

Universidad de Valencia

Facultad de Medicina y Odontología

Departamento de Fisiología



**PERFILES DE EXPRESIÓN
GENÓMICA: NUEVO ENFOQUE
DIAGNÓSTICO PARA LA SEPSIS
EN RECIÉN NACIDOS
MENORES DE 1500 GRAMOS.**

Memoria realizada por Dña. María Cernada Badía
para optar al grado de Doctor en Medicina.

Dirigida por los doctores: Federico V. Pallardó
Calatayud, Eva Serna García y Máximo Vento Torres.

Los doctores D. Federico V. Pallardó Calatayud, Dña. Eva Serna García y D. Máximo Vento Torres, como directores de tesis, certifican que Dña. María Cernada Badía ha realizado el estudio titulado: “Perfiles de expresión genómica: Nuevo enfoque diagnóstico para la sepsis en recién nacidos menores de 1500 gramos”. Siguiendo el programa de Doctorado en Fisiología, y en cumplimiento de la normativa vigente de la Universidad de Valencia; la doctoranda presenta el presente proyecto para optar al grado de Doctor en Medicina.

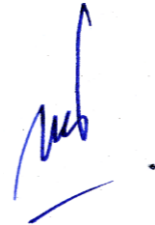
Directores:



Dr. Federico V. Pallardó

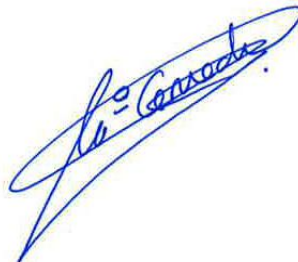


Dra. Eva Serna



Dr. Máximo Vento

Doctoranda:



María Cernada Badía

A mi familia, mi roca y mi refugio.

Gracias a mis padres, por su amor incondicional y su ejemplo constante del sentido de la responsabilidad y el valor del esfuerzo.

A Fran, por ser el mejor marido que haya podido desear. Por compartir conmigo su vida, por nuestra familia en común, por su apoyo en lo personal y lo profesional, por anteponer mis deseos a los suyos propios y conseguir que cada día sea un poco más feliz.

A mis hijos, Carlos, Fran y el pequeño Miguel que está en camino, por su alegría, por sus preguntas, por sus sonrisas y sus abrazos, por hacer que quiera ser mejor persona por y para ellos.

A mi hermana Rosa, por compartirlo todo conmigo, por estar siempre pendiente de mí y de todas mis necesidades, por ser el mejor regalo que me han hecho mis padres.

Gracias a mis directores de tesis por su impagable ayuda. A Max Vento por su fe ciega en mí y en este proyecto, por su confianza y su constante buen humor, por apoyarme, aconsejarme y ayudarme en todas y cada una de mis ideas. A Eva Serna por el soporte técnico y científico, por todas las mañanas dedicadas al

análisis de los resultados, por su entusiasmo y toda la energía positiva que transmite, pero sobre todo por su amistad. A Federico Pallardó por su sencillez y cercanía, por los buenos consejos y las ideas innovadoras.

A Asun Mejías y Octavio Ramilo por su calidísima acogida durante nuestra estancia en EEUU, por hacernos sentir como en casa. Gracias por descubirme el magnífico mundo de los microarrays y ayudarme a gestar este proyecto.

A Chema Millán, por su mano siempre tendida y dispuesta a ayudar.

A Julia Kuligowski y José Manuel Morales por su dedicación y total disponibilidad para explicarme los modelos predictivos. A Jessica Portero por las mañanas pipeteando en el laboratorio, por su trabajo concienzudo y ordenado.

A M. Carmen Collado, Gaspar Pérez y Christine Bauerl, por su ayuda y sus ideas, por abrir nuevos horizontes en cada reunión.

A todos los neonatólogos del hospital La Fe, compañeros y amigos, por los almuerzos, por las risas y por el buen ambiente que generan. Gracias por hacer que ir a trabajar cada mañana sea

un placer. A Antonio y Elena por las mañanas en el “despacho” de los becarios, por el apoyo logístico y humano, por disfrutar conmigo del periodo investigador. A Carmen, Marta, Ester y María por su amistad dentro y fuera del hospital.

Al personal de enfermería del servicio de Neonatología, por su colaboración desinteresada en la recogida de muestras, por su paciencia y buena disposición para todos los estudios.

Gracias a nuestros pequeños pacientes, grandes luchadores, y a sus familias. Gracias por mostrarnos cada día la grandeza de la vida y el inmenso poder del amor y del esfuerzo. Todo nuestro trabajo es siempre por y para ellos.

**Pedid, y se os dará; buscad, y hallaréis; llamad, y se os
abrirá. Porque el que pide, recibe; el que busca, halla; y al
que llama, se le abrirá.**

(Mateo 7:7-8)

ÍNDICE

RESUMEN	31
I. INTRODUCCIÓN	39
I. 1. Definiciones. Importancia	41
I.1.1. Etiología.....	45
I.1.2. Factores de riesgo.....	47
I.1.3. Manifestaciones clínicas.....	47
I.1.4. Diagnóstico.....	48
I.1.5. Pronóstico.....	50
I. 2. Marcadores de sepsis	52
I.2.1. Fisiopatología de la sepsis bacteriana.....	52
I.2.2. Utilidad clínica de los marcadores de sepsis.....	63
I. 3. Perfiles de expresión genómica	72
I.3.1. Biochips de expresión.....	74
I.3.2. Proceso de hibridación.....	77
I.3.3. Utilidad en el diagnóstico de la sepsis.....	81
II. HIPÓTESIS Y OBJETIVOS	83
II. 1. Hipótesis	85
II. 2. Objetivos	85
II. 2.1. Objetivo principal.....	85
II. 2.2. Objetivos secundarios.....	86

III. POBLACIÓN Y METODOLOGÍA	87
III. 1. Diseño y población	89
III. 1.1. Criterios de selección.....	90
III. 1.2. Criterios diagnósticos de sepsis.....	91
III. 1.3. Variables a estudio. Recogida de datos.....	93
III. 2. Análisis de las muestras	95
III. 2.1. Extracción de ARN total de sangre periférica.....	95
III. 2.2. Control de calidad, pureza e integridad ARN total.....	103
III.2.2.1. Integridad de la muestra de ARN total.....	103
III.2.2.2. Concentración de ARN total.....	109
III. 2.3. Estudio de la expresión genómica a escala global.....	109
III. 2.3.1. <i>GeneChip</i> [®] <i>Human Gene 1.0. ST array</i>	111
III. 2.3.2. Pre-procesado de los datos de intensidad brutos..	113
III. 2.4. Validación mediante RT-PCR en tiempo real.....	116
III. 3. Análisis estadístico	124
III. 3.1. Cálculo del tamaño muestral.....	124
III. 3.2. Análisis de variables clínicas y demográficas.....	124
III. 3.3. Análisis de los perfiles expresión genómica.....	125
III. 3.4. Modelo predictivo: regresión de mínimos cuadrados parciales con análisis discriminante.....	128

III. 3.5. Análisis RT-PCR.....	131
III. 3.6. Cohorte de validación independiente.....	133
IV. RESULTADOS.....	135
IV. 1. Cohorte inicial.....	138
IV. 1.1. Datos clínicos y demográficos.....	138
IV. 1.2. Microbiología.....	140
IV. 1.3. Análisis transcriptómico.....	141
IV. 1.3.1. Extracción del ARN total.....	141
IV. 1.3.2. Análisis de componentes principales.....	143
IV. 1.3.4. Agrupación jerárquica no supervisada.....	146
IV. 1.4. Modelo predictivo: regresión de mínimos cuadrados parciales con análisis discriminante.....	147
IV. 1.5. Genes significativos, reguladores y rutas de señalización.....	152
IV. 2. Validación por RT-PCR.....	159
IV. 3. Cohorte de validación independiente.....	162
IV. 3.1. Datos clínicos y demográficos.....	162
IV. 3.2. Microbiología.....	164
IV.3.3. Análisis de componentes principales.....	164
IV. 3.4. Modelo predictivo: regresión de mínimos cuadrados parciales con análisis discriminante.....	166
IV. 4. Análisis de las sepsis clínicas.....	167

IV. 5. Estudio diferencial según el agente causal de la sepsis.....	170
IV. 5.1. Datos clínicos y demográficos.....	170
IV.5.2. Análisis transcriptómico.....	172
IV. 5.2.1. Análisis de componentes principales.....	172
IV. 5.2.2. Agrupación jerárquica no supervisada.....	174
IV. 5.3. Modelo predictivo: regresión de mínimos cuadrados parciales con análisis discriminante.....	176
IV. 5.4. Genes significativos, reguladores y rutas de señalización.....	187
V. DISCUSIÓN Y DESARROLLO ARGUMENTAL.....	189
VI. CONCLUSIONES.....	215
VII. BIBLIOGRAFÍA.....	221
VIII. ANEXOS.....	249
VIII. 1. Consentimiento informado y aprobación del Comité Ético.....	251
VIII. 2. Listado de los genes significativos que diferencian sepsis de controles.....	257

ÍNDICE DE TABLAS

Tabla 1. Incidencia y mortalidad según el origen de la sepsis neonatal.....	44
Tabla 2. Resumen de estudios que evalúan marcadores de sepsis neonatal vertical.....	69
Tabla 3. Resumen de estudios que evalúan marcadores de sepsis neonatal nosocomial.....	70
Tabla 4. Datos clínicos y demográficos.....	95
Tabla 5. Valores de concentración de ARN total expresados en ng/μl y de integridad expresados por el valor de integridad RIN (<i>RNA Integrity Number</i>) en función de las distintas concentraciones de sangre y solución estabilizadora de ARN contenida en los Tubos Tempus®.....	99
Tabla 6. Productos requeridos para la realización del protocolo <i>GeneChip® Human Gene 1.0 ST</i>	112
Tabla 7. Esquema de trabajo para la realización del protocolo <i>GeneChip® Human Gene 1.0 ST</i>	113
Tabla 8. Características clínicas y demográficas de la cohorte inicial al nacimiento.....	139

ÍNDICE

Tabla 9. Características clínicas y demográficas de la cohorte inicial en el momento de la recogida de la muestra.....	140
Tabla 10. Valores de ARN total en cuanto a integridad expresados en RIN (<i>RNA Integrity Number</i>), pureza en ratio de absorbancia (A260/280) y concentraciones en ng/μl.....	142
Tabla 11. Listado de los 9 genes correspondientes con las 10 variables con mayor poder discriminativo en el PLS-DA.....	150
Tabla 12. Rutas de señalización más significativas de las sepsis respecto a los controles.....	152
Tabla 13. Reguladores más significativos de las sepsis respecto a los controles.....	153
Tabla 14. Genes validados por RT-PCR.....	159
Tabla 15. Valores de cambio de expresión ($2^{-\Delta\Delta C_t}$).....	160
Tabla 16. Características clínicas y demográficas de la cohorte de validación al nacimiento.....	162
Tabla 17. Características clínicas y demográficas de la cohorte de validación en el momento de la recogida de la muestra.....	163

Tabla 18. Características clínicas y demográficas al nacimiento y en el momento de la recogida de la muestra de los 13 pacientes con sepsis clínicas.....168

Tabla 19. Características clínicas y demográficas al nacimiento de los recién nacidos con sepsis en función del tipo de bacteria causal.....171

Tabla 20. Características clínicas y demográficas en el momento de la recogida de la muestra de los recién nacidos con sepsis en función del tipo de bacteria causal.....172

Tabla 21. Genes con diferencias de expresión significativas entre las sepsis por bacterias Gram positivas y bacterias Gram negativas, obtenidos con el modelo predictivo PLS-DA.....178

Tabla 22. Rutas de señalización más significativas en las sepsis por bacterias Gram positivas frente a las Gram negativas.....187

Tabla 23. Reguladores más significativos en las sepsis por bacterias Gram positivas frente a las Gram negativas.....188

ÍNDICE DE FIGURAS

Figura 1. Recién nacido prematuro de muy bajo peso.....	44
Figura 2. Gérmenes en las sepsis verticales.....	46
Figura 3. Gérmenes en las sepsis nosocomiales.....	46
Figura 4. Problemática de la sepsis.....	49
Figura 5. Activación de las células inmunes centinela.....	54
Figura 6. Fisiopatología de la respuesta inmune neonatal.....	61
Figura 7. Integración de la cascada inflamatoria y su impacto en la función de los distintos órganos y el pronóstico.....	62
Figura 8. Área bajo la curva (AUC) de IL-6 y PCR con intervalo de confianza (IC) 95%.....	67
Figura 9. Concentraciones de los marcadores TNF- α , IL-6, PCT y PCR tras el estímulo de una endotoxina infecciosa y su evolución a lo largo del tiempo.....	71
Figura 10. Dogma central de la Biología Molecular modificado de la propuesta inicial de Crick.....	73
Figura 11. Biochips diseñados por la empresa Affymetrix®	77

Figura 12. Complementariedad entre cadenas de nucleótidos.....	78
Figura 13. Proceso de tinción de las moléculas unidas al chip tras la hibridación.....	79
Figura 14. Resumen del proceso de hibridación con los biochips de oligonucleótidos de Affymetrix®	80
Figura 15. Agrupación jerárquica no supervisada de genes en función de su expresión.....	81
Figura 16. Dinámica de reclutamiento. Diagrama de flujo.....	89
Figura 17. Protocolo de extracción del ARN con el kit de aislamiento MagMAX™	98
Figura 18. Electroferogramas y geles de electroforesis de ARN obtenidos de las muestras de sangre utilizadas para la optimización del método de extracción de ARN.....	106
Figura 19. Gráfica de la absorbancia medida a longitudes de ondas de 260 y 280 de los ácidos nucleicos y de las proteínas.....	108
Figura 20. <i>GeneChip</i> ® Human Gene 1.0 ST Array.....	110
Figura 21. Imagen digital obtenida del biochip (.DAT) tras el escaneado.....	115

ÍNDICE

Figura 22. Imagen digital obtenida del biochip (.CEL).....	115
Figura 23. Curvas de amplificación del gen de estudio por RT-PCR.....	123
Figura 24. Esquema de reclutamiento.....	138
Figura 25. Distribución de los gérmenes aislados en la cohorte inicial.....	141
Figura 26. Análisis de Componentes Principales (PCA: <i>Principal component analysis</i>) tridimensional basado en todo el genoma en la cohorte inicial.....	143
Figura 27. Agrupación jerárquica no supervisada en la cohorte inicial.....	147
Figura 28. Regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) de la cohorte inicial. <i>VIP score</i>	149
Figura 29. Regresión de mínimos cuadrados parciales con análisis discriminante de la cohorte inicial (PLS-DA).....	151
Figura 30. Área bajo la curva (AUC) Característica Operativa del Receptor (ROC) de la cohorte inicial.....	151
Figura 31. Redes de los principales reguladores en las sepsis respecto a los controles.....	155

Figura 32. Integración de los procesos celulares, clases funcionales y enfermedades relacionadas con los genes obtenidos en el modelo predictivo.....158

Figura 33. Validación por RT-PCR.....161

Figura 34. Distribución de los gérmenes aislados en la cohorte de validación.....164

Figura 35. Análisis de Componentes Principales (PCA: *Principal component analysis*) tridimensional basado en todo el genoma en la cohorte de validación.....165

Figura 36. Regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) en la cohorte de validación.....167

Figura 37. Regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) en las sepsis clínicas.....169

Figura 38. Análisis de Componentes Principales (PCA: *Principal component analysis*) tridimensional basado en todo el genoma en los casos de sepsis.....173

Figura 39. Agrupación jerárquica no supervisada de la cohorte de sepsis.....175

Figura 40. Regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) en la cohorte de sepsis. *VIP score*.....177

Figura 41. Regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) en la cohorte de sepsis.....179

ÍNDICE

Figura 42. Área bajo la curva (AUC) Característica Operativa del Receptor (ROC) en el modelo predictivo según el agente causal de la sepsis.....	180
Figura 43. Test de validación cruzada: <i>Number of missclassifications</i> (NMC).....	183
Figura 44. Test de validación cruzada.....	184
Figura 45. Test de validación cruzada: área bajo la curva ROC (AUROC).....	185
Figura 46. Test de validación cruzada: Q^2 2cv.....	186
Figura 47. Perfiles de expresión genética en pacientes pediátricos con neumonía.....	197

ABREVIATURAS

ADN: Ácido desoxirribonucleico.

ADNc: Ácido desoxirribonucleico complementario.

ANOVA: Análisis de la varianza.

ARN: Ácido ribonucleico.

ARNm: Ácido ribonucleico mensajero.

AUC (*Area under curve*): Área bajo la curva.

CD177: Molécula CD177.

Ct (*Threshold cycle*): Ciclo umbral.

DAMPs (*Damage-associated molecular patterns*): Patrones moleculares asociados a daño o peligro.

EG: Edad gestacional.

FC: Frecuencia cardiaca.

FiO₂: Fracción inspiratoria de oxígeno.

GSTM1 (*Glutathion S transferase mu 1*): Glutathion S-transferasa mu 1.

HMGB-1 (*High mobility group box-1 protein*): Grupo de proteínas de alta movilidad de la caja 1.

HP (*Haptoglobin*): Haptoglobina.

Hsps: Proteínas del shock caliente.

IL: Interleuquina.

lpm: Latidos por minuto.

LCN 2 (*Lipocalin 2*): Lipocalina 2

ÍNDICE

LPS: Lipopolisacáridos.

MAPK (*Mitogen activated protein kinases*): Proteínas quinasas activadas por mitógenos.

MMP8 (*Matrix metalloproteinase 8*): Metaloproteínasa 8.

NF- κ B: Factor de transcripción nuclear kappa B.

NMC (*Number of missclassifications*): Número de clasificaciones erróneas.

NOD: Dominio de oligomerización unido a nucleótido.

OLFM4 (*Olfactomedin 4*): Olfactomedina 4.

PAMPs (*Pathogen-associated molecular patterns*): Patrones moleculares asociados a patógenos.

PBS: Solución de fosfato salino tamponada.

PCA: *Principal Component Analysis*): análisis de componentes principales.

PCR: Proteína C-Reactiva.

PCT: Procalcitonina.

PLS-DA: *Partial Least Squares Discriminant Analysis*): Regresión de mínimos cuadrados parciales con análisis discriminante.

PN: Peso al nacimiento.

PRRs: Receptores de patrones de reconocimiento.

RIN (*RNA Integrity Number*): Número de integridad del ARN.

RNMBP: Recién nacido de muy bajo peso.

ROC (*Receiver Operating Characteristic*): Curva característica operativa del receptor.

RT-PCR (*Real-time reverse-transcription polymerase chain reaction*): Reacción en cadena de la polimerasa en tiempo real con reacción de retro-transcripción.

SGB: *Streptococcus grupo B*.

SRIS: Síndrome de respuesta inflamatoria sistémica.

Th: Linfocitos T helper.

TLR (*Toll-like receptor*): Receptor *Toll-like*.

TNF- α : Factor de necrosis tumoral alfa.

VIP (*Variable importance in projection*) scores: Importancia de la variable proyectada.

VPP: Valor predictivo positivo.

VPN: Valor predictivo negativo.

RESUMEN

Background: Bacterial sepsis is associated with high morbidity and mortality in preterm infants. However, diagnosis of sepsis and identification of the causative agent remains challenging. Genome-wide expression profiles have been successfully harnessed for the diagnosis of sepsis and patient stratification based on the severity of septic shock in the pediatric and adult population. To our knowledge, no transcriptomic studies for the diagnosis of bacterial sepsis in preterm infants have been previously conducted.

Objective: Our aim was to determine genome wide expression profiles of very low birth weight (VLBW) infants (less than 1500 grams) with and without bacterial sepsis and assess differences.

Methods: Prospective observational double cohort study conducted in VLBW infants with culture-positive bacterial sepsis and non-septic matched controls. Venous blood (0.5 mL) was obtained from eligible patients before starting antibiotics, and from matched controls, mixed with RNA stabilizing solution. Patients were diagnosed with sepsis when a microorganism was isolated from blood culture. Total RNA was isolated from each sample and integrity and quality were assessed

RESUMEN

prior to hybridization into Human 1.0 ST GeneChip Array (Affymetrix®). Genome-wide expression profiles were obtained using Partek Genomic 6.6 software (Partek Inc®) and descriptive statistical analysis was completed using SPPS 17.0 (SPSS Inc®). A Partial Least Squares Discriminant Analysis was performed to obtain a predictive model was performed and the most significant genes were validated by real-time reverse-transcription polymerase chain reaction (RT-PCR) using TaqMan® probes (Applied Biosystems). Furthermore, the signature was subsequently validated in an independent validation set of patients.

Results: 17 septic VLBW infants and 19 matched controls were enrolled. Microorganisms isolated were: *Coagulase negative Staphylococcus* (10), *E. coli* (3), *E. faecalis* (2), *S. aureus* (1) y *M. morgagnii* (1). A three-dimensional unsupervised principal component analysis based on the entire genome identified three clusters of patients based on gene expression patterns: Gram positive sepsis, Gram negative sepsis and non-infected controls. Furthermore, these groups were confirmed using analysis of variance (ANOVA), which identified a transcriptional signature of 554 of differentially expressed genes with Fold Discovery Rate ≤ 0.05 between the groups.

Predictive model using PLS-DA yielded 9 genes with high discriminative power: CD177, MMP8, OLFM4, HP, GSTM1, LCN2, ANKRD22, CEACAM1 y GPR84. The most significantly overexpressed pathways in septic neonates related with innate immune and inflammatory responses and the top master regulators based on p-value and numbers of genes were: tumor necrosis factor (TNF), cytokine and nuclear factor kappa B light chain enhancer of activated B cells (NF-kB). Differences between sepsis caused by Gram positive and Gram negative bacteria were assessed. Unsupervised analysis showed 2 independent groups according to gene expression and 719 significant genes derived from ANOVA. Predictive model (PLS-DA) revealed 23 discriminant genes. Results were validated using a multiple permutation test with double cross validation.

Conclusions: Our results suggest that Genome-wide expression profiles early discriminate septic from non-septic VLBW infants. The most significant genes have been identified and validated. Differences in Genome-wide expression profiles between Gram positive and Gram negative bacterial sepsis have been described.

RESUMEN

Introducción: La disponibilidad de marcadores de sepsis precoces y fiables que permitan anticiparse al resultado del hemocultivo continúa siendo un reto. Los perfiles de expresión genómica determinan los cambios de expresión de todo el genoma a partir de una pequeña cantidad de sangre usando una biochip. Aunque en pacientes adultos y pediátricos se han obtenido resultados prometedores con genes candidatos para marcadores de sepsis, no hay estudios equivalentes en la sepsis neonatal.

Objetivo: Determinar si existen perfiles de expresión genómica específicos que diferencien recién nacidos de muy bajo peso (RNMBP) infectados de controles no infectados.

Población y Metodología: Estudio de doble cohorte, observacional y prospectivo en RNMBP. A RNBP con sospecha clínica de sepsis se les extrajo hemocultivo y 0,5 ml de sangre que se conservó en solución estabilizadora de ácido ribonucleico (ARN) previo inicio de antibiótico. Consideramos sepsis aquéllos en los que se aisló un microorganismo en hemocultivo. Se seleccionaron controles apareados por características clínicas y demográficas, a los que también se extrajeron 0,5 ml sangre. Se aisló el ARN a partir de las muestras y se

hibridaron sólo las muestras con una integridad y pureza de ARN adecuadas usando la plataforma de microarrays Human Gene 1.0 ST (Affymetrix®). Los perfiles de expresión se obtuvieron con el software Partek Genomic 6.6 (Partek Inc®) y el análisis estadístico descriptivo con SPSS 17.0 (SPSS Inc®). Se realizó un modelo predictivo con regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA: *Partial Least Squares Discriminant Analysis*, MatLab®) y los genes más significativos se validaron con reacción en cadena de la polimerasa a tiempo real combinada con una reacción de retro-transcripción (RT-PCR) usando sondas TaqMan® (Applied Biosystems). Se utilizó además una cohorte de validación externa con nuevos casos y controles.

Resultados: Se analizaron 17 casos y 19 controles con características clínicas y demográficas similares. Se aislaron: *S. coagulasa negativo* (10), *E. coli* (3), *E. faecalis* (2), *S. aureus* (1) y *M. morgagnii* (1). Se observaron diferencias de expresión entre los dos grupos con 554 genes significativos con Fold Discovery Rate ≤ 0.05 en el ANOVA. El análisis no supervisado (*Principal component analysis*) objetivó diferencias de expresión entre: controles, sepsis por bacterias Gram positivas y sepsis por bacterias Gram negativas. El modelo predictivo

RESUMEN

con PLS-DA mostró 9 genes con una muy buena capacidad de clasificación: CD177, MMP8, OLFM4, HP, GSTM1, LCN2, ANKRD22, CEACAM1 y GPR84. Los procesos biológicos más significativamente implicados fueron la respuesta inmune innata y la respuesta inflamatoria y los reguladores cuyas redes explicaron más genes: el factor de necrosis tumoral (TNF), las citoquinas y el factor de transcripción nuclear kappa B (NF- κ B).

Se analizaron las diferencias entre las sepsis por bacterias Gram positivas y por bacterias Gram negativas en la cohorte de sepsis. El análisis no supervisado mostró 2 grupos diferentes en función de la expresión a partir de 719 genes significativos y el modelo predictivo 23 genes con una gran capacidad de clasificación. Se validaron los resultados con un test de permutaciones múltiples y doble validación cruzada.

Conclusiones: Los perfiles de expresión genómica permiten discriminar a los RNBP con sepsis de los controles no infectados. Se han identificado y validado los genes más significativos que podrían ser futuros marcadores. Se han observado diferencias de expresión genómica entre las sepsis por bacterias Gram positivas y Gram negativas.

I. INTRODUCCIÓN

I.1. DEFINICIONES. IMPORTANCIA

Se conoce por período neonatal por consenso internacional aquel que comprende desde el momento del nacimiento hasta el día 28 después del mismo. Se define una gestación a término aquella que tiene una duración entre 37 y 42 semanas postconcepción. En el mundo desarrollado la mortalidad de los recién nacidos a término es inferior a 3-4 casos por cada mil nacidos vivos. Además, la morbilidad de los mismos es asimismo escasa. Sin embargo, en el caso de los nacidos pretérmino, es decir, aquellos cuya gestación tiene una duración inferior a 37 semanas las tasas de mortalidad y morbilidad son mucho más elevadas y serán tanto mayores cuanto menor sea la edad de gestación. La viabilidad en el momento actual se sitúa en torno a las 23 semanas de gestación. Por otra parte, la tasa de prematuridad también está relacionada con el nivel de desarrollo económico y social de los países; siendo las cifras más bajas en torno al 7-8% en Europa, mientras que en países en vías de desarrollo pueden llegar al 30%-40%¹.

Nuestras aproximaciones diagnósticas y terapéuticas parecen haber alcanzado un techo y estas cifras no han experimentado variaciones significativas en el mundo industrializado en la última década. Es más,

INTRODUCCIÓN

recientes publicaciones hacen hincapié en la necesidad urgente de aplicar los nuevos avances científicos en la mejoría de la mortalidad pero, especialmente, de la morbilidad neonatal². Una de las principales causas no sólo de mortalidad sino también de morbilidad neonatal son las infecciones, especialmente en los prematuros de bajo peso³.

Se define **infección** como el síndrome o condición clínica causada por la presencia de un microorganismo. Ante una agresión como es la infección, el huésped puede responder generando un **síndrome de respuesta inflamatoria sistémica (SRIS)**. Exige la presencia de temperatura rectal $\geq 38.5^{\circ}\text{C}$ o $< 36^{\circ}\text{C}$ o leucocitosis/neutropenia/ $>10\%$ formas jóvenes, unido a otro criterio como: taquicardia o bradicardia en ausencia dolor o drogas, taquipnea o necesidad de ventilación mecánica. Hablamos de **sepsis** cuando el SRIS es secundario a una infección sospechada o comprobada⁴. Hay distintos tipos de sepsis en el recién nacido en función del origen de los gérmenes:

Sepsis vertical: sepsis causada por gérmenes localizados en el canal genital materno y que contaminan al feto por vía ascendente o por contacto directo con secreciones contaminadas en el canal del parto.

Sepsis horizontal nosocomial: sepsis adquirida en el contexto de un ingreso hospitalario, causada por gérmenes que colonizan al recién nacido (RN) a través del personal sanitario (habitualmente las manos) o de materiales diagnósticos y/o terapéuticos contaminados.

La sepsis neonatal constituye un problema importante no sólo por su incidencia, que oscila entre 1.5-1.9% en el caso de la sepsis vertical y el 20% en la horizontal, sino también por sus altas tasas de mortalidad que se sitúan en torno al 18%⁵⁻⁶, sobre todo en los recién nacidos de muy bajo peso (RNMBP). Es más, con la disminución del límite de la viabilidad a las 23 semanas, se observa una tendencia creciente en las cifras de sepsis en recién nacidos pretérmino que precisan cuidados intensivos durante largos periodos de tiempo, están sometidos a una estrecha monitorización y a multitud de procedimientos invasivos (figura 1) que incrementan el riesgo de infección⁷. Según el origen de la sepsis: vertical o nosocomial se han descrito distintas tasas de incidencia y mortalidad^{5, 8} (tabla 1).

INTRODUCCIÓN



Figura 1. Recién nacido prematuro de muy bajo peso. Se trata de un recién nacido de 24 semanas de edad gestacional, 9 días de vida y 610 gramos de peso. Se encuentra intubado y conectado a ventilación mecánica. Está monitorizado con electrocardiograma y pulsioxímetro y es portador de una sonda nasogástrica y de una vía central. La cinta métrica marca la longitud en centímetros. (Fotografía cedida por la familia).

Tabla 1. Incidencia y mortalidad según el origen de la sepsis neonatal^{5, 8}.

	VERTICAL	NOSOCOMIAL
Incidencia global	1.1/1000 recién nacidos vivos	3.5/ 100 ingresos
Incidencia menores 1500 g	17.5/1000 recién nacidos vivos	28/ 100 ingresos
Mortalidad global	8.7/100 infectados	11.8/ 100 infectados
Mortalidad menores 1500 g	30.6/100 infectados	17.3/ 100 infectados

I. 1.1. ETIOLOGÍA

En cuanto a la etiología, en las sepsis verticales los gérmenes más comúnmente aislados son las bacterias Gram negativas, encabezadas por *Escherichia coli*, alcanzando cifras cercanas al 60% (figura 2). Las bacterias Gram positivas, en especial el *Streptococcus* del grupo B han sufrido una importante reducción desde la implantación de la profilaxis intraparto desde cifras de 1.7/1000 recién nacidos vivos hasta 0.3 casos/ 1000 recién nacidos vivos en la actualidad⁹.

Sin embargo, en las sepsis nosocomiales los gérmenes Gram positivos superan el 60% de los casos, en especial *Staphylococcus coagulasa negativo* y *Staphylococcus aureus*, seguidos por un 25-30% de Gram negativos y en último lugar un 10-15% de infecciones por *Candida*. En las figuras 2 y 3 se representan la distribución por etiologías de las sepsis verticales y nosocomiales respectivamente en los últimos 5 años en el servicio de Neonatología de Hospital Universitario y Politécnico La Fe, Valencia (HUiP La Fe), que son muy similares a las descritas en la bibliografía¹⁰.

INTRODUCCIÓN

GÉRMENES SEPSIS VERTICALES (S. Neonatología H. La Fe 2010-2014)

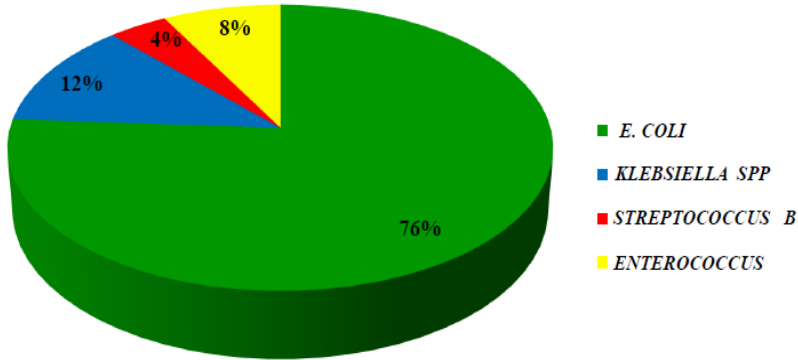


Figura 2. Gérmenes en las sepsis verticales. Distribución de gérmenes en las sepsis verticales diagnosticadas en el Servicio de Neonatología del Hospital Universitario y Politécnico La Fe desde el año 2010 al año 2014 ambos inclusive.

GÉRMENES SEPSIS NOSOCOMIALES (S. Neonatología H. La Fe 2010-2014)

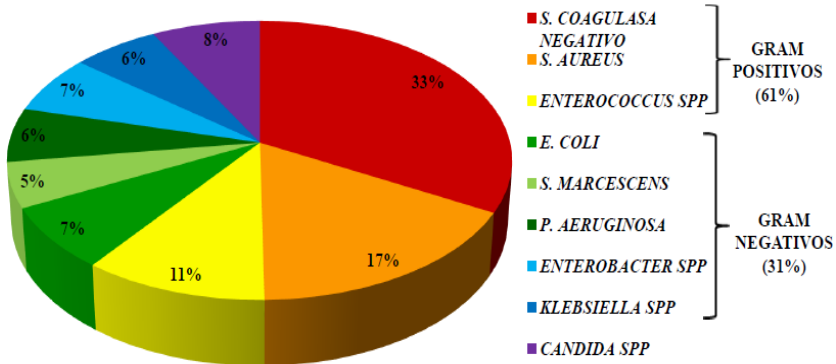


Figura 3. Gérmenes en las sepsis nosocomiales. Distribución de gérmenes en las sepsis nosocomiales diagnosticadas en el Servicio de Neonatología del Hospital Universitario y Politécnico La Fe desde el año 2010 al año 2014 ambos inclusive.

I. 1.2. FACTORES DE RIESGO

Se han descrito distintos factores de riesgo que pueden favorecer el desarrollo de sepsis. Entre las verticales destacan el antecedente de corioamnionitis materna, la colonización materna por *Streptococcus* del grupo B, la rotura prematura de membranas más allá de las 18 horas, el parto pretérmino o la gestación múltiple. Mientras que los factores de riesgo descritos en las sepsis nosocomiales son la prematuridad, el bajo peso al nacimiento, el uso prolongado de catéteres intravenosos y/o de antibióticos de amplio espectro, los procedimientos invasivos o la conexión a ventilación mecánica¹¹.

I. 1.3. MANIFESTACIONES CLÍNICAS

Las manifestaciones clínicas de sepsis en los neonatos, y sobre todo en los menores de 1500 gramos de peso al nacimiento, son muy sutiles e inespecíficas y, a menudo, indistinguibles de las propias de la transición a la vida extrauterina. Estos signos incluyen fiebre o hipotermia, dificultad respiratoria, cianosis, apneas, intolerancia digestiva, distensión abdominal, vómitos, hepatomegalia, ictericia no filiada, mala perfusión periférica letargia, irritabilidad o convulsiones¹². El impacto de la sepsis va más allá del compromiso clínico e incluye las complicaciones derivadas de la infección (peor

INTRODUCCIÓN

evolución en el neurodesarrollo, riesgo de mortalidad), los posibles efectos secundarios del tratamiento antibiótico en cuanto a nefro y ototoxicidad, el estrés que supone para la familia el empeoramiento clínico pudiendo afectar a la producción de leche materna tan necesaria para la buena evolución de estos niños, el dolor que se les inflige por las venopunciones y contantes analíticas para monitorizar los niveles de antibiótico o la evolución de la infección y el coste económico por el aumento de la estancia hospitalaria, del gasto sanitario y de las complicaciones a largo plazo¹³.

I. 1.4. DIAGNÓSTICO

La prueba diagnóstica de elección (*gold standard*) se ha basado clásicamente en la confirmación microbiológica mediante el aislamiento de un germen en el hemocultivo⁴. Sin embargo, los resultados del hemocultivo tardan un mínimo de 48-72 horas y son poco sensibles para detectar neonatos infectados cuyas madres han recibido tratamiento antibiótico durante el parto, arrojando a menudo resultados que son falsos negativos¹⁴⁻¹⁵. Por otra parte, la escasa volemia en los RNMBP, limita en muchas ocasiones la obtención de la cantidad necesaria de sangre para la correcta interpretación de los hemocultivos, obteniendo volúmenes de 0.5-1 ml en lugar de 3 ml, lo

que disminuye la sensibilidad pudiendo condicionar así mismo falsos negativos¹⁶, ya que hasta el 60-70% de los neonatos presentan un bajo grado de bacteriemia en el momento del diagnóstico¹⁷. Es por ello, que en la práctica habitual, los neonatos con factores de riesgo infeccioso y manifestaciones clínicas compatibles, aunque inespecíficas, se diagnostican como “sepsis clínicas” y son tratados con antibióticos aunque el resultado del hemocultivo sea negativo. Esto conlleva, en ocasiones, a un uso indiscriminado de antibióticos, no exentos de efectos secundarios, y que no sólo aumenta los costes de forma innecesaria si no que produce cambios en la microbiota intestinal¹⁸ y favorece el desarrollo y el aumento en el número de resistencias antibacterianas (figura 4).



Figura 4. Problemática de la sepsis. ATB: Antibiótico.

I. 1.5. PRONÓSTICO

La evolución clínica depende de numerosos factores que incluyen la edad gestacional, estado general del niño en el momento del diagnóstico y presencia o no de otras condiciones perinatales que agravan la infección. La mortalidad en general es del 20%, con cifras cercanas al 40% en los casos de sepsis verticales, en las sepsis acompañadas de meningitis o en presencia de otras complicaciones concomitantes. En relación a la etiología, la mortalidad en las sepsis verticales parece ser independiente del germen; sin embargo en las nosocomiales las infecciones por *Pseudomonas spp*, *Candida spp* o *Listeria spp*, alcanzan cifras de mortalidad entre el 35 y el 100% de los casos¹⁹. Aun cuando no conduzcan a la muerte, las infecciones y la respuesta inflamatoria que desencadenan contribuyen al desarrollo de complicaciones a largo plazo en el recién nacido pretérmino²⁰: desarrollo de displasia broncopulmonar, daño del sistema nervioso central sobre todo a expensas de lesión en la sustancia blanca²¹, y por consiguiente una peor evolución en el neurodesarrollo motor y cognitivo²²⁻²³. En los menores de 1000 gramos de peso al nacimiento, las infecciones se relacionan no sólo con problemas en el neurodesarrollo, sino también con el crecimiento⁶. La sepsis constituye uno de los factores que más influyen en la evolución a

50

largo plazo del prematuro; siendo un factor de mal pronóstico neurológico a los 2 años en los menores de 28 semanas, de manera independiente de otros factores de riesgo neurológico²⁴. Aunque las infecciones por bacterias Gram negativas se ha asociado con mayor gravedad clínica y tasas de mortalidad más elevadas¹⁹, las infecciones por bacterias Gram positivas cuadruplican el riesgo de parálisis cerebral y duplican el riesgo de discapacidad²⁴. Es más, incluso las infecciones por el germen nosocomial más frecuente, el *Staphylococcus coagulasa negativo*, que presentan escasa virulencia y bajas tasas de mortalidad en comparación con otras infecciones bacterianas, parecen asociarse a una mayor probabilidad de retraso cognitivo en los prematuros menores de 29 semanas²⁵.

El pronóstico de la sepsis neonatal depende tanto de la administración precoz como de la efectividad del tratamiento antibiótico escogido, pudiendo presentar un curso complicado con metástasis sépticas, coagulación intravascular diseminada, insuficiencia cardíaca congestiva y shock, en los casos de diagnóstico tardío, por lo que es esencial un diagnóstico fiable y precoz¹². En este sentido, la búsqueda de marcadores diagnósticos fiables y precoces continúa siendo un reto para los neonatólogos.

I.2. MARCADORES DE SEPSIS

I.2.1. FISIOPATOLOGÍA DE LA SEPSIS BACTERIANA

Reconocimiento de antígenos

La respuesta inmune comienza con el reconocimiento del patógeno, que en el caso de la sepsis se trata de un microorganismo, por parte de las células inmunes centinela (monocitos, macrófagos) a través de la activación de receptores de patrones de reconocimiento (PRRs) entre los que destacan los receptores *Toll-like* (TLRs). Estos receptores se encuentran en múltiples tipos celulares y reconocen los productos microbianos conocidos como patrones moleculares asociados a patógenos (*Pathogen-associated molecular patterns*: PAMPs), desencadenando la cascada inflamatoria y activando la respuesta inmune. El paradigma de estos PAMPs son las endotoxinas de lipopolisacáridos (LPS) que se encuentran en la superficie de los gérmenes Gram negativos, y el ácido lipoteicoico de la pared de los Gram positivos. Los LPS activan los TLR4 en conjunción con las proteínas de superficie CD14 y de diferenciación mieloide MD2, mientras que la señal del ácido lipoteicoico activa los TLR2. Sin embargo, en ocasiones, los microorganismos activan más de un TLR

simultáneamente con el fin de desencadenar una respuesta inmune del huésped específica para cada patógeno²⁶.

Además de los TLRs, existen otros PRR intracelulares, entre los que se incluyen aquellos con un dominio de oligomerización unido a nucleótido (NOD): los receptores NOD-like (NLRs), que detectan los peptidoglicanos de las bacterias Gram positivas, y los receptores Rig-1-like (RLRs). Una vez acoplados a los patógenos, los PRRs inician la respuesta inmune a través de una serie de segundos mensajeros, estimulando la producción de citoquinas a través de las *mitogen activated protein kinases* (MAPK) y el factor de transcripción nuclear kappa B (NF-κB), con los que se amplifica la respuesta inmune²⁷ (figura 5).

Los TLRs pueden ser activados por patrones moleculares asociados a daño o peligro (*Damage-associated molecular patterns: DAMPs*) además de por PAMPs, a través de la liberación de proteínas o mediadores intracelulares dañados o muertos. Estos DAMPs, al igual que las proteínas del tipo High mobility group box-1 (HMGB-1), están implicados en la progresión de sepsis a shock séptico. HMGB-1 la producen macrófagos y células endoteliales estimuladas por LPS o

INTRODUCCIÓN

factor de necrosis tumoral alfa (TNF- α) a través de TLR2 y TLR4. Sus funciones incluyen la producción de citoquinas, la activación de la coagulación y el reclutamiento de neutrófilos²⁸. Existen otros DAMPs como proteínas del shock caliente (Hsps) o el ácido úrico, que activan la señal proinflamatoria y contribuyen al desarrollo del shock séptico como coadyuvantes inmunes. Estas moléculas, no han sido estudiadas en la sepsis neonatal²⁹.

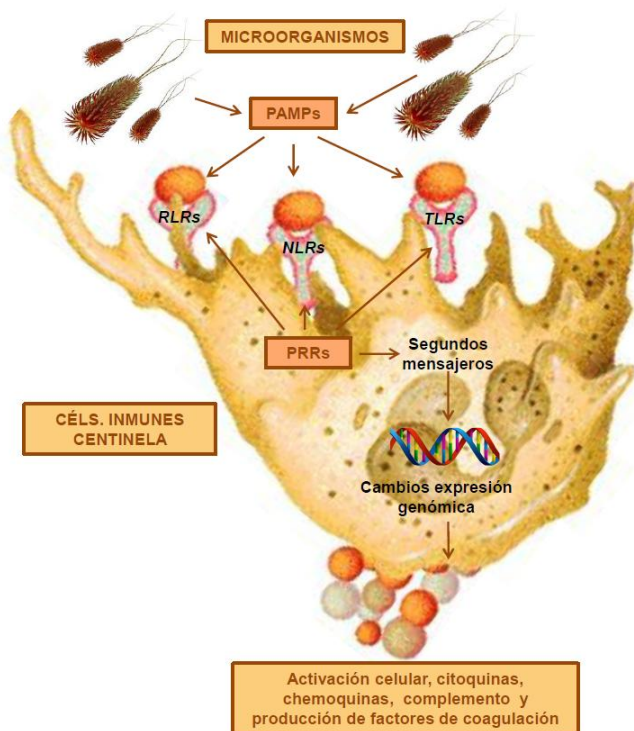


Figura 5. Activación de las células inmunes centinela. PAMPs: Patrones moleculares asociados a patógenos, PRRs: Receptores de patrones de reconocimiento, NLRs: Receptores NOD-like, RLRs: Receptores Rig-1-like.

Producción de citoquinas y activación de la cascada inflamatoria

Tras la estimulación de los PRR, se inicia la producción de Óxido Nítrico (NO), leucotrienos, prostaglandinas, factor activador de plaquetas, complemento, citoquinas y quemoquinas proinflamatorias como interleuquinas (citoquinas producidas en células hematopoyéticas): IL-1 β , IL-6, IL-8, IL-12, IL-18, interferón gamma (INF- γ) y TNF- α , amplificando la respuesta inmune innata contra los microorganismos invasores³⁰ y generando una respuesta inflamatoria caracterizada por un aumento de la permeabilidad vascular y el reclutamiento de células inflamatorias.

Las citoquinas son proteínas cuya acción fundamental es la regulación de la inflamación, actuando como mediadores de la comunicación intercelular, induciendo la activación de receptores específicos de membrana e interviniendo en funciones de proliferación y diferenciación celular, quimiotaxis, crecimiento y modulación de la secreción de inmunoglobulinas. Son producidas fundamentalmente por los linfocitos y los macrófagos activados, aunque también pueden ser producidas por otras células como: leucocitos polimorfonucleares, adipocitos, células endoteliales, epiteliales, y del tejido conjuntivo. Sin embargo, en comparación con los adultos, los neonatos producen

INTRODUCCIÓN

menor cantidad de IL-1 β , TNF- α , INF- γ e IL-12, debido a una menor producción de mediadores de la señal de los TLRs como el factor de diferenciación mieloide 88 (MyD88), el factor regulador de interferon 5 (IRF5) y el p38 en relación inversa a la edad gestacional ³¹.

La producción de citoquinas proinflamatorias conduce a la activación de células endoteliales y la sobreexpresión de moléculas de adhesión celular (CAM): ICAM, VCAM, L-selectina, P-selectina y E-selectina y CD11b/CD18, que facilitan el reclutamiento de leucocitos y la diapedesis (migración extravascular hacia el foco de infección). La quimioataxis de leucocitos se ve favorecida, además de por la sobreexpresión de las CAM, por la acción de gradientes de quemoquinas producidas por células endoteliales y macrófagos locales. Entre las sustancias quimio-atractivas se encuentran: la proteína quimioatrayente de monocitos (MCP-1), la proteína inflamatoria de macrófagos MIP-1, la proteína inducida por INF- γ (IP-10), IL-8, quemoquina regulada tras la activación normal de las células T expresadas y secretadas (RANTES), proteínas del complemento como C3a y C5a y proteínas de defensa del huésped como catelicidinas y defensinas.

En la respuesta inflamatoria del recién nacido se ha evidenciado una pobre quimiotaxis celular, que no parece relacionada con una baja concentración sérica de citoquinas³², sino con una insuficiente regulación al alza del receptor del complemento y una inhibición por los productos bacterianos entre otras causas³³.

Los leucocitos polimorfonucleares de los recién nacidos en relación con los del adulto, presentan déficits cuantitativos y cualitativos. Durante la infección, se produce una rápida depleción de las reservas medulares de leucocitos polimorfonucleares sobre todo en los casos de sepsis por gérmenes Gram negativos³⁴. Por otra parte tienen una apoptosis retardada y una capacidad aumentada para activar funciones citotóxicas³⁵. La deformabilidad está reducida lo que predispone a una agregación irreversible en el espacio intravascular, que se traduce en una menor diapedesis y en el aumento de la depleción de las reservas medulares, liberándose formas inmaduras de neutrófilos (bandas), que todavía presentan mayor disfunción que las formas maduras³⁶. Todo ello hace del recién nacido un individuo especialmente susceptible a la infección.

INTRODUCCIÓN

Respuesta antiinflamatoria

La respuesta inflamatoria es beneficiosa cuando las citoquinas se producen en cantidad adecuada; pero puede ser nociva, e incluso letal, si se producen en exceso³⁷. Es por ello, que además de citoquinas proinflamatorias, existen citoquinas antiinflamatorias: IL-4, IL-10, IL-11, IL-13, factor de crecimiento transformante beta (TGF- β), que se encargan de contrarrestar esta respuesta mediante la supresión de la activación de macrófagos y de la producción de TNF- γ , IL-1, IL-6, IL-8 e IFN- γ ³⁰. Estos mediadores, en especial la IL-10 liberada por células Th2, B y macrófagos en respuesta a la presencia de TNF α , bloquean la activación de las células fagocíticas y la fiebre, modifican la expresión del factor de coagulación y disminuyen la expresión de especies reactivas intermedias de oxígeno y nitrógeno y otros mediadores vasoactivos³⁸. Sin embargo, una sobreproducción de estos mediadores, puede conducir a la supresión de la función inmune. Es por ello, que en las sepsis de recién nacidos pretérmino de muy bajo peso, se ha observado una activación de tanto citoquinas proinflamatorias, como antiinflamatorias, en relación con controles no infectados, produciéndose una disminución de estos valores a las 48 horas de instaurar un adecuado tratamiento antibiótico³⁰. Además de las citoquinas antiinflamatorias, hay citoquinas solubles y receptores

antagonistas como sIL-6R, sIL-2, IL-1ra o TNFR2, que también modulan la acción pro-inflamatoria y cuya elevación se ha documentado en casos de sepsis neonatal tratados con éxito³⁹.

Papel del complemento

El complemento es un componente fundamental de la inmunidad innata precoz que facilita la opsonización (proceso por el que se marca a un patógeno para su ingestión y destrucción por un fagocito) y la actividad microbicida. Posee actividad quimiotáctica y anafiláctica que incrementa la agregación leucocitaria y la permeabilidad vascular local. Activa de manera recíproca otros procesos como la coagulación, la producción de citoquinas proinflamatorias y la activación de leucocitos. En los recién nacidos, la opsonización a través del complemento está limitada y es tanto más pobre cuanto menor es la edad gestacional. Esto, sumado a unos niveles bajos de proteínas del complemento y una función deprimida, hace a los neonatos prematuros especialmente susceptibles a la infección⁴⁰.

Activación linfocitaria y de la respuesta inmune

El patrón de citoquinas circulantes, así como las características del patógeno, determinan la diferenciación del precursor de linfocitos T

INTRODUCCIÓN

helper (Th0) hacia linfocitos helper 1 (Th1) o 2 (Th2). Los Th1 producen $\text{INF}\gamma$, IL2 y $\text{TNF}\beta$ que promueven la inmunidad celular y la actividad fagocítica, especialmente en el caso de infecciones intracelulares. Los Th2 producen IL4, IL5, IL6, IL9, IL10, IL13 desencadenando la inmunidad humoral y la producción de anticuerpos⁴¹.

Reactantes de fase aguda y opsoninas

Además de la respuesta inflamatoria inicial, la presencia de microorganismos aumenta la producción de proteínas innatas con una importante función inmunológica y que reducen la carga bacteriana. Son reactantes de fase aguda, que se producen principalmente en el hígado, como las proteínas que aumentan la permeabilidad bactericida, colectinas, lactoferrina, haptoglobina, fosfolipasa A2, procalcitonina, proteína C-Reactiva (PCR) (que actúa como opsonina), amiloide A sérico (interviene en el reclutamiento celular).

Por otra parte, la sepsis genera un aumento de otras proteínas séricas con función de opsonización como la fibronectina y anticuerpos naturales, sobre todo IgM, que producen linfocitos B circulantes. Sin embargo, el plasma de los recién nacidos presenta una actividad de

opsonización disminuida y una deficiente producción de proteínas que aumentan la permeabilidad bactericida por parte de los neutrófilos, lo que les confiere una mayor susceptibilidad a la infección⁴².

En la figura 6 se muestra de manera esquemática la fisiopatología de la respuesta inmune neonatal.

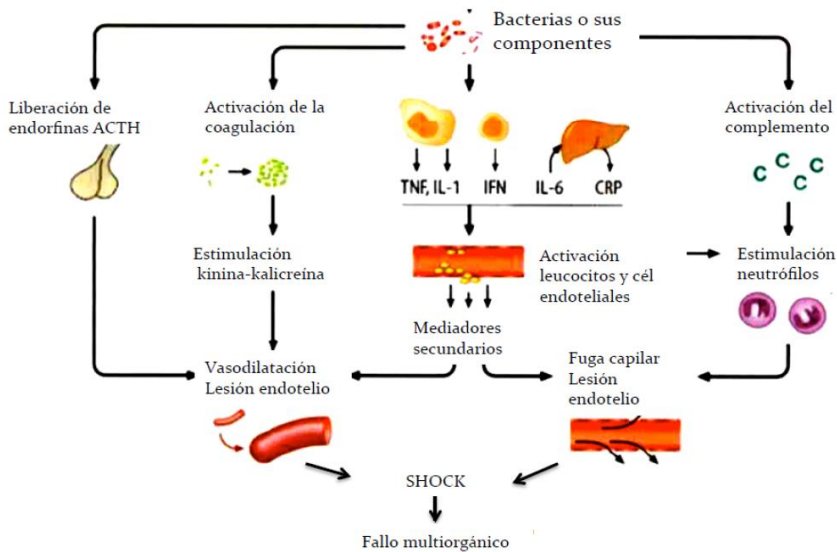


Figura 6. Fisiopatología de la respuesta inmune neonatal.

La respuesta inmune y la cascada inflamatoria que genera tienen un impacto en una serie de órganos como cerebro, riñón, pulmón, sistema cardiovascular, hígado, intestino y en la microcirculación, generando

INTRODUCCIÓN

la sintomatología propia de la sepsis. Si finalmente se consigue un control efectivo de la causa, se normalizarán los biomarcadores, se resolverá la disfunción orgánica y el sujeto se recuperará. Por el contrario, si la respuesta no es efectiva, persistirá la anormalidad en los marcadores, que conducirán al fallo multiorgánico y la muerte (figura 7).

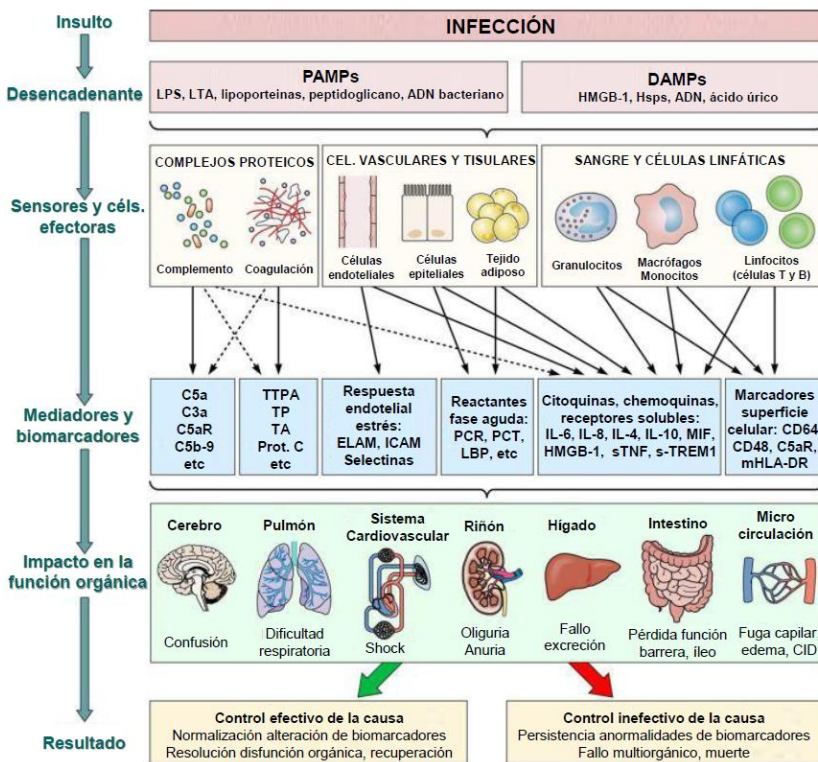


Figura 7. Integración de la cascada inflamatoria y su impacto en la función de los distintos órganos y el pronóstico. Reproducida con permiso de: <https://rjaramillo.wordpress.com/2012/11/09/sepsis>.

I.2.2. UTILIDAD CLÍNICA DE LOS MARCADORES DE SEPSIS

Aunque la prueba diagnóstica de elección continua siendo el aislamiento de un germen, el hemocultivo presenta importantes limitaciones por lo tardío de sus resultados⁴³ y la posibilidad de falsos negativos¹⁴⁻¹⁵, haciendo necesarios marcadores diagnósticos fiables y precoces. La mayor parte de estudios se han dirigido a investigar diversos mediadores de la cascada inflamatoria, ya que la sepsis no está confinada a ningún órgano en concreto. Asumiendo que todos los marcadores van a presentar un porcentaje de falsos negativos y falsos positivos, dada la importante mortalidad de la sepsis y la morbilidad asociada, el marcador ideal debería ser precoz, tener una gran sensibilidad (proporción de niños con infección que presentan un test diagnóstico positivo) y un valor predictivo negativo (proporción de niños no infectados sobre el total de test diagnósticos negativos) lo más cercanos posible al 100%. Esto no debería ser a expensas de perder especificidad (proporción de test negativos entre los controles) ni valor predictivo positivo (proporción de casos de infección entre los test diagnósticos positivos), que deberían sobrepasar el 85%⁴⁴.

Pese a que se han investigado múltiples marcadores analíticos para el diagnóstico de sepsis neonatal, no hay consenso sobre la superioridad

INTRODUCCIÓN

de un marcador sobre otro y ninguno ofrece una sensibilidad y una especificidad adecuadas. Tradicionalmente se han utilizado el recuento de leucocitos totales y la ratio inmaduros/totales, pese a su baja especificidad y la posibilidad de verse afectado por diversas patologías asociadas en el recién nacido⁴⁵. Posteriormente, diferentes estudios han evaluado otros marcadores que se pensó eran más específicos e igualmente implicados en la respuesta inmune, entre los que se encuentran por su accesibilidad y su utilidad clínica: la PCR, la IL-6 y la Procalcitonina (PCT).

Proteína C Reactiva (PCR)

Es una proteína producida en el hígado tras la activación de la cascada inflamatoria y en respuesta a la presencia de la IL6 en plasma, entre otras citoquinas. Actúa como opsonina. Su síntesis comienza a las 8 horas de producirse la exposición a un agente infeccioso; sin embargo su vida media es de 12-24 horas. Ha demostrado mayor sensibilidad y especificidad que el recuento de neutrófilos⁴⁵ y la normalización de sus valores de utilidad en la monitorización de la respuesta al tratamiento antibiótico⁴⁶, sin embargo como por su lenta elevación en la fase inicial, la sensibilidad en el momento de la sospecha de sepsis

puede llegar a descender hasta el 60%⁴⁶ por lo que se puede considerar un marcador específico pero tardío de sepsis⁴⁷.

Procalcitonina (PCT)

La procalcitonina es un péptido precursor de la calcitonina sin actividad hormonal y cuyos niveles son independientes de ésta. Actúa como reactante de fase aguda, y aunque no se ha identificado su lugar de producción con exactitud, se postulan los monocitos y hepatocitos como las principales fuentes. Sus concentraciones séricas se elevan más precozmente que las de la PCR⁴⁸⁻⁵⁰, comenzando a las 4 horas de la exposición al agente infeccioso y permaneciendo elevadas al menos 24 horas con una vida media de 25-30 horas⁵¹. Parece tener valores de sensibilidad y especificidad superiores a los de la PCR, y ha demostrado ser útil a la hora de diagnosticar precozmente la infección⁵², determinar su severidad y pronóstico⁵⁰, así como para monitorizar la respuesta al tratamiento antibiótico⁵³. No obstante, la elevación fisiológica de los niveles PCT en las primeras 48 horas de vida descrita en algunos estudios⁵⁴⁻⁵⁵ podría limitar su valor diagnóstico en los casos de sepsis vertical. Esta elevación parece deberse no tanto a un paso transplacentario como a una producción endógena, dado que los valores de PCT son mayores en la sangre del

INTRODUCCIÓN

recién nacido que en sangre materna y las diferencias aumentan a las 24 y 48 h tras el parto⁵⁶.

Interleuquina 6 (IL-6)

En los últimos años se ha estudiado el papel de la determinación de citoquinas, especialmente la interleuquina-6 (IL-6), como marcadores precoces de sepsis neonatal⁵⁷⁻⁶⁰. La IL6 es una citoquina proinflamatoria pleiotrópica, producida fundamentalmente por el sistema monocito/macrófago en respuesta a la producción de IL1 y TNF tras el reconocimiento de los PAMPs por parte de los PRRs. Entre sus múltiples funciones se encuentra la regulación de la respuesta del huésped a la infección, actuando como una señal para la activación de las células T, promoviendo la secreción de anticuerpos por parte de las células B, la diferenciación de las células T citotóxicas y activando la liberación de otras citoquinas. Estimula los hepatocitos para iniciar la producción de PCR, siendo su elevación por tanto mucho más precoz que la de ésta. Sus valores aumentan de manera inmediata en las primeras 2 horas tras el estímulo infeccioso con un pico máximo entre las 4 y 6 horas; sin embargo su vida media es muy corta siendo prácticamente indetectable tras 24 horas⁴⁶.

La IL-6, por su cinética, parece ser un marcador más precoz y fiable que la PCR en el diagnóstico de la sepsis vertical. En un estudio de cohortes prospectivo realizado por nuestro grupo, determinamos IL-6 y PCR en sangre de cordón umbilical de todos los recién nacidos con factores de riesgo infeccioso prenatal. De 128 neonatos reclutados, 10 de ellos desarrollaron sepsis, siendo los valores predictivos (VP) de la IL-6 (VP positivo: 7.14, VP negativo: -0.11) superiores a los de la PCR (VP positivo: 2.86, VP negativo: -0.51)⁶¹ (figura 8).

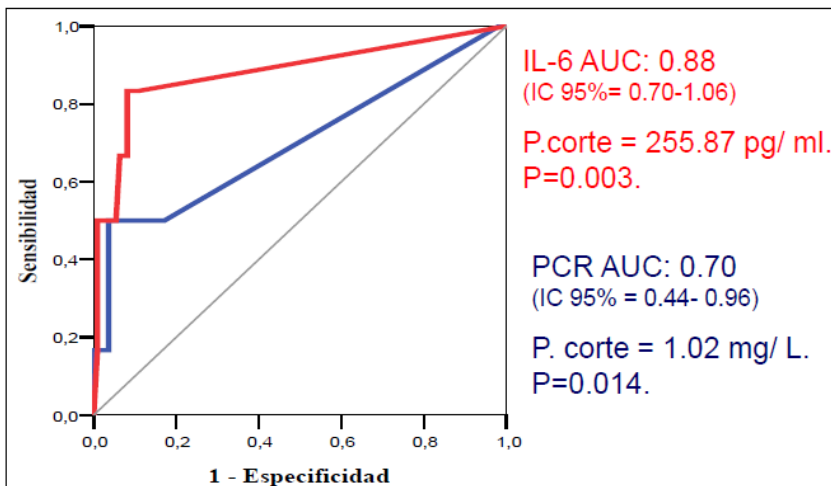


Figura 8. Área bajo la curva (AUC) de IL-6 y PCR con intervalo de confianza (IC) 95%⁶¹.

INTRODUCCIÓN

Sin embargo, otros estudios encuentran valores de especificidad menos óptimos e incluso elevaciones significativas en recién nacidos con patología grave no infecciosa^{48, 62-63}.

En esta línea, se han seguido buscando marcadores focalizados en una respuesta inmune más concreta y precoz, como el factor de necrosis tumoral alfa (TNF- α)⁶⁴⁻⁶⁶, otras citoquinas como IL8, IL-1 β ⁶⁶⁻⁶⁸ o IL-10⁶⁹⁻⁷⁰, RANTES⁶⁹, IP10⁷¹, P y E-Selectina⁷²⁻⁷³ que parecen no variar en neonatos con sepsis en otro trabajo donde destaca como marcador más sensible y específico sICAM1⁷⁴, los antígenos de superficie de los neutrófilos CD 64⁷⁵⁻⁷⁷, CD 14⁷⁸⁻⁷⁹, los TLR⁸⁰ o la sustancia Amiloide A⁸¹ que siguen sin demostrar, por sí mismos, superioridad a los tradicionalmente usados PCR, PCT e IL6.

Los distintos estudios difieren no sólo en su diseño y en los criterios diagnósticos de sepsis, sino también en los valores que toman como puntos de corte, arrojando en ocasiones resultados contradictorios^{48-49, 64, 82-86} (tablas 2 y 3). De manera aislada ninguno ha demostrado ser suficientemente sensible y específico, recomendándose el uso combinado de varios de ellos⁴⁴.

Tabla 2. Resumen de estudios que evalúan marcadores de sepsis neonatal vertical. (S: Sensibilidad, E: Especificidad, VPP: Valor predictivo positivo, VPN: Valor Predictivo Negativo. PCR: Proteína C-Reactiva, PCT: Procalcitonina, IL-6: Interleuquina-6). La sepsis fue considerada con cultivo positivo y/o clínica compatible. N hace referencia al número de pacientes incluidos en el estudio.

AUTOR	SEPSIS	PUNTO CORTE	S (%)	E (%)	VPP (%)	VPN (%)
Joram ⁸² (N= 197)	Cultivo	PCR \geq 5 mg/L PCT \geq 0.5 ng/mL	50 87	97 91	67 87	94 99
Pérez-Solís ⁸³ (N=133)	Cultivo Clínica	PCR \geq 5.5 mg/L PCT \geq 3.3 ng/mL	63 68	36 82	54 50	96 91
Resch ⁸⁴ (N = 76)	Cultivo Clínica	PCR \geq 8 mg/L PCT \geq 2 ng/mL IL-6 \geq 10 pg/mL	49 83 71	100 61 65	100 76 76	58 70 60
Chiesa ⁴⁸ (N=134)	Cultivo Clínica	PCR \geq 4 mg/L PCT \geq 1 ng/mL IL-6 \geq 200 pg/mL	74 79 74	83 95 89	42 71 52	95 96 95

Tabla 3. Resumen de estudios que evalúan marcadores de sepsis neonatal nosocomial. (S: Sensibilidad, E: Especificidad, VPP: Valor predictivo positivo, VPN: Valor Predictivo Negativo. PCR: Proteína C-Reactiva, PCT: Procalcitonina, IL-6: Interleuquina-6). La sepsis fue considerada con cultivo positivo y/o clínica compatible. N hace referencia al número de pacientes incluidos en el estudio.

AUTOR	SEPSIS	PUNTO CORTE	S (%)	E (%)	VPP (%)	VPN (%)
Caldas ⁶⁴ (N= 82)	Cultivo	PCR \geq 5.1 mg/L	79	87	92	70
	Clínica	IL6 \geq 3.1 pg/mL	77	87	91	69
Khassawneh ⁸⁵ (N= 77)	Cultivo	PCR \geq 5 mg/L	96	89	79	98
	Clínica	IL6 \geq 20 pg/mL	83	50	41	87
Galetto-Lacour ⁸⁶ (N= 99)	Cultivo	PCR \geq 40 mg/L	79	79	61	90
		PCT \geq 0.5 ng/mL	93	74	60	96
		IL6 \geq 0.1 pg/mL	36	80	38	77
Blommendahl ⁴⁹ (N= 169)	Cultivo	PCR \geq 30 mg/L	58	84	24	94
		PCT \geq 1 ng/mL	77	62	16	97

Importancia de la cinética

Estas diferencias en cuanto a los estudios con mediadores de la cascada inflamatoria están en íntima relación con su cinética. El momento de la recogida de la muestra es fundamental para la correcta interpretación de los resultados^{58, 60} tal y como se muestra en la figura 9. En las fases más iniciales de la infección el uso de IL6 sola o en combinación con PCT puede ser más útil; sin embargo, pasadas 24 horas la IL6 pierde poder diagnóstico y la PCR empieza a ser más rentable.

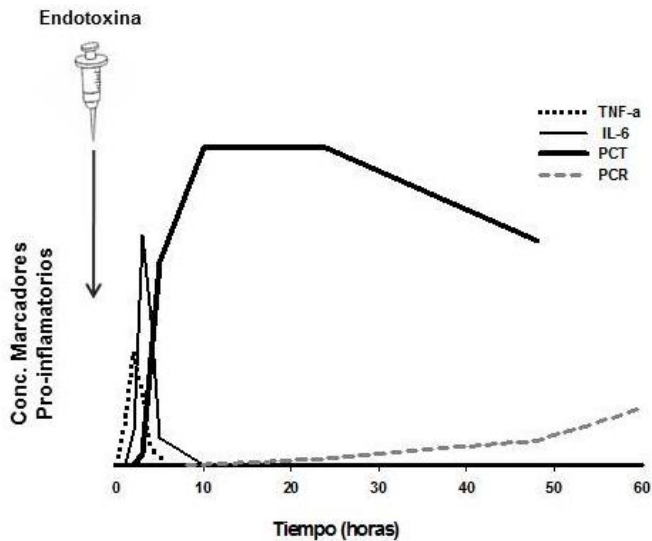


Figura 9. Concentraciones de los marcadores TNF- α , IL-6, PCT y PCR tras el estímulo de una endotoxina infecciosa y su evolución a lo largo del tiempo⁶⁰.

No obstante, a la hora de diseñar estudios clínicos, la dificultad se acentúa ya que la muestra se suele extraer cuando se objetiva clínica de sepsis y el umbral de sospecha presenta variabilidad entre los médicos.

Hibridación del genoma bacteriano

Los estudios de microbiología molecular están basados en la hibridación y amplificación del genoma bacteriano. La detección del patógeno se basa fundamentalmente en métodos de amplificación a partir de reacciones en cadena de la polimerasa que amplifican

INTRODUCCIÓN

regiones específicas del genoma microbiano, ya que los métodos de hibridación no se han estudiado en la población neonatal. Aunque no proporcionan información sobre susceptibilidad antibiótica. Parecen ser capaces de diagnosticar la sepsis hasta en 12 horas con sensibilidades iguales o superiores al hemocultivo tradicional según algunos autores⁸⁷⁻⁸⁸. Sin embargo, un reciente meta-análisis que incluye estos estudios entre otros, no encuentra una sensibilidad superior de estas técnicas como para reemplazar al hemocultivo tradicional como diagnóstico de elección en la sepsis neonatal⁸⁹.

I.3. PERFILES DE EXPRESIÓN GENÓMICA

Ante la falta de especificidad de un marcador por sí mismo, la combinación de varios marcadores se postula como la herramienta más eficaz para el diagnóstico precoz de sepsis neonatal⁹⁰. Una forma de valorar una respuesta global y visualizar todos los marcadores posibles a la vez son los estudios transcriptómicos a gran escala basados en el dogma central de la biología molecular. Según este dogma, descrito por Crick en 1970⁹¹, la información genética procedente del ácido desoxirribonucleico (ADN) se encuentra contenida en el ácido ribonucleico mensajero (ARNm). Para la

formación de una proteína, los genes del ADN son expresados y la información es transferida por medio de la transcripción a ARNm. Así, todo el ARNm transcrito bajo unas circunstancias específicas forma el transcriptoma. Estas moléculas serán más tarde usadas para sintetizar proteínas a partir de un proceso de traducción (figura 10).



Figura 10. Dogma central de la Biología Molecular modificado de la propuesta inicial de Crick. Adaptado de: <http://pendientedemigracion.ucm>)

Mediante mecanismos de corte y ajuste (*splicing*) alternativos y modificaciones post-traduccionales, cada gen puede producir diferentes proteínas. Además, se conoce de la existencia de genes que en vez de codificar proteínas, producen moléculas de ARN funcionales⁹², denominados genes no codificantes. De esta forma, se han identificado hasta la fecha más de 10.000 transcritos no codificantes, aunque sólo una pequeña parte de estos ARN tiene anotación funcional hasta la fecha. Sin embargo, avances recientes han evidenciado la asociación de estos ARN no codificantes con un

INTRODUCCIÓN

amplio rango de funciones celulares (regulación de la transcripción, regulación de modificaciones transcripcionales, activación y transporte de factores transcripcionales...). De este modo, actualmente, un análisis transcripcional no sólo se centra en el estudio del nivel de expresión de genes codificantes de proteínas, sino también en la distinción entre variantes de *splicing* y el estudio de la expresión de genes no codificantes.

I.3.1. BIOCHIPS DE EXPRESIÓN (*MICROARRAYS*)

Para analizar la expresión diferencial de todos los genes que componen el genoma, los estudios transcripcionales o de perfiles genómicos utilizan una herramienta llamada biochips que los monitorizan simultáneamente en un solo ensayo. Un biochip, o más conocido como *microarray* en inglés, es una superficie sólida sobre la que se encuentran insertadas y ordenadas un gran número de moléculas de material genético. Básicamente hay dos tipos de biochips de expresión: los biochips de ADN complementario (ADNc) y los biochips de oligonucleótidos.

Los **biochips de ADNc** están formados por una colección de secuencias de genes, normalmente productos de PCR (100-2000 pares

de bases), procedentes de secuencias de ADNc y de secuencias de expresión (*Expression Sequence Tags*), que se encuentran colocados individualmente sobre una matriz sólida. Con este tipo de biochips, se pueden comparar dos tipos de muestra en un mismo ensayo. Para ello, tras la preparación de ambas muestras, se lleva a cabo el marcaje con diferentes fluoróforos y la hibridación en un único array. Existen varios inconvenientes de estos chips. Por una parte, el estado del ADN de doble cadena en el biochip está mal definido, de forma que podría haber restricción de contactos con la matriz y podrían producirse enlaces cruzados que afectarían a la hibridación. Por otra parte, la fabricación del biochip depende de los conjuntos de clones, con lo cual incluso utilizando sets muy bien mantenidos, estos son propensos a la contaminación, condicionando los resultados del análisis.

Los **biochips de oligonucleótidos**, se fabrican mediante la síntesis *in situ* de sondas formadas por cadenas de oligonucleótidos sobre una matriz sólida, en base a la información de secuencias anotadas en bases de datos. A diferencia de los biochips de ADNc, sólo se puede hibridar una muestra en un array, por lo que las muestras se analizan en ensayos individuales (figura 11). Aunque la fabricación de estos biochips necesita el uso de un equipamiento especializado, no siendo

INTRODUCCIÓN

posible la fabricación por parte del investigador, presentan un gran número de ventajas respecto a los biochips de ADNc. En los biochips de oligonucleótidos las sondas están basadas enteramente en información de secuencias de bases de datos, por lo que no conlleva problemas a la hora de la preparación de set de clones, y por tanto tampoco problemas de contaminación. Además, el uso de secuencias cortas minimiza la hibridación cruzada, aumentando así la especificidad. Por otro lado, para cada transcrito se usan múltiples secuencias cortas (25 pares de bases) que representan partes significativamente únicas dentro del genoma, lo que permite distinguir variantes de *splicing* y genes estrechamente relacionados, así como aumentar la sensibilidad. Por tanto, aunque las secuencias de corta longitud podrían producir una disminución de la sensibilidad en comparación con los biochips de ADNc, el uso de múltiples sondas para cada transcrito permite conseguir finalmente una alta sensibilidad.

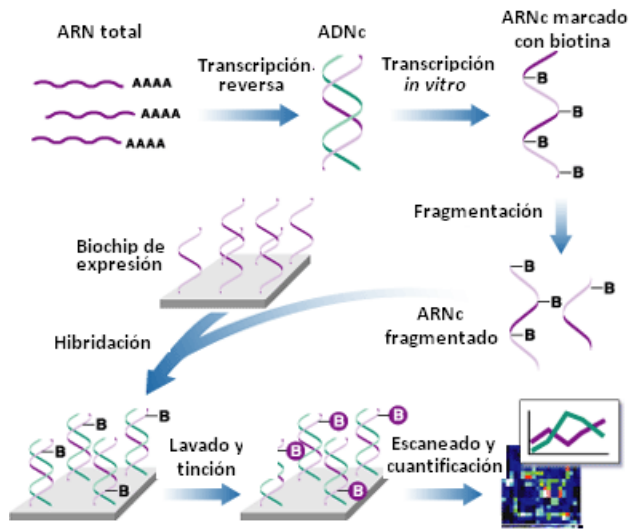


Figura 11. Biochips diseñados por la empresa Affymetrix. Imagen modificada de Sánchez & Ruíz de Villa, A Tutorial Review of Microarray Data Analysis, 2008.

I. 3.2. PROCESO DE HIBRIDACIÓN

Los ácidos nucleicos extraídos a partir de las muestras a analizar, se marcan por diversos métodos y se incuban sobre el panel de sondas, permitiendo la hibridación (reconocimiento y unión por complementariedad entre bases nitrogenadas) si ese gen está representado, o tiene una cantidad de número de copias de ARNm suficiente para ser detectado. Dicha hibridación entre la sonda y la muestra, se basa en el modelo de pares de bases de Watson and Crick y proporciona una medida cuantitativa de la abundancia de cada una

INTRODUCCIÓN

de las secuencias (figura 12). Cuando se produce la hibridación entre la sonda específica y la molécula diana en la muestra biológica se genera una señal fluorescente. El número de genes representado depende del biochip usado. A mayor número de copias, y por tanto mayor representación de un gen determinado, más intensa es esa señal de fluorescencia.

Tras la hibridación, se realizan una serie de lavados del chip para eliminar hibridaciones no específicas de ningún gen, así como para eliminar todo el ARN/ADNs no unido a ninguna sonda. Durante estos lavados también se lleva a cabo el paso de tinción mediante el cual se producirá la unión de las moléculas fluorescentes a las cadenas de ADN/ARN unidas en el chip (figura 13).



Figura 12. Complementariedad entre cadenas de nucleótidos (Affymetrix® 2005).

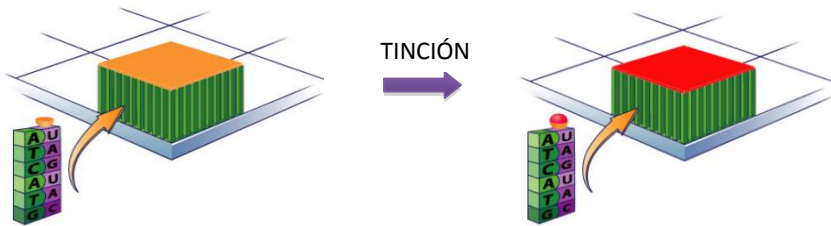


Figura 13. Proceso de tinción de las moléculas unidas al chip tras la hibridación (Affymetrix, 2005).

El último paso del procedimiento es el escaneo y lectura del chip. El escáner contiene un láser que incide sobre el array causando la fluorescencia de la molécula fluorescente añadida durante la tinción, permitiéndonos cuantificar el nivel de expresión de cada transcrito. Esta información es capturada digitalmente y sometida a varios análisis para extraer información biológica. La comparación de patrones de hibridación mediante herramientas informáticas permite la identificación de los ARNm que difieren en abundancia entre diferentes muestras. El resumen de todo el proceso de hibridación se encuentra representado en la figura 14. Tras la lectura del chip se obtienen los datos en crudo de la expresión de todos los genes que hay representados en el chip.

INTRODUCCIÓN

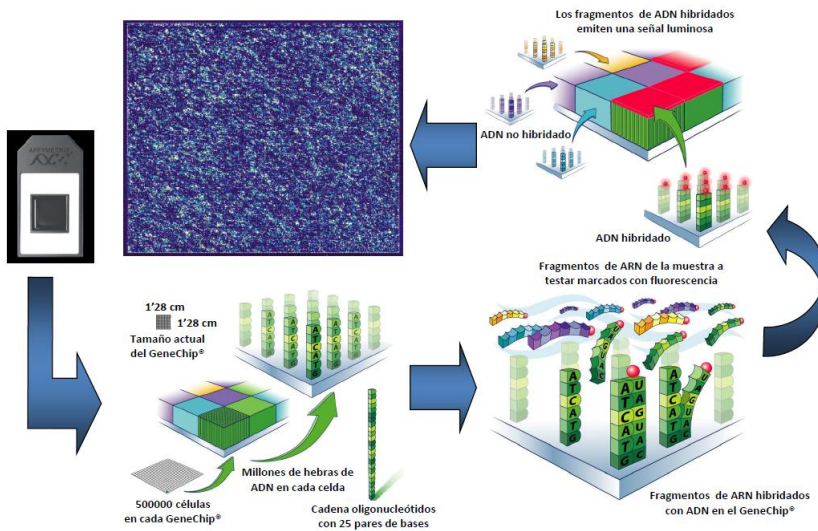


Figura 14. Resumen del proceso de hibridación con los biochips de oligonucleótidos de Affymetrix®

Posteriormente, se realiza un análisis estadístico a partir de múltiples comparaciones para obtener los genes con diferencias estadísticamente significativas en la intensidad de la expresión (información más ampliada en material y métodos) y se representan gráficamente. Las formas más habituales de representación son las agrupaciones jerárquicas no supervisadas (*hierarchical clustering*) en forma de mapas de calor (*heatmap*) o árboles de genes (*gene-tree*) que se obtienen a partir de los genes que presentan una expresión diferencial significativamente distinta en la población a estudio. Los genes se agrupan de manera no supervisada en función de su nivel de

expresión, y quedan representados por una gradación de color en función de si se encuentran sobre o infraexpresados en relación a los controles.

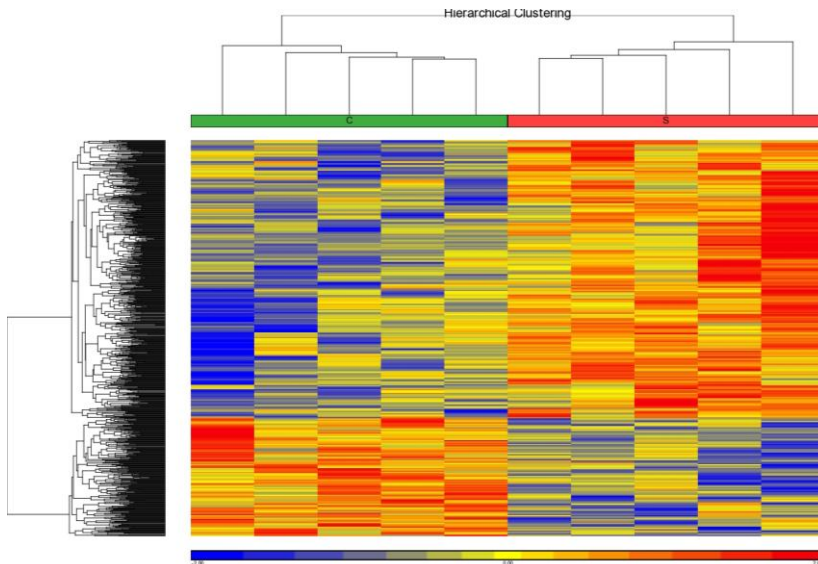


Figura 15. Agrupación jerárquica no supervisada de genes en función de su expresión. Cada columna representa un paciente y cada línea un gen. Los genes sobreexpresados respecto a los controles están representados en rojo y los infraexpresados en azul. La intensidad del color está directamente relacionada con la intensidad de la expresión.

I. 3.3. UTILIDAD EN EL DIAGNÓSTICO DE LA SEPSIS

El estudio global de la expresión genómica a partir de ARN total aislado de sangre periférica proporciona una excelente visión exhaustiva de la respuesta inmune del huésped frente a la infección, ya que la sangre es una fuente rica en células involucradas en la respuesta

INTRODUCCIÓN

inmune como leucocitos, monocitos o macrófagos. Sin embargo, no hay publicados estudios transcriptómicos en prematuros de muy bajo peso al nacimiento con sepsis.

Nuestro objetivo global es caracterizar los perfiles de expresión genómica de los recién nacidos con sepsis como marcadores diagnósticos precoces. Proponemos la hipótesis que estos patrones de expresión genómica nos permitirán diferenciar los neonatos con sepsis de los que no la presentan.

II. HIPÓTESIS Y OBJETIVOS

II. 1. HIPÓTESIS

Los recién nacidos menores de 1500 gramos de peso al nacimiento con sepsis documentada microbiológicamente tienen un patrón de expresión genómica significativamente distinto a la de neonatos menores de 1500 gramos de peso al nacimiento con características clínicas y demográficas similares, pero no infectados.

El desarrollo de esta técnica podría constituir una futura herramienta diagnóstica que nos permitiera discriminar a los neonatos con sepsis verdadera de aquellos con manifestaciones clínicas compatibles, pero que no presentan sepsis confirmada microbiológicamente.

II. 2. OBJETIVOS

II. 2.1. OBJETIVO PRINCIPAL

Determinar si existen perfiles de expresión genómica específicos que diferencien recién nacidos menores de 1500 gramos de peso al nacimiento con sepsis documentada microbiológicamente de controles no infectados con características clínicas y demográficas similares.

II. 2.2. OBJETIVOS SECUNDARIOS

- Selección de genes candidatos a marcadores de infección y validación por RT-PCR.

- Creación de un modelo predictivo de sepsis y validación del mismo con una cohorte externa.

- Clasificar los perfiles de expresión genómica de los neonatos diagnosticados de “sepsis clínica” según su semejanza con aquellos diagnosticados de sepsis con confirmación microbiológica o con los controles no infectados.

III. POBLACIÓN Y METODOLOGÍA

III. 1. DISEÑO Y POBLACIÓN

El estudio fue íntegramente realizado en el Servicio de Neonatología (HUiP La Fe). Se trata de un estudio prospectivo, observacional, de doble cohorte, realizado en pacientes prematuros que presentaron un peso al nacer menor de 1500 gramos (RNMBP) ingresados entre abril de 2011 y septiembre de 2012 en nuestro centro. El proyecto de investigación fue aprobado por el Comité Ético de Investigación Clínica del HUiP La Fe (anexo I). La dinámica de reclutamiento se explica en el diagrama de flujo de la figura 16. Todos los padres de los recién nacidos firmaron el consentimiento informado y cumplieron los criterios de selección que se indican a continuación.

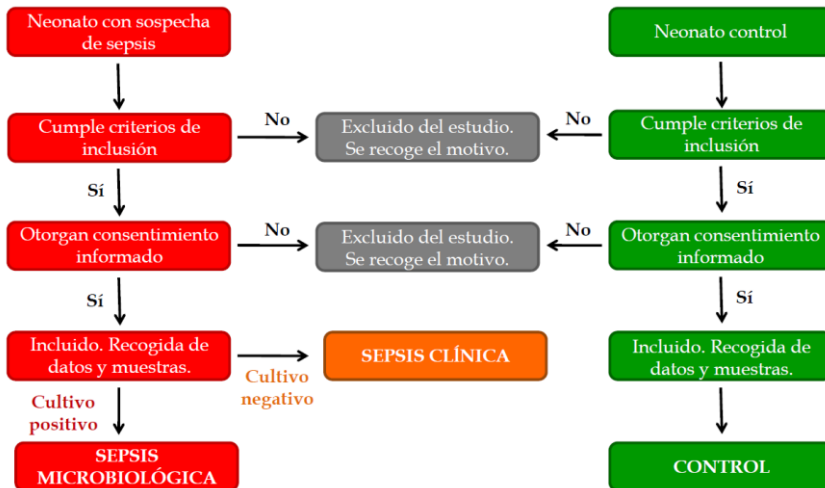


Figura 16. Dinámica de reclutamiento. Diagrama de flujo.

III. 1.1. CRITERIOS DE SELECCIÓN

Criterios de inclusión

- a) Peso al nacimiento menor de 1500 g.
- b) Edad postnatal menor de 30 días de vida tras alcanzar las 40 semanas de edad de gestación corregida, que se corresponde con el periodo neonatal si el paciente hubiera nacido a término.
- c) Sospecha de sepsis según los criterios que se definen a continuación. La cohorte de pacientes no infectados fue homogénea respecto a la infectada excepto por los criterios de infección que no estaban presentes.

Criterios de exclusión

- a) Cromosomopatías.
- b) Malformaciones congénitas mayores.
- c) Depresión neonatal severa con ausencia de signos vitales detectables (Apgar al minuto con puntuación de 0).
- d) Historia materna o hijos previos de inmunodeficiencia.
- e) Patología crónica en la madre sujeta a tratamientos que pudieran alterar la respuesta inmune tanto materna como fetal: asma, diabetes gestacional, patologías autoinmunes, etc.

- f) Infección congénita (STORCH: Sífilis, Toxoplasma, Rubeola, Citomegalovirus, Herpes, Parvovirus B19).
- g) Rechazo de los padres/tutores a participar.

III. 1.1. CRITERIOS DIAGNÓSTICOS DE SEPSIS

Se consideró sospecha de sepsis en aquellos pacientes que tuvieran factores de riesgo de sepsis^{6, 8, 93} y manifestaciones clínicas de sepsis⁹⁴.

La cohorte de RNMBP no infectados se reclutaron entre los neonatos ingresados sin clínica de sepsis emparejándolos según los siguientes parámetros: edad gestacional, peso nacimiento, sexo, raza, tipo de parto, administración previa de corticoides prenatales, días de vida a la recogida de la muestra y estado clínico (portador de vía central, soporte respiratorio).

A todos los pacientes con sospecha clínica y factores de riesgo de sepsis se les extrajeron 0'5 ml de sangre por punción venosa periférica tras desinfección de la piel con clorhexidina acuosa al 2% para hemocultivo (BacT/Alert® PF; Biomérieux®, Durham, NC; USA)⁹⁵. Se envió el hemocultivo, según el protocolo clínico de la unidad, para su análisis al servicio de Microbiología del Hospital Universitario y

POBLACIÓN Y METODOLOGÍA

Politécnico La Fe. Según el resultado del hemocultivo, se distinguieron 2 grupos: aquellos sin documentación microbiológica (Sepsis clínica) y los pacientes con documentación microbiológica (Sepsis microbiológica). Para considerar documentación microbiológica se requirió un hemocultivo positivo, excepto cuando la bacteria aislada fue *S. coagulasa negativo* que se requirieron dos hemocultivos positivos según los criterios actualmente aceptados⁴.

Factores de riesgo de sepsis^{6, 8, 93}:

Vertical

Sospecha de *corioamnionitis*: Definida por fiebre materna (temperatura axilar igual o mayor de 38°C) y 2 ó más de los siguientes criterios:

- Taquicardia materna (frecuencia cardiaca (FC) mayor de 100 latidos por minuto (lpm)).
 - Taquicardia fetal (FC mayor de 160 lpm).
 - Hipersensibilidad y/o dolor uterino.
 - Leucocitosis materna (más de 15000 leucocitos/ mmc).
 - Olor fétido del líquido amniótico.
- *Streptococcus grupo B (SGB) no testado en cultivos maternos rectales-vaginales, y factores de riesgo (Edad gestacional menor*

de 37 semanas, más de 18 horas de bolsa rota o fiebre materna entendida como temperatura axilar igual o superior a 38°C intraparto), si la madre no ha recibido profilaxis correcta (al menos una dosis de antibiótico 4 horas previas al parto).

- *SGB aislado en cultivos maternos vaginales-rectales, bacteriuria por SGB o hijo previo con SGB, si la madre no ha recibido profilaxis correcta* (al menos una dosis de antibiótico 4 horas previas al parto).
- *Rotura prematura de membranas* más de 18 horas.
- Presencia de gérmenes Gram negativos en frotis o cultivos vaginales maternos.

Nosocomial

Paciente que en el momento de la sospecha de sepsis, o en las 72 horas previas, cumpla alguno de los siguientes requisitos:

- Portador de *vía periférica o central*.
- Portador de *válvula de derivación ventrículo-peritoneal*.
- Portador de *tubo endotraqueal*.
- Portador de *sonda uretral*.
- Portador de *drenaje torácico*.
- *Intervención quirúrgica*.

Criterios clínicos⁹⁴

- *Temperatura*: T^a rectal $\geq 38^{\circ}\text{C}$ o $\leq 36^{\circ}\text{C}$.
- *Respiratorios*: polipnea (frecuencia respiratoria igual o superior a 70 respiraciones por minuto), tiraje, quejido, cianosis, apneas.
- *Cardiocirculatorios*: hipotensión (tensión arterial inferior al percentil 5 para su edad), taquicardia (FC mayor o igual a 180 lpm) o bradicardia (FC menor o igual a 100 lpm), hipoperfusión o piel reticulada.
- *Neurológicos*: convulsiones clínicas o eléctricas, hipotonía, letargia.
- *Gastrointestinales*: rechazo de tomas, mala tolerancia (restos gástricos superiores al 50% de la ingesta), vómitos o distensión abdominal marcada.

III. 1.3. VARIABLES A ESTUDIO. RECOGIDA DE DATOS

En todos los neonatos se recogieron datos demográficos y clínicos en el momento del reclutamiento (Tabla 4).

Tabla 4. Datos clínicos y demográficos.

DATOS DEMOGRÁFICOS	DATOS CLÍNICOS
<ul style="list-style-type: none"> - Edad gestacional (semanas) - Peso nacimiento (gramos) <ul style="list-style-type: none"> - Sexo - Raza - Apgar₁ y Apgar₅ - Edad (días) a la inclusión - Peso (gramos) a la inclusión <ul style="list-style-type: none"> - Corticoides prenatales 	<ul style="list-style-type: none"> - Factores de riesgo - Patología de base - Tratamiento recibido - Soporte respiratorio - Soporte hemodinámico

III. 2. ANÁLISIS DE LAS MUESTRAS

III. 2.1. EXTRACCIÓN DE ARN TOTAL DE SANGRE PERIFÉRICA

Se extrajo una muestra de sangre de 0'5 ml por venopunción en el momento de la sospecha de sepsis, previo al inicio de antibióticos en la cohorte de casos. En la cohorte de controles, se obtuvo la muestra cuando se reunieron las características para ser emparejados con un caso en cuanto a datos demográficos y clínicos, aprovechando una extracción de sangre por otro motivo. Esta muestra se introdujo en una solución estabilizadora (TempusTM Blood RNA tubes, Applied Biosystems[®], Foster City, CA, EE.UU.) para que no se degradase el

ARN total, se agitó la muestra hasta mezclarla homogéneamente y se guardó a menos 80°C hasta la posterior extracción conjunta del ARN total.

Optimización del protocolo extracción del ARN total

El fabricante de los tubos Tempus[®] que contienen 6 ml de solución estabilizadora de ARN recomienda introducir 3 ml de sangre total. Extraer este volumen es inviable en un RNMBP, teniendo en cuenta que su volemia es de 80-90 ml/Kg, lo que supone un volumen total de sangre que en muchas ocasiones no llega a los 100 ml⁹⁶. Por este motivo, realizamos una optimización del método para poder extraer ARN con una cantidad de sangre total de 0.5 ml y 1 ml de solución estabilizadora del tubo Tempus[®], manteniendo la proporción sangre:solución estabilizadora 1:2. Para ello alicuotamos, en condiciones de esterilidad bajo campana de flujo laminar, cada tubo Tempus[®] de 6 ml en 6 tubos Eppendorf Blenos[®] de 2 ml de capacidad con 1 ml de solución estabilizadora.

Se empleó el protocolo de extracción del ARN total de cada muestra usando el kit de aislamiento MagMAX[™] (Ambion/Applied Biosystems[®], Foster City, CA, EE.UU.), que se encuentra

96

esquemático en la figura 17. Utilizamos la centrífuga Thermo Scientific Heraeus® Multifuge 3SR Plus®. Dado que el volumen de partida eran 1.5 ml (0.5 ml sangre y 1 ml solución estabilizadora ARN) en lugar de los 9 ml que marca la casa comercial, optimizamos el protocolo del kit de aislamiento MagMAX™ para extracción de ARN con el fin de obtener más concentración y calidad de la muestra de ARN total (tabla 5).

El protocolo original contemplaba añadir 3 ml de solución de fosfato salino tamponada (PBS) a los 9 ml de sangre y solución estabilizadora, centrifugar en un primer paso durante 15 minutos a 4 °C y 5000 g y añadir 4 ml de solución de lavado *Pre-Digestion Wash* para centrifugar de nuevo en un segundo paso durante 10 minutos a la misma velocidad y temperatura. Se probaron distintas cantidades de PBS y de solución de lavado *Pre-Digestion Wash*. Se redujo la velocidad y se aumentó el tiempo de centrifugado para obtener una mayor concentración de ARN, dada la escasa cantidad de sangre utilizada. Para ello utilizamos 7 muestras de sangre de cordón de recién nacidos sanos (edad gestacional media 39 ± 2 semanas, con un peso al nacimiento de 3298 ± 192 gramos y Apgar al minuto 10 (9-10) y a los 5 minutos de 10 (10-10).

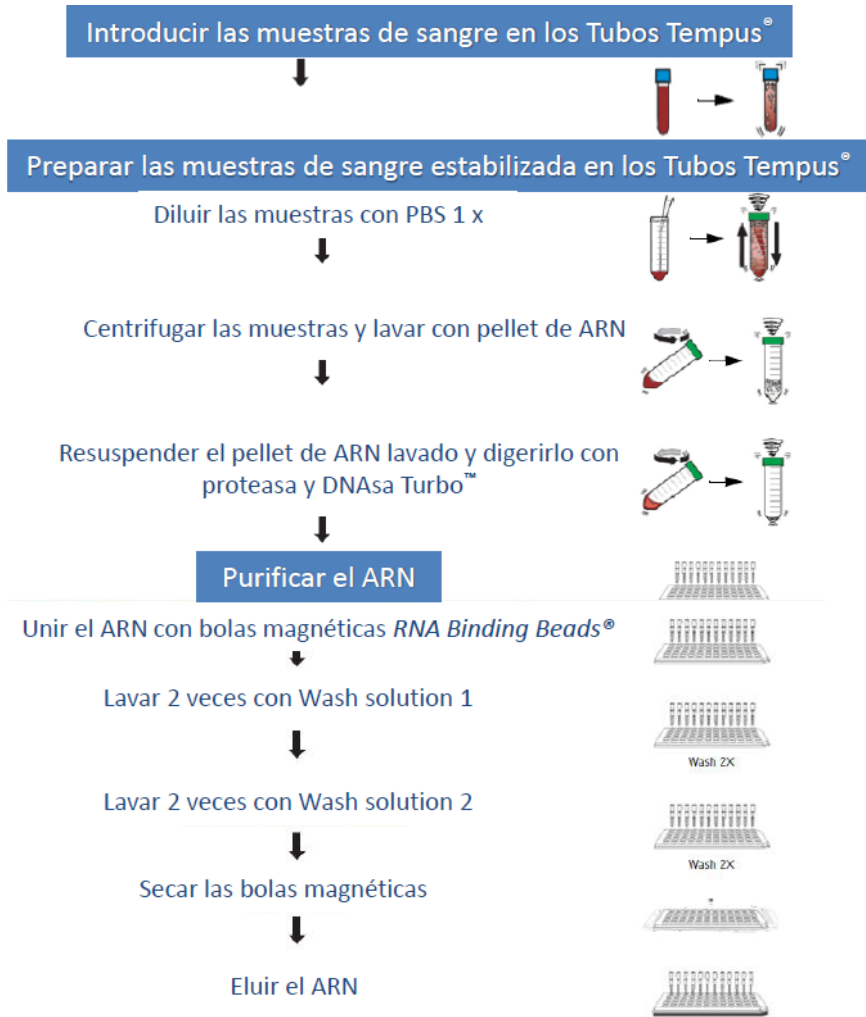


Figura 17. Protocolo de extracción del ARN con el kit de aislamiento MagMAX™ (Ambion/Applied Biosystems®, Foster City, CA, EE.UU.).

Tal y como se describe en el punto III.2.2., se utilizó un sistema de electroforesis capilar, el Bioanalizador 2100 (Agilent, Palo Alto, CA, USA) para comprobar la integridad del ARN total de cada muestra

(RIN: *RNA Integrity Number*), y se midió la concentración de ARN total por espectrofotometría. Los resultados obtenidos están descritos en la tabla 5.

Tabla 5. Valores de concentración de ARN total expresados en ng/μl y de integridad expresados por el valor de integridad RIN (*RNA Integrity Number*) en función de las distintas concentraciones de sangre y solución estabilizadora de ARN contenida en los Tubos Tempus[®]. PD Wash: *Solution Pre-Digestion Wash*

Nº	PBS (ml)	PD WASH (ml)	CENTRIFUGADO	ARN (ng/μl)	RIN
1	0,5	1	1º: 5000 g 4°C 20 min 2º: 5000 g 4°C 10 min	164	5,1
2	0,5	1	1º: 5000 g 4°C 20 min 2º: 5000 g 4°C 10 min	365	3,1
3	1	1	1º: 5000 g 4°C 20 min 2º: 5000 g 4°C 10 min	252	3,6
4	2	2	1º: 5000 g 4°C 20 min 2º: 5000 g 4°C 10 min	99	2,3
5	0,5	4	1º: 3000 g 4°C 40 min 2º: 3000 g 4°C 30 min	273	6,5
6	0,5	4	1º: 3000 g 4°C 60 min 2º: 3000 g 4°C 40 min	563	5,7
7	0,5	4	1º: 3000 g 4°C 60 min 2º: 3000 g 4°C 40 min	450	7,8

POBLACIÓN Y METODOLOGÍA

Tras los resultados obtenidos decidimos utilizar las cantidades, velocidades y tiempos de las muestras 6 y 7, quedando el protocolo final tal y como se describe a continuación de manera detallada.

PROTOCOLO FINAL

1º. Diluir las muestras con solución de fosfato salino tamponada (PBS: Phosphate Buffered Saline)

- Estabilizar en hielo 10-30 minutos.
- Introducir en tubo cónico y añadir 0,5 ml PBS. Mezclar en agitador de tubos 30 s a alta velocidad.

2º. Centrifugar las muestras

- Centrifugar a 3000 g y 4°C durante 60 min. Verter el sobrenadante dejando al final el supuesto precipitado que es transparente).
- Añadir al precipitado 4 ml de Solución *Tempus Pre-digestion Wash*[®] y mezclar en agitador de tubos 5 s a velocidad moderada.
- Centrifugar a 3000 g y 4°C durante 40 min. Verter el sobrenadante y dejar los tubos sobre papel absorbente 2 minutos para eliminar el líquido residual.
- Introducir los tubos en hielo.

3º. Re-suspensión del precipitado ARN

- Preparar re-suspensión: 117,5 μL de *Resuspension Solution*[®] y 2,5 μL de Proteinasa.
- Introducir en un tubo Eppendorf[®] de 1.5 ml y mezclar en agitador de tubos 2-3 s a velocidad moderada.
- Centrifugar 2-3 s para recoger mezcla de resuspensión en la parte inferior del tubo y guardar la mezcla en hielo hasta su uso.
- Añadir a cada tubo (con el precipitado) 120 μL de la mezcla de resuspensión.
- Mezclar en agitador de tubos a velocidad ligera 5-10 s.
- Añadir a cada Eppendorf 10 μL de DNAsa y mezclar con agitador de tubos al 7 durante 10 min.

4°. Purificación del ARN

- Agitar vigorosamente las bolas magnéticas (*RNA Binding Beads*[®]) para resuspenderlas.
- Mezclar con 50 μL *Binding solution*[®] y 20 μL de *RNA Binding Beads*[®].
- Añadir 70 μL de la mezcla a cada muestra y agitar con agitador de tubos al 5 un minuto.
- Añadir a cada muestra 200 μL de Isopropanol 100% y mezclar con agitador de tubos al 5 durante 3 min.
- Dejar los tubos en la placa magnética durante 5 min.

POBLACIÓN Y METODOLOGÍA

Lavar 2 veces con Wash Solution[®] 1

- Añadir a cada muestra 150 μ L de *Wash Solution[®] 1*
- Mezclar con agitador de tubos al 5 durante 1 min.
- Dejar los tubos en la placa magnética durante 5 min.
- Sacar los tubos de la placa y eliminar el sobrenadante.

Lavar 2 veces con Wash Solution[®] 2

- Añadir a cada muestra 150 μ L de *Wash Solution[®] 2*
- Mezclar con agitador de tubos al 6 durante 1 min.
- Dejar los tubos en la placa magnética durante 5 min.
- Sacar los tubos de la placa y eliminar el sobrenadante.

Secar

- Sacar los tubos de la placa y eliminar el sobrenadante. Invertir los tubos sobre papel absorbente 2 minutos a temperatura ambiente para eliminar cualquier resto de líquido.

5°. Eluir el RNA

- Añadir en cada muestra al precipitado que ha quedado 20 μ l de Buffer de lisis.
- Mezclar con agitador de tubos al 7 durante 4 min.
- Dejar los tubos en la placa magnética durante 5 min.
- Coger el sobrenadante sin coger las bolas magnéticas.

III. 2.2. CONTROL DE CALIDAD, PUREZA E INTEGRIDAD DEL ARN TOTAL

La integridad de la muestra se determinó usando un sistema de electroforesis capilar (Bioanalizador 2100 Agilent, Palo Alto, CA, USA) utilizando el kit RNA 6000 Nano LabChip®. Cada chip contiene microcanales interconectados que son usados para la separación electroforética de ácidos nucleídos en función del tamaño. Se midió la concentración de ARN total por espectrofotometría.

El electroferograma que deriva del sistema de electroforesis capilar proporciona un valor de integridad (RIN: *RNA Integrity Number*). Se hibridaron las muestras con una concentración de ARN mayor de 200 ng/μl y que alcanzaron un RIN superior a 6.

III. 2.2.1. Integridad de la muestra de ARN total

La determinación de la integridad de la muestra es fundamental para asegurar un buen análisis posterior, es esencial que el ARN total de las muestras tenga una buena calidad e integridad, ya que si el ARN fuera de baja calidad, la transcripción reversa daría lugar a ADNc carente de algunas regiones codificantes.

POBLACIÓN Y METODOLOGÍA

Los reactivos necesarios para la preparación de los LabChip[®] se dejaron 30 min a temperatura ambiente antes de su utilización. Se preparó el gel utilizando una columna de 550 μ l, que se centrifugó a 1500 g durante 10 min. El gel filtrado se dividió en alícuotas de 65 μ l. Tras la adicción de 1 μ l de colorante al gel, la mezcla gel-colorante se inyectó dentro del sistema de canales. Se añadieron 5 μ l del marcador endógeno (ADN fragmentado en un tamaño de 50 pares de bases (pb) y usado para alinear todas las muestras) tanto en el pocillo del marcador de pesos moleculares como en los pocillos de las muestras. Se introdujo en cada uno de los pocillos de muestra 1 μ l de ARN total de las correspondientes muestras a analizar y 1 μ l de marcador de pesos moleculares (mezcla de ARN de diferente tamaños a distinta concentración) en su correspondiente pocillo.

Las muestras de ARN y el marcador de pesos moleculares fueron anteriormente preparados mediante una desnaturalización a 70°C durante 2 minutos. Finalmente, los chips se agitaron durante 1 minuto a máxima potencia y se introdujeron en el Agilent 2100 Bioanalyzer. Para llevar a cabo la visualización de los resultados, se utilizó el software 2100 Expert v. B.01.02.sI136.

Para la determinación de la calidad e integridad del ARN total, el bioanalizador 2100 de Agilent usa fluorescencia entre 670 nm y 700 nm. Para cada muestra, el software calcula el valor de integridad *RIN* mediante un algoritmo que evalúa diferentes parámetros de integridad en las medidas electroforéticas de RNA (Agilent Technologies, 2005), de manera que la integridad de la muestra no solo es determinada por el ratio de las bandas ribosomales 18S y 28S, sino por todas las características del electroferograma, incluyendo la presencia o ausencia de productos de degradación. Este valor, que tomará valores entre 1 (gran degradación) y 10 (muestra intacta), nos permite también comparar directamente muestras de ARN y garantizar la reproducibilidad de los experimentos⁹⁷.

La figura 18 muestra algunos de los electroferogramas obtenidos a partir de las muestras utilizadas en la optimización del método de extracción de ARN.

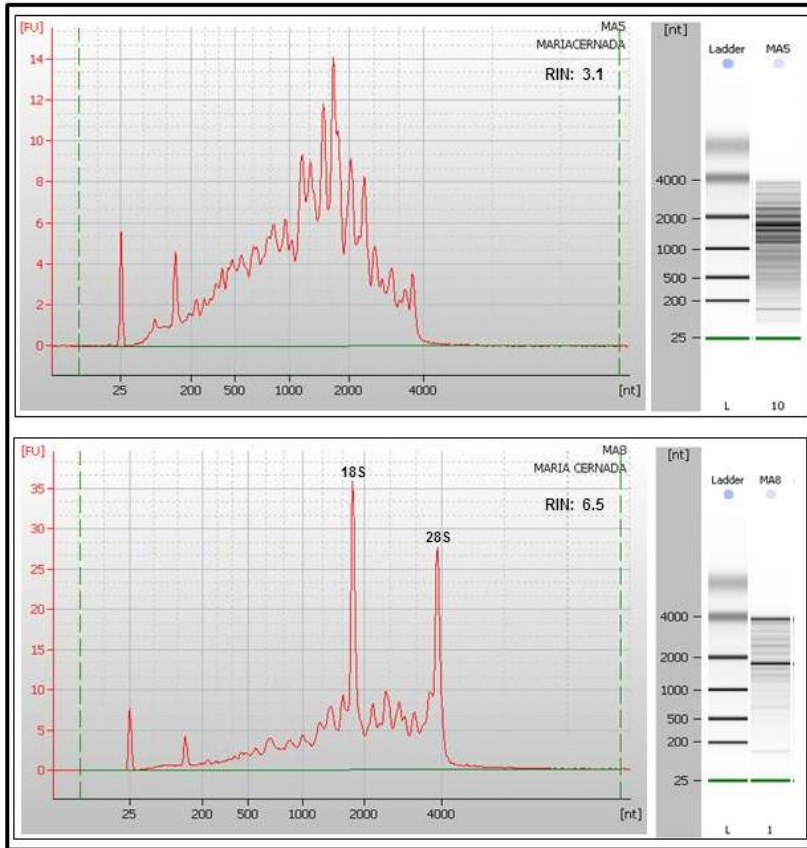


Figura 18. Electroferogramas y geles de electroforesis de ARN obtenidos de las muestras de sangre utilizadas para la optimización del método de extracción de ARN. A la izquierda se encuentran los electroferogramas, en el eje de abscisas se representan las unidades de fluorescencia (FU) y en el de ordenadas los nucleótidos (nt). A la derecha se encuentran las bandas de los geles de electroforesis. La muestra superior (MA5) muestra un perfil degradado en el que no se identifican los picos de las bandas ribosomales 18S y 28S, obteniéndose un RIN de 3,1; mientras que en la muestra inferior (MA8) el perfil muestra una buena integridad, se identifican los picos de las bandas ribosomales 18S y 28S y el RIN es de 6,5.

III. 2.2.2. Concentración de ARN total

Tras comprobar el estado de integridad del ARN total extraído de las muestras, y antes de proceder a realizar el protocolo de preparación de las muestras para el análisis en los biochips, se realizó una medida de la concentración de ARN total, y se estimó la pureza del mismo, la cual puede afectar a la efectividad de los procesos posteriores. Para ello se utilizó una técnica espectrofotométrica. Los ácidos nucleídos absorben eficientemente luz ultravioleta debido a la presencia en sus cadenas de bases aromáticas nitrogenadas. La absorción de radiación ultravioleta del ADN y del ARN es una característica específica de estas moléculas, y por tanto es usada eficientemente para determinar su concentración. Así, para calcular la cantidad de ARN en una muestra se utiliza la siguiente conversión: 1 unidad de absorbancia a 260 nm de longitud de onda (A_{260}) de ARN de cadena simple es igual a 40 $\mu\text{g/mL}$ (1 unidad A_{260} de ssRNA = 40 $\mu\text{g/mL}$). Es muy importante tener en cuenta la dilución del ARN realizada, ya que los resultados se tienen que ajustar a los límites de medida del espectrofotómetro. Una vez inyectamos la dilución de ARN en el espectrofotómetro, éste automáticamente realiza los cálculos comentados y nos proporciona directamente la concentración de ARN presente en cada muestra.

Por otro lado, para comprobar la pureza, se mide espectrofotométricamente la absorbancia a 260 nm (A260) y a 280 nm (A280), ya que mientras los ácidos nucleídos tienen su máximo de absorción a 260 nm, las proteínas lo tienen a 280 nm, tal y como muestra la figura 19. De esta manera, observando la relación entre ambos valores de absorbancia (A260/A280) podremos saber si nuestro ARN total extraído tiene una alta pureza o, si por lo contrario, está contaminado con restos de proteínas. Un ratio A260/A280 cercano a 2,0 indica una preparación de gran pureza.

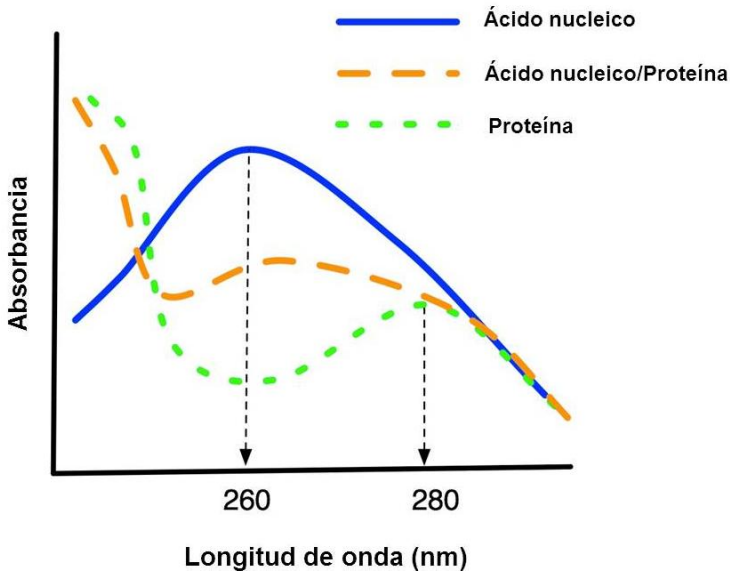


Figura 19. Gráfica de la absorbancia medida a longitudes de ondas de 260 y 280 de los ácidos nucleicos y de las proteínas.

Para llevar a cabo tanto la cuantificación como la determinación del grado de pureza de las muestras se utilizó el miniespectrofotómetro GeneQuant Pro™ (GE Healthcare).

Se realizó una disolución 1:3 del ARN total (1 µl de la muestra obtenida y 2 µl de agua libre de RNAsas) para cada muestra. Las diluciones obtenidas se introdujeron en capilares de cuarzo para realizar la medición en el espectrofotómetro. La medida de referencia o blanco se realizó con 3 µl de la misma agua tratada con agua libre de RNAsas utilizada para la disolución.

III. 2.3. ESTUDIO DE LA EXPRESIÓN GENÓMICA A ESCALA GLOBAL

La hibridación de las muestras se realizó usando un biochip de oligonucleótidos por sus ventajas respecto a los biochips de ADNc descritas en la introducción. Un gran número de empresas disponen de biochips con este formato, aunque la empresa líder en este campo es Affymetrix®, por lo que utilizamos su biochip *GeneChip*® Human Gene 1.0 ST Array (Affymetrix®, Santa Clara, CA, USA), que integra 28.869 genes mediante 764.885 sondas distintas de 25 pares de bases

de longitud, con una mediana de 26 sondas por gen representando unos 47.000 genes bien caracterizados (figura 20).



Figura 20. GeneChip® Human Gene 1.0 ST Array (Affymetrix®)

Utiliza secuencias de transcritos anotados procedentes de las bases de datos Reference Sequence (RefSeq) (Noviembre 3, 2006) y Ensembl (Octubre 2006, versión 41.36c), así como secuencias de transcritos anotados procedentes de la base de datos GeneBank, siempre que tengan la región codificante (CDS) completa (Affymetrix®, 2007). Por otra parte, estos biochips son fabricados de manera que puedan aportar información del rendimiento del procesado de la muestra, de la hibridación y del escaneado de la imagen. Con este propósito, cada uno de ellos contiene un conjunto de sondas que hacen el papel de

controles de calidad de la muestra. El valor de la intensidad que se obtiene de este conjunto de sondas es utilizado para generar los estadísticos del control de calidad del biochip.

III. 2.3.1. *GeneGhip*[®] *Human Gene 1.0 ST array*

El protocolo se lleva a cabo mediante una serie de procesos que incluyen la síntesis del ADNc de cadena simple con sentido a partir del ARN total extraído de las muestras, su fragmentación y marcaje, y finalmente la hibridación y lectura del biochip. Durante todo el proceso se necesitan una serie de instrumentación, software y reactivos específicos que se muestran en la tabla 6.

El esquema de trabajo a seguir para llevar a cabo el protocolo completo se dividió en tres días (Tabla 7), tras los cuales se obtuvo una imagen (archivo .DAT) que procesamos para obtener los datos de expresión.

Tabla 6. Productos requeridos para la realización del protocolo *GeneChip® Human Gene 1.0 ST*

INSTRUMENTACIÓN	
Escáner	<i>GeneChip® Scanner 3000 7G</i>
Estación de fluidos	<i>GeneChip® Fluidics Station 450</i>
Horno de hibridación	<i>GeneChip® Hybridization Oven640</i>
SOFTWARE	
Adquisición y análisis de imágenes	<i>Software Affymetrix GeneChip® Command Console® (AGCC)</i>
REACTIVOS	
Preparación de las muestras	<i>Kit Ambion® WT Expression</i>
Fragmentación y marcaje final	<i>Kit GeneChip® WT Terminal Labeling</i>
Control del marcaje	<i>Kit GeneChip® Poly-A RNA Control</i>
Control de la hibridación	<i>Kit GeneChip® Hybridization Control</i>
Hibridación, lavado y tinción	<i>Kit GeneChip® Hybridization, Wash, and Stain</i>

Tabla 7. Esquema de trabajo para la realización del protocolo *GeneChip® Human Gene 1.0 ST*.

DÍA 1
<ul style="list-style-type: none"> • Síntesis de la primera cadena de ADNc • Síntesis de la segunda cadena de ADNc • Síntesis del ARNc mediante transcripción in vitro
DÍA 2
<ul style="list-style-type: none"> • Purificación del ARNc • Evaluación de la concentración del ARNc • Segundo ciclo de síntesis de ADNc • Hidrólisis con ARN asa H • Purificación del ADNc de simple cadena • Evaluación de la concentración del ADNc • Fragmentación y marcaje del ADNc de simple cadena • Hibridación de las muestras en los biochip (comienzo)
DÍA 3
<ul style="list-style-type: none"> • Hibridación de la muestra en el biochip (finalización) • Lavados y tinción • Escaneo de los biochips

III. 2.3.2. Pre-procesado de los datos de intensidad brutos

Tras el escaneo de los biochips, obtuvimos una imagen del biochip, (almacenada en un archivo .DAT). Posteriormente, mediante el

software de Affymetrix *GeneChip*[®] *Command Console*, se llevó a cabo un procesado de dicha imagen, dando lugar a un archivo .CEL que contiene, para cada celda, un valor de intensidad de fluorescencia. Este fichero es el resultado de la segmentación de la imagen obtenida y contiene la media y desviación estándar de los niveles de gris así como la localización de la sonda dentro del *array*. Sin embargo, en los chips Affymetrix *GeneChip*[®] cada valor de intensidad no corresponde a la expresión de un gen, ya que hay múltiples valores (sondas) por cada gen. Así, estos valores deben ser combinados para cada set de sondas (*probeset*), generando un valor de expresión por cada gen presente en el chip (*Summarization*). Para llevar a cabo esta *Summarization* se necesita no solo la información que aporta el archivo .CEL sino también conocer la correspondencia entre sondas y nombres de los genes. Esta información aparece en el fichero CDF.

Por tanto, la información que se maneja se puede considerar a tres niveles. En primer lugar, tras el escaneo y lectura de los datos obtenemos la imagen digital bruta (.DAT) (figura 21). Sobre esta imagen se aplican técnicas de procesado de imagen digital que directamente permiten obtener a partir de esta imagen bruta otra imagen normalizada donde cada pixel contiene una cuantificación media de la cantidad de hibridación para cada sonda, datos a nivel de

sonda (.CEL) (figura 22). Finalmente, tras el pre-procesado de los datos a nivel de sonda obtenemos los datos de expresión para cada gen contenido en el biochip. A partir de aquí se trabaja para analizar la posible actividad diferenciada de un gen o un conjunto de genes.

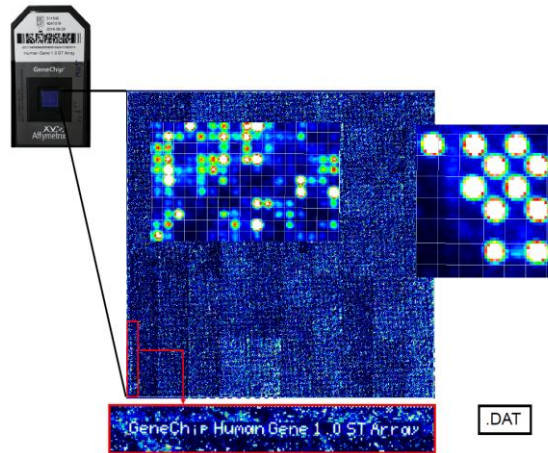


Figura 21. Imagen digital obtenida del biochip (.DAT) tras el escaneado.

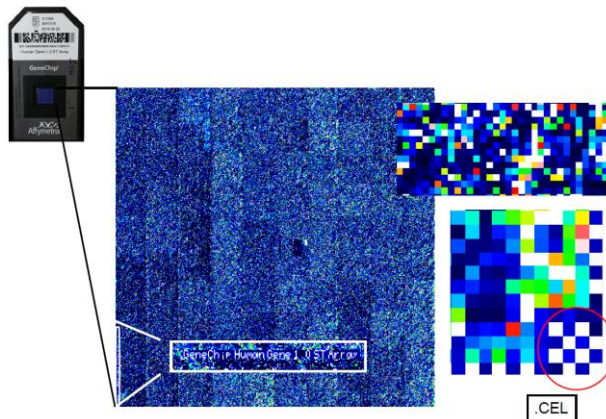


Figura 22. Imagen digital obtenida del biochip (.CEL).

III. 2.4. VALIDACIÓN MEDIANTE PCR-RT EN TIEMPO REAL

Los genes que de manera estadísticamente significativa fueron capaces de discriminar entre RNMBP con sepsis y RNMBP controles, se validaron usando la reacción en cadena de la polimerasa (PCR) a tiempo real combinada con una reacción de retro-transcripción o RT (RT-PCR) (RT-PCR: *real-time reverse-transcription polymerase chain reaction*) usando sondas de hidrólisis TaqMan[®] Gene Expression Assays probes (Applied Biosystems[®], Foster City, CA, EE.UU.).

La reacción en cadena de la polimerasa en tiempo real (PCR-RT) es actualmente el método más sensible y preciso para determinar la cantidad de ARNm de una muestra mediante una cuantificación relativa, ya que compara entre diferentes muestras la cantidad relativa o relación del ARNm de un gen específico respecto a la cantidad de ARNm de un gen constitutivo (control endógeno).

En la cuantificación de la expresión del gen de estudio, es necesario un control interno para corregir la posible degradación de ARN en el proceso o el posible problema de sensibilidad de la síntesis de ADNc.

Esto se denomina **normalización** de la expresión del gen específico, o normalizar respecto a la diferente concentración de ARN total de las muestras, ya que si la cantidad de control endógeno varía es debido a cambios en la cantidad de ARN total. Frecuentemente a este control interno se le llama gen de mantenimiento *housekeeping gene*, llamado de esta forma porque su síntesis ocurre en todos los tipos celulares nucleados, y son necesarios para la supervivencia celular. Estos genes son expresados constitutivamente por todo tipo de células, y no deberían estar alterados en condiciones fisiopatológicas. La principal condición que debe cumplir este gen es mantener constante su expresión tanto en las condiciones control como en los casos enfermos, para así poder evaluar los cambios de expresión de nuestro gen de estudio respecto a la expresión del *housekeeping gene*. Habitualmente se utilizan: β -actina, gliceraldehido-3-fosfato deshidrogenasa (GAPDH), hipoxantina, 18S, 28S, etc. En nuestro caso, utilizamos como control endógeno rRNA 18S⁹⁸.

Este método, permite la detección directa del producto de amplificación durante la fase exponencial de la reacción, a través del empleo de compuestos con propiedades fluorescentes, de modo que

POBLACIÓN Y METODOLOGÍA

determinando el incremento de fluorescencia se puede determinar la cantidad de producto formado.

Una vez sintetizado el ADNc, se amplificaron de forma selectiva secuencias específicas de ADN por medio de la reacción en cadena de la polimerasa (PCR). La PCR utiliza dos fragmentos cortos de ADN (oligonucleótidos) como cebadores (*primers*) de la síntesis, que se unen específicamente a secuencias que flanquean la región a amplificar, uno en cada una de las cadenas de ADN, y permiten la amplificación selectiva. El proceso básico se desarrolla en tres pasos que se repiten sucesivamente:

- Desnaturalización: separación de las cadenas complementarias del ADN
- Unión o *annealing* de los *primers* específicos a sus secuencias complementarias
- Extensión: síntesis de la hebra complementaria a partir del *primer* respectivo

La repetición de este ciclo un determinado número de veces produce un aumento exponencial de la cantidad de la región diana de ADN,

que viene dado por la expresión 2^n (siendo n el número de ciclos) hasta que se agotan los reactivos y la reacción deja de ser exponencial.

El principio básico de la PCR-RT es la detección de los productos de la PCR mediante la generación de señales fluorescentes. La señal de fluorescencia en este trabajo está generada por unos fluoróforos unidos a sondas que hibridan en una región interna específica de los productos de amplificación, mediante el sistema Taqman[®] (Applied Biosystems). El sistema Taqman[®] desarrollado por Perkin-Elmer, permite monitorizar la reacción ciclo a ciclo, de modo que podemos escoger la región de la curva donde el incremento es lineal. Este sistema incorpora a la reacción de PCR un oligonucleótido denominado sonda Taqman[®] que hibrida en una región interna del producto de la PCR (secuencia diana), entre el cebador directo y el cebador indirecto. Esta sonda incorpora en el extremo 5' un fluoróforo donante que emite fluorescencia al ser excitado y en el extremo 3' una molécula apagadora no fluorescente que absorbe la fluorescencia liberada por el donante. Cuando la sonda está intacta, la proximidad entre ambas moléculas resulta en la supresión de la fluorescencia del donante, debido al fenómeno de transferencia de energía de resonancia de Förster. Este fenómeno se basa en que la proximidad espacial de

POBLACIÓN Y METODOLOGÍA

ambas moléculas y en que el espectro de absorción de la molécula donante se solapa con el espectro de absorción de la molécula receptora. Al carecer la molécula apagadora de señal fluorescente, la medida de la señal de fluorescencia aportada por el fluoróforo donante cuando se libera es muy precisa.

Mientras la sonda está intacta, la fluorescencia emitida por el donante es absorbida por el aceptor. Sin embargo, durante la amplificación del ADN diana, el desplazamiento a lo largo de la cadena de la ADN polimerasa, que tiene actividad 5' exonucleasa, hidroliza el extremo libre 5' de la sonda que ha hibridado con su cadena complementaria, produciéndose la liberación del fluoróforo donante. Como donante y aceptor están ahora alejados, la fluorescencia emitida por el primero ya no es absorbida por el aceptor y es captada por el lector. Este incremento de fluorescencia ocurre solamente si la secuencia diana es complementaria a la sonda y esta se amplifica durante la PCR, por lo que las amplificaciones no específicas no se detectan. La polimerización de la hebra continúa, pero no se produce amplificación de la sonda durante la PCR puesto que su extremo 3' se encuentra bloqueado. Este proceso ocurre en cada ciclo y no interfiere con la acumulación exponencial del producto de amplificación.

Condiciones experimentales

Este termociclador incorpora un lector de fluorescencia para llevar a cabo la PCR-RT y está diseñado para poder medir, en cualquier momento, la fluorescencia emitida en cada uno de los viales donde se realiza la amplificación. El sistema de detección por fluorescencia empleado en la RT-PCR en tiempo real utiliza cebadores y sondas específicas de longitud entre 50-150 pb marcadas con fluoróforos previamente diseñados, validados y vendidos comercialmente como *Inventoried TaqMan[®] Gene Expression Assays* (Applied Biosystems). El fluoróforo FAM se utilizó tanto para los genes de estudio como para el control endógeno. El programa de temperaturas utilizado en el termociclador fue el siguiente: 50°C durante 2 minutos, 95°C durante 10 minutos, y 40 ciclos de 95°C durante 15 segundos seguidos de 60°C durante 1 minuto. La incubación a 50°C fue necesaria para activar la polimerasa. Típicamente una reacción de PCR se ajusta a la ecuación:

$X_n = X_0 (1 + E_x)^n$, donde:

X_n = Número de copias del gen de estudio en el ciclo n

X_0 = Número de copias iniciales del gen de estudio

E_x = Eficacia de amplificación del gen de estudio (X)

n = Número de ciclos

POBLACIÓN Y METODOLOGÍA

La información generada por el termociclador en cada uno de los ciclos es captada y enviada a un ordenador con el programa informático Sequence Detection Software (SDS) v 2.4 (Applied Biosystems) que construye las curvas de amplificación a partir de los datos de emisión de fluorescencia recogidos durante la reacción en tiempo real. Finalmente se obtiene una representación en la que el eje de ordenadas es el logaritmo de la intensidad de la fluorescencia y el eje de abscisas es el número de ciclos transcurridos (figura 23).

El parámetro de medida de la expresión de un determinado gen es el ciclo en el que la amplificación comienza a ser exponencial. Este ciclo se denomina ciclo umbral (Ct: *threshold cycle*), a partir del cual la intensidad de fluorescencia empieza a ser apreciable y proporcional a la cantidad inicial de ADNc. De este modo, los valores de ciclo umbral decrecerán linealmente conforme aumenta la cantidad de ADNc de partida. Cuantas más copias de ARNm de partida del gen estudiado, más ADNc se obtendrá en la transcripción inversa, y antes comenzará la amplificación a ser exponencial. El valor umbral en todos nuestros experimentos de RT-PCR se estableció en 0,01.

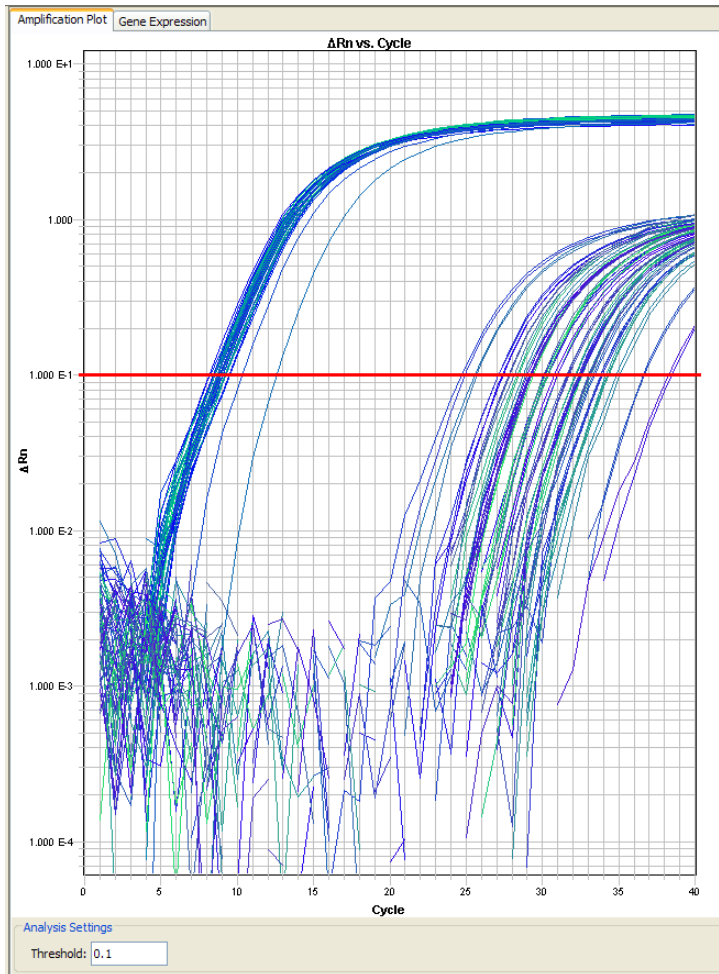


Figura 23. Curvas de amplificación del gen de estudio por PCR-RT. La línea horizontal de color rojo indica el umbral (*threshold*) a partir del cual empieza la linealidad, se establece como umbral el valor de 0,1 para todos los grupos de estudio. ΔRn = Incremento de fluorescencia. Cycle: número de ciclo en que se encuentra la reacción.

III. 3. ANÁLISIS ESTADÍSTICO

III. 3.1. CÁLCULO DEL TAMAÑO MUESTRAL

No existen publicaciones previas con este tipo de pacientes. Se utilizaron las diferencias en los niveles de expresión publicados en el estudio de niños con neumonía, con una razón entre grupos de 0.5 (ya que el rendimiento del cultivo está en torno al 50%), un poder del 90% y un intervalo de confianza del 95%. Para encontrar diferencias estadísticamente significativas, el tamaño muestral necesario sería de 8 pacientes sin sepsis y 4 casos de sepsis. Estimando las pérdidas que pudiera haber y los errores en el procesamiento de las muestras, se estimó un tamaño de 30 pacientes sin sepsis y 15 con sepsis.

III. 3.2. ANÁLISIS DE VARIABLES CLÍNICAS Y DEMOGRÁFICAS

Se realizó el análisis Kolgomorov–Smirnov para testar la distribución de los datos. Las variables continuas con una distribución normal se expresan como media \pm DS, y las no normales como mediana y rango intercuartílico. Las variables continuas se compararon usando el test T-Student o el Análisis de la varianza (ANOVA) en las que presentaron distribución normal y el test U-Mann Whitney en el resto. Las variables categóricas se compararon usando el test Chi-cuadrado o

exacto de Fisher (dos colas) según el número encontrado en cada casilla. El análisis se realizó usando el programa SPSS[®] versión 17.0 (SPSS Inc., Chicago, IL, EE.UU.). Se consideró estadísticamente significativo un valor de $p \leq 0.05$.

III. 3.3. ANÁLISIS DE LOS PERFILES EXPRESIÓN GENÓMICA

Se realizó el análisis usando una muestra por chip. Los archivos se capturaron usando el Affymetrix[®] GeneChip Scanner 3000 7G y se analizaron y filtraron estadísticamente con el programa Partek Genomic Suite 6.6 software (Partek Inc., St. Louis, MO, USA). Se normalizaron los datos con el algoritmo “robust multiple-array average” (RMA)⁹⁹.

Análisis multivariante basado en métodos de proyección

Se utilizó un análisis multivariante basado en la proyección de la matriz de datos original en un nuevo sistema de coordenadas. Los métodos de proyección tienen una serie de ventajas cuando se dispone de un conjunto de variables: permiten condensar la información en un número reducido de nuevas variables, calculadas a partir de las iniciales, de manera que expliquen una parte significativa de la

POBLACIÓN Y METODOLOGÍA

varianza total inicial. Estas nuevas variables (i.e. componentes principales o variables latentes), permiten eliminar ruido de los datos originales y, al mismo tiempo, analizar visualmente relaciones entre muestras y/o variables iniciales para, por ejemplo, establecer asociaciones entre grupos de muestras, discriminar entre poblaciones o clasificar nuevos individuos.

En este trabajo se seleccionaron dos estrategias de análisis multivariante basados en métodos de proyección: análisis de componentes principales (PCA: *Principal Component Analysis*) y regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA: *Partial Least Squares Discriminant Analysis*). En primer lugar se aplicó PCA, para un análisis exploratorio no supervisado inicial de los datos. Posteriormente, se generó un modelo predictivo mediante PLS-DA. Este segundo abordaje supervisado, permite llevar a cabo el cálculo de un modelo de clasificación e identificación de variables discriminantes. Los resultados obtenidos se compararon con los obtenidos a partir de un análisis univariante mediante ANOVA para la selección de potenciales marcadores diagnósticos.

Análisis de Componentes Principales (PCA)

Se realizó un Análisis de Componentes Principales (PCA) para determinar las fuentes significativas de varianza en los conjuntos de datos. PCA se puede utilizar como método no supervisado de clasificación para detectar la presencia de agrupaciones o *clusters*, y por otra parte para la detección de *outliers*.

Un modelo PCA se obtiene a partir de una serie de combinaciones lineales de las variables originales, ortogonales entre sí, que busca explicar la máxima varianza en \mathbf{X} de manera decreciente. Mientras que los primeros componentes principales son capaces de explicar la mayor parte de la varianza encontrada en la matriz de datos original, el resto de los componentes principales contienen fundamentalmente ruido. De esta manera, el PCA actúa como un potente método de reducción de la complejidad de los datos y simplifica la tarea de identificar patrones y fuentes de variabilidad a través de representaciones gráficas¹⁰⁰. La proyección de las muestras en el nuevo sistema multidimensional calculado mediante PCA permite establecer similitudes entre ellas utilizando p.ej. la distancia euclídea entre pares de muestras.

ANOVA

En todas las muestras se realizó un test de ANOVA. Los genes con diferencias de expresión estadísticamente significativas frente a la cohorte de no infectados, se identificaron usando un análisis de la varianza con un FDR (*fold discovery rate*) < 0.05 . Se trata de una medida mucho más restrictiva que el p-valor, considerando significativos valores de p en torno a 10^{-6} . Los genes significativos obtenidos a partir del ANOVA se ordenaron de acuerdo a su nivel de expresión en una agrupación jerárquica no supervisada (*unsupervised hierarchical clustering*). Finalmente los genes diferencialmente expresados se introdujeron en el programa Pathway Studio versión 9 (Ariadne Genomics® software, Elsevier® Inc, Rockville, MD, EE.UU.) para identificar los reguladores, rutas de señalización y procesos biológicos más significativamente implicados.

III. 3.4. MODELO PREDICTIVO: REGRESIÓN DE MÍNIMOS CUADRADOS PARCIALES CON ANÁLISIS DISCRIMINANTE.

A partir de las muestras de la cohorte inicial (en las que se habían comparado sepsis documentadas microbiológicamente con controles) se realizó una regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA: *Partial Least Squares Discriminant Analysis*)

con el programa MATLAB[®] 2014 (MathWorks[®] Natick, Massachusetts, EE.UU.).

La regresión de mínimos cuadrados parciales (PLS) es un método estadístico supervisado. Los modelos supervisados se calculan a partir de datos multivariantes recogidos en una cohorte de sujetos clasificados como p.ej. con/sin sepsis, con dos objetivos principales: clasificar el estado de futuros pacientes a partir del análisis de muestras biológicas; o identificar las variables responsables de la discriminación entre las dos clases. PLS es un modelo bilineal que ajusta los valores de la matriz de datos \mathbf{X} y el vector de clasificación \mathbf{y} (en nuestro caso RNMBP con sepsis vs no infectados) y los reestructura en un nuevo sistema de coordenadas. Cada eje ortogonal en PLS se llama variable latente (LV: *latent variable*). La carga o *loading* correspondiente a cada LV contiene el peso de cada variable en el modelo. Se utiliza para encontrar las relaciones fundamentales entre \mathbf{X} e \mathbf{y} , es decir, un enfoque de variable latente para modelar la estructura de covarianza en estos dos espacios. El modelo de PLS trata de encontrar el sentido multidimensional en el espacio de \mathbf{X} que explica la dirección de la máxima varianza multidimensional en el

espacio \mathbf{y} . Es capaz de conectar la matriz de datos \mathbf{X} y el vector de clasificación \mathbf{y} para generar resultados fácilmente interpretables¹⁰¹.

La principal diferencia entre PLS y el PCA es que la selección del eje se basa en la regresión de \mathbf{X} frente a \mathbf{y} . El modelo PLS trata de maximizar las varianzas de ambos componentes (\mathbf{X} e \mathbf{y}), mientras que el PCA, al ser un modelo no supervisado, sólo tiene en cuenta la matriz de datos \mathbf{X} . Cuando el vector \mathbf{y} es binario (en este caso: sepsis y no sepsis), se calcula un modelo PLS de análisis discriminante (PLS-DA: *Partial least squares discriminant analysis*). El PLSDA es un método que encuentra una combinación lineal de características que se pueden usar para clasificar datos en diferentes grupos. El modelo queda definitivamente validado al utilizar una cohorte de datos independientes que no se utilizaron para generar el modelo, llamado conjunto de validación externo.

En este estudio se utilizaron los *VIP scores (variable importance in projection)* para seleccionar los genes con mayor influencia en el modelo discriminante. Utilizando este subconjunto de genes discriminantes, se generó un segundo modelo predictivo y se calculó el área bajo la curva, el error de clasificación y los valores de

sensibilidad, especificidad, valores predictivos positivo y negativo. Se comprobó que esos genes formaban parte del listado de genes significativos que había proporcionado el ANOVA y se validaron por RT-PCR.

Como parte de uno de los objetivos secundarios, se introdujeron en el modelo predictivo las muestras diagnosticadas como sepsis clínicas para comprobar si desde un punto de vista transcriptómico se comportaban como sepsis, aunque no se hubiera aislado ningún germen en el hemocultivo; o si por el contrario, su perfil era similar al de los controles y el diagnóstico clínico fue erróneo.

III. 3.5. ANÁLISIS RT-PCR

Para el análisis de los resultados de RT-PCR usamos el método semi-cuantitativo denominado $\Delta\Delta Ct$, en el que se comparan directamente los Cts del gen testado y gen de referencia en cada muestra, y posteriormente se comparan los Ct de la muestras experimental con respecto a la muestra control¹⁰². Este método permite referir la cantidad de ARNm que existe en una muestra frente a otra de referencia. En líneas generales consiste en calcular ΔCt , que es la diferencia entre el Ct del gen que se quiere cuantificar y el Ct del

POBLACIÓN Y METODOLOGÍA

estándar interno, en nuestro caso el 18S. Comparando el ΔCt frente al ΔCt de una muestra control podemos calcular cuánto más o menos se expresa este gen problema en dicha muestra frente al control. Una vez conocidos los valores de Ct para el gen problema y para gen de referencia interno, tanto en nuestra muestra control como en la muestra problema, tan sólo basta aplicar la siguiente fórmula para conocer la expresión relativa:

Expresión relativa: $2^{-\Delta\Delta Ct}$

Donde $\Delta\Delta Ct$: [Ct de X gen problema - Ct gen de control interno] - [Ct X control - Ct gen de control interno].

Para que la determinación a través de esta fórmula aritmética sea válida, debe cumplirse que la eficiencia de la reacción para el gen problema y para el gen de referencia sea aproximadamente igual a 1.

Tratamiento estadístico de los datos

Cada muestra se estudió por triplicado haciendo tres réplicas de cada muestra de ADNc y por cada gen en estudio. Aplicamos la ecuación de comparación de Ct mencionada anteriormente para el cálculo de la diferencia de expresión génica entre genes a validar en el grupo de sepsis y el grupo de controles. El grupo control toma el valor de $2^{-\Delta\Delta Ct}$

como 1. En cada grupo de experimentos, n indica el número de pacientes a estudio.

Para la comparación de medias múltiples se realizó el análisis de la varianza (ANOVA), correspondiente a un diseño experimental complementario aleatorio de un solo factor. Cuando el ANOVA dio diferencias significativas se utilizó el test de Newman-Keuls. Para la aplicación de este test estadístico se empleó el programa GraphPad InStat[®] (GraphPad[®] Software, EE.UU.). Las diferencias entre grupos fueron estimadas por el test de la t-Student. Las diferencias se consideraron estadísticamente significativas para un valor de $p \leq 0,05$.

III. 3.6. COHORTE DE VALIDACIÓN INDEPENDIENTE

Se utilizó una cohorte independiente de muestras para validar los resultados obtenidos en la primera cohorte. En esta cohorte, se tomaron las muestras, se les extrajo el ARN y se hibridaron siguiendo la misma metodología que en la cohorte inicial. Se realizó un análisis no supervisado con un análisis de componentes principales (PCA) para comprobar que seguía habiendo diferencias entre las sepsis y los controles.

POBLACIÓN Y METODOLOGÍA

Posteriormente, se introdujeron las muestras de la cohorte de validación en el modelo generado PLS-DA con las muestras de la cohorte inicial para comprobar si el modelo las clasificaba correctamente.

IV. RESULTADOS

Este capítulo se estructura en varias partes:

1. Resultados de la cohorte inicial:

Datos poblacionales.

Análisis transcriptómico.

Modelo predictivo PLS-DA.

Rutas y procesos biológicos más significativos.

2. Validación por RT-PCR de los genes más significativos

3. Cohorte independiente para validar el modelo inicial:

Datos poblacionales.

Análisis transcriptómico global (PCA).

Capacidad diagnóstica del PLS-DA sobre las nuevas muestras.

4. Análisis de las sepsis clínicas:

Datos poblacionales.

Capacidad diagnóstica del PLS-DA sobre las nuevas muestras.

5. Estudio diferencial según la causa de la sepsis:

Datos poblacionales.

Análisis transcriptómico.

Modelo predictivo PLS-DA.

Validación del modelo predictivo: test de múltiples permutaciones con doble validación cruzada.

Rutas y procesos biológicos más significativos.

IV. 1. COHORTE INICIAL

IV. 1.1. DATOS CLÍNICOS Y DEMOGRÁFICOS

En la cohorte inicial, hubo 54 sospechas de sepsis, 13 de ellas no se reclutaron porque se inició el tratamiento antibiótico antes de la extracción de la muestra; por lo que se incluyeron 41 RNMBP con sospecha de sepsis y 27 controles. Entre las muestras iniciales se excluyeron 11 en el grupo control y 8 en los casos por falta de integridad del ARN total. De los 30 con sospecha de sepsis, se confirmó el diagnóstico con hemocultivo positivo en 17; así que finalmente se incluyeron 17 casos y 19 controles (figura 24). En 3 casos de sepsis, los controles fueron sus propios hermanos gemelos.

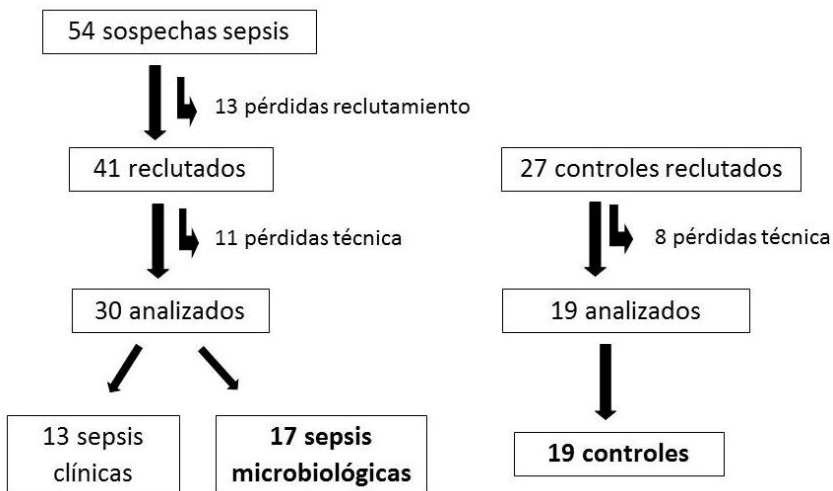


Figura 24. Esquema de reclutamiento.

No se observaron diferencias estadísticamente significativas en los datos clínicos y demográficos entre los 2 grupos ni al nacimiento (tabla 8) ni en el momento de la recogida de la muestra (Tabla 9).

Tabla 8. Características clínicas y demográficas de la cohorte inicial al nacimiento. Las variables cuantitativas con distribución normal se expresan en media y desviación estándar, y las cualitativas en número absoluto y porcentaje. La edad gestacional (EG) se expresa en semanas (s) y el peso al nacimiento (PN) en gramos (g). Los corticoides hacen referencia a las dosis de corticoides antenatales. ^{\$}Test T-Student ^{\$\$}Test exacto de Fisher (dos colas).

VARIABLES	SEPSIS (N = 17)	CONTROLES (N = 19)	P-VALOR
EG (s)	27 ± 2	28 ± 2	0.137 ^{\$}
SEXO (n, %)			0.813 ^{\$\$}
Hombre	11 (65%)	13 (68%)	
Mujer	6 (35%)	6 (32%)	
PN (g)	1030 ± 231	1130 ± 329	0.295 ^{\$}
RAZA (n, %)			0.152 ^{\$\$}
Caucásica	12 (70%)	17 (89%)	
Negra	3 (18%)	0 (0%)	
Hispano	2 (12%)	2 (11%)	
PARTO (n, %)			0.082 ^{\$\$}
Eutócico	9 (53%)	4 (21%)	
Cesárea	8 (47%)	15 (79%)	0.671 ^{\$}
APGAR ₁	6 ± 3	6 ± 2	0.664 ^{\$}
APGAR ₅	8 ± 2	9 ± 1	0.881 ^{\$\$}
CORTICOIDES			
1 dosis	3 (18%)	3 (16%)	
2 dosis	14 (82%)	16 (84%)	

RESULTADOS

Tabla 9. Características clínicas y demográficas de la cohorte inicial en el momento de la recogida de la muestra. Las variables cuantitativas cuya distribución fue normal se expresan en media y desviación estándar, y las que no en mediana y rango intercuartílico. Las variables cualitativas se expresan en número absoluto y porcentaje. La edad se expresa días de vida (d) y el peso en gramos (g). FiO₂: Fracción inspiratoria de oxígeno. V. central: vía central. S. Respiratorio: soporte respiratorio. ^{\$}Test T-Student ^{\$\$}Test exacto de Fisher (dos colas) ^{\$\$\$}Test U Mann-Whitney

VARIABLES	SEPSIS (N = 17)	CONTROLES (N = 19)	P-VALOR
EDAD (d)	11 (6-23)	14 (2-26)	0.876 ^{\$\$\$}
PESO (g)	1184 ± 360	1230 ± 401	0.720 ^{\$}
FiO ₂	0.21 (0.21-0.27)	0.21 (0.21-0.21)	0.156 ^{\$\$\$}
V. CENTRAL (n, %)	10 (59%)	10 (53%)	0.709 ^{\$\$}
S. RESPIRATORIO			0.346 ^{\$\$}
Sin soporte	5 (29%)	8 (42%)	
No invasivo	7 (42%)	9 (47%)	
Invasivo	5 (29%)	2 (11%)	

IV. 1.1.2. MICROBIOLOGÍA

Se aislaron los siguientes gérmenes en los hemocultivos: *Staphylococcus coagulasa negativo*: n=10 (59%); *Escherichia coli*: n=3 (17%); *Enterococcus faecalis*: n=2 (12%); *Staphylococcus aureus*: n=1 (6%) y *Morganella morgagnii*: n=1 (6%) (figura 25).

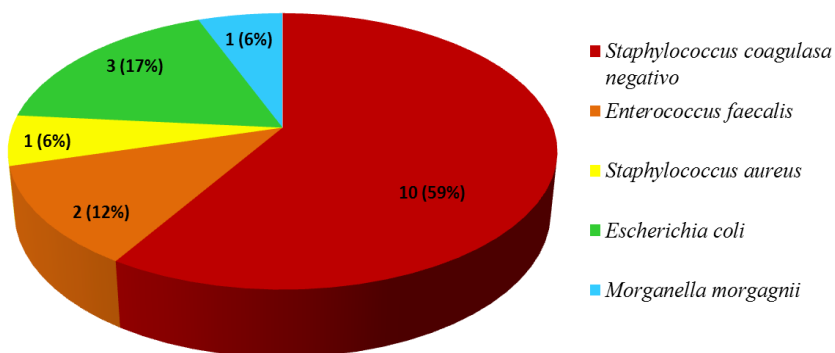


Figura 25. Distribución de los gérmenes aislados en la cohorte inicial. Se representa el número total de casos de cada germen y el porcentaje sobre el total de casos.

IV. 1.3. ANÁLISIS TRANSCRIPTÓMICO

IV. 1.3.1. Extracción del ARN total de sangre periférica

Tras la optimización del método de extracción del ARN total de sangre periférica, se obtuvo una media de valor de integridad (RIN) de 6.4 ± 0.1 , con concentraciones de 944 ± 135 ng/ ml.

En la tabla 10 se muestran los valores de integridad (RIN), las concentraciones de cada una de las 36 muestras y la pureza medida en ratio de absorbancia (A 260/280).

RESULTADOS

Tabla 10. Valores de ARN total en cuanto a integridad expresados en RIN (*RNA Integrity Number*), pureza en ratio de absorbancia (A260/280) y concentraciones en ng/μl. En las dos últimas filas están representados la media y desviación estándar (DS) para cada uno de los valores.

Muestra	RIN	Concentración ARN total (ng/μl)	Ratio (A 260/280)
1	6,3	331	2,12
2	6,1	475	2,15
4	6,3	492	2,16
6	6,7	168	2,26
7	6,5	1294	1,98
8	6,5	398	2,15
9	6,3	739	2,11
14	6,3	739	2,11
16	6,2	912	2,11
17	6,6	230	2,18
19	6,1	768	2,12
20	6,8	310	2,12
22	6,1	569	2,15
24	6,9	1831	2,10
28	6,3	1697	2,09
29	6,1	665	2,09
30	6,1	929	2,01
32	6,6	295	2,08
35	6,4	283	2,14
39	6,1	631	2,10
40	6,5	926	2,07
41	6,8	922	2,07
42	6,7	293	2,14
44	6,2	698	2,12
49	6,3	758	2,12
50	6,4	1894	2,10
52	6,2	1778	2,11
53	6,5	413	2,02
55	6,9	907	2,09
56	8,1	895	2,10
58	7,1	1584	2,08
59	6,6	4553	2,08
60	6,1	691	2,08

Muestra	RIN	Concentración ARN total (ng/μl)	Ratio (A 260/280)
61	6,1	1200	2,08
65	6,4	2114	2,07
66	6,5	602	2,08
Media	6,5	944	2,10
DS	0,42	817	0,05

IV. 1.3.2. Análisis de componentes principales

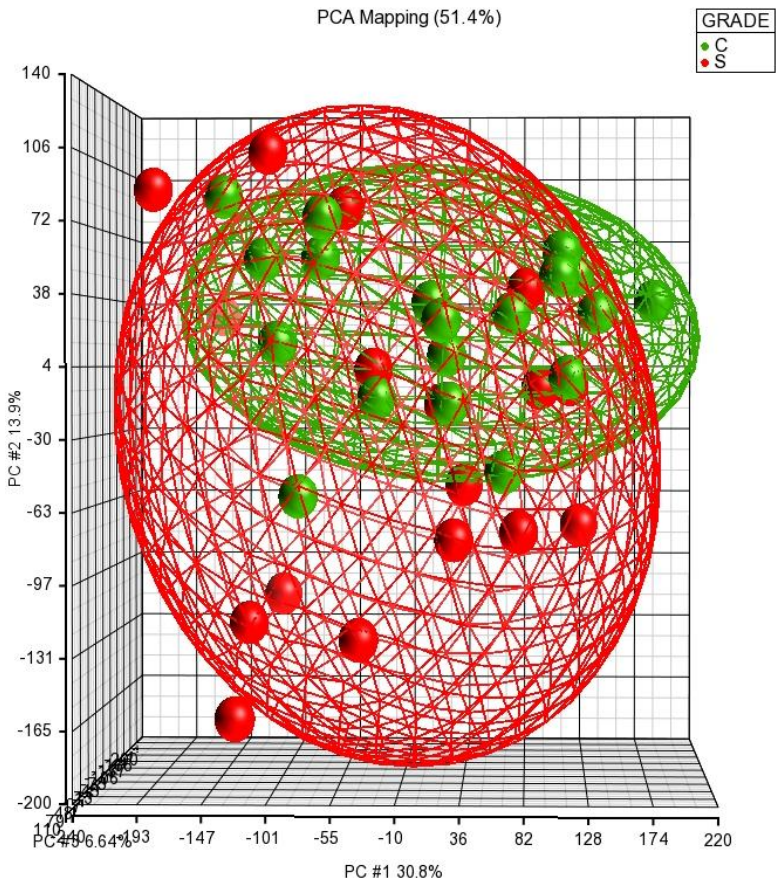
Analizamos simultáneamente todos los pacientes para identificar perfiles transcripcionales que permitieran distinguir a los RNMBP con sepsis bacteriana comparada con controles no infectados. Realizamos un análisis de componentes principales tridimensional no supervisado (PCA: *Principal Component Analysis*) basado en el análisis de todo el genoma (33.297 genes) de los 17 casos y los 19 controles (figura 26). El PCA identificó 2 grupos independientes de acuerdo a la expresión genómica, uno en sentido vertical, que se correspondió con las sepsis y otro en horizontal en el que se agrupaban los controles (figura 26-A). Sin embargo, en el grupo de las sepsis se observaron 2 grupos: uno con 4 muestras aisladas en la zona superior y las restantes 13 muestras en la zona más inferior. La diferencia básica entre estos 2 grupos fue el germen causal de la sepsis, las 4 muestras se correspondieron con sepsis por bacterias Gram negativas, y las 13

RESULTADOS

muestras restantes que fueron causadas por bacterias Gram positivas.

La figura 26-B muestra el PCA con los 3 grupos independientes: sepsis por bacterias Gram positivas, sepsis por bacterias Gram negativas y controles.

26-A



26-B

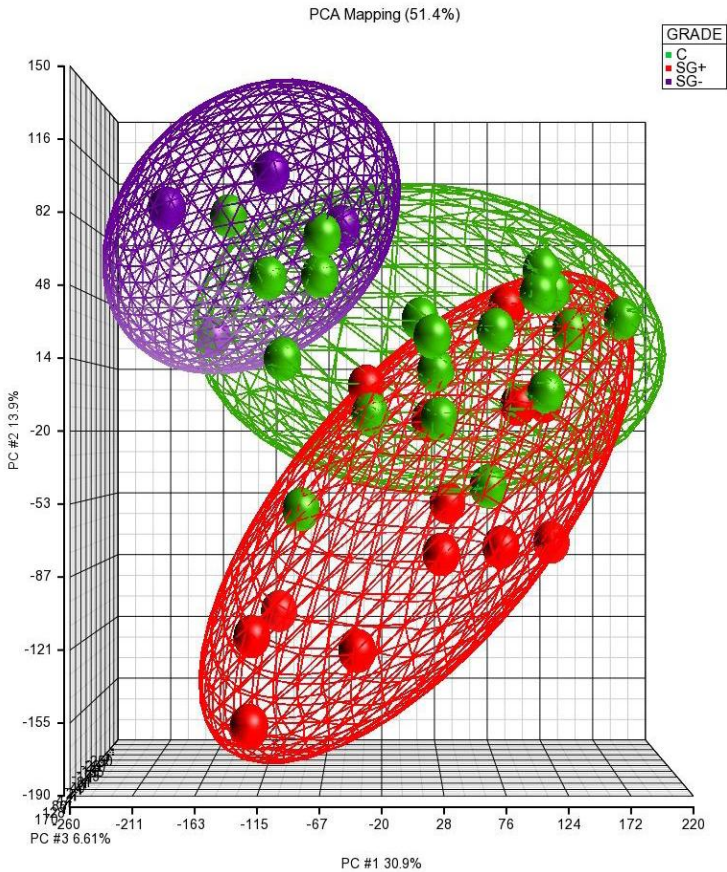


Figura 26. Análisis de Componentes Principales (PCA: *Principal component analysis*) tridimensional basado en todo el genoma en la cohorte inicial. Cada paciente (representado por una esfera) queda posicionado en función de expresión global a lo largo de los tres ejes. El panel 26-A muestra en sentido horizontal 19 controles (C: esferas verdes) y en sentido vertical 17 sepsis (S: esferas rojas). El panel 26-B es el mismo PCA pero distinguiendo las sepsis en función de su germen causal: 19 controles (C: esferas verdes), 13 sepsis por gérmenes Gram positivos (SG+: esferas rojas) y 4 sepsis por gérmenes Gram negativos (SG-: esferas moradas).

RESULTADOS

IV. 1.3.3. Agrupación jerárquica no supervisada

Posteriormente se llevaron a cabo múltiples comparaciones estadísticas usando el test ANOVA entre los casos y los controles encontrando 554 genes con una significación estadística de $FDR < 0.05$, de los cuales 305 (55%) estaban sobreexpresados y 249 (45%) infraexpresados en los RNMBP infectados.

Realizamos una agrupación jerárquica no supervisada (*unsupervised hierarchical clustering*) que mostró 3 grupos de pacientes: el grupo 1 formado por 13 controles, el grupo 2 que fue mixto formado por 2 RNMBP con sepsis por *E.coli*, 1 por *E.faecalis* y 4 con sepsis por *S.coagulasa negativo*, mezclados con 6 controles; y finalmente un grupo 3 que incluyó un control, 1 RNMBP con sepsis por *M. morganii* y 8 con sepsis por Gram positivos incluyendo: 6 *S.coagulasa negativo*, 1 *E. faecalis* y 1 *S.aureus* (Figura 27).

La señal de sobreexpresión en los RNMBP con sepsis por bacterias Gram positivas se relacionó con las funciones moleculares del factor de necrosis tumoral (TNF) y las citoquinas, y la señal de infraexpresión con el receptor de las células T.

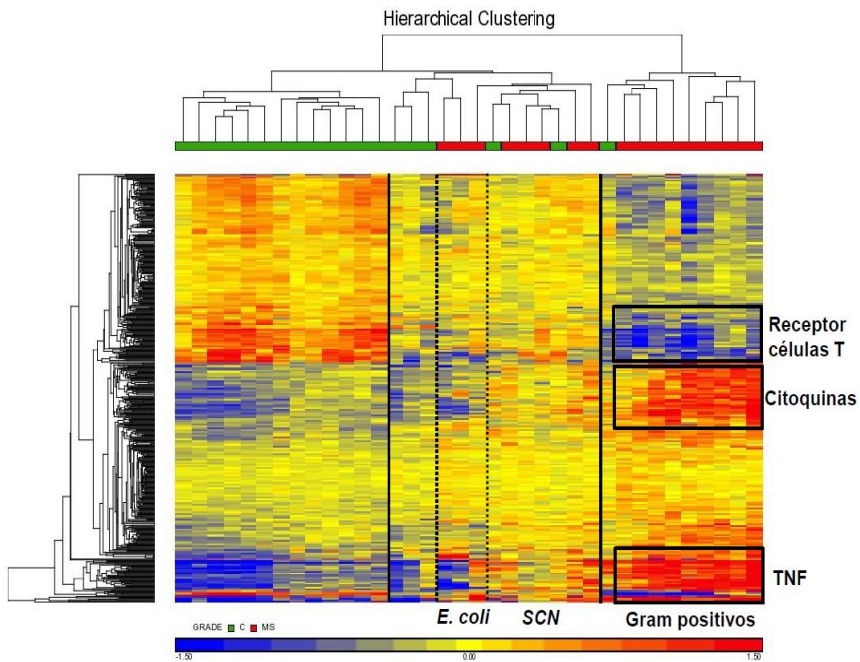


Figura 27. Agrupación jerárquica no supervisada de la cohorte inicial. Representa en sentido horizontal los 17 neonatos con sepsis y los 19 controles y en sentido vertical los 554 genes significativos en el análisis de la varianza. Cada columna representa un paciente y cada línea un gen. Los genes sobreexpresados se representan en rojo y los infraexpresados se representan en azul. Las barras verdes representan a los controles (n=19) (C) y las rojas a los pacientes con sepsis (n=17). TNF: Factor de necrosis tumoral. *SCN*: *Staphylococcus coagulasa negativo*.

IV. 1.4. MODELO PREDICTIVO: REGRESIÓN DE MÍNIMOS CUADRADOS PARCIALES CON ANÁLISIS DISCRIMINANTE

Se creó un modelo predictivo utilizando un *Partial Least Square*

Discriminant Analysis (PLS-DA), con un área bajo la curva (AUC)

Característica Operativa del Receptor (ROC: *Receiver Operating*

RESULTADOS

Characteristic) de 0.96, un error de clasificación del 11.1%, sensibilidad de 90% y Especificidad del 88%.

A partir de este modelo, se eligieron las variables con un VIP score superior a 60 (figura 28), que fueron los que tenían un mayor poder discriminativo. Se obtuvieron 10 variables que se correspondieron con 9 genes, ya que la sonda del gen CD177 estaba representada en 2 ocasiones: *Matrix metalloproteinase 8* (MMP8), *CD 177 molecule* (CD177), *Haptoglobin* (HP), *Olfactomedin 4* (OLFM4), *Glutathion S transferasa mu 1* (GSTM1), *Lipocalin 2* (LCN 2), *Carcinoembryonic antigen-related cell adhesion molecule 1* (CEACAM1), *Ankyrin repeat domain 22* (ANKRD22), *G protein-coupled receptor 84* (GPR84).

Todos estos genes formaban parte de la lista de 554 genes proporcionada por el ANOVA dentro de los más significativos. En la tabla 11 se representan los 9 genes del modelo predictivo, con los valores de p y de cambio de expresión que proporcionó el ANOVA.

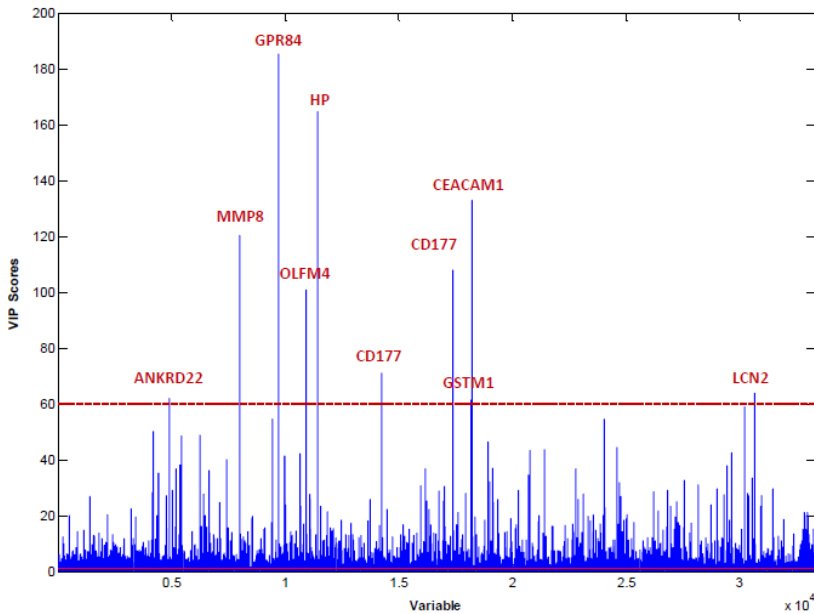


Figura 28. Regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) de la cohorte inicial. VIP score. La gráfica muestra en el eje de abscisas las variables y en el de ordenadas los valores de VIP score para cada una de ellas. Se especifican los genes que se corresponden con las 10 variables con VIP score superior a 60: *Matrix metalloproteinase 8* (MMP8), *CD177 molecule* (CD177), *Haptoglobin* (HP), *Olfactomedin 4* (OLFM4), *Glutathion S transferasa mu 1* (GSTM1), *Lipocalin 2* (LCN 2), *Carcinoembryonic antigen-related cell adhesion molecule 1* (CEACAM1), *Ankyrin repeat domain 22* (ANKRD22), *G protein-coupled receptor 84* (GPR84).

El cambio de expresión es una medida que describe cuánto cambia una cantidad desde un valor inicial a un valor final. En nuestro caso, se calcula como el ratio del valor de intensidad en los casos de sepsis entre valor de los controles. En el caso en que el cambio es hacia la disminución en lugar de al aumento, se puede representar igualmente

RESULTADOS

con el cociente obteniéndose un número entre 0 y 1, o bien calculando el negativo de su inverso; es decir, el negativo del ratio del valor de los controles entre el valor de las sepsis.

Tabla 11. Listado de los 9 genes correspondientes con las 10 variables con mayor poder discriminativo en el PLS-DA. ID representa el símbolo de identificación del gen y CE el valor del cambio de expresión.

ID GEN	NOMBRE	P-VALOR	CE
MMP8	<i>Matrix metalloproteinase 8</i>	$2,08 \times 10^{-6}$	7,34
CD177	<i>CD177 molecule</i>	$7,63 \times 10^{-6}$	6,80
GPR84	<i>G protein-coupled receptor 84</i>	$2,81 \times 10^{-7}$	5,13
ANKRD22	<i>Ankyrin repeat domain 22</i>	$3,11 \times 10^{-6}$	6,41
CEACAM1	<i>Carcinoembryonic antigen-related cell adhesion molecule 1</i>	$2,03 \times 10^{-6}$	3,79
OLFM4	<i>Olfactomedin 4</i>	$2,08 \times 10^{-5}$	7,11
HP	<i>Haptoglobin</i>	$8,20 \times 10^{-5}$	4,17
LCN2	<i>Lipocalin 2</i>	$4,74 \times 10^{-5}$	3,55
GSTM1	<i>Glutathione S transferase mu 1</i>	$3,19 \times 10^{-4}$	2,12

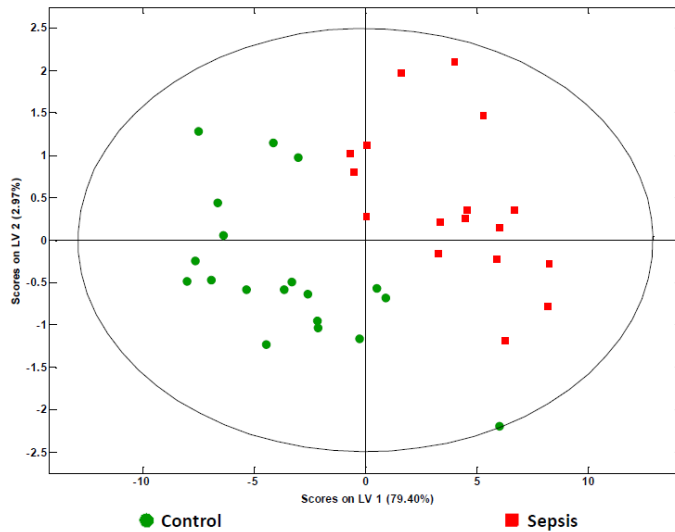


Figura 29. Regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) de la cohorte inicial. La gráfica muestra la distribución espacial de las sepsis (cuadrados rojos) que se agrupan en la zona de la derecha, y los controles (círculos verdes) agrupados en el lado izquierdo de la gráfica.

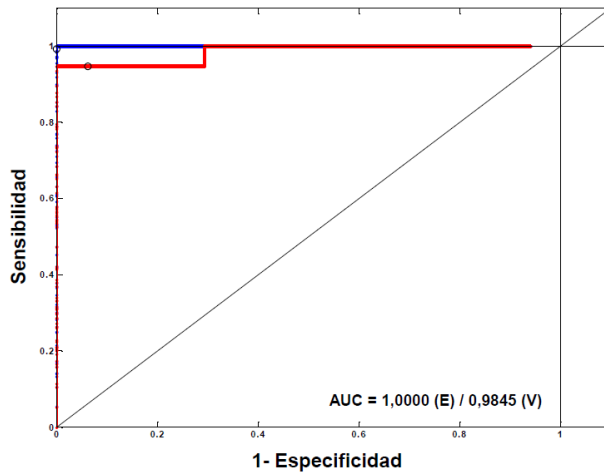


Figura 30. Área bajo la curva (AUC) Característica Operativa del Receptor (ROC) de la cohorte inicial. La gráfica muestra en el eje de abscisas 1-Especificidad y en el de ordenadas la Sensibilidad. La

RESULTADOS

curva Característica Operativa del Receptor (ROC) en rojo es la curva validada (V) con el modelo y el área que queda bajo la curva (AUC) es igual a 0.98. La curva ROC en azul es la curva estimada (E) que dejaría un AUC de 1.

IV. 1.5. GENES SIGNIFICATIVOS, REGULADORES Y RUTAS DE SEÑALIZACIÓN

Para entender el significado biológico de los 554 genes significativos analizamos los procesos biológicos más relevantes, que fueron: la respuesta inmune innata (p-valor= 4.46×10^{-9}) y la respuesta inflamatoria (p-valor = 1.24×10^{-7}). En la tabla 12 se representan las rutas de señalización más significativas en términos de p-valor y número de genes que integran.

Tabla 12. Rutas de señalización más significativas en las sepsis respecto a los controles de los 554 genes significativos usando Pathway Studio. Genes representa el número de genes implicados en cada regulador o ruta de señalización.

RUTAS DE SEÑALIZACIÓN	P-VALOR	GENES
Señal ligada al receptor de las células de superficie	8.07×10^{-6}	13
Receptor de células T	1.09×10^{-5}	8
Señal mediada por Interferón gamma	2.32×10^{-5}	7
Señal de la regulación negativa del receptor de insulina	5.21×10^{-5}	4
Señal mediada por citoquinas	5.73×10^{-5}	11

En la tabla 13 se muestran los reguladores más significativos en términos de p-valor y número de genes que integran. La ruta del TNF- α resultó la más significativa y la más enriquecida en cuanto a número de genes.

Tabla 13. Reguladores más significativos en las sepsis respecto a los controles de los 554 genes significativos usando Pathway Studio. Genes representa el número de genes implicados en cada regulador o ruta de señalización, perteneciente a la lista de los 554 genes significativos. TNF: Factor de necrosis tumoral, NF- κ B: Nuclear factor kappa-light-chain-enhancer of activated B cells, SP1: Specificity Protein 1, IFNG: Interferon gamma.

REGULADORES	P-VALOR	GENES
TNF	3.56×10^{-8}	67
Citoquinas	6.51×10^{-8}	57
NF-κB	4.39×10^{-8}	51
SP1	2.32×10^{-7}	53
IFNG	5.00×10^{-5}	45

La figura 31 muestra las redes de los 3 reguladores más importantes desde el punto de vista de significación estadística y por el número de genes que explican.

RESULTADOS

La figura 31-A muestra la red del TNF, que representa 67 de los 554 genes significativos (12%) incluyendo 5 de los 9 genes del modelo predictivo: MMP8, CD177, HP, LCN2 y GPR84.

La figura 31-B representa la red de las Citoquinas, que contiene 56 de los 554 genes (10%) e incluye 4 genes del modelo predictivo: MMP8, CD177, HP y LCN2.

Finalmente la figura 31-C muestra la red del NF-kB, que representa 51 de los 554 genes significativos (9%) incluyendo 3 de los 9 genes del modelo predictivo: MMP8, LCN2 y OLFM4.

La figura 32 representa una integración de los procesos celulares, clases funcionales y enfermedades relacionadas con los genes obtenidos en el modelo predictivo.

31-A TNF

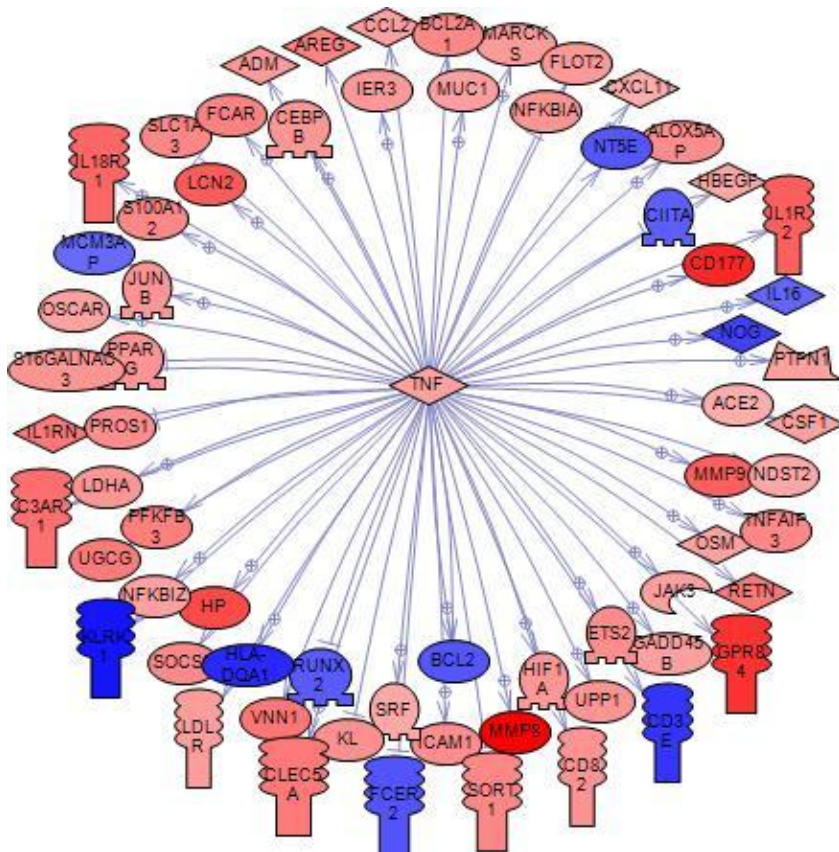
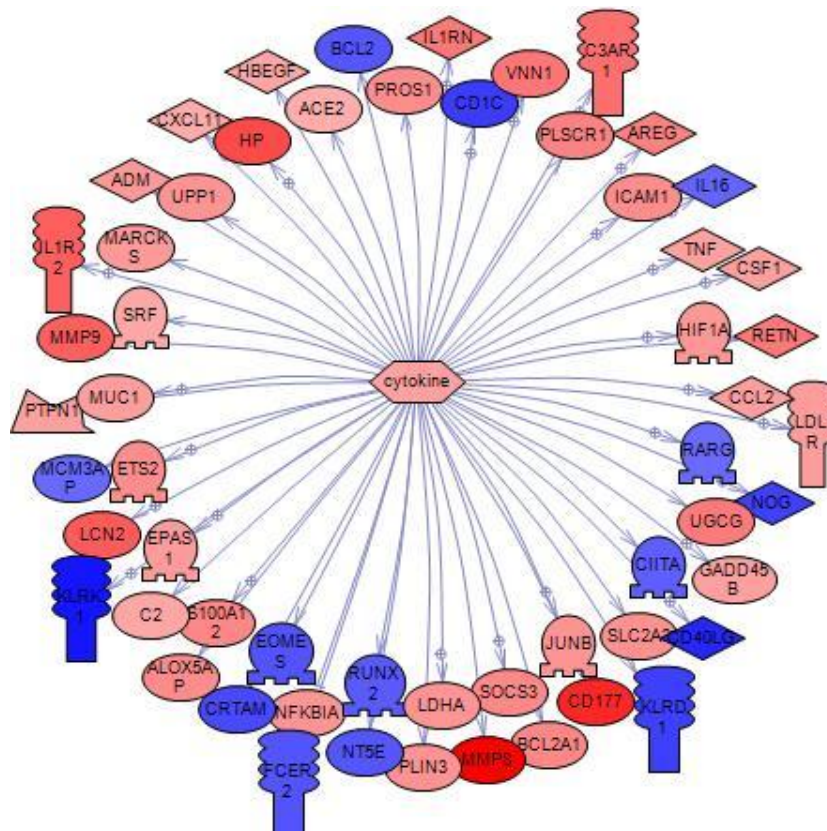
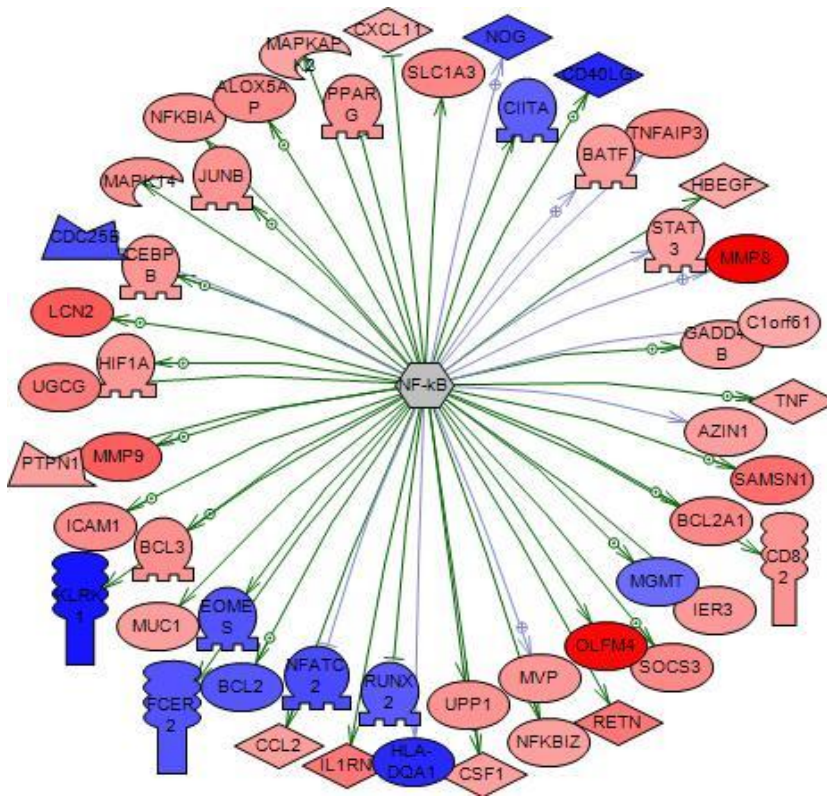


Figura 31. Redes de los principales reguladores en las sepsis respecto a los controles. Cada ruta representa el regulador y todos los genes del listado de 554 genes que explica. En azul se representan los genes infraexpresados y en rojo los sobreexpresados. La intensidad del color es directamente proporcional a la intensidad del cambio de expresión. Las flechas en verde indican activación y en gris regulación. La figura 31-A muestra la red del TNF (Factor de necrosis tumoral), la 31-B la red de las Citoquinas y la 31-C la del NF-kB (Nuclear factor kappa-light-chain-enhancer of activated B cells).

31-B CITOQUINAS



31-C NF-kB



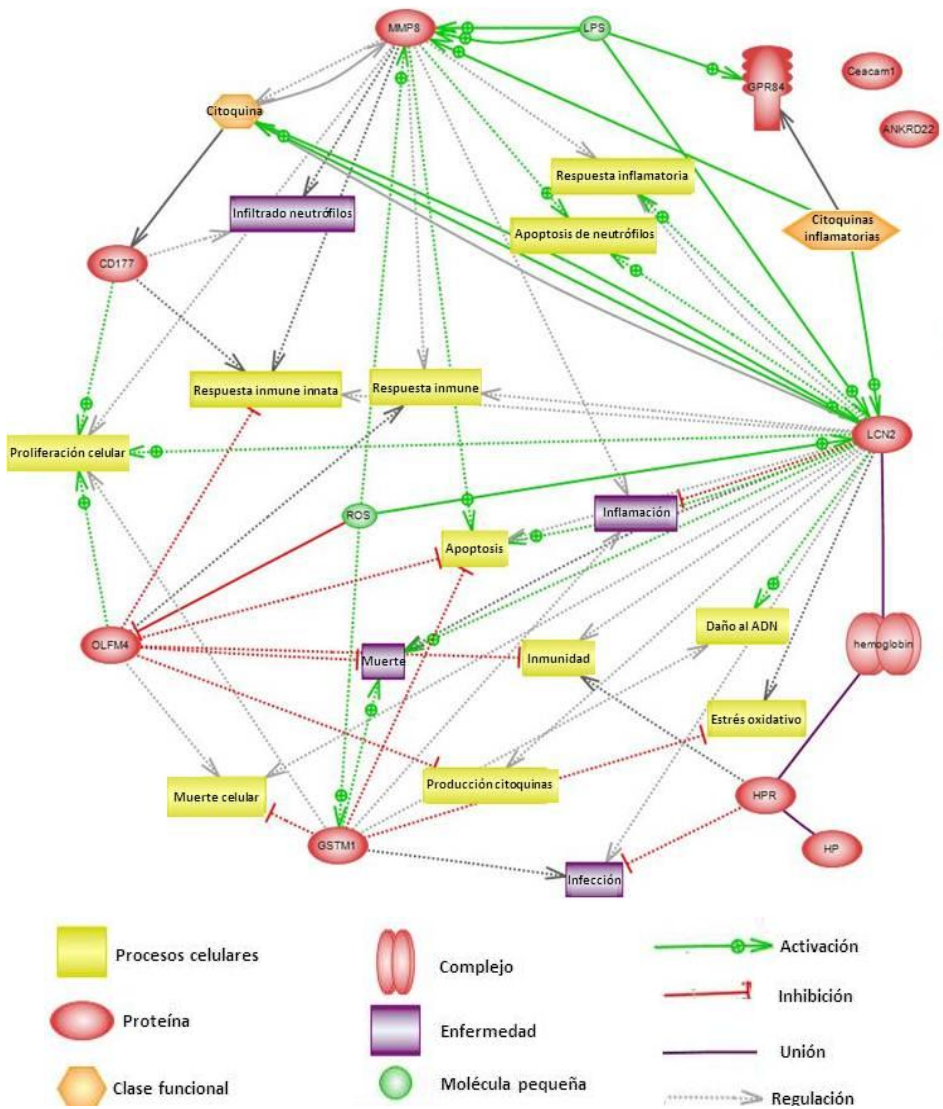


Figura 32. Integración de los procesos celulares, clases funcionales y enfermedades relacionadas con los genes obtenidos en el modelo predictivo. Matrix metallopeptidase 8 (MMP8), CD177 molecule (CD177), Haptogloblin (HP), Olfactomedin 4 (OLFM4), Glutation S transferase mu 1 (GSTM1) y Lipocalin 2 (LCN 2).

IV. 2. VALIDACIÓN POR RT-PCR

Los genes más significativamente expresados desde un punto de vista estadístico y funcional se validaron por RT-PCR: *Matrix metalloproteinase 8* (MMP8), *CD177 molecule* (CD177), *Haptoglobin* (HP), *Olfactomedin 4* (OLFM4), *Glutation S transferase mu 1* (GSTM1), *Lipocalin 2* (LCN 2) (tabla 14). Los cambios de expresión fueron significativamente superiores en los casos de sepsis respecto a los controles. En todos los casos estas diferencias alcanzaron la significación estadística utilizando el test T-Student (tabla 15).

Tabla 14. Genes validados por RT-PCR. Están representados los genes que se validaron, el símbolo del gen, la referencia de la sonda Taqman para ese gen y la longitud del amplicón en pares de bases (pb).

Nombre	Símbolo	Referencia	Longitud
<i>Matrix metalloproteinase 8</i>	<i>MMP8</i>	Hs01029057_m1	67 pb
<i>CD177 molecule</i>	<i>CD177</i>	Hs00360669_m1	66 pb
<i>Haptoglobin</i>	<i>HP</i>	Hs00605928_g1	90 pb
<i>Olfactomedin 4</i>	<i>OLFM4</i>	Hs00197437_m1	85 pb
<i>Glutation S transferase mu 1</i>	<i>GSTM1</i>	Hs01683722_gH	80 pb
<i>Lipocalin 2</i>	<i>LCN 2</i>	Hs01008571_m1	61 pb
18S	<i>Gen control</i>	Hs99999901_s1	187 pb

RESULTADOS

Tabla 15. Valores de $2^{-\Delta\Delta Ct}$. Valores de media y error estándar de $2^{-\Delta\Delta Ct}$ para cada uno de los genes validados en el grupo de sepsis y en el de los controles: *Matrix metalloproteinase 8* (MMP8), *CD177 molecule* (CD177), *Haptoglobin* (HP), *Olfactomedin 4* (OLFM4), *Glutathion S transferase mu 1* (GSTM1) y *Lipocalin 2* (LCN 2). Valor de p calculado con el test t-Student.

	CONTROLES	SEPSIS	P-VALOR
MMP8	1,00 ± 0,95	7,80 ± 4,42	0,0293
CD 177	1,00 ± 1,30	9,20 ± 5,37	0,0045
HP	1,00 ± 0,25	2,67 ± 3,29	0,0370
OLFM4	1,00 ± 0,70	17,90 ± 17,73	0,0068
GSTM1	1,00 ± 2,07	29,34 ± 18,90	0,0039
LCN2	1,00 ± 0,54	6,35 ± 5,40	0,0431

En la figura 33 se representan los valores de RT-PCR para cada uno de estos genes en forma de diagrama de cajas. Las diferencias que se observan fueron estadísticamente significativas en todos los casos.

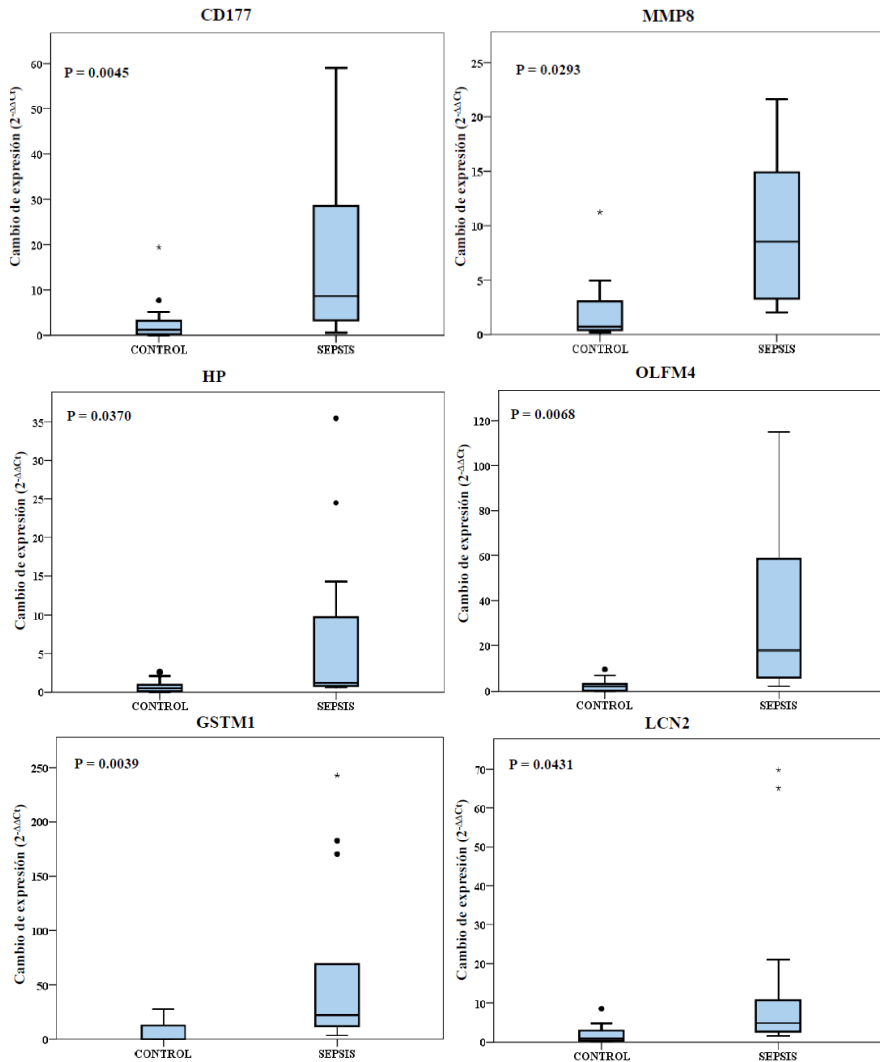


Figura 33. Validación por RT-PCR. En el diagrama de cajas se representan la mediana y el rango intercuartílico del cambio de expresión de mRNA ($2^{-\Delta\Delta C_t}$) en los controles y las sepsis para cada gen: *CD177* molecule (*CD177*), *Matrix metalloproteinase 8* (*MMP8*), *Haptoglobin* (*HP*), *Olfactomedin 4* (*OLFM4*), *Glutation S transferase mu 1* (*GSTM1*) y *Lipocalin 2* (*LCN 2*). El valor de p se calculó con el test T-Student.

IV. 3. COHORTE DE VALIDACIÓN INDEPENDIENTE

IV. 3.1. DATOS CLÍNICOS Y DEMOGRÁFICOS

En la cohorte de validación se incluyeron 18 RNMBP, 8 casos de sepsis y 10 controles.

Tabla 16. Características clínicas y demográficas de la cohorte de validación al nacimiento. Las variables cuantitativas se expresan en media y desviación estándar ya que su distribución fue normal, y las cualitativas en número absoluto y porcentaje. La edad gestacional (EG) se expresa en semanas (s) y el peso al nacimiento (PN) en gramos (g). Los corticoides hacen referencia a las dosis de corticoides antenatales. ^{\$}Test T-Student ^{\$\$}Test exacto de Fisher (dos colas) ^{\$\$\$}Test Chi cuadrado.

VARIABLES	SEPSIS (N = 8)	CONTROLES (N = 10)	P- VALOR
EG (s)	28 ± 4	27 ± 2	0.357 ^{\$}
SEXO (n, %)			
Hombre	7 (88%)	6 (60%)	0.314 ^{\$\$}
Mujer	1 (35%)	4 (40%)	
PN (g)	1052 ± 175	879 ± 226	0.094 ^{\$}
RAZA (n, %)			0.493 ^{\$\$\$}
Caucásica	6 (75%)	8 (80%)	
Negra	1 (12.5%)	0 (0%)	
Hispano	1 (12.5%)	2 (20%)	
PARTO (n, %)			0.066 ^{\$\$}
Eutócico	1 (12%)	6 (60%)	
Cesárea	7 (88%)	4 (40%)	
APGAR ₁	7 ± 2	6 ± 2	0.309 ^{\$}
APGAR ₅	9 ± 1	8 ± 1	0.060 ^{\$}
CORTICOIDE			0.477 ^{\$\$}
1 dosis	0 (0%)	2 (20%)	
2 dosis	8 (100%)	8 (80%)	

No se observaron diferencias estadísticamente significativas en los datos clínicos y demográficos entre los 2 grupos ni al nacimiento (tabla 16) ni en el momento de la recogida de la muestra (Tabla 17). Las características demográficas y clínicas fueron similares a las de la cohorte inicial.

Tabla 17. Características clínicas y demográficas de la cohorte de validación en el momento de la recogida de la muestra. Las variables cuantitativas cuya distribución fue normal se expresan en media y desviación estándar, y las que no siguieron distribución normal en mediana y rango intercuartílico. Las variables cualitativas se expresan en número absoluto y porcentaje. La edad se expresa días de vida (d) y el peso en gramos (g). FiO₂: Fracción inspiratoria de oxígeno. V. central: vía central. S. Respiratorio: soporte respiratorio. ^{\$}Test T-Student ^{\$\$}Test exacto de Fisher (dos colas) ^{\$\$\$}Test Chi cuadrado ^{\$\$\$\$}Test u-Mann Whitney.

VARIABLES	SEPSIS (N = 8)	CONTROLES (N = 10)	P- VALOR
EDAD (d)	15 (2-41)	23 (2-30)	0.697 ^{\$\$\$\$}
PESO (g)	1245 ± 71	1010 ± 325	0.062 ^{\$}
FiO ₂	0.22 (0.21-0.35)	0.21 (0.21-0.21)	0.194 ^{\$\$\$\$}
V. CENTRAL (n, %)	3 (37%)	2 (30%)	0.608 ^{\$\$}
S. RESPIRATORIO			0.286 ^{\$\$\$}
Sin soporte	0 (0%)	1 (10%)	
No invasivo	5 (63%)	8 (80%)	
Invasivo	3 (37%)	1 (10%)	

RESULTADOS

IV. 3.2. MICROBIOLOGÍA

Se aislaron los siguientes gérmenes en los hemocultivos: *Staphylococcus aureus*: $n=3$ (38%) *Staphylococcus coagulasa negativo*: $n=1$ (12.5%); *Serratia marcescens*: $n=1$ (12.5%); *Enterobacter cloacae*: $n=1$ (12.5%); *Klebsiella pneumoniae*: $n=1$ (12.5%); y *Klebsiella oxytoca*: $n=1$ (12.5%) (figura 34).

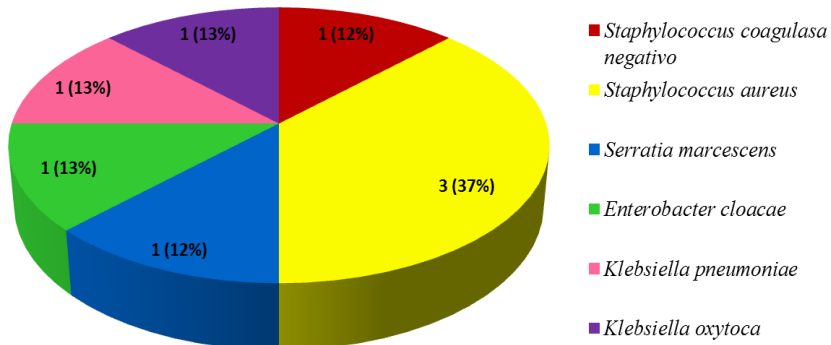


Figura 34. Distribución de los gérmenes aislados en la cohorte de validación. Se representa el número total de casos de cada germen y el porcentaje sobre el total de casos.

IV.3.3. ANÁLISIS DE COMPONENTES PRINCIPALES

En la cohorte de validación realizamos un PCA de las 8 sepsis y los 10 controles (Figura 35) para comprobar si se seguían distinguiendo los 3 grupos. El PCA identificó nuevamente 3 grupos independientes de

acuerdo a la expresión genómica que se correspondieron con: sepsis por bacterias Gram positivas, sepsis por bacterias Gram negativas y controles.

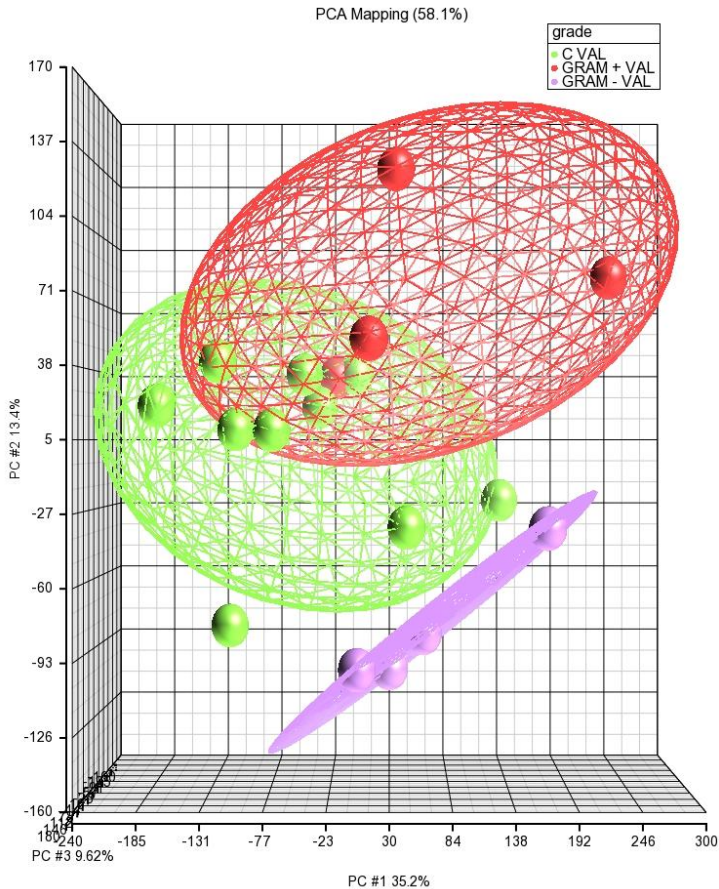


Figura 35. Análisis de Componentes Principales (PCA: Principal component analysis) tridimensional basado en todo el genoma en la cohorte de validación. Cada paciente (representado por una esfera) queda posicionado en función de sus respectivas posiciones a lo largo de los tres ejes. El PCA muestra 10 controles (C: esferas verdes) y 8 sepsis: 4 sepsis por gérmenes Gram positivos (GRAM +: esferas rojas) y 4 sepsis por gérmenes Gram negativos (GRAM -: esferas moradas).

IV. 3.4. MODELO PREDICTIVO: REGRESIÓN DE MÍNIMOS CUADRADOS PARCIALES CON ANÁLISIS DISCRIMINANTE

Se introdujeron las 18 muestras de la cohorte de validación en el modelo predictivo definitivo (figura 29) creado a partir de las 10 variables que se obtuvieron en el PLS-DA de la cohorte inicial.

Están representadas junto con las 36 muestras de la cohorte inicial en la figura 36, siguiendo el mismo código de color pero con tonos más claros. Se observa que todos los controles quedan agrupados junto a los controles de la cohorte inicial y únicamente 2 sepsis quedan en la zona intermedia entre las sepsis y los controles de la cohorte inicial. Asumiendo que estas 2 sepsis estuvieran mal clasificadas, el modelo presenta un error de clasificación del 11% en la cohorte de validación independiente, con lo que podemos afirmar que el modelo queda bien validado.

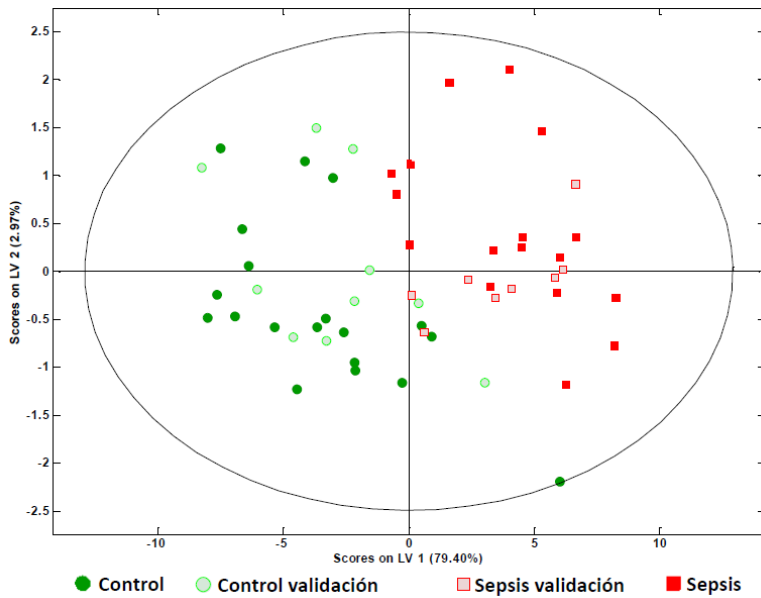


Figura 36. Regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) de la cohorte de validación. La gráfica muestra la distribución espacial de la cohorte inicial con las sepsis bacterianas (en cuadrados rojos) agrupadas en la zona de la derecha, y los controles (círculos verdes) agrupados en el lado izquierdo de la gráfica. Las muestras de la cohorte de validación están representadas en los mismos colores en tono más claro. Los controles en círculos verdes claros, agrupados junto a los controles de la cohorte inicial y las sepsis en cuadrados rojos claros, agrupados junto a las sepsis de la cohorte inicial.

IV. 4. ANÁLISIS DE LAS SEPSIS CLÍNICAS

Como parte de los objetivos secundarios se analizaron las 13 sepsis clínicas; es decir, aquellas en las que no se aisló ningún germen en el hemocultivo. Los datos clínicos y demográficos al nacimiento y en el

RESULTADOS

momento de la recogida de la muestra fueron similares a los pacientes de las cohortes inicial y de validación (tabla 18).

Tabla 18. Características clínicas y demográficas al nacimiento y en el momento de la recogida de la muestra de los 13 pacientes con sepsis clínicas. Las variables cuantitativas con distribución normal se expresan en media y desviación estándar, y las que no siguieron distribución normal en mediana y rango intercuartílico. Las variables cualitativas en número absoluto y porcentaje. La edad gestacional se expresa en semanas (s), la edad en el momento de la recogida de la muestra en días de vida (d) y el peso en gramos (g). FiO₂: Fracción inspiratoria de oxígeno.

VARIABLES AL NACIMIENTO (N =13)	
EDAD GESTACIONAL (s)	28 ± 2
SEXO (n, %)	
Hombre	10 (77%)
Mujer	3 (23%)
PESO NACIMIENTO (g)	1021 ± 285
RAZA (n, %)	
Caucásica	13 (100%)
PARTO (n, %)	
Eutócico	4 (31%)
Cesárea	9 (69%)
APGAR₁	7 ± 2
APGAR₅	9 ± 1
VARIABLES A LA RECOGIDA DE LA MUESTRA (N =13)	
EDAD (d)	7 (3-14)
PESO (g)	1049 ± 294
FiO₂	0.22 ± 0.02
VÍA CENTRAL (n, %)	8 (62%)
SOPORTE RESPIRATORIO	
Ventilación no invasiva	7 (54%)
Ventilación invasiva	6 (46%)

Estas muestras se introdujeron en el modelo predictivo con el fin de determinar si su expresión genómica era similar a las sepsis microbiológicas y por tanto podíamos afirmar que eran sepsis en las que no se había conseguido aislar ningún germen en el hemocultivo, o bien se comportaban transcriptómicamente como los controles y eran “falsos diagnósticos” de sepsis. En la figura 37 se observa como 7 (54%) de ellas se comportaron como auténticas sepsis, mientras que las otras 6 (46%) se agruparon junto a los controles.

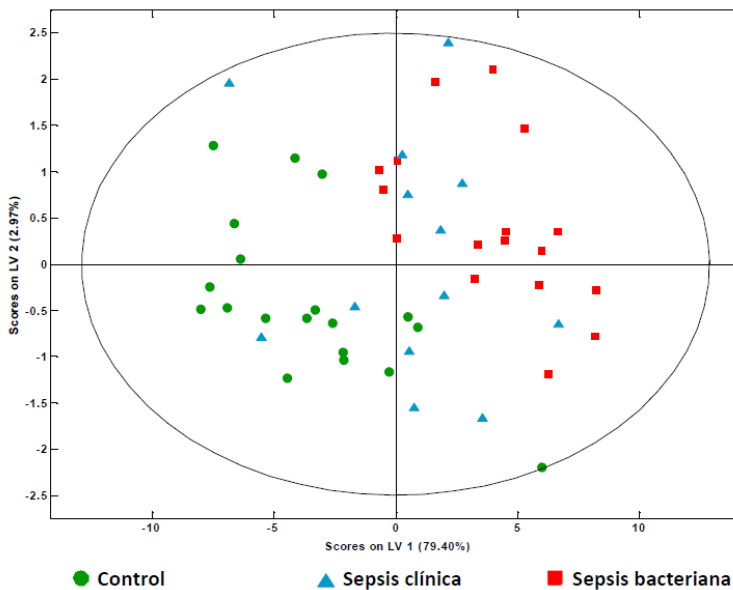


Figura 37. Regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) en las sepsis clínicas. La gráfica muestra la distribución espacial de las sepsis bacterianas (en cuadrados rojos) que quedan agrupadas en la zona de la derecha, y los controles (círculos verdes) que se agrupan en el lado izquierdo de la gráfica y en la zona intermedia las sepsis clínicas (triángulos azules).

IV. 5. ESTUDIO DIFERENCIAL SEGÚN EL AGENTE CAUSAL DE LA SEPSIS

Aunque el objetivo principal del estudio era establecer las diferencias en los perfiles de expresión genómica entre RNMBP con sepsis y los controles, encontramos que dentro del grupo de las sepsis había notables diferencias entre las producidas por bacterias Gram positivas y las producidas por bacterias Gram negativas. Sin embargo, en la cohorte inicial disponíamos sólo de 17 casos de sepsis, de los cuáles únicamente 4 fueron por Gram negativos. Para confirmar estos resultados, utilizamos las muestras de sepsis de la cohorte de validación junto con las de la cohorte inicial y realizamos un estudio diferencial entre las sepsis en función del tipo de germen causal.

IV. 5.1. DATOS CLÍNICOS Y DEMOGRÁFICOS

Se incluyeron las 25 sepsis microbiológicas analizadas: 17 causadas por gérmenes Gram positivos y 8 por gérmenes Gram negativos. No se observaron diferencias estadísticamente significativas entre los 2 grupos al nacimiento (tabla 19). Sin embargo, en el momento de la recogida de la muestra hubo un mayor porcentaje de pacientes sometidos a ventilación invasiva en las sepsis Gram negativos (tabla 20). Esto puede deberse a que las sepsis por gérmenes Gram negativos

suelen condicionar un mayor deterioro clínico que obliga en muchas ocasiones a la intubación y ventilación mecánica.

Tabla 19. Características clínicas y demográficas al nacimiento de los recién nacidos con sepsis en función del tipo de bacteria causal. Las variables cuantitativas se expresan en media y desviación estándar ya que su distribución fue normal, y las cualitativas en número absoluto y porcentaje. La edad gestacional (EG) se expresa en semanas cumplidas (s), el peso al nacimiento (PN) en gramos (g). Corticoides hace referencia a la administración de corticoide antenatal. ^{\$}Test t-Student ^{\$\$}Test exacto de Fisher (dos colas) ^{\$\$\$}Test Chi cuadrado.

VARIABLES	GRAM POSITIVOS (N = 17)	GRAM NEGATIVOS (N = 8)	P- VALOR
EG (s)	28 ± 2	27 ± 3	0.850 ^{\$}
SEXO (n, %)			
Hombre	11 (65%)	7 (88%)	0.362 ^{\$\$}
Mujer	6 (35%)	1 (12%)	
PN (g)	1032 ± 215	1048 ± 218	0.861 ^{\$}
RAZA (n, %)			0.320 ^{\$\$\$}
Caucásica	11 (65%)	7 (88%)	
Negra	4 (23%)	0 (0%)	
Hispano	2 (12%)	1 (12%)	
PARTO (n, %)			0.861 ^{\$\$}
Eutócico	7 (41%)	3 (37%)	
Cesárea	10 (59%)	5 (63%)	
APGAR ₁	7 ± 2	5 ± 3	0.053 ^{\$}
APGAR ₅	9 ± 1	8 ± 2	0.304 ^{\$}
CORTICOIDES			0.743 ^{\$\$}
1 dosis	3 (18%)	1 (12%)	
2 dosis	14 (82%)	7 (88%)	

RESULTADOS

Tabla 20. Características clínicas y demográficas en el momento de la recogida de la muestra de los recién nacidos con sepsis en función del tipo de bacteria causal. Las variables cuantitativas con distribución normal se expresan en media y desviación estándar, y las que no siguieron distribución normal en mediana y rango intercuartílico. Las variables cualitativas se expresan en número absoluto y porcentaje. La edad se expresa días de vida (d) y el peso en gramos (g). FiO₂: Fracción inspiratoria de oxígeno. V. central: vía central. S. Respiratorio: soporte respiratorio. ^{\$}Test t-Student ^{\$\$}Test exacto de Fisher (dos colas) ^{\$\$\$}Test Chi cuadrado ^{\$\$\$\$}Test U-Mann Whitney.

VARIABLES	GRAM POSITIVOS (N = 17)	GRAM NEGATIVOS (N = 8)	P- VALOR
EDAD (d)	12 (7-23)	10 (0-38)	0.552 ^{\$\$\$\$}
PESO (g)	1205 ± 348	1199 ± 168	0.964 ^{\$}
FiO ₂	0.21 (0.21-0.24)	0.25 (0.21-0.46)	0.262 ^{\$\$\$\$}
V. CENTRAL (n, %)	12 (71%)	5 (63%)	0.637 ^{\$\$}
S.RESPIRATORIO			0.006^{\$\$\$}
Sin soporte	4 (23%)	0 (0%)	
No invasivo	11 (65%)	2 (25%)	
Invasivo	2 (12%)	6 (80%)	

IV. 5.2. ANÁLISIS TRANSCRIPTÓMICO

IV. 5.2.1. Análisis de componentes principales

Realizamos un PCA de las 17 sepsis causadas por bacterias Gram positivas y las 8 sepsis por bacterias Gram negativas 10 controles. El

PCA identificó 2 grupos independientes de acuerdo a la expresión genómica que se correspondieron con: sepsis por bacterias Gram positivas y sepsis por bacterias Gram negativas (figura 38).

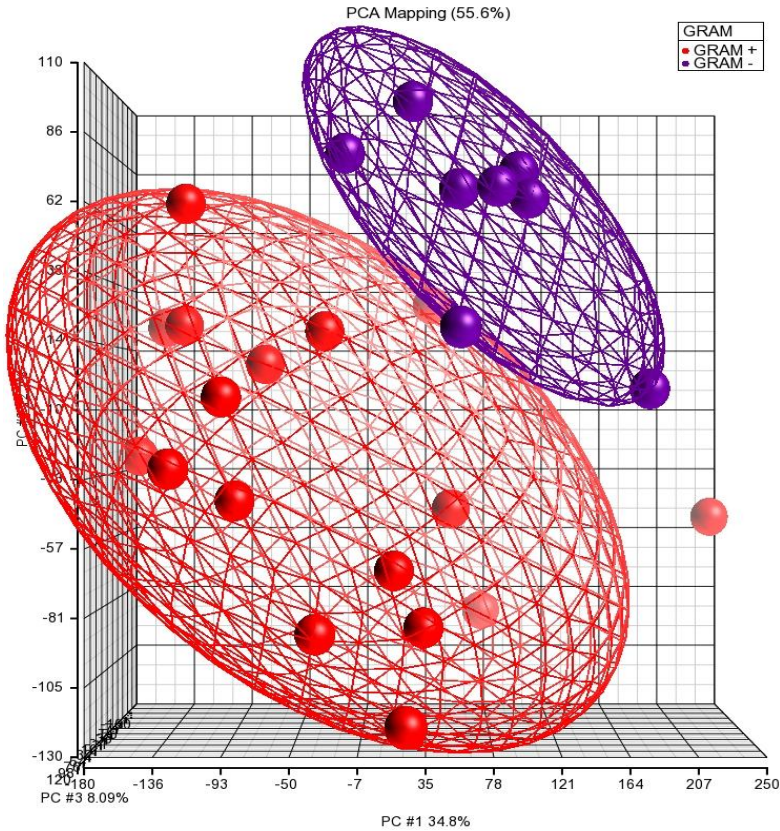


Figura 38. Análisis de Componentes Principales (PCA: Principal component analysis) tridimensional basado en todo el genoma en los casos de sepsis. Cada paciente (representado por una esfera) queda posicionado en función de sus respectivas posiciones a lo largo de los tres ejes. El PCA muestra 17 sepsis por gérmenes Gram positivos (GRAM +: esferas rojas) y 8 sepsis por gérmenes Gram negativos (GRAM -: esferas moradas).

RESULTADOS

IV. 5.2.2. Agrupación jerárquica no supervisada

Se llevaron a cabo múltiples comparaciones estadísticas usando el test ANOVA entre los RNMBP con sepsis por bacterias Gram positivas y RNMBP con sepsis por bacterias Gram negativas. Se identificaron 719 genes con una significación estadística de $FDR < 0.05$, de los cuales 433 (60%) estaban sobreexpresados y 286 (40%) infraexpresados en los RNMBP infectados.

Realizamos una agrupación jerárquica no supervisada (*unsupervised hierarchical clustering*) que mostró 2 grupos de pacientes (figura 39). A la izquierda quedaron agrupadas las sepsis por bacterias Gram negativas y a la derecha la sepsis por Gram positivas.

Dentro de ambos grupos se observa un subgrupo con la misma señal pero menos intensa. En el caso de las sepsis por Gram negativos se trata de 2 pacientes y en el de las Gram positivas de 9 pacientes. No se han observado diferencias clínicas, ni de evolución ni en cuanto al tipo de germen causal entre estos subgrupos.

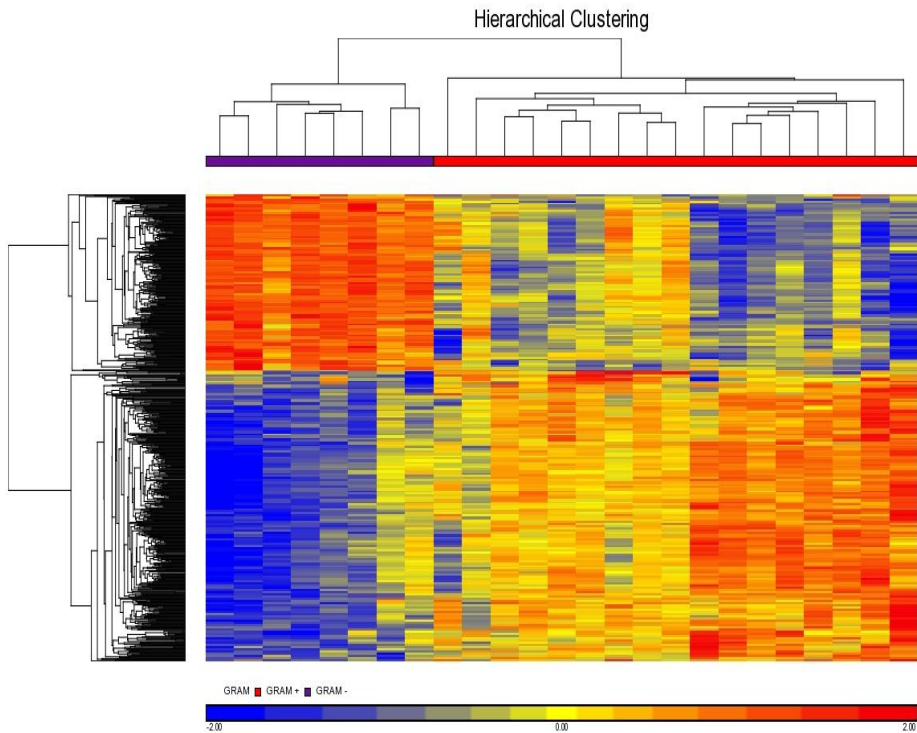


Figura 39. Agrupación jerárquica no supervisada de la cohorte de sepsis. Representados 17 neonatos con sepsis por bacterias Gram positivas y 8 neonatos con sepsis por bacterias Gram negativas (sentido horizontal) y los 554 genes (sentido vertical) significativos en el análisis de la varianza. Cada columna representa un paciente y cada línea un gen. Los genes sobreexpresados se representan en rojo y los infraexpresados en azul. Las barras rojas representan a los pacientes con sepsis por Gram positivos (GRAM +) (n=17) y las moradas a los pacientes con sepsis por Gram negativos (GRAM -) (n=8).

IV.5.3. MODELO PREDICTIVO: REGRESIÓN DE MÍNIMOS CUADRADOS PARCIALES CON ANÁLISIS DISCRIMINANTE

Se creó un modelo predictivo utilizando un *Partial Least Square Discriminant Analysis* (PLS-DA), con un área bajo la curva (AUC) Característica Operativa del Receptor (ROC: *Receiver Operating Characteristic*) de 1, un error de clasificación del 0%, sensibilidad de 100% y especificidad del 100%. A partir de este modelo, se eligieron las variables con un VIP score superior a 5 (figura 40), que fueron los que tenían un mayor poder discriminativo. Se obtuvieron 23 variables que se correspondieron con 21 genes, ya que una de las sondas no tiene un gen asignado, y la sonda del gen *CCDC72* estaba representada en 2 ocasiones.

De éstos 21 genes, 14 formaban parte de la lista de 719 genes proporcionada por el ANOVA. En la tabla 21 se representan estos genes, proporcionándose los valores de p y de cambio de expresión en los 14 casos que mostraron diferencias estadísticamente significativas en el ANOVA. Los cambios de expresión positivos indican una sobreexpresión de esos genes en las sepsis por bacterias Gram positivas respecto a las Gram negativas, mientras que los cambio de expresión negativos indican infraexpresión.

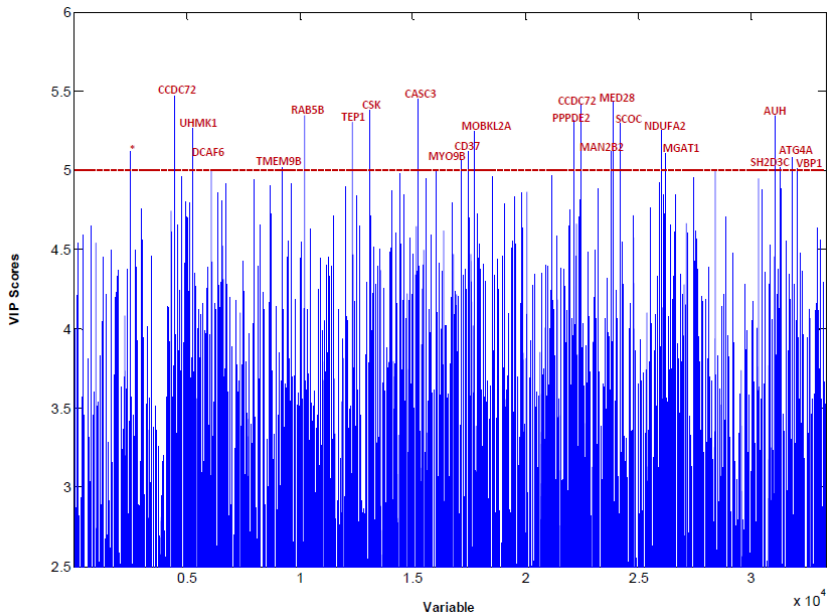


Figura 40. Regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) en la cohorte de sepsis. VIP score. La gráfica muestra en el eje de abscisas las variables y en el de coordenadas los valores de VIP score para cada una de ellas. Se especifican los genes que se corresponden con las 23 variables con VIP score superior a 5: * No hay gen asignado, CCDC72 (*coiled-coil domain containing 72*), UHMK1 (*U2AF homology motif (UHM) kinase 1*), DCAF6 (*DDB1 and CUL4 associated factor 6*), TMEM9B (*TMEM9 domain family, member B*), RAB5B (*RAB5B, member RAS oncogene family*), TEP1 (*Telomerase-associated protein 1*), CSK (*c-src tyrosine kinase*), CASC3 (*cancer susceptibility candidate 3*), MYO9B (*myosin IXB*), CD37 (*CD37 molecule*), MOBKL2A (*MOB1, Mps One Binder kinase activator-like 2A (levadura)*), PPPDE2 (*PPPDE peptidase domain containing 2*), MAN2B2 (*mannosidase, alpha, class 2B, member 2*), MED28 (*mediator complex subunit 28*), SCOC (*short coiled-coil protein*), NDUFA2 (*NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 2, 8kDa*), MGAT1 (*mannosyl (alpha-1,3-)-glycoprotein beta-1,2-N-acetylglucosamin*), AUH (*AU RNA binding protein/enoyl-CoA hydratase*), SH2D3C (*SH2 domain containing 3C*), ATG4A (*ATG4 autophagy related 4 homologue*), VBP1 (*von Hippel-Lindau binding protein 1*).

RESULTADOS

Tabla 21. Genes con diferencias de expresión significativas entre las sepsis por bacterias Gram positivas y bacterias Gram negativas, obtenidos con el modelo predictivo PLS-DA. ID representa el símbolo de identificación del gen y CE el valor de cambio de expresión.

ID GEN	NOMBRE	P-VALOR	CE
CCDC72	<i>Coiled-coil domain containing 72</i>		
UHMK1	<i>U2AF homology motif (UHM) kinase 1</i>	1.09×10^{-3}	-218
DCAF6	<i>DDB1 and CUL4 associated factor 6</i>	6.55×10^{-2}	-207
RAB5B	<i>RAB5B, member RAS oncogene family</i>		
TMEM9B	<i>TMEM9 domain family, member B</i>	7.05×10^{-3}	-240
TEP1	<i>Telomerase-associated protein 1</i>	1.09×10^{-6}	220
CSK	<i>C-src tyrosine kinase</i>	1.96×10^{-3}	180
CASC3	<i>Cancer susceptibility candidate 3</i>	5.69×10^{-1}	142
CD37	<i>Molécula CD37</i>		
MYO9B	<i>Myosin IXB</i>		
MOBKL2A	<i>MOB1, Mps One Binder kinase activator-like 2A (yeast)</i>	3.36×10^{-5}	184
PPPDE2	<i>PPPDE peptidase domain containing 2</i>	8.42×10^{-4}	-200
MAN2B2	<i>Mannosidase, alpha, class 2B, member 2</i>		
MED28	<i>Mediator complex subunit 28</i>	6.42×10^{-4}	-205
SCOC	<i>Short coiled-coil protein</i>	1.47×10^{-7}	-184
VBPI	<i>Von Hippel-Lindau binding protein 1</i>	4.15×10^{-1}	-231
NDUFA2	<i>NADH dehydrogenase (ubiquinone) 1 alpha subcomplex</i>	5.93×10^{-3}	-253
AUH	<i>AU RNA binding protein/enoyl-CoA hydratase</i>	8.05×10^{-5}	-227

ID GEN	NOMBRE	P-VALOR	CE
MGAT1	<i>Mannosyl (alpha-1,3)-glycoprotein beta-1,2-N-acetylglucosamin</i>	4.19×10^{-4}	160
SH2D3C	<i>SH2 domain containing 3C</i>		
ATG4A	<i>autophagy related 4 homologue ATG4</i>		

A partir de estas 23 variables se creó el modelo predictivo definitivo (figura 41).

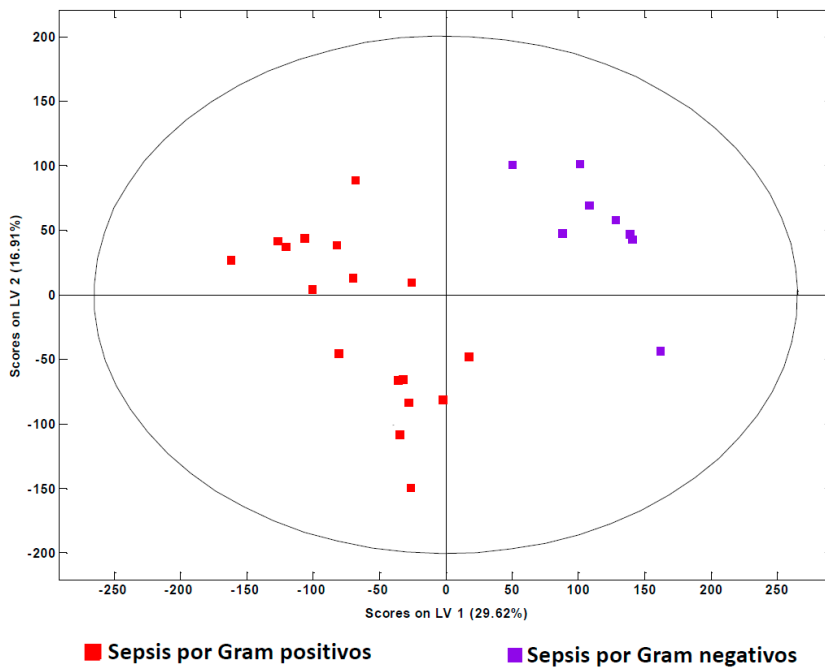


Figura 41. Regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) en la cohorte de sepsis. La gráfica muestra la distribución espacial de las sepsis por bacterias Gram positivas (en cuadrados rojos) que quedan agrupadas en la zona de la izquierda, y las sepsis por bacterias Gram negativas (en cuadrados morados) que se agrupan en el lado derecha de la gráfica.

RESULTADOS

Se trata de un modelo con un AUC que seguía siendo 1, un error de clasificación del 0% y unas sensibilidad de 100% y especificidad de 100% (figura 42).

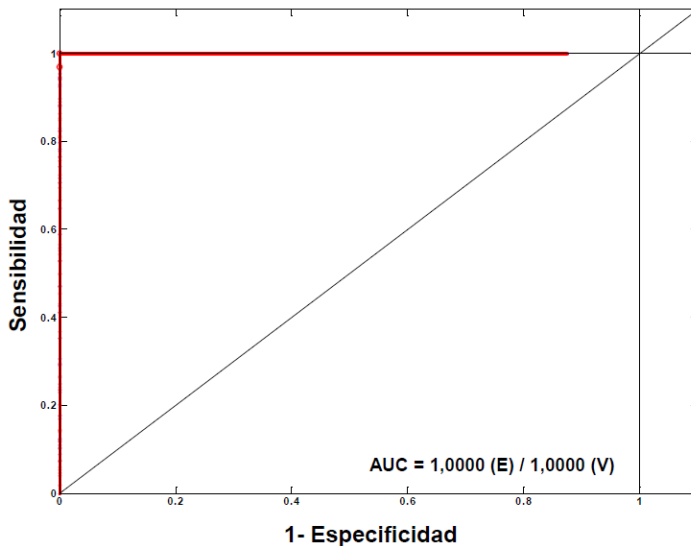


Figura 42. Área bajo la curva (AUC) Característica Operativa del Receptor (ROC) en el modelo predictivo según el agente causal de la sepsis. La gráfica muestra en el eje de abscisas 1-Especificidad y en el de ordenadas la Sensibilidad. La curva Característica Operativa del Receptor (ROC) en rojo es la curva validada (V) con el modelo, que en este caso al dejar un AUC de 1 se superpone con la curva azul que es la curva estimada (E).

Validación del modelo: test de múltiples permutaciones con doble validación cruzada

Dado que los hallazgos de las diferencias entre bacterias Gram positivas y Gram negativas no estaban previstos en el diseño del estudio, no se dispuso de una cohorte de validación independiente.

Para validar los resultados del estudio diferencial según el tipo de

sepsis se utilizó un método de doble validación cruzada utilizando un test de múltiples permutaciones aleatorias¹⁰³⁻¹⁰⁴. El test de permutaciones permite evaluar si la clasificación de los individuos en los 2 grupos designados es significativamente mejor que otra clasificación aleatoria en 2 grupos arbitrarios. Consiste en cambiar de manera aleatoria la asignación de las muestras manteniendo la proporción original en cada uno de los grupos. Estas permutaciones se repitieron 1000 veces utilizando todas las variables.

La representación gráfica del modelo se hace utilizando los siguientes parámetros en la clase permutada: número de clasificaciones equivocadas (NMC: *Number of misclassifications*), área bajo la curva (AUC: *Area under curve*) y el error de predicción del modelo expresada en Q^2 , en función del número de veces que se repite cada uno de ellos (Fx).

Se define DQ^2 con doble validación cruzada (DQ^2_{2cv}) como 1 menos la ratio de la suma de cuadrados del error de predicción, sobre el total de la suma de cuadrados de la respuesta del vector y .

RESULTADOS

$$Q^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2}$$

Donde \hat{y}_i representa la predicción real en la que la muestra i no se usó para la construcción del modelo.

Como el denominador es una constante, el numerador debería cuantificar la exactitud con la que se clasifican las muestras ¹⁰⁵.

La figura 43 muestra la distribución del número de clasificaciones equivocadas en la clase permutada (NMC: *Number of misclassifications*) en función del número de veces que se repite cada clasificación errónea. El modelo ideal debería mostrar valores cercanos a 0 mientras que el test de permutaciones, al ser un modelo aleatorio, se debería equivocar la mitad de las veces; es decir, en nuestro caso que contamos con 25 muestras, la distribución debería agruparse en torno a 12.5. El número de veces que se repite cada NMC, representado por $F(x)$, depende del número de permutaciones utilizadas en el modelo. Nuestro modelo a validar queda representado por un punto rojo, que en ningún momento se solapa con el histograma de la clase permutada, lo que significa que el modelo con la clase real se diferencia de un modelo con una distribución aleatoria

182

de las clases ($p < 0.001$). Cada muestra se testó 5 veces (figura 44) y se evaluó la clasificación del vector y . En únicamente una muestra hubo error de clasificación en 4 de los 5 test (0.8/25), por tanto hablamos de un error de clasificación del 3%.

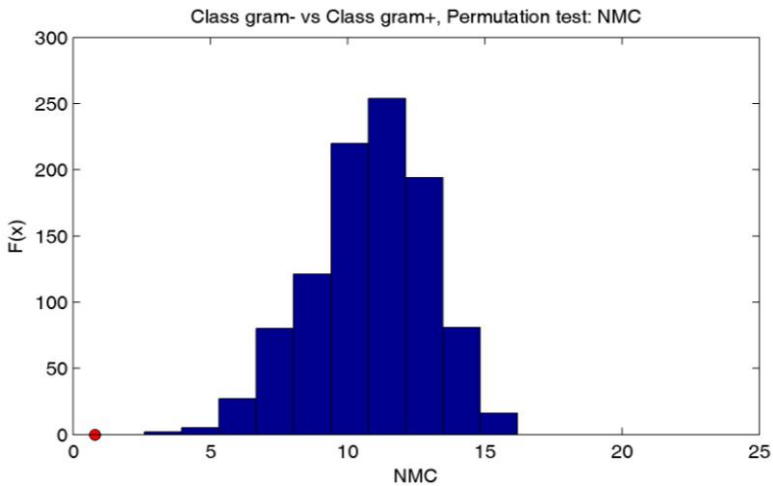


Figura 43. Test de validación cruzada: *Number of misclassifications* (NMC). La gráfica muestra la distribución de la clase permutada frente al resultado de la clase real. En el eje de abscisas el número de clasificaciones equivocadas (NMC: *Number of misclassifications*) y en el de ordenadas $F(x)$, que indica el número de veces que se repite cada NMC. El punto rojo representa el valor de nuestro modelo a validar.

RESULTADOS

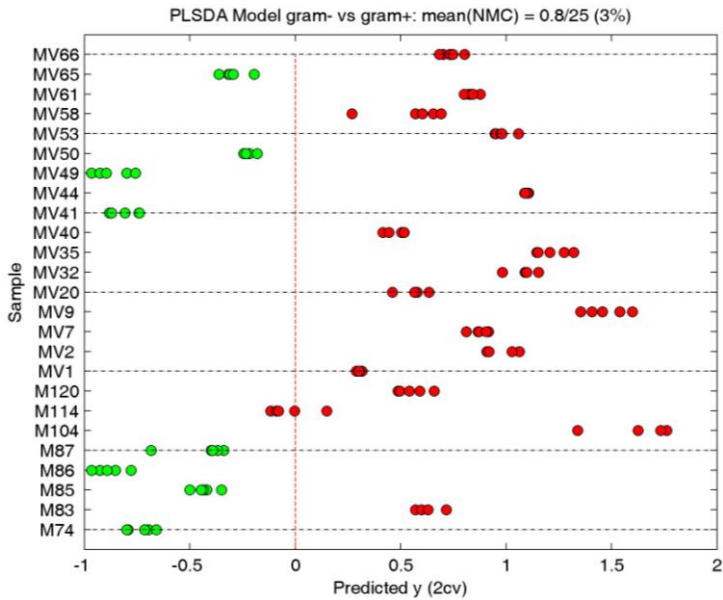


Figura 44. Test de validación cruzada. La gráfica muestra la distribución de la clase real tras aplicar doble validación cruzada (2 CV). En el eje de abscisas está representado el valor del vector y para cada una de los 5 test que se hicieron en cada muestra. En el eje de ordenadas se representa cada una de las muestras. Las sepsis por bacterias Gram positivas están representadas en rojo y las bacterias Gram negativas en verde.

La figura 45 muestra el área bajo la curva (AUC) en función del número de veces que se repite cada valor de AUC. El modelo ideal debería tener un valor cercano a 1 (que implicaría sensibilidades y especificidades cercanas al 100%), mientras que el aleatorio debería distribuirse en torno a 0.5.

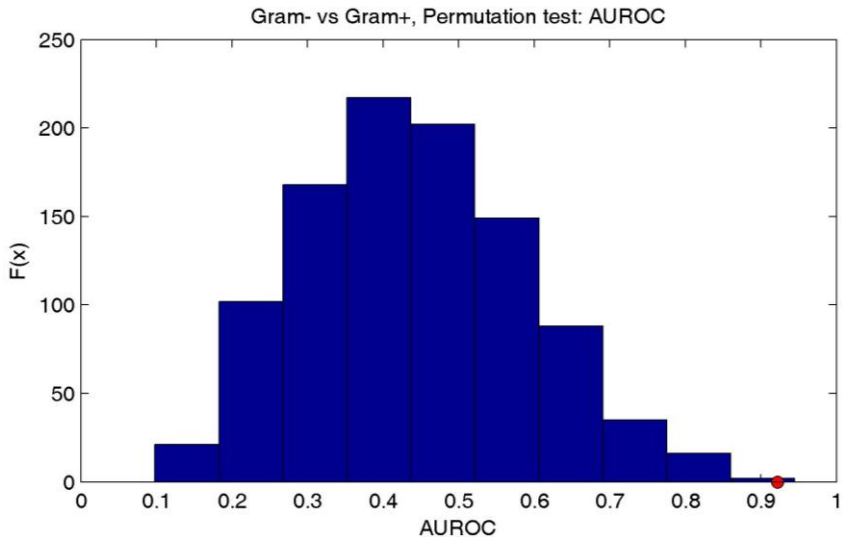


Figura 45. Test de validación cruzada: área bajo la curva ROC (AUROC). La gráfica muestra la distribución de la clase permutada frente al AUROC de la clase real. En el eje de abscisas el AUROC y en el de ordenadas $F(x)$, que indica el número de veces que se repite cada AUROC. El punto rojo representa el valor de nuestro modelo a validar.

El valor de nuestro modelo, que en este caso aparece como 0.92 ya que en la validación lo calcula con el modelo global de las 33.297 variables, se solapa ligeramente con la gráfica de las clases permutadas, dando un p-valor de 0.002. Es decir, la predicción que ofrece nuestro modelo no es fruto del azar, ya que se diferencia de la distribución aleatoria.

RESULTADOS

Por último, la figura 46 muestra el error Q^2 2cv en función del número de veces que se repite cada valor de Q^2 2cv. Al tratarse de un error el modelo ideal debería mostrar un error de 0 a 1, mientras que el modelo aleatorio se debe distribuir en torno a -0.5, como ocurre en nuestro caso con un p-valor < 0.001 .

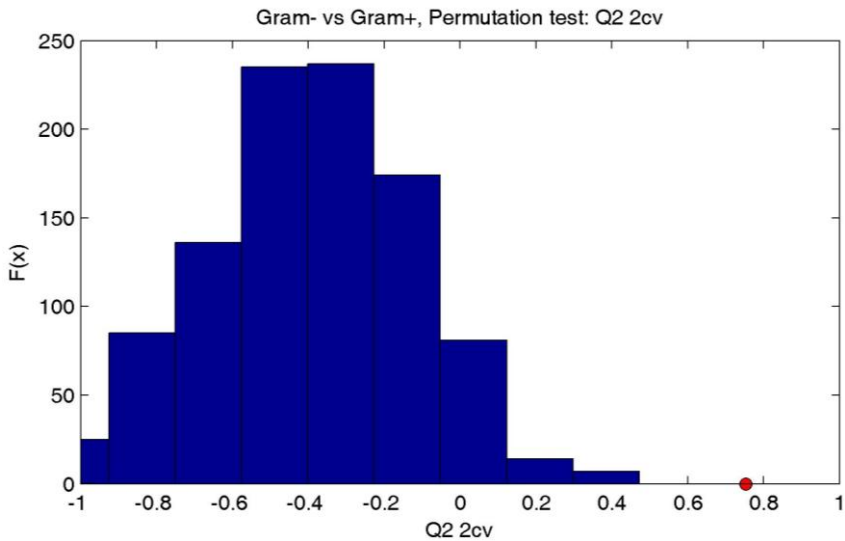


Figura 46. Test de validación cruzada: Q^2 2cv. La gráfica muestra la distribución de la clase permutada frente al resultado obtenido utilizando la clase real. En el eje de abscisas el error Q^2 2cv y en el de ordenadas $F(x)$, que indica el número de veces que se repite cada Q^2 2cv. El punto rojo representa el valor de nuestro modelo a validar.

IV.5.4. GENES SIGNIFICATIVOS, REGULADORES Y RUTAS DE SEÑALIZACIÓN

Para entender el significado biológico de los 719 genes significativos entre las sepsis por bacterias Gram positivas comparadas con las Gram negativas analizamos los procesos biológicos más relevantes. En la tabla 22 se representan las rutas de señalización.

Tabla 22. Rutas de señalización más significativas en las sepsis por bacterias Gram positivas frente a las Gram negativas (en términos de p-valor y en número de genes) de los 719 genes significativos usando Pathway Studio. Genes representa el número de genes implicados en cada regulador o ruta de señalización.

RUTAS DE SEÑALIZACIÓN	P-VALOR	GENES
Procesos metabólicos	8.90×10^{-12}	84
Señal del receptor del factor de crecimiento nervioso	3.97×10^{-9}	19
Apoptosis	1.37×10^{-8}	36
Transporte	2.77×10^{-8}	61
Señal mediada por citoquinas	3.10×10^{-8}	18
Señal de la transducción intracelular	8.40×10^{-8}	23
Respuesta inmune	1.08×10^{-7}	27
Regulación de la actina del citoesqueleto	4.64×10^{-7}	38
Respuesta a lipopolisacáridos	5.34×10^{-7}	15
Señal de la transducción intracelular	4.16×10^{-6}	91

RESULTADOS

En la tabla 23 se muestran los reguladores más significativos en términos de p-valor y en número de genes explicados.

Tabla 23. Reguladores más significativos en las sepsis por bacterias Gram positivas frente a las Gram negativas (en términos de p-valor y en número de genes) de los 719 genes significativos (resultantes de la comparación entre sepsis y controles) usando Pathway Studio. Genes representa el número de genes implicados en cada regulador o ruta de señalización, perteneciente a la lista de los 719 genes significativos: SPI1: Proteína específica 1, IFNG: Interferón gamma, IL3: Interleuquina 3, IL2: Interleuquina 2, TNF: Factor de necrosis tumoral, IL4: Interleuquina 4, CSF2: Factor 2 estimulante de colonias granulocitos-macrófagos, MAP2K1: Proteína quinasa activada por mitógenos 1, MIR1-1: microRNA 1-1.

REGULADORES	P-VALOR	GENES
Complejo proteosómico de la endopeptidasa	0.0006	43
SPI1	0.0033	15
IFNG	0.0064	48
IL2	0.0091	19
IL3	0.0091	13
TNF	0.0115	61
IL4	0.0216	28
CSF2	0.0267	18
MAP2K1	0.0309	12
MIR1-1	0.0383	10

**V. DISCUSIÓN Y
DESARROLLO
ARGUMENTAL**

La sepsis neonatal continúa siendo una importante causa de morbi-mortalidad en el mundo. Los neonatos y especialmente los RNMBP tienen un sistema inmune funcionalmente inmaduro, que condiciona tanto la presentación clínica como la respuesta a la sepsis. En ocasiones, las manifestaciones clínicas son inespecíficas o extremadamente sutiles; no obstante, puede presentar una evolución fulminante¹⁰⁶. El diagnóstico rápido y fiable es esencial para mejorar las tasas de supervivencia; sin embargo, los resultados del hemocultivo son tardíos y los parámetros infecciosos clásicos en ocasiones ofrecen resultados son ambiguos y variables.

Se han publicado distintos estudios, la mayoría unicéntricos o retrospectivos, que han validado el uso de biomarcadores como herramientas diagnósticas de sepsis. Estos estudios son difícilmente valorables ya que tanto los criterios diagnósticos de sepsis como los puntos de corte de los marcadores, varían de unos a otros¹⁰⁷⁻¹⁰⁸. Ante la inconsistencia de estos resultados, diferentes autores insisten en la necesidad de encontrar marcadores específicos de sepsis¹⁰⁹⁻¹¹⁰.

El estudio de los perfiles de expresión genómica utilizando biochips de expresión, ofrece una visión general de una enfermedad tan

DISCUSIÓN Y DESARROLLO ARGUMENTAL

compleja y heterogénea como las sepsis, con la ventaja de disminuir sesgos y aumentar la capacidad de hallar marcadores, ya que todos los genes son potencialmente evaluados. Este trabajo constituye el primer estudio de los perfiles de expresión genómica global realizado exclusivamente en RNMBP con sepsis bacteriana con el fin de evaluar la utilidad de estos perfiles como herramienta diagnóstica¹¹¹.

Diseño e inclusión de pacientes

Los estudios transcriptómicos de sepsis en pacientes pediátricos presentan un tamaño muestral reducido, ya que las diferencias entre grupos son tan elevadas que se pueden poner de manifiesto con pocos pacientes. Por otra parte, el acceso a una metodología tan novedosa es limitado y la necesidad de laboratorios especializados encarece las determinaciones. Nuestro estudio es hasta la fecha el que más número de RNMBP incluye, con un total de 67 pacientes: 54 repartidos en dos cohortes: inicial y validación, y 13 en un grupo paralelo de sepsis clínicas.

A pesar de que los pacientes incluidos en el estudio representa un grupo heterogéneo de RNMBP con distintos espectros clínicos, enfermedades concomitantes intrínsecas a la prematuridad y sepsis por

diferentes patógenos, hemos sido capaces de demostrar la existencia de perfiles de expresión genómica relacionados con procesos biológicos y funciones moleculares biológicamente relevantes¹¹¹. Hay que resaltar que los controles fueron estrictamente seleccionados en función de características clínicas y demográficas para que fueran comparables con los pacientes con sepsis, siendo en 3 ocasiones los hermanos gemelos de los casos. Esta estrategia, unida al estudio discriminativo con PLS-DA nos permite atribuir a la infección, y no a otros factores o enfermedades concomitantes, las diferencias encontradas en el estudio transcriptómico. Las muestras de sangre se extrajeron en cuanto se observaron los signos clínicos y previamente al inicio de los antibióticos, con el fin de evitar cambios en la respuesta inmune relacionados con la medicación.

Se trata de un estudio prospectivo que permitió la recogida de la muestra en las mismas condiciones, previo al inicio del tratamiento antibiótico. Esto generó un porcentaje de pérdidas de cerca del 20% puesto que una vez iniciado el antibiótico las muestras ya no se utilizaron.

DISCUSIÓN Y DESARROLLO ARGUMENTAL

La falta de consenso internacional en la definición de sepsis neonatal nos llevó a diseñar unos criterios diagnósticos bastante restrictivos que incluyeran no solamente manifestaciones clínicas, sino también factores de riesgo y aislamiento de un microorganismo en el hemocultivo, tal y como recomienda el CDC ¹¹². De esta manera sólo se incluyeron las sepsis con diagnóstico microbiológico confirmado para establecer las diferencias con los controles, eliminando las sepsis clínicas de las cuales no podíamos estar seguros que fueran realmente sepsis.

Método

Uno de los resultados más novedosos de este trabajo es la obtención de PEG utilizando ARN extraído a partir de 0.5 ml de sangre total. En nuestro caso utilizamos solución estabilizadora de ARN en Tubos Tempus[®] que facilitaban la conservación de la muestra a temperatura ambiente. Esto es de vital importancia en un estudio clínico puesto que la sospecha de sepsis puede ocurrir en cualquier momento y los clínicos no estar disponibles para realizar el procesamiento de la muestra. Sin embargo, los kits comerciales de extracción se basaban en la obtención de 3 ml de sangre total a diluir en los 6 ml de solución estabilizadora. Esta cantidad de sangre es inviable dada la escasa

194

volemia de estos recién nacidos y tuvimos que optimizar el método de extracción para no extraer más de 0.5 ml de sangre total. Los escasos estudios que han utilizado ARN extraído de sangre total en pacientes neonatales, formaban parte de una población global de pacientes pediátricos y en ningún momento hacen referencia a la cantidad de sangre extraída en este subgrupo. Muy recientemente se ha publicado un estudio de perfiles de expresión genómica realizado exclusivamente en recién nacidos en el que refieren utilizar 0.4-0.5 ml de sangre total, obteniendo valores de RIN similares a los obtenidos en las muestras del presente trabajo, pero que fueron procesadas inmediatamente tras la extracción¹¹³.

Algunos autores promulgan el uso de células mononucleares de la sangre periférica (PBMCs)¹¹⁴ para la obtención de ARN, en lugar de muestras de sangre total con el fin de caracterizar mejor las diferencias en las poblaciones de la serie blanca¹¹⁵. Sin embargo, los perfiles de expresión a partir de sangre total tienen el potencial de aportar una visión más general sin perder información transcriptómica biológicamente relevante procedente de células no incluidas en los PBMCs, como los neutrófilos que juegan un papel tan importante en la sepsis. El ARN se extrajo de sangre total, lo que tiene la ventaja de

DISCUSIÓN Y DESARROLLO ARGUMENTAL

una estabilización rápida de los ácidos nucleicos, facilita su uso en estudios clínicos y su uso ha sido ampliamente validado con un potencial futuro en el diagnóstico rápido¹¹⁶⁻¹¹⁹.

Las muestras se procesaron para evitar sesgos de manera conjunta por la misma persona en 2 fases, la primera para la cohorte inicial y la segunda para la cohorte de validación.

Utilidad de los perfiles de expresión genómica

La identificación de los diferentes perfiles de expresión genómica del huésped frente a la infección, ha demostrado ofrecer nuevos marcadores no sólo diagnósticos sino también pronósticos y de gravedad de la enfermedad en niños y adultos con shock séptico¹¹⁷⁻¹²⁰. Estos estudios, de forma global muestran patrones que evidencian una sobreexpresión de los genes que codifican PRRs con una activación de las rutas de señalización no sólo de reconocimiento del patógeno sino también de la transcripción: NF- κ B, MAPK y la proteína activadora de la transcripción (STAT)¹²⁰. En uno de los estudios de perfiles de expresión genómica como herramientas de diagnóstico de infección en población pediátrica más ambiciosos y mejor diseñados, el Dr. Ramilo y sus colaboradores identificaron 137 genes que permiten discriminar

196

entre niños con neumonías bacterianas y víricas, e incluso entre diferentes tipos de bacterias, con un 95% de especificidad¹²¹ (figura 47), validando los resultados con PCR cuantitativa.

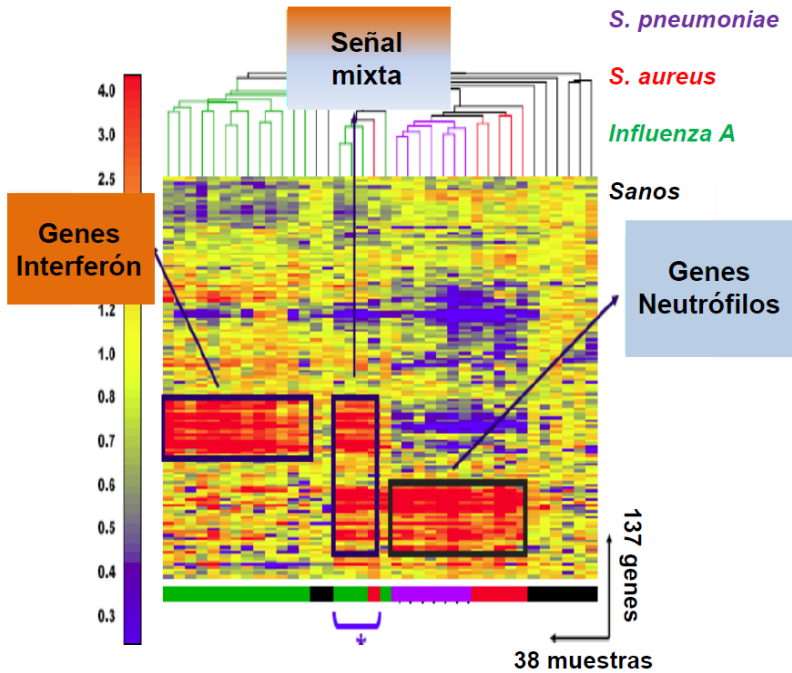


Figura 47. Perfiles de expresión genética en pacientes pediátricos con neumonía¹²¹ (reproducida con permiso). Cada columna representa una muestra de un paciente y cada línea un gen. Se identificaron cuatro perfiles de expresión: controles sanos (negro) se distinguen fácilmente de los niños con infecciones agudas. Pacientes con infección grave por Influenza A (verde) que mostró un aumento de la expresión de genes de interferon-inducible, infecciones bacterianas por *S.aureus* (rojo) y *S.pneumoniae* (violeta), que mostraron sobreexpresión de genes asociados a los neutrófilos. Tres muestras del grupo de Influenza A y una del de *S.aureus*, presentaron un cuarto perfil (marcado con *) que combina elementos de los previos sugiriendo la posibilidad de una co-infección por patógenos víricos y bacterianos.

DISCUSIÓN Y DESARROLLO ARGUMENTAL

Pese a su enorme potencial diagnóstico, hasta el momento no se habían realizado estudios de este tipo exclusivamente en neonatos prematuros de muy bajo peso. La inmadurez del sistema inmune de los RNMBP indudablemente influye en la respuesta transcriptómica a la infección, lo cual podría parcialmente explicar las diferencias observadas en la cohorte de prematuros con sepsis incluida en el presente estudio en comparación con las publicadas en adultos y niños más mayores¹²². Se han realizado estudios de perfiles de expresión genómica in vitro que evalúan la respuesta inmune tras la inyección de LPS en células mononucleares de cordón umbilical frente a sangre periférica de adultos. Se encontraron diferencias de expresión en más de 100 genes además de 5 genes que mostraron una interacción significativa entre el tipo celular y la estimulación con LPS¹²³. Esto podría explicar las diferencias en clínica tan agresiva y fulminante de la sepsis por bacterias Gram negativas en neonatos respecto a pacientes adultos.

Por todo esto, no debemos extrapolar los datos obtenidos en adultos y en los niños mayores a los RNMBP, ya que se trata de poblaciones con una respuesta inmune completamente distinta.

Genes y rutas que distinguen sepsis de controles

En este estudio se observan diferencias entre los perfiles de expresión genómica de RNMBP con sepsis comparados con controles no infectados¹¹¹. Es más, hemos encontrado funciones moleculares y procesos biológicos con importantes diferencias de expresión entre los neonatos infectados por bacterias Gram positivas y los infectados por Gram negativos. La sobreexpresión de las vías del TNF condiciona una expresión diferencial de citoquinas^{44, 118, 124}, que ha demostrado tener un papel importante en la sepsis.

Sin embargo, las diferencias en cuanto a las rutas inflamatorias parecen menos consistentes. Mientras algunos estudios, principalmente los realizados en población pediátrica, muestran una sobreexpresión precoz y transitoria de los mediadores pro-inflamatorios o patrones mixtos con mediadores pro y antiinflamatorios^{120, 125}, otros estudios evidencian un predominio de la sobreexpresión de genes antiinflamatorios^{114, 118}.

En nuestro estudio los procesos biológicos más significativamente expresados en pacientes sépticos son los relacionados con la repuesta inmune innata y la respuesta inflamatoria y en cuanto a las funciones

DISCUSIÓN Y DESARROLLO ARGUMENTAL

moleculares, el receptor de las células T y las citoquinas, de acuerdo con datos publicados en otras poblaciones^{120, 122}. En estos estudios, recién nacidos a término con sepsis mostraron una menor expresión de receptor de las células B y de los genes relacionados con TREM y NF- κ B en comparación con lactantes y niños mayores con shock séptico. En nuestro trabajo, el receptor de las células T también se encontró infraexpresado, mientras que el NF- κ B estaba sobreexpresado. Estas diferencias podrían explicarse parcialmente por la baja edad gestacional de nuestros pacientes, y por el hecho de que hemos comparado RNMBP con sepsis con RNMBP controles no infectados, mientras que los autores del estudio previo usaron para las comparaciones grupos de niños con distintas edades con shock séptico.

Los genes con mayor valor predictivo y con una mayor significación desde un punto de vista estadístico y biológico fueron: MMP8, CD177, OLFM4, GSTM1, HP, LCN2, CEACAM1, GPR84, ANKRD22. Estos resultados están en consonancia con los hallazgos de Wynn en una cohorte de recién nacidos (a término y pretérmino) con sospecha de sepsis, que muestra entre los 10 genes más

significativamente sobre-expresados 4 de los genes descritos en el presente trabajo: MMP8, OLFM4, ANKRD22, GPR84¹¹³.

MMP8: Matrix metalloproteinase 8 (neutrophil collagenase)

La familia de las metalopeptidasas de la matriz (MMP) está implicada en procesos de degradación de la matriz extracelular y remodelamiento de un importante número de tejidos, y se asocian a patologías como la artritis o las metástasis. La mayor parte de las MMP son secretadas como proproteínas inactivas y son activadas al ser adheridas por proteínas extracelulares. En el caso de la MMP8, también conocida como colagenasa-2, el enzima que codificado se almacena en gránulos en los neutrófilos y es activado por escisión autolítica. Se encarga de degradar los colágenos tipo I, II y III, α_2 macroglobulina, α_1 inhibidor de la proteinasa, Clq, fibrinógeno y sustancia P¹²⁶, activando a los PMN para iniciar la respuesta inflamatoria inicial y promover la respuesta tisular¹²⁷.

Se ha demostrado que las MMP's están reguladas al alza durante las infecciones, especialmente en aquellas causadas por gérmenes Gram negativos a través del estímulo de LPS y de las citoquinas proinflamatorias, que producen un aumento en la expresión de

DISCUSIÓN Y DESARROLLO ARGUMENTAL

MMP8, lo que confirma su importante papel en la patogénesis de la endotoxemia.

La MMP8 fue uno de los genes con mayor sobreexpresión en los RNMBP con sepsis bacteriana. Estos datos están en concordancia con estudios previos, en niños con shock séptico^{120, 125}, y en adultos con neumonía¹²⁸. Es más, el uso de inhibidores de las MMP's, tanto genéticos como farmacológicos, confiere protección contra el shock séptico in vitro y en modelos animales¹²⁹, aumentando significativamente las tasas de supervivencia¹³⁰ y pudiendo constituir en un futuro cercano una relevante alternativa terapéutica.

CD 177: CD177 molecule

El gen más significativamente sobreexpresado en RNMBP con sepsis comparados con los controles fue el CD177, también conocido como HNA-2a y NB1. Se trata de una glicoproteína de superficie celular unida a glicosil-fosfatidil-inositol (GPI) que se expresa selectivamente en los neutrófilos. Fue descrita por primera vez en un caso de neutropenia neonatal aloinmune¹³¹. Posteriormente se ha identificado en otros trastornos hematológicos como trombopenias o enfermedades mieloproliferativas¹³²⁻¹³⁵. Su expresión se ha visto

202

aumentada en individuos con infecciones bacterianas severas¹³⁶; sin embargo, esta es la primera vez que se describe como biomarcador de sepsis neonatal. Aunque no se conoce su función exacta se ha identificado su unión a moléculas de adhesión celular como *platelet endothelial cell adhesion molecule-1 (PEACAM-1)*, participando en la migración de neutrófilos con una función granulocito específica, contribuyendo con su sobreexpresión a la activación de los neutrófilos y por tanto a la respuesta inmune innata¹³⁷.

OLFM4: Olfactomedin 4

Este gen se clonó originariamente a partir de mieloblastos humanos y se expresa selectivamente en el epitelio del colon inflamado. La proteína que codifica es una glucoproteína de la matriz extracelular que facilita la adhesión celular. Su función exacta es desconocida. Algunos trabajos indican que actúa como factor antiapoptótico promoviendo el crecimiento tumoral¹³⁸. Mientras que en otros estudios parece actuar como un importante regulador negativo de la actividad bactericida de los neutrófilos a nivel intestinal, regulando la respuesta inmune innata mediante la inhibición de la producción de citoquinas y la actividad del NK-kB mediada por NOD1 y NOD2^{75, 139}.

GSTM1: Glutathione S-transferase mu 1

Este gen codifica una glutatión S-transferasa que pertenece a la clase mu, que es una de las 8 isoformas de glutatión S-transferasa que se han descrito. Se encarga de la detoxificación de componentes electrofílicos como carcinógenos, fármacos, toxinas ambientales y productos de estrés oxidativo mediante la conjugación con Glutathione. Variaciones genéticas en GSTM1 pueden condicionar susceptibilidades individuales a carcinógenos así como afectar la toxicidad o la eficacia de ciertos fármacos. Es más, mutaciones del GSTM1 que se traducen en ausencia de transcripción a ARN o de traducción a una proteína, se han relacionado con aumento de infecciones bacterianas severas en niños con leucemia linfoblástica aguda¹⁴⁰.

En estudios previos de nuestro grupo, hemos podido comprobar cómo en la transición fetal neonatal, especialmente si hay suplementación de oxígeno y generación de un importante estrés oxidativo, se produce una activación de la Glutathione-S-transferasa que facilita la reducción de GSSG a GSH¹⁴¹. Por lo tanto, existe una capacidad de respuesta de los enzimas antioxidantes y en concreto de GSH-S-Tr frente a la agresión oxidativa.

En el caso de infecciones, especialmente bacterianas, se generan especies reactivas (ROS y NOS) por parte de polimorfonucleares y macrófagos con liberación de peróxido de hidrógeno y ácido hipocloroso. En estas circunstancias pro-oxidantes se consumen importantes cantidades de GSH que se transforma en glutatión sulfonamida (GSA) ¹⁴². Hemos podido comprobar en nuestros estudios, que GSA se produce abundantemente en la neumonía asociada al respirador reflejando un aflujo de macrófagos al pulmón, que se evidencian en el análisis del fluido obtenido con el lavado broncoalveolar ¹⁴³. Por lo tanto, nuestros hallazgos demuestran que la activación de las enzimas antioxidantes capaces de reducir el GSSG a GSH y establecer el estado redox intracelular son una respuesta precoz a la infección y podrían utilizarse como marcador de septicemia en el prematuro con sepsis.

HP: Haptoglobin

Este gen codifica una pre-proteína, que se procesa para producir cadenas alfa y beta, que posteriormente se combinan como un tetrámero para producir haptoglobina. La haptoglobina es una proteína plasmática que se une a la hemoglobina libre en plasma y permite por una parte que los enzimas degradantes puedan acceder a ella, mientras

DISCUSIÓN Y DESARROLLO ARGUMENTAL

que a la vez previene la pérdida de hierro a través del riñón protegiéndolo del daño de la hemoglobina. Este gen se ha relacionado con la nefropatía diabética cardiopatía isquémica en la diabetes tipo 1, enfermedad de Crohn, colangitis esclerosante primaria o enfermedad de Parkinson. Actúa como reactante de fase aguda y mediador de la respuesta inflamatoria, ya que induce la producción de IL6 jugando un importante papel en el equilibrio entre las poblaciones de linfocitos Th₁ y Th₂¹⁴⁴.

LCN2: lipocalin 2

También se conoce como NGAL: *neutrophil gelatinase-associated lipocalin*. Los ligandos de la LCN2 son los sideróforos férricos bacterianos. En ausencia de hierro soluble, los microorganismos generan sideróforos, que constituyen los quelantes de hierro más fuertes que se conocen y permiten a las bacterias absorber el hierro que precisan para su crecimiento. La unión de los sideróforos a la LCN2 priva a las bacterias de hierro, de manera que actúa como una proteína bacteriostática¹⁴⁵.

La expresión de LCN2 es inducida por LPS y TNF- α a través de la vía del NF- κ B tras el estímulo desencadenado por la producción de citoquinas proinflamatorias. Se almacena en gránulos específicos de

los neutrófilos y es excitada cuando éstos migran al lugar de la infección¹⁴⁶⁻¹⁴⁷. En modelos animales ha demostrado proteger frente a la infección por *E. coli*¹⁴⁸.

CEACAM1: carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)

Este gen codifica a un miembro de la familia del antígeno carcinoembrionario (CEA) que pertenece a la superfamilia de las inmunoglobulinas. La proteína que traduce fue descrita originariamente en los conductos biliares como una glicoproteína biliar. Actúa como molécula de adhesión celular en leucocitos, epitelio y endotelio. Participa en la angiogénesis, apoptosis, supresión tumoral, metástasis y en la modulación de la respuesta inmune innata y adaptativa. La activación de las células T por parte de citoquinas como IL2, IL7 e IL15 genera una sobreexpresión de CEACAM1 que persiste varios días. Esta unión del CEACAM1 a las células T induce posteriormente una cascada que conduce a la supresión de la producción y proliferación de células T por parte de las citoquinas¹⁴⁹. Esta regulación al alza del CEACAM1 en las células T medido por citometría de flujo se ha confirmado en el shock séptico en niños y en una pequeña muestra de RNMBP con sepsis¹⁵⁰.

GPR84: G protein-coupled receptor 84

Este gen codifica una superfamilia de 7 receptores transmembrana cuyas funciones y regulación no están bien establecidas. Su expresión parece inducirse por endotoxinas mediadas parcialmente por citoquinas proinflamatorias, TNF e IL1¹⁵¹.

ANKRD22: ankyrin repeat domain 22

Este gen codifica una proteína transmembrana sin función biológica conocida.

Modelo predictivo

Con el fin de trasladar nuestros resultados a la práctica clínica habitual, se generó un modelo predictivo utilizando un método estadístico supervisado conocido como regresión de mínimos cuadrados parciales con análisis discriminante (PLS-DA) que permite identificar las variables responsables de la discriminación entre las dos clases (sepsis y controles) y clasificar el estado de futuros pacientes a partir del análisis de muestras biológicas. El modelo obtenido presentó un bajo error de clasificación y un área bajo la curva ROC cercana al 1, con un número de genes, 9, relativamente pequeño como para poder ser utilizado como kit diagnóstico en la práctica clínica.

Validación

Hemos validado nuestros resultados utilizando 3 estrategias diferentes, coincidiendo las 3 aproximaciones en la consistencia de los mismos. Por una parte utilizamos el método de validación ideal utilizando una cohorte de pacientes independiente, con variables clínicas y demográficas similares a la cohorte inicial. Aplicando el modelo predictivo sobre esta cohorte, sólo en uno de los casos la clasificación fue dudosa, en el resto las muestras se agruparon perfectamente en su categoría. Por otra parte, en el estudio transcriptómico se utilizaron herramientas estadísticas muy restrictivas y estrategias como el FDR para corregir por las múltiples comparaciones. Finalmente, los genes del modelo predictivo con mayor significación estadística y funcional fueron validados por RT-PCR.

Estudio diferencial según el germen causal de la sepsis

Es más, los perfiles de expresión genómica han sido capaces de distinguir infecciones agudas causadas por virus, bacterias Gram positivas y Gram negativas^{121, 152-153}.

Nuestros datos muestran que la respuesta inmune del huésped basada en los perfiles de expresión genómica es diferente entre niños

DISCUSIÓN Y DESARROLLO ARGUMENTAL

infectados por bacterias Gram positivas y los infectados por bacterias Gram negativas; de acuerdo con los datos previamente publicados por Ramilo et al¹²¹, Feezor et al¹⁵² y Yu et al¹⁵³. Al contrario que Tang et al¹⁵⁴ que no encontraron diferencias significativas en adultos críticamente enfermos. El único estudio en recién nacidos que hace referencia al origen de la sepsis es el recientemente publicado por Wynn¹¹³, que encuentra diferencias en los perfiles de expresión genómica entre las sepsis precoces y las tardías. Hay que tener en cuenta que las sepsis precoces suelen ser de origen vertical, y por tanto producidas con más frecuencia por bacterias Gram negativos, y las tardías suelen ser nosocomiales producidas por bacterias Gram positivas. Sin embargo, los autores no mencionan en ningún momento los microorganismos causantes de las sepsis.

No obstante, hay que tener precaución en la interpretación de estos datos dada la escasa muestra de niños con sepsis por Gram negativos, y siempre teniendo en cuenta que este es un hallazgo secundario y que el objetivo principal del trabajo, y por tanto el diseño del estudio, no estaba orientado a profundizar en las diferencias de los perfiles de expresión genómica según el germen causal de la sepsis.

Estudio sepsis clínicas

Como parte de los objetivos secundarios se analizaron las sepsis clínicas; es decir aquellas en las que no se aisló ningún germen en el hemocultivo. Tal y como supusimos en el diseño del estudio, no todas las sepsis clínicas se comportaron desde un punto de vista transcriptómico como las microbiológicas. De hecho, únicamente la mitad de ellas se agruparon con las sepsis en el modelo predictivo, mientras que el resto fueron mucho más similares a los controles. Estas diferencias también se han observado en otros estudios, que a pesar de encontrar genes comunes en sepsis clínicas y confirmadas microbiológicamente (como la MMP8), observaron diferencias hasta en el 60% de los genes más significativos¹¹³.

Aplicabilidad y relevancia

Las septicemias y meningitis en el período neonatal y, especialmente, en el recién nacido prematuro, constituyen una de las causas de mortalidad y morbilidad más importantes en el momento actual. Es más, en las últimas décadas, apenas se han producido avances significativos en su diagnóstico y tratamiento. La utilización de esta nueva metodología propuesta y que nunca anteriormente había sido aplicada en neonatología, nos permite evaluar mejor la respuesta

DISCUSIÓN Y DESARROLLO ARGUMENTAL

inmune de los recién nacidos ante una agresión infecciosa. Utilizando los genes candidatos podríamos diseñar un test diagnóstico en un pequeño biochip, y así disponer de una herramienta extraordinariamente eficaz en el diagnóstico de la sepsis neonatal que diferenciara con mayor precisión aquellos neonatos infectados de los que no lo están con un tiempo de respuesta operativo. La consecuencia sería la utilización más precoz de combinaciones antibióticas más eficaces en aquellos pacientes infectados disminuyendo la morbilidad y mortalidad de los mismos. Asimismo, se evitaría el uso indiscriminado de antibióticos en aquellos pacientes que no estuviesen infectados evitando los efectos indeseables de los mismos (infecciones fúngicas; alteraciones de la microbiota) y el coste derivado de la realización de analíticas innecesarias y prolongación de la estancia hospitalaria.

En definitiva, el presente estudio aporta información sobre la utilidad y la aplicabilidad de los perfiles de expresión genómica en el diagnóstico de la sepsis bacteriana en los RNMBP. A partir de un enfoque traslacional y dirigido exclusivamente a la sepsis en RNMBP, nuestro estudio confirma y amplía los resultados obtenidos por otros grupos en pacientes adultos y pediátricos. Es más, la posibilidad de

212

utilizar un biochip que nos permitiera distinguir el tipo de germen causal de la sepsis, sería de una gran trascendencia clínica, puesto que podríamos indicar antibióticos con un espectro más dirigido y con ello mejorar el pronóstico y reducir las resistencias antibióticas. Esta aplicación de los resultados obtenidos precisa de más estudios para validar el uso clínico generalizado de los genes candidatos. Con todo esto, en un futuro esperamos que no muy lejano, podríamos mejorar considerablemente el diagnóstico y el manejo de la sepsis neonatal.

VI. CONCLUSIONES

1. Genome-wide expression profiles are described for the first time in very low birth weight infants with proven bacterial sepsis. Our results suggest that Genome-wide expression profiles could be used to early discriminate septic newborn infants from non-septic ones.
2. We were able to reduce to 0.5 mL the whole blood volume needed for Genome-wide expression profiles rendering it more suitable for the newborn period.
3. The most significant over-expressed genes in septic patients compared with controls were related to innate immune and inflammatory responses: MMP8, CD177, OLFM4, GSTM1, HP, LCN2, CEACAM1, GPR84, ANKRD22. Moreover, over-expressed TNF- α network included differentially expressed cytokines that play an important role in sepsis.
4. We found that the host immune response assessed using gene expression patterns may be different in neonates infected with Gram positive versus Gram negative bacteria.

CONCLUSIONES

5. Our study provides an initial insight to the applicability of Genome-wide expression profiles for the diagnosis of bacterial sepsis in preterm infants using a translational approach. Specific candidate genes could be used for the diagnosis and management of neonatal sepsis.

1. En este estudio se describen, por primera vez, los perfiles de expresión genómica en recién nacidos de muy bajo peso con sepsis bacterianas. Nuestros resultados sugieren que los perfiles transcriptómicos se podrían utilizar como marcadores precoces que pudieran distinguir a los recién nacidos infectados de los que o lo están.

2. Hemos sido capaces de reducir a 0,5 ml el volumen total de sangre necesario para obtener los perfiles de expresión genómica, haciéndolos más aptos para el uso en el periodo neonatal.

3. Los genes que estuvieron más significativamente sobreexpresados en los pacientes con sepsis comparados con los controles estuvieron relacionados con las respuestas inmune innata e inflamatoria: MMP8, CD177, OLFM4, GSTM1, HP, LCN2, CEACAM1, GPR84, ANKRD22. Es más, la sobreexpresión de la ruta del TNF- α incluyó distintas citoquinas que podrían jugar un importante papel en la sepsis.

4. La respuesta inmune del huésped según los perfiles de expresión genómica podría ser diferente en recién nacidos infectados por

CONCLUSIONES

bacterias Gram positivas con respecto a los infectados por bacterias Gram negativas.

5. Nuestro estudio proporciona una visión inicial de la aplicabilidad de los perfiles de expresión genómica en el diagnóstico de la sepsis bacteriana en los recién nacidos prematuros utilizando un abordaje traslacional. Los genes candidatos obtenidos podrían ser útiles para el diagnóstico y manejo de la sepsis neonatal.

VII. BIBLIOGRAFÍA

1. Larroque B, Ancel PY, Marret S, et al. Neurodevelopmental disabilities and special care of 5-year-old children born before 33 weeks of gestation (the EPIPAGE study): a longitudinal cohort study. *Lancet* 2008;371:813-20.
2. Vento M, Cheung PY, Aguar M. The first golden minutes of the extremely-low-gestational-age neonate: a gentle approach. *Neonatology* 2009;95:286-98.
3. Mitha A, Foix-L'Helias L, Arnaud C, et al. Neonatal infection and 5-year neurodevelopmental outcome of very preterm infants. *Pediatrics* 2013;132:e372-80.
4. Goldstein B, Giroir B, Randolph A. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr Crit Care Med* 2005;6:2-8.
5. Lopez Sastre JB, Coto Cotallo D, Fernandez Colomer B. Neonatal sepsis of nosocomial origin: an epidemiological study from the "Grupo de Hospitales Castrillo". *J Perinat Med* 2002;30:149-57.
6. Stoll BJ, Hansen N, Fanaroff AA, et al. Late-onset sepsis in very low birth weight neonates: the experience of the NICHD Neonatal Research Network. *Pediatrics* 2002;110:285-91.

BIBLIOGRAFÍA

7. Bizzarro MJ, Raskind C, Baltimore RS, Gallagher PG. Seventy-five years of neonatal sepsis at Yale: 1928-2003. *Pediatrics* 2005;116:595-602.
8. Lopez Sastre JB, Coto Cotallo GD, Fernandez Colomer B. Neonatal sepsis of vertical transmission: an epidemiological study from the "Grupo de Hospitales Castrillo". *J Perinat Med* 2000;28:309-15.
9. Verani JR, McGee L, Schrag SJ. Prevention of perinatal group B streptococcal disease--revised guidelines from CDC, 2010. *MMWR Recomm Rep* 2010;59:1-36.
10. Polin RA. Management of neonates with suspected or proven early-onset bacterial sepsis. *Pediatrics* 2012;129:1006-15.
11. Camacho-Gonzalez A, Spearman PW, Stoll BJ. Neonatal infectious diseases: evaluation of neonatal sepsis. *Pediatr Clin North Am* 2013;60:367-89.
12. Shah BA, Padbury JF. Neonatal sepsis: an old problem with new insights. *Virulence* 2014;5:170-8.
13. Payne NR, Carpenter JH, Badger GJ, Horbar JD, Rogowski J. Marginal increase in cost and excess length of stay associated with nosocomial bloodstream infections in surviving very low birth weight infants. *Pediatrics* 2004;114:348-55.

14. Towers CV, Carr MH, Padilla G, Asrat T. Potential consequences of widespread antepartal use of ampicillin. *Am J Obstet Gynecol* 1998;179:879-83.
15. De Cueto M, Sanchez MJ, Sampedro A, Miranda JA, Herruzo AJ, Rosa-Fraile M. Timing of intrapartum ampicillin and prevention of vertical transmission of group B streptococcus. *Obstet Gynecol* 1998;91:112-4.
16. Connell TG, Rele M, Cowley D, BATTERY JP, Curtis N. How reliable is a negative blood culture result? Volume of blood submitted for culture in routine practice in a children's hospital. *Pediatrics* 2007;119:891-6.
17. Hornik CP, Benjamin DK, Becker KC, et al. Use of the complete blood cell count in early-onset neonatal sepsis. *Pediatr Infect Dis J* 2012;31:799-802.
18. Collado MC, Cernada M, Neu J, Perez-Martinez G, Gormaz M, Vento M. Factors influencing gastrointestinal tract and microbiota immune interaction in preterm infants. *Pediatr Res* 2015;77:726-31.
19. Hornik CP, Fort P, Clark RH, et al. Early and late onset sepsis in very-low-birth-weight infants from a large group of neonatal intensive care units. *Early Hum Dev* 2012;88 Suppl 2:S69-74.

BIBLIOGRAFÍA

20. Adams-Chapman I. Long-term impact of infection on the preterm neonate. *Semin Perinatol* 2012;36:462-70.
21. Shah DK, Doyle LW, Anderson PJ, et al. Adverse neurodevelopment in preterm infants with postnatal sepsis or necrotizing enterocolitis is mediated by white matter abnormalities on magnetic resonance imaging at term. *J Pediatr* 2008;153:170-5.
22. Chau V, Brant R, Poskitt KJ, Tam EW, Synnes A, Miller SP. Postnatal infection is associated with widespread abnormalities of brain development in premature newborns. *Pediatr Res* 2012;71:274-9.
23. Van Vliet EO, de Kieviet JF, Oosterlaan J, van Elburg RM. Perinatal infections and neurodevelopmental outcome in very preterm and very low-birth-weight infants: a meta-analysis. *JAMA Pediatr* 2013;167:662-8.
24. Schlapbach LJ, Aebischer M, Adams M, et al. Impact of sepsis on neurodevelopmental outcome in a Swiss National Cohort of extremely premature infants. *Pediatrics* 2011;128:e348-57.
25. Alshaikh B, Yee W, Lodha A, Henderson E, Yusuf K, Sauve R. Coagulase-negative staphylococcus sepsis in preterm infants and long-term neurodevelopmental outcome. *J Perinatol* 2014;34:125-9.

26. Wynn JL, Wong HR. Pathophysiology and treatment of septic shock in neonates. *Clin Perinatol* 2010;37:439-79.
27. Kawai T, Akira S. The roles of TLRs, RLRs and NLRs in pathogen recognition. *Int Immunol* 2009;21:317-37.
28. Lotze MT, Tracey KJ. High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal. *Nat Rev Immunol* 2005;5:331-42.
29. Kono H, Rock KL. How dying cells alert the immune system to danger. *Nat Rev Immunol* 2008;8:279-89.
30. Ng PC, Li K, Wong RP, et al. Proinflammatory and anti-inflammatory cytokine responses in preterm infants with systemic infections. *Arch Dis Child Fetal Neonatal Ed* 2003;88:F209-13.
31. Sadeghi K, Berger A, Langgartner M, et al. Immaturity of infection control in preterm and term newborns is associated with impaired toll-like receptor signaling. *J Infect Dis* 2007;195:296-302.
32. Sullivan SE, Staba SL, Gersting JA, et al. Circulating concentrations of chemokines in cord blood, neonates, and adults. *Pediatr Res* 2002;51:653-7.
33. Hickey MJ, Kubes P. Intravascular immunity: the host-pathogen encounter in blood vessels. *Nat Rev Immunol* 2009;9:364-75.

BIBLIOGRAFÍA

34. Engle WA, McGuire WA, Schreiner RL, Yu PL. Neutrophil storage pool depletion in neonates with sepsis and neutropenia. *J Pediatr* 1988;113:747-9.
35. Koenig JM, Stegner JJ, Schmeck AC, Saxonhouse MA, Kenigsberg LE. Neonatal neutrophils with prolonged survival exhibit enhanced inflammatory and cytotoxic responsiveness. *Pediatr Res* 2005;57:424-9.
36. Linderkamp O, Ruef P, Brenner B, Gulbins E, Lang F. Passive deformability of mature, immature, and active neutrophils in healthy and septicemic neonates. *Pediatr Res* 1998;44:946-50.
37. Sriskandan S, Altmann DM. The immunology of sepsis. *J Pathol* 2008;214:211-23.
38. Opal SM, Esmon CT. Bench-to-bedside review: functional relationships between coagulation and the innate immune response and their respective roles in the pathogenesis of sepsis. *Crit Care* 2003;7:23-38.
39. Sikora JP, Chlebna-Sokol D, Krzyzanska-Oberbek A. Proinflammatory cytokines (IL-6, IL-8), cytokine inhibitors (IL-6sR, sTNFRII) and anti-inflammatory cytokines (IL-10, IL-13) in the pathogenesis of sepsis in newborns and infants. *Arch Immunol Ther Exp (Warsz)* 2001;49:399-404.

40. Drossou V, Kanakoudi F, Diamanti E, et al. Concentrations of main serum opsonins in early infancy. *Arch Dis Child Fetal Neonatal Ed* 1995;72:F172-5.
41. Mosmann TR. Cytokine secretion patterns and cross-regulation of T cell subsets. *Immunol Res* 1991;10:183-8.
42. Levy O. Innate immunity of the newborn: basic mechanisms and clinical correlates. *Nat Rev Immunol* 2007;7:379-90.
43. Kumar Y, Qunibi M, Neal TJ, Yoxall CW. Time to positivity of neonatal blood cultures. *Arch Dis Child Fetal Neonatal Ed* 2001;85:F182-6.
44. Ng PC. Diagnostic markers of infection in neonates. *Arch Dis Child Fetal Neonatal Ed* 2004;89:F229-35.
45. Berger C, Uehlinger J, Ghelfi D, Blau N, Fanconi S. Comparison of C-reactive protein and white blood cell count with differential in neonates at risk for septicaemia. *Eur J Pediatr* 1995;154:138-44.
46. Ng PC, Cheng SH, Chui KM, et al. Diagnosis of late onset neonatal sepsis with cytokines, adhesion molecule, and C-reactive protein in preterm very low birthweight infants. *Arch Dis Child Fetal Neonatal* 1997;77:F221-7.

BIBLIOGRAFÍA

47. Mathers NJ, Pohlandt F. Diagnostic audit of C-reactive protein in neonatal infection. *Eur J Pediatr* 1987;146:147-51.
48. Chiesa C, Pellegrini G, Panero A, et al. C-reactive protein, interleukin-6, and procalcitonin in the immediate postnatal period: influence of illness severity, risk status, antenatal and perinatal complications, and infection. *Clin Chem* 2003;49:60-8.
49. Blommendahl J, Janas M, Laine S, Miettinen A, Ashorn P. Comparison of procalcitonin with CRP and differential white blood cell count for diagnosis of culture-proven neonatal sepsis. *Scand J Infect Dis* 2002;34:620-2.
50. Chiesa C, Pacifico L, Rossi N, Panero A, Matrunola M, Mancuso G. Procalcitonin as a marker of nosocomial infections in the neonatal intensive care unit. *Intensive Care Med* 2000;26 Suppl 2:S175-7.
51. Dandona P, Nix D, Wilson MF, et al. Procalcitonin increase after endotoxin injection in normal subjects. *J Clin Endocrinol Metab* 1994;79:1605-8.
52. Auriti C, Fiscarelli E, Ronchetti MP, et al. Procalcitonin in detecting neonatal nosocomial sepsis. *Arch Dis Child Fetal Neonatal* 2012;97:F368-70.

53. Stocker M, Fontana M, El Helou S, Wegscheider K, Berger TM. Use of Procalcitonin-Guided Decision-Making to Shorten Antibiotic Therapy in Suspected Neonatal Early-Onset Sepsis: Prospective Randomized Intervention Trial. *Neonatology* 2010;97:165-74.
54. Sachse C, Dressler F, Henkel E. Increased serum procalcitonin in newborn infants without infection. *Clin Chem* 1998;44:1343-4.
55. Chiesa C, Panero A, Rossi N, et al. Reliability of procalcitonin concentrations for the diagnosis of sepsis in critically ill neonates. *Clin Infect Dis* 1998;26:664-72.
56. Assumma M, Signore F, Pacifico L, Rossi N, Osborn JF, Chiesa C. Serum procalcitonin concentrations in term delivering mothers and their healthy offspring: a longitudinal study. *Clin Chem* 2000;46:1583-7.
57. Krueger M, Nauck MS, Sang S, Hentschel R, Wieland H, Berner R. Cord blood levels of interleukin-6 and interleukin-8 for the immediate diagnosis of early-onset infection in premature infants. *Biol Neonate* 2001;80:118-23.
58. Procianoy RS, Silveira RC. The role of sample collection timing on interleukin-6 levels in early-onset neonatal sepsis. *J Pediatr (Rio J)* 2004;80:407-10.

BIBLIOGRAFÍA

59. Beceiro Mosquera J, Sivera Monzo CL, Oria de Rueda Salguero O, Olivas Lopez de Soria C, Herbozo Nory C. Usefulness of a rapid serum interleukin-6 test combined with C-reactive protein to predict sepsis in newborns with suspicion of infection. *An Pediatr (Barc)* 2009;71:483-8.
60. Cernada Badía M, Roqués Serradilla V, Vento Torres M. Interleukin-6 and diagnosis of neonatal sepsis: some clarifications. *An Pediatr (Barc)* 2010;73 (2): 104-5.
61. Cernada M, Badia N, Modesto V, et al. Cord blood interleukin-6 as a predictor of early-onset neonatal sepsis. *Acta Paediatr* 2012;101:e203-7.
62. Chiesa C, Pellegrini G, Panero A, et al. Umbilical cord interleukin-6 levels are elevated in term neonates with perinatal asphyxia. *Eur J Clin Invest* 2003;33:352-8.
63. Santana C, Guindeo MC, Gonzalez G, Garcia-Munoz F, Saavedra P, Domenech E. Cord blood levels of cytokines as predictors of early neonatal sepsis. *Acta Paediatr* 2001;90:1176-81.
64. Caldas JP, Marba ST, Blotta MH, Calil R, Morais SS, Oliveira RT. Accuracy of white blood cell count, C-reactive protein, interleukin-6 and tumor necrosis factor alpha for diagnosing late neonatal sepsis. *J Pediatr (Rio J)* 2008;84:536-42.

65. Rite Gracia S, Grasa Ullrich JM, Ruiz de la Cuesta Martin C, et al. Interleukin-6 and tumor necrosis factor-alpha as markers of vertically-transmitted neonatal bacterial infection. *An Pediatr (Barc)* 2003;59:246-51.
66. Martin H, Olander B, Norman M. Reactive hyperemia and interleukin 6, interleukin 8, and tumor necrosis factor-alpha in the diagnosis of early-onset neonatal sepsis. *Pediatrics* 2001;108:E61.
67. Barug D, Goorden S, Herruer M, Muller M, Brohet R, de Winter P. Reference values for interleukin-6 and interleukin-8 in cord blood of healthy term neonates and their association with stress-related perinatal factors. *PLoS One* 2014;9:e114109.
68. Kurt AN, Aygun AD, Godekmerdan A, Kurt A, Dogan Y, Yilmaz E. Serum IL-1beta, IL-6, IL-8, and TNF-alpha levels in early diagnosis and management of neonatal sepsis. *Mediators Inflamm* 2007;2007:31397.
69. Ng PC, Li K, Leung TF, et al. Early prediction of sepsis-induced disseminated intravascular coagulation with interleukin-10, interleukin-6, and RANTES in preterm infants. *Clin Chem* 2006;52:1181-9.
70. Sherwin C, Broadbent R, Young S, et al. Utility of interleukin-12 and interleukin-10 in comparison with other cytokines

BIBLIOGRAFÍA

and acute-phase reactants in the diagnosis of neonatal sepsis. *Am J Perinatol* 2008;25:629-36.

71. Ng PC, Li K, Chui KM, et al. IP-10 is an early diagnostic marker for identification of late-onset bacterial infection in preterm infants. *Pediatr Res* 2007;61:93-8.

72. Kingsmore SF, Kennedy N, Halliday HL, et al. Identification of diagnostic biomarkers for infection in premature neonates. *Mol Cell Proteomics* 2008;7:1863-75.

73. Zaki Mel S, el-Sayed H. Evaluation of microbiologic and hematologic parameters and E-selectin as early predictors for outcome of neonatal sepsis. *Arch Pathol Lab Med* 2009;133:1291-6.

74. Figueras-Aloy J, Gomez-Lopez L, Rodriguez-Miguel JM, et al. Serum soluble ICAM-1, VCAM-1, L-selectin, and P-selectin levels as markers of infection and their relation to clinical severity in neonatal sepsis. *Am J Perinatol* 2007;24:331-8.

75. Liu W, Yan M, Liu Y, McLeish KR, Coleman WG, Jr., Rodgers GP. Olfactomedin 4 inhibits cathepsin C-mediated protease activities, thereby modulating neutrophil killing of *Staphylococcus aureus* and *Escherichia coli* in mice. *J Immunol* 2012;189:2460-7.

76. Bhandari V, Wang C, Rinder C, Rinder H. Hematologic profile of sepsis in neonates: neutrophil CD64 as a diagnostic marker. *Pediatrics* 2008;121:129-34.
77. Groselj-Grenc M, Ihan A, Pavcnik-Arnol M, Kopitar AN, Gmeiner-Stopar T, Derganc M. Neutrophil and monocyte CD64 indexes, lipopolysaccharide-binding protein, procalcitonin and C-reactive protein in sepsis of critically ill neonates and children. *Intensive Care Med* 2009;35:1950-8.
78. Berner R, Furll B, Stelter F, Drose J, Muller HP, Schutt C. Elevated levels of lipopolysaccharide-binding protein and soluble CD14 in plasma in neonatal early-onset sepsis. *Clin Diagn Lab Immunol* 2002;9:440-5.
79. Pavcnik-Arnol M, Hojker S, Derganc M. Lipopolysaccharide-binding protein, lipopolysaccharide, and soluble CD14 in sepsis of critically ill neonates and children. *Intensive Care Med* 2007;33:1025-32.
80. Viemann D, Dubbel G, Schleifenbaum S, Harms E, Sorg C, Roth J. Expression of toll-like receptors in neonatal sepsis. *Pediatr Res* 2005;58:654-9.

BIBLIOGRAFÍA

81. Arnon S, Litmanovitz I, Regev RH, Bauer S, Shainkin-Kestenbaum R, Dolfin T. Serum amyloid A: an early and accurate marker of neonatal early-onset sepsis. *J Perinatol* 2007;27:297-302.
82. Joram N, Boscher C, Denizot S, et al. Umbilical cord blood procalcitonin and C reactive protein concentrations as markers for early diagnosis of very early onset neonatal infection. *Arch Dis Child Fetal Neonatal Ed* 2006;91:F65-6.
83. Perez Solis D, Lopez Sastre JB, Coto Cotallo GD, Dieguez Junquera MA, Deschamps Mosquera EM, Crespo Hernandez M. Procalcitonin for the diagnosis of neonatal sepsis of vertical transmission. *An Pediatr (Barc)* 2006;64:341-8.
84. Resch B, Gusenleitner W, Muller WD. Procalcitonin and interleukin-6 in the diagnosis of early-onset sepsis of the neonate. *Acta Paediatr* 2003;92:243-5.
85. Khassawneh M, Hayajneh WA, Kofahi H, Khader Y, Amarin Z, Daoud A. Diagnostic markers for neonatal sepsis: comparing C-reactive protein, interleukin-6 and immunoglobulin M. *Scand J Immunol* 2007;65:171-5.
86. Galetto-Lacour A, Zamora SA, Gervais A. Bedside procalcitonin and C-reactive protein tests in children with fever

without localizing signs of infection seen in a referral center. *Pediatrics* 2003;112:1054-60.

87. Shang S, Chen G, Wu Y, Du L, Zhao Z. Rapid diagnosis of bacterial sepsis with PCR amplification and microarray hybridization in 16S rRNA gene. *Pediatr Res* 2005;58:143-8.

88. Wu YD, Chen LH, Wu XJ, et al. Gram stain-specific-probe-based real-time PCR for diagnosis and discrimination of bacterial neonatal sepsis. *J Clin Microbiol* 2008;46:2613-9.

89. Pammi M, Flores A, Leeflang M, Versalovic J. Molecular assays in the diagnosis of neonatal sepsis: a systematic review and meta-analysis. *Pediatrics* 2011;128:e973-85.

90. Lam HS, Ng PC. Biochemical markers of neonatal sepsis. *Pathology* 2008;40:141-8.

91. Crick F. Central dogma of molecular biology. *Nature* 1970;227:561-3.

92. Eddy SR. Non-coding RNA genes and the modern RNA world. *Nat Rev Genet* 2001;2:919-29.

93. Cernada M, Aguar M, Brugada M, et al. Ventilator-associated pneumonia in newborn infants diagnosed with an invasive bronchoalveolar lavage technique: a prospective observational study. *Pediatr Crit Care Med* 2013;14:55-61.

BIBLIOGRAFÍA

94. Tollner U. Early diagnosis of septicemia in the newborn. Clinical studies and sepsis score. *Eur J Pediatr* 1982;138:331-7.
95. Jawaheer G, Neal TJ, Shaw NJ. Blood culture volume and detection of coagulase negative staphylococcal septicaemia in neonates. *Arch Dis Child Fetal Neonatal Ed* 1997;76:F57-8.
96. Usher R, Shephard M, Lind J. The Blood Volume of the Newborn Infant and Placental Transfusion. *Acta Paediatr* 1963;52:497-512.
97. Mueller O, Lightfoot S, Schroeder A. RNA Integrity Number (RIN). Standardization of RNA Quality Control. . Agilent Application Note 2004:1-8.
98. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002;3 (7):1-11.
99. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 2003;19:185-93.
100. I.T. J, ed. *Principal Component Analysis*. 2nd ed. NY: Springer; 2002.

101. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* 2001;58:109-30.
102. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001;29:e45.
103. Kuligowski J, Perez-Guaita D, Escobar J, et al. Evaluation of the effect of chance correlations on variable selection using Partial Least Squares-Discriminant Analysis. *Talanta* 2013;116:835-40.
104. Westerhuis JA, Hoefsloot HCJ, Smit S, et al. Assessment of PLS-DA cross validation. *Metabolomics* 2008;4:81-9.
105. Westerhuis JA, van Velzen EJJ, Hoefsloot HCJ, Smilde AK. Discriminant Q2 for improved discrimination in PLS-DA models. *Metabolomics* 2008;4:293-6.
106. Edmond K, Zaidi A. New approaches to preventing, diagnosing, and treating neonatal sepsis. *PLoS Med* 2010;7:e1000213.
107. Arnon S, Litmanovitz I. Diagnostic tests in neonatal sepsis. *Curr Opin Infect Dis* 2008;21:223-7.
108. Vouloumanou EK, Plessa E, Karageorgopoulos DE, Mantadakis E, Falagas ME. Serum procalcitonin as a diagnostic marker for neonatal sepsis: a systematic review and meta-analysis. *Intensive Care Med* 2011;37:747-62.

BIBLIOGRAFÍA

109. Hofer N, Zacharias E, Muller W, Resch B. An update on the use of C-reactive protein in early-onset neonatal sepsis: current insights and new tasks. *Neonatology* 2012;102:25-36.
110. Ng PC, Lam HS. Biomarkers in neonatology: the next generation of tests. *Neonatology* 2012;102:145-51.
111. Cernada M, Serna E, Bauerl C, Collado MC, Perez-Martinez G, Vento M. Genome-wide expression profiles in very low birth weight infants with neonatal sepsis. *Pediatrics* 2014;133:e1203-11.
112. Dellinger RP, Levy MM, Rhodes A, et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. *Crit Care Med* 2013;41:580-637.
113. Wynn JL, Guthrie SO, Wong HR, et al. Post-natal age is a critical determinant of the neonatal host response to sepsis. *Mol Med* 2015.
114. Tang BM, McLean AS, Dawes IW, Huang SJ, Lin RC. Gene-expression profiling of peripheral blood mononuclear cells in sepsis. *Crit Care Med* 2009;37:882-8.
115. Russell JA. Gene expression in human sepsis: what have we learned? *Crit Care* 2011;15:121.

116. Johnson SB, Lissauer M, Bochicchio GV, Moore R, Cross AS, Scalea TM. Gene expression profiles differentiate between sterile SIRS and early sepsis. *Ann Surg* 2007;245:611-21.
117. Wong HR, Cvijanovich N, Lin R, et al. Identification of pediatric septic shock subclasses based on genome-wide expression profiling. *BMC Med* 2009;7:34.
118. Prucha M, Ruryk A, Boriss H, et al. Expression profiling: toward an application in sepsis diagnostics. *Shock* 2004;22:29-33.
119. Tang BM, Huang SJ, McLean AS. Genome-wide transcription profiling of human sepsis: a systematic review. *Crit Care* 2010;14:R237.
120. Wong HR, Shanley TP, Sakthivel B, et al. Genome-level expression profiles in pediatric septic shock indicate a role for altered zinc homeostasis in poor outcome. *Physiol Genomics* 2007;30:146-55.
121. Ramilo O, Allman W, Chung W, et al. Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood* 2007;109:2066-77.
122. Wynn JL, Cvijanovich NZ, Allen GL, et al. The influence of developmental age on the early transcriptomic response of children with septic shock. *Mol Med* 2011;17:1146-56.

BIBLIOGRAFÍA

123. Koch L, Linderkamp O, Ittrich C, Benner A, Poeschl J. Gene expression profiles of adult peripheral and cord blood mononuclear cells altered by lipopolysaccharide. *Neonatology* 2008;93:87-100.
124. Markel TA, Crisostomo PR, Wang M, Herring CM, Meldrum DR. Activation of individual tumor necrosis factor receptors differentially affects stem cell growth factor and cytokine production. *Am J Physiol Gastrointest Liver Physiol* 2007;293:G657-62.
125. Wong HR, Cvijanovich N, Allen GL, et al. Genomic expression profiling across the pediatric systemic inflammatory response syndrome, sepsis, and septic shock spectrum. *Crit Care Med* 2009;37:1558-66.
126. Herman MP, Sukhova GK, Libby P, et al. Expression of neutrophil collagenase (matrix metalloproteinase-8) in human atheroma: a novel collagenolytic pathway suggested by transcriptional profiling. *Circulation* 2001;104:1899-904.
127. Tester AM, Cox JH, Connor AR, et al. LPS responsiveness and neutrophil chemotaxis in vivo require PMN MMP-8 activity. *PLoS One* 2007;2:e312.
128. Schaaf B, Liebau C, Kurowski V, Droemann D, Dalhoff K. Hospital acquired pneumonia with high-risk bacteria is associated

with increased pulmonary matrix metalloproteinase activity. *BMC Pulm Med* 2008;8:12.

129. Vanlaere I, Libert C. Matrix metalloproteinases as drug targets in infections caused by gram-negative bacteria and in septic shock. *Clin Microbiol Rev* 2009;22:224-39.

130. Solan PD, Dunsmore KE, Denenberg AG, Odoms K, Zingarelli B, Wong HR. A novel role for matrix metalloproteinase-8 in sepsis. *Crit Care Med* 2012;40:379-87.

131. Lalezari P, Murphy GB, Allen FH, Jr. NB1, a new neutrophil-specific antigen involved in the pathogenesis of neonatal neutropenia. *J Clin Invest* 1971;50:1108-15.

132. Bux J. Molecular nature of antigens implicated in immune neutropenias. *Int J Hematol* 2002;76 Suppl 1:399-403.

133. Stroncek DF. Neutrophil-specific antigen HNA-2a, NB1 glycoprotein, and CD177. *Curr Opin Hematol* 2007;14:688-93.

134. Meyerson HJ, Osei E, Schweitzer K, Blidaru G, Edinger A, Balog A. CD177 expression on neutrophils: in search of a clonal assay for myeloid neoplasia by flow cytometry. *Am J Clin Pathol* 2013;140:658-69.

BIBLIOGRAFÍA

135. Pliyev BK, Menshikov M. Comparative evaluation of the role of the adhesion molecule CD177 in neutrophil interactions with platelets and endothelium. *Eur J Haematol* 2012;89:236-44.
136. Gohring K, Wolff J, Doppl W, et al. Neutrophil CD177 (NB1 gp, HNA-2a) expression is increased in severe bacterial infections and polycythaemia vera. *Br J Haematol* 2004;126:252-4.
137. Bayat B, Werth S, Sachs UJ, Newman DK, Newman PJ, Santoso S. Neutrophil transmigration mediated by the neutrophil-specific antigen CD177 is influenced by the endothelial S536N dimorphism of platelet endothelial cell adhesion molecule-1. *J Immunol* 2010;184:3889-96.
138. Zhang X, Huang Q, Yang Z, Li Y, Li CY. GW112, a novel antiapoptotic protein that promotes tumor growth. *Cancer Res* 2004;64:2474-81.
139. Barnich N, Hisamatsu T, Aguirre JE, Xavier R, Reinecker HC, Podolsky DK. GRIM-19 interacts with nucleotide oligomerization domain 2 and serves as downstream effector of anti-bacterial function in intestinal epithelial cells. *J Biol Chem* 2005;280:19021-6.
140. Marino S, Verzegnassi F, Tamaro P, et al. Response to glucocorticoids and toxicity in childhood acute lymphoblastic

leukemia: role of polymorphisms of genes involved in glucocorticoid response. *Pediatr Blood Cancer* 2009;53:984-91.

141. Vento M, Asensi M, Sastre J, Lloret A, Garcia-Sala F, Vina J. Oxidative stress in asphyxiated term infants resuscitated with 100% oxygen. *J Pediatr* 2003;142:240-6.

142. Harwood DT, Darlow BA, Cheah FC, McNeill N, Graham P, Winterbourn CC. Biomarkers of neutrophil-mediated glutathione and protein oxidation in tracheal aspirates from preterm infants: association with bacterial infection. *Pediatr Res* 2011;69:28-33.

143. Cernada M, Brugada M, Golombek S, Vento M. Ventilator-associated pneumonia in neonatal patients: an update. *Neonatology* 2014;105:98-107.

144. Arredouani M, Matthys P, Kasran A, Baumann H, Ceuppen JL. Haptoglobin and the Th1/Th2 balance: hints from in vitro and in vivo studies. *Redox Rep* 2001;6:369-71.

145. Flo TH, Smith KD, Sato S, et al. Lipocalin 2 mediates an innate immune response to bacterial infection by sequestering iron. *Nature* 2004;432:917-21.

146. Borregaard N, Sorensen OE, Theilgaard-Monch K. Neutrophil granules: a library of innate immunity proteins. *Trends Immunol* 2007;28:340-5.





BIBLIOGRAFÍA

147. Shen F, Hu Z, Goswami J, Gaffen SL. Identification of common transcriptional regulatory elements in interleukin-17 target genes. *J Biol Chem* 2006;281:24138-48.
148. Wu H, Santoni-Rugiu E, Ralfkiaer E, et al. Lipocalin 2 is protective against *E. coli* pneumonia. *Respir Res* 2010;11:96.
149. Nagaishi T, Iijima H, Nakajima A, Chen D, Blumberg RS. Role of CEACAM1 as a regulator of T cells. *Ann N Y Acad Sci* 2006;1072:155-75.
150. Van der Flier M, Sharma DB, Estevao S, et al. Increased CD4(+) T cell co-inhibitory immune receptor CEACAM1 in neonatal sepsis and soluble-CEACAM1 in meningococcal sepsis: a role in sepsis-associated immune suppression? *PLoS One* 2013;8:e68294.
151. Bouchard C, Page J, Bedard A, Tremblay P, Vallieres L. G protein-coupled receptor 84, a microglia-associated protein expressed in neuroinflammatory conditions. *Glia* 2007;55:790-800.
152. Feezor RJ, Oberholzer C, Baker HV, et al. Molecular characterization of the acute inflammatory response to infections with gram-negative versus gram-positive bacteria. *Infect Immun* 2003;71:5803-13.

153. Yu SL, Chen HW, Yang PC, et al. Differential gene expression in gram-negative and gram-positive sepsis. *Am J Respir Crit Care Med* 2004;169:1135-43.
154. Tang BM, McLean AS, Dawes IW, Huang SJ, Cowley MJ, Lin RC. Gene-expression profiling of gram-positive and gram-negative sepsis in critically ill patients. *Crit Care Med* 2008;36:1125-8.

VIII. ANEXOS

VIII. 1. APROBACIÓN DEL COMITÉ ÉTICO Y CONSENTIMIENTO INFORMADO


 GENERALITAT VALENCIANA CONSELLERIA DE SANITAT	 AGÈNCIA VALENCIANA DE SALUT	 La Fe DEPARTAMENT DE SALUT	 Fundación de Investigación Hospital La Fe
--	--	---	--

Comisión de Investigación

D. José Vte. Castell Ripoll, Presidente de la Comisión de Investigación del Hospital Universitario La Fe de Valencia,

INFORMA:

Que el Proyecto de Investigación titulado: "*Perfiles de expresión genética: nuevo enfoque diagnóstico para la sepsis en recién nacidos menores de 1500 gramos*" que presenta el/la Dr./Dra. **Máximo Vento Torres** del Servicio de Neonatología del Hospital La Fe de Valencia, contiene elementos objetivos suficientes en cuanto a la Hipótesis, Planteamientos y Plan de Trabajo que, a juicio de esta Comisión, permiten pronunciarse **positivamente** en cuanto a su viabilidad.



Valencia a 5 de Mayo de 2011.

Comisión de Investigación. Escuela de Enfermería – despacho 619.- Hospital Universitario La Fe.
 Avda. Campanar, 21. 46009-Valencia. Spain
 Tel.: (+34) 96 1973313; Fax: (+34) 96 3494416. e-mail: cominvest_lafe@gva.es

HOJA INFORMATIVA AL PACIENTE

TÍTULO DEL ESTUDIO: PERFILES DE EXPRESIÓN GENÓMICA: NUEVO ENFOQUE DIAGNÓSTICO PARA LA SEPSIS EN MENORES DE 1500 GRAMOS.
INVESTIGADOR PRINCIPAL: María Cernada Badia. Pediatra, Servicio de Neonatología. Hospital Universitario La Fe.
mariacernada@gmail.com
CENTRO: Hospital Universitario La Fe.

1. INTRODUCCIÓN

Nos dirigimos a Ud. para informarle sobre un estudio, aprobado por el Comité Ético de Investigación Biomédica del Hospital Universitario La Fe de Valencia, en el que a su hijo/a se le invita a participar.

Nuestra intención es tan sólo que usted reciba la información correcta y suficiente para que pueda evaluar y juzgar, si quiere que su hijo/a participe o no en este estudio.

Para ello le ruego lea esta hoja informativa con atención, pudiendo consultar con las personas que considere oportuno, quedando a su disposición para aclararle las dudas que pudieran surgirle.

2. PARTICIPACIÓN VOLUNTARIA

Debe saber que la participación de su hijo/a en este estudio es voluntaria, y que puede decidir no participar, o cambiar su decisión y retirar su consentimiento en cualquier momento, sin que por ello se altere la relación con su médico ni produzca perjuicio alguno en su tratamiento.

3. DESCRIPCIÓN GENERAL DEL ESTUDIO

Se trata de un trabajo de investigación dirigido a recién nacidos con peso inferior a los 1500 gramos, con sospecha de presentar infección. Los recién nacidos prematuros o con muy bajo peso tienen más riesgo de infectarse por su inmadurez y por los procedimientos a los que son sometidos durante su estancia hospitalaria. Estos niños, por el hecho de haber nacido antes de tiempo, no presentan un sistema inmune totalmente desarrollado, y en el momento actual desconocemos cómo se comporta ante la presencia de agentes infecciosos. Los métodos de los que disponemos para diagnosticarlos no son suficientemente sensibles en estos niños, por lo que en muchas ocasiones nos basamos en las manifestaciones clínicas, y ante la más mínima duda, iniciamos tratamiento antibiótico sin tener la absoluta certeza que haya una infección activa.

El objetivo de este trabajo es conocer cómo se comporta el sistema inmune del recién nacido de bajo peso y analizar si la expresión de los genes podría ser una herramienta diagnóstica más fiable que las que se utilizan en la actualidad.

Para ello se estudiará a los recién nacidos menores de 1 mes y con un peso al nacimiento inferior a los 1500 gramos, que presenten signos de infección. Aprovechando otra extracción, y sin que sea necesario un pinchazo adicional, se extraerán 0.2 mililitros de sangre que se utilizarán para analizar la expresión genética. Se trata de un análisis no de los genes en sí mismos sino de cómo se modifica su expresión en función de si los niños están infectados o no.

El estudio tiene una duración de tres años. Durante este tiempo se irán incluyendo pacientes y se seguirá su evolución clínica. Si los resultados así lo aconsejan, podría existir una prolongación temporal todavía por definir.

El paciente incluido en el estudio no precisará ninguna visita extraordinaria al hospital. El protocolo de trabajo se basa en la inclusión de unos 100 pacientes, todos ellos menores de 1 mes de vida y con peso al nacimiento inferior a 1500 gramos, durante el tiempo que permanezca abierto el estudio.

En ningún caso la inclusión en el estudio supondrá variación alguna en cuanto al tratamiento o seguimiento de los pacientes.

4. BENEFICIOS Y RIESGOS DERIVADOS DE SU PARTICIPACIÓN EN EL ESTUDIO.

La participación en el estudio no supone ninguna modificación de tratamiento ni de seguimiento; por lo que los pacientes incluidos no afrontarán ningún peligro adicional para su salud. La única diferencia entre los pacientes incluidos y no incluidos en el estudio, es que los primeros se extraerá, aprovechando otra extracción, una cantidad adicional de 0'2 mililitros de sangre.

El proyecto persigue un mayor conocimiento de la respuesta del sistema inmune de los recién nacidos menores de 1500 gramos a la infección. No se estudiarán los genes directamente, sino la forma en la que estos genes se expresan. Los pacientes incluidos no obtendrán un beneficio directo de los resultados de este estudio. No obstante, podrán conocer cómo responde su sistema inmune a la infección; además de colaborar con el progreso científico en este campo, contribuyendo al beneficio de otros pacientes en su misma situación.

El estudio podrá incluir a todos los recién nacidos menores de 1500 gramos menores de 1 mes de vida, que presenten signos de infección, y no tengan enfermedades subyacentes que pudieran afectar a normal la respuesta inmunológica.

5. TRATAMIENTOS ALTERNATIVOS

Como se ha comentado en apartados anteriores, los pacientes incluidos no recibirán ningún tipo de tratamiento por el hecho de estar incluidos en el estudio.

6. CONFIDENCIALIDAD Y TRATAMIENTO DE DATOS

El tratamiento, la comunicación y la cesión de los datos de carácter personal de todos los sujetos participantes se ajustará a lo dispuesto en la Ley Orgánica 15/99 de 13 de diciembre de protección de datos de carácter personal.

De acuerdo a lo que establece la legislación mencionada, usted puede ejercer los derechos de acceso, modificación, oposición y cancelación de datos, para lo cual se deberá dirigir a su médico del estudio.

Los datos recogidos para el estudio estarán identificados mediante un código, siguiendo un mecanismo de disociación de modo que sólo el médico responsable del registro (Dra. María Cernada Badía) podrá relacionar dichos datos con Usted y con su historia clínica. Por lo tanto, su identidad no será revelada a persona alguna salvo caso de urgencia médica o requerimiento legal.

El acceso a su información personal quedará restringido al médico del estudio / colaboradores, autoridades sanitarias (Agencia Española del Medicamento y Productos Sanitarios), al Comité Ético de Investigación Clínica y personal autorizado por el promotor, cuando lo precisen para comprobar los datos y procedimientos del estudio, pero siempre manteniendo la confidencialidad de los mismos de acuerdo a la legislación vigente en nuestro país.

La entrega de muestras genéticas no tiene carácter de propiedad permanente, por lo que la donación está limitada a los objetivos del estudio, durante un tiempo máximo de 4 años, durante el cual el promotor tendrá la oportunidad y ventaja de realizar el estudio sobre dichas muestras con exclusividad y sin competencias; así como desarrollar tantas patentes como pueda y le sean autorizadas.

Transcurrido este tiempo:

- Los datos de la información genética derivados del estudio deben ser públicos y/o podrán ser fácilmente consultados.
- El material genómico dejará ser propiedad del promotor y el destino del mismo debe ser decidido y/o regulado por los Organismos Gubernamentales del país de procedencia de las muestras, o en su caso, ser destruido.
- Bajo ningún concepto y en ningún momento las muestras serán motivo de lucro directo, bien sea por la venta del material o de los derechos para realizar estudios sobre los mismos.

NORMATIVA INTERNA QUE REGULA LOS ESTUDIOS GENÉTICOS
--

ASPECTOS GENERALES QUE DEBEN SER CONSIDERADOS

Este CEIC entiende que la entrega de este tipo de muestras no tiene carácter de propiedad permanente, que la donación debe estar limitada a unos objetivos claros y concretos, a un tiempo de estudio razonable (no superior a los 6 años) durante el cual el promotor tendrá la oportunidad y ventaja de realizar el estudio sobre dichas muestras con exclusividad y sin competencias; así como desarrollar tantas patentes como pueda y le sean autorizadas.

Transcurrido este tiempo:

- Los datos de la información genética derivados del estudio deben ser públicos y/o podrán ser fácilmente consultados.
- El material genómico dejará ser propiedad del promotor y el destino del mismo debe ser decidido y/o regulado por los Organismos Gubernamentales del país de procedencia de las muestras, o en su caso, ser destruido.
- Bajo ningún concepto y en ningún momento las muestras serán motivo de lucro directo, bien sea por la venta del material o de los derechos para realizar estudios sobre los mismos.

Todo lo anteriormente expuesto debe ser recogido en los protocolos de ensayos clínicos y en la hoja de información al paciente.

**MODELO DE CONSENTIMIENTO INFORMADO
PARA REPRESENTANTE LEGAL**

Título del estudio: PERFILES DE EXPRESIÓN GENÓMICA: NUEVO ENFOQUE
DIAGNÓSTICO PARA LA SEPSIS EN RECIÉN NACIDOS MENORES DE 1500
GRAMOS

Yo,..... en calidad
de (relación con el participante)

He leído la hoja de información que se me ha entregado.

He podido hacer preguntas sobre el estudio.

He recibido suficiente información sobre el estudio.

He hablado con.....

Comprendo que su participación es voluntaria.

Comprendo que puede retirarse del estudio:

- Cuando quiera.
- Sin tener que dar explicaciones.
- Sin que esto repercuta en sus cuidados médicos.

Y presto mi conformidad con queparticipe en el
estudio.

Fecha:

Firma del representante:

Firma del investigador:

VIII. 2. LISTADO DE GENES SIGNIFICATIVOS QUE DIFERENCIAN SEPSIS DE CONTROLES

Gene Name	Gene Symbol	P-value	Fold Change
Matrix metallopeptidase 8 (neutrophil collagenase)	MMP8	$2,08 \times 10^{-6}$	7,34
Olfactomedin 4	OLFM4	$2,08 \times 10^{-5}$	7,11
CD177 molecule	CD177	$7,63 \times 10^{-6}$	6,80
Ankyrin repeat domain 22	ANKRD22	$3,11 \times 10^{-6}$	6,41
CD177 molecule	CD177	$2,29 \times 10^{-6}$	5,61
G protein-coupled receptor 84	GPR84	$2,81 \times 10^{-7}$	5,13
Haptoglobin	HP	$8,20 \times 10^{-5}$	4,17
Carcinoembryonic antigen-related cell adhesion molecule 1	CEACAM1	$2,03 \times 10^{-6}$	3,79
Lipocalin 2	LCN2	$4,74 \times 10^{-5}$	3,55
Interleukin 1 receptor, type II	IL1R2	$4,31 \times 10^{-5}$	3,40
C-type lectin domain family 4, member D	CLEC4D	$2,10 \times 10^{-4}$	3,39
Matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase)	MMP9	$4,84 \times 10^{-4}$	3,38
Transcobalamin I (vitamin B12 binding protein)	TCN1	$2,19 \times 10^{-4}$	3,32
Interleukin 18 receptor 1	IL18R1	$3,65 \times 10^{-6}$	3,15
6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2	PFKFB2	$1,70 \times 10^{-6}$	3,10
Glycogenin 1	GYG1	$4,04 \times 10^{-5}$	3,00
SAM domain, SH3 domain and nuclear localization signals 1	SAMSN1	$1,34 \times 10^{-5}$	2,86
Complement component 3a receptor 1	C3AR1	$9,22 \times 10^{-5}$	2,84
BMX non-receptor tyrosine kinase	BMX	$5,17 \times 10^{-5}$	2,63
Solute carrier family 26, member 8	SLC26A8	$2,02 \times 10^{-4}$	2,63
Vanin 1	VNN1	$7,67 \times 10^{-4}$	2,62
N-terminal EF-hand calcium binding protein 1	NECAB1	$2,70 \times 10^{-5}$	2,60
Interleukin 1 receptor antagonist	IL1RN	$1,38 \times 10^{-4}$	2,58

Gene Name	Gene Symbol	P-value	Fold Change
Resistin	RETN	$1,32 \times 10^{-5}$	2,55
C-type lectin domain family 5, member A	CLEC5A	$2,50 \times 10^{-4}$	2,53
UDP-glucose ceramide glucosyltransferase	UGCG	$3,82 \times 10^{-5}$	2,46
Family with sequence similarity 20, member A	FAM20A	$2,17 \times 10^{-5}$	2,45
Proline-serine-threonine phosphatase interacting protein 2	PSTPIP2	$3,71 \times 10^{-5}$	2,41
UDP-GlcNAc:betaGal beta-1,3N-acetylglucosaminyltransferase 5	B3GNT5	$5,47 \times 10^{-5}$	2,40
Chromosome 19 open reading frame 59	C19orf59	$1,75 \times 10^{-4}$	2,39
TRAF-interacting protein with forkhead-associated domain	TIFA	$4,19 \times 10^{-4}$	2,37
6-phosphofructo-2-kinase/fructose-2,6-biphosphatase	PFKFB3	$1,13 \times 10^{-5}$	2,33
Egf-like module containing, mucin-like, hormone receptor-like 1	EMR1	$5,02 \times 10^{-4}$	2,31
Fc fragment of IgA, receptor for Amphiregulin	FCAR	$1,07 \times 10^{-4}$	2,29
Amphiregulin	AREG	$1,09 \times 10^{-4}$	2,28
Phospholipid scramblase 1	PLSCR1	$4,80 \times 10^{-4}$	2,25
ATPase, class II, type 9A	ATP9A	$1,86 \times 10^{-4}$	2,23
Proteoglycan 3 pseudogene	SLED1	$2,85 \times 10^{-4}$	2,22
G protein-coupled receptor 97	GPR97	$4,54 \times 10^{-4}$	2,18
Dehydrogenase/reductase (SDR family) member 9	DHRS9	$7,84 \times 10^{-5}$	2,18
Peptidyl arginine deiminase, type IV	PADI4	$2,19 \times 10^{-4}$	2,17
TBC1 domain family, member 8 (with GRAM domain)	TBC1D8	$1,30 \times 10^{-4}$	2,17
Membrane-spanning 4-domains, subfamily A, member 4	MS4A4A	$2,18 \times 10^{-4}$	2,15
Glutathione S-transferase mu 1	GSTM1	$3,19 \times 10^{-4}$	2,12
Sortilin 1	SORT1	$3,46 \times 10^{-4}$	2,11
Solute carrier family 1 (glial high affinity glutamate transport)	SLC1A3	$5,91 \times 10^{-5}$	2,10

Gene Name	Gene Symbol	P-value	Fold Change
Tumor necrosis factor, alpha-induced protein 3	TNFAIP3	$3,56 \times 10^{-4}$	2,10
BCL2-related protein A1	BCL2A1	$3,85 \times 10^{-4}$	2,10
S100 calcium binding protein A12	S100A12	$3,44 \times 10^{-4}$	2,09
Suppressor of cytokine signaling 3	SOCS3	$1,58 \times 10^{-6}$	2,06
Feline leukemia virus subgroup C cellular receptor family, member 2	FLVCR2	$5,11 \times 10^{-4}$	2,06
Ectonucleoside triphosphate diphosphohydrolase 7	ENTPD7	$6,66 \times 10^{-4}$	2,06
Tribbles homolog 1 (Drosophila)	TRIB1	$6,33 \times 10^{-6}$	2,03
NOP2/Sun domain family, member 7	NSUN7	$3,66 \times 10^{-4}$	2,01
Heat shock 70kDa protein 1A	HSPA1A	$1,33 \times 10^{-5}$	1,99
Heat shock 70kDa protein 1A	HSPA1A	$1,31 \times 10^{-5}$	1,99
Calcium channel, voltage-dependent, R type, alpha 1E subunit	CACNA1E	$6,90 \times 10^{-5}$	1,98
Heat shock 70kDa protein 1B	HSPA1B	$2,30 \times 10^{-5}$	1,98
Heat shock 70kDa protein 1B	HSPA1B	$2,30 \times 10^{-5}$	1,98
Heat shock 70kDa protein 1B	HSPA1B	$2,84 \times 10^{-5}$	1,97
Arachidonate 5-lipoxygenase-activating protein	ALOX5AP	$5,73 \times 10^{-4}$	1,96
Intercellular adhesion molecule 1	ICAM1	$3,82 \times 10^{-4}$	1,94
Potassium voltage-gated channel, Isk-related family, member 1	KCNE1	$1,52 \times 10^{-4}$	1,94
Protein S (alpha)	PROS1	$7,74 \times 10^{-5}$	1,94
Chromosome 6 open reading frame 150	C6orf150	$4,76 \times 10^{-4}$	1,94
V-ets erythroblastosis virus E26 oncogene homolog 2 (avian)	ETS2	$2,07 \times 10^{-4}$	1,93
Flotillin 1	FLOT1	$2,40 \times 10^{-6}$	1,93
Flotillin 1	FLOT1	$2,40 \times 10^{-6}$	1,93
Flotillin 1	FLOT1	$2,40 \times 10^{-6}$	1,93
FK506 binding protein 5	FKBP5	$9,35 \times 10^{-5}$	1,92
Free fatty acid receptor 3	FFAR3	$1,63 \times 10^{-4}$	1,92

Gene Name	Gene Symbol	P-value	Fold Change
CD63 molecule	CD63	$3,22 \times 10^{-5}$	1,91
S100 calcium binding protein P	S100P	$5,74 \times 10^{-4}$	1,89
Leucine-rich alpha-2-glycoprotein 1	LRG1	$5,11 \times 10^{-4}$	1,87
Multiple C2 domains, transmembrane	MCTP2	$3,80 \times 10^{-4}$	1,86
B-cell CLL/lymphoma 3	BCL3	$2,06 \times 10^{-6}$	1,86
Klotho	KL	$4,48 \times 10^{-5}$	1,85
Growth factor receptor-bound protein 10	GRB10	$7,63 \times 10^{-4}$	1,85
Sulfide quinone reductase-like	SQRDL	$2,01 \times 10^{-5}$	1,85
RAB20, member RAS oncogene family	RAB20	$6,02 \times 10^{-6}$	1,84
Isopentenyl-diphosphate delta isomerase 1	IDI1	$6,35 \times 10^{-5}$	1,84
Cytidine and dCMP deaminase domain containing 1	CDADC1	$5,92 \times 10^{-5}$	1,83
Solute carrier family 2 member 3 (facilitated glucose transporter)	SLC2A3	$1,54 \times 10^{-5}$	1,82
Tudor domain containing 9	TDRD9	$3,39 \times 10^{-4}$	1,82
Mitogen-activated protein kinase 14	MAPK14	$8,18 \times 10^{-4}$	1,81
Free fatty acid receptor 3	FFAR3	$1,05 \times 10^{-4}$	1,81
Methylenetetrahydrofolate dehydrogenase (NADP ⁺ dependent) 2	MTHFD2	$1,22 \times 10^{-5}$	1,81
WD repeat and SOCS box-containing 1	WSB1	$1,46 \times 10^{-4}$	1,78
CD82 molecule	CD82	$3,41 \times 10^{-4}$	1,78
Peroxisome proliferator-activated receptor gamma	PPARG	$5,06 \times 10^{-4}$	1,77
---		$6,35 \times 10^{-4}$	1,77
Perilipin 3	PLIN3	$2,51 \times 10^{-6}$	1,76
Organic solute transporter alpha	OSTalpha	$3,87 \times 10^{-5}$	1,75
---		$2,25 \times 10^{-5}$	1,75
UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide	B4GALT5	$4,31 \times 10^{-4}$	1,75

Gene Name	Gene Symbol	P-value	Fold Change
Jun B proto-oncogene	JUNB	$1,10 \times 10^{-5}$	1,75
7-dehydrocholesterol reductase	DHCR7	$1,39 \times 10^{-4}$	1,74
ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)	ST6GALNAC3	$1,72 \times 10^{-4}$	1,74
Thrombospondin 1	THBS1	$4,17 \times 10^{-43}$	1,74
Uridine phosphorylase 1	UPP1	$2,56 \times 10^{-5}$	1,73
Nuclear factor of kappa light polypeptide gene enhancer in B-cel	NFKBIA	$4,27 \times 10^{-4}$	1,73
Cytoskeleton-associated protein 4	CKAP4	$3,44 \times 10^{-4}$	1,72
Methyltransferase like 7B	METTL7B	$1,98 \times 10^{-6}$	1,72
Lactate dehydrogenase A	LDHA	$6,39 \times 10^{-6}$	1,71
Formin 1	FMN1	$2,75 \times 10^{-4}$	1,70
ADAM metallopeptidase domain 9	ADAM9	$7,67 \times 10^{-4}$	1,70
---		$6,03 \times 10^{-4}$	1,70
Torsin family 1, member B (torsin B)	TOR1B	$7,54 \times 10^{-4}$	1,70
CCAAT/enhancer binding protein	CEBPB	$7,13 \times 10^{-4}$	1,69
Amphiregulin	AREG	$3,17 \times 10^{-4}$	1,69
Guanine nucleotide binding protein (G protein), alpha 15	GNA15	$8,03 \times 10^{-6}$	1,68
Hypoxia inducible factor 1, alpha subunit (basic helix-loop-helix)	HIF1A	$3,09 \times 10^{-5}$	1,67
Major vault protein	MVP	$1,54 \times 10^{-6}$	1,65
Protein phosphatase 1, regulatory (inhibitor) subunit 3B	PPP1R3B	$5,45 \times 10^{-5}$	1,65
Kringle containing transmembrane protein 1	KREMEN1	$7,70 \times 10^{-4}$	1,65
Teashirt zinc finger homeobox 3	TSHZ3	$7,01 \times 10^{-4}$	1,64
Netrin G2	NTNG2	$8,05 \times 10^{-4}$	1,64
Interleukin-1 receptor-associated kinase 2	IRAK2	$5,78 \times 10^{-4}$	1,63
---		$1,52 \times 10^{-5}$	1,63
Sprouty-related, EVH1 domain containing 2	SPRED2	$3,06 \times 10^{-5}$	1,62

Gene Name	Gene Symbol	P-value	Fold Change
Strawberry notch homolog 2 (Drosophila)	SBNO2	$3,35 \times 10^{-6}$	1,62
Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2	MTHFD2	$1,15 \times 10^{-5}$	1,62
Flotillin 2	FLOT2	$0,000186185$	1,62
Neuritin 1	NRN1	$9,48 \times 10^{-5}$	1,61
Adrenomedullin	ADM	$8,17 \times 10^{-4}$	1,60
Lamin B1	LMNB1	$2,63 \times 10^{-4}$	1,60
Immediate early response 3	IER3	$1,28 \times 10^{-4}$	1,60
Myristoylated alanine-rich protein kinase C substrate	MARCKS	$6,49 \times 10^{-4}$	1,59
Tctex1 domain containing 1	TCTEX1D1	$5,89 \times 10^{-4}$	1,58
Exoribonuclease 1	ERI1	$6,49 \times 10^{-4}$	1,58
GRAM domain containing 1A	GRAMD1A	$4,27 \times 10^{-4}$	1,58
Solute carrier family 2 (facilitated glucose transporter), memb	SLC2A14	$1,21 \times 10^{-5}$	1,56
PR domain containing 5	PRDM5	$2,09 \times 10^{-4}$	1,56
SH3 and PX domains 2B	SH3PXD2B	$3,49 \times 10^{-4}$	1,56
Endothelial PAS domain protein 1	EPAS1	$1,30 \times 10^{-4}$	1,55
Glutamate receptor, ionotropic, N-methyl D-aspartate-associated	GRINA	$4,70 \times 10^{-4}$	1,55
Galactosidase, alpha	GLA	$6,98 \times 10^{-5}$	1,55
Yip1 domain family, member 1	YIPF1	$7,16 \times 10^{-5}$	1,54
Basic leucine zipper transcription factor, ATF-like	BATF	$1,28 \times 10^{-7}$	1,54
Transmembrane emp24 protein transport domain containing 8	TMED8	$6,21 \times 10^{-5}$	1,53
Ras homolog gene family, member U	RHOA	$2,14 \times 10^{-4}$	1,53
DCN1, defective in cullin neddylation 1, domain containing 3	DCUN1D3	$1,09 \times 10^{-4}$	1,52
Engulfment and cell motility 2	ELMO2	$4,91 \times 10^{-4}$	1,52
Ets variant 6	ETV6	$1,66 \times 10^{-4}$	1,52

Gene Name	Gene Symbol	P-value	Fold Change
Immediate early response 3	IER3	$3,18 \times 10^{-4}$	1,52
Immediate early response 3	IER3	$3,18 \times 10^{-4}$	1,52
Complement component 1, q subcomponent, C chain	C1QC	$3,77 \times 10^{-4}$	1,52
Low density lipoprotein receptor	LDLR	$6,67 \times 10^{-5}$	1,51
Centrosomal protein 55kDa	CEP55	$2,07 \times 10^{-4}$	1,50
Phospholipase C, eta 1	PLCH1	$1,79 \times 10^{-4}$	1,50
Chromosome 16 open reading frame 7	C16orf7	$7,48 \times 10^{-4}$	1,49
Family with sequence similarity 89, member A	FAM89A	$2,12 \times 10^{-4}$	1,49
Antizyme inhibitor 1	AZIN1	$5,26 \times 10^{-4}$	1,48
G0/G1switch 2	G0S2	$5,98 \times 10^{-4}$	1,48
Signal transducer and activator of transcription 3	STAT3	$4,72 \times 10^{-6}$	1,48
Dual specificity phosphatase 3	DUSP3	$6,86 \times 10^{-4}$	1,48
AT rich interactive domain 5A (MRF1-like)	ARID5A	$2,42 \times 10^{-6}$	1,47
Hypothetical LOC100049716	LOC100049716	$1,93 \times 10^{-4}$	1,47
Chemokine (C-C motif) ligand 2	CCL2	$2,31 \times 10^{-4}$	1,47
Mucin 1, cell surface associated	MUC1	$2,12 \times 10^{-4}$	1,47
Janus kinase 3	JAK3	$2,27 \times 10^{-4}$	1,46
Transient receptor potential cation channel, subfamily M, member 2	TRPM2	$5,48 \times 10^{-4}$	1,46
---		$5,69 \times 10^{-4}$	1,46
G protein-coupled receptor 108	GPR108	$4,28 \times 10^{-4}$	1,45
---		$1,63 \times 10^{-4}$	1,44
Nuclear factor of kappa light polypeptide gene enhancer in B-cel	NFKBIZ	$1,65 \times 10^{-4}$	1,44
Growth arrest and DNA-damage-inducible, beta	GADD45B	$5,04 \times 10^{-5}$	1,44
Exosome component 4	EXOSC4	$4,77 \times 10^{-4}$	1,44
Inositol 1,4,5-triphosphate receptor interacting protein	ITPRIP	$1,22 \times 10^{-4}$	1,44

Gene Name	Gene Symbol	P-value	Fold Change
Lemur tyrosine kinase 2	LMTK2	6,86 x 10 ⁻⁴	1,43
Neurotensin receptor 1 (high affinity)	NTSR1	3,01 x 10 ⁻⁵	1,42
Colony stimulating factor 1 (macrophage)	CSF1	2,70 x 10 ⁻⁴	1,42
Tumor protein p53 inducible protein 11	TP53I11	6,44 x 10 ⁻⁴	1,42
Protein tyrosine phosphatase, non-receptor type 1	PTPN1	1,19 x 10 ⁻⁵	1,42
---		3,49 x 10 ⁻⁴	1,42
SH3-domain binding protein 5 (BTK-associated)	SH3BP5	1,24 x 10 ⁻⁴	1,41
Eukaryotic translation initiation factor 4 gamma, 3	EIF4G3	7,49 x 10 ⁻⁴	1,41
Oncostatin M	OSM	1,39 x 10 ⁻⁴	1,41
EH-domain containing 1	EHD1	1.16 x 10 ⁻⁴	1,41
Cysteine-serine-rich nuclear protein 1	CSRNP1	1,62 x 10 ⁻⁴	1,40
Glucokinase (hexokinase 4) regulator	GCKR	1,46 x 10 ⁻⁵	1,40
Glycosyltransferase-like domain containing 1	GTDC1	6,87 x 10 ⁻⁴	1,40
Tumor necrosis factor	TNF	3,87 x 10 ⁻⁴	1,40
Tumor necrosis factor	TNF	3,87 x 10 ⁻⁴	1,40
Tumor necrosis factor	TNF	3,87 x 10 ⁻⁴	1,40
Osteoclast associated, immunoglobulin-like receptor	OSCAR	3,88 x 10 ⁻⁴	1,39
Zinc finger, DHHC-type containing 19	ZDHHC19	4,87 x 10 ⁻⁴	1,39
Ral guanine nucleotide dissociation stimulator-like 1	RGL1	2,37 x 10 ⁻⁴	1,38
Chromosome 17 open reading frame 62	C17orf62	2,35 x 10 ⁻⁴	1,38
NM_177401 // MIDN // midnolin	MIDN	5,93 x 10 ⁻⁴	1,38
Complement component 1, q subcomponent, B chain	C1QB	7,79 x 10 ⁻⁴	1,38
Mitogen-activated protein kinase kinase kinase 8	MAP3K8	2,46 x 10 ⁻⁴	1,38
MAX dimerization protein 3	MXD3	6,15 x 10 ⁻⁴	1,38

Gene Name	Gene Symbol	P-value	Fold Change
Mitochondrial translational release factor 1-like	MTRF1L	$2,73 \times 10^{-4}$	1,36
Chromosome 13 open reading frame 31	C13orf31	$7,32 \times 10^{-4}$	1,36
Lysophosphatidylcholine acyltransferase 3	LPCAT3	$7,29 \times 10^{-5}$	1,35
---		$3,71 \times 10^{-4}$	1,34
Growth arrest-specific 8	GAS8	$8,21 \times 10^{-5}$	1,34
---		$1,76 \times 10^{-4}$	1,32
Phosphoglycerate kinase 1	PGK1	$3,23 \times 10^{-4}$	1,32
Mitogen-activated protein kinase-activated protein kinase 2	MAPKAPK2	$7,86 \times 10^{-4}$	1,32
Chromosome 7 open reading frame 16	C7orf16	$2,11 \times 10^{-4}$	1,32
Triosephosphate isomerase 1	TPI1	$3,09 \times 10^{-4}$	1,31
Dual specificity phosphatase 5	DUSP5	$2,34 \times 10^{-4}$	1,31
Torsin family 1, member A (torsin A)	TOR1A	$4,12 \times 10^{-4}$	1,30
Proline rich 11	PRR11	$7,39 \times 10^{-4}$	1,30
Metallothionein 1 pseudogene 3	MT1P3	$1,00 \times 10^{-4}$	1,30
Protein phosphatase 1, regulatory (inhibitor) subunit 14B	PPP1R14B	$5,90 \times 10^{-4}$	1,30
Chromosome 17 open reading frame 53	C17orf53	$4,96 \times 10^{-4}$	1,29
Heparin-binding EGF-like growth factor	HBEGF	$2,40 \times 10^{-4}$	1,29
CD55 molecule, decay accelerating factor for complement	CD55	$1,05 \times 10^{-4}$	1,28
Poliovirus receptor-related 2 (herpesvirus entry mediator B)	PVRL2	$7,66 \times 10^{-4}$	1,28
N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 2	NDST2	$6,06 \times 10^{-4}$	1,28
Exonuclease 1	EXO1	$2,52 \times 10^{-4}$	1,28
Complement component 2	C2	$2,72 \times 10^{-4}$	1,28
Complement component 2	C2	$2,72 \times 10^{-4}$	1,28
Histone cluster 1, H3g	HIST1H3G	$2,51 \times 10^{-4}$	1,27

Gene Name	Gene Symbol	P-value	Fold Change
Replication protein A4, 30kDa	RPA4	$1,58 \times 10^{-5}$	1,27
RAP2C, member of RAS oncogene family	RAP2C	$7,06 \times 10^{-4}$	1,27
Valosin containing protein (p97)/p47 complex interacting protein	VCPIP1	$5,41 \times 10^{-4}$	1,27
Jumonji domain containing 6	JMJD6	$2,48 \times 10^{-4}$	1,26
Histone cluster 1, H2bi	HIST1H2BI	$4,86 \times 10^{-4}$	1,26
Cell division cycle associated 8	CDCA8	$2,80 \times 10^{-4}$	1,25
LIM domain only 2 (rhombotin-like 1)	LMO2	$5,45 \times 10^{-5}$	1,25
Complement component 2	C2	$4,97 \times 10^{-4}$	1,25
---		$4,80 \times 10^{-4}$	1,25
Hypothetical protein LOC100130433	LOC100130433	$4,86 \times 10^{-4}$	1,25
Mitochondrial ribosomal protein L28	MRPL28	$6,78 \times 10^{-4}$	1,25
DnaJ (Hsp40) homolog, subfamily C, member 5	DNAJC5	$5,16 \times 10^{-4}$	1,25
WW domain binding protein 5	WBP5	$5,08 \times 10^{-4}$	1,24
Serum response factor (c-fos serum response element-binding)	SRF	$2,64 \times 10^{-5}$	1,24
CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A)	CTDP1	$1,21 \times 10^{-5}$	1,24
---		$7,24 \times 10^{-5}$	1,24
MAD2 mitotic arrest deficient-like 2 (yeast)	MAD2L2	$1,94 \times 10^{-4}$	1,24
---		$1,14 \times 10^{-4}$	1,23
Microtubule-associated protein 1B	MAP1B	$4,67 \times 10^{-5}$	1,23
Activin A receptor type II-like 1	ACVRL1	$3,72 \times 10^{-4}$	1,23
Chromatin modifying protein 2A	CHMP2A	$7,45 \times 10^{-4}$	1,23
---		$7,06 \times 10^{-4}$	1,22
---		$1,73 \times 10^{-4}$	1,22
Triosephosphate isomerase 1 pseudogene 2	TPI1P2	$2,51 \times 10^{-4}$	1,22
---		$6,75 \times 10^{-4}$	1,22

Gene Name	Gene Symbol	P-value	Fold Change
ACN9 homolog (<i>S. cerevisiae</i>)	ACN9	$8,15 \times 10^{-4}$	1,21
---		$8,17 \times 10^{-5}$	1,21
G-2 and S-phase expressed 1	GTSE1	$2,70 \times 10^{-4}$	1,20
---		$1,19 \times 10^{-4}$	1,20
Cholinergic receptor, nicotinic, alpha 5	CHRNA5	$9,59 \times 10^{-6}$	1,20
G protein-coupled receptor 85	GPR85	$1,50 \times 10^{-4}$	1,20
TNFAIP3 interacting protein 2	TNIP2	$6,32 \times 10^{-4}$	1,20
Gliomedin	GLDN	$6,05 \times 10^{-4}$	1,20
---		$4,14 \times 10^{-4}$	1,19
Chromosome 1 open reading frame 61	C1orf61	$8,11 \times 10^{-4}$	1,18
---		$3,26 \times 10^{-4}$	1,18
Olfactory receptor, family 52, subfamily N, member 5	OR52N5	$4,02 \times 10^{-4}$	1,17
Aldehyde dehydrogenase 1 family, member L2	ALDH1L2	$3,27 \times 10^{-4}$	1,17
Coiled-coil domain containing 9	CCDC9	$5,10 \times 10^{-4}$	1,17
Chemokine (C-X-C motif) ligand 11	CXCL11	$4,46 \times 10^{-5}$	1,17
ATPase, H ⁺ transporting, lysosomal 9kDa, V0 subunit e1	ATP6V0E1	$5,62 \times 10^{-4}$	1,17
---		$8,20 \times 10^{-5}$	1,17
ADAM-like, decysin 1	ADAMDEC1	$1,52 \times 10^{-4}$	1,17
Kalirin, RhoGEF kinase	KALRN	$1,67 \times 10^{-4}$	1,16
---		$1,62 \times 10^{-4}$	1,16
---		$4,64 \times 10^{-4}$	1,16
---		$2,59 \times 10^{-4}$	1,16
Cholinergic receptor, nicotinic, epsilon	CHRNE	$5,42 \times 10^{-4}$	1,15
MicroRNA 7-2	MIR7-2	$6,28 \times 10^{-4}$	1,15
CLK4-associating serine/arginine rich protein	CLASRP	$3,56 \times 10^{-4}$	1,15
Family with sequence similarity 115, member A	FAM115A	$5,73 \times 10^{-4}$	1,15
---		$4,01 \times 10^{-4}$	1,15

Gene Name	Gene Symbol	P-value	Fold Change
Cytochrome P450, family 27, subfamily C, polypeptide 1	CYP27C1	6,39 x 10 ⁻⁴	1,15
Endomucin	EMCN	4,36 x 10 ⁻⁴	1,15
Gamma-glutamyltransferase 3 pseudogene	GGT3P	7,27 x 10 ⁻⁴	1,15
---		6,10 x 10 ⁻⁴	1,14
---		1,02 x 10 ⁻⁴	1,14
T, brachyury homolog (mouse)	T	2,95 x 10 ⁻⁴	1,14
Retinal pigment epithelium-specific protein 65kDa	RPE65	3,96 x 10 ⁻⁴	1,14
---		3,19 x 10 ⁻⁴	1,14
Chromosome 12 open reading frame 37	C12orf37	6,41 x 10 ⁻⁴	1,13
Coiled-coil domain containing 129	CCDC129	4,99 x 10 ⁻⁴	1,13
---		5,18 x 10 ⁻⁴	1,13
---		4,39 x 10 ⁻⁴	1,13
---		3,71 x 10 ⁻⁴	1,13
Olfactory receptor, family 6, subfamily C, member 6	OR6C6	5,54 x 10 ⁻⁴	1,13
Solute carrier family 6 (amino acid transporter), member 14	SLC6A14	1,30 x 10 ⁻⁴	1,13
Olfactory receptor, family 5, subfamily D, member 14	OR5D14	5,66 x 10 ⁻⁴	1,12
Angiotensin I converting enzyme (peptidyl-dipeptidase A) 2	ACE2	7,13 x 10 ⁻⁴	1,12
---		4,81 x 10 ⁻⁴	1,12
Leber congenital amaurosis 5-like	LCA5L	1,40 x 10 ⁻⁴	1,12
Solute carrier family 15, member 5	SLC15A5	1,15 x 10 ⁻⁵	1,12
Chromosome 1 open reading frame 137	C1orf137	8,14 x 10 ⁻⁴	1,12
Glycoprotein (transmembrane) nmb	GPNMB	5,23 x 10 ⁻⁴	1,12
---		8,19 x 10 ⁻⁴	1,11
Annexin A13	ANXA13	5,01 x 10 ⁻⁴	1,11
---		7,49 x 10 ⁻⁴	1,11

Gene Name	Gene Symbol	P-value	Fold Change
---		$7,60 \times 10^{-5}$	1,11
SPANX family, member N4	SPANXN4	$6,95 \times 10^{-4}$	1,11
---		$1,39 \times 10^{-4}$	1,11
Zinc finger protein 679	ZNF679	$4,44 \times 10^{-4}$	1,11
Chromosome 16 open reading frame 78	C16orf78	$7,34 \times 10^{-4}$	1,11
---		$8,65 \times 10^{-6}$	1,11
---		$3,93 \times 10^{-4}$	1,10
Glycine receptor, alpha 3	GLRA3	$4,03 \times 10^{-4}$	1,10
---		$9,01 \times 10^{-5}$	1,10
---		$3,33 \times 10^{-4}$	1,09
---		$4,63 \times 10^{-4}$	1,08
Methenyltetrahydrofolate synthetase domain containing	MTHFSD	$6,39 \times 10^{-4}$	-1,11
EP400 N-terminal like	EP400NL	$3,49 \times 10^{-4}$	-1,11
O-6-methylguanine-DNA methyltransferase	MGMT	$2,46 \times 10^{-5}$	-1,15
Molybdenum cofactor synthesis 3	MOCS3	$2,26 \times 10^{-4}$	-1,155
Ribosome binding factor A (putative)	RBFA	$4,00 \times 10^{-4}$	-1,16
Trypsin domain containing 1	TYSND1	$4,45 \times 10^{-4}$	-1,17
PTEN induced putative kinase 1	PINK1	$2,51 \times 10^{-4}$	-1,17
LysM, putative peptidoglycan-binding, domain containing 4	LYSMD4	$5,81 \times 10^{-4}$	-1,17
Retinoic acid receptor, gamma	RARG	$2,17 \times 10^{-4}$	-1,17
Chibby homolog 1 (Drosophila)	CBY1	$5,07 \times 10^{-4}$	-1,18
Protein kinase C, zeta	PRKCZ	$2,08 \times 10^{-4}$	-1,18
Vesicle-associated membrane protein 2 (synaptobrevin 2)	VAMP2	$1,86 \times 10^{-4}$	-1,18
BAI1-associated protein 3	BAIAP3	$2,26 \times 10^{-4}$	-1,18
Translocase of outer mitochondrial membrane 7 homolog (yeast)	TOMM7	$3,82 \times 10^{-4}$	-1,19
Ribosomal protein L19	RPL19	$3,16 \times 10^{-4}$	-1,19
Peptidylprolyl isomerase A (cyclophilin A) pseudogene	LOC342541	$6,97 \times 10^{-4}$	-1,19

Gene Name	Gene Symbol	P-value	Fold Change
Pyridoxal-dependent decarboxylase domain containing 2	PDXDC2P	4,84 x 10 ⁻⁴	-1,19
Small nuclear ribonucleoprotein polypeptide N pseudogene	LOC100129534	2,82 x 10 ⁻⁴	-1,19
Pyridoxal-dependent decarboxylase domain containing 2	PDXDC2P	6,06 x 10 ⁻⁴	-1,19
Chromosome 4 open reading frame 39	C4orf39	2,74 x 10 ⁻⁴	-1,19
Minichromosome maintenance complex component 3 associated protein	MCM3AP	2,48 x 10 ⁻⁴	-1,19
SWI/SNF related, matrix associated, actin dependent regulator	SMARCC2	7,15 x 10 ⁻⁴	-1,19
Component of oligomeric golgi complex 1	COG1	6,09 x 10 ⁻⁴	-1,19
Proline rich 12	PRR12	2,26 x 10 ⁻⁴	-1,19
Pyridoxal-dependent decarboxylase domain containing 2	PDXDC2P	6,27 x 10 ⁻⁴	-1,20
---		7,55 x 10 ⁻⁴	-1,20
Polycystic kidney disease 1 (autosomal dominant) pseudogene 1	PKD1P1	1,37 x 10 ⁻⁴	-1,20
Tectonic family member 1	TCTN1	7,73 x 10 ⁻⁴	-1,21
B-cell translocation gene 1, anti-proliferative	BTG1	3,88 x 10 ⁻⁴	-1,21
SH3-binding domain kinase 1	SBK1	3,85 x 10 ⁻⁴	-1,21
---		4,07 x 10 ⁻⁵	-1,21
Hypothetical LOC93622	LOC93622	5,07 x 10 ⁻⁴	-1,21
ADP-ribosylation factor-like 5A	ARL5A	1,97 x 10 ⁻⁴	-1,21
RAB39B, member RAS oncogene family	RAB39B	5,60 x 10 ⁻⁴	-1,21
COX11 cytochrome c oxidase assembly homolog (yeast)	COX11	7,11 x 10 ⁻⁴	-1,22
Protein O-linked mannose beta1,2-N-acetylglucosaminyltransferase	POMGNT1	6,11 x 10 ⁻⁴	-1,22

Gene Name	Gene Symbol	P-value	Fold Change
Stathmin-like 3	STMN3	$1,07 \times 10^{-4}$	-1,23
MyoD family inhibitor domain containing	MDFIC	$8,89 \times 10^{-5}$	-1,23
BTB (POZ) domain containing 6	BTBD6	$5,73 \times 10^{-5}$	-1,23
SEC31 homolog A (S. cerevisiae)	SEC31A	$7,37 \times 10^{-4}$	-1,23
FLJ41327 protein	FLJ41327	$2,56 \times 10^{-5}$	-1,23
Relaxin 1	RLN1	$1,11 \times 10^{-4}$	-1,23
Family with sequence similarity 43, member A	FAM43A	$2,25 \times 10^{-4}$	-1,24
Ras association (RalGDS/AF-6) domain family member 1	RASSF1	$6,10 \times 10^{-4}$	-1,24
Transmembrane protein 177	TMEM177	$6,90 \times 10^{-4}$	-1,24
---		$2,68 \times 10^{-4}$	-1,24
Mitochondrial ribosomal protein S7	MRPS7	$3,75 \times 10^{-4}$	-1,24
AarF domain containing kinase 3	ADCK3	$6,42 \times 10^{-4}$	-1,24
---		$1,62 \times 10^{-4}$	-1,24
Chromosome 7 open reading frame 70	C7orf70	$7,96 \times 10^{-4}$	-1,25
Plexin A3	PLXNA3	$5,77 \times 10^{-4}$	-1,25
EPH receptor B6	EPHB6	$2,68 \times 10^{-5}$	-1,25
Dipeptidase 3	DPEP3	$6,87 \times 10^{-4}$	-1,25
---		$3,83 \times 10^{-4}$	-1,25
Lysophosphatidic acid receptor 5	LPAR5	$4,87 \times 10^{-4}$	-1,26
Interleukin 16 (lymphocyte chemoattractant factor)	IL16	$6,10 \times 10^{-4}$	-1,26
Chromosome 14 open reading frame 182	C14orf182	$7,39 \times 10^{-4}$	-1,26
F-box protein 25	FBXO25	$4,05 \times 10^{-4}$	-1,27
Tetratricopeptide repeat domain 9	TTC9	$7,25 \times 10^{-4}$	-1,27
Nucleolar protein family 6 (RNA-associated)	NOL6	$7,60 \times 10^{-4}$	-1,27
Small EDRK-rich factor 1A (telomeric)	SERF1A	$5,07 \times 10^{-4}$	-1,27
Small EDRK-rich factor 1A (telomeric)	SERF1A	$5,07 \times 10^{-4}$	-1,27

Gene Name	Gene Symbol	P-value	Fold Change
Small EDRK-rich factor 1A (telomeric)	SERF1A	$5,07 \times 10^{-4}$	-1,27
General transcription factor IIIH, polypeptide 4, 52kDa	GTF2H4	$4,24 \times 10^{-5}$	-1,27
WD repeat domain 27	WDR27	$7,24 \times 10^{-4}$	-1,27
Family with sequence similarity 113, member A	FAM113A	$2,29 \times 10^{-4}$	-1,27
Poly(A) binding protein interacting protein 2B	PAIP2B	$7,06 \times 10^{-4}$	-1,28
Fat storage-inducing transmembrane protein 2	FITM2	$4,42 \times 10^{-4}$	-1,28
---		$2,09 \times 10^{-4}$	-1,28
---		$3,00 \times 10^{-4}$	-1,28
Guanine nucleotide binding protein (G protein), gamma 7	GNG7	$2,56 \times 10^{-4}$	-1,29
Zinc finger protein 64 homolog (mouse)	ZFP64	$5,89 \times 10^{-4}$	-1,29
Chromosome 16 open reading frame 74	C16orf74	$7,90 \times 10^{-5}$	-1,29
Chromosome 3 open reading frame 57	C3orf57	$6,35 \times 10^{-5}$	-1,29
Ribosomal protein S25	RPS25	$4,98 \times 10^{-4}$	-1,30
Polycystic kidney disease 1 (autosomal dominant) pseudogene 1	PKD1P1	$1,97 \times 10^{-5}$	-1,30
Polycystic kidney disease 1 (autosomal dominant) pseudogene 1	PKD1P1	$3,76 \times 10^{-5}$	-1,30
Coenzyme Q10 homolog A (S. cerevisiae)	COQ10A	$6,63 \times 10^{-4}$	-1,30
Nuclear pore complex interacting protein	NPIP	$6,24 \times 10^{-6}$	-1,30
Protein arginine methyltransferase 7	PRMT7	$6,89 \times 10^{-4}$	-1,30
Granzyme M (lymphocyte met-ase 1)	GZMM	$3,38 \times 10^{-4}$	-1,30
Olfactomedin 2	OLFM2	$4,75 \times 10^{-4}$	-1,30
Nuclear pore complex interacting protein	NPIP	$5,90 \times 10^{-6}$	-1,31
B-cell CLL/lymphoma 9-like	BCL9L	$4,04 \times 10^{-4}$	-1,31

Gene Name	Gene Symbol	P-value	Fold Change
Zinc finger protein 90	ZNF90	4,90 x 10 ⁻⁴	-1,311
---		1,48 x 10 ⁻⁴	-1,311
StAR-related lipid transfer (START) domain containing 9	STARD9	7,42 x 10 ⁻⁴	-1,311
ALS2 C-terminal like	ALS2CL	1,31 x 10 ⁻⁴	-1,32
Polycystic kidney disease 1 pseudogene 1	PKD1P1	6,17 x 10 ⁻⁵	-1,32
Eukaryotic translation initiation factor 3, subunit H	EIF3H	5,06 x 10 ⁻⁴	-1,32
Forkhead box P1	FOXP1	2,09 x 10 ⁻⁴	-1,32
O-linked N-acetylglucosamine (GlcNAc) transferase	OGT	3,53 x 10 ⁻⁴	-1,33
---		4,91 x 10 ⁻⁵	-1,33
---		2,95 x 10 ⁻⁴	-1,33
Class II, major histocompatibility complex, transactivator	CIITA	4,99 x 10 ⁻⁴	-1,333
Tripartite motif-containing 32	TRIM32	6,79 x 10 ⁻⁴	-1,33
Nuclear factor of activated T-cells, cytoplasmic, calcineurin-de	NFATC3	4,48 x 10 ⁻⁴	-1,33
Hypothetical FLJ38379	FLJ38379	6,15 x 10 ⁻⁴	-1,33
Ribosomal protein L18a	RPL18A	1,23 x 10 ⁻⁴	-1,33
EPH receptor A1	EPHA1	5,87 x 10 ⁻⁵	-1,33
Zinc finger protein 41	ZNF41	8,18 x 10 ⁻⁴	-1,34
Major histocompatibility complex, class II, DQ alpha 2	HLA-DQA2	5,04 x 10 ⁻⁴	-1,34
RGD motif, leucine rich repeats, tropomodulin domain	RLTPR	1,28 x 10 ⁻⁴	-1,34
Phospholipase D family, member 4	PLD4	2,50 x 10 ⁻⁴	-1,34
Fumarylacetoacetate hydrolase domain containing 2A	FAHD2A	1,00 x 10 ⁻⁴	-1,34
Ribosomal protein L17	RPL17	5,18 x 10 ⁻⁴	-1,35
Intermediate filament family orphan 2	IFFO2	1,41 x 10 ⁻⁵	-1,35
Ribosomal protein L17	RPL17	8,13 x 10 ⁻⁴	-1,35
Methylthioribose-1-phosphate isomerase homolog (S. cerevisiae)	MRI1	2,26 x 10 ⁻⁴	-1,35

Gene Name	Gene Symbol	P-value	Fold Change
Ataxin 7-like 1	ATXN7L1	$7,58 \times 10^{-4}$	-1,35
Major histocompatibility complex, class II, DQ alpha 2	HLA-DQA2	$6,21 \times 10^{-4}$	-1,35
---		$3,72 \times 10^{-4}$	-1,35
F-box and WD repeat domain containing 8	FBXW8	$1,82 \times 10^{-4}$	-1,35
Ribosomal protein S18 pseudogene 9	RPS18P9	$2,83 \times 10^{-4}$	-1,35
Ribosomal protein L18a	RPL18A	$7,75 \times 10^{-4}$	-1,36
Zinc finger family member 767	ZNF767	$6,47 \times 10^{-4}$	-1,36
EF-hand domain containing 1	EFHC1	$2,96 \times 10^{-4}$	-1,36
---		$2,75 \times 10^{-4}$	-1,37
Zinc finger protein 135	ZNF135	$4,89 \times 10^{-4}$	-1,37
---		$5,65 \times 10^{-4}$	-1,37
Zinc finger protein 37B, pseudogene	ZNF37BP	$7,52 \times 10^{-4}$	-1,38
---		$2,55 \times 10^{-4}$	-1,38
RNA binding motif protein 14	RBM14	$5,68 \times 10^{-4}$	-1,39
Syntaxin binding protein 4	STXBP4	$6,37 \times 10^{-4}$	-1,39
Chromosome 7 open reading frame 58	C7orf58	$1,82 \times 10^{-4}$	-1,39
Rho/Rac guanine nucleotide exchange factor (GEF) 18	ARHGEF18	$3,28 \times 10^{-4}$	-1,39
---		$4,71 \times 10^{-4}$	-1,39
Myosin phosphatase Rho interacting protein	MPRIP	$4,65 \times 10^{-5}$	-1,39
ADAM metallopeptidase with thrombospondin type 1 motif, 10	ADAMTS10	$5,23 \times 10^{-4}$	-1,39
Ribosomal protein L39	RPL39	$3,07 \times 10^{-4}$	-1,39
NOP2/Sun domain family, member 5 pseudogene 2	NSUN5P2	$2,27 \times 10^{-4}$	-1,40
Chromosome 10 open reading frame 88	C10orf88	$5,25 \times 10^{-4}$	-1,40
Ribosomal protein S10	RPS10	$3,25 \times 10^{-4}$	-1,40
Small nucleolar RNA, C/D box 105	SNORD105	$4,10 \times 10^{-4}$	-1,40
---		$5,63 \times 10^{-4}$	-1,40

Gene Name	Gene Symbol	P-value	Fold Change
Runt-related transcription factor 2	RUNX2	4,00 x 10⁻⁴	-1,40
Ribosomal protein L39	RPL39	3,05 x 10⁻⁴	-1,40
---		1,29 x 10 ⁻⁵	-1,41
Ribosomal protein S8	RPS8	6,67 x 10⁻⁴	-1,42
Junction plakoglobin	JUP	7,00 x 10⁻⁴	-1,42
Polycystic kidney disease 1 (autosomal dominant)	PKD1	6,11 x 10⁻⁵	-1,42
Patched 1	PTCH1	1,63 x 10⁻⁴	-1,42
Histone deacetylase 9	HDAC9	5,31 x 10⁻⁴	-1,42
Ligatin	LGTN	3,84 x 10⁻⁴	-1,42
Tet oncogene 1	TET1	9,02 x 10⁻⁵	-1,42
NLR family, CARD domain containing 3	NLRC3	6,33 x 10⁻⁵	-1,43
---		7,00 x 10 ⁻⁴	-1,43
---		1,07 x 10 ⁻⁴	-1,43
Glucuronidase beta pseudogene 3	GUSBP3	7,45 x 10⁻⁴	-1,44
G protein-coupled receptor 114	GPR114	4,02 x 10⁻⁴	-1,44
Integrin alpha FG-GAP repeat containing 2	ITFG2	7,22 x 10⁻⁴	-1,44
Glucuronidase, beta pseudogene	SMA5	6,75 x 10⁻⁴	-1,44
---		6,35 x 10 ⁻⁴	-1,44
---		7,42 x 10 ⁻⁴	-1,45
Glucuronidase, beta pseudogene 1	GUSBP1	4,98 x 10⁻⁴	-1,45
Glucuronidase, beta pseudogene	SMA5	6,13 x 10⁻⁴	-1,45
RNA binding motif protein 14	RBM14	4,68 x 10⁻⁴	-1,45
B-cell CLL/lymphoma 2	BCL2	5,75 x 10⁻⁴	-1,45
5'-nucleotidase, ecto (CD73)	NT5E	8,03 x 10⁻⁶	-1,46
---		5,43 x 10 ⁻⁴	-1,46
Yippee-like 1 (Drosophila)	YPEL1	8,37 x 10⁻⁶	-1,46
Inhibitor of DNA binding 3, dominant negative helix-loop-helix	ID3	5,09 x 10⁻⁵	-1,46
Glucuronidase beta pseudogene 3	GUSBP3	7,04 x 10⁻⁴	-1,46

Gene Name	Gene Symbol	P-value	Fold Change
Progesterin and adipoQ receptor family member VIII	PAQR8	$1,01 \times 10^{-4}$	-1,46
Glucuronidase beta pseudogene 3	GUSBP3	$6,92 \times 10^{-4}$	-1,46
T cell receptor beta constant 1	TRBC1	$5,17 \times 10^{-4}$	-1,46
Inositol 1,4,5-trisphosphate 3-kinase B	ITPKB	$3,22 \times 10^{-4}$	-1,46
Galactose-3-O-sulfotransferase 4	GAL3ST4	$3,98 \times 10^{-5}$	-1,47
NOP2/Sun domain family, member 5 pseudogene 2	NSUN5P2	$1,74 \times 10^{-4}$	-1,47
Chromosome 2 open reading frame 89	C2orf89	$7,45 \times 10^{-4}$	-1,47
Single immunoglobulin and toll-interleukin 1 receptor (TIR) d	SIGIRR	$1,24 \times 10^{-4}$	-1,47
Neurocalcin delta	NCALD	$3,66 \times 10^{-5}$	-1,48
Fc fragment of IgE, low affinity II, receptor for (CD23)	FCER2	$2,47 \times 10^{-4}$	-1,48
Glucuronidase, beta pseudogene	SMA5	$6,09 \times 10^{-4}$	-1,48
---		$6,23 \times 10^{-4}$	-1,48
Ribosomal protein L34	RPL34	$1,22 \times 10^{-4}$	-1,48
Eomesodermin	EOMES	$7,08 \times 10^{-6}$	-1,49
Purinergic receptor P2Y, G-protein coupled, 8	P2RY8	$3,84 \times 10^{-4}$	-1,49
Purinergic receptor P2Y, G-protein coupled, 8	P2RY8	$3,84 \times 10^{-4}$	-1,49
Nuclear pore complex interacting protein	NPIP	$7,62 \times 10^{-6}$	-1,49
Tetratricopeptide repeat domain 21B	TTC21B	$7,82 \times 10^{-4}$	-1,50
Amino-terminal enhancer of split	AES	$3,27 \times 10^{-4}$	-1,50
Mannosidase, alpha, class 1C, member 1	MAN1C1	$07,31_4 \times 10^{-7}$	-1,51
---		$5,83 \times 10^{-4}$	-1,51
---		$1,37 \times 10^{-4}$	-1,51
Chromosome 11 open reading frame 61	C11orf61	$9,63 \times 10^{-6}$	-1,51
Transmembrane protein 135	TMEM135	$3,53 \times 10^{-4}$	-1,52
Small Cajal body-specific RNA 17	SCARNA17	$7,42 \times 10^{-4}$	-1,53

Gene Name	Gene Symbol	P-value	Fold Change
Family with sequence similarity 159, member A	FAM159A	$3,19 \times 10^{-4}$	-1,54
---		$8,35 \times 10^{-5}$	-1,56
---		$1,26 \times 10^{-4}$	-1,56
B-cell CLL/lymphoma 11A (zinc finger protein)	BCL11A	$6,01 \times 10^{-4}$	-1,57
TRAF3 interacting protein 3	TRAF3IP3	$8,14 \times 10^{-4}$	-1,577
Tubulin polymerization-promoting protein family member 3	TPPP3	$3,35 \times 10^{-4}$	-1,59
---		$3,54 \times 10^{-4}$	-1,59
---		$1,35 \times 10^{-4}$	-1,60
SET and MYND domain containing 2	SMYD2	$4,32 \times 10^{-4}$	-1,60
Glutaredoxin-like protein YDR286C homolog	FLJ44606	$1,90 \times 10^{-4}$	-1,62
Cell division cycle 25 homolog B (S. pombe)	CDC25B	$5,12 \times 10^{-4}$	-1,62
Nuclear factor of activated T-cells, cytoplasmic	NFATC2	$5,78 \times 10^{-4}$	-1,63
RUN domain containing 2C	RUNDC2C	$1,86 \times 10^{-4}$	-1,63
Cytotoxic and regulatory T cell molecule	CRTAM	$1,79 \times 10^{-4}$	-1,64
Zeta-chain (TCR) associated protein kinase 70kDa	ZAP70	$4,84 \times 10^{-4}$	-1,65
Cerebellar degeneration-related protein 2, 62kDa	CDR2	$1,09 \times 10^{-5}$	-1,65
Major histocompatibility complex, class II, DM beta	HLA-DMB	$7,06 \times 10^{-4}$	-1,65
Fibroblast growth factor binding protein 2	FGFBP2	$4,62 \times 10^{-4}$	-1,65
Septin 1	SEPT1	$7,37 \times 10^{-4}$	-1,66
T cell receptor alpha variable 8-3	TRAV8-3	$3,91 \times 10^{-4}$	-1,66
---		$3,19 \times 10^{-4}$	-1,67
Transmembrane protein 204	TMEM204	$3,77 \times 10^{-4}$	-1,70
B-cell CLL/lymphoma 11B (zinc finger protein)	BCL11B	$2,66 \times 10^{-4}$	-1,70
Solute carrier family 7 (cationic amino acid transporter, y+ sys	SLC7A6	$8,16 \times 10^{-4}$	-1,70

Gene Name	Gene Symbol	P-value	Fold Change
Family with sequence similarity 117, member B	FAM117B	$2,69 \times 10^{-4}$	-1,72
Non-protein coding RNA 287	NCRNA00287	$1,54 \times 10^{-4}$	-1,72
RUN domain containing 2C	RUNDC2C	$2,37 \times 10^{-4}$	-1,72
Fc receptor-like 6	FCRL6	$2,30 \times 10^{-4}$	-1,73
Signal-regulatory protein gamma	SIRPG	$3,84 \times 10^{-4}$	-1,74
GRB2-related adaptor protein-like	GRAPL	$6,31 \times 10^{-4}$	-1,74
T cell receptor alpha joining 17	TRAJ17	$1,42 \times 10^{-4}$	-1,77
Noggin	NOG	$3,78 \times 10^{-5}$	-1,77
CD160 molecule	CD160	$4,88 \times 10^{-5}$	-1,80
Killer cell lectin-like receptor subfamily D, member 1	KLRD1	$5,53 \times 10^{-4}$	-1,80
TBC1 domain family, member 9	TBC1D9	$1,19 \times 10^{-4}$	-1,81
ATP-binding cassette, sub-family D (ALD), member 2	ABCD2	$7,17 \times 10^{-4}$	-1,84
CD1c molecule	CD1C	$2,25 \times 10^{-4}$	-1,85
KIAA0748	KIAA0748	$5,00 \times 10^{-4}$	-1,85
Lymphocyte antigen 9	LY9	$9,28 \times 10^{-5}$	-1,89
Prostaglandin D2 receptor (DP)	PTGDR	$1,27 \times 10^{-4}$	-1,91
Oxysterol binding protein-like 10	OSBPL10	$5,53 \times 10^{-4}$	-1,921
Killer cell lectin-like receptor subfamily B, member 1	KLRB1	$2,01 \times 10^{-4}$	-1,93
CD3e molecule, epsilon (CD3-TCR complex)	CD3E	$7,83 \times 10^{-4}$	-1,96
GRB2-related adaptor protein	GRAP	$4,46 \times 10^{-5}$	-1,96
Ring finger protein 125	RNF125	148×10^{-4}	-2,00
---		$1,49 \times 10^{-4}$	-2,026
---		$7,42 \times 10^{-4}$	-2,06
Zinc finger protein 860	ZNF860	$5,27 \times 10^{-4}$	-2,10
Major histocompatibility complex, class II, DQ alpha 1	HLA-DQA1	$6,22 \times 10^{-4}$	-2,16
Major histocompatibility complex, class II, DQ alpha 1	HLA-DQA1	$6,22 \times 10^{-4}$	-2,16
Sterile alpha motif domain containing 3	SAMD3	$3,17 \times 10^{-4}$	-2,19

Gene Name	Gene Symbol	P-value	Fold Change
CD40 ligand	CD40LG	6,90 x 10⁻⁵	-2,21
Adenylate kinase 5	AK5	1,14 x 10⁻⁴	-2,27
Killer cell lectin-like receptor subfamily G, member 1	KLRG1	4,79 x 10⁻⁴	-2,28
Killer cell lectin-like receptor subfamily F, member 1	KLRF1	3,40 x 10⁻⁴	-2,29
NEL-like 2 (chicken)	NELL2	6,42 x 10⁻⁴	-2,29
---		5,71 x 10⁻⁵	-2,33
---		3,63 x 10⁻⁴	-2,37
Granzyme K (granzyme 3; tryptase II)	GZMK	2,86 x 10⁻⁴	-2,39
Killer cell lectin-like receptor subfamily K, member 1	KLRK1	5,87 x 10⁻⁴	-2,50
Leucine rich repeat neuronal 3	LRRN3	1,19 x 10⁻⁴	-2,87