e 2602,16 a diversi tempi di ritenzione (RT) che potrebbero corrispondere al segnale a m/z 2603 rilevato nello spettro MALDI-TOF in modalità lineare, come precedentemente descritto. Entrambi i peptidi risultavano sotto-rappresentati nei pazienti con tumore. Il peptide a Mr 2601,30 è un frammento del propeptide di F13A, il peptide attivo rilasciato dalla Trombina, mentre il secondo è un frammento della Serglicina (SRGN), un proteoglicano con sequenze ripetute di un dipeptide costituito da serina e glicina e caratterizzato da molte catene laterali di glicosamminoglicani. Il *core* della Serglicina è costituito da 158 residui amminoacidici che formano tre domini funzionali e il peptide identificato in questo studio fa parte del dominio di glicosilazione²¹⁰. Ueda *et al.* hanno già riportato in letteratura che il livello serico di questi due peptidi, così come di altri peptidi inclusi nei clusters discriminanti, sono alterati in pazienti affetti da adenocarcinoma del polmone, ma non hanno fornito informazioni circa la loro concentrazione²¹¹. Lo ione a m/z 4089

rilevato in MALDI-LM è stato identificato come la catena leggera della Trombina, tagliata dal fattore Xa. Il complesso formato dal fattore Xa, dal fattore Va, dai fosfolipidi e dagli ioni calcio viene chiamato protrombinasi in quanto permette la trasformazione della protrombina in trombina²¹². Il C1 inibitore (o inibitore della C1 esterasi, C1-INH) è un inibitore delle proteasi seriniche (SERum Protease INhibitors, SERPIN) che blocca l'attività di C1r, C1s, MASP-1 e MASP-2 nel sistema del complemento, fattore XII e callicreina nel sistema di contatto, fattore XI e trombina nel sistema della coagulazione, attivatore tissutale del plasminogeno (tPA) e plasmina nel sistema fibrinolitico. In letteratura è stato già dimostrato come la carenza di questo inibitore è connesso a patologie con quadro clinico caratterizzato da angioedema²¹³.

Tra gli altri cinque segnali, non implicati nei processi di coagulazione del sangue o nel sistema del complemento, solo per uno ione, a m/z 4052 in MALDI-LM, è stata ottenuta una possibile identificazione come frammento di AKA11 (*score* di 31, sopra la soglia di omologia pari a 26). Altre due sequenze amminoacidiche sono state fornite dal database di ricerca con uno *score* più basso e possibili modificazioni variabili (ZC12C con uno *score* di 26 e UBP2 con uno *score* di 25). Le AKAP (*A-Kinase Anchor Proteins*) sono un gruppo di proteine citoplasmatiche capaci di legare la sub unità regolatoria della *Protein Chinasi A* (PKA) e AKA11 risulta essere espressa in molti organi tra cui il rene²¹⁴.

5.3 Espressione proteica

I valori dei *ratios* dei loro peptidi determinati mediante approccio *label-free*, identificati con uno *score* maggiore della soglia di identità o di estesa omologia (>37) hanno permesso di individuare alcune proteine i cui frammenti presentano valori di abbondanza coerenti tra loro, ovvero tutti up o down rappresentati. Per questo sembrerebbe ragionevole ipotizzare che il trend di espressione peptidica possa riflettere la concentrazione serica della proteina e in alcuni casi essere correlata al loro significato biologico.

Attraverso i risultati proposti da Villanueva *et al.* sia da quelli ottenuti dalle analisi oggetto di tesi si potrebbero in questo modo valutare le tendenze di espressione di alcune proteine e la generazione di uno specifico cluster di frammenti per ogni tipo di tumore²¹⁵.

Accession (UNIPROT)	Description	Protein ratio (patients/controls)
SDPR_HUMAN	Serum deprivation-response protein OS=Homo sapiens GN=SDPR PE=1 SV=3	0.56
THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	9.39
ZYX_HUMAN	Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	0.31
SRGN_HUMAN	Serglycin OS=Homo sapiens GN=SRGN PE=1 SV=3	0.52
TMSL3_HUMAN	Thymosin beta-4-like protein 3 OS=Homo sapiens GN=TMSL3 PE=2 SV=1	0.47

Tabella 17. *Ratio* dell'espressione differenziale delle proteine, non citate da Villanueva *et al.*, aventi un *trend* di espressione coerente con i loro derivanti peptidi endogeni determinati mediante approccio label-free. L'identificazione dei peptidi è stata condotta sulla base degli *score* di *Mascot* maggiori di 37 (soglia per l'identità o l'omologia estesa). Il *ratio* è stato calcolato tra i due pools di siero di 80 controlli e 62 pazienti ccRCC sommati a 18 pazienti non-ccRCC.

Delle proteine elencate nella tabella 18 solo tre di queste, il *ratio* proteico è in accordo con quanto riportato in letteratura.

In particolare, la Protrombina (THRB) ha mostrato un'elevata over-espressione e quasi tutti i nove peptidi, identificati e quantificati come frammenti della Protrombina mediante approccio LF, avevano un *ratio* peptidico molto alto, che correla con un maggiore valore di concentrazione nei pazienti coerentemente con il relativo *ratio* proteico (Tab. 18).

La protrombina (fattore II) è una glicoproteina di 72 KDa è il precursore inattivo della Trombina ed è formata da una unica catena polipeptidica. Questa proteina è costituita da quattro regioni: il frammento 1 (Ala44- Arg198), il frammento 2 (Ser199-Arg327), una catena leggera (Thr328-Arg363) ed unacatena pesante (Ile364-Glu622). La trasformazione in trombina avviene attraverso una serie di reazioni catalizzate dal complesso costituito dal fattore Xa, dal fattore Va, dai fosfolipidi e dagli ioni calcio. Il fattore Xa scinde la protrombina in due parti: una contenete i frammenti 1 (Ala44- e 2 l'altra costituita dalla pretrombina 2. La trombina deriva dalla protrombina 2 quando questa viene scissa dal fattore Xa in due polipeptidi^{217,212}.

Tutti i nove frammenti identificati in questo studio coprono solo la regione proteica compresa tra Arg314 e Arg363. Questo è in accordo con i possibili siti di taglio del processo di attivazione della Protrombina. È risaputo che sulla superficie della membrana fosfolipidica cellulare che lega la sequenza N-terminale della Protrombina, il Fattore Xa rimuove il peptide di attivazione e taglia la parte rimanente in catena leggera e pesante a livello di Arg327 e Arg363. Tuttavia non è ancora chiaro se gli altri peptidi più piccoli, con taglio dopo Arg314, possano essere rilasciati nel naturale processo di coagulazione del sangue. Nello studio oggetto di tesi sono stati trovati frammenti over-espresi nel siero dei pazienti affetti da tumore renale, alcuni dei quali coprono quasi l'intera catena leggera. Inoltre altri quattro frammenti tra quelli presenti in maggiore concentrazione nei pazienti, coprono l'intera sequenza del frammento 2, da Arg314 a Arg327. È da evidenziare anche il fatto che livelli sierici aumentati del frammento di Protrombina con sequenza T315ATSEYQTFNPR327 (segnale a *m/z* 1561,77) sono stati già descritti in tumori

della mammella al IV stadio. Inoltre, in un recente lavoro pubblicato è stata riportata una concentrazione aumentata nelle urine di pazienti affetti da carcinoma della vescica della forma intatta della proteina, che è stata inclusa in un pannello proteico diagnostico a causa della sua grande efficacia nella discriminazione della popolazione dei pazienti da quella dei controlli.

Anche la proteina citosolica SDPR (*Serum Deprivation-Response Factor*) è stata individuata come sotto-espressa (Tab. 18) come riportato in diversi studi di neoplasia. SDPR si lega alla fosfatidilserina è la principale proteina di legame presente nelle piastrine. Inoltre, SDPR lega specificamente PKC e svolge un ruolo importante nel fissare le PKC alle caveolae²¹⁸.

Il gene che codifica per questa proteina è stato associato alla capacità delle cellule normali di controllare la crescita di cellule tumorali adiacenti. È stato inoltre dimostrato che SDPR è espresso in maniera estesa in tessuti umani, ma risulta essere diminuito in molte cellule tumorali, comprese quelle della prostata, del rene e dell'epitelio luminale della mammella. In aggiunta, il gene SDPR si trova sul cromosoma 2q33246, che ha un'alta probabilità di perdita di eterozigosità in alcuni tumori del polmone²¹⁹. In conclusione, la proteina SRBC (*SDR-related gene product that binds c-kinase*) è anch'essa indotta dalla deprivazione del siero ed associata alla Protein Chinasi C248. L'espressione di SRBC risulta essere soppressa nei carcinomi del polmone e della mammella apparentemente a causa di una delezione genica o di una metilazione. Inoltre dati recenti suggeriscono il potenziale uso di SDPR come biomarcatore che porta alla diminuzione di Fhl1: è stato trovato infatti che l'espressione genica di SDPR e di FHL1 è ridotta in maniera significativa in tumori della mammella ($p < 0,001$ e $p < 0,02$), del rene ($p < 0,01$) e della prostata ($p < 0,05$). La soppressione dell'espressione di SDPR e Fhl1 è guidata da Src in maniera indipendente da Protein Chinasi Mitogeno-Attivate (MAPK). Infatti gli autori hanno stabilito che Fhl1 induce l'espressione di SDPR in cellule Src-trasformate e che, come Fhl1, Src utilizza Cas (*Crk-associated substrate*) per sopprimere l'espressione di SDPR. Questi risultati sono coerenti con altri lavori pubblicati in letteratura^{220,221}.

In accordo con le conoscenze attuali relativamente alla sua espressione nei tumori anche per la proteina Zyxin può essere riportata una sotto-espressione sulla base della concentrazione dei suoi sei frammenti identificati (Tab. 18).

Zyxin è un componente del complesso adesione focale e svolge un ruolo centrale nella polimerizzazione del filamento di actina in cellule di mammifero²²². Varie evidenze dimostrano che zyxin può funzionare per reclutare componenti necessari per l'assemblaggio dell'actina a siti specifici della cellula e stimolare la polimerizzazione dell'actina stessa.

È interessante notare che zyxin mostra un segnale funzionale che parte dal nucleo e arriva ai siti di adesione cellulare. Hervy *et al.* hanno inoltre suggerito che tale proteina abbia un ruolo nella promozione dell'apoptosi attraverso il legame al ciclo cellulare e la Proteina 1 regolatrice dell'apoptosi (CARP-1) tuttavia il suo ruolo nella regolazione dell'apoptosi è controverso²²³.

La ridotta espressione di questa proteina è stata associata con la progressione del cancro della vescica, ed è stato mostrato che Zyxin agisce come un soppressore tumorale nel sarcoma di Ewing e in cellule di cancro alla prostata. Sebbene questi risultati indicano un'attività soppressiva del tumore, le vie di segnalazione e i meccanismi molecolari connessi alla soppressione del tumore rimangono in gran parte non chiari²²⁴. Un lavoro recente condotto mediante RT-PCR, *Western-Blot* riporta che Zyxin è un target genico diretto di miR-16, un miRNA particolarmente up-regolato nelle cellule del tumore alla laringe quando confrontate con tessuto normale. MiR-16, legando Zyxin, potrebbe avere una profonda influenza sull'adesione e sulla migrazione cellulare. Questi risultati sono in accordo con il fatto che l'ablazione genetica della Zyxin causa la mislocalizzazione di Mena (*Mannheimia succiniciproductens*) e VASP (*vasodilator stimulated phosphoprotein*), l'aumentata motilità e il deficit nel rimodellamento dell'actina²²⁵.

E' stato studiato che Zyxin partecipa alle dinamiche dell'actina legando le VASP, un'interazione che avviene attraverso le sequenze ripetute ricche di Prolina dell'N terminale dell'ActA. Dato che il legame di Zyxin a molte molecole, attraverso i domini LIM, necessita di fosforilazione, sembra probabile che la fosforilazione possa alterare l'interazione testa-coda e, quindi, l'attività di Zyxin. I mutanti puntiformi di Zyxin ad un sito di fosforilazione noto, la serina, presentano un'alterazione dell'abilità del frammento della proteina di legare direttamente un altro frammento del dominio LIM. Questi dati suggeriscono che la fosforilazione di Zyxin a livello della Ser142 porti alla perdita dell'interazione testa-coda, causando un cambiamento nell'attività di Zyxin nei contatti cellula-cellula²²⁶. È interessante notare che sono stati identificati, solo nei campioni di siero dei controlli, due frammenti endogeni di Zyxin contigui nella sequenza e che sono stati tagliati a livello del sito di fosforilazione Ser142-Ser143: un frammento a Mr 3426,70 corrispondente alla sequenza E11EIFSPPPPPEEEGGPEAPIPPPPQPREKVS142, identificato con uno *score* di 28 e sopra la soglia di omologia ed un secondo a Mr 3008,50 avente sequenza S143IDLEIDSLSSLLDDMTKNDFKARVS169. Recentemente Sansing *et al.* hanno studiato gli effettori delle integrine nel carcinoma orale²²⁷ utilizzando un approccio proteomico basato sulla spettrometria di massa. I suddetti autori hanno determinato che una riduzione dell'espressione di Zyxin porta ad una modulazione della migrazione cellulare, una ridotta resistenza nelle linee cellulari tumorali al cisplatino e una diminuzione della proliferazione. Tra

gli effettori delle integrine studiati in questo lavoro²²⁷, la riduzione mirata della Talina ha prodotto il maggiore impatto sui processi adesione-dipendenti, inclusi l'invasione, la migrazione e la chemioresistenza. In accordo con i dati ottenuti, è stato identificato il frammento F409GLEGDEESTMLEDSVSPKKSTVLQ433 (Mr 2725,30), presente solo nei campioni dei controlli, appartenente alla Talina-1 (TLN1).

Per altre due proteine identificate riportate nella Tabella 18, invece i livelli di espressione sono in disaccordo con quanto riportato in letteratura.

Il segnale a Mr 2602 Da è stato identificato come frammento della proteina Serglicina ed è stata trovata sotto-espressa mostrando un *ratio* in disaccordo con i dati presenti in letteratura. La Serglicina (SRGN) è una proteina facente parte della classe dei proteoglicani meglio nota come granulo di cellule ematopoietiche²²⁷. È stato trovato che la serglicina è sintetizzata da cellule endoteliali, è localizzato nelle vescicole citoplasmatiche ed è secreto costitutivamente²²⁸.

La struttura di questa proteina presenta la ripetizione di dipeptidi Ser-Gly e viene modificata con diversi tipi di catene laterali glicosaminoglicani. La Serglicina intracellulare è responsabile della conservazione e della secrezione di proteasi, chemochine, citochine legandosi a questi fattori nei granuli secretori²²⁹. La Serglicina extracellulare è stata trovata sovra-espressa in diversi tipi di cellule tumorali umane, ed è in grado di promuovere la metastasi di cellule di carcinoma nasofaringeo. Diversi studi hanno confermato che Serglycin promuove la metastasi del cancro a seguito della glicosilazione della core-protein, che può essere ottenuto mediante meccanismi di secrezione autocrina e paracrina.

L'altra proteina identificata come *Timosin β -4-Like Protein* (TMSL3) è stata trovata sotto-espressa: anche questo risultato è in contrasto con quanto pubblicato in un recente studio. È stato dimostrato infatti, che l'aumento di TMSL3 promuove la tumorigenicità ed il potenziale metastatico, inoltre sembra indurre la farmacoresistenza delle cellule maligne²³⁰.

Questi dati discordanti possono essere spiegati con il fatto che siccome poiché il presente lavoro è stato svolto a livello peptidomico, queste differenze potrebbero essere ascritte alla presenza di peptidi condivisi tra isoforme proteiche multiple aventi diversa abbondanza o alla presenza di modificazioni post-traduzionali della proteina che portano ad una diversa concentrazione del peptide modificato^{231,232}.

In conclusione, i risultati di questo lavoro di tesi mostrano la possibilità di costruire dei modelli discriminativi sulla base di cluster di peptidi serici di elevata accuratezza che permettano di distinguere non solo i pazienti dai controlli ma soprattutto i tumori renali benigni da quelli maligni e dai soggetti sani.

La caratterizzazione mediante metodica nLC-ESI MS/MS del siero purificato con biglie magnetiche ha permesso la costruzione di un database di peptidi endogeni che consente di aumentare la conoscenza e l'informazione circa l'identità delle specie presenti.

Inoltre l'approccio LC-MS *label-free* si è rivelato un valido metodo per la verifica dei risultati ottenuti mediante strategia di *profiling* MALDI-TOF applicata alla ricerca di biomarcatori. E' stata possibile, inoltre, l'individuazione di proteine con espressione alterata che potrebbero fornire importanti informazioni sulla patogenesi tumorale.

6.BIBLIOGRAFIA

1. Anatomia dell'uomo; AA.VV; Casa Editrice Edi-Ermes, Milano 2004.
2. Anatomia Umana; Martini F.H./Timmons M.S./Tallitsch R.B. Edises,;2010
3. Anatomia Umana Saladin Kenneth S./De Caro R.; Piccin; 2011
4. Fisiologia Ed. Terza Klink R.; Pape H.; Kurtz A.; Silbernagl S., Edises, 2012
5. I principi di Biochimica di Lehninger; Zanichelli; 2006
6. Le Basi Farmacologiche Della Terapia; Goodman & Gilman; Bruton L. / Chabner B.A. / Knollmann B.C., Zanichelli, 2010
7. Elementi di patologia generale; Pontieri G.M. Piccin, 2011,
8. Patologia generale, AA.VV; Casa Editrice Idelson-Gnocchi, 2012
9. Ferlay J, Bray F, Pisani P, Parkin DM. GLOBOCAN 2002 Cancer Incidence, Mortality and Prevalence Worldwide IARC Cancer Base No. 5, version 2.0 IARC Press, Lyon 2004
10. AIRT Working Group. Italian cancer figure - Report 2006 - Incidence, mortality and estimates. Epidemiologia e prevenzione 2006; 30(1)
11. Raspollini MR, Messerini L, Taddei GL. Long-term outcome of renal transplantation from older donors. New Engl J Med 2006; 354:2071-2072
12. Hock LM, Lynch J, Balaji KC. Increasing incidence of all stages of kidney cancer in the last 2 decades in the United States: an analysis of surveillance, epidemiology and end results program data. J Urol 2002; 167(1):57-60
13. Lopez-Beltran, Scarpelli, M.; Montironi; Kirkali; 2004 WHO classification of the renal tumors of the adults. Eur Urol 2006, 49, (5), 798-805
14. Zbar, B.; Lerman, M., Inherited carcinomas of the kidney Adv Cancer Res 1998, 75, 163-201.
15. Eble JN, Sauter G, Epstein JI, Sesterhenn IA. WHO: Tumours of the Urinary System and Male Genital Organs. Lyon: IARC Press; 2004
16. Suzigan S, Lopez-Beltran A, Montironi R, Drut R, Romero A, Ayashi T, Gentili AL, Fonseca PS, deTorres I, Billis A, Japp LC, Bollito E, Algaba F, Requena-Tapias MJ. Multilocular cystic renal cell carcinoma: a report of 45 cases of a kidney tumor of low malignant potential. Am J Clin Pathol 2006 Feb; 125(2):217-22
17. Delahunt B, Eble JN. History of the development of the classification of renal cell neoplasia. Clin Lab Med. Jun 2005;25(2):231-46
18. Decastro GJ, McKiernan JM. Epidemiology, clinical staging, and presentation of renal cell carcinoma. Urol Clin North Am. Nov 2008;35(4):581-92;

19. Pascual D, Borque A. Epidemiology of kidney cancer. *Adv Urol*. 2008.
20. Patard JJ, Leray E, Rioux-Leclercq N, et al. Prognostic value of histologic subtypes in renal cell carcinoma: a multicenter experience. *J Clin Oncol*. Apr 20 2005;23(12):2763-71
21. HuJ, Mao Y, White K; Canadian Cancer Registries Epidemiology Research Group. Overweight and obesity in adults and risk of renal cell carcinoma in Canada. *Soz Praventivmed* 2003
22. PrestiJC Jr, Rao PH, Chen Q, et al. Histopathological, cytogenetic and molecular characterization of renal cortical tumors. *Cancer Res*1991; 51: 1544–1552.
23. PatardJJ, Leray E, Rioux-Leclercq N, et al. Prognostic value of histologic subtypes in renal cell carcinoma a multicenter experience. *J Clin Oncol* 2005
24. DelahuntB, Eble JN, McCredie MR, Bethwaite PB, Stewart JH, Bilous AM Morphologic typing of papillary renal cell carcinoma: comparison of growth kinetics and patient survival in 66 cases
25. KiyoshimaK, Oda Y, Nakamura T, et al. Multi-centri c papillary renal cell carcinoma associated with renal adenomatosis. *Pathol* 2004
26. Amin MB, Paner GP, Alvarado-Cabrero , Alvarado-Cabrero I, Young AN, Stricker HJ, Lyles RH, Moch H: Chromophobe Renal Cell Carcinoma: Histomorphologic Characteristics and Evaluation of Conventional Pathologic prognostic Parameters in 145 Cases. *Am J Surg Pathol* 2008
27. Beck SDW, Manish I, Patel IM, Snyder ME, Kattan MW, Motzer RJ, Reuter VE, Russo P: Effect of Papillary and Chromophobe Cell Type on Disease-Free Survival After Nephrectomy for Renal Cell Carcinoma *Ann of Surg Oncol*; 2004,
28. Cheville JC, Lohse CM, Zincke H, Weaver H, Blute AL, Michael L: Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma *Am J Surg Pathol* 2003.
29. Wu SL, Fishman IJ, Shanon RL: Chromophobe Renal Cell Carcinoma With Extensive Calcification and Ossification. *Ann of Diag Pathol* 2002
30. Dhote R et al. Risk factors for adult renal cell carcinoma. *Urol Clin N Am* 2004;
31. World Cancer Research Fund and American Institute for Cancer Research. Food, nutrition, physical activity, and prevention of cancer: a global perspective. American Institute of Cancer Research, Washington, 2007
32. AIRT Working Group. Italian cancer figure - Report 2007 - Survival. *Epidemiologia e prevenzione* anno 2007
33. Dickman PW, Hakulinen T, Luostarinen T, Pukkala E, Sankila R, Söderman B, Teppo L.

- Survival of cancer patients in Finland 1955-1994. *Acta Oncologica* 1999; 38 Supplemento 12:1-103
34. Gago-Dominguez M, Yuan J-M, Castela J E, Ross RK, Yu MC. Regular use of analgesics is a risk factor for renal cell carcinoma *Br J Cancer* 1999; 81(3):542-8
 35. McLaughlin JK, Lipworth L. Epidemiologic aspects of renal cell cancer. *Semin Oncol* 2000; 27(2):115-23
 36. Schlehofer B et al. International renal-cell cancer study. VI. The role of medical and family history. *J Int J Cancer* 1996;
 37. Cohen HT, McGovern FJ. Renal-cell carcinoma. *N Engl J Med* 2005; 353(23):2477-90
 38. Pathology For The Physical Therapist Assistant Goodman C.C. , Fuller K.S. , O'shea R.K. Saunders, 2011
 39. Linehan WM, Vasselli J, Srinivasan R, et al. Genetic basis of cancer of the kidney: disease-specific approaches to therapy *Clin Cancer Res.* 2004
 40. Lee YS, Vortmeyer AO, Lubensky IA, et al. Coexpression of erythropoietin and erythropoietin receptor in von Hippel-Lindau disease-associated renal cysts and renal cell carcinoma *Clin Cancer Res.* 2005
 41. Klatte T, Seligson DB, Riggs SB, et al. Hypoxia-inducible factor 1 alpha in clear cell renal cell carcinoma *Clin Cancer Res.* 2007
 42. Nyhan MJ, O'Sullivan GC, McKenna SL. Role of the VHL (von Hippel-Lindau) gene in renal cancer:a multifunctional tumor suppressor. *Biochem Soc Trans.* 2008
Calzada MJ, Esteban MA, Feijoo-Cuaresma M, et al von Hippel-Lindau tumor suppressor protein regulates the assembly of intercellular junctions in renal cancer cells through hypoxia-inducible factor-independent mechanisms *Cancer Res.* 2006
 43. Ferrara N, Gerber HP. The role of vascular endothelial growth factor in angiogenesis *Acta Haematol.*2001
 44. Rivet J, Mourah S, Murata H, et al. VEGF and VEGFR-1 are coexpressed by epithelial and stromal cells of renal cell carcinoma. *Cancer.* 2008
 45. Schips L, Dalpiaz O, Lipsky K, et al. Serum levels of vascular endothelial growth factor (VEGF) and endostatin in renal cell carcinoma patients compared to a control group *Eur Urol.*2007
 46. Wang X, Hawk N, Yue P, et al. Overcoming mTOR inhibition induced paradoxical activation of survival signaling pathways enhances mTOR inhibitors' anticancer efficacy *Cancer Biol Ther.* 2008
 47. Costa LJ, Drabkin HA. Renal cell carcinoma: new developments in molecular biology and

- potential for targeted therapies. *Oncologist*.2007
48. Russell RC, Ohh M. The role of VHL in the regulation of Ecadherin a new connection in an old pathway. *Cell Cycle* 2007
 49. Haase VH, Glickman JN, Socolovsky M, Jaenisch R. Vascular tumors in livers with targeted inactivation of the von Hippel-Lindau tumor suppressor *Proc Natl Acad Sci USA* 2001
 50. Brunelli M, Eble JN, Zhang S, Martignoni G, Cheng L. Gains of chromosomes 7, 17, 12, 16, and 20 and loss of Y occur early in the evolution of papillary renal cell neoplasia: a fluorescent in situ hybridization study. *Mod Pathol* 2003 Oct
 51. Jiang F, Richter J, Schraml P, Bubendorf L, Gasser T, Sauter G et al. Chromosomal imbalances in papillary renal cell carcinoma: genetic differences between histological subtypes. *Am J Pathol* 1998 Nov
 52. Stewart L, Glenn G, Toro JR. Cutaneous leiomyomas: a clinical marker of risk for hereditary leiomyomatosis and renal cell cancer. *Dermatol Nurs* 2006 Aug
 53. Brunelli M, Eble JN, Zhang S, Martignoni G, Delahunt B, Cheng L. Eosinophilic and classic chromophobe renal cell carcinomas have similar frequent losses of multiple chromosomes from among chromosomes 1, 2, 6, 10, and 17, and this pattern of genetic abnormality is not present in renal oncocytoma. *Mod Pathol* 2005 Feb
 54. Pavlovich CP, Walther MM, Eyler RA, Hewitt SM, Zbar B, Linehan WM et al. Renal tumors in the Birt-Hogg-Dube syndrome. *Am J Surg Pathol* 2002 Dec
 55. Warren MB, Torres-Cabala CA, Turner ML, Merino MJ, Matrosova VY, Nickerson ML et al. Expression of Birt-Hogg-Dube gene mRNA in normal and neoplastic human tissues. *Mod Pathol* 2004 Aug
 56. Minardi D, Lucarini G, Filosa A, Milanese G, Zizzi A, Di Primio R, Montironi R, Muzzonigro G. Prognostic role of global DNA-methylation and histone acetylation in pT1a clear cell renal carcinoma in partial nephrectomy specimens. *J Cell Mol Med*. 2009 Aug;13
 57. Seligson DB, Horvath S, McBrien MA, Mah V, Yu H, Tze S, Wang Q, Chia D, Goodglick L, Kurdiani SK. Global levels of histone modifications predict prognosis in different cancers. *Am J Pathol*. 2009 May;
 58. McDougal WS, Garnick MB. Clinical signs and symptoms of renal cell carcinoma. In NJ Vogelzang, PT Scardino, WU Shipley and DS Coffey, Editors. *Comprehensive Textbook of Genitourinary Oncology*. Williams and Wilkins Co, Baltimore 1995;
 59. Lee CT, Katz J, Fearn PA, Russo P. Mode of presentation of renal cell carcinoma provides prognostic information. *Urol Oncol* 2002; 7(4):135-40
 60. Kim HL, Belldgrun AS, Freitas DG, Bui MTH, Han KR, Dorey FJ, Figlin RA. Paraneoplastic

- signs and symptoms of renal cell carcinoma: implications for prognosis. *Clin Urol* 2003
61. Sugimura K, Ikemoto SI, Kawashima H, Nishisaka N, Kishimoto T. Microscopic hematuria as a screening marker for urinary tract malignancies. *Int J Urol* 2001
 62. Ou Yc, Yang CR, Ho HC, Cheng CL, Kao YL, Su CK, Chiu KY, Chen WM. The symptoms of renal carcinoma related to patients' survival. *J Clin Med Assoc* 2003; 66:537-543
 63. Odell WD. Paraendocrine syndromes of cancer. *Adv Intern Med* 1989; 34:325-352
 64. Boxer RJ, Waisman J, Lieber MM, et al. Non-metastatic hepatic dysfunction associated with renal carcinoma. *J Urol* 1978; 119: 468-471
 65. Bukowski RM. Prognostic factors for survival in metastatic renal cell carcinoma: update 2008. *Cancer* 2009; 115: 2273-2281
 66. Okubo Y, Yonese J, Kawakami S, Yamamoto S, Komai Y, Takeshita H, Ishikawa Y, Fukui I. Obstinate cough as a sole presenting symptom of non -metastatic renal cell carcinoma. *Int J Urol* 2007
 67. Chasan SA, Pothel RL, Huben RP. Management and prognostic significance of hypercalcemia in renal cell carcinoma. *Urology* 1989; 33:167-171
 68. Swan SK, Keane WF. Clinical evaluation of renal function. In Greenberg A ed: *Primer on kidney diseases*. 3rd edition. New York, National Kidney Foundation 2001
 69. Aben KK, Luth TK, Janssen-Heijnen ML, et al. No improvement in renal cell carcinoma survival: a population-based study in the Netherlands. *Eur J Cancer* 2008; 44: 1701-1709
 70. Smith SJ, Bosniak MA, Megibow AJ, Hulnick DH, Horii SC, Raghavendra BN. Renal cell carcinoma: earlier discovery and increased detection. *Radiology* 1989 Mar; 170(3 Pt 1):699-703
 71. Tammela TL, Leinonen AS, Kontturi MJ. Comparison of excretory urography, angiography, ultrasound and computed tomography for T category staging of renal cell carcinoma. *Scand J Urol Nephrol* 1991; 25(4):283-6
 72. Francis IR. Detection, staging and surveillance in renal cell carcinoma. *Cancer Imaging* 2006 Nov 8; 6:168-74
 73. Kitamura H, Fujimoto H, Tobisu K, Mizuguchi Y, Maeda T, Matsuoka N, Komiyama M, Nakagawa T, Kakizoe T. Dynamic computed tomography and color Doppler ultrasound of renal parenchymal neoplasms: correlations with histopathological findings. *Jpn J Clin Oncol* 2004 Feb; 34(2):78-81
 74. Jamis-Dow CA, Choyke PL, Jennings SB, Linehan WM, Thakore KN, Walther MM. Small (≤ 3 cm) renal masses: detection with CT versus US and pathologic correlation. *Radiology* 1996 Mar; 198(3):785-8

75. Coll, DM, Smith RC. Update on radiological imaging of renal cell carcinoma. *BJU Int* 2007; 99(5b):1217-22
76. Heidenreich A, Ravery V. Preoperative imaging in renal cell cancer. *World J Urol* 2004; 22:307-315
77. Ho VB, Allen SF, Hood MN, Choyke PL. Renal masses: quantitative assessment of enhancement with dynamic MR imaging. *Radiology* 2002; 224:695-700
78. Reznick RH. Imaging in the staging of renal cell carcinoma. *Eur Radiol* 1996; 6(2):120-8
79. Greene FL, Page D, Fleming ID et al. (eds) (2002) *AJCC cancer staging manual*. 6th ed. Springer-Verlag, New York
80. Ficarra V, Galfano A, Mancini M, Martignoni G, Artibani W. TNM staging system for renal-cell carcinoma: current status and future perspectives. *Lancet Oncol* 2007
81. Ficarra V, Guille F, Schips L, de la Taille A, Prayer Galetti T, Tostain J, Cindolo L, Novara G, Zigeuner R, Bratti E, Li G, Altieri V, Abbou CC, Zanolla L, Artibani W, Patard JJ. Proposal for revision of the TNM classification system for renal cell carcinoma. *Cancer* 2005; 104(10):2116-23
82. Frank I, Blute ML, Leibovich BC, Chevillet JC, Lohse CM, Kwon ED, Zincke H. pT2 classification for renal cell carcinoma. Can its accuracy be improved? *J Urol* 2005; 173(2):380-4
83. Klatt T, Patard JJ, Goel RH, Kleid MD, Guille F, Lobel B, Abbou CC, De La Taille A, Tostain J, Cindolo L, Altieri V, Ficarra V, Artibani W, Prayer-Galetti T, Allhoff EP, Schips L, Zigeuner R, Figlin RA, Kabbinavar FF, Pantuck AJ, Belldegrun AS, Lam JS. Prognostic impact of tumor size on pT2 renal cell carcinoma: an international multicenter experience. *J Urol* 2007; 178(1):35-40
84. Ficarra V, Novara G, Iafrate M, Cappellaro L, Bratti E, Zattoni F, Artibani W. Proposal for reclassification of the TNM staging system in patients with locally advanced (pT3-4) renal cell carcinoma according to the cancer-related outcome. *Eur Urol* 2007;
85. Terrone C, Cracco C, Guercio S, Bollito E, Poggio M, Scoffone C, Tarabuzzi R, Porpiglia F, Scarpa RM, Fontana D, Rocca Rossetti S. Prognostic value of the involvement of the urinary collecting system in renal cell carcinoma. *Eur Urol* 2004 Oct; 46(4):472-6
86. Klatt T, Chung J, Leppert JT, Lam JS, Pantuck AJ, Figlin RA, Belldegrun AS. Prognostic relevance of capsular involvement and collecting system invasion in stage I and II renal cell carcinoma. *BJU Int* 2007
87. Delahunt B, Sika-Paotonu D, Bethwaite PB, McCredie MR, Martignoni G, Eble JN, Jordan TW. Fuhrman grading is not appropriate for chromophobe renal cell carcinoma. *Am J Surg*

Pathol 2007; 31(6):957-60

88. Karakiewicz PI, Trinh QD, Bhojani N, Bensalah K, Salomon L. et al. Renal cell carcinoma with nodal metastases in the absence of distant metastatic disease: prognostic indicators of diseasespecific survival. *Eur Urol* 2007
89. Palapattu GS, Pantuck AJ, Dorey F, Said JW, Figlin RA, Belldegrun AS. Collecting system invasion in renal cell carcinoma: impact on prognosis and future staging strategies. *J Urol* 2003
90. Sorbellini M, Kattan MW, Snyder ME, Reuter V, Motzer R, Goetzl M, McKiernan J, Russo P. A postoperative prognostic nomogram predicting recurrence for patients with conventional clear cell renal cell carcinoma. *J Urol* 2005
91. Lam JS, Shvarts O, Said JW, Pantuck AJ, Seligson DB, Aldrige ME et al. Clinicopathologic and molecular correlations of necrosis in the primary tumor of patients with renal cell carcinoma. *Cancer* 2005; 103(12):2517-25
92. Bui MH, Seligson D, Han KR, Pantuck AJ, Dorey FJ, Huang Y, Horvath S, Leibovich BC, Chopra S, Liao SY, Stanbridge E, Lerman MI, Palotie A, Figlin RA, Belldegrun AS. Carbonic anhydrase IX is an independent predictor of survival in advanced renal clear cell carcinoma: implications for prognosis and therapy. *Clin Cancer Res* 2003; 9(2):802-11
93. Frank I, Blute ML, Cheville JC, Lohse CM, Weaver AL, Zincke H. An outcome prediction model for patients with clear cell renal cell carcinoma treated with radical nephrectomy based on tumor stage, size, grade and necrosis: the SSIGN score. *J Urol* 2002; 168(6):2395-400
94. Sengupta S, Lohse CM, Leibovich BC, Frank I, Thompson RH, Webster WS, Zincke H, Blute ML, Cheville JC, Kwon ED. Histologic coagulative tumor necrosis as a prognostic indicator of renal cell carcinoma aggressiveness. *Cancer* 2005; 104(3):511-20
95. Robson CJ, Churchill BM, Anderson W. The results of radical nephrectomy for renal cell carcinoma. *J Urol* 1969; 101:297-301
96. Ljungberg B, Hanbury DC, Kuczyk MA, Merseburger AS, Mulders PFA, Patard J, Sinescu IC. Renal Cell Carcinoma Guideline. *Eur Urol* 2007; 51(6):1502-10
97. Wotkowicz C, Libertino A. Renal cell cancer: radical nephrectomy. *BJU int* 2007; 99:1231-1238
98. Makhoul B, De La Taille A, Vordos D, Salomon L, Sebe P, Audet JF, Ruiz L, Hoznek A, Antiphon P, Cicco A, Yiou R, Chopin D, Abbou CC. Laparoscopic radical nephrectomy for T1 renal cancer: the gold standard? A comparison of laparoscopic vs open nephrectomy. *BJU Int* 2004; 93(1):67-70

99. Carmignani G, Simonato A, Corbu C, Traverso T, Ruggiero G, Pizzorno P, Vota V, Martorana G, Belgrano E, Romagnoli E. Open Radical Nephrectomy: 35 Years of Experience at the "Luciano Giuliani" Urological Department of the University of Genoa. *Eur Urol Suppl* 2006; 5(8):593-599
100. Wilson D, Hiller L, Gray L, Grainger M, Stirling A, James N. The effect of biological effective dose on time to symptom progression in metastatic renal cell carcinoma. *Clin Oncol* 2003; 15:400-07
101. Harris DT. Hormonal therapy and chemotherapy of renal-cell carcinoma. *Semin Oncol* 1983; 10:422-430
102. Ruiz JA, Raftopoulos H, Petrylak DP. Chemotherapy for metastatic renal-cell carcinoma (RCC). In Bukoski RM, Novick AC, editors. *Renal cell carcinoma: molecular biology, immunology, and clinical management*. Totowa, NJ. Humana Press 2000; 283-300
103. Amato RJ. Renal cell carcinoma: review of novel single-agent therapeutics and combination regimens. *Ann Oncol* 2005; 16:7-15
104. George CM, Stadler WM. The role of systemic chemotherapy in the treatment of kidney cancer. *Cancer Treat Res* 2003; 116:173-182
105. Boccardo F, Guglielmini P, Tacchini L. Medical treatment of advanced renal cell carcinoma: present options and future directions. *Eur Urol suppl* 2006; 5:619-626
106. Hartmann JT, Bokemeyer C. Chemotherapy for renal cell carcinoma. *Anticancer Res* 1999; 19:1541-1543
107. De Mulder PH, van Herpen CM, Mulders PA. Current treatment of renal cell carcinoma. *Ann Oncol* 2004; 15:319-328
108. Ljungberg B, Hanbury DC, Kuczyk MA, Merseburger AS, Mulders PF, Patard JJ, Sinescu IC; European Association of Urology Guideline Group for renal cell carcinoma. Renal cell carcinoma guideline. *Eur Urol* 2007; 51:1502-1510
109. Lawrie LC, Fothergill JE, Murray GI. Spot the differences proteomics in cancer research. *Lancet Oncol* 2001; 2:270-7
110. Cutler P. Protein arrays: the current state-of-the-art *Proteomics* 2003 ; 3:3-18
111. Wilkins MR, et al. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev* 1996;
112. Alessio M, et al. *Atti del Convegno Proteomica: tecnologia per la caratterizzazione del proteoma e lo studio delle interazioni proteina-proteina*, 2001
113. Weissinger, E. M.; Wittke, S.; Kaiser, T.; Haller, H.; Bartel, S.; Krebs, R.; Golovko, I.; Rupprecht, H. D.; Haubitz, M.; Hecker, H.; Mischak, H.; Fliser, D. Proteomic patterns

- established with capillary electrophoresis and mass spectrometry for diagnostic purposes. *Kidney Int* 2004, 65, (6), 2426-34
114. Tremolada L, et al. Characterization of heat shock protein 27 phosphorylation sites in renal cell carcinoma. *Proteomics* 2005; 5: 788
115. Theodorescu D, et al. Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis. *Lancet Oncol* 2006; 7:230
116. Stannard, C. et al. New paradigms in cellular function and the need for top-down proteomics analyses. *Curr. Proteom.* 2004; 1:13–25
117. Link, A. J. Multidimensional peptide separations in proteomics *Trends Biotechnol.* 2002 20: S8–13
118. Godovac-Zimmermann, J., et al. Perspectives in spicing up proteomics with splicing. *Proteomics* 2005; 5:699–709
119. Y. Ge, et al. Top down characterization of larger proteins (45 kDa) by electron capture dissociation mass spectrometry. *J. Am. Chem. Soc.* 2002; 124: 672– 678
120. Kleiner, O. et al. Proteomics of signal transduction pathways. In: *Signalling Pathways in Liver Diseases*. Springer Verlag, Heidelberg, Germany, 2005; 417–431
121. Guerrero, I.C. et al. Application of Mass Spectrometry in Proteomics. *Bioscience Reports* Vol. 25, Nos. 1/2, February/April 2005
122. E. Nägele, et al. Improved 2D Nano-LC/MS for Proteomics Applications: A Comparative Analysis Using Yeast Proteome. *J of Biomol. Tech.* 2004; 15, (2), 134-143
123. Magni F, et al. Expanding the proteome two-dimensional gel electrophoresis reference map of human renal cortex by peptide mass fingerprinting. *Proteomics* 2005; 5: 816
124. Lin Dayin, et al. Large-scale protein identification using mass spectrometry. *Biochimica et Biophysica Acta* 2003;1646.
125. *Analisi chimico farmaceutica* F.Savelli O.Bruno Ed. Piccin, 2005
126. *Fondamenti di chimica analitica*, Skoog, West, Holler, Ed. Edises
127. *Foye's principi di chimica farmaceutica* Lemke T.L. / Williams, D.A., Piccin, 2009
128. *Chimica e propedeutica biochimica* Binaglia, L. / Giardina, B., McGraw-Hill Italia, 2006
129. Merchant M., Weinberger SR. Recent advancements in surface enhanced laser desorption/ionization-time of flight-mass spectrometry. *Electrophoresis* 2000; 21: 164
130. Petricoin EF, et al., SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. *Curr Opin Biotechnol* 2004; 15: 24
131. Girault S, Chassaing G, Blais JC et al. Coupling of MALDI-TOF mass analysis to the separation of biotinylated peptides by magnetic streptavidin beads *Anal Chem* 1996; 68(13):

132. Pusch W, Flocco MT, Leung SM et al. Mass spectrometry-based clinical proteomics. *Pharmacogenomics* 2003;4(4): 463-76
133. Villanueva J, Philip J, Entenberg D et al. Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. *Anal Chem* 2004;
134. Callesen AK, DePont Christensen R, Madsen JS et al. Reproducibility of serum protein profiling by systematic assessment using solid-phase extraction and matrix-assisted laser desorption/ionization mass spectrometry. *Rapid communications in mass spectrometry* 2008;22: 291
135. Jimenez CR, El Filali Z, Knol JC et al. Automated serum peptide profiling using novel magnetic C18 beads offline coupled to MALDI-TOF-MS. *Proteomics Clin Appl* 2007; 1: 1-7
136. Baumann S, Ceglarek U, Fiedler GM et al. Standardize approach to proteome profiling of human serum based on magnetic bead separation and Matrix-Assisted Laser Desorption/Ionization Time-of-Flight mass spectrometry. *Clinical Chemistry* 2005; 51(6): 973-80
137. Righetti PG, Castagna A, Antonioli P et al. Prefractionation techniques in proteome analysis: the mining tools of the third millennium. *Electrophoresis* 2005; 26: 297-319
138. Boschetti E, Righetti PG. The proteoMiner in the proteomics arena: a non-depleting tool for discovering low-abundance species. *Journal of Proteomics* 2008;71: 255-64
139. What is MASS SPECTROMETRY?, ASMS (American Society for Mass Spectrometry)
140. Overberg A., Hassenburger A., Hillenkamp F., *Mass Spectrometry in the Biological Sciences*, Doerdecht: Kluwer Academic, 1999
141. Hanton S.D., Paula A. Cornelio Clark, Owens K.G., *American Society for Mass Spectrometry*, 1999, 10, 104
142. Gaskell Simon J., *Journal of Mass Spectrometry*, 1997, 32, 677
143. Wilm M.S., Mann M., *International Journal of Mass Spectrometry Ion Processes*, 1994, 136, 167
144. Aston, F. W., The mass-spectra of chemical elements. *Phil Mag* 1920, 39: 611-625
145. Yates, J. R., 3rd, Mass spectrometry and the age of the proteome. *J Mass Spectrom* 1998, 33, (1), 1-19
146. Shen, Y.; Smith, R. D., Advanced nanoscale separations and mass spectrometry for sensitive high-throughput proteomics. *Expert Rev Proteomics* 2005, 2, (3), 431-47
147. Jonscher, K. R.; Yates, J. R., 3rd, The quadrupole ion trap mass spectrometer--a small

- solution to a big challenge. *Anal Biochem* 1997, 244, (1), 1-15
148. Marshall, A., Fourier transform ion cyclotron resonance detection: principles and experimental configurations. *International J of Mass Spectrometry* 2002, 215, 59
149. Marvin, L. F.; Roberts, M. A.; Fay, L. B., Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in clinical chemistry. *Clin Chim Acta* 2003, 337, (1-2), 11-21
150. Huy A. Bui, R. Graham Cooks, *Journal of Mass Spectrometry*, 1998, 33, 297
151. R.M. Caprioli, A. Malorni and G. Sindona, *Selected Topics and Mass Spectrometry in the Biomolecular Sciences*, Kluwer Academic Publisher 1997, 103-119
152. Geno P.W., *Mass Spectrometry in the Biological Science*, Kluwer Academic Dordrecht, 1992
153. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. *Science* 1989, 246, (4926), 64-71
154. Wilm, M.; Mann, M., Analytical properties of the nanoelectrospray ion source *Anal Chem* 1996, 68, (1), 1-8
155. Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T., Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal Chem* 1991, 63, (24), 1193A-1203°
156. Katta, V.; Chow, D. T.; Rohde, M. F., Applications of in-source fragmentation of protein ions for direct sequence analysis by delayed extraction MALDI-TOF mass spectrometry. *Anal Chem* 1998, 70, (20), 4410-6
157. Glückmann M, K. M., The initial ion velocity and its dependence on matrix analyte and preparation method in ultraviolet matrix-assisted laser desorption/ionization. *J Mass Spectrom* 1999, 34, 467-477
158. Bakhtiar, R.; Nelson, R. W., Electrospray ionization and matrix-assisted laser desorption ionization mass spectrometry. *Emerging technologies in biomedical sciences. Biochem Pharmacol* 2000, 59, (8), 891-905
159. Marvin, L. F.; Roberts, M. A.; Fay, L. B., Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in clinical chemistry. *Clin Chim Acta* 2003, 337, (1-2), 11-21
160. Wittal, R. M.; Li, L., High resolution matrix-assisted laser desorption/ionization in a linear time-of-flight mass spectrometer. *Anal Chem* 1995, 67, 1950-1954
161. Beavis, R. C.; Chait, B. T., High-accuracy molecular mass determination of proteins using matrix-assisted laser desorption mass spectrometry. *Anal Chem* 1990, 62, (17), 1836-40
162. Cotter, R. J., Time-of-flight mass spectrometry: an increasing role in the life sciences. *Biomed Environ Mass Spectrom* 1989, 18, (8), 513-32

163. Chaurand, P.; Luetzenkirchen, F.; Spengler, B., Peptide and protein identification by matrix-assisted laser desorption ionization (MALDI) and MALDI- postsorce decay time-of-flight mass spectrometry. *J Am Soc Mass Spectrom* 1999, 10, (2), 91-103
164. Dooley, K. C., Tandem mass spectrometry in the clinical chemistry laboratory *Clin Biochem* 2003, 36, (6), 471-81
165. Guerrero, I. C.; Kleiner, O., Application of mass spectrometry in proteomics *Biosci Rep* 2005, 25, (1-2), 71-93
166. R.S. Johnson, S.A. Martin, K. Biemann, *Int. J. Mass Spectrom Ion Process.* 86 (1988) 137–154
167. K. Hakansson, M.J. Chalmers, J.P. Quinn, M.A. McFarland C.L. Hendrickson, A.G. Marshall, *Anal. Chem.* 75 (2003) 3256-3262
168. G. Hopfgartner, E. Varesio, V. Tschappat, C. Grivet, E. Bourgoigne. A. Leuthold, *J. Mass Spectrom.* 39 (2004) 845–855
169. I.V. Chernushevich, A.V. Loboda, B.A. Thomson, *J. Mass Spectrometry* 36 (2001) 849–865
170. Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J., Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 2004, 3, (12), 1154-69
171. Yao, X.; Freas, A.; Ramirez, J.; Demirev, P. A.; Fenselau, C., Proteolytic ¹⁸O labeling for comparative proteomics: model studies with two serotypes of adenovirus *Anal Chem* 2001, 73, (13), 2836-42
172. Bridges, S. M.; Magee, G. B.; Wang, N.; Williams, W. P.; Burgess, S. C.; Nanduri, B., ProtQuant: a tool for the label-free quantification of MudPIT proteomic data. *BMC Bioinformatics* 2007, 8 Suppl 7, S24
173. Ong, S. E.; Mann, M., Stable isotope labeling by amino acids in cell culture for quantitative proteomics. *Methods Mol Biol* 2007, 359, 37-52
174. Bantscheff, M.; Schirle, M.; Sweetman, G.; Rick, J.; Kuster, B., Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* 2007, 389
175. Gafken, P. R.; Lampe, P. D., Methodologies for characterizing phosphoproteins by mass spectrometry. *Cell Commun Adhes* 2006, 13, (5-6), 249-256
176. Asara, J. M.; Christofk, H. R.; Freimark, L. M.; Cantley, L. C., A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in proteomics screen. *Proteomics* 2008, 8, (5), 994-9

177. Sechi, S., *Quantitative Proteomics by Mass Spectrometry*. Totowa, 2007
178. Mueller, L. N.; et al., e., An Assessment of Software Solutions for the Analysis of Mass Spectrometry Based Quantitative Proteomics Data. *J Proteome Res* 2007, 7, 51-61
179. Mueller, L. N.; et al., e., An Assessment of Software Solutions for the Analysis of Mass Spectrometry Based Quantitative Proteomics Data. *J Proteome Res* 2007, 51-61
180. Magni, F.; Van Der Burgt, Y. E.; Chinello, C.; Mainini, V.; Gianazza, E.; Squeo.; Deelder, A. M.; Kienle, M. G., Biomarkers discovery by peptide and protein profiling in biological fluids based on functionalized magnetic beads purification and mass spectrometry. *Blood Transfus* 2010, 8 Suppl 3, s92-7
181. Tsou, C. C.; Tsai, C. F.; Tsui, Y. H.; Sudhir, P. R.; Wang, Y. T.; Chen, Y. J. Chen, J. Y.; Sung, T. Y.; Hsu, W. L., IDEAL-Q, an automated tool for label-free quantitation analysis using an efficient peptide alignment approach and spectral data validation. *Mol Cell Proteomics* 2010, 9, (1), 131-44
182. Chia-Li Han, Jinn-Shiun, CheErr-Cheng, ChanChien-Peng Wu, An Informatics-assisted Label-free Approach for Personalized Tissue Membrane Proteomics: Case Study on Colorectal Cancer *Mol. Cell. Proteomics* 2011 10: M110.003087
183. Cohen HT, McGovern FJ. Renal Cell Carcinoma. *New Engl J Med* 2005; 353: 2477-2490
184. World cancer report. Edited by Stewart B, Kleihues P. IARC Press, International Agency for Research on Cancer 2003. Laber DA. Risk factors, classification and staging of renal cell cancer. *Med Oncol* 2006; 23(4): 443-454
185. Mathew A, Devesa SS, Fraumeni JF Jr et al. Global increases in kidney cancer incidence, 1973-1992. *Eur J Cancer Prev* 11: 171-178, 2002
186. Parkin CM, Whelan SL, Frelay J et al. Cancer Incidence in five continents, Volume VIII. IARC Scientific Publication No. 155. Lyon, France International Agency for Research on Cancer 2002
187. Jemal A, Murray T, Ward E et al. Cancer statistics, 2005. *CA Cancer J Clin* 2005 ; 55 : 10-30
188. Linehan WM, Yang JC, Bates SE. Cancer of the kidney. In: DeVita SE, Hellman S, Rosenberg SA, editors. *Cancer Principle and practice of oncology*. Philadelphia: Lippincott Williams and Wilkins: 2005; 1139-1168.
189. Linehan WM, Pinto PA, Srinivasan R et al. Identification of the genes for kidney cancer: opportunity for disease specific targeted therapeutics. *Clin Cancer Res* 2007; 13(2 Suppl): 671-679
190. Eble JN, Sauter G, Epstein J et al. Pathology and Genetics of Tumours of the Urinary

Systems and Male Genital Organs. Lyon, IARC Press, 2004

191. Lill, J. Proteomic tools for quantitation by mass spectrometry *Mass Spectrom. Rev.* 2003 22:182–194
192. Cotter RJ. Time-of-flight mass spectrometry: an increasing role in the life sciences. *Biomed Environ Mass Spectrom* 1989; 18: 513
193. Swets, J. A., Measuring the accuracy of diagnostic systems. *Science* 1988;240, (4857)
194. Chaurand P, et al. Proteomics in diagnostic pathology: profiling and imaging proteins directly in tissue sections. *Am J Pathol*, 2004. 165(4):1057-68
195. Cotter RJ. Time-of-flight mass spectrometry: an increasing role in the life sciences. *Biomed Environ Mass Spectrom* 1989; 18: 513
196. Spengler B, et al. Mass spectrometric sequencing of linear peptides by product-ion analysis in a reflectron time-of-flight mass spectrometer using matrix-assisted laser desorption ionization. *Rapid Commun Mass Spectrom* 1993; 7: 902
197. Cheng AJ, et al. Oral cancer plasma tumor marker identified with beadbased affinity-fractionated proteomic technology. *Clin Chem* 2005; 51: 2236
198. de Noo ME, et al. Detection of colorectal cancer using MALDI-TOF serum protein profiling. *Eur J Cancer* 2006; 42: 1068
199. Lill, J. Proteomic tools for quantitation by mass spectrometry *Mass Spectrom. Rev.* 2003 22:182–194
200. Svensson M, Skold K, Svenningsson P, Andren PE. Peptidomics-based discovery of novel neuropeptides. *J Proteome Res.* 2003;2:213-9
201. Schenk S, Schoenhals GJ, de Souza G, Mann M. A high confidence, manually validated human blood plasma protein reference set. *BMC Med Genomics.* 2008;1:41.
202. Harper RG, Workman SR, Schuetzner S, Timperman AT, Sutton JN. Low molecular-weight human serum proteome using ultrafiltration, isoelectric focusing, and mass spectrometry. *Electrophoresis.* 2004;25:1299-306.
203. Tirumalai RS, Chan KC, Prieto DA, Issaq HJ, Conrads TP, Veenstra TD. Characterization of the low molecular weight human serum proteome. *Mol Cell Proteomics.* 2003;2:1096-103.
204. Altelaar AF, Mohammed S, Brans MA, Adan RA, Heck AJ. Improved identification of endogenous peptides from murine nervous tissue by multiplexed peptide extraction methods and multiplexed mass spectrometric analysis. *J Proteome Res.* 2009;8:870-6.
205. Villanueva, J.; Shaffer, D. R.; Philip, J.; Chaparro, C. A.; Erdjument-Bromage, H.; Olshen, A. B.; Fleisher, M.; Lilja, H.; Brogi, E.; Boyd, J.; Sanchez-Carbayo, M.; Holland, E. C.; Cordon-Cardo, C.; Scher, H. I.; Tempst, P., Differential exoprotease activities confer tumor-

- specific serum peptidome patterns. *J Clin Invest* 2006, 116, 271-84.
206. Tiss, A., Smith, C. J. and Cramer, R. (2011) MALDI MS-based biomarker profiling of blood samples.. *PLoS One*. 2011;6:1-12.
207. Villanueva J, Nazarian A, Lawlor K, Yi SS, Robbins RJ, Tempst P. A sequence specific exopeptidase activity test (SSEAT) for "functional" biomarker discovery. *Mol Cell Proteomics*. 2008;7:509-18.
208. van den Elsen, J.M., Martin, A., Wong, V., Clemenza, L., Rose, D.R., Isenman, D.E. X-ray crystal structure of the C4d fragment of human complement component C4. *J.Mol.Biol.*, v.322, pp.1103 - 1115, 2002
209. Yu CY, Belt KT, Giles CM, et al. (1987). "Structural basis of the polymorphism of human complement components C4A and C4B: gene size, reactivity and antigenicity.". *EMBO J*. 5 (11): 2873–81
210. Schick BP (2000). "Regulation of expression of megakaryocyte and platelet proteoglycans.". *Stem Cells* 14 Suppl 1: 220–31
211. Ueda, K.; Saichi, N.; Takami, S.; Kang, D.; Toyama, A.; Daigo, Y.; Ishikawa N.; Kohno, N.; Tamura, K.; Shuin, T.; Nakayama, M.; Sato, . A.; Nakamura, Y. comprehensive peptidome profiling technology for the identification of early detection biomarkers for lung adenocarcinoma. *PLoS One* 2011, 6,
212. Davie, E.W., Fujikawa, K., Kurachi, K. and Kisiel, W. The role of serine proteases in the blood coagulation cascade. *Adv. Enzymol.* 48
213. MacGillivray, R.T.A. and Davie, E.W. Characterization of bovine prothrombin mRNA and its translation product. *Biochemistry* 23 (1984) 1626–1634
214. Schwartz JH (November 2001). "The many dimensions of cAMP signaling". *Proc. Natl. Acad. Sci. U.S.A.* 98 (24): 13482–4.
215. Villanueva, J.; Martorella, A. J.; Lawlor, K.; Philip, J.; Fleisher, M.; Robbins J.; Tempst, P., Serum peptidome patterns that distinguish metastatic thyroid carcinoma from cancer-free controls are unbiased by gender and age. *Mol Cell Proteomics* 2006, 5, (10), 1840-52.
216. Mann, K.G. Prothrombin. *Methods Enzymol.* 45 (1976) 123–156. Entrez PubMed
217. SF Abcouwer, PL Marjon, RK Loper Response of VEGF expression to amino acid deprivation and inducers of endoplasmic reticulum stress
218. S Gstincich, C Schneide Serum deprivation response gene is induced by serum starvation but not by contact inhibition.
219. Xu XL, Wu LC, Du F, Davis A, Peyton M, Tomizawa Y, et al. Inactivation of human SRBC, located within the 11p15.5-p15.4 tumor suppressor region, in breast and lung cancers.

- Cancer Res. 2001;61:7943-9.
220. Gustincich S, Schneider C. Serum deprivation response gene is induced by serum starvation but not by contact inhibition. *Cell Growth Differ.* 1993;4:753-60.
221. Izumi Y, Hirai S, Tamai Y, Fujise-Matsuoka A, Nishimura Y, Ohno S. A protein kinase Cdelta-binding protein SRBC whose expression is induced by serum starvation. *J Biol Chem.* 1997;272:7381-9.
222. Amsellem V, Kryszke MH, Hervy M, Subra F, Athman R, Leh H, et al. The actin cytoskeleton-associated protein zyxin acts as a tumor suppressor in Ewing tumor cells. *Exp Cell Res.* 2005;304:443-56.
223. Hervy M, Hoffman LM, Jensen CC, Smith M, Beckerle MC. The LIM Protein Zyxin Binds CARP-1 and Promotes Apoptosis. *Genes Cancer.* 2010;1:506-15
224. Wu H, Liu T, Wang R, Tian S, Liu M, Li X, et al. MicroRNA-16 targets zyxin and promotes cell motility in human laryngeal carcinoma cell line HEP-2. *IUBMB Life.* 2011;63:101-8.
225. Hoffman LM, Jensen CC, Kloeker S, Wang CL, Yoshigi M, Beckerle MC. Genetic ablation of zyxin causes Mena/VASP mislocalization, increased motility, and deficits in actin remodeling. *J Cell Biol.* 2006;172:771-82.
226. Call GS, Chung JY, Davis JA, Price BD, Primavera TS, Thomson NC, et al. Zyxin phosphorylation at serine 142 modulates the zyxin head-tail interaction to alter cell-cell adhesion. *Biochem Biophys Res Commun.* 2011;404:780-4
227. Sansing HA, Sarkeshik A, Yates JR, Patel V, Gutkind JS, Yamada KM, et al. Integrin alpha5beta1, alpha6beta1, alpha6beta2 effectors p130Cas, Src and talin regulate carcinoma invasion and chemoresistance. *Biochem Biophys Res Commun.* 2011;406:171-6.
228. Beyer-Sehlmeyer G, Hiddemann W, Wormann B, Bertram J. Suppressive subtractive hybridisation reveals differential expression of serglycin, sorcin, bone marrow proteoglycan and prostate-tumour-inducing gene I (PTI-1) in drug-resistant and sensitive tumour cell lines of haematopoietic origin. *Eur J Cancer.* 1999;35:1735
229. Li XJ, Ong CK, Cao Y, Xiang YQ, Shao 1088 JY, Ooi A, et al. Serglycin is a theranostic target in nasopharyngeal carcinoma that promotes metastasis. *Cancer Res.* 1090 2011;71:3162-72.
230. Wang ZH, Ding KF, Yu JK, Zhai XH, Ruan SQ, Wang SW, et al. Proteomic analysis of primary colon cancer-associated fibroblasts using the SELDI-ProteinChip platform. *J Zhejiang Univ Sci B.* 2012;13:159-67.
231. Nanni P, Mezzanotte L, Roda G, Caponi A, Levander F, James P, et al. Differential proteomic analysis of HT29 Cl.16E and intestinal epithelial cells by LC ESI/QTOF mass

spectrometry. *J Proteomics*. 2009;72:865-73.

232. Nesvizhskii AI, Aebersold R. Interpretation of shotgun proteomic data: the protein inference problem. *Mol Cell Proteomics*. 2005;4:1419-40.

RINGRAZIAMENTI

A conclusione di questo lavoro di tesi colgo l'occasione per ringraziare chi mi ha sostenuto durante tutto questo percorso.

Il primo grazie va al Professor Fulvio Magni per la sua disponibilità e comprensione. Non è stata solo una guida per lo sviluppo del progetto ma anche e soprattutto una persona con la quale parlare apertamente di qualunque problema. Un grazie anche alla professoressa Del Puppo per la sua gentilezza e disponibilità.

Un ringraziamento speciale va a tutte le ragazze del laboratorio ed in particolare

...Clizia che mi ha seguita e istruita pazientemente passo dopo passo,

...Erica che mi ha insegnato ad usare il "robottino" e ad affrontare ogni suo capriccio ,

...Marta una "guida" nella stesura di questo testo

...Veronica per avermi ascoltato tutte le volte che mi trovavo in crisi e che spero diventi presto una "socia" per il beach volley.

...Laura per avermi aiutato con tutta la burocrazia e grazie a questa è nata una splendida amicizia

E ancora non posso che ringraziare Simone: il suo sostegno è stato fondamentale in questi anni non sono con l'amore e i sorrisi che mi regala ogni giorno, ma anche con le critiche costruttive che, anche se odio , sono una guida nelle scelte che faccio quotidianamente.

Un ringraziamento anche ai miei genitori che hanno creduto in me fino al punto di sostenermi per qualunque scelta io facessi e, nonostante sia capitato che i problemi della vita ci abbiano allontanato, mi hanno dimostrato di essere sempre presenti.

Non posso non ringraziare, inoltre, anche Laura che, dopo ormai 15 anni d'amicizia, considero ormai come una sorella.

Infine ringrazio tutte quelle persone che mi hanno aiutato, sostenuto e compreso durante tutte questo tempo.....

DI NUOVO GRAZIE A TUTTI DAL CUORE