

SUPPLEMENTAL MATERIAL

How linked selection shapes the diversity landscape in *Ficedula* flycatchers

Agnes Rettelbach, Alexander Nater & Hans Ellegren

Alternative explanations to hitchhiking

1. Density of functional sites. The most obvious reason why we overestimate B is that there are more sites under purifying selection than just the conserved sites used for the calculations. For example, Rands et al. (2014) estimated 8% of the human genome to be constrained, in contrast to 5% judged from sequence conservation. However, if twice as many sites are assumed to be targets of selection in the valley regions, it is still not sufficient to explain the difference between baseline π and valleys (Fig. S2).
2. Mutation rate. We used an overall mutation rate but there may be some variation across the genome. Mutation rate variation has two contrasting effects. When calculating BGS, a lower mutation rate leads to higher B_c (see equation. 1). On the other hand, π_0 decreases with lower mutation rate and thus our reference B_d increases, as $B = \frac{\pi}{\pi_0}$. The second effect is larger than the first one (factor versus power), i.e. we need a lower mutation rate in diversity valleys in order to make the gap between B_c and B_d smaller. As an example, if we consider a window with low π that has $B_d=0.12$ and $B_c=0.79$ and assume an eightfold lower mutation rate, $B_d=0.96$ and $B_c=0.97$. Although both values increase, the difference becomes smaller. We did a rough estimate of mutation rate variation along the chromosome using divergence between collared and pied flycatcher

and found that there was not much change in B_d to the case with a constant rate, see Fig. S3.

3. Recombination rate. Our estimates for recombination rate may be too high. A 10 times lower recombination rate in the valley regions could explain the difference, see Fig. S4. While it is indeed more difficult to estimate recombination rate precisely in regions of very low recombination, it is more likely to be underestimated due to the low probability of observing a recombination event in the pedigree, making this explanation improbable.
4. Demography. Flycatchers have undergone population expansion since the last ice age. As shown in Comeron (2017), the effect of population bottlenecks on diversity can vary across the genome. However, a stronger reduction in diversity is expected for regions with lower strength of background selection. In order to explain the valley depth, a stronger reduction in regions with high amount of BGS would be necessary.

REFERENCES

- Comeron, J. M. (2017). Background selection as null hypothesis in population genomics: insights and challenges from drosophila studies. *Philosophical Transactions of the Royal Society of London B*, 372(1736):20160471.
- Rands, C. M., Meader, S., Ponting, C. P., and Lunter, G. (2014). 8.2% of the human genome is constrained: variation in rates of turnover across functional element classes in the human lineage. *PLoS Genetics*, 10(7):e1004525.

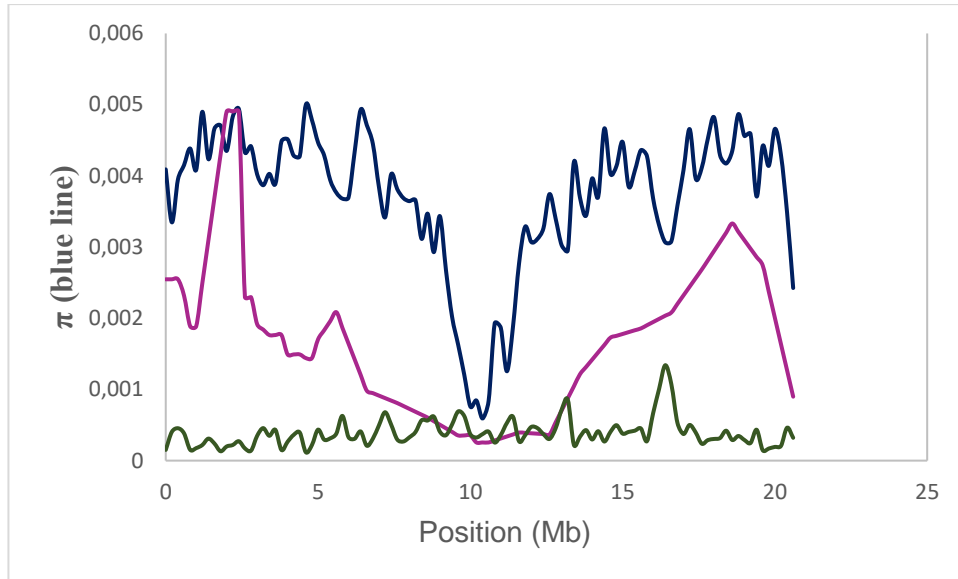


Figure S1: π (blue), recombination rate (purple), and density of conserved sites (green) along chromosome 12, which has a typical landscape. Recombination rate and dcs are scaled to fit in the graph, so no values are given.

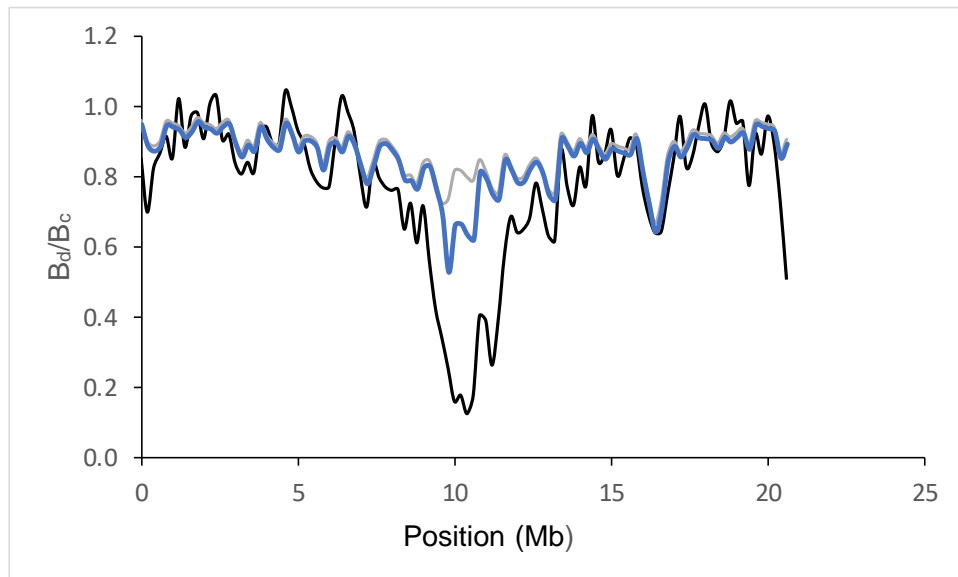


Figure S2: Robustness to functional sites and deleterious mutation rate. The dark blue line shows B_c with a deleterious mutation rate two times higher for 10-12 Mb.

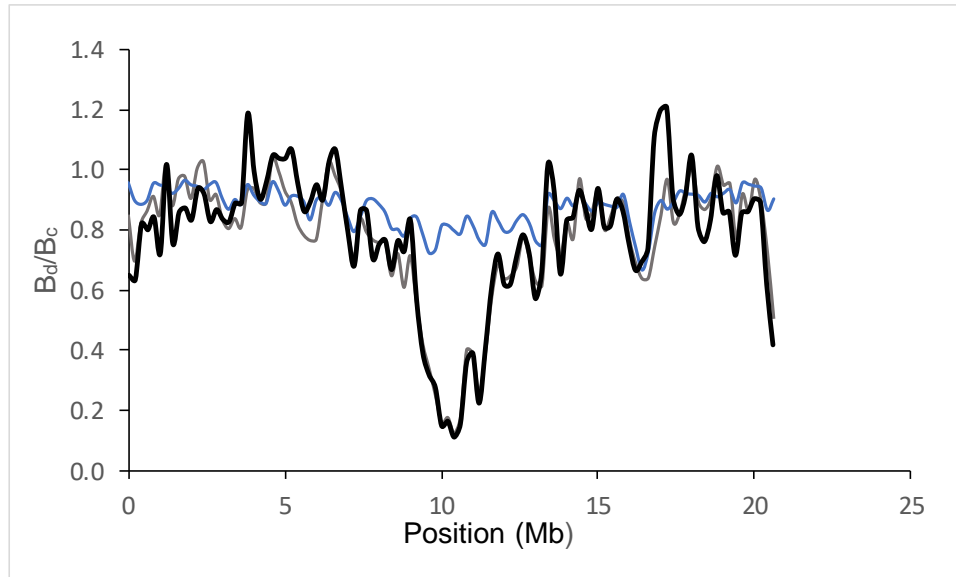


Figure S3: Robustness to mutation rate. As the main effect is on B from the data, in this figure, the black line is B_d with variable mutation rate. The grey line denotes assuming a constant mutation rate and the blue line is model 1. All parameters as in Fig. 2.

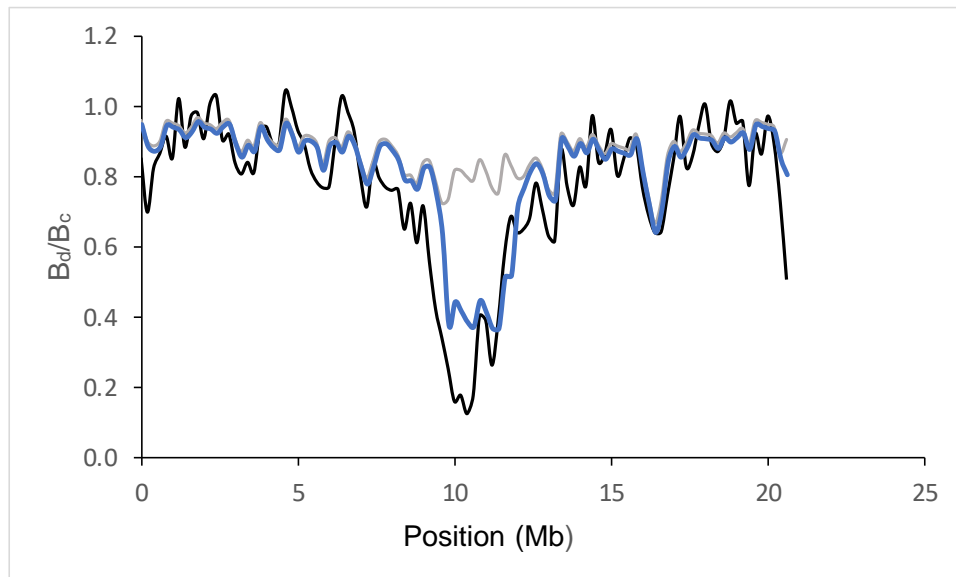


Figure S4: Robustness to recombination: The blue line shows B_c with a 10 times lower recombination rate for 10-12 Mb..

Table S1: Correlations (r^2) between π from data and B_c . Additionally, the correlations between π and the density of conserved sites and recombination rate are shown.

Chromosome	Model 1 (BGS)	Model 2 (recent sweeps)	Model 3 (recent and ancient sweeps)	dcs	Recombination rate
1	0.109	0.360	0.316	0.172	0.020
1A	0.504	0.672	0.474	0.095	0.078
2	0.221	0.221	0.090	0.229	0.012
3	0.348	0.462	0.156	0.309	0.001
4	0.058	0.152	0.151	0.032	0.026
4A	0.533	0.740	0.706	0.067	0.417
5	0.194	0.194	0.246	0.127	0.065
6	0.423	0.563	0.212	0.375	0.123
7	0.314	0.325	0.219	0.253	0.122
8	0.490	0.757	0.045	0.278	0.235
9	0.397	0.397	0.368	0.255	0.286
10	0.656	0.706	0.656	0.061	0.414
11	0.672	0.689	0.663	0.308	0.383
12	0.548	0.846	0.765	0.158	0.417
13	0.260	0.723	0.610	0.043	0.324
14	0.672	0.846	0.812	0.127	0.454
15	0.144	0.314	0.313	0.060	0.229
17	0.533	0.533	0.615	0.109	0.364
18	0.640	0.774	0.052	0.078	0.773
19	0.005	0.003	0.474	0.001	0.078
20	0.063	0.423	0.538	0.001	0.365
21	0.384	0.490	0.607	0.199	0.355
22	0.397	0.490	0.664	0.296	0.289
23	0.336	0.336	0.426	0.255	0.116
24	0.462	0.624	0.766	0.299	0.563
25	0.689	0.689	0.626	0.656	0.462
26	0.176	0.102	0.337	0.008	0.476
27	0.563	0.250	0.287	0.533	0.164
28	0.281	0.423	0.368	0.119	0.423