Designing Effective Figures

20.109 Communication Workshop 1

Please find your seat next to your partner(s)
and get started on the individual brainstorming activity.
Write down a few responses to the broad questions on the sheet.

Sean Clarke and Prerna Bhargava

MIT BE BIOLOGICAL ENGINEERING | Communication Lab

be.mit.edu/communicationlab

Why care about scientific communication?

What makes you feel that any communication has been successful?

As a receiver?

As a sender?

When you think of successful scientific communication, what words/phrases come to mind?

What can you do after reading a good paper or viewing a good figure?

What makes you feel that communication has been successful?

As a receiver?

Clear message, logic flows, you can find your way around, visual appeal

As a sender?

Reward (citation, grade, funding), good feedback: questions or criticism

We often blame ourselves for struggling to understand talks or papers…

"I got stuck here. I feel like there was a huge logical leap I couldn't follow."

"There's way too much going on in this plot. What am I supposed to be looking at?"

but poor communication is often the barrier, not your scientific understanding.

In these workshops, we'll turn your instincts as a reader of science into tools for identifying…

- WHEN scientific communication is confusing
- WHY it's confusing
- HOW to fix the problem

…and start applying these tools to your 20.109 communication tasks.

How the workshops will go

1. Discuss an example from the field
2. Derive principles and strategies
3. Practice strategies
4. Go home with a checklist/rubric

Practice with a fellow at the MIT BE BIOLOGICAL ENGINEERING | Communication Lab !

The 8 communication criteria

WHY

Strategic purpose
Identified a strategic goal for your communication and achieved it

WHO

Audience
Understood your audience's needs and optimized the efficiency with which they can process your information

WHAT

Context
Included the appropriate context and motivation for the problem you are addressing; explained who would care and why

Ideas
Presented scientifically accurate information; application knowledge is creative and novel

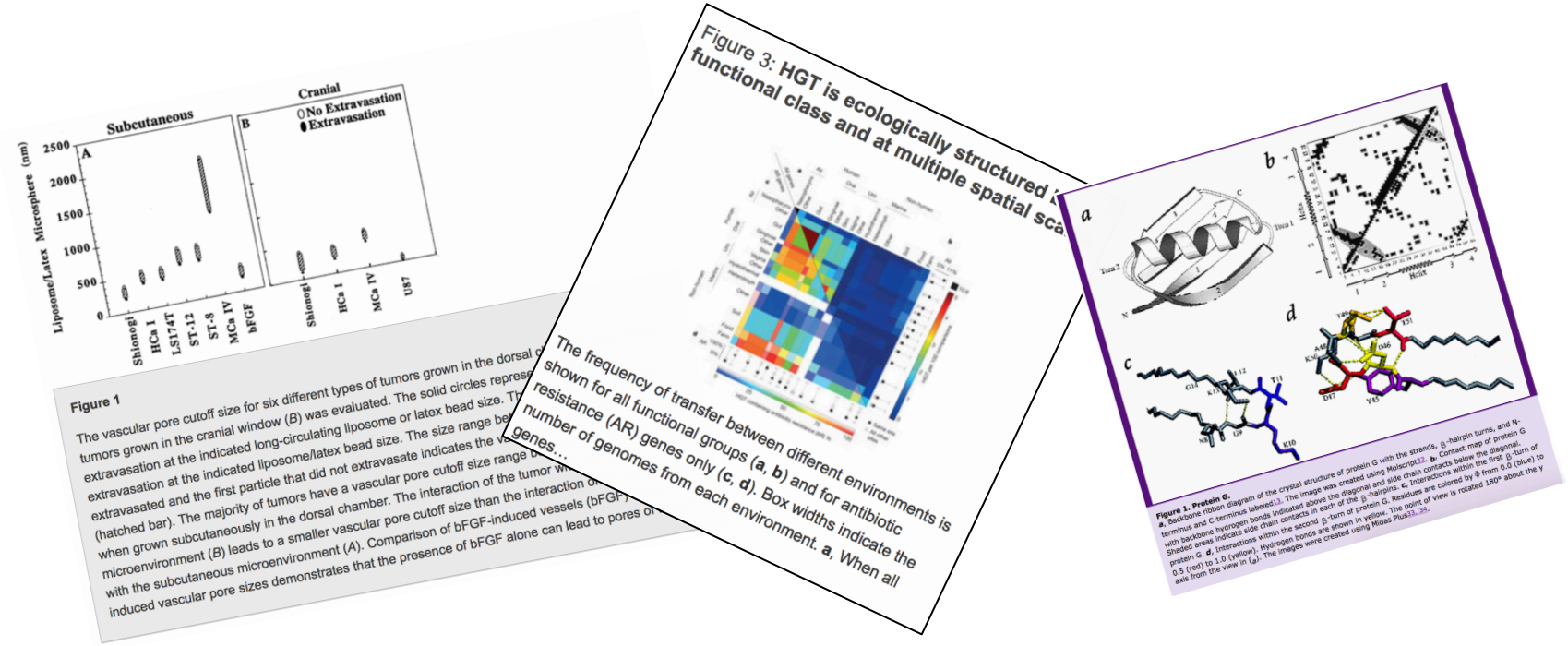
HOW

Organizational structure
Chosen an effective structure for your purpose and content

Flow
Developed smooth segues between ideas

Visual impact
Designed visual elements that support your message

Synergy
Created a streamlined yet effective redundancy of information



Figures (and captions)

Why start here?

This Article

Accepted manuscript posted online 6 August 2010, doi: 10.1128/JB.00524-10

J. Bacteriol. **October 2010** vol. 192 no. 19 **5103-5114**

Abstract ***Free***

» Figures

Full Text

PDF

Science 06 Feb 2009:
Vol. 323, Issue 5915, pp. 741-746
DOI: 10.1126/science.1159388

Article | **Figures & Data** | Info & Metrics | eLetters | PDF

Abstract | Full Text | Authors & Info | Figures | Metrics | Related Content | PDF

Figures must convince your audience of your data's impact and credibility.

- Expert audiences may ONLY READ your title, abstract, and FIGURES.
- Hold your "naked" data up to be judged.
- Help tell your story compellingly AND honestly.

Identify the basic figure components

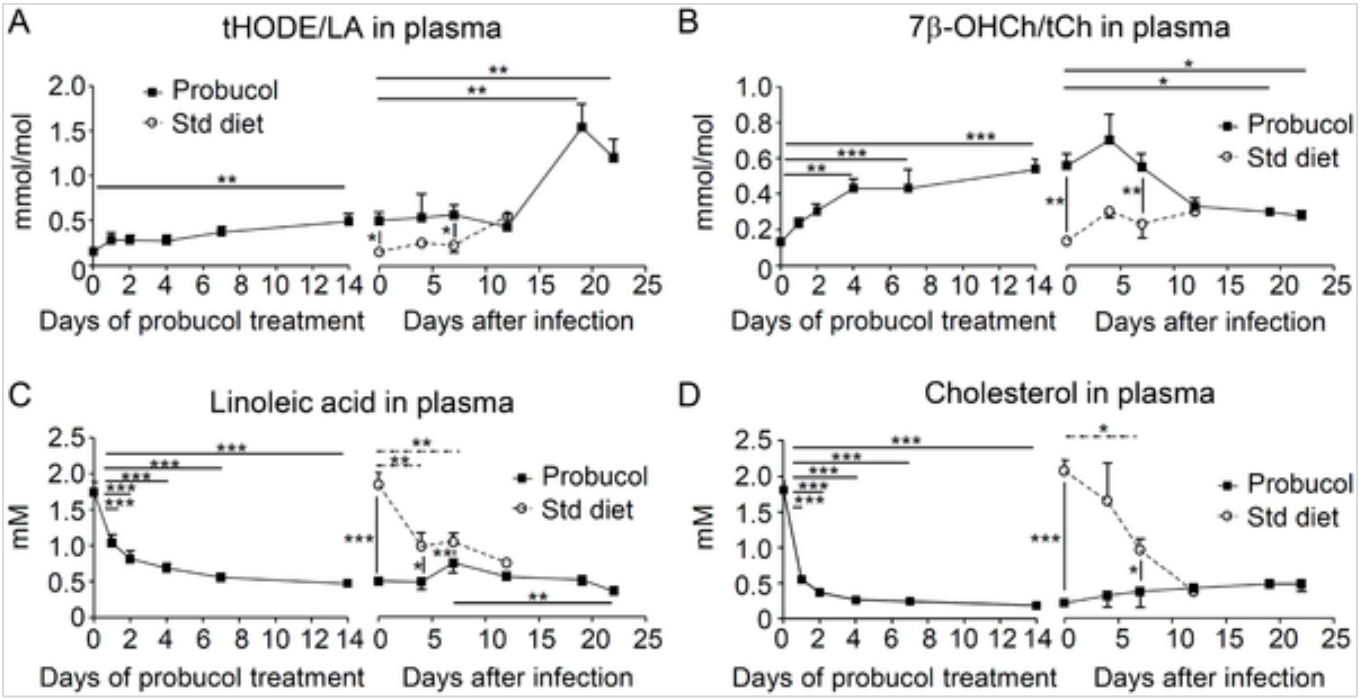


Fig 4. The ratios of lipid peroxidation products to parent lipids in plasma increased after probucol pre-treatment. Six-week-old C57BL/6J mice were treated with 1% w/w probucol in the diet for 2 weeks and then infected with 0.2 mL of 1 × 10^5 erythrocytes /mL infected with *Plasmodium yoelii* XL-17. Plasma samples were obtained at day 0, 1, 2, 4, 7, and 14 after starting the probucol diet (n = 5 per group) and at day 0, 4, 7, 12, 19, and 22 post-infection (n = 2 to 7). The ratio of total hydroxyoctadecadienoic acid (HODE), a peroxidation product of linoleic acid (LA), to linoleic acid (tHODE/LA) in plasma (A) and the ratio of 7β-hydroxycholesterol (7β-OHCh), a peroxidation product of cholesterol, to total cholesterol (7β-OHCh/tCh) in plasma (B) were measured. The concentration of LA (C) and tCh (D) were measured by using gas chromatography-mass spectrometry (GC-MS). All data are expressed as mean ± SE. Statistical analysis was carried out by analysis of variance (ANOVA). *$p < 0.05$, **$p < 0.025$, and ***$p < 0.001$. The solid bars indicate the significant changes in probucol-treated groups and the dotted bars indicate the significant changes in standard (Std) diet-fed mice.

Basic figure components

Figure = message + data

Choice of data

Presentation choices

Title

Caption

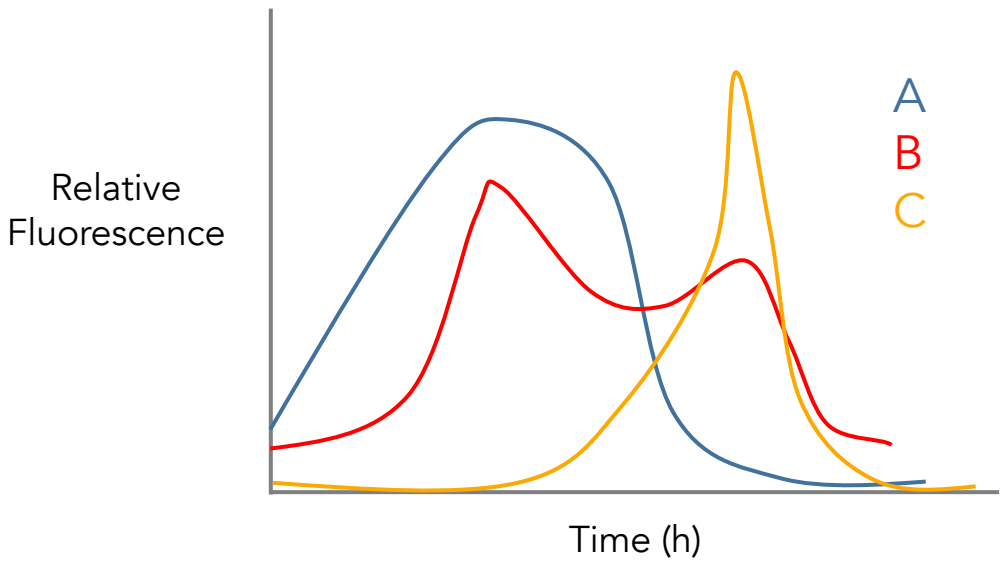


Fig. 1: A, B, and C have different dynamics under Condition X. A, B, and C were sampled using Method 1 and their fluorescence quantified with Method 2. Fluorescence data normalized to negative control.

Basic figure components

Figure = message + data

Choice of data

Presentation choices

Title

Caption

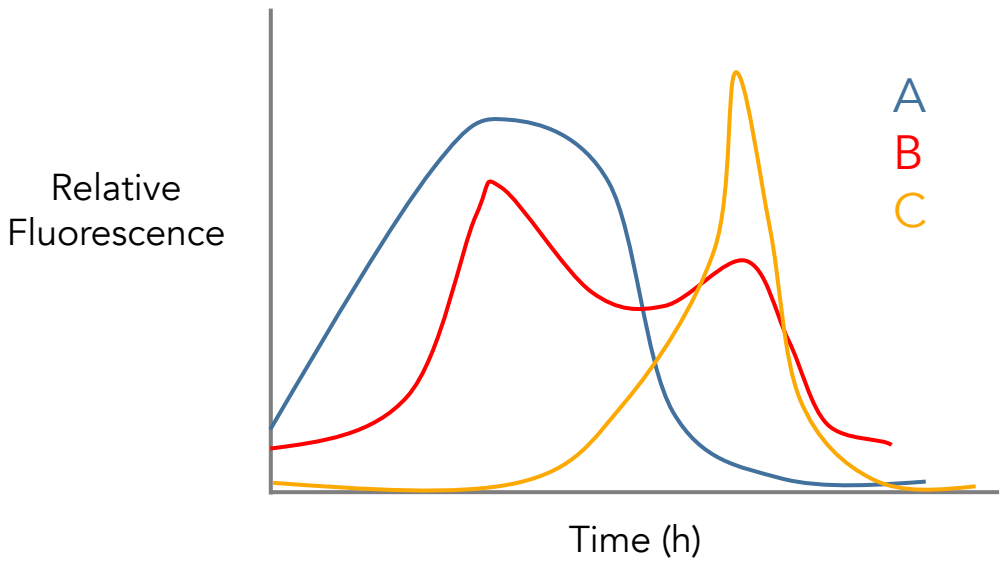


Fig. 1: A, B, and C have different dynamics under Condition X. A, B, and C were sampled using Method 1 and their fluorescence quantified with Method 2. Fluorescence data normalized to negative control.

Basic figure components

Figure = message + data

Choice of data

Presentation choices

Title

Caption

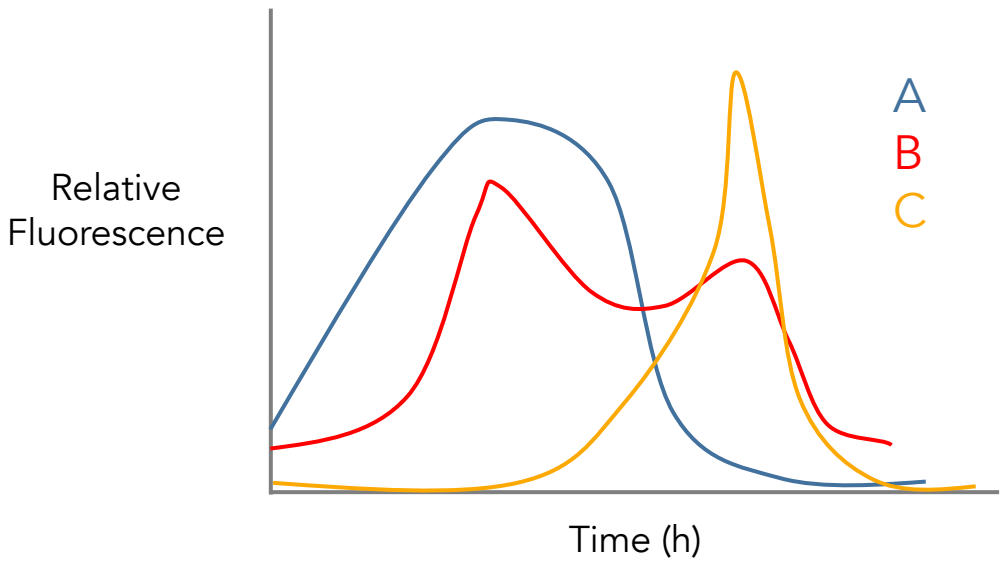


Fig. 1: A, B, and C have different dynamics under Condition X. A, B, and C were sampled using Method 1 and their fluorescence quantified with Method 2. Fluorescence data normalized to negative control.

Basic figure components

Figure = message + data

Choice of data	• Only data that are critical to the conclusion
Presentation choices	• Type of graph or display, legends & labeling, design choices • Uncluttered • Allow quick evaluation of conclusions, without referring to legend or caption.
Title	• Take-home ***message*** of the figure • What conclusion should the reader evaluate when looking at the figure?
Caption	• Descriptive, not explanatory/interpretive. • Only enough methodological detail to make it clear how results were obtained.

All the figures you make need all of these components

Schematics

Diagrams

Photos

…count as figures too.

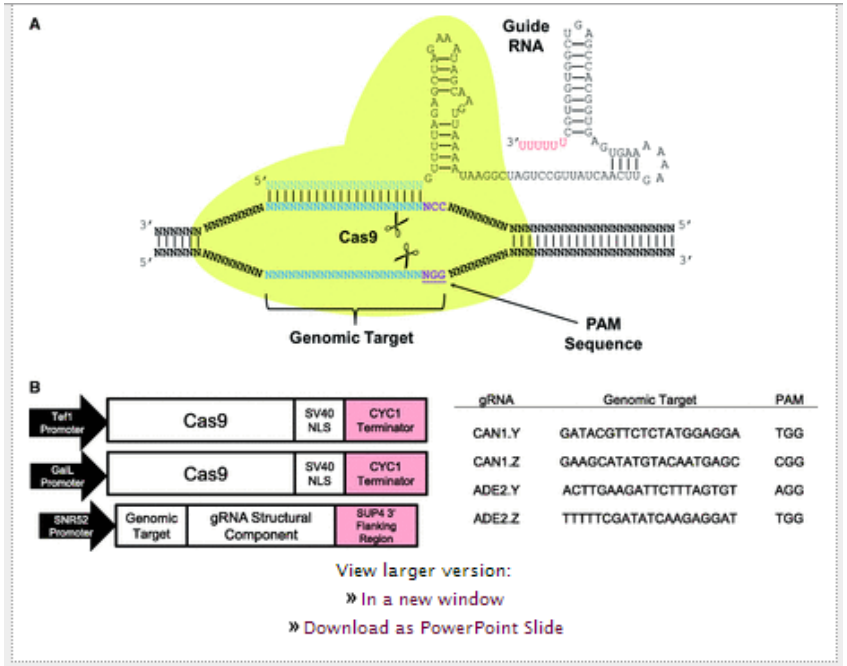
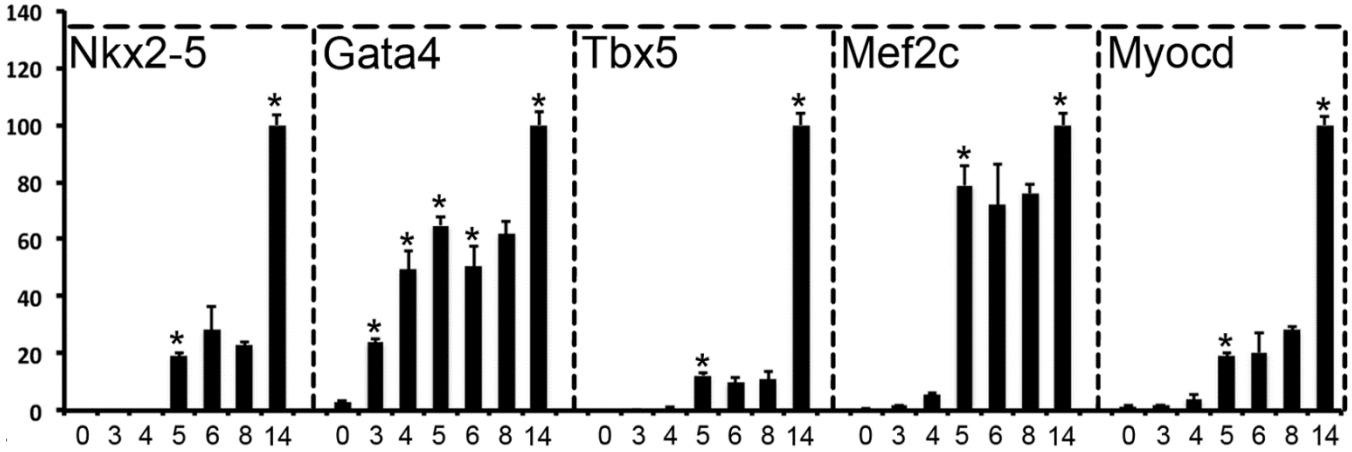


Figure 1.
Diagram of Cas9 complex and schematic of genetic constructs. (**A**) Illustration of Cas9 protein interacting with CRISPR gRNA to direct endonuclease activity proximal to the PAM sequence. (**B**) Design of the Cas9 and gRNA constructs. Cas9 gene contained a SV40 nuclear localization signal and was expressed under the Gal-L inducible promoter in CAN1 experiments and the TEF1 constitutive promoter in ADE2 experiments. The gRNA was

DiCarlo et al., 2013 Nuc. Acids. Res.

Use titles to state a figure's message, not the method



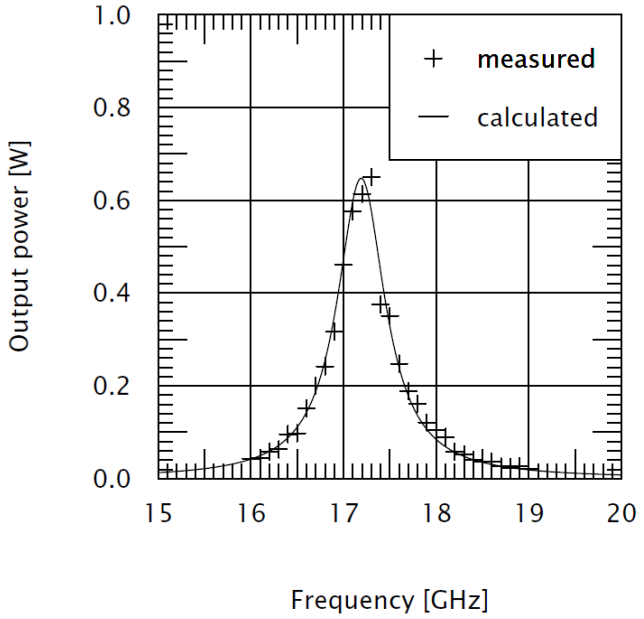
Gene expression analysis performed on differentiating mouse iPS cells

Expression of early cardiac transcription factors increases over time in differentiating mouse iPS cells

Christoforou N, *et al.* (2013) PLoS ONE 8(6): e65963.
doi:10.1371/journal.pone.0065963

Maximize signal-to-noise:

State your message.
Eliminate anything that distracts from it.



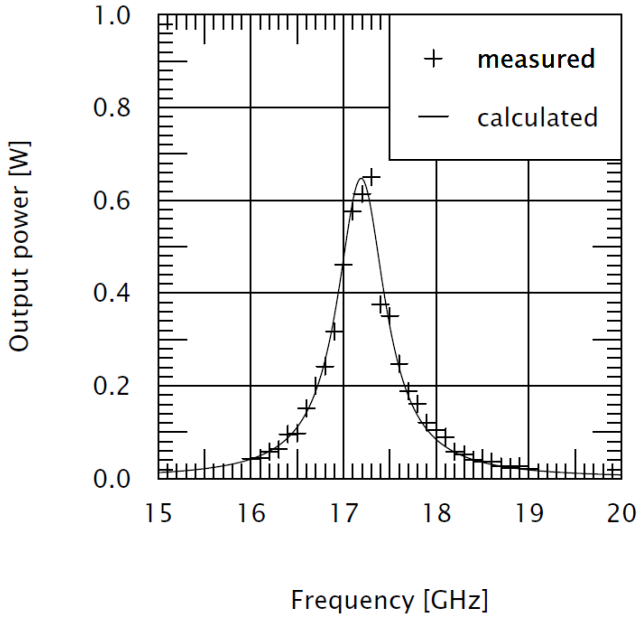
Low signal-to-noise ratio

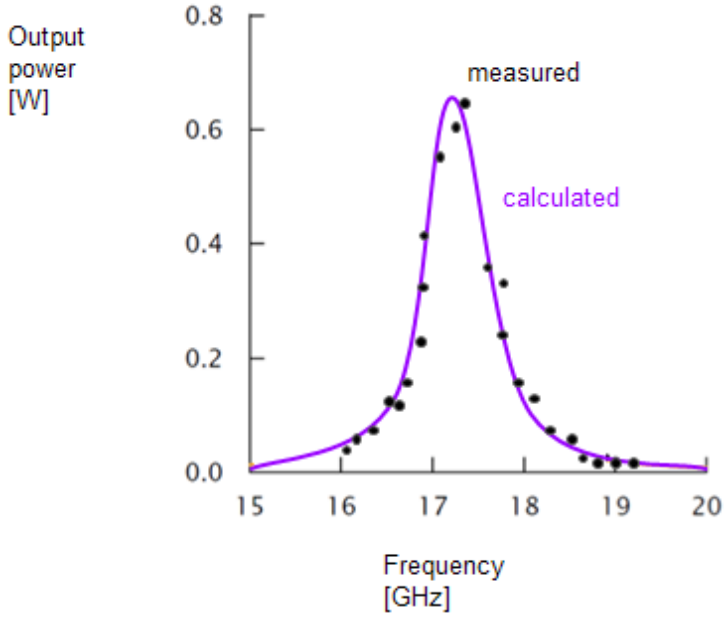
The background interferes with the data.

Adapted from Jean-luc Doumont, *Trees, Maps, and Theorems*

Maximize signal-to-noise:

State your message.
Eliminate anything that distracts from it.





Low signal-to-noise ratio

The background interferes with the data.

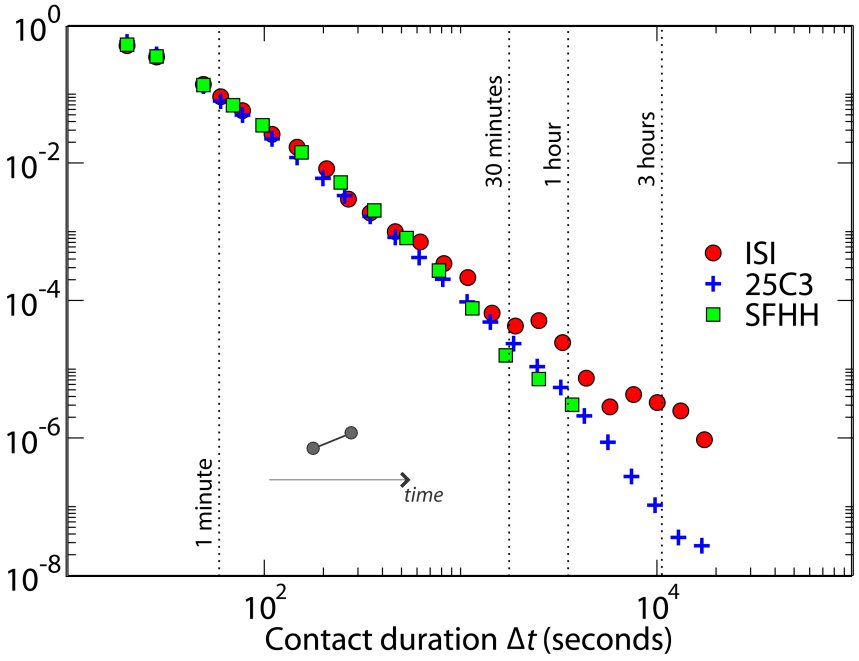
High signal-to-noise ratio

Only the necessary information is shown.

Adapted from Jean-luc Doumont, *Trees, Maps, and Theorems*

Only show as much data as you need to convey your message.

Probability distribution of human interactions at 3 conferences



At a poster session, 50% of your audience walks away after 20 sec.

Time talking	Probability listening
0 sec	100%
20 sec	50%
1 min	10%
2 min	5%
5 min	<1%

Cattuto et al., 2010, PLoS ONE

Activity: Improve this published figure

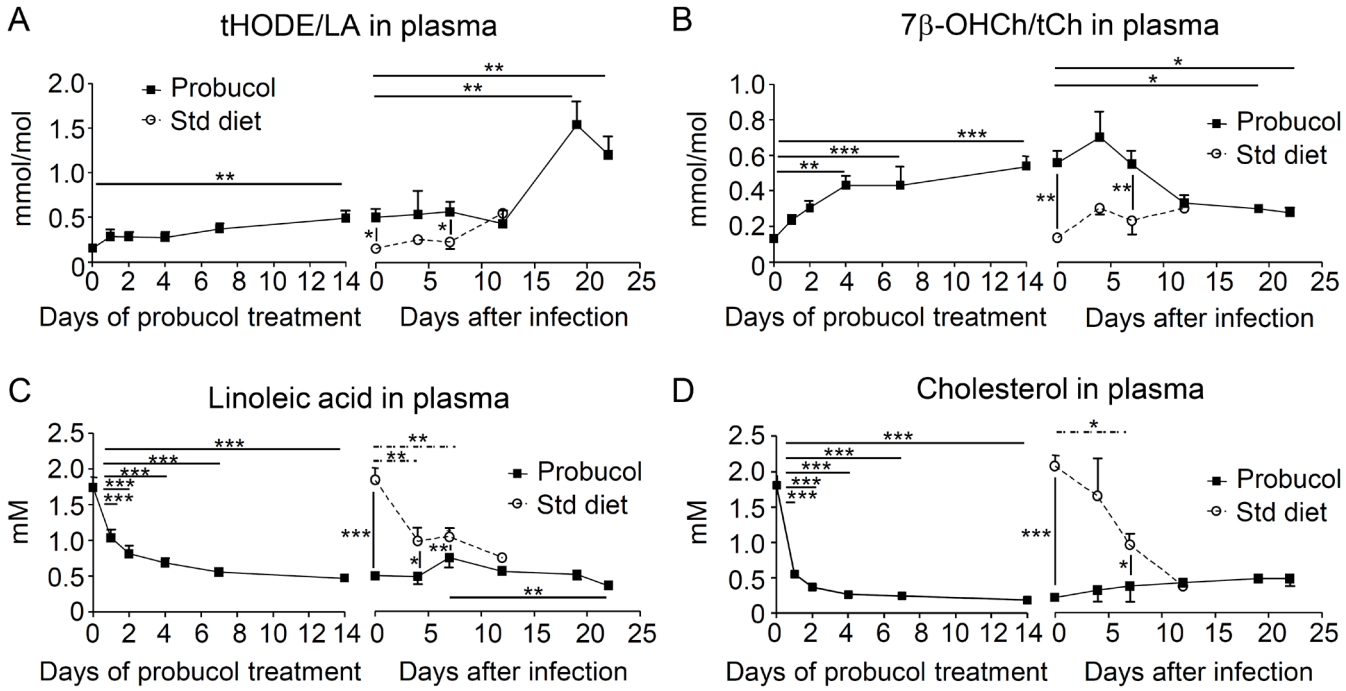


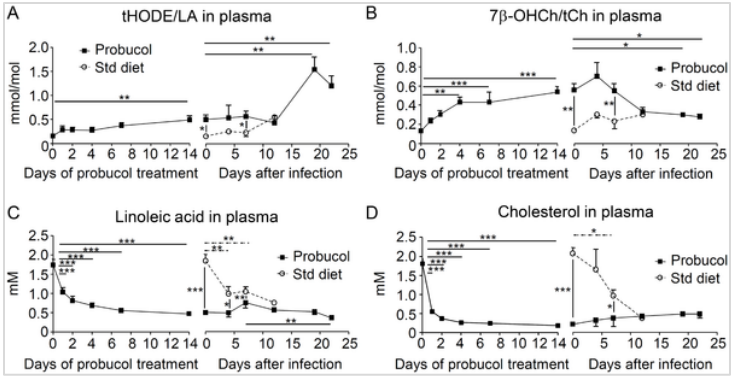
Fig 4. The ratios of lipid peroxidation products to parent lipids in plasma increased after probucol pre-treatment.
Six-week-old C57BL/6J mice were treated with 1% w/w probucol in the diet for 2 weeks and then infected with 0.2 mL of 1 × 10^5 erythrocytes /mL infected with *Plasmodium yoelii* XL-17. Plasma samples were obtained at day 0, 1, 2, 4, 7, and 14 after starting the probucol diet (n = 5 per group) and at day 0, 4, 7, 12, 19, and 22 post-infection (n = 2 to 7). The ratio of total hydroxyoctadecadienoic acid (HODE), a peroxidation product of linoleic acid (LA), to linoleic acid (tHODE/LA) in plasma (A) and the ratio of 7β-hydroxycholesterol (7β-OHCh), a peroxidation product of cholesterol, to total cholesterol (7β-OHCh/tCh) in plasma (B) were measured. The concentration of LA (C) and tCh (D) were measured by using gas chromatography-mass spectrometry (GC-MS). All data are expressed as mean ± SE. Statistical analysis was carried out by analysis of variance (ANOVA). *$p < 0.05$, **$p < 0.025$, and ***$p < 0.001$. The solid bars indicate the significant changes in probucol-treated groups and the dotted bars indicate the significant changes in standard (Std) diet-fed mice.

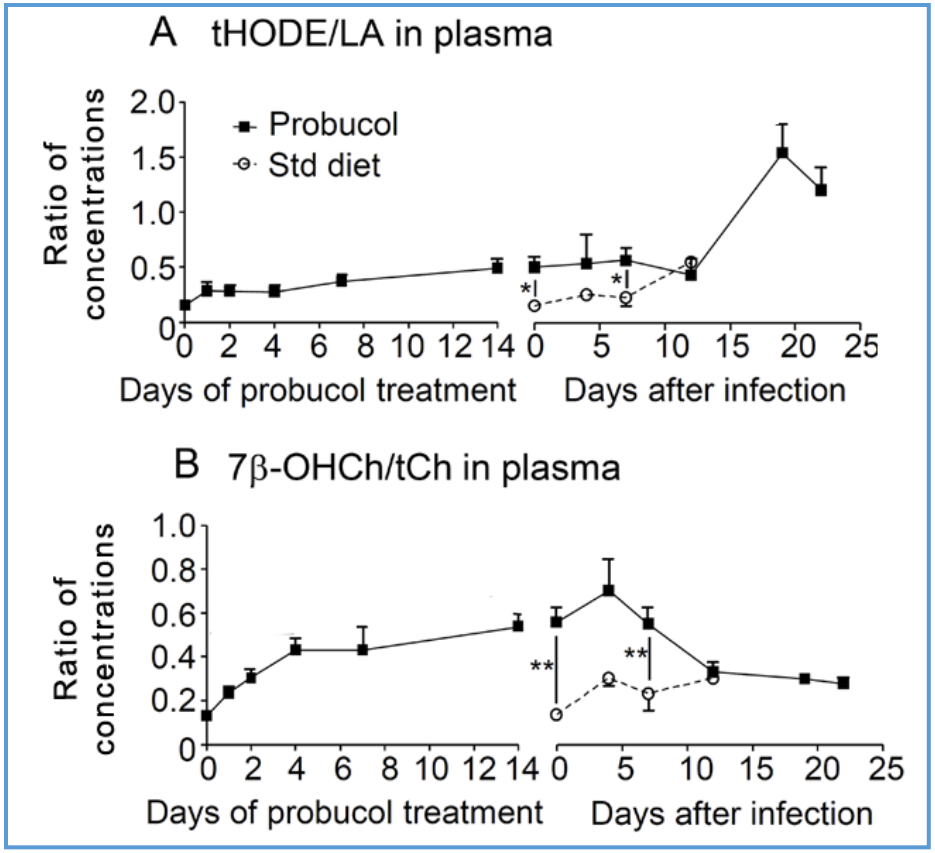
Herbas MS *et al* (2015) **PLoS ONE** 10(8): e0136014.

Evaluating figure choices

- Which data are irrelevant?
- Are there any data/labels missing?
- What could be done to better highlight the most important data?
- Is there a better way to present the data?
- Do the statistics actually add anything here?

Only include the minimum information necessary to draw a conclusion.





Next up: Improve your figure.

Don't get bogged down in the details.
Approximations and sketches are perfect for this exercise.
Try to draft all the parts we have talked about.

Title	Take-home message of the figure. What conclusion should reader evaluate when looking at the figure?
Choice of data	Only data that are critical to the conclusion.
Presentation choices	Type of graph or display, legends & labeling, design choices. Uncluttered; allow quick evaluation of conclusions, without referring to legend or caption.
Caption	Descriptive only, not explanatory/interpretive. Only enough methodological detail to make it clear how results were obtained. "…include only the most relevant aspects of the methods, such as the names of the diagnostic enzymes, a clear description of any normalization or statistics done on the flow cytometry data, etc." *(Mod. 1 Wiki)*

These are our next steps

- This presentation and rubric will be on the wiki for you to reference.
- Put these tips to work on your 109 figures today and all semester
- Tell us: How can these workshops be better? saclarke@mit.edu and bhargavp@mit.edu
- Next week… Abstracts!

Optimize your figures with these reminders

High-level questions

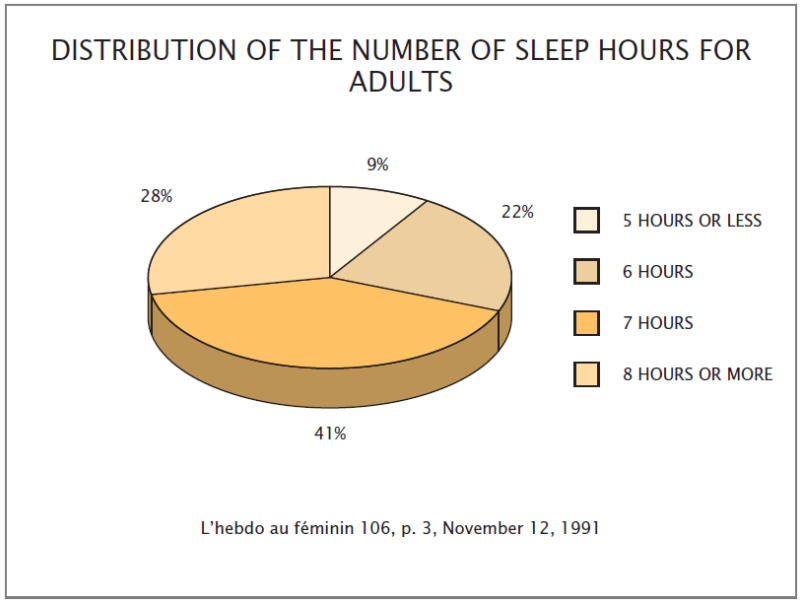
- *Strategic purpose:*
  - What do you want to convey?
  - How will you and/or your audience use this figure?
- *Organizational structure:*
  - Where does this figure fit into the communication?
  - Why?

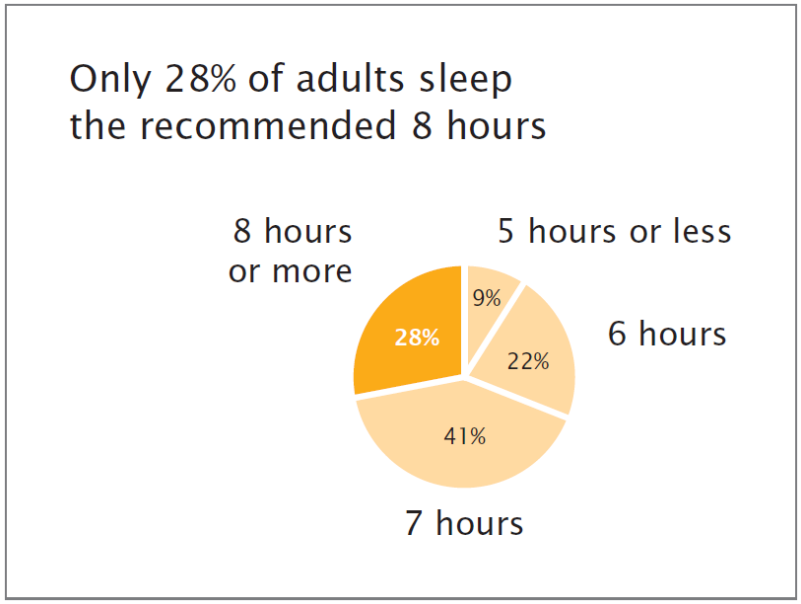
Checklist

- ❑ Choice of data
- ❑ Title/caption
  - Can figure stand alone?
- ❑ Consistent layout
  - Fonts, spacing, colors
- ❑ Text amount and placement
- ❑ Scale, axes, tick marks
- ❑ Error analysis
- ❑ Ink-to-whitespace ratio

Maximize signal-to-noise:

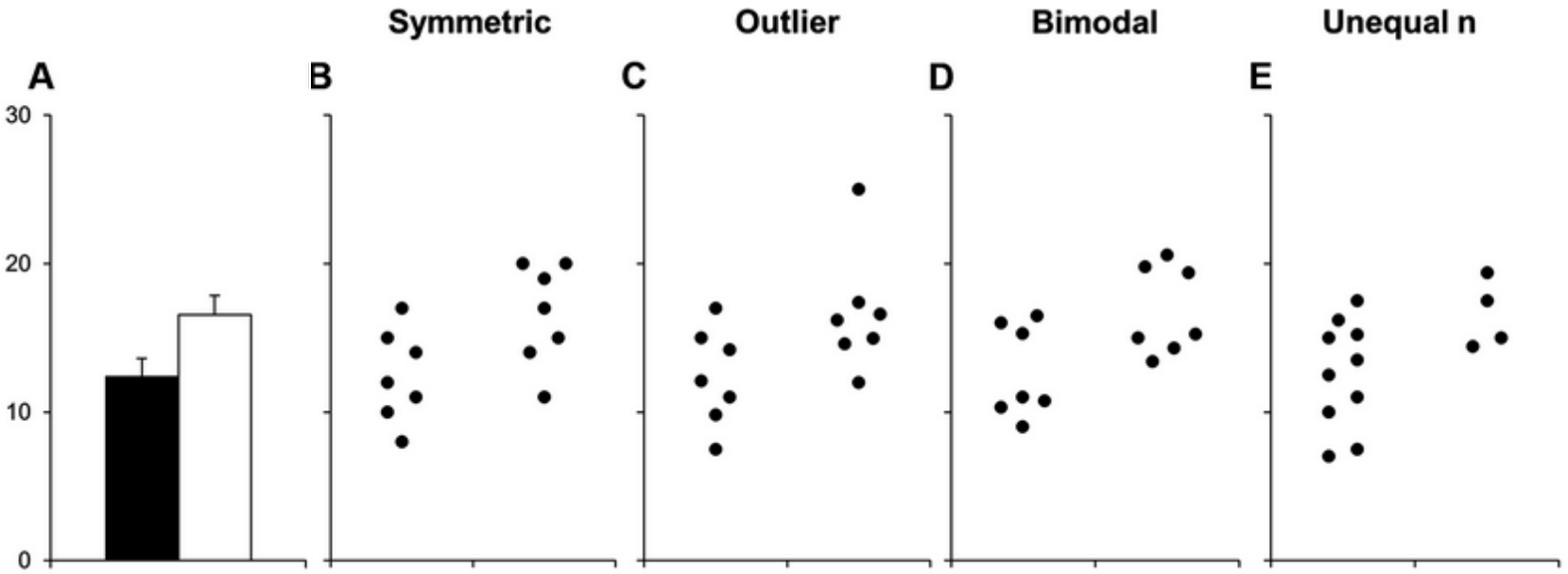
State your message.
Eliminate anything that distracts from it.





Jean-luc Doumont, *Trees, Maps, and Theorems*

Consider which plot type best allows the reader to evaluate your conclusion.



Bar charts assume a Gaussian distribution.

Scatter plots allow reader to evaluate the true distribution.

Weissgerber et al. (2015), Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm. *PloS Biology*