Teacher's Guide to Clustering Exercises

Anastasios Koutsos
Alexandra Manaia
Julia Willingale-Theune

Version 2.3

English version

ELLS – European Learning Laboratory for the Life Sciences

EMBL
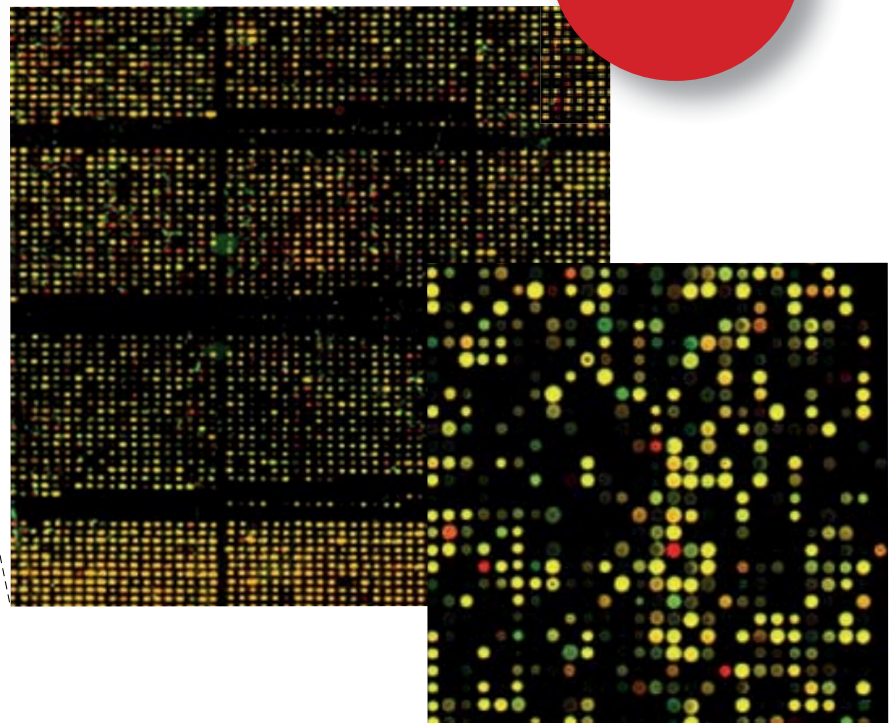
Anastasios Koutsos, Alexandra Manaia and Julia Willingale-Theune

Teacher's Guide to Clustering Exercises

Version 2.3

Teacher's Guide to Clustering Exercises_

In this section, you will need the number of green and red torches that you recorded for the virtual microarray experiment. Below is a table summarizing all the information:

Gene	Green	Red
james watson	0	0
francis crick	9	9
rosalind franklin	1	1
maurice wilkins	3	3
thomas morgan	4	8
alexander fleming	0	3
barbara mcclintock	1	6
john kendrew	2	1
jacques monod	4	0
leo szilard	9	3
Total	**33**	**34**

We now need to think of ways of grouping these genes into clusters. For example, cluster genes that have more green torches, more red torches, or an equal number.

More red torches means that the gene is **up regulated**, more green means that the gene is **down regulated**. The clustering could be as follows:

Cluster	Genes	Green	Red	Regulation
1	james watson	0	0	Identical
	francis crick	9	9	Identical
	rosaslind franklin	1	1	Identical
	maurice wilkins	3	3	Identical
2	thomas morgan	4	8	2-fold up regulation
	barbara mcclintock	1	6	6-fold up regulation
3	jacques monod	4	0	disappearing
	alexander fleming	0	3	appearing
4	leo szilard	9	3	3-fold down regulation
	john kendrew	2	1	2-fold down regulation

Now that we have performed the clustering, it is time to see what these genes have in common. The genes are all named after famous scientists who have worked in the field of biology.

•   *Search in the library or browse the Internet to find information about these scientists Try to find out if they have something in common[1].*

As a clue, most of them have received a Nobel Prize in Physiology or Medicine for their research. A good starting point is The Nobel e-museum website (http://www.nobel.se) which provides a lot of information about Nobel laureates and their discoveries. You will find short facts about their life and research, as well as articles and educational activities that you can use in class.

However, if you are still puzzled, here are two clustering schemes. You can of course devise your own ones.

Let's take **Cluster 1** first. This consists of four genes (james watson, francis crick, maurice wilkins and rosalind franklin), which show equal amounts of mRND in control and test cell lines. You probably already know that James Watson, Francis Crick, Maurice Wilkins and Rosalind Franklin all played a role in the discovery of the double helix structure of DNA and the way that it stores information in the sequence of its nucleotides. James Watson, Francis Crick and Maurice Wilkins received the Nobel Prize in Physiology or Medicine in 1962. Unfortunately, Rosalind Franklin's role in the discovery was not fully recognized until after her death.

Cluster 2 shows two genes that are **up regulated**. Thomas Morgan pioneered work on fruit flies (*Drosophila melanogaster*), which are still a favourite model organism for many biologists. He received the Nobel Prize of Physiology or Medicine in 1933. Barbara McClintock received the Nobel Prize in 1983 for her discovery of genetic mobile elements in maize; pieces of DNA that are able to move from one part of the genome to another.

Cluster 3 shows two genes that light up in one colour only. Alexander Fleming received the Nobel Prize for Physiology or Medicine in 1945 for the discovery of penicillin. Jacques Monod received the Nobel Prize in 1965 for biochemical studies on enzymes. So the common denominator is that both of these scientists conducted research on micro-organisms.

[1] To help in the Internet search, the real spelling of the Barbara mcclintock is Barbara McClintock.

Cluster 4 shows two genes that are down regulated. This cluster is a little bit trickier to decipher. Leo Szilárd was a physicist and was among a group who were the first to discover atomic energy. He was also one of the scientists who fought forcefully against its use as a weapon. Gradually his interest shifted from physics to biology. In 1962, Leo Szilárd, Victor F Weisskopf, James Watson and John Kendrew met in Geneva to discuss the possibility of establishing an international laboratory for molecular biology. In 1963, the European Molecular Biology Organization was founded, and in 1974, the European Molecular Biology Laboratory, situated in Heidelberg, became a legal entity. John Kendrew, apart from winning the Nobel Prize in Chemistry in 1962 for the first crystal structure of a protein, became the first Director General of EMBL.

In this example, you can clearly see that the similarities in gene activity are directly linked to some "functional" similarities shared by the genes (scientists), i.e. in cluster 1 they all worked on the structure of DNA, in cluster 2 they are both genetisists, in cluster 3 they work on mircoorganisms, and the link between Leo Szilárd, and John Kendrew is EMBL!!

However, when interpreting microarray results, you must be careful. Sometimes, genes that are contained in the same cluster, may not share any similarities at all. Sometimes this has to do with the actual genes, but sometimes it also has to do with the way that the clustering is done. For example, some of your students may make the following clustering:

Cluster	Genes	Green	Red	Result
A	james watson	0	0	no regulation
	francis crick	9	9	no regulation
	rosaslind franklin	1	1	no regulation
	maurice wilkins	3	3	no regulation
B	thomas morgan	4	8	2-fold up regulation
	barbara mcclintock	1	6	6-fold up regulation
	alexander fleming	0	3	up regulation
C	leo szilard	9	3	3-fold down regulation
	john kendrew	2	1	2-fold down regulation
	jacques monod	4	0	down regulation

Acknowledgements

ELLS employs creative commons copyrights to protect material produced for ELLS LLABs which will subsequently be used by teachers and other institutions. The copyright symbols also appear on the ELLS TeachingBASE website and in the downloadable pdfs/docs/ppts.

Attribution Non-commercial Share Alike

This license lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms. Others can download and redistribute your work just like the by-nc-nd license, but they can also translate, make remixes, and produce new stories based on your work. All new work based on yours will carry the same license, so any derivatives will also be non-commercial in nature.

Furthermore, the author of the derivative work may not imply that the derivative work is endorsed or approved by the author of the original work.

Explanation of the copyright symbols

   to Share

   to Remix

   Attribution

   Noncommercial

   Share alike

For further details, see http://creativecommons.org