



---

# Causal Protein - Signaling Networks Derived from Multiparameter Single-Cell Data

## Redes Bayesianas

**Profesor:** Dr. Max Chacón Pacheco

**Ayudante:** Mg. Felipe Bello Robles

**Alumnos**

*Ignacio Ibáñez Aliaga*

*Joaquín Villagra Pacheco*

---

# Tabla de contenido



## 1. Contextualización

- a. Estudio previo basal
- b. Generación del conjunto de datos
- c. Definición de moléculas

## 2. Estudio realizado

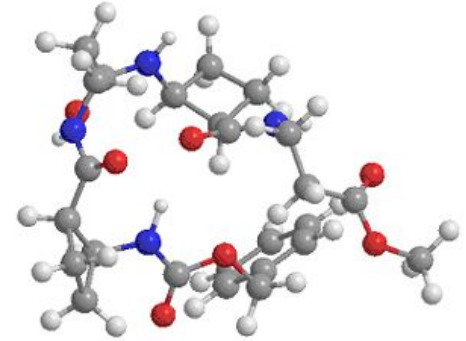
- a. Método
- b. Datos entregados
- c. Red Bayesiana obtenida por aplicación del método estudiado
- d. Resultados
- e. Primera iteración de mejora del método
- f. Segunda iteración de mejora del método

## 3. ¿Aspectos a mejorar o investigar?

# Estudio previo basal



- ❑ Se basó en la medición de componentes de proteínas y fosfolípidos múltiples en miles de células del sistema inmunológico humano primario.
- ❑ Perturbando estas células con intervenciones moleculares condujo a ordenar las conexiones entre componentes de la vía, en los que los métodos computacionales de la red bayesiana dilucidaron la mayoría de las relaciones de señalización tradicionalmente reportadas y prediciendo la red de causalidades, que fueron verificadas experimentalmente.



# Generación del conjunto de datos



- ❑ A partir del análisis de las perturbaciones, se generó un modelo genético-proteico.
- ❑ Con el modelo anterior se generan 10.000 ejemplos que componen el conjunto de datos estudiado [Conjunto disponibilizado por el equipo de docentes del curso].
- ❑ Se miden 11 moléculas en base al estudio literario de la influencia de una molécula proteica sobre otra.

# Definición de moléculas

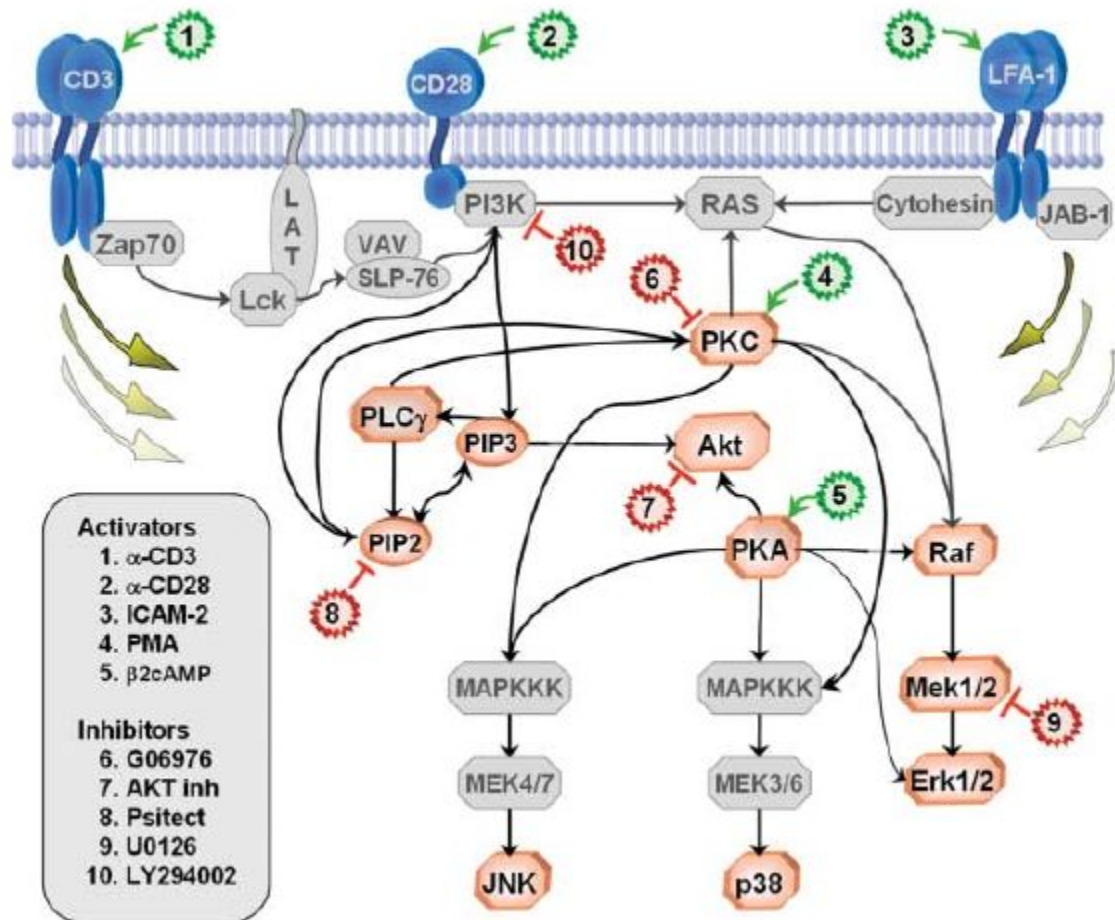


Molécula medida	Especificidad del anticuerpo
Raf	Fosforilación en S259
Erk1 and Erk2	Fosforilación en T202 y Y204
p38	Fosforilación en T180 y Y182
Jnk	Fosforilación en T183 y Y185
Akt	Fosforilación en S473
Mek1 y Mek2	Fosforilación en S217 y S221

Molécula medida	Especificidad del anticuerpo
PKA	Detecta proteínas y péptidos que contienen un residuo phospho-Ser / Thr
PKC	Detecta fosforilación PKCx
PLCg	Fosforilación en Y783
PIP2	Detecta PIP2
PIP3	Detecta PIP3

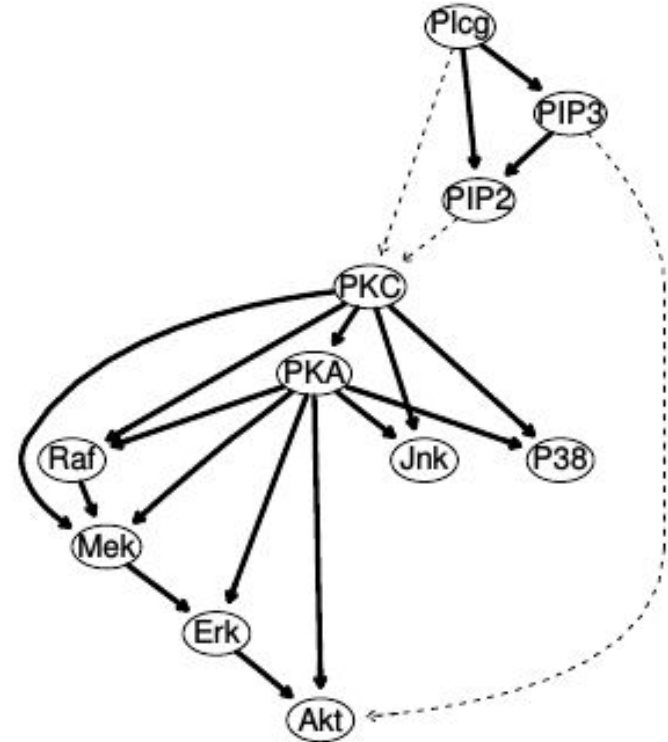


Connection	Influence path	Citation
PKC→Raf	PKC→Ras→RafS259	(12-16)
PKC→Mek	PKC→RafS497/S499→Mek	(17)
PKC→Jnk	PKC→→MKKs→Jnk	(18)
PKC→p38	PKC→→MKKs→p38	(18)
PKC→PKA	PKC→cAMP→PKA	(19)
PKA→Raf	PKA→RafS259	(20)
PKA→Mek	PKA→RafS621→Mek	(21)
PKA→Erk	Unknown	
PKA→Jnk	PKA→→MKKs→Jnk	(22)
PKA→p38	PKA→→MKKs→p38	(23)
Raf→Mek	direct phosphorylation	(24-26)
PKA→Akt	PKA→CaMKK→AktT308→AktS473	(27-29)
Mek→Erk	direct phosphorylation	(30)
Plcγ→PIP2	direct hydrolysis to IP3	(31, 32)
Plcγ→PIP3	recruitment leading to phosphorylation (reversed edge)	(33)
PIP3→PIP2	precursor-product	(33)
Erk→Akt	direct or indirect	(34)



# Método

El proceso que realizan es inicializar con diferentes grafos aleatorios y repetir el proceso 500 veces, asignando una mejor puntuación a aquellos modelos que están más cerca de la distribución empírica de los datos, con aquellos modelos de mejor puntuación realizar el promedio.





# Datos entregados



Corresponden a 10000 simulaciones de la interacción entre las proteínas por medio del modelo presentado antes, de forma discretizada.

- 1 = bajo
- 2 = medio
- 3 = alto

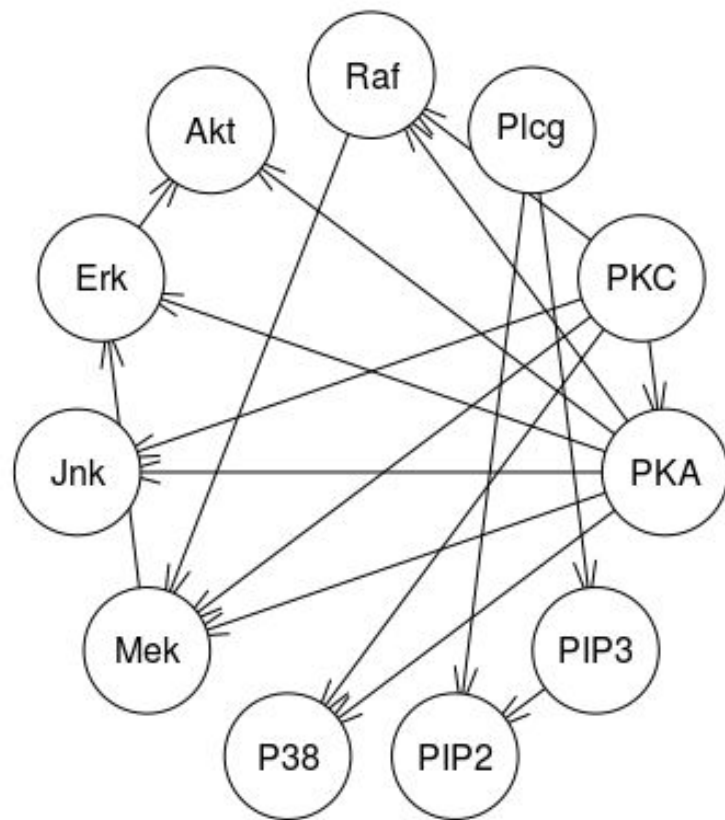


# Método y Resultados obtenidos

- Tp: Número de arcos idénticos en el modelo original y el generado.
- Fp: Número de arcos que no están presentes en el modelo original, pero si en el modelo generado.
- Fn: Número de arcos que están presentes en el modelo original y en el modelo generado pero con cambios de sentido.

	tp	fp	fn	BIC
HC	5	13	12	-72605,79
mmhc	6	7	11	-75357,33
mmpc	0	13	17	error

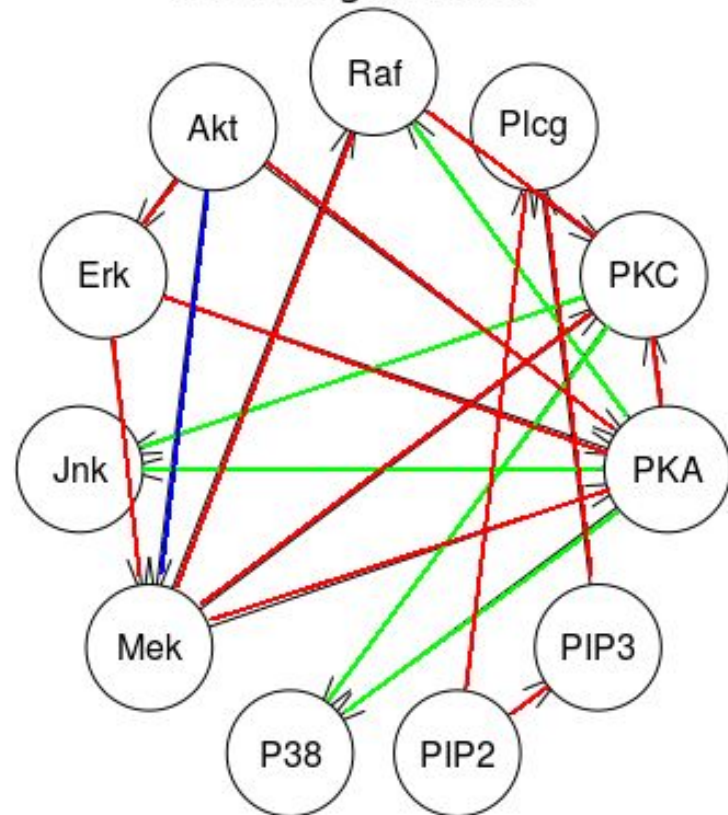
## Modelo original



TP = 5

FP = 13

### Modelo generado



FN = 12



# Primera mejora



## Cambio de columnas

- Akt,Erk,Jnk,Mek,P38,PIP2,PIP3,PKA,PKC,Plcg,Raf
- Raf,Mek,Plcg,PIP2,PIP3,Erk,Akt,PKA,PKC,P38,Jnk

	tp	fp	fn	BIC
HC	11	6	6	-72448.71
mmhc	9	6	8	-75009.38
mmpc	0	15	17	error



# Segunda mejora

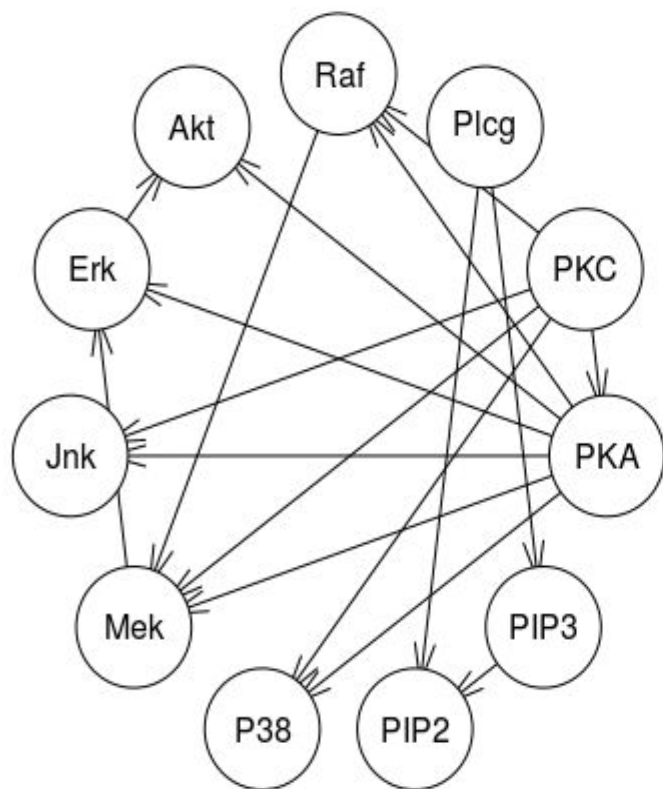


Considerando que el método que se arroja los mejores resultados es HC, se continuó el estudio con su uso variando el valor el **set.seed()** quedando definido en 6 y variando el parámetro **restart=** quedando definido en 35.

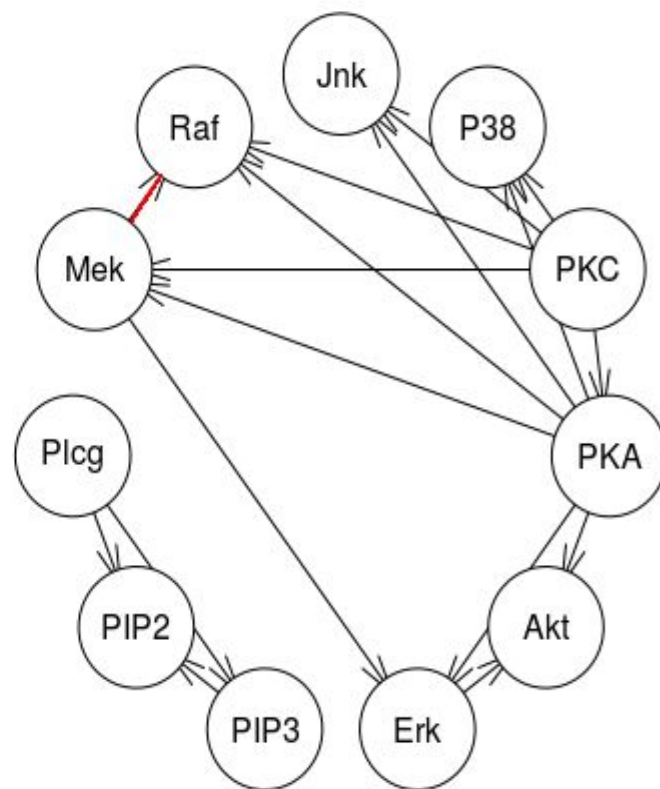
Set.seed(6)				
hc(data , restart = 35)				
	tp	fp	fn	BIC
HC	16	1	1	-72448,71

Modelo original (BIC)
-72448.71

Modelo original



Modelo generado



TP = 16  
FP = 1

FN = 1



# Rendimiento del modelo



Raf	Mek	Plcg	PIP2	PIP3	Erk
71.16%	63.2%	81.4%	88.22%	47.24%	72.24%

Akt	PKA	PKC	P38	Jnk
79.98%	69.6%	49.64%	80%	65.92%

Accuracy
69.86%





## Modelo Original

Parameters of node Plcg (multinomial distribution)

Conditional probability table:

	1	2	3
	0.8154	0.0820	0.1026

## Modelo Generado

Parameters of node Plcg (multinomial distribution)

Conditional probability table:

	1	2	3
	0.8154	0.0820	0.1026

Conditional probability table:

, , PKC = 1, Raf = 1

Modelo Original

PKA			
Mek	1	2	3
1	0.769911504	0.760928962	0.996511628
2	0.230088496	0.239071038	0.003488372
3	0.000000000	0.000000000	0.000000000

, , PKC = 2, Raf = 1

PKA			
Mek	1	2	3
1	0.719696970	0.726674887	0.978723404
2	0.280303030	0.273325113	0.021276596
3	0.000000000	0.000000000	0.000000000

, , PKC = 3, Raf = 1

PKA			
Mek	1	2	3
1	0.894736842	0.803526448	0.551724138
2	0.105263158	0.196473552	0.448275862
3	0.000000000	0.000000000	0.000000000

, , PKC = 1, Raf = 2

PKA			
Mek	1	2	3
1	0.397683398	0.353909465	1.000000000
2	0.119691120	0.635802469	0.000000000
3	0.482625483	0.010288066	0.000000000

Conditional probability table:

, , PKC = 1

Modelo Generado

PKA			
Mek	1	2	3
1	0.329177057	0.668313774	0.997029703
2	0.037406484	0.323038913	0.002970297
3	0.633416459	0.008647313	0.000000000

, , PKC = 2

PKA			
Mek	1	2	3
1	0.617647059	0.518110236	0.893939394
2	0.356209150	0.473565804	0.106060606
3	0.026143791	0.008323960	0.000000000

, , PKC = 3

PKA			
Mek	1	2	3
1	0.894736842	0.720000000	0.516129032
2	0.105263158	0.280000000	0.483870968
3	0.000000000	0.000000000	0.000000000



, , PKC = 2, Raf = 2

## Modelo Original

PKA			
Mek	1	2	3
1	0.283950617	0.261142857	1.000000000
2	0.716049383	0.732000000	0.000000000
3	0.000000000	0.006857143	0.000000000

, , PKC = 3, Raf = 2

PKA			
Mek	1	2	3
1	0.094339623	0.000000000	
2	0.905660377	1.000000000	
3	0.000000000	0.000000000	

, , PKC = 1, Raf = 3

PKA			
Mek	1	2	3
1	0.274350649	0.880299252	1.000000000
2	0.002435065	0.097256858	0.000000000
3	0.723214286	0.022443890	0.000000000

, , PKC = 2, Raf = 3

PKA			
Mek	1	2	3
1	0.763440860	0.297709924	0.000000000
2	0.150537634	0.606870229	1.000000000
3	0.086021505	0.095419847	0.000000000

, , PKC = 3, Raf = 3

PKA			
Mek	1	2	3
1			
2			
3			



# Método y Resultados obtenidos



```
> cpquery(fittedbn, event = (Jnk == "LOW"),  
evidence = (PKA=="HIGH" & PKC=="HIGH"))
```

```
[1] 0.2258065
```

```
> cpquery(fittedbn, event = (Jnk == "AVG"),  
evidence = (PKA=="HIGH" & PKC=="HIGH"))
```

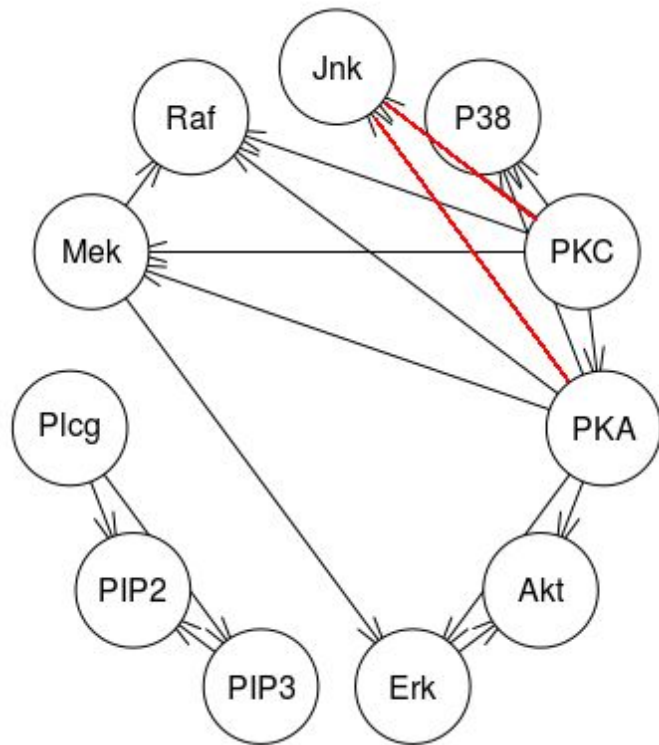
```
[1] 0.7741935
```

```
> cpquery(fittedbn, event = (PKA=="HIGH" &  
PKC=="HIGH"), evidence = (Jnk == "LOW"))
```

```
[1] 0.001278072
```

```
> cpquery(fittedbn, event = (PKA=="HIGH" &  
PKC=="HIGH"), evidence = (Jnk == "AVG"))
```

```
[1] 0.006329114
```



# Método y Resultados obtenidos

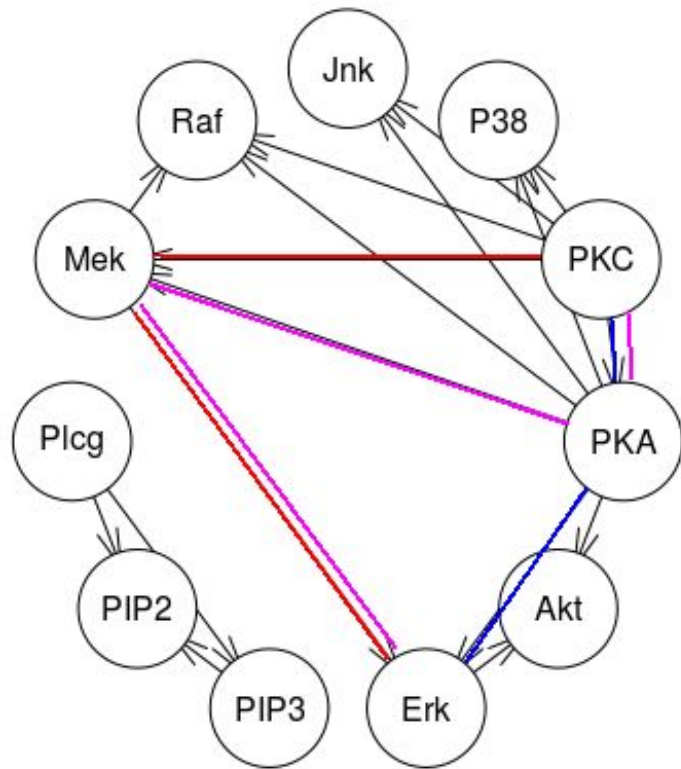


```
>cpquery(fittedbn, event = (Erk=="LOW"), evidence =  
(PKC == "HIGH"))  
[1] 0.0907173
```

```
> cpquery(fittedbn, event = (Erk=="HIGH"), evidence =  
(PKC == "HIGH"))  
[1] 0.228903
```

```
> cpquery(fittedbn, event = (Erk=="AVG"), evidence =  
(PKC == "HIGH"))  
[1] 0.6803797
```

```
> cpquery(fittedbn, event = (PKC == "LOW"), evidence =  
((Erk == "AVG") | (Erk == "LOW") | (Erk == "HIGH")))  
[1] 0.424
```



# Aspectos a mejorar o investigar



- ❑ Dado los problemas y la complejidad del método aplicado sobre este conjunto de datos tan particular, ¿utilizamos el modelo generado por nuestro estudio o el original del Estudio basal señalado ?
- ❑ ¿Qué otro aspecto se considera relevante y no hemos tomado en cuenta?



# Muchas Gracias!

