Errata *Bioinformatics for Evolutionary Biologists. A Problems Approach*, 2nd Edition

Bernhard Haubold & Angelika Börsch-Haubold

February 16, 2024

1. p. 14, Problem 1.60: The URL `sn.pub/dy6S42` is not resolved by `git`. So the line

   ```
   git clone sn.pub/dy6S42
   ```

   should be

   ```
   https://github.com/sn-code-inside/biobox.git
   ```

2. p. 41, Problem 2.54: "were we invest" → where we invest

3. p. 78, Problem 3.66: "of" → to

4. p. 83, Problem 3.88: Add "(We count the string itself as the first rotation.)"

5. p. 88, Problem 3.110: A byte occupies eight bits, not four.

6. p. 91, Problem 3.121: "plot the Huffman tree for the proteome" → plot the Huffman tree for the concatenated sequences of the proteome

7. p. 97, Problem 4.16: "OH157H7" → O157H7

8. p. 107, Problem 4.57: "prgram" → program

9. p. 222, Answer 1.114: The expression

$$580076/540447 \times 100 \approx 93\%$$

   should be
$$540447/580076 \times 100 \approx 93\%$$

10. p. 224, Answer 1.124: "8289 lines" → 8288 lines

11. p. 232, Answer 2.14: "$16 \times 3 + 2 = 51$" → $17 \times 3 = 51$

12. p. 239, Answer 2.45: What we give here are just the *ends* of the three possible alignments. The actual alignments are

    ```
    -A          A-          A
    T-          -T          T
    ```

13. p. 244, Answer 2.62: "Notice the two backslashes in the `plotSeg` command, which continue the lines across the carriage returns." This sentence should move to the first occurrence of the backslash notation in Answer 1.109 on p. 220.

14. p. 245, Answer 2.66: The result

```
0a1
> 99   -1836  88   -1847
```

should be

```
0a1
> 99    1836  88    1847
```

15. p. 248, Answer 2.83: The matrix

	–	A	C	G
–	0	$\leftarrow -1$	$\leftarrow -2$	$\leftarrow -3$
A	$\uparrow -1$	$\nwarrow 1$	$\leftarrow 0$	$\leftarrow -1$
G	$\uparrow -2$	$\uparrow 0$	$\uparrow -1$	$\nwarrow 1$

should be

	–	A	C	G
–	0	$\leftarrow -1$	$\leftarrow -2$	$\leftarrow -3$
A	$\uparrow -1$	$\nwarrow 1$	$\leftarrow 0$	$\leftarrow -1$
G	$\uparrow -2$	$\uparrow 0$	$\nwarrow 0$	$\nwarrow 1$

16. p. 248, Answer 2.84: The matrix

```
     -   A   C   G
 -   0 <-1 <-2 <-3
 A ^-1  \1  <0 <-1
 G ^-2  ^0 ^-1  \1
```

should be

```
     -   A   C   G
 -   0 <-1 <-2 <-3
 A ^-1  \1  <0 <-1
 G ^-2  ^0  \0  \1
```

17. p. 280, Answer 3.109: The alphabet in the decoding command should be `CG` rather than `GC`.

18. p. 304, Answer 7.17: "a_1" $\rightarrow m_1$ and "a_2" $\rightarrow m_2$

19. p. 213, Answer 1.69: "We've extracted 47 data files, of which 27 are FASTA files." $\rightarrow$ We've extracted 35 data files, of which 22 are FASTA files.