

A classification of tasks for the systematic study of immune response using functional genomics data

Supplemental data file 2 – Classification of questions/tasks

C HEDELER^{1,2*}, N. W. PATON¹, J. M. BEHNKE³, J. E. BRADLEY³, M. G. HAMSHERE³, K. J. ELSE²

¹*School of Computer Science, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK*

²*Faculty of Life Sciences, Michael Smith Building, The University of Manchester, Oxford Road, Manchester, M13 9PT, UK*

³*School of Biology, Nottingham University, Nottingham, NG7 2RD, UK*

Email: C Hedeler - chedeler@cs.manchester.ac.uk;

*Corresponding author

1 Classification of questions

Abbreviations:

- IID: Information identifier
- QID: Question identifier

1.1 Low level - simple questions

1.1.1 Genome

QID	Requ. inf.	Questions
Sequence		
G-SEQ-Q1	G-SEQ-PP	Which primary polypeptides have a similar sequence compared to a particular primary polypeptide?

Table 1. Questions - Genome - Sequence

QID	Requ. inf.	Questions
Location		
<i>Chromosome</i>		
G-LOC-Q1	G-LOC-CHR	Which genes are on a particular chromosome?
G-LOC-Q2		On which chromosome is a particular gene?
G-LOC-Q3		On which chromosomes is a set of genes (distribution)?
<i>Marker</i>		
G-LOC-Q4	G-LOC-MARK	Which genes are between two markers?
G-LOC-Q5		Between which markers is a particular gene?
G-LOC-Q6		Between which markers is a set of genes (distribution, are they close to each other)?
<i>bp-position; cM position</i>		
G-LOC-Q7	G-LOC-BP; G-LOC-CM	Which genes are next/close to a particular position (bp, cM)?
G-LOC-Q8		What is the position of a particular gene (bp, cM)?
G-LOC-Q9		What is the position of a set of genes (bp, cM) (distribution, are they close to each other)?

Table 2. Questions - Genome - Location

QID	Requ. inf.	Questions
Strain		
<i>SNP</i>		
G-STR-Q1	G-STR-SNP	Which genes have SNPs?
G-STR-Q2		Which SNPs does a particular gene have (incl. strain information)?
G-STR-Q3		Which SNPs does a set of genes have (incl. strain information)?
<i>Haplotype</i>		
G-STR-Q4	G-STR-HAP	Which genes have haplotypes?
G-STR-Q5		What are the known haplotypes of a particular gene (incl. strain information)?
G-STR-Q6		What are the known haplotypes of a set of genes (incl. strain information)?
G-STR-Q7		How many different haplotypes does a gene have (incl. strain information)?
G-STR-Q8		How many different haplotypes does a set of genes have (incl. strain information)?

Table 3. Questions - Genome - Strain

QID	Requ. inf.	Questions
Functional annotation		
<i>Gene Ontology</i>		
G-FA-Q1	G-FA-GO	Which genes have a particular Gene Ontology annotation?
G-FA-Q2		Which Gene Ontology annotation does a particular gene have?
G-FA-Q3		Which Gene Ontology annotation does a set of genes have (distribution)?
G-FA-Q4	G-FA-G)-HS	Which Gene Ontology annotation at a particular level of the hierarchy does a gene have?
G-FA-Q5		Which Gene Ontology annotation at a particular level of the hierarchy does a set of genes have (distribution)?
<i>Enzyme numbers</i>		
G-FA-Q6	G-FA-EC	Which genes have a particular Enzyme number?
G-FA-Q7		Which Enzyme number does a particular gene have?
G-FA-Q8		Which Enzyme numbers does a set of genes have (distribution)?
<i>Protein families</i>		
G-FA-Q9	G-FA-PF	Which proteins belong to a particular protein family?
G-FA-Q10		To which protein family does a particular protein belong?
G-FA-Q11		To which protein families does a set of proteins belong (distribution)?

Table 4. Questions - Genome - Functional annotation

1.1.2 Transcriptome

QID	Requ. inf.	Questions
Microarray description		
T-MD-Q1	T-MD-GL	Which genes are on a particular microarray?

Table 5. Questions - Transcriptome - Microarray description

QID	Requ. inf.	Questions
Experimental condition		
<i>Data sets</i>		
T-ECOND-Q1	T-ECOND	Which data sets have been produced under a particular experimental condition?
<i>Expression levels of genes</i>		
T-ECOND-Q2	T-ECOND; T-MD-GL; T-R-EL	Under which conditions (in which data sets) is a particular gene up-regulated?
T-ECOND-Q3		Under which conditions (in which data sets) is a particular gene down-regulated?
T-ECOND-Q4		Under which conditions (in which data sets) is a set of genes up-regulated?
T-ECOND-Q5		Under which conditions (in which data sets) is a set of genes down-regulated?

Table 6. Questions - Transcriptome - Experimental conditions

QID	Requ. inf.	Questions
Expression levels		
<i>One data set</i>		
T-EL-Q1	T-R-EL	Which genes are up-regulated in a particular data set?
T-EL-Q2		Which genes are down-regulated in a particular data set?
<i>Two data sets</i>		
T-EL-Q3	T-R-EL; T-R-TC	Which genes are differentially expressed in two data sets?
T-EL-Q4		Which genes are similarly expressed (co-regulated) in two data sets?
T-EL-Q5		Which genes have changed similarly between two data sets?
T-EL-Q6		Which genes have changed differently between two data sets?
<i>More than two data sets</i>		
T-EL-Q7	T-R-EL; T-R-TC	Which genes are differently expressed in more than two data sets?
T-EL-Q8		Which genes are similarly expressed (co-regulated) in more than two data sets?
T-EL-Q9		Which genes have changed similarly in more than two data sets?
T-EL-Q10		Which genes have changed differently in more than two data sets?

Table 7. Questions - Transcriptome - Expression levels in data sets

QID	Requ. inf.	Questions
Expression levels of a set of genes		
<i>One data set</i>		
T-ELGS-Q1	T-MD-GL; T-R-EL	What are the expression levels of a particular gene in a particular data set?
T-ELGS-Q2		What are the expression levels of a set of genes in a particular data set?
T-ELGS-Q3		Which genes of a set of genes are up-regulated in a particular data set?
T-ELGS-Q4		Which genes of a set of genes are down-regulated in a particular data set?
<i>Two data sets</i>		
T-ELGS-Q5	T-MD-GL; T-R-EL; T-R-TC	What are the expression levels of a particular gene in two data sets?
T-ELGS-Q6		What are the expression levels of a set of genes in two data sets?
T-ELGS-Q7		Which genes of a set of genes are differentially expressed in two data sets?
T-ELGS-Q8		Which genes of a set of genes are similarly expressed (co-regulated) in two data sets?
T-ELGS-Q9		Which genes of a set of genes have changed similarly between two data sets?
T-ELGS-Q10		Which genes of a set of genes have changed differently between two data sets?
<i>More than two data sets</i>		
T-ELGS-Q11	T-MD-GL; T-R-EL; T-R-TC	What are the expression levels of a particular gene in more than two data sets?
T-ELGS-Q12		What are the expression levels of a set of genes in more than two data sets?
T-ELGS-Q13		Which genes of a set of genes are differentially expressed in more than two data sets?
T-ELGS-Q14		Which genes of a set of genes are similarly expressed (co-regulated) in more than two data sets?
T-ELGS-Q15		Which genes of a set of genes have changed similarly between more than two data sets?
T-ELGS-Q16		Which genes of a set of genes have changed differently between more than two data sets?

Table 8. Questions - Transcriptome - Expression levels of a set of genes

1.1.3 Interactome

QID	Requ. inf.	Questions
Protein-protein interaction		
I-PPI-Q1	I-PPI	In which protein-protein interactions is a particular protein involved?
I-PPI-Q2		In which protein-protein interactions is a set of proteins involved (distribution)?
I-PPI-Q3		Which proteins are involved in a particular protein-protein interaction?
I-PPI-Q4		Which proteins are involved in a set of protein-protein interactions (distribution)?
Protein complex		
I-COMPL-Q1	I-COMPL	In which protein complexes is a particular protein involved?
I-COMPL-Q2		In which protein complexes is a set of proteins involved (distribution)?
I-COMPL-Q3		Which proteins are involved in a particular protein complex?
I-COMPL-Q4		Which proteins are involved in a set of protein complexes (distribution)?

Table 9. Questions - Interactome - Protein protein interactions

1.1.4 Metabolome - Control

QID	Requ. inf.	Questions
Pathways		
MC-P-Q1	M-MP-LG; C-CMRP-LP	Which genes/proteins are involved in a particular pathway?
MC-P-Q2		In which pathways is a particular gene involved?
MC-P-Q3		In which pathways is a set of genes involved (distribution)?

Table 10. Questions - Metabolome, Control - Pathways

1.1.5 Cluster algorithms

QID	Requ. inf.	Cluster
Cluster algorithms		
CL1	T-R-EL	Cluster a set of genes according to their expression levels (using different cluster algorithms).
CL2	G-FA	Cluster a set of genes according to their functional annotation.
CL3	G-FA-GO-SS	Cluster a set of genes according to their Gene Ontology annotation using semantic similarities.
CL4	G-LOC	Cluster a set of genes according to their location.
CL5	G-FA-RIF; (Jenssen <i>et al.</i> , 2001), (Masys <i>et al.</i> , 2001)	Cluster a set of genes according to references in the literature.

Table 11. Cluster algorithms

1.1.6 Comparison of results

QID	Requ. inf.	Comparison
Combination		
C1		Comparison of expression levels for a set of genes in two (or more) data sets (affy, glass).
C2		Comparison of two distributions of the same feature.
C3		Comparison of two data sets (union, intersection, difference).
C4		Comparison of clusters.

Table 12. Comparison

1.1.7 Examples of combinations of simple questions

QID	Used Questions	Question
Combination		
<i>Expression levels and location</i>		
CombQ1	T-ECOND-Q1; T-EL-Q1/T-EL-Q10; G-LOC-Q3/G-LOC-Q6/G-LOC-Q9	Which genes are up-/down-regulated under a certain condition and where are they?
CombQ2	G-LOC-Q1; T-ECOND-Q1; T-ELGS-Q3/T-ELGS-Q4/T-ELGS-Q8/T-ELGS-Q14	Which genes on a particular chromosome are up-/down-regulated under a certain condition?
CombQ3	G-LOC-Q1/G-LOC-Q4/G-LOC-Q7; T-ECOND-Q1; C3; T-ELGS-Q3/T-ELGS-Q4/T-ELGS-Q7/T-ELGS-Q8/T-ELGS-Q13/T-ELGS-Q14	Which of the genes in a particular region (chromosome, cM-position, bp-position, markers) are up-/down-/co-/differently regulated under a particular/several conditions?
CombQ4	T-ECOND-Q1; T-EL-Q3/T-EL-Q4/T-EL-Q7/T-EL-Q8; G-LOC-Q3/G-LOC-Q6/G-LOC-Q9	Looking at similar/differently expressed genes under certain conditions and their location: are they at different locations, are they close to each other?

Table 13. Combination of questions - Host - Expression levels and location

QID	Used Questions	Question
Combination		
<i>Expression levels and strain</i>		
CombQ5	T-ECOND-Q1; C3; T-EL-Q3/T-EL-Q4/T-EL-Q7/T-EL-Q8; G-STR-Q3/G-STR-Q6; G-STR-Q8	Which genes are co-/differently regulated under different conditions and do they have SNPs/haplotypes? How many different haplotypes are there?
CombQ6	G-STR-Q1/G-STR-Q4; T-ECOND-Q1; C3; T-ELGS-Q7/T-ELGS-Q13	Which genes with SNPs/different haplotypes are differently expressed under different conditions?

Table 14. Combination of questions - Host - Expression levels and strain

QID	Used Questions	Question
Combination		
<i>Expression levels and functional annotation</i>		
CombQ7	T-MD-Q1; G-FA-Q3; C2	What is the distribution of Gene Ontology annotation for a particular microarray? How does the distribution differ from another microarray?
CombQ8	T-EL-Q1; G-FA-Q3; T-EL-Q2; G-FA-Q3; T-MD-Q1; G-FA-Q3; C2	What is the distribution of Gene Ontology annotation for the up-/down-regulated genes in a particular data set? Compare the distributions for up-regulated genes with the distribution for down-regulated genes. Compare both distributions with the distribution for the whole microarray.
CombQ9	T-ECOND-Q1; C3; T-EL-Q3/T-EL-Q4/T-EL-Q7/T-EL-Q8; G-FA-Q3	Looking at similar/differently regulated genes under certain conditions and their Gene Ontology annotation: do they have different annotations/do they share annotations?
CombQ10	T-ECOND-Q1; C3; T-EL-Q3/T-EL-Q7; G-FA-Q11	To which protein families do the proteins of genes that are differentially expressed under certain conditions belong?
CombQ11	G-FA-Q1/G-FA-Q6; T-ECOND-Q1; C3; T-ELGS-Q3/T-ELGS-Q4/T-ELGS-Q7/T-ELGS-Q8/T-ELGS-Q13/T-ELGS-Q14	Which of the genes with a particular Enzyme number/Gene Ontology annotation are up-/down-/co-/differently regulated under a particular/several conditions?

Table 15. Combination of questions - Host - Expression levels and functional annotation

QID	Used Questions	Question
Combination		
<i>Expression levels of genes under certain conditions</i>		
CombQ12	T-ECOND-Q1; C3; T-ELGS-Q11;	What does a particular gene do under a particular/several conditions?
CombQ13	T-ECOND-Q1; C3; T-ELGS-Q12;	What does a set of genes do under a particular/several conditions?
CombQ14	T-ECOND-Q1; C3; T-ELGS-Q11; C1	How does the expression level of a particular gene change between two/several different conditions?
CombQ15	T-ECOND-Q1; C3; T-ELGS-Q12; C1	How does the expression levels of a set of genes change between two/several different conditions?
CombQ16	T-ECOND-Q1; C3; T-EL-Q7/T-EL-Q8	Which genes are co-/differently regulated under certain (different/similar) conditions?
CombQ17	T-ECOND-Q1; C3; T-EL-Q9/T-EL-Q10	Which genes have changed differently/similarly between certain conditions?

Table 16. Combination of questions - Host - Expression levels of genes under certain conditions

QID	Used Questions	Question
Combination		
<i>Expression levels and protein-protein interaction</i>		
CombQ18	T-ECOND-Q1; C3; T-EL-Q3/T-EL-Q4/T-EL-Q7/T-EL-Q8; I-PPI-Q2/I-COMPL-Q2; C3	Which of the genes that are co-/differently regulated under certain conditions have proteins that are involved in protein-protein interactions? Are they involved in the same interactions or different interactions?
CombQ19	I-COMPL-Q3; T-ECOND-Q1; C3; T-ELGS-Q3/T-ELGS-Q4/T-ELGS-Q7/T-ELGS-Q8/T-ELGS-Q13/T-ELGS-Q14	Which proteins can be found in a particular protein complex and which of the genes, that produce these proteins, are up-/down-/co-/differently regulated under a particular/several conditions?

Table 17. Combination of questions - Host - Expression levels and protein-protein interaction

QID	Used Questions	Question
Combination		
<i>Expression levels and pathways</i>		
CombQ20	T-ECOND-Q1; C3; T-EL-Q3/T-EL-Q4/T-EL-Q7/T-EL-Q8; MC-P-Q3; C2/C3	Looking at co-/differently regulated genes under certain conditions: in which pathways are these genes involved?
CombQ21	MC-P-Q1; T-ECOND-Q1; C3; T-ELGS-Q7/T-ELGS-Q8/T-ELGS-Q13/T-ELGS-Q14	Which genes, that are involved in a particular pathway, are co-/differently regulated under certain conditions?

Table 18. Combination of questions - Host - Expression levels and pathways

1.2 Medium level - Complex Questions

1.2.1 What happens to the host when it is infected?

QID	Questions
Pathways	
H-P-CQ1	Which pathways are differently activated in susceptible and resistant mice?
H-P-CQ2	Are there pathways that differ only partially in susceptible and resistant mice? At what stage do they differ?
H-P-CQ3	Which pathways are involved over time course of an infection in susceptible and resistant mice? At what degree do they differ?

Table 19. Complex Questions - Host - Pathways

QID	Questions
Expression patterns	
H-E-CQ1	How do the expression patterns differ between susceptible and resistant mice?
H-E-CQ2	How do the expression patterns differ between infections by different pathogens?
H-E-CQ3	How do the expression patterns change over times course of an infection?
H-E-CQ4	How do the expression patterns differ in different tissues?
H-E-CQ5	Is this observed expression pattern unique to this experiment/condition/infection or has it been observed in another experiment/under different conditions?

Table 20. Complex Questions - Host - Expression patterns

QID	Questions
Involved genes	
H-G-CQ1	Which genes are differently expressed in susceptible and resistant mice?
H-G-CQ2	Where are the genes that are differently expressed in susceptible and resistant mice? Do they form clusters?
H-G-CQ3	Which functional annotation do the genes that are differently expressed in resistant and susceptible mice have? Do they form clusters according to their functional annotation?
H-G-CQ4	Compare the set of genes that are up-/down-regulated in data sets of two/several experiments with the same tendency for a Th1/Th2 response: what is the intersection/difference/union of these sets (experiments might have been carried out using different microarrays with a different set of genes on it)?

Table 21. Complex Questions - Host - Involved genes

1.2.2 What happens to the infecting agent on infecting?

QID	Questions
Proteins produced by pathogen	
P-PR-CQ1	Which proteins does the pathogen produce that are similar to proteins produced by the host?

Table 22. Complex Questions - Infecting agent - Proteins produced by pathogen

1.2.3 What makes an individual susceptible?

QID	Questions
Strain differences	
S-SD-CQ1	Which strain differences (SNPs, haplotypes) can be found between susceptible and resistant mice or fast and slow responders?
S-SD-CQ2	Are these SNPs functional, e.g., leading to different cytokine levels/different expression levels?
S-SD-CQ3	Which strain differences (SNPs, haplotypes) can be found in genes that are involved in immune response and differently/co-regulated in susceptible and resistant mice?

Table 23. Complex Questions - Strain differences

1.3 Higher level questions - Biological Lessons Queries

1.3.1 What happens to the host when it is infected?

QID	Questions
Pathways	
H-P-BLQ1	Which of the differently activated pathways in susceptible and resistant mice are significant?
H-P-BLQ2	Which of the differences in partially different pathways in susceptible and resistant mice are significant?

Table 24. Biological Lessons Questions - Pathways

QID	Questions
Expression patterns	
H-E-BLQ1	Is an observed expression pattern tissue-specific?
H-E-BLQ2	Is an observed expression pattern pathogen-specific?
H-E-BLQ3	Is an observed expression pattern specific to a particular life cycle stage of the pathogen?
H-E-BLQ4	Is an observed expression pattern cell type specific?
H-E-BLQ5	Is an observed expression pattern strain-specific?
H-E-BLQ6	Is an observed expression pattern dose level specific?

Table 25. Biological Lessons Questions - Expression patterns

QID	Questions
Involved genes	
H-G-BLQ1	Which of the genes that are differently expressed in resistant and susceptible mice are significant?
H-G-BLQ2	Which genes are responsible for resistance and susceptibility?
H-G-BLQ3	Which genes are involved in/important for a particular immune response?
H-G-BLQ4	Which are the tissue-specific genes involved in a particular immune response?
H-G-BLQ5	Is a particular set of genes tissue-specific (infection with a different pathogen that infects the same tissue in the same or a similar way) or is it pathogen-specific (are there pathogens that can infect different tissues?)
H-G-BLQ6	Which genes are responsible for which trait, such as worm burden, production of a particular antibody, cytokine production...? Which of these genes are different in susceptible and resistant mice or slow and fast responders? Which of the differences are significant?

Table 26. Biological Lessons Questions - Involved genes

1.3.2 What happens to the infecting agent on infecting?

QID	Questions
Proteins produced by pathogen	
P-PR-BLQ1	Which proteins, produced by the pathogen and similar to proteins produced by the host, are significant to ensure the pathogens survival?
P-PR-BLQ2	Which proteins, produced by the pathogen and similar to proteins produced by the host, are made to trigger a particular immune response in the host?
P-PR-BLQ3	Which proteins, produced by the pathogen and similar to proteins produced by the host, are made to hide the pathogen from the immune system of the host?

Table 27. Biological Lessons Questions - Infecting agent - Proteins produced by pathogen

1.3.3 What makes an individual susceptible?

QID	Questions
Strain differences	
S-SD-BLQ1	Which of the strain differences (SNPs, haplotypes) found between susceptible and resistant mice are significant for differences in gene regulation?
S-SD-BLQ2	Which of the strain differences (SNPs, haplotypes) found between susceptible and resistant mice are significant for susceptibility/resistance?
S-SD-BLQ3	Which of the strain differences (SNPs, haplotypes) found in genes that are involved in immune response and differently regulated in susceptible and resistant mice are significant?
S-SD-BLQ4	Are the same SNPs/haplotypes significant for susceptibility/resistance to another pathogen/a particular group of pathogens?
S-SD-BLQ5	At what stage of immune response are the strain differences important?

Table 28. Biological Lessons Questions - Strain differences

1.4 General questions

1.4.1 What happens to the host when it is infected?

QID	Questions
Immune response	
H-IR-GQ1	Why has a mouse become infected in the first place?
H-IR-GQ2	What did the immune system do?
H-IR-GQ3	What did the immune system of the susceptible mouse do inappropriately?
H-IR-GQ4	How does the immune system decide which immune response to mount?
H-IR-GQ5	What is the result of the immune response; can it throw off the infection; is this the appropriate immune response?
H-IR-GQ6	If it is the wrong immune response, why did the mouse mount the wrong one? Was it because it is a Th1-ish or Th2-ish mouse and this happens to be the inappropriate immune response?
H-IR-GQ7	Which part of the immune response is different (recognition, innate immune response, adaptive immune response)? If there are several differences, which one are significant? What are the reasons for these differences?
H-IR-GQ8	What are the reasons for different expression patterns? At what stage of the immune responses do they differ in resistant/susceptible mice?
H-IR-GQ9	How can the different stages of immune response be identified (is it possible to identify them by expression pattern)?

Table 29. General Questions - Host - Immune response

QID	Questions
Expression pattern	
H-EP-GQ1	What are the reasons for different expression patterns in resistant and susceptible mice?
H-EP-GQ2	What are the reasons for changes in expression patterns between different conditions?
H-EP-GQ3	What are the reasons for similar expression patterns under different conditions?

Table 30. General Questions - Host - Expression pattern

QID	Questions
Treatment	
H-TR-GQ1	How does treatment influence an infection?
H-TR-GQ2	Does treatment affect the production of memory cells, i.e. if a mouse that would normally mount an inappropriate immune response is treated, and can throw off the infection because of the treatment, does it produce the right memory?
H-TR-GQ3	How does treatment influence the immune response?

Table 31. General Questions - Host- Treatment

QID	Questions
Dose level	
H-DL-GQ1	Do different dose levels cause different immune responses? If yes, at which stage do the immune responses differ and what are the reasons?
H-DL-GQ2	How do differences in dose level affect the mounting of immune response; why can a different dose level lead to a different immune response?

Table 32. General Questions - Host - Dose level

1.4.2 What happens to the infecting agent on infecting?

QID	Questions
Pathogen	
P-P-GQ1	Did the pathogen trigger the immune response of the host in the wrong direction to ensure its own survival?
P-P-GQ2	How do pathogens evolve?
P-P-GQ3	What do pathogens do to survive?

Table 33. General Questions - Pathogen

1.4.3 What makes an individual susceptible?

QID	Questions
Strain differences	
S-SD-GQ1	Why is one mouse susceptible and the other resistant?
S-SD-GQ2	Why are some strains of mice Th1-ish and others Th2-ish?
S-SD-GQ3	Do the Th1-ish mice always mount a Th1 response or are there situations when they mount a Th2 response? If yes, what are the reasons to mount the other response?
S-SD-GQ4	How do the strain differences affect the immune response and at what stage of the immune response?
S-SD-GQ5	If the location for a particular trait in an infection with a particular pathogen and in a comparison of two strains has been identified, does this location have a different function in an infection with a different pathogen, is it important at all, are there differences in comparison with other strains (similar reaction in an infection with this particular pathogen, different reaction)?
S-SD-GQ6	Which haplotypes are responsible for a particular trait, such as worm burden or production of a particular antibody?
S-SD-GQ7	What causes the different regulation of genes that are differently regulated in resistant and susceptible mice but don't have different haplotypes?
S-SD-GQ8	What causes the co-regulation of genes that are co-regulated in resistant and susceptible mice but have different haplotypes?
S-SD-GQ9	What is known about resistance or susceptibility of a particular strain to infection by a particular pathogen?

Table 34. General Questions - Strain differences

References

- Jenssen, T.-K., Lægreid, A., Komorowski, J., and Hovig, E. (2001). A literature network of human genes for high-throughput analysis of gene expression. *Nature Genetics* **28**, 21–28.
- Masys, D. R., Welsh, J. B., Fink, J. L., Gribskov, M., Klacansky, I., and Corbeil, J. (2001). Use of keyword hierarchies to interpret gene expression patterns. *Bioinformatics* **17**, 319–326.