

## First model

We make the following assumptions

1. Complete divergence at divergently selected sites (as in Aeschbacher *et al.* (2017)).
2. The number of selected sites is  $\nu$ .
3. The number of selected sites in window  $i$  is a random variable  $X_i$  which is Poisson distributed with mean  $\nu L_i / L$ , where  $L_i$  is the map length of the  $i$ th window, and  $L$  is the total map length of the genomic element.
4. Selected loci are equally spaced within a window.
5. The recombination rate between selected loci in window  $j$  and window  $i$  where  $j \neq i$  is equal to the recombination rate between the midpoints of window  $i$  and  $j$ ,  $\bar{r}_{ij}$ .
6. All selected loci have the same selection coefficient.

With these assumptions, we approximate the gff in window  $i$  by

$$g_i = g_{ii} \prod_{j \neq i} g_{ij}$$

where

$$g_{ii} = \frac{1}{L_i} \int_0^{L_i} \exp \left( - \sum_{j=1}^{X_i} \frac{s}{s + m + r(x_j, x)} \right) dx$$

and

$$g_{ij} = \exp \left( - \frac{s X_j}{s + m + \bar{r}_{ij}} \right)$$

Treating the gIMble-inferred effective migration rates as observed data ( $y$ ), we have the following generative model for the ‘data’

$$\begin{aligned} \nu &\sim \Gamma(\alpha_1, \theta_1) \\ m &\sim \Gamma(\alpha_2, \theta_2) \\ s &\sim \Gamma(\alpha_3, \theta_3) \\ X_i | \nu &\sim \text{Poisson}(\nu L_i / L) \\ \log y_i | X, s, m &\sim \text{Normal}(\log mg_i, \sigma) \end{aligned}$$

where  $\sigma$  is a tuning parameter determining how close we fit the ‘observed’  $m_e$  profile.

We devise a sampling scheme which makes use of the conjugacy of the Poisson and Gamma distributions. Specifically, we use the following Gibbs sampler

$$\begin{aligned} s | m, \nu, X, y \\ m | s, \nu, X, y \\ \nu | X, y \\ X_i | \nu, s, m, y \end{aligned}$$

where for the first two sampling steps we use a Metropolis-Hastings update, for the third we sample from the analytically available posterior (exploiting conjugacy) and for the fourth we

calculate the posterior probabilities for  $X_i = \{0, 1, \dots, X_{\max}\}$ , where  $X_{\max}$  is chosen so that  $\Pr\{X_i > X_{\max} | \nu, s, m\} < \epsilon$  for some suitably chosen  $\epsilon$ .

This sampler is reasonably efficient. Instead of fitting the gIMble inferred  $m_e$  profile, one could fit  $F_{ST}$ , but this would not take any variation in  $N_e$  into account and would rely heavily on equilibrium assumptions.

Fig. 1 and fig. 2 show an example fit to chromosome 18 of *Heliconius* (data from the gIMble paper), assuming  $\sigma = 0.5$ . The inferred selection density per basepair is very similar to the one inferred in Laetsch *et al.* (2023) under the Aeschbacher model (they inferred  $\nu s = 2.77 \times 10^{-9}$  (95% CI:  $2.44 \times 10^{-9}, 3.09 \times 10^{-9}$ ). Very roughly, we estimate that, under this model, this selection density corresponds to about 10 to 100 selected sites with a selection coefficient of about 0.1%. Note that for the estimated values of  $N_e$  (roughly  $10^6$ ), this would correspond to very strong selection (i.e.  $100 < N_e s < 10000$ ), so that assuming complete divergence at selected sites is probably reasonable.

If we increase  $\sigma$ , this amounts to treating the ‘observed’  $m_e$ -profile from gIMble as more noisy. When we do so, we infer a lower rate of migration, and lower selection density. What is inferred in the  $\sigma = \sqrt{1/4}$  analysis as a reduction in gene flow due to selection is now inferred (in the  $\sigma = \sqrt{1/2}$  analysis) as a reduction in gene flow due to reduced migration. The joint posterior distribution for  $\log m$  and  $\log s$  has a very marked ‘boomerang’ shape: with this assumed  $\sigma$  the data is compatible both with any strength of selection. The opposite holds for the analysis with  $\sigma = \sqrt{1/8}$ .

**Table 1:** Summary of the posterior distribution for the analysis of the first model with  $\sigma = \sqrt{1/4}$ .

| parameter                    | mean     | 2.5%     | 97.5%    |
|------------------------------|----------|----------|----------|
| $m$                          | 1.70e-06 | 1.04e-06 | 3.09e-06 |
| $s$                          | 1.43e-03 | 4.74e-04 | 2.93e-03 |
| $\nu$                        | 37.5     | 10.3     | 96.6     |
| $\nu \times s$ [bp $^{-1}$ ] | 2.64e-09 | 1.07e-09 | 5.04e-09 |

**Table 2:** Results for  $\sigma = \sqrt{1/2}$ .

| parameter                    | mean     | 2.5%     | 97.5%    |
|------------------------------|----------|----------|----------|
| $m$                          | 1.10e-06 | 6.64e-07 | 2.02e-06 |
| $s$                          | 1.12e-03 | 1.48e-13 | 3.46e-03 |
| $\nu$                        | 25.4     | 0.8      | 92.2     |
| $\nu \times s$ [bp $^{-1}$ ] | 1.24e-09 | 3.51e-20 | 3.49e-09 |

**Table 3:** Results for  $\sigma = \sqrt{1/8}$ .

| parameter | mean     | 2.5%     | 97.5%    |
|-----------|----------|----------|----------|
| $m$       | 3.32e-06 | 1.81e-06 | 6.50e-06 |

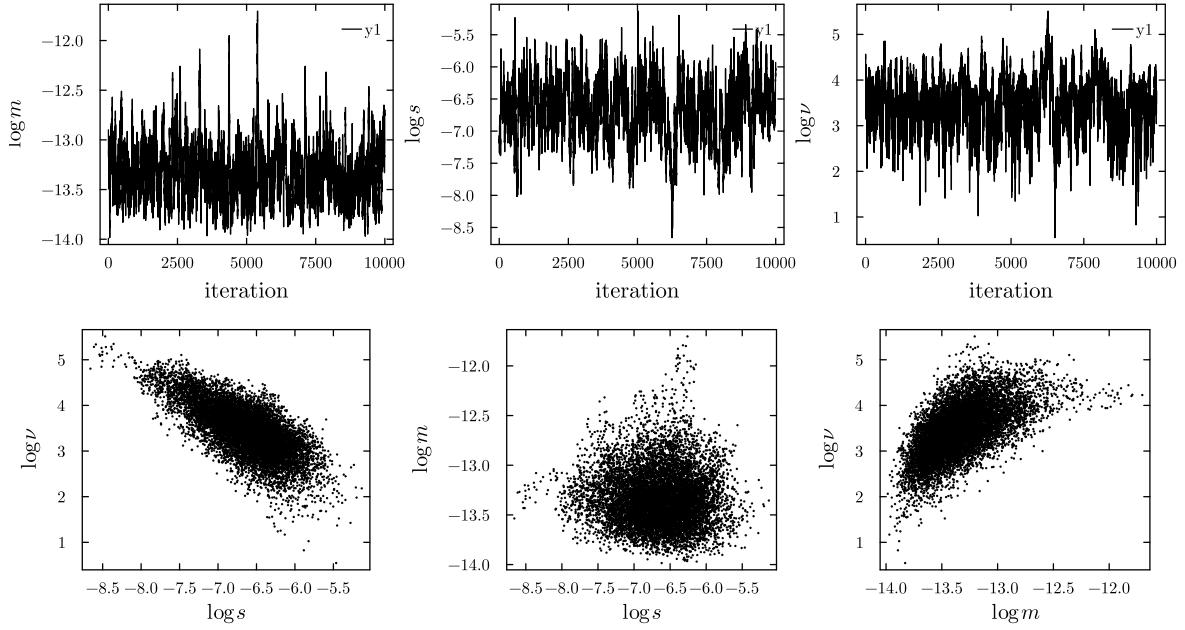
| parameter                    | mean     | 2.5%     | 97.5%    |
|------------------------------|----------|----------|----------|
| $s$                          | 1.03e-03 | 4.79e-04 | 1.97e-03 |
| $\nu$                        | 83.6     | 34.4     | 178      |
| $\nu \times s$ [bp $^{-1}$ ] | 4.52e-09 | 2.63e-09 | 6.95e-09 |

## Notes

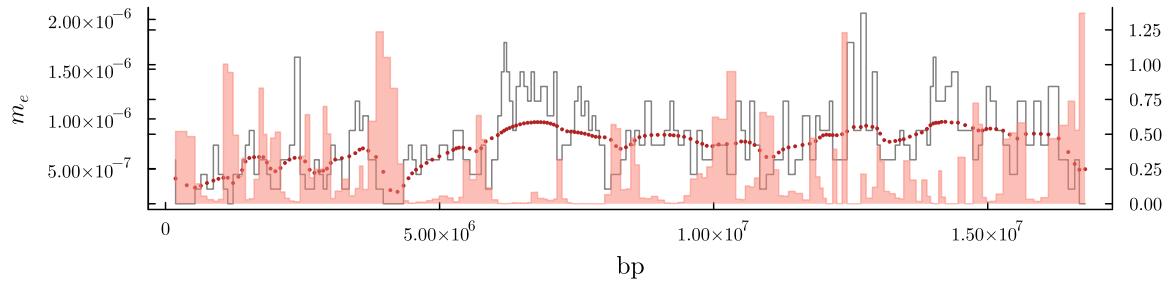
1. Taking  $\nu = 37$  and  $s = 1.4 \times 10^{-3}$  at face value, this would imply that, when the loci are equally spaced along the entire chromosome,  $r/s \approx 9$ , which is fine. However, when they are uniformly scattered along the chromosome, the expected  $r/s$  between neighboring sites is similar, but 5% of the pairs of loci have  $r/s < 0.7$  or so, which should be considered as tight linkage (tight coupling). Considering the non-uniform distribution of selected sites across the chromosome, this suggests we will have tightly coupled sites for which the theory breaks down. Note also that these parameters imply  $Ls \approx 0.05$  (for chromosome 18), i.e. very weak divergence.
2. For the *Heliconius* data, we won't gain anything from considering the partial divergence model etc. See fig. 7. The actual situation is even less distinguishable, because  $N_e$  is likely about a million (instead of 200.000).

## References

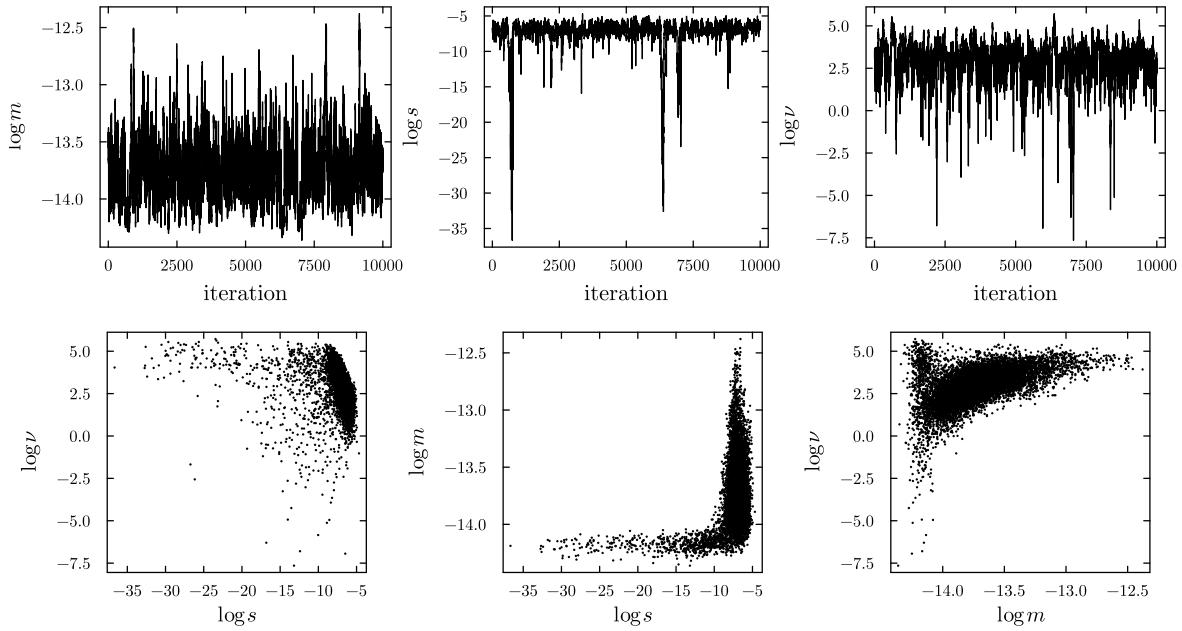
- Aeschbacher S., and R. Bürger, 2014 The effect of linkage on establishment and survival of locally beneficial mutations. *Genetics* 197: 317–336.
- Aeschbacher S., J. P. Selby, J. H. Willis, and G. Coop, 2017 Population-genomic inference of the strength and timing of selection against gene flow. *Proceedings of the National Academy of Sciences* 114: 7061–7066.
- Laetsch D. R., G. Bisschop, S. H. Martin, S. Aeschbacher, D. Setter, *et al.*, 2023 Demographically explicit scans for barriers to gene flow using gIMble. *PLoS genetics* 19: e1010999.
- Zwaenepoel A., H. Sachdeva, and C. Fraïsse, 2024 The genetic architecture of polygenic local adaptation and its role in shaping barriers to gene flow. *Genetics* iyae140.



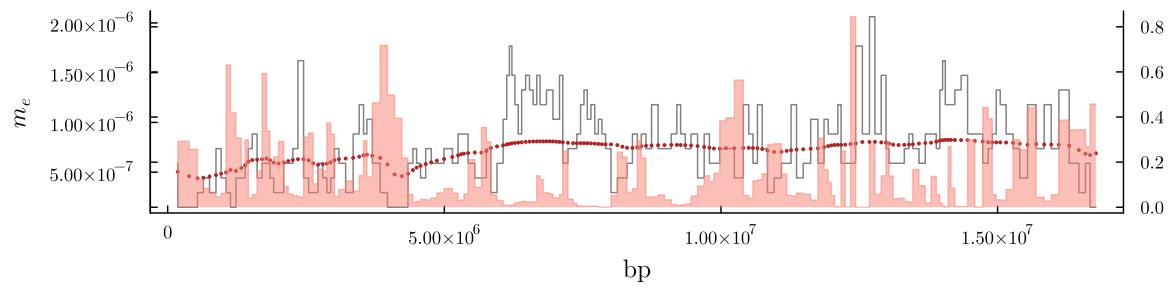
**Figure 1:** Sampler and joint posterior distribution for  $\sigma = 0.5$ . The sampler was run for 110,000 iterations, discarding the first 10,000 and keeping every 10th iterate.



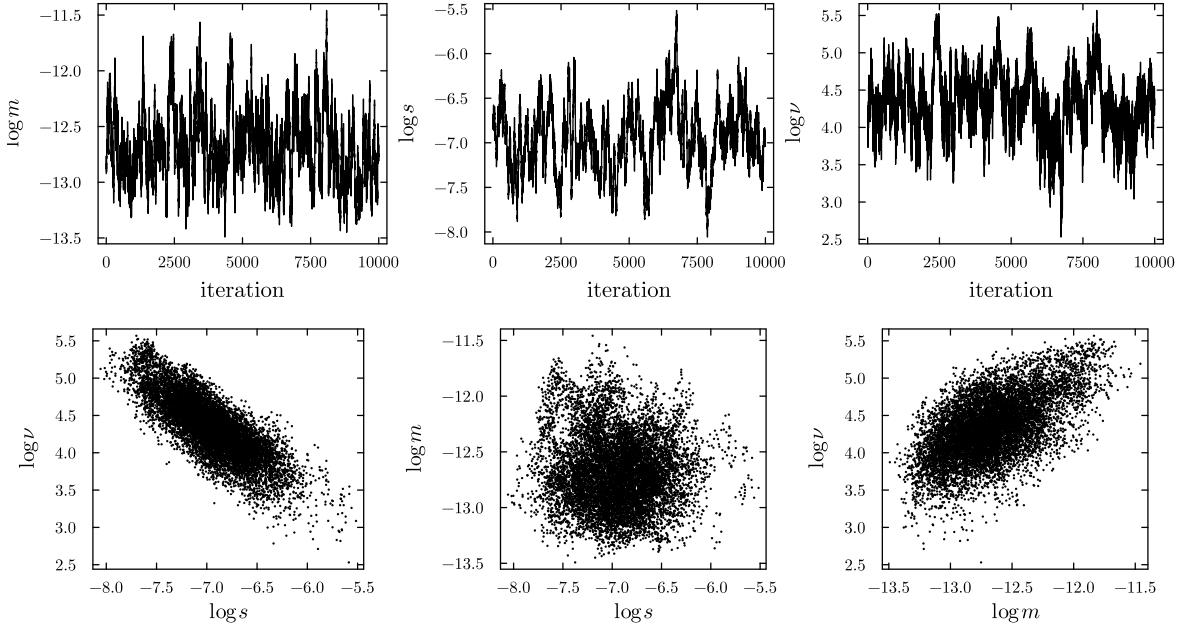
**Figure 2:** Posterior  $m_e$  and number of selected sites in each window for  $\sigma = 0.5$ . The grey line is the gIMble  $m_e$  profile, the red dots are the marginal posterior mean  $m_e$  predictions for each window, and the red bars show the marginal posterior mean number of selected sites in each window (y-axis on the right hand side).



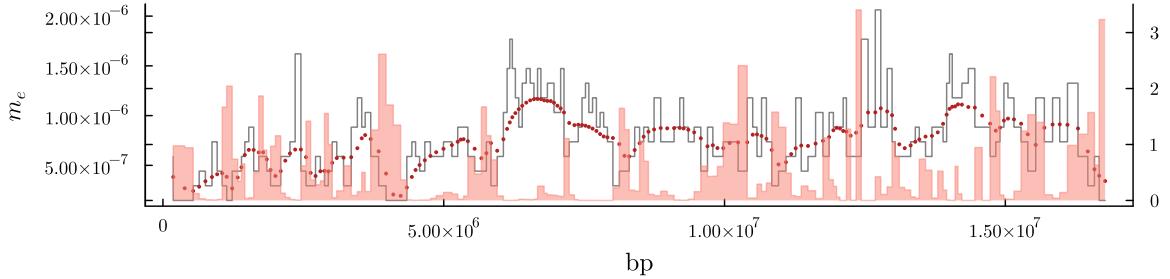
**Figure 3:** As in fig. 1 but with  $\sigma = \sqrt{1/2}$ .



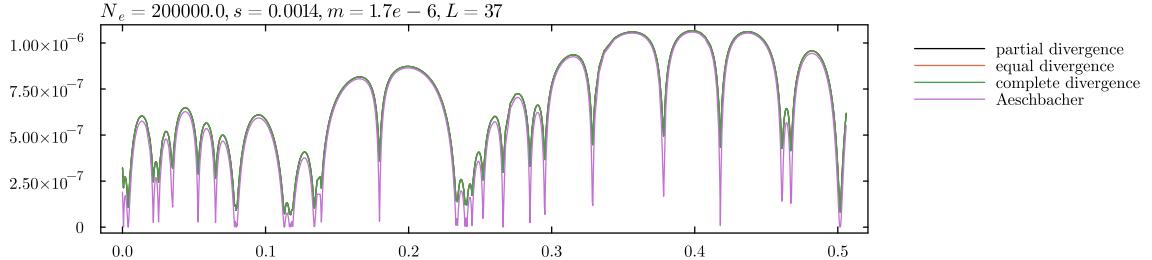
**Figure 4:** As in fig. 2 but with  $\sigma = \sqrt{1/2}$ .



**Figure 5:** As in fig. 1 but with  $\sigma = \sqrt{1/8}$ .



**Figure 6:** As in fig. 2 but with  $\sigma = \sqrt{1/8}$ .



**Figure 7:** Comparison of various  $m_e$  predictions for the *Heliconius* data, assuming plausible parameter values. Partial divergence: this is the full  $m_e$  prediction assuming Zwaenepoel *et al.* (2024). Equal divergence: here we assume all selected loci are at the same equilibrium frequency (the average across the chromosome). Complete divergence: this is the prediction assuming complete divergence (but using our RV-based expressions). Aeschbacher: this is the computation using Aeschbacher and Bürger (2014).