## **Manuscript Title**

This manuscript (<u>permalink</u>) was automatically generated from <u>malariagen/manubot-test@9d5818c</u> on May 30, 2020.

### **Authors**

- John Doe

Department of Something, University of Whatever  $\cdot$  Funded by Grant XXXXXXXX

- Jane Roe

Department of Something, University of Whatever; Department of Whatever, University of Something

#### **Abstract**

This manuscript is a template (aka "rootstock") for <u>Manubot</u>, a tool for writing scholarly manuscripts. Use this template as a starting point for your manuscript.

The rest of this document is a full list of formatting elements/features supported by Manubot. Compare the input (.md files in the /content directory) to the output you see below.

### **Basic formatting**

**Bold text** 

Semi-bold text

Centered text

Right-aligned text

Italic text

Combined italics and bold

#### Strikethrough

- 1. Ordered list item
- 2. Ordered list item
  - a. Sub-item
  - b. Sub-item
    - i. Sub-sub-item
- 3. Ordered list item
  - a. Sub-item
- List item
- · List item
- · List item

subscript: H<sub>2</sub>O is a liquid

superscript: 2<sup>10</sup> is 1024.

unicode superscripts 0123456789

#### unicode subscripts 0123456789

A long paragraph of text. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Putting each sentence on its own line has numerous benefits with regard to <u>editing</u> and <u>version</u> control.

Line break without starting a new paragraph by putting two spaces at end of line.

### **Document organization**

Document section headings:

# **Heading 1**

### **Heading 2**

**Heading 3** 

**Heading 4** 



#### Horizontal rule:

Heading 1's are recommended to be reserved for the title of the manuscript.

Heading 2's are recommended for broad sections such as Abstract, Methods, Conclusion, etc.

Heading 3's and Heading 4's are recommended for sub-sections.

### Links

Bare URL link: <a href="https://manubot.org">https://manubot.org</a>

<u>Long link with lots of words and stuff and junk and bleep and blah and stuff and other stuff and more stuff yeah</u>

Link with text

Link with hover text

Link by reference

### **Citations**

Citation by DOI.<sup>1</sup>

Citation by PubMed Central ID.<sup>2</sup>

Citation by PubMed ID.<sup>3</sup>

Citation by Wikidata ID.4

Citation by ISBN.<sup>5</sup>

Citation by URL. 6

Citation by alias. <sup>7</sup>

Multiple citations can be put inside the same set of brackets.  $\frac{1}{2}$ , Manubot plugins provide easier, more convenient visualization of and navigation between citations.  $\frac{2}{2}$ ,  $\frac{3}{2}$ ,  $\frac{7}{2}$ ,  $\frac{8}{2}$ 

Citation tags (i.e. aliases) can be defined in their own paragraphs using Markdown's reference link syntax:

### Referencing figures, tables, equations

Figure 1

Figure 2

```
Figure 3

Figure 4

Table 1

Equation 1

Equation 2
```

### **Quotes and code**

Quoted text

Quoted block of text

Two roads diverged in a wood, and I—I took the one less traveled by, And that has made all the difference.

Code in the middle of normal text, aka inline code.

Code block with Python syntax highlighting:

```
from manubot.cite.doi import expand_short_doi

def test_expand_short_doi():
    doi = expand_short_doi("10/c3bp")
    # a string too long to fit within page:
    assert doi == "10.25313/2524-2695-2018-3-vliyanie-enhansera-copia-i-
        insulyatora-gypsy-na-sintez-ernk-modifikatsii-hromatina-i-
        svyazyvanie-insulyatornyh-belkov-vtransfetsirovannyh-geneticheskih-
        konstruktsiyah"
```

Code block with no syntax highlighting:

```
Exporting HTML manuscript
Exporting DOCX manuscript
Exporting PDF manuscript
```

### **Figures**



**Figure 1:** A square image at actual size and with a bottom caption. Loaded from the latest version of image on GitHub.



**Figure 2:** An image too wide to fit within page at full size. Loaded from a specific (hashed) version of the image on GitHub.



Figure 3: A tall image with a specified height. Loaded from a specific (hashed) version of the image on GitHub.



**Figure 4:** A vector .svg image loaded from GitHub. The parameter sanitize=true is necessary to properly load SVGs hosted via GitHub URLs. White background specified to serve as a backdrop for transparent sections of the image.

### **Tables**

**Table 1:** A table with a top caption and specified relative column widths.

Bowling Scores	Jane	John	Alice	Bob
Game 1	150	187	210	105
Game 2	98	202	197	102
Game 3	123	180	238	134

**Table 2:** A table too wide to fit within page.

	Digits	Digits 1-33 Digits 34-66		Digits 1-33 Digits 34-66 Digits 67-99		Ref.
p	011 110	9265358979323 338327950	28841971693993751 0582097494459230	78164062862089986 2803482534211706	piday.org	
е		8182845904523 8747135266	24977572470936999 5957496696762772	40766303535475945 7138217852516642	nasa.gov	

 Table 3: A table with merged cells using the attributes plugin.

	Colors					
Size	Text Color	Background Color				
big	blue	orange				
small	black	white				

### **Equations**

A LaTeX equation:

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2} \tag{1}$$

An equation too long to fit within page:

$$x = a + b + c + d + e + f + g + h + i + j + k + l + m + n + o + p + q + r + s + t + u + v + w + x + y + z + 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9$$
 (2)

### **Special**

▲ WARNING The following features are only supported and intended for .html and .pdf exports. Journals are not likely to support them, and they may not display correctly when converted to other formats such as .docx .

LINK STYLED AS A BUTTON

Adding arbitrary HTML attributes to an element using Pandoc's attribute syntax:

Manubot Manubot Manubot Manubot Manubot. Manubot Manubot Manubot Manubot. Manubot Manubot Manubot. Manubot Manubot. Manubot.

Adding arbitrary HTML attributes to an element with the Manubot attributes plugin (more flexible than Pandoc's method in terms of which elements you can add attributes to):

Manubot Manubo

Available background colors for text, images, code, banners, etc:

white lightgrey grey darkgrey black lightred lightyellow lightgreen lightblue lightpurple red orange yellow green blue purple

Using the **Font Awesome** icon set:



**Light Grey Banner** useful for *general information* - <u>manubot.org</u>

### **1** Blue Banner

useful for important information - manubot.org

**♦ Light Red Banner** useful for *warnings* - <u>manubot.org</u>

### Tables (misc. test)

1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10

### At 6pt by wrapping in <div> with a style:

1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10

## References (misc. test)

URL citation overridden by manual-references.json.

### **Test table from CSV hard-coded into Markdown content**

Using the pantable Pandoc filter.

### Awesome Markdown Table

First row	defaulted to be header row	can be disabled
1	cell can contain <b>markdown</b>	It can be aribrary block element:  • following standard markdown syntax • like this
2	Any markdown syntax, e.g.	$E=mc^2$

### **Test table from CSV file source**

- Using the pantable Pandoc filter with the include option, loading from content/csv/by\_country.csv.
- Alignment has been specified: LRRRRRR.
- 'Samples By Country' is the specified caption.
- dialect: unix is passed to csv.reader() by way of csv-kwargs

### Samples By Country

country	is_arabie nsis	is_coluzz ii	is_gam biae	ye ar	latitude	longitude	contributor
Angola	0	81	0	20 09	-8.8840000000 00011	13.3020000000 00008	Joao Pinto
Burkina Faso	3	135	157	20 14	11.2482567567 56797	-4.2819290540 5405	Nora Besansky
Cameroon	2	26	416	20 13	5.64027477477 4801	13.5362229729 72988	Nora Besansky
Central African Republic	0	18	55	19 94	4.36700000000 0001	18.5830000000 00002	Alessandra della Torre
Cote d'Ivoire	0	80	0	20 12	5.89800000000 00095	-4.8229999999 99997	David Weetman
Democratic Republic of Congo	0	0	76	20 15	4.28299999999 9999	21.0170000000 00035	David Weetman
Equatorial Guinea	0	0	10	20 02	3.7	8.7000000000 0001	Igor Sharakhov
Gabon	0	0	69	20 00	0.38400000000 000023	9.45500000000 0005	Joao Pinto
Gambia, The	0	169	69	20 12	13.5528458781 36198	-15.294989247 311792	Martin Donnelly
Ghana	0	64	36	20 12	5.57405	-0.9347799999 999995	David Weetman
Guinea	0	11	123	20 12	8.84470588235 294	-9.7579411764 70567	Ken Vernick
Guinea-Bissau	0	0	29	20 10	11.9427425742 57442	-15.496168316 831659	Joao Pinto
Kenya	13	0	28	20 12	-3.5110000000 000015	39.9090000000 0008	Janet Midega
Lab Cross	0	0	0	-1	53.4089999999 997	-2.9690000000 000367	Martin Donnelly
Malawi	41	0	0	20 15	-15.932999999 999996	34.7550000000 0003	Martin Donnelly
Mali	2	91	131	20 14	12.0143111111 1115	-7.8246222222 22243	Nora Besansky
Mayotte	0	0	23	20 11	-12.857086956 52174	45.1374347826 08686	Igor Sharakhov
Mozambique	0	0	74	20 04	-23.715999999 99996	35.2989999999 99985	Joao Pinto

country	is_arabie nsis	is_coluzz ii	is_gam biae	ye ar	latitude	longitude	contributor
Tanzania	225	0	68	20 15	-2.5083233333 33342	33.8503566666 6663	David Weetman
Uganda	82	0	207	20 12	0.26649655172 41366	32.5942758620 6877	Martin Donnelly

- See that column headings are carried onto the next page.
- <u>Pantable</u> uses [Panflute], e.g. https://github.com/ickc/pantable/blob/master/pantable/csv\_to\_table\_markdown.py
- If we needed to apply custom formatting during the build, e.g. to round those floats, we could probably write our own Pandoc filter.

### **Test figures from image sources**

• May need to use secno=1 on all figures, e.g. {#fig:id secno=1} to avoid mislabelling/miscounting bug? https://github.com/tomduck/pandoc-fignos/issues/76

### 1) Separate figures with alt text ("captions"), no resizing, refs linked by default

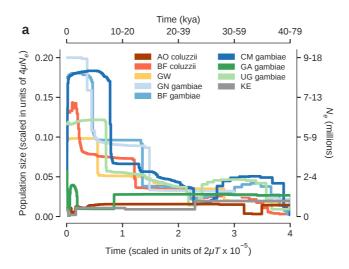


Figure 5: Stairway Plot

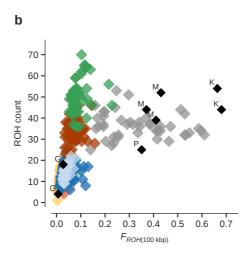


Figure 6: ROH Scatter Plot

- 5, Stairway plot of changes in population size over time. Absolute values of time and Ne are shown on alternative axes as a range of values, assuming lower and upper limits for the mutation rate  $\mu$  as 2.8 × 10–9 and 5.5 × 10–9, respectively, and t = 11 generations per year. ka, thousand years ago. 6, Runs of homozygosity (ROH) in individual mosquitoes, highlighting recent inbreeding in Kenyan (grey) and colony (black) mosquitoes. G, Ghana; K, Kisumu; M, Mali; P, Pimperena.
- The "Figure {n}:" label before the alt text "captions", e.g. Figure 6s: ROH Scatter Plot, is added automatically.
- Btw, {@fig:figure-test-1b} appears as {???}: when in "quotes", e.g. "{???}", which is unexpected, possibly a bug
- Curly brackets are to separate fig ref from adjacent non-whitespace strings, e.g. @fig:figure-test-1b}-

### 2) Separate figures with no alt text (no "captions"), no resizing, refs linked by default

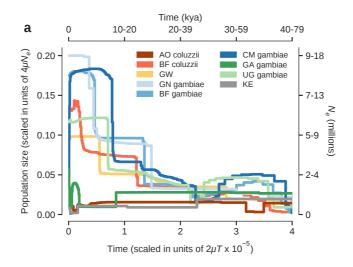


Figure 7:

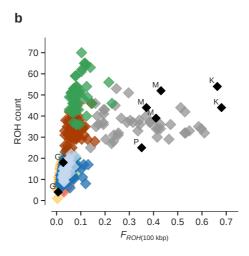


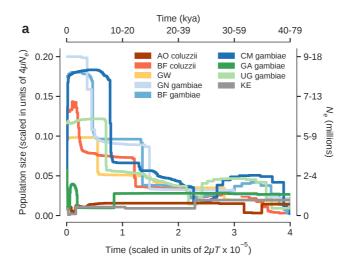
Figure 8:

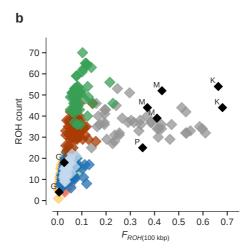
**7**, Stairway plot of changes in population size over time. Absolute values of time and Ne are shown on alternative axes as a range of values, assuming lower and upper limits for the mutation rate  $\mu$  as 2.8 × 10–9 and 5.5 × 10–9, respectively, and t = 11 generations per year. ka, thousand years ago. **8**, Runs of homozygosity (ROH) in individual mosquitoes, highlighting recent inbreeding in Kenyan (grey) and colony (black) mosquitoes. G, Ghana; K, Kisumu; M, Mali; P, Pimperena.

- The "Figure {n}:" label before the empty alt text "captions", e.g. Figure 8:, is added automatically.
  Want a way to override this label's template
- The colon after the Figure 8:, which anticipates text, looks odd/confusing when there is no text. It looks like it applies to the content that follows after that label, but it doesn't/shouldn't.

### 3) Separate subfigures using Pandoc "divs", with no alt text, no layout, no resizing

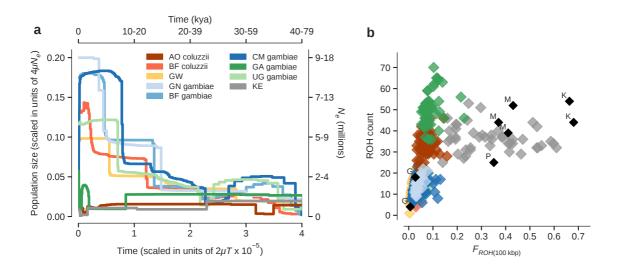
Figure 1





- **a**, Stairway plot of changes in population size over time. Absolute values of time and Ne are shown on alternative axes as a range of values, assuming lower and upper limits for the mutation rate  $\mu$  as 2.8 × 10–9 and 5.5 × 10–9, respectively, and t = 11 generations per year. ka, thousand years ago. **b**, Runs of homozygosity (ROH) in individual mosquitoes, highlighting recent inbreeding in Kenyan (grey) and colony (black) mosquitoes. G, Ghana; K, Kisumu; M, Mali; P, Pimperena.
- Pandoc divs let you group content under one Manubot figure reference, in this case two simple images (not Manubot figures in their own right)
- Putting the Markdown for the figures on separate paragraphs (as required by Pandoc) causes a vertical layout by default, which is undesirable here
- Figure references for **1a** and **1b** are just references/links to **1** plus manually managed suffixes
- Can switch off the link for figure references, e.g. @fig:figure-test3{nolink=True} gives 1 rather than 1
- The figure label was added in manually, i.e. \*\*Figure @fig:figure-test3{nolink=True}\*\*,
   otherwise there would be no visible label for this figure, even though it has a reference
- Note figure-test3 has reset to "1", rather than 8 + 1. Bug? https://github.com/tomduck/pandoc-fignos/issues/76
  - Using secno=1 here, i.e. {#fig:figure-test3 secno=1} has made no difference, but perhaps need to use secno=1 on **all** figures in order to workaround. (?!)

### 4) Breaking of pandoc-fignos when figures are not in their own Markdown paragraph



??, Stairway plot of changes in population size over time. Absolute values of time and Ne are shown on alternative axes as a range of values, assuming lower and upper limits for the mutation rate  $\mu$  as  $2.8 \times 10-9$  and  $5.5 \times 10-9$ , respectively, and t = 11 generations per year. ka, thousand years ago. ??? {nolink=True}, Runs of homozygosity (ROH) in individual mosquitoes, highlighting recent inbreeding in Kenyan (grey) and colony (black) mosquitoes. G, Ghana; K, Kisumu; M, Mali; P, Pimperena.

- See broken refs as "??" and ???{nolink=True}.
- · Reports errors:
  - o pandoc-fignos: Bad reference: @fig:figure-test4a.
  - Generating csl\_item for 'figure-test3b' failed due to a NotImplementedError: Manubot does not know how to generate a csl\_item for 'figure-test3b'

### Pandoc requires:

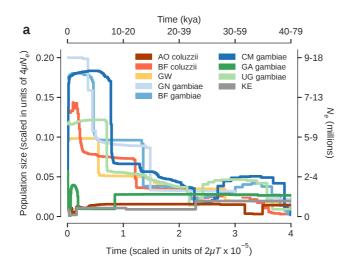
```
![Stairway Plot](images/stairway_plot.svg){#fig:figure-test4a}
![ROH Scatter Plot](images/roh_scatter.svg){#fig:figure-test4b}
```

#### Instead of the above:

```
![Stairway Plot](images/stairway_plot.svg){#fig:figure-test4a}
![ROH Scatter Plot](images/roh_scatter.svg){#fig:figure-test4b}
```

 This constraint prevents figures from being placed next to eachother by a simple inline method (as above) without breaking the managed references.

### 5) Separate figures with HTML table-based layout, with tags, no alt text, no resizing



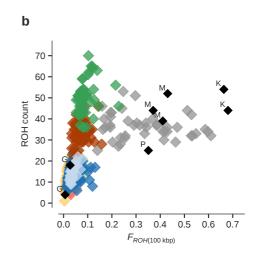
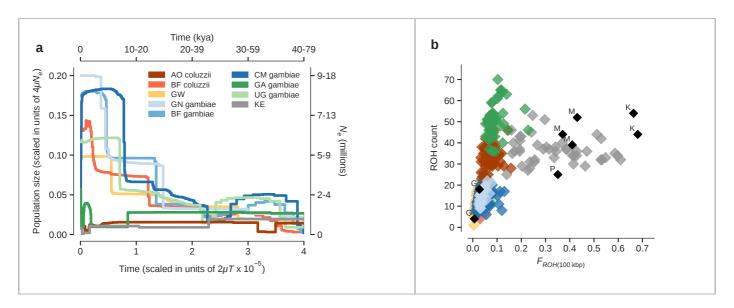


Figure a: Figure b:

**a**, Stairway plot of changes in population size over time. Absolute values of time and Ne are shown on alternative axes as a range of values, assuming lower and upper limits for the mutation rate  $\mu$  as 2.8 × 10–9 and 5.5 × 10–9, respectively, and t = 11 generations per year. ka, thousand years ago. **b**, Runs of homozygosity (ROH) in individual mosquitoes, highlighting recent inbreeding in Kenyan (grey) and colony (black) mosquitoes. G, Ghana; K, Kisumu; M, Mali; P, Pimperena.

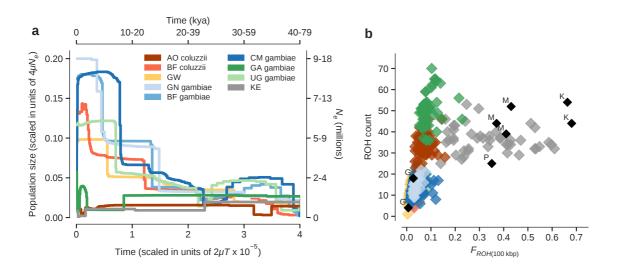
- If the HTML contains any preceding whitespace in the markup, e.g. indentation, you will see broken HTML markup
- HTML table-based layout could be used to place the figures next to each other while still meeting the Pandoc separate-paragraph rule, but it's a bit of a hack
- Table cell borders removed using
- Manubot shows a figure label for each figure, even when no alt/caption text. Here the figure label/number is overidden usign specified "tags"
- Tags can be used to override the automatic managed figure number, e.g. here <code>@fig:figure-test5a</code> is labelled as "a" instead of an automatically incremented figure number. In other words, normally the figure label would be a number, which would be incremented automatically, except for bug/feature: https://github.com/tomduck/pandoc-fignos/issues/76, but here we have overridden that figure label.
- Again, the colon after the figure label, which anticipates text, looks odd/confusing when there is no text. It looks like it applies to the content that follows after that label, but it doesn't/shouldn't.

# 6) Breaking of separate figures with Markdown table-based layout, no alt text, no resizing



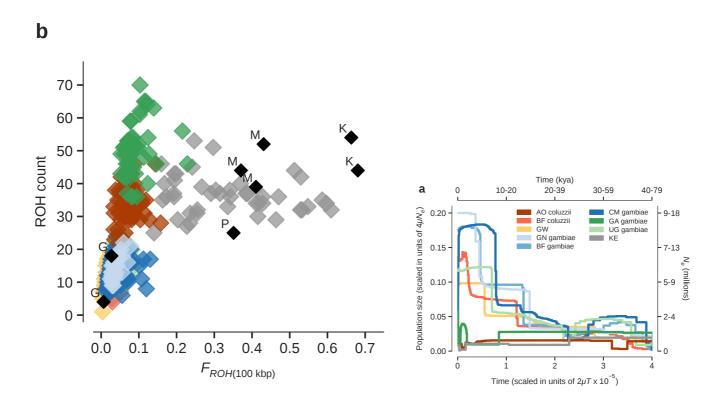
- ??, Stairway plot of changes in population size over time. Absolute values of time and Ne are shown on alternative axes as a range of values, assuming lower and upper limits for the mutation rate  $\mu$  as  $2.8 \times 10-9$  and  $5.5 \times 10-9$ , respectively, and t = 11 generations per year. ka, thousand years ago. ??, Runs of homozygosity (ROH) in individual mosquitoes, highlighting recent inbreeding in Kenyan (grey) and colony (black) mosquitoes. G, Ghana; K, Kisumu; M, Mali; P, Pimperena.
- See broken figure references as "??", and no figure labels under the figures, compared to (5)
  - Generates errors pandoc-fignos: Bad reference: @fig:figure-test6a. and pandoc-fignos: Bad reference: @fig:figure-test6b.
- Want to remove visible table cell borders
- Markdown table-based layout cannot be used to place the figures next to each other without breaking the Pandoc separate-paragraph rule

### 7) Separate figures using images inline (not Manubot figures), no resizing



- **a**, Stairway plot of changes in population size over time. Absolute values of time and Ne are shown on alternative axes as a range of values, assuming lower and upper limits for the mutation rate  $\mu$  as 2.8 × 10–9 and 5.5 × 10–9, respectively, and t = 11 generations per year. ka, thousand years ago. **b**, Runs of homozygosity (ROH) in individual mosquitoes, highlighting recent inbreeding in Kenyan (grey) and colony (black) mosquitoes. G, Ghana; K, Kisumu; M, Mali; P, Pimperena.
- images can be used to simulate figures, at the cost of all the figure management features of Manubot (assuming the features worked as intended)
- references for figures would have to be managed manually. A figure reference has not been added here, and would have to be added and maintained manually (outside of Manubot's system, by hard-coding the figure identifier/label, e.g. Figure 1, in the Markdown)

# 8) Separate figures using images inline (not Manubot figures), resized to 3in and 4in width, rearranged



- **a**, Stairway plot of changes in population size over time. Absolute values of time and Ne are shown on alternative axes as a range of values, assuming lower and upper limits for the mutation rate  $\mu$  as 2.8 × 10–9 and 5.5 × 10–9, respectively, and t = 11 generations per year. ka, thousand years ago. **b**, Runs of homozygosity (ROH) in individual mosquitoes, highlighting recent inbreeding in Kenyan (grey) and colony (black) mosquitoes. G, Ghana; K, Kisumu; M, Mali; P, Pimperena.
- Labels a and b in the description are manually coded (these are not Manubot-managed figures, only Markdown images)
- See that subfigure **b** appears before (left of) subfigure **a**, intentionally swapped around
- See that the **b** label is larger than the **a** label, because the whole **b** image has been made larger than **a**, intentionally
- embedding subfigure labels in the image means that image resizing and rearranging is frustrated
   new images would need to be generated with new labels, or the SVG code modified by hand

# 9) Separate figures using images (not Manubot figures) with HTML table-based layout, resized to 3in and 4in width

If the labels were not embedded in the image, the labels would naturally need be added to the Markdown in some way, perhaps as figure "captions".

Imagine that the two images below did not contain their **a** and **b** labels in their top-left corners:

Figure 42

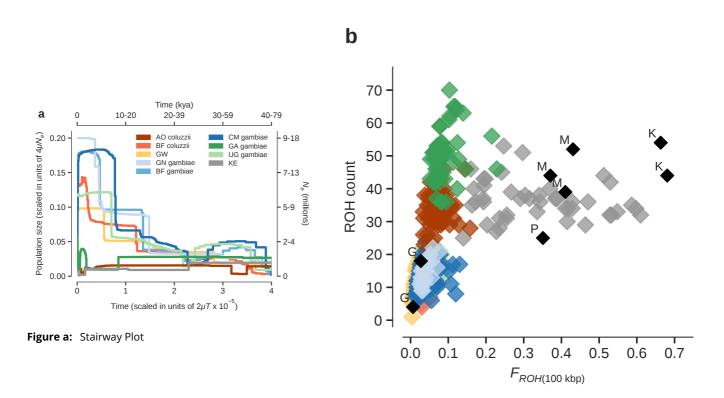


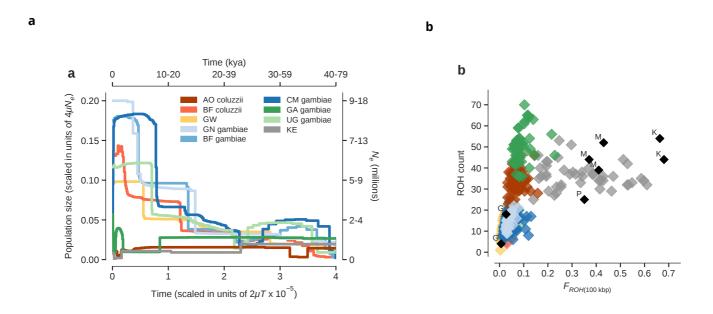
Figure b: ROH Scatter Plot

- **a**, Stairway plot of changes in population size over time. Absolute values of time and Ne are shown on alternative axes as a range of values, assuming lower and upper limits for the mutation rate  $\mu$  as 2.8 × 10–9 and 5.5 × 10–9, respectively, and t = 11 generations per year. ka, thousand years ago. **b**, Runs of homozygosity (ROH) in individual mosquitoes, highlighting recent inbreeding in Kenyan (grey) and colony (black) mosquitoes. G, Ghana; K, Kisumu; M, Mali; P, Pimperena.
- This is similar to (1), except with the layout approach of (5)
- The **a** and **b** in the descriptions are taken from the Manubot figure tag labels, specified in Markdown.
- The label **Figure 42** was added manually. Only the subfigure identifiers/labels are managed by Manubot. There is no "Manubot figure 42", so that label would have to be managed/maintained manually.

### 10) Single figure with subfigures as labelled images in a HTML table-based layout

Imagine that the following two images do not contain labels **a** and **b** in their top-left corners:

Figure 1



**1a**, Stairway plot of changes in population size over time. Absolute values of time and Ne are shown on alternative axes as a range of values, assuming lower and upper limits for the mutation rate  $\mu$  as  $2.8 \times 10-9$  and  $5.5 \times 10-9$ , respectively, and t = 11 generations per year. ka, thousand years ago. **1b**, Runs of homozygosity (ROH) in individual mosquitoes, highlighting recent inbreeding in Kenyan (grey) and colony (black) mosquitoes. G, Ghana; K, Kisumu; M, Mali; P, Pimperena.

- This is similar to (3), except with the layout approach of (5)
- The label Figure 1 was added manually, but the 1 part is automatic.
- The subfigure labels **a** and **b** outside the image were added manually.
- In the description, the 1 part (of 1a and 1b) is automatic, whereas the "a" and "b" parts are added manually, i.e. hard-coded in the Markdown.

### References

- 1. Himmelstein, D. S. et al. Sci-Hub provides access to nearly all scholarly literature. eLife 7, (2018).
- 2. Beaulieu-Jones, B. K. & Greene, C. S. Reproducibility of computational workflows is automated using continuous analysis. *Nat Biotechnol* **35**, 342–346 (2017).
- 3. Heaven, D. Bitcoin for the biological literature. *Nature* **566**, 141–142 (2019).
- 4. S, c. Plan S: Accelerating the transition to full and immediate Open Access to scientific publications. (2018).
- 5. Suber, P. Open access. (MIT Press, 2012).
- 6. Himmelstein, D. S. *et al. Open collaborative writing with Manubot*. <a href="https://greenelab.github.io/meta-review/">https://greenelab.github.io/meta-review/</a> (2020).
- 7. Ching, T. *et al.* Opportunities and obstacles for deep learning in biology and medicine. *Journal of The Royal Society Interface* **15**, 20170387 (2018).
- 8. Himmelstein, D. S. *et al.* Open collaborative writing with Manubot. *PLOS Computational Biology* **15**, e1007128 (2019).
- 9. Himmelstein, D. manubot/rootstock GitHub repository. *GitHub* <a href="https://github.com/manubot/rootstock">https://github.com/manubot/rootstock</a> (2019).