The Capra Lab manuscript template

Your Name Here1 and John A. Capra2,3,\*

1Place, City, XX

2Bakar Computational Health Sciences Institute, University of California, San Francisco, CA

3Department of Epidemiology and Biostatistics, University of California, San Francisco, CA

\*Correspondence to tony@capralab.org

# Abstract

**Background:** (No more than three sentences.) Motivate your work by stating the problem you addressed or the hypothesis you tested. The last sentence should clearly state the gap in knowledge.

**Methods:** (One or two optional sentences; this can be combined with results if it is not a methods focused paper. *To address this gap, we applied X to Y.* or *we synthesized V and W to reveal Z.*

**Results:** (Most of the rest of the abstract) *Here, we show…* What did you do and what did you find? This should summarize the main analyses and results of your work. Keep it high level; it is ok if some results are not mentioned.

**Conclusions:** (No more than two sentences.) What did you learn from your studies? How does this change the field? What do your results say about the problem/hypothesis identified in the Background section?

# Introduction

**First paragraph(s)** What is the problem? Why do we care? First sentence can be a general truth or general limitation that your audience accepts. E.g., *While differences between archaic and modern humans are well described, there is a poor understanding of the genetic mechanisms underlying such phenotypic differences.* or *Most aspects of archaic hominin biology cannot be directly studied due to their lack of preservation in fossils.* Don’t waste the first sentence with a definition.

**Middle paragraph(s)** What have others found about this specific problem? What are the current hypotheses and support for each? and/or what data have been previously analyzed? What are the gaps? E.g., *One such mechanism may be alternative splicing, but we cannot obtain such data from extinct taxa.* This part of the Intro may require a few paragraphs, but you do not need to fully review the history of work on the question. Focus on the current state of the field.

**Last paragraph** Overview of the major questions addressed here and why they are important. *Here, we investigated X using dataset Y and dataset Z. We tested the A hypothesis. We also considered this other obscure hypothesis. We predicted x, y, and z.* E.g., *Here, we leverage a new deep neural network that can predict splice altering variants from sequence alone to examine such variants in archaic hominins.*

# Results

**Optional paragraph 0** Concise overview of what the reader needs to know about the methods or context (but don’t take too long to get to the real results). *The goal of this work is to demonstrate “X”*. If you use a lot of complicated or easily confused terminology, you can define it here to set the stage.

## Declarative result statement

**Context** Set up the context and logic for the experiment/analysis. Clearly state the question and hypothesis. Sometimes this will already have been set up by the previous section and can be skipped.

**Approach** What was the experiment/analysis, briefly (cite methods section if needed)? Describe key factors of the approach, especially details that apply specifically to this result but not to others using the same method (i.e. specific simulation parameters that vary across results sections).

**Result** Clearly state the MAIN result of the experiment/analysis and any relevant statistics. Give a quantification of the effect size, and if you performed a significance test, give the quantification of significance (e.g., p-value or q-value) and test used in parentheses. E.g., *The mean expected 3D divergence is 78% higher than the observed 3D divergence (P* = 1*.*8× 10−48*, t-test)*. Reference relevant figure panels.

**Details** Elaborate on any results that modulate or demonstrate the robustness of your main result. These usually reference supplemental figures that support your main result. **Conclusion** Conservative conclusions drawn from the result. What do these data say about your hypothesis? *This suggests…*

**Example** *Given X, we hypothesized that Tony is great. To evaluate/quantify Tony’s greatness, we did X. We found Y (big picture overview of result) (Main text figure reference). We controlled for Z. Tony’s greatness was robust even when considering W (Supplemental figure(s) reference). This suggests that Tony’s greatness knows no bounds.*

## Example results section to demonstrate figure references and placement

You should insert main text figures and tables into the document as soon as possible after the first mention in the text. We suggest you do this for initial submission even if the journal requests that you put your figures and captions at the end in a separate section. Your reviewers will thank you! Use a “code name” to reference figures while you prepare your manuscript! This will save you a headache later when you have to rearrange and subsequently renumber dozens of figures. At the very end, you can replace all the code names with numbers to look something like the following. You might want to reference a full figure in-text (**Figure 1**) or just a sub-panel (**Figure 1A**). You might also want to reference two figures at the same time like this (**Figures 1**, **S1**). Sometimes you might reference a range of figures like these pair of *Capra* shown in **Figures S1**–**S2**. You can also reference a range of subpanels like **Figures 1A**–**1B**. You can do the same thing for tables (**Table S1**).

**Text

Description automatically generatedFigure 1: Making good figures requires attention to detail.**

**(A)** General principals of making and labelling a good figure. These should help facilitate understanding and reproducibility. **(B)** Some more detailed technical specifics about making figures and their sizes. Inkscape is a great free vector figure editor and alternative to Adobe Illustrator.

## Example results section to demonstrate references

Use a reference manager to cite things parenthetically like this (Einstein, 1905; Kent et al., 2002). You can also cite things in-text like this: Kent et al. (2002) did so many useful things!

# Discussion

**First paragraph** Brief summary of the paper that sets up the discussion of the major contributions. You can start by restating the major gap that you addressed (*Despite X, we previously had a limited understanding of Y. Here, we show/demonstrate Z. Together, these findings highlight/provide…*) This can parallel the last paragraph of your intro (where you outline the big picture questions and their context for the field) or parallel the results, but this is not required.

**Middle Paragraph(s)** (1-4 paragraphs) Major contribution/finding/conclusion and its context. What does your result/contribution mean? Who does it speak to and why do they care? What are the implications? Were your hypotheses correct, surprising, contradictory? If the results are not conclusive, why not? How would you perform further work to get to a more conclusive result?

**Penultimate paragraph(s)** (1-2 paragraphs) Limitations. Identify the broad categories/ themes of limitations and don’t get too particular about the details of each limitation. (i.e., need more samples/ data, model misspecification or over-simplification). Keep it succinct, don’t apologize, but show that you have thought about caveats. Should segue into the future directions. *Although our approach provides many novel benefits, it also has limitations that we hope future work will address.*

**Last paragraph** Future directions and final big picture conclusion. This could be broken out into a separate conclusions section. *As X happens, we anticipate that Y will facilitate understanding of Z which allows for W*. Why is this exciting for your audience? Where are the next steps? Note that this is could possibly result in scooped, but even if you are thinking of working on a next step, consider putting it out there, because science for all!! *We encourage future studies to address X*. Be positive about specific future opportunities inspired by your work. It leaves an impression).

# Methods

## Some overarching part of your results

### Data source 1

Methods about Data source 1 go here.

### **Analysis 1: descriptive title**

Methods about Analysis 1 go here.

## Next methods section

### Next sub-methods section

### Next sub-methods section

## Data availability

The publicly available data used for analysis are available in the following repositories:

The custom datasets we generated are available in the repository “[name here]” available here [link][citation][accession number].

## Code availability

The publicly available code for analysis are available in the following repositories:

The custom code/software we generated are available in the repository “[name here]” available here [link][citation][accession number].

### Acknowledgements

Acknowledge any help you received from people not in your group. List funding sources. Ask collaborators for information they want to include here.

This work was supported by the National Institutes of Health (NIH) General Medical Sciences award [award number] to [initials].

Acknowledge your computing resources (check their resources): “This work was conducted in part using the resources of the Advanced Computing Center for Research and Education (ACCRE) at Vanderbilt University, Nashville, TN.”

## Author Contributions

Summarize the contributions of each author to the project. This section must be approved by all group members. Use the categories, criteria, and style outlined here: [http://www.cell. com/pb/assets/raw/shared/guidelines/CRediT-taxonomy.pdf](http://www.cell.com/pb/assets/raw/shared/guidelines/CRediT-taxonomy.pdf)

## Competing interests

The authors declare no competing interests.

# References

Einstein, Albert (1905). “Zur Elektrodynamik bewegter Körper. (German) [On the electrodynamics of moving bodies]”. In: *Annalen der Physik* 322.10, pp. 891–921. DOI: [http:// dx.doi.org/10.1002/andp.19053221004](https://doi.org/http://dx.doi.org/10.1002/andp.19053221004).

Kent, W. James et al. (June 2002). “The Human Genome Browser at UCSC”. In: *Genome Research* 12.6, pp. 996–1006. ISSN: 1088-9051. DOI: [10.1101/gr.229102](https://doi.org/10.1101/gr.229102).

# Supplementary Information

## Supplemental Text

Some supplemental text that doesn’t fit in the main text or provides extra detail can go here.

## Supplemental Figures



**Figure S1: This is a picture of Tony Capra.** Nobody puts Tony in a supplement…



**Figure S2: This cute sideways goat wishes you good luck with your manuscript.**

Sometimes you might want to rotate figures to fit.

## Supplemental Tables

|  |  |  |  |
| --- | --- | --- | --- |
| Models | Metric 1 |  | Metric 2 |
| precision recall | F-score | R@10 |
| model 1 | 0.67 0.8 | 0.729 | 0.75 |
| model 2 | 0.8 0.9 | 0.847 | 0.85 |

**Table S1: Example table.** Small tables can be here but store big tables in a supplementary (e.g. excel) file.