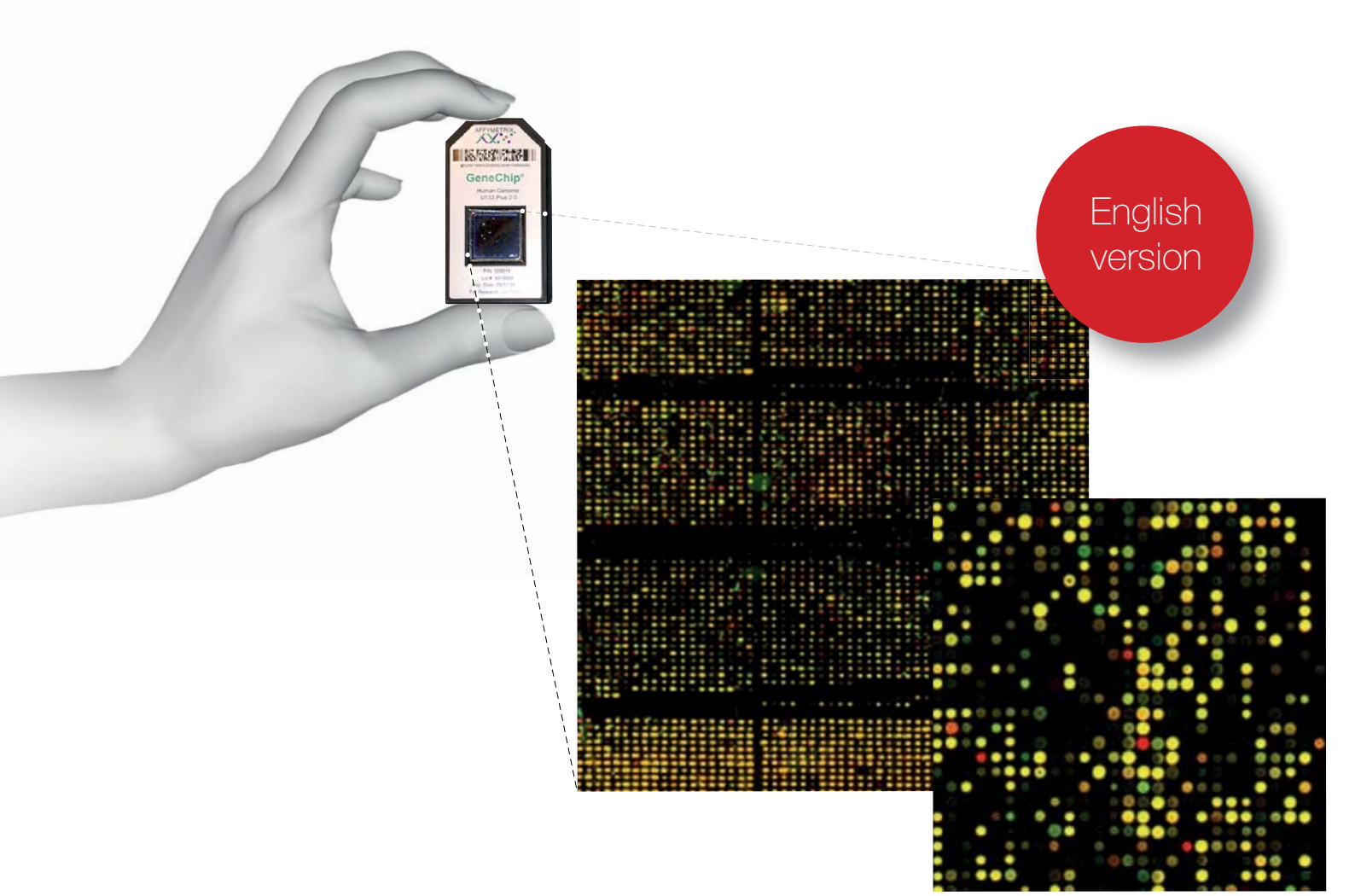
Reading Club: a Look at Scientific Papers

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

Reading Club: a Look at Scientific Papers

Version 2.3

Reading Club: a Look at Scientific Papers

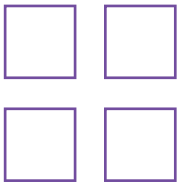
Now that you understand how microarrays work (at least this is what we hope!), it is time to see how researchers have used them to get answers to biological problems.

We are actually going to take a look at a piece work that researchers have completed and published in a **scientific journal**—scientists commonly refer to this as a "paper". Papers are essential in science. They allow information to be shared among researchers, so that people avoid doing experiments that have been done before.

To the layman, scientific papers are full of technical jargon. Sometimes, only scientists who work in the same field can understand it. But so that you do not feel intimidated, we have provided you with a summary of each paper containing all the important facts.

Before you start

Before reading the "papers", it might be useful to know some things about them. Usually, papers follow the

Title	**Materials and**	Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla	**Acknowledgements (f)**
Abstract (a)	**Methods (c)**	Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla
Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla
Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla	**References (f)**
Introduction (b)	Bla Bla Bla Bla Bla Bla		Bla Bla Bla Bla Bla Bla	1 Bla Bla Bla Bla Bla Bla
Bla Bla Bla Bla Bla Bla	**Results (d)**		**Discussion (e)**	2 Bla Bla Bla Bla Bla Bla
Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla		Bla Bla Bla Bla Bla Bla	3 Bla Bla Bla Bla Bla Bla
Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla		Bla Bla Bla Bla Bla Bla	4 Bla Bla Bla Bla Bla Bla
Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla		Bla Bla Bla Bla Bla Bla	5 Bla Bla Bla Bla Bla Bla
Bla Bla Bla Bla Bla Bla	Bla Bla Bla Bla Bla Bla		Bla Bla Bla Bla Bla Bla	6 Bla Bla Bla Bla Bla Bla

IMRAD format (**I**ntroduction, **M**aterial and Methods, **R**esults and **D**iscussion):

a) A small paragraph immediately after the title of the paper and the author's(') name(s), which summarizes the work that has been done in the paper. This is called "**abstract**", and might be missing on rare occasions.

b) **Introduction**: This part introduces the reader to the biological question addressed in the work. People

usually describe work that has been done before (by them or by others) in the same field and, at the end of the introduction, they state what they have done and what will be presented in the paper.

c) **Materials and methods**: This section is mostly technical and usually describes the way the experiments have been done. This section is very important. Scientists can see that by adapting to techniques, they get different results.

d) **Results**: This section, which is the usually the longest, describes the results of the experiment.

e) **Discussion**: In this section, scientists usually discuss how the results answer the biological question they have posed in the introduction. Sometimes they also talk about future experiments they could do, in order to answer the question in more detail.

f) **Acknowledgements**: Although this is not an official section, more and more scientists include it nowadays. Here, the authors thank people that have helped in their work. They also mention the funding sources that have enabled them to do their research.

g) **References**: If you have tried to read one of the papers (in spite of telling you not to do so yet!) you might have noticed strange numbers and/or names appearing in parenthesis every now and then. These numbers and/or names refer to work of other authors that is relevant to the current work. At the end of the paper, all the details about these papers appear, so that it is easy to find them.

My papers do not have some of the sections mentioned above. Does this mean that they are no good?

There is no reason to worry about this. This does not reflect the quality of the paper. It is just that different journals have different styles. If you read some of these articles, you will realize that there are parts that serve the purpose of the introduction, materials and methods, etc. without explicitly stating so.

Who decides which work can appear in a journal?

Actually, this is a very good question. The process of publishing papers is a science on its own!

First of all, the experiments must provide new information about something (Usually you can not publish a paper just by repeating experiments that others have done in the past). Then, it is time for the researchers to write a text explaining the scientific

background, their idea, their experiments and the new information they want to share. They do this by writing a text with the sections discussed above.

Then, the scientists send this text, along with their pictures, to a scientific journal: this process is referred to as “paper submission”. The journal will then give this text to some other scientists, commonly called “referees”, who will judge the quality of the work, the clarity of the assumptions and the conclusions. This evaluation process is called peer-review and is the most critical step, because sometimes the referees may not be satisfied with the quality of the work, and reject a paper for publication.

In other circumstances, the referees may be satisfied with the overall quality of the work, but may also ask the scientists to perform some additional experiments, in order to make their conclusions more concrete. In such a case, the scientists do the additional experiments, and return the revised text to the referees.

In the end, if the referees are satisfied with the work, they give their consent for the paper to be published in that particular scientific journal. This means that the “paper has been accepted”. This is also the time of celebration in the lab, with food, and champagne!

Now that you have learned the essentials about scientific papers, here are the summaries of selected papers...

Reading Club 1

Chips trace gene switches during normal life cycle:

a story from the yeast

De Risi et al. (1997) Science 278:680-686

When a gene is active, or as we say, expressed, its DNA sequence is copied (transcribed) into mRNA. Then, the ribosomes read the mRNA sequence and assemble the corresponding chain of amino acids, to make up the respective protein (translation).

Gene expression can be modified according to particular contexts. To understand the behaviour of a cell, it would be very useful to be able to identify the genes that are tuned on and off in the presence and the absence of hormones, toxins, environmental signals, etc.

One way chips can be used is to allow scientists to track, simultaneously, "minute–to-minute" changes of a large number of genes in particular contexts.

John DeRisi and his colleagues wanted to understand the molecular basis of spore production in yeast. The yeast life cycle involves a metabolic shift from anaerobic (also known as fermentation, by which alcohol is produced) to aerobic conditions (also known as respiration, in which carbon dioxide is produced). In presence of a rich sugar medium the yeast preferential metabolism is anaerobic. When sugar is exhausted the metabolism switches to aerobic and there is production of resistant spores. This switch from anaerobic to aerobic conditions involved the switching on and switching off of many genes.

What the researchers did first was to make the chips; in other words to print 6400 spots: each spot corresponding to the sequence of a single yeast gene.

Then, they cultured yeast cells in sugar-rich medium, took samples of the culture at different time points and isolated mRNA from these samples. Using a red fluorescent dye, they stained the mRNAs and mixed this mRNA with a reference sample that they had taken at the beginning of the experiment, which had been labelled with a green fluorescent dye. They flooded this mixture over the surface of the chip and several mRNAs bound to the spots.

In this way they obtained a DNA chip corresponding to every culture time-point analysed. Each red dot corresponded to genes that had been turned on. Yellow dots meant that gene expression had not changed (after all, they knew that equal amount of fluorescent green and red make yellow). Green dots corresponded to turned off genes.

This allowed them to identity the set of genes that were active in resting conditions and how their expression evolved during the culture growth and progressive sugar exhaustion.

The significance of this work goes way beyond yeast physiology. This is the very first work using a DNA microarray and it established the basis of using chips to see how genes work in concert to affect a cell's behaviour.

Reading Club 2

Chips trace gene switches during normal life cycle:

a story from the fruit fly

Arbeitman MN et al . (2002) Science 297: 2270-227

Research in Drosophila melanogaster, commonly called the fruit fly, laid the basis for modern developmental biology. Amongst the different invaluable data about development regulation we learned from fruit flies, they taught us how an organism's body plan is set up. The three developmental biologists involved in this discovery were awarded the Nobel Prize in Medicine and Physiology in 1995[1].

Until recently we were limited to trace individual or a small set of gene activities during the normal fly life cycle: we were able to identity when and where they took place within the organism. But we didn't have a picture of global processes like: which genes are active during particular developmental stages and how they interact, and which genes are active in a particular tissue or organ and how they interact.

Using DNA chips allowed us to extend the single gene analysis to the genome level by measuring the activity of thousands of genes simultaneously.

In the Arbeitman et al. paper the activity of a total of 4028 genes (approximately one-third of all fruit fly genes) were analysed for 40 days—a span of time stretching from fertilization of the egg to the "old age" . During the first 10 days the egg transforms to a larva (these slimy little caterpillars we usually see in fruit), a pupa and an adult; the other 30 days correspond to adulthood (here, males and females were sampled separately).

The 40-day period was broken into 66 smaller time windows. Since changes happen very rapidly in the first stages, those stages were examined in overlapping one-hour window time frames, whereas later stages had much longer time windows. Each experimental sample was compared to a reference sample representing all stages of life cycle, allowing comparison of the different developmental stages and establishment of the relative activity of each gene.

This study gathered an incredible amount of information, which enabled scientists to:

- Quantify the global activity of genes through development
- Identify the global activity pattern of genes (waves of expression)

[1] 1995 Nobel prize winners were Christiane Nusslein-Volhard, Edward Lewis, Eric Wieschaus. Nusslein-Volhard and Wieschaus conducted their experiments at the EMBL.

- Identify which sets of genes are active during different developmental stages
- Identify which sets of genes are active in particular tissues or organs at different developmental stages
- Recognize sets of genes which have a similar activity pattern (waves of expression) during common developmental stages; these are genes that could potentially work together
- Point out differences in the sets of genes which are active in males and female development.

Reading Club 3

Genome expression analysis of Anopheles gambiae: responses to injury, bacterial challenge, and malaria infection.

Dimopoulos G et al. (2002) PNAS 99:8814-8819

Malaria is the most common and deadly parasitic disease in the world. Each year, the world over, malaria destroys, through premature death and disability, the equivalent of at least 35 million years of healthy, productive human life. Africa is terribly affected, and accounts for over 90% of reported cases of malaria.

It is caused by Plasmodium, a parasitic protozoa that infects human red blood cells[2]. Plasmodium[3] parasites have a complex life cycle. In order to live, they need to have both a human and a mosquito host. The mosquito host has to be of the genus Anopheles. The mosquito picks up the malaria parasites from the blood of an infected human when it feeds. The malaria parasite reproduces itself in the gut of the Anopheles mosquito. Then, the mosquito passes the malaria parasites to the human in its saliva upon subsequent blood feeding.

Scientists are interested in getting a picture of how the mosquito deals with the infection by Plasmodium. If they understand what happens during the infection, they will eventually develop tools to block it at some point

In the paper of Dimopoulos et al., microarrays were used to identify the mosquito's genes that are activated/ "inactivated" in response to infection by Plasmodium.

Sequences corresponding to 6000 red cell mosquito genes were spotted on a slide.

To make a profile of the mosquito's genes that triggered during infections in general, they extracted mRNA from mosquito cells infected with different types of bacteria, stained them with fluorescent dye and allowed to them to hybridize to the genes on the mosquito chip. Then, they did same procedure using, this time, mRNAs extracted from mosquito Plasmodium infected cells. By comparison they could identify the mosquito genes, which are triggered specifically during infection by Plasmodium.

[2] Protozoa are unicellular organisms that are as sophisticated as a human cell.

[3] There are four different species of malaria parasites (Plasmodium falciparum, Plasmodium vivax, Plasmodium malaria and Plasmodium ovale) which cause types of malaria that are somewhat different from each other. The worst type is caused by Plasmodium falciparum which kills approximately 1-2% of those infected. Falciparum malaria is characterized by fever, headache, and weakness. Complications of falciparum malaria include cerebral malaria, in which the brain is infected; severe malaria, in which the parasitic infection essentially "runs out of control"; and placental malaria, in which falciparum interferes with pregnancy and can led to death. Each of these complications is very serious and often fatal.

Reading Club 4

Chips help classify and predict cancer

Golub TR et al. (1999) Science 286: 535-537

Besides basic biological research, DNA chips can be used in a more applied fashion. Lander's group's paper demonstrates the use of DNA chips as a powerful tool in clinical diagnostics, in particular for cancer identification (classification) and cancer type prediction. They did this by using acute leukemias as a test case and focused on two types of leukaemia: acute lymphoblastic leukemia (ALL) and acute myeloide leukaemia (AML).

Acute leukemia is a disease of white blood cells and of their precursors. It is caused by malfunction of blood cell formation leading to the production of immature cells which proliferate and replace normal blood cells in the bone marrow, peripheral blood and frequently in the liver, spleen and lymph nodes. In ALL, the early immature abnormal cells population is lymphoid, whereas in AML the malignant is myeloid.

Until recently, the distinction between ALL and AML was established through the interpretation of a series tests (histological, immuno-histochemistry, cytogenetics, molecular biology), each performed in separate in a highly specialized laboratory. Although generally accurate, leukaemia classification remained imperfect and errors occur. It is crucial to be able to make an accurate prediction of a patient's leukaemia type in order to prescribe the most appropriate treatment.

Lander and colleagues wanted to come up with a unique, extremely and accurate test, so that it could later be used by clinicians, for the diagnosis of these two types of cancer.

First, they improved the classification of the two types of leukaemia. They gathered around 40 samples of bone marrow samples of patients already known to have either ALL or AML and extracted their RNA. They allowed each of these samples to bind with already available DNA chips (from a biotech company), containing around 700 human genes. After statistical analysis they were able to identify the global set of genes that correlated with ALL and AML situations. In other words, they were able to establish those genes that were specifically involved in each of the leukemias.

What they did next was to use these patterns of gene expression to see if they could correctly predict which type of cancer other patients had.

In a near future, doctors will be able to use this test to decide which is the best treatment for new leukaemia patients. Researchers also plan to develop similar test.

Acknowledgements

We would like to thank the following individuals for their advice and help during the making of this activity:

Udo Ringeisen and the entire staff of the EMBL Photolaboratory for printing the virtual microarray mat and the ‚lite' version, for use in the classroom;

Thomas Sandmann, PhD student at EMBL, Heidelberg, for helpful discussions and suggestions, and for drawing our attention to the excellent educational material of the NIH Office of Science Education supported by the Office of Research on Women's Health called ‚Snapshots of Science and Medicine';

Russ Hodge of the Office of Information and Public Affairs (OIPA) at EMBL Heidelberg, as well as the European Learning Laboratory for the Life Sciences (ELLS) staff, for helpful discussions, suggestions and never-ending encouragement;

Dr. Giovanni Frazzetto, Mehrnoosh Rayner and Vassiliki Koumandou for reading through the first version of the virtual microarray teacher's guides;

Friends and staff of EMBL Heidelberg with whom we shared our ideas, enthusiasm and concerns;

The microarray exercise has been adapted from ‚Snapshots of Science and Medicine' which can be found on the following web page science-education.nih.gov/snapshots;

The cover image by André-Pierre Olivier;

Layout design by Nicola Graf;

Edited by Corinne Kox.

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