

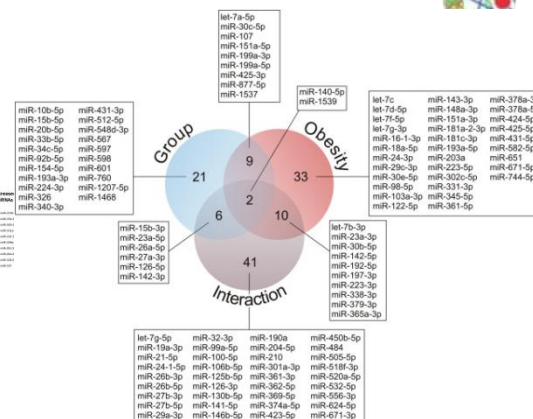
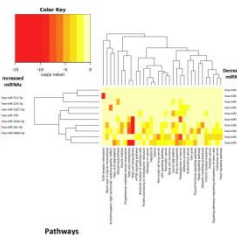
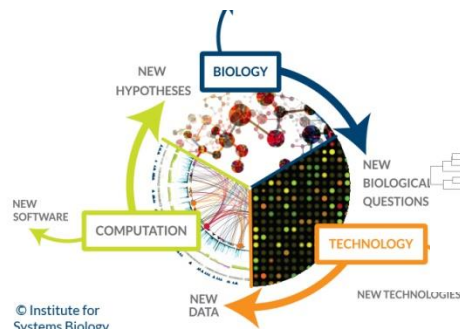
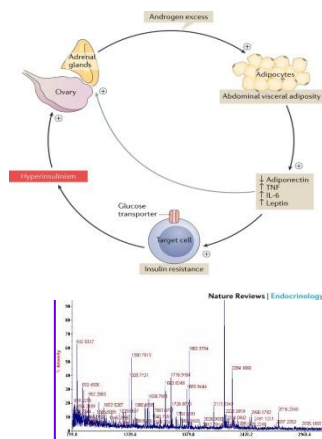
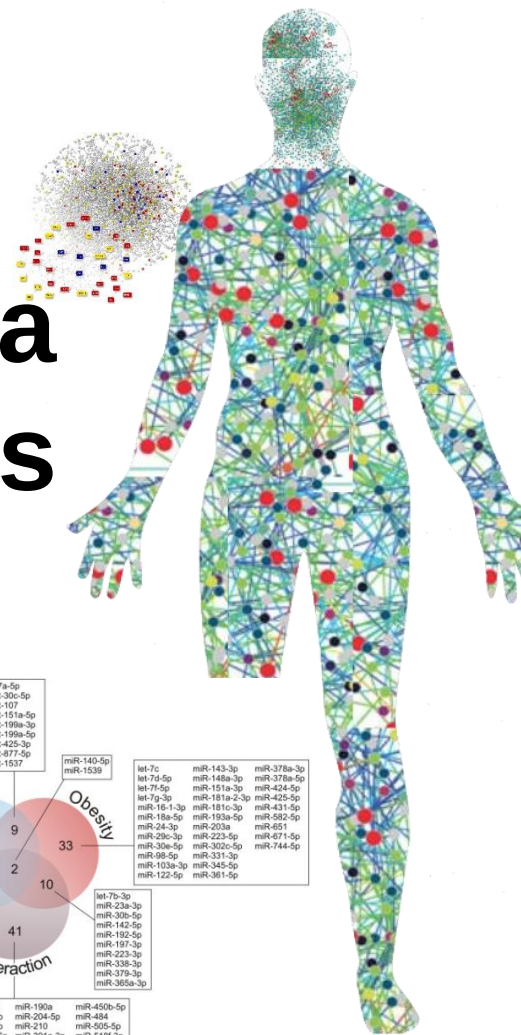


Doctorado en Ingeniería Informática

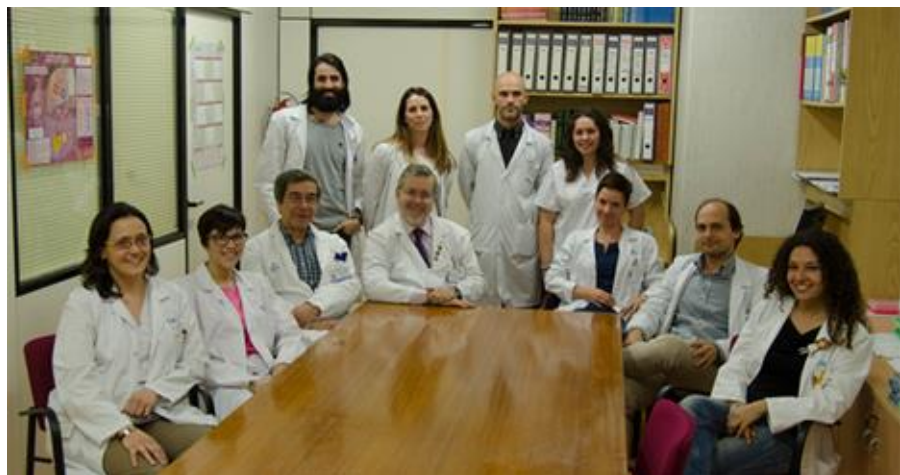
Curso 2018-2019

Aplicación de técnicas estadísticas e inteligencia artificial en enfermedades metabólicas complejas

María Insenser Nieto



AGRADECIMIENTOS



Héctor Escobar-Morreale
Manuel Luque Ramírez
M^a Ángeles Martínez
Elena Fernández
Samuel Moncayo
Verónica de Dios
Rafael Montes Nieto
Mora Murri
Miriam Ojeda



Grupo Diabetes, Obesidad y Reproducción Humana



María Guijarro Mata-García
Víctor Javier Cerquera Parra
Juan Jiménez Castellamos



Grupo Ingeniería de Sistemas, Control, Automatización y Robótica

 **INTRODUCCIÓN A LA INVESTIGACIÓN BIOMÉDICA**

 **ENFERMEDADES METABÓLICAS COMPLEJAS**

 **APLICACIÓN DE TÉCNICAS ESTADÍSTICAS E**

INTELIGENCIA ARTIFICIAL

INVESTIGACIÓN BIOMÉDICA



Profesionales sanitarios, investigadores básicos e investigadores clínicos.

Objetivos

- Mejorar el diagnóstico, tratamiento y prevención de enfermedades.
- Mejorar de la calidad de vida.
- Mejorar la innovación en salud.

Institutos de Investigación Sanitaria

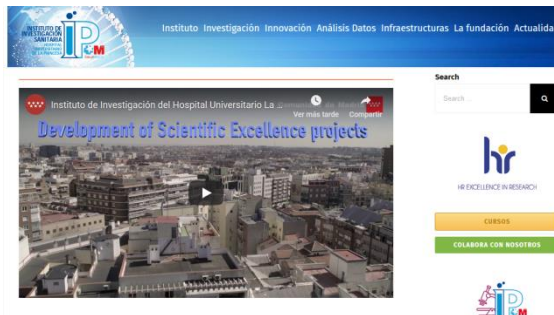


Colaboración e investigadores

Formación y docencia

Difusión de resultados

Ayudas y convocatorias



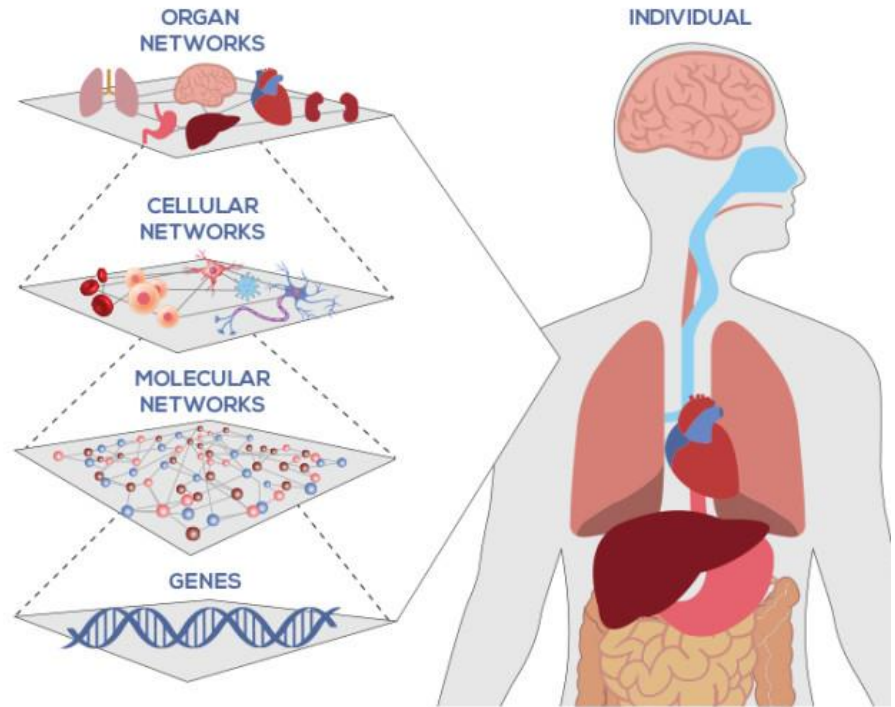
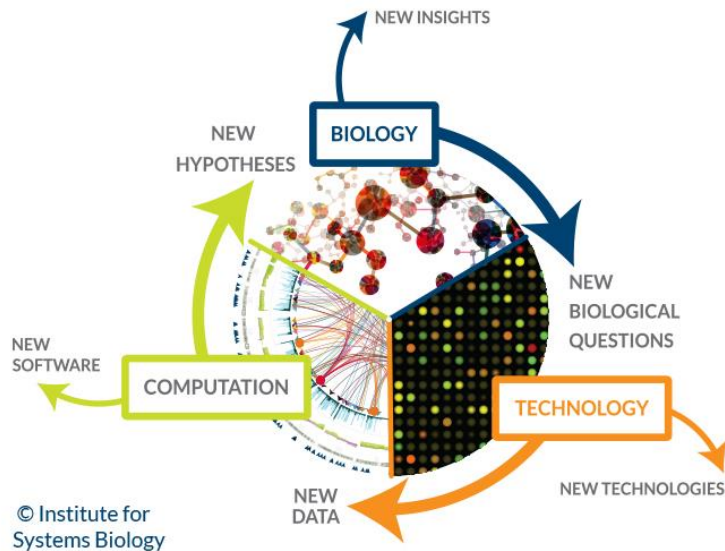
Centro de Investigación Biomédica en Red

The screenshot shows the website of the Centro de Investigación Biomédica en Red (CIBER ISCIII). The header includes the logo 'ciber isciii' and the text 'Centro de Investigación Biomédica en Red'. A search bar is located on the right. Below the header is a navigation menu with links: QUIÉNES SOMOS, ÁREAS TEMÁTICAS, EMPLEO, PERFIL CONTRATANTE, CONVOCATORIAS, TRANSFERENCIA TECNOLÓGICA, and COMUNICACIÓN. The main section is titled 'Áreas Temáticas' and displays 12 thematic areas in a grid:

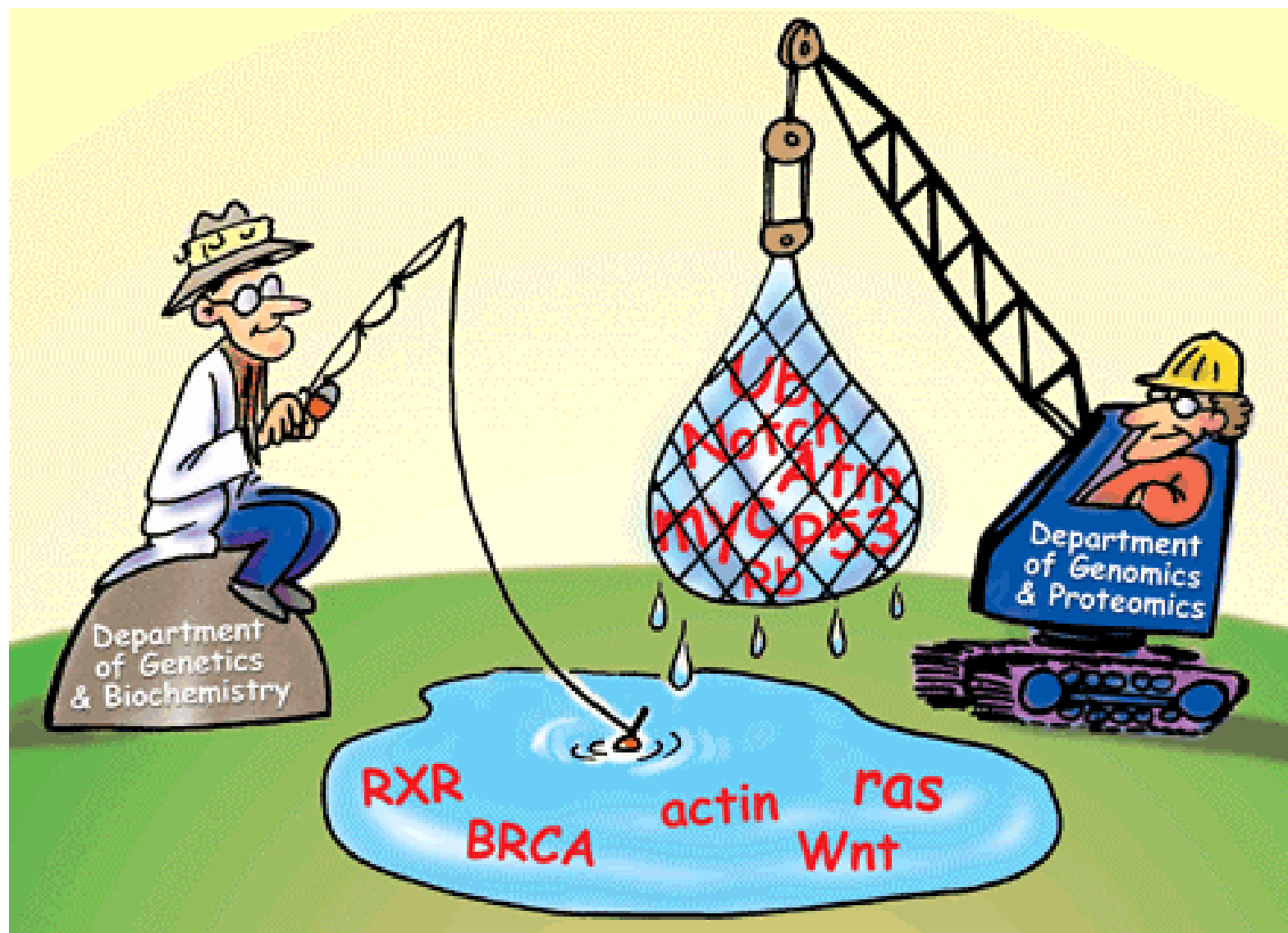
- ciber-bbn**: El CIBER en su área temática de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) está formado actualmente por 47 grupos de investigación... [Seguir leyendo](#)
- ciber-cv**: El CIBER en su área temática Enfermedades Cardiovasculares (CIBERCV) lo forman 40 grupos de investigación seleccionados sobre la base de su excelencia científica pertenecientes... [Seguir leyendo](#)
- ciber-dem**: El CIBER en su área temática de Diabetes y Enfermedades Metabólicas asociadas (CIBERDEM) tiene como principal objetivo liderar el esfuerzo investigador de excelencia en diabetes y enfermedades metabólicas asociadas... [Seguir leyendo](#)
- ciber-ehd**: El CIBER en su área temática de Enfermedades Hepáticas y Digestivas (CIBEREHd) tiene como finalidad la promoción y protección de la salud... [Seguir leyendo](#)
- ciber-er**: El área temática de Enfermedades Raras (CIBERER) creada para servir de referencia, coordinar y potenciar la investigación sobre las enfermedades... [Seguir leyendo](#)
- ciber-es**: En el área temática de Enfermedades Respiratorias (CIBERES), de carácter multidisciplinar y multi-institucional, su finalidad es combatir las enfermedades... [Seguir leyendo](#)
- ciber-es**: El CIBER en su área temática de Epidemiología y Salud Pública (CIBERESP) centra sus actividades en dos aspectos clave: conocer la magnitud y la distribución de los problemas de salud pública e... [Seguir leyendo](#)
- ciber-fes**: El CIBERFES, área temática CIBER de Fragilidad y Envejecimiento Saludable ha sido creada a finales de 2016 con el objetivo de entender, evaluar y paliar, en... [Seguir leyendo](#)
- ciber-obn**: El CIBER en su área temática de Fisiopatología de la Obesidad y Nutrición (CIBEROBN) es un consorcio... [Seguir leyendo](#)
- ciber-onc**: El área temática de Cáncer (CIBERONC) creada a finales de 2016 está formada por 50 grupos de investigación... [Seguir leyendo](#)
- ciber-sam**: El área temática de Salud Mental (CIBERSAM) está formada por 23 grupos de investigación clínica, preclínica y... [Seguir leyendo](#)

Consortio público de investigación creado por iniciativa del Instituto de Salud Carlos III (ISCIII) para impulsar la investigación de excelencia en Biomedicina y Ciencias de la Salud que se realiza en el Sistema Nacional de Salud y en el Sistema de Ciencia y Tecnología.

Biología de Sistemas



Systems biology is collaborative, integrating many scientific disciplines – biology, computer science, engineering, bioinformatics, physics and others – to predict how these systems change over time and under varying conditions, and to develop solutions to the world's most pressing health and environmental issues. <https://www.systemsbiology.org/about/what-is-systems-biology/>₇



Biología de Sistemas

What can happen

GENOME

Bioinformatics

What appears to be happening

TRANSCRIPTOME

What makes it happen

PROTEOME

What actually happens

Systems Biology

METABOLOME

PHENOTYPE

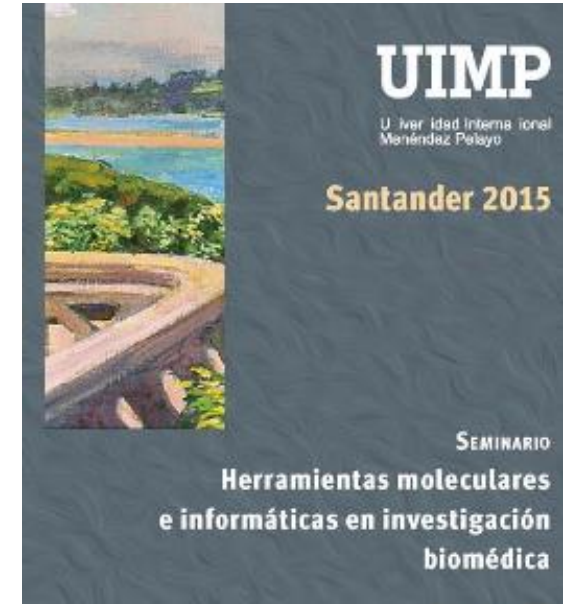
WHY WE CARE!

Systems Biology is the unified interconnection of multi-omics data to statistically model and predict, or dynamically simulate, and visualize functional gene-environment interactions.

La medicina personalizada y el big data clínico centran unas jornadas en el Instituto de Biomedicina de Sevilla



Terapias innovadoras, investigación traslacional y bioinformática perfilan el futuro de las enfermedades raras



YouTube

Buscar



Bioinformática: de la eSalud a la Inteligencia Artificial Convergencia Salud

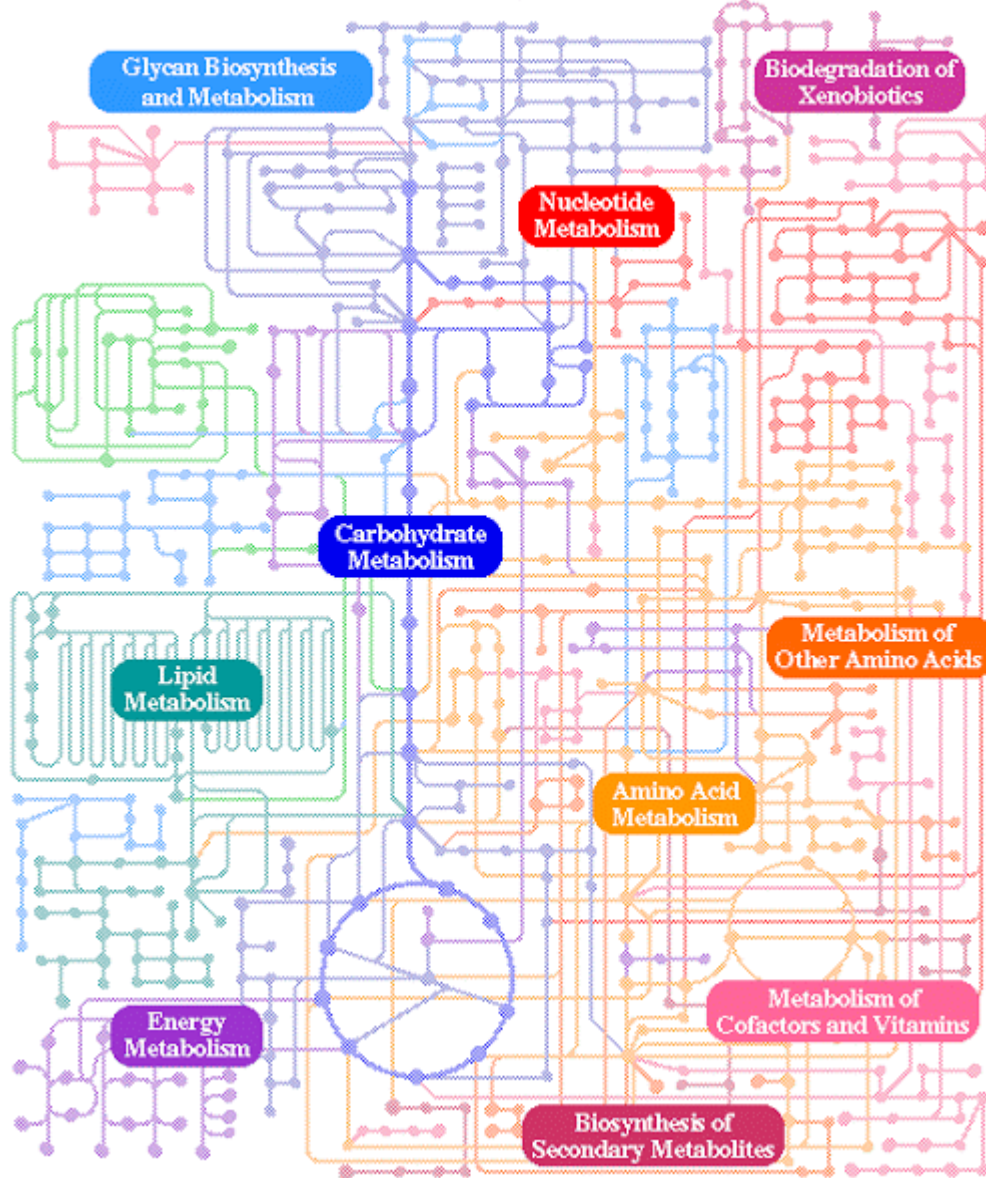
 INTRODUCCIÓN A LA INVESTIGACIÓN BIOMÉDICA

 **ENFERMEDADES METABÓLICAS COMPLEJAS**

 APLICACIÓN DE TÉCNICAS ESTADÍSTICAS E

INTELIGENCIA ARTIFICIAL

METABOLIC PATHWAYS



01100 5/31/04

Conjunto de reacciones bioquímicas y procesos fisicoquímicos que ocurren en una célula y en el organismo.

- Regulación energética
- Estructural

Un trastorno metabólico: alteración del equilibrio

Obesidad

Prevalencia de obesidad
en España 15% (35-55
años)



$\text{BMI} \geq 30 \text{ Kg/m}^2$

Población mayor
(> 55 años) 25%
Población pediátrica
(2-17 años) 14%.

El gasto sanitario de las
patologías asociadas
con la obesidad
(diabetes, hipertensión
y enfermedad
cardiovascular) será de
un 80% en los próximos
10 años

Una pérdida de peso
del 5-10% disminuye
el riesgo
cardiovascular un
50%.

Los pacientes obesos tienen una mayor incidencia de diabetes tipo 2, hipertensión, dislipidemia, enfermedad cardiovascular, apnea del sueño y un incremento en la incidencia de ciertos tipos de cáncer. También tienen una mortalidad mayor comparados con individuos de su misma edad y sexo.¹³



Figures



References



Related



Details

Journal of Computational Biology, Ahead of Print | Research Article

Machine Learning-Based Method for Obesity Risk Evaluation Using Single-Nucleotide Polymorphisms Derived from Next-Generation Sequencing

Hsin-Yao Wang, Shih-Cheng Chang, Wan-Ying Lin, Chun-Hsien Chen, Szu-Hsien Chiang, Kai-Yao Huang, Bo-Yu Chu, Jang-Jih Lu , and Tzong-Yi Lee 

Published Online: 8 Sep 2018 | <https://doi.org/10.1089/cmb.2018.0002>

**VOLUME 0, ISSUE 0****Information**

Copyright 2018, Mary Ann Liebert, Inc., publishers

Obstetric Medicine

View all
publication
partners
▼

[Journal Home](#)[Browse Journal](#) ▼[Submit Paper](#) [About](#) ▼[Subscribe](#) 

Article Menu

Close 

Download PDF



Full Article

Visceral fat mass as a novel risk factor for predicting gestational diabetes in obese pregnant women

J Balani , SL Hyer, H Shehata, more...

[Show all authors](#) ▼

First Published March 14, 2018 | Research Article |



<https://doi.org/10.1177/1753495X17754149>

[Article information](#) ▼

0



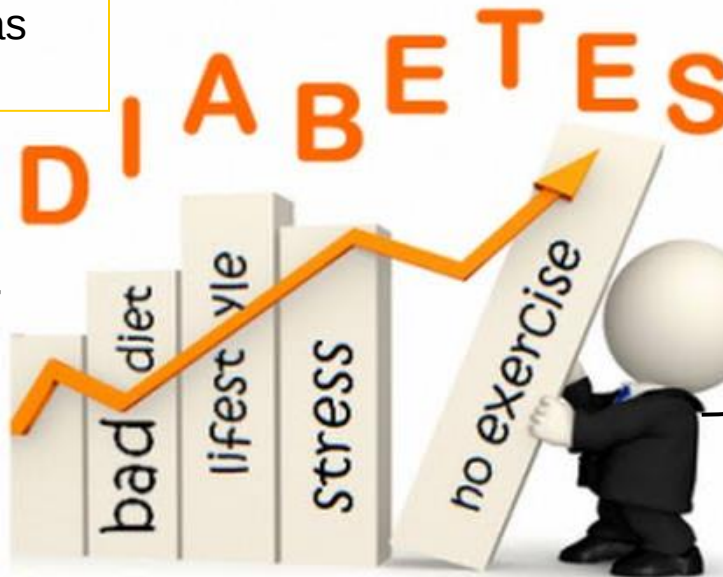
Diabetes

415 millones de adultos tienen diabetes

Complicaciones asociadas con la diabetes: enfermedad renal, retinopatía, pie diabético.

50% to 80% de las muertes son debidas a causas cardiovasculares

Se prevee que la diabetes será la séptima causa de muerte a nivel mundial en 2030



Diabetes mellitus, o simplemente diabetes, es una enfermedad crónica que aparece cuando el páncreas no es capaz de producir insulina o cuando el organismo no puede utilizar la insulina que produce.

Consorcio Español en Páncreas Artificial y Tecnologías para la Diabetes



Grupos » Grupo de Investigación

Grupo de Investigación

BONDIA COMPANY, JORGE - CB17/08/00004



- Modelado y simulación ante incertidumbre y variabilidad intra-paciente.
- Procesamiento de señales, optimización y control. (páncreas artificial y monitorización continua de glucosa)
- Inteligencia Artificial (IA): implementación de técnicas de IA, algoritmos de aprendizaje automático y big data para diseñar sistemas que automáticamente sean capaces de detectar, diagnosticar o predecir efectos de relevancia clínica en diabetes.

<https://tecnodiabetes.ai2.upv.es/>

Síndrome del Ovario Poliquístico (SOP)



Es el síndrome endocrinológico más prevalente en mujeres en edad fértil (6.5% en España). Se caracteriza por un exceso de andrógenos y está asociado a enfermedades metabólicas como la obesidad, la resistencia a la insulina y la diabetes.

Síndrome del ovario poliquístico (SOP)

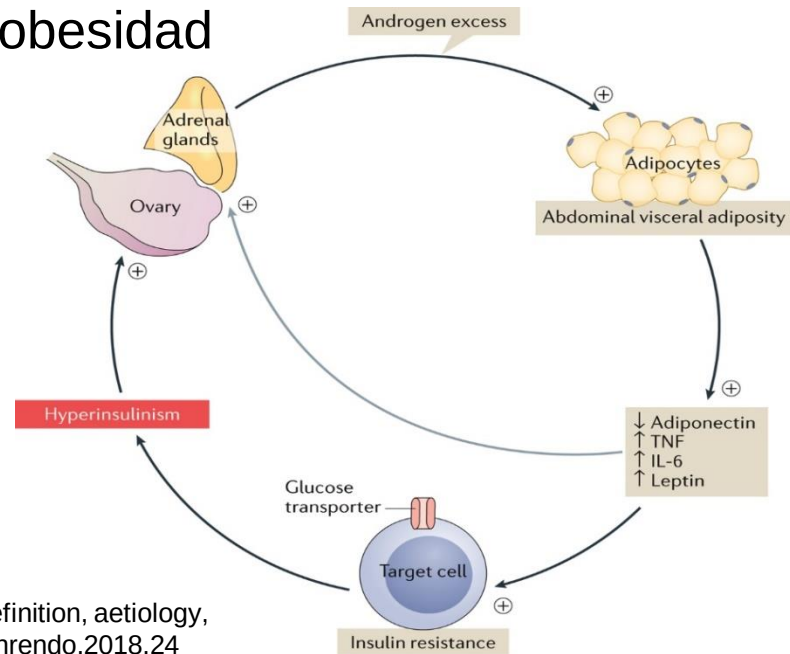
Desorden endocrino metabólico

- hiperandrogenismo clínico y/o bioquímico
- disfunción ovulatoria
- morfología poliquística de los ovarios

Asociado a la resistencia a la insulina y a la obesidad

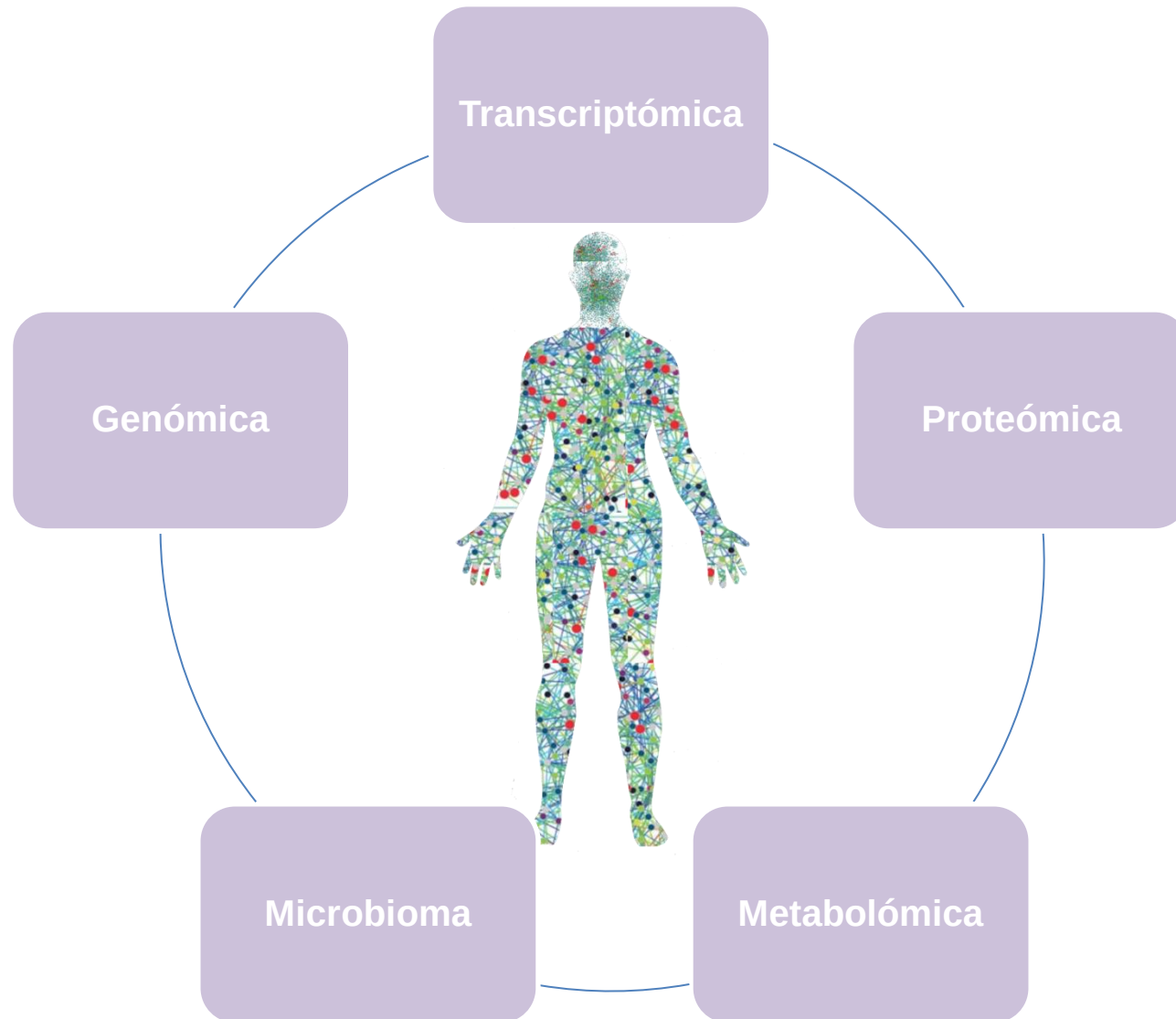
Etiología multifactorial y compleja

- factores genéticos
- factores ambientales



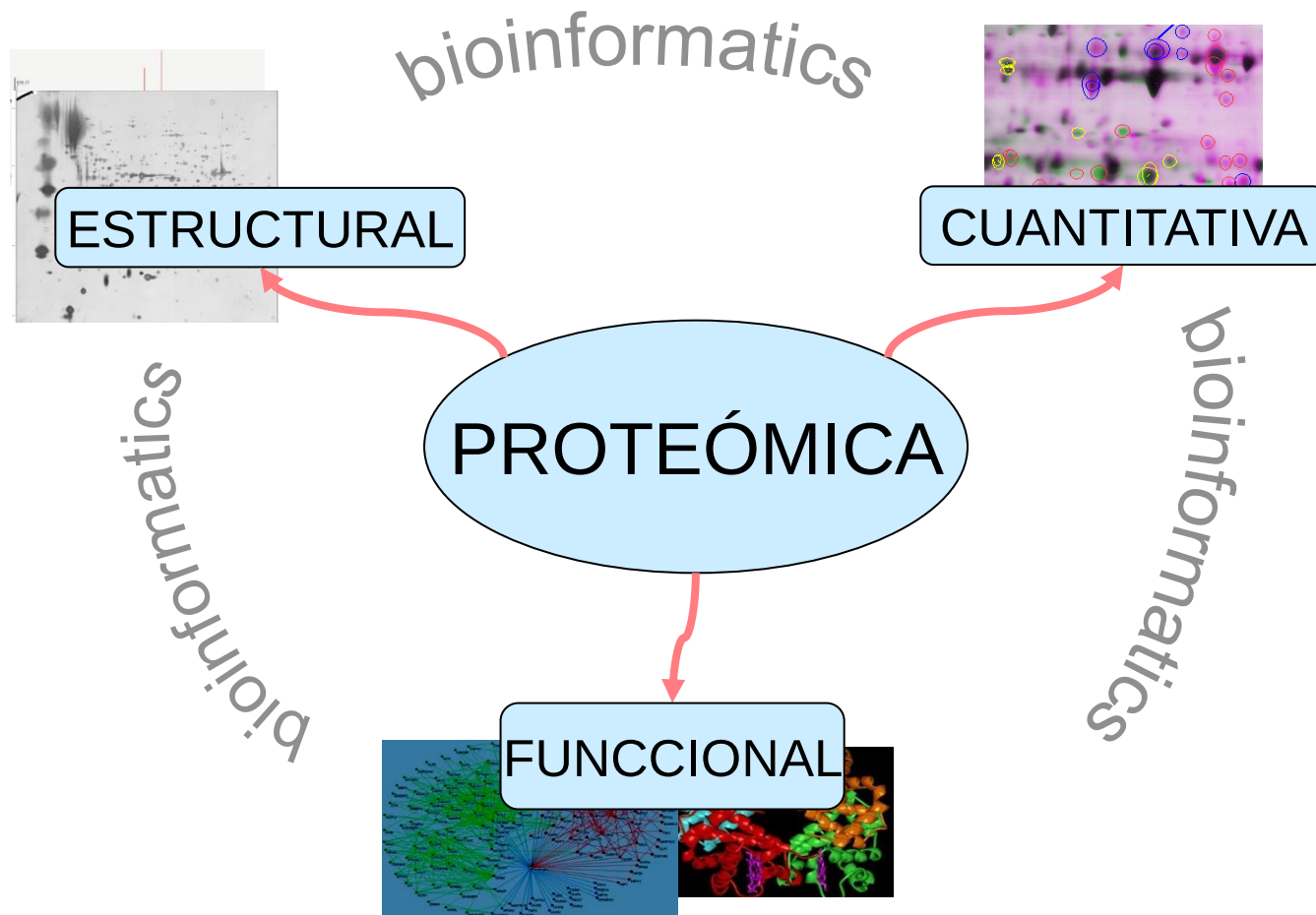
Escobar-Morreale, H. F. (2018) Polycystic ovary syndrome: definition, aetiology, diagnosis and treatment *Nat. Rev. Endocrinol.* doi:10.1038/nrendo.2018.24

Aplicación tecnologías ÓMICAS en SOP



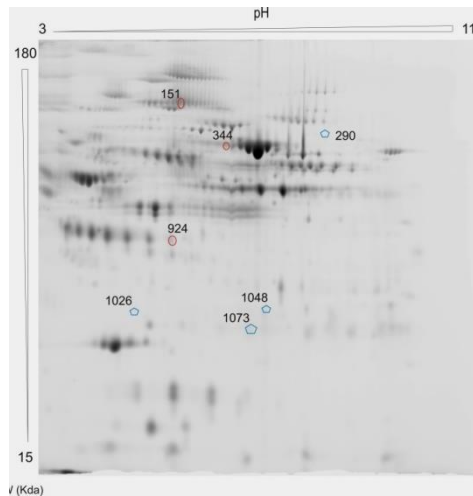
Estudio del proteoma en SOP

- PROTEOMA: conjunto de proteínas codificadas por un genoma de un organismo en un determinado momento.
- PROTEÓMICA: Estudio del proteoma y sus funciones.



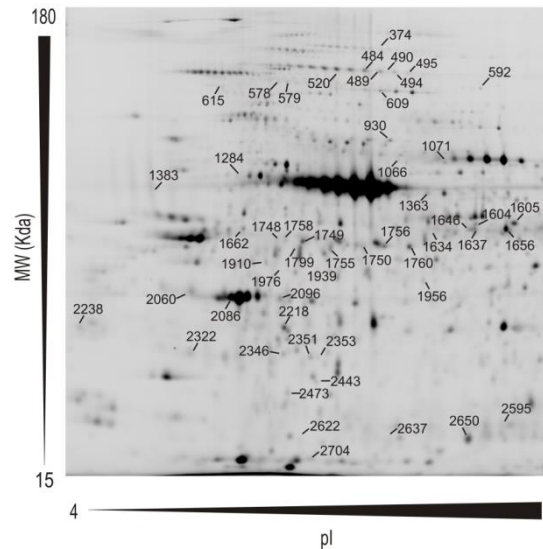
Estudio del proteoma en SOP

Proteínas plasmáticas



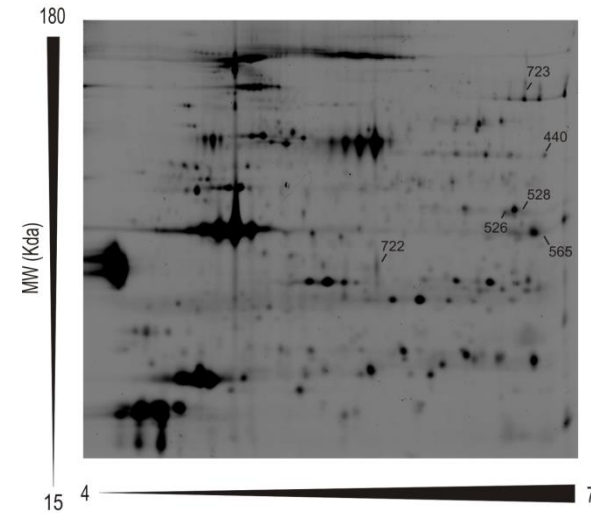
Spots per gel: 1800
 Spots present on all gels: 900
 Differential Spots: 7
 Identified Spots: 6
 Proteins: 4

Proteínas tejido adiposo



Spots per gel: 2600
 Spots present on all gels: 1500
 Differential Spots: 63
 Identified Spots: 53
 Proteins: 34

Proteínas músculo

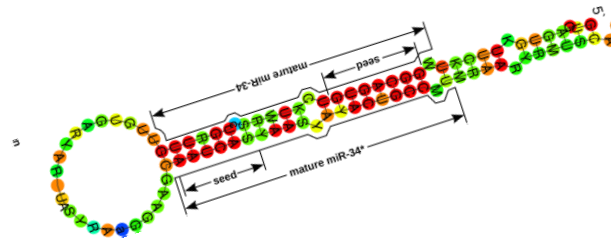
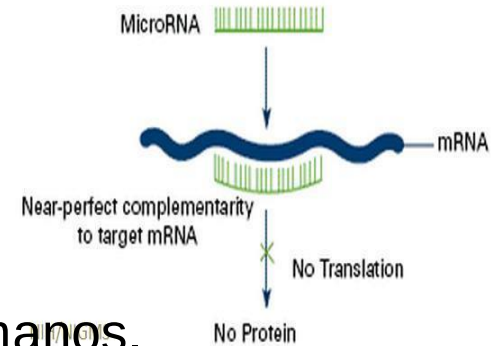


Spots per gel: 920
 Spots present on all gels: 750
 Differential Spots: 7
 Identified Spots: 6
 Proteins: 6

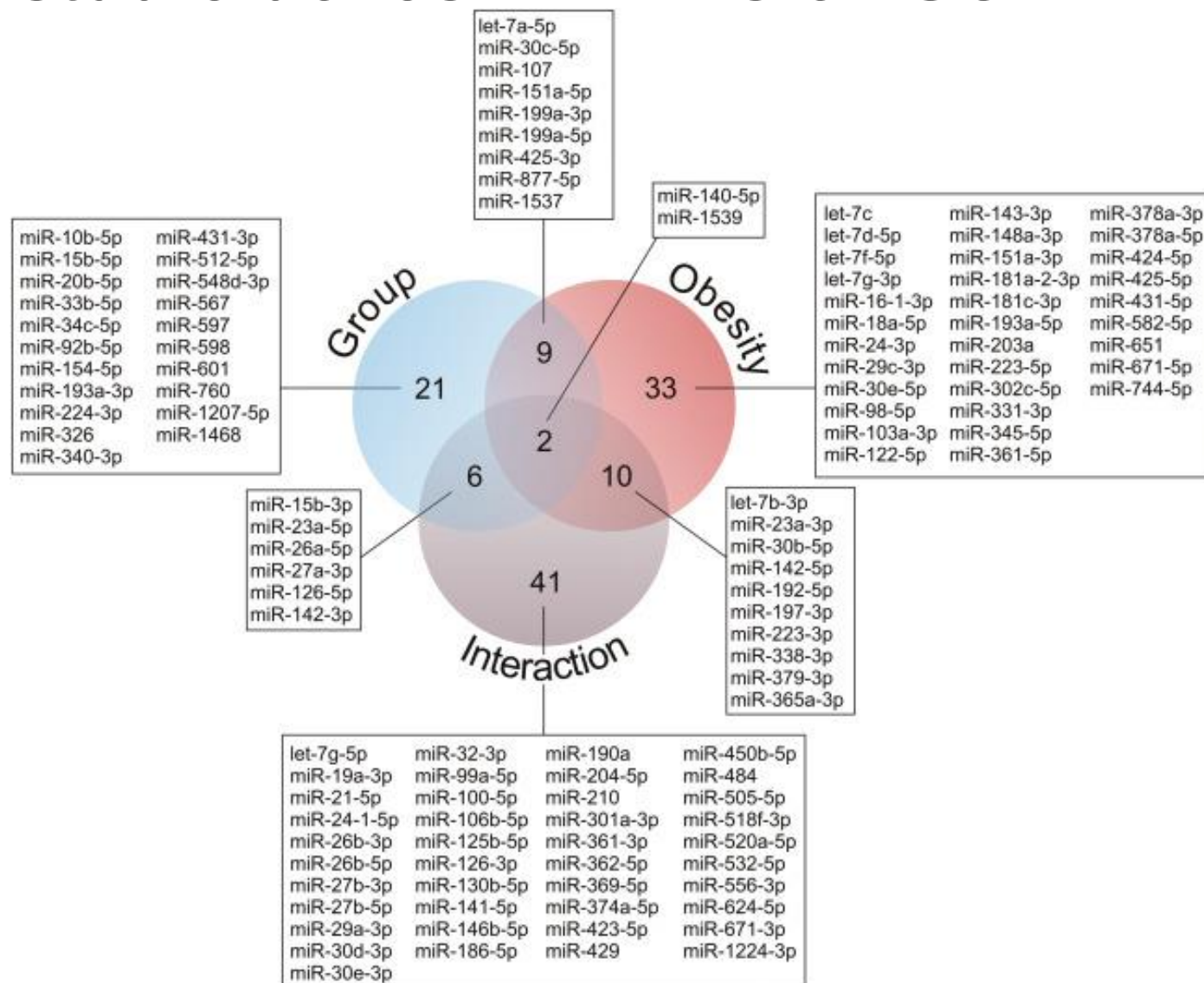
Insenser M, et al. A nontargeted study of muscle proteome in severely obese women with androgen excess compared with severely obese men and nonhyperandrogenic women. *Eur J Endocrinol*. 2016 Mar;174(3):389-98.
 Montes-Nieto R et al. A nontargeted proteomic study of the influence of androgen excess on human visceral and subcutaneous adipose tissue proteomes. *J Clin Endocrinol Metab*. 2013 Mar;98(3):E576-85.
 Insenser M, et al. Proteomic analysis of plasma in the polycystic ovary syndrome identifies novel markers involved in iron metabolism, acute-phase response, and inflammation. *J Clin Endocrinol Metab*. 2010 Aug;95(8):3863-70.

Estudio de los miRNAs en SOP

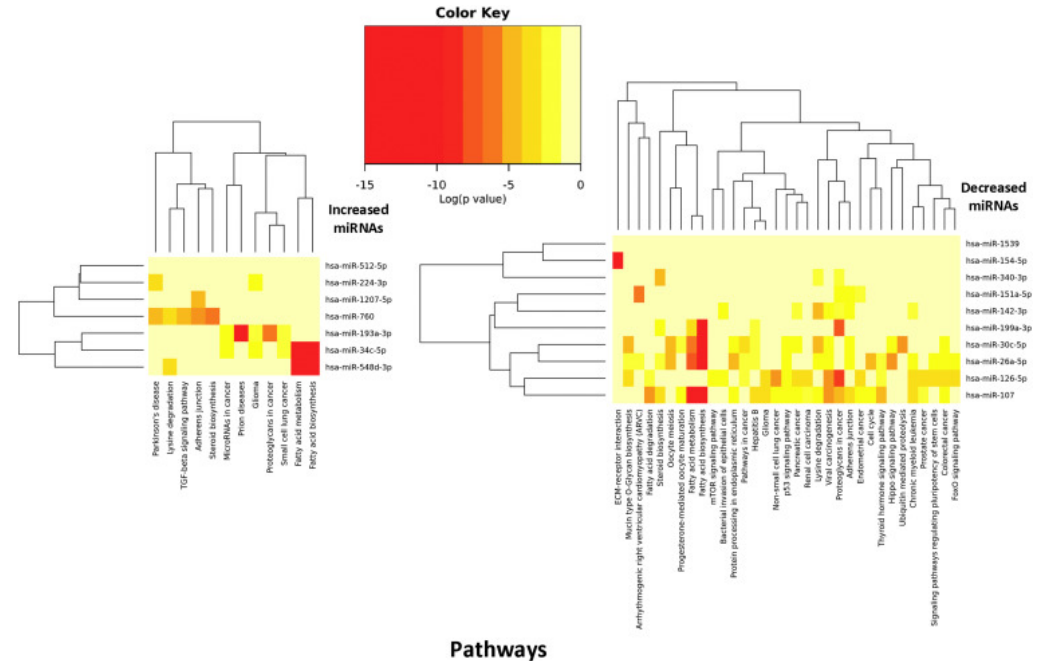
- ▶ Moléculas pequeñas de RNA no codificante
- ▶ Reguladores post-transcripcionales del mRNA
- ▶ Regulan al menos la tercera parte de los genes humanos.
- ▶ Existen 2603 microRNAs humanos descritos (<http://www.mirbase.org>)
- ▶ La alteración de los microRNAs tiene un papel relevante en la patogénesis del cáncer, enfermedades infecciosas, diabetes, enfermedades cardiovasculares e inflamatorias...



Estudio de los miRNAs en SOP



Heatmaps of pathways predicted for miRNAs over- or underexpressed in patients with PCOS compared with control women regardless of obesity.



Heatmaps of pathways predicted for miRNAs over- or underexpressed in patients with PCOS compared with control women regardless of obesity.

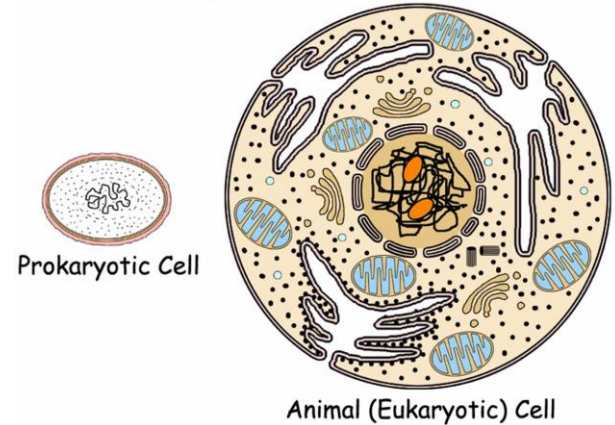
24

Estudio del microbiota en SOP

Composición celular del cuerpo humano



Microbioma

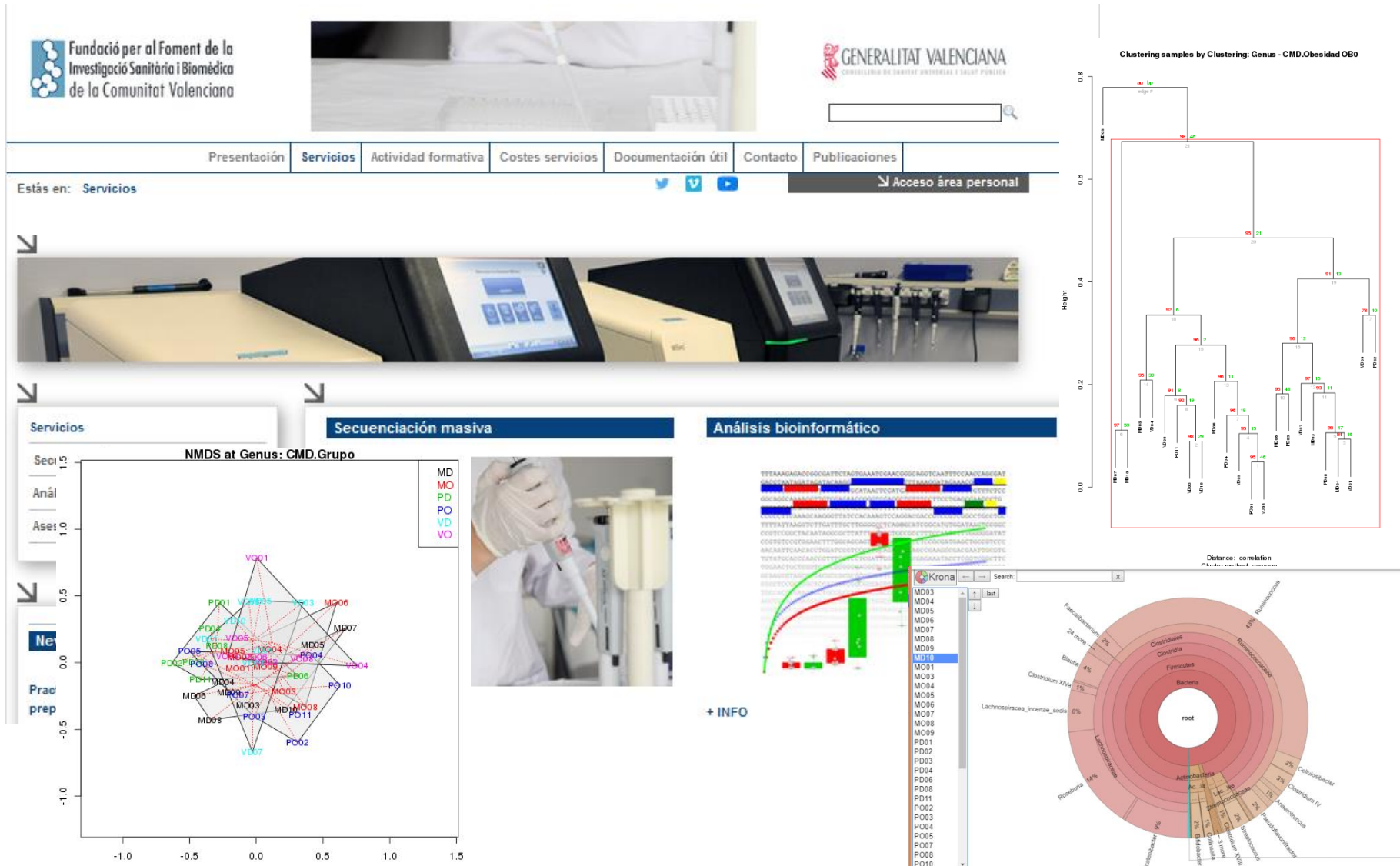


Conjunto de microorganismos que residen en el cuerpo humano.

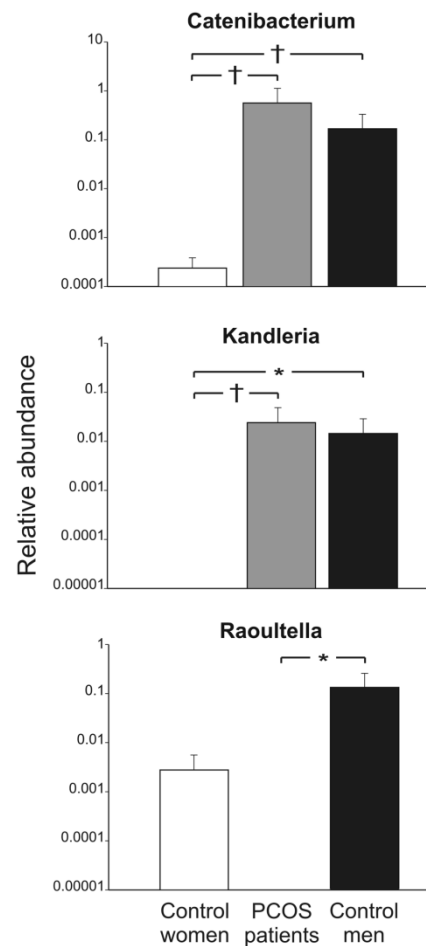
Es importante para el funcionamiento del organismo y tiene un impacto sobre la salud

"We are never really alone"

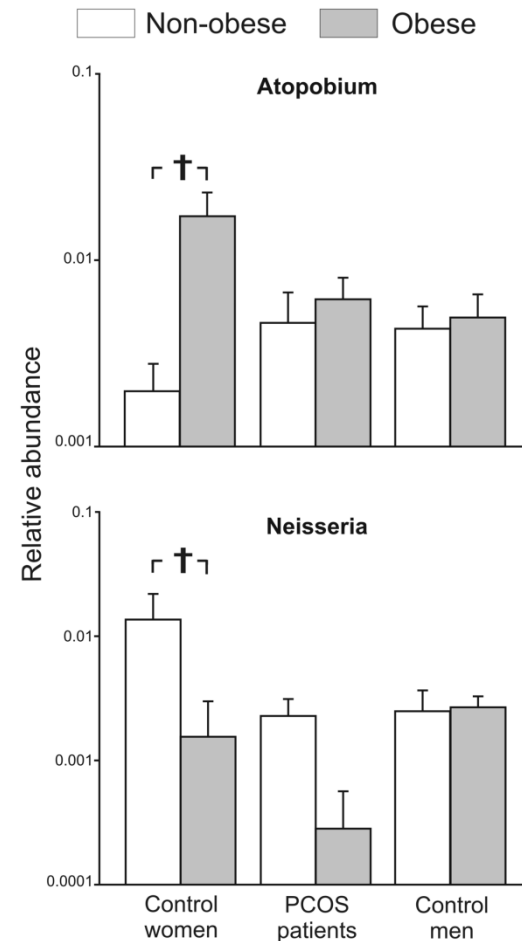
Estudio del microbiota en SOP



Estudio del microbiota en SOP



Differences in bacterial relative abundance at the genus level among groups.



Interactions between study group and obesity in bacterial relative abundance at the genus level.

 INTRODUCCIÓN A LA INVESTIGACIÓN BIOMÉDICA

 ENFERMEDADES METABÓLICAS COMPLEJAS

 **APLICACIÓN DE TÉCNICAS ESTADÍSTICAS E
INTELIGENCIA ARTIFICIAL**

Proyecto: respuesta hormonal, metabólica, inflamatoria y oxidativa a los diferentes macronutrientes de la dieta: influencia de los esteroides sexuales.

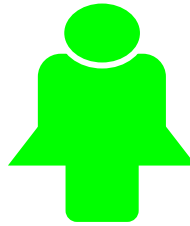
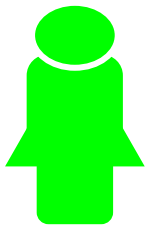
Hipótesis: Las mujeres con SOP pueden presentar una respuesta a los macronutrientes de la dieta diferente a la de las mujeres sin exceso androgénico.

Objetivo: contrastar la hipótesis nula de que no existen diferencias entre mujeres con SOP, mujeres sin exceso androgénico y hombres en respuesta a los diferentes macronutrientes de la dieta.

Diseño experimental

Mujeres

↓↓ andrógenos
↑ estrógenos



BMI < 30 N = 9 BMI ≥ 30 N = 8

SOP

↑ andrógenos
↑ estrógenos



BMI < 30 N = 9 BMI ≥ 30 N = 8

Hombres

↑↑ andrógenos
↓↓ estrógenos



BMI < 30 N = 10 BMI ≥ 30 N = 9

Sobrecarga Glucosa

Sobrecarga lípidos

Sobrecarga proteínas

N = 53 pacientes ³⁰

Diseño experimental

► GLUCOSA



Sobrecarga glucosa

0' 60' 120'

► LÍPIDOS



Sobrecarga lípidos

0' 120' 240'

► PROTEÍNA



Sobrecarga proteínas

0' 60' 120'

Hormonas
Metabolitos
Estrés oxidativo
Adipokinas
Enterokinas

477 Muestras de suero

Variables estudiadas

GROUP
OBESE
EDAD
PESO
TALLA
BMI (Indice masa corporal)
Grasa
TAS
TAD
CINTURA
CADERA
Menstruación
Proteína C reactiva
FERRITINA
Monocitos
Eosinófilos
Basófilos
Leucocitos
Neutrofilos
Linfocitos
Glucosa
Insulina
HOMA IR
ISI
Hemoglobina
Hemoglobina glicada
Fibrinógeno
GOT
GPT
GGT
FA

FSH
TSH
Prolactina
Testosterona total
Testosterona libre
SHBG

Estradiol
17-hidroxiprogesterona
Androstendiona
FERRIMAN (hirsutismo)

CD4inicial
CD4final
CD8inicial
CD8final
CH50 inicial
C3 inicial
C3 final
C4 inicial
C4 final

Glucosa
Insulina
Colesterol
Triglicéridos
HDL
LDL

Adiponectina
Lipocalin
Adipsin
PAI
GIP
GLP
PYY
TNFa
IL1b
IL6
IL10
IL10
Omentin
sLeptin
Vaspin
TBARs
Pentraxin
Galectin
FGF23
FGF21
PON1

Análisis de las muestras

Luminex®

America's 25 Fastest-Growing Tech Companies (FORBES)



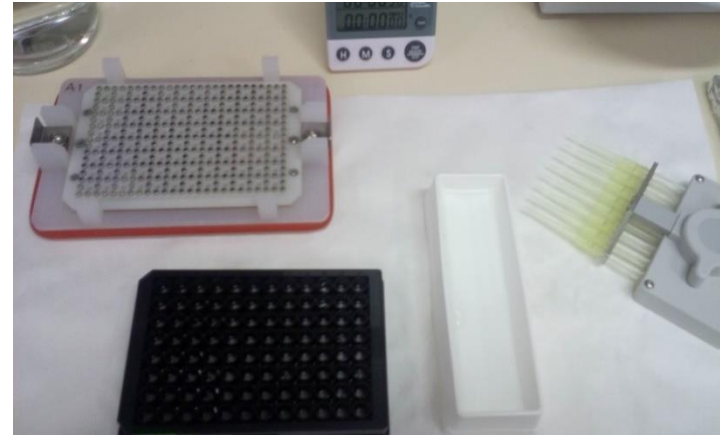
La tecnología Luminex's xMAP® (Multi-Analyte Profiling) es una tecnología multiplex que permite la determinación simultánea de hasta 500 analitos en un volumen muy pequeño, utilizando bolitas microscópicas de poliestireno (**microesferas**).



Análisis de las muestras



Preparación de las muestras



Protocolo y lavados



Medida en el Magpix

Análisis de los resultados

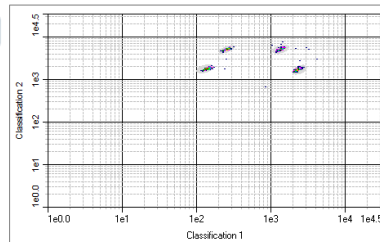
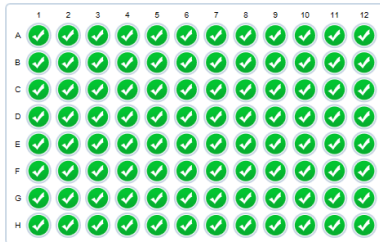
"Placa 8 - Adipoquinas", created 4/14/2016 9:50:48 AM.

Instructions View the results for this batch. Select the Information icon for more information on results options.

Results

Statistic: Median Analyte: Select an Analyte Displayed Well: 1,A1

Well	Sample	Run Sta...	Adiponectin	Lipocalin-2/...	Adipsin	PAI-1 (total)
1,A1	Background0	Ok	11	35	25	28
1,B1	Standard1	Ok	12	75	29	57
1,C1	Standard2	Ok	20	89	51	186
1,D1	Standard3	Ok	58	168	224.5	884.5
1,E1	Standard4	Ok	267	771	1455	3398



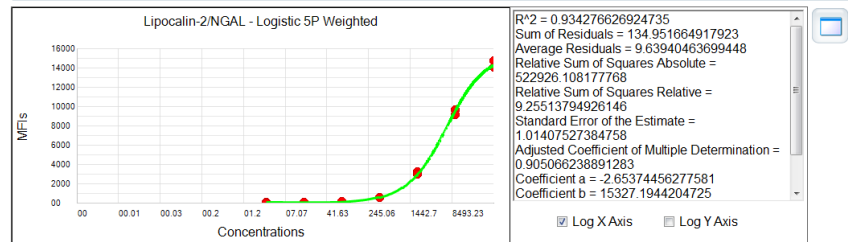
"Placa 12 Adipokinas", created 4/20/2016 11:40:32 AM.

Instructions View the results for this batch. Select the Information icon for more information on results options.

Results

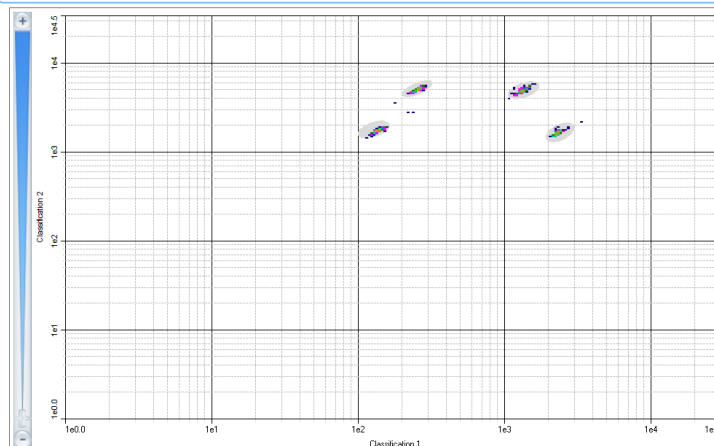
Statistic: Select a Statistic Analyte: Lipocalin-2/NGAL(56) Displayed Well: 1,A1

Well	Sample	Run Sta...	Median	Test Result	Range	Comments	Count	Net MFI (M...	Mean
1,A1	Background0	Ok	21				363	0	21.37
1,B1	Standard1	Ok	23	4.709			348	2	23.12
1,C1	Standard2	Ok	36	16.031			331	15	36.97
1,D1	Standard3	Ok	109	72.178			388	88	111.13
1,E1	Standard4	Ok	604	406.459			320	583	588.98



"Placa 12 Adipokinas", created 4/20/2016 11:40:32 AM.

Instructions View the results for this batch. Select the Information icon for more information on results options.

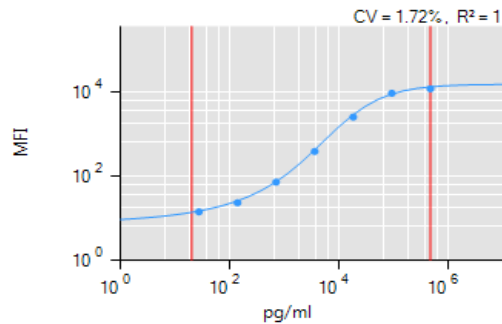


Análisis de los resultados

5-P.log(2.02, 9.53, 0.41, 7.94, 0.89)

DC=(20.21, 471009) Chi = 0.055%

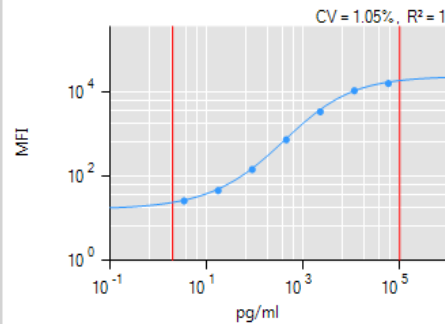
Adiponectin



5-P.log(2.76, 9.96, 0.56, 5.81, 0.62)

DC=(1.98, 100991) Chi = 0.036%

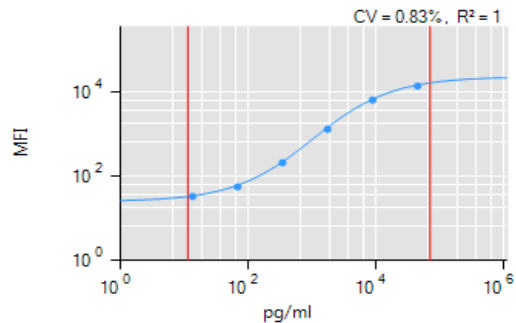
Lipocalin-2/NGAL



5-P.log(3.18, 9.94, 0.71, 6.87, 0.71)

DC=(11.39, 70419) Chi = 0.022%

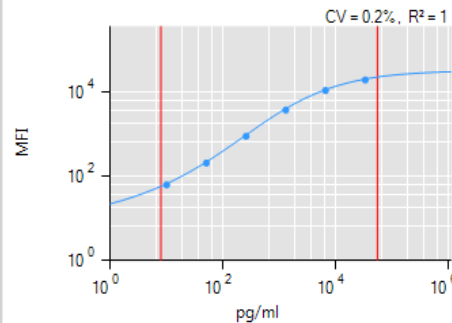
Adipsin



5-P.log(1.67, 10.28, 0.27, 4.63, 0.57)

DC=(7.96, 55237) Chi = 0.0060%

PAI-1 (total)



Resultados 12 placas

53 pacientes

Tres sobrecargas

Tres tiempos por sobrecarga

AUC de cada sobrecarga

Notes:

CV-The Coefficient of Variation of standard curve replicates at each dilution level.

Chi-The Chi-Square test statistic of the distance between observed concentrations with expected concentrations.

UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA INFORMÁTICA



Máster Universitario en
Ingeniería de Sistemas y de Control

BIOINFORMATICS: USING STATISTICAL
METHODS AND ARTIFICIAL INTELLIGENCE
FOR COMPLEX METABOLIC DISORDERS
UNDERSTANDING

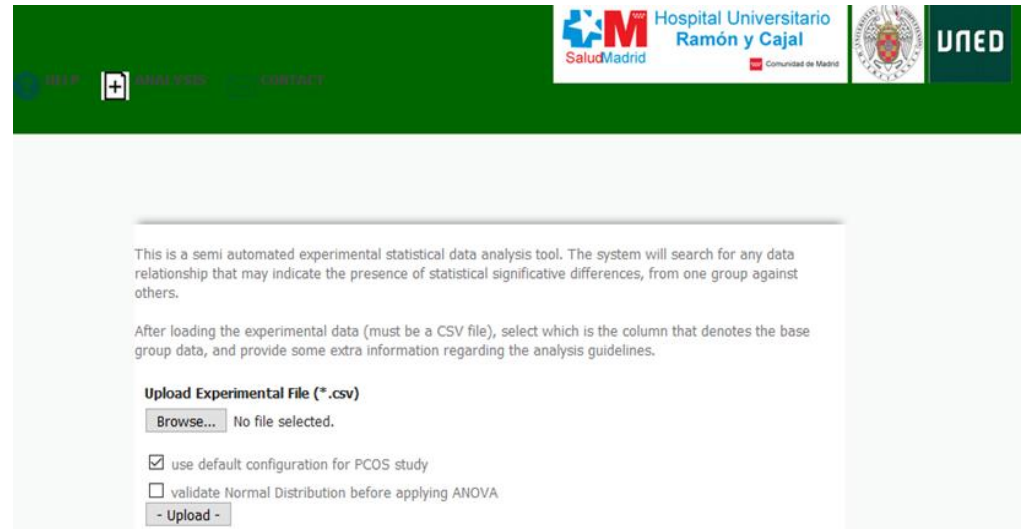
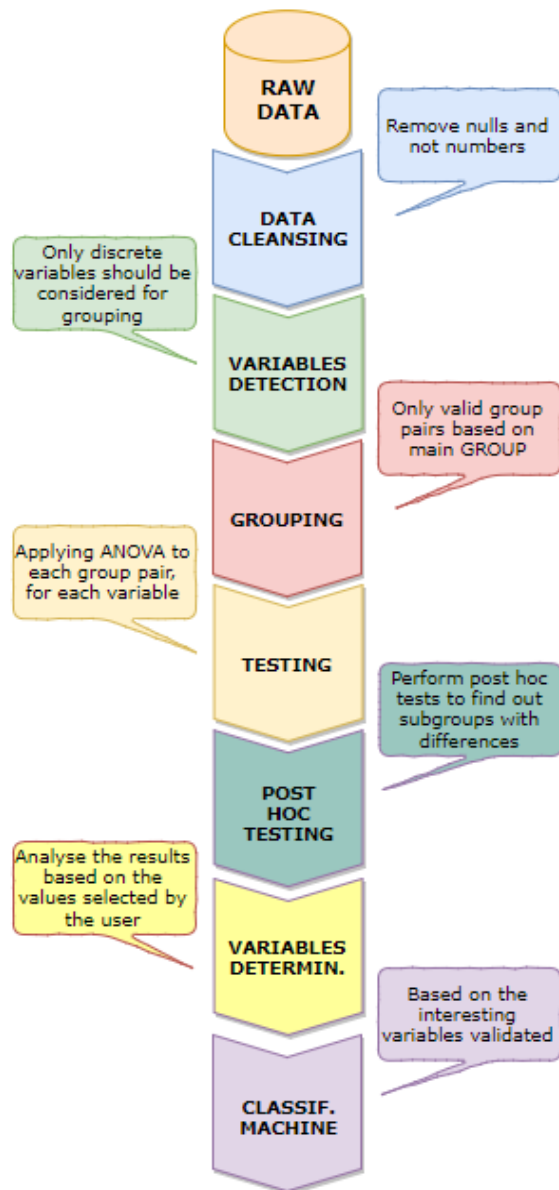
Author: Víctor Javier Cerquera Parra

Director(s): María Guijarro Mata-García

Juan Jiménez Castellanos

María Insenser Nieto

Diseño herramienta de análisis



<http://vjctfm.ddns.net:5000/>

Diseño herramienta de análisis

Significative groups configuration

GROUP	
Value Combination	Relevant
WOMEN	☐
PCOS	☒
MEN	☐

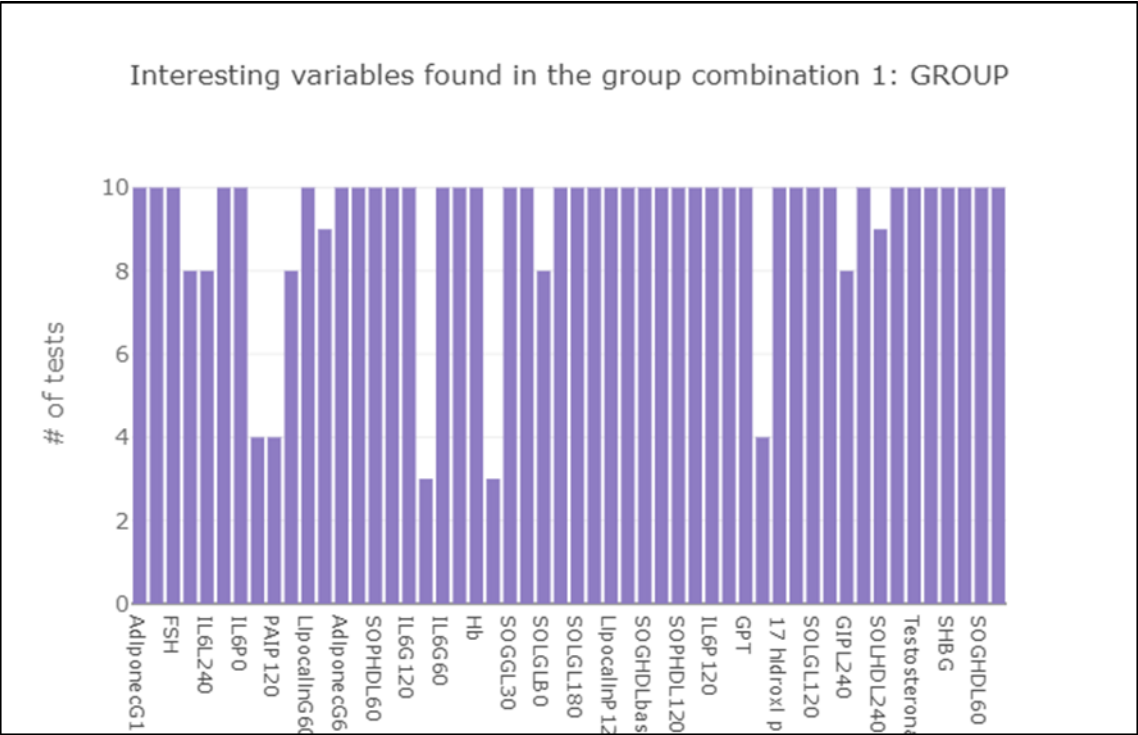
Significative groups configuration

GROUP	
GROUP_OBESE	
Value Combination	Relevant
WOMEN - NONOBESE	☐
WOMEN - OBESE	☐
PCOS - NONOBESE	☒
PCOS - OBESE	☒
MEN - NONOBESE	☐
MEN - OBESE	☐

Validation Post Hoc tests to include

Post Hoc Test	Include in Analysis
Tukey	☒
Bonferroni	☒
Holm Bonferroni	☒
Holm Sidak	☒
Simes Hochberg	☒
Hommel	☒
FDR Benjamini / Hochberg	☒
FDR Benjamini / Yekutieli	☒
Benjamini / Hochberg two step FDR correction	☒
Benjamini / Krieger / Yekutieli two step FDR correction	☒
FDR adaptive Gavrilov-Benjamini-Sarkar	☒

Resultados



Detected interesting variables combination GROUP

List of variables and case which they satisfy on deep analysis, combination GROUP

Variables	Case
AdiponecG120	6
AdiponecG60	6
Fibrinógeno	6
IL6P0	6
IL6P120	6
LipocalinG60	6
normAUCsinGIPL	6
SOGGL30	6
SOGGLB0	6
SOLAUCglutotal	6
SOLGL120	6
SOLGL60	6
SOLGLB0	6
SOPAUCHDLtotal	6
SOPGL120	6
SOPGL30	6
SOPGL60	6
SOPGL90	6
AdipsinG0	4
AUCsinGIPL	2
GIPL240	2
Androstendiona0	3
GIPP120	3
TBARsG0	3
IL6P60	1
Testosterona libre	1

Resultados

SVM Machine Training

Training will only be done on variables detected on the base group analysis.

Minimum number of tests passed

Variable must pass at least: 5 Tests

- Train SVM -

Best parameters set found on development set:

```
{'kernel': 'poly', 'C': 1, 'degree': 2}
```

Detailed classification report:

The model is trained on the full development set

	precision	recall	f1-score	support
MEN	1.00	1.00	1.00	12
PCOS	1.00	1.00	1.00	12
WOMEN	1.00	1.00	1.00	12
avg / total	1.00	1.00	1.00	36

The scores are computed on the full evaluation set

	precision	recall	f1-score	support
MEN	1.00	1.00	1.00	2
PCOS	0.50	0.50	0.50	2
WOMEN	0.00	0.00	0.00	1
avg / total	0.60	0.60	0.60	5

Final results:

Conclusiones

- ▶ La bioinformática es una herramienta metodológica de gran impacto en la investigación biomédica.
- ▶ Gracias a la aplicación de tecnologías ómicas en enfermedades metabólicas se están generando una gran cantidad de datos de datos útiles no solo para conocer la fisiopatología de estas enfermedades sino también para mejorar el diagnóstico y tratamiento de las mismas.
- ▶ La aplicación de la inteligencia artificial en la investigación biomédica es fundamental para el desarrollo de estrategias y herramientas que permitan validar resultados, generar nuevas hipótesis y orientar la investigación hacia una medicina personalizada.

