



PRINCIPE FELIPE

CENTRO DE INVESTIGACION

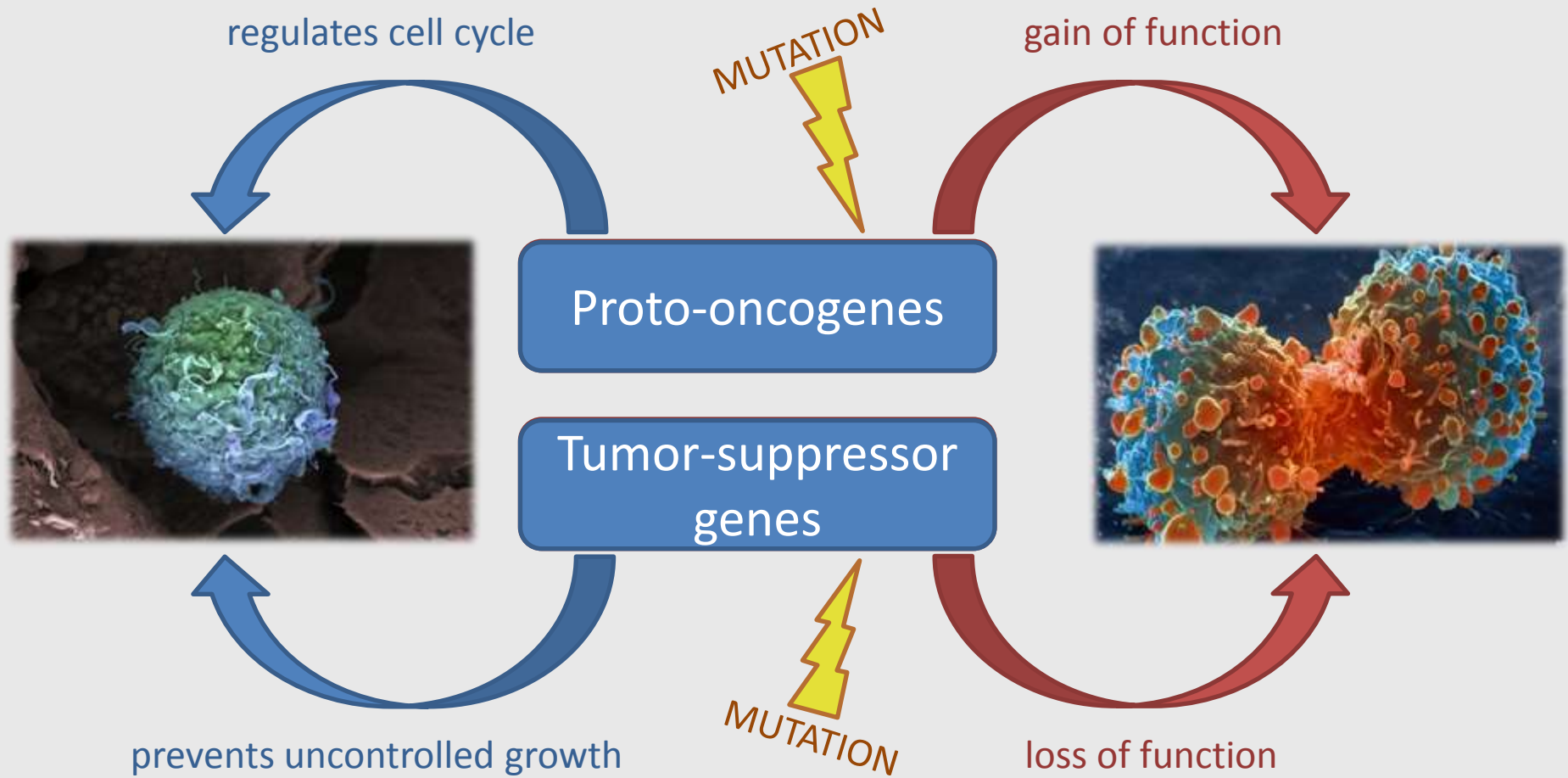
Computational Genomics Laboratory

The pan-cancer pathological regulatory landscape

Matías Marín Falco

Joaquín Dopazo

José Carbonell Caballero



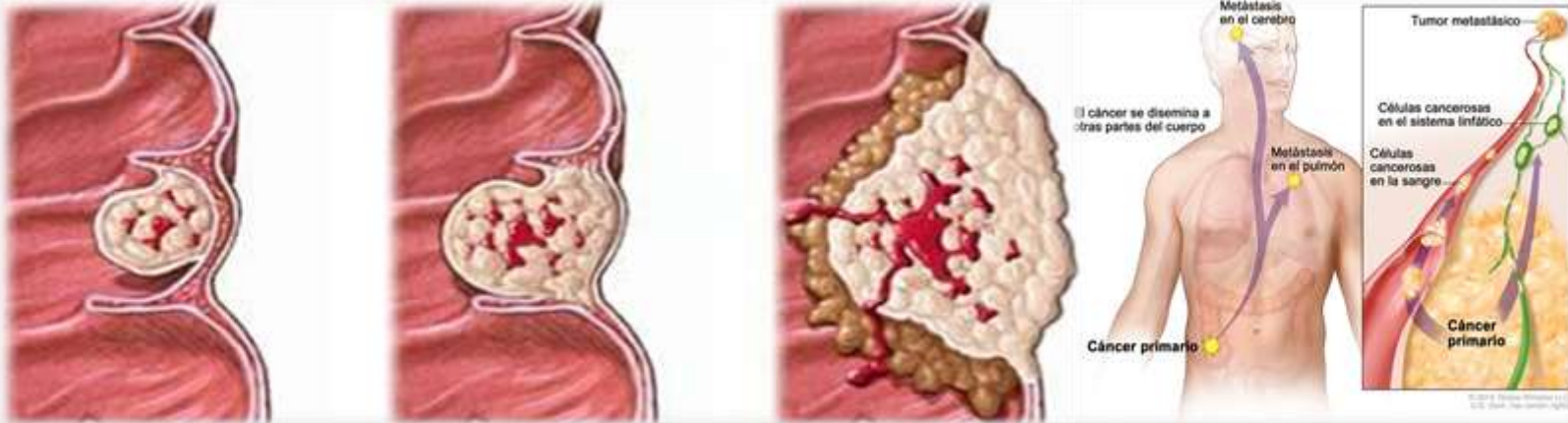
Cancer

Stage I

Stage II

Stage III

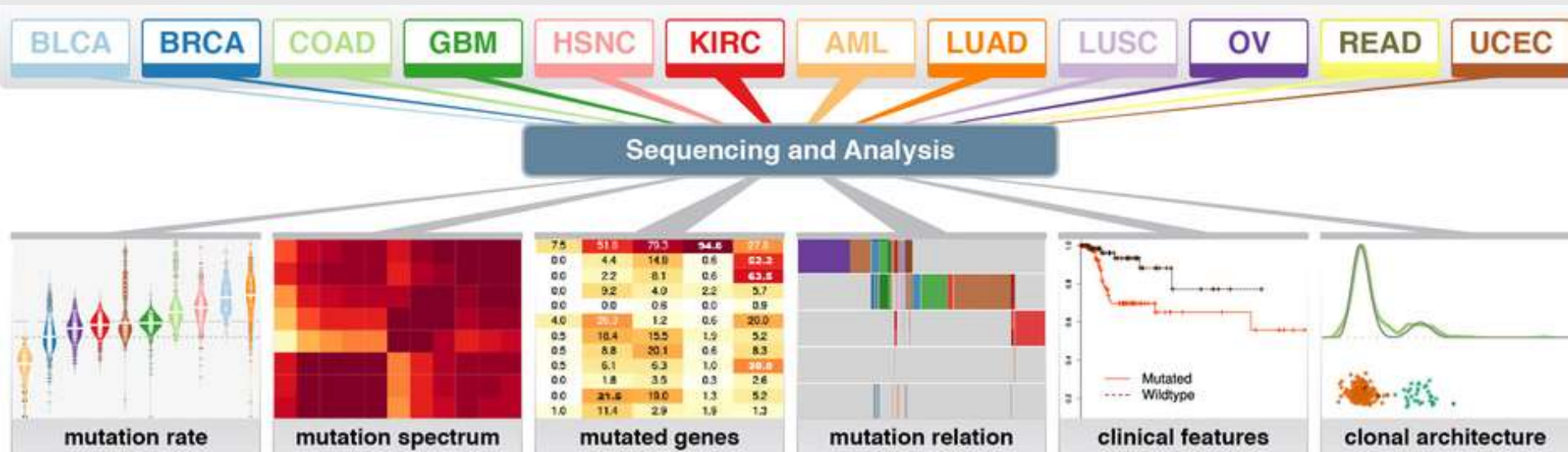
Stage IV



- Stage I: tumor are localized to one part of the body. Stage I cancer can be surgically removed if small enough.
- Stage II: tumor are locally advanced. Stage II cancer can be treated by chemo, radiation, or surgery.
- Stage III: tumors are also locally advanced. Tumor may have spread to nearby tissues. Stage III can be treated by chemo, radiation, or surgery.
- Stage IV: cancers have often metastasized, or spread to other organs or throughout the body.

Pan-cancer

Pan-cancer analysis aims to examine the similarities and differences among the genomic and cellular alterations found across diverse tumor types.



International
Cancer Genome
Consortium



OBJECTIVES

It was proposed to perform a pan-cancer analysis and observe the relations between the alteration of some Transcription Factors and some clinical endpoints

Selecting cancer types

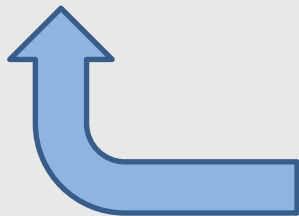
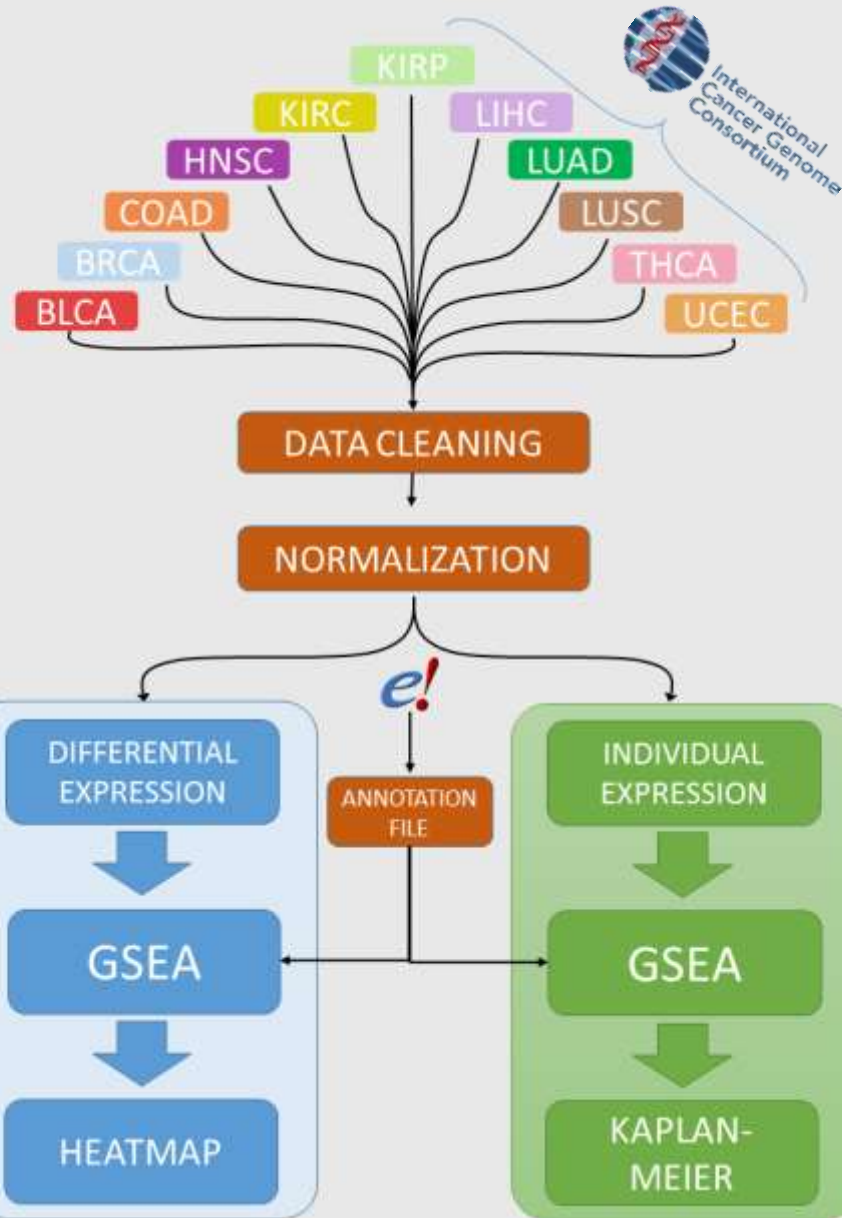
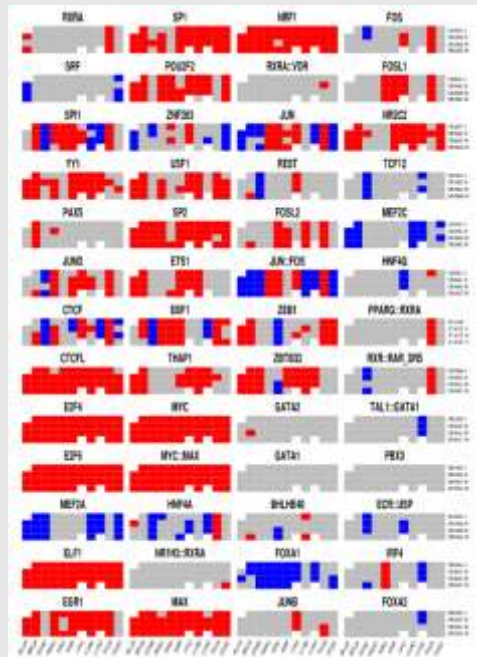
Retrieve data from selected cancers and process it

Perform statistical analysis in different cancer types

Analyze the results as a whole, not individualizing

Perform a reproducible research

METHODS

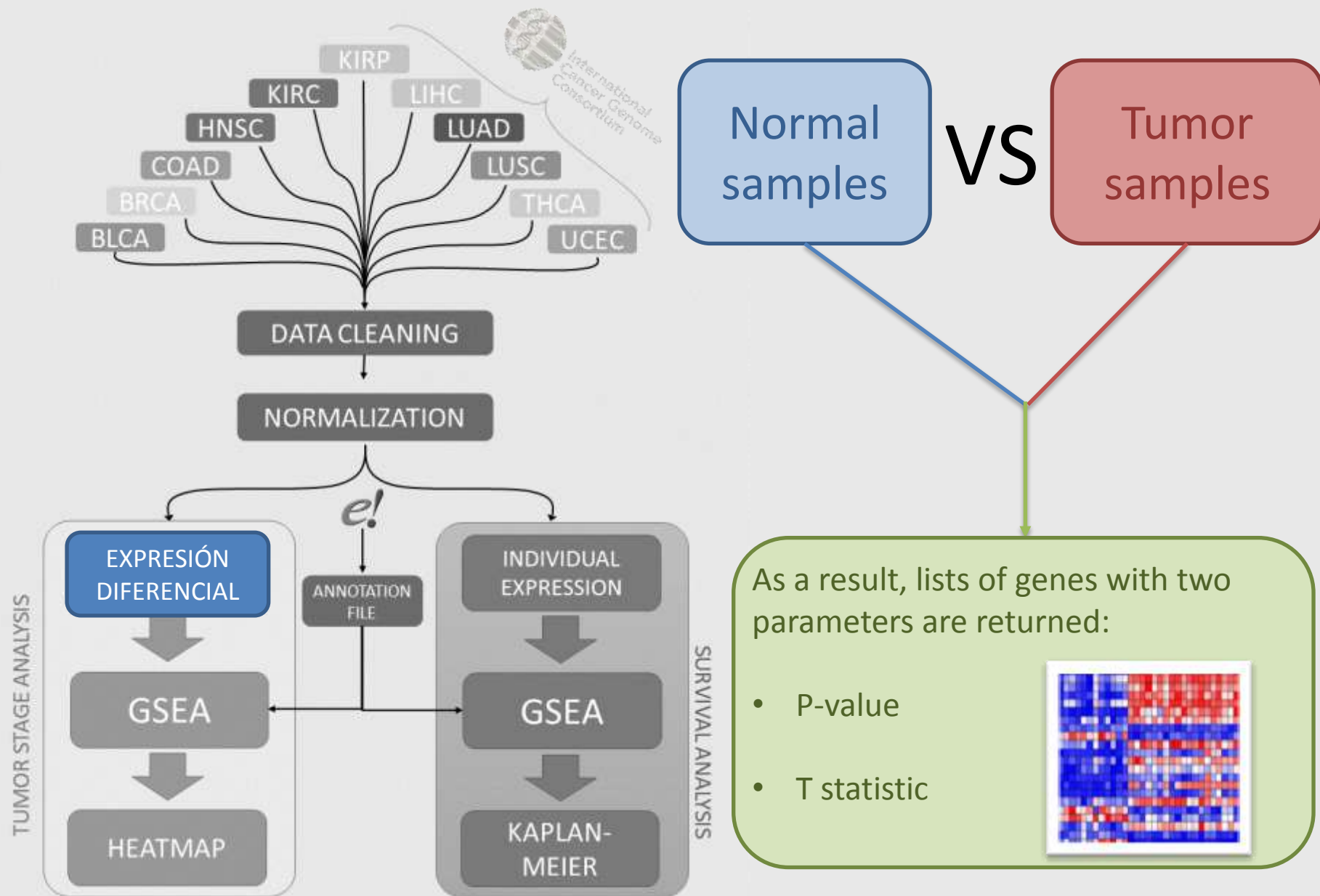


TUMOR STAGE ANALYSIS



SURVIVAL ANALYSIS

Differential Expression



Transcription Factor Target Enrichment Analysis (TFTEA)

GSEA

The GSEA is a computational method that determines whether an a priori defined block of genes shows statistically significant, concordant differences between two biological states.

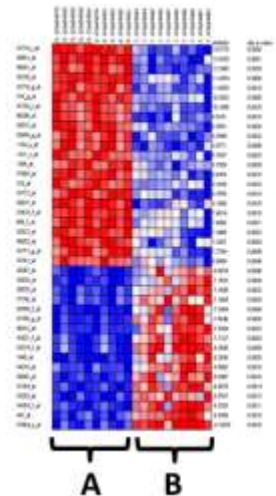
Blocks

Annotation file

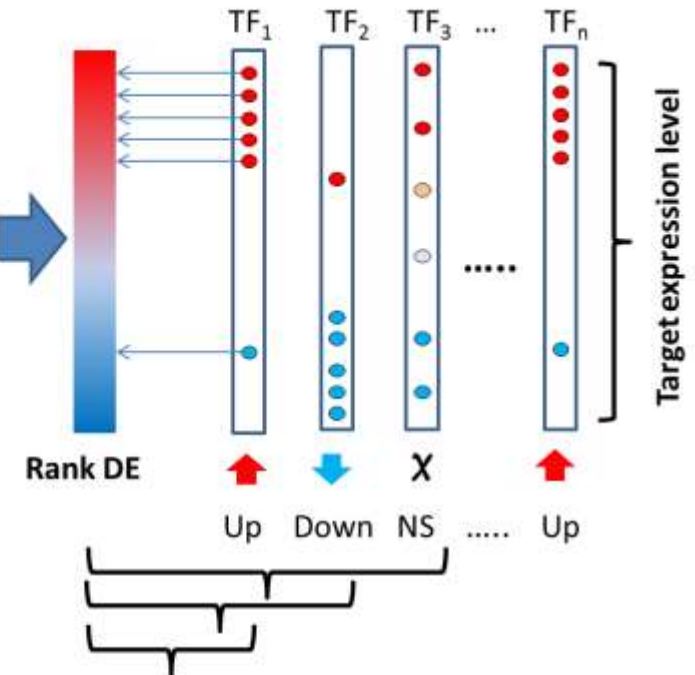
Set of genes

Differential expression

Ranked list of genes



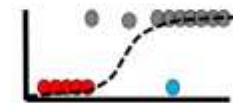
Blocks of genes



TF target expression in the rank



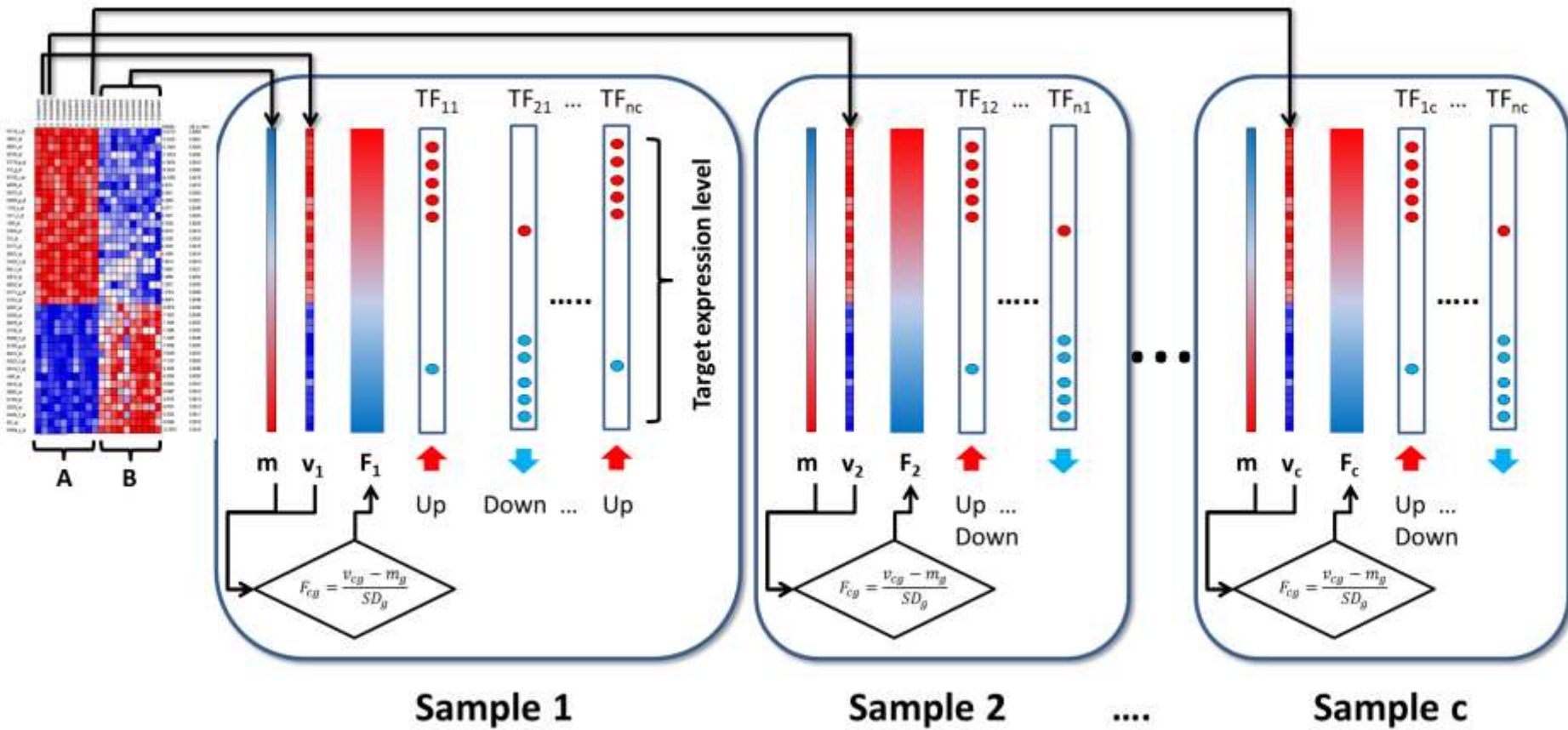
$$\ln \left(\frac{P(g \in F)}{P(g \notin F)} \right) = K + \alpha X$$



Survival analysis

DEFINITION

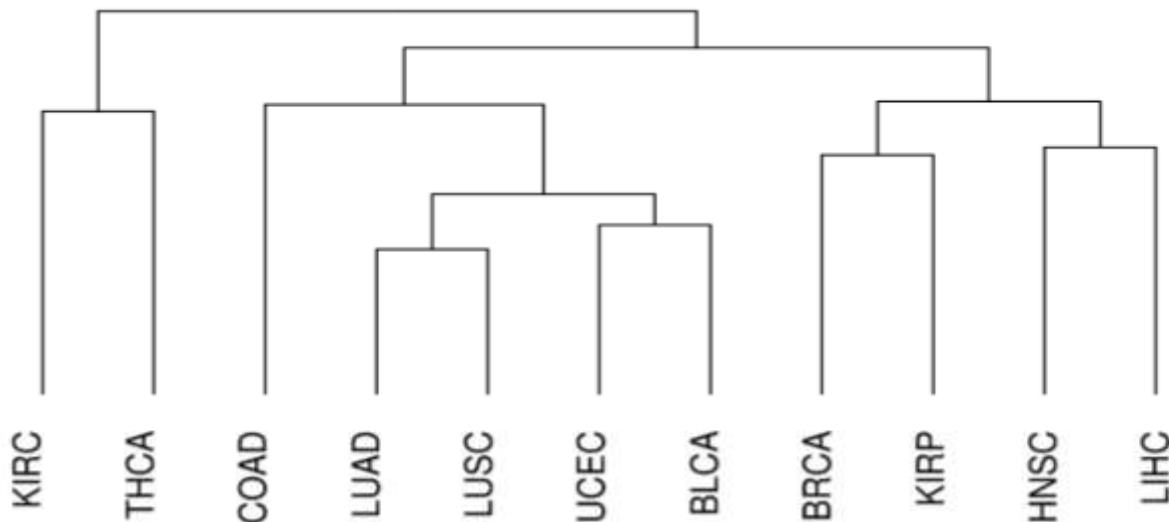
Survival analysis is a set of methods for analyzing data where the outcome variable is the time until the occurrence of an event of interest, in this case, disease.



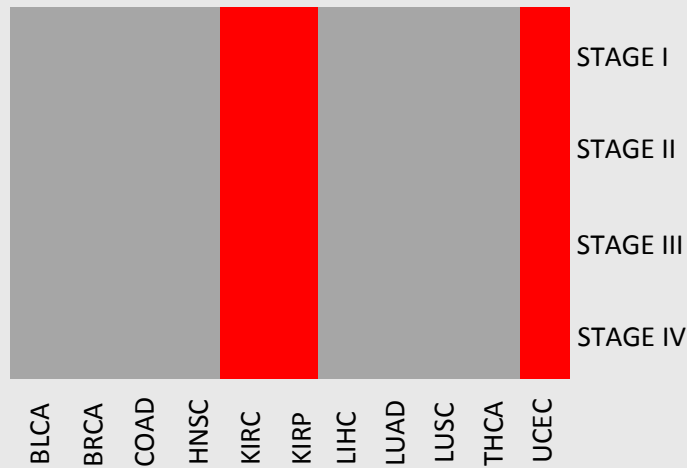
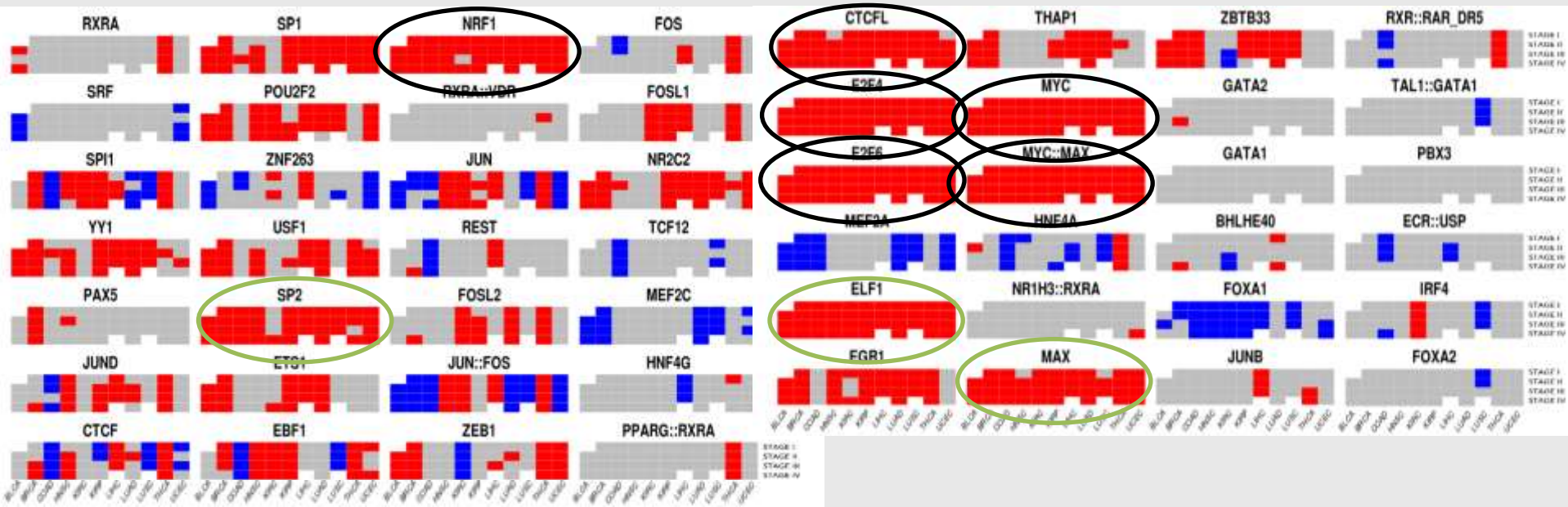
RESULTS AND DISCUSSION

NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	E2F6
High	High	Medium	Medium	Medium	Medium	Medium	Medium	High	NA	High	E2F4
Medium	Medium	Low	High	High	Low	Medium	Medium	Medium	NA	Medium	MYC
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	MYC::MAX
Medium	Medium	Medium	Low	Low	Medium	Medium	Medium	Medium	NA	ND	NRF1
Medium	High	Medium	High	High	Medium	High	High	Medium	NA	ND	MAX
ND	ND	ND	ND	ND	ND	ND	ND	ND	NA	ND	CTCF
Medium	Low	Medium	Low	Low	High	High	Medium	Medium	NA	Low	EGR1
Medium	Medium	Medium	Medium	Medium	Medium	High	Medium	Medium	NA	Low	ELF1
Medium	Low	Medium	Medium	Medium	Medium	High	Low	Medium	NA	ND	SP1
Medium	Medium	Medium	Medium	Medium	Medium	High	Medium	Medium	NA	Low	YY1
Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	NA	Medium	USF1
Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	NA	Low	SP2
High	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	NA	Low	ZBTB33
High	ND	High	ND	ND	ND	ND	High	High	NA	High	NR2C2
Medium	High	High	High	High	High	High	High	Medium	NA	Medium	HNF4A
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	RXRA
Medium	Low	Medium	Low	Low	High	High	ND	Medium	NA	Medium	RXR::RAR_DR5
Medium	Medium	Medium	Medium	Medium	High	High	Medium	Medium	NA	Low	FOS
High	Medium	High	Medium	Medium	High	High	High	High	NA	Low	REST
High	Medium	High	High	High	High	High	High	High	NA	Low	JUNB
Medium	Low	Medium	Low	Low	Low	Low	Medium	Medium	NA	Medium	CTCF
Low	ND	ND	Low	Low	ND	Low	ND	Low	NA	ND	ZNF263
High	Low	High	Low	Low	Low	Low	Low	Low	NA	ND	FOSL1
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	FOSL2
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	EBF1
ND	ND	ND	ND	ND	ND	ND	ND	ND	NA	NA	THAP1
ND	ND	ND	ND	ND	ND	ND	ND	ND	NA	ND	ETS1
High	Medium	High	High	High	High	High	High	High	NA	Medium	PAX5
Medium	Low	Medium	High	High	High	High	High	High	NA	Medium	TCF12
Low	ND	Medium	Low	Low	Medium	Medium	ND	Medium	NA	ND	MEF2A
High	Medium	High	Medium	Low	Low	Low	ND	Low	NA	ND	MEF2C
NA	NA	NA	NA	NA	NA	NA	NA	High	NA	Low	SRF
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	RYR2::ROR

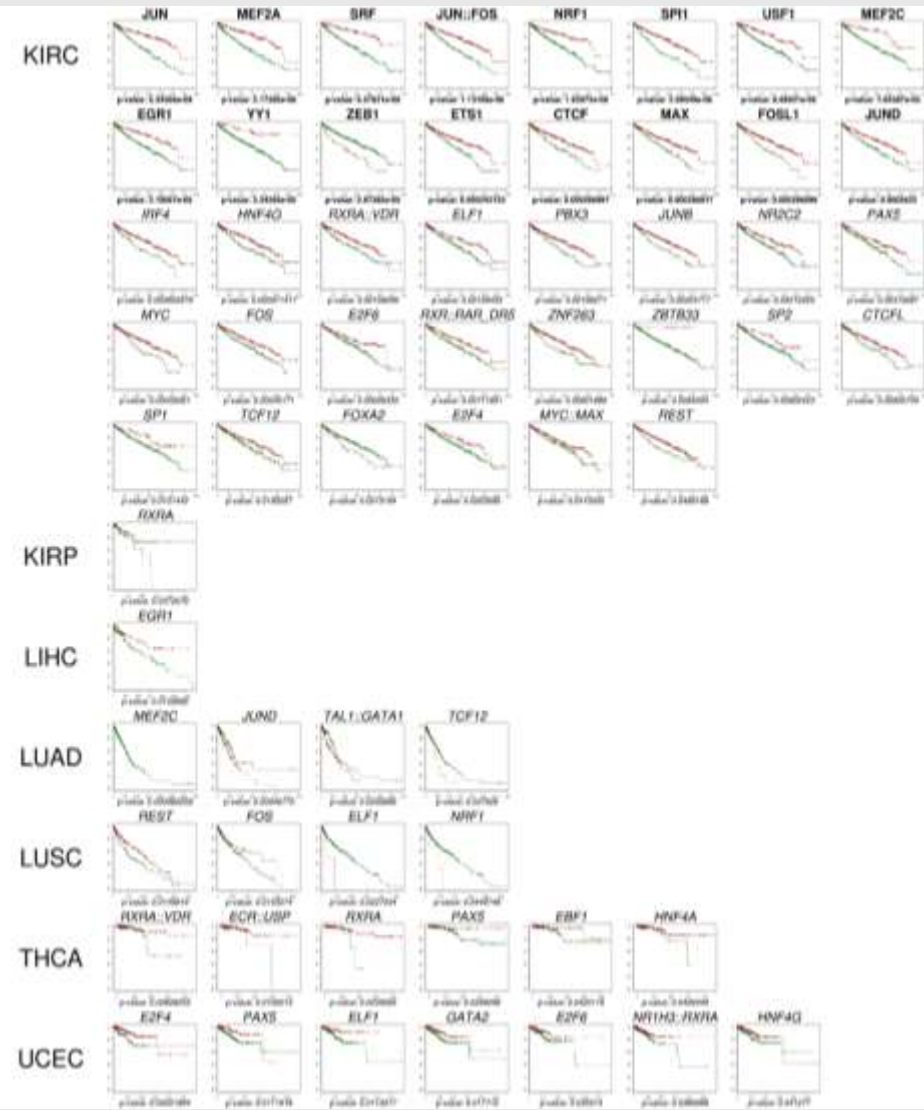
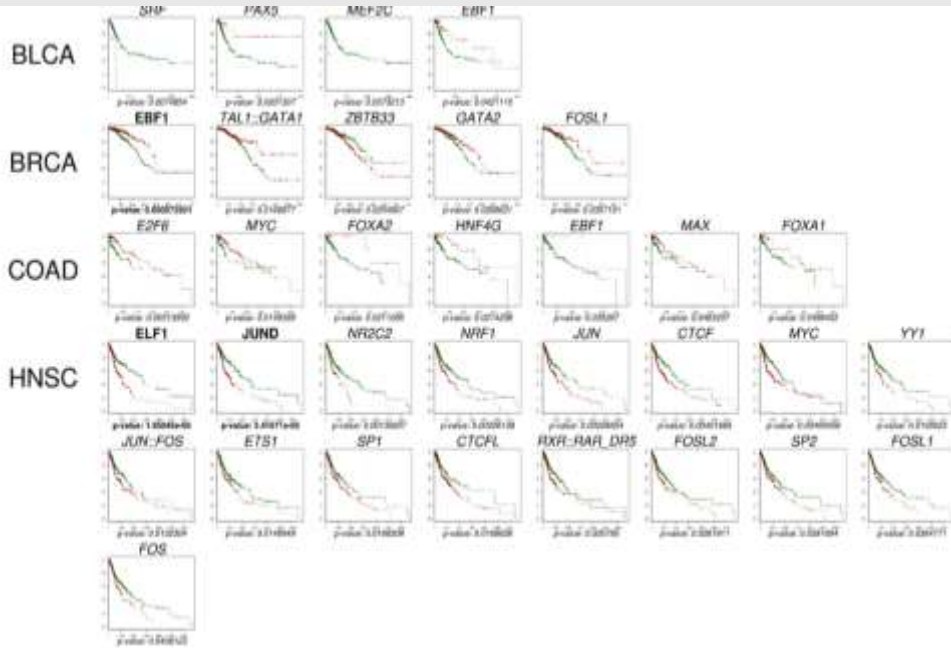
E2F6, MYC, E2F4,
MYC::MAX



Tumor stage analysis



Survival analysis



CONCLUSIONS

The availability of relevant clinical data makes the analysis of pan-cancer Big Data repositories especially compelling. Nevertheless, more data is needed to extract more conclusions

This analysis was made under some data limitations. However, we have been able to create a comprehensive description of the alterations of some transcription factors in all cancers analyzed

Any of the analyzed transcription factors could not be associated to a certain stage. In fact, the alteration of a transcription factor usually occurs in all stages

The alterations of several transcription factors were related with patient survival

This approaches may stablish the basis for the election of prognostic markers. The determination of alteration patterns could also constitute a tool for diagnostic and personalized treatment of patient



PRINCIPE FELIPE

CENTRO DE INVESTIGACION

Computational Genomics Laboratory



**Any
question???**

**Matías Marín Falco
Joaquín Dopaz
José Carbonell Caballero**