UNIVERSITAT DE VALÈNCIA

Doctorate in Medicine

Instituto de Investigación Sanitaria La Fe

Identification of *omic* profiles for diagnosis and monitoring of bladder cancer

Ph.D. THESIS

Submitted by

Alba Loras Monfort

Ph.D. Supervisors:

José Luis Ruiz-Cerdá
Guillermo Quintás Soriano
Salvador Gil Grau

April, 2019

JOSÉ LUIS RUIZ-CERDÁ, Ph.D. in Medicine and Surgery and Professor at the *Universitat de València*, SALVADOR GIL GRAU, Ph.D. in Chemistry and Professor at the *Universitat de València*, and GUILLERMO QUINTÁS SORIANO, Ph.D. in Chemistry and researcher at Instituto de Investigación Sanitaria La Fe de Valencia.

CERTIFY:

That the work entitled *"Identification of omic profiles for diagnosis and monitoring of bladder cancer"* has been developed by Alba Loras Monfort under their supervision at the Instituto de Investigación Sanitaria La Fe de Valencia, as a Thesis Project in order to obtain the degree of Ph.D. in Medicine at the *Universitat de València*.

Valencia, 05 de Abril 2019.

Dr. José Luis Ruiz-Cerdá Dr. Salvador Gil Grau Dr. Guillermo Quintás Soriano

A la meua família

Agraïments

La redacció d'aquestes línes suposa la culminació d'una etapa científica marcada per l'aprenentatge i el companyerisme. I es que qui anava a dir-me que durant aquests tres anys els companys de treball acabaríen sent els meus amics i confidents? Aquestes paraules d'agraïment van destinades a ells:

Al Dr. José Luis Ruiz-Cerdá per confiar en mi des del primer moment, ensenyar-me tot el que ara sé sobre l'urologia i recolzar-me en cada decisió que he hagut que prendre. Gràcies per tot, però sobretot per ser el meu amic. Al Dr. Guillermo Quintás i el Dr. Salvador Gil, per compartir cada un dels vostres coneixements, la vostra professionalitat i estar disposats a solventar cadascun dels meus dubtes. Moltes gràcies als tres per acceptar ser els meus tutors de tesi.

A tot el personal del Servici d'Urologia, però especialement als residents Leo i Laura i a l'equip d'infermeria. Gràcies per acollir-me tan amablement i fer-me sentir una més. Per descomptat, moltes gràcies a tots els pacients que de forma altruista han col·laborat en els nostres estudis. Sense vosaltes aquesta tesi no s'haguera realitzat.

A Raquel Amigo i tot l'equip del Biobanc la Fe. Hem vau integrar al vostre grup només arribar com si fos una més. Ha segut un plaer treballar amb vosaltres.

A les meues companyes de laboratori: Nati, Mª Carmen, Sara, Àngela i recentment a Lluís. Amb vosaltes les hores passen més apressa, els problemes es solucionen i els èxits es celebren. Nati, hem compartit molts moments durant aquest anys. Et dessitge molta sort en la teua nova etapa. Mª Carmen, gràcies per tantes coses... Hem estat treballant juntes dia a dia, minut a minut, aprenent i equivocant-mos, però tot l'esforç ha valgut la pena. Gràcies per estar sempre aquí quan t'he necessitat.

A Ramón Martinez, per confiar en els nostres projectes, fer més competent la nostra unitat i estar disposat a treballar conjuntament.

A Alegría Montoro per ser com és, una amiga i companya en qui confiar.

A Alfredo, per estar disposat a solucionar cadascun dels meus dubtes i donar-me ànims per seguir endavant.

A tots els membres del laboratori d'Oncología Molecular del CIEMAT però especialment a Jesús Maria Paramio i Cristian. La vostra amabilitat, hospitalitat i paciència va ser essencial en un moment clau de la meua vida. Vos vau comportar com part de la meua família. Vaig aprendre molt de vosaltres.

A Imma, per acceptar dissenyar la portada de la tesi. Qui millor que tu!

Als meus pares, pel seu esforç en donar-me uns estudis, confiar i recolzar cada decisió que he pres i valorar el meu treball. Gràcies per animar-me en els moments difícils. Gràcies per l'educació i els valors que m'heu inculcat. Vos vull!

A Lucia, la meua germana menuda. No haguera pogut tindre una germana millor. Hem compartit moltes vivències juntes. I les que ens queden!

A Flavia i Juan Antonio, per estar disposats sempre a tot. La vostra ajuda ha segut imprescindible per poder acabar la tesis. Gràcies.

Al meu amic pelut Roc. Quina alegría arribar a casa i trobar-te sempre feliç! Els teus abraços em carreguen les piles.

A Román, per creure incondicionalment en mí. Per donar positivisme a la meua vida, per fer-me sentir orgullosa de mí mateixa i del meu treball, per ensenyar-me a donar importancia a les coses que realment ho mereixen. Per estimar-me tant.

A Pol, per ser la força que em mou dia a dia. Eres el millor que m'ha passat.

En definitiva, moltes gràcies a tots els que heu fet possible que aquesta tesi es fera realitat.

“És simplement la possibilitat de realitzar un somni
el que fa que la vida siga interessant”

Abstract

The present Thesis entitled *Identification of omic profiles for diagnosis and monitoring of bladder cancer* is focused on identifying non-invasive urinary metabolomic biomarkers of diagnosis and monitoring of bladder cancer (BC). In order to achieve this objective, two analytical strategies based on Nuclear Magnetic Resonance (^{1}H NMR) and Ultraperformance Liquid Chromatography–Mass Spectrometry (UPLC-MS) have been used for the analysis of urine samples. Besides, bladder tissue samples have been analyzed by High-Resolution Magic Angle Spinning NMR (HRMAS NMR) technique to get further insight into altered metabolic pathways in BC and assess their link with altered urinary metabolomic profiles. On the other hand, transcriptomic analysis has been carried out in bladder tissues to identify key metabolic genes in BC. Additionally, integrative studies using metabolomic and transcriptomic data have been performed to study the gene-metabolite networks in BC and its association with the altered urinary metabolome.

Initially, a metabolic profile able to distinguish BC tissues from non-tumor tissues with a sensitivity and specificity of 100%, independently of stage and grade of the tumor, is presented in chapter four. Moreover, the metabolites that take part of this profile are showed, as well as, the disturbed metabolic pathways linked to BC carcinogenesis. On the other hand, the transcriptomic analysis performed in these same tissues is described, indicating principally that metabolic genes are downregulated in bladder tumors and that transcriptional repressors, histone marks, and alternatively splicing processes may be regulating those genes. Additionally, an integrative analysis between metabolomic and transcriptomic data is detailed, showing concordance between the results obtained through these two techniques that represent different levels of molecular regulation.

Finally, a ^{1}H NMR-based urinary metabolic profile capable of distinguishing BC urines from control urines (collected after surgery) with significant sensitivity (90.9%) and specificity (76.9%) is shown. Urine and tissue samples were collected from the same patients, so at the end of chapter, the connections between the perturbed metabolic pathways in tissues and urines are described.

The following two chapters of the thesis are focused principally on searching non-invasive biomarkers of BC in urine samples for monitoring this disease by means of two analytical techniques.

In order to validate the urinary profile as a biomarker for monitoring, a study with additional urinary samples collected from patients with NMIBC was carried out. Urinary samples were collected monthly during a follow-up period. The urinary ^{1}H NMR metabolic profile showed a sensitivity and specificity around 85% classifying BC urines. Moreover, tumor recurrences were detected by the metabolic profile in an early stage of disease, anticipating in some cases to the BC visualization by cystoscopy. The altered metabolic pathways in the urinary metabolome were also identified.

Finally, the sixth chapter of the thesis exhibits the results of an investigative clinical study carried out in a large number or urinary samples collected from NMIBC patients before and after surgery, as well as during the subsequent surveillance. In this case, the urine samples were analyzed through UPLC-MS and perturbed metabolic pathways were assessed. Some urinary samples were common to those analyzed by ^{1}H NMR, and data were in agreement. The analysis of the longitudinal trajectories of the metabolic biomarker discriminating between BC and control samples allowed a preliminary

evaluation of its potential utility to monitor NMIBC relapse in patients undergoing surveillance for tumor recurrence.

On the whole, the results presented in this thesis give support to the hypothesis of the existence of a urinary metabolic signature linked with tumor alterations in BC tissues able to detect and predict recurrences during the surveillance period of patients with NMIBC. Moreover, the good results obtained and the concordance between the urinary analyses by ^{1}H NMR and UPLC–MS highlight the metabolomics as a competitive *omic* for searching biomarkers, since offers robust and dynamic information about the biology of the tumor.

Resumen

La presente Tesis titulada *Identificación de perfiles ómicos para el diagnóstico y la monitorización del cáncer de vejiga* se centra en la identificación de biomarcadores metabolómicos urinarios no invasivos para el diagnóstico y la monitorización del cáncer de vejiga (CaV). Con este fin, se han utilizado dos plataformas analíticas: la Resonancia Magnética Nuclear (*Nuclear Magnetic Resonance, ^{1}H NMR*) y la Cromatografía Líquida de alta resolución acoplada a la Espectrometría de Masas (*Ultraperformance Liquid Chromatography–Mass Spectrometry, UPLC-MS*). Además, se han analizado tejidos vesicales mediante la técnica de Resonancia Magnética Nuclear de Alta Resolución con Giro de Ángulo Mágico (*High-Resolution Magic Angle Spinning NMR, HRMAS NMR*) para obtener más información sobre las vías metabólicas alteradas en el CaV y evaluar su relación con los perfiles metabólicos urinarios alterados. Por otro lado, se han realizado análisis transcriptómicos en tejidos vesicales para identificar genes metabólicos clave en el CaV. Finalmente, se han llevado a cabo estudios integradores con los datos metabolómicos y transcriptómicos para estudiar las conexiones entre genes y metabolitos y establecer su asociación con el metaboloma urinario.

Inicialmente, en el capítulo cuarto se presenta un perfil metabólico capaz de distinguir los tejidos tumorales de los tejidos no tumorales con una sensibilidad y especificidad del 100%, independientemente del estadio y el grado del tumor. Además, se muestran los metabolitos que forman parte de este perfil, así como las vías metabólicas alteradas asociadas a la carcinogénesis vesical. Por otro lado, tras la realización de análisis transcriptómicos en esos mismos tejidos, se detalla como los genes metabólicos están siendo regulados a la baja en los tumores de vejiga mediante la acción de represores transcripcionales, marcas de

histonas y procesos de *splicing* alternativo. Además, mediante un análisis integrativo entre los datos metabolómicos y transcriptómicos, se muestra la concordancia entre los resultados obtenidos a través de estas dos técnicas que representan diferentes niveles de regulación molecular. Finalmente, se muestra un perfil metabólico urinario identificado mediante ^{1}H *NMR* capaz de distinguir orinas con CaV de orinas control (recogidas después de la cirugía) con una significativa sensibilidad (90,9%) y especificidad (76,9%). Como las muestras de orina y de tejido se recogieron de los mismos pacientes, al final del capítulo se describen las conexiones encontradas entre las rutas metabólicas alteradas en tejidos y orinas

Los siguientes dos capítulos de la tesis se centran principalmente en la búsqueda de biomarcadores no invasivos de CaV en muestras de orina para el seguimiento de esta enfermedad a través de dos técnicas analíticas.

Con el objetivo de validar el perfil urinario como biomarcador de monitorización de CaV, se llevó a cabo un estudio con muestras urinarias adicionales de pacientes con CaV no músculo-invasivo (CVNMI). Las muestras urinarias se recogieron mensualmente durante un período de seguimiento activo. El perfil metabolómico urinario detectado mediante ^{1}H *NMR* presentó sensibilidades y especificidades alrededor del 85% en la clasificación de las orinas tumorales, e incluso detectó las recidivas tumorales en un estado temprano de su desarrollo, anticipándose en algunos casos a la visualización de éstas mediante cistoscopia. Este quinto capítulo también detalla los metabolitos discriminantes que forman parte de este perfil metabolómico y su relación con las rutas bioquímicas alteradas en el CaV.

Finalmente, el sexto capítulo de la tesis muestra los resultados de un estudio clínico llevado a cabo con un gran número de muestras de orina recolectadas de pacientes con CVNMI antes y después de la cirugía, así como durante un periodo de seguimiento posterior. En este caso, las muestras urinarias se analizaron mediante *UPLC-MS* y se estudiaron las vías metabólicas perturbadas vinculadas al CaV. Algunas muestras de orina fueron comunes a las analizadas por ^{1}*H NMR*, y en general, los datos de los dos estudios fueron concordantes. El análisis de las trayectorias longitudinales del biomarcador metabólico urinario capaz de discriminar las muestras tumorales de las controles permitió una evaluación preliminar de su utilidad como biomarcador de seguimiento para la detección de las recurrencias en pacientes con CVNMI.

En general, los resultados presentados en esta tesis respaldan la hipótesis de la existencia de una huella metabólica urinaria vinculada a las alteraciones tumorales presentes en los tejidos vesicales, capaz de detectar y predecir las recurrencias durante el período de vigilancia en pacientes con CVNMI. Además, los buenos resultados obtenidos y la concordancia entre ambos estudios urinarios (^{1}*H NMR y UPLC-MS*) posicionan la metabolómica al frente de las técnicas ómicas para la búsqueda de biomarcadores robustos y dinámicos que reflejen la biología del tumor.

Resum

La present Tesi titulada *Identificació de perfils òmics per al diagnòstic i el seguiment del càncer de bufeta* es centra en la identificació de biomarcadors metabolòmics urinaris no invasius per al diagnòstic i el seguiment del càncer de bufeta (CB). Amb aquesta finalitat, s'han utilitzat dues plataformes analítiques: la Ressonància Magnètica Nuclear (*Nuclear Magnetic Resonance,* ^{1}*H NMR*) i la Cromatografia Líquida d'Alta Resolució acoblada a l'Espectrometria de Masses (*Ultraperformance Liquid Chromatography–Mass Spectrometry, UPLC-MS*). D'altra banda, teixits vesicals s'han analitzat mitjançant la tècnica de Ressonància Magnètica Nuclear d'Alta Resolució amb Gir d'Àngle Màgic (*High-Resolution Magic Angle Spinning NMR, HRMAS NMR*) per tal de conèixer les vies metabòliques que estan alterades en els tumors de bufeta i la seva relació amb el metaboloma urinari. A més s'han realitzat anàlisis transcriptòmics en teixits vesicals per conèixer quins gens metabòlics estan alterats i quins processos moleculars estarien regulant-los. Finalment, s'han realitzat estudis integradors amb les dades derivades dels estudis metabolòmics i transcriptòmics realitzats en teixits, per estudiar les connexions entre gens i metabòlits i establir la seva vinculació amb el metaboloma urinari.

Inicialment, en el quart capítol es presenta un perfil metabòlic capaç de distingir els teixits tumorals dels teixits no tumorals amb una sensibilitat i especificitat del 100%, independentment de l'estadi i el grau del tumor. A més, es mostren els metabòlits que formen part d'aquest perfil, així com les vies metabòliques alterades associades a la carcinogènesi vesical. D'altra banda, es descriu l'anàlisi transcriptòmic realitzat amb aquests mateixos teixits, indicant que principalment els gens metabòlics estan sent regulats a la baixa en els tumors de bufeta i que els repressors transcripcionals, les marques d'histones i els processos de *splicing* alternatiu podrien estar regulant aquests gens. A més, es detalla un

anàlisi integratiu entre les dades metabolòmiques i transcriptòmiques que mostra la concordança entre els resultats obtinguts a través d'aquestes dues tècniques, que representen diferents nivells de regulació molecular. Finalment, es mostra un perfil metabòlic urinari identificat mitjançant *^{1}H NMR* capaç de distingir orines tumorals d'orines control (recollides després de la cirurgia) amb una significativa sensibilitat (90,9%) i especificitat (76,9%). Com les mostres d'orina i de teixit es van recollir dels mateixos pacients, al final del capítol es descriuen les connexions trobades entre les rutes metabòliques alterades en teixits i orines.

Els següents dos capítols de la tesi es centren principalment en la recerca de biomarcadors no múscul-invasius de CB (CBNMI) en mostres d'orina per a la monitorització d'aquesta malaltia a través de dues tècniques analítiques.

Amb l'objectiu de validar el perfil urinari com a biomarcador de seguiment de CB, es va dur a terme un estudi amb mostres urinàries de pacients amb CBNMI. Les mostres urinàries es van recol·lectar mensualment durant un període de seguiment actiu. El perfil metabolòmic urinari identificat mitjançant *^{1}H NMR*, va mostrar sensibilitats i especificitats al voltant del 85% per a la classificació de les orines tumorals, i fins i tot va detectar les recidives tumorals en un estat primerenc del seu desenvolupament, anticipant-se en alguns casos a la visualització d'aquestes mitjançant cistoscòpia. Aquest cinquè capítol també detalla els metabòlits discriminants que formen part d'aquest perfil metabolòmic i la seva relació amb les rutes bioquímiques alterades en el CB.

Finalment, el sisè capítol de la tesi mostra els resultats d'un estudi clínic dut a terme amb un gran nombre de mostres d'orina recollides de pacients amb CBNMI abans i després de la cirurgia, així com durant un període de seguiment posterior. En aquest cas, les mostres urinàries es van analitzar mitjançant *UPLC-*

MS i es van estudiar les vies metabòliques vinculades al CB. Algunes mostres d'orina van ser comunes a les analitzades per *^{1}H NMR*, i en general, les dades dels dos estudis van ser concordants. L'anàlisi de les trajectòries longitudinals del biomarcador metabòlic urinari discriminaren les mostres tumorals de les control i van permetre una avaluació preliminar de la seva utilitat com a biomarcador de seguiment per a la detecció de les recurrències en pacients amb CBNMI.

En general, els resultats presentats en aquesta tesi donen suport a la hipòtesi de l'existència d'una signatura metabòlica urinària relacionada amb les alteracions tumorals en teixits vesicals, capaç de detectar i predir les recurrències durant el període de vigilància en pacients amb CBNMI. A més, els bons resultats obtinguts i la concordança entre els dos estudis urinaris (*^{1}H NMR y UPLC-MS*) posicionen la metabolòmica al capdavant de les tècniques òmiques per a la recerca de biomarcadors robusts i dinàmics que reflexen la biologia del tumor.

Resumen ampliado de la tesis

La presente Tesis titulada *Identificación de perfiles ómicos para el diagnóstico y la monitorización del cáncer de vejiga* se centra en la identificación de biomarcadores metabolómicos urinarios no invasivos para el diagnóstico y la monitorización del CaV, y pretende esclarecer que genes y rutas metabólicas podrían ser claves en este tipo de tumores. Para esto se realizaron estudios metabolómicos basados en ^{1}*H NMR* y *UPLC-MS* en muestras urinarias; y estudios de *HRMAS* ^{1}*H NMR* y transcriptómicos (*microarrays*) en muestras tisulares vesicales. También se llevaron a cabo estudios integradores con los datos metabolómicos y transcriptómicos para estudiar las conexiones entre genes y metabolitos en los tejidos tumorales y establecer su asociación con el metaboloma urinario.

Los primeros tres capítulos de la tesis engloban la introducción, los objetivos y el planteamiento del problema, y los capítulos 4, 5 y 6 los diferentes estudios realizados.

Capítulo 4

La primera parte de la tesis se centró en la identificación de perfiles metabolómicos de ^{1}*H NMR*, tanto tisulares como urinarios, para utilizarse como biomarcadores de CaV. Además pretendió identificar metabolitos y genes metabólicos clave en los tumores vesicales con el fin de comprender mejor la biología del CaV y las conexiones entre genes y metabolitos. Finalmente se focalizó en el estudio de las conexiones entre las vía metabólicas alteradas en los tejidos tumorales y el metaboloma urinario perturbado.

Metodología

En este estudio se incluyeron 21 pacientes diagnosticados de CaV (14 hombres and 7 mujeres) que estaban en lista de espera para ser intervenidos. De

cada paciente se recogieron muestras de tejido y orina. Las muestras de tejido tumorales (n=22) y no tumorales (n=22) se tomaron durante la resección transuretral (RTU) y la clasificación de los tumores en CVNMI o cánceres de vejiga músculo invasivos (CVMI) se realizó según criterios anatomopatológicos. Las muestras de orina se recolectaron antes de la RTU "tumorales" (n=22) y un mes después de la RTU "controles" (n=13) solamente en pacientes en los que la RTU fue completa, el diagnóstico CVNMI y el resultado de la cistoscopia tras la cirugía negativo. Las orinas y los tejidos se analizaron en un espectrómetro Bruker Avance 600 MHz. En el caso de los tejidos, se utilizó una sonda de *HRMAS*. Tras la adquisición de los espectros monodimensionales (1D) ^{1}H NOESY y bidimensionales (2D), las muestras de tejido se volvieron a congelar y se descongelaron de nuevo para la realización de los análisis transcriptómicos.

Tras la adquisición de los espectros metabolómicos, tanto en tejidos como orinas, estos fueron pre-procesados utilizando MestReNova (versión 6.0.2). En tejidos se corrigió la fase, la línea base, y el desplazamiento químico se referenció al singlete de la creatina (3.03 ppm) y al doblete de la alanina (1.48 ppm). A continuación, se asignaron los principales metabolitos en la región de 0,8 a 9,5 ppm del espectro. Para ello se utilizaron bases de datos como *Human Metabolome Database* (HMDB) o *Metabolomics: main*. Antes de realizar los análisis estadísticos las señales de los espectros fueron alineadas utilizando el algoritmo (*Icoshift*) y los datos se normalizaron respecto a los mg de tejido y se autoescalaron. En los tejidos, la intensidad de las resonancias asignadas a cada metabolito se transfirió a la plataforma online *MetaboAnalyst 3.0. La prueba* U-Mann Whitney determinó las diferencias significativas entre los valores medios de intensidad entre los tejidos tumorales y controles. El análisis de la varianza (ANOVA) y las pruebas *posthoc* de Tukey evaluaron las diferencias entre los CVNMI, CVMI y los tejidos no tumorales; y entre los estadios Ta, T1, T2 y los tejidos no tumorales. Por otro lado, se realizaron análisis estadísticos

multivariantes de tipo *Partial Least Square-Discriminant Analysis* (*PLS-DA*) utilizando el software PLS_Toolbox Solo 8.0 (Eigenvector Research, Inc., Manson, WA, EE. UU.). Se incluyeron las regiones de 0.5-4.8 y 5.2-9.5 ppm. La región espectral del agua (4.79-5.2 ppm) se excluyó. Para la realización del modelo *PLS-DA*, las muestras se separaron aleatoriamente en dos grupos: calibración (n=34) y validación (n=10) y se calcularon dos modelos. El primero consideró el conjunto inicial de características espectrales (11698); y el segundo se creó tras una selección de las variables más importantes del modelo (*Variable Importance in Projection* (*VIP*>1)) (4800 características retenidas). La selección de características durante la optimización del modelo se realizó con el fin de mejorar el rendimiento predictivo y facilitar su interpretación. La identificación de las señales de metabolitos altamente discriminantes se llevó a cabo en el modelo VIP>1. La lista de metabolitos obtenida se usó luego para realizar análisis de rutas de enriquecimiento en la plataforma *MetaboAnalyst 3.0*.

Para la realización de los análisis transcriptómicos los tejidos se descongelaron y a continuación se llevó a cabo la extracción del ARN. Tras comprobar la integridad del ARN en un Bioanalyzer, se seleccionaron 18 muestras (10 no tumorales y 8 tumorales) para analizarse en los chips "*Human Transcriptome Array 2.0 (HTA 2.0*)". La expresión de los datos se normalizó y el efecto *Bach* se corrigió utilizando el análisis robusto de transformación de espacio de señal (SST-RMA, por sus siglas en inglés) implementado en el software TAC 4.0. Se llevó a cabo un Análisis de Componentes Principales (*PCA*) y un *Headmap* para observar las diferencias entre los tejidos tumorales y no tumorales sobre la base de todo el transcriptoma. Se consideró como criterio de selección un *Fold-Change* de al menos 2 o -2 y una tasa de descubrimiento falso (*FDR*) de 0,05. Análisis de *Gene Ontology Biological Processes* (GOBP), ChEA y *Encode Histone Roadmap* se realizaron *in silico* utilizando la herramienta web Enrich. Los genes metabólicos con una expresión diferencial significativa entre muestras tumorales y no

tumorales se seleccionaron desde las plataformas GOBP y GeneCards. Se realizaron análisis de ChEA y Encod*e Histone Roadmap* para identificar la supuesta unión de los factores de transcripción y las marcas de histonas a esos genes metabólicos. Además, los genes identificados se utilizaron para realizar un agrupamiento jerárquico no supervisado usando los datos del Atlas del Genoma del Cáncer (TCGA) utilizando como herramienta MeV. Los análisis integradores entre los resultados metabolómicos y transcriptómicos en tejidos se realizaron utilizando: *MetaboAnalys*t, *Wikipathways* y la base de datos *Small Molecule Pathway Database* (SMPDB).

En las orinas, tras la adquisición de los espectros 1D ^{1}H NOESY y 2D se corrigió la fase, la línea base, y el desplazamiento químico, que en este caso se referenció a la señal del DSS (0,00). Antes de la realización de los análisis estadísticos multivariantes realizó un *binning* de 0.003 en los espectros 1D y los datos se normalizaron (1-norm) y autoescalaron. A continuación, se asignaron los principales metabolitos en la región de 0,8 a 9,5 ppm del desplazamiento químico. Para ello se utilizaron bases de datos como HMDB o *Metabolomics: main*. Se realizó un modelo *PLS-DA* considerando las regiones espectrales de 0,8-4,5 y 6,5-9,0 ppm (2068 características). Las orinas se clasificaron como control (orina recolectada después de la RTU) o tumor (orina recolectada antes de la RTU). La validación cruzada (*Cross Validation* (CV)) se realizó en el nivel más alto de jerarquía de muestreo, que en este caso fue el voluntario, para estimar el rendimiento predictivo del modelo. La significación estadística del modelo estimada por CV se evaluó mediante la prueba de permutación con 100 interacciones. Los metabolitos más importantes en el modelo estadístico (*VIP*>1) se identificaron y se utilizaron para realizar un análisis de enriquecimiento de las vías metabólicas en la plataforma *MetaboAnalyst*.

Resultados

Estudios metabolómicos en tejidos

Los espectros de tejidos no tumorales mostraron una gran presencia de lípidos en la región alifática prácticamente ausente en las muestras de tumor. Por el contrario, en los espectros tumorales destacaron las señales de pequeños metabolitos como el glutatión (GSH), la tirosina (Tyr) y la citidina difosfato (CDP) en la región aromática.

La comparación de medias de intensidad entre las muestras tumorales y no tumorales mostró mayores niveles de colina (Cho), CDP, mioinositol (mI), uridina difosfato (UDP) y GSH en los tejidos tumorales. La intensidad para el fragmento lipídico -(n)CH_2-CH_2-CH_2-COOH (denominado lípido (c)) fue mayor en las muestras de control. La mayoría de los metabolitos que mostraron diferencias significativas entre los tumores y los tejidos control (valor de $p < 0,05$) también presentaron mayores intensidades en los CVNMI y CVMI en comparación con las muestras control. Teniendo en cuenta las diferencias entre los diversos estadios tumorales y las muestras no tumorales, los fragmentos lipídicos –CH_2-CH_2-CH_2-CH_3 (lípido (a)) y –CH=CH-CH_2-CH=CH-CH_2-CH_2-(n)-CH_2- (lípido (b), Tyr, Cho, CDP, mI, UDP y GSH mostraron cambios significativos entre todos los grupos. Concretamente, los tumores T2 (los más agresivos considerados en este estudio) presentaron la mayor intensidad de Cho y GSH; y junto con los T1 mostraron niveles elevados de mI y Tyr. Por otro lado, los tumores Ta (los tumores menos agresivos) mostraron un incremento significativo de UDP y CDP. En el caso del lactato, las señales a 1,33 ppm y 4,1 ppm estaban solapadas con las señales de lípidos en las muestras no tumorales. Por lo tanto, la intensidad de lactato se comparó solamente entre los diferentes estadios tumorales. Se observó una mayor intensidad de lactato en los tumores invasivos (T2) que en los tumores con estadios T1 o Ta.

El modelo *PLS-DA* inicial (n=44, variables latentes (LVs)=2) proporcionó una sensibilidad del 82,4% y una especificidad del 88,2% para el set de calibración. El análisis de las muestras incluidas en el conjunto de validación externa presentó valores de sensibilidad y selectividad del 100% para los dos modelos (conjunto original de variables y *VIP*>1). Los metabolitos más discriminantes identificados en el modelo (*VIP*>1) fueron: lípidos (fragmentos a, b y c), treonina (Thr), lactato, alanina (Ala), glutamato (Glu), prolina (Pro), glutamina (Gln), GSH, creatina, Cho, fosfocolina (PCho), glicerofosfocolina (GPCho), taurina, metanol, mI, glicina, glicerol, UDP, Tyr y CDP. Este conjunto de metabolitos reflejó alteraciones en las vías metabólicas relacionadas con el metabolismo de aminoácidos, GSH, taurina e hipotaurina, y glicerolípidos.

Estudios transcriptómicos en tejidos

El *PCA* explicó el 66% de la varianza de los datos y mostró claras diferencias entre las muestras tumorales y no tumorales considerando todo el transcriptoma. Un total de 4409 transcritos diferenciaron las muestras tumorales y las no tumorales, estando predominantemente *down*-regulados en los tumores (3112 transcritos). El análisis de GOBP reveló que esos genes *down*-regulados en tumores estaban asociados con la regulación del metabolismo de aminoácidos y aminas, los procesos biosintéticos de purinas y en genes vinculados a la fosforilación oxidativa (OXPHOS) y a la cadena de transporte mitocondrial. Por el contrario, entre los genes *up*-regulados en los tumores, no se encontraron categorías significativas relacionadas con el metabolismo.

Utilizando la información proporcionada por las bases de datos GOBP y GeneCards se identificaron un total de 364 genes codificantes vinculados con procesos metabólicos, de los cuales solamente 20 estaban *up*-regulados. Curiosamente, algunos de los genes *up*-regulados ya se habían relacionado con el cáncer como: el factor de transcripción PPARG involucrado en la regulación de la biogénesis de los lípidos; la hexoquinasa (HK2) y el transportador de glucosa

SLC2A1, ambos relacionados con la glucólisis; la ribosa 5 fosfato isomerasa A (RPIA) vinculada a la ruta de las pentosas fosfato; y genes que pertenecen al citocromo P450 (CYP450) (*CYP2J2, CYP2C9, CYP4F11*). Por otro lado, se encontraron genes metabólicos clave *down*-regulados en tejidos tumorales vinculados al: 1) metabolismo del piruvato: piruvato deshidrogenasas (*PDHA1, PDHB, PDHX*) y piruvato deshidrogenasa quinasa 4 (*PDK4*); 2) ciclo de los ácidos tricarboxílicos (TCA): *FH, IDH3A, IDH3B, MDH1, MDH2, ACO1, OGDH y SUCLG1*; 3) metabolismo de las poliaminas: *AMD1, SMS, ODC1, SAT2, AOC3*; 4) metabolismo de aminoácidos: *GLS, GOT2, MUT, ASS1, MAOA, MAOB*; 5) estado redox: *GPX3 y GPX4*, catalasa, *SOD1* y glutaredoxina (*GLRX3*). Es importante destacar que, aunque nuestra serie estuvo enriquecida con CVNMI, la expresión de los genes relacionados con el metabolismo *up*- and *down*-regulados mostró un comportamiento similar en múltiples muestras de tumores presentes en la cohorte TCGA (enriquecida en CVMI) sin discriminación de estadio, grado de tumor o diferentes subtipos tumorales identificados.

Por otra parte, los análisis *in silico* en la plataforma Enrich revelaron que varios factores de transcripción, actualmente relacionados con el CaV, se unían a múltiples genes metabólicos y podrían actuar como represores transcripcionales: ETS1, TTF2, E2F1, YY1, FLI1, ASH2L, E2F4, VDR, GABP, JARID1A, CTCF, KLF4 y CHD1. Se conoce que algunos de estos factores modulan la expresión génica a través de la remodelación directa de la cromatina, como GABP, JARID1A, CTCF y CHD1. Consecuentemente, utilizamos Encode Histone Roadmap para determinar si los genes metabólicos *down*-regulados estaban asociados con marcas de histonas particulares y encontramos marcas activadoras como H3K79me2, H3K79me3 o H3K27ac y marcas represoras como H3K36me3, H4K20me1 o H3K9me3.

Respecto a los análisis de empalme alternativo, se encontraron variantes significativas de *splicing* entre los tejidos tumorales y los controles en los siguientes genes metabólicos: *ASS1, GOT2, RARS, OAT, AOC3, HIBADH, DLD, FH,*

AUH, MUT, GLS, MAOA, ADH1C, ADH5, OGDH, SUCLG1, MDH1, MDH2, SMS, ODC1, FARSB, ACADM, PPM1L, PDK4, PDHX, PDHA1.

Finalmente, el análisis integrativo de los datos metabolómicos y transcriptómicos reveló alteraciones en múltiples rutas bioquímicas, tales como: el ciclo TCA, la OXPHOS y el metabolismo de aminoácidos y poliaminas, entre otros.

Estudios metabolómicos en orinas

El modelo *PLS-DA* (n=35, LVs=5) proporcionó una sensibilidad del 90,9%, una especificidad del 76,9%, un valor predictivo negativo (VPN) del 83,0%, un valor predictivo positivo (VPP) del 86,9% y un área bajo la curva Roc (AUROC)= 0,9. La prueba de permutación (100 permutaciones) proporcionó un valor de *p valor*<0,05 y evaluó la significación estadística del rendimiento predictivo del modelo mediante CV. Los metabolitos urinarios discriminantes identificados en el modelo (VIP>1) incluyeron: Val, ácido metilsuccínico, lactato, Ala, lisina, ácido N-acetilneuramínico, Gln, ácido succínico, citrato, creatinina, trimetilamina-N-óxido, metanol, taurina, sacarosa, ácido hipúrico, fenilalanina (Phe), pseudouridina y trigonelina. Estos metabolitos se vincularon con el metabolismo perturbado de la Ala, aspartato (Asp) y Glu, taurina e hipotaurina, ciclo TCA y biosíntesis de ARNt en los CaV. Algunas de las rutas metabólicas alteradas en orinas fueron comunes a aquellas alteradas en los tejidos.

Conclusiones

La integración de los datos metabolómicos y transcriptómicos indicó que la reprogramación metabólica en los tumores vesicales se produce principalmente a través de una regulación a la baja de los genes metabólicos relacionados con el ciclo TCA, la OXPHOS y el metabolismo de aminoácidos, entre otros. Estas alteraciones producidas en los tejidos tumorales se reflejan en el metaboloma urinario del mismo paciente y proporcionan un perfil metabólico específico capaz

de detectar el CaV con una elevada sensibilidad y especificidad mediante una aproximación no invasiva.

Capítulo 5

Con el objetivo de validar el perfil urinario encontrado en el capítulo 4 como un biomarcador de monitorización de CaV, se llevó a cabo un estudio con muestras urinarias adicionales de pacientes CVNMI. Se observó si el biomarcador urinario basado en perfiles de ^{1}H *NMR* permitía la detección de recurrencias durante el periodo de seguimiento activo.

Metodología

En este estudio, se incluyeron 20 hombres y 8 mujeres con diagnóstico de CaV y en lista de espera para ser intervenidos. Los criterios de inclusión para la selección de pacientes fueron: hombres y mujeres de 20 a 90 años diagnosticados de CVNMI, tumores únicos o múltiples, tumores primarios o recurrentes. Los criterios de exclusión fueron: pacientes con catéter urinario, CVMI (T2-T4), diagnóstico patológico de papiloma, diagnóstico único de *carcinoma in situ* (CIS) determinado mediante anatomía patológica (AP).

Los pacientes se clasificaron según los criterios de la EORTC en diferentes grupos de riesgo de recurrencia, y de forma mensual a todos se les recogió una muestra de orina durante un período de 18 meses. En total se recogieron 153 orinas, que fueron procesadas y almacenadas a -80 °C. Las muestras se clasificaron como CaV (n=70) cuando la cistoscopia fue positiva y la AP confirmó la presencia de un tumor tras la RTU; y un mes antes de la cistoscopia positiva con tumores ≥3 cm. Las muestras no tumorales recogidas después de la RTU del tumor se dividieron en dos grupos: CTRL (n=29) y MONITOR (n=38). El grupo CTRL incluyó las orinas recolectadas dentro de las 2-4 semanas después de la RTU; y el grupo MONITOR incluyó las orinas recolectadas durante el período de monitorización de los pacientes, considerando aquellas con AP negativas (T0),

orinas con cistoscopia negativa en el momento del muestreo y orinas recolectadas durante el período de monitorización entre cistoscopias negativas. Si no se disponía de cistoscopia en el momento del muestreo o después, las orinas se clasificaron como "evaluación cistoscópica no disponible" (NA) (n=16).

Las orinas se analizaron en un espectrómetro Bruker Avance II 500 MHz y se adquirieron espectros 1D de presaturación de agua y espectros 2D ^{1}H-^{13}CHSQC. Los espectros urinarios y los datos fueron pre-procesados de la misma forma que se indicó en el capítulo 4.

La intensidad de las resonancias asignadas se transfirió a *MetaboAnalyst 3.0*. La prueba de U-Mann Whitney determinó las diferencias significativas entre las orinas control y las CaV. Kruskal-Wallis evaluó las diferencias entre el control, los tumores primarios y las recurrencias; y comparó la media entre las muestras tumorales con estadios Ta, T1 y las muestras control.

Para el desarrollo del modelo *PLS-DA* se consideró la región 0.8-4.5 y 6.5-9.0 ppm del desplazamiento químico del espectro, y el conjunto de datos se dividió en dos subconjuntos: calibración y validación. El conjunto de calibración se utilizó para desarrollar el modelo e incluyó 69 muestras de 24 pacientes (48 CaV y 21 CTRL). El conjunto de validación se utilizó para evaluar el rendimiento predictivo del modelo e incluyó orinas de 7 pacientes con diferentes evoluciones clínicas. En primer lugar, el conjunto de validación se realizó considerando 22 CaV y 8 orinas CTRL. Después, se añadieron al conjunto de validación las muestras MONITOR (n=38) y NA (n=16) para evaluar el rendimiento del modelo en la detección de recurrencias durante un período de vigilancia. Al agregar las muestras MONITOR, el número de muestras de control aumentó a 46. Los metabolitos más importantes en el modelo estadístico (*VIP*>1) se identificaron y se utilizaron para realizar un análisis de enriquecimiento de las vías metabólicas en la plataforma *MetaboAnalyst 3.0*.

Resultados

Los niveles de ácido hipúrico y Ala mostraron diferencias significativas entre los grupos CaV y control (*p valor* <0,05). Ambos metabolitos presentaron mayores intensidades en las muestras control. El ácido hipúrico también presentó diferencias significativas entre los tumores primarios y las recidivas, y entre los estadios Ta, T1 y las muestras control.

Para el conjunto de validación (CaV *vs* CTRL), el modelo PLS-DA proporcionó una sensibilidad del 86,4%, una especificidad del 87,5% y un AUROC=0,96. Cuando se consideraron las muestras MONITOR dentro del subconjunto de validación, el valor de sensibilidad se mantuvo, pero la especificidad fue del 80,4%. Aun así, la AUROC fue de 0,89 y el VPN 92.7%. Los metabolitos urinarios discriminantes identificados en el modelo (*VIP*>1) incluyeron: Val, Ala, Lys, Gln, citrato, dimetilamina, creatinina, trimetilamina N-óxido, taurina, sacarosa, creatina, ácido hippúrico, histidina, Phe y trigonelina. El análisis de las vías realizado en *MetaboAnalys*t, vinculó estos metabolitos con alteraciones en el metabolismo de la taurina e hipotaurina y diferentes vías relacionadas con el metabolismo de aminoácidos: Ala, Asp, Glu, Arg y Pro, y Phe.

Los análisis longitudinales del perfil metabolómico urinario se evaluaron considerando las evoluciones clínicas de 7 de los 28 pacientes incluidos en el estudio. La trayectoria longitudinal del perfil metabolómico de estos pacientes permitió evaluar su utilidad en la detección de recidivas tumorales durante el periodo de seguimiento activo, dado que su comportamiento concordó con los resultados de la cistoscopia y la AP tras la RTU. Además, el perfil metabólico no se vio afectado por procesos inflamatorios como la cistitis, ya que el modelo clasificó correctamente las orinas recolectadas durante estas circunstancias.

Conclusiones

El presente estudio muestra por primera vez un perfil metabólico urinario, altamente dinámico, y basado en ^{1}H *NMR*, que cambia de un fenotipo tumoral a

uno control después de la extirpación del CVNMI y regresa a la condición maligna cuando ocurre una recurrencia. Este hecho potencia la metabolómica como herramienta para identificar biomarcadores no invasivos, que podrían aplicarse en la clínica para mejorar el manejo de los pacientes con CVNMI, ya sea: disminuyendo el número de cistoscopias durante el período de seguimiento; detectando lesiones no visibles por cistoscopia (displasias, hiperplasias, CIS); o detectando recurrencias tempranas, RTU incompletas o quizás tumores en el tracto urinario superior.

Capítulo 6

El último capítulo de la tesis tuvo como objetivo la identificación de un perfil metabolómico urinario basado en *UPLC-MS* que permitiese detectar los cambios metabolómicos producidos tras la extirpación del CVNMI, y que consecuentemente permitiese la detección de recurrencias en pacientes con CVNMI sometidos a un periodo de seguimiento activo. Dado que parte de las muestras analizadas en este estudio fueron comunes a aquellas inlcuidas en el estudio presentado en el capítulo 5, también se pretendió aumentar la cobertura de los metabolitos vinculados al CVNMI y determinar si los resultados obtenidos mediante ambas técnicas eran similares.

Metodología

En este estudio se incluyeron 31 pacientes con diagnóstico de CaV y en lista de espera para ser operados. Los criterios de inclusión y exclusión, y la codificación de las muestras fue la misma que la explicada en el capítulo 5, pero en este caso el número de muestras analizadas fue mayor con 316 orinas (CaV=68, CTRL=29, MONITOR=166, NA=53).

A 100 µL de orina y se agregaron 200 µL de HCOOH 0,1% v/v en H_2O. La muestra se homogeneizó y se centrifugó a 10000 x g (4 °C, 10 min). 100 µL del sobrenadante se transfirieron a una placa de 96 pocillos donde a cada muestra se

le añadieron 5 µL de los siguientes patrones internos: fenilalanina-D_5, cafeína-D_9, leucina encefalina, y reserpina en H_2O:CH_3OH (1:1, 0.1% v/v HCOOH) en una concentración final de 1 µM. Los blancos se prepararon reemplazando la orina por H_2O. Se prepararon muestras de control de calidad (QC) mezclando 5 µL de cada muestra. Todas muestras de analizaron mediante *UPLC-QTOF-MS* en dos tandas (*bach 1* y *bach 2*) para reducir el tiempo en que las muestras permanecian el inyector.

Los datos de *UPLC-TOF-MS* sin procesar se convirtieron a formato mzXML utilizando *ProteoWizard* antes de generar las tablas de picos usando el software XCMS. El método centWave se utilizó para la detección de picos con los siguientes parámetros: ppm: 15, ancho de picos: (5, 20), umbral de la señal de ruido: 6. Se seleccionó una diferencia mínima en m/z de 5 mDa para picos con tiempos de retención (RT) superpuestos. Los valores m/z ponderados por intensidad de cada característica se calcularon utilizando la función wMean. La agrupación de picos se llevó a cabo usando el método "más cercano" considerando RT=1 y tolerancias de RT y m/z de 6 s y 5 mDa, respectivamente. Después de la agrupación de picos, se aplicó el método *fillPeaks* con los parámetros predeterminados para completar los datos de picos faltantes. Las tolerancias de RT y m/z utilizadas para la generación de la tabla de picos y la alineación de las características en los *baches* se basaron en la variación observada en cinco metabolitos seleccionados (Phe, Trp, kynurenine, hydroxykynurenine y phenylacetylglutamine) y los estándares internos indicados anteriormente. La precisión de integración se evaluó comparando los resultados de integración manual y automatizada para los estándares internos. Se encontraron un total de 4299 características en el *batch* 1 y 4416 características en el *batch* 2 después de la detección de picos. Los blancos se utilizaron para identificar y eliminar las características de fondo que surgen de la presencia de contaminantes, plastificantes, o impurezas en los solventes. La eliminación de efectos dentro del *batch* se realizó ajustando funciones no lineales

dependientes del tiempo a los QC inyectados, seguido de una normalización de los datos utilizando la función QC-Suport Vector Regresión (SVRC) y la función "*radial basis function kernel*".

Los efectos entre *batches* se eliminaron al escalar la intensidad de cada característica metabólica en cada muestra utilizando un factor definido como la relación entre la intensidad mediana en QC en el *batch* correspondiente y la intensidad mediana entre *batches*. Finalmente, las características metabólicas que mostraron una desviación estándar relativa (RSD%)>15 en las muestras QC se consideraron no confiables y se eliminaron, lo que dejó 2006 caracteristicas para el análisis de datos.

Para la creación de los diferentes modelos *PLS-DA*, el conjunto inicial de muestras se dividió en los subconjuntos: calibración y validación. El conjunto de calibración se utilizó para el desarrollo del modelo *PLS-DA* y la selección de características. El conjunto de validación se utilizó exclusivamente para la evaluación del rendimiento predictivo del modelo. El escalado de datos incluyó la corrección de dispersión multiplicativa con las medianas de los QC como referencia, seguido de un escalado Pareto.

La identificación de los metabolitos se realizó comparando los valores m/z con las bases de datos HMDB y METLIN considerando una precisión de 5. Las fórmulas moleculares se estimaron por MassHunter Workstation Software-Qualitative Analysis (Agilent). La adquisición de datos y la integración manual de los picos de los estándares internos y los metabolitos seleccionados se realizaró utilizando MassHunter (Agilent). Los análisis de tipo *PLS-DA* se llevaron a cabo utilizando PLS_Toolbox 8.0. y MATLAB (Mathworks Inc., Natick, MA, EE. UU.) Para el análisis de rutas metabólicas se utilizó *MetaboAnalyst* 3.0.

Resultados

Para facilitar los análisis de las diferencias en los perfiles metabólicos entre las muestras CaV, CTRL y MONITOR se consideraron tres modelos *PLS-DA*

independientes en los que los grupos se compararon por pares (es decir, CaV *vs* CTRL, BC *vs* MONITOR y CTRL *vs* MONITOR). El modelo CaV *vs* CTRL proporcionó una clasificación precisa de 27/33 CaV y 10/11 muestras CTRL (sensibilidad: 82% y especificidad: 91%). El segundo modelo, CaV vs MONITOR, se comportó peor en la clasificación de las muestras tumorales (sensibilidad: 70% y especificidad: 75%). Finalmente, el análisis de las diferencias entre los grupos CTRL y MONITOR proporcionó rendimientos predictivos no significativos, menor sensibilidad (45%) y especificidad (76%), de acuerdo con una mayor superposición de las muestras CTRL.

Vistos los resultados previos, el modelo CaV *vs* CTRL se optimizó seleccionando un conjunto de 128 características discriminantes considerando una VIP>3. Este modelo clasificó correctamente 29/33 CaV y 11/11 orinas CTRL en el conjunto de validación, lo que proporcionó una AUROC=0.96 y una sensibilidad ligeramente mejorada (87.9%, especificidad 100%). Ajustando por probabilidad previa de recurrencia (15%, 24%, 28% y 61%) los VPN para los grupos de riesgo bajo, intermedio bajo, intermedio alto, y alto fueron 96,5%, 94,0%, 92,9% y 76,1% respectivamente. Los metabolitos discriminantes identificados en el modelo CaV *vs* CTRL con una VIP>1 reflejaron alteraciones en las rutas metabólicas de la Arg, Pro, los ácidos grasos, la Phe, las purinas y pirimidinas y el Trp, entre otras.

Por otro lado, el análisis de las trayectorias longitudinales del biomarcador metabólico desarrollado para la discriminación entre las muestras CaV y CTRL permitió evaluar su utilidad en la detección de recurrencias. Estas trayectorias se observaron en 6 pacientes diferentes con evoluciones clínicas. De la misma forma que en capitulo 5, la trayectoria longitudinal del perfil metabolómico durante el periodo de seguimiento activo mostró un cambio gradual de un fenotipo control a uno tumoral asociado a la aparición de la recurrencia. Este hecho fue consistente con los resultados confirmatorios de cistoscopia y la AP positiva tras la RTU del tumor. Sin embargo, en algunos pacientes, los resultados mostraron trayectorias

inconsistentes debido a un número reducido de muestras mal clasificadas. El análisis de orina es un reto debido a la variación en la composición química y la amplia variedad de factores de confusión que pueden afectar al metaboloma urinario (dieta, consumo de agua, exposición ambiental, ingesta de medicamentos). Por lo tanto, se necesitan más estudios que evalúen las fuentes de variabilidad en la orina y permitan aumentar la robustez de las pruebas para que los análisis metabolómicos puedan trasladarse a la práctica clínica.

Conclusiones

Los resultados de este estudio clínico revelaron un cambio estadísticamente significativo en el perfil metabólico urinario antes y después de la RTU, y respaldaron la hipótesis de un perfil metabólico basado UPLC-MS capaz de detectar recurrencias en pacientes bajo seguimiento activo.

En general, los resultados presentados en esta tesis respaldan la hipótesis de la existencia de una huella metabólica urinaria vinculada a las alteraciones tumorales presentes en los tejidos vesicales, capaz de detectar y predecir las recurrencias durante el período de vigilancia en pacientes con CVNMI. Además, los buenos resultados obtenidos y la concordancia entre ambos estudios urinarios (^{1}*H NMR y UPLC-MS*) posicionan la metabolómica al frente de las técnicas ómicas para la búsqueda de biomarcadores robustos y dinámicos que reflejen la biología del tumor.

Abbreviations and Acronyms

A

AA	Amino acid
Acetyl CoA	Acetyl coenzyme A
ACC	Accuracy
ACLY	ATP citrate lyase
ACS	Acyl-CoA synthetases
ADP	Adenosine diphosphate
Ala	Alanine
Arg	Arginine
AS	Alternative splicing
ASL	Argininosuccinate lyase
Asn	Asparagine
Asp	Aspartate
ASS-1	Argininosuccinate synthetase 1
ATP	Adenosine 5'-triphosphate
AUROC	Area Under the Receiver Operating Characteristic

B

BC	Bladder cancer
BCAAs	Branched chain amino acids
BCATc	Cytosolic branched chain aminotransferase
BCG	Bacillus Calmette-Guérin
1,3biPG	1,3-biphosphoglycerate

C

CDP	Cytidine diphosphate
CE	Capillary Electrophoresis
cDNA	Complementary DNA
ChEA	Chip Enrichment Analysis
Cho	Choline
CIS	Carcinoma *in situ*
CTRL	Control
COSY	Correlation spectroscopy

CPT1	Carnitine palmitoyl transferase 1
CTCs	Circulating tumor cells
ctDNA	Circulating tumor DNA
CUETO	Club Urologico Español de Tratamiento Oncológico
CV	Cross validation
Cys	Cysteine

D

1D	One-dimensional
2D	Two-dimensional
DG	Diacylglycerides
DNA	Desoxiribonucleic acid
DNMT	DNA methyltransferases
DSS	4,4-Dimethyl-4-silapentane-1-sulfonic acid

E

EMT	Epithelial-mesenchymal transition
EORTC	European Organization for the Research and Treatment
ESI	Electrospray ionization

F

F1,6biP	Fructose-1,6-bisphosphate
FA	Fatty acid
$FADH_2$	Flavin adenine dinucleotide
FASN	Fatty acids synthase
FC	Fold change
FDA	Food and Drug Administration
FH	Fumarate hydratase
FID	Free induction decay

G

GA3P	Glyceraldehyde 3-phosphate
GC-MS	Gas Chromatograpy-Mass Spectrometry
GEO	Gene expression omnibus
GOBP	Gene Ontology of Biological Processes
Gln	Glutamine
GLS	Glutaminase

Glu	Glutamate
Gly	Glycine
GPCho	Glycerophosphocholine
GPT2	Alanine aminotrasferase 2
GPx	Glutathione peroxidases
GSEA	Gene Set Enrichment Analysis
GSH	Glutathione
GSR	Glutathione reductase

H

HAT	Histone acetyltransferases
HDM	Hystone demethylases
HG	High grade
2-HG	2-Hydroxyglutarate
HIF	Hypoxia inducible factor
His	Histidine
HK	Hexokinase
HMDB	Human Metabolome Data Base
HMT	Histone methyltransferases
HRMAS NMR	High Resolution Magic Angle Spinning Nuclear Magnetic Resonance
HSQC	Heteronuclear Single-Quantum Correlation
5-HTP	5-Hydroxytryptophan

I

IDH1	Isocitrate dehydrogenase 1
IDO	Indoleamine-2,3-dioxygenase
Ile	Isoleucine
ISs	Internal standards
ISUP	International Society of Urological Pathology

K

KEGG	Kyoto Encyclopedia of Genes and Genomes
α-KG	Alpha-ketoglutarate
KHK	Ketohexokinase
α-KIC	Alpha-ketoisocaproate

L

Leu	Leucine
LG	Low grade
Lys	Lysine

M

MAG	Monoacylglycerols
MCT	Monocarboxylate transporters
Met	Methionine
mI	Myo-insitol
MIBC	Muscle-invasive bladder cancer
miRNA	microRNA
MS	Mass Spectrometry
MSEA	Metabolite Set Enrichment Analysis
mTHF	5-Methyltetrahydrofolate
mTOR	Complex 1 of the mammalian target of rapamycin

N

NA	Non available
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
ncRNAs	Non-coding RNAs
NGS	Next generation sequencing
NMIBC	Non muscle-invasive bladder cancer
NMR	Nuclear Magnetic Resonance
NO	Nitric oxide
NOESY	Nuclear Overhauser Effect Spectroscopy
NLR	Negative likelihood ratio
NPV	Negative predictive value

O

OAA	Oxalacetate
OGDH	Oxoglutarate dehydrogenase
OXPHOS	Oxidative phosphorylation

P

PA	Pathological anatomy
PC	Principal Component

PCA	Principal Component Analysis
PCho	Phosphocholine
PCR	Polymerase chain reaction
PEP	Phosphoenolpyruvate
PFK	Phosphofructokinase
3-PG	3-Phosphoglycerate
Phe	Phenylalanine
PHGDH	Phosphoglycerate dehydrogenase
3-PHP	3-phosphohydroxy pyruvate
PI3K	Phosphatidylinositol 3-kinase
PK	Pyruvate kinase
PKM	Pyruvate kinase muscle isoenzyme
PLs	Phospholipids
PLR	Positive likelihood ratio
PLS-DA	Partial Least Square Discriminant Analysis
PPARs	Peroxisome-proliferator-activated receptors
Pro	Proline
PPP	Pentose phosphate pathway
PPV	Positive predictive value
PtdCho	Phosphatidylcholine
PtdE	Phosphatidylethanolamine
PUNLMP	Papillary urothelial neoplasm of low malignant potential.

Q

QC	Quality control
QTOF	Quadrupole time of flight

R

RC	Radical cystectomy
RMSECV	Root Mean Square Error for Cross-Validation
RNA	Ribonucleic acid
RNA-Seq	RNA-Sequencing
RTs	Retention times
ROS	Reactive oxygen species
RSD	Relative standard deviation
RT-PCR	Reverse transcription-polymerase chain reaction

RT-qPCR	Reverse transcription-quantitative polymerase chain reaction

S

SAM	S-adenosyl methionine
SDH	Succinate dehydrogenase
Ser	Serine
SHMT1	Serine hydroxymethyl transferase 1
SLC2A1	Solute carrier family 2, facilitated glucose transporter member 1
SMPDB	Small Molecule Pathway Database
SNPs	Single nucleotide polymorphisms
SOD	Superoxide dismutase
SVRC	Support Vector Regression

T

TCA	Tricarboxylic acid
TCGA	The Cancer Genome Atlas
TDO	Tryptophan-2,3-dioxygenase
TFs	Transcription factors
TG	Triacylglycerides
THF	Tetrahydrofolate
Thr	Threonine
TNM	Tumor-Node-Metastasis system
TOCSY	Total correlation spectroscopy
TOF	Time of flight
TPH	Tryptophan hydroxylase-1
Trp	Tryptophan
TSP	3-Trimethylsilylpropionate
TUR	Transurethral resection
Tyr	Tyrosine

U

UDP	Uridine diphosphate
UGT	5'-diphosfoglucuronosil transferase
UK	Unknown
UPLC-MS	Ultra Performance Liquid Chromathography-Mass Spectrometry
UTUC	Upper tract urothelial carcinoma

V

Val	Valine
VEGF	Vascular endothelial growth factor
VIP	Variable Importance in Projection

W

WHO	World Health Organization

Table of Contents

Chapter 1: Introduction

1.1 Epidemiology of bladder cancer and risk factors

Bladder cancer (BC) is the most common cancer of the urinary tract and represents an unsolved clinical, social and economic problem[1,2]. BC is more frequent in men than women with a ratio of 3:1. Specifically in men, it is the 7th most common cancer worldwide with 330380 new cases per year and it is among the top 10 most deadly cancers[3]. In women, BC presents a lower incidence rate with 99413 cases per year[3].

The highest rates of incidence and prevalence of this tumor are observed in developed countries. Each year in the European Union approximately 110500 men and 70000 women are diagnosed and 38200 patients die from BC[2]. Spain has one of the highest incidences in Europe. According to both sexes, the incidence of BC rank 5th and the mortality 7th in Spain[3] (**Figure 1.1**), but considering only men, Spain shows the highest BC incidence rates among European countries and this tumor occupies 2nd place in terms of mortality[4].

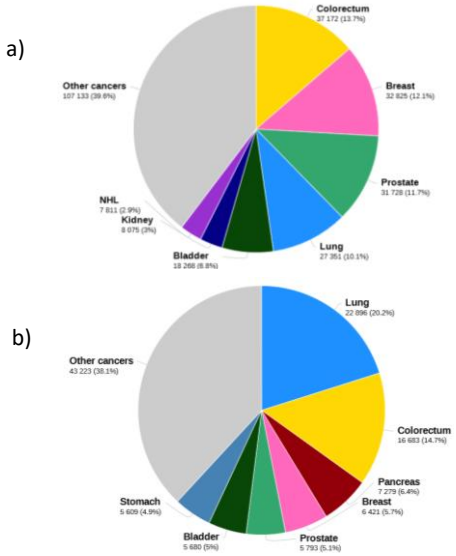


Figure 1.1 Estimated number of cancer cases (a) and deaths (b) in Spain considering both sexes[5].

The gender discrepancy and the high incidence rates in Spain have been associated with BC risk factors including chronic exposure to carcinogens (both, environmental and associated to life habits such as smoking), age, sex, and genetic susceptibility related to altered metabolic enzymes involved in the detoxification of xenobiotics[2,6–11].

1.2 Clinical management of bladder cancer

1.2.1 Pathological criteria for BC diagnosis

BCs are diagnosed as NMIBC or muscle-invasive (MIBC) according to their stage, which is determined by Tumor-Node-Metastasis (TNM) system[12]. TNM system considers the extent of tumor invasion (Tis-T4) in base on pathological criteria (**Figure 1.2**). NMIBCs include papillary tumors confined to the mucosa (Ta) or invading the lamina propria (T1) and flat high-grade lesions confined to the mucosa classified as carcinoma *in situ* (CIS)[13]. MIBCs include the stage of tumors ≥T2. The stage designations T_X indicate that the primary tumor could not be assessed, while T_0 indicates that there was no evidence of a tumor. The tumor stage is the most important factor for treatment selection[14].

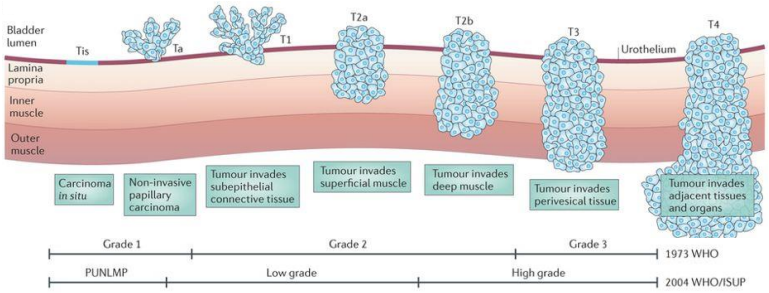


Figure 1.2 Bladder cancer staging according to TNM sytem and grading according to 1973 WHO; and 2004 WHO and International Society of Urological Pathology (ISUP)[6]. Note: PUNLMP: Papillary urothelial neoplasm of low malignant potential.

The grade of tumors is determined according to the degree of cellular atypia, growth pattern, and mitotic activity. The grade is also considered an important prognostic indicator because it reflects the inherent aggressiveness of the tumor. Particularly in NMIBC, where a high histological grade is associated with a higher risk of recurrence or progression to MIBC[14,15].

In clinical practice, two parallel systems are used for grading tumors: the "World Health Organisation (WHO) 1973" and the "WHO/ISUP 2004/2016" systems. While the WHO 1973 grading system uses a three-tiered system (Grade 1-3) where increasing grade reflects the degree of cellular abnormality, the "WHO/ISUP 2004/2016" separate tumors into only two categories: low grade (LG) and high grade (HG), considering as HG both grade 2 and 3 in the WHO 1973 system (**Figure 1.2**). The purpose of the "WHO/ISUP 2004/2016" system is to reduce the inter-observer variability that was seen in 1973 system, which was criticized for ambiguous and poorly defined separation criteria, due to only very low-risk cases were classified as grade 1 and only those with huge cellular abnormality as grade 3, being the majority classified as grade 2[16,17].

- *Carcinoma in situ*

At diagnosis, approximately 10% of patients with BC present CIS[18]. CIS is a flat high grade urothelial carcinoma that increases the risk of recurrence and progression. Between 40%-83% of patients with CIS develop muscle invasion if untreated, especially if it is associated with papillary tumors[19]. Often it is not visible by cystoscopy, so a biopsy is needed for its diagnosis. Even so, CIS diagnosis varies significantly according to the pathologist, showing concordance in 70–78% of cases[20].

At the same way that other types of bladder tumors, CIS can be multifocal and to be present either in the bladder as in the upper urinary tract, prostatic ducts or

prostatic urethra[13]. This characteristic is known as pan-urothelial disease and makes more difficult the management of the BC.

- *Non muscle-invasive bladder cancer*

Approximately 75–85% of tumors are NMIBC, being Ta tumors the most frequent with 70% of cases. The rest of them are high grade T1 tumors (5-20%)[19]. The staging and grading between Ta and T1 tumors present a 50%-60%[21–23] of concordance between specialists. Usually, NMIBCs are associated with good prognosis. However, they present elevated rates of recurrence (50-70%) and moderate rates of progression (10-15%), being the 5-year survival ~90%[6].

Ta tumors are confined to the urothelium, have a papillary appearance and are usually of LG. Recurrence is common when there is multiplicity but progression is rare. The most important factor for progression is the grade of tumor[19].

Many T1 tumors are understated, so tumor extension is greater than the transurethral resection (TUR) specimen. Specifically, 33% of patients diagnosed with NMIBC presented muscle invasion during cystectomy[19]. This fact is important and should be controlled because T1 HG tumor has been definite as an aggressive and potentially lethal disease[24]. In fact, T1 HG cancer has a 69% to 80% recurrence rate and a probability of progression of 33% to 48% after the first TUR[24].

- *Muscle-invasive bladder cancer*

MIBC is a complex disease. About 15% of patients diagnosed present previously a non-invasive carcinoma. However, 80-90% of the MIBCs are primary and approximately 5% of those present metastasis at diagnosis. MIBCs have less favorable prognosis than NMIBC with 5-year survival <50% and common progression to metastasis[6].

1.2.2 Diagnostic techniques

One of the most prevalent symptoms of BC is macroscopic or microscopic blood in the urine (hematuria). However, several inflammatory benign conditions that affect the bladder can be associated with this symptom. Therefore, hematuria only has a 5% specificity in its diagnostic role[25].

Urinary cytology is one of the principal method used for BC diagnosis for more than 50 years. It presents advantages such as: is a non-invasive method, is not affected by hematuria or chemical confounders such as pH or salt concentrations; has reasonable in-house cost, and its preparation is relatively simple. However, urinary cytology presents the following drawbacks: cytological interpretation is user-dependent[25] and a normal cytology outcome does not exclude the existence of a tumor. In fact, urinary cytology has a high sensitivity in HG tumors (84%) but low sensitivity in LG and early tumors (16%)[13]. Hence, cystoscopy is still the standard method for BC diagnosis and monitoring. Nevertheless, this not means that cystoscopy is an ideal technique since it has limitations. Cystoscopy is an invasive approach that requires specialized personnel and equipment, making BC one of the most expensive tumors for health systems[26]. Moreover, the sensitivity and specificity of cystoscopy are operator-dependent (62-84%, 43-98%, respectively)[27]; pre-neoplasic lesions (dysplasias and hyperplasia), as well as, CIS are imperceptible, and areas of inflammation can be confused with tumors. These limitations are important, especially the overlooking of CIS at diagnosis since this type of tumors are related to high risk of progression[18].

Fluorescence cystoscopy has been performed using violet light after intravesical instillation of 5-aminolevulinic acid or hexaminolevulinic acid to improve the detection of malignant tumors, specifically CIS[28]. But the value of fluorescence cystoscopy for improvement of outcome in relation to progression

rate and survival has not been demonstrated yet. Therefore, currently, CIS is diagnosed by a combination of cystoscopy, urine cytology, and histologic evaluation of multiple bladder biopsies.

1.2.3 Risk prediction models for non-muscle invasive bladder cancer

The European Organisation for the Research and Treatment of Cancer (EORTC) and Club Urologico Español de Tratamiento Oncologico (CUETO)[15] are two well-studied risk assessment models that are used for prognostication of NMIBC.

Both systems incorporate the WHO 1973 tumor grade together with multiple clinical variables in order to predict the recurrence and progression risk of a patient[14]. Each variable has a weight based on its predictive value, and by summing the score of all variables the patients are categorized into low, intermediate and high-risk groups, which present different probabilities of recurrence and progression[14].

The EORTC criteria have been well established in the urological context and are used to know the prognosis of patients with NMIBC and to select the optimal treatment for each one.

1.2.4 Bladder cancer treatment

- *Transurethral resection of bladder tumor*

In patients with BC diagnosis, a TUR of bladder tumor followed by pathology examination of the obtained specimen is recommended as a diagnostic procedure and initial treatment step. In NMIBC the TUR allows to know the pathological diagnosis and extirpate all visible lesions. A second TUR is necessary after incomplete initial TUR if there is no muscle in the specimen after initial resection

(with exception of Ta G1 tumors and primary CIS) and in all T1 tumors or tumors with HG (except primary CIS)[13].

- *Treatment for non-muscle-invasive bladder cancers*

Low-risk bladder cancer

Ta LG papillary tumors have the most favorable oncologic outcomes. In these tumors, an immediate single postoperative instillation of chemotherapy within 6-24h after complete TUR reduces the risk of recurrence for the first 2 years[29]. Mitomycin C, epirubicin and doxorubicin are used as chemotherapeutic drugs[19]. The instillation intends to destroy circulating tumor cells after TUR and residual tumor cells at the resection site that could lead to future recurrences. High variability in the 3-month recurrence rate has been observed, which could indicate that the prior TUR was incomplete[19].

Intermediate- and high-risk bladder cancer

Intravesical *Bacillus Calmette–Guérin* (BCG) immunotherapy is indicated in intermediate and high-risk BC. BCG is an attenuated mycobacterium developed as a vaccine for tuberculosis. Its mechanism of action remains unknown but involves immune activation. BCG is internalized by urothelial cells and unchains a cascade of cytokine production, including interleukins, tumor necrosis factor, interferon, and granulocyte macrophage colony-stimulating factor. Subsequently, neutrophils, macrophages, dendritic cells, and $CD4^+$ lymphocytes are recruited to the site, resulting in cytotoxicity to BC cells[30].

Although the complete response rate to BCG therapy in high-risk NMIBC patients is high (80%), some patients have recurrences[24]. When this occurs, other conservative secondary treatments[19] or early radical cystectomy (RC) can be applied[24].

In the histological diagnosis of CIS, no consensus exists about whether conservative therapy (intravesical BCG instillations) or aggressive therapy by RC should be done, especially in combination with concurrent HG papillary tumors[19].

- *Treatment for muscle-invasive bladder cancers*

MIBC requires radical management, often in older patients with comorbid conditions. RC with bilateral pelvic lymph node dissection is the standard procedure, which increases overall survival in non-metastatic patients across all tumor stages[31]. In some cases, neoadjuvant chemotherapy with cisplatin can be administered before RC. Timely diagnosis and early surgery are crucial to improving patient outcomes.

1.3 Biomarkers of cancer

In 1998, the National Institutes of Health Biomarkers Definitions Working Group defined a biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention". In the specific context of cancer, several types of biological samples coming from patients (blood, urine, saliva, semen) are being used with the purpose of identifying dynamic and non-invasive biomarkers of disease[32,33]. The TNM staging provides a basis to predict survival, choice initial treatment, stratification of patients in clinical trials, and uniform reporting of the end result of cancer management[34]. However, these data offer static information about a disease and require a piece of the tumor to be evaluated pathologically. Consequently, the use of biomarkers could be useful in clinical practice, both to complement information about cancer staging or grading, as for non-invasive diagnosis and monitoring of the disease. Besides, molecular

target therapies are becoming more common and they need their associated biomarkers to inform about the clinical response, independently of initial TNM stage[34]. For all this, the field of biomarker research is booming and different approaches have been developed considering different levels of molecular regulation (genome, epigenome, transcriptome, proteome, and metabolome, among others).

In the last years, the development of high-throughput techniques has evidence that novel biomarkers can be discovered through several emerging technologies. The analyses of biologic samples by technologies such as MS, NMR, microarrays and next-generation sequencing (NGS), among others, have provided opportunities to identify discriminant profiles (signatures), to be used as biomarkers of diagnosis, prognosis, and monitoring. Nevertheless, limitations related to study design, inter- and intraindividual variability and low diagnostic sensitivity and specificity have prevented the translation of the biomarker to the clinical routine. The ideal biomarker assay should be sensitive, specific, cost-effective, fast, preferably non-invasive, and robust against inter-operator variability.

1.3.1 Biomarkers of bladder cancer

In the last decade, advanced technology has utilized mostly urine of patients with BC to identify non-invasive biomarkers for screening, early diagnosis, surveillance, staging, and prognosis[35]. To date, many candidate biomarkers based on proteins, circulating tumor cells (CTCs), circulating tumor DNA (ctDNA), microRNAs (miRNAs), histone marks, exosomes, and metabolic profiles have been reported in the context of BC, but only six urinary diagnostic tests have been approved by the FDA for clinical use: BTA-STAT, BTA-TRAK and immunocytology

(uCyt+) for BC surveillance, and NMP22 BC test kit, NMP22 BladderChek and UroVysion for BC diagnosis and surveillance[35–37]. Other two tests based on mRNA profiles (CxBladder, and Xpert Bladder Cancer) are also commercially available for BC diagnosis and monitoring[38,39]. Even so, none of them outperforms cystoscopy and have been implemented into the clinic.

1.4 Biology of cancer

1.4.1 Hallmarks of cancer

The definition of cancer encompasses a set of very complex diseases characterized principally by an uncontrolled cell proliferation that leads to tumor development. During this sequential process, the cells acquire capacities that allow them to survive in a hostile environment. These capabilities represent the hallmarks of cancer and are related to sustaining proliferative signaling, insensitivity to antigrowth signals, resisting apoptosis, enabling replicative immortality, inducing angiogenesis, and enhancing invasion and metastasis[40]. Underlying these capabilities different cellular processes are involving such as genome instability, inflammation, reprogramming of energy metabolism and evading immune[40].

1.4.2 Cancer metabolism

The studies performed in the last years have indicated that metabolic reprogramming not only maintains the proliferation and growth of tumors through supplying nutrients but also affects the tumor microenvironment and has effect in immune system[41]. This reprogramming of cellular metabolism can be

dependent on oncogenic mutations but epigenetic and environmental processes can influence.

Aided by emerging molecular tools, studies in cancer cell metabolism have increased the knowledge about tumor-associated metabolic alterations in different stages of its development. In particular, these metabolic alterations have been grouped in several hallmarks that consider: the deregulated uptake of glucose and amino acids (AA); the use of glycolysis/tricarboxylic acid (TCA) cycle intermediates for biosynthesis and nicotinamide adenine dinucleotide phosphate (NADPH) production; the increased demand for nitrogen; metabolic interactions with the microenvironment and immune system; the altered fatty acid metabolism; and the redox status[42].

- *Metabolic hallmarks of cancer*

Glucose metabolism

Cancer cells have to increase the uptake of nutrients from the environment to supply bioenergetic and biosynthetic demands of proliferation and growth. The two main molecules to support survival are glucose and glutamine (Gln), since their carbon skeletons can be oxidated allowing capture reducing power in the form of nicotinamide adenine dinucleotide (NADH) and flavin adenine dinucleotide ($FADH_2$) that is used in electron transport chain in the mitochondria as fuel to ATP generation; or in the form of NADPH, which is used in biosynthetic reactions but also to maintain cellular redox state[42].

Glycolysis is an oxygen-independent metabolic pathway that uses glucose as initial substrate to create pyruvate and other carbon intermediates, which can be used in a wide variety of metabolic pathways. For example, the intermediate dihydroxyacetone phosphate that serves as a backbone precursor for diverse phospholipids (PLs), or the glucose 6-phosphate that is used into pentose

phosphate pathway (PPP) to generate reducing power, pentoses but also ribose-5-phosphate a precursor for nucleotide synthesis[42].

On the other hand, the pyruvate obtained in glycolysis can be derivated towards two processes: one anaerobic (lactic fermentation), and other aerobic, which involves the generation of acetyl-Coenzyme A (acetyl CoA), its oxidation in TCA cycle and the use of the reducing power obtained to create ATP in the electron transport chain (oxidative phosphorylation (OXPHOS)). Although OXPHOS offers a higher energy performance than lactic fermentation (36 ATP *versus* 2 ATP), in 1927 Otto Warburg observed that cancer cells preferred to oxidize glucose through anaerobic respiration, even in oxidative conditions[43] (**Figure 1.3**). Why proliferative cancer cells would prefer a pathway less beneficial? Tumor cells require energy but also precursor molecules and reducing equivalents coming from glucose catabolism. Although TCA cycle offers these two demands, it is the major negative regulator of glucose metabolism[42]. Therefore, by deriving the excess of pyruvate to lactate production, the tumor cells would prevent the accumulation of NADH and ATP, promoting the metabolism of cytosolic glucose free of repressive feedback.

Regarding glucose regulation, this metabolic pathway is controlled by itself through the activity of three key allosteric enzymes; hexokinase (HK), phosphofructokinase (PFK) and pyruvate kinase (PK), which converts phosphoenolpyruvate (PEP) in pyruvate in the last step of glycolysis[44]. Moreover, signaling pathways such as phosphatidylinositol 3-kinase pathway (PI3K/Akt) also acts as a master regulator of glucose uptake. Genetic alterations of *PFK* isoforms has been related with an invasive and more aggressive state of the disease (MIBC)[45,46]; and the upexpression of the variant PKM2 (an oncogene) with poor prognosis for BC patients[47]. Unlike PKM1 that is more efficient promoting glycolysis, PKM2 is ineffective promoting glycolysis but provides an advantage to

tumor cells since allows carbohydrate metabolites to enter other subsidiary pathways, including the hexosamine pathway, uridine diphosphate (UDP)-glucose and glycerol synthesis, and PPP[48] (**Figure 1.3**). Its high expression in various solid tumors renders it a potential biomarker of tumor development and invasion.

Amino acid metabolism

Moreover to the well-established role for altered glucose metabolism in tumors, recent investigations highlight the involvement of AA metabolism in cancer, especially Gln. Gln plays important functions in cells due to they can use it for energy production, as a source of reducing power, but also as a source of carbon and nitrogen needed to generate nucleotides and other non-essential AAs. Besides, Gln maintains the pool of intermediate metabolites such as acetyl CoA or α-ketoglutarate (α-KG), both important for the anaplerotic reactions of TCA cycle and is the substrate for the generation of glutathione (GSH). Glutaminase (GLS) is the enzyme that converts Gln to glutamate (Glu), which is used by alanine aminotrasferase (GPT2) to produce α-KG (**Figure 1.3**). GLS has been demonstrated to be regulated by oncogenes and to support tumor cell growth[49].

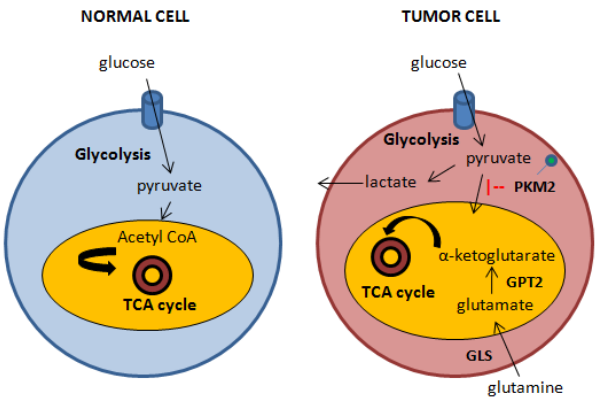


Figure 1.3 Metabolic changes in cancer cells.

Serine (Ser) can be synthesized through a glycolytic intermediate, the 3-phosphoglycerate. Ser has a unique metabolic role in the cell as a major substrate for one-carbon metabolism, which is involved in several cell processes: epigenetics, redox status, genome maintenance, protein translation, and biosynthesis of lipids and nucleotides for cell proliferation[50]. The role of one-carbon as a driver of tumor pathogenesis and tumor maintenance, including genome integrity and epigenetic maintenance, has also been described[50].

On the other hand, many AA, enzymes, and metabolites have been described as immunosuppressive in the tumor microenvironment and have been postulated for cancer therapy[51]. For example, the enzymes related to tryptophan (Trp) and arginine (Arg) metabolism.

Arg is a nonessential AA for healthy humans, but under certain conditions of disease, it becomes essential. Arg is an important precursor for protein, urea and creatine synthesis, as well as for the synthesis of important molecules (e.g. Glu, nitric oxide (NO) and agmatine) related with signaling processes[51] (**Figure 1.4**). Moreover, Arg plays an important role in immune regulation by affecting the immune response and inflammation[52].

Trp is another nonessential AA related to the regulation of immune tolerance and anti-tumor immune responses. Trp is catabolized to kynurenine by the action of two enzymes, indoleamine-2,3-dioxygenase (IDO) and tryptophan-2,3-dioxygenase (TDO)[51] (**Figure 1.4**). Several studies have suggested the potential role of these metabolites and enzymes in cancer and their link with an immunosuppressive microenvironment[53,54].

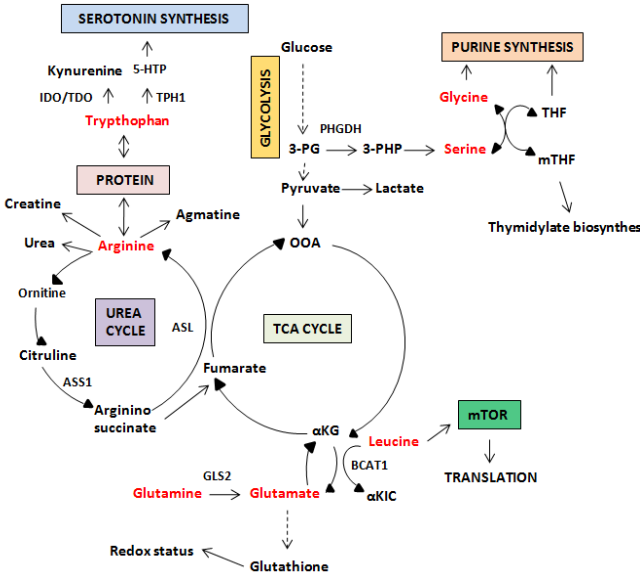


Figure 1.4 Integrated representation of amino acid metabolic pathways in cancer cells.

Note: ASL: Argininosuccinate lyase; ASS1: Argininosuccinate synthetase 1; BCATc: Cytosolic branched chain aminotransferase; α-KIC: alpha-ketoisocaproate; 5-HTP: 5-hydroxytryptophan; mTHF: Methyltetrahydrofolate; OAA: Oxaloacetate; 3-PG: 3-phosphoglycerate; PHGDH: Phosphoglycerate dehydrogenase; 3-PHP: 3-phosphohydroxy pyruvate; THF: Tetrahydrofolate; TPH1: Tryptophan hydroxylase-1.

The role of the branched chain AAs (BCAAs) such as leucine (Leu), isoleucine (Ile) and valine (Val)) has also been studied in cancer. BCAAs constitute ~40% of the essential AAs requirements of healthy individuals and they have an important role for protein synthesis but also as nitrogen donors for alanine (Ala) and Gln production. In addition, the role of Leu has been related to with the regulation of the signaling pathway mTOR, which regulates protein translation, cell growth, proliferation, and autophagy (**Figure 1.4**). T cells are very sensitive to mTOR pathway to develop correctly their functions. Therefore, Leu is presented as an important nutrient signal that is sensed by the immune cells via mTOR pathway and it is critical for their proliferation[51].

Fatty acid metabolism

The fatty acid (FA) metabolism is required for essential cellular processes such as energy storage, membrane proliferation, generation of signaling molecules, redox balance and oxidative stress balance[55,56] (**Figure 1.5**). Then, understanding how cancer cells modulate it could be useful in the development of target therapies but also in biomarker discovery approaches.

Lipids include a diverse group of molecules such as triacylglycerides (TG), PLs, lyso-PLs, sterols, and sphingolipids. FAs are the main building blocks for the synthesis of TG that are used for energy storage. PLs include phosphatidylcholine (PtdCho) and phosphatidylethanolamine (PtdE) and together with sterols, lyso-PLs and sphingolipids, are the major structural components of biological membranes[55,56]. Other several lipids can also have important roles in signaling, functioning as second messengers and as hormones[56](**Figure 1.5**).

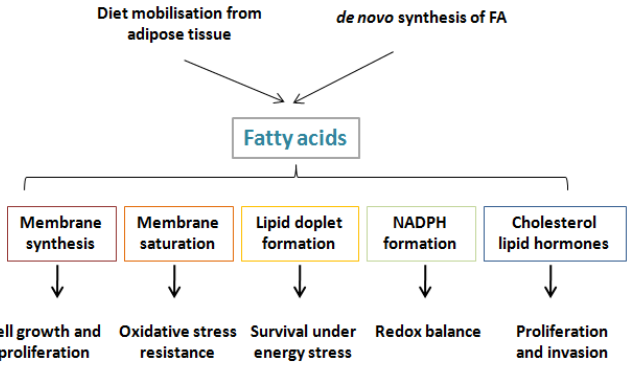


Figure 1.5 Roles of lipids in cancer cells.

Citrate is an intermediate in the TCA cycle that together with the enzyme ATP citrate lyase (ACLY) act as a nexus between glucose and FA metabolism. When citrate is in mitochondria is lead into the TCA cycle, and when it is in

cytoplasm it is converted to acetyl CoA by ACLY. Then, acetyl CoA is used by fatty acid synthase (FASN) and acyl-CoA synthetases (ACS) to produce FA CoA[55] (**Figure 1.6**). FA CoA can be driven to different destinations: can lead to the formation of reserve molecules such as TG; can be lead to diacylglycerol synthesis, which is the precursor of PLs; or can be imported back into the mitochondria to obtain energy in the β-oxidation process. β-oxidation is the metabolic pathway where the FA are degraded producing acetyl CoA, which can be derived towards the TCA cycle to produce reducing equivalents for OXPHOS. The FA CoAs are transported from the cytosol across the outer mitochondrial membrane and afterward, they are converted to FA carnitines by carnitine palmitoyl transferase 1 (CPT1). Currently, it is unclear if high rates of FA degradation block proliferation or contrary provide increasing ATP levels for cellular division[55].

Redox status

Oxidative stress is produced from an imbalance in the production of reactive oxygen species (ROS) and the cell's own antioxidant defenses. Mitochondrial OXPHOS is the major cellular source of ROS production[43]. ROS deregulate the redox homeostasis and promote tumor development through initiating an aberrant induction of signaling networks that cause tumorigenesis. Cells have several defense systems that comprise detoxification enzymes such as glutathione-S-transferases, NADP(H) quinone oxidoreductases, glutathione peroxidases (GPx), catalase, superoxide dismutases (SODs), uridine 5'-diphospho-glucuronosyltransferases (UGTs), among many others[57], but also the tripeptide GSH. GSH is an important cellular antioxidant that cells synthesize from glycine (Gly), cysteine (Cys) and Glu. In cells, GSH appears mainly in its reduced state due to the enzyme glutathione reductase (GSR), which is constitutively active. GSH

reduces the presence of ROS, but also play a role in processes of innate and adaptive immune system[58].

- *Oncometabolites as a cancer biomarkers*

Mutations or deregulated expression of genes encoding enzymes with key functions in cellular metabolism produce the abnormal accumulation of metabolites that can trigger metabolic dysregulation and a potential transformation to malignancy. These metabolites have been termed "oncometabolites". The main oncometabolites that have been postulated as biomarkers of cancer and have been linked with oncogenesis, come from loss-of-mutations in genes encoding TCA cycle enzymes such as fumarate hydratase (FH) and succinate dehydrogenase (SDH). These mutations cause the accumulation of fumarate and succinate, respectively. On the other hand, the gain-of-function isocitrate dehydrogenase (IDH) mutations increase levels of 2-hydroxyglutarate (2-HG)[59,60]. The accumulation of succinate, fumarate, and 2-HG promote cancer progression and position TCA cycle as mitochondrial custodian of the methylome[61]. Recent data suggest that both oncogenic mutations as microenvironmental successes (e.g. hypoxia) could affect the metabolism reprogramming of cells and then, the abundance of oncometabolites[41].

1.4.3 Genetic and epigenetic regulation of cancer

In the last years, there has been a rise in the study of metabolism and gene expression, two bidirectionally connected and tightly coordinated cell processes that allow tumors to adapt to the dynamic and changing environment through the acquisition of pro-survival traits[61]. Concerning genes regulation, several cellular mechanisms can act to determine the expression or inhibition of genes: packing of

DNA and histones in chromatin or epigenetic remodeling, transcription factors (TFs) and alternative splicing (AS)[61,62].

- *DNA and histone methylation and acetylation*

One of the main cellular regulatory processes that determines which genes are activated is the post-translational modification of histones (predominantly acetylation and methylation). These modifications alter the chromatin structure, helping or preventing the recruitment of TF complexes that will ultimately regulate gene expression.

Cancers usually display global DNA hypomethylation but hypermethylation of CpG islands in genomic regions where tumor suppressor genes are located[61]. These histone methyl marks can either activate or repress gene expression, although hypermethylation is more related to repression. Methylation is linked to the intermediate metabolism by means of S-adenosyl methionine synthetase (SAM), an essential substrate involved in the transfer of methyl groups generated in the folate and methionine (Met) cycles. The intracellular levels of SAM depend on Ser and Met availability and control the activity of histone methyltransferases (HMT) and DNA methyltransferases (DNMT) (**Figure 1.6**). Consequently, the deprivation of these essential AAs induces reversible and rapid changes in histone and DNA methylation, affecting the transcriptional landscape of cancer cells. Is for this that Met cycle and one-carbon metabolism gene networks has been postulated as the major determinants of DNA methylation status in human cancer[61]. Indeed, dysregulation of histone methylation in concrete chromatin regions is a principal selective force for tumor progression and metastatic potential[61].

On the other hand, the enzymatic removal of methyl groups by histone demethylases (HDM) should also be considered. HDMs are modulated by the oncometabolites α-KG, fumarate, and succinate[61]. The levels of these oncometabolites depend on the TCA cycle enzymes, so mutations in FH, SDH or IDH would produce the accumulation of them. α-KG is a positive co-factor of the HDM activity, unlike fumarate and succinate that are competitive inhibitors of several HDM[61] (**Figure 1.6**). The contribution of fumarate to tumorigenesis has been shown in the context of FH loss. Fumarate accumulation triggers epigenetic changes in a regulatory region of the antimetastatic miRNA cluster mir-200ba429. The repression of miR-200 family leads to the expression of epithelial-mesenchymal transition (EMT)-related TFs and the enhancement of migratory and metastatic processes[61]. The deficient activity of SDH is also associated with global DNA methylation changes and a migratory phenotype, and mutations in IDH genes have been linked with oncogenic properties because they favor the formation of 2-HG[61]. 2-HG accumulation inhibits DNA demethylation and primes cancer cells for transformation. 2-HG has been presented as an immunometabolite that links the environmental with the immune system through a metabolic–epigenetic network[61]. Finally, acetyl CoA levels affect histone acetyltransferases (HAT) and therefore, to global levels of nuclear histone acetylation (**Figure 1.6**)[61].

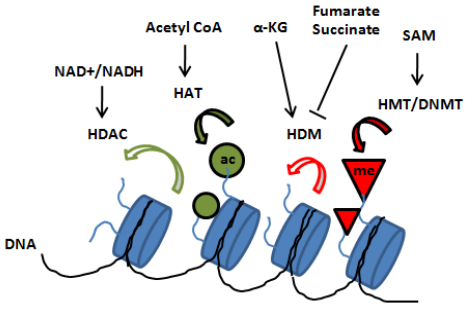


Figure 1.6 Crosstalk beween metabolism and chromatin.

In brief, chromatin-associated enzymes detect levels of intermediary metabolites and create dynamic chromatin modifications that ultimately regulate adaptive transcriptional programs associated with oncogenic pathways.

- *Transcriptional regulation of metabolic pathways*

The changes in gene expression can also be derived from TFs activity. TFs are proteins that control the rate of transcription of genes to mRNA through binding to a specific DNA sequence (promoter or gene's enhancer response element). TFs are downstream of signal transduction pathways and after signal reception, they up- or down-regulate the expression of specific genes in order to make sure that they are expressed in the right cell at the right time and in the right amount according to cell needs[63] (**Figure 1.7**).

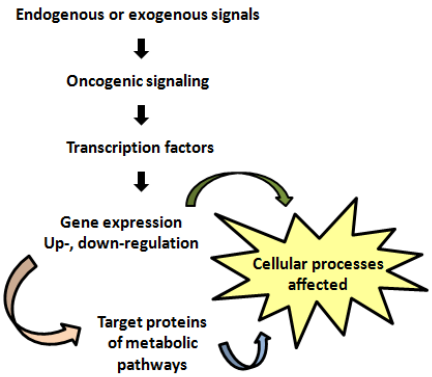


Figure 1.7 Scheme of the mechanism of transduction of signals.

In the context of cancer, several TFs have been related to oncogenic signaling pathways. They include: Tp53, c-Myc, TWIST, E2F, GATA, PPARs, SREBP, Nrf2, FOS, JUN, GABP, CTCF, HIF-1, among many others[61,64–68]. However, the role of PPARs,

HIF-1, Tp53, and c-Myc has been linked with metabolic processes such as: lipid sensing, glucose homeostasis and glucose metabolism (glycolysis)[61,69,70].

Peroxisome-proliferator-activated receptors (PPARα, PPARβ, PPARγ) are nuclear receptors that respond to specific ligands by altering gene expression[71]. PPARs control the intracellular FA levels, acting as biological sensors of altered lipid metabolism but also regulate gene expression programs that impact on proliferation, differentiation, and survival in several tumor tissues[61].

The effect of hypoxia-inducible factor-1 (HIF-1), MYC, and p53 in cancer cells is depicted in **Table 1.1**:

Table 1.1 TFs related to anaerobic glycolysis in cancer cells.

Molecule	Function	Activity in cancer	Effect
HIF-1	Hypoxia-inducible; TF	Adaptative or constitutive expression; gain of function.	Up-expression of glycolytic genes, vascular endothelial growth factor (VEGF) and other proteins involved in hypoxic adaptation[69]
MYC	Oncogene; TF	Constitutive expression; gain of function	Up-expression of lactate dehydrogenase, increases glycolysis and lactate production[69]
p53	Tumor suppressor; TF	Mutated; loss of function	Represses lactate (MCT) and glucose (GLUT) transporters and enhances aerobic glycolysis and OXPHOS[72]

- *The role of RNA alternatively splicing in gene regulation*

Approximately 95% of the human genes coding for proteins undergo AS, a regulated process of gene expression that greatly diversifies the proteome by creating multiple proteins from a single gene through the use of alternative promoters or the inclusion or exclusion of different exons or parts of exons in

mRNA[62]. This process is essential for normal biological processes but an unbalanced of AS can drive to tumorigenesis by means of increasing the production pro-carcinogenic isoforms and the reduction of anti-carcinogenic isoforms. Aberrations in AS are produced by mutations or deletions of regulatory sequences or by dysregulated expression of RNA-binding proteins participating in splicing[62].

In the context of cancer cells metabolism, several metabolic enzymes present AS as a mechanism of regulation of metabolic pathways but few have been well studied (e.g. PKM2, GLS, and oxoglutarate dehydrogenase (OGDH)[47,62,73].

1.5 Metabolomics

Metabolomics evaluates, identifies and quantifies endogenous and exogenous metabolites that are present in a biological sample. Fiehn defined this term as a "comprehensive and quantitative analysis of all metabolites in a system"[33]. Another term liked with metabolomics is metabonomics. Nicholson et al. defined metabonomics as "quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification"[33]. Nevertheless, both terms are commonly used interchangeably.

The metabolome includes the complete set of small molecular weight organic molecules (<1500 Da) within a biological system. The metabolome is highly dynamic and reflects the perturbations of genome, transcriptome, and proteome (**Figure 1.8**), but it is also influenced by environment and "in-vironment" (i.e. gut microbiota[74]). Therefore, metabolome is considered the closest representation of the phenotype[75]. The nearness of the metabolism to an organism's phenotypes

indicates that it will be affected by a pathology and therefore, its metabolites or metabolic profiles can be used as biomarkers of a disease[76].

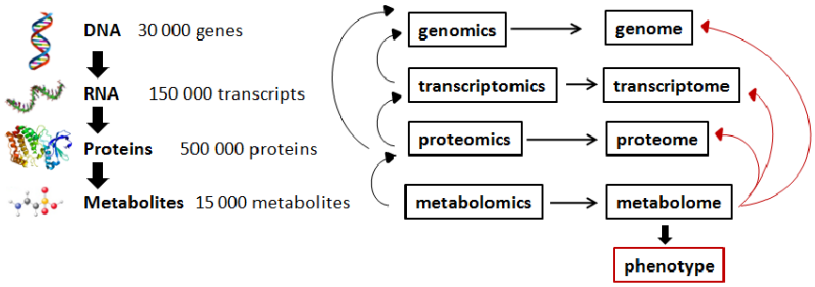


Figure 1.8 The "omics" cascade

Nowadays, there are several public metabolomics databases available. One of the largest organism-specific database is the Human Metabolome Database (HMDB)[77]. HMDB contains detailed information about metabolites in several samples collected from the human body (urine, tissue, blood...). Specifically, it contains more than 41000 metabolites that include both water-soluble and lipid-soluble molecules as well as abundant and relatively rare metabolites[77].

1.5.1 Basic structure of a metabolomic study

Currently, three approaches are used for metabolomic studies: untargeted, targeted and semi-targeted[78–80]. On the one hand, targeted and semi-targeted metabolomics studies focus on accurate identification and quantitation of a previously defined set of metabolites in biological samples, so they may not be considered as a true '*omic*' approaches[79]. On the other hand, untargeted metabolomics aims to characterize the maximum number of metabolites in a given biological sample. It aims at the unbiased profiling of the metabolome to

link specific patterns of metabolites to a given treatment or condition (e.g. healthy *vs* diseased)[81]. This approach is usually linked with biomarker discovery.

A basic workflow of a untargeted metabolomic study is depicted in **Figure 1.9**.

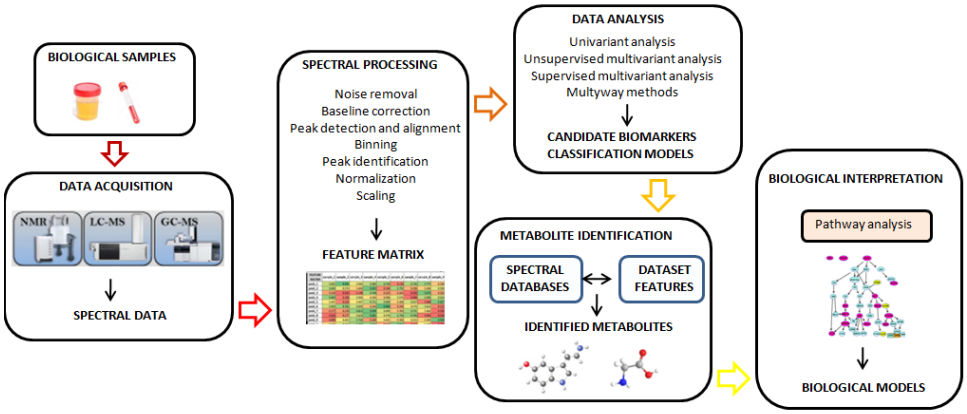


Figure 1.9 Analysis workflow in untargeted metabolomic studies.

In an untargeted metabolomics study, the first step is to establish the hypothesis of work and define the objectives. According to this, the study is designed, and the samples of interest, the sample size and additional experimental conditions including: the analytical platform for the analysis of samples (e.g. NMR, HRMAS NMR, MS), the pre-processing of samples, data acquisition parameters, and quality analysis and quality control (QC) criteria.

After data acquisition, spectral processing and data analysis, metabolite identification has to be performed. This step is indispensable to give a biological meaning to the associated features in a metabolomic study[82]. The use of reference spectral databases is usual. Once identified the principal metabolites of the study, a pathway analysis can be interesting to know the relationship between metabolites and metabolic pathways (**Figure 1.9**). Biological databases such as

Kyoto Encyclopedia of Genes and Genomes (KEGG)[83], small molecule pathway database (SMPDB)[84], WikiPathways[85], and MetaCyc[86] offer exhaustive information of a large number of metabolic pathways. The availability of this data is essential in metabolomics studies to understand cell metabolism. These methods are referred to as metabolite set enrichment analysis (MSEA) and are methodologically based on the gene set enrichment analysis (GSEA) approach, designed for pathway analysis of gene-expression data[82].

1.5.2 Techniques for the study of metabolome

In order to acquire reliable and valid metabolomic data, it is critical the use of comprehensive techniques to provide a high metabolomic coverage in biological samples, in a reproducible way and preferably at a low cost. The comprehensive analysis of the complete set of metabolites in biological samples is a technological challenge due to the wide range in physicochemical properties and concentration ranges of the metabolites. Current analytical platforms, including gas chromatography (GC), liquid chromatography (LC), capillary electrophoresis coupled to MS, and NMR spectroscopy present advantages and limitations, but none of them enables a complete qualitative or quantitative detection of the whole metabolome. Therefore, the combined use of different analytical approaches is required to increase the coverage of obtained metabolomic profiles that otherwise cannot be achieved by single-analysis techniques[87]. Although a comprehensive comparison of the main characteristics of each technique is out of the scope of the Ph.D. Thesis, it is worth to review their main advantages and disadvantages.

1.5.2.1 *Nuclear Magnetic Resonance*

NMR spectroscopy was developed in the late 1940s to study atomic nuclei and is described as "…the study of molecular structure through measurement of the interaction of an oscillating radiofrequency electromagnetic field with a collection of nuclei immersed in a strong external magnetic field"[87]. Its use in metabolomics is based on the study of nuclei magnetically active such as ^{1}H, ^{13}C, ^{19}F, and ^{31}P. Usually, in NMR metabolomics studies, the nucleus most used is the proton (^{1}H), since it is present in 99.9% of biomolecules.

In presence of an external magnetic field, these atomic nuclei undergo a splitting of energy levels: low (α) and high (β). The application of a radiofrequency pulse, perpendicular to the direction of magnetic field, causes energetic transitions. The energy of radiation necessary to produce this shift depends on the type of nucleus, its chemical environment, the type of nuclei present on its vicinity and the strength of the applied magnetic field. When the radiofrequency pulse ends, the excited nuclei release the excess energy returning to the state of equilibrium (relaxation). In this process, the magnetization disappears giving rise to what is known as Free Induction Decay (FID), which provides information about the irradiated sample. The FID is a time-dependent function that can be transformed into a signal spectrum dependent on the frequency by using the Fourier-Transform (**Figure 1.10**)

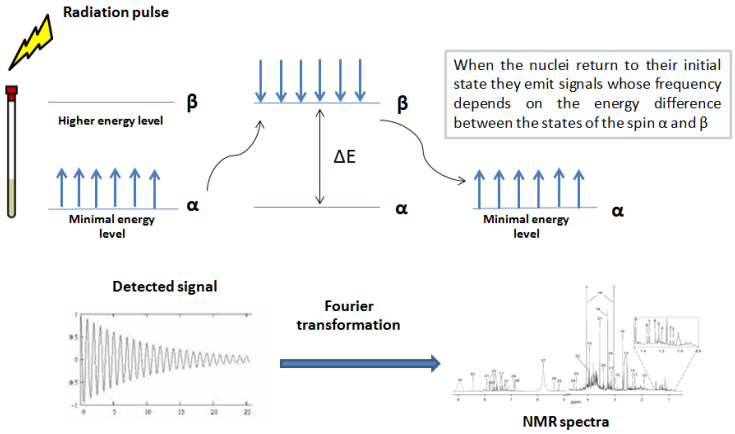


Figure 1.10 Scheme of NMR spectra adcquisition

NMR spectroscopy provides both, quantitative and qualitative information. On the one hand, the area under each peak can be related to the concentration of a given metabolite in the sample. On the other hand, metabolite annotation may also be carried out based on chemical shifts observed in an NMR spectrum. The position of each peak in the NMR spectrum is defined by the chemical shift. It is expressed in parts per million (ppm) relative to an added reference compound, usually 3-trimethylsilyl propionate (TSP) or 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) or relative to internal metabolites (e.g. Ala, creatine), whose signals appear well defined in all samples and their chemical shift is insensitive to their specific characteristics.

NMR-based experiments

The spectral data obtained with NMR techniques can be referenced to one or two frequency axes. To date, one-dimensional (1D) ^{1}H NMR spectra is the most commonly used method in high-throughput metabolomics studies, especially in disease diagnosis, prognosis or monitoring. However, other nuclei that take part of important biomolecules such as ^{31}P or ^{15}N have been considered

for the study of PLs or the study of proteins, RNA or DNA, repectively[87]. The most used pulse sequences in ^{1}H NMR experiment are 1D NOESY and CPMG.

1D NOESY sequence contains a small mixing period that improves the suppression of the water signal, the phase, and the baseline. This experiment provides a simple, highly reproducible and robust method for the acquisition of high-quality NMR spectra in aqueous solutions. This technique is undoubtedly the most widely used in terms of water suppression[88] and is applied in a sequence of pulses known as noesy-presat[89], where the water signal is saturated through selective irradiation.

CPMG (Carr Purcell Meiboom and Gill) sequence is based on the relaxation proprieties of the nuclei and is the most used in the analysis of plasma and serum samples since these samples have a high level of proteins. This sequence allows the elimination of the wide signals in the spectrum coming from these biomolecules, thus improving the resolution of low molecular weight signals.

^{1}H NMR provides the most sensitive signal but, due to its small chemical shift range, an overlapping of peaks results from a mixture of biochemical species (e.g. proteins, lipids, and low-molecular-weight metabolites) making the spectral assignment a challenging task. To solve this problem two-dimensional (2D) NMR spectroscopy is applied. The second dimension allows to separate otherwise overlapping spectral peaks and, therefore, offers additional information on the chemical properties of the metabolite. Globally, 2D NMR spectra can be classified into homonuclear (i.e., ^{1}H–^{1}H-NMR) and heteronuclear (i.e., ^{1}H–^{13}C or ^{1}H–^{15}N)[82]. Most commonly used 2D NMR experiments include the homonuclear through-bond correlation methods correlation spectroscopy (COSY), the total correlation spectroscopy (TOCSY) and nuclear overhauser effect spectroscopy (NOESY); and

the heteronuclear single-quantum correlation spectroscopy (HSQC)[74]. The choice of the type of experiment will depend on the requirements of the study.

Finally, NMR-based metabolomics analysis in semisolid biospecimens such as tissue biopsies, cells, and organisms has been possible by the use of HRMAS NMR spectroscopy[90–92]. The molecules within the tissue have restricted mobility and their nuclei are subject to static anisotropic NMR interactions, resulting in broad lines of the NMR spectra. The use of rapid spinning of the sample (typically 5 kHz) around an axis inclined 54.7° (the magic angle) to the direction of the static magnetic field reduce line broadening. Consequently, HRMAS mimics a liquid solution state, yielding spectral resolutions close to that of extracts[74]. HRMAS NMR provides an efficient way to identify the metabolites present in tissue without pre-preparation steps such as extraction. Specifically, HRMAS NMR spectroscopy is the only technique available for the analysis of intact tissue[93]. Besides this, HRMAS NMR does not destroy the samples, so they can be used in posterior studies (e.g. MS-based metabolomics, transcriptomics, and proteomics).

NMR-based studies: advantages and drawbacks

Compared with other analytical platforms NMR spectroscopy may be considered as a fast and reproducible method in metabolomics research[93] since through flow-injection NMR probes is possible to screen hundreds of samples in a day. Moreover, this technique has an unequaled cross-laboratory reproducibility[94,95]; allows the analysis of samples collected from non-invasive way (e.g. urine, blood, and seminal fluid)[87,95]; requires small sample amounts and minimal or no sample preparation[95]; and it is no destructive[95], which make possible further analyses with the same samples (**Table 1.2**). This is also true for HRMAS NMR spectroscopy[74]. Additionally, the non-invasive nature of NMR

spectroscopy makes it an ideal technique to be used for *in vivo* studies, referred to as magnetic resonance spectroscopy or magnetic resonance imaging.

Its low sensitivity is the main disadvantage of NMR spectroscopy since only metabolites in mM range concentrations will be detected[87,93]. Nonetheless, if NMR can identify a biomarker by the analysis of a relatively small number metabolites, this may be considered an advantage, due to profiling thousands of metabolites would be a complicated task.

1.5.2.2 *Mass Spectrometry*

Ultraperformance Liquid Chromatography–Mass Spectrometry (UPLC-MS) is one of the most powerful and widely used technique for untargeted metabolomic analysis in biomedical research. Recent developments in chromatographic instrumentation such as core-shell particles, narrow bore columns packed with sub-2 µm particles and LC instruments operating at pressures up to ~1400 bar have dramatically enhanced the chromatographic efficiency enabling higher separation performances in less analysis time[96]. Besides, the availability of high-resolution MS instrumentation such as time of flight (TOF) and the development of improved interfaced for the online hyphenation of chromatographic separations and MS ionization has extended its application to almost every area of biomedical research[97].

The main advantages of MS-based metabolomics include its high sensitivity, several orders of magnitude lower than NMR; its high resolution and selectivity; and the possibility of performing fragmentation analysis thus confirming the identity of the detected metabolites and the identification of unknown and unexpected compounds[98]. All the strengths and weakness of this technique respect to NMR are detailed in the following table:

Table 1.2 Strengths and drawbacks of NMR and MS

	Strengths	Weakness
MS	-High sensitivity (detection limit reach nano-picomolar) - Can be used for selective and nonselective analysis (targeted, untargeted analysis) -Wide detection range -Easy metabolite identification-databases availability -Possibility to couple with separation techniques - Need a small amount of sample	-Destructive -More demanding sample preparation: optimization of ionization conditions -Tissue analysis require extraction
NMR	-Rapid -Non-destructible -High-throughput -Minimal sample manipulation -Using HRMAS NMR tissue samples are analyzed directly -Very high reproducibility -All metabolites at NMR concentration level can be detected in one measurement - Allows *in vivo* studies	-Limited sensitivity -Low specificity -High sample volume requirements

In this context, UPLC-MS was selected in this Thesis as a complementary analytical technique to NMR spectroscopy, using electrospray (ESI) for the ionization of analytes prior to MS detection.

In ESI a high voltage is applied to a liquid sample to assist the transfer of ions from solution into the gaseous phase before MS analysis. This process is carried out through three steps that include: dispersal of a fine spray of charged droplets, followed by solvent evaporation and ion ejection from the highly charged droplets. These ions are passing down towards the mass analyzer, in our case a TOF, where they are accelerated by an electric field and separated and subsequently detected according to their m/z ratio[99].

Metabolite annotation in UPLC-MS can be based on accurate mass (i.e. m/z), retention times (RT) and MS/MS fragmentation patterns (see **Figure 1.11**). Depending on the type of information employed for the annotation, several confidence levels can be defined. Level 0 includes stereochemistry discrimination; Level 1 requires the analysis of a standard in the same instrumental conditions and at least two orthogonal techniques such as m/z and RT; Level 2 involved the identification supported by the analysis of a class-specific standard; Level 3 is based on a single parameter, typically m/z and finally, Level 4 is feature level without annotation (i.e. m/z-RT)[100].

So, tandem MS (MS/MS) analysis improves the identification of compounds and facilitates the determination of the structure of unknown compounds based on MS/MS fragments originating from the same molecule. However, MS/MS analysis in untargeted metabolomics also presents several limitations. Each molecule presents optimal fragmentation energy which is unknown in advance. Small molecules may generate a few characteristic fragments hindering the interpretation of MS/MS spectra. Also, a high percentage of molecules in metabolomics UPLC-MS data sets are of low abundance, reducing the quality of acquired MS/MS spectra. On top of that, the MS/MS analysis reduces the sensitivity of UPLC-MS and because of that, the analysis is normally repeated a number of times. This strategy increases the amount and quality of MS/MS data but also the sample requirements.

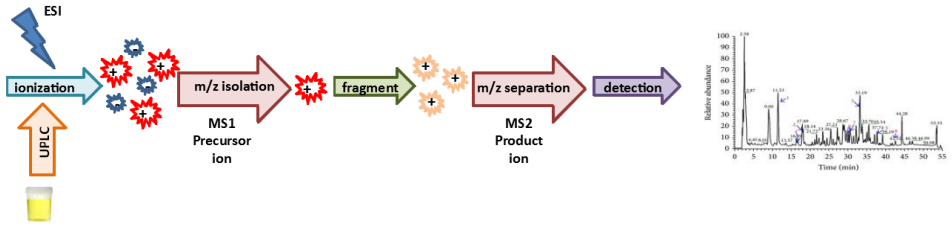


Figure 1.11 Scheme of UPLC-MS/MS analysis

Quality control procedures in UPLC-MS analysis

UPLC-MS generates information-rich data sets. However, data must be appropriately handled to generate consistent, meaningful and reproducible results. A crucial aspect to increase the reproducibility of metabolomic studies and to report the quality requirements after data acquisition is the implementation of QC procedures during design and analysis, typically supported by the injection of blanks and pooled QC samples.

The analysis of biological extracts (e.g. diluted urines) by UPLC-MS leads to a cumulative build-up of matrix components and metabolites in the analytical system that causes non-linear drifts in chromatographic and MS performances affecting the RT, signal intensity, and mass accuracy. To overcome this potential pitfall, UPLC-MS metabolomic experiments are normally designed including an initial system pre-conditioning by repeated injections of a pooled QC samples to reach a system steady state. Therefore, the use of QC samples allows conditioning the analytical platform, performing intra-study reproducibility measurements (QC) and to correct mathematically systematic errors[80]. The data pre-processing pipeline also must include the elimination of background and contaminant features in the data set arising from non-biological sources (e.g. solvents impurities and stabilizers, plasticizers, reagents used during sample clean-up, cumulative carryover contamination)[80].

1.5.3 Data analysis in metabolomics

After careful data pre-treatment, the resulting UPLC-MS or NMR peak tables are high-dimensional data matrices with relatively few observations (samples) compared to variables (peaks). Multivariate analytical tools reduce the complexity or dimensionality of the data sets into a more manageable number of dimensions,

considering all the metabolomic features simultaneously and thereby identifying relationship patterns between them are required[101]. These pattern-recognition methods may be classified into two groups: unsupervised and supervised methods[82]. Unsupervised techniques do not require foreknowledge of the classes of the samples included in the data. Supervised techniques aim at getting insight into underlying biochemical differences associated with previous information about the considered samples (e.g. collected from healthy or disease volunteers).

- *Unsupervised data analysis*

Principal component analysis (PCA) is the most commonly used unsupervised analysis for detecting trends or patterns in the data, such as subgroups or outliers[74,102]. PCA is based on the linear transformation of the features into a set of linearly uncorrelated variables that are called Principal Components (PCs). This transformation maximizes the variance explained by the first component while the subsequent components explain increasingly reduced amounts of variance[82]. The results of a PCA are shown graphically in terms of 'scores' and 'loadings'. The scores plot contains the information about the samples of the original matrix, where each point corresponds to an observation and summarizes the relationship between them; while the loading plot contains the information about the variables. In this case, each point represents one point of the spectrum and its distribution explains the relationship between them. The loading plot allows to interpret the spatial distribution of the samples in the score plot and identify the variables responsible for the observed patterns. The information from both plots can be jointly analyzed[103].

- *Supervised data analysis*

Unlike unsupervised methods, supervised strategies exploit previous information to identify metabolomic profiles associated with a phenotype of interest[74,82]. Several methods have been applied in metabolomics to achieve this purpose including univariate (e.g. t-test, one-way ANOVA, Wilcoxon rank sum test), fold changes or on multivariate strategies such as linear discriminant analysis, principal component discriminant analysis, support vector machines, random forests, artificial neural networks or partial least squares discriminant analysis (PLS-DA). Among them, PLS-DA is arguably the most popular[102].

PLS-DA is able to find the fundamental relations between two matrices: an X matrix containing independent variables (e.g. spectral intensity values) from the samples and a Y matrix containing dependent variables (i.e. class, gender, stage of the tumor, grade). Therefore, PLS-DA is used to identify profiles able to discriminate two groups of samples (e.g healthy vs diseased) with predictive capacity[104].

Unlike PCA, PLS components do not maximize the explained dataset variance but the covariance between the variable of interest and the metabolomic data. Hence, the feature coefficients (loadings) of PLS components represent a measure of the relative contribution of a given feature to the observed discrimination between the sample groups[82].

When a classifier is developed, its performance should be assessed to provide an estimate of its prediction accuracy beyond the training data set and ensure the lack of model overfitting. The ideal design would include the split of initial data into three independent subsets that were representative of the whole population and considered the different sources of variation. These sets would include a calibration set to adjust the model, a validation set to select the optimal model

parameters such as the number of latent variables (LVs) in PLS-DA, and a validation set to estimate its predictive accuracy. Nevertheless, when the sample size is small another feasible strategy should be applied. In these cases, internal validation using resampling methods such as cross-validation (CV) is often applied[105].

CV approaches estimate the predictive performance of a classifier using an iterative approach. At each round of CV, samples are split into exclusive subsets. Then, a single subset (validation set) is used to evaluate the model performance, which has been created using the remaining sets of samples. This procedure is repeated several times so that all the samples have been used once as a validation group. Averaging these results, an unbiased estimate of the performance of the predictor is obtained[82]. Then, a permutation test is carried out to assess the significance of CV.

In a permutation test, the class labels are randomly assigned to samples and a new classifier is trained and CV using the relabeled data set. This process is repeated hundreds or thousands of times and the accuracy estimates got using random class labels is contrasted to the accuracy estimate got using true class labels. After this, a *p*-value is calculated considering the fraction of values where the classifier showed better classification performance in the random than in the initial set. Considering that the null distribution assumpts that no differences exist between sample groups, if the p-value is lower than a user-selected threshold (e.g. 0.05) the null hypothesis would be dismissed and it would be assumed that the classifier is significant[105].

1.6 Transcriptomics

Transcriptomics is the study of the whole transcriptome, which includes the complete set of RNA transcripts that are produced by the genome under specific circumstances or in a specific cell. Transcriptomics uses high-throughput methods to identify genes that are differentially expressed, but also informs about other characteristics linked with alterations or modifications in DNA sequence. At the same way that metabolome, the information provided by RNA transcripts is highly dynamic and reflects the cellular phenotype associated with molecular alterations in genes (e.g. cancer).

Cancer cells are characterized by altered protein function and aberrant transcriptional patterns, which are the consequence of somatic mutations and epigenetic alterations. Therefore, the transcriptomic analysis allows analyzing the heterogeneity of tumors to discover new biomarkers or therapeutic strategies[106]. Although the capacity to describe the molecular alterations of cancer was historically one of the primary applications of transcriptomics, the development of techniques for whole-transcriptome sequencing has provided data about chemical modifications, interactions, somatic mutations including single nucleotide variants, gene fusions and amplification of the transcripts studied. Hence, the study of alternative splicing, RNA editing, post-transcriptional modifications, and various non-coding RNAs (ncRNAs) is currently a crucial aspect of transcriptomics[106].

1.6.1 Techniques for the study of transcriptome

Continued advances in transcript profiling over the last four decades have increased our understanding of the genetics and molecular biology of cancer. The development of new experimental methodologies, together with advances in bioinformatics, have led to important discoveries in cancer. Behind that progress was the constant push for an increase in the breadth, depth or fidelity of measuring cellular RNA[106].

The study of the transcriptome was initiated in 1977 with the development of the northern blot and the first sequences of cloned complementary DNAs (cDNAs). The later development of reverse transcription polymerase chain reaction (RT-PCR) made possible to identify cellular mRNAs in an unbiased way. However, these low-performance techniques only allowed the qualitative or semiquantitative analysis of a few candidate genes for each analysis. Subsequently, the development of the quantitative technique RT-qPCR improved the methodology in many aspects and increased the performance, since several transcripts could be analyzed at the same time, although this was still far from a large-scale coverage of the whole transcriptome. With the appearance of the microarrays, the characterization of expression levels of thousands of transcripts could be achieved simultaneously and a cost-effective way, and more recently the development of NGS techniques have revolutionized the world of research due to their applications in molecular diagnostics. NGS makes possible the sequencing of DNA but also the sequencing of all RNA molecules present in a sample (i.e. complete transcriptome). This technology is called RNA-Sequencing (RNA-Seq)[106]. Recently, NanoString, another technique that allows the measurement of gene expression with high levels of precision and sensitivity, is being applied in clinical practice for disease diagnosis. Although this technology is still in the early stages

of clinical use, it is expected that in the future it may have important functions in the context of cancer, both in patients' stratification as in predicting the response to therapy[107].

1.6.1.1 RT-qPCR

When the starting material is RNA, either total RNA or mRNA, the RT-qPCR transcribes this into cDNA by reverse transcription. Reverse transcriptase is the enzyme that makes DNA from RNA. The cDNA is then used as the template for the PCR reaction. Some reverse transcriptases have RNase activity to degrade the RNA strand in the RNA-DNA hybrid during the first cycles of PCR. If an enzyme does not possess RNase activity, an RNaseH may be added for better qPCR efficiency[108].

The quality control of RT-qPCR process is performed at first time by primer design. The optimum primers are those that reduce the risk of false positives from the amplification of any contaminating genomic. Hence, internal control of genomic DNA or PCR product from a previous run can be used as negative control to observe if there is amplification.

Applications of RT-qPCR

RT-qPCR is used in several applications including gene expression analysis, microarray validation, pathogen detection, genetic testing, biomarker analysis, and disease research[109,110].

1.6.1.2 Microarrays

Microarrays have been applied most intensively to the field of cancer research. The basic premise behind a microarray is that thousands of fragments of DNA (the probes) representing various genes (or their fragments such as exons or exon junctions) are attached to a small surface of an inert material (often referred

to as a "chip"), which provides a genome-wide view of gene expression[111]. Since the location of the gene probes is predetermined, the relative or absolute amount of RNA for each of the genes on the chip can be calculated[112]. The workflow in microarrays experiments is depicted in **Figure 1.12**.

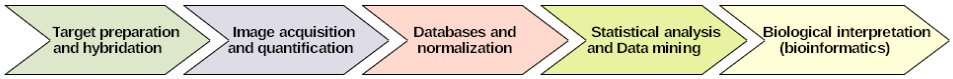


Figure 1.12 Workflow in microarrays experiments

Previous to perform microarray analysis, the quality and quantity RNA is assessed by the use of microcapillary-based devices such as the Agilent Bioanalyzer (Agilent Technologies)[113]. After, the RNA is converted into a more stable cDNA form by a reverse transcriptase. cDNA is labeled by fluorescent dyes, and by complementarity, it hybridizes to the DNA probes attached to the array. After removal of non-selectively bound fluorescent material, the microarray is scanned by passing a laser beam. Microarray image is then analyzed to identify the spots, calculate their associated signal intensities, and assess local background noise. The ratio of the fluorescence intensity for each spot represents the relative abundance of the corresponding DNA sequence[113].

Most image acquisition software packages also have basic filtering tools to mark extremely low-intensity or damaged spots. The image acquisition provides TIFF image pairing and a quantified but not normalized data file. Therefore, data have to be normalized in order to remove some sources of variation (e.g. differences in labeling efficiencies, amounts of starting RNA materials, fluctuating hybridization conditions), which affect the measured gene expression levels[114]. After data normalization, statistical analyses are carried out to find genes differentially expressed. Among others, Limma package has proven a popular choice for the

analysis of gene expression data from microarray experiments[115]. Limma provides the ability to analyze comparisons between many RNA targets simultaneously, and it has especially been designed for analyzing complex experiments with a variety of experimental conditions and predictors. Recently, the capabilities of limma have expanded significantly because the package allows to carry out both differential expression and differential splicing analyses of RNA-seq data. These capabilities allow users to analyze both RNA-seq and microarray data with very similar pipelines. On the other hand, limma is now able to go past the traditional gene-wise expression analyses in a variety of ways, analyzing expression profiles in terms of co-regulated sets of genes or in terms of higher-order expression signatures. This provides enhanced possibilities for biological interpretation of gene expression differences[116].

Finally, bioinformatic processes are crucial to facilitate the manipulation and interpretation of biological data. Open access databases as *The Cancer Genome Atlas* (TCGA), which include the characterization over 20000 primary cancer and matched normal samples spanning 33 cancer types can be used to compare the obtained results with the available information about a specific tumor[117]; and cloud-based applications such as MeV4 (Multiple Experiment Viewer) are used for the analysis, visualization, and stratification of large genomic data, particularly for RNA-Seq and microarray data[118]. Besides this, other web tools allow performing enrichment analysis such as GSEA, which determines if a set of genes is enriched between two biological groups of samples (e.g., tumor vs control)[119]; or *Chip Enrichment Analysis* (ChEA) a web-based interactive application where the users can input lists of gene symbols for which the program computes over-representation of TF targets from the chip database[120]. Finally, *Gene Ontology Biological Processes* (GOBP) allows to relate deregulated genes to biological processes[121].

Applications of microarrays in cancer

Microarrays have been used in the field of cancer for the identification of SNPs, ncRNAs, miRNAs, differential splicing isoforms, mutations; identification of cancer biomarkers, such as gene signatures; identification of genes associated with chemoresistance, and in processes of drug discovery[111,112].

1.6.1.3 RNA-sequencing

RNA-seq makes possible to completely quantify the total RNA molecules with high resolution and reproducibility, including coding and ncRNAs, miRNAs or long-ncRNAs. RNA-Seq has overcome some drawbacks of previously used technologies. Among its strengths spotlight: the ability to detect new transcripts, wide dynamic range, high specificity and sensitivity and simple detection of rare and low-abundance transcripts[122–127]. At the same way that microarrays, processes of quality control, alignment, quantification, and differential expression have to be performed after sample analysis.

The advantages and disadvantages of RNA-seq *versus* microarrays are summarized in **Table 1.3.**

Table 1.3 Comparative advantages and disadvantages of microarrays versus RNA-seq

	Advantages	Disadvantages
Microarray	-Well-defined protocols for hybridization -Well-defined analysis pipelines -Standardized approaches for data submission -Relatively low cost -Allows the study of samples with poor RNA quality (e.g. formalin-fixed paraffin-embedded samples)	-Analysis only for pre-defined sequences -Dynamic range limited by scanner -Relies on hybridization -Hybridization potentially non-specific -Might not give paralogue information -High variance for low expressed genes -Requires technical replication to increase statistical robustness
RNA-seq	-Not reliant on previous sequence information -High dynamic range -Direct sequence alignment, no hybridization -AS detected if aligned to the genome -Paralogous genes can be defined -Can be used for SNPs identification -Has lower technical variation and higher levels of reproducibility.	-Protocols still not fully optimized -High cost -Requires high power computer facilities -High step-up costs if carried out in-house -Complex analysis of splice variants -Requires high RNA quality to perform the analysis

Applications of RNA-Seq in cancer

RNA-Seq has been used extensively in cancer research. It has emerged as a powerful transcriptomics technique for gene expression profiling, to study of AS events associated with cancer, identification of allele-specific expression, disease-associated SNPs and gene fusions contributing to our understanding about disease causal variants in cancer[128–131].

Chapter 2: Rationale

BCs are the most common tumors of urinary and include two groups, MIBC and MIBC, which have different management and clinical outcomes. NMIBCs represent the 75-85% of diagnosed cases and due to its recurrent nature, lifelong surveillance is recommended after TUR of the primary tumor, leading to frequent cystoscopies for an early recurrence detection to avoid progression to invasive disease. Although urinary cytology and cystoscopy are the goal standards techniques for diagnosis and monitoring of BC, both have limitations. Cystoscopy is expensive, operator-dependent, invasive and overlooks CIS. Urine cytology is considered the most important non-invasive method in the detection of BC and has demonstrated utility adjunct to cystoscopy when HG tumors or CIS are present. Nevertheless, cytological interpretation is user-dependent and a normal cytology outcome does not exclude the existence of a tumor, specifically in LG tumors where the sensitivity decreases to 16%. Consequently, to reduce morbidity and costs associated with cystoscopy, improved monitoring is necessary, ideally via non-invasive urinary analysis.

To date, several urinary biomarkers have been approved by the FDA for diagnosis or BC surveillance with sensitivities and specificities in the 50-80% range but poor accuracy for low stage and LG tumors, so none of them have been implemented in the clinical routine. Therefore, the non-invasive detection of BC through urine analysis remains a challenge.

Metabolism reprogramming represents a hallmark of cancer. Tumors reprogram pathways of nutrient acquisition and metabolism to supply their bioenergetic, biosynthetic and redox demands. Then, metabolic profiles or single metabolites can be used as biomarkers. Previous studies performed by MS and NMR have identified metabolic profiles associated with colon, stomach, lung, prostate but also BC, opening the possibility to identify novel non-invasive urinary biomarkers. In the context of BC, prior studies have shown limitations in the use

of a small sample size, lack of external validation, a limited reproducibility among the compounds identified as potential metabolic biomarkers and inadequate experimental design to identify biomarkers for monitoring disease progression. Besides, previous metabolic studies have been focused mainly on identifying BC diagnostic biomarkers comparing differences between urinary metabolome of healthy *versus* BC patients. However, this experimental design might not be adequate to identify biomarkers for BC monitoring, which should be highly dynamic and should consider that after TUR some metabolites may not necessarily return to normality as it would be the case in healthy patientes. Add to this, the normal appearance urothelium can contain not observable pre-neoplasic lesions by cystoscopy, such as hyperplasia or dysplasia, which may be responsible of metabolic changes.

On the whole, these facts show the current lack of new dynamic and non-invasive biomarkers for BC diagnosis and follow-up and give support to the use of metabolomics as a technique for biomarker discovery.

Chapter 3: Hypothesis and objectives

The hypotheses underlying this doctoral Thesis are:

- The metabolic reprogramming of bladder tumors triggers alterations in metabolic pathways.
- The tissue metabolome can be used to identify biomarkers of diagnosis and aggressiveness of BC.
- The urinary metabolome reflects the genomic and metabolic alterations produced in bladder tumors.
- Urinary metabolome provides dynamic information about tumor biology.
- The urinary metabolome can be used to identify biomarkers for diagnosis and monitoring of patients with NMIBC.

The general objective of this Thesis is to identify a non-invasive urinary metabolomic profile for the diagnosis and monitoring of BC, which in turn is related to the metabolic and transcriptomic alterations found in bladder tumor tissues. This objective can be summarized in five general goals:

- To identify by HRMAS ^{1}H NMR analysis and chemometric analyses a tissue metabolic signature to be used as a biomarker of diagnosis and aggressiveness of BC.
- To identify by HRMAS ^{1}H NMR and pathway analysis the main disturbed metabolic pathways in BC tissues.
- To know what metabolic genes are being involved in BC and the processes that are responsible for their regulation.
- To integrate metabolomics and transcriptomics data coming from tissue analysis with the purpose of understanding better the biology of the bladder tumor and the genes-metabolites networks.
- To identify by ^{1}H NMR and UPLC-MS and chemometric analysis urinary metabolomic profiles to be used as a biomarker of diagnosis and monitoring of NMIBC.
- To analyze the link between the altered metabolic pathways in BC tissues and the perturbed urinary metabolome.

Chapter 4: Integrative metabolomic and transcriptomic analysis for the study of bladder cancer

Introduction

The limitations of diagnostic tools for the correct management of BC patients have driven to the search of non-invasive approaches to find new biomarkers within the field of "omics" (i.e. genomics, transcriptomics, proteomics, metabolomics). Nevertheless, in the context of BC none biomarker have been implemented into clinic routine[132]. Alterations in genes and metabolic pathways have been considered hallmarks of cancer. Bearing this in mind, integrative transcriptomic and metabolomic studies show high potential in biomarker research for the implementation of precision medicine.

In this line, technological developments in MS and NMR have fostered the study of cancer metabolism. Several metabolomic studies have been performed in tissues or biofluids, such as urine or serum to identify BC biomarkers for diagnosis or monitoring of patients[53,133]. However, further research and validation are required to demonstrate their clinical usefulness. Additionally, although several genomic analyses have been carried out in BC showing altered expression or mutations in metabolic enzymes, there are no reported integrative transcriptomic and metabolomic studies in BC tissues or liquid biopsies as far as we know.

Here, we report tissue and urinary metabolic signatures as biomarkers of BC through the study of gene-metabolite networks and the integration of metabolomics and transcriptomics data.

Material and Methods

Patient selection and sample collection

This study was approved by the Ethics Committee for Biomedical Research of the Instituto de Investigación Sanitaria Hospital Universitario y Politécnico La Fe (Valencia, Spain).

21 BC patients (14 males and 7 females) were recruited in the Urology Service of the Hospital Universitario y Politécnico La Fe (Valencia, Spain). All the patients underwent TUR. Tissues and urines were collected from each patient. The samples were stored by the Biobanco La Fe (PT13/0010/0026) and processed following standard operating procedures.

Tumor (n=22) and adjacent non-tumor tissues (n=22) were collected during TUR, and were immediately placed into cryo-vials, immersed in liquid N_2 and stored at -80˚C**.** Adjacent pieces of the tumor tissues underwent routine histopathological examination. BC presence, grade and tumor stage were determined and tumors were classified as NMIBC (Ta-T1) or MIBC (≥T2). Two tumor episodes (primary tumor and recurrence) were studied from one patient. Therefore, 22 tumor samples were collected from 21 patients.

Urines were collected before TUR (pre-TUR: BC, n=22) and one month after TUR (post-TUR: control, n=13) and were stored at -80˚C. Nine urines post-TUR, from MIBC patients that underwent RC and NMIBC with incomplete TUR or taking antibiotics at the time of urine collection, were discarded. BC absence in post-TUR samples was confirmed by cystoscopy.

The clinical and demographic data are shown in **Table 4.1**, and the scheme of sample collection is shown in **Figure 4.1**.

Table 4.1 Clinical and demographic data of patients included in this study.

	TISSUES		URINES
PLS-DA models (BC *vs* Control)	**Calibration (CV)**	**Validation**	**Calibration (CV)**
Male/female patients	17 (10/7)	5 (4/1)	21 (14/7)
Mean age (standard deviation)	71 (9)	63 (11)	69 (10)
Total samples (male/female)	34 (20/14)	10 (8/2)	35 (23/11)
Tumor samples (BC)	17	5	22
Non-tumor samples (Control)	17	5	13
Primary/Recurrent BC	15/2	4/1	19/3
Tumor stage (Ta,T1,T2)	7/6/4	3/0/2	10/6/6
Tumor grade (High/Low)	13/4	5/0	18/4
Recurrence risk group 5 yr (EORTC)[a]: L/ L-I/ H-I/ H	2/4/6/1	0/0/3/0	2/4/9/1
Progression risk group 5 yr (EORTC)[a]: L/L-I/H-I/H	4/2/5/2	0/1/2/0	4/3/7/2

Note: [a]Probability of recurrence and progression in NMIBC (n=16) according to the EORTC risk tables total score; L: Low risk, L-I: Low-Intermediate risk, H-I: High-Intermediate risk, H: High risk.

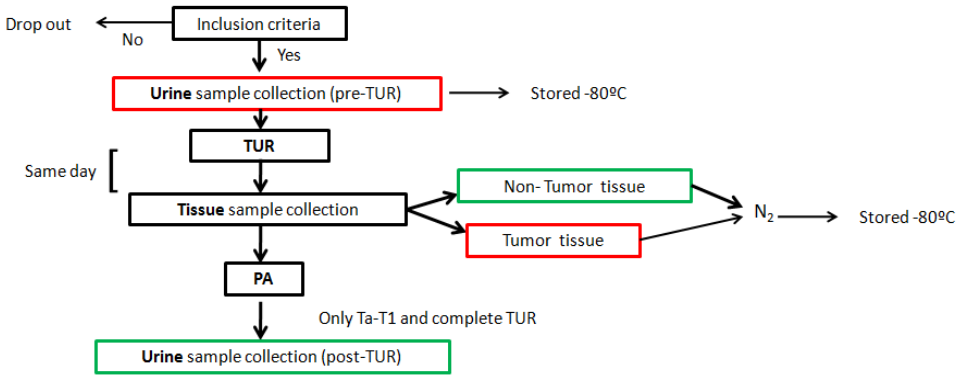


Figure 4.1 Sample collection scheme.

Tissue NMR experiments

Tissues were prepared for measurements according to already published data[90]. Briefly, 8-40 mg tissue sample was weighted and introduced into a disposable KelF insert (external diameter: 3 mm, internal diameter: 2.6 mm, total internal volume: 30 µL, Bruker, Billerica, MA, USA). 10 µL of D_2O were added to each insert and them it was sealed with a small screw cap. The insert containing the bladder tissue was placed in a standard 4mm ZrO_2 rotor, and this was transferred to the precooled at 0˚C NMR probe.

Experiments were acquired at a real temperature of 277˚K. NMR spectra were acquired using a Bruker Avance DRX 600 spectrometer (Bruker GmbH, Rheinstetten, Germany) operating at a ^1H frequency of 600.13 MHz. The instrument was equipped with a 4 mm triple resonance $^1H/^{13}C/^{15}N$ HRMAS probe with magnetic field gradients aligned with the magic angle axis. A Bruker cooling unit was used to control the temperature by cooling down the bearing air flowing into the probe. For all NMR experiments, samples were spun at 5 kHz in order to keep the rotation sidebands out of the acquisition window. For each tissue sample, 1D ^1H NOESY spectrum with water presaturation was acquired in 5 min using a 1.14 s acquisition time, 128 transients, a 12 ppm (7211 Hz) spectral width, a mixing time of 100 ms and a relaxation delay of 1 s. In addition, 2D ^1H-^1H TOCSY using a DIPSI2 sequence for mixing were acquired with a 142 ms acquisition time, 50 ms spin lock duration, 7211 Hz spectral width, and a 1.5 s relaxation delay. Sixteen transients were averaged for each of the 256 increments during t1, corresponding to a total acquisition time of 2 h 33 min. After NMR study, the disposable inserts containing the tissues were frozen and preserved at -80˚C until the transcriptomic analysis.

Tissue transcriptomic experiments

Frozen tissues after ^{1}H HRMAS NMR experiments were thawed and underwent microarray experiments. Total RNA was extracted with the miRNeasy Mini Kit, DNA was digested and RNA integrity checked on an Agilent 2100 Bioanalyzer. After quality filtering, 8 primary tumors (Ta n=3; T1 n=2; T2 n=3) and 10 non-tumor tissues were analyzed. cDNAs from total RNA (30 ng) were generated, fragmented, biotinylated, and hybridized to the GeneChip Human Transcriptome Array 2.0 (HTA 2.0) (Affymetrix, Thermofisher, Watham, Massachusetts, USA). The arrays were washed and stained on a GeneChip Fluidics Station 450 (Affymetrix); scanning was carried out with the GeneChip Scanner 3000 7G; and image analysis with the AffymetrixGeneChip Command Console software.

Urine NMR experiments

Urine samples were thawed at room temperature and were prepared following the established procedures for urine samples[134]. Phosphate buffer (pH 7.4) was prepared by weighing Na_2HPO_4 (28.85 g), NaH_2PO_4 (5.25 g), NaN_3 (0.195 g) and DSS (0.218 g) into a 1 L volumetric flask. 200 mL of D_2O were added and the flask was filled to 1 L with water. To 500 µL of urine 200 µL of phosphate buffer were added. This mix was centrifuged at 10000 rpm for 5 min at 5˚C. After this, 550 µL of the supernatant were transferred to a 5 mm NMR tube for analysis.

The experiments were recorded at 298˚K. Spectra were acquired using a Bruker Avance DRX 600 spectrometer (Bruker GmbH, Rheinstetten, Germany) operating at a ^{1}H frequency of 600.13 MHz. The instrument was equipped with a 5 mm triple resonance $^{1}H/^{13}C/^{31}P$ probe. For each urine sample, 1D ^{1}H NOESY spectra using water presaturation were acquired in 3 min using a 3.91 s acquisition time, 32 transients, a 14 ppm (8370 Hz) spectral width, a mixing time

of 100 ms and a relaxation delay of 1 s. Moreover, 2D ^{1}H-^{13}CHSQC spectra were acquired, to assess the assignments of the overlapped signals in 1D ^{1}H spectra.

Tissue NMR data pre-processing and analysis

Spectra processing and handling were performed with *MestReNova* (version 6.0.2). Spectra were Fourier-transformed, baseline, and phase corrected, and chemical shift referenced to the creatine singlet (3.03 ppm) and Ala doublet (1.48 ppm). The main metabolites in the region from 0.8 to 9.5 ppm were assigned, according to the bibliography and NMR databases (i.e HMDB and Metabolomics: main)[77,92,135–137].

Mean metabolite intensities were compared between groups. The intensity of the assigned resonances was transferred to *MetaboAnalyst 3.0*[138]. U-Mann Whitney test determined the significant differences between control and BC tissues. ANOVA (Tukey's posthoc test) evaluated the differences between NMIBC, MIBC, and control tissue, and the changes between Ta, T1, T2 stages and control tissue.

Data analysis was performed on 1D spectra. Regions from 0.5-4.8 and 5.2-9.5 ppm were included. Spectral region of water (4.79-5.2 ppm) was excluded. PLS-DA was performed using the software *PLS_Toolbox Solo 8.0* (Eigenvector Research, Inc., Manson, WA, USA). Previously, the spectra were peak aligned using the *icoshift*[139] algorithm to correct for minor changes in chemical shift due to differences in pH. Finally, the spectra were normalized using the sample weight as normalization factor and autoscaled. The set of tissue spectra was split into calibration (n=34) and validation (n=10) subsets randomly selected. Two PLS-DA models were calculated including the original set of 11698 spectral features and after a feature selection based on the variable importance in projection (VIP) score using VIP=1 (4800 features retained) as threshold value. Features selection

during model optimization aimed at improving the predictive model performance and facilitating its interpretation. The selection of the VIP threshold values was based on the analysis of the evolution of CV-figures of merit (dQ^2 and mean classification error of CV) as a function of the threshold value.

Metabolite identification was carried out on the selected subset of highly discriminant features (i.e. VIP>1 in the PLS-DA model). The obtained metabolite list was then used to perform a pathway enrichment and topology analysis using a global test and a relative betweenness centrality measure in *MetaboAnalyst.*

Tissue transcriptomic data analysis

The expression of data was normalized and background and batch corrected using the Signal Space Transformation-Robust Multi-Chip Analysis (SST-RMA) implemented in the Transcriptome Analysis Console software version 4.0 (TAC 4.0). Data were deposited in Gene Expression Omnibus (GEO) (GSE121711). A PCA and a heatmap analysis were carried out to observe differences between BC and non-tumor tissues on the basis of the whole transcriptome. A fold change of at least 2 or -2 and a false discovery rate of 0.05 was considered as selection criteria. Gene Ontology of Biological processes (GOBP), ChEA, and Encode Histone Roadmap were performed *in silico* using Enrich webtool[140]. Metabolism related genes showing deregulation between tumors and non-tumor samples were selected from GOBP and GeneCards. ChEA and Encode Histone Roadmap analysis were performed to identify the putative TFs binding and histone marks in these metabolic genes. Moreover, the identified genes were used to perform non-supervised hierarchical clustering of the Cancer Genome Atlas (TCGA) data according to stage, grade, and mRNA subtypes[45] using MeV[141].

Integrated metabolic pathway analysis on results obtained from combined metabolomics and transcriptomics analyses in tissues was carried out using the

information of: joint pathway analysis of platform *MetaboAnalyst,* the tool *Wikipathways*[85]*,* and the SMPDB[84].

Urine NMR data pre-processing and analysis

Spectra were Fourier-transformed, baseline, and phase corrected, and the chemical shift was referenced to the DSS singlet (0.0 ppm). The main metabolites in the region from 0.8 to 9.5 ppm were identified, according to already published data and NMR databases[77,92,135–137]. The spectra were binned into 0.003 ppm buckets using *MestReNova*.

A PLS-DA analysis was performed considering the spectral regions from 0.8-4.5 and 6.5-9.0 ppm (2068 features). Water (4.5-5.1 ppm), urea (6.1-5.55 ppm) and regions lower 0.8 ppm and greater 9.00 ppm were excluded from the analysis to avoid interferences arising from differences in water suppression and variability from urea and DSS signals. Prior to data analysis, urinary spectra were normalized to the sum of all variables (1-norm) and autoscaled. Urines were classified as control (urines collected post-TUR) or BC (urines collected pre-TUR). CV data splitting was performed at the highest level of sampling hierarchy, which in this case was the volunteer. The statistical significance of PLS-DA figures of merit estimated by CV was assessed by permutation test with 100 iterations. The most important metabolites in the statistical model (VIP>1) were identified and used to perform a pathway enrichment and topology analysis using *MetaboAnalyst*.

Results

Tissue NMR profile

Representative tissue spectra from a non-tumor sample and three different stages of BC (Ta, T1, and T2) are displayed in **Figure 4.2**. Non-tumor tissues spectra are dominated by intense lipid signals in the aliphatic region that were practically absent in tumors. By contrast, in tumor spectra highlight the signals of small metabolites such as GSH, Tyr and cytidine diphosphate (CDP) in the aromatic region. Assignment of signals in a 1D ^{1}H NMR spectra of bladder tissues is displayed in **Table 4.2**.

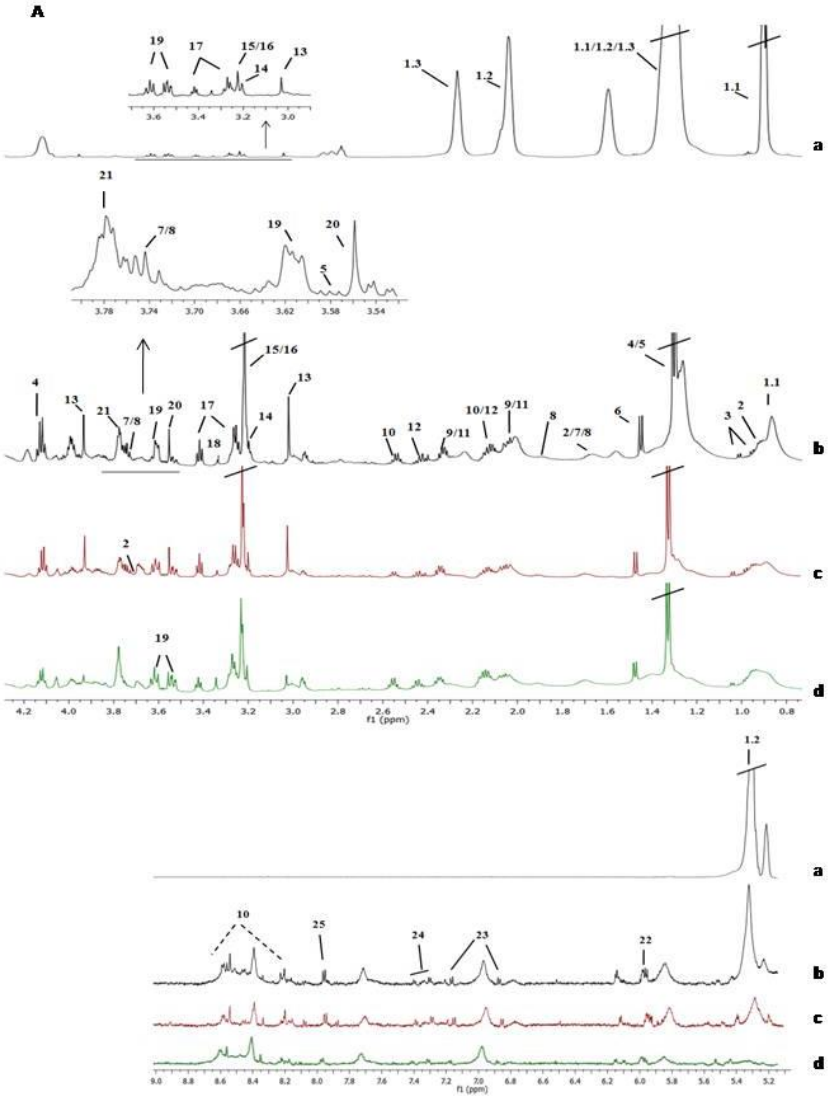


Figure 4.2 Assignment of the main signals in 1D ^{1}H NMR spectra of tissue samples. Representative NMR spectra of non-tumor tissue (a), and various BC pathologic stages Ta/T1 (c and d) and T2 (b). NMR spectra were shifted along the Y-axis and the spectral region of water was removed from the figure for a better visualization.

Table 4.2. Assignment of the metabolites and metabolic signals identified in 1D ^{1}H NMR tissue spectra.

Nº	Metabolite	Group	Chemical shift (ppm)
1.1	Lipid fragment –(n)CH_2-CH_2-CH_2-CH_3 (a)	CH_3	0.90
2	Leucine	δCH_3	0.93
3	Valine	γCH_3	0.98
3	Valine	γCH_3	1.04
1.1	Lipid fragment –(n)CH_2-CH_2-CH_2-CH_3 (a)	(n) CH_2	1.29
1.1	Lipid fragment –(n)CH_2-CH_2-CH_2-CH_3 (a)	(2) CH_2	1.29
1.2	Lipid fragment –CH=CH-CH_2-CH=CH-CH_2-CH_2-(n)CH_2 (b)	(n) CH_2	1.29
1.3	Lipid fragment –(n)CH_2-CH_2-CH_2-COOH (c)	(n) CH_2	1.29
1.1	Lipid fragment –(n)CH_2-CH_2-CH_2-CH_3 (a)	(1) CH_2	1.33
1.2	Lipid fragment –CH=CH-CH_2-CH=CH-CH_2-CH_2-(n)CH_2 (b)	(1) CH_2	1.33
4	Lactate	CH_3	1.33
5	Threonine	γCH_3	1.33
6	Alanine	βCH_3	1.46
1.3	Lipid fragment –(n)CH_2-CH_2-CH_2-COOH (c)	(2) CH_2	1.59
8	Arginine	γCH_2	1.68
2	Leucine	βCH_2	1.70
2	Leucine	γCH	1.70
7	Lysine	βCH_2	1.71
8	Arginine	βCH_2	1.90
1.2	Lipid fragment –CH=CH-CH_2-CH=CH-CH_2-CH_2-(n)CH_2 (b)	(2) CH_2	2.03
9	Glutamate	βCH_2	2.05
11	Proline	βCH_2	2.06
12	Glutamine	βCH_2	2.13
10	Glutathione	CH_2	2.15
1.3	Lipid fragment –(n)CH_2-CH_2-CH_2-COOH (c)	(1) CH_2	2.26
9	Glutamate	γCH_2	2.33
11	Proline	βCH_2	2.34
12	Glutamine	γCH_2	2.44
10	Glutathione	CH_2	2.54
13	Creatine	CH_3	3.03
14	Choline	-N^+- $(CH_3)_3$	3.18
15	Phosphocholine	-N^+- $(CH_3)_3$	3.20
16	Glycerophosphocholine	-CH_2-NH_3^+	3.21

17	Taurine	$-CH_2-NH_3^+$	3.25
18	Methanol	CH_3	3.34
17	Taurine	$-CH_2-SO_3^-$	3.42
19	Myo-inositol	C1',3' H	3.52
20	Glycine	αCH	3.55
19	Myo-inositol	C4',6' H	3.62
7	Lysine	αCH	3.74
21	Glycerol	-CH-(OH)-	3.78
13	Creatine	CH_2	3.92
4	Lactate	CH	4.10
1.2	Lipid fragment $–CH=CH-CH_2-CH=CH-CH_2-CH_2-(n)CH_2$ (b)	=CH-	5.30
22	UDP-sugars	__	5.99
23	Tyrosine	C3',5' H	6.89
23	Tyrosine	C2',6' H	7.17
24	Phenylalanine	C2',6' H	7.32
24	Phenylalanine	C3',5' H	7.42
25	Cytidine diphosphate	CH	7.97
10	Glutathione	NH	8.15

Mean comparison of metabolites detected in tissue

From the mean comparison of metabolites intensities from HRMAS in tumor and control samples, a higher intensity of choline (Cho), CDP, myo-inositol (mI), UDP-sugars, and GSH was found in tumors tissues (**Figure 4.3; Table annexed 4.3A**). Intensity for lipid fragment $–(n)CH_2-CH_2-CH_2-COOH$ (named lipid (c) as in[92]) was higher in control samples. Most of the metabolites showing significant differences between tumor and control tissues (p-value<0.05) were also elevated in both tumor groups (NMIBC and MIBC) compared to control samples (**Figure 4.3**; **Table annexed 4.3B**)**.**

Considering the differences among the diverse tumors stages and controls, lipid fragments $–CH_2-CH_2-CH_2-CH_3$ (lipid (a) as in[92]) and$–CH=CH-CH_2-CH=CH-CH_2-CH_2-(n)CH_2-$ (lipid (b) as in[92]), Tyr, Cho, CDP, mI, UDP-sugars and GSH showed significant changes between all groups. Specifically, T2 tumors (the most aggressive here considered) presented the highest intensity of Cho and GSH. T2

and T1 stages showed elevated mI and Tyr. On the other hand, Ta tumors (the least aggressive tumors) showed a significant increment of UDP-sugars and CDP (**Figure 4.3**; **Table annexed 4.3C**). Lactate signals at 1.33 ppm and 4.1 ppm overlapped with lipid signals. The overlap hinders the discrimination between lactate and lipid signals in control samples. Therefore, the lactate intensity was compared only between the different tumor stages and not respect to the control group. Higher intensity of lactate in T2 than in T1 or Ta was observed.

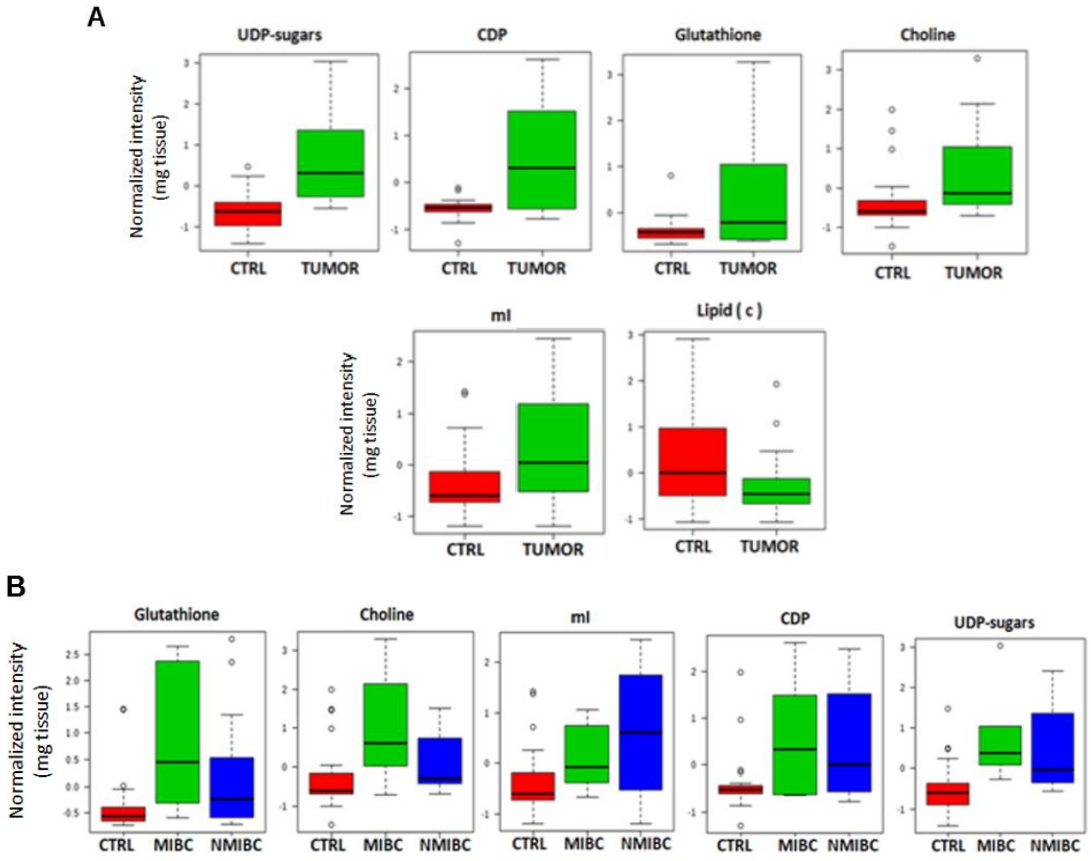


Figure 4.3 Box and whisker plots illustrating discrimination between: A, tumor and non-tumor tissues; B, controls, NMIBC and MIBC; C, different stages of BC and control tissues.

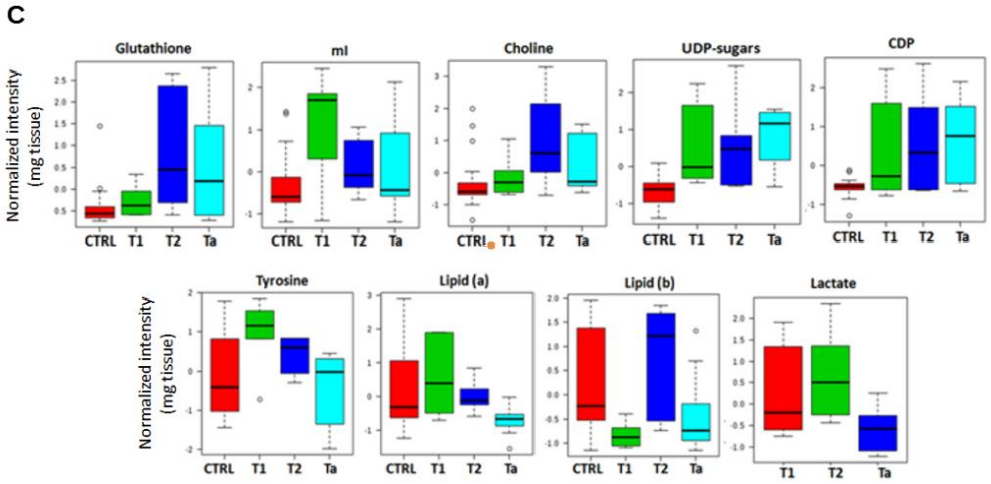


Figure 4.3 Box and whisker plots illustrating discrimination between: A, tumor and non-tumor tissues; B, controls, NMIBC and MIBC; C, different stages of BC and control tissues.

PLS-DA analysis in NMR tissue data

A chemometric analysis was performed on 1D HRMAS tissue spectra including the original set of variables, and after a feature selection (VIP>1). A VIP threshold of 1 was selected as a compromise between the model complexity and the predictive performance of the PLS-DA model (**Figure 4.4**). The predictive performance of the initial PLS-DA model (n=44, LV=2) was assessed by CV providing a sensitivity of 82.4% and a specificity of 88.2% (confidence intervals are provided in **Table 4.4**). Permutation test (100 permutations) provided a p-value<0.05. The analysis of the samples included in the external validation set presented sensitivity and selectivity values of 100% for the two PLS-DA models (original set of variables and VIP>1) (**Table 4.4**).

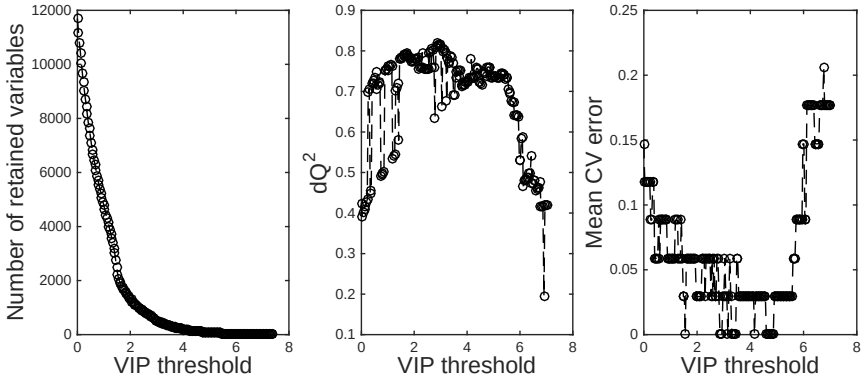


Figure 4.4 Evolution of three indicators (number of features, discriminant Q^2 and mean CV-classification error) of the discriminant performance of the PLS-DA model as a function of the VIP cutoff value used for the elimination of features in an initial PLS-DA model.

Table 4.4 Indices of test validity estimated for the evaluation of predictive performance of PLS-DA models between tumor and non-tumor tissue samples with LVs=2.

Indices test Validiy	Calibration set (CV)	Validation set
	Estimation (95% CIs)	Estimation (95% CIs)
Sensitivity	82.4 (62.3-97.9)%	100 (46.3-98.1)%
Specificity	88.2 (55.8-95.3)%	100 (46.3-98.1)%
PPV[a]	87.5 (60.4-97.8)%	100 (46.3-98.1)%
NPV[b]	83.3 (57.7-95.6)%	100 (46.3-98.1)%
ACC[c]	85.3%	100%
MCC[d]	0.707	1
PT[e] (p-value)	--	0.015
AUROC[f]	--	1

Note: [a]Positive predictive value; [b]Negative predictive value; [c]Diagnostic accuracy; [d]Matthew's correlation coefficient; [e]Permutation test; [f]Area Under the Receiver Operating Characteristic.

Predicted "y values" for calibration and validation subsets considering the PLS-DA model performed in bladder tissues after an initial feature selection (VIP>1) are showed in **Figure 4.5**. The most discriminant metabolites between tumor and non-tumor tissues identified in this PLS-DA model included: lipids (a, b and c fragments), threonine (Thr), lactate, Ala, Glu, proline (Pro), Gln, GSH, creatine, Cho, phosphocholine (PCho), glycerophosphocholine (GPCho), taurine, methanol, mI, Gly, glycerol, UDP-sugars, Tyr, and CDP. This set of metabolites reflected alterations in metabolic pathways related to the metabolism of several AA pathways and also with GSH, taurine and hypotaurine or glycerolipids (**Table 4.5; Figure 4.6**;).

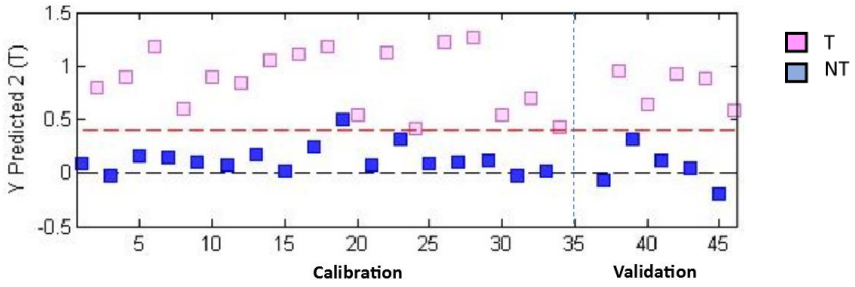


Figure 4.5 Predicted y values for calibration and validation subsets considering the PLS-DA model performed in bladder tissues after an initial feature selection (VIP>1).

Table 4.5. Identified metabolites and associated altered metabolic pathways in bladder tissues.

Altered pathways in BC	Metabolites	p-value	Impact
Alanine, aspartate, glutamate	Ala, Glu, Gln	6.5E-4	0.44
Taurine and hypotaurine	Taurine, Ala	9.0E-3	0.36
Aminoacyl-tRNA biosynthesis	Pro, Gly, Ala, Gln, Thr, Glu	1.0E-5	0.06
Methane	Gly, methanol	2.6E-2	0.02
Glycine, serine, threonine	Gly, creatine, Cho, Thr	3.5E-4	0.29
Glutathione	GSH, Gly, Glu	2.0E-3	0.25
Glycerolipid	FAs, glycerol	2.3E-2	0.19
Arginine and proline	Glu, Gln, Pro, creatine	2.0E-3	0.17
Glutamine and glutamate	Glu, Gln	3.0E-3	0.14
Glycerophospholipid	Cho, GPCho, PCho	3.0E-3	0.08

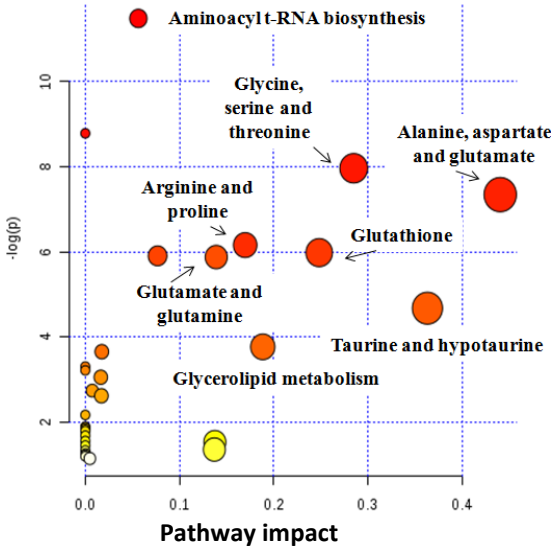


Figure 4.6 Analysis of altered metabolic pathways in bladder tissues.

Note: the color and the size of each circle indicate its p-value and pathway impact value, respectively.

Transcriptomic analysis in tissues

PCA and heatmap analyses were performed to search transcriptomic differences between BC tissues and non-tumor tissues. A 3 principal component model explaining 66% of the data variance showed clear differences between tumor and non-tumor samples on the basis of the whole transcriptome (**Figure 4.7A**). Following a fold change of at least 2 or -2 and a false discovery rate of 0.05 as selection criteria, a total of 4409 transcripts differentiated tumor and non-tumor samples, being predominantly downregulated in tumors (3112 transcripts) (**Figure 4.7B**).

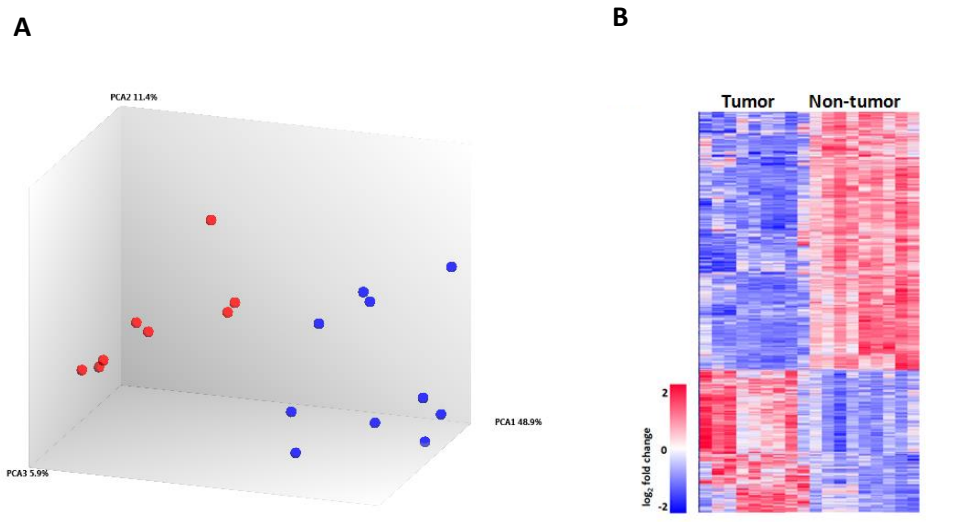


Figure 4.7 A, PCA using the overall transcriptome; **B, Heatmap showing the hierarchical clustering of tumor and non-tumor samples, as well as the number of downregulated and upregulated transcripts.** Note: In PCA, tumors are shown in red and non-tumors in blue.

The GOBP analysis revealed that different metabolism-related genes were significantly downregulated in tumors (**Figure 4.8**), including various mitochondrial RNAs and genes related with the regulation of AAs and amine metabolism, purine biosynthetic processes and OXPHOS. On the contrary, among the upregulated transcripts in tumors, no significant metabolism-related categories were found.

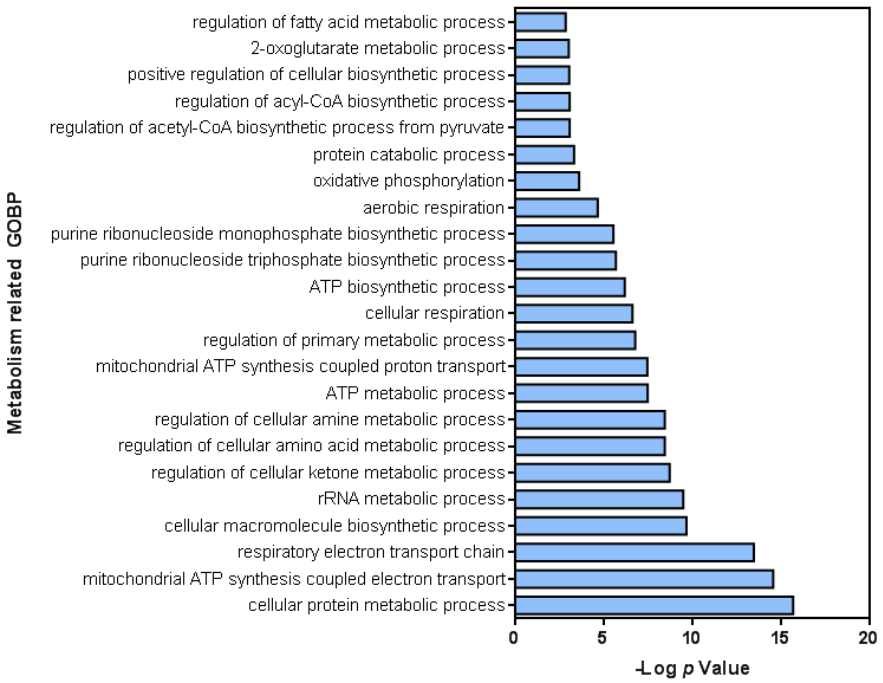
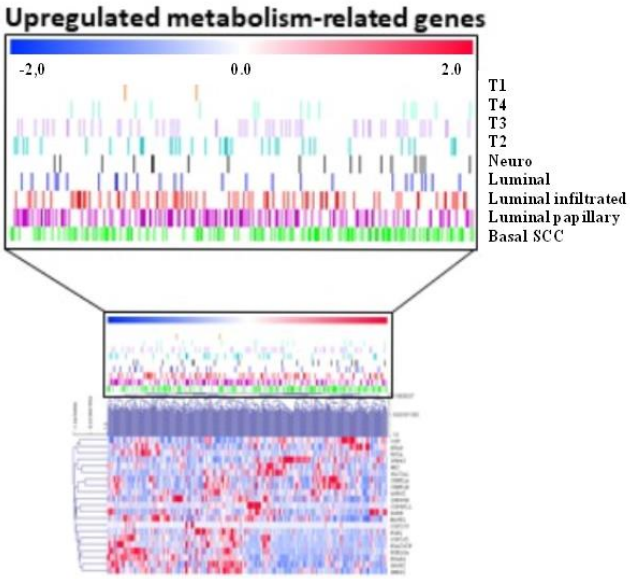


Figure 4.8 Summary of GOBP analysis: main metabolism-related biological processes altered in tumor samples are shown.

A total of 364 unique genes were identified considering only protein-coding genes linked to metabolic processes from GOBP and GeneCards databases (**Table annexed 4.6**), from which 20 genes were overregulated. Interestingly, some of the overregulated genes have already been linked to cancer such as:

PPARG involved in lipid biogenesis regulation[61]; *HK2* and solute carrier family 2 facilitated glucose transporter member 1 (*SLC2A1)*[142], both them related to glycolysis; the ribose-5-phosphate isomerase A (*RPIA*) linked to PPP; and genes belonging to cytochrome P450[143] (CYP450) (C*YP2J2, CYP2C9, CYP4F11*). On the other hand, key metabolic genes were found downregulated in BC tissues, including genes related to: 1) pyruvate metabolism: pyruvate dehydrogenases (*PDHA1, PDHB, PDHX)* and pyruvate dehydrogenase kinase *4 (PDK4);* 2) TCA cycle: *FH, IDH3A, IDH3B, MDH1, MDH2, ACO1, OGDH,* and *SUCLG1*; 3) polyamine metabolism: *AMD1, SMS, ODC1, SAT2, AOC3;* 4) AA metabolism: *GLS, GOT2 , MUT, ASS1, MAOA, MAOB;* 5) redox status: *GPX3* and *GPX*4, catalase, *SOD1*, and glutaredoxin (*GLRX3*). Importantly, although our series is enriched in NMIBC, the expression of the upregulated and downregulated metabolism-related genes also showed similar behavior in multiple tumor samples present in the TCGA cohort (enriched in MIBC) without discrimination of stage, grade of previously identified mRNA tumor subtypes (**Figure 4.9**).



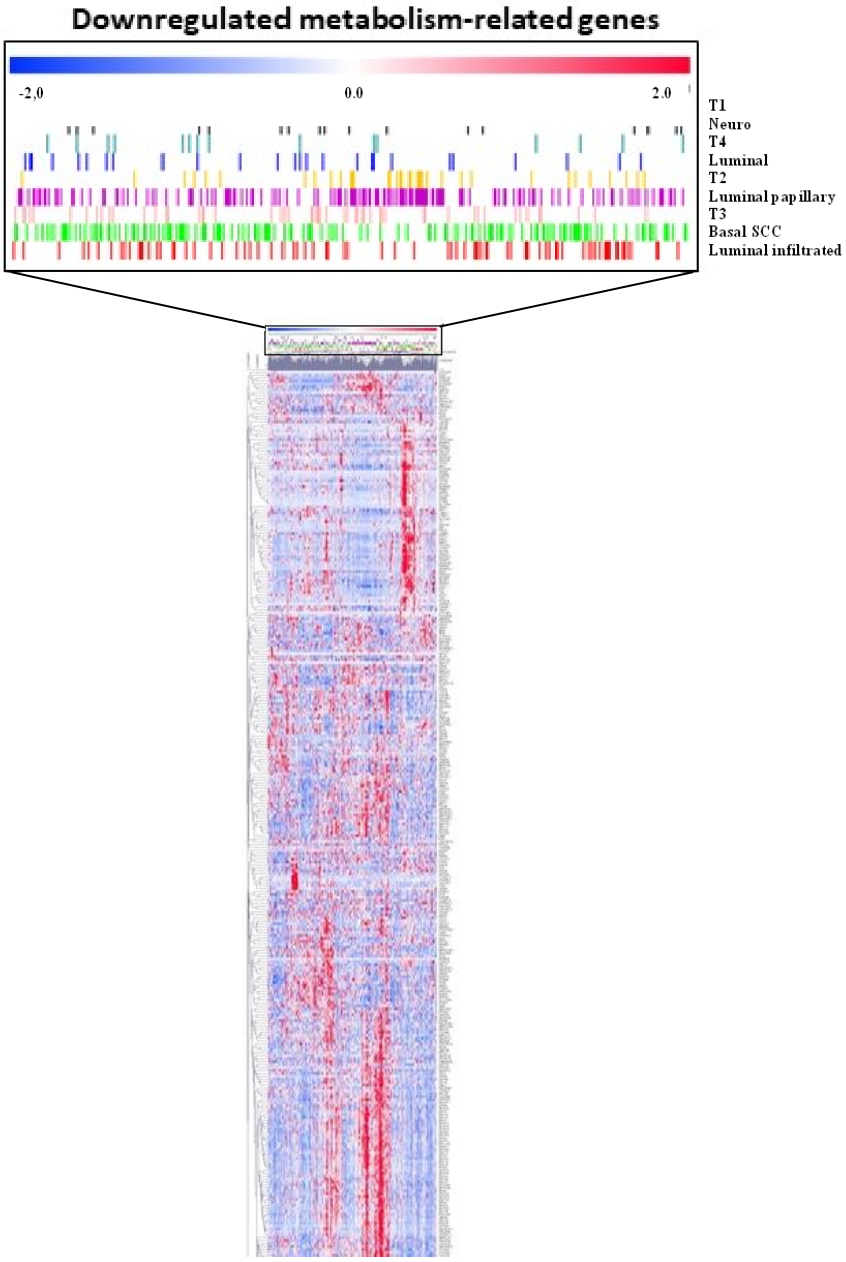


Figure 4.9. Comparison between upregulated and downregulated metabolism-related genes with gene expression of tumor samples present in the TCGA cohort.

In silico analysis revealed that multiple genes were bound by TFs currently known as potentially involved in BC pathogenesis (**Table annexed 4.7**). Remarkably, these could also act as transcriptional repressors, such as ETS1, TTF2, E2F1, YY1, FLI1, ASH2L, E2F4, VDR, GABP, JARID1A, CTCF, KLF4, and CHD1 (**Figure 4.10**). Besides, some of these factors are known to modulate gene expression through direct chromatin remodeling, such as GABP, JARID1A, CTCF, and CHD1.

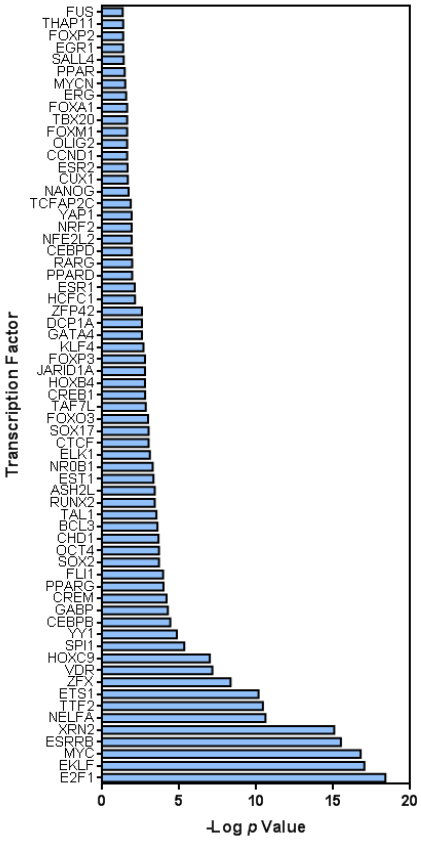


Figure 4.10. **Summary of putative binding motif enrichment analysis using the Enrich webtool, showing the relative relevance of various transcription factors**

Accordingly, we used the Encode Histone Roadmap[140] to determine whether the downregulated metabolism-related genes were specifically associated with particular histone marks. This analysis revealed that active histone marks such as H3K79me2, H3K79me3, or H3K27ac, as well as repressive marks such as H3K36me3, H4K20me1 or H3K9me3, were present in the downregulated metabolic genes (**Figure 4.11**; **Table annexed 4.8**).

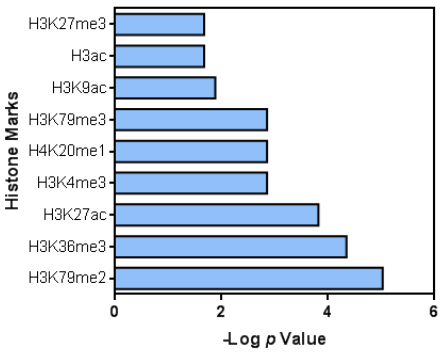


Figure 4.11. Summary of putative binding motif enrichment analysis using the Enrich webtool, showing the relative relevance of various histone marks associated with the differentially expressed transcripts in BC tissues.

On the other hand, the transcriptome analysis revealed the presence of specific splicing variants of some metabolic genes (*ASS1, GOT2, RARS, OAT, AOC3, HIBADH, DLD, FH, AUH, MUT, GLS, MAOA, ADH1C, ADH5, OGDH, SUCLG1, MDH1, MDH2, SMS, ODC1, FARSB, ACADM, PPM1L, PDK4, PDHX, PDHA1*) in tumors compared to non-tumor samples. Some of them are shown in **Figure 4.12**.

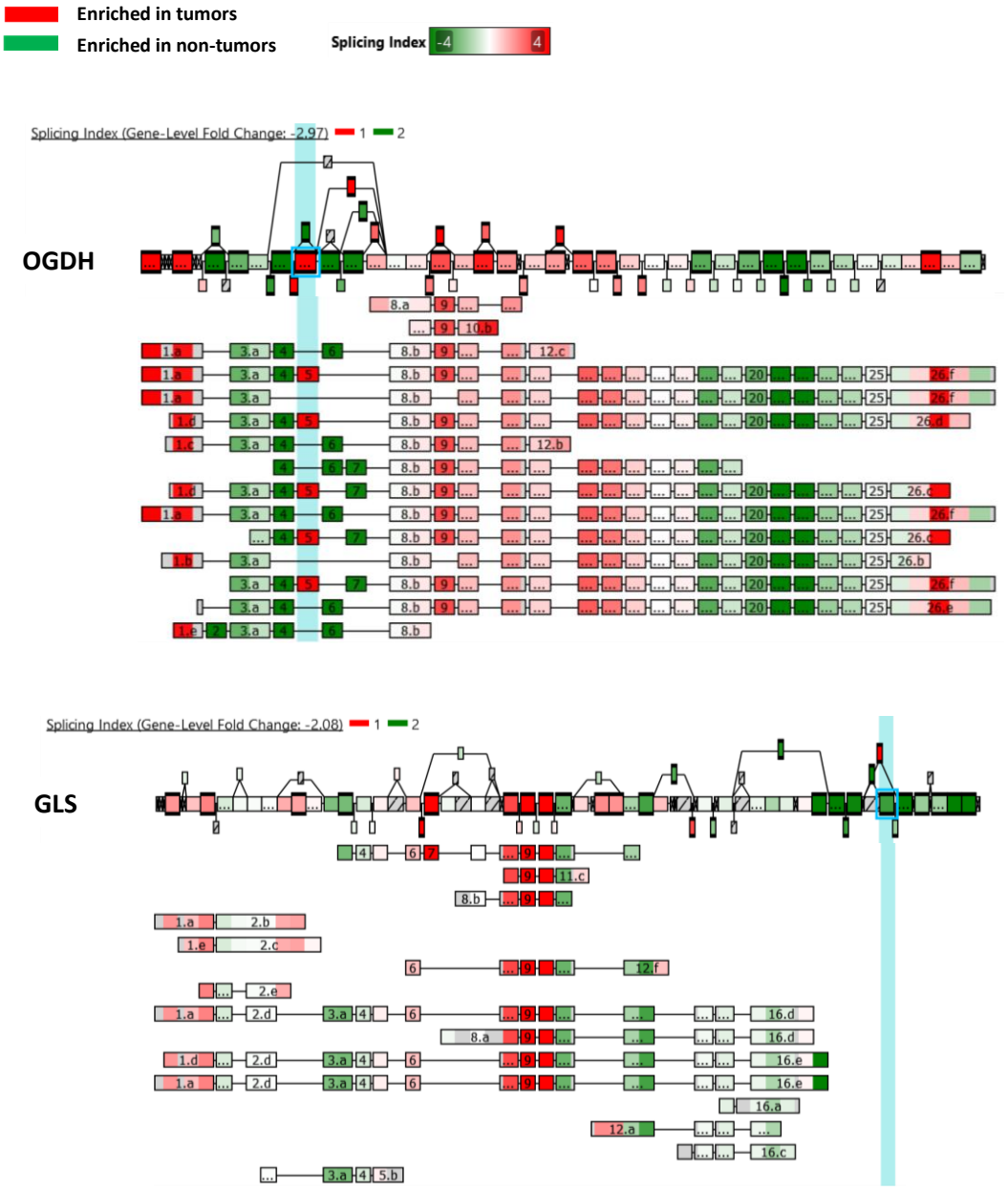


Figure 4.12 Splicing variant analyses in some metabolic genes.

Note: *FARSB*: Phenylalanyl-tRNA Synthetase Subunit Beta; *GLS*: Glutaminase; *MAOA*: Monoamine Oxidase A; *OGDH*: Oxoglutarate Dehydrogenase.

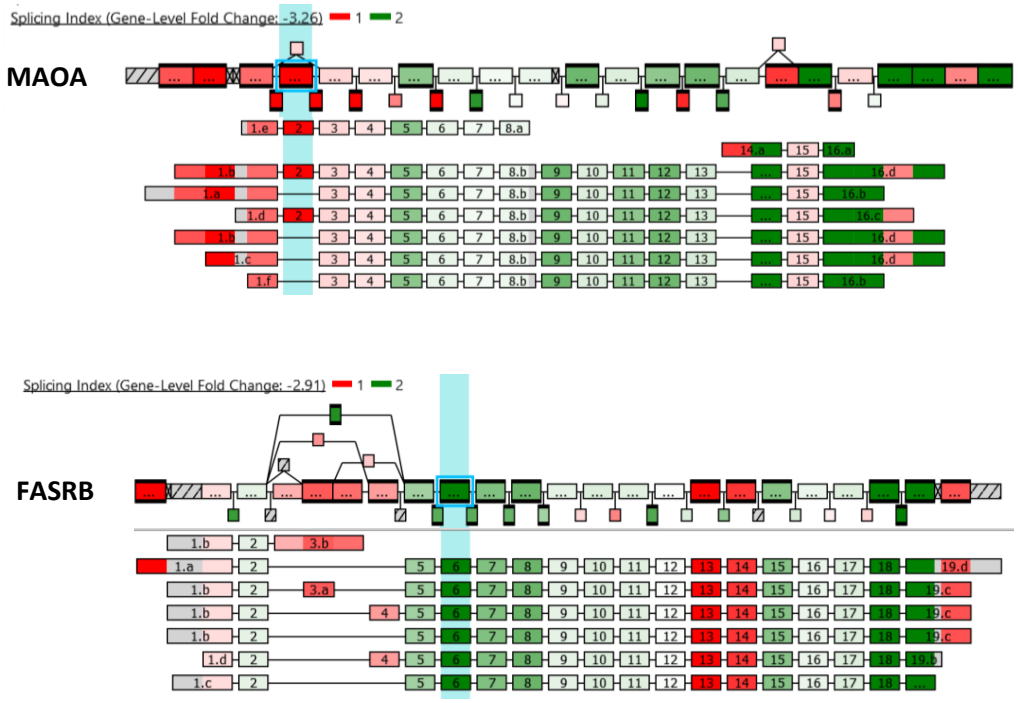


Figure 4.12 Splicing variant analyses in some metabolic genes.

Note: *FARSB*: Phenylalanyl-tRNA Synthetase Subunit Beta; *GLS*: Glutaminase; *MAOA*: Monoamine Oxidase A; *OGDH*: Oxoglutarate Dehydrogenase.

Finally, the integrative analysis using metabolomics and transcriptomics data revealed the deregulated activity of multiple biochemical pathways such as: TCA cycle, OXPHOS and polyamine, and AA metabolism, among others (**Figure 4.13**).

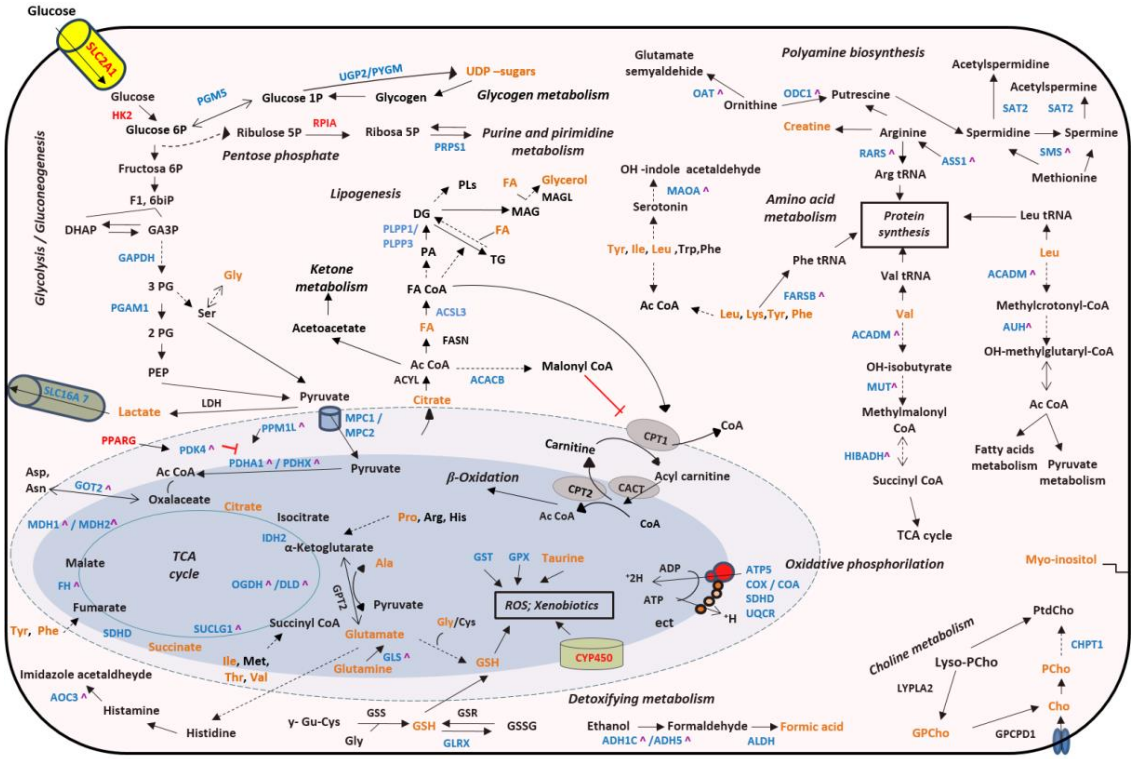


Figure 4.13 Integrated altered metabolic pathways in bladder tumors according to our metabolomics and transcriptomics data. Enzymes shown in red indicate significant overexpression in BC tissue; enzymes shown in blue indicate significant downexpression; metabolites shown in orange indicate that they were identified in our studies; ^ Indicate that these genes presented a differential alternative splicing in our tumor tissues; Dashed lines indicate that the reaction is not direct. ADP: adenosine diphosphate; Asn: asparagine; Asp: aspartate; 1;3 biPG: 1,3-bisphosphoglycerate; DG: diacylglycerides; DHAP: dihydroxyacetone phosphate; F1,6biP: fructose-1,6-bisphosphate; GA3P: glyceraldehyde 3-phosphate; His: histidine; Lys: lysine; MAG: monoacylglycerols ; NO: nitric oxide; PA: phosphatidic acid ; PEP: phosphoenolpyruvate; 2PG: 2-bisphosphoglycerate; 3PG: 3-bisphosphoglycerate; Phe: phenylalanine; tRNA: transfer ribonucleic acid. The names of the significant metabolic genes are detailed in the (**Table annexed 4.6**).

Urine NMR profile

Assignment of metabolites identified in BC urines and controls is displayed in **Figure 4.14** and in **Table 4.9**. In all samples the resonances of urea, creatinine, trimethylamine-*N*-oxide, hippuric acid, and citrate were the most predominant.

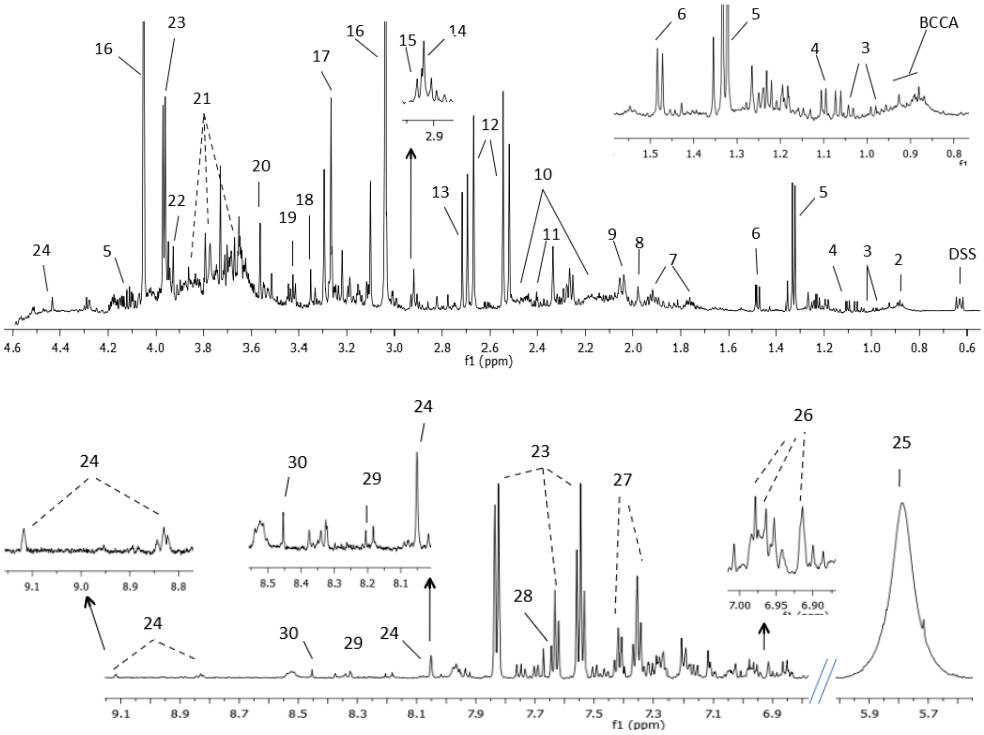


Figure 4.14 Assignment of the main signals in 1D ^{1}H NMR spectra in a BC urine sample. The vertical scale of all the spectra was kept constant and the spectral region of water was removed from the figure. The intensity of peaks in the chemical shift region 5.0 – 9.0 ppm was scaled equally in all the spectra to show lower abundant metabolites

Table 4.9. Assignment of the metabolites identified in 1D ^{1}H NMR urine spectra.

Nº	Metabolite	Group	Chemical shift (ppm)
3	Valine	γCH_3	0.98
3	Valine	γCH_3	1.04
4	Methylsuccinic acid	αCH_3	1.08
5	Lactate	CH_3	1.33
6	Alanine	βCH_3	1.46
7	Lysine	βCH_2	1.71
8	Acetic acid	CH_3	1.97
7	Lysine	βCH_2	1.91
9	N-acetylneuraminic acid	CH_3	2.04
10	Glutamine	βCH_2	2.13
11	Succinic acid	2,3 CH_2	2.39
10	Glutamine	γCH_2	2.44
12	Citrate	CH_2	2.52
12	Citrate	CH_2	2.66
13	Dimethylamine	CH_3	2.71
14	Trimethylamine	CH_3	2.93
15	Dimethylglycine	CH_3	2.96
16	Creatinine	CH_3	3.03
19	Taurine	$-CH_2-NH_3^+$	3.25
17	Trimethylamine N-oxide	CH_3	3.28
18	Methanol	CH_3	3.34
19	Taurine	$-CH_2-SO_3^-$	3.42
20	Glycine	αCH	3.55
21	Sucrose	C6' H_2	3.82
21	Sucrose	C5' H	3.87
22	Creatine	CH_2	3.92
23	Hippuric acid	αCH_2	3.97
16	Creatinine	CH_2	4.05
5	Lactate	CH	4.10
24	Trigonelline	CH_3	4.43
25	Urea	$-NH_2$	5.80
26	3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA)	C4' H	6.93
26	3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA)	C6' H	6.97
27	Phenylalanine	C2,6' H	7.32

27	Phenylalanine	C3,5' H	7.42
23	Hippuric acid	C3,5' H	7.60
28	Pseudouridine	CH	7.66
23	Hippuric acid	C2,6' H	7.82
29	Hypoxanthine	C7' H	8.20
30	Formic acid	CH	8.45
24	Trigonelline	C3,5' H	8.82
24	Trigonelline	C1' H	9.10

PLS-DA analysis in NMR urine data

A PLS-DA analysis was performed using urine NMR data to classify BC and control urines. The PLS-DA model (n=35, LV=5) provided a sensitivity of 90.9%, specificity of 76.9%, a NPV of 83.0% and a PPV of 86.9% and an AUROC=0.9 (**Figure 4.15A**). Permutation test (100 permutations) provided a p-value<0.05 and assessed the statistical significance of the CV-predictive performance estimates of the model. Predicted y values for CV in the PLS-DA model performed in bladder urines are shown in **Figure 4.15B**.

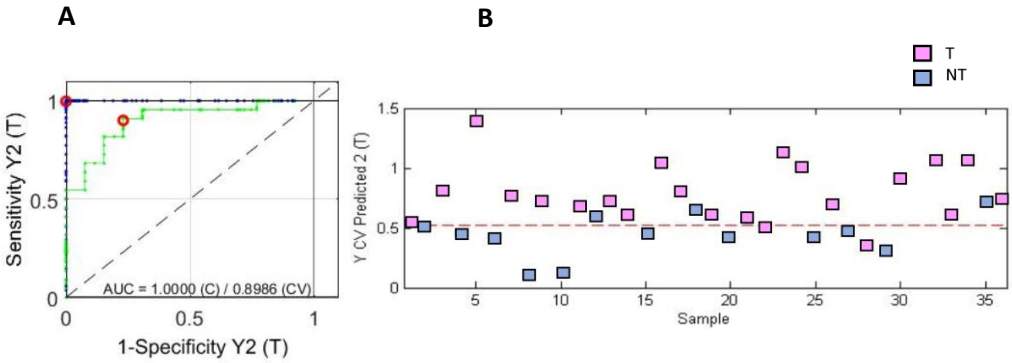


Figure 4.15 A, AUROC calculated for CV through a PLS-DA model performed in urine samples. B, Predicted *y* values for CV considering the PLS-DA model performed in bladder urines.

The most discriminant metabolites identified in urine (VIP>1) included: Val, methylsuccinic acid, lactate, Ala, Lys, N-acetylneuraminic acid, Gln, succinic acid, citrate, creatinine, trimethylamine-*N*-oxide, methanol, taurine, sucrose, hippuric acid, Phe, pseudouridine and trigonelline. These metabolites were linked with the following disturbed metabolic pathways: Ala, Asp and Glu metabolism, taurine and hypotaurine, TCA cycle, and aminoacyl tRNA biosynthesis. Some of them were common to tissue altered metabolic pathways (**Figure 4.16; Table 4.10).**

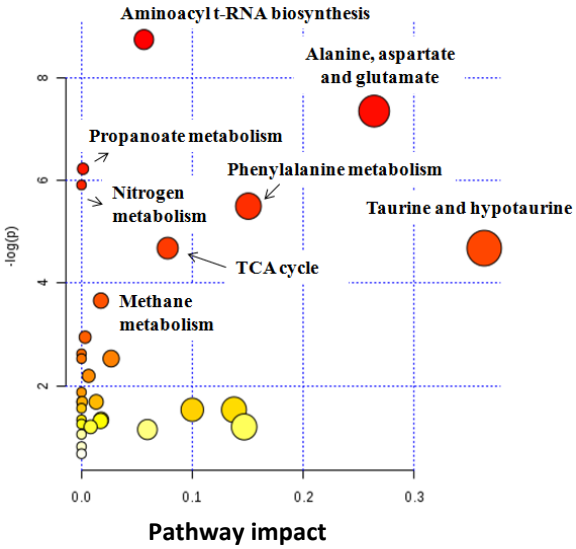


Figure 4.16 Analysis of altered metabolic pathways in bladder urines.

Note: the color and the size of each circle indicate its p-value and pathway impact value, respectively.

Table 4.10. Identified metabolites and associated metabolic pathways in urines and tissues.

Altered pathways in BC		Metabolites	p_value	Impact
Alanine, aspartate, glutamate	Tissue	Ala, Glu, Gln	6.5E-4	0.44
	Urine	Ala, Gln, succinate	6.5E-4	0.26
Taurine and hypotaurine	Tissue	Taurine, Ala	9.0E-3	0.36
	Urine	Taurine, Ala	9.0E-3	0.36
Aminoacyl-tRNA biosynthesis	Tissue	Pro, Gly, Ala, Gln, Thr,Glu	1.0E-5	0.06
	Urine	Phe, Gln, Val, Ala, Lys	1.6E-4	0.56
Methane	Tissue	Gly, methanol	2.6E-2	0.02
	Urine	Trimethylamine N-oxide, methanol	2.5E-2	0.02
Glycine, serine, threonine	Tissue	Gly, creatine, Cho, Thr	3.5E-4	0.29
Glutathione	Tissue	GSH, Gly, Glu	2.0E-3	0.25
Glycerolipid	Tissue	FAs, glycerol	2.3E-2	0.19
Arginine and proline	Tissue	Glu, Gln, Pro, creatine	2.0E-3	0.17
Glutamine and glutamate	Tissue	Glu, Gln	3.0E-3	0.14
Glycerophospholipid	Tissue	Cho, GPCho, PCho	3.0E-3	0.08
Citrate cycle (TCA cycle)	Urine	Citrate, succinate	9.0E-3	0.08
Phenylalanine	Urine	Succinate, Phe, hippuric acid	4.0E-3	0.07
Nitrogen metabolism	Urine	Phe, taurine, Gln	2.7E-3	0.05
Propanoate	Urine	Succinate, lactic acid, Val	2.0E-3	0.05

Discussion

Cancer cells undergo genetic reprogramming of their metabolism to fulfill the increased energetic and biosynthetic demands for cell proliferation. Therefore, a better understanding of altered genetic basis of the metabolic pathways and some altered metabolites could be useful to find non-invasive diagnostic, prognostic and surveillance biomarkers but also to develop new therapeutic strategies in the context of BC. With this aim, a comprehensive study focused on knowing better the relationship between the metabolome and transcriptome of BC, and its link with urinary metabolome has been here carried out.

Integrated metabolomic and transcriptomic data showed perturbations in several tissue metabolic pathways, reflecting how the activity of metabolic genes and enzymes would be regulating the demands of BC. Moreover, the upregulated and downregulated metabolism-related genes found in our series enriched in NMIBC showed similar behavior in multiple tumor samples present in the TCGA cohort (enriched in MIBC), indicating that the deregulation of these genes is probably a common feature of BC pathogenesis.

According to our data, BC cancer cells would increase the uptake of glucose and AAs from the environment, especially Glu, Gln, and Tyr. The signals of these AAs were found important in our statistical model to classify BC but also high levels of Tyr were observed in T1 and T2 tumors compared to Ta or control tissues. These data would suggest that the more invasive tumors could consume Tyr and Gln for energy generation, as a source of reducing power (NADH, $FADH_2$, NADPH), as a donor of carbon and nitrogen for generating nucleotides and other AAs, and also to maintain the pool of intermediate metabolites such as acetyl CoA and α-KG, both of them important for the anaplerotic reactions of TCA cycle, but also for epigenetic processes due to effect that they have in HAT and HDM.

Besides, we found a downregulation in all TCA cycle genes, in addition to *GLS* gene. GLS converts Gln to Glu, which is used by GPT2 to produce α-KG. These findings are in concordance with previous studies in which frequently mutated or deregulated TCA cycle genes are described in human tumors[144]. Moreover, our data support a downregulation of OXPHOS process and the overexpression of glucose metabolic genes (*SLC2A1*, *HK2,* and *RPIA*) linked to high levels of lactate, mostly in T2 BC. These data agree with the Warburg effect, which describes that tumor cells oxidize glucose to obtain energy and precursor molecules through anaerobic respiration (lactic fermentation) instead of obtaining ATP through a more efficient process in the mitochondria (i.e. OXPHOS)[42]; and show a mechanism by which tumor cells avoid the repressive feedback of TCA about glucose metabolism in presence of high levels of NADH and ATP[42].

On the other hand, they are in concordance with previous studies about BC, that showed a correlation between the expression of *SLC2A1* with disease progression and poor survival[145] and high levels of HK2 with an acceleration in glucose flux through glycolysis towards pyruvate[146] and lactic fermentation. The high levels of lactate in aggressive tumors were also observed in previous studies where the progression of BC from a less to a highly invasive stage was associated with increased production of lactate[147]. It has been reported that lactate accumulation is essential for the promotion of angiogenesis by means of VEGF from tumor-associated stromal cells, but it also contributes to tumor invasiveness due to stimulating hyaluronic acid production by fibroblasts. Tumor invasion also is enhanced through the action of matrix metalloproteinases and cathepsins that degrade the extracellular matrix. The action of these proteins is stimulated by the extracellular acidification, which is caused by the secretion of lactate into extracellular space coupled to the co-transport of proton[42].

Finally, we found *PDK4* downregulated in BC tissues and several genes encoding for pyruvate dehydrogenases (PDH). The downexpression of PDH pyruvate would be in concordance with high levels of lactate since pyruvate could be used by LDH producing lactate but would be in discordance with the downregulation of PDK4. PDK4 mediates the activity of the pyruvate dehydrogenases through inhibitory phosphorylation and has been presented as a potential regulator of BC tumorigenesis and chemoresistance[148]. Therefore, more studies in this field are necessary to knowing better the regulation of these metabolic pathways in BC.

The nexus between glucose and fatty acid metabolism is carried out by the transcriptional factor PPARG. PPARG regulates *PDK4* and is involved in the regulation of adipogenesis, cell proliferation, angiogenesis, and immune surveillance[149]. We found an overexpression of *PPARG*, as occurs in other tumors (colon, lung, prostate)[61,150] and elevated levels of lipid (b) in T2 tumors and lipid (c) in control samples. Glycerol also was identified as important in our statistical model. Glycerol and FAs are released during the degradation of TG. These data suggests a key role of lipid metabolism in BC[151,152], which could use glycerol to obtain glucose through gluconeogenesis in the liver; and FA to obtain energy (β-oxidation), produce signaling molecules, but also for the biosynthesis of PLs, the most abundant lipids of cell membranes.

Regarding choline and inositol metabolism, the signals of Cho, PCho, GPCho, CDP, and mI were relevant in our statistical model. Besides, significant high intensities of Cho, CDP, and mI were found in BC tissues, specifically in invasive and T1 tumors. These data agree with previous studies, which present the abnormal cholinic phenotype as a hallmark of tumor development and progression[153]. Cho and PCho are precursors of PtdCho, one of the main constituents of the lipid bilayers of cell membranes together with PtdE. CDP also

participates in PtdCho biosynthesis, at the same way that mI, which has been associated with PtdCho turnover and modulation of PLs. *De novo* synthesis of PtdCho and PtdE is referred as Kennedy pathway and it has been proposed as a chemotherapeutic target against cancer[154]. On whole, data obtained indicate that bladder tumors overexpress pathways that enable a higher synthesis of metabolites related with lipids that take part in cell membranes, what would be related to high rates of cell proliferation.

Polyamine metabolism is frequently deregulated in tumors[155] and is related to AA metabolism through Arg, Met, and ornithine. Arg participates in urea cycle, at the same way that Arginosuccinate Synthase 1 (ASS1). We found downexpressed *ASS1* in BC tissues, what is in agreement with previous studies that present this enzyme as a tumor suppressor. In fact, the lack of ASS1 in tumors has been related to an aggressive phenotype and negative prognostic impact in patients[51]. On the other hand, we observed a downregulation in polyamine enzymes, either biosynthetic (*ODC1, AMD1, SMS*) or catabolic (*SAT2, AOC3*), but we have not identified any metabolite as discriminant related to this pathway. However, studies performed by MS in urines and BC cell lines have found higher levels of spermidine, Arg, Met and ornithine in BC compared to control samples[133]. Polyamines can act in the regulation of gene expression through altering chromatin structure, signal transduction pathways and interactions between DNA-protein or protein-protein[156].

Maintaining the redox status is key for cancer cells. At low levels, ROS increase cell proliferation and survival but at high levels ROS cause DNA damage and trigger apoptosis[48]. Therefore, the role of antioxidant molecules is essential. In our study, BC tissues presented high levels of GSH (specifically MIBC), but a downregulated expression of other antioxidant enzymes involved in ROS detoxification. High levels of ROS have been related to apoptosis[48] and in the

context of BC, these have been linked with chemotherapy resistance and bladder tumorigenesis[157]. On the other hand, some genes of CYP450 family were found upregulated in our BC tissues (CYP2C9, CYP2J2, CYP4F11), but another downregulated (CYP1B1) like in the study performed by Putluri et al.[158]. These authors confirmed the role of methylation as an alternate mechanism in regulating CYP1A1 and CYP1B1 expression in BC tumors. CYP450-dependent monooxygenases have an essential function in the metabolism of chemical carcinogens. Polymorphisms in CYP450 genes correlate with cancer susceptibility, including BC[143]. Therefore, during carcinogenesis BC cells could enhance antioxidant and detoxifying pathways to control the redox state and to modulate the immune system with the purpose of avoiding apoptosis and enhancing proliferation.

Immunotherapy is currently an area of great interest in BC. In this sense, more insight into the connection between immune system and metabolism could be interesting to understand how BC cells modulate local levels of nutrients to alter immune cell function. Metabolites and TFs identified as important in our study such as: AAs, lactate, FA, GSH, and PPARG, would have an immunometabolic role in the BC tumor microenvironment. The relationship among extracellular lactate coming from cancer cells, tumor microenvironment and immune cell action has been widely studied. Increased lactate levels attenuate dendritic and T cells and the migration of monocyte leading to immunosuppression[42]. Gln uptake is critical for various T-cell metabolic processes (TCA cycle, nucleotide synthesis, and detoxification of ROS)[159]. Therefore, in an often hostile tumor microenvironment, Gln may be limited and could induce immunosuppression in BC. On the other hand, the role of GSH as anti-inflammatory molecule and its role in processes of innate and adaptive immune system have been also reported[58]. Finally, the function of PPARG has also been

linked to the mechanism of action by which BCG inhibits bladder tumor growth[149]. PPARG would repress T-cell effects by means of transrepression of pro-inflammatory processes that result in cytokine production. For all this, PPARG has been proposed as a promising therapeutical target for cancer treatment[160].

In silico analysis of transcriptome data indicated that metabolic genes are mostly downregulated in BC by the action of transcriptional repressors and histone marks. Two examples are GABP and E2F1, already related to bladder carcinogenesis but not with its metabolism. GABP modulates gene expression of housekeeping genes[161] but also would regulate the expression of respiratory chain genes, so it could be a key regulator of energy metabolism in BC. E2F1 has been related to progression of NMIBC to MIBC [162] but according to our results, it could regulate several metabolic pathways such as: OXPHOS, TCA cycle, pyruvate, polyamine, and AA metabolism. Particular active and repressive transcription histone marks were also associated with downregulated metabolism-related genes, indicating that epigenetic processes are important in BC regulation. On the other hand, the role of disturbed AS promoting pro-tumorigenic isoforms of metabolic enzymes[62] should be considered. We found a significant AS in several metabolic genes in BC tissues that may contribute to the observed metabolic changes. Nevertheless, the impact of these isoforms on tumor progression and cancer metabolism has not well studied. Only the role of AS in the gene *OGDH* has been reported in colorectal cancer, where an upregulation of *OGDH* alternative mRNA transcript has linked with Gln metabolism[163] and an increase in energy.

Regarding urines as a non-invasive alternative to detect BC, our study revealed the presence of several common altered pathways in BC mainly related with the metabolism of taurine and AA, and also showed their own urinary metabolic pathways, different to those found in the tissue study. The metabolization of some tissue metabolites in other derivatives and a large

number of substances present in urine coming from other sources could explain these results. Even so, our data reinforce the idea that some tumor metabolic alterations (disturbed metabolic pathways for Ala, Asp and Glu metabolism, taurine and hypotaurine, TCA cycle and aminoacyl t-RNA biosynthesis) are reflected in the urinary metabolome (changes in Val, methylsuccinic acid, lactate, Ala, Lys, N-acetylneuraminic acid, Gln, succinic acid, citrate, creatinine, trimethylamine-N-oxide, methanol, taurine, sucrose, hippuric acid, Phe, pseudouridine and trigonelline), suggesting that urine can be used to find a non-invasive approach for BC diagnosis and follow-up. Although our tissue metabolic profile distinguished BC from non-tumor tissues even at early pathologic stage (Ta) with a sensibility and specificity of 100%, our metabolic urinary signature also showed a high sensitivity (91%) and a good specificity (77%) classifying BC and control urines. Considering the limitations of cystoscopy (being expensive, operator-dependent, invasive, and overlooking CIS) and the urinary cytology (low sensitivity in LG tumors) the metabolic urinary signature could be used to detect and monitor the dynamic changes in the disease, and for diagnosis and maybe follow-up of BC patients, reducing the number of cystoscopies in NMIBC patients and identifying early tumor development or recurrences.

Conclusions

In summary, our integrated metabolomic and transcriptomic study indicated that metabolic reprogramming in BC is produced mainly through the downregulation of metabolic genes related with TCA cycle, OXPHOS, and AAs metabolic pathways. These tumor tissue metabolic alterations are reflected in the urinary metabolome of the same patient and provide a specific NMR-metabolic profile able to detect BC with a significant sensitivity and specificity from a non-invasive approach.

Annexed tables

Table 4.3 Significant metabolites in bladder tissues using U-Mann-Whitney test between tumor and non-tumor tissues (Table A); ANOVA test among NMIBC, MIBC and CTRL tissues (Table B); and ANOVA test among Ta, T1, T2 and CTRL tissues (Table C).

Table A

Mol. Fmla.	Metabolites	HMDB	p-val	T/NT
--	UDP-sugars	--	<0.05	>1
$C_9H_{15}N_3O_{11}P_2$	CDP	HMDB01546	<0.05	>1
$C_{10}H_{17}N_3O_6S$	GSH	HMDB00125	<0.05	>1
$C_5H_{14}NO$	Cho	HMDB00097	<0.05	>1
$C_6H_{12}O_6$	mI	HMDB00211	<0.05	>1
$-(CH_2)_n-CH_2-CH_2-COOH$	Lipid (c)	--	<0.05	<1

Table B

Mol. Fmla.	Metabolites	HMDB	p-val	Tukey's Post-hoc tests
--	UDP-sugars	--	<0.05	MIBC:CTRL NMIBC:CTRL
$C_9H_{15}N_3O_{11}P_2$	CDP	HMDB01546	<0.05	MIBC:CTRL NMIBC:CTRL
$C_{10}H_{17}N_3O_6S$	GSH	HMDB00125	<0.05	MIBC:CTRL NMIBC:CTRL
$C_5H_{14}NO$	Cho	HMDB00097	<0.05	MIBC:CTRL NMIBC:CTRL
$C_6H_{12}O_6$	mI	HMDB00211	<0.05	NMIBC:CTRL

Table C

Mol. Fmla.	Metabolites	HMDB	p-val	Tukey's Post-hoc tests
$C_5H_{14}NO$	Cho	HMDB0097	<0.05	T2:T1;T2:Ta;T1:Ta;T2:CTRL;T1:CTRL;Ta:CTRL
$C_9H_{11}NO_3$	L-Tyr	HMDB00158	<0.05	T2:T1;T2:Ta;T1:Ta;T2:CTRL;T1:CTRL;Ta:CTRL
$C_6H_{12}O_6$	mI	HMDB00211	<0.05	T2:T1 T2:Ta T1:Ta T2:CTRL T1:CTRL Ta:CTRL
$C_{10}H_{17}N_3O_6S$	GSH	HMDB00125	<0.05	T2:T1 T2:Ta T1:Ta T2:CTRL T1:CTRL Ta:CTRL
$-(n)CH_2\text{-}CH_2\text{-}CH_2\text{-}CH_3$	Lipid (a)	--	<0.05	T2:T1 T2:Ta T1:Ta T2:CTRL T1:CTRL Ta:CTRL
---	UDP-sugars	--	<0.05	T2:T1 T2:Ta T1:Ta T2:CTRL T1:CTRL Ta:CTRL
$C_9H_{15}N_3O_{11}P_2$	CDP	HMDB01546	<0.05	T2:T1 T2:Ta T1:Ta T2:CTRL T1:CTRL Ta:CTRL
$-CH{=}CH\text{-}CH_2\text{-}CH{=}CH\text{-}CH_2\text{-}CH_2\text{-}(CH_2)_n\text{-}$	Lipid (b)	--	<0.05	T2:T1 T2:Ta T1:Ta T2:CTRL T1:CTRL Ta:CTRL

Table 4.6 Metabolic genes significantly altered in bladder tumors.

Gene	Description	Regulation	FC	p-val	FDR p-val
AASS	aminoadipate-semialdehyde synthase	DOWN	-2.55	2.00E-03	8.90E-03
ABL1	ABL proto-oncogene 1, non-receptor tyrosinekinase	DOWN	-5.80	6.63E-06	2.00E-04
ACAA2	acetyl-CoA acyltransferase 2	DOWN	-5.13	2.10E-06	6.80E-05
ACACB	acetyl-CoA carboxylase beta	DOWN	-8.43	5.27E-07	2.36E-05
ACADM	acyl-CoA dehydrogenase, C-4 to C-12 straightchain	DOWN	-2.31	1.95E-07	1.12E-05
ACAT1	acetyl-CoA acetyltransferase 1	DOWN	-2.60	4.68E-06	1.00E-04
ACO1	aconitase 1, soluble	DOWN	-4.27	1.38E-08	1.54E-06
ACSL3	acyl-CoA synthetase long-chain family member 3	DOWN	-2.09	8.00E-04	4.70E-03
ADARB1	adenosinedeaminase, RNA-specific, B1	DOWN	-14.27	1.67E-07	9.90E-06
ADH1B	alcohol dehydrogenase 1B (class I), beta polypeptide	DOWN	-97.23	4.94E-06	1.00E-04
ADH1C	alcohol dehydrogenase 1C (class I), gamma polypeptide	DOWN	-3.84	7.78E-06	2.00E-04
ADH5	alcohol dehydrogenase 5 (class III), chi polypeptide	DOWN	-3.58	1.65E-10	9.70E-08
AK2	adenylatekinase 2	DOWN	-2.40	1.00E-04	1.30E-03
AK4	adenylatekinase 4	DOWN	-12.83	1.25E-06	4.61E-05
ALDH1B1	Aldehyde dehydrogenase 1 family, member B1	DOWN	-7.33	3.33E-08	2.85E-06
ALDH2	Aldehyde dehydrogenase 2 family (mitochondrial)	DOWN	-5.62	1.19E-07	7.65E-06
ALDH6A1	Aldehyde dehydrogenase 6 family, member A1	DOWN	-2.39	2.47E-07	1.32E-05
ALDH9A1	Aldehyde dehydrogenase 9 family, member A1	DOWN	-2.04	1.31E-09	3.21E-07
AMD1	Adenosylmethionine decarboxylase 1	DOWN	-2.13	4.00E-04	3.10E-03
AOC3	amineoxidase, copper containing 3	DOWN	-6.48	7.01E-07	2.98E-05
APLP2	amyloid beta (A4) precursor-likeprotein 2	DOWN	-5.67	9.07E-05	1.10E-03
APP	amyloid beta (A4) precursor protein	DOWN	-4.43	7.61E-07	3.15E-05
ASS1	Argininosuccinate synthase 1	DOWN	-2.07	2.23E-05	4.00E-04
ATIC	5-aminoimidazole-4-carboxamide ribonucleotideformyl transferase/IMP cyclohydrolase	DOWN	-2.48	4.57E-06	1.00E-04

ATP5A1	ATP synthase, H+ transporting, mitochondrial F1 complex, alphasubunit 1, cardiac muscle	DOWN	-2.18	1.84E-05	3.00E-04
ATP5B	ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide	DOWN	-9.72	4.72E-09	7.48E-07
ATP5C1	ATP synthase, H+ transporting, mitochondrial F1 complex, gamma polypeptide 1	DOWN	-3.83	2.29E-07	1.26E-05
ATP5D	ATP synthase, H+ transporting, mitochondrial F1 complex, delta subunit	DOWN	-2.95	2.39E-09	4.62E-07
ATP5E	ATP synthase, H+ transporting, mitochondrial F1 complex, epsilonsubunit	DOWN	-4.53	1.55E-06	5.43E-05
ATP5F1	ATP synthase, H+ transporting, mitochondrial Fo complex subunit B1	DOWN	-2.05	3.19E-08	2.77E-06
ATP5G2	ATP synthase, H+ transporting, mitochondrial Fo complex subunit C2 (subunit 9)	DOWN	-41.60	4.29E-10	1.63E-07
ATP5G3	ATP synthase, H+ transporting, mitochondrial Fo complex subunit C3 (subunit 9)	DOWN	-19.75	1.29E-08	1.48E-06
ATP5H	ATP synthase, H+ transporting, mitochondrial Fo complex subunit D	DOWN	-4.29	7.01E-06	2.00E-04
ATP5I	ATP synthase, H+ transporting, mitochondrial Fo complex subunit E	DOWN	-10.48	3.33E-08	2.85E-06
ATP5J	ATP synthase, H+ transporting, mitochondrial Fo complex subunit F6	DOWN	-2.06	1.20E-08	1.40E-06
ATP5L	ATP synthase, H+ transporting, mitochondrial Fo complex subunit G	DOWN	-2.31	1.00E-04	1.30E-03
ATP5O	ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit	DOWN	-17.06	2.97E-08	2.62E-06
ATP6AP1	ATPase, H+ transporting, lysosomalaccessoryprotein 1	DOWN	-6.97	6.39E-05	8.00E-04
ATP6V0A1	ATPase, H+ transporting, lysosomal V0 subunit a1	DOWN	-2.78	2.80E-05	5.00E-04
ATP6V0E1	ATPase, H+ transporting, lysosomal 9kDa, V0 subunit e1	DOWN	-6.95	4.73E-08	3.71E-06
ATP6V1E1	ATPase, H+ transporting, lysosomal 31kDa, V1 subunit E1	DOWN	-2.32	1.82E-07	1.06E-05
ATP6V1F	ATPase, H+ transporting, lysosomal 14kDa, V1 subunit F	DOWN	-3.98	3.16E-06	9.25E-05
ATP8B2	ATPase, aminophospholipid	DOWN	-2.10	5.04E-05	7.00E-04

	transporter, class I, type 8B, member 2				
AUH	AU RNA binding protein/enoyl-CoA hydratase	DOWN	-2.39	7.82E-05	1.00E-03
B2M	beta-2-microglobulin	DOWN	-2.36	3.30E-02	6.82E-02
BACE1	beta-site APP-cleavingenzyme 1	DOWN	-2.58	5.00E-04	3.30E-03
BCHE	butyrylcholinesterase	DOWN	-3.66	7.93E-06	2.00E-04
C3	complement component 3	DOWN	-2.35	6.20E-03	1.94E-02
C4A	complement component 4A (Rodgersbloodgroup)	DOWN	-4.18	1.30E-03	6.50E-03
CALM2	calmodulin 2 (phosphorylasekinase, delta)	DOWN	-9.27	1.25E-10	8.38E-08
CALM3,2,1	calmodulin 3); calmodulin 1); calmodulin 2	DOWN	-7.15	2.87E-09	5.23E-07
CALU	calumenin	DOWN	-5.32	2.83E-07	1.47E-05
CAT	catalase	DOWN	-3.44	1.88E-05	4.00E-04
CDC42BPA	CDC42 binding proteinkinase alpha (DMPK-like)	DOWN	-5.89	3.74E-09	6.25E-07
CHPT1	cholinephosphotransferase 1	DOWN	-22.00	7.78E-11	6.49E-08
CHRDL1	chordin-like 1	DOWN	-3.73	1.00E-04	1.40E-03
CHRM2	cholinergic receptor, muscarinic 2	DOWN	-4.70	3.00E-06	8.90E-05
CHRM3	cholinergic receptor, muscarinic 3	DOWN	-12.79	1.47E-06	5.21E-05
CIT; MIR1178	citron rho-interacting serine/threonine kinase; microRNA 1178	UP	2.45	2.70E-05	5.00E-04
CKB	creatinekinase, brain	DOWN	-2.95	4.32E-06	1.00E-04
CKMT1A; CKMT1B	creatinekinase, mitochondrial 1A; creatinekinase, mitochondrial 1B	UP	2.77	8.00E-04	4.60E-03
COA5	cytochrome c oxidase assembly factor 5	DOWN	-2.09	6.00E-04	4.00E-03
COX10	COX10 heme A:farnesyltransferase cytochrome c oxidase assembly factor	DOWN	-2.57	5.22E-07	2.35E-05
COX11	COX11 cytochrome c oxidase copper chaperone	DOWN	-5.49	6.72E-09	9.62E-07
COX15	cytochrome c oxidase assembly homolog 15 (yeast)	DOWN	-2.67	3.25E-08	2.81E-06
COX20	COX20 cytochrome c oxidase assembly factor	DOWN	-4.40	1.45E-08	1.58E-06
COX4I1	cytochrome c oxidase subunit IV isoform 1	DOWN	-4.84	3.14E-11	5.04E-08
COX5A	cytochrome c oxidase subunit Va	DOWN	-3.13	1.50E-07	9.14E-06
COX5B	cytochrome c oxidase subunitVb	DOWN	-5.14	1.28E-07	8.10E-06
COX6A1	cytochrome c oxidase subunit VI	DOWN	-4.01	2.00E-04	2.00E-03

	apolypeptide 1				
COX6C	cytochrome c oxidase subunit VIc	DOWN	-2.31	1.51E-08	1.63E-06
COX7A1	cytochrome c oxidase subunit VII apolypeptide 1 (muscle)	DOWN	-8.45	3.59E-10	1.52E-07
COX7A2	cytochrome c oxidase subunit VII apolypeptide 2 (liver)	DOWN	-5.00	2.93E-08	2.61E-06
COX7B	cytochrome c oxidase subunit VIIb	DOWN	-4.85	1.94E-07	1.12E-05
COX7C; MIR3607	cytochrome c oxidase subunit VIIc; microRNA 3607	DOWN	-3.74	9.68E-10	2.71E-07
COX8A	cytochrome c oxidase subunit VIIIA (ubiquitous)	DOWN	-82.96	5.73E-11	5.92E-08
CRYM	crystallinmu	DOWN	-2.56	8.16E-07	3.33E-05
CSRP2	Cysteine and glycine-richprotein 2	DOWN	-8.27	4.03E-08	3.32E-06
CST3	cystatin C	DOWN	-2.44	2.07E-07	1.16E-05
CYCS	cytochrome c, somatic	DOWN	-2.44	1.20E-03	6.10E-03
CYP1B1	cytochrome P450, family 1, subfamily B, polypeptide 1	DOWN	-2.10	4.10E-03	1.45E-02
CYP2C9	cytochrome P450, family 2, subfamily C, polypeptide 9	UP	2.18	2.00E-04	1.60E-03
CYP2J2	cytochrome P450, family 2, subfamily J, polypeptide 2	UP	7.98	1.59E-02	3.85E-02
CYP4F11	cytochrome P450, family 4, subfamily F, polypeptide 11	UP	2.03	2.52E-02	5.44E-02
CYP4V2	cytochrome P450, family 4, subfamily V, polypeptide 2	DOWN	-2.68	4.61E-08	3.65E-06
CYR61	cysteine-rich, angiogenicinducer, 61	DOWN	-4.12	1.27E-02	3.25E-02
DARS	aspartyl-tRNA synthetase	DOWN	-2.27	1.18E-05	3.00E-04
DGKB	diacylglycerolkinase, beta 90kDa	DOWN	-2.95	3.00E-04	2.30E-03
DGKG	diacylglycerolkinase gamma	DOWN	-3.82	1.65E-06	5.69E-05
DGKH	diacylglycerolkinase, eta	UP	-2.12	3.14E-02	6.45E-02
DLAT	dihydrolipoamide S-acetyltransferase	DOWN	-2.03	1.01E-06	3.92E-05
DLD	Dihydrolipoamide dehydrogenase	DOWN	-3.22	1.31E-05	3.00E-04
DNAJC3	DnaJ (Hsp40) homolog, subfamily C, member 3	DOWN	-11.87	1.51E-07	9.20E-06
DPYD	Dihydropyrimidine dehydrogenase	DOWN	-2.67	2.00E-04	1.90E-03
ECH1	enoyl-CoAhydratase 1, peroxisomal	DOWN	-5.42	6.34E-07	2.74E-05
ECI2	enoyl-CoA delta isomerase 2	DOWN	-3.14	1.12E-08	1.34E-06
EEF1A1	Eukaryotic translation elongation factor 1 alpha 1	DOWN	-4.04	4.33E-07	2.03E-05
EEF1G	Eukaryotic translation elongation factor	DOWN	-4.68	1.95E-05	4.00E-04

	1 gamma; microRNA 3654				
EIF4EBP2	Eukaryotic translation initiation factor 4E bindingprotein 2	DOWN	-2.19	8.25E-05	1.00E-03
ENTPD1	Ectonucleoside triphosphated iphosphohydrolase 1	DOWN	-9.40	6.96E-09	9.80E-07
ESD	esterase D	DOWN	-8.14	4.76E-10	1.73E-07
FARSB	phenylalanyl-tRNA synthetase beta subunit	DOWN	-2.91	5.27E-07	2.36E-05
FAXDC2	Fatty acid hydroxylase domaincontaining 2	DOWN	-5.15	3.24E-12	1.37E-08
FBN1	fibrillin 1	DOWN	-3.44	1.00E-04	1.40E-03
FH	fumaratehydratase	DOWN	-2.79	4.34E-09	7.01E-07
FN1	fibronectin 1	DOWN	-55.97	7.47E-08	5.30E-06
FNBP1	formin bindingprotein 1	DOWN	-11.83	1.62E-09	3.68E-07
FSTL1	Follistatinlike 1; microRNA 198	DOWN	-3.21	3.89E-06	1.00E-04
GAPDH	glyceraldehyde-3-phosphate dehydrogenase	DOWN	-3.10	1.32E-05	3.00E-04
GAS6	growth arrest-specific 6	DOWN	-2.80	7.92E-08	5.56E-06
GATA6	GATA bindingprotein 6	DOWN	-3.50	1.20E-05	3.00E-04
GATM	Glycineamidino transferase (L-arginine:glycineamidino transferase)	DOWN	-6.37	1.00E-04	1.30E-03
GLRX3	glutaredoxin 3	DOWN	-2.09	9.45E-05	1.10E-03
GLS	glutaminase	DOWN	-2.08	1.03E-06	3.94E-05
GMPS	Guaninemonophosphate synthase	DOWN	-2.67	2.48E-05	4.00E-04
GNPDA2	glucosamine-6-phosphate deaminase 2	DOWN	-2.41	9.93E-07	3.86E-05
GOT2	glutamic-oxaloacetic transaminase 2, mitochondrial	DOWN	-3.15	5.58E-06	1.00E-04
GPX3	glutathioneperoxidase 3	DOWN	-2.77	7.20E-07	3.03E-05
GPX4	glutathioneperoxidase 4	DOWN	-2.83	2.77E-08	2.50E-06
GSN	gelsolin	DOWN	-3.52	2.88E-07	1.49E-05
GSTA4	glutathione S-transferasealpha 4	DOWN	-2.60	8.28E-07	3.37E-05
GSTO1	glutathione S-transferase omega 1	DOWN	-2.20	2.32E-08	2.19E-06
GSTO2	glutathione S-transferase omega 2	UP	2.38	5.20E-03	1.71E-02
GSTT1	glutathione S-transferase theta 1	DOWN	-3.00	2.20E-03	9.50E-03
H3F3A	H3 histone, family 3A	DOWN	-3.60	1.00E-04	1.20E-03
HADHA	hydroxyacyl-CoAdehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoAhydratase , alphasubunit	DOWN	-2.46	1.78E-05	3.00E-04
HADHB	hydroxyacyl-CoAdehydrogenase/3-	DOWN	-2.49	3.00E-04	2.60E-03

	ketoacyl-CoA thiolase/enoyl-CoAhydratase , beta subunit				
HIBADH	3-hydroxyisobutyrate dehydrogenase	DOWN	-2.89	1.52E-07	9.23E-06
HK2	hexokinase 2	UP	2.25	2.57E-05	4.00E-04
IDH2	isocitratedehydrogenase 2 (NADP+), mitochondrial	DOWN	-2.26	5.00E-04	3.50E-03
IDH3A	isocitratedehydrogenase 3 (NAD+) alpha	DOWN	-2.20	3.44E-06	9.89E-05
IDH3B	isocitratedehydrogenase 3 (NAD+) beta	DOWN	-2.53	6.93E-05	9.00E-04
IGF1	insulin-likegrowth factor 1 receptor	DOWN	-2.20	3.20E-03	1.22E-02
IGFBP5	Insulinlike growth factor binding protein 5	DOWN	-12.84	2.06E-06	6.70E-05
IGFBP6	Insulin like growth factor binding protein 6	DOWN	-2.28	5.61E-07	2.48E-05
IGFBP7	Insulin like growth factor binding protein 7	DOWN	-20.02	2.26E-10	1.14E-07
IMPA2	inositol(myo)-1(or 4)-monophosphatase 2	DOWN	-2.62	5.41E-06	1.00E-04
INPP4B	inositol polyphosphate-4-phosphatase type II B	UP	2.11	6.00E-03	1.90E-02
IP6K2	inositol hexakisphosphatekinase 2	UP	2.37	2.33E-02	5.14E-02
ITM2B	integral membraneprotein 2B	DOWN	-7.06	6.36E-07	2.74E-05
ITPK1	inositol-tetrakisphosphate 1-kinase	DOWN	-2.49	1.74E-06	5.93E-05
ITPKB	inositol-trisphosphate 3-kinase B	DOWN	-5.52	1.79E-06	6.05E-05
ITPR1	inositol 1,4,5-trisphosphate receptor, type 1	DOWN	-5.27	4.35E-05	6.00E-04
ITPR3	inositol 1,4,5-trisphosphate receptor, type 3	UP	2.31	3.00E-04	2.60E-03
ITPRIPL2	inositol 1,4,5-trisphosphate receptor interactingprotein-like 2	DOWN	-2.01	1.30E-03	6.50E-03
KDSR	3-ketodihydrosphingosine reductase	DOWN	-13.58	7.27E-09	1.00E-06
LAMB2	laminin, beta 2 (laminin S)	DOWN	-2.83	1.22E-06	4.54E-05
LAMC1	laminin, gamma 1 (formerly LAMB2)	DOWN	-2.39	1.15E-05	2.00E-04
LGALS1	lectin, galactoside-binding, soluble, 1	DOWN	-20.98	1.03E-07	6.85E-06
LMCD1	LIM andcysteine-richdomains 1	DOWN	-3.98	1.50E-08	1.62E-06
LTBP1	latent transforminggrowth factor beta bindingprotein 1	DOWN	-8.63	1.89E-09	4.02E-07
MAOA	monoamineoxidase A	DOWN	-3.26	2.94E-06	8.79E-05
MAOB	monoamineoxidase B	DOWN	-9.60	4.75E-10	1.73E-07
MDH1	malatedehydrogenase 1	DOWN	-2.44	1.11E-07	7.24E-06

MDH2	malatedehydrogenase 2	DOWN	-2.42	2.12E-07	1.18E-05
MFGE8	milk fat globule-EGF factor 8 protein	DOWN	-2.07	2.68E-05	5.00E-04
MGST1	microsomalglutathione S-transferase 1	UP	5.90	2.50E-03	1.01E-02
MGST3	microsomalglutathione S-transferase 3	DOWN	-2.46	1.78E-06	6.03E-05
MIA3	melanoma inhibitoryactivityfamily, member 3	DOWN	-2.29	3.87E-07	1.86E-05
MMP2	matrixmetallopeptidase 2	DOWN	-3.14	2.00E-04	1.50E-03
MPC1	Mitochondrial pyruvate carrier 1	DOWN	-3.19	4.94E-11	5.64E-08
MPC2	Mitochondria lpyruvate carrier 2	DOWN	-2.08	5.12E-07	2.31E-05
MRPL11	Mitochondrial ribosomal protein L11	DOWN	-5.02	2.31E-05	4.00E-04
MRPL24	Mitochondrial ribosomal protein L24	DOWN	-2.52	9.58E-09	1.21E-06
MRPL3	Mitochondria lribosomal protein L3	DOWN	-2.60	8.28E-05	1.00E-03
MRPL35	Mitochondrial ribosomal protein L35	DOWN	-3.08	2.85E-05	5.00E-04
MRPL42	Mitochondrial ribosomal protein L42	DOWN	-2.56	4.97E-05	7.00E-04
MRPS14	Mitochondrial ribosomal protein S14	DOWN	-6.50	4.00E-10	1.59E-07
MRPS18A	Mitochondrial ribosomal protein S18A	DOWN	-2.40	2.00E-04	2.00E-03
MRPS18C	Mitochondrial ribosomal protein S18C	DOWN	-2.20	6.90E-06	2.00E-04
MRPS21	Mitochondrial ribosomal protein S21	DOWN	-3.18	1.95E-05	4.00E-04
MRPS36	Mitochondrial ribosomal protein S36	DOWN	-2.07	2.81E-05	5.00E-04
MRPS7	Mitochondrial ribosomal protein S7	DOWN	-3.55	5.58E-06	1.00E-04
MSRB1	Methioninesulfoxide reductase B1	DOWN	-2.42	1.34E-07	8.40E-06
MSRB2	Methioninesulfoxide reductase B2	DOWN	-4.80	3.39E-10	1.50E-07
MSRB3	Methioninesulfoxide reductase B3	DOWN	-5.39	6.34E-08	4.69E-06
MTR	5-methyltetrahydrofolate-homocysteine methyltransferase	DOWN	-2.39	3.00E-04	2.70E-03
MUT	methylmalonyl-CoAmutase	DOWN	-2.30	4.24E-06	1.00E-04
NDUFA1	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 1, 7.5kDa	DOWN	2.38	2.69E-02	5.73E-02
NDUFA11	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 11, 14.7kDa	DOWN	-6.23	1.32E-10	8.68E-08
NDUFA12	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 12	DOWN	-5.33	3.40E-09	5.85E-07
NDUFA2	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 2, 8kDa	DOWN	-5.29	6.40E-06	2.00E-04
NDUFA4	NDUFA4, mitochondrial complex associated	DOWN	-18.25	1.07E-08	1.30E-06
NDUFA6	NADH dehydrogenase (ubiquinone) 1 alphasubcomplex, 6, 14kDa	DOWN	-7.61	7.74E-11	6.49E-08

NDUFA9	NADH dehydrogenase (ubiquinone) 1 alphasubcomplex, 9, 39kDa	DOWN	-2.03	2.00E-04	1.80E-03
NDUFAB1	NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1, 8kDa	DOWN	-9.44	1.48E-07	9.05E-06
NDUFAF1	NADH dehydrogenase (ubiquinone) complex I, assembly factor 1	DOWN	-2.07	4.11E-07	1.95E-05
NDUFAF2	NADH dehydrogenase (ubiquinone) complex I, assembly factor 2	DOWN	-10.19	6.06E-11	5.92E-08
NDUFAF2	NADH dehydrogenase (ubiquinone) complex I, assembly factor 2	DOWN	-4.36	1.39E-08	1.55E-06
NDUFAF4	NADH dehydrogenase (ubiquinone) complex I, assembly factor 4	DOWN	-2.40	1.47E-05	3.00E-04
NDUFB10	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 10, 22kDa	DOWN	-3.46	2.16E-08	2.08E-06
NDUFB3	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 3, 12kDa	DOWN	-2.81	2.93E-06	8.77E-05
NDUFB5	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 5, 16kDa	DOWN	-3.40	2.38E-06	7.47E-05
NDUFB6	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 6, 17kDa	DOWN	-5.17	4.87E-07	2.22E-05
NDUFB7	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 7, 18kDa	DOWN	-4.61	2.20E-06	7.03E-05
NDUFB9	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 9, 22kDa	DOWN	-2.73	1.98E-06	6.53E-05
NDUFS3	NADH dehydrogenase (ubiquinone) Fe-S protein 3, 30kDa (NADH-coenzyme Q reductase)	DOWN	-2.35	5.04E-07	2.28E-05
NDUFS4	NADH dehydrogenase (ubiquinone) Fe-S protein 4, 18kDa (NADH-coenzyme Q reductase)	DOWN	-3.95	8.79E-08	6.03E-06
NDUFS5	NADH dehydrogenase (ubiquinone) Fe-S protein 5, 15kDa (NADH-coenzyme Q reductase)	DOWN	-12.64	4.16E-09	6.78E-07
NDUFV2	NADH dehydrogenase (ubiquinone) flavoprotein 2, 24kDa	DOWN	-2.15	2.13E-06	6.88E-05
NEDD8	neural precursor cell expressed, developmentallydown-regulated 8	DOWN	-2.05	2.71E-06	8.26E-05
NME4	NME/NM23 nucleoside diphosphate kinase 4	DOWN	-2.09	5.97E-09	8.79E-07
OAT	Ornithine aminotransferase	DOWN	-2.47	2.92E-05	5.00E-04
OAZ1	Ornithine decarboxylase antizyme 1	DOWN	-3.37	2.05E-06	6.68E-05
ODC1	Ornithine decarboxylase 1	DOWN	-2.21	1.50E-03	7.20E-03
OGDH	oxoglutarate (alpha-ketoglutarate)	DOWN	-2.97	1.99E-07	1.13E-05

	dehydrogenase (lipoamide)				
OXCT1	3-oxoacid CoA-transferase 1	DOWN	-2.22	8.00E-04	4.90E-03
PAM	peptidylglycinealpha-amidatingmonooxygenase	DOWN	-4.59	1.22E-06	4.54E-05
PCMT1	protein-L-isoaspartate (D-aspartate) O-methyltransferase	DOWN	-2.51	2.46E-07	1.32E-05
PDE10A	phosphodiesterase 10A	UP	2.38	1.86E-02	4.34E-02
PDE1C	phosphodiesterase 1C, calmodulin-dependent 70kDa	DOWN	-5.91	1.24E-06	4.58E-05
PDGFC	Platelet derived growth factor C	DOWN	-2.13	1.64E-05	3.00E-04
PDHA1	Pyruvate dehydrogenase (lipoamide) alpha 1	DOWN	-2.52	2.47E-08	2.27E-06
PDHB	Pyruvate dehydrogenase (lipoamide) beta	DOWN	-3.46	4.74E-08	3.71E-06
PDHX	Pyruvate dehydrogenase complex, component X	DOWN	-2.13	4.69E-08	3.70E-06
PDK4	Pyruvate dehydrogenasekinase, isozyme 4	DOWN	-11.76	9.57E-07	3.75E-05
PELO	pelotahomolog (Drosophila); integrin alpha 1	DOWN	-2.50	2.41E-07	1.30E-05
PFAS	Phosphoribosyl formylglycinamidine synthase	UP	2.09	1.01E-05	2.00E-04
PGAM1	Phosphoglycerate mutase 1 (brain)	DOWN	-3.62	1.53E-06	5.36E-05
PGM2L1	Phosphogluco mutase 2-like 1	DOWN	-2.73	4.74E-05	7.00E-04
PGM5	phosphoglucomutase 5	DOWN	-28.42	1.63E-09	3.69E-07
PHKB	Phosphorylasekinase, beta	DOWN	-2.77	1.63E-06	5.65E-05
PICALM	Phosphatidyl inositol binding clathrin assembly protein	DOWN	-2.63	5.67E-07	2.50E-05
PIGP	Phosphatidyl inositol glycan anchor biosynthesis class P	DOWN	-4.09	3.61E-07	1.76E-05
PIGT	Phosphatidyl inositol glycan anchor biosynthesis class T	DOWN	-2.26	8.46E-05	1.00E-03
PIGY; PYURF	Phosphatidyl inositol glycan anchor biosynthesis class Y; PIGY upstream reading frame	DOWN	-2.10	4.06E-06	1.00E-04
PIK3C2A	phosphatidylinositol-4-phosphate 3-kinase, catalytic subunittype 2 alpha	DOWN	-2.32	3.07E-06	9.07E-05
PIK3CA	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunitalpha	DOWN	-4.42	1.35E-08	1.51E-06
PIP4K2A	Phosphatidylinositol-5-phosphate 4-kinase, type II, alpha	DOWN	-4.14	2.12E-11	4.33E-08
PITPNC1	Phosphatidyl inositol transfer protein,	DOWN	-2.44	8.00E-03	2.33E-02

	cytoplasmic 1				
PLA2G12A	phospholipase A2, group XIIA	DOWN	-2.08	2.21E-05	4.00E-04
PLA2G16	phospholipase A2, group XVI	DOWN	-2.08	7.30E-03	2.18E-02
PLA2G2F	phospholipase A2, group IIF	UP	2.18	1.47E-02	3.62E-02
PLA2G4A	phospholipase A2, group IVA (cytosolic, calcium-dependent)	DOWN	-3.23	2.11E-07	1.17E-05
PLB1	phospholipase B1	UP	2.29	8.10E-03	2.34E-02
PLCB4	phospholipase C, beta 4	DOWN	-7.50	7.80E-09	1.06E-06
PLCD4	phospholipase C, delta 4	DOWN	-2.42	8.40E-05	1.00E-03
PLCL1	phospholipase C-like 1	DOWN	-2.23	6.32E-06	2.00E-04
PLOD2	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2	DOWN	-2.11	2.64E-02	5.65E-02
PLPP1	phospholipidphosphatase 1	DOWN	-3.81	5.14E-05	7.00E-04
PLPP3	phospholipidphosphatase 3	DOWN	-4.09	4.72E-07	2.17E-05
PLSCR4	phospholipidscramblase 4	DOWN	-4.43	1.00E-08	1.24E-06
PMPCB	peptidase (mitochondrial processing) beta	DOWN	-3.52	1.18E-08	1.39E-06
PPA1	pyrophosphatase (inorganic) 1	DOWN	-2.77	3.00E-04	2.40E-03
PPARG	Peroxisome proliferator-activated receptor gamma	UP	6.47	6.72E+00	4.50E-03
PPM1K	Protein phosphatase, Mg2+/Mn2+ dependent, 1K	DOWN	-8.62	2.92E-09	5.26E-07
PPM1L	Protein phosphatase, Mg2+/Mn2+ dependent, 1L	DOWN	-17.00	2.36E-09	4.60E-07
PPP2CA	Protein phosphatase 2, catalytic subunit, alphaisozyme	DOWN	-2.27	2.95E-06	8.81E-05
PPP2CB	Protein phosphatase 2, catalytic subunit, beta isozyme	DOWN	-3.28	6.33E-09	9.17E-07
PPP2R1A	Protein phosphatase 2, regulatory subunit A, alpha	DOWN	-3.06	5.00E-04	3.50E-03
PRKAA2	Protein kinase, AMP-activated, alpha 2 catalytic subunit	DOWN	-19.76	3.28E-07	1.64E-05
PRKACB	Protein kinase, cAMP-dependent, catalytic, beta	DOWN	-2.08	5.23E-07	2.35E-05
PRKCA	Protein kinase C, alpha	DOWN	-5.10	2.00E-03	8.60E-03
PRKCB	Protein kinase C, beta	DOWN	-2.31	8.20E-06	2.00E-04
PRPS1	Phosphoribosyl pyrophosphate synthetase 1	DOWN	-2.59	3.84E-08	3.21E-06
PRSS23	protease, serine, 23	DOWN	-2.46	2.00E-04	1.60E-03
PTEN	Phosphatase and tensin homolog	DOWN	-2.12	7.24E-06	2.00E-04

PWP1	PWP1 homolog, endonuclein	DOWN	-2.09	2.73E-07	1.43E-05
PYGB	phosphorylase, glycogen; brain	DOWN	-33.99	1.55E-08	1.65E-06
PYGM	phosphorylase, glycogen, muscle	DOWN	-8.47	1.42E-07	8.78E-06
RARS	arginyl-tRNA synthetase	DOWN	-2.00	1.00E-04	1.40E-03
RPIA	ribose 5-phosphate isomerase A	UP	2.95	3.00E-04	2.20E-03
RPL12	Ribosomal protein L12	DOWN	-3.44	1.71E-09	3.82E-07
RPL13	Ribosomal protein L13	DOWN	-6.57	2.22E-10	1.14E-07
RPL13A	Ribosomal protein L13a	DOWN	-22.61	1.62E-08	1.71E-06
RPL14	Ribosomal protein L14	DOWN	-3.07	1.52E-07	9.20E-06
RPL15	Ribosomal protein L15	DOWN	-2.05	5.24E-06	1.00E-04
RPL17	Ribosomal protein L17	DOWN	-2.03	7.29E-06	2.00E-04
RPL18	Ribosomal protein L18	DOWN	-4.18	9.53E-08	6.41E-06
RPL18A	Ribosomal protein L18a	DOWN	-10.32	1.38E-08	1.54E-06
RPL19	Ribosomal protein L19	DOWN	-2.53	1.25E-07	7.94E-06
RPL21	Ribosomal protein L21	DOWN	-23.51	4.55E-08	3.61E-06
RPL24	Ribosomal protein L24	DOWN	-4.48	1.13E-06	4.24E-05
RPL26	Ribosomal protein L26	DOWN	-4.31	2.30E-07	1.26E-05
RPL27	Ribosomal protein L27	DOWN	-6.93	3.60E-06	1.00E-04
RPL30	Ribosomal protein L30	DOWN	-2.60	8.22E-07	3.35E-05
RPL35	Ribosomal protein L35	DOWN	-2.36	1.63E-06	5.65E-05
RPL35A	Ribosomal protein L35a	DOWN	-2.11	1.23E-08	1.42E-06
RPL36	Ribosomal protein L36	DOWN	-2.06	2.41E-06	7.54E-05
RPL37	Ribosomal protein L37	DOWN	-5.54	1.35E-05	3.00E-04
RPL38	Ribosomal protein L38	DOWN	-2.40	3.78E-08	3.17E-06
RPL4	Ribosomal protein L4	DOWN	-8.15	9.61E-07	3.76E-05
RPL41	Ribosomal protein L41	DOWN	-2.38	2.73E-07	1.43E-05
RPL5	Ribosomal protein L5	DOWN	-3.09	4.50E-11	5.43E-08
RPL6	Ribosomal protein L6	DOWN	-9.13	4.72E-09	7.48E-07
RPL8	Ribosomal protein L8	DOWN	-6.80	4.51E-09	7.24E-07
RPL9	Ribosomal protein L9	DOWN	-2.40	4.17E-08	3.40E-06
RPLP0	Ribosomal protein, large, P0	DOWN	-2.76	2.02E-06	6.62E-05
RPLP1	Ribosomal protein, large, P1	DOWN	-3.80	5.20E-12	1.95E-08
RPS10	Ribosomal protein S10	DOWN	-2.08	1.25E-06	4.61E-05
RPS11	Ribosomal protein S11	DOWN	-10.30	6.91E-07	2.94E-05
RPS13	Ribosomal protein S13	DOWN	-51.68	7.28E-11	6.49E-08

RPS14	Ribosomal protein S14	DOWN	-2.41	1.36E-07	8.50E-06
RPS15	Ribosomal protein S15	DOWN	-10.26	1.68E-08	1.76E-06
RPS15A	Ribosomal protein S15a	DOWN	-5.23	8.35E-10	2.49E-07
RPS16	Ribosomal protein S16	DOWN	-25.84	3.33E-08	2.85E-06
RPS18	Ribosomal protein S18	DOWN	-3.28	1.34E-05	3.00E-04
RPS19	Ribosomal protein S19	DOWN	-7.24	6.01E-09	8.82E-07
RPS2	Ribosomal protein S2	DOWN	-4.29	1.08E-10	7.95E-08
RPS20	Ribosomal protein S20	DOWN	-2.58	2.99E-07	1.53E-05
RPS21	Ribosomal protein S21	DOWN	-3.32	1.08E-06	4.10E-05
RPS23	Ribosomal protein S23	DOWN	-2.94	4.00E-10	1.59E-07
RPS25	Ribosomal protein S25	DOWN	-5.52	1.95E-07	1.12E-05
RPS26	Ribosomal protein S26	DOWN	-2.48	2.35E-05	4.00E-04
RPS27	Ribosomal protein S27	DOWN	-2.14	9.87E-09	1.23E-06
RPS28	Ribosomal protein S28	DOWN	-2.14	7.60E-08	5.37E-06
RPS29	Ribosomal protein S29	DOWN	-2.50	2.43E-07	1.31E-05
RPS3	Ribosomal protein S3	DOWN	-3.68	9.09E-09	1.16E-06
RPS3A	Ribosomal protein S3a	DOWN	-4.37	6.14E-10	2.02E-07
RPS4X	Ribosomal protein S4, X-linked	DOWN	-3.90	2.04E-08	2.00E-06
RPS5	Ribosomal protein S5	DOWN	-12.72	5.37E-09	8.21E-07
RPS9	Ribosomal protein S9	DOWN	-34.45	1.10E-09	2.91E-07
RSL1D1	Ribosomal L1 domain containing 1	DOWN	-2.35	3.91E-07	1.87E-05
RSL24D1	Ribosomal L24 domain containing 1	DOWN	-4.60	2.77E-08	2.50E-06
S1PR3	sphingosine-1-phosphate receptor 3	DOWN	-2.73	9.14E-07	3.64E-05
SARS	seryl-tRNA synthetase	DOWN	-3.05	8.97E-08	6.12E-06
SAT2	spermidine/spermine N1-acetyltransferase family member 2	DOWN	-2.43	2.68E-07	1.41E-05
SDC2	syndecan 2	DOWN	-3.34	1.70E-09	3.82E-07
SDHC	Succinate dehydrogenase complex, subunit C, integral membrane protein, 15kDa	DOWN	-2.31	2.10E-03	9.00E-03
SDHD	Succinate dehydrogenase complex subunit D, integral membrane protein	DOWN	-5.21	7.62E-08	5.38E-06
SDPR	Serum deprivation response	DOWN	-13.78	7.14E-09	9.94E-07
SERINC1	Serine incorporator 1	DOWN	-3.61	2.78E-09	5.18E-07
SIAH1	siah E3 ubiquitin prote inligase 1	DOWN	-3.23	2.00E-04	2.00E-03
SLC16A7	Solute carrier family 16 (monocarboxylate transporter),	DOWN	-2.01	1.93E-05	4.00E-04

	member 7				
SLC25A5	Solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 5	DOWN	-2.62	2.67E-06	8.16E-05
SLC25A6	solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 6	DOWN	-2.08	3.76E-06	1.00E-04
SLC2A1	solute carrier family 2 (facilitated glucose transporter), member 1	UP	2.66	4.30E-03	1.49E-02
SLC2A12	solute carrier family 2 (facilitated glucose transporter), member 12	DOWN	-2.96	1.42E-05	3.00E-04
SLC2A3	solute carrier family 2 (facilitated glucose transporter), member 3	DOWN	-3.22	1.21E-06	4.51E-05
SLC2A4	solute carrier family 2 (facilitated glucose transporter), member 4	DOWN	-2.13	9.12E-06	2.00E-04
SLC35A1	solute carrier family 35 (CMP-sialicacid transporter), member A1	DOWN	-3.55	5.55E-07	2.46E-05
SMPD2	sphingomyelinphosphodiesterase 2, neutral membrane (neutral sphingomyelinase)	UP	2.29	7.68E-05	1.00E-03
SMS	Spermine synthase	DOWN	-2.57	1.04E-05	2.00E-04
SOD1	Superoxide dismutase 1, soluble	DOWN	-40.66	8.52E-11	6.77E-08
SPARCL1	SPARC like 1	DOWN	-127.65	1.30E-08	1.48E-06
SPCS2	signalpeptidase complex subunit 2	DOWN	-3.68	1.85E-05	3.00E-04
SPCS3	signalpeptidase complex subunit 3	DOWN	-2.32	1.00E-04	1.20E-03
SPPL2A	signalpeptidepeptidaselike 2A	DOWN	-3.20	3.97E-05	6.00E-04
SUCLG1	succinate-CoAligase, alphasubunit	DOWN	-3.09	2.42E-08	2.24E-06
TALDO1	transaldolase 1	DOWN	-5.41	2.09E-08	2.04E-06
TMEM208	transmembraneprotein 208	DOWN	-3.65	6.90E-07	2.94E-05
TNC	tenascin C	DOWN	-2.33	3.10E-02	6.39E-02
TXN	thioredoxin	DOWN	-13.15	1.96E-07	1.12E-05
TXNRD1	thioredoxinreductase 1	DOWN	-2.77	1.88E-06	6.29E-05
UBA52	ubiquitin A-52 residueribosomalprot. fusionproduct 1	DOWN	-5.37	2.23E-07	1.23E-05
UBB	ubiquitin B	DOWN	-2.68	5.99E-11	5.92E-08
UBE2L3	Ubiquitin-conjugating enzyme E2L 3	DOWN	-2.48	1.26E-05	3.00E-04
UBE2M	ubiquitin-conjugating enzyme E2M	DOWN	-3.21	2.00E-04	1.60E-03
UBL5	ubiquitin-like 5	DOWN	-7.60	1.46E-08	1.59E-06
UGDH	UDP-glucose 6-dehydrogenase	DOWN	-2.14	5.94E-07	2.60E-05

UGP2	UDP-glucosepyrophosphorylase 2	DOWN	-2.57	2.23E-09	4.45E-07
UQCR10	ubiquinol-cytochrome c reductase, complex III subunit X	DOWN	-2.30	1.17E-02	3.06E-02
UQCR11	ubiquinol-cytochrome c reductase, complex III subunit XI	DOWN	-2.85	8.56E-07	3.46E-05
UQCRB	ubiquinol-cytochrome c reductase binding protein	DOWN	-2.80	3.01E-10	1.38E-07
UQCRC2	ubiquinol-cytochrome c reductase core protein II	DOWN	-2.82	9.14E-07	3.64E-05
UQCRFS1	ubiquinol-cytochrome c reductase, Rieskeiron-sulfur polypeptide 1	DOWN	-3.45	6.00E-07	2.62E-05
UQCRH	ubiquinol-cytochrome c reductase hingeprotein	DOWN	-2.48	1.54E-05	3.00E-04
UQCRHL	ubiquinol-cytochrome c reductase hingeprotein like	DOWN	-3.20	2.00E-04	2.00E-03
USP25	Ubiquitin specific peptidase 25	DOWN	-2.18	3.07E-05	5.00E-04
USP9X	Ubiquitin specific peptidase 9, X-linked	DOWN	-2.42	7.18E-08	5.13E-06
WFS1	Wolfram syndrome 1 (wolframin)	DOWN	-3.02	7.59E-07	3.15E-05
YME1L1	YME1-like 1 ATPase	DOWN	-2.56	1.64E-07	9.76E-06
YWHAB	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta	DOWN	-2.58	1.00E-04	1.20E-03
YWHAE	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon	DOWN	-3.71	8.10E-06	2.00E-04

NOTE: FC= Fold Change; p-val= p-value; FDR p-val= FDR p-value.

Table 4.7 Putative binding motif enrichment analysis using the Enrich webtool and showing the relative relevance of various transcription factors in the deregulated transcripts.

Term	p-val	Adj. p-val	Genes
E2F1	6.3E-22	4.0E-19	RPL4,RPL30,ACAA2,NDUFA12,ATP5C1,ADARB1,RPL8,PWP1,RPL6,UBE2L3,RPS15,RPS14,RPS16,RPL18A,RPS18,PPP2R1A,PMPCB,RPL35,RPL37,RPS11,ATP6V1E1,RPS10,PELO,B2M,PDHX,RPS9,RPL21,FNBP1,RPS5,APLP2,SARS,MRPS18A,SDHC,EEF1A1,SMS,RPL24,RPL27,SUCLG1,RPL26,UQCRC2,UQCRB,NDUFB10,COX15,MAOA,GMPS,AK2,ITPR1,COX7A2,ATP5O,OAZ1,UQCRH,MRPL11,ATP5L,ATP5B,ATP5E,ATP5D,COX11,IGFBP7,NDUFV2,ATP6V1F,DARS,RPL41,PDHA1,IDH2,GOT2,H3F3A,FN1,MRPL24,RPS28,RPS27,CALU,OGDH,RPS20,SERINC1,CALM1,RPS21,CALM2,ITM2B,RPS23,FARSB,PICALM,YWHAE,COX7B,MRPS14,RPLP1,PIGP,PTEN,SLC2A1,MTR,SLC2A3,LAMC1,ATP5G3,ATP5G2,COX7C,MRPL35,PITPNC1,CYR61,HK2,PIGY,MRPL3,ATIC,CSRP2,ALDH2,RARS,ACADM,DLAT,IDH3A,COX8A,ATP6AP1,ITPK1,RPL13A,RPS3A,NME4,ATP5F1,EEF1G,HADHB,PLCB4,CAT,NDUFS3,SLC25A5,ALDH9A1,PRPS1,OAT,NDUFB7,MRPS36,RPL12,ATP5A1,ODC1,NDUFB3,COX5B,COX5A,FSTL1,PPP2CA,CST3,PPP2CB,RPS15A,RPS3,CKB,RPL15,RPS2,RPL18,RPL17,RPL19,BCHE,NDUFA6,MDH1,MDH2,USP9X,NDUFA4,NDUFA2,NDUFA1,COX6C,ASS1,SOD1,SPCS3,RSL1D1,SPCS2,MSRB2
EKLF	2.9E-20	9.3E-18	RPL4,HIBADH,MRPS14,NDUFA12,RPLP1,SLC2A1,SLC2A4,ATP5G2,PFAS,PWP1,RPL6,RPS4X,MRPL3,RPS14,ATIC,LGALS1,RPS16,ALDH2,RPS18,IDH3B,PMPCB,RPL35,RPS11,COX8A,RPS9,GSTO1,ECH1,SARS,RPL13A,SDHC,NME4,EEF1A1,EEF1G,CAT,RPL24,RPL27,SUCLG1,SLC25A5,ALDH9A1,OAT,UQCRB,NDUFB10,MGST3,NDUFB3,GMPS,ATP5J,NEDD8,ATP5O,MRPL11,ACAT1,PPP2CA,CST3,OXCT1,RPS3,EIF4EBP2,RPL15,RPL19,MDH1,MDH2,NDUFA4,TXNRD1,GOT2,H3F3A,IDH2,RPIA,RPS26,RPS25,RSL1D1,RPS28,SPCS2,NDUFAB1,CALM3,ITM2B,UBE2M,RPS23
MYC	7.8E-20	1.7E-17	YWHAE,RPL4,RPL5,COX7B,RPL30,RPLP1,ATP5C1,RPL8,ATP5G2,COX7C,PWP1,PIGY,RPL6,GLS,RPS15,MRPL3,RPS14,ATIC,RPS19,RPL18A,RPS18,RARS,RPL36,RPL35,RPL38,RPL37,RPS13,COX8A,RPS9,RPL21,RPS5,SARS,NME4,RPS3A,ATP5F1,EEF1A1,EEF1G,NDUFS4,NDUFS3,YME1L1,RPL27,RPL26,NDUFB5,RPL12,ATP5A1,ODC1,MRPL11,ATP5B,RPS15A,ATP5E,RPS3,EIF4EBP2,RPL13,RPL15,RPS2,RPL18,RPL17,COX10,ATP6V1F,DARS,NDUFA4,GOT2,H3F3A,MRPL24,SOD1,RPS25,RSL1D1,RPS27,RPS29,RPS20,RPS21,FARSB,RPS23
ESRRB	2.5E-18	3.1E-16	HIBADH,YWHAB,NDUFA12,COX4I1,ATP5C1,PYGM,ATP5G3,COX6A1,ATP5G2,COX7C,PIGY,GLS,RPS15,CSRP2,AUH,IDH3B,UQCRFS1,DLAT,AASS,IDH3A,COX8A,PLA2G12A,GPX4,PGAM1,APLP2,SDHC,SDHD,NDUFS5,NDUFS4,NDUFS3,SUCLG1,UQCRC2,GAS6,DLD,NDUFB9,

			PRKAA2,UQCRB,NDUFB10,NDUFB5,MRPS36,MGST3,ATP5A1,ODC1,GMPS,AK2,COX7A2,PPM1K,PDHB,ATP5O,COX5B,COX5A,COX7A1,LTBP1,UQCRH,ACAT1,ATP5B,OXCT1,ATP5E,PDGFC,ATP5D,COX11,CKB,NDUFV2,COX10,NDUFA9,NDUFA6,MDH1,MDH2,NDUFA2,COX6C,NDUFAB1,OGDH,ACO1,CALM3,CHPT1,CALM2,ITM2B, UBE2M
XRN2	8.1E-18	8.5E-16	YWHAE,RPL4,RPL5,NDUFA11,MRPS14,RPLP0,AMD1,RPL8,MRPL35,RPL9,RPL6,UBE2L3,RPS4X,RPS14,RPS19,RPS18,RPL36,IDH3B,RPL35,RPL38,RPL37,B2M,RPS10,RPS9,RPS5,SARS,MRPS21,RPL13A,SDHC,MRPS7,EEF1A1,NDUFS5,NDUFS3,RPL24,RPL27,RPL26,GAPDH,FH,OAT,NDUFB6,RPL12,ITPR1,ATP5J,TXN,OAZ1,ACAT1,ATP5B,UGP2, RPS15A,UBB,RPL14,COX11,RPL13,RPL15,CKB,RPS2,RPL18,RPL17,IP6K2,RPL19,RPL41,MDH1,NDUFA2,RPL35A,MRPL24,SOD1,RPS26,RPS28,RPS27,GSTA4,RPS29,CYCS,CALM3,SERINC1,RPS20,CALM1,CALM2,RPS21,UBE2M,FARSB
NELFA	3.5E-13	2.5E-11	YWHAE,PYGB,ACAA2,HIBADH,NDUFA12,AMD1,PTEN,SLC2A1,ATP5C1,MTR,SLC2A3,ATP5G3,PIK3C2A,RPL8,RPL6,GLS,RPS15,PCMT1,CSRP2,ALDH2,RPS18,IMPA2,ACADM,RPL37,DLAT,RPS11,RPS13,PDHX,RPS9,RPL21,GPX4,FNBP1,APLP2,RPL13A,SDHC,NME4,SDHD,CIT,EEF1A1,ITPKB,UGDH,PPA1,CAT,UQCRC2,RPL26,GAPDH,UQCRB,NDUFB6,RPL12,ATP5A1,GMPS,ITPR1,ACACB,ADH5,PPP2CA,ATP5B, UGP2,UBB,RPS3,EIF4EBP2,RPL13,CKB,RPS2,ATP6V0A1,RPL19,SLC35A1,DARS,RPL41,NDUFA6,NDUFA4,H3F3A,FN1,SOD1,RPS26,RSL1D1,RPS28,SPCS2,RPS27,CALU,OGDH,RPS20,CALM1,RPS21,CALM2
TTF2	5.8E-13	3.7E-11	YWHAE,RPL4,RPL5,MRPS14,RPLP1,RPLP0,AMD1,COX7C,RPL9,RPL6,RPS15,RPS4X,MRPL3,RPS14,RPS18,RPL36,RPL35,RPL38,RPL37,RPS11,PRKACB,RPS13,RPS9,RPS3A,ACSL3,MRPS7,EEF1A1,EEF1G,NDUFS4,NDUFS3,RPL24,UQCRC2,RPL26,GAPDH,UQCRB,NDUFB10,RPL12,MGST3,ITPR1,COX7A2,PDHB,RPS15A,UBB,RPL14,RPS3,RPL13, RPL15,RPS2,RPL18,RPL17,RPL19,RPL41,NDUFA4,H3F3A,MRPL24,SOD1,RSL1D1,RPS28,RPS27,NDUFAF4,CYCS,CALM3,CHPT1,SERINC1,RPS20,CALM1,CALM2,PAM,UBE2M,RPS23
ETS1	1.2E-12	7.1E-11	RPL5,COX7B,COX4I1,RPLP0,MTR,RPL8,COX7C,RPL9,RPL6,RPS15,UBL5,RPS14,MRPL42,MRPL3,RPS16,RPS19,RARS,IDH3B,RPL36,UQCRFS1,RPL38,RPL37,ATP6V1E1,RPS11,RPS13,RPS9,RPL21,RPS5,RPL13A,RPS3A,SDHD,PGM2L1,MRPS18C,RPL24,YME1L1,SUCLG1,RPL27,UQCRC2,RPL26,UBA52,NDUFB7,NDUFB10,RPL12,MGST3,NDUFB3,ITPR1,COX7A2,ATP5O,MRPL11,RPS15A,ATP5D,RPL14,RPS3,TMEM208,RPL13,ATP6V1F,NDUFA9,SLC35A1,NDUFA4,PHKB,MRPL24, RPS26,RPIA,RPS25,RPS27,NDUFAB1,CYCS,RPS20,CALM1,RPS21,FARSB,RPS23
ZFX	8.6E-11	4.6E-09	RPL30,NDUFA12,ATP5C1,ADARB1,RPL8,PWP1,RPL6,UBE2L3,RPS14,PCMT1,RPS16,RPS19,RPL18A,RPS18,PPP2R1A,RPL36,RPL37,RPS11,PDHX,FNBP1,RPS5,MRPS18A,SDHD,EEF1A1,RPL24,RPL27,RPL26,UQCRC2,UBA52,NDUFB10,GMPS,AK2,OAZ1,LTBP1,COX7A1,ADH

			5,ACAT1,MRPL11,ATP5L,ATP5B,UBB,ATP5E,ESD,ATP5D,ABL1,NDUFV2,DARS,RPS27,RPS29,OGDH,RPS20,ACO1,CALM1,RPS21,ITM2B,RPS23,UBE2M,HIBADH,YWHAB,RPLP1,PIGP,PTEN,MTR,ATP5G2,COX7C,MRPL3,ATIC,AUH,IMPA2,RARS,UQCRFS1,DLAT,IDH3A,COX8A,GSTO1,ITPK1,NME4,UGDH,NDUFS3,SLC25A5,DLD,RPL12,ATP5A1,PDHB,COX5B,COX5A,PPP2CA,PPP2CB,RPS15A,RPL14,RPS3,RPL13,RPL15,RPS2,RPL18,SLC35A1,NDUFA6,MDH1,LAMB2,USP9X,NDUFA4,TXNRD1,COX6C,SOD1,RPIA,RSL1D1,GNPDA2,NDUFAB1,MSRB2
VDR	1.5E-09	7.1E-08	YWHAE,RPL5,COX7B,NDUFA11,MRPS14,NDUFA12,AMD1,RPLP0,PTEN,SLC2A3,MRPL35,PFAS,UBE2L3,RPS15,RPS14,MRPL3,RPS19,AUH,RPL36,PMPCB,RPL38,RPL37,RPS11,RPS9,GSTO2,RPS5,ECH1,MRPS21,RPL13A,RPS3A,MRPS18C,HADHB,RPL24,YME1L1,SUCLG1,RPL27,UQCRC2,RPL26,UBA52,NDUFB7,NDUFB6,NDUFB5,ODC1,NDUFB3,ATP5I,NEDD8,COX7A2,PPM1K,COX5B,ACAT1,ATP5L,PPP2CA,UBB,OXCT1,ATP5E,TMEM208,RPL13,RPL19,ATP6V1F,NDUFA9,RPL41,NDUFA6,MDH1,RPL35A,COX6C,MRPL24,RPS26,RPIA,SPCS3,RPS27,NDUFAB1,CALU,CYCS,RPS20,CALM3,NDUFAF1,RSL24D1,RPS21,CALM2,FARSB,RPS23
HOXC9	2.3E-09	1.1E-07	COX7B,RPL30,PIGT,HIBADH,MRPS14,NDUFA12,RPLP1,SLC2A3,ATP5G3,COX7C,MRPL35,PIGY,RPS15,UBL5,MRPL3,ATIC,LGALS1,RPS16,RPS19,IDH3B,RPL38,RPL37,ATP6V1E1,B2M,RPS9,APLP2,MRPS21,MRPS18A,RPS3A,SDHD,ACSL3,MUT,COX6B1,HADHA,CAT,NDUFS4,RPL24,YME1L1,UQCRC2,RPL26,MFGE8,DLD,OAT,NDUFB7,MRPS36,NDUFB5,ITPR1,ATP5I,PPM1K,COX5B,UQCRH,RPS15A,UBB,ATP5E,ATP5D,ESD,TMEM208,RPL15,RPS2,RPL17,IP6K2,NDUFA9,NDUFA6,MDH2,COX6C,RPS26,RPS25,RSL1D1,RPS28,RPS29,NDUFAF4,NDUFAB1,NDUFAF2,CALU,NDUFAF1
SPI1	1.1E-07	4.7E-06	YWHAE,RPL30,PIGT,ACAA2,MRPS14,YWHAB,ATP5C1,SLC2A3,ATP5G3,COX6A1,RPL8,PITPNC1,RPL6,GLS,RPS15,RPS14,ATIC,RPS16,RPS19,RPL18A,IMPA2,RARS,IDH3B,RPL35,RPL38,RPS11,IDH3A,GSTO2,RPL21,PGAM1,RPL13A,PLA2G4A,MRPS18A,RPS3A,EEF1G,ITPKB,NDUFS4,SUCLG1,RPL27,GAPDH,NDUFB9,NDUFB5,AK2,ITPR1,NEDD8,PDHB,COX5B,COX7A1,ACAT1,PPP2CA,CST3,ITPRIPL2,RPL14,RPS3,EIF4EBP2,RPL13,RPL17,COX10,NDUFA9,DARS,ATP8B2,GOT2,RPS26,RPS27,SERINC1,RSL24D1,RPS21,CALM2,UBE2M
YY1	3.7E-07	1.5E-05	YWHAE,RPL4,APP,COX7B,RPL30,MRPS14,AMD1,LAMC1,ATP5G2,MRPL35,COX7C,UBE2L3,MRPL42,RPL36,IDH3B,RPS11,COX8A,PDHX,ECH1,RPL13A,MRPS18A,SDHC,MRPS7,MRPS18C,EEF1A1,INPP4B,DPYD,NDUFS4,YME1L1,UBA52,DGKH,UQCRB,RPL12,ATP5A1,PDHB,OAZ1,LTBP1,UQCRH,MRPL11,ATP5L,UBB,RPS3,RPL18,COX10,ATP6V0A1,NDUFA9,NDUFA2,RPL35A,MRPL24,RPS26,RPS28,SPCS2,RPS27,OGDH,ITM2B
CEBPB	1.0E-06	3.9E-05	COX7B,NDUFA11,PIGP,SAT2,ATP5G3,PIK3C2A,CYR61,RPL6,RPS15,RPS4X,UBL5,RPS14,LGALS1,CSRP2,RPS19,UQCRFS1,RPL38,ATP6V1

			E1,RPS10,PELO,COX8A,PDHX,RPS9,GSTO2,FNBP1,ECH1,MRPS21,SDHD,DNAJC3,SUCLG1,RPL27,SLC25A5,UQCRB,ATP5A1,ODC1,ITPR1,ATP5H,UQCR10,COX5B,ATP5O,COX7A1,MRPL11,ITPRIPL2,PPP2CB,UBB,RPL14,RPL15,COX10,IP6K2,RPL19,USP25,ATP8B2,NDUFA4,WFS1,NDUFA2,H3F3A,NDUFA1,PHKB,RPS26,RPIA,RPS25,RPS27,RPS29,NDUFAF4,CYCS,RPS20,PICALM
GABP	1.6E-06	5.5E-05	RPL30,PIGT,NDUFB7,NDUFB5,NDUFB3,COX5B,COX6A1,COX7C,UQCRH,RPL6,MRPL3,UBL5,RPS14,RPS18,RARS,IDH3B,PMPCB,ATP6V1E1,COX10,COX8A,RPL41,MDH1,RPS5,MRPS21,RPL35A,MRPS18A,SDHC,RPS3A,SDHD,MRPL24,MRPS18C,RPS25,RPS27,CALU,NDUFS3,YME1L1,RPL27,SERINC1,RPS20,RPL26,UBE2M
CREM	2.0E-06	6.8E-05	RPL5,RPL30,ACAA2,ATP5C1,COX6A1,RPL8,PWP1,RPL6,GLS,RPS14,PCMT1,RPS16,RPS19,RPL18A,RPS18,PPP2R1A,RPL36,SLC16A7,CYP1B1,RPL38,RPS11,PRKACB,PELO,RPS13,RPS9,RPL21,IGFBP5,RPS5,APLP2,SDHC,ACSL3,EEF1A1,RPL24,RPL27,SUCLG1,RPL26,UBA52,NDUFB10,MAOA,GMPS,AK2,ATP5J,COX7A2,ATP5O,OAZ1,ADH5,ACAT1,ATP5B,UBB,ATP5D,ABL1,COX11,EIF4EBP2,NDUFV2,ATP6V1F,ATP6V0A1,DARS,RPL41,H3F3A,FN1,RPS26,RPS27,RPS29,NDUFAF4,NDUFAF2,CALU,OGDH,RPS20,CALM3,SERINC1,CALM1,CALM2,UBE2M,FARSB,PICALM,YWHAE,COX4I1,RPLP0,AMD1,PTEN,MTR,LAMC1,PIK3C2A,ATP5G2,PFAS,PRSS23,CYR61,AUH,IMPA2,RARS,IDH3B,UQCRFS1,DLAT,IDH3A,COX8A,PLA2G12A,GSTO2,GPX4,PGAM1,RPL13A,RPS3A,PGM2L1,CIT,EEF1G,DNAJC3,HADHB,ITPKB,HADHA,UGDH,NDUFS3,PPARG,GAPDH,DLD,OAT,NDUFB6,PPM1L,MRPS36,ATP5A1,ODC1,PDHB,COX5B,COX5A,PPP2CA,PPP2CB,UGP2,RPL14,RPS3,RPL13,TMEM208,CKB,RPL18,SLC35A1,USP25,NDUFA6,MDH1,MDH2,USP9X,TXNRD1,RPL35A,COX6C,CDC42BPA,SOD1,SPCS3,RSL1D1,SPCS2,PDE10A,MSRB3,CHPT1
PPARG	3.4E-06	1.0E-04	ACAA2,COX15,YWHAB,ECH1,MGST1,AK2,SDHC,ATP5G3,COX5A,SOD1,C3,HADHB,RPS19,SDPR,CYCS,PPARG,CHPT1,ACADM
FLI1	3.7E-06	1.1E-04	RPL4,CHRM3,RPL30,NDUFA12,TNC,ATP5C1,COX6A1,PWP1,PLB1,RPL6,UBE2L3,RPS15,RPS14,PCMT1,LGALS1,RPS16,RPS19,RPL18A,RPL36,RPL38,RPL37,RPS11,ATP6V1E1,RPL21,IGFBP5,FNBP1,PRKCB,RPS5,APLP2,SARS,SDHC,PRKCA,SDHD,COX6B1,MRPS18C,PLA2G16,INPP4B,SDPR,RPL24,SPARCL1,YME1L1,RPL27,RPL26,UBA52,ITPR1,ATP5J,ATP5I,ITPR3,ATP5H,ATP5O,LTBP1,UQCRH,ADH5,ATP5L,C3,EIF4EBP2,IGFBP7,NDUFV2,COX10,DARS,RPL41,GOT2,H3F3A,PHKB,IGF1,MRPL24,RPS26,RPS25,RPS27,RPS29,CALU,OGDH,CALM3,SERINC1,NDUFAF1,CALM1,RPS21,RSL24D1,RPS23,UBE2M,PICALM,YWHAE,PIGT,MRPS14,YWHAB,RPLP1,COX4I1,AMD1,PIGP,SLC2A1,MIA3,SLC2A3,ATP5G2,COX7C,UBL5,MRPL3,ATIC,IMPA2,RARS,CYP4V2,IDH3B,UQCRFS1,COX8A,GSTO2,GLRX3,GPX4,GSTO1,MRPS21,RPS3A,MUT,EEF1G,ITPKB,HADHA,UGDH,PIK3CA,NDUFS5,NDUFS3,MFGE8,GAPDH,DLD,ALDH9A1,SLC25A6,NDUFB9,NDUFB6,NDUFB5,RPL12,NDUFB3,COX5B,PPP2CA,CST3,RPS3,TMEM208,RPL15,RPL19,

			GSN,NDUFA6,MDH1,MDH2,PLCL1,TXNRD1,NDUFA2,CDC42BPA, ASS1,RPIA,RSL1D1,PDE10A,GNPDA2,NDUFAB1
SOX2	7.4E-06	2.1E-04	YWHAE,COX7B,COX4I1,AMD1,PTEN,SLC2A3,ATP5G2,COX7C,PITPN C1,CYR61,HK2,UBL5,RPS14,RPS19,ALDH2,RPS18,PMPCB,RPS11,ID H3A,COX8A,ENTPD1,RPL21,FNBP1,RPS5,RPL13A,SARS,SDHC,RPS3 A,PGM2L1,CIT,EEF1A1,EEF1G,BACE1,UGDH,SMS,RPL27,GAPDH,PR PS1,ITPR1,PDHB,ATP5H,COX5A,OAZ1,FSTL1,COX7A1,CST3,UBB,PD GFC,ATP5D,CKB,NDUFV2,RPL18,ATP8B2,FN1,IGF1,COX6C,MRPL24 ,RPS26,RPIA,GSTA4,CALM1,RPS21,PLCD4,RPS23
OCT4	7.5E-06	2.1E-04	RPL5,SLC2A1,SLC2A3,RPL8,RPL9,CYR61,HK2,GLS,RPS4X,RPS14,RPS 16,RPS19,ALDH2,RPS18,RPS11,PRKACB,ENTPD1,RPL21,FNBP1,RPS 5,RPL13A,SARS,RPS3A,CIT,EEF1A1,BACE1,UGDH,PLCB4,SMS,SPAR CL1,RPL27,GAPDH,PRPS1,ITPR1,NEDD8,COX7A2,COX5B,COX5A,O AZ1,COX7A1,CST3,ATP5B,UBB,PDGFC,RPS3,ABL1,CKB,RPS2,RPL18, ATP6V1F,AOC3,ATP8B2,GOT2,FN1,COX6C,ASS1,RPS26,RPIA, GSTA4,RPS20,CALM1,RPS21,UBE2M,FARSB
CHD1	8.5E-06	2.5E-04	YWHAE,RPL5,APP,MRPS14,SPPL2A,SLC2A1,ATP5C1,MIA3,LAMC1, ATP5G2,COX7C,MRPL35,CYR61,GLS,UBL5,CSRP2,IMPA2,CKMT1B, ACADM,RPS13,COX8A,GSTO2,GLRX3,PGAM1,SARS,MRPS18A,RPS 3A,ACSL3,ITPKB,PPA1,UQCRC2,ALDH9A1,PPM1L,COX15,SDC2,MG ST3,AK2,UQCR11,UQCRH,ATP5L,UBB,CKB,NDUFV2,IP6K2,USP25,A TP8B2,USP9X,LAMB2,CKMT1A,TXNRD1,GOT2,CDC42BPA,RPS26,R PS25,RPS28,RPS27,RPS29,OGDH,MSRB2,CHPT1,CALM1,RPS21, CALM2,PICALM
BCL3	1.0E-05	2.6E-04	YWHAE,RPL5,NDUFB10,NDUFB5,ATP5A1,PIGP,AMD1,ATP5J,PYGM ,UQCR11,LAMC1,RPL8,COX5A,CYR61,HK2,MRPL11,ATP5B,PCMT1, UGP2,RPL18A,UBB,ATP5D,IDH3B,PELO,RPL17,PDHX,RPL41,RPS5, ECH1,MRPS18A,SOD1,SPCS2,CYCS,RPS21,ITM2B
TAL1	1.3E-05	3.1E-04	RPL4,NDUFA11,MRPS14,COX4I1,RPLP0,PTEN,SLC2A1,ATP5G3,RPL 8,RPL6,RPS15,RPS14,ATIC,RPS19,PPP2R1A,RPS18,RPL36,RPS10,B2 M,RPS9,GLRX3,GPX4,RPS5,MRPS7,EEF1A1,RPL27,RPL26,UBA52,AL DH9A1,NDUFB9,NDUFB10,NDUFB6,RPL12,ATP5A1,ODC1,NDUFB3 ,ITPR1,UQCR11,COX5A,PPP2CA,ATP5B,UBB,RPL14,RPS3,COX11,RP L13,RPL18,RPL17,RPL19,DARS,RPL41,NDUFA6,MDH2,TXNRD1,RPL 35A,PHKB,RPS26,RPS25,RPS27,PDE10A,RPS20,CALM1,RPS21, UBE2M,RPS23
RUNX2	1.7E-05	3.9E-04	PIGT,TNC,ATP5C1,PITPNC1,CYR61,HK2,GLS,LGALS1,RPS19,RPL18A ,ALDH2,CYP1B1,DLAT,RPS10,PELO,LMCD1,IGFBP5,FNBP1,GSTO1,P GAM1,MMP2,APLP2,ITPK1,MRPS18A,NME4,MRPS7,CIT,EEF1A1,E EF1G,BACE1,UGDH,PLA2G16,PLCB4,RPL26,GAS6,PRKAA2,OAT,CO X15,AK2,ATP5H,UQCR10,ACACB,LTBP1,ADH5,MRPL11,UGP2,PDGF C,ATP6V0A1,RPL19,SLC35A1,DARS,GSN,MDH1,ATP8B2,LAMB2, TXNRD1,GOT2,FN1,CDC42BPA,RSL1D1,RPS28,RPS20,CALM1
ASH2L	1.7E-05	3.9E-04	RPL5,COX7B,HIBADH,MRPS14,NDUFA12,PTEN,LAMC1,COX6A1,RP L8,MRPL35,RPL9,PITPNC1,RPL6,RPS15,RPS4X,UBL5,RPS14,SMPD2,

			RPS16,CSRP2,RPS18,RPL36,IDH3B,UQCRFS1,RPL38,DLAT,RPS10,LMCD1,RPS13,FNBP1,RPS5,ITPK1,TALDO1,SDHC,RPS3A,SDHD,PGM2L1,DNAJC3,EEF1A1,HADHA,PLCB4,NDUFS4,RPL24,SPARCL1,SUCLG1,GAS6,PRPS1,OAT,NDUFB7,UQCRB,NDUFB6,ATP5A1,COX7A2,PDHB,ATP5H,ATP5O,UQCRH,ADH5,ATP5L,RPS15A,ALDH1B1,PDGFC,ABL1,COX10,ATP6V1F,ATP6V0A1,NDUFA9,BCHE,DARS,RPL41,MDH1,NDUFA4,IDH2,NDUFA2,H3F3A,RPL35A,COX6C,MRPL24,SOD1,RPS26,RPS25,RSL1D1,GSTA4,MSRB2,CALU,RPS20,CHPT1,SERINC1,NDUFAF1,CALM1,CALM2,ITM2B,FARSB
EST1	2.1E-05	4.7E-04	NDUFB5,NDUFB3,ITPR1,NEDD8,MTR,ATP5O,COX7C,RPL9,UQCRH,RPL6,MRPL11,ATP5L,MRPL42,RPS14,RPS18,IDH3B,COX10,RPL41,RPS9,MDH1,RPS5,ECH1,MRPS21,RPL35A,COX6C,MRPL24,PGM2L1,RPIA,RPS26,RPS25,SPCS2,RPS27,RPS29,NDUFS5,YME1L1,RPL27,SERINC1,UQCRC2
NR0B1	2.6E-05	5.3E-04	NDUFA12,ATP5C1,PYGM,SLC2A3,ATP5G3,CYR61,RPS14,IMPA2,RPL36,ACADM,RPS10,PGAM1,ITPK1,SARS,SDHC,SDHD,HADHB,HADHA,SMS,SLC25A5,GAPDH,PRPS1,OAT,NDUFB6,MGST3,ATP5A1,GATA6,ATP5J,PDHB,PPM1K,COX5B,ATP5O,COX5A,ATP5L,CST3,ATP5B,ALDH1B1,ATP5E,PDGFC,CKB,NDUFV2,RPL17,COX10,RPL19,GOT2,FN1,COX6C,ASS1,RPS26,GSTA4,CALU,OGDH,MSRB2,ACO1,CHPT1
ELK1	4.0E-05	8.0E-04	RPL30,MRPS14,NDUFB3,MGST1,COX7A2,UQCRH,CYR61,GLS,RPS14,CSRP2,RPS19,PPP2R1A,RPS18,RPL38,PRKACB,RPL19,COX8A,PDHX,DARS,RPL41,NDUFA4,ECH1,SIAH1,MRPS21,RPL35A,HADHB,DNAJC3,RPS27,NDUFS5,RPL24,MSRB3,PPARG,SUCLG1,RPS20,NDUFAF1
CTCF	5.3E-05	9.9E-04	RPL5,APP,HIBADH,MRPS14,SLC2A1,SLC2A3,SLC2A4,ATP5G2,RPS15,RPS14,RPS19,RPL18A,RARS,RPL36,PMPCB,CYP1B1,RPL35,RPL38,RPL37,FNBP1,RPS5,ITPK1,RPL13A,MRPS18A,RPS3A,ATP5F1,RPL27,RPL26,NDUFB7,UQCRB,NDUFB10,PDE1C,COX15,MGST3,GATA6,COX5A,LTBP1,UQCRH,MRPL11,PPP2CB,OXCT1,ATP5D,NDUFV2,NDUFA6,USP9X,GOT2,ASS1,RPS26,RPS29,NDUFAB1,RPS20
SOX17	5.1E-05	9.9E-04	APP,COX7B,MRPS14,COX4I1,SLC2A1,LAMC1,SLC2A4,PIK3C2A,ATP5G2,MRPL35,CYR61,HK2,UBL5,RPS16,RARS,PMPCB,RPL38,ACADM,RPL37,RPS11,COX8A,GSTO1,GPX3,RPS5,APLP2,ECH1,RPL13A,PLA2G4A,SDHC,MRPS18C,EEF1A1,DPYD,NDUFS4,PRPS1,PPM1L,ODC1,GATA6,ITPR3,COX7A2,PDHB,PPM1K,ATP5O,UQCRH,MRPL11,CST3,UBB,RPS3,ATP6V1F,RPL41,NDUFA6,USP9X,IGF1,SOD1,RPS26,RPS25,RPS27,GSTA4,MSRB3,RPS20,CALM1
FOXO3	5.9E-05	1.1E-03	PIGT,YWHAB,ATP5G3,LAMC1,PIK3C2A,RPL9,RPL6,RPL18A,AUH,ALDH2,UQCRFS1,RPL37,COA5,RPS9,GSTO2,GSTO1,MRPS18A,SDHC,SDHD,EEF1A1,YME1L1,UQCRC2,RPL26,OAT,NDUFB7,NDUFB6,MRPS36,COX15,NDUFB5,COX5B,ADH5,MRPL11,ACAT1,ATP5L,RPL14,TMEM208,RPL15,RPS2,NDUFV2,COX10,ATP6V1F,NDUFA9,USP25,NDUFA6,MDH2,NDUFA4,IDH2,RPL35A,GATM,SPCS2,RPS27,GNPDA

			2,CYCS,ACO1,CALM3,SERINC1,RPS21,CALM2,ITM2B,FARSB,RPS23
TAF7L	8.4E-05	1.5E-03	ACAA2,NDUFB10,COX15,PTEN,MGST1,ATP5J,PDHB,ATP5G3,COX5A,CST3,SMPD2,LGALS1,RPS19,UBB,PDK4,RPL38,ACADM,DLAT,LMCD1,AOC3,GSTO1,NDUFA2,ECH1,RPS3A,MRPS7,ALDH6A1,SDPR,CAT,CALU,CYCS,RPL27,CALM1,CALM2,ALDH9A1
CREB1	9.6E-05	1.7E-03	PRPS1,RPL5,UQCRB,NDUFB6,MRPS14,RPL12,AK2,MTR,PWP1,ATP5B,RPS14,RPS19,RPL18A,RPS18,RPL37,RPL18,AASS,RPS13,NDUFA6,PDHA1,MDH1,RPS5,MRPS21,SDHC,MRPL24,MUT,MRPS18C,RPS28,RPS29,CALU,OGDH,RPS20,DLD,CALM2,RPS21
HOXB4	1.0E-04	1.7E-03	COX7B,RPLP1,AMD1,PTEN,COX6A1,PITPNC1,UBE2L3,UBL5,MRPL3,PCMT1,RPS19,RPL18A,PPP2R1A,RPS18,RPL38,PRKACB,ECH1,MRPS21,RPL13A,PLA2G4A,ACSL3,EEF1G,HADHB,NDUFS3,RPL24,SUCLG1,MFGE8,NDUFB9,UQCRB,PPM1L,NDUFB3,COX7A2,PDHB,PPM1K,COX5B,OAZ1,COX7A1,UQCRH,ATP5L,PPP2CA,ATP5B,UBB,EIF4EBP2,RPL13,ATP6V0A1,NDUFA9,NDUFA4,GOT2,PHKB,MRPL24,RPS26,RPIA,CALU,MSRB2,RPS21,PICALM
JARID1A	1.0E-04	1.7E-03	RPL4,PIGT,COX4I1,RPLP0,ATP5C1,MRPL35,RPL9,PWP1,UBL5,MRPL3,RPS16,RPL18A,AUH,RPS18,IDH3B,UQCRFS1,DLAT,PRKACB,COX8A,RPS5,MRPS21,MRPS18A,SDHC,SDHD,COX6B1,MRPS18C,NDUFS4,NDUFS3,YME1L1,RPL27,DLD,NDUFB9,NDUFB10,NDUFB6,COX15,NDUFB5,MGST3,NDUFB3,NEDD8,COX7A2,ATP5H,COX5B,ATP5D,RPL14,TMEM208,RPS2,RPL18,COX10,ATP6V0A1,RPL19,NDUFA9,LAMB2,NDUFA2,NDUFA1,COX6C,RSL1D1,ALDH6A1,RPS28,SPCS2,NDUFAF2,SERINC1,NDUFAF1,RPS21,FARSB
FOXP3	1.0E-04	1.7E-03	NDUFB6,MRPS36,RPL12,MGST3,RPLP0,NDUFB3,PTEN,ATP5J,MIA3,SLC2A3,ATP5G3,UQCRH,MRPL11,MRPL3,RPS14,RPS15A,RPS19,UBB,RARS,TMEM208,RPL38,RPL15,RPL37,RPS2,RPL18,NDUFA9,MDH1,RPL21,NDUFA4,ECH1,COX6C,MRPL24,MRPS18C,EEF1A1,RPIA,ITPKB,RPS26,NDUFS4,NDUFS3,RPL24,SUCLG1,RPS20,UQCRC2,NDUFAF1,RPL26,FARSB
KLF4	1.4E-04	2.1E-03	RPL4,RPL5,COX7B,NDUFB7,RPL12,GATA6,PTEN,ITPR3,SLC2A3,RPL8,ADH5,RPS15A,RPL18A,UBB,PPP2R1A,RPS18,ATP5E,PDGFC,ABL1,RPL38,ACADM,RPL15,RPL37,RPS11,RPL17,RPL19,RPL41,RPS9,RPL21,IGFBP5,RPS5,ECH1,ITPK1,FN1,RPL13A,MUT,ITPKB,RPS25,SPCS2,PDE10A,NDUFS4,SMS,RPL27,RPS20,GAPDH,UBA52,PICALM,RPS23
GATA4	1.8E-04	2.7E-03	ACAA2,MRPS14,NDUFA12,COX4I1,COX6A1,MRPL35,PRSS23,PITPNC1,HK2,RPS14,CSRP2,PDK4,PELO,IGFBP5,GPX4,FNBP1,RPS5,ITPK1,RPL13A,PGM2L1,COX6B1,HADHB,HADHA,PLCB4,SDPR,PLSCR4,UQCRC2,GAS6,ALDH9A1,OAT,ODC1,GMPS,PPM1K,COX7A1,LTBP1,UQCRH,ADH5,ACAT1,CST3,ATP5B,UBB,ATP6V0A1,ATP6V1F,NDUFA9,SLC35A1,DARS,H3F3A,IGF1,ASS1,SOD1,RPIA,SPCS3,RSL1D1,CYCS,CALM3,CALM1,PAM,CALM2,FARSB,PICALM
DCP1A	1.9E-04	2.7E-03	UQCRB,RPLP0,SLC2A1,RPL8,OAZ1,ATP5L,RPS16,LGALS1,CYP4V2,RPL37,RPL21,MDH1,NDUFA4,PGAM1,GOT2,RPL13A,RPL35A,MRPS7

			,MRPL24,RPS26,NDUFS3,CYCS,CALM3,RPS20,CALM1,MFGE8, CALM2,UBE2M,PICALM
ZFP42	1.9E-04	2.7E-03	YWHAE,RPL4,RPL5,PIGT,UQCRB,NDUFB5,MAOA,RPL12,UQCRH,ATP5L,UBE2L3,ATP5B,RPS14,UBB,ALDH2,PPP2R1A,RPS18,RPL36,ACADM,RPL15,CKB,RPS11,PELO,NDUFA9,PDHX,RPS9,NDUFA6,GSTO2, PGAM1,NDUFA2,ECH1,MRPS21,RPL13A,RPL35A,MRPS18A,PHKB, MRPS7,MRPL24,ASS1,EEF1A1,HADHB,HADHA,GSTA4,RPL24, YME1L1,ITM2B,RPS23
HCFC1	6.0E-04	7.4E-03	MRPS14,MDH2,NDUFB5,ATP5F1,UQCRH,MRPL11,RPIA,RPS28, RPS19,RPS18,COX11,RPL27,SUCLG1,RPL26,RPS23
ESR1	6.4E-04	7.8E-03	PRKAA2,PDHA1,ACAA2,NDUFA4,GOT2,GATA6,MGST1,ASS1,ATP5L ,SOD1,C3,PCMT1,UBB,ALDH2,PDK4,RPL15,B2M,PELO,PICALM
PPARD	9.9E-04	1.1E-02	ITPKB,HADHB,HADHA,CSRP2,OXCT1,ITPR1,PLA2G4A,PRKCA,SDHD, MFGE8,ALDH9A1
RARG	1.0E-03	1.1E-02	PIGT,PPM1L,NDUFB10,NDUFA4,MMP2,FN1,RPL13A,LAMC1,FSTL1 ,LTBP1,PIGY,PITPNC1,ESD,IDH3B,CALU,ACADM,RPL19
CEBPD	1.0E-03	1.2E-02	CHRM2,HIBADH,RPLP0,PTEN,PRSS23,CYR61,HK2,RPS14,LGALS1,AUH,PDK4,CYP1B1,RPS13,RPS9,GSTO2,GSTO1,MMP2,MRPS18A,RPS3A,NME4,PLCB4,CAT,NDUFS4,PPARG,GAS6,SLC25A5,UBA52,MGST3,ITPR1,MGST1,ATP5O,FSTL1,LTBP1,ATP5L,UGP2,UBB,PDGFC,ABL1,COX10,ATP6V0A1,ATP6V1F,USP25,FN1,COX6C,ASS1,RPIA,CALU ,CALM2,ITM2B,FBN1
NFE2L2	1.2E-03	1.2E-02	RPL5,PYGB,CHRM3,OAT,MAOA,NDUFB3,TNC,LAMC1,ATP5G3,ACACB,RPS15A,PDGFC,ESD,CYP1B1,IGFBP7,S1PR3,LMCD1,FNBP1,TXNRD1,FN1,PLA2G4A,TALDO1,MRPS18A,IGF1,COX6C,CDC42BPA, SPCS3,PIK3CA,GSTA4,SDPR,DPYD,CAT,CHPT1,PAM
NRF2	1.2E-03	1.2E-02	RPL5,PYGB,CHRM3,OAT,MAOA,NDUFB3,TNC,LAMC1,ATP5G3,ACACB,RPS15A,PDGFC,ESD,CYP1B1,IGFBP7,S1PR3,LMCD1,FNBP1,TXNRD1,FN1,PLA2G4A,TALDO1,MRPS18A,IGF1,COX6C,CDC42BPA, SPCS3,PIK3CA,GSTA4,SDPR,DPYD,CAT,CHPT1,PAM
YAP1	1.2E-03	1.2E-02	YWHAE,DGKG,CHRM3,APP,TNC,SLC2A1,SLC2A3,ATP5G3,RPL8,COX7C,CHRDL1,PITPNC1,RPL6,GLS,PCMT1,AUH,PDK4,PGM5,RPL38,RPL37,RPS9,GLRX3,PRKCB,RPS5,APLP2,RPL13A,PLA2G4A,PRKCA,EEF1A1,ITPKB,INPP4B,DPYD,NDUFS4,SPARCL1,PPARG,SUCLG1,DLD,PPM1L,PDE1C,COX5A,ACACB,PPP2CA,ATP5B,UGP2,UBB,RPS3,IGFBP7,S1PR3,COX10,LAMB2,IDH2,GOT2,FN1,PHKB,COX6C,SOD1, SPCS3,PDE10A,GNPDA2,NDUFAB1,OGDH,ACO1,CALM2
TCFAP2C	1.3E-03	1.4E-02	PYGB,PIGT,YWHAB,RPLP1,COX4I1,AMD1,SLC2A1,LAMC1,COX6A1, PIK3C2A,RPL8,PRSS23,PITPNC1,CYR61,HK2,RPL6,UBL5,PCMT1,LGALS1,RPS18,CYP1B1,RPL35,RPL38,ACADM,RPS11,LMCD1,IGFBP5, GPX4,FNBP1,APLP2,ECH1,ITPK1,SDHC,RPS3A,SDHD,EEF1G,HADHB ,UGDH,MFGE8,PRPS1,OAT,PPM1L,ATP5A1,ODC1,GATA6,ITPR1,NEDD8,ITPR3,COX7A2,COX5A,ATP5L,PPP2CA,UGP2,UBB,OXCT1,ATP6V0A1,ATP6V1F,LAMB2,TXNRD1,IDH2,H3F3A,FN1,SOD1,RPIA,CALU

			,MSRB3,CALM1,PAM,RPS21,PICALM
NANOG	2.0E-03	2.0E-02	PRPS1,RPL30,YWHAB,SLC2A3,ATP5G2,RPL9,CYR61,HK2,MRPL11,CST3,CSRP2,RPS19,UBB,ALDH1B1,PMPCB,PIP4K2A,RPL15,RPL18,NDUFA9,PDHA1,APLP2,IDH2,PHKB,IGF1,GSTA4,SMS,CALM3,RPS23
CUX1	2.2E-03	2.2E-02	CHRM2,AMD1,SLC2A1,MTR,SLC2A3,ATP5G3,COX7C,HK2,RPS4X,C4A,UBL5,MRPL42,PCMT1,RPS16,RPS19,RPS18,RARS,PMPCB,RPL37,DLAT,RPS11,B2M,RPS13,PDHX,RPL21,MRPS21,SARS,MRPS18A,SDHC,SDHD,MUT,EEF1A1,HADHB,DNAJC3,HADHA,SDPR,PPARG,SUCLG1,RPL27,UQCRC2,PRPS1,NDUFB9,FH,OAT,UQCRB,PPM1L,MGST3,GATA6,GSTT1,TXN,LTBP1,UQCRH,MRPL11,PPP2CA,ATP5B,ESD,RPL14,RPL15,RPL18,RPL17,RPL19,ATP6V1F,USP25,GSN,MDH1,NDUFA2,H3F3A,COX6C,SOD1,RPIA,RPS28,PDE10A,GSTA4,RPS29,PAM,CALM2,UBE2M
ESR2	2.4E-03	2.3E-02	PRPS1,ATP6V0E1,WFS1,MAOA,SLC2A1,PLOD2,LAMC1,PITPNC1,BACE1,UGDH,PLA2G16,RPL13,PIP4K2A,CYP1B1,S1PR3,CHPT1,LMCD1
CCND1	2.8E-03	2.4E-02	YWHAE,RPL4,APP,PIGT,MRPS14,RPLP1,PTEN,LAMC1,SLC2A4,ATP5G2,PITPNC1,CYR61,UBL5,PCMT1,RPS16,RPS19,ALDH2,PPP2R1A,RPL36,PMPCB,ATP6V1E1,PELO,PLA2G12A,RPS9,IGFBP5,RPS5,TALDO1,SARS,MUT,PGM2L1,HADHB,HADHA,OAT,NDUFB10,NDUFB6,MGST3,AK2,ITPR1,PPM1K,ATP5O,OAZ1,ATP5L,ATP5B,UBB,ATP5D,RPL18,RPL19,ATP6V1F,WFS1,GOT2,NDUFA2,FN1,RPS25,ALDH6A1,OGDH,MSRB2,SERINC1
OLIG2	2.7E-03	2.4E-02	CHRM3,APP,PIGT,PTEN,PLOD2,MTR,PRSS23,PLB1,AUH,RARS,PDK4,SLC16A7,RPL38,DLAT,PRKACB,LMCD1,COA5,PDHX,ATP6V0E1,GLRX3,SARS,PRKCA,HADHA,INPP4B,PLCB4,SDPR,NDUFS5,PLSCR4,DPYD,NDUFS4,PRPS1,PDE1C,MAOB,MAOA,GATA6,ITPR3,OXCT1,IGFBP7,DARS,USP25,NDUFA4,PLCL1,PHKB,CDC42BPA,ASS1,SOD1,RPIA,GATM,RPS27,GNPDA2,RPS29,CALU,MSRB2,PAM,PLCD4
FOXM1	2.8E-03	2.4E-02	HIBADH,COX4I1,RPLP0,TNC,SLC2A1,PLOD2,ATP5G3,LAMC1,ATP5G2,PITPNC1,CYR61,GLS,MRPL42,ATIC,SLC16A7,PIP4K2A,CYP1B1,ACADM,COX8A,PDHX,IGFBP5,SLC2A12,PLA2G4A,ACSL3,PGM2L1,CIT,EEF1G,UGDH,INPP4B,PLCB4,SDPR,NDUFS4,SMS,SPARCL1,PPARG,SUCLG1,NDUFB9,PPM1L,PDE1C,ODC1,ITPR1,MGST1,PPM1K,COX5B,LTBP1,PPP2CA,RPS15A,UGP2,BCHE,MDH1,PLCL1,TXNRD1,CDC42BPA,CALM2
TBX20	2.8E-03	2.4E-02	YWHAE,DGKG,PYGB,RPLP0,SLC2A1,MTR,ATP5G2,RPL8,RPL9,CYR61,HK2,ATIC,IMPA2,PDK4,LMCD1,GSTO2,GPX3,PRKCA,CIT,EEF1A1,DNAJC3,PLCB4,PIK3CA,PPARG,RPL27,GAS6,SLC25A5,PRKAA2,NDUFB7,PDE1C,SDC2,MGST3,ATP5A1,GATA6,ITPR1,PDHB,ATP5H,COX5B,ACACB,ACAT1,CST3,RPS15A,UGP2,ABL1,EIF4EBP2,IGFBP7,KDSR,IP6K2,GSN,MDH1,IDH2,GOT2,FN1,FBN1
FOXA1	2.8E-03	2.4E-02	PYGB,RPL5,MRPS14,RPLP1,SPPL2A,LAMC1,PRSS23,CYR61,MRPL42,ATIC,ACADM,LMCD1,COX8A,GLRX3,GPX4,GSTO1,PGAM1,APLP2,T

			ALDO1,MRPS21,ATP5F1,PGM2L1,UGDH,PLA2G16,PPA1,NDUFS5,DPYD,RPL26,GAPDH,ALDH9A1,PRKAA2,ADH1C,MGST3,AK2,UQCR11,PPM1K,UQCRH,ATP5L,UGP2,COX11,KDSR,IGFBP6,ATP6V0A1,CYP2J2,SLC35A1,RPL41,H3F3A,CDC42BPA,RPS26,CYP2C9,RPS27, NDUFAF1,CALM2,PICALM
ERG	3.3E-03	2.8E-02	YWHAE,RPL4,RPL5,PIGT,NDUFA11,RPLP0,RPS15,ATIC,RPS16,IMPA2,IDH3B,PMPCB,IDH3A,RPL21,GPX4,GSTO1,RPS5,TALDO1,PGM2L1,EEF1G,PLSCR4,RPL24,YME1L1,SUCLG1,RPL26,ALDH9A1,PRPS1,ATP5A1,ODC1,GMPS,AK2,ITPR1,ITPR3,COX5B,UBB,OXCT1,ATP5D,RPL14,COX11,TMEM208,RPL13,RPL17,ATP6V1F,MDH2,TXNRD1, RPL35A,SOD1,CALU,OGDH,MSRB2,ACO1,RPS21,RPS23
MYCN	3.8E-03	3.2E-02	YWHAE,ACAA2,HIBADH,PIGP,LAMC1,RPL9,PITPNC1,GLS,RPS4X,PCMT1,RPS16,RPL38,RPS11,PGAM1,RPL13A,MRPS18A,PGM2L1,PPA1,UBA52,PPM1L,COX15,NDUFB5,MAOA,ODC1,ITPR1,ITPR3,PDHB, COX5A,UQCRH,ADH5,ATP5L,PPP2CB,UGP2,UBB,ATP5D,RPL14,RPL19,RPL41,WFS1,TXNRD1,H3F3A,FN1,RPL35A,PHKB,ASS1,SOD1,RPS26,ALDH6A1,RPS28,RPS29,NDUFAB1,RPS20,CALM3,NDUFAF1, CALM1,PAM,RPS21,CALM2,PICALM
PPAR	4.1E-03	3.4E-02	PYGB,FH,MAOA,MGST3,GATA6,AK2,PLOD2,ATP5O,FSTL1,LTBP1,HK2,MRPL3,AUH,PDGFC,RPS3,PDK4,CYP1B1,BCHE,TXNRD1,ECH1,SIAH1,SARS,MRPS18A,ACSL3,UQCRHL,EEF1A1,HADHB,HADHA, INPP4B,NDUFS5,CAT,ACO1,RPL26,PICALM,FBN1
SALL4	4.9E-03	4.0E-02	NDUFB6,NDUFB5,MRPS14,MGST3,AMD1,PTEN,ATP5I,SLC2A3,ATP5G3,PRSS23,PCMT1,RPS15A,RPS16,UBB,ABL1,IDH3B,IGFBP7,IGFBP6,NDUFV2,RPS11,PELO,B2M,SLC35A1,ENTPD1,RPL41,RPS9,MDH1,NDUFA4,MMP2,RPS5,H3F3A,SARS,PLA2G4A,HADHA,PLSCR4, CYCS,RPL27,RPS20,RPL26,RPS23
EGR1	5.4E-03	4.3E-02	HADHB,NDUFB9,H3F3A,ITPK1,GMPS,PIP4K2A,ATP5I,KDSR,ATP5H, CYP4F11,CRYM,PRSS23
FOXP2	5.5E-03	4.3E-02	APP,HIBADH,YWHAB,PIGP,ADARB1,PLB1,PCMT1,UGP2,AUH,UBB, RARS,PDGFC,PDK4,EIF4EBP2,PIP4K2A,NDUFV2,USP25,ATP6V0E1, GOT2,H3F3A,PLA2G4A,MUT,RPIA,ITPKB,SPCS2,GSTA4,CYCS
THAP11	5.5E-03	4.4E-02	YWHAE,NDUFB7,ITPR1,ATP5J,NEDD8,MTR,ATP5O,ATP5G2,COX7C, RPS18,ATP5E,RPL36,RPS3,RPL15,RPL18,ATP6V1F,MDH1,GOT2,PGM2L1,RPIA,BACE1,RPS26,SPCS2,SUCLG1,RPS20,RPL26,RPS23
FUS	6.1E-03	4.7E-02	PRKAA2,MAOA,PLA2G4A,PHKB,MTR,ATP5G3,RPL9,FSTL1,COX7C, MUT,CYR61,ACAT1,EEF1A1,RPS4X,RARS,ATP5E,RPL15,PAM, GAPDH

NOTE: FC= Fold Change; p-val= p-value; FDR p-val= FDR p-value.

Table 4.8 Putative binding motif enrichment analysis using the Enrich webtool and showing the relative relevance of various histone marks in the deregulated transcripts.

Term	p-val	Adj. p-val	Genes
H3K79me2	7.2E-24	3.0E-21	NDUFA11;NDUFA12;ATP5C1;COX6A1;RPL9;RPL6;RPS15;LGALS1;RPS16;RPL18A;RPS18;RPL35;PMPCB;RPL37;RPS11;RPS10;RPS13;PDHX;RPS9;IGFBP5;RPL21;SDHC;SDHD;MRPS18C;COX6B1;EEF1A1;RPL27;RPL26;UBA52;UQCRB;SDC2;ITPR3;ATP5O;OAZ1;UQCRH;ADH5;ATP5B;UBB;ESD;NDUFV2;COX10;ATP6V0A1;DARS;RPL41;ATP8B2;MRPL24;RPS26;RPS25;RPS28;NDUFAF4;CALU;CYCS;RPS20;ACO1;RPS21;RSL24D1;RPS23;COX20;MRPS14;RPLP1;RPLP0;ATP5G3;PFAS;CYR61;RPS4X;UBL5;MRPL3;CSRP2;ATIC;IDH3B;COA5;GPX4;GSTO1;MRPS21;RPS3A;ATP5F1;MRPS7;CIT;EEF1G;NDUFS5;NDUFS4;DLD;NDUFB7;NDUFB6;RPL12;ATP5A1;NDUFB3;UQCR10;PDHB;COX5B;FSTL1;PPP2CB;RPS3;TMEM208;RPL15;RPS2;RPL18;RPL17;USP25;NDUFA6;LAMB2;NDUFA2;RPL35A;RSL1D1;NDUFAB1
H3K36me3	6.4E-18	6.5E-16	COX7B;RPL30;MRPS14;YWHAB;NDUFA12;RPLP1;COX4I1;ECI2;SPPL2A;RPLP0;ATP5G3;COX7C;CYR61;RPS15;MRPL3;RPS14;ATIC;RPS16;ALDH2;RPS19;RPS18;RPL36;RPL35;RPL38;ACADM;RPL37;DLAT;RPS11;PELO;RPS10;RPS13;COA5;COX8A;RPL21;GSTO2;GPX4;ATP6AP1;ECH1;MRPS21;SDHD;ATP5F1;MRPS18C;EEF1A1;EEF1G;PIK3CA;RPL24;RPL27;RPL26;UQCRC2;SLC25A5;GAPDH;UBA52;NDUFB10;COX15;RPL12;ATP5A1;MGST1;ITPR1;ATP5J;PDHB;ATP5H;ATP5O;COX5B;ACACB;ADH5;C3;ATP5B;ATP5E;ATP5D;RPL14;EIF4EBP2;RPL13;TMEM208;RPL15;RPL18;RPL17;RPL19;FN1;IGF1;COX6C;SOD1;RPS26;SPCS3;RPS25;RPS28;RPS27;RPS29;NDUFAB1;CYCS;CHPT1;RPS20;RSL24D1;COX20;RPS23
H3K79me3	6.4E-18	6.5E-16	MRPS14;NDUFA12;RPLP1;COX4I1;ATP5C1;PLOD2;COX6A1;COX7C;RPL9;RPL6;RPS4X;RPS15;MRPL42;PCMT1;LGALS1;RPS16;RPS19;RPL18A;RPS18;RPL36;PMPCB;ATP6V1E1;RPS11;RPS10;RPS13;RPS9;GPX4;GLRX3;RPS5;RPS3A;SDHC;SDHD;MRPS7;CIT;COX6B1;EEF1G;EEF1A1;PPA1;NDUFS5;NDUFS4;NDUFS3;RPL27;RPL26;SLC25A5;DLD;UBA52;GAPDH;OAT;NDUFB6;UQCRB;NDUFB5;NDUFB3;COX7A2;COX5B;ATP5O;OAZ1;ADH5;UBB;ATP5E;ESD;RPS3;TMEM208;IGFBP7;RPL15;RPS2;NDUFV2;RPL18;RPL17;RPL19;NDUFA9;RPL41;NDUFA6;MDH1;LAMB2;NDUFA2;H3F3A;RPL35A;RPS25;RSL1D1;RPS28;SPCS2;RPS29;NDUFAF4;NDUFAB1;NDUFAF2;CALU;CYCS;RPS20;CALM3;NDUFAF1;CALM1;RSL24D1;RPS21;RPS23
H4K20me1	4.0E-09	7.5E-08	RPL4;RPL5;ACAA2;RPLP1;ECI2;RPLP0;ATP5C1;SLC2A3;ADARB1;ATP5G3;RPL8;CYR61;UBE2L3;RPS15;RPS14;RPL18A;RPL36;RPL35;SLC16A7;RPL38;RPL37;RPS11;RPS9;RPL21;FNBP1;RPS5;MMP2;RPL13A;RPS3A;PRKCA;MRPS7;EEF1A1;UGDH;PPA1;YME1L1;RPL24;SUCLG1;RPL27;UQCRC2;RPL26;UBA52;GAPDH;OAT;MAOB;PPM1L;R

			PL12;ATP5J;UQCR10;PPM1K;OAZ1;ADH5;ITPRIPL2;ATP5B;UGP2;UBB;ABL1;RPL13;RPS2;RPL15;NDUFV2;RPL18;RPL17;ATP6V0A1;IP6K2;RPL19;MDH2;GOT2;RPS26;RPS28;MSRB3;RPS20;CALM3;ACO1;FBN1
H3K27ac	3.0E-07	4.0E-06	CHRM2;ACAA2;NDUFA11;NDUFA12;SPPL2A;COX6A1;PWP1;PCMT1;RPS19;RPL36;PDK4;ATP6V1E1;PELO;PRKACB;RPS10;RPS9;RPL21;IGFBP5;RPS5;MRPS18A;SDHC;ACSL3;SUCLG1;UQCRC2;DGKH;COX15;ITPR1;GSTT1;ITPR3;COX7A2;UQCRH;ITPRIPL2;UBB;ATP5E;ESD;COX11;IGFBP7;COX10;SIAH1;PHKB;RPS26;RPS25;RPS27;RPS29;NDUFAF4;OGDH;SERINC1;NDUFAF1;CALM1;PAM;FBN1;RPS23;FARSB;PYGB;YWHAB;RPLP1;AMD1;SLC2A1;MIA3;MRPL35;PITPNC1;SMPD2;IDH3B;PGM5;AASS;IDH3A;PLA2G12A;GSTO2;SLC2A12;PLA2G4A;MRPS21;MRPS7;CIT;HADHB;HADHA;UGDH;CAT;NDUFS3;NDUFB9;NDUFB5;NDUFB3;ODC1;UQCR10;COX5B;PPP2CB;RPS15A;OXCT1;TMEM208;RPL13;RPL15;RPS2;RPL19;NDUFA9;BCHE;NDUFA6;MDH1;USP9X;NDUFA4;TXNRD1;NDUFA2;RPL35A;RPIA;RSL1D1;SPCS2;MSRB3;PLCD4
H2AFZ	1.1E-04	9.8E-04	YWHAE;DGKG;PIGT;HIBADH;NDUFA12;MRPS14;PIGP;SLC2A1;MIA3;ATP5G3;ATP5G2;RPL6;RPS14;CSRP2;ATIC;ALDH2;PPP2R1A;CYP4V2;RPL36;PMPCB;UQCRFS1;PGM5;RPL37;PELO;RPS10;COX8A;PLA2G12A;GPX3;SDHC;ACSL3;MRPS7;ATP5F1;CIT;INPP4B;UGDH;PPA1;SMS;SUCLG1;UBA52;PRPS1;NDUFB9;NDUFB7;NDUFB5;NDUFB3;AK2;COX7A2;PDHB;ATP5L;RPS15A;ATP5D;RPL14;EIF4EBP2;COX10;MDH2;PLCL1;NDUFA2;RSL1D1;RPS20;SERINC1;RSL24D1
H3K4me3	6.1E-04	4.3E-03	CHRM2;RPL4;PYGB;SPPL2A;ECI2;SAT2;ATP5G2;CHRDL1;PCMT1;MRPL3;MRPL42;ATIC;ALDH2;PDK4;UQCRFS1;ACADM;DLAT;ATP6V1E1;B2M;AASS;IDH3A;RPS9;GSTO2;ITPK1;ECH1;MRPS7;DPYD;NDUFS4;MFGE8;GAPDH;DLD;PDE1C;NDUFB6;MAOB;NDUFB10;COX15;SDC2;MGST3;ITPR1;FSTL1;ESD;TMEM208;IP6K2;PDHA1;NDUFA4;FN1;NDUFA1;RPS26;RSL1D1;ALDH6A1;RPS27;GSTA4;ACO1;NDUFAF1;PLCD4;FARSB;MSRB1
H3K9me1	1.0E-03	6.3E-03	RPL4;RPL30;NDUFA11;RPLP1;RPLP0;ATP5C1;ATP5G3;ATP5G2;GLS;LGALS1;ATIC;RPL36;PMPCB;RPL35;RPL37;RPS11;RPL21;ATP5F1;COX6B1;PPA1;UQCRC2;RPL26;MFGE8;UBA52;SLC25A6;FH;MAOB;COX15;MAOA;RPL12;ACACB;ACAT1;ITPRIPL2;RPS15A;UGP2;ALDH1B1;ATP5E;RPL14;TMEM208;EIF4EBP2;RPL13;RPS2;RPL15;RPL18;RPL17;ATP6V0A1;GSN;ATP8B2;MDH1;MDH2;FN1;UQCRHL;RPS28;GNPDA2;CHPT1;FARSB
H3ac	2.8E-03	1.4E-02	NDUFA11;NDUFA12;COX4I1;RPLP1;SLC2A1;ATP5C1;PLOD2;PRSS23;MRPL35;CSRP2;MPC1;IDH3B;ACADM;PRKACB;PELO;RPS10;RPS13;PLA2G12A;GSTO1;APLP2;ITPK1;RPS3A;NME4;SDHD;MRPS7;HADHB;PPA1;NDUFS5;NDUFS3;UQCRC2;PRKAA2;PPM1L;NDUFB5;UQCR11;NEDD8;COX7A2;FSTL1;MRPL11;COX11;TMEM208;NDUFV2;SLC35A1;DARS;USP25;NDUFA6;NDUFA4;RSL1D1;NDUFAF4;NDUFAF2;CYCS;MSRB3;ACO1;PICALM;COX20

H3K9 me3	1.1E-02	4.5E-02	RPL5;PIGT;NDUFA12;YWHAB;RPLP1;RPLP0;ATP5C1;ATP5G2;RPL6 ;RPS14;RPS16;AUH;RPL18A;ALDH2;RPS19;RPL35;ACADM;RPS10; RPS13;RPL21;RPL13A;RPS3A;SDHD;MRPS7;EEF1A1;HADHB;HADH A;CAT;NDUFS4;RPL27;UQCRC2;UBA52;MAOA;AK2;COX7A2;TXN; ATP5B;RPS3;RPL13;RPS2;RPL18;ATP6V0A1;RPL41;NDUFA6;COX6 C;RSL1D1;MSRB2;RPS20;CALM3;RPS21;FBN1

Note: FC= Fold Change; p-val= p-value; FDR p-val= FDR p-value.

Chapter 5: Search of non-invasive metabolic biomarkers for diagnosis and monitoring of BC by ^{1}H NMR

Introduction

Due to the recurrent nature and the risk of progression towards invasive stages, NMIBC requires long-term surveillance by cystoscopy, an invasive, unspecific and expensive technique[13,164]. Could the patients be monitored by the use of non-invasive and dynamic urinary biomarkers, thus avoiding the use of cystoscopies?

In recent years the use of non-invasive approaches is gaining significant attention in the biomarker discovery field[165–167]. Its application to BC management would allow reducing the morbidity and the costs associated with cystoscopy, but also could provide information about tumor phenotype, which cystoscopy does not.

The direct contact between the tumor and the urine is a major characteristic of the urological cancers, since molecules released from bladder cancer cells may be enriched in urine samples[168]. Besides, urine can be obtained non-invasively, so it could be used in active BC surveillance programs without causing patient discomfort.

^{1}H NMR analysis requires minimal sample preparation, are rapid and non-destructive, reproducible and cost-effective. Therefore, considering the advantages of ^{1}H NMR studies and the current lack of biomarkers for the follow-up of patients with NMIBC, we developed an experimental design pre-TUR (BC) vs post-TUR (control) with the purpose of searching a non-invasive biomarker capable to detect recurrences in NMIBC patients undergoing a surveillance program.

Material and Methods

Patient selection and sample collection

In this study 20 males and 8 females diagnosed with BC undergoing planned TUR were included from the Urology Service of the Hospital Universitario y Politécnico La Fe (Valencia, Spain). The study was approved by the Ethics Committee for Biomedical Research of the Instituto de Investigación Sanitaria La Fe (Valencia, Spain) and all patients gave written informed consent to participate in it. Inclusion criteria for patient selection were: male or female from 20-90 years old with NMIBC diagnosed, single or multiple tumors and primary or recurrent tumors. Exclusion criteria were: patients with a urinary catheter, invasive tumor diagnosed (T2-T4), papilloma pathological diagnosis or only CIS diagnosis by pathological anatomy (PA).

After recurrent risk group classification according to the EORTC criteria, patients were included in a monthly monitoring group to collect urine samples in case of a new recurrence. Patients were followed for a period of 18 months in which a total of 153 urine samples were collected. Urines were processed and stored at -80˚C by the Biobanco La Fe. Samples were considered as BC (n=70) when cystoscopy was positive and PA confirmed the presence of a tumor; and one month before positive cystoscopy with tumors ≥3 cm. Non-tumor samples collected after TUR of tumor were split into two groups: CTRL (n=29) and MONITOR (n=38). CTRL group included urines collected from NMIBC patients within 2-4 weeks after TUR; and MONITOR group included urines collected during monitoring period with negative PA (T0), urines with negative cystoscopy at the time of sampling, and urines collected during monitoring period between negative cystoscopies. If no cystoscopy was available at the time of sampling or after, urines were classified as 'non available cystoscopic evaluation' (NA) (n=16).

Table 5.1 summarizes clinical-pathological and demographic data of patients and samples included in the study.

Table 5.1 Clinical-pathological and demographic data of patients with NMIBC included in this study.

PLS-DA model	Calibration (CV)	Validation CTRL+MONITOR
Male/ female patients	24 (18/6)	7 (5/2)
Mean age (standard deviation)	70 (11.15)	63 (5.13)
Total samples	69	84
CTRL	21	8
MONITOR	0	38
NA	0	16
BC	48	22
Primary/Recurrent BC	13/35	4/18
Tumor stage (Ta,T1,Tx)	33/13/2	17/4/1
Tumor grade (High/Low)	32/15 (UK:1)	20/1 (UK:1)
Tumor size (≥3/<3)	11/35 (UK:2)	7/15
Tumor number (1/2-7/≥8)	18/28/0 (UK:2)	14/8/0

Note: In the validation set 16 NA samples were included for view the trajectory of urinary profile but not were considerate to create the PLS-DA model. 10 MONITOR samples from 3 patients included in the calibration set were included in the validation set. UK: unknown.

Sample preparation and H^1 NMR acquisition

Urine samples were thawed at room temperature and were processed in the same way that was described above (Chapter 4)[134]. Briefly, 200 µL phosphate buffer (pH 7.4) were added to 500 µL urine. The mixture was then homogenized (vortex) and centrifuged at 10000 rpm for 5 min at 5˚C and transferred to an NMR tube. For each urine sample, 1D ^{1}H water presaturation spectra were acquired at 300˚ K using a Bruker Avance II 500 MHz spectrometer in 16 min: 46 s and the following spectral acquisition parameters: acquisition time: 2.72 s; transients: 128;

spectral width: 12 ppm (6000 Hz) and relaxation delay: 5 s. Moreover, 2D ^{1}H-^{13}CHSQC spectra were acquired, to assess the assignments of the overlapped signals in 1D ^{1}H spectra.

^{1}H NMR spectra pre-processing and metabolite assignment

After 1D ^{1}H water presaturation spectra acquisition, the FID was Fourier-transformed, phase and baseline corrected and chemical shift referenced to DSS at 0.0 ppm using *MestReNova* version 6.0.2 (Mestrelab Research SL, Santiago de Compostela, Spain). Metabolites were identified and assigned according to the published data[92,135,136,169] and open NMR databases[77,137] considering a peak tolerance ±0.02 ppm. The spectra were binned into 0.003 ppm buckets using *MestReNova*. For the statistical analysis, the 0.8-4.5 and 6.5-9.0 ppm chemical shift spectral regions were considered, excluding thus the regions containing water and urea resonances.

Data analysis

The intensity of the assigned resonances was transferred to *MetaboAnalyst 3.0*[138]. Mean metabolite intensities were normalized by the sum (1-norm) and compared between groups. U-Mann Whitney test determined the significant differences between control and BC urines. Kruskal-Wallis evaluated differences among control, primary tumors and recurrence tumors; and compared the mean among Ta, T1 and control samples.

Multivariate statistical analysis was carried out in first-derivative row-wise normalized 1D spectra using autoscaling as data pre-processing. This was performed using the *PLS_Toolbox Solo 8.0* (Eigenvector Research Inc., Manson, WA, USA). A PCA of the set of BC, CTRL, MONITOR and NA samples was carried out to detect potential outliers. After, the data set was split into a calibration and

validation subsets. 69 samples from 24 patients were included in the calibration set (48 BC and 21 CTRL) and were used to develop a PLS-DA model. The validation set was used to evaluate the model predictive performance of the PLS-DA model, and included urines from 7 patients with different clinical evolutions. Firstly, the validation set was performed considering (22 BC and 8 CTRL). After, MONITOR (n=38) and NA (n=16) samples (not included in the calibration set) were added to the validation set to evaluate the performance of the model to diagnose BC recurrence during a surveillance period **(See table 5.2)**. Adding MONITOR samples the number of control samples was increased to 46.

Table 5.2. **Patients and samples used in the calibration and validation set of PLS-DA model.**

	Calibration		Validation		
Patient	**BC**	**CTRL**	**BC**	**CTRL**	**MONITOR**
1	3	0	--	--	--
2	1	0	--	--	--
3	3	1	--	--	--
4	3	1	--	--	--
5	1	0	--	--	--
6	2	1	--	--	--
7	2	2	--	--	--
8	1	1	--	--	--
9	1	1	--	--	--
10	1	0	--	--	--
11	1	1	--	--	--
12	2	1	--	--	--
13	2	1	--	--	--
14	1	0	--	--	--
15	1	0	--	--	--
16	1	1	--	--	--
17	3	1	--	--	--
18	2	1	--	--	--
19	2	1	--	--	--
20	3	2	--	--	--
21	2	1	--	--	--
22	--	--	2	0	10

23	--	--	8	5	6
24	6	2	2	0	5
25	2	2	0	0	5
26	--	--	1	2	9
27	2	0	--	--	--
28	--	--	9	1	3

The optimal number of latent variables (LVs=3) was selected according to the RMSECV. A permutation test (100 permutations) was carried out to assess the statistical significance of PLS-DA figures of merit and the probability of a chance correlation. The relative importance of each metabolic feature in the PLS-DA model was determined using the VIP scores vector. The most important metabolites in the statistical model (VIP>1) were identified and used to perform a pathway enrichment and topology analysis using *MetaboAnalyst*.

Results

Urinary metabolomic profile in BC patients

Figure 5.1 displays a representative urine ^{1}H NMR spectrum from a BC patient and **Table 5.3** summarizes the assignment of the most relevant identified metabolites. Among them, amino acids (Ala, Phe, Gly, Lys), benzenoids (hippuric acid), organic nitrogen compounds (trimethylamine-*N*-oxide) and organic acids (lactate, citrate), creatinine and urea are highlighted.

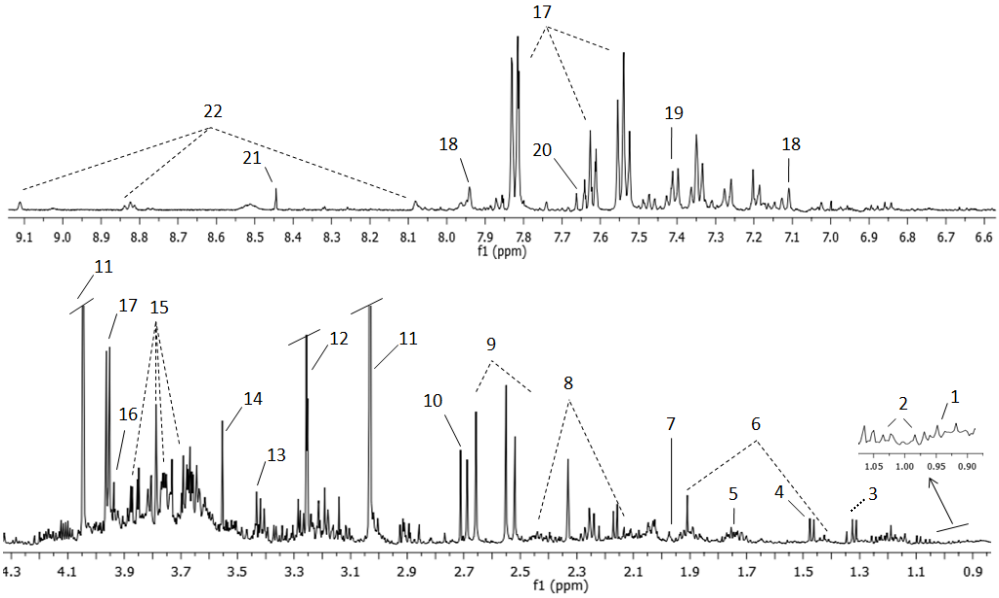


Figure 5.1 Representative 1D ^{1}H NMR spectrum and assignment of a urine sample from a NMIBC patient. The vertical scale of the entire spectrum was kept constant and spectral region of water and urea was removed from the figure. The intensity of peaks in the chemical shift region 6.6 – 9.1 ppm was enhanced to show the low abundant metabolites. Assigned metabolites: 1.Leu, 2.Val, 3.Lactate, 4.Ala, 5.DSS, 6.Lys, 7.Acetic acid, 8.Gln, 9.Citrate, 10.Dimethylamine, 11.Creatinine, 12.Trimethylamine N-oxide, 13.Taurine, 14.Gly, 15.Sucrose, 16.Creatine, 17.Hippuric acid, 18.His, 19.Phe, 20.Pseudouridine, 21.Formic acid, 22.Trigonelline.

Table 5.3 Assignment of the metabolites identified in 1D ^{1}H NMR urine spectra.

Nº	Metabolite	Chemical Formula	Chemical shift ppm
1	Leucine	CH_3	0.94
2	Valine	γCH_3	0.98
2	Valine	γCH_3	1.03
3	Lactate	CH_3	1.32
6	Lysine	γCH_2	1.43
4	Alanine	βCH_3	1.47
6	Lysine	βCH_2	1.71
7	Acetic acid	CH_3	1.97
6	Lysine	βCH_2	1.91
8	Glutamine	βCH_2	2.13
8	Glutamine	γCH2	2.44
9	Citrate	CH_2	2.51
9	Citrate	CH_2	2.65
10	Dimethylamine	CH_3	2.71
11	Creatinine	CH_3	3.03
13	Taurine	$-CH_2-NH_3^+$	3.25
12	Trimethylamine N-oxide	CH_3	3.29
13	Taurine	$-CH_2-SO_3^-$	3.42
14	Glycine	αCH	3.55
15	Sucrose	C6’ H_2	3.81
15	Sucrose	C5’H	3.87
16	Creatine	CH_2	3.92
17	Hippuric acid	αCH_2	3.96
11	Creatinine	CH_2	4.05
18	Histidine	CH	7.09
19	Phenylalanine	C2’6 H	7.33
19	Phenylalanine	C3’5 H	7.41
17	Hippuric acid	C3’5 H	7.63
20	Pseudouridine	CH	7.66
17	Hippuric acid	C2’6 H	7.82
18	Histidine	CH	7.93
21	Formic acid	CH	8.45
22	Trigonelline	C3’5 H	8.82
22	Trigonelline	C1’ H	9.11

Mean comparison of metabolites detected in urine

Hippuric acid and alanine levels showed significant differences between BC and control groups (*p*-value<0.05). Both metabolites presented higher intensities in control samples. Hippuric acid also presented significant differences between primary tumors and recurrent tumors and between the Ta and T1 stages of tumors and control samples **(Figure 5.2)**.

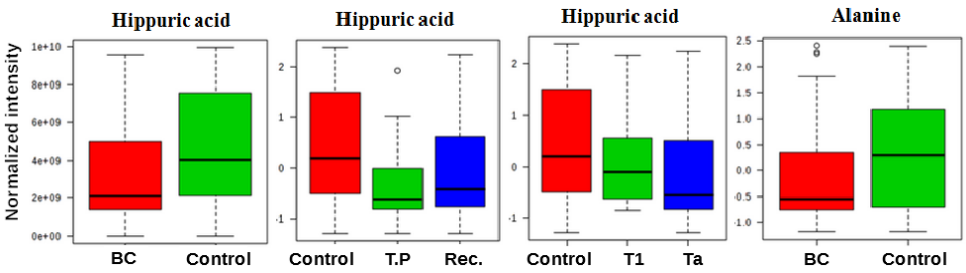


Figure 5.2 Box and whisker plots illustrating discrimination between: BC and control urines; control, tumor primary (T.P) and recurrences (Rec); and differences among stages (Ta and T1) of BC and control urines. The vertical axis shows the normalized intensity (1-norm).

Comparisons between BC and CTRL samples

Figure 5.3 depicts the PC1 vs PC2 scores of a PCA build for the unsupervised analysis of the initial data set. No significant clustering of BC and CTRL samples was observed in the PC1-PC2 scores space.

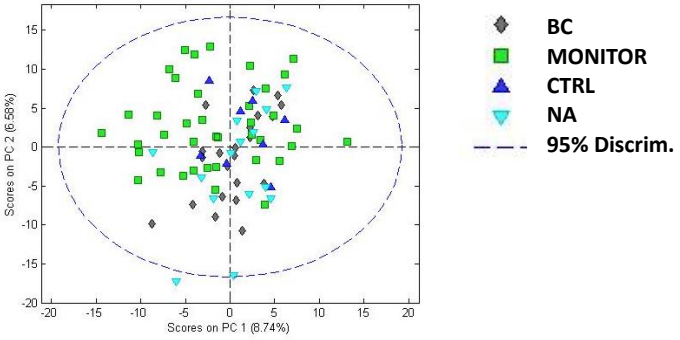


Figure 5.3 PCA of samples included in the validation set.

PCA provides an unbiased overview of the data structure but it might not reveal a significant between-groups clustering if the class is not among the main sources of variance in the data set. To overcome this potential limitation, a PLS-DA model was developed. **Figure 5.4** shows PLS-DA scores for the calibration set (LVs=3). Although PLS-DA scores should not be used to assess a difference between classes, this plot showed no sub-clustering neither in BC or CTRL sub-groups.

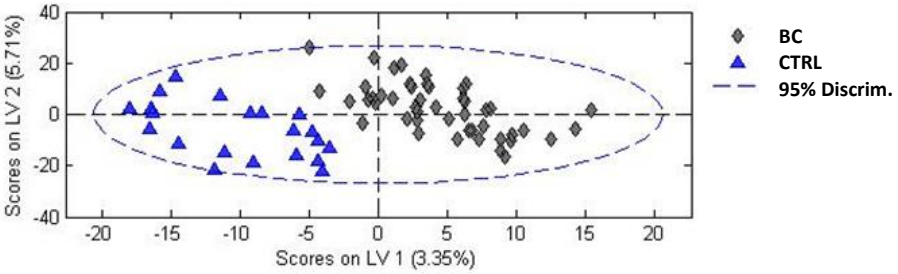


Figure 5.4 Scores plot for calibration set (BC *vs* CTRL).

For validation sets, PLS-DA scores plots and predicted values are shown in **Figure 5.5** and all statistical parameters evaluated are summarized in **Table 5.4.** Confusion tables obtained from the evaluation of the predictive performance of PLS-DA are depicted in **Table 5.5.**

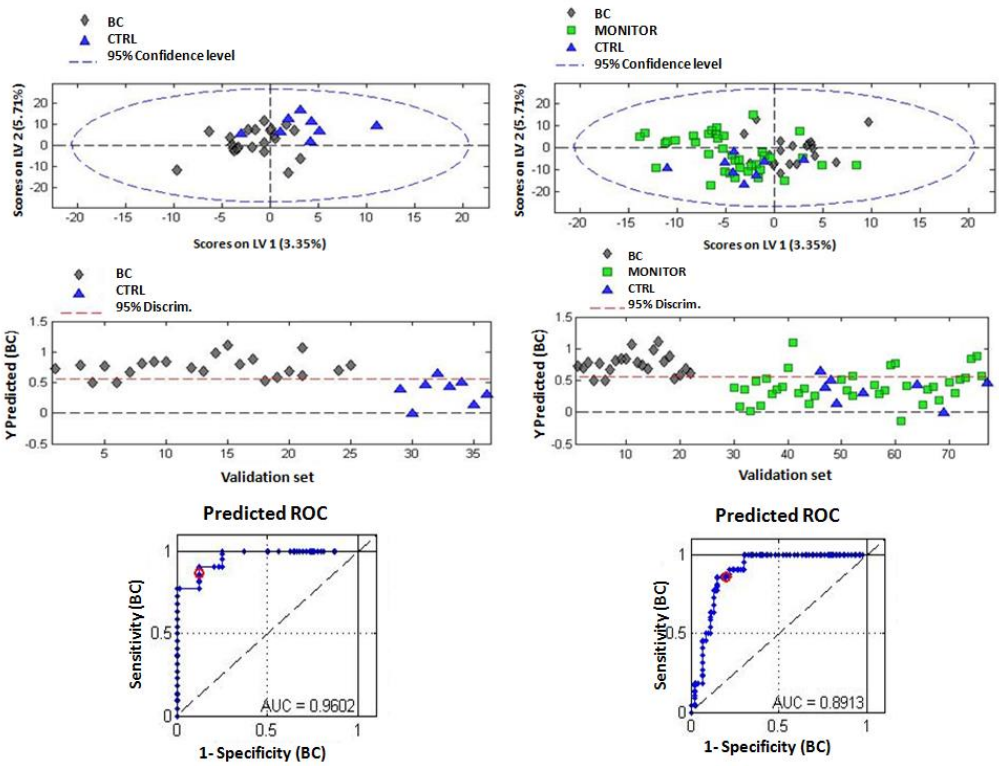


Figure 5.5 Discriminant analysis of BC, CTRL and MONITOR samples. (Left) Scores plot, PLS-DA predicted *y* values and AUROC for the first validation set (BC vs CTRL); (Right) Scores plot, PLS-DA predicted *y* values and AUROC for the second validation set (BC vs CTRL+MONITOR as control samples). Number of LVs: 3.

Table 5.4 PLS-DA figures of merit for the discrimination between BC and control samples in the calibration and validation sets.

Indices Test Validity	Calibration set (BC vs CTRL)	Validation set (BC vs CTRL)	Validation set (BC *vs* MONITOR+CTRL)
	Estimation (95% CIs)	Estimation (95% CIs)	Estimation (95% CIs)
True prevalence	69.6%	73.3%	31.9%
Sensitivity	81.3% (68.1%-89.8%)	86.4% (66.7%-95.3%)	86.4% (66.7%-95.3%)
Specificity	66.7% (45.4%-82.8%)	87.5% (52.9%-97.8%)	80.4% (67.5%-89.6%)
PPV[a]	84.8% (71.8%-92.4%)	95.0% (76.4% -9.1%)	67.9% (49.3% -2.1%)
NPV[b]	60.9% (40.8%-77.8%)	70.0 % (39.7%-9.2%)	92.7% (80.6%-97.5%)
ACC[c]	76.8 % (65.6%-5.2%)	86.7 % (70.3%-4.7%)	82.6 % (72.0%-9.8%)
PLR[d]	2.44(1.31-4.53)	6.91(1.10-43.54)	4.51(2.45-8.3)
NLR[e]	0.28 (0.15-0.52)	0.16 (0.05-0.45)	0.17 (0.06-0.49)

Note: [a]Positive predictive value; [b]Negative predictive value; [c]Diagnostic accuracy; [d]Positive Likelihood Ratio; [e]Negative Likelihood Ratio. PLS-DA model: LVs=3.

Considering only BC and CTRL samples in the validation set, the PLS-DA model provided a sensitivity of 86.4%, a specificity of 87.5% and an AUROC=0.96 with an ACC of 86.7%. When also were considered MONITOR samples, the sensitivity value was maintained but the specificity was 80.4%. Even so, AUROC was 0.89 and the NPV and NLR were 92.7% and 0.17 respectively. Permutation test (100 permutations) showed a statistically significant *p*-value<0.05.

Table 5.5 Confusion tables obtained from the evaluation of the predictive performance of PLS-DA models between BC *vs* CTRL and BC *vs* CTRL and MONITOR samples in the calibration and validation sets.

	Calibration (CV)		Validation		
	BC	**CTRL**	**BC**	**CTRL**	**CTRL+MONITOR**
Predicted as BC	39	7	19	1	9
Predicted as CTRL	9	14	3	7	37

The VIP>1 scores of this PLS-DA model identified, among the selected NMR features, the following discriminant metabolites: Val, Ala, Lys, Gln, citrate, dimethylamine, creatinine, trimethylamine N-oxide, taurine, sucrose, creatine, hippuric acid, histidine, Phe and trigonelline. The pathways analysis performed in *MetaboAnalyst*, linked these metabolites with alterations in taurine and hypotaurine; Ala, Asp, Glu; Arg and Pro; and Phe metabolic pathways (*p*-value <0.05) **(see Figure 5.6**; **Table 5.6).**

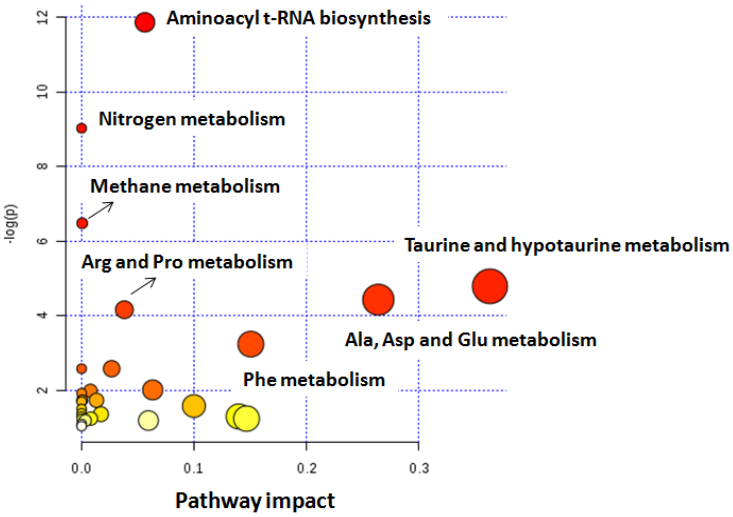


Figure 5.6 Analysis of altered metabolic pathways in bladder urines.

Table 5.6 Identified metabolites and associated metabolic altered pathways in urines.

Altered pathways in BC	Metabolites	*p*_value	Impact
Alanine, Aspartate and Glutamate	Ala, Gln, succinate	6.5E-4	0.26
Taurine and hypotaurine	Taurine, Ala	9.0E-3	0.36
Aminoacyl-tRNA biosynthesis	Phe, Gln, Val, Ala, Lys	1.6E-4	0.56
Methane	Trimethylamine N-oxide, methanol	2.5E-2	0.02
Arginine and Proline	Citrate, succinate	9.0E-3	0.08
Phenylalanine	Succinate, Phe, hippuric acid	4.0E-3	0.07
Nitrogen metabolism	Phe, taurine, Gln	2.7E-3	0.05

Metabolic changes occurred during the follow-up period of patients with NMIBC

The PLS-DA model correctly classified 19/22 BC and 37/46 as non-BC (i.e. CTRL or MONITOR) samples. Out of the total patients (n=28), 7 undergoing active follow-up had different clinical evolutions. The longitudinal trajectory of the metabolomic profile of these patients allowed assessing its utility to detect tumor recurrences during a surveillance period (e.g. patients 23*, 24*, 25*, 28* in **Figure 5.7**) given that its behavior was in agreement with the results of cystoscopy and PA after TUR. In the patients 23*, 24*, 25*, 28* a shift in the metabolomic profile, from control to tumor, was observed along with its detection by positive cystoscopy and it was confirmed by PA. The phenotype returned to control phenotype after the tumor removal. In patients who did not develop many recurrences the metabolic profile remained in a non-tumor phenotype after TUR during a monitoring period, and tumor absence was confirmed through negative cystoscopies (see patients 22*, 26* 25* in **Figure 5.7**). Besides, the metabolic profile was not affected by inflammatory processes (i.e. cystitis), since classified correctly non-tumor urines collected during these circumstances (see patient 23*) **(Figure 5.7)**. In patient 27* most of urinary samples were coded as NA due to the lack of clinical information associated with these.

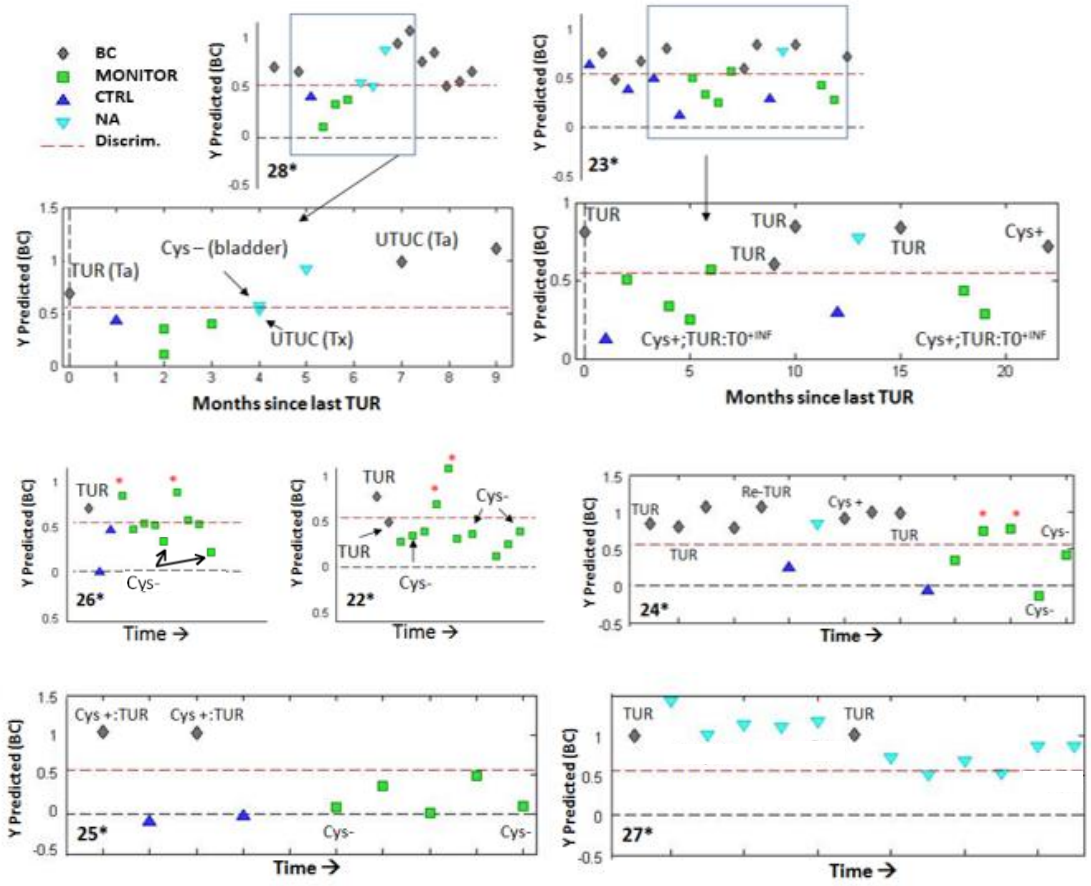


Figure 5.7 Analysis of longitudinal trajectories after TUR of BC. Predicted *y* PLS-DA values in 7 patients during the follow-up period. Note:(*) indicates a MONITOR sample showing an inconsistent trajectory; (Cys+ or Cys-) indicates a positive and negative cystoscopy; (TUR) mean transurethral resection; (T0) indicates tumor absence by PA evaluation; (Tx) indicates that pathologist did not confirm the presence of a tumor; (INF) indicates cystitis and (UTUC) means upper tract urothelial carcinoma.

Discussion

NMIBCs require complex clinical management due to the high heterogeneity of tumors. Patients with equal PA have completely different evolutions, indicating a unique biology for each tumor. Tumors represent dynamic entities that are constantly changing their molecular programs to adapt to the microenvironment conditions. That ability to adapt depends, in part, on metabolic reprogramming, a phenomenon recognized as an emerging hallmark of cancer[170]. Therefore, to analyze the metabolic changes produced in biological samples can offer opportunities for biomarker discovery.

With an experimental design pre-TUR (BC) *vs* post-TUR (control), we developed a PLS-DA model that provided an elevated sensitivity (86.4%), specificity (80.9%) and NPV (92.7%) classifying BC from control samples collected during the NMIBC surveillance after TUR. These results indicated that our metabolic profile could be very useful in clinical practice to detect recurrences since it would ensure a very low probability of having false negatives during a monitoring period. This fact was assessed observing the trajectories of the metabolic profile in 7 different clinical cases of patients with NMIBC undergoing active follow-up. The metabolic profile of patient 25* changed from BC towards control phenotype after tumor removal by TUR in the two developed tumor episodes. Negative cystoscopies confirmed the absence of recurrences in agreement with the metabolic profile trajectories. A similar profile was observed in patients 26* and 22*; although, in these cases 4 MONITOR samples were misclassified. Probably confounding factors (e.g. drugs, diet, treatment) affected the correct classification of urinary samples. The metabolic profile of patient 24* remained in a BC phenotype between the first complete TUR and a re-TUR performed 5 months later. The re-TUR was programmed due to the size of the

tumor (5-7cm) and the AP (Ta G3) of the previous tumor. During re-TUR three tumors were found with a PA (T1 G2). The metabolic signature was in agreement with an early detection of the recurrence. The patient 23* had a large number of recurrences since the diagnosis of the primary tumor. The classification of BC samples was in concordance with positive cystoscopy and PA of tumor after TUR (see samples collected in the months 9, 10 and 22). The correct classification of CTRL and MONITOR samples was in concordance with the negative results of PA (T0) (see samples in months 5 and 19). In these cases, the patient had been operated on suspicion of tumor (positive cystoscopy). However, the diagnosis of tumor specimen was cystitis. Remarkably, the patient 28* had a recurrence in the bladder (see month 0) and another in the upper tract (UTUC) with the same PA (see month 7). Lately, the metabolic signature detected the UTUC development, which was confirmed by PA in 7th month. Negative cystoscopies performed in the month 4 and 7 showed the absence of tumor in bladder epithelium.

In general, the obtained results indicated that the metabolic profile could be useful in very diverse clinical situations linked with NMIBC. The profile proved to be highly dynamic and sensitive in the detection of bladder recurrences not only in very early stages of their development but also in the detection of recurrences in the upper urinary tract. This fact has relevance considering the BC as a pan-urothelial disease. Moreover, in some cases, the metabolic profile was able to detect incipient tumors not detected by cystoscopy. From the clinical point of view, if these results were confirmed in further studies, they would be important since would allow personalizing the follow-up schemes to each patient, controlling better the disease and avoiding possible future progressions to MIBC. Besides, the application of this metabolic profile in the detection of incomplete TUR would be essential since would allow choosing the best option of treatment for each patient: a strict control by urinary cytology or cystoscopy, a re-TUR or a new regimen of chemo- immuno-therapy. By contrast, if the urinary signature

would indicate a control phenotype after TUR, a re-TUR could be avoided in patients who for clinical routine have to be operated again. Surgeries could be also prevented in those cases where cystoscopy confounds cystitis with tumors, but after TUR the PA shows tumor absence (i.e. T0). In brief, these data reinforce the idea that urinary metabolome reflects tumor biology and can be used to study tumor development. Nevertheless, although the obtained results are hopeful, a better understanding of how several intrinsic and extrinsic factors such as chemotherapy or BCG, drugs, and inflammatory processes affect the urinary metabolome is essential to improve the robustness of metabolic tests so they can be implemented into clinical routine.

On the other hand, knowing which metabolites take part in this profile is important to establish the link between altered metabolic pathways and the tumor phenotype, which indirectly reflects the events developed at genomic and transcriptomic levels. A better knowledge of metabolic regulation would allow detecting key metabolic enzymes that could be a target for the development of new BC therapies.

Among the metabolites identified as discriminants between BC and control urines (VIP>1) we spotlight the Gln, Glu, citrate, hippuric acid, and taurine. These metabolites have been described as relevant in several tumors (e.g. prostate, breast, ovarian) and their roles in cellular metabolism are quite known[171–173]. Gln and Glu have been described as two important metabolites in BC[133]. Cancer cells used Gln as a source of energy but also for nucleotide biosynthesis or the synthesis of other AAs[42]. GLS converts Gln to Glu, which can be used by GLT2 to produce intermediates of TCA cycle or be used for the synthesis of GSH, a tripeptide that acts as an important antioxidant in cells[174]. Gln uptake also links with the immune system, since several T cell metabolic processes require it[159]. Citrate is a key intermediate of the TCA cycle and has been closely

related to an increase in fatty acids β-oxidation to support cancer cell proliferation. Although we did not find significant levels of citrate between tumor and non-tumor urines, low concentrations of this metabolite have been reported in bladder tumors, suggesting the role of β-oxidation as source of energy[53]. Additionally, alterations in taurine metabolism have been showed in NMIBC[175,176], what is in concordance with our results. The lower levels of hippuric acid found in our BC urines compared to control samples are in agreement with other works[53,175]. However, we also identified differences in the levels of hippuric acid among urines collected from controls and patients with primary and recurrent tumors, suggesting that the metabolism of both types of tumors is not equal.

Overall, our results are supported by prior references that describe the importance of AA metabolism in BC[177] and are in concordance with the study that will be presented in the Chapter 6, in which a great part of the same set of urinary samples were analyzed by UPLC-MS[53]. Nevertheless, NMR-based metabolic profiling has advantages respect MS that include: minimal sample preparation, non-destructive analysis, higher reproducibility, and cost-effectiveness.

Conclusion

In summary, the present study shows for the first time a dynamic ^{1}H NMR-based urinary metabolic profile associated with NMIBC that changes from a tumor to a control phenotype after tumor removal and return to the malignant condition when a recurrence occurs. This fact highlights metabolomics as tool for searching non-invasive biomarkers, which could be applied in clinic to improve the management of BC patients by: 1) decreasing the performance of unnecessary cystoscopies during the follow-up period; 2) detecting lesions not visible by cystoscopy such as dysplasias, hyperplasias, CIS but also 3) detecting early recurrences, incomplete TUR or maybe UTUC.

Chapter 6: Search of non-invasive metabolic biomarkers for diagnosis and monitoring of BC by UPLC-MS

Introduction

Given the high rate of recurrence of NMIBC, the limitations of cystoscopy and cytology and the lack of non-invasive approaches applied intro clinical routine for diagnosis, but especially, for the follow-up of patients with this type of tumors, the identification of a surveillance biomarker is crucial.

Since the metabolome provides a direct meaningful readout of the dynamic biochemical status of a biological system, metabolomics is now considered as a highly relevant approach to explore individual phenotypes in systems biology of cancer[178]. NMR and LC-MS have been widely used to detect the majority of metabolites in a given biological sample. Each analytical technique has its advantages and disadvantages. However, its combined use allows attaining a coverage of detected metabolites that can not be achieved by single-analysis techniques[87]

Bearing this in mind, we conducted an investigative study for the analysis of urinary metabolome changes in NMIBC patients before and after TUR, as well as in a monitoring period, using ultraperformance liquid chromatography combined with time of flight mass spectrometry (UPLC-TOFMS). One of the purposes was to expand the study presented in Chapter 5 increasing the number of samples and using UPLC-MS as analytical technique, in order to increase the coverage of detected metabolites linked with NMIBC and detect surveillance biomarkers.

Material and Methods

Patient selection and sample collection

The study was approved by the Ethics Committee for Biomedical Research of the Instituto de Investigación Sanitaria Hospital Universitario y Politécnico La Fe (Valencia, Spain) (2012/0186). Urine samples were prospectively collected from patients that gave written informed consent to participate in the study. Patients with bladder tumor diagnosed and TUR surgery planned were invited to participate in the study. Inclusion criteria for patient selection were: male or female from 20-90 years old with NMIBC diagnosed, single or multiple tumors and primary or recurrent tumors. Exclusion criteria were: patients with a urinary catheter, invasive tumor diagnosed (T2-T4), papilloma pathological diagnosis or only CIS diagnosis by pathological anatomy (PA).

After recurrent risk group classification according to the EORTC criteria, some patients were included in a monthly monitoring group to collect urine samples in case of a new recurrence. In this study, we analyzed metabolomic profiles of 316 urine samples collected from 31 patients. **Table 6.1** summarizes the main features and pathological data of patients included in the study. Urine samples collected from patients diagnosed of BC by cystoscopy and tissue pathology were coded as BC. Those collected from NMIBC patients within 2-4 weeks after TUR were coded as CTRL. Urines collected after TUR with negative cystoscopy at the time of sampling and those collected in the course of regular visits to the urologist between negative cystoscopies were classified as MONITOR. If no cystoscopy was available at the time of sampling or after, samples were classified as NA cystoscopic evaluation and were not included in the estimation of figures of merit.

Table 6.1 Clinical-pathological overview of recruited patients.

	Calibration	Validation
Patients (male / female)	18 (13 / 5)	28 (23 / 6)
Samples (male / female)	53 (38 / 15)	210 (169 / 41)
PLS-DA model BC *vs* CTRL		
Samples pre-TURBT (BC)	35	33
Samples post-TURBT (CTRL)	18	11
Samples surveillance (MONITOR)	0	166*
Primary / Recurrent BC	8 / 27	7 / 23 (3: UK)
Tumor stage (Tx, Ta, T1)	1 / 21 / 13	0 / 21 / 3 (9: UK)
Tumor grade (I / II / III)	9 / 20 / 6	7 / 14 / 3 (9: UK)
PLS-DA model BC *vs* MONITOR		
Samples pre-TURBT (BC)	35	33
Samples post-TURBT (CTRL)	0	29*
Samples surveillance (MONITOR)	82	84
Primary / Recurrent BC	8 / 27	7 / 23 (3: UK)
Tumor stage (Tx, Ta, T1)	1 / 21 / 13	0 / 21 / 3 (9: UK)
Tumor grade (I / II / III)	9 / 20 / 6	7 / 14 / 3 (9: UK)
PLS-DA model CTRL *vs* MONITOR		
Samples pre-TURBT (BC)	0	68*
Samples post-TURBT (CTRL)	18	11
Samples surveillance (MONITOR)	82	84
Primary / Recurrent BC	0	15 / 50 (3:UK)
Tumor stage (Tx, Ta, T1)	0	1 / 42 / 16 (9: UK)
Tumor grade (I / II / III)	0	16 / 34 / 9 (9: UK)

Note: * indicates that these samples were not used for the estimation of the discriminant performance in that particular model. UK: unknown.

Sample preparation

Urine samples stored at -80°C were thaw at room temperature on ice, homogenized (vortex, 10 s) and centrifuged at 10000 x *g* (4°C, 10 min). Then, 100

µL of supernatant were withdrawn and 200 µL of HCOOH 0.1% v/v in H_2O were added. Sample was homogeneized (vortex, 10 s) and centrifuged at 10000 x *g* (4°C, 10 min). 100 µL of the supernatant was transferred to a 96 well plate where each sample was spiked with 5 µL of a solution of Phenylalanine- D_5 (Cambridge Isotopes Laboratory Inc., Andover, MA, USA), caffeine-D_9 (Toronto Research Chemicals, Toronto, Ontario, Canada), leukine enkephalin (Sigma-Aldrich Química SA, Madrid, Spain) and reserpine (Sigma-Aldrich Química) in H_2O:CH_3OH (1:1, 0.1% v/v HCOOH) providing a final concentration of 1 µM. Blanks were prepared by replacing urine by H_2O. A QC sample was prepared by mixing 5 µL of each sample. All solvents were of LC-MS grade and were purchased from Scharlau (Barcelona, Spain). Ultra-pure water was generated with a Milli-Q water purification system (Merck Millipore, Darmstadt, Germany). Formic acid (≥95%) was obtained from Sigma-Aldrich Química.

UPLC-TOF-MS sample analysis

Chromatographic analysis was performed on an Agilent 1290 Infinity UPLC chromatograph using an UPLC BEH C_{18} (100 x 2.1 mm, 1.7 µm, Waters, Wexford, Ireland) column. Autosampler and column temperatures were set to 4°C and 55°C, respectively and the injection volume was 4 µL. A gradient elution was performed at a flow rate of 400 µL min^{-1} as follows: initial conditions of 98% of mobile phase A (0.1% HCOOH in H_2O, v/v) were kept for 0.5 min, followed by a linear gradient from 2% to 20% of mobile phase B (CH_3CN (0.1% v/v HCOOH)) in 3.5 min and from 20% to 95% B in 4 min. 95% B was held for 1 min and then, a 0.25 min gradient was used to return to the initial conditions, which were held for 2.75 min. Full scan MS data from 70 to 1700 m/z with a scan frequency of 6 Hz (1274 transients/spectrum) was collected on a QTOF Agilent 6550 spectrometer (Agilent Technologies, CA, USA) in the TOF-MS mode. The following positive ESI

parameters were selected: gas T, 200°C; drying gas, 14 l/min; nebulizer, 37 psig; sheath gas T, 350°C; sheath gas flow, 11 l/min. Automatic MS spectra recalibration during analysis was carried out introducing a mass reference standard into the source via a reference sprayer valve using the 149.02332 (background contaminant), 121.050873 (Purine) and 922.009798 (HP-0921) m/z as references. Sample acquisition was randomized and the QC sample was analyzed every 5 injections to monitor and correct changes in the instrument response. Eight replicates of the QC were injected at the beginning of each batch for column conditioning. Data acquired during conditioning was excluded from the analysis. The sample set included 316 urine samples, 56 QCs and 4 blanks. Sample analysis was carried out in two batches to reduce the time that samples are kept in the autosampler during analysis. Batch 1 included 224 injections of 187 urine samples (35 BC, 19 CTRL, 105 MONITOR, 28 NA), 4 Blanks and 33 QCs. Batch 2 included 152 injections of 129 urine samples (33 BC, 10 CTRL, 61 MONITOR, 23 NA) and 23 QCs. All samples from the same patient were analyzed in the same batch. Patient distribution between batches was randomized.

Peak table generation and data quality assessment

Centroid raw UPLC-TOF-MS data was converted into mzXML format using ProteoWizard (http://proteowizard.sourceforge.net/) before generating peak tables using XCMS software[179]. The centWave method was used for peak detection with the following parameters: ppm: 15, peak width: (5, 20), signal to noise threshold: 6. A minimum difference in m/z of 5 mDa was selected for peaks with overlapping RTs. Intensity weighted m/z values of each feature were calculated using the wMean function. Peak limits used for integration were found through descent on the Mexican hat filtered data. Peak grouping was carried out using the 'nearest' method using mzVs RT=1 and RT and m/z tolerances of 6 s and

5 mDa, respectively. After peak grouping, the fillPeaks method with the default parameters was applied to fill missing peak data. RT and m/z tolerances used for peak table generation and alignment of features across batches was based on the observed variation in five selected metabolites (Phe, Trp, kynurenine, hydroxykynurenine and phenylacetylglutamine) and spiked internal standards (ISs) (phenylalanine-D_5, caffeine-D_9, leukine enkephalin and reserpine) (see in results **Figure 6.1**). Peak integration accuracy was assessed by comparing automated and manual integration results for internal standards (see in results **Figure 6.1**). A total of 4299 and 4416 features found after peak detection, integration chromatographic de-convolution in batches 1 and 2, respectively. Alignment of features detected across batches lead to 3226 metabolic features. Blank samples were used to identify and remove background features arising from e.g. source contaminants, plasticizers, or solvent impurities. Within-batch effect elimination was performed by fitting time dependent non-linear functions to the injected QCs followed by a normalization of the data to this function using QC-Suport Vector Regresion (SVRC) and a radial basis function kernel, as described elsewhere[180]. The ε-insensitive loss parameter, the error penalty C and the kernel parameter γ used for the fitting of the SVR functions were selected using the 10-fold RMSECV as estimates of the expected generalization error. The ε-insensitive loss parameter for each metabolic feature was selected as the expected instrumental precision (i.e. ±2.5% of the median value observed in QCs). The error penalty C was calculated as the median value of the responses in QCs[180]. The kernel parameter γ providing the lowest RMSECV for each variable in the [2^{-3}, 2^{-2},… , 2^{9}] range was selected. Between-batch effects were eliminated by scaling the intensity of each metabolic feature in each sample using a factor defined as the ratio between the median intensity in QCs in the corresponding batch and the median intensity across batches. Finally, metabolic features showing relative

standard deviation (RSD%)>15 in QCs were considered unreliable and removed, leaving 2006 features for data analysis. Batch effects affecting the number of missing values due to e.g. instability of the chromatographic separation or wrong feature alignments were not considered in this work.

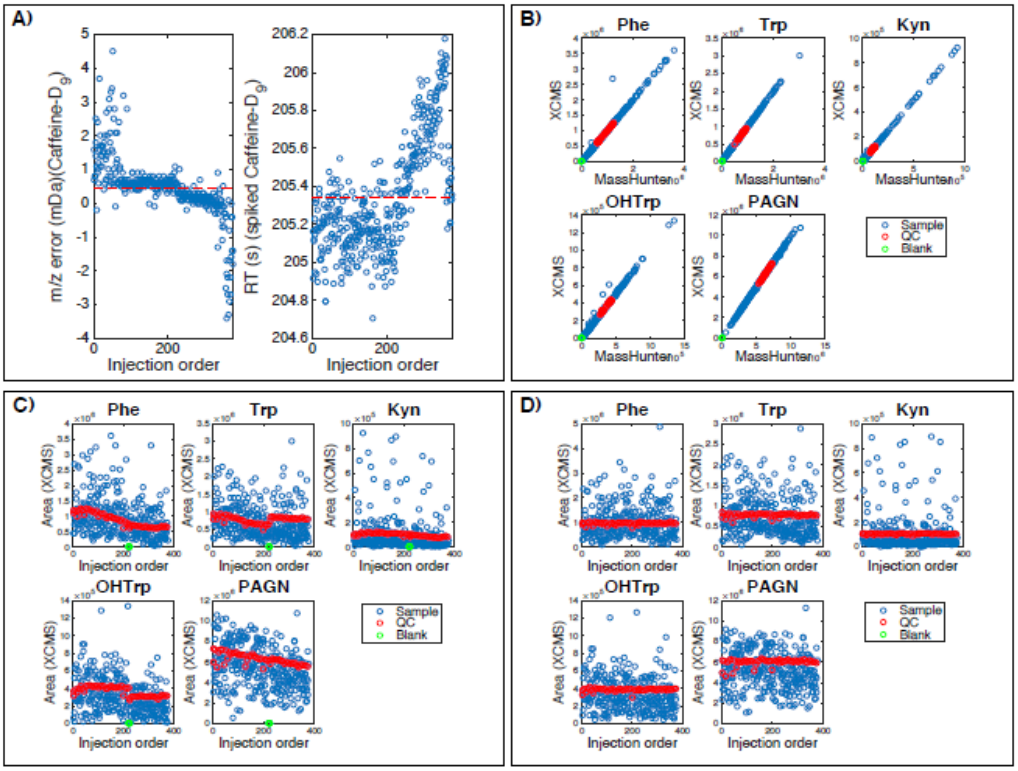


Figure 6.1 Quality control of instrument performance and metabolomic data pretreatment. A, Variation of the m/z accuracy and RT of a spiked internal standard (Caffeine-D_9) as a function of the injection order; **B**, peak area values of five selected metabolites: phenylalanine (Phe), tryptophan (Trp), kynurenine (Kyn), hydroxykynurenine (OHKyn) and phenylacetylglutamine (PAGN) using manual (MassHunter) and automated (XCMS) integration; **C**, intensity of peak area values of five selected metabolites as a function of the injection order in raw data; **D**, intensity of peak area values of five selected metabolites as a function of the injection order after batch effect correction.

Chemometric and statistical analysis

The data set was initially split into a calibration and validation subsets. The calibration set was used for PLS-DA model development and feature selection. The validation set was exclusively used for the evaluation of the model predictive performance. Data scaling included multiplicative scatter correction with the median QC as reference followed by pareto scaling. The selection of the optimal number of PLS-DA LVs was carried out using the RMSECV and a leave-*one patient*-out CV strategy. The classification accuracy, the AUROC as well as the sensitivity, selectivity and NLR, PLR were employed as PLS-DA figures of merit.

Identification of metabolites was carried out by matching m/z values against the HMDB (http://www.hmdb.ca) and METLIN databases (http://metlin.scripps.edu/) with 5 accuracy. Molecular formulae were estimated by MassHunter Workstation Software-Qualitative Analysis (Agilent). Data acquisition and manual integration of peaks of IS and selected metabolites were carried out using MassHunter workstation (Agilent). PLS-DA was carried out using *PLS_Toolbox 8.0* (Eigenvector Research Inc., Wenatchee, USA) and in-house written MATLAB (Mathworks Inc., Natick, MA, USA) scripts. Support Vector Regression was carried out in MATLAB using the LIBSVM library[181]. Pathway analysis and topology analysis were carried out using a global test and a relative betweenness centrality measure using *MetaboAnalyst 3.0*[182].

Results and Discussion

Data overview and quality assessment

The replicate analysis of a QC sample throughout the batch enables a straight forward evaluation of the instrument performance[183]. Under optimal conditions, technical variation should lead to random variation in intensities across QC replicates. However, the plot of the peak areas in QCs as a function of the injection order showed trends both within and between batches, as well as heterocedastic variance across batches (see **Figure 6.1**). Likewise, cumulative distribution functions of the relative standard deviations in QCs (RSD_{QC}) in raw data depicted in **Figure 6.2** showed a significantly better instrument performance in terms of repeatability (i.e. lower RSD_{QC}) in batch 2. **Figure 6.2** depicts PC1-PC3 scores of a PCA model for QC replicates as a function of the injection order, showing a significant between-batch effect in PC1 and within-batch effects in PC2 and PC3. Batch effects difficult the identification of underlying trends in the data and so, an initial batch effect correction was carried out as described above. After, within- and between-batch effect correction, the number of metabolic features showing RSD_{QC}<15% increased from 652 up to 2006, and the median RSD_{QC} decreased from 20.2% down to 8.7%. PCA scores after batch effect correction depicted in **Figure 6.2** showed no association with the injection or batch order in agreement with results depicted in **Figure 6.1** where the corrected intensities of the previous set of 5 metabolites as a function of the injection order are depicted.

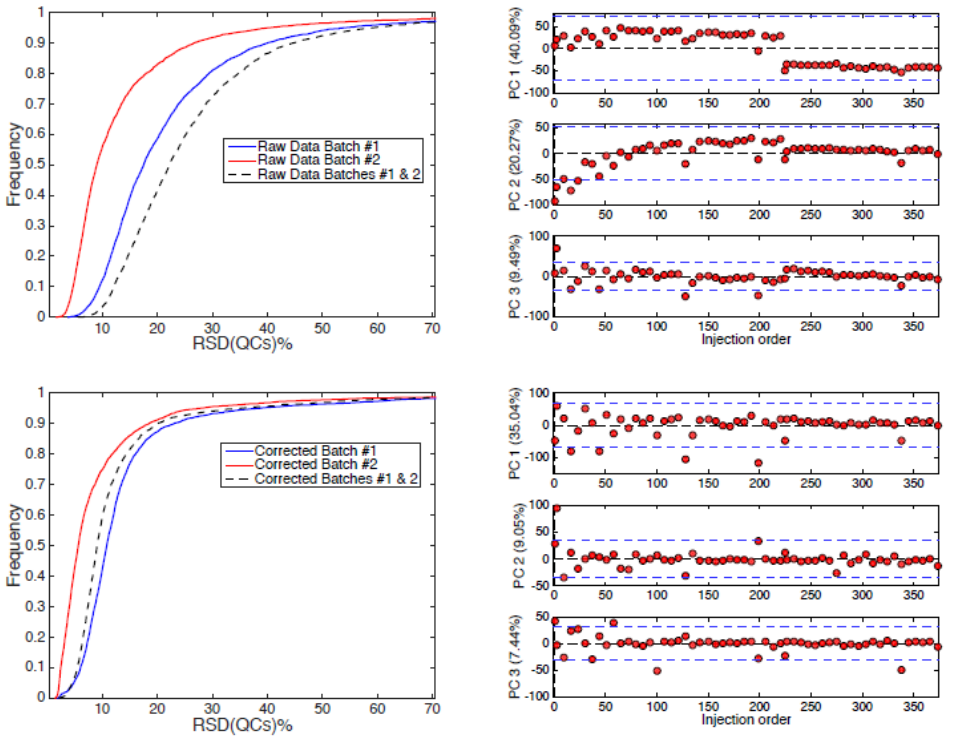


Figure 6.2 Distribution of RSD (QCs)% in batches 1 and 2 (left) and PCA scores as a function of injection order (right) in raw data (top) and after batch effect correction (bottom).

A PCA model of the set of BC, MONITOR and CTRL samples was calculated. The PC1 vs PC2 scores plots obtained from the PCA models of the set of BC, MONITOR and CTRL samples after batch effect correction showed a high overlap of BC, MONITOR and CTRL samples (see **Figure 6.3**). No clustering among the groups was observed using higher PCs (data not shown). PCA did not reveal a specific structure related to BC progression. Nonetheless, the PCA model was used to assess the absence of outlying samples based on their relative position to the 95% confidence limit.

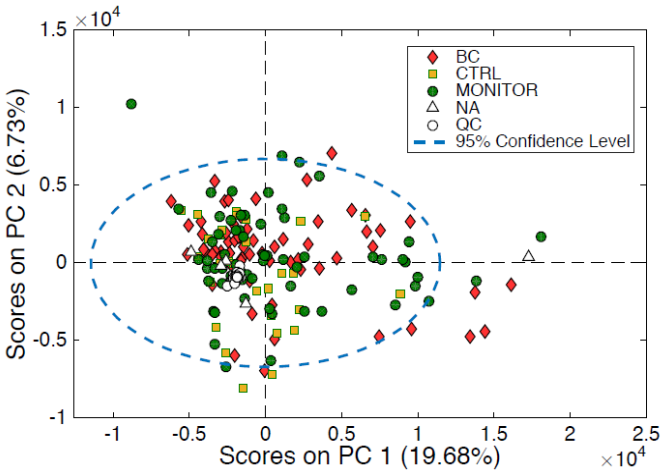


Figure 6.3 PCA scores plot using the set of QC, BC, CTRL and MONITOR samples.

Discriminant analysis among BC, CTRL and MONITOR samples

To facilitate the analysis of the differences in the metabolic profiles among BC, CTRL and MONITOR samples, three independent PLS-DA models were considered in which the groups were compared pairwise (i.e. BC *vs* CTRL, BC *vs* MONITOR and CTRL *vs* MONITOR). Calibration and validation sets selected for the three models are summarized in **Table 6.1**. PLS-DA scores plots and predicted values for the three models depicted in **Figure 6.4** and the figures of merit calculated for the validation sets summarized in **Table 6.2** showed a statistically significant shift in the urinary metabolic profiles after TUR. The BC *vs* CTRL model provided an accurate classification of 27/33 BC and 10/11 CTRL samples (sensitivity: 82% and specificity: 91%). The second model build for the discrimination between BC *vs* MONITOR samples performed worse in classifying BC samples (sensitivity: 70% and specificity: 75%). Finally, the analysis of the differences between CTRL *vs* MONITOR groups provided non-significant predictive performances, lower sensitivity (45%) and specificity (76%) values, in agreement with the higher overlapping of CTRL and MONITOR samples depicted in **Figure 6.4**.

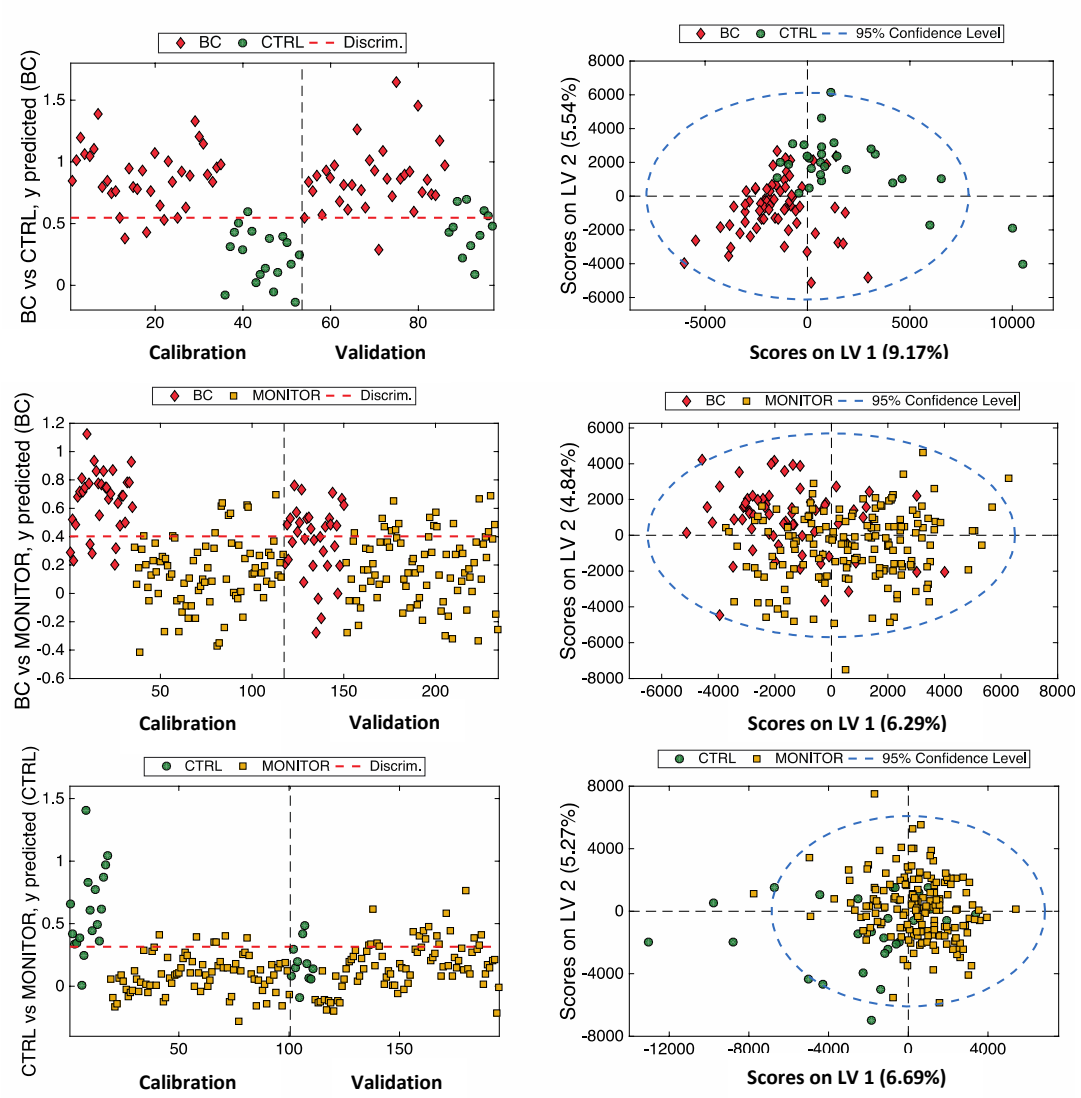


Figure 6.4. Discriminant analysis of BC, CTRL and MONITOR samples. Left) PLS-DA predicted *y* values for the calibration (auto-prediction) and validation subsets; Right) scores plot for the calibration and validation sets. Number of LVs: 3.

Table 6.2 Indices of test validity estimated for the evaluation of the predictive performance of PLS-DA models between BC *vs* CTRL, BC *vs* MONITOR and MONITOR *vs* CTRL samples in the validation set.

	PLS-DA model		
	BC *vs* CTRL	**BC *vs* MONITOR**	**CTRL *vs* MONITOR**
LVs	3	3	3
AUROC	0.94	0.75	0.53
Sensitivity	81.8 (64.5-93.0)%	69.7 (51.3-84.4)%	45.4 (16.7-76.6)%
Specificity	90.9 (58.7-99.8)%	75.0 (64.4-83.8)%	76.0 (66.6-83.8)%
PLR	9.0 (1-4-58.7)	2.8 (1.8-4.3)	1.9 (0.9-3.9)
NLR	0.2 (0.00-0.42)	0.4 (0.2-0.7)	0.7 (0.4-1.2)

The relative importance of each metabolic feature in the projection used in PLS-DA models was evaluated using the VIP scores estimate[184]. The BC *vs* CTRL model was used to screen an initial set of 128 discriminant features using a VIP>3 as threshold (**Figure 6.5**). This set of metabolic features associated to the effect of tumor removal by TUR in NMBIC patients, was used to build an optimized model (3 LVs), which correctly classified 29/33 BC and 11/11 CTRL samples of the validation set, providing an AUROC= 0.96 and slightly improved sensitivity (87.9 (71.8-96.6)%), specificity (100 (71.5-100)%) and the NLR 0.1 (0.05-0.3). Adjusting by prior probability of recurrence per risk grouping at 15%, 24%, 28% and 61%, the negative predictive values for low, low-intermediate and high-intermediate and high-risk groups were 96.5%, 94.0%, 92.9% and 76.1% respectively.

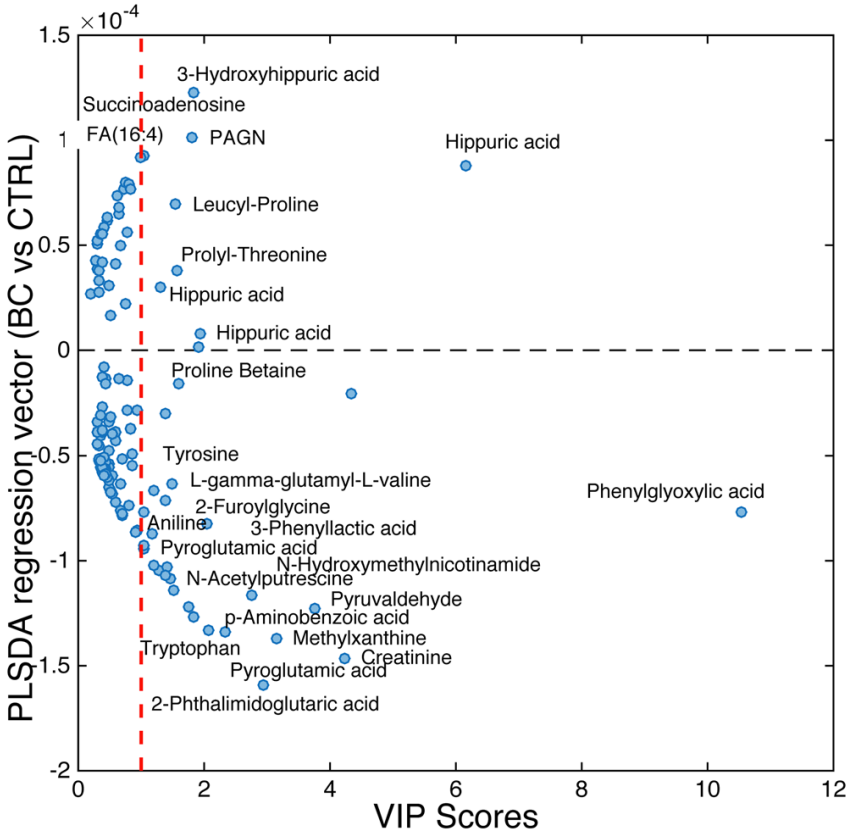


Figure 6.5 Discriminant metabolites between BC and CTRL samples. VIP scores as a function of the value in the PLS-DA regression vector in a model build using 128 selected metabolic features.

On the other hand, putatively identified discriminant metabolites showing a VIP>1 in the BC *vs* CTRL model summarized in **Table annexed 6.3** reflected alterations in the metabolic pathways of Arg, Pro, FAs, Phe, purine, pyrimidine and Trp, among others (see **Table 6.4**).

Table 6.4 Putatively identified metabolites and associated pathways.

Metabolic Pathways	Metabolites
Aminobenzoate degradation; microbial metabolism	**quinone**
Arg and Pro metabolism	*Creatine, creatinine,* **guanidinobutanoic acid**, *oxoarginine*, *gamma-glutamyl-putrescine*, *spermine*, *citrulline*
Arg, purine, pyrimidine[185–187], Ala, Asp[186], Glu[188] metabolism	*n-acetylglutamine,* **thymine**, *dihydrothymine*
Biosynthesis of secondary metabolites[185]	*methylxanthine***, hydroxyphenylalanine**
Citrate cycle [175,185,187,189,190]	*citric acid* [175,185,187]
Energy metabolism	*Carnitine*[189], **acetylcarnitine**[191], *o-isobutyryl-carnitine*, *3-methylglutarylcarnitine, propionylcarnitine*
FAs metabolism[186,189,191]	*Carnitine*[189], *furoylglycine*, **aminohippuric acid**, **hydroxyhippuric acid**
GSH metabolism	**pyroglutamic acid**
Phe metabolism[191]	**hydroxyhippuric acid**, *hippuric acid*[175,191,192] **phenylacetylglutamine**[191], *phenyllactic*, *hydrocinnamic acid*, **homophenylalanine**, *phenylacetylglycine*, *aminosalicyluric acid*, **phenylglyoxylic acid**, **Tyr**[188]
Primary degradation product of tRNA	*dimethylguanosine*
Purine metabolism[186,188,190]	*Hypoxanthine*[188], *methylhypoxanthine*, *adenosine*[193], **xanthine**, **uric acid**[188]
Trp metabolism[185,188,189]	**Tyr**[188], **hydroxyindole**, *hydroxyanthranilic acid*, **anthranilic acid**[185], **indolelactic acid**, *methyltryptamine*, **Trp**[188], *hydroxyindoleacetic acid*, **kynurenine**, *hydroxyindolepyruvic acid*, **hydroxytryptophan**

Note: References indicate previous clinical urinary metabolomic studies of BC in which the metabolites were selected as discriminant and/or dysregulated pathways reviewed in[194].
Note: Metabolites found at higher levels before TUR are highlighted in bold. Pathways highlighted in bold were found dysregulated (p-value<0.05).

Pathway analysis was used to extract biological information within relevant networks of metabolic pathways integrating MSEA and pathway topology analysis of BC and CTRL profiles. These analyses were carried out excluding unidentified or without matching HMDB ID metabolic features (VIP>1). Results obtained are depicted in **Figure 6.6**, where the color and the size of each circle indicate its *p*-value and pathway impact value, respectively. Phe, Arg, Pro and Trp pathways were found significantly altered (*p*-value<0.05), which is in concordance with the results obtained in Chapter 5.

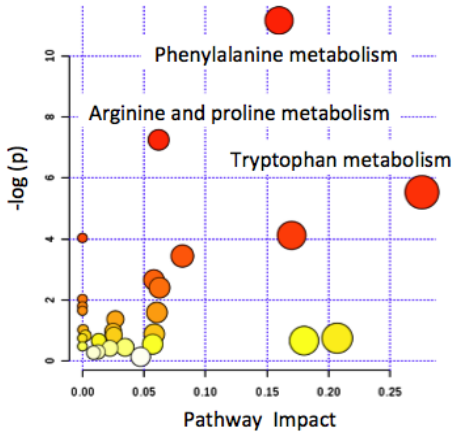


Figure 6.6 Pathway analysis of the urinary metabolic shift after TUR.

Moreover, this observation was in agreement with recent results reporting increased levels of four Trp metabolites (kynurenine, acetyl-N-formyl-5-methoxykynurenamine, indoleacetic acid and indolelactic acid) in serum samples of BC patients compared to healthy controls, and previous studies in BC tissue[158] and urine[185] that suggested the potential role of kynurenine in the malignancy BC associated to IDO and IDO2, two Trp-metabolizing enzymes that control the Trp catabolism-signaling pathway. The generation of kynurenine and other Trp metabolites can modulate T-cell immunity via activation of suppressive regulatory

T-cells and activation of aryl hydrocarbon receptor, thus promoting cancer cell survival[195]. Higher levels of pyroglutamic acid and lower levels of hippuric acid before TUR were in also agreement with previous studies reporting results from the analysis of urine samples collected from BC patients and reference healthy groups[175,177,191] but also were in concordance with our study performed by NMR (Chapter 5). Phenylacetylglutamine is synthesized in the liver from Gln and phenylacetyl-CoA and it is also a known microbial metabolite[196]. Altered levels of phenylacetylglutamine might indicate a deregulation of the Phe or Gln metabolism, observed in a previous urinary metabolomic study involving BC patients and healthy controls, that attributed this deregulation to the increased energy demands of cancer cells for growing and proliferation[197]. Citrate was found at lower concentrations before TUR. This metabolite is a key intermediate in TCA cycle and its deregulation has been repeatedly associated with an increased conversion into FAs required for ß-oxidation to support cancer cell proliferation. Carnitine and several carnitine metabolites were also among the most discriminant metabolites. Carnitine is an essential metabolite for the transport of long-chain FAs into the mitochondria and for the regulation of the intramitochondrial ratio of acetyl-CoA to free CoA. Hence, results may support higher levels of fatty acids ß-oxidation deregulation associated to the BC tumor[186,188,190]. Altered pyrimidine and purine metabolism has been previously attributed to enhanced cancer cells cycle activity[185,187].

In summary, results support the hypothesis of a urinary metabolic profile associated with a macroscopic NMIBC tumor. Besides, the observed metabolic shift after TUR is well aligned with previous studies aiming at the identification of metabolic biomarkers carried out in BC patients and healthy controls, thereby supporting the hypothesis that the metabolic shift might be useful for the surveillance of cancer recurrence after TUR in NMIBC patients.

Longitudinal analysis of metabolic changes during surveillance for recurrence

The analysis of longitudinal trajectories of the metabolic biomarker developed for the discrimination between BC and CTRL samples, in patients undergoing surveillance for recurrence allowed a preliminary evaluation of its utility as a monitoring system to diagnose NMIBC recurrence in a pre-clinical stage or to adjust follow-up protocols according to NMIBC risk. **Figure 6.7** shows the predicted *y* PLS-DA values in 6 patients with multiple episodes of BC recurrence. Results show CTRL and BC samples were accurately classified (90.1% overall accuracy) and, for some patients (see patients *23, *28, and *29), the longitudinal trajectory during surveillance indicated a gradual shift of the metabolic profile towards a BC profile. This fact, was in agreement with the depelopment of new recurrences in the bladder or in the upper urinary tract (see patient *28), and was consistent with the confirmatory results of cystoscopy and PA. Nevertheless, in some patients, results showed inconsistent trajectories due to a reduced number of misclassified samples (see e.g. MONITOR samples from patients *23 and *24 in **Figure 6.7**). The fact that urine samples can be easily obtained, and that they are in contact with the tumor in NMIBC patients are two *a priori* advantageous features for the development of a non-invasive metabolomic analysis. However, urine analysis is challenging due to the variation in chemical composition and concentrations across and within individuals. A wide range of potential confounding factors such as individual genotype, diet, water consumption, environmental exposure or drug intake that affect urine metabolome may be responsible for these sample misclassification. Further research is needed to assess the sources of short-term variability in urine and increase the robustness of metabolic tests in exploratory studies to facilitate the validation and translation of biomarker discoveries into clinical practice.

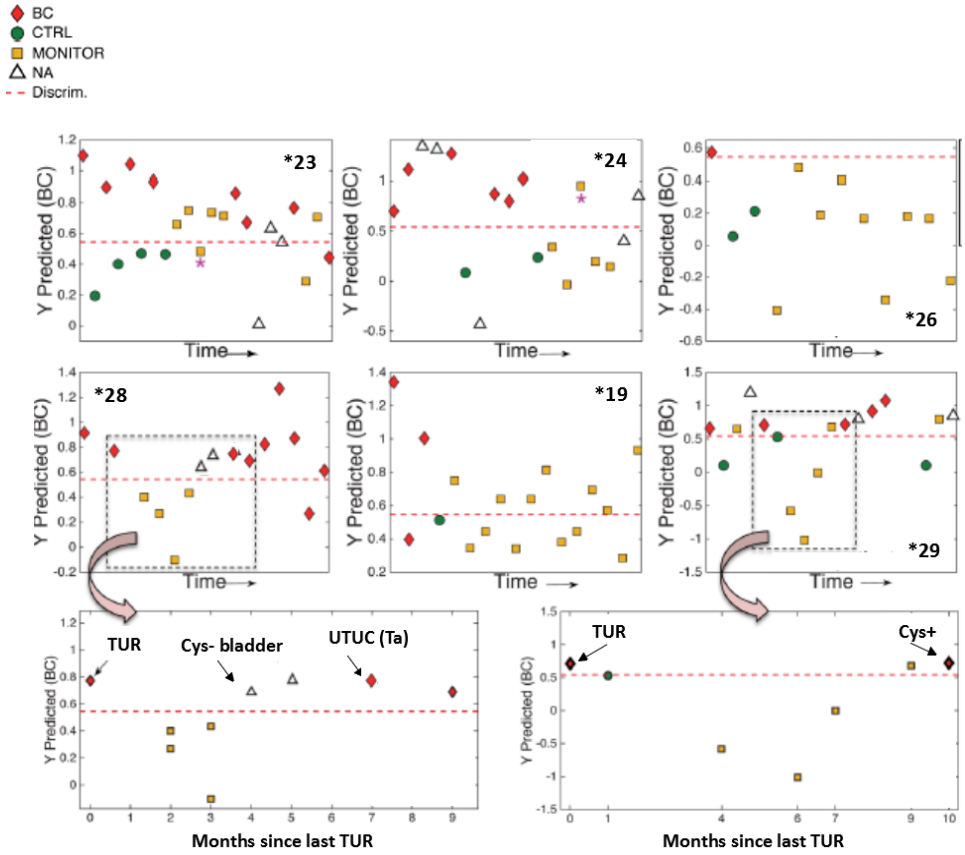


Figure 6.7 Analysis of longitudinal trajectories after TUR. Predicted *y* PLS-DA values in 6 patients during surveillance of BC recurrence. Note: (*) indicates a MONITOR sample showing an inconsistent trajectory with a gradual progression of the disease after TUR. (Cys+ or Cys-) indicates a positive and negative cystoscopy; (TUR) mean transurethral resection; and (UTUC) means upper tract urothelial carcinoma**.** BC and CTRL samples from patients *24 and *29 were included in the calibration set.

Conclusions

Results from this exploratory clinical study disclosed a statistically significant shift in the urinary metabolic profile in NMIBC before and after TUR and support the hypothesis of a specific urinary metabolic profile associated to the occurrence of a NMIBC tumor. Moreover, follow-up of the urine metabolome of those patients undergoing surveillance for risk of recurrence indicated a gradual shift in the metabolic profile towards the BC profile that was further confirmed by cystoscopy. Taken together, these results provide a strong basis for the development of a metabolomic biomarker analysis as a non-invasive monitoring system to detect NMIBC recurrence in an early stage and eventually adjust follow-up protocols according to NMIBC risk. However, larger population sizes need to be studied for the clinical validation of the proposed biomarkers.

Annexed tables

Table 6.3. Urinary metabolites putatively identified among the set of metabolic features with VIP>1 in the initial BC *vs* CTRL PLSDA model.

ID	Mol. Formula	Error (ppm)	Score	Putative ID	HMDB/KEGG	t-test	BC/CTRL
109.0285-3.37	$[C6H5O2]^+$	-3	99	Quinone	HMDB03364	<0.05	>1
115.0693-0.67	$[C4H8N3O]^+$	2	98	Creatinine	HMDB00562	>0.05	<1
146.0921-0.83	$[C5H12N3O3]^+$	2	90	Guanidino-butanoic	HMDB03464	>0.05	>1
174.0873-0.76	$[C6H12N3O3]^+$	0.9	85	Oxoarginine	HMDB04225	>0.05	<1
182.1291-1.12	$[C9H16N3O]^+$	0.9	87	g-glutamyl-putrescine	HMDB12230	<0.05	<1
203.2231-052	$[C10H27N4]^+$	-1	87	Spermine	HMDB01256	<0.05	<1
176.1033-0.78	$[C6H14N3O3]^+$	0.1	90	Citrulline	HMDB00904	>0.05	>1
171.0761-0.70	$[C7H11N2O3]^+$	-3	83	Acetylglutamine	HMDB06029	>0.05	<1
127.0488-1.95	$[C5H7N2O2]^+$	-0.1	NA	Thymine	HMDB00262	<0.05	>1
129.0655-0.67	$[C5H9N2O2]^+$	-3	87	Hydrouracil	HMDB00079	<0.05	<1
167.0567-1.77	$[C6H7N4O2]^+$	-4	92	Methylxanthine	HMDB01991	>0.05	<1
182.0817-3.89	$[C9H12NO3]^+$	-4	82	Hydroxy-phenylalanine	C19712	<0.05	>1
215.0167-0.86	$[C6H8O7Na]^+$	-5	85	Citric acid	HMDB00094	>0.05	<1
220.0642 -1.55	$[C11H11NO4]^+$	-6	83	Hydroxyindole-pyruvate	C05646	>0.05	<1
162.1123-0.66	$[C7H16NO3]^+$	0.6	82	Carnitine	HMDB00062	>0.05	<1
204.1239-0.91	$[C9H18NO4]^+$	-8	86	Acetylcarnitine	HMDB00201	>0.05	<1
232.1551-2.75	$[C11H22NO4]^+$	-3	93	Isobutyryl carnitine	HMDB62606	>0.05	<1
290.1613-2.50	$[C13H24NO6]^+$	-5	92	Methyl-glutarylcarnitine	HMDB00052	>0.05	<1

170.0452-2.19	$[C_7H_8NO_4]^+$	-5	93	Furoylglycine	HMDB00439	<0.05	>1
245.0928-5.19	$[C_{11}H_{14}N_2O_3Na]^+$	-3	95	Aminohippuric acid	HMDB01867	<0.05	>1
218.0430-2.29	$[C_9H_9NO_4Na]^+$	-4	86	Hydroxyhippuric acid	HMDB06116	>0.05	<1
180.0660-3.51	$[C_9H_{10}NO_3]^+$	-4	97	Hippuric acid	HMDB00714	<0.05	<1
265.1185-3.67	$[C_{13}H_{17}N_2O_4]^+$	1	98	Phenylacetyl-glutamine	HMDB06344	>0.05	<1
130.0506-1.31	$[C_5H_8NO_3]^+$	6	97	Pyroglutamic acid	HMDB00267	>0.05	>1
151.0753-5.60	$[C_9H_{11}O_2]^+$	-2	86	Hydrocinnamic acid	HMDB00764	>0.05	>1
180.1017-6.09	$[C_{10}H_{14}NO_2]^+$	0.1	88	Homophenyl-alanine	NA	<0.05	<1
194.0812-4.04	$[C_{10}H_{11}NO_3]^+$	0.9	75	Phenylacetyl-glycine	HMDB00821	<0.05	<1
211.0718-1.90	$[C_9H_{11}N_2O_4]^+$	-2	97	Aminosalicyluric acid	HMDB61683	>0.05	<1
151.0393-2.85	$[C_8H_7O_3]^+$	-4	96	Phenylglyoxylic acid	HMDB60026	<0.05	>1
182.0790-2.23	$[C_9H_{12}NO_3]^+$	-4	95	Tyrosine	HMDB00158	>0.05	>1
312.1306-2.44	$[C_{12}H_{18}N_5O_5]^+$	-3	96	Dimethyl-guanosine	HMDB04824	>0.05	<1
137.0460-0.88	$[C_5H_5N_4O]^+$	-3	91	Hypoxanthine	HMDB00157	>0.05	>1
151.0618-1.89	$[C_6H_7N_4O]^+$	-4	88	Methyl-hypoxanthine	HMDB13141	>0.05	<1
268.1044-1.07	$[C_{10}H_{14}B_5O_4]^+$	-2	91	Adenosine	HMDB00050	>0.05	<1
153.0407-0.89	$[C_5H_4N_4O_2]^+$	-6	85	Xanthine	HMDB00292	>0.05	>1
169.0360-0.80	$[C_5H_4N_5O_3]^+$	-5	82	Uric acid	HMDB00289	>0.05	>1
134.0600-3.33	$[C_8H_8NO]^+$	-2	87	Hydroxyindole	HMDB59805	>0.05	>1
154.0501-1.13	$[C_7H_8NO_3]^+$	-4	83	Hydrox-yanthranilic	HMDB01476	>0.05	<1

138.0549-0.67	$[C_7H_8NO_2]^+$	-6	94	Anhranillic acid	HMDB01123	>0.05	>1
175.1237-2.65	$[C_{11}H_{15}N_2]^+$	-5	93	Methyltryptamine	HMDB04370	<0.05	<1
205.0980-2.86	$[C_{11}H_{13}N_2O_2]^+$	-5	96	Tryptophan	HMDB00929	>0.05	>1
192.0659-1.99	$[C_{10}H_{10}NO_3]^+$	-3	85	Hydroxyindole-acetic acid	HMDB00763	>0.05	<1
209.0925-1.99	$[C_{10}H_{13}N_2O_3]^+$	-2	87	Kynurenine	HMDB00684	>0.05	>1
221.0925-1.50	$[C_{11}H_{13}N_2O_3]^+$	-3	83	Hydroxy-tryptophan	HMDB00472	>0.05	>1
132.0765-0.68	$[C_6H_{14}N_2O]^+$	-2	86	Creatine	HMDB00687	>0.05	>1
200.1281-2.41	$[C_{10}H_{18}NO_3]^+$	-4	82	Propionyl-carnitine	HMDB00824	<0.05	>1
129.0655-0.67	$[C_5H_9N_2O_2]^+$	-2	87	Dihydrothymine	HMDB00079	<0.05	<1
126.0657-0.88	$[C_5H_8N_3O]^+$	-6	83	Methylcytosine	HMDB02894	<0.05	<1

Note: t-test: t-test *p*-value; BC/CTRL: ratio mean value in BC over CTRL samples.

Chapter 7: Conclusions and future perspectives

This Ph.D. Thesis has tried to improve the knowledge about BC biology through metabolomic and transcriptomic studies but has also provided non-invasive approaches for the detection of this tumor. The conclusions obtained are detailed following:

- HRMAS ^{1}H NMR-based metabolomic analyses are able to classify correctly BC tissues independently of tumor stage and grade, and they provide information about tumor aggressiveness.
- Metabolism of AAs, Cho, and GSH would be especially important in more aggressive bladder tumors (T1-T2), since they presented the highest levels of these metabolites.
- Metabolic reprogramming of bladder tumors would be producing mainly through the downregulation of metabolic genes. Transcription factors, histone marks, and alternative splicing processes would be regulating these metabolic genes.
- Anaerobic metabolism could be essential in BC since tumor tissues showed an overexpression of glycolytic genes (e.g. *HK2* and *SLC21A*), elevated levels of lactate and a downexpression of genes related to TCA cycle and OXPHOS.
- The regulation of metabolic pathways related with glycolysis, polyamines, OXPHOS, AAs and TCA cycle would play essential roles in bladder carcinogenesis since integrated metabolomics and transcriptomics studies showed perturbation in metabolites and genes coding for enzymes that take part of these pathways.

- Molecular alterations of tumor tissues are reflexed in the urinary metabolome of the same patient, so the urine would be the ideal sample to search non-invasive BC biomarkers based on metabolomics analysis.
- Our ^{1}H NMR-based urinary metabolic profile showed a better sensitivity and specificity than standard urinary cytology in the detection of NMIBC and MIBC (91%, 77% respectively). Therefore, its transference to clinical practice would improve the management of BC by means of a non-aggressive approach.
- Urinary metabolomic profiles could be implemented for BC diagnosis but also for the monitoring of NMIBC patients since our ^{1}H NMR- and UPLC-MS-based metabolomic studies showed high values of sensitivity and specificity in the detection of tumor recurrences in patients undergo a surveillance period.
- Several of the metabolic alterations found in BC urines through ^{1}H NMR and UPLC-MS studies were common (e.g. disturbed Phe, Arg, and Pro metabolism). The robustness obtained from both studies enhances the use of metabolomic techniques in the field of biomarker discovery and supports their transfer to clinical practice.
- Urinary-based metabolomic analyses are highly dynamic and represent tumor biology. Its transference into clinical routine could improve the management of BC and consequently the quality of life of patients, especially in those included in the high risk of progression group.

In summary, these conclusions have opened new horizons of study in the field of BC, so in the coming years we will focus on: 1) validating the results obtained in larger patient cohorts; 2) studying in detail the effect that confounding factors (i.e. treatments, drugs, inflammatory processes, hematuria) can have on the urinary metabolome; 3) studying the role of certain metabolic enzymes in tumor metabolism and the possibility of developing future target therapies against them; 4) studying if the obtained metabolic profiles could be applied for UTUC diagnosis.

References

1. Leal, J., Luengo-Fernandez, R., Sullivan, R. & Witjes, J. A. Economic Burden of Bladder Cancer Across the European Union. *Eur. Urol.* **69,** 438–447 (2016).
2. Burger, M. *et al.* Epidemiology and Risk Factors of Urothelial Bladder Cancer. *Eur. Urol.* **63,** 234–241 (2013).
3. Globocan Project International Agency for Research on Cancerv 2016. *[10.10.2014].,* *Http://globocan.iarc.fr./factsheet.asp* doi:http://dx.doi.org/10.1016/j.eururo.2012.09.063
4. Lopez-Abente, G. *et al.* Municipal distribution of bladder cancer mortality in Spain: Possible role of mining and industry. *BMC Public Health* **6,** 17 (2006).
5. Globocan 2012 - Home. Available at: http://globocan.iarc.fr/Default.aspx.
6. Knowles, M. A. & Hurst, C. D. Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity. *Nat. Rev. Cancer* **15,** 25–41 (2015).
7. Dobruch, J. *et al.* Gender and Bladder Cancer: A Collaborative Review of Etiology, Biology, and Outcomes. *Eur. Urol.* **69,** 300–310 (2016).
8. Gu, J. & Wu, X. Genetic susceptibility to bladder cancer risk and outcome. *Per. Med.* **8,** 365–374 (2011).
9. Gu, J., Liang, D., Wang, Y., Lu, C. & Wu, X. Effects of N-acetyl transferase 1 and 2 polymorphisms on bladder cancer risk in Caucasians. *Mutat. Res. Toxicol. Environ. Mutagen.* **581,** 97–104 (2005).
10. Zhao, Y. *et al.* Association between MYC rs9642880[T] allele and bladder cancer risk: a meta-analysis. *Genet. Mol. Res.* **14,** 14745–14751 (2015).
11. Urist, M. J. *et al.* Loss of p63 expression is associated with tumor progression in bladder cancer. *Am. J. Pathol.* **161,** 1199–206 (2002).
12. Wiley: TNM Classification of Malignant Tumours, 7th Edition - James D. Brierley, Mary K. Gospodarowicz, Christian Wittekind.
13. Babjuk, M. *et al.* EAU Guidelines on Non–Muscle-invasive Urothelial Carcinoma of the Bladder: Update 2016. *Eur. Urol.* **71,** 447–461 (2017).
14. Sylvester, R. J. *et al.* Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. *Eur. Urol.* **49,** 466-465; discussion 475-477 (2006).
15. Fernandez-Gomez, J. *et al.* Predicting Nonmuscle Invasive Bladder Cancer Recurrence and Progression in Patients Treated With Bacillus Calmette-Guerin: The CUETO Scoring Model. *J. Urol.* **182,** 2195–2203 (2009).

16. Kiemeney, L. A., Witjes, J. A., Verbeek, A. L., Heijbroek, R. P. & Debruyne, F. M. The clinical epidemiology of superficial bladder cancer. Dutch South-East Cooperative Urological Group. *Br. J. Cancer* **67,** 806–12 (1993).
17. Bostwick, D. G. & Mikuz, G. Urothelial papillary (exophytic) neoplasms. *Virchows Arch.* **441,** 109–116 (2002).
18. Tang, D. H. & Chang, S. S. Management of carcinoma in situ of the bladder: best practice and recent developments. *Ther. Adv. Urol.* **7,** 351–64 (2015).
19. Anastasiadis, A. & de Reijke, T. M. Best practice in the treatment of nonmuscle invasive bladder cancer. *Ther. Adv. Urol.* **4,** 13–32 (2012).
20. Murphy, W. M., Takezawa, K. & Maruniak, N. A. Interobserver discrepancy using the 1998 World Health Organization/International Society of Urologic Pathology classification of urothelial neoplasms: practical choices for patient care. *J. Urol.* **168,** 968–72 (2002).
21. Van Rhijn, B. W. G. *et al.* Pathological stage review is indicated in primary pT1 bladder cancer. *BJU Int.* **106,** 206–211 (2009).
22. Compérat, E. *et al.* An interobserver reproducibility study on invasiveness of bladder cancer using virtual microscopy and heatmaps. *Histopathology* **63,** 756–766 (2013).
23. May, M. *et al.* Prognostic Accuracy of Individual Uropathologists in Noninvasive Urinary Bladder Carcinoma: A Multicentre Study Comparing the 1973 and 2004 World Health Organisation Classifications. *Eur. Urol.* **57,** 850–858 (2010).
24. Nepple, K. G. & O'Donnell, M. A. The optimal management of T1 high-grade bladder cancer. *Can. Urol. Assoc. J.* **3,** S188-92 (2009).
25. Rosser, C. J., Urquidi, V. & Goodison, S. Urinary biomarkers of bladder cancer: an update and future perspectives. *Biomark. Med.* **7,** 779–90 (2013).
26. Svatek, R. S. *et al.* The economics of bladder cancer: Costs and considerations of caring for this disease. *Eur. Urol.* **66,** 253–262 (2014).
27. Cauberg Evelyne, C., de Reijke, T. & de la Rosette, J. M. C. . Emerging optical techniques in advanced cystoscopy for bladder cancer diagnosis: A review of the current literature. *Indian J. Urol.* **27,** 245 (2011).
28. Mowatt, G. *et al.* Photodynamic diagnosis of bladder cancer compared with white light cystoscopy: Systematic review and meta-analysis. *Int. J. Technol. Assess. Health Care* **27,** 3–10 (2011).
29. Falke, J. & Witjes, J. A. Contemporary management of low-risk bladder cancer. *Nat. Rev. Urol.* **8,** 42–49 (2011).
30. Gontero, P. *et al.* The Role of Bacillus Calmette-Guérin in the Treatment of Non–Muscle-Invasive Bladder Cancer. *Eur. Urol.* **57,** 410–429 (2010).
31. Scarpato, K. R., Morgans, A. K. & Moses, K. A. Optimal management of muscle-invasive bladder cancer - a review. *Res. reports Urol.* **7,** 143–51

(2015).
32. Castro-Giner, F. *et al.* Cancer Diagnosis Using a Liquid Biopsy: Challenges and Expectations. *Diagnostics (Basel, Switzerland)* **8,** (2018).
33. Bujak, R., Struck-Lewicka, W., Markuszewski, M. J. & Kaliszan, R. Metabolomics for laboratory diagnostics. *J. Pharm. Biomed. Anal.* **113,** 108–120 (2015).
34. Ludwig, J. A. & Weinstein, J. N. Biomarkers in Cancer Staging, Prognosis and Treatment Selection. *Nat. Rev. Cancer* **5,** 845–856 (2005).
35. Kamat, A. M. *et al.* ICUD-EAU International Consultation on Bladder Cancer 2012: Screening, diagnosis, and molecular markers. *Eur. Urol.* **63,** 4–15 (2013).
36. Santoni, G., Morelli, M. B., Amantini, C. & Battelli, N. Urinary Markers in Bladder Cancer: An Update. *Front. Oncol.* **8,** 362 (2018).
37. Urquidi, V., Rosser, C. J. & Goodison, S. Molecular diagnostic trends in urological cancer: biomarkers for non-invasive diagnosis. *Curr. Med. Chem.* **19,** 3653–63 (2012).
38. Miyake, M., Owari, T., Hori, S., Nakai, Y. & Fujimoto, K. Emerging biomarkers for the diagnosis and monitoring of urothelial carcinoma. 251–261 (2018).
39. Mbeutcha, A., Lucca, I., Mathieu, R., Lotan, Y. & Shariat, S. F. Current Status of Urinary Biomarkers for Detection and Surveillance of Bladder Cancer. *Urol. Clin. North Am.* **43,** 47–62 (2016).
40. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144,** 646–674 (2011).
41. Sciacovelli, M. & Frezza, C. Oncometabolites: Unconventional triggers of oncogenic signalling cascades. *Free Radic. Biol. Med.* **100,** 175–181 (2016).
42. Pavlova, N. N. & Thompson, C. B. The Emerging Hallmarks of Cancer Metabolism. *Cell Metab.* **23,** 27–47 (2016).
43. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324,** 1029–33 (2009).
44. Hasawi, N. Al, Alkandari, M. F. & Luqmani, Y. A. Phosphofructokinase: A mediator of glycolytic flux in cancer progression. *Crit. Rev. Oncol. Hematol.* **92,** 312–321 (2014).
45. Liu, J. *et al.* An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* **173,** 400–416 (2018).
46. Sun, C.-M. *et al.* Genetic alteration in phosphofructokinase family promotes growth of muscle-invasive bladder cancer. *Int. J. Biol. Markers* **31,** e286-93 (2016).
47. Huang, C. *et al.* Expression of pyruvate kinase M2 in human bladder cancer and its correlation with clinical parameters and prognosis. *Onco. Targets.*

Ther. **11,** 2075–2082 (2018).
48. Cairns, R. A., Harris, I. S. & Mak, T. W. Regulation of cancer cell metabolism. *Nat. Publ. Gr.* **11,** (2011).
49. Mates, J. M. *et al.* Glutaminase Isoenzymes as Key Regulators in Metabolic and Oxidative Stress Against Cancer. *Curr. Mol. Med.* **13,** 514–534 (2013).
50. Locasale, J. W. Serine, glycine and one-carbon units: cancer metabolism in full circle. *Nat. Rev. Cancer* **13,** 572–83 (2013).
51. Ananieva, E. Targeting amino acid metabolism in cancer growth and anti-tumor immune response. *World J. Biol. Chem.* **6,** 281–9 (2015).
52. Kang, K., Shu, X.-L., Zhong, J.-X. & Yu, T.-T. Effect of L-arginine on immune function: a meta-analysis. *Asia Pac. J. Clin. Nutr.* **23,** 351–9 (2014).
53. Loras, A. *et al.* Bladder cancer recurrence surveillance by urine metabolomics analysis. *Sci. Rep.* **8,** 9172 (2018).
54. Amobi, A., Qian, F., Lugade, A. A. & Odunsi, K. in *Advances in experimental medicine and biology* **1036,** 129–144 (2017).
55. Currie, E., Schulze, A., Zechner, R., Walther, T. C. & Farese, R. V. Cellular Fatty Acid Metabolism and Cancer. *Cell Metab.* **18,** 153–161 (2013).
56. Santos, C. R. & Schulze, A. Lipid metabolism in cancer. *FEBS J.* **279,** 2610–2623 (2012).
57. Acharya, A., Das, I., Chandhok, D. & Saha, T. Redox regulation in cancer: a double-edged sword with therapeutic potential. *Oxid. Med. Cell. Longev.* **3,** 23–34 (2010).
58. Ghezzi, P. Role of glutathione in immunity and inflammation in the lung. *Int. J. Gen. Med.* **4,** 105–13 (2011).
59. Yang, M., Soga, T. & Pollard, P. J. Oncometabolites: linking altered metabolism with cancer. *J. Clin. Invest.* **123,** 3652–8 (2013).
60. Collins, R. R. J., Patel, K., Putnam, W. C., Kapur, P. & Rakheja, D. Oncometabolites: A New Paradigm for Oncology, Metabolism, and the Clinical Laboratory. *Clin. Chem.* **63,** 1812–1820 (2017).
61. Martín-Martín, N., Carracedo, A. & Torrano, V. Metabolism and Transcription in Cancer: Merging Two Classic Tales. *Front. cell Dev. Biol.* **5,** 119 (2017).
62. Kozlovski, I., Siegfried, Z., Amar-Schwartz, A. & Karni, R. The role of RNA alternative splicing in regulating cancer metabolism. *Hum. Genet.* **136,** 1113–1127 (2017).
63. Nebert, D. W. Transcription factors and cancer: an overview. *Toxicology* **181–182,** 131–141 (2002).
64. Bhagwat, A. S. & Vakoc, C. R. Targeting Transcription Factors in Cancer. *Trends in cancer* **1,** 53–65 (2015).
65. Cabrera Je, E. The Role of Transcription Factor TWIST in Cancer Cells. *J. Genet. Syndr. Gene Ther.* **04,** 1–7 (2013).

66. Nagaraju, G. P. & Bramhachari, P. V. *Role of transcription factors in gastrointestinal malignancies*.
67. Oh, S., Oh, C. & Yoo, K. H. Functional roles of CTCF in breast cancer. *BMB Rep.* **50,** 445–453 (2017).
68. Bell, R. J. A. *et al.* Cancer. The transcription factor GABP selectively binds and activates the mutant TERT promoter in cancer. *Science* **348,** 1036–9 (2015).
69. Dang, C. V & Semenza, G. L. Oncogenic alterations of metabolism. *Trends Biochem. Sci.* **24,** 68–72 (1999).
70. Krishnan, A., Nair, S. & Pillai, M. Biology of PPARγ in Cancer: A Critical Review on Existing Lacunae. *Curr. Mol. Med.* **7,** 532–540 (2007).
71. Burns, K. A. & Vanden Heuvel, J. P. Modulation of PPAR activity via phosphorylation. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1771,** 952–960 (2007).
72. Coller, H. A. Is cancer a metabolic disease? *Am. J. Pathol.* **184,** 4–17 (2014).
73. Hirschey, M. D. *et al.* Dysregulated metabolism contributes to oncogenesis. *Semin. Cancer Biol.* **35 Suppl,** 129–150 (2015).
74. Editor, J. T. B. & Walker, J. M. *Metabonomics Series Editor*.
75. Smolinska, A., Blanchet, L., Buydens, L. M. C. & Wijmenga, S. S. NMR and pattern recognition methods in metabolomics: from data acquisition to biomarker discovery: a review. *Anal. Chim. Acta* **750,** 82–97 (2012).
76. Patel, S. & Ahmed, S. Emerging field of metabolomics: big promise for cancer biomarker identification and drug discovery. *J. Pharm. Biomed. Anal.* **107,** 63–74 (2015).
77. Wishart, D. S. *et al.* HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* **46,** 608–617 (2018).
78. Liesenfeld, D. B., Habermann, N., Owen, R. W., Scalbert, A. & Ulrich, C. M. Review of mass spectrometry-based metabolomics in cancer research. *Cancer Epidemiol. Biomarkers Prev.* **22,** 2182–201 (2013).
79. Bingol, K., Bingol & Kerem. Recent Advances in Targeted and Untargeted Metabolomics by NMR and MS/NMR Methods. *High-Throughput* **7,** 9 (2018).
80. Broadhurst, D. *et al.* Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* **14,** 72 (2018).
81. León, Z., García-Cañaveras, J. C., Donato, M. T. & Lahoz, A. Mammalian cell metabolomics: Experimental design and sample preparation. *Electrophoresis* **34,** n/a-n/a (2013).
82. Alonso, A. methods in untargeted metabolomics: state of the art in 2015., Marsal, S. & Julià, A. Analytical methods in untargeted metabolomics: state of the art in 2015. *Front. Bioeng. Biotechnol.* **3,** 23 (2015).

83. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28,** 27–30 (2000).
84. Frolkis, A. *et al.* SMPDB: The Small Molecule Pathway Database. *Nucleic Acids Res.* **38,** 480–487 (2010).
85. Kelder, T. *et al.* WikiPathways: building research communities on biological pathways. *Nucleic Acids Res.* **40,** D1301-7 (2012).
86. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* **44,** D471–D480 (2016).
87. Zhang, A., Sun, H., Wang, P., Han, Y. & Wang, X. Modern analytical techniques in metabolomics analysis. *Analyst* **137,** 293–300 (2012).
88. Hoult, D. . & I., D. Solvent peak saturation with single phase and quadrature fourier transformation. *J. Magn. Reson.* **21,** 337–347 (1976).
89. Mckay, R. T. How the 1D-NOESY suppresses solvent signal in metabonomics NMR spectroscopy: An examination of the pulse sequence components and evolution. *Concepts Magn. Reson. Part A* **38A,** 197–220 (2011).
90. Martínez-Bisbal, M. C. *et al.* Determination of metabolite concentrations in human brain tumour biopsy samples using HR-MAS and ERETIC measurements. *NMR Biomed.* **22,** 199–206 (2009).
91. Esteve, V., Celda, B. & Martínez-Bisbal, M. C. Use of 1H and 31P HRMAS to evaluate the relationship between quantitative alterations in metabolite concentrations and tissue features in human brain tumour biopsies. *Anal. Bioanal. Chem.* **403,** 2611–2625 (2012).
92. Martínez-Bisbal, M. C. *et al.* 1H and 13C HR-MAS spectroscopy of intact biopsy samples ex vivo and in vivo 1H MRS study of human high grade gliomas. *NMR Biomed.* **17,** 191–205 (2004).
93. Emwas, A.-H. M., Salek, R. M., Griffin, J. L. & Merzaban, J. NMR-based metabolomics in human disease diagnosis: applications, limitations, and recommendations. *Metabolomics* **9,** 1048–1072 (2013).
94. Armitage, E. G. & Barbas, C. Metabolomics in cancer biomarker discovery: Current trends and future perspectives. *J. Pharm. Biomed. Anal.* **87,** 1–11 (2014).
95. Emwas, A.-H. *et al.* Standardizing the experimental conditions for using urine in NMR-based metabolomic studies with a particular focus on diagnostic studies: a review. *Metabolomics* **11,** 872–894 (2015).
96. Ortega, N. *et al.* Comparative study of UPLC–MS/MS and HPLC–MS/MS to determine procyanidins and alkaloids in cocoa samples. *J. Food Compos. Anal.* **23,** 298–305 (2010).
97. Theodoridis, G. A., Gika, H. G., Want, E. J. & Wilson, I. D. Liquid chromatography-mass spectrometry based global metabolite profiling: A

review. *Anal. Chim. Acta* **711,** 7–16 (2012).
98. Nassar, A. F., Wu, T., Nassar, S. F. & Wisnewski, A. V. UPLC–MS for metabolomics: a giant step forward in support of pharmaceutical research. *Drug Discov. Today* **22,** 463–470 (2017).
99. Ho, C. S. *et al.* Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin. Biochem. Rev.* **24,** 3–12 (2003).
100. Chaleckis, R., Meister, I., Zhang, P. & Wheelock, C. E. Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics. *Curr. Opin. Biotechnol.* **55,** 44–50 (2019).
101. Patel, S. & Ahmed, S. Emerging field of metabolomics: Big promise for cancer biomarker identification and drug discovery. *J. Pharm. Biomed. Anal.* **107,** 63–74 (2015).
102. Johan Trygg, †, Elaine Holmes, ‡ and & Torbjörn Lundstedt*, §,‖. Chemometrics in Metabonomics. (2006). doi:10.1021/PR060594Q
103. Eriksson, L. *Multi- and megavariate data analysis : Part I: Basic principles and applications*. (Umetrics, 2006).
104. Brereton, R. G. & Lloyd, G. R. Partial least squares discriminant analysis: Taking the magic away. *J. Chemom.* **28,** 213–225 (2014).
105. Pérez-Guaita, D., Kuligowski, J., Garrigues, S., Quintás, G. & Wood, B. R. Assessment of the statistical significance of classifications in infrared spectroscopy based diagnostic models. *Analyst* **140,** 2422–7 (2015).
106. Cieślik, M. & Chinnaiyan, A. M. Cancer transcriptome profiling at the juncture of clinical translation. *Nat. Rev. Genet.* **19,** 93–109 (2017).
107. Tsang, H.-F. *et al.* NanoString, a novel digital color-coded barcode technology: current and future applications in molecular diagnostics. *Expert Rev. Mol. Diagn.* **17,** 95–103 (2017).
108. Bustin, S. A. *A-Z of quantitative PCR*. (International University Line, 2004).
109. Devonshire, A. S. *et al.* Application of next generation qPCR and sequencing platforms to mRNA biomarker analysis. *Methods* **59,** 89–100 (2013).
110. Wilkerson, M. D. *et al.* Prediction of lung cancer histological types by RT-qPCR gene expression in FFPE specimens. *J. Mol. Diagn.* **15,** 485–97 (2013).
111. Virtanen, C. & Woodgett, J. Clinical uses of microarrays in cancer research. *Methods Mol. Med.* **141,** 87–113 (2008).
112. Govindarajan, R., Duraiyan, J., Kaliyappan, K. & Palanisamy, M. Microarray and its applications. *J. Pharm. Bioallied Sci.* **4,** S310-2 (2012).
113. Macgregor, P. F. & Squire, J. A. Application of microarrays to the analysis of gene expression in cancer. **1177,** 1170–1177 (2002).
114. Park, T. *et al.* Evaluation of normalization methods for microarray data. *BMC Bioinformatics* **4,** 33 (2003).
115. Smyth, G. K., Ritchie, M. & Thorne, N. Linear Models for Microarray Data

User's Guide. *Bioinformatics* (2011).
116. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43,** e47–e47 (2015).
117. Gao, Q. *et al.* Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. *Cell Rep.* **23,** 227–238.e3 (2018).
118. Saeed, A. I. *et al.* TM4: A Free, Open-Source System for Microarray Data Management and Analysis. *Biotechniques* **34,** 374–378 (2003).
119. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102,** 15545–15550 (2005).
120. Lachmann, A. *et al.* ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* **26,** 2438–2444 (2010).
121. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25,** 25–9 (2000).
122. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10,** 57–63 (2009).
123. Wilhelm, B. T. & Landry, J.-R. RNA-Seq—quantitative measurement of expression through massively parallel RNA-sequencing. *Methods* **48,** 249–257 (2009).
124. Zhao, S., Fung-Leung, W.-P., Bittner, A., Ngo, K. & Liu, X. Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells. *PLoS One* **9,** e78644 (2014).
125. Liu, Y. *et al.* RNA-Seq identifies novel myocardial gene expression signatures of heart failure. *Genomics* **105,** 83–89 (2015).
126. Wang, C. *et al.* The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. *Nat. Biotechnol.* **32,** 926–932 (2014).
127. Li, J. *et al.* Comparison of microarray and RNA-Seq analysis of mRNA expression in dermal mesenchymal stem cells. *Biotechnol. Lett.* **38,** 33–41 (2016).
128. Reynders, K. *et al.* RNA-sequencing in non-small cell lung cancer shows gene downregulation of therapeutic targets in tumor tissue compared to non-malignant lung tissue. *Radiat. Oncol.* **13,** 131 (2018).
129. Yang, X. *et al.* Identification of circular RNA signature in bladder cancer. *J. Cancer* **8,** 3456–3463 (2017).
130. Ozsolak, F. & Milos, P. M. RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet.* **12,** 87–98 (2011).
131. Han, Y., Gao, S., Muegge, K., Zhang, W. & Zhou, B. Advanced Applications of RNA Sequencing and Challenges. *Bioinform. Biol. Insights* **9,** 29–46 (2015).

132. Tilki, D. *et al.* Urine Markers for Detection and Surveillance of Non–Muscle-Invasive Bladder Cancer. *Eur. Urol.* **60,** 484–492 (2011).
133. Cheng, Y. *et al.* Metabolomics in bladder cancer: a systematic review. *Int. J. Clin. Exp. Med.* **8,** 11052–63 (2015).
134. Beckonert, O. *et al.* Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Protoc.* **2,** 2692–2703 (2007).
135. Bouatra, S. *et al.* The Human Urine Metabolome. *PLoS One* **8,** e73076 (2013).
136. Cao, M., Zhao, L., Chen, H., Xue, W. & Lin, D. NMR-based metabolomic analysis of human bladder cancer. *Anal. Sci.* **28,** 451–6 (2012).
137. Lundberg,P., Vogel, T., Malusek, A., Lundquist P.-O., Cohen, L., Dahlqvist, O. MDL - The Magnetic Resonance Metabolomics Database. (2005).
138. Xia, J., Sinelnikov, I. V, Han, B. & Wishart, D. S. MetaboAnalyst 3.0--making metabolomics more meaningful. *Nucleic Acids Res.* **43,** 251–7 (2015).
139. Savorani, F., Tomasi, G. & Engelsen, S. B. icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. *J. Magn. Reson.* **202,** 190–202 (2010).
140. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44,** 90–97 (2016).
141. Howe, E. A., Sinha, R., Schlauch, D. & Quackenbush, J. RNA-Seq analysis in MeV. *Bioinformatics* **27,** 3209–10 (2011).
142. Yu, M. *et al.* The prognostic value of GLUT1 in cancers: a systematic review and meta-analysis. *Oncotarget* **8,** 43356–43367 (2017).
143. Hu, J. *et al.* Heterogeneity of tumor-induced gene expression changes in the human metabolic network. *Nat. Biotechnol.* **31,** 522–529 (2013).
144. Sajnani, K., Islam, F., Smith, R. A., Gopalan, V. & Lam, A. K.-Y. Genetic alterations in Krebs cycle and its impact on cancer pathogenesis. *Biochimie* **135,** 164–172 (2017).
145. Lew, C. R., Guin, S. & Theodorescu, D. Targeting glycogen metabolism in bladder cancer. *Nat. Rev. Urol.* **12,** 383–391 (2015).
146. Patra, K. C. *et al.* Hexokinase 2 Is Required for Tumor Initiation and Maintenance and Its Systemic Deletion Is Therapeutic in Mouse Models of Cancer. *Cancer Cell* **24,** 213–228 (2013).
147. Conde, V. R. *et al.* The progression from a lower to a higher invasive stage of bladder cancer is associated with severe alterations in glucose and pyruvate metabolism. *Exp. Cell Res.* **335,** 91–98 (2015).
148. Woolbright B, Choudhary D, Mikhalyuk A, Trammel C, S. S. et al. The Role of Pyruvate Dehydrogenase Kinase-4 (PDK4) in Bladder Cancer and Chemoresistance. *Mol. Cancer Ther.* **17,** 2004–2012 (2018).
149. Inamoto, T., Shah, J. B. & Kamat, A. M. Friend or foe? Role of peroxisome proliferator-activated receptor-γ in human bladder cancer. *Urol. Oncol.*

Semin. Orig. Investig. **27,** 585–591 (2009).

150. Annicotte, J.-S., Culine, S. & Fajas, L. Rôle de PPARγ dans la croissance du cancer de la prostate : de nouvelles pistes thérapeutiques. *Bull. Cancer* **94,** 135–137 (2007).
151. Piyarathna, D. W. B. *et al.* Distinct Lipidomic Landscapes Associated with Clinical Stages of Urothelial Cancer of the Bladder. *Eur. Urol. Focus* **17,** 30107–30114 (2017).
152. Tripathi, P. *et al.* HR-MAS NMR Tissue Metabolomic Signatures Cross-Validated by Mass Spectrometry Distinguish Bladder Cancer from Benign Disease NIH Public Access. *J. Proteome Res.* **5,** 3519–3528 (2013).
153. Glunde, K., Bhujwalla, Z. M. & Ronen, S. M. Choline metabolism in malignant transformation. *Nat. Rev. Cancer* **11,** 835–48 (2011).
154. Gibellini, F. & Smith, T. K. The Kennedy pathway-De novo synthesis of phosphatidylethanolamine and phosphatidylcholine. *IUBMB Life* **62,** 414–422 (2010).
155. Nowotarski, S. L., Woster, P. M., Casero, R. A. & Jr. Polyamines and cancer: implications for chemotherapy and chemoprevention. *Expert Rev. Mol. Med.* **15,** e3 (2013).
156. Thomas, T. & Thomas, T. J. Polyamine metabolism and cancer. *J. Cell. Mol. Med.* **7,** 113–126 (2003).
157. Pendyala, L. *et al.* Translational studies of glutathione in bladder cancer cell lines and human specimens. *Clin. Cancer Res.* **3,** 793–8 (1997).
158. Putluri, N. *et al.* Metabolomic profiling reveals potential markers and bioprocesses altered in bladder cancer progression. *Cancer Res.* **71,** 7376–86 (2011).
159. Singer, K., Cheng, W.-C., Kreutz, M., Ho, P.-C. & Siska, P. J. Immunometabolism in cancer at a glance. *Dis. Model. Mech.* **11,** 1–10 (2018).
160. Woolbright, B. L., Ayres, M. & Taylor, J. A. Metabolic changes in bladder cancer. *Urol. Oncol. Semin. Orig. Investig.* **36,** 327–337 (2018).
161. Rosmarin, A. G., Resendes, K. K., Yang, Z., McMillan, J. N. & Fleming, S. L. GA-binding protein transcription factor: a review of GABP as an integrator of intracellular signaling and protein–protein interactions. *Blood Cells, Mol. Dis.* **32,** 143–154 (2004).
162. Lee, J.-S. *et al.* Expression signature of E2F1 and its associated genes predict superficial to invasive progression of bladder tumors. *J. Clin. Oncol.* **28,** 2660–7 (2010).
163. Snezhkina, A. V. *et al.* Differential expression of alternatively spliced transcripts related to energy metabolism in colorectal cancer. *BMC Genomics* **17,** 1011 (2016).
164. Rosser, C. J., Urquidi, V. & Goodison, S. Urinary biomarkers of bladder

cancer: an update and future perspectives. *Biomark. Med.* **7,** 779–790 (2013).

165. Lodewijk, I. *et al.* Liquid Biopsy Biomarkers in Bladder Cancer: A Current Need for Patient Diagnosis and Monitoring. *Int. J. Mol. Sci.* **19,** (2018).
166. Di Meo, A., Bartlett, J., Cheng, Y., Pasic, M. D. & Yousef, G. M. Liquid biopsy: a step forward towards precision medicine in urologic malignancies. *Mol. Cancer* **16,** 80 (2017).
167. Bardelli, A. & Pantel, K. Liquid Biopsies, What We Do Not Know (Yet). *Cancer Cell* **31,** 172–179 (2017).
168. Shao, C.-H. *et al.* Metabolite marker discovery for the detection of bladder cancer by comparative metabolomics. *Oncotarget* **8,** 38802–38810 (2017).
169. Moestue, S., Sitter, B., Frost Bathen, T., Tessem, M.-B. & Susann Gribbestad, I. HR MAS MR Spectroscopy in Metabolic Characterization of Cancer. *Curr. Top. Med. Chem.* **11,** 2–26 (2011).
170. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–674 (2011).
171. Giskeødegård, G. F. *et al.* Spermine and Citrate as Metabolic Biomarkers for Assessing Prostate Cancer Aggressiveness. *PLoS One* **8,** e62375 (2013).
172. El Agouza, I. M., Eissa, S. S., El Houseini, M. M., El-Nashar, D. E. & Abd El Hameed, O. M. Taurine: a novel tumor marker for enhanced detection of breast cancer among female patients. *Angiogenesis* **14,** 321–330 (2011).
173. Bishnu, A. *et al.* Long term treatment of metformin impedes development of chemoresistance by regulating cancer stem cell differentiation through taurine generation in ovarian cancer cells. *Int. J. Biochem. Cell Biol.* **107,** 116–127 (2019).
174. Balendiran, G. K., Dabur, R. & Fraser, D. The role of glutathione in cancer. *Cell Biochem. Funct.* **22,** 343–352 (2004).
175. Srivastava, S. *et al.* Taurine - a possible fingerprint biomarker in non-muscle invasive bladder cancer: A pilot study by 1H NMR spectroscopy. *Cancer Biomark.* **6,** 11–20 (2010).
176. Gamagedara, S., Shi, H. & Ma, Y. Quantitative determination of taurine and related biomarkers in urine by liquid chromatography–tandem mass spectrometry. *Anal. Bioanal. Chem.* **402,** 763–770 (2012).
177. Kim, J.-W. *et al.* Metabolomic screening and star pattern recognition by urinary amino acid profile analysis from bladder cancer patients. *Metabolomics* **6,** 202–206 (2010).
178. Frantzi, M. *et al.* Development and Validation of Urine-based Peptide Biomarker Panels for Detecting Bladder Cancer in a Multi-center Study. *Clin. Cancer Res.* **22,** 4077–4086 (2016).
179. Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R. & Siuzdak, G. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using

Nonlinear Peak Alignment, Matching, and Identification. *Anal. Chem.* **78,** 779–787 (2006).

180. Kuligowski, J., Sánchez-Illana, Á., Sanjuán-Herráez, D., Vento, M. & Quintás, G. Intra-batch effect correction in liquid chromatography-mass spectrometry using quality control samples and support vector regression (QC-SVRC). *Analyst* **140,** 7810–7817 (2015).
181. Chang, C.-C. & Lin, C.-J. LIBSVM. *ACM Trans. Intell. Syst. Technol.* **2,** 1–27 (2011).
182. Xia, J., Sinelnikov, I. V., Han, B. & Wishart, D. S. MetaboAnalyst 3.0—making metabolomics more meaningful. *Nucleic Acids Res.* **43,** 251-W257 (2015).
183. Dunn, W. B., Wilson, I. D., Nicholls, A. W. & Broadhurst, D. The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. *Bioanalysis* **4,** 2249–2264 (2012).
184. Chong, I. G. & Jun, C. H. Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* **78,** 103–112 (2005).
185. Pasikanti, K. K. *et al.* Urinary metabotyping of bladder cancer using two-dimensional gas chromatography time-of-flight mass spectrometry. *J. Proteome Res.* **12,** 3865–73 (2013).
186. Shen, C. *et al.* Developing urinary metabolomic signatures as early bladder cancer diagnostic markers. *OMICS* **19,** 1–11 (2015).
187. Pasikanti, K. K. *et al.* Noninvasive urinary metabonomic diagnosis of human bladder cancer. *J. Proteome Res.* **9,** 2988–95 (2010).
188. Alberice, J. V. *et al.* Searching for urine biomarkers of bladder cancer recurrence using a liquid chromatography-mass spectrometry and capillary electrophoresis-mass spectrometry metabolomics approach. *J. Chromatogr. A* **1318,** 163–70 (2013).
189. Jin, X. *et al.* Diagnosis of bladder cancer and prediction of survival by urinary metabolomics. *Oncotarget* **5,** 1635–45 (2014).
190. Wittmann, B. M. *et al.* Bladder cancer biomarker discovery using global metabolomic profiling of urine. *PLoS One* **9,** e115870 (2014).
191. Huang, Z. *et al.* Bladder cancer determination via two urinary metabolites: a biomarker pattern approach. *Mol. Cell. Proteomics* **10,** M111.007922 (2011).
192. Gamagedara, S., Shi, H. & Ma, Y. Quantitative determination of taurine and related biomarkers in urine by liquid chromatography–tandem mass spectrometry. *Anal. Bioanal. Chem.* **402,** 763–770 (2012).
193. Nuhn, P. *et al.* External Validation of Postoperative Nomograms for Prediction of All-Cause Mortality, Cancer-Specific Mortality, and

Recurrence in Patients With Urothelial Carcinoma of the Bladder. *Eur. Urol.* **61,** 58–64 (2012).

194. Rodrigues, D. *et al.* Biomarkers in bladder cancer: A metabolomic approach using in vitro and ex vivo model systems. *Int. J. cancer* **139,** 256–68 (2016).
195. Opitz, C. A. *et al.* An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. *Nature* **478,** 197–203 (2011).
196. Li, M. *et al.* Symbiotic gut microbes modulate human metabolic phenotypes. *Proc. Natl. Acad. Sci.* **105,** 2117–2122 (2008).
197. Tan, G. *et al.* Three serum metabolite signatures for diagnosing low-grade and high-grade bladder cancer. *Sci. Rep.* **7,** 46176 (2017).

Tables

Figures

Gràcies a la Generalitat Valenciana per concedir-me un contracte d'investigació pre-doctoral (Vali+d ACIF2015/115) per a la realització d'aquesta tesi.