

# PAC2 Desenvolupament el treball - Fase 1

Vasyl Druchkiv

Estudiant de Màster de Bioestadística i Bioinformàtica

18 de Març 2019

## Índice

<b>1</b>	<b>Identificació del treball i data de l'informe</b>	<b>1</b>
<b>2</b>	<b>Descripció de l'avanç del projecte</b>	<b>1</b>
2.1	Grau de compliment dels objectius i resultats previstos en el pla de treball. . . . .	2
2.2	Justificació dels canvis en cas necessari . . . . .	3
<b>3</b>	<b>Relació de les activitats realitzades</b>	<b>3</b>
3.1	Activitats previstes en el pla de treball . . . . .	3
3.2	Activitats no previstes i realitzades o programes . . . . .	3
<b>4</b>	<b>Relació de les desviacions en la temporització i accions de mitigació si escau i actualització del cronograma si escau</b>	<b>3</b>
<b>5</b>	<b>Llistat dels resultats parcials obtinguts fins al moment (lliurables que s'adjunten)</b>	<b>3</b>
<b>6</b>	<b>Comentaris del vostre director particular si ho considereu necessari.</b>	<b>3</b>

L'estructura recomanada per realitzar l'informe de seguiment és la següent:

## Identificació del treball i data de l'informe

## Descripció de l'avanç del projecte

A la data d'avui he desenvolupat l'aplicació d'anàlisi de les rutes. L'aplicació es completament funcional localment i ofereix l'anàlisi a partir de les bases de dades GO, KEGG i Reactome. Com estava previst, l'usuari indica l'especie, puja l'arxiu amb els gens i els LogRatios provinents d'estudi de microarrays o NGS.

L'aplicació està dividida doncs en 4 parts substancials:

1. Entrada de les dades;

2. Anàlisi GO;
3. Anàlisi KEGG;
4. Anàlisi Reactome.

Figure 1: Pàgina d'entrada

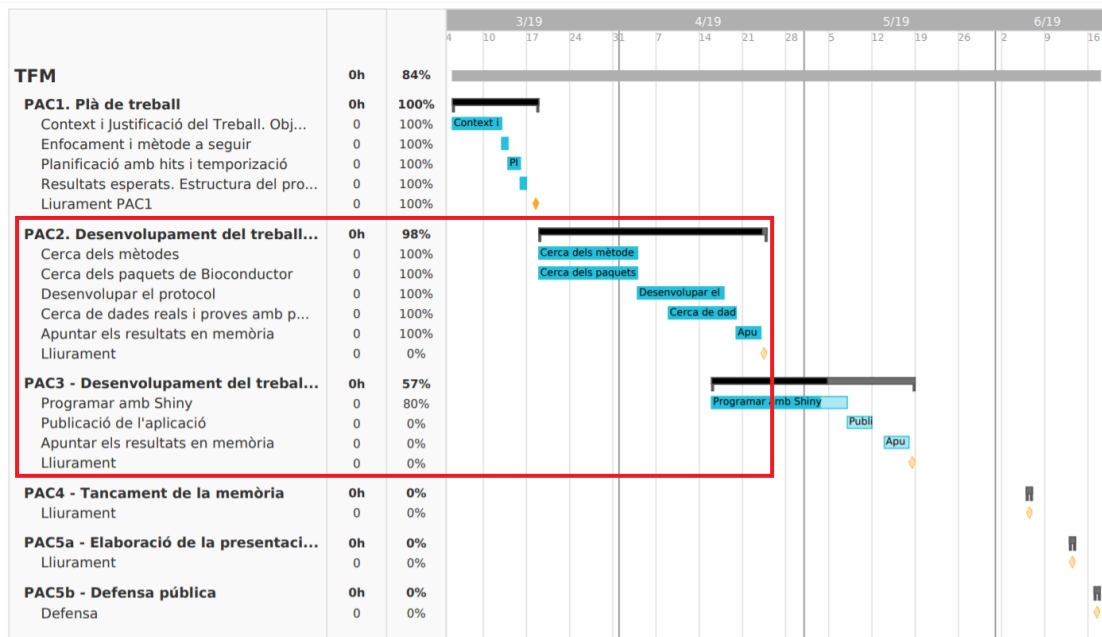
The screenshot shows the 'Pathway analysis' web application. On the left is a dark sidebar with a menu: 'Data input' (with an upload icon), 'GO Analysis', 'KEGG Analysis', and 'Reactome Analysis'. The main area has a light blue background. At the top, there's a 'Specie:' label and a dropdown menu currently showing 'Homo Sapiens'. Below that is a 'File with all genes' section with a 'Browse...' button and a status 'No file selected'. At the bottom, a note states: 'Note: The file should contain two columns, one for gene ID (no duplicated ID allowed) and another one for fold change.'

L'aplicació ofereix dos mètodes d'anàlisi: d'una banda es pot fer ORA (Over-Representation Analysis) i d'altra banda l'anàlisi GSEA (Gene Set Enrichment Analysis). Recordem que l'ORA consisteix en seleccionar els gens diferencialment expressats i basant-se en GO, KEGG o Reactome comprobar si una de les agrupacions de gens suggerides per aquestes bases de dades està sobre o sotraexpressada en els gens seleccionats. Per dur a terme l'ORA l'usuari té opció de definir un *cut-off* de Log-Ratio per formal el conjunt dels gens que s'hi utilitzara (*gene set*). ORA és una bona eina per veure els efectes grans però els efectes petits li escapen. Els efectes petits derivats dels gens individuals poden acumular-se en un efect conjunt substancial el qual ORA no serà capaç de detectar. És aquí on GSEA mostra la seva utilitat.

## Grau de compliment dels objectius i resultats previstos en el pla de treball.

Primer recordem les tasques proposats a la PAC1:

Figure 2: Calendari



Justificació dels canvis en cas necessari

Relació de les activitats realitzades

Activitats previstes en el pla de treball

Activitats no previstes i realitzades o programes

Relació de les desviacions en la temporització i accions de mitigació si escau i actualització del cronograma si escau

Llistat dels resultats parcials obtinguts fins al moment (lliurables que s'adjunten)

Comentaris del vostre director particular si ho considereu necessari.