

2684-1

# Identificar els gens que promouen els canvis fenotípics

Realitzat per  
Carles Hernández

Dirigit per  
Jordi González  
(CVC)

Codirigit per  
Mario Huerta  
(IBB)








Universitat  
Autònoma  
de Barcelona



escola  
d'enginyeria



# Índex

-  Introducció
-  Estat de l'art
-  Objectius
-  Desenvolupament
-  Conclusions

# Introducció

- ✿ L'Institut de Biotecnologia i de Biomedicina de la UAB té l'objectiu de promoure i facilitar la investigació biològica de caràcter bàsic i fonamental.
- ✿ 7 àrees d'expertesa: Bionformàtica, Biologia cel·lular, Biologia estructural, Genòmica, Immunologia, Microbiologia i Proteòmica.
- ✿ Dins d'un dels grups centrats en bioinformàtica (Bioinformàtica de biologia molecular), hi ha una línia d'investigació dedicada a l'anàlisi de microarrays.

<http://revolutionresearch.uab.es>

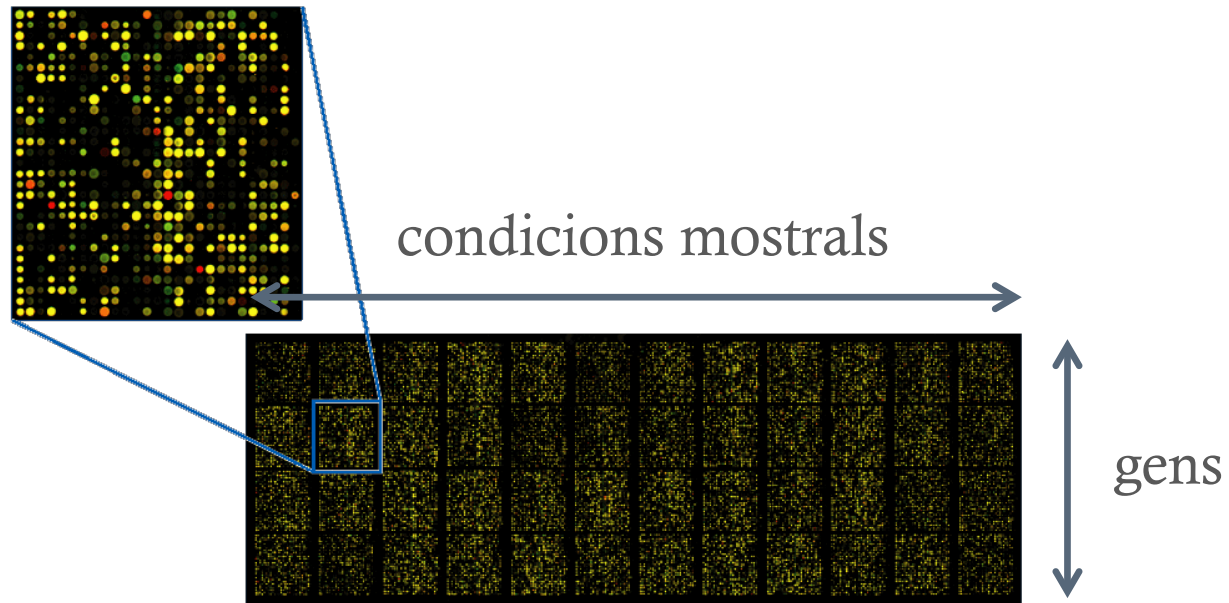
<http://devresearch.uab.es>

<http://wallace.uab.es>



# Estat de l'art

Microarray: Matriu d'expressió gènica



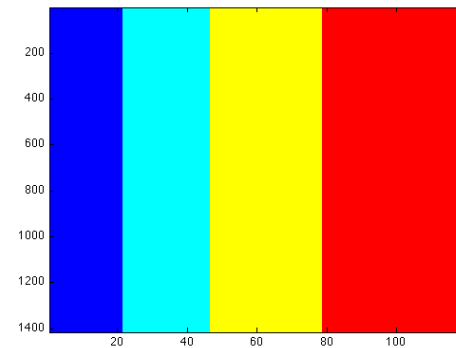
# Estat de l'art

## Clustering

- 📈 Hierarchical Cl.
- 📈 Point Accepted Mutation
- 📈 Self-Organizing Map (SOM)
- 📈 Self-Organizing Tree Alg.

1416 gens

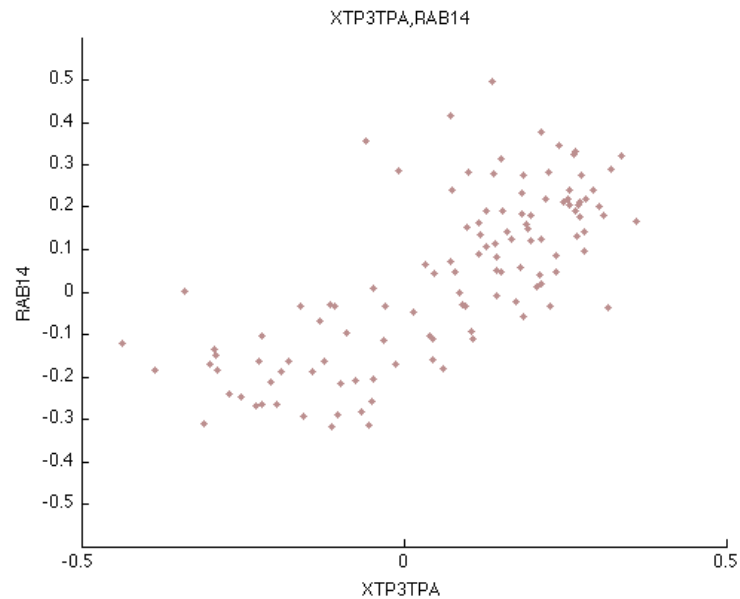
118 condicions mostrals



Microarray (17)  
SOM(k=4)

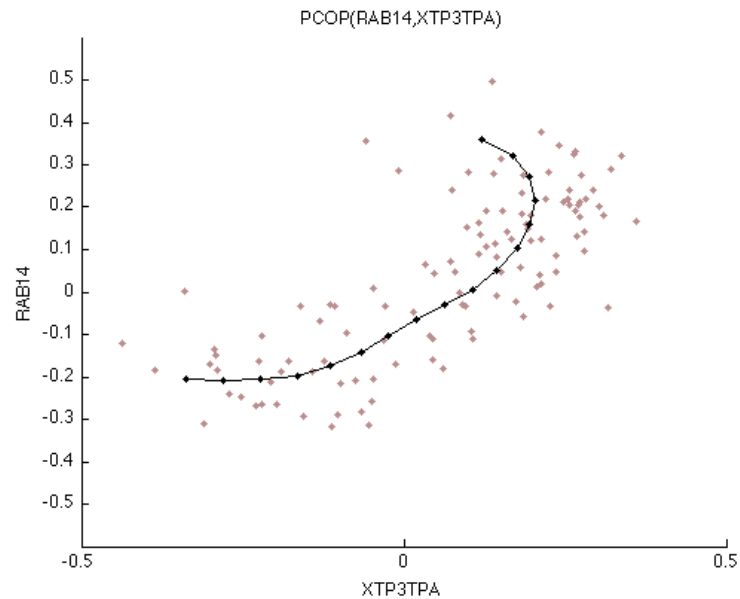
# Estat de l'art

Principal Curves of Oriented Points [PCOP]: Relació d'expressió no lineal entre 2 gens



# Estat de l'art

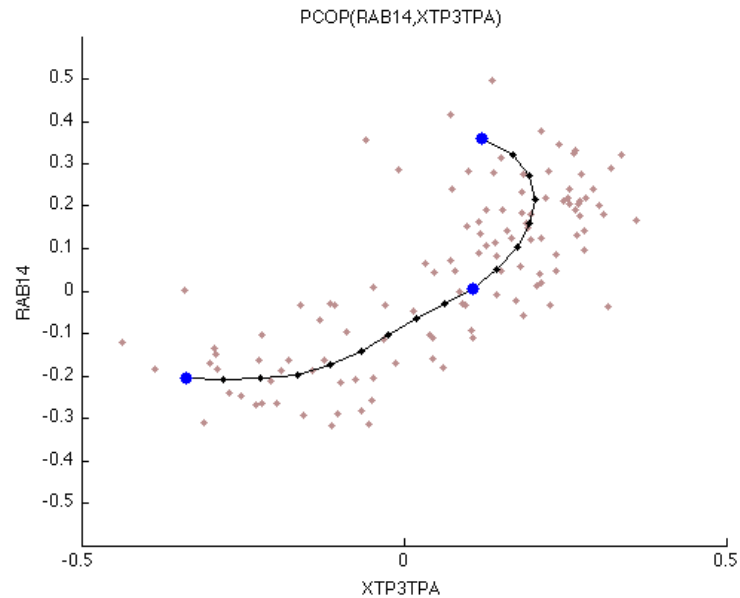
Principal Curves of Oriented Points [PCOP]: Relació d'expressió no lineal entre 2 gens





# Estat de l'art

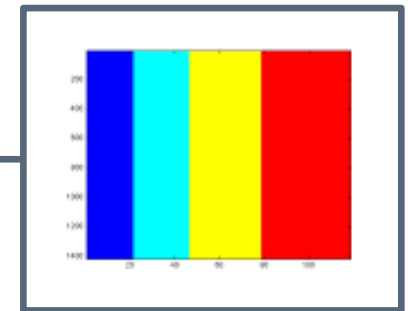
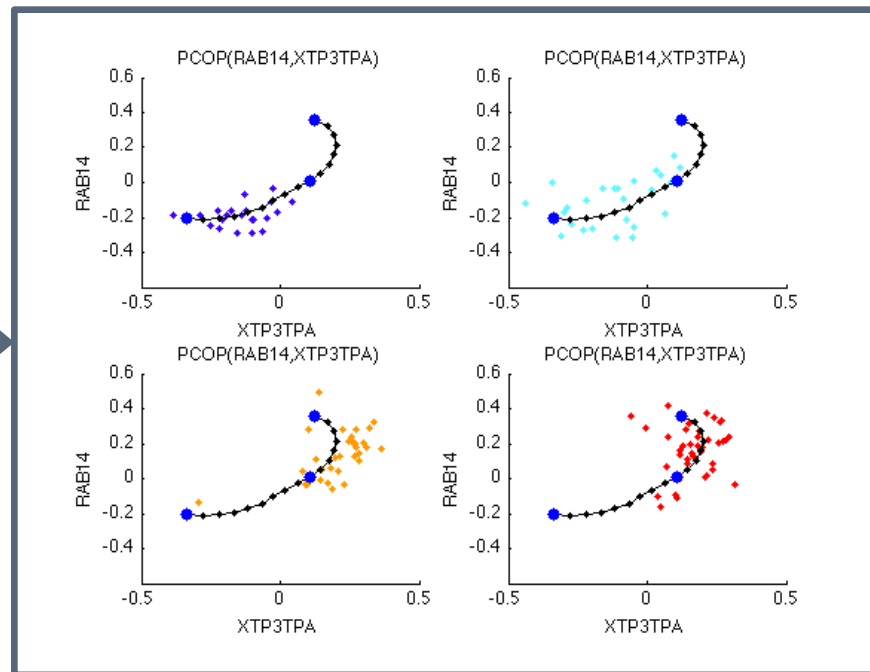
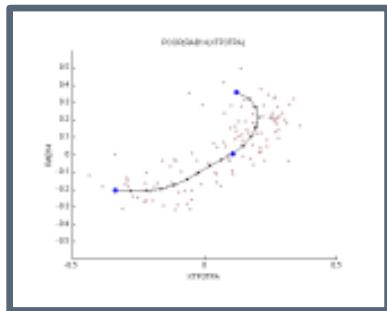
Principal Curves of Oriented Points [PCOP]: Relació d'expressió no lineal entre 2 gens





# Objectius

Aplicar sobre les PCOP les distribucions de clusters i determinar si són capaces de diferenciar entre clusters (fenotips).



# Objectius

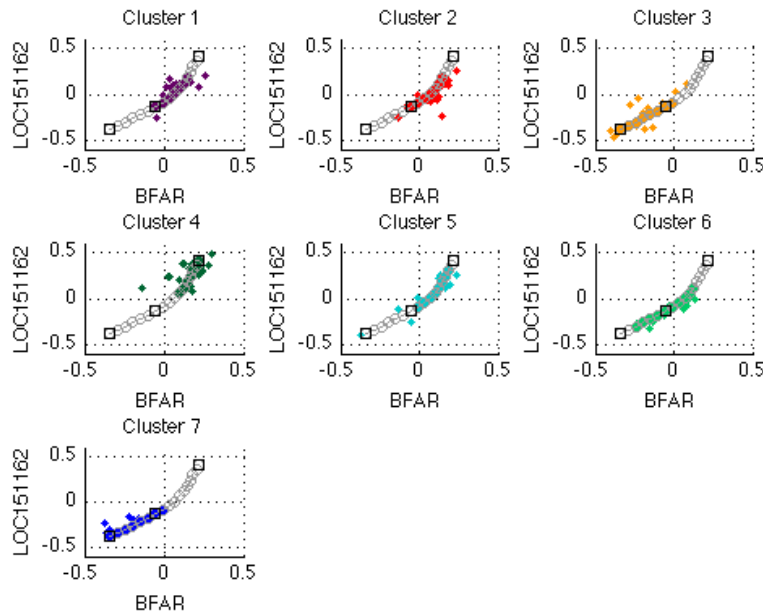
- ❧ Procés de creuament de dades entre:
  - ❧ Distribucions de clusters obtingudes de l'aplicació dels mètodes de clustering sobre la microarray.
  - ❧ Les PCOPs obtingudes de l'anàlisi de les relacions d'expressió entre els gens de la microarray.
  - ❧ Filtrar les PCOPs sobre les que s'aplica una distribució de clusters per descartar aquelles que no ens són útils.
  - ❧ Agrupar jeràrquicament les PCOPs segons la similitud de la transició entre fenotips que descriuen.

# Objectius

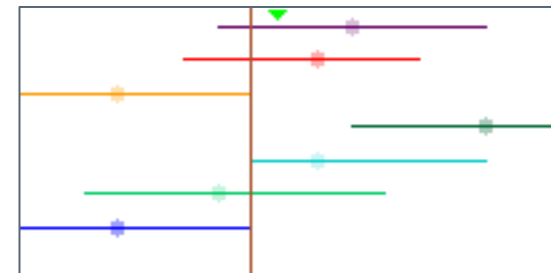
- ✧ Generació d'un gràfic per a la representació d'una PCOP sobre la qual s'aplica una distribució de clusters.
- ✧ Aplicació web per a l'anàlisi dels resultats obtinguts:
  - ✧ Amb una interfície general que permeti comparar de manera visual les distribucions de clusters.
  - ✧ Una interfície de detall amb el llistat de PCOPs obtingut i ordenat jeràrquicament segons la similitud de la transició entre fenotips que descriuen.

# Desenvolupament

Representació gràfica d'una PCOP sobre la qual s'aplica una distribució de clusters.



PCOP(BFAR,LOC151162)



Mds Pam Descarte (k=7)

# Desenvolupament

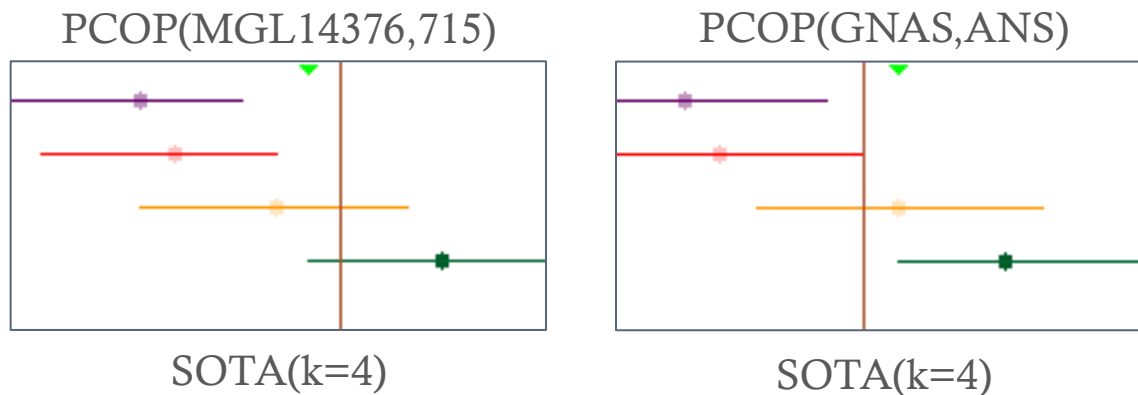
## Creuament de dades

- Relacionar els nivells d'expressió dels gens que formen la PCOP amb els fenotips (clusters) d'una distribució de clusters.
- En diem aplicar una distribució de clusters sobre una PCOP.
- Un total de 475 distribucions de clusters.
- PCOPs classificades segons correlació i curvatura (en total més de 10.000 PCOPs).

# Desenvolupament

Creuament de dades: Filtratge

🐉 Filtrar les PCOPs per número de clusters vàlids

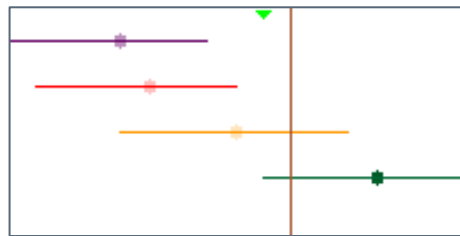


🐉 Filtrar les PCOPs per clusters en intervals

# Desenvolupament

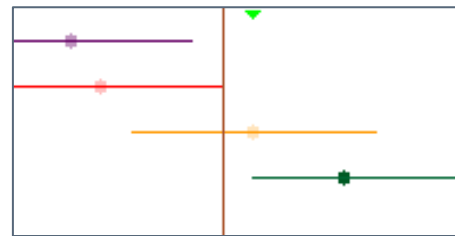
## Creuament de dades: Agrupació jeràrquica

PCOP(MGL14376,715)



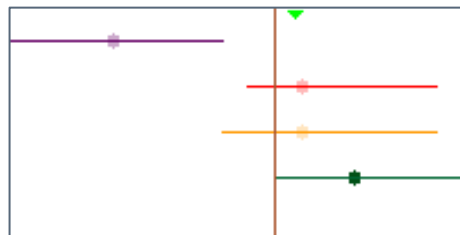
SOTA(k=4)

PCOP(GNAS,ANS)



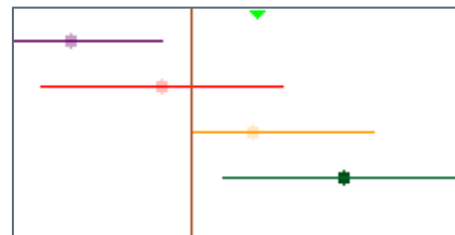
SOTA(k=4)

PCOP(1,HNRPLL)



SOTA(k=4)

PCOP(IFITM2,SLC36A4)



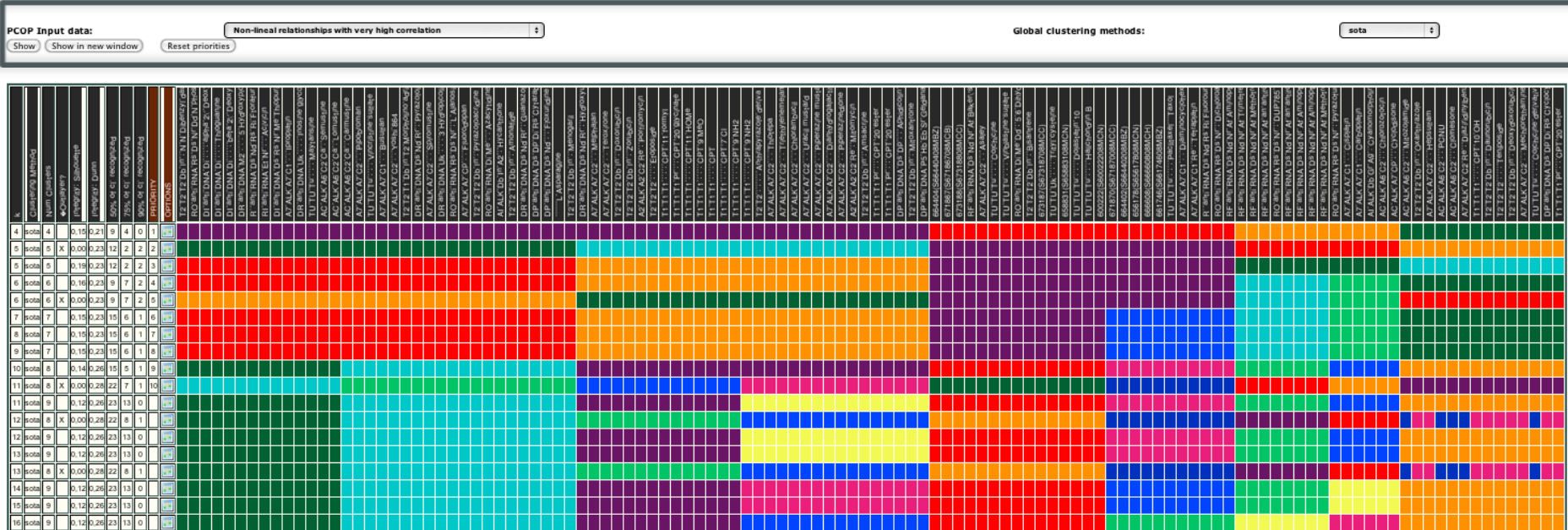
SOTA(k=4)





# Desenvolupament

## CrossingClusters (aplicació web) [vista general - SOTA]

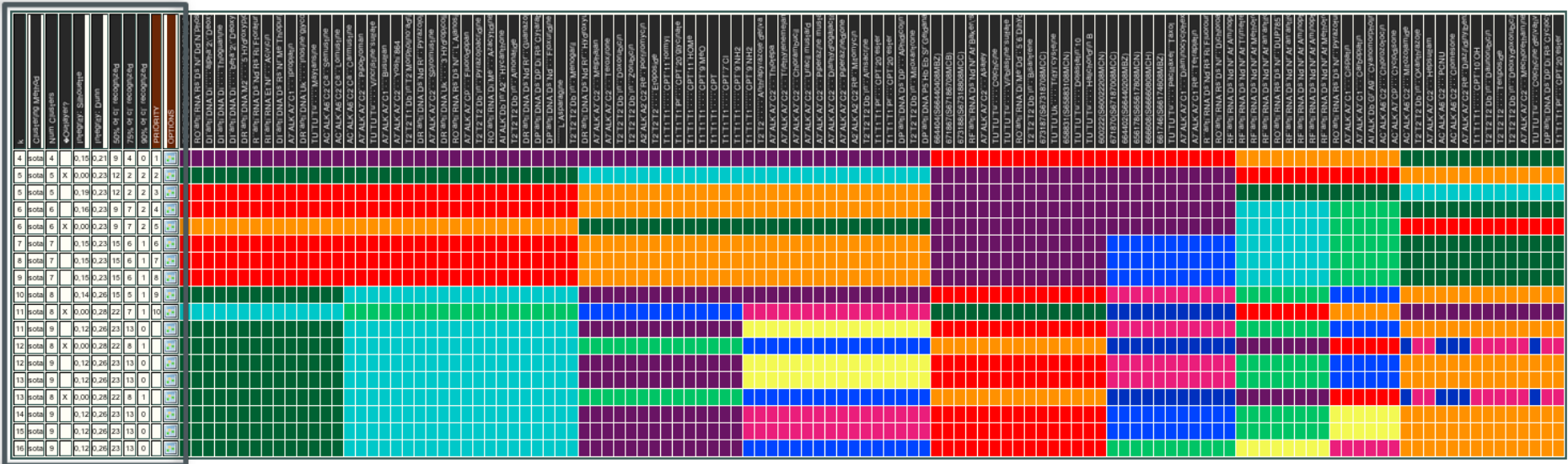


# Desenvolupament

## CrossingClusters (aplicació web) [vista general - SOTA]

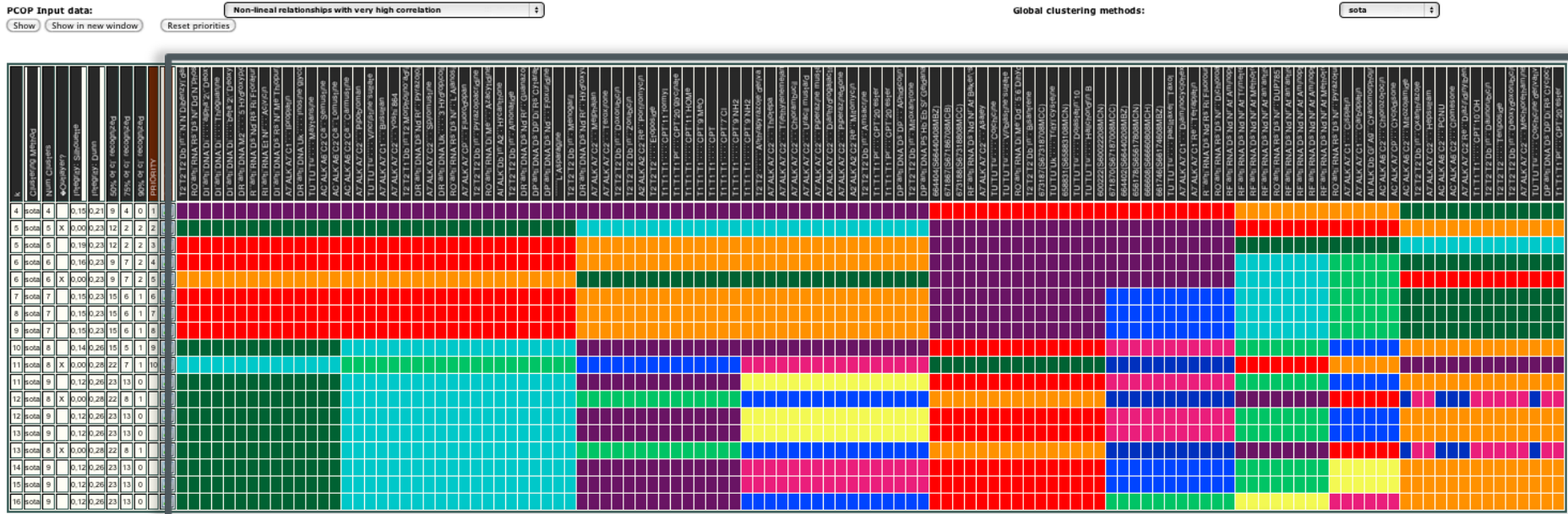
PCOP Input data: Non-linear relationships with very high correlation Reset priorities

Global clustering methods: sota



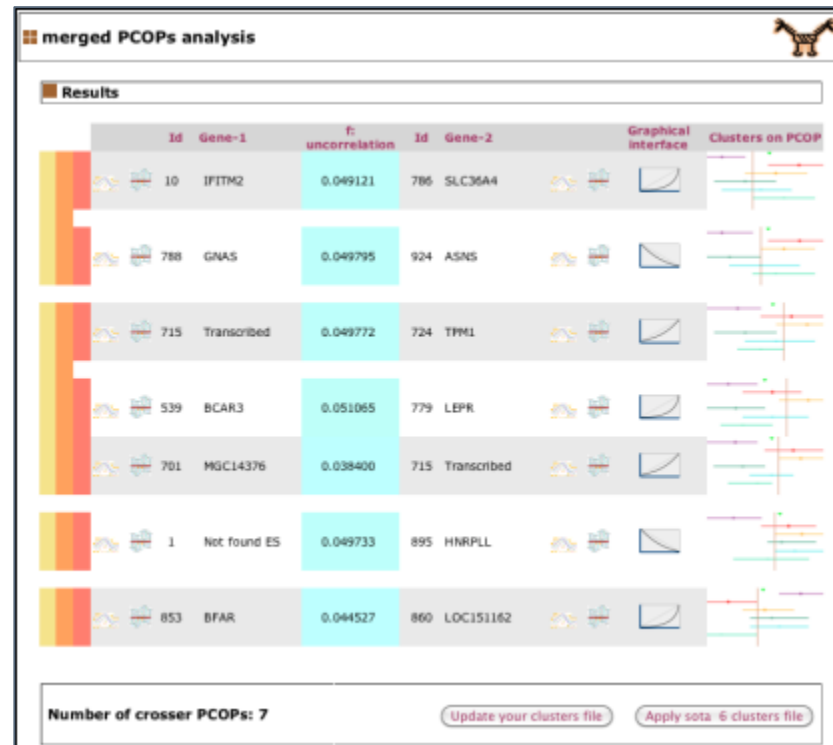
# Desenvolupament

## CrossingClusters (aplicació web) [vista general - SOTA]



# Desenvolupament

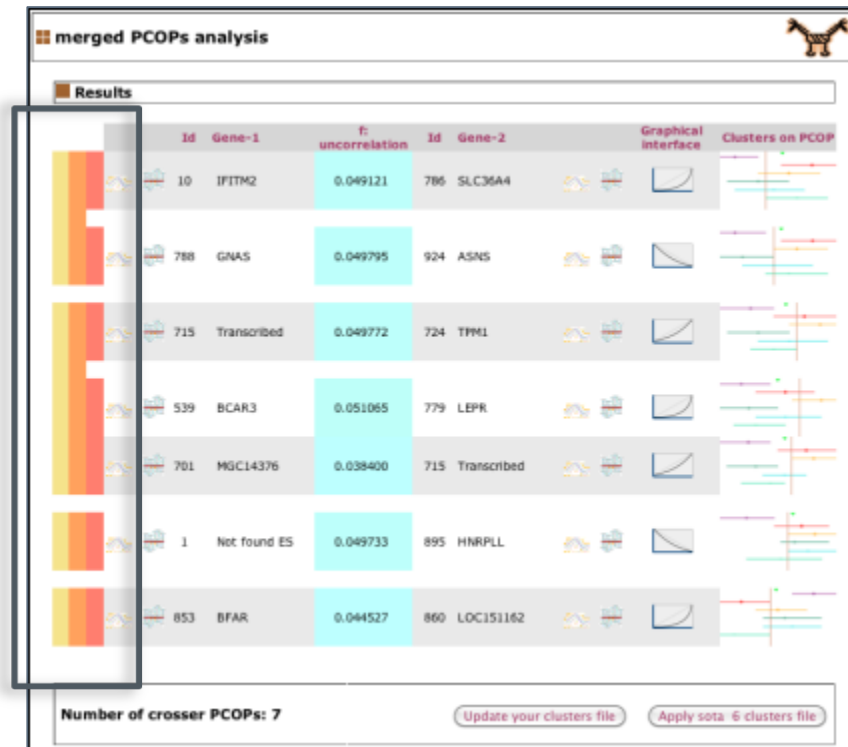
CrossingClusters (aplicació web) [vista detall - SOTA (k=6) 65%]





# Desenvolupament

CrossingClusters (aplicació web) [vista detall - SOTA (k=6) 65%]



# Desenvolupament

CrossingClusters (aplicació web) [vista detall - SOTA (k=6) 65%]

**merged PCOPs analysis**

**Results**

Id	Gene-1	r: uncorrelation	Id	Gene-2	Graphical interface	Clusters on PCOP
10	IFITM2	0.049121	786	SLC35A4		
788	GNAS	0.049795	924	ASNS		
715	Transcribed	0.049772	724	TPM1		
539	BCAR3	0.051065	779	LEPR		
701	MGC14376	0.038400	715	Transcribed		
1	Not found ES	0.049733	895	HNRPLL		
853	BFAR	0.044527	860	LOC151162		

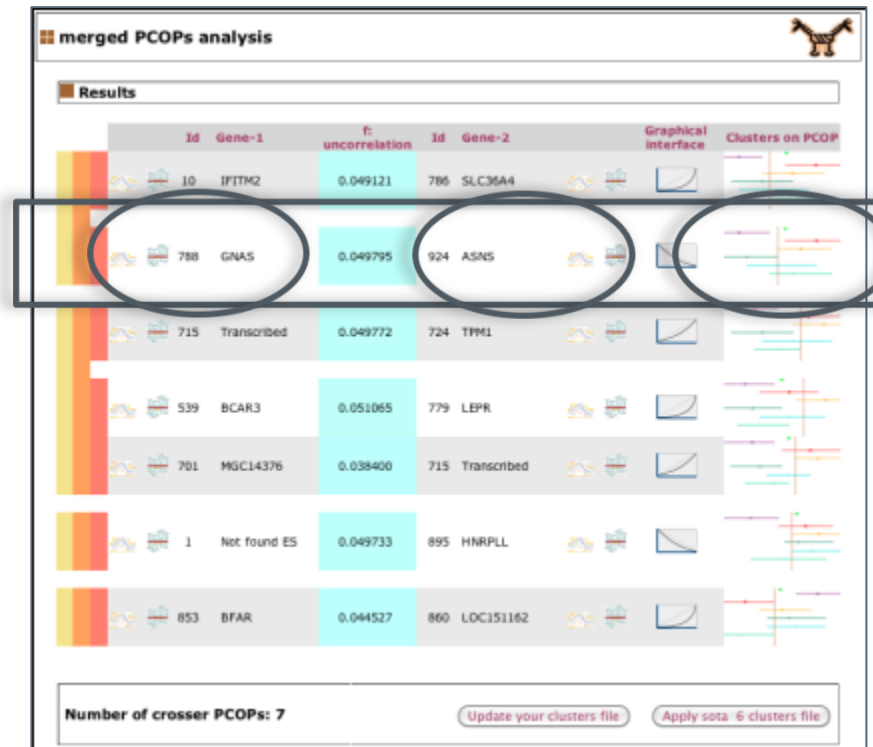
Number of crosser PCOPs: 7

[Update your clusters file](#) [Apply sota - 6 clusters file](#)



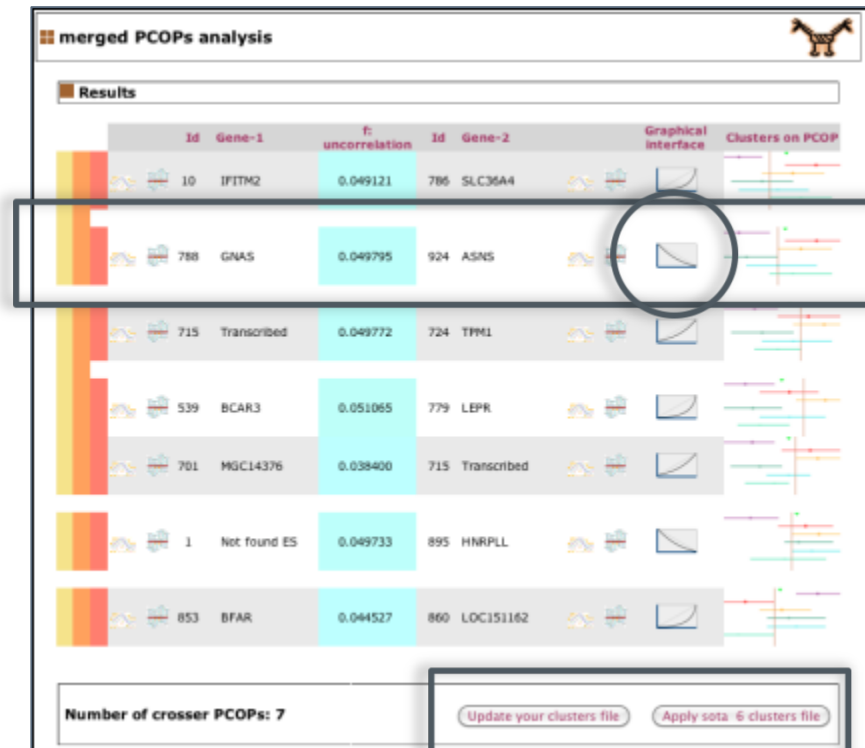
# Desenvolupament

CrossingClusters (aplicació web) [vista detall - SOTA (k=6) 65%]



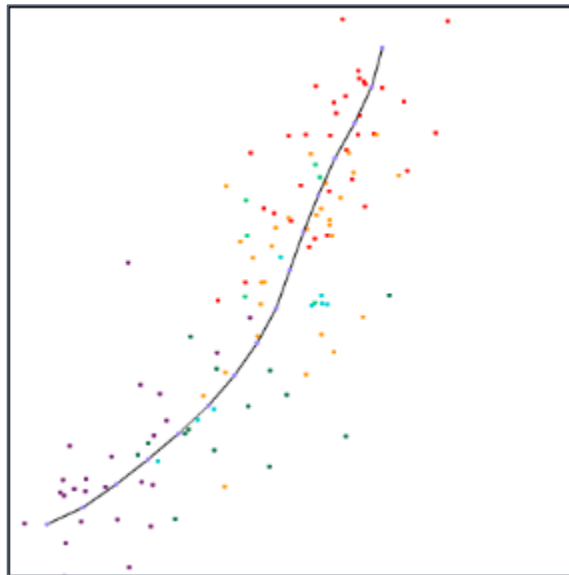
# Desenvolupament

CrossingClusters (aplicació web) [vista detall - SOTA (k=6) 65%]



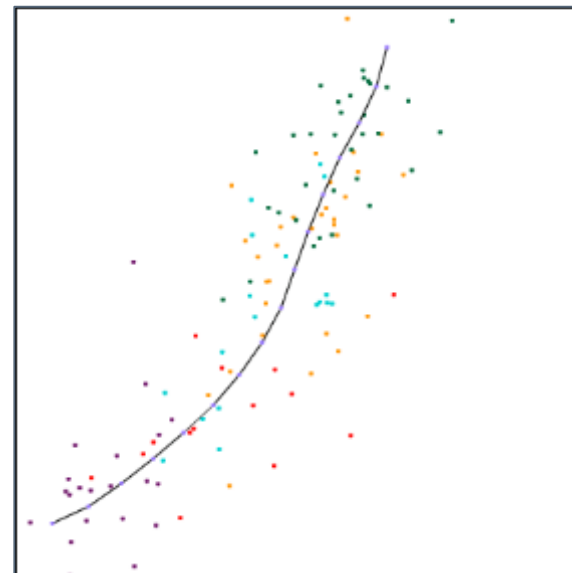
# Desenvolupament

CrossingClusters (aplicació web) [vinculació amb estudi en detall de PCOPs]  
SOTA (k=6)



PCOP(IFITM2,SLC36A4)

USER ( PAM (k=5) )



PCOP(IFITM2,SLC36A4)

# Conclusions

- ✎ S'ha generat un procés de creuament de dades entre distribucions de clusters i PCOPs que agrupa les PCOPs de forma jeràrquica segons la transició entre fenotips que porten a terme.
- ✎ S'ha creat un procés que permet la generació d'un gràfic per a la representació d'una PCOP sobre la qual s'aplica una distribució de clusters.
- ✎ S'ha construït una aplicació web per a l'anàlisi dels resultats obtinguts.

**S'ha complert amb tots els objectius**

# Bibliografía

Delicado, P.(2001) Another look at principal curves and surfaces. Journal of Multivariate Analysis, 77, 84-116.

Delicado, P. and Huerta, M. (2003): 'Principal Curves of Oriented Points: Theoretical and computational improvements'. Computational Statistics 18, 293-315.

Cedano J, Huerta M, Estrada I, Ballllosera F, Conchillo O, Delicado P, Querol E. (2007) A web server for automatic analysis and extraction of relevant biological knowledge. Comput Biol Med. 37:1672-1675.

Huerta M, Cedano J, Querol E. (2008) Analysis of nonlinear relations between expression profiles by the principal curves of oriented-points approach. J Bioinform Comput Biol. 6:367-386.

# Bibliografía

[Cedano J, Huerta M, Querol E. \(2008\) NCR-PCOPGene: An Exploratory Tool for Analysis of Sample-Classes Effect on Gene-Expression Relationships Advances in Bioinformatics, vol. 2008.](#)

[Huerta M, Cedano J, Peña D, Rodriguez A, Querol E. \(2009\) PCOPGene-Net: holistic characterisation of cellular states from microarray data base on continuous and non-continuous analysis of gene-expression relationships. BMC Bioinformatics 2009 May 9;10:138.](#)

Huerta M, Fernández-Márquez J, Cabello JL, Medrano A, Querol A, Cedano J (2011) Studying glucocorticoids' Dual Behaviour and Other Tumour-Progression Paradoxes by means of Exhaustive Analysis of Phenotypic Interdependences, Nature Oncogene [Accepted]

2684-1

# Identificar els gens que promouen els canvis fenotípics

Realitzat per  
Carles Hernández

Dirigit per  
Jordi González  
(CVC)

Codirigit per  
Mario Huerta  
(IBB)



Universitat  
Autònoma  
de Barcelona



escola  
d'enginyeria

