



## Tutorial Questions

### Exercise 1

From the homepage, search for 'Signaling by EGFR'. Click on the top pathway hit. This will open it in the Pathway Browser. Ignoring the diagram for now, look at the Pathways tab on the left.

1. How many sub-pathways does this pathway have?
2. How many reactions are associated with the "EGFR interacts with phospholipase C-gamma" sub-pathway?
3. What reaction follows 'EGFR dimerization'?

*Hint: If it's not visible, open the Details pane at the bottom of the page by clicking on the blue triangle.*

4. What happens if you click on the sub-pathway 'Grb2 events in EGFR signaling'?

*Hint: If you don't see any change, use the navigator feature at the top left of the page to 'zoom out'.*

### Exercise 2

Find the reaction 'Activated type I receptor phosphorylates R-SMAD directly'.

1. What pathway does it belong to?
2. In which cellular compartment does this reaction take place?
3. What is the associated GO molecular function?
4. What references verify this reaction?
5. Is this reaction predicted to occur in *Canis familiaris*? In *Saccharomyces cerevisiae*?

### Exercise 3

Open the pathway diagram for 'Netrin-1 Signaling.'

1. Find the protein SIAH2 (centre of the cytosol). Right click on it and select 'Display Interactors'. How many are there?
2. How many times has the interaction between SIAH2 and DRPLA been documented? *Hint: This detail is not in Reactome.*
3. Find the protein SRC (up and to the left of SIAH2). Display interactors for this protein. How many are there? Can you get a list of them?
4. Display interactors for UNC5B (bottom left of the cytosol). What happens and why?
5. What is the easiest way to remove interactors?

#### Exercise 4

On the homepage, click the button Pathway Analysis. When the submission form appears click on the Example button. Select the radio button for Overrepresentation analysis. Click the button marked Analyse.

1. What is the most significantly over-represented top-level pathway for this dataset?
2. How many genes are in this pathway, and how many were represented in the dataset?
3. Why is the top-level pathway Chromosome Maintenance higher in the list than Signalling by Wnt when the latter has a more significant probability score? (*Hint – use the Open All button*)
4. For the reaction 'Raf activation', which identifiers were represented in the dataset?

#### Exercise 5

Launch the Species Comparison and select the species *Rattus norvegicus*. When the results are displayed, open the pathway Complement Cascade.

1. Find Complement factor B (centre left in the diagram) - what colour is it? What does that mean?
2. What other species is this protein inferred to be present in? *Hint: You can answer this question without rerunning Species comparison.*
3. Find Complement factor 2 (slightly up and to the right of Complement factor B) – why is it blue?
4. Find C3b (left and slightly up from Complement factor 2) – Why is it black? How many proteins contribute to this object? Are they predicted to exist in *Rattus*?
5. Why is Calcium grey?

#### Exercise 6

Launch Expression Analysis and load the example dataset. Click Analyse. When the results are displayed, find the pathway Nucleotide Excision Repair.

1. How many proteins are in this pathway?
2. How many had expression data?
3. Click on the View button to see this pathway in the Pathway Browser. Use the Experiment Browser toolbar to cycle through the timepoints.
4. Which protein has the greatest change of expression?
5. Find the complex 'Active Pol II complex with repaired DNA template:mRNA hybrid' (top right of the diagram). Which component of the complex has the highest expression at 24h?
6. What was the probe ID used to measure expression of this component?

#### Exercise 7

Go to the Reactome BioMart page (Accessible from the Navigation bar under 'Tools').

1. How would you select the "pathway" dataset?

2. If you were interested in the protein “Nucleolar transcription factor 1” (UniProt ID P17480), how would you identify pathways in Reactome involving that protein? How many pathways does your query find?
3. How would you find the UniProt IDs of the other proteins in the first of the pathways (*Hint: Pathway Stable ID: REACT\_2232*) that you discovered?