



MULTIOMICS



Ciencia de datos biomédicos

1

INTRO AL DOMINIO

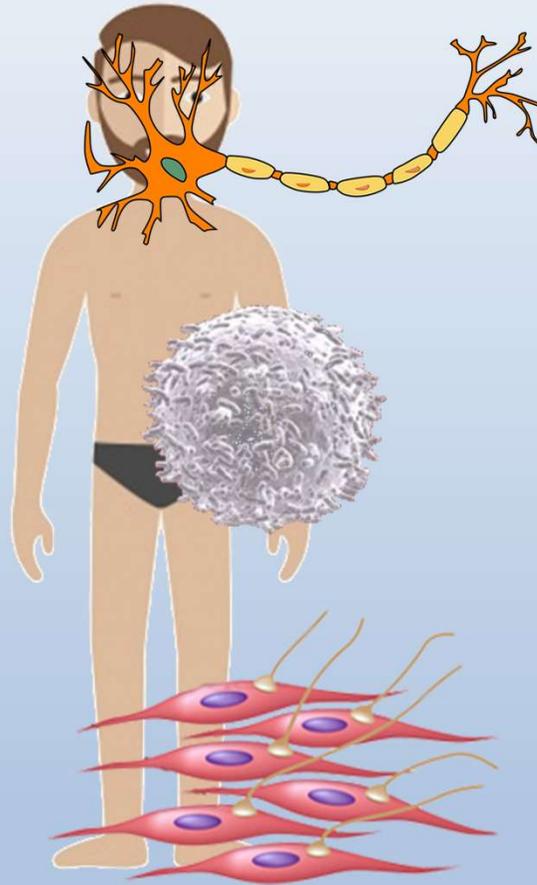
2

BIOPLAT: BIOMARCADORES
EN CÁNCER

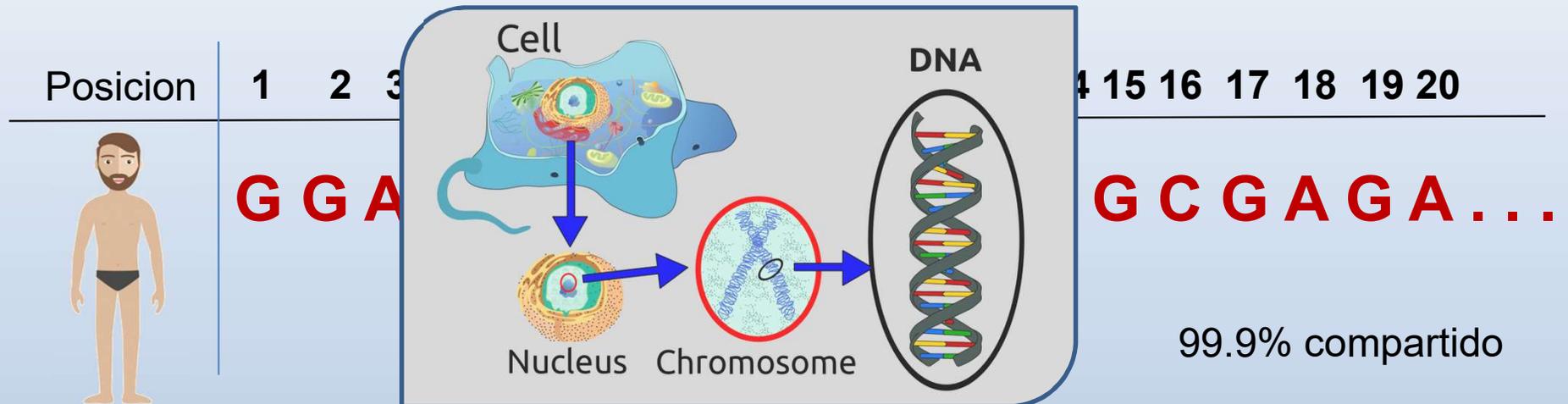
3

genomIT - Zoigen

Nivel celular



ADN



Membrana
celular

Citoplasma

Núcleo

Del ADN a las Proteínas

GGACCTGGAATATGGCGAGAAAAAACTCAAACCTCACGAAAATGAAGGCTTAGA. . . .

GENE 1

GENE 2

GENE 3

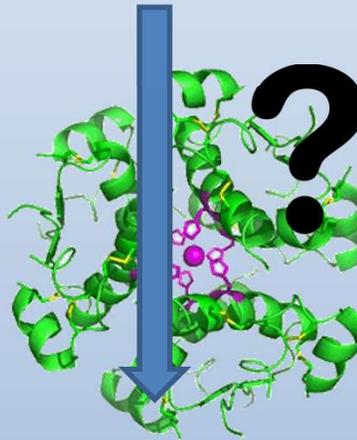


DNA

Proteínas ?

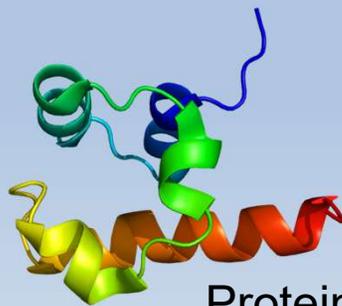
- estructurales
- contractiles
- enzimáticas
- inmunológicas
- transducción de señales

...



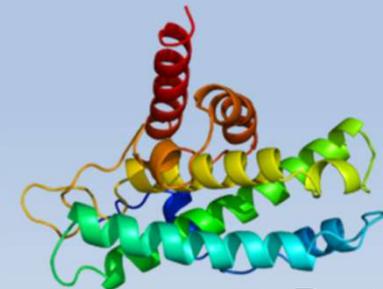
Proteína. Ej: Insulina

?



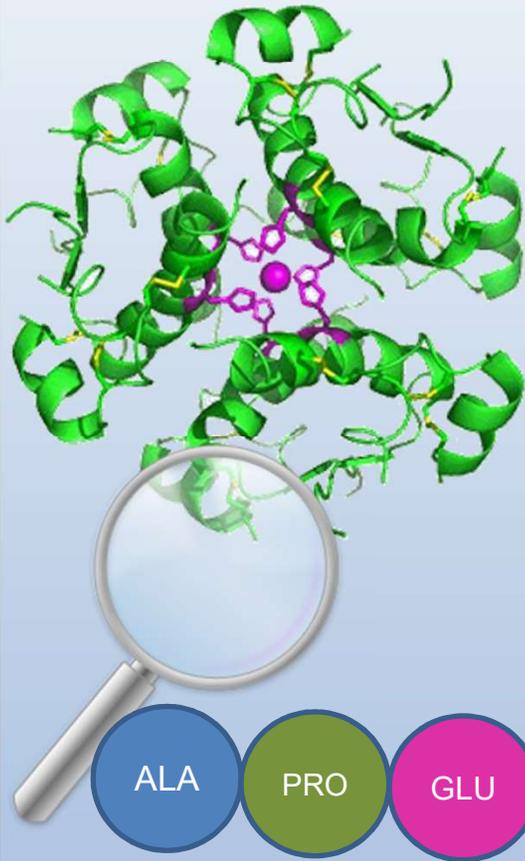
Proteina 1

Proteina 2



Proteina 3

Estructura de las Proteínas



AMINOÁCIDOS

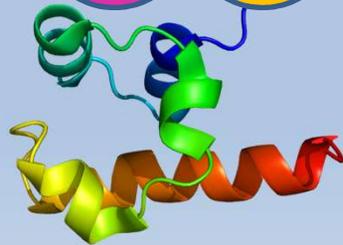
| Nombre | 3 letras |
|-------------------------|------------|
| Alanina | Ala |
| Arginina | Arg |
| Asparagina | Asn |
| Ácido Aspártico | Asp |
| Cisteína | Cys |
| Ácido Glutámico | Glu |
| Glutamina | Gln |
| Glicina | Gly |
| Histidina | His |
| Homoserina | Hse |
| Isoleucina | Ile |
| Leucina | Leu |
| Lisina | Lys |
| Metionina | Met |
| Metionina sulfóxido | Met (O) |
| Metionina metilsulfonio | Met (S-Me) |
| Norleucina | Nle |
| Fenilalanina | Phe |
| Prolina | Pro |
| Serina | Ser |
| Treonina | Thr |
| Triptófano | Trp |
| Tirosina | Tyr |
| Valina | Val |

Del ADN a las Proteínas

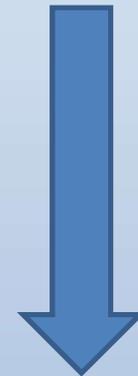
GENE 1

G C A C C A G A A A A G A T G G ...

?



ADN
(4 elementos)

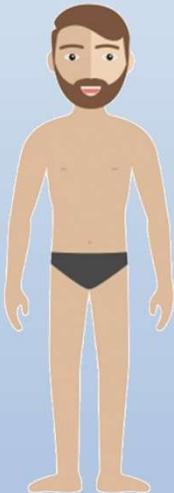


Proteínas
(20 elementos)

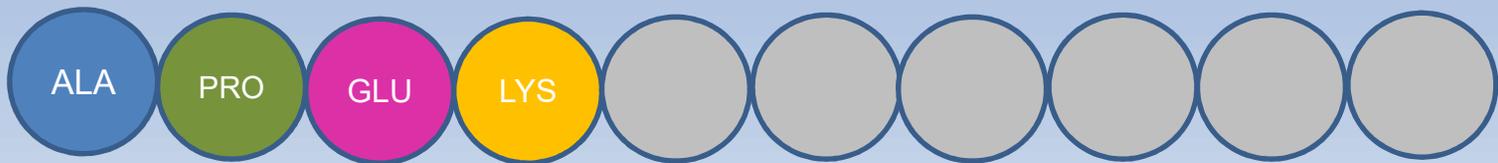
Del ADN a las proteínas

CAA

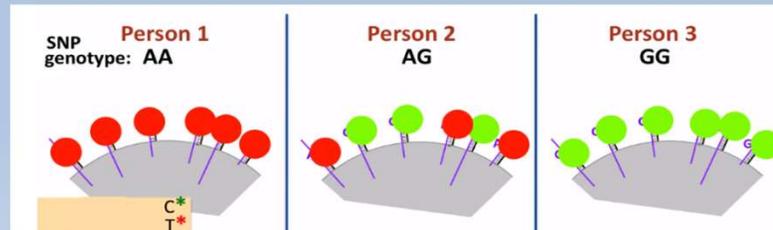
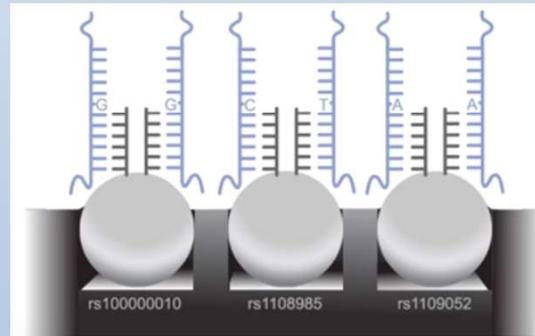
G C A C C A G A A
A A G A T G G ...



| | | Second letter | | | | Third letter |
|---|-----------|---------------|-----------|-----------|------------------|--------------|
| | | U (T) | C | A | G | |
| U | UUU } Phe | UCU } Ser | UAU } Tyr | UGU } Cys | U C A G | |
| | UUC } Leu | UCC } Ser | UAC } Tyr | UGC } Cys | | |
| | UUA } Leu | UCA } Ser | UAA Stop | UGA Stop | | |
| | UUG } Leu | UCG } Ser | UAG Stop | UGG Trp | | |
| C | CUU } Leu | CCU } Pro | CAU } His | CGU } Arg | U C A G | |
| | CUC } Leu | CCC } Pro | CAC } His | CGC } Arg | | |
| | CUA } Leu | CCA } Pro | CAA } Gln | CGA } Arg | | |
| | CUG } Leu | CCG } Pro | CAG } Gln | CGG } Arg | | |
| A | AUU } Ile | ACU } Thr | AAU } Lys | AGU } Ser | U C A G | |
| | AUC } Ile | ACC } Thr | AAC } Lys | AGC } Ser | | |
| | AUA } Met | ACA } Thr | AAA } Lys | AGA } Arg | | |
| | AUG } Met | ACG } Thr | AAG } Lys | AGG } Arg | | |
| G | GUU } Val | GCU } Ala | GAU } Asp | GGU } Gly | U C A G | |
| | GUC } Val | GCC } Ala | GAC } Asp | GGC } Gly | | |
| | GUA } Val | GCA } Ala | GAA } Glu | GGA } Gly | | |
| | GUG } Val | GCG } Ala | GAG } Glu | GGG } Gly | | |



1-Secuenciación del ADN



2-Genoma de referencia

Posicion

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Referencia

G G A C C T G G A A T A T G A C G A G A . . .



G G C C C T A G A A T A T G A C G T G A . . .

2-Proyecto genoma humano



NCBI Resources How To

Genome Data Viewer

Homo sapiens: GRCh38.p12 (GCF_000001405.38) Chr 17 (NC_000017.11)

Ideogram View

Unplaced/unlocalized scaffolds: 168
Alt loci/patches: 401

Region: TP53 (Gene) NM_000546.5 (Transcript)

NC_000017.11

7,677,960 7,677,970

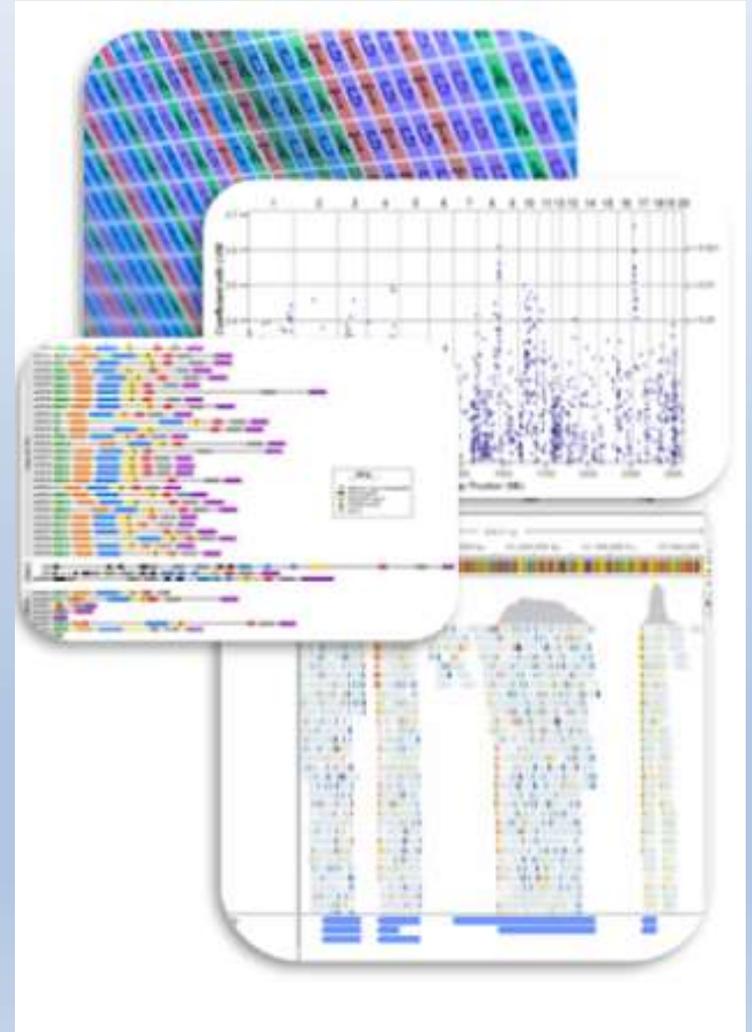
ATG

Genes, NCBI Homo sapiens Annotation Release 109, 2018-03-27

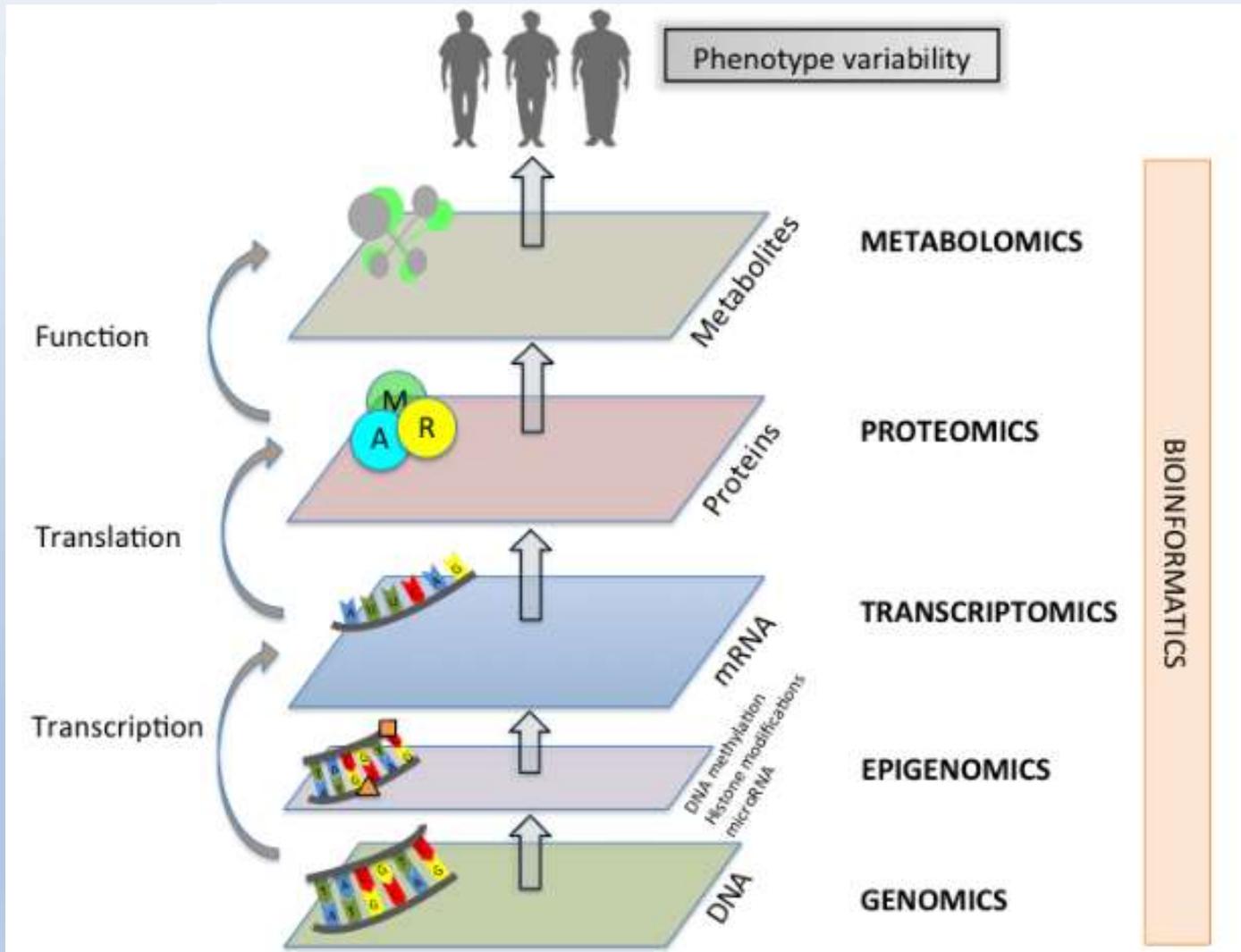
The screenshot displays the NCBI Genome Data Viewer interface. It shows a genomic ideogram for chromosome 17, a track for the TP53 gene and transcript (NM_000546.5), and a sequence viewer for the TP53 gene region (NC_000017.11) with coordinates 7,677,960 to 7,677,970. The sequence viewer shows the DNA sequence: ATG... and the corresponding protein sequence: M... The interface includes navigation controls and a search bar.

Poder de cómputo

Algoritmos para el Análisis
de la información molecular



Múltiples tipos de datos

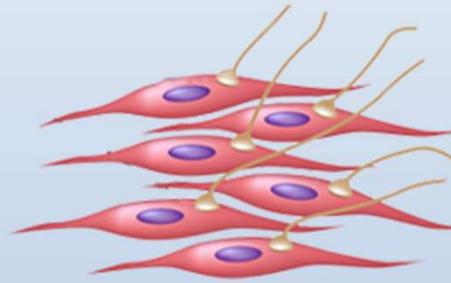
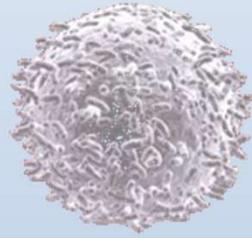
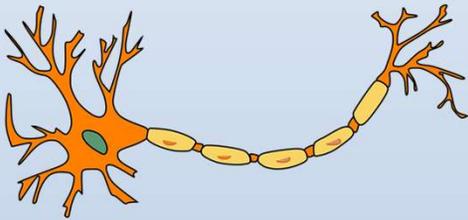


Microbioma



Actividad génica - Transcriptómica

Conocer solo la secuencia el ADN no siempre alcanza → Perfiles de expresión génica



GGACCTGG AATATGGCGAGA...

Actividad génica: Transcriptoma

Samples



G
e
n
e
s

| sampleId | Sample 14E | Sample 15E | Sample 1 | Sample 1 | Sample 1 | Sample E | Sample 1 | Sample E | Sample 1 | Sample E |
|----------|------------|------------|----------|----------|----------|----------|----------|----------|----------|----------|
| 4886 | -1.26 | -1.4 | -1.54 | -1.24 | -1.59 | 0.66 | -1.5 | -0.13 | -1.29 | -0.1 |
| 8821 | -0.17 | -0.1 | -0.14 | -0.04 | -0.14 | 0.12 | -2.0 | 0.12 | -0.31 | 0.27 |
| 57758 | -0.74 | -0.78 | -0.71 | -0.66 | -1.32 | -0.1 | -0.86 | 0.71 | -0.64 | 0.7 |
| 596 | -0.21 | -0.1 | -0.04 | -0.05 | -0.04 | 0.07 | -0.09 | 0.07 | -0.13 | 0.09 |
| 4085 | -0.37 | -0.31 | -0.41 | -0.38 | -0.32 | 0.03 | -0.58 | 0.3 | -0.35 | 0.53 |
| 3488 | -0.27 | -0.33 | -0.36 | -0.34 | -0.3 | -0.09 | -0.57 | -0.12 | -0.4 | 0.36 |
| 2099 | -0.89 | -1.03 | -1.25 | -1.01 | -1.41 | 0.34 | -1.19 | 0.38 | -1.13 | 0.35 |
| 6241 | -0.36 | -0.29 | -0.17 | -0.33 | -0.26 | 0.18 | -0.19 | 0.49 | -0.5 | 0.12 |
| 2064 | -0.84 | -0.83 | -0.84 | -0.84 | -0.88 | 0.63 | -0.85 | 0.45 | -0.92 | 0.89 |
| 3872 | -0.6 | -0.37 | -0.4 | -0.37 | -0.45 | 1.21 | -0.61 | 0.46 | -0.66 | 0.35 |
| 2335 | -0.72 | -0.2 | -0.19 | -0.37 | -0.35 | 0.36 | -0.61 | 0.25 | -0.37 | 0.34 |
| 2296 | -0.59 | -0.77 | -0.61 | -0.61 | -0.56 | 0.67 | -0.71 | -0.04 | -0.71 | -0.31 |
| 22974 | -0.27 | -0.24 | -0.24 | -0.27 | -0.26 | 0.22 | -0.18 | 0.27 | -0.38 | 0.29 |
| 3667 | -0.44 | -0.53 | -0.52 | -0.42 | -0.44 | 0.09 | -0.4 | 0.08 | -0.68 | 0.26 |
| 10950 | -0.27 | -0.37 | -0.14 | -0.36 | -0.6 | 0.38 | -0.1 | 0.24 | -0.34 | 0.38 |
| 891 | -0.28 | -0.59 | -0.34 | -0.31 | -0.49 | 0.32 | -0.17 | 0.34 | -0.15 | 0.27 |
| 890 | -0.13 | -0.1 | -0.22 | -0.1 | -0.24 | 0.16 | -0.1 | 0.05 | -0.19 | 0.09 |
| 5214 | -0.64 | -0.63 | -0.64 | -0.38 | -0.45 | 0.37 | -0.76 | 0.4 | -0.41 | 0.38 |
| 51514 | -0.23 | -0.23 | -0.17 | -0.09 | -0.19 | 0.06 | -0.29 | 0.26 | -0.19 | 0.11 |
| 79682 | -0.09 | -0.22 | -0.17 | -0.22 | -0.15 | 0.08 | -0.19 | 0.02 | -0.26 | 0.01 |
| 6790 | -0.37 | -0.15 | -0.4 | -0.15 | -0.31 | 0.28 | -0.49 | 0.56 | -0.2 | 0.05 |
| 7272 | -0.17 | -0.21 | -0.18 | -0.24 | -0.23 | 0.36 | -0.18 | 0.19 | -0.19 | 0.28 |
| 4602 | -0.27 | -0.29 | -0.21 | -0.2 | -0.25 | 0.05 | -0.51 | 0.34 | -0.26 | 0.27 |
| 2625 | -0.53 | -0.23 | -0.14 | -0.25 | -0.52 | 0.36 | -0.65 | 0.34 | -0.15 | 0.43 |
| 9319 | -0.15 | -0.16 | -0.41 | -0.48 | -0.39 | 0.08 | -0.22 | 0.05 | -0.27 | 0.49 |
| 9156 | -0.33 | -0.28 | -0.27 | -0.28 | -0.32 | 0.4 | -0.28 | 0.03 | -0.41 | 0.09 |
| 9787 | -0.26 | -0.1 | -0.28 | -0.08 | -0.09 | 0.25 | -0.12 | 0.29 | -0.25 | 0.17 |
| 1033 | -0.14 | -0.1 | -0.15 | -0.32 | -0.28 | 0.11 | -0.21 | 0.12 | -0.21 | 0.01 |
| 9493 | -0.1 | -0.07 | -0.25 | -0.07 | -0.27 | 0.02 | -0.18 | 0.28 | -0.15 | 0.22 |

Genómica del cancer



Identificación y validación de
biomarcadores con
poder pronóstico/predictivo en cáncer

Bioplat

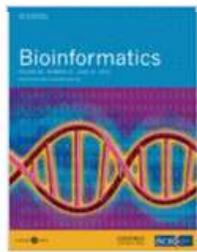


Equipo interdisciplinario



Investigadores y becarios CAETI
Investigadores Conicet
Tesis Maestría UBA
Egresados, tesis y doctorando UNLP





Impact factor 7.307
(2016)

BioPlat: a software for human cancer biomarker discovery

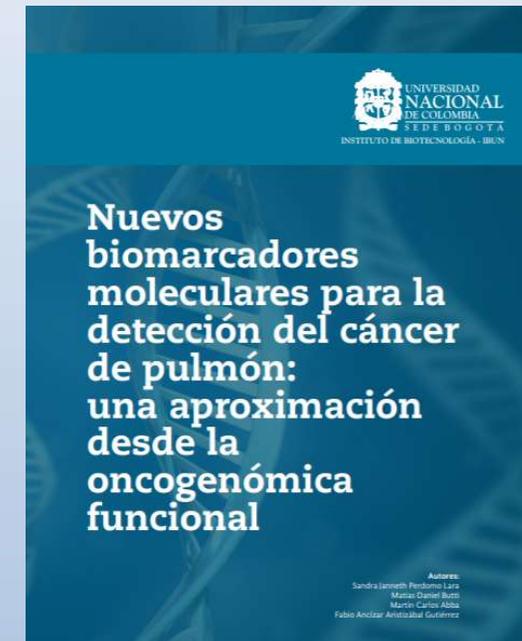
Matias D. Butti, Hernan Chanfreau, Diego Martinez, Diego García, Ezequiel Lacunza, Martin C. Abba ✉

Bioinformatics, Volume 30, Issue 12, 15 June 2014, Pages 1782–1784,
<https://doi.org/10.1093/bioinformatics/btu111>



Breast Cancer Biomarker Discovery

in the Functional Genomic Age: A Systematic Review of 42 Gene Expression Signatures.



Beca Instituto Nacional de Cáncer

1 tesis de grado UNLP, 1 tesis posgrado UBA

En curso: 1 tesis de posgrado UBA, 1 tesis doctoral UNLP, 1 publicación

Premio CIITI estudiantil

Recorrido: Silicon Valley



- Ganadora de Concurso de innovación en 2017
- Presentación en Silicon Valley incluyendo
 - ✓ GenenTech
 - ✓ Genomic Health
 - ✓ Stanford university Tech transfer office
 - ✓ UCSF Medical Center – Ted Goldstein
 - ✓ Folgarty Institute
 - ✓ Inmunoprecise
 - ✓ Tekteam
 - ✓ Miroculus



Genómica del cancer

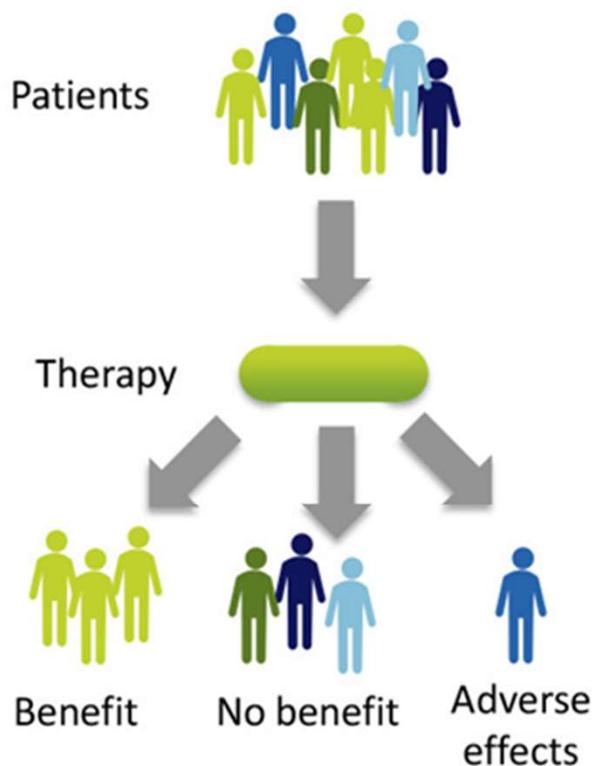


Identificación y validación de
**biomarcadores con
poder pronóstico/predictivo** en cáncer

Objetivo de clasificar a los pacientes con biomarcadores moleculares

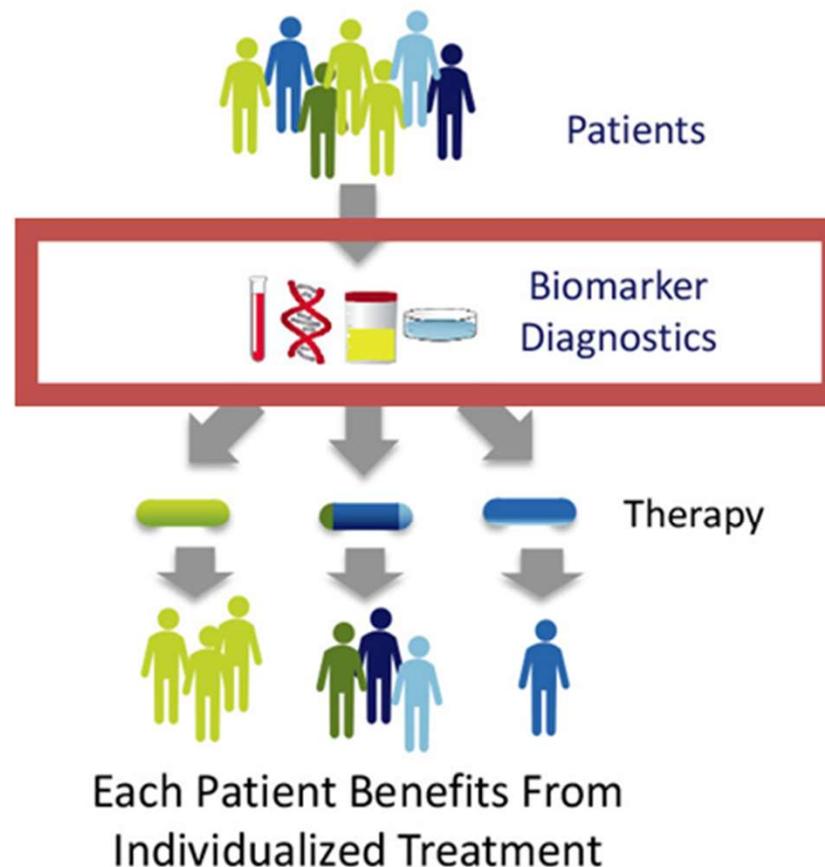
Without Personalized Medicine:

Some Benefit, Some Do Not



With Personalized Medicine:

Each Patient Receives the Right Medicine For Them

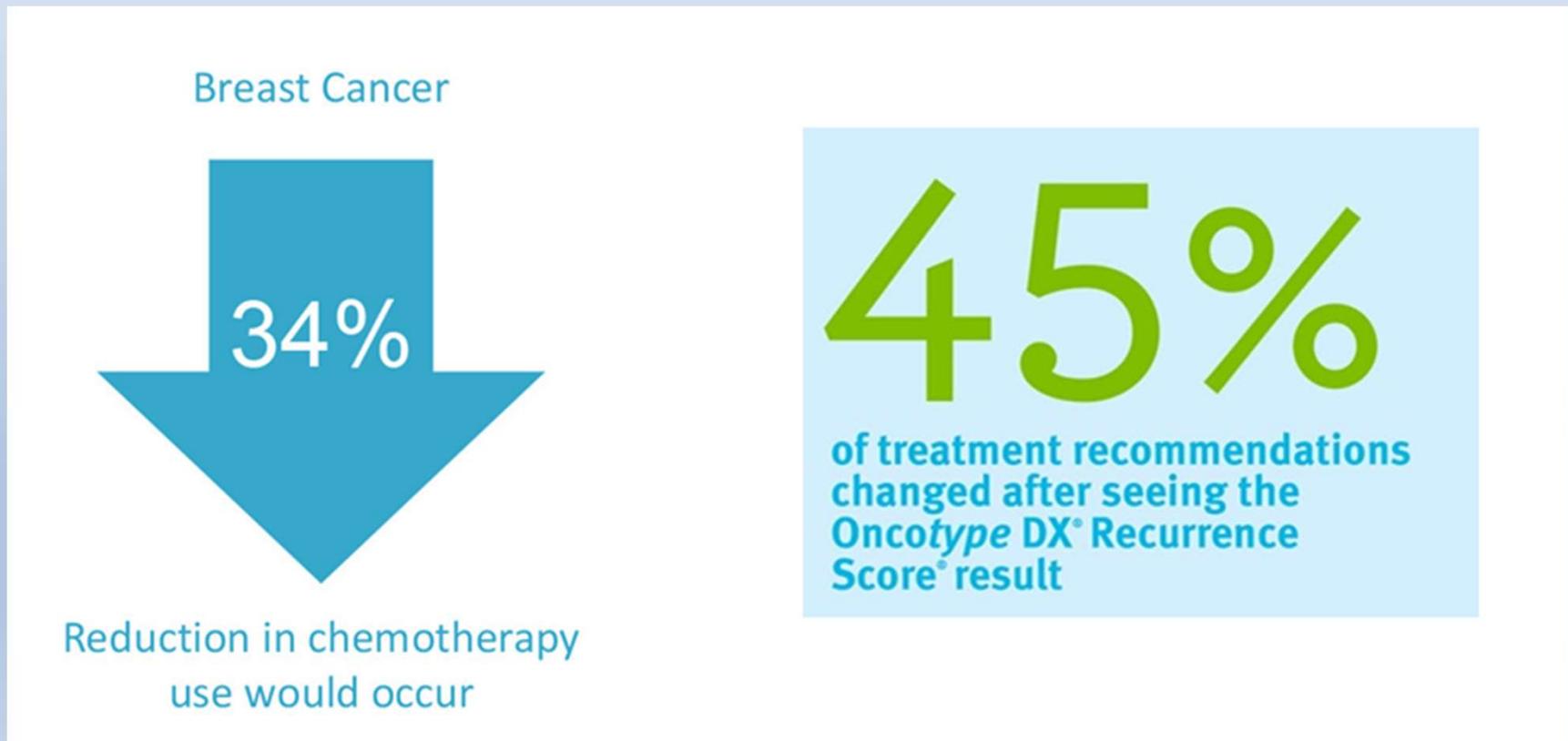


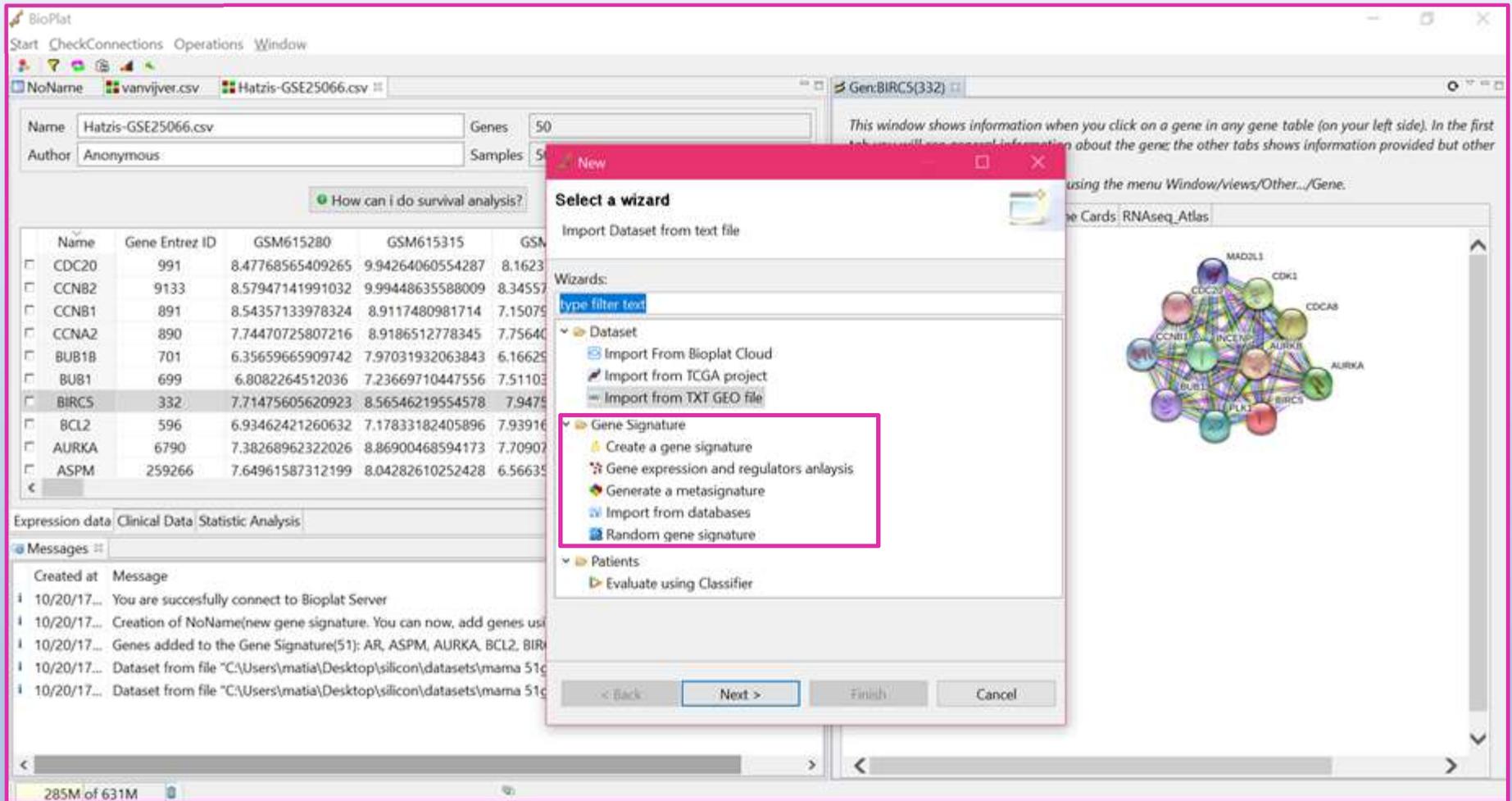
Objetivo de clasificar a los pacientes con biomarcadores moleculares

- **1- Marcadores pronóstico: Capacidad pronostica de supervivencia. Útil para predecir recidiva de un tumor.**
- **2-Marcadores predictivos (en respuesta a tratamientos oncológicos). Companion diagnostics. Predecir si es candidato para una terapia.**
- 3-Entender los mecanismos moleculares de progresión tumoral, con el objetivo de encontrar nuevos blancos terapéuticos específicos para cada clase de tumor.
- ...

Biomarcadores Pronóstico en cáncer

- 1-Marcador pronóstico: Ejemplo en la clínica





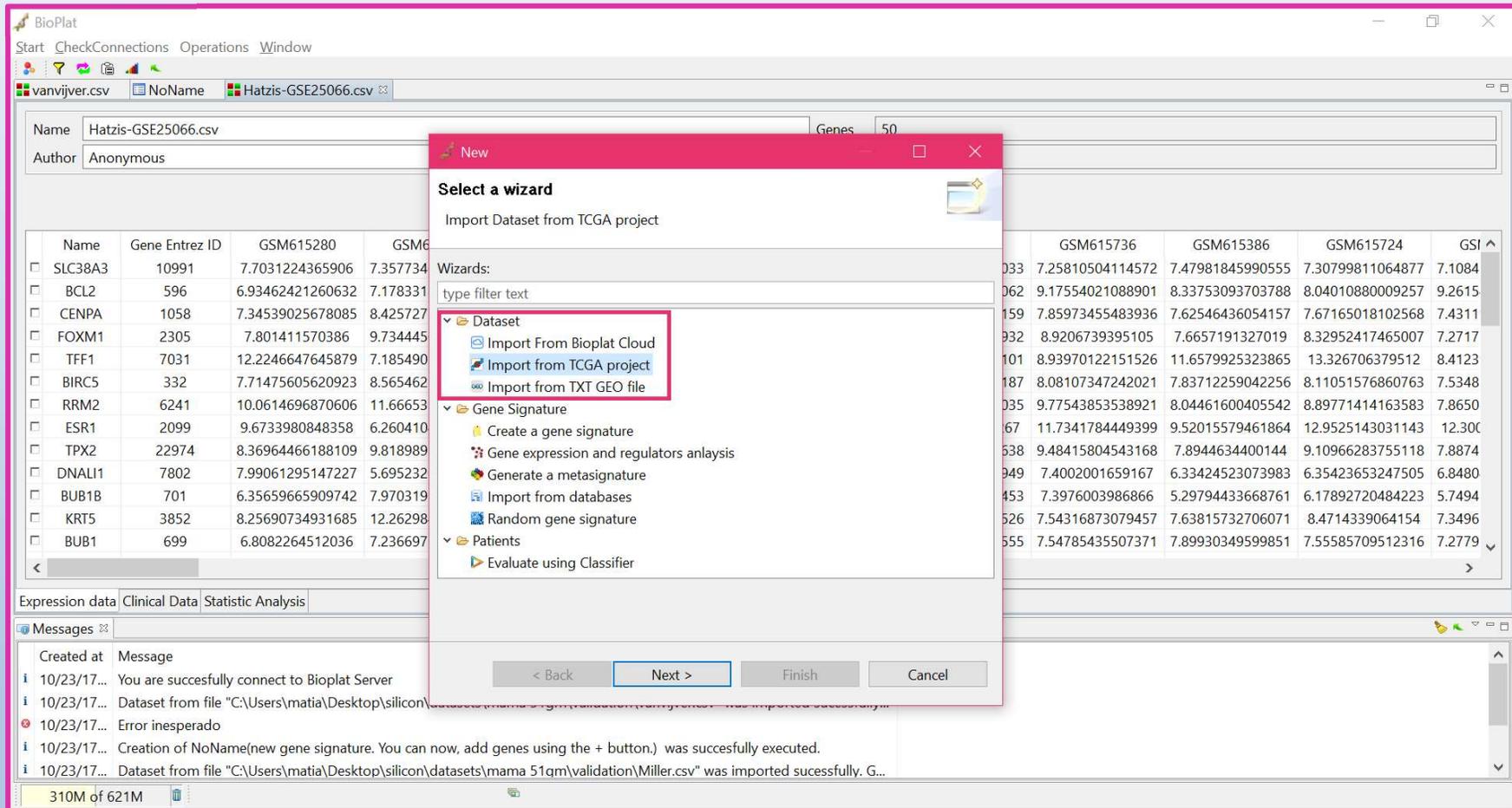
The screenshot shows the BioPlat2 software interface. A 'New' dialog box is open, displaying a 'Select a wizard' menu. The 'Gene Signature' section is highlighted with a pink box, containing the following options:

- Create a gene signature
- Gene expression and regulators analysis
- Generate a metasignature
- Import from databases
- Random gene signature

In the background, a table of gene expression data is visible. The table has columns for Name, Gene Entrez ID, and several GSM IDs. The gene BIRCS (Entrez ID 332) is highlighted in the table.

| Name | Gene Entrez ID | GSM615280 | GSM615315 | GSM615315 |
|--------------------------------|----------------|------------------|------------------|-----------|
| <input type="checkbox"/> CDC20 | 991 | 8.47768565409265 | 9.94264060554287 | 8.1623 |
| <input type="checkbox"/> CCNB2 | 9133 | 8.57947141991032 | 9.99448635588009 | 8.34557 |
| <input type="checkbox"/> CCNB1 | 891 | 8.54357133978324 | 8.9117480981714 | 7.15079 |
| <input type="checkbox"/> CCNA2 | 890 | 7.74470725807216 | 8.9186512778345 | 7.75640 |
| <input type="checkbox"/> BUB1B | 701 | 6.35659665909742 | 7.97031932063843 | 6.16629 |
| <input type="checkbox"/> BUB1 | 699 | 6.8082264512036 | 7.23669710447556 | 7.51103 |
| <input type="checkbox"/> BIRCS | 332 | 7.71475605620923 | 8.56546219554578 | 7.9479 |
| <input type="checkbox"/> BCL2 | 596 | 6.93462421260632 | 7.17833182405896 | 7.93916 |
| <input type="checkbox"/> AURKA | 6790 | 7.38268962322026 | 8.86900468594173 | 7.70907 |
| <input type="checkbox"/> ASPM | 259266 | 7.64961587312199 | 8.04282610252428 | 6.56639 |

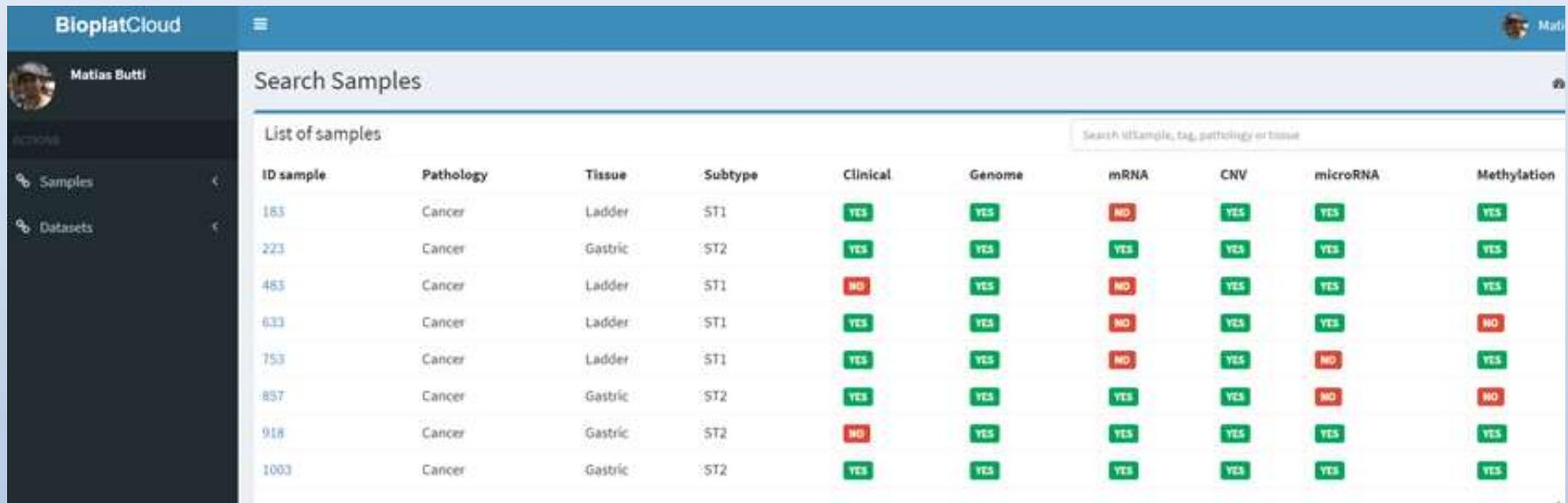
1-Get your initial gene list



The screenshot shows the BioPlat application window with a 'New' wizard dialog box open. The dialog box is titled 'Select a wizard' and has a subtitle 'Import Dataset from TCGA project'. It contains a list of wizards under the 'Dataset' category, with 'Import from TCGA project' highlighted by a red box. Other wizards include 'Import From Bioplat Cloud', 'Import from TXT GEO file', 'Gene Signature', 'Patients', and 'Evaluate using Classifier'. The background shows a data table with columns for Name, Gene Entrez ID, GSM615280, GSM615736, GSM615386, GSM615724, and GSI. The message pane at the bottom shows a successful import message: 'Dataset from file "C:\Users\matia\Desktop\silicon\datasets\mama_51qm\validation\Miller.csv" was imported successfully. G...'

| Name | Gene Entrez ID | GSM615280 | GSM615736 | GSM615386 | GSM615724 | GSI |
|----------------------------------|----------------|------------------|-----------|-----------|-----------|-----|
| <input type="checkbox"/> SLC38A3 | 10991 | 7.7031224365906 | 7.357734 | | | |
| <input type="checkbox"/> BCL2 | 596 | 6.93462421260632 | 7.178331 | | | |
| <input type="checkbox"/> CENPA | 1058 | 7.34539025678085 | 8.425727 | | | |
| <input type="checkbox"/> FOXM1 | 2305 | 7.801411570386 | 9.734445 | | | |
| <input type="checkbox"/> TFF1 | 7031 | 12.2246647645879 | 7.185490 | | | |
| <input type="checkbox"/> BIRC5 | 332 | 7.71475605620923 | 8.565462 | | | |
| <input type="checkbox"/> RRM2 | 6241 | 10.0614696870606 | 11.66653 | | | |
| <input type="checkbox"/> ESR1 | 2099 | 9.6733980848358 | 6.260410 | | | |
| <input type="checkbox"/> TPX2 | 22974 | 8.36964466188109 | 9.818989 | | | |
| <input type="checkbox"/> DNALI1 | 7802 | 7.99061295147227 | 5.695232 | | | |
| <input type="checkbox"/> BUB1B | 701 | 6.35659665909742 | 7.970319 | | | |
| <input type="checkbox"/> KRT5 | 3852 | 8.25690734931685 | 12.26298 | | | |
| <input type="checkbox"/> BUB1 | 699 | 6.8082264512036 | 7.236697 | | | |

2-Import datasets from different sources



The screenshot shows the BioPlatCloud interface. On the left is a navigation sidebar with 'Samples' and 'Datasets'. The main area is titled 'Search Samples' and contains a table with the following data:

| ID sample | Pathology | Tissue | Subtype | Clinical | Genome | mRNA | CNV | microRNA | Methylation |
|-----------|-----------|---------|---------|----------|--------|------|-----|----------|-------------|
| 183 | Cancer | Ladder | ST1 | YES | YES | NO | YES | YES | YES |
| 223 | Cancer | Gastric | ST2 | YES | YES | YES | YES | YES | YES |
| 483 | Cancer | Ladder | ST1 | NO | YES | NO | YES | YES | YES |
| 633 | Cancer | Ladder | ST1 | YES | YES | NO | YES | YES | NO |
| 753 | Cancer | Ladder | ST1 | YES | YES | NO | YES | NO | YES |
| 857 | Cancer | Gastric | ST2 | YES | YES | YES | YES | NO | NO |
| 918 | Cancer | Gastric | ST2 | NO | YES | YES | YES | YES | YES |
| 1003 | Cancer | Gastric | ST2 | YES | YES | YES | YES | YES | YES |

2-Import datasets from different sources
Bioplat cloud

The screenshot shows the BioPlat software interface. The main window displays a table with columns for 'Experiment', 'Concordance Index', 'Log-Rank Test P-value', and 'ROC AUC'. Two experiments are listed: 'vanvijver.csv-C2' and 'Hatzis-GSE25066.csv-C2'. A 'Statistic and Graphics' dialog box is open, showing options for 'Log Rank Test Chi-Squared', 'Concordance Index', and 'ROC'. Below the dialog, a large data table is visible with columns for 'OS_MONTHS', 'OS_STATUS', and various gene expression values (sample1 to sample11). The 'OS_STATUS' column has values like 'RECURRENT', 'NON REC.', and 'RECURRENT'. The gene expression values range from -0.966 to 0.7049. At the bottom, there is a list of 'Optimized 51gmsin...' files with associated IDs and values.

| Experiment | Concordance Index | Log-Rank Test P-value | ROC AUC | Statistic and |
|------------------------|--------------------|-----------------------|--------------------|---------------|
| vanvijver.csv-C2 | 0.7609882217753519 | 2.052325808188371E-7 | 0.6540143161451668 | |
| Hatzis-GSE25066.csv-C2 | 0.7903985507246377 | 5.3842610370224975... | 0.6513784550086983 | |

| OS_MONTHS | 1 | 50 | 51 | 4 | 48 | 47.27 | 5 | 52 | 55 | 52 | 7 |
|-----------|-----------|----------|----------|-----------|-----------|----------|-----------|-----------|----------|----------|-----------|
| OS_STATUS | RECURRENT | NON REC. | NON REC. | RECURRENT | RECURRENT | NON REC. | RECURRENT | RECURRENT | NON REC. | NON REC. | RECURRENT |
| sampleId | Sample1 | sample2 | sample3 | sample4 | sample5 | sample6 | sample7 | sample8 | sample9 | sample10 | sample11 |
| CD68 | -0.3522 | -0.2522 | -0.1522 | -0.0522 | -0.0422 | -0.0322 | -0.0222 | -0.0122 | -0.0022 | 0.0078 | 0.0178 |
| PGR | -0.6815 | -0.5815 | -0.4815 | -0.3815 | -0.3715 | -0.3615 | -0.3515 | -0.3415 | -0.3315 | -0.3215 | -0.3115 |
| GRB7 | -0.9102 | -0.8102 | -0.7102 | -0.6102 | -0.6002 | -0.5902 | -0.5802 | -0.5702 | -0.5602 | -0.5502 | -0.5402 |
| MKI67 | 0.27 | 0.3700 | 0.4700 | 0.5700 | 0.5800 | 0.5900 | 0.6000 | 0.6100 | 0.6200 | 0.6300 | 0.6400 |
| BAG1 | -0.1552 | -0.0552 | 0.0448 | 0.1448 | 0.1548 | 0.1648 | 0.1748 | 0.1848 | 0.1948 | 0.2048 | 0.2148 |
| BIRC5 | 0.0843 | 0.1843 | 0.2843 | 0.3843 | 0.3943 | 0.4043 | 0.4143 | 0.4243 | 0.4343 | 0.4443 | 0.4543 |
| MYBL2 | 0.3349 | 0.4349 | 0.5349 | 0.6349 | 0.6449 | 0.6549 | 0.6649 | 0.6749 | 0.6849 | 0.6949 | 0.7049 |
| SCUBE2 | -0.2055 | -0.1055 | -0.0055 | 0.0945 | 0.1045 | 0.1145 | 0.1245 | 0.1345 | 0.1445 | 0.1545 | 0.1645 |
| BCL2 | -0.966 | -0.8660 | -0.7660 | -0.6660 | -0.6560 | -0.6460 | -0.6360 | -0.6260 | -0.6160 | -0.6060 | -0.5960 |
| GSTM1 | 0.1682 | 0.2682 | 0.3682 | 0.4682 | 0.4782 | 0.4882 | 0.4982 | 0.5082 | 0.5182 | 0.5282 | 0.5382 |

3-Statistical validation

| OS_MONTHS | 1 | 50 | 51 | 4 | 48 | 47.27 | 5 | 52 | 55 | 52 | 7 |
|-----------|------------|----------|----------|------------|------------|----------|------------|------------|----------|----------|------------|
| OS_STATUS | RECURRENCE | NON REC. | NON REC. | RECURRENCE | RECURRENCE | NON REC. | RECURRENCE | RECURRENCE | NON REC. | NON REC. | RECURRENCE |
| sampleId | Sample1 | sample2 | sample3 | sample4 | sample5 | sample6 | sample7 | sample8 | sample9 | sample10 | sample11 |
| CD68 | -0.3522 | -0.2522 | -0.1522 | -0.0522 | -0.0422 | -0.0322 | -0.0222 | -0.0122 | -0.0022 | 0.0078 | 0.0178 |
| PGR | -0.6815 | -0.5815 | -0.4815 | -0.3815 | -0.3715 | -0.3615 | -0.3515 | -0.3415 | -0.3315 | -0.3215 | -0.3115 |
| GRB7 | -0.9102 | -0.8102 | -0.7102 | -0.6102 | -0.6002 | -0.5902 | -0.5802 | -0.5702 | -0.5602 | -0.5502 | -0.5402 |
| MKI67 | 0.27 | 0.3700 | 0.4700 | 0.5700 | 0.5800 | 0.5900 | 0.6000 | 0.6100 | 0.6200 | 0.6300 | 0.6400 |
| BAG1 | -0.1552 | -0.0552 | 0.0448 | 0.1448 | 0.1548 | 0.1648 | 0.1748 | 0.1848 | 0.1948 | 0.2048 | 0.2148 |
| BIRC5 | 0.0843 | 0.1843 | 0.2843 | 0.3843 | 0.3943 | 0.4043 | 0.4143 | 0.4243 | 0.4343 | 0.4443 | 0.4543 |
| MYBL2 | 0.3349 | 0.4349 | 0.5349 | 0.6349 | 0.6449 | 0.6549 | 0.6649 | 0.6749 | 0.6849 | 0.6949 | 0.7049 |
| SCUBE2 | -0.2055 | -0.1055 | -0.0055 | 0.0945 | 0.1045 | 0.1145 | 0.1245 | 0.1345 | 0.1445 | 0.1545 | 0.1645 |
| BCL2 | -0.966 | -0.8660 | -0.7660 | -0.6660 | -0.6560 | -0.6460 | -0.6360 | -0.6260 | -0.6160 | -0.6060 | -0.5960 |
| GSTM1 | 0.1682 | 0.2682 | 0.3682 | 0.4682 | 0.4782 | 0.4882 | 0.4982 | 0.5082 | 0.5182 | 0.5282 | 0.5382 |

3-Statistical validation

BioPlat

Start CheckConnections Operations Window

ALIX50 Datamining Regulators 51gm Vanvijver.csv Hatzis-GSE25066.csv brca_tcga_3way_complete Miller.csv

| Experiment | Concordance Index | Log-Rank Test P-value | ROC AUC | Statistic and... | View Used... | Open validating ... | Export Gene Signature... | Copy R Sc... | Preprocess... |
|--|--------------------|-----------------------|--------------------|------------------|--------------|---------------------|--------------------------|--------------|---------------|
| <input type="checkbox"/> Hatzis-GSE25066.cs... | 0.7903985507246377 | 5.3842610370224975... | 0.6513784550086983 | | | | | | |
| <input type="checkbox"/> Miller.csv-C2 | 0.69788835458917 | 0.0051804004033707... | 0.5479884556492252 | | | | | | |
| <input type="checkbox"/> Vanvijver.csv-C2 | 0.7609882217753519 | 2.052325808188371E-7 | 0.6540143161451668 | | | | | | |

Optimum

Genomic space

Particle Swarm Optimization

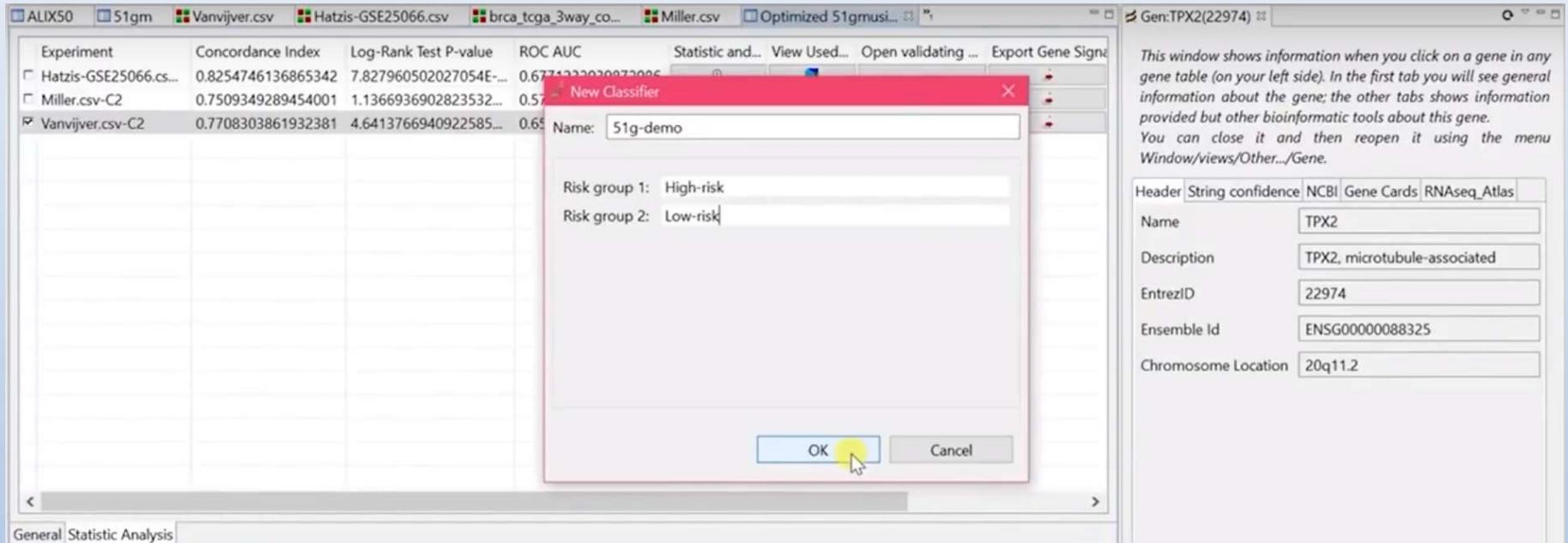
General | **Statistic Analysis**

Messages **Top Ten PSO Result**

| Gene Signature name | Number ... | Training (Concordance index) | Testing (Concordance index) | Validation (Concordance index) |
|--|------------|------------------------------|-----------------------------|--------------------------------|
| <input checked="" type="checkbox"/> Optimized 51gmusing PSO (23 genes) | 23 | 0.7708303861932381 | 0.8254746136865342 | 0.7509349289454001 |
| <input type="checkbox"/> Optimized 51gmusing PSO (28 genes) | 28 | 0.747072951065746 | 0.8232538147431765 | 0.6563328033916269 |
| <input type="checkbox"/> Optimized 51gmusing PSO (25 genes) | 25 | 0.7265434251831183 | 0.8070151981102643 | 0.6467611336032388 |
| <input type="checkbox"/> Optimized 51gmusing PSO (28 genes) | 28 | 0.7390265750970438 | 0.7868611243346173 | 0.6960172744721689 |
| <input type="checkbox"/> Optimized 51gmusing PSO (25 genes) | 25 | 0.72043757596805 | 0.7847819307702588 | 0.6995327102803738 |

585M of 708M

4-Optimization



The screenshot shows the BioPlat2 software interface. A 'New Classifier' dialog box is open in the center, with the following fields:

- Name: 51g-demo
- Risk group 1: High-risk
- Risk group 2: Low-risk

At the bottom of the dialog box are 'OK' and 'Cancel' buttons. A mouse cursor is pointing at the 'OK' button.

In the background, a table displays experimental data:

| Experiment | Concordance Index | Log-Rank Test P-value | ROC AUC |
|--|--------------------|------------------------|--------------------|
| <input type="checkbox"/> Hatzis-GSE25066.cs... | 0.8254746136865342 | 7.827960502027054E-... | 0.6771222020972006 |
| <input type="checkbox"/> Miller.csv-C2 | 0.7509349289454001 | 1.1366936902823532... | 0.57... |
| <input checked="" type="checkbox"/> Vanvijver.csv-C2 | 0.7708303861932381 | 4.6413766940922585... | 0.65... |

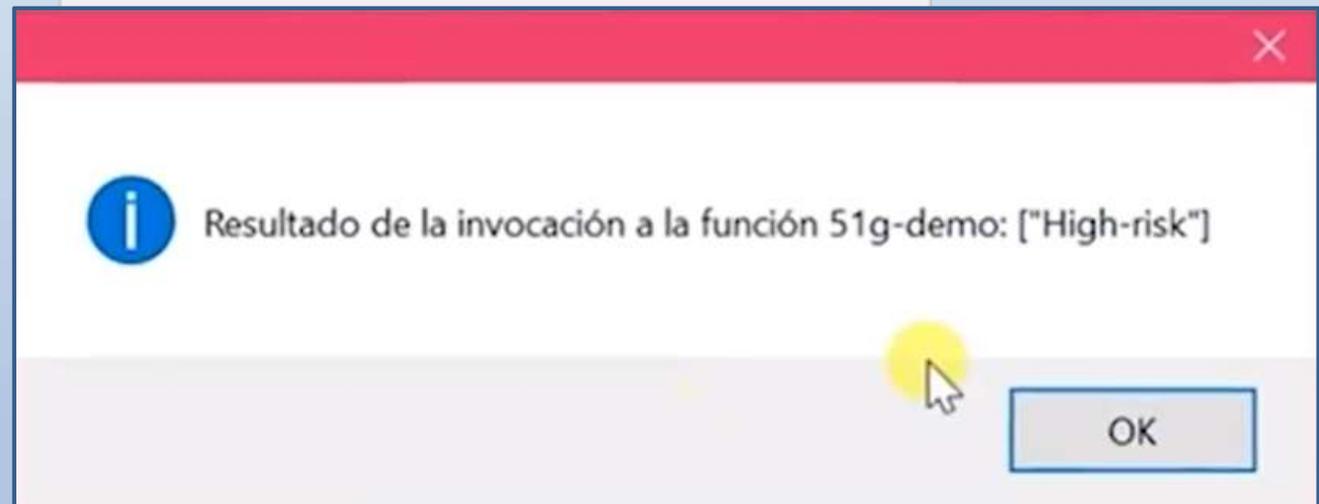
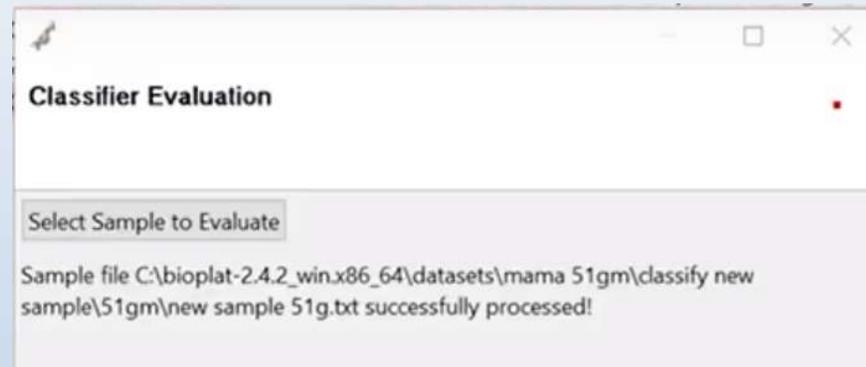
On the right side, a gene information panel for 'Gen:TPX2(22974)' is visible. It contains the following information:

- Name: TPX2
- Description: TPX2, microtubule-associated
- EntrezID: 22974
- Ensemble Id: ENSG00000088325
- Chromosome Location: 20q11.2

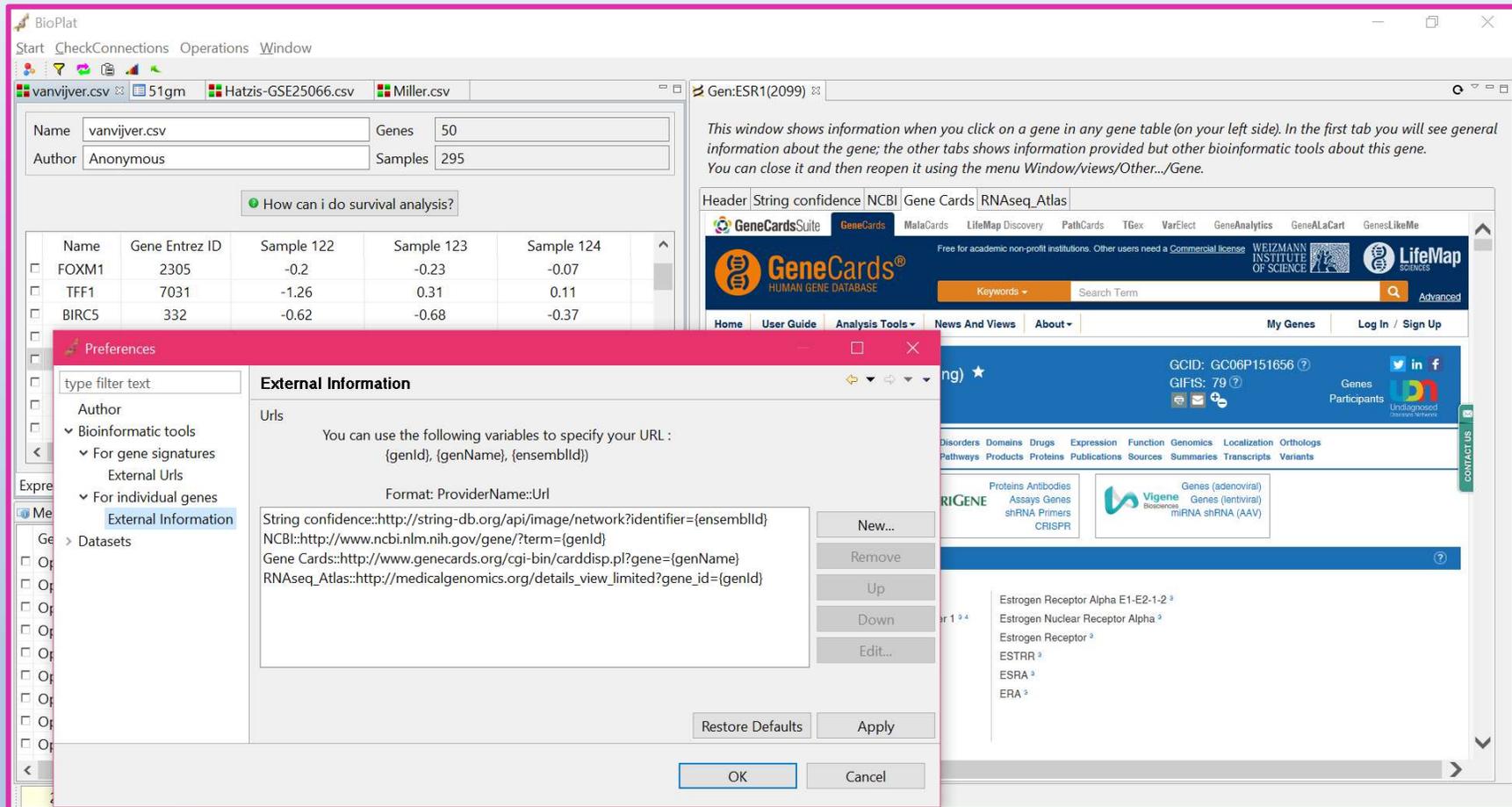
Below the gene information, there are tabs for 'String confidence', 'NCBI', 'Gene Cards', and 'RNAseq Atlas'. The 'General' and 'Statistic Analysis' tabs are visible at the bottom of the main window.

5-Create classifier

| | |
|---------|-------|
| AURKA | -0.41 |
| CCNA2 | -0.24 |
| CCNB1 | -0.3 |
| CCNB2 | -0.58 |
| CDC20 | -0.18 |
| CEP55 | -0.26 |
| DUSP4 | 0.63 |
| ESR1 | -0.41 |
| FOS | 0.32 |
| FOXA1 | -0.1 |
| KPNA2 | -0.47 |
| MELK | -0.32 |
| MYB | -0.15 |
| MYBL2 | -0.14 |
| NEK2 | -0.3 |
| PRC1 | -0.41 |
| RRM2 | -0.63 |
| SCUBE2 | -0.16 |
| SLC38A3 | -0.02 |
| SLC39A6 | 0.35 |
| TPX2 | -0.15 |
| TRIP13 | -0.26 |
| VAV3 | 0.34 |



6-Evaluate classifier



The screenshot displays the BioPlat software interface. On the left, a table shows gene expression data for samples 122, 123, and 124. A 'Preferences' dialog box is open, showing the 'External Information' section with a list of URLs for external data sources. On the right, a web browser window shows the GeneCards website for the gene ESR1, displaying various data points and links to related resources.

Table Data:

| Name | Gene Entrez ID | Sample 122 | Sample 123 | Sample 124 |
|-------|----------------|------------|------------|------------|
| FOXM1 | 2305 | -0.2 | -0.23 | -0.07 |
| TFF1 | 7031 | -1.26 | 0.31 | 0.11 |
| BIRC5 | 332 | -0.62 | -0.68 | -0.37 |

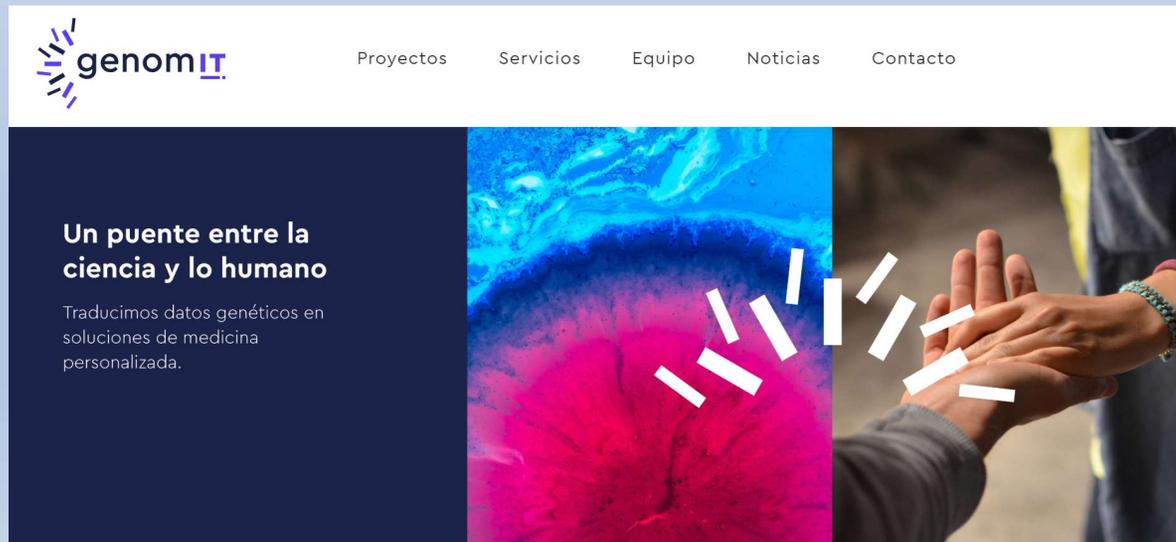
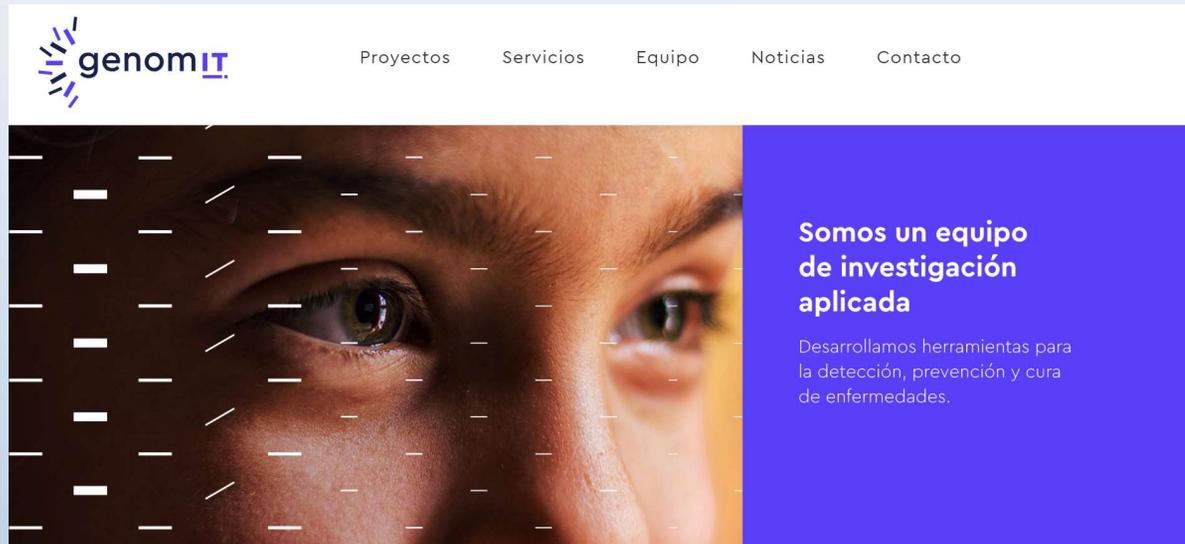
External Information URLs:

- String confidence: <http://string-db.org/api/image/network?identifier={ensemblid}>
- NCBI: <http://www.ncbi.nlm.nih.gov/gene/?term={genld}>
- Gene Cards: <http://www.genecards.org/cgi-bin/carddisp.pl?gene={genName}>
- RNAseq Atlas: http://medicalgenomics.org/details_view_limited?gene_id={genld}

Companion tools



Recorrido



Recorrido

Equipo médico

Equipo biotecnológico

Equipo de ciencia de datos biomédicos

Desarrollo de software



CONICET



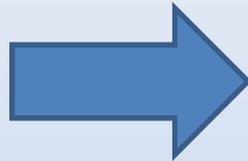
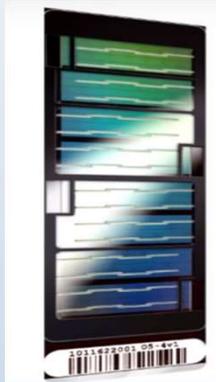
EBT

Recorrido

- ✓ Servicios de bioinformática y ciencia de datos biológicos.
- ✓ Actualmente trabajando con 3 hospitales en problemáticas que necesitaban una solución genómica. Próspera colaboración con el hospital Escuela de UAI.
- ✓ Apoyamos desde la empresa el grupo de investigación de “bioinformática en oncogenómica funcional”.



- ✓ Diseño de estudios genéticos para médicos derivantes o pacientes directamente.
- ✓ Paneles de genes de terceros: cáncer y fertilidad con la interpretación y el asesoramiento genético pertinente.
- ✓ Octubre 2018: Expansión de Zoigen a Chile.
- ✓ 2019: Capitales privados.



Proceso Bioinformático

- ✓ Preprocesamiento para cada una de las tecnologías de secuenciación que utilizamos.
- ✓ Imputación de genotipos



Algoritmo de Interpretación

- ✓ Algoritmo definido de forma interdisciplinaria para interpretación de información genómica
- ✓ Generación del reporte para el medico que realiza la devolución



Machine learning

- ✓ Base de datos genómica
- ✓ Machine learning sobre datos genómicos y clínicos



Desarrollo web

Sistemas de gestión interna
Sistemas web para medicos
Sistemas web para pacientes



Recorrido



CARRIER GEN

Se analizan cientos de genes relacionados con enfermedades recesivas. El 50% de las personas es portadora de al menos una mutación de este tipo, y 2 a 5% de las parejas comparten mutaciones en el mismo gen.

➤ Realizar estudio

CorionGen TEST DEL MATERIAL DE ABORTO

—

Estudio en material de aborto.

Determinación de anomalías cromosómicas en el producto del embarazo para evitar estudios y tratamientos innecesarios.

➤ Saber más

ScreenGen TEST POSTNATAL

—

Estudio posnatal.

Detección de enfermedades hereditarias en neonatos pasibles de tratamiento precoz para evitar la aparición de síntomas.

➤ Saber más

BabyGen TEST PRENATAL NO INVASIVO

—

ADN fetal en sangre materna.

Estudio prenatal no invasivo. Detección de anomalías cromosómicas fetales durante el embarazo.

➤ Saber más

TEST DE PATERNIDAD NO INVASIVO

—

Estudio prenatal no invasivo

para definición de la paternidad durante el embarazo.

➤ Saber más

EmbryoGen TEST PREIMPLANTATORIO

—

Estudio preimplantatorio (PGD).

Detección de alteraciones genéticas en embriones logrados por fecundación in vitro para optimizar los resultados reproductivos y reducir riesgos para la descendencia.

➤ Saber más

Recorrido



FIBROSIS QUÍSTICA

Indicado para pacientes con antecedentes familiares de la enfermedad o varones con falta de producción de espermatozoides de causa obstructiva.

➤ Realizar estudio



FRAGILIDAD DEL X

—

Estudio molecular para Fragilidad del X.
Determinación de expansiones en gen FMR1 mediante métodos de alta sensibilidad.

⊕ Saber más



DELECCIONES DEL CROMOSOMA Y

—

Estudio molecular para microdeleciones del cromosoma Y.
Determinación de la presencia o ausencia de genes AZF.

⊕ Saber más



—

Test de fertilidad masculino

Busca causas genéticas por la que vos o tu pareja están teniendo dificultades reproductivas.

⊕ Saber más



—

Test de fertilidad femenino

Busca causas genéticas por la que vos o tu pareja están teniendo dificultades reproductivas.

⊕ Saber más



—

Estudio molecular del cariotipo.

Análisis del cariotipo para descartar anomalías cromosómicas.

⊕ Saber más

Recorrido



[Servicios](#)

[Nosotros](#)

[Contacto](#)

[Registrarme](#) | [Mi Perfil](#) 

CONOCÉ TU ADN. DESCUBRITE.

Con Zoigen podés conocer tu perfil genético y mejorar tu calidad de vida. Accedé a nuestros estudios y alcanzá tus objetivos.

[➔ Comprar](#)



z.fit

Conocé tu adn,
alcanzá tu
potencial.



¡Hola! - ¿Te puedo ayudar?



[Mirá el video de Zoigen-Z.fit](#)



Gracias por su atención!



| genomit.com.ar |



| zoigen.com |



| [@zoigen.Argentina](https://www.facebook.com/zoigen.Argentina) |



| [@zoigen](https://www.instagram.com/zoigen)

matias@zoigen.com

sebastian@zoigen.com