

FICHA DEL TRABAJO FINAL

Título del trabajo:	Implementació d'una eina en R/Shiny per a l'anàlisi de significació biològica utilitzant l'anàlisi de les rutes
Nombre del autor:	Vasyl Druchkiv
Nombre del consultor/a:	Alexandre Sánchez Pla
Nombre del PRA:	<i>Nombre y dos apellidos</i>
Fecha de entrega (mm/aaaa):	Juny/2019
Titulació::	<i>Bioinformàtica I bioestadística</i>
Àrea del Trabajo Final:	<i>Anàlisi de dades òmiques</i>
Idioma del trabajo:	<i>Català</i>
Palabras clave	<i>Pathway analysis, R, Shiny</i>
<p>Resumen del Trabajo (máximo 250 palabras): <i>Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.</i></p>	
<p>L'objectiu de treball és trobar paquets de Bioconductor i desenvolupar una aplicació Shiny per dur a terme l'anàlisi de les rutes (<i>Pathway analysis</i>). Un <i>Pathway</i> és el conjunt de gens relacionats amb una funció biològica i descriu la relació entre els gens. Primer s'identifiquen els mètodes teòrics actualment presents per trobar i visualitzar les rutes diferencialment expressades. També s'identifiquen les bases de dades per anotar les rutes i els paquets de Bioconductor específics per fer l'anàlisi. D'aquesta manera es crea una aplicació Shiny que ofereix l'anàlisi ORA (Over-representation Analysis), GSEA (Gene Set Enrichment Analysis) i l'anàlisi de la topografia de les rutes. Els paquets elegits per a l'anàlisi de les rutes de Bioconductor són: <i>clusterProfiler</i>, <i>ReactomPA</i> i <i>patview</i>. L'anotació de gens es fa finalment via tres bases de dades: GO (Gene Ontology), KEGG (Kyoto Encyclopedia of Genes and Genomes) i Reactome.</p> <p>L'usuari pot seleccionar les ontologies de GO, especificar el nivell de significació i el mètode d'ajustament. Es obté per cada base de dades una taula per a l'anàlisi ORA i l'altra per GSEA. Els resultats es visualitzen via Bar plot, Dot plot, enrichment map, gene-concept network, GO plot, KEGG pathway i Reactome pathway. Es fa possible la descàrrega de les taules en format de .csv i la de les imatges en format .png. L'usuari pot especificar la resolució i altres atributs d'imatges.</p> <p>La creació de l'aplicació està seguida per la seva validació, on es presenta la seva funcionalitat i es compara el resultat amb l'estudi original.</p>	

Abstract (in English, 250 words or less):

Objective of the thesis is to find Bioconductor packages for pathway analysis and implement them using Shiny. A pathway is a set of genes related to a specific biological function and describes a relation between that genes. First I identify theoretic methods available at the moment for finding and visualising the differentially expressed pathways. Furthermore I identify data bases for gene annotation and specific Bioconductor packages for performing analysis. Second, I develop Shiny application which offers ORA (Over-representation Analysis), GSEA (Gene Set Enrichment Analysis) and topologic analysis of the pathways. The packages selected for the analysis are *clusterProfiler*, *ReactomPA* i *patview*. Gene annotation is done using three data bases: GO (Gene Ontology), KEGG (Kyoto Encyclopedia of Genes and Genomes) and Reactome.

User can select GO ontologies, specify significance level and adjustment method. For each annotation data base a table with results is generated. Additionally results are visualized with bar plot, dot plot, enrichment map

L'usuari pot seleccionar les ontologies de GO, especificar el nivell de significació i el mètode d'ajustament. Es obté per cada base de dades una taula per a l'anàlisi ORA i l'altra per GSEA. Els resultats es visualitzen via Bar plot, Dot plot, enrichment map, gene-concept network, GO plot, KEGG pathway i Reactome pathway. Es fa possible la descarga de les taules en format de .csv i la de les imatges en format .png. L'usuari pot especificar la resolució i altres atributs d'imatges.

La creació de l'aplicació està seguida per la seva validació, on és presenta la seva funcionalitat i es compara el resultat amb l'estudi original.

