

# Analyse du protéome et de l'interactome : Exercices série 3

Anaïs Barry. M2 BCD

Le but de cette série d'exercice est d'analyser un réseau d'interaction donné à l'aide de Cytoscape et DAVID afin d'en ressortir les tendances et clusters majeurs ayant un impact dans ce réseau. Nous analyserons notamment quelques uns de ces clusters afin d'identifier leur fonction.

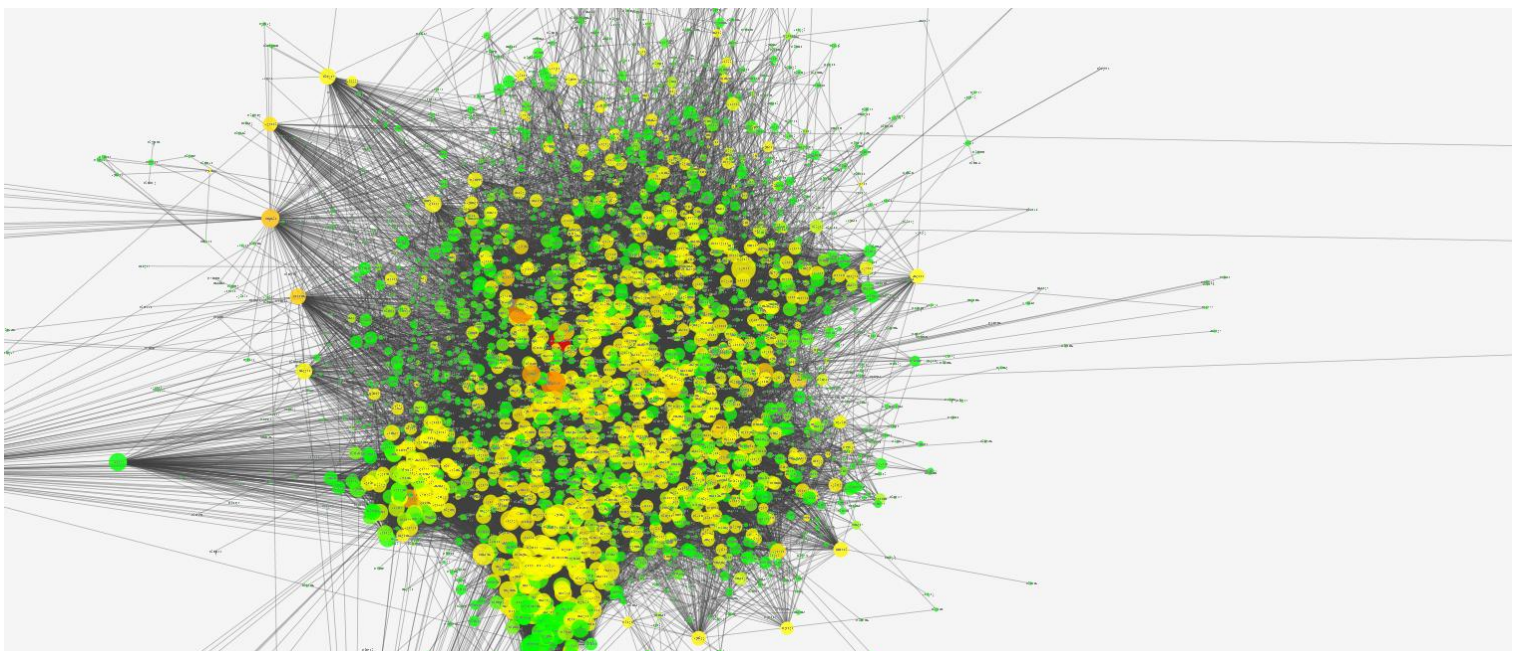
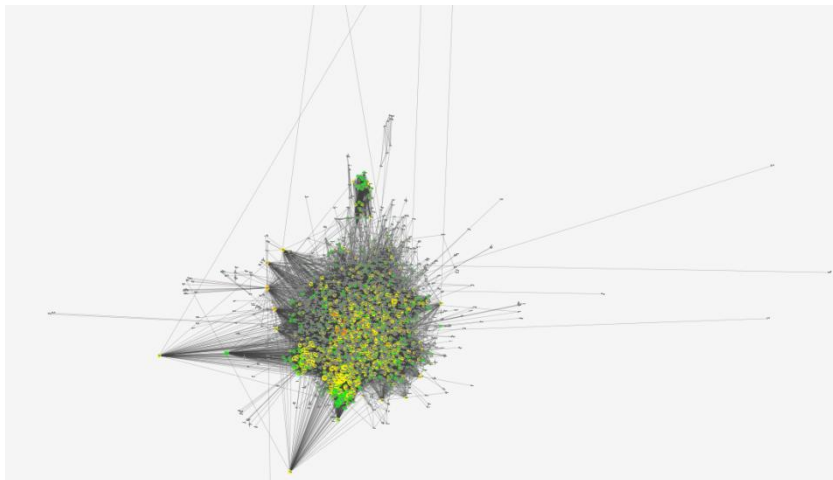
Pour cela nous utilisons un fichier d'interaction de protéines et un deuxième fichier attribuant un nom d'accension UNIPROT à ces protéines.

## *1. Exercice 1 : Analyse d'un réseau d'interaction de protéines avec Cytoscape.*

Paramètres utilisés :

- layout : organic
- node : symbol
- node color : betweenness
- size : node

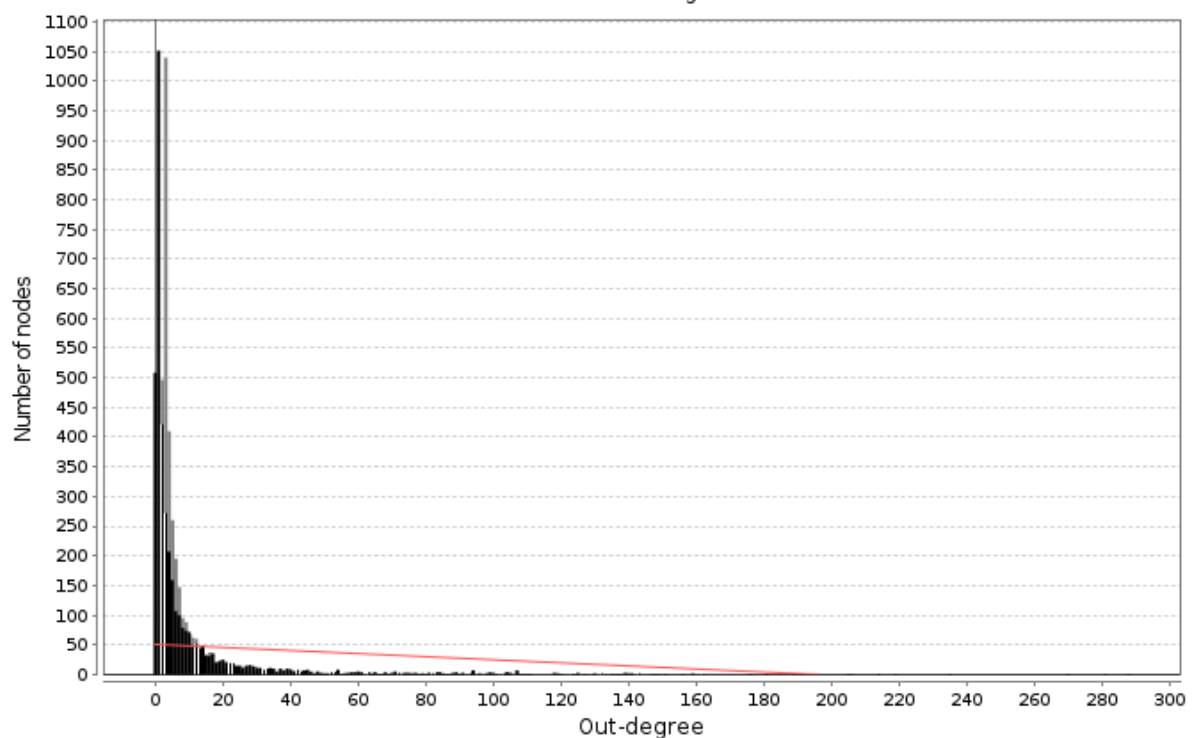
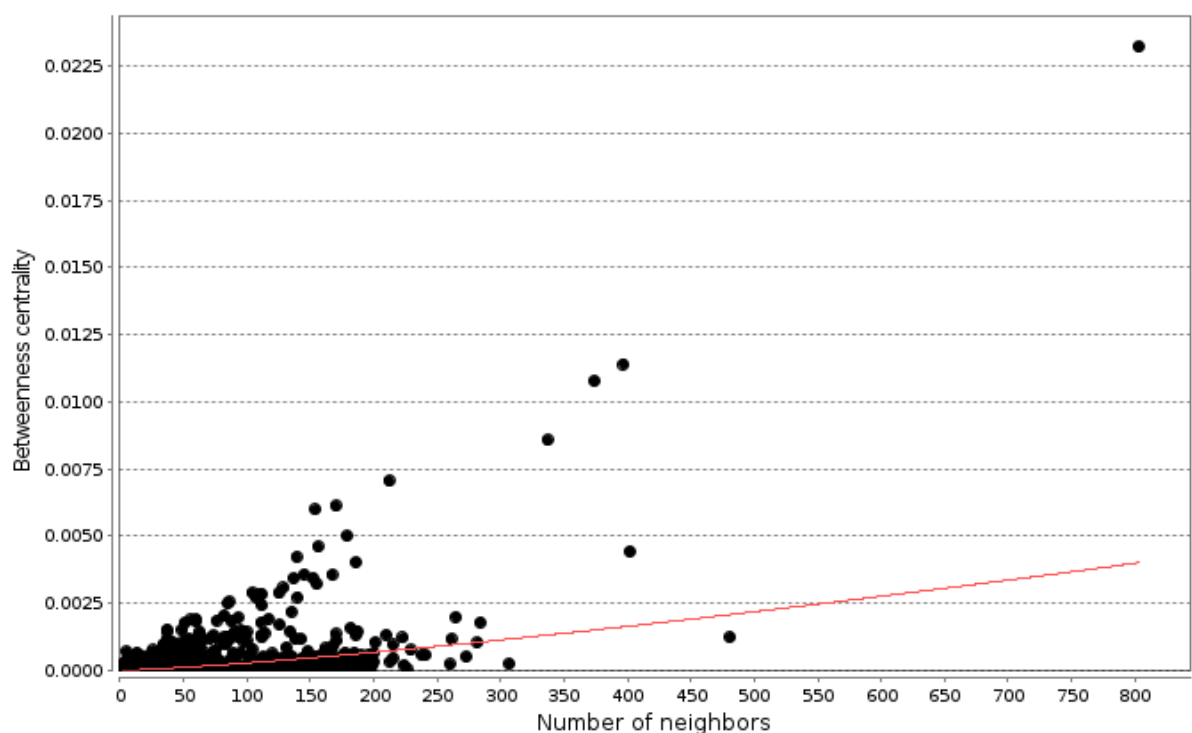
Ci-dessous figurent deux captures d'écran du réseau après la Network Analysis.



Nous distinguons des nœuds de tailles et couleurs différentes, qui dépendent de leur implication dans le réseau :

- La *betweenness centrality* d'un nœud équivaut au nombre de plus courts chemins allant d'un sommet à un autre passant par ce nœud. Les nœuds oranges/rouges signifient que de nombreux chemins y passent, nous déduisons que ces protéines ont une grande influence dans les interactions du réseau.
- Il s'agit d'un graphe orienté, et nous nous sommes ensuite intéressés au *out-degree*, c'est-à-dire au nombre d'arêtes partant d'un nœud. Le *out-degree* reflète le nombre d'interaction qu'une protéine crée avec d'autres, interactions venant d'elle. Dans nos précédentes images, plus un nœud a un grand *out-degree*, plus il sera gros.

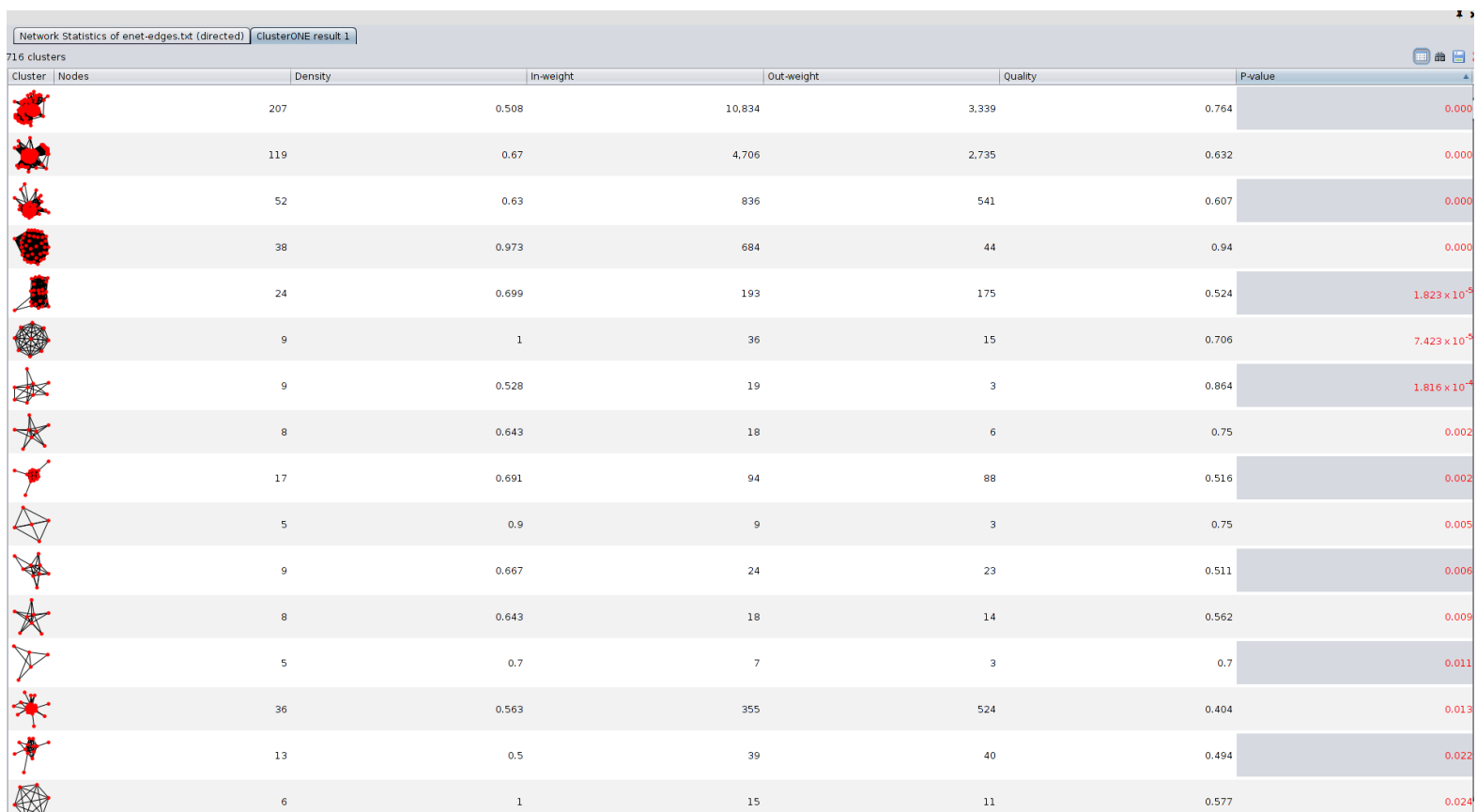
Les deux graphes suivants mettent en évidence que peu de nœuds se distinguent selon ces critères.



















## 2. Exercice 2 : Analyse de clusters avec ClusterOne et DAVID

Avec le plugin ClusterOne, nous obtenons une prédiction de 716 clusters. Ces clusters sont ensuite triés par p-value croissante : ce score représente la probabilité d'obtenir un résultat au hasard, plus la valeur est proche de 0 plus l'obtention de ce résultat est significative.

La capture d'écran suivante montre l'interface Cytoscape suite à l'analyse ClusterOne avec les paramètres par défaut et avec un tri par p-value croissante.



Cluster	Nodes	Density	In-weight	Out-weight	Quality	P-value
	207	0.508	10,834	3,339	0.764	0.000
	119	0.67	4,706	2,735	0.632	0.000
	52	0.63	836	541	0.607	0.000
	38	0.973	684	44	0.94	0.000
	24	0.699	193	175	0.524	$1.823 \times 10^{-5}$
	9	1	36	15	0.706	$7.423 \times 10^{-5}$
	9	0.528	19	3	0.864	$1.816 \times 10^{-4}$
	8	0.643	18	6	0.75	0.002
	17	0.691	94	88	0.516	0.002
	5	0.9	9	3	0.75	0.005
	9	0.667	24	23	0.511	0.006
	8	0.643	18	14	0.562	0.009
	5	0.7	7	3	0.7	0.011
	36	0.563	355	524	0.404	0.013
	13	0.5	39	40	0.494	0.022
	6	1	15	11	0.577	0.024

La prochaine étape est de sélectionner la liste de gènes correspondant à chaque cluster et de prédire leur fonction à l'aide de DAVID. Nous nous intéressons ici à la fonction Biological Process définie dans la Gene Ontology.

Les deux tables suivantes correspondent à l'analyse DAVID selon ces critères pour les deux premiers clusters. Les résultats sont ici aussi ordonnés par p-value croissante.


















## Functional Annotation Chart

[Help and Manual](#)
**Current Gene List:** clusterone\_converted

**Current Background:** Homo sapiens

**207 DAVID IDs**
☒ **Options**


 [Download File](#)
**78 chart records**

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">RNA splicing</a>	RT		163	78,7	2,8E-263	1,7E-260
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">mRNA processing</a>	RT		165	79,7	2,8E-256	8,7E-254
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">mRNA metabolic process</a>	RT		165	79,7	1,3E-242	2,7E-240
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">RNA processing</a>	RT		170	82,1	1,2E-219	1,8E-217
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">RNA metabolic process</a>	RT		171	82,6	8,3E-177	1,0E-174
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">RNA splicing, via transesterification reactions with bulged adenosine as nucleophile</a>	RT		106	51,2	5,4E-171	5,5E-169
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">nuclear mRNA splicing, via spliceosome</a>	RT		106	51,2	5,4E-171	5,5E-169
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">RNA splicing, via transesterification reactions</a>	RT		106	51,2	5,4E-171	5,5E-169
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">gene expression</a>	RT		177	85,5	9,2E-97	8,1E-95
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">nucleobase, nucleoside, nucleotide and nucleic acid metabolic process</a>	RT		177	85,5	5,3E-87	4,1E-85
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">cellular nitrogen compound metabolic process</a>	RT		177	85,5	2,0E-81	1,4E-79
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">nitrogen compound metabolic process</a>	RT		177	85,5	3,1E-79	1,9E-77
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">cellular macromolecule metabolic process</a>	RT		182	87,9	1,1E-61	5,9E-60
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">macromolecule metabolic process</a>	RT		182	87,9	8,7E-55	4,5E-53
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">ribonucleoprotein complex assembly</a>	RT		38	18,4	1,4E-51	6,6E-50
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">cellular metabolic process</a>	RT		182	87,9	1,7E-43	7,3E-42
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">primary metabolic process</a>	RT		182	87,9	2,3E-40	9,3E-39



















## Functional Annotation Chart

[Help and Manual](#)
**Current Gene List:** clusterone\_2\_converted

**Current Background:** Homo sapiens

**122 DAVID IDs**
☒ **Options**


 [Download File](#)
**96 chart records**

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">translational elongation</a>	RT		83	68,0	1,3E-167	6,7E-165
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">translation</a>	RT		85	69,7	4,0E-113	1,0E-110
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">gene expression</a>	RT		104	85,2	1,4E-53	2,4E-51
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">cellular protein metabolic process</a>	RT		90	73,8	2,9E-45	3,7E-43
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">cellular macromolecule biosynthetic process</a>	RT		94	77,0	2,4E-43	2,4E-41
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">macromolecule biosynthetic process</a>	RT		94	77,0	4,4E-43	3,8E-41
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">protein metabolic process</a>	RT		90	73,8	8,6E-39	6,3E-37
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">cellular macromolecule metabolic process</a>	RT		109	89,3	1,0E-35	6,4E-34
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">cellular biosynthetic process</a>	RT		94	77,0	1,1E-35	6,3E-34
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">biosynthetic process</a>	RT		94	77,0	1,3E-34	6,6E-33
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">ribosome biogenesis</a>	RT		29	23,8	3,4E-33	1,6E-31
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">macromolecule metabolic process</a>	RT		109	89,3	1,1E-31	4,9E-30
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">ribonucleoprotein complex biogenesis</a>	RT		30	24,6	1,3E-29	5,3E-28
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">cellular metabolic process</a>	RT		109	89,3	4,5E-25	1,7E-23
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">rRNA processing</a>	RT		22	18,0	6,2E-25	2,1E-23
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">rRNA metabolic process</a>	RT		22	18,0	1,7E-24	5,3E-23
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">primary metabolic process</a>	RT		109	89,3	3,1E-23	9,4E-22
<input type="checkbox"/>	GOTERM_BP_ALL	<a href="#">ncRNA processing</a>	RT		23	18,9	1,8E-19	5,3E-18

Nous pouvons interpréter que les gènes du premier cluster (207 gènes) sont très probablement impliqués dans des fonctions liées à l'ARN tel que l'épissage ou le traitement de l'ARNm. Le second cluster (122 gènes) est constitué de nombreux gènes impliqués dans des mécanismes de traduction et de biosynthèse.