The Genetic Switch Regulating Activity of Early Promoters of… Baceriophage TP901-1
Madsen et al (1999), Journal of Bacteriology 181:7430-7438
A tour

I. Why read this article

Well, you're reading this tour, so evidently you've already made your decision (unless you think, quite wrongly, that reading the tour can substitute for reading the article). But why? At this point, a reasonable response might be "Because the calendar said so", but that answer doesn't have much legs: (a) In a couple of years, there will be no one to tell you what to read – what then? and (b) Even sooner, in a couple of weeks, you'll need to find articles pertinent to your project – how to decide which you should spend time on?

The fact is, there are too many research articles to read – tens of thousands per year – and reading even one constitutes a sizable commitment of time. You need a reason to do it. That reason is not the same as you might have for picking up a novel ("Entertain me!") or a text book ("Teach me!"). Research articles are not designed to do either. They uniquely are designed to present experimental truth – what was actually observed and under what conditions. They are ***not*** designed to fill your open cranium with insight. If you approach a research article as a passive receptacle, you will be overwhelmed, confused, and disappointed. For you to gain from reading a research article, you need to take control. What truth are you looking for? Does this article promise to give you the truth you seek? No? Then toss it over your shoulder. Maybe? Then examine it further, throwing out the extraneous parts and focusing on the parts that address your needs.

Even that entails a sizable commitment of time, but at least you'll often get what you asked for.

This early in the semester, however, you probably don't have a burning need for a specific truth that this article can offer, so, for the purpose of illustration, let me provide one.

> *You have a newly discovered interest in bacteriophage and are on the brink of assuming a role as an agent of phage discovery. You will be examining poorly assimilated phage DNA sequences, trying to discern from them what manner of phage they came from. One critical distinction you want to make is whether the phage has a lysogenic or purely lytic life style. What in the DNA sequence can provide you with clues? You believe that any lysogenic phage should have a repressor protein that blocks the expression of lytic genes when the phage is persisting in the lysogenic state. If you find such a protein encoded in a phage sequence, that phage is more likely to be capable of lysogeny.*
>
> *How can you decide if a phage gene encodes a bona fide repressor? Many phage repressors have been identified in the past… how was that done? What kind of evidence has convinced others that a certain phage protein is indeed a repressor? What kind of evidence might you be able to produce yourself, with the tools you have or will have in hand?*

SQ1. Drawing on your knowledge of phage repressors, explain what you are looking for, in terms that you would have been able to understand a year ago (before you took this course). What kinds of evidence can you imagine might be useful in identifying a phage repressor?

With that firmly in mind, let's take a look at the article.

II. First pass through the article

First, you need to be looking at the article. The reference to it is in the title of these notes, and that should be enough. Do you have the article? If so, you can skip the next paragraph.

> You need the article – the full text, not just the abstract. These notes will be worthless without it, so don't waste your time. If the article is in a major journal, as this one is, Google or similar may be sufficient to lead you to the article. If not, then you can go through an article database, like PubMed or Web of Science, as described on the course web page (look at the calendar for the first day of class). You don't have to be on campus to get to full length articles. Do it now!

II.A. Title and abstract

Your first task is to judge whether this article will meet your needs: Does it address the question whose answer you seek? The ability to scan dozens of articles rapidly greatly increases your chances of finding the One. In the bad old days, all we had to go on were titles in tables of contents, and you could see people in the library flipping madly through the beginnings of journals looking for titles that jumped out at them. They could be excused for blowing past titles like "*On the Topology of the Genetic Fine Structure*", which gave little clue of their importance (that one was the one of the first to describe in molecular terms what a gene is). Now with search engines, titles are of less importance, so don't be too perturbed if a title that emerges from your search makes no sense to you.

SQ2. What in the title of the article by Madsen et al leads you to believe that the article might or might not address your question?

SQ3. By the way, what IS that question?

Abstracts are often more confusing than titles, as they generally are written within tight word limits and may be more designed to attract hits by search engines than to elicit understanding by humans. I generally spend no more than a couple of seconds on an abstract, looking for something that will immediately confirm the article as worth reading (i.e. able to answer my question) or help me realize that the search engine was fooled and that I should move on to another article.

It's generally easy to get titles and abstracts, more work to get to the full-text article. Titles and abstracts are useful to attract the attention of computers and to help us decide quickly whether to fetch the full-text. The decision is often arbitrary, based on imperfect information, but supposing you chose to get the full-text (***and I hope you did!***), let's move on.

II.B. INTRODUCTION

Once at the full-text, I generally do two things to judge whether I should read the article: (1) Skim the Introduction, and (2) skim the results, particularly the section headers, figures, and tables. The Introduction tests the ability of the authors to tell a story, one that leads the reader from a description of a broad problem of obvious importance to the enunciation of the specific, much more obscure question addressed by the work. Most authors are not up to the task, and that's a shame, but ultimately nothing you should concern yourself with. There are poorly written articles that contain gems hidden within them and well written articles that say nothing you want to know about. For now, skim the Introduction, looking to see whether you recognize the focus of the article (and that it's the focus you're looking for) and whether you can pick out the central

question addressed by the authors (generally at the end of the Introduction). There are two major schools of thought on how to end an Introduction: (1) End it with the central question(s), and (2) End it with the major result(s).

SQ4. Do you recognize ideas expressed in the Introduction?

SQ5. Does anything in the Introduction persuade you that the article is going to address your question? If so, what?

SQ6. Which school of thought concerning Introduction endings do the authors belong to? Can you discern the central question addressed by the article?

II.C. MATERIALS AND METHODS

The Materials and Methods section… don't bother with it, unless your goal is to find out something about a specific technique. When you judge the article is worth reading AND you find an experiment within it you want to understand, THEN, maybe, it will be time to read part of this section.

II.D. RESULTS

My goal is to find results that, if I understood them, would help me reach an answer to my question. My goal at this point isn't to achieve that understanding but to see if I've found a good article for my purposes. So for now, I scan:

Tables 1 and 2

Just a list of strains and plasmids. Boring. Forget about them.

Orf4 is a negative regulator of p_L and p_R

SQ7. What do you think of this last section?

"P_L" and "P_R"? That sounds pertinent, since the lambda repressor acts on something like that. "Negative regulator"? The lambda repressor is one of those. Also, the section refers to Figure 1, the top of which looks strikingly like the *cI-cro* region of the lambda genome. The rest of the figure is mysterious, but that's OK for now. I mark this section as a Yes.

Antirepressing activity of ORF5

SQ8. What do you think of this section?

I'm not interested in antirepressors at the moment, just repressors. I mark this section as a No.

SQ9. What about the remaining five sections of the Results?

ORF4 and ORF5 are *trans* acting

No clue, and the first sentence doesn't help. It's associated with Table 3, which is equally obscure. I mark this section as a Neutral, Probably Not.

The state of p_L can be changed

State of p_L? I'm not sure. The first sentence speaks of "clonal variation", which doesn't burst with meaning. It's associated with Figure 2, which has an appealing shape but no obvious meaning. I'll mark this as Neutral, Probably Not.[1]

Study of the regulation of promoters by primer extension

I know that the lambda repressor regulates promoters. That's good. Primer extension? Let that go. The section refers to Figure 3, which looks just like a sequencing gel! I mark this section as a Yes.

Effect of ORF4 and ORF5 on phage immunity

I seem to recall that one consequence of lysogeny is making the host immune to further phage infection. Maybe this section has something to do with that. If so, then it might help me see how to identify a repressor. I'll mark this is as a Probably.

Induction of mitomycin C is dependent on RecA and ORF5

"Induction"? "Mitomycin C"? "RecA"? None of this jumps out at me. On the other hand, the first sentence talks about lambda and the maintenance of lysogeny and the binding of the . I'll mark this as a Maybe.

I end up with two sections I definitely want to read, one probably, one maybe, and the rest no.

II.E. DISCUSSION

I usually don't bother with this section until I've read what I want to read from the Results and have some curiosity as to what the authors thought about it.

II.F. REFERENCES

You might think it's silly to read the references, but sometimes that's the most useful part of an article. If the article is close to what you want but not quite, it might refer to another article that's closer yet. Better to let someone who knows something about the are do the work of finding articles for you.

III. Thorough reading of portions of the article

III.A. Orf4 is a negative regulator of p_L and p_R

Just as this tour is of no use unless you're looking at the article, the article is of no use unless you're looking at the data. A research article is, at root, a tour of the results of experiments. So my first task, after deciding I'm going to read seriously this section of the paper, is to figure out what data the authors are taking me through. Skimming the section, I see that everything is an explanation of Figure 1. So I need to understand Figure 1.

SQ10. Actually, before we go to Figure 1, try reading the first paragraph of this section. Make any sense of it? (I don't mean the conclusions but the actual results they're describing)

[1] Actually, the results reported in this section are to my mind the most fascinating in the article, a clear example of epigenetics (non-DNA-based inheritance), a very important topic. Look it up.

Pretty rough going, in my opinion. Now put that paragraph aside and go for Figure 1.

SQ11. Which part of Figure 1 is actual experimental result and which is merely information about the experiment?

Surely the strains and plasmids aren't experimental results, although I may have to figure out what they're about before I understand the experiment. Likewise, I think the map at the top is just to help me understand, not a result. Maybe the lines are supposed to be a graph? But "Specific β-gal. activity"… numbers! I will almost certainly need to understand what they mean. I hope the figure legend or text will help.

SQ12. Does the figure legend for Fig. 1 or the text in this section say anything informative about the significance of "β-gal. activity"?

SQ13. What do you think high β-gal activity signifies? Low activity?

I can get a general idea from the figure legend and text what β-galactosidease activity is supposed to tell us, but I'd be more comfortable if I knew how the experiment worked. So now, for the first time, I look through the Materials and Methods section for any clues.

SQ14. Can you find any part of the Materials and Methods section that describes the nature of the experiment that produced β-galactosidase activity?

Unfortunately that strategy leads to a dead end. I'm sent to reference 21, but that's a book. Maybe it's online, but not likely.[2] I could go to the library, but before going that far, maybe I can find something else pertinent on the web.

SQ15. What is β-galactosidase? Why is it used in this experiment?

SQ16. How did the authors measure β-galactosidase activity, and what are Miller units?

By now I've gotten the idea that the experiment tried to measure transcription directed by promoters p_R and p_L, determined by fusing one promoter or the other to genes *lacLM*, which encode β-galactosidase and measuring the resulting enzyme activity. That seems to be the basic experiment, but then you'd think there would be just two numbers: one for p_R and one for p_L. Instead there are 12, six in Fig. 1A and six in Fig. 1B.

SQ17. Why so many numbers?

Each enzyme activity measurement has on the same line an associated strain designation and plasmid designation, plus a line (except for two). I recall that there were tables that explained the strains and plasmids. Maybe they'll help here. Considering the top line:

SQ18. What do the authors tell us about strain PM75? (Note: The nomenclature "X/pY" means that the strain is a derivative of strain X, differing from it in that it carries the plasmid pY. A strain with no description is probably just a wild-type strain.)

[2] In the interest of full disclosure, I should say that since I'm an old guy, I believe I happen to have that book, but never mind.

Well, that seemed a waste of time! Evidently, all that's relevant about the strain is that it is a strain of *Lactococcus lactis* subsp. *cremoris*, not surprising, since ***every*** strain in this article is of that bacterium, and that it carries pPM126, which we already knew from the second column of Fig. 1. Maybe the other table will be more informative.

SQ19. What do the authors tell us about plasmid pPM126? (Note: The nomenclature "pX::Y PCR (n1-n2)" means that the plasmid is a derivative of plasmid pX, differing from it in that it carries a fragment of DNA from coordinate n1 to coordinate n2 – the footnote says from "sequence Y14232" – amplified by Polymerase Chain Reaction, and cloned into restriction site Y of pX)

SQ20. From the coordinates given for pPM126, can you make any sense of its associated line shown in Fig. 1A?

SQ21. What in the world is "sequence Y14232"? It's really irritating that the authors didn't explain this. You might be able to figure it out from the context. If not, then pay a visit to GenBank (http://www.ncbi.nlm.nih.gov/) and search for Y14232 within the DNA database.

SQ22. The figure legend talks about "fragments cloned upstream of the *lacLM* genes". If you were to put in the figure an arrow representing *lacLM*, where would you put it? Which direction would the arrow be pointing?

SQ23. What's the meaning of the vertical line associated with pPM138 and some other plasmids? What's the significance with regards to the experiments?

SQ24. Pointing to specific results, explain how the authors arrived at the last sentence of the paragraph ("*The levels of repression are 62-…*").

SQ25. Pointing to specific results, explain why the authors chose to entitle this section "*ORF4 is a negative regulator of p_L and p_R*".

III.B. Study of the regulation of promoters by primer extension

The gene fusion experiments we've just gone through give indirect information about the regulation of transcription. It's possible that β-galactosidase activity in the experiments is regulated at another level altogether besides transcriptional initiation (though I certainly wouldn't bet on it). Does ORF4 regulate transcription initiating in the region between ORF4 and ORF5? To address that question, the authors get closer to the actual event – transcription – measuring the product of transcription, RNA. How do they do it?

SQ26. Where is the data that serves as the basis for this section?

"Primer extension" in the header for this section, "primer extension" in the text, and "primer extension" in the main figure. No doubt about it, we will make no headway until we figure out what "primer extension" is all about.

SQ27. Take time out and educate yourself on primer extension. Do you have an idea what it does and how it works? Do you understand why Fig. 3 looks like a sequencing gel? Do you understand the nature of the three lanes that *don't* look like they're part of a sequencing gel?

The sequence at the right of Fig. 3A and Fig. 3B seem to say something very important about p_L and p_R.

SQ28. Those numbers to the right of the sequence… you've seen "-10" before. Where? What could "+1" mean?

It's time to look at the relevant sequence in context. Fortunately, you've done this before.

SQ29. Using the tricks you learned to analyze the lambda *cI-cro* region, find the region in phage TP901-1 between ORF4 and ORF5. Can you locate the sequences shown in Fig. 3A and 3B?

SQ30. Copy the intergenic region to your favorite word processor and underline the two promoter sequences. What strikes you about the region *in between* the two promoters?

SQ31. Using another trick you learned in the analysis of lambda, bring up a map of the genes of TP901-1. Identify in the map where ORF4 and ORF5 are. Do you see any interesting generalities?

SQ32. From what you discern from the sequence, compare the ORF4-ORF5 intergenic region with that of *cI-cro* -- not the exact sequences but the pattern of features.

SQ33. Taking together everything you've gotten from the article and the sequence, what arguments can you muster for or against the proposition that ORF4 is the general repressor of lytic function in phage TP901-1?