Title: On the importance of time scales when studying adaptive evolution

Running title: Time scale of bill length evolution

Authors: Charles PERRIER[1] & Anne CHARMANTIER[1]

Affiliations: [1]CEFE UMR 5175, CNRS, Université de Montpellier, Université Paul-Valery Montpellier, EPHE 1919, route de Mende - 34293 Montpellier cedex 05 – France

Corresponding authors: Charles.perrier@cefe.cnrs.fr

Abstract:

Long-term field studies coupled with quantitative genomics offer a powerful mean to understand the genetic bases underlying quantitative traits and their evolutionary changes. However, analyzing and interpreting the time scales at which adaptive evolution occurs is challenging. First, while evolution is predictable on the short term, with strikingly rapid phenotypic changes in data series, it remains unpredictable on the long term. Second, while the temporal dynamics of some loci with large effect on phenotypic variation and fitness have been characterized, this task can be complicated notably in cases of highly polygenic trait' architecture implicating numerous small effect size loci, or when statistical tests are sensitive to the heterogeneity of some key characteristics of the genome, like recombination rate. After introducing these aforementioned challenges, we discuss a recent investigation of the genomic architecture and spatio-temporal variation in great tit bill length, which has been supposedly linked to the recent use of bird feeders. We discuss how this case study illustrates the importance of considering different temporal scales and evolutionary mechanisms both while analyzing trait temporal trends and when searching for and interpreting the signals of putative genomic footprints of selection. This commentary overall discusses stimulating challenges to unravel the time scale at which adaptive traits evolve, and their genomic bases.

Main Text:

Longitudinal field studies have brought invaluable insight for the understanding of evolutionary processes (Clutton-Brock & Sheldon 2010). Long-term studies notably allowed characterizing the temporality of trait variation and the strength and directionality of natural selection underlying such variation. Some of these long-term examinations of key phenotypic traits detected strikingly fast phenotypic variations, driven by prompt ecological changes (Grant & Grant 2006). In contrast, many long-term studies failed to reveal micro-evolutionary changes and response to selection (Merila *et al.* 2001; Pujol *et al.* 2018). While examining longitudinal data looking for both long-term trends and embedded prompt anomalies has the potential to shed light on traits' evolution and evolutionary trajectories in natural populations, our ability to understand and more notably predict evolution remains limited outside the laboratory. Quantitative genetics was initially developed with the embedded aim of predicting evolutionary change, based on estimates of selection and additive genetic variation (Falconer book 1996). Yet when spatio-temporal ecological heterogeneity is involved, evolution in the wild remains largely unpredictable (Pemberton 2010; Pujol et al. 2018).

Coupling such long-term studies with genomic tools is a powerful way to improve our understanding of the genetic bases underlying evolutionary changes in response to environmental variation. Rapid and recent monogenic adaptations based on denovo mutations are often used as examples, for instance the rise of the melanic morph of the peppered moth *Biston betularia* following the industrial revolution (van't Hof et al. 2016). Similarly, there are famous examples of rapid parallel monogenic or oligogenic adaptation based on long lasting standing genetic variation segregating in heterogeneous environments, for example the coloration in deer mice *Peromyscus* and the armor plates in stickleback *Gasterosteus aculeatus* (Barret & Hoekstra 2011; Nelson & Cresko 2018). However, these genomic analyses are facing several challenges when it comes to concluding on polygenic adaptation and quantitative trait evolution. First, loci effect sizes are often

small, requiring thousands of SNPs and of individuals for genome wide association study (GWAS) to reveal significant effects. Such an investigation requires technical commitments and recent advances (Wellenreuther & Hansson 2016; Gienapp et al. 2017a), but also a conceptual revolution consisting in silencing our desire to find large effect alleles that are theoretically rarely responsible for quantitative variation (Rockman 2011). Second, genome characteristics and especially variations in recombination rate along the genome can cause variations of the effect of background selection (*eg* Nordborg et al. 2009), and hence of linkage disequilibrium (LD), diversity and differentiation that can confound footprints of positive selection and bias our comprehension of the timing and nature of the patterns of differentiation along the genomes (Roesti et al. 2012; Burri et al. 2015; Berner & Roesti 2017; Burri 2017; Comeron 2017; Delmore et al. 2018). Nevertheless, several studies have begun to decipher the polygenic mechanisms of rapid evolution of quantitative trait.

Avian bill morphology has played a prominent role in empirical studies of evolution and natural selection (Lack 1947; Grant 1999), perhaps because the size and shape of bills show large variations across and within bird species, and are shaped by strong selective forces since they directly determine foraging efficiency on various food sources. For instance, the emblematic study of Darwin's finches on the Galapagos island of Daphne Mayor aimed at capturing evolutionary changes in bill size (Boag & Grant 1981; Grant & Grant 1993). From 1977 to 1978, bill size increased markedly after a severe drought in 1977. This analysis clearly demonstrated that extreme climatic events such as an El niño event are strong drivers of bill size evolution in this species. After 30 years of perspective, however, Grant & Grant (2002) concluded that while evolution of bill length was predictable as a rapid response to strong selection, it remained unpredictable on a slightly longer microevolutionary scale. Although the question of predictability of evolution across time scales remains challenging, even in the genomic era (Nosil et al 2018), genomic tools did provide insights on evolution of bill size in the context of the rapid diversification of Darwin's finches. A handful of genes were found significantly associated to bill size and shape in the medium ground finch *Geospiza fortis*

(Lamichhaney et al. 2016), among which a major locus has been shown influencing bill dimensions in the Darwin finches' entire radiation (Lamichhaney et al. 2015). These genomic analyses hence cracked the genomic architecture of this trait variation at both small and large microevolutionary scale, with the predominant control of a few large effect loci.

In a recent study, Bosse et al. (2017) investigated the genetic architecture of bill length in the Great tit *Parus major*, using a tremendous amount of data from long-term research programs in Wytham woods in UK and in Netherlands (NL). Applying modern analyses of both population and quantitative genomics using 500k single nucleotide polymorphisms, the authors provide insight into the signatures of divergent selection in the studied populations and the genomic architecture of variation in bill length. In line with the quantitative nature of variation in bill length and with quantitative genetics theory (Lynch & Walsh 1998), the authors showed that the genetic architecture of bill length was highly polygenic. Specifically, the authors showed, using a mixture analysis fitting all the SNPs simultaneously, that 3009 SNPs explained collectively 31% of bill length phenotypic variation. None of these SNPs reached genome wide significance in the GWAS with bill length, revealing small effect sizes of individual variants. Accordingly with recent quantitative genomic findings, the proportion of variance in bill length explained by each chromosome amazingly scaled with its size, showing that the many SNPs additively explaining bill length are distributed throughout the genome. The polygenic analysis also predicted the difference in bill length observed between UK and the NL, further illustrating the polygenic nature of this trait variation. This evidence for a polygenic control of bill length with no large effect SNP is hence very different from the previous example on Darwin finches' bill with large effect loci. Nevertheless, Bosse et al. (2017) also showed that variation at one gene, col4A5, was associated with use of feeders and that this might have driven the evolution of larger bills in UK compared to NL. Among the arguments pointing to feeders as drivers of longer bill lengths in UK great tits, Bosse et al. (2017) reported that col4A5 was associated to bill length in UK but not in NL, was highly differentiated between UK and NL, and was

associated with greater reproductive success and higher activity at feeding sites in UK. In addition, they reported that bills were longer in UK compared to mainland Europe (Figure 4A in Bosse et al. 2017) and increased from 1982 to 2007 in UK (Figure 4B in Bosse et al. 2017).

We argue in this comment that the speculation on the role of feeder on bill length evolution is not well supported by these arguments. Based on both phenotypic and genomic data, we propose instead that differences in bill length between UK and NL might have been evolving on a longer term than the contemporary use of feeders. First, an examination of bill length monitoring shows puzzling temporal and geographic patterns that may not incriminate the use of feeders. Second, the genomic patterns found at the region containing col4A5 might be compatible with the combined effect of structural variation in recombination rate and background selection, hence questioning the putative recent role of feeders as agents of positive selection.

Bill length trend inferred from the long-term monitoring is highly dependent on the time scale considered. Inspired by the readings of papers such as the one of Grant & Grant (2006) relating the effects of rare and extreme events on the evolution of phenotypic traits, we carefully inspected the evolution of bill length during the studied period, looking for particularly prompt changes. We reanalyzed the data using a breakpoint computations method (coin R-package) aiming at localizing such striking change. The best cutpoint was found between 1986 and 1987 (maxT = 7.41, p-value < 2.2E-16), which corresponds to a conspicuous change in bill length. Measures from the five years preceding 1987 differed from subsequent years (t-test: t = -7.28, df = 674.53, p-value = 9.1E-13). Although a linear regression though the entire period, from 1982 to 2007, as implemented by Bosse et al., yielded a positive slope, removing the five measures prior to 1987, bill length significantly decreased from 1987 to 2007 (Figure 1A, Linear model from 1987 to 2007 taking into account or not tarsus length, F = 7.13, p-value = 8.2E-4; F = 13.2, p-value: 2.9E-4, respectively). A LOESS model confirms the decrease in bill length during the second part of the record. More generally, slopes of linear models linking bill length to time, using all of the possible combinations of 10 to 25

consecutive years (simply removing 1 to 16 years at the beginning, or the end, or both, of the data-series), were often positive when including one or more data points collected between 1982 and 1987, while negative slopes were often observed when excluding these years (Figure 2), further illustrating that both time periods yield opposite patterns. Besides, tarsus length also decreased significantly both over the 1987-2007 and the 1982-2007 periods (Figure 1B, F = 6.61, p-value = 0.01 & F = 24.07, p-value = 9.9E-07, respectively). Correspondingly, the bill length / tarsus length ratio did not change significantly from 1987 to 2007 (Figure 1C, F = 1.51, p-value = 0.28), indicating no change in bill length, when scaled to tarsus length, over this period. The origin of this cutpoint is important to investigate, since it could constitute an evidence for a response, either genetic or plastic, of both bill and tarsus lengths to a prompt abiotic or biotic factor or to a methodological change. Yet, it seems at present difficult to decipher whether bill length or tarsus length or both, if any, were targeted by selection. Indeed, both traits are commonly phenotypically and genetically positively correlated in passerines (Teplitsky et al. 2014; Poissant et al. 2016), hence should often evolve together. Overall, these results much likely rule out the possibility of a contemporary (1982-2007) positive effect of feeders on bill length.

One could nevertheless argue that bill length could have increased at a longer time scale than the period 1982-2007, but still recent enough to incriminate the use of feeders. Figure 3 however shows the absence of a clear pattern of bill length increase in museum specimens collected in UK. Although more data are needed to confirm the pattern, it suggests that larger bill length in UK seems to have evolved on a larger time period than the one during which feeders have been used. Moreover, bill length across Europe is not displaying a clear dichotomy between UK and mainland Europe but rather smooth spatial variations (Figure 4) with an ANOVA showing a significant effect of country (p = 3.75e-07) but not of sex (p = 0.345), with indeed UK birds having longer bills compared to three other countries (France, Italy, and NL). This spatial variation challenges, once again, the

149 supposition by Bosse et al.'s (2017) of a recent increase of bill length in British great tits that would
150 have been caused by feeders.

151 We then question whether the evolution of the genomic region containing the gene col4A5,
152 which was the corner stone linking bill length, activity at feeders, reproductive success and divergent
153 selection, could have been influenced by recombination rate variation and background selection,
154 rather than recent positive selection via feeders. While the decay of LD is very fast in great tits
155 (marginal after 2kb, as shown in Figure S1 in Bosse et al. 2017 and in Figure 5A here (see
156 supplementary material 1 for details about the methods)), col4A5 was found in a large (>1Mb)
157 genomic region with high long distance (20 to 200Kb) LD not only in UK but also in NL (Figure 5B).
158 The eigenGWAS was also significant in both UK and NL (Figure 5C). This large region was previously
159 identified by Laine et al. (2016) as showing signature of selective sweeps and reduced diversity (Sixth
160 chromosome in Figure 2 in Laine et al. 2016) at the scale of the species entire distribution. Overall, it
161 seems difficult that positive selection in UK could explain these large genomic variations given the
162 omnipresence of this large genomic region with reduced variation and increased LD. This same
163 region was also identified as showing elevated differentiation in several lineages and populations of
164 *Ficedula* flycatchers (Ellegren et al. 2012; Figure 1C in Burri et al. 2015). Burri et al. (2015)
165 determined that such high differentiation shared across lineages was due to the effect of linked
166 selection in interaction with low recombination and concluded on a crucial warning: "scans are likely
167 to identify recombination-mediated elevations of differentiation not necessarily attributable to
168 selective sweeps". Accordingly, we propose that low recombination and background selection in the
169 region containing col4A5 in great tits could have resulted in reduced genetic diversity, increased
170 differentiation and increased LD that could have altogether mimicked signatures of positive
171 selection. In addition, neglecting the effects of variation in recombination and LD along the genome
172 might have resulted both in false positive footprints of selection in coldspots of recombination but
173 also in false negatives in regions with larger recombination rate (Berner & Roesti 2017).

174 Furthermore, the large sliding window used in Bosse et al. (200kb while the LD is highly reduced
175 after 2Kb) probably worsened the problem raised above, by capturing principally elevated
176 differentiation in large regions with large LD and by diluting narrow peaks of elevated differentiation
177 in regions of elevated recombination and low LD. Unfortunately, this bias holds independently from
178 the nature of selection and is a purely mathematical problem of mismatch between the sliding
179 window length and the extent of LD variation along the genome. In fact, such large sliding windows
180 might be rather used as neutral local baselines to take into account the effects of variation in LD and
181 recombination along the genome, with local outliers detected in inspecting local residuals to such
182 baselines (Roesti 2012; Burri 2017). Correspondingly, the col4A5 FST peak only emerged after using a
183 smoothing method, in this region that clearly displayed high LD and low marker density (Figure 5D).
184 Similarly, the col4A5 GWAS peak only emerged when using a smoothing method (Figure 5E).
185 However, this GWAS peak emerged only in UK and not in NL (Figure 5F). All these considerations
186 suggest that, although there is indeed an association between col4A5 and bill length in UK, the role
187 for the combination of background selection, recombination rate variation and invariably large
188 averaging window while long distance LD is variable, should be more comprehensively tested.

189 To illustrate the general challenge raised here in differentiating recombination rate variation
190 combined with background selection from positive selection, we present a simple simulation of a
191 20Mb chromosome (hence comparable to chromosome 4A) containing a region of 1.6Mb with highly
192 reduced recombination, with random occurrence of neutral and deleterious mutations but no
193 beneficial ones (see supplementary material 2 for details about the methods). We show that for an
194 average LD decay comparable to what was found in great tits (Figure 6A), long distance LD (Figure
195 6B), eigenGWAS (Figure 6C) and integrated extended haplotype homozygosity (Figure 6D) are all
196 three much larger in the recombination coldspot than elsewhere in the chromosome. Therefore, we
197 illustrate that such pattern as observed in Bosse et al. is compatible with the combined action of

large-scale variation in recombination and background selection, challenging the putative role of the feeders as a contemporary selective agent on bill length.

Opportunely, this context offers exciting avenues of research to unravel the determinants of both recent and long-term as well as spatial variations in quantitative traits in the Great tit but also in other emblematic species displaying quantitative traits variations. Spatial traits variation could be unraveled in multi-trait GWAS, using polygenic frameworks, which could be inspired from a pioneering study on great tit breeding phenology (Gienapp et al. 2017b). Besides, a formal selection analysis relating bill length to over-winter survival when birds are most likely to benefit from food provided by people, is required, to elucidate the nature of the evolutionary forces behind bill length variation. Finally, taking into account the effect of variation in recombination rate, background selection and LD along the genome to draw local neutral envelops of genomic differentiation and modulate the local width of sliding windows, will likely help identifying further candidate loci (Roesti et al. 2012; 2013; Burri 2017; Comeron 2017; Delmore et al. 2018). Providentially, the genomic analyses performed in other populations across Europe could also help determining the timing and the nature of selection, by helping to take into account the temporal dynamics of differentiation landscape (Burri 2017).

Acknowledgements

We thank two anonymous native English speakers for comments on the English syntax. We thank several colleagues from the Institute of Evolutionary Sciences of Montpellier and from the Center for Functional and Evolutionary Ecology, especially Paul Jay, for discussion regarding the impact of recombination rate variation on landscape of differentiation. We thank the associate editor and four anonymous reviewers for constructive comments on previous versions of the manuscript. Both authors are presently funded by a European Research Council Starting grant – *Selection in Heterogeneous Environment* (ERC-2013-StG-337365-SHE).

References and Notes

Barrett RD & Hoekstra HE (2011). Molecular spandrels: tests of adaptation at the genetic level. *Nature Reviews Genetics*, *12*, 767.

Berner D, Roesti M (2017) Genomics of adaptive divergence with chromosome-scale heterogeneity in crossover rate. *Molecular Ecology*, **26**, 6351–6369.

Boag PT, Grant PR (1981) Intense natural selection in a population of Darwin's finches (Geospizinae) in the Galapagos. *Science*, **214**, 82–85.

Bosse M, Spurgin LG, Laine VN *et al.* (2017) Recent natural selection causes adaptive evolution of an avian polygenic trait. *Science (New York, N.Y.)*, **358**, 365–368.

Burri R (2017) Interpreting differentiation landscapes in the light of long-term linked selection. *Evolution Letters*, **1**, 118–131.

Burri R, Nater A, Kawakami T *et al.* (2015) Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedulaflycatchers. *Genome Research*, **25**, 1656–1665.

Clutton-Brock T, Sheldon BC (2010) Individuals and populations: the role of long-term, individual-based studies of animals in ecology and evolutionary biology. *Trends in Ecology & Evolution*, **25**, 562–573.

Comeron JM (2017) Background selection as null hypothesis in population genomics: insights and challenges from Drosophila studies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **372**, 20160471–13.

Delmore KE, Lugo Ramos JS, Van Doren BM *et al.* (2018) Comparative analysis examining patterns of genomic differentiation across multiple episodes of population divergence in birds. *Evolution Letters*, **2**, 76–87.

Ellegren H, Smeds L, Burri R *et al.* (2012) The genomic landscape of species divergence in Ficedula flycatchers. *Nature*, **491**, 756–760.

Gienapp P, Fior S, Guillaume F *et al.* (2017a) Genomic Quantitative Genetics to Study Evolution in the Wild. *Trends in Ecology & Evolution*, 1–12.

Gienapp P, Laine VN, Mateman AC, Van Oers K, Visser ME (2017b) Environment-Dependent Genotype-Phenotype Associations in Avian Breeding Time. *Frontiers in Genetics*, **8**, 1392–9.

Grant BR, Grant PR (1993) Evolution of Darwin's Finches Caused by a Rare Climatic Event. *Proceedings of the Royal Society B-Biological Sciences*, **251**, 111–117.

Grant PR (1999) *Ecology and evolution of Darwin's finches*. Princeton University Press.

Grant PR, Grant BR (2002) Unpredictable Evolution in a 30-Year Study of Darwin's Finches. *Science*, **296**, 707–711.

Grant PR, Grant BR (2006) Evolution of character displacement in Darwin's finches. *Science (New York, N.Y.)*, **313**, 224–226.

Haller BC & Messer PW (2017) SLiM 2: Flexible, interactive forward genetic simulations. *Molecular Biology and Evolution 34*, 230–240.

Lack D (1947) *Darwin's finches*. Cambridge University Press, London.

Laine VN, Gossmann TI, Schachtschneider KM *et al.* (2016) Evolutionary signals of selection on cognition from the great tit genome and methylome. *Nature Communications*, **7**, 1–9.

Lamichhaney S, Berglund J, Almén MS *et al.* (2015) Evolution of Darwin's finches and their beaks revealed by genome sequencing. 1–16.

Lamichhaney S, Han F, Berglund J, Wang C, 2016 A beak size locus in Darwin's finches facilitated character displacement during a drought. *science.sciencemag.org*

Lynch M, Walsh B (1998) *Genetics and Analysis of Quantitative Traits*. Sinauer Associates Incorporated.

Merila J, Sheldon BC, Kruuk LE (2001) Explaining stasis: microevolutionary studies in natural populations. *Genetica*, **112-113**, 199–222.

Nelson TC & Cresko WA (2018). Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. *Evolution Letters*, *2*, 9-21.

Nordborg M, Charlesworth B, & Charlesworth D (1996). The effect of recombination on background selection. *Genetics Research*, *67*, 159-174.

Poissant J, Morrissey MB, Gosler AG, Slate J, Sheldon BC (2016) Multivariate selection and intersexual genetic constraints in a wild bird population. *Journal of Evolutionary Biology*, **29**, 2022–2035.

Pemberton JM (2010) Evolution of quantitative traits in the wild: mind the ecology. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *365*, 2431-2438.

Pujol B, Blanchet S, Charmantier A *et al.* (2018) The Missing Response to Selection in the Wild. *Trends in Ecology & Evolution*, 1–10.

Rockman MV (2011) The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution*, **66**, 1–17.

Roesti M, Hendry AP, Salzburger W, Berner D (2012) Genome divergence during evolutionary diversification as revealed in replicate lake–stream stickleback population pairs. *Molecular Ecology*, **21**, 2852–2862.

Roesti M, Moser D, Berner D (2013) Recombination in the threespine stickleback genome-patterns and consequences. *Molecular Ecology*, **22**, 3014–3027.

Teplitsky C, Tarka M, Møller AP *et al.* (2014) Assessing Multivariate Constraints to Evolution across Ten Long-Term Avian Studies. **9**, e90444–15.

van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA [...] & Saccheri IJ (2016). The industrial melanism mutation in British peppered moths is a transposable element. *Nature*, *534*, 102.

Wellenreuther M, Hansson B (2016) Detecting Polygenic Evolution: Problems, Pitfalls, and Promises. *Trends in Genetics*, **32**, 155–164.

Figures legends

Figure 1. Decrease of mean bill length (A) and tarsus length (B) but conservation of the allometry between both traits (C), from 1987 to 2007 (red lines) in Wytham great tits. Regressions from 1982 to 2007 as in the original paper by Bosse et al. are showed by the blue lines. Dots and bars illustrate means and standard errors, respectively, for each year for each variable. Shaded areas represent 95% standard error around regressions lines.

Figure 2. Slopes of linear models of bill length temporal evolution, including every possible combination of 10 to 25 consecutive years of data, with or without one to several years of data from 1982 to 1986.

Figure 3. Variation of mean bill length in UK from 1850 to 2007, based on museum records used in Bosse et al. 2017.

Figure 4. Variation in bill length across European countries for both female (in blue) and male (in green) great tit specimens from museums. Letters A and B illustrate significant differences between sites and 'ns' refers to non-significant differences (Tukey HSD).

Figure 5. A) Smoothed linkage disequilibrium (LD) decay with genomic distance in UK and NL great tits. Along the chromosome 4A for UK (in red) and NL (in blue) populations, smoothed long distance (between SNPs distant from 20 to 200Kb) LD variation (B), smoothed significativity of eigenGWAS (C), SNP and smoothed FST between UK and NL (D), SNP and smoothed significativity of GWAS for bill length in UK (E) and NL (F). The grey area indicates the region in which col4A5 is located.

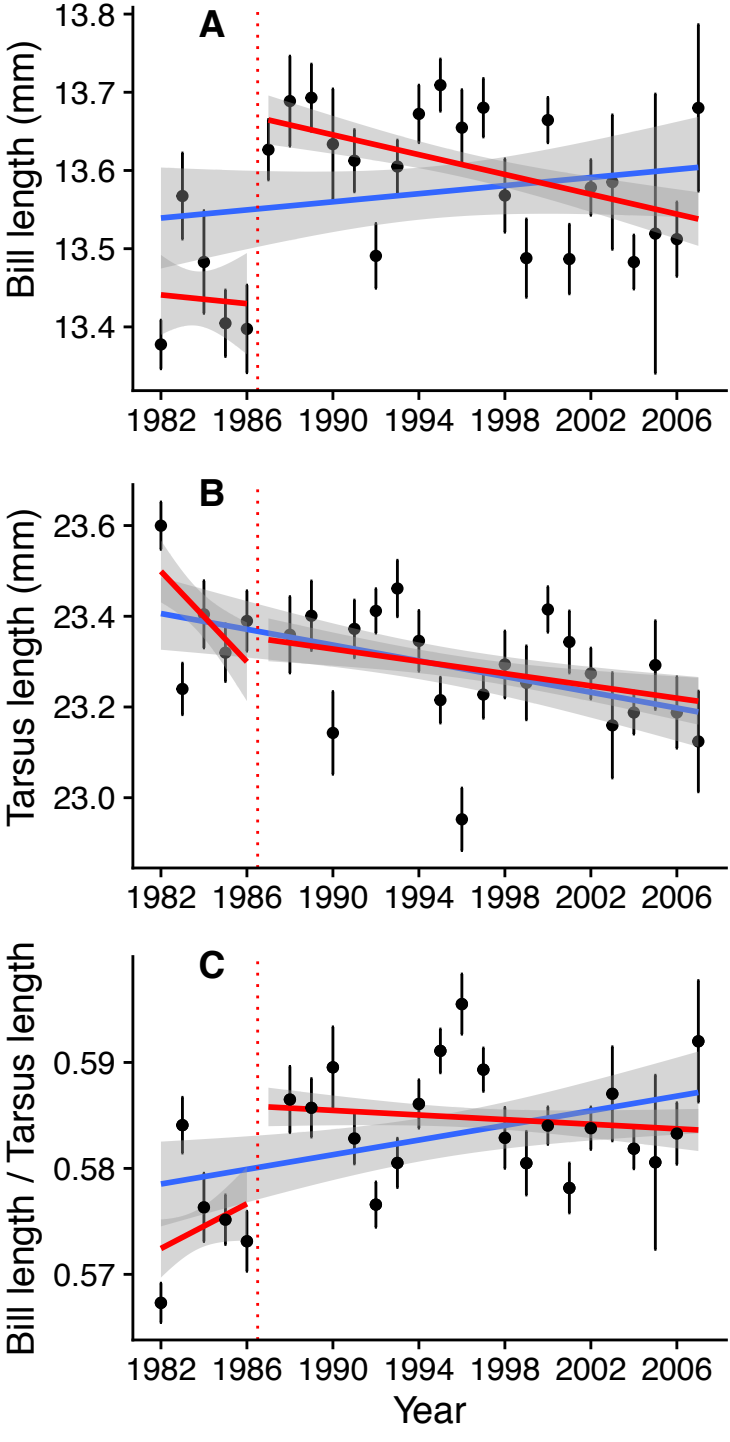
Figure 6. Genomic analysis of a 20Mb chromosome evolving in a single population, with a 1.6Mb region of reduced recombination (indicated by a grey area), and random occurrence of neutral and deleterious mutations and absence of beneficial mutations, simulated in SLiM. A) Smoothed

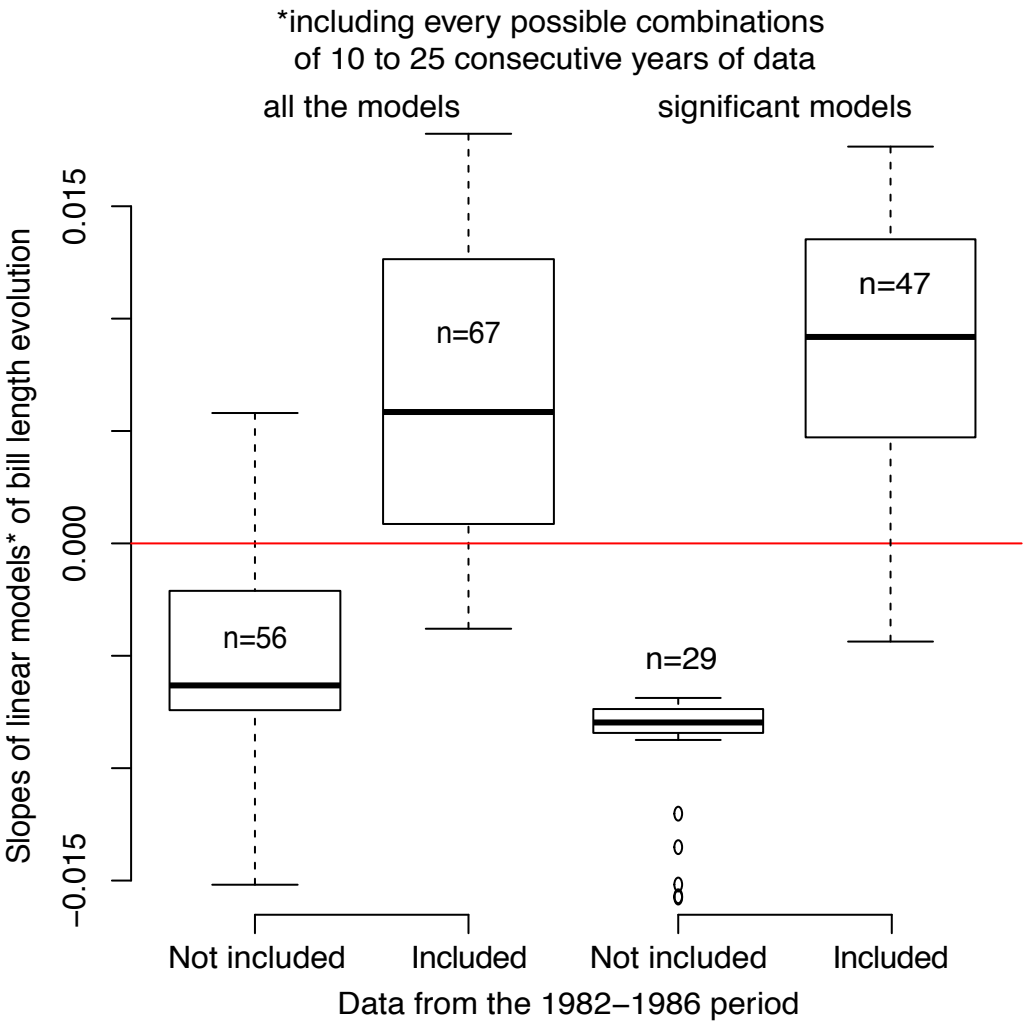
LD decay with genomic distance. B) Smoothed long distance (between sites distant from 20 to 200Kb) LD variation. C) smoothed significativity of eigenGWAS. D) Smoothed integrated Extended Haplotype Homozygosity (iES).

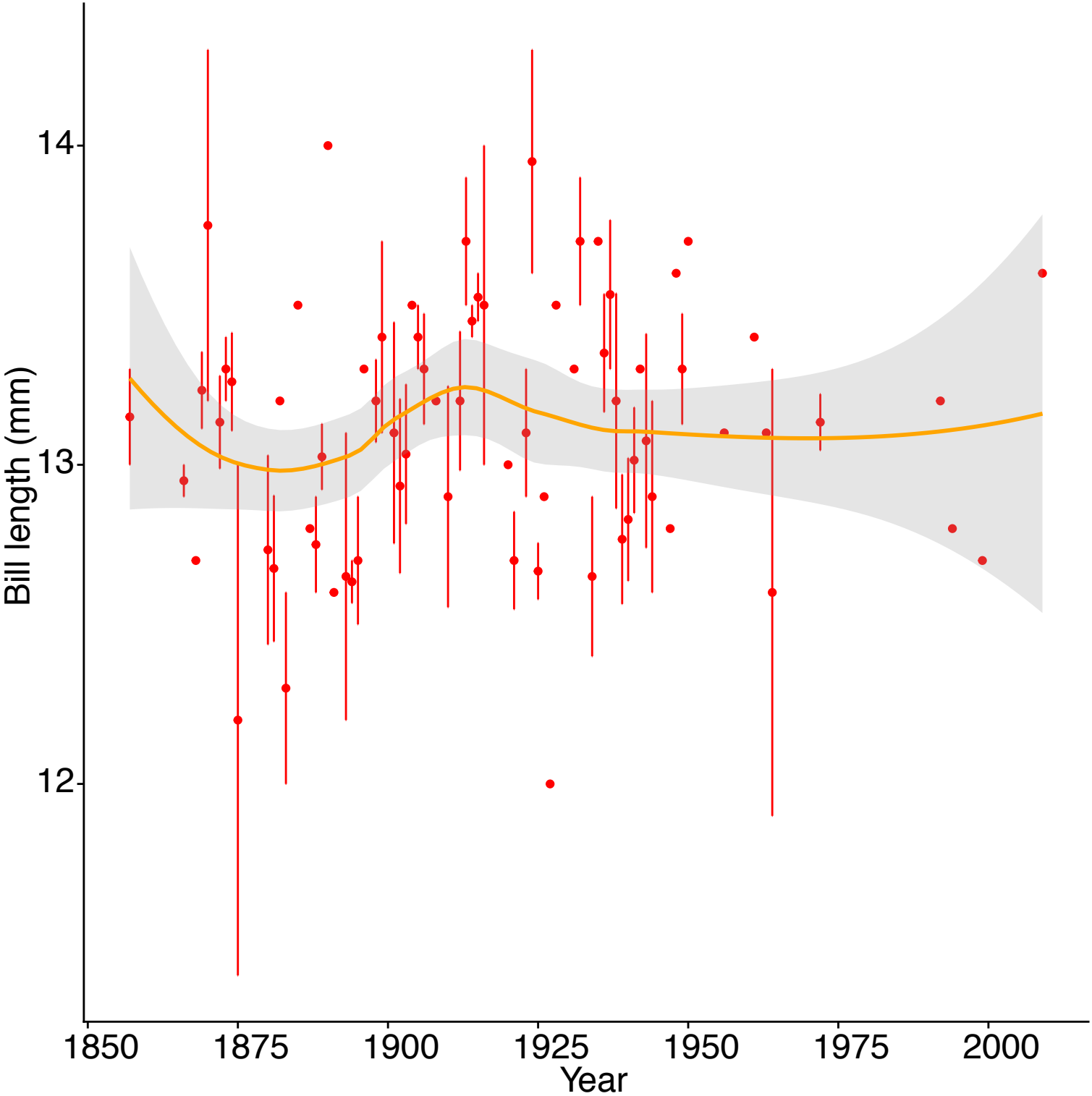
Supplementary material 1. Reanalysis of the COL4A5 region. We downloaded the SNP dataset from Bosse et al. 2017 and used vcftools (Danecek et al. 2011) to estimate LD across the genome for the entire dataset. We then estimated LD for both populations separately along the chromosome 4A. To represent long distance LD, we kept LD values between each pairs of markers distant from 20kb to 200kb. We performed eigenGWAS tests (Chen et al. 2017) in UK and NL. We used Genabel (Aulchenko et al. 2007) to perform GWAS for bill length in both UK and NL. We smoothed statistics using the loess function native to R.

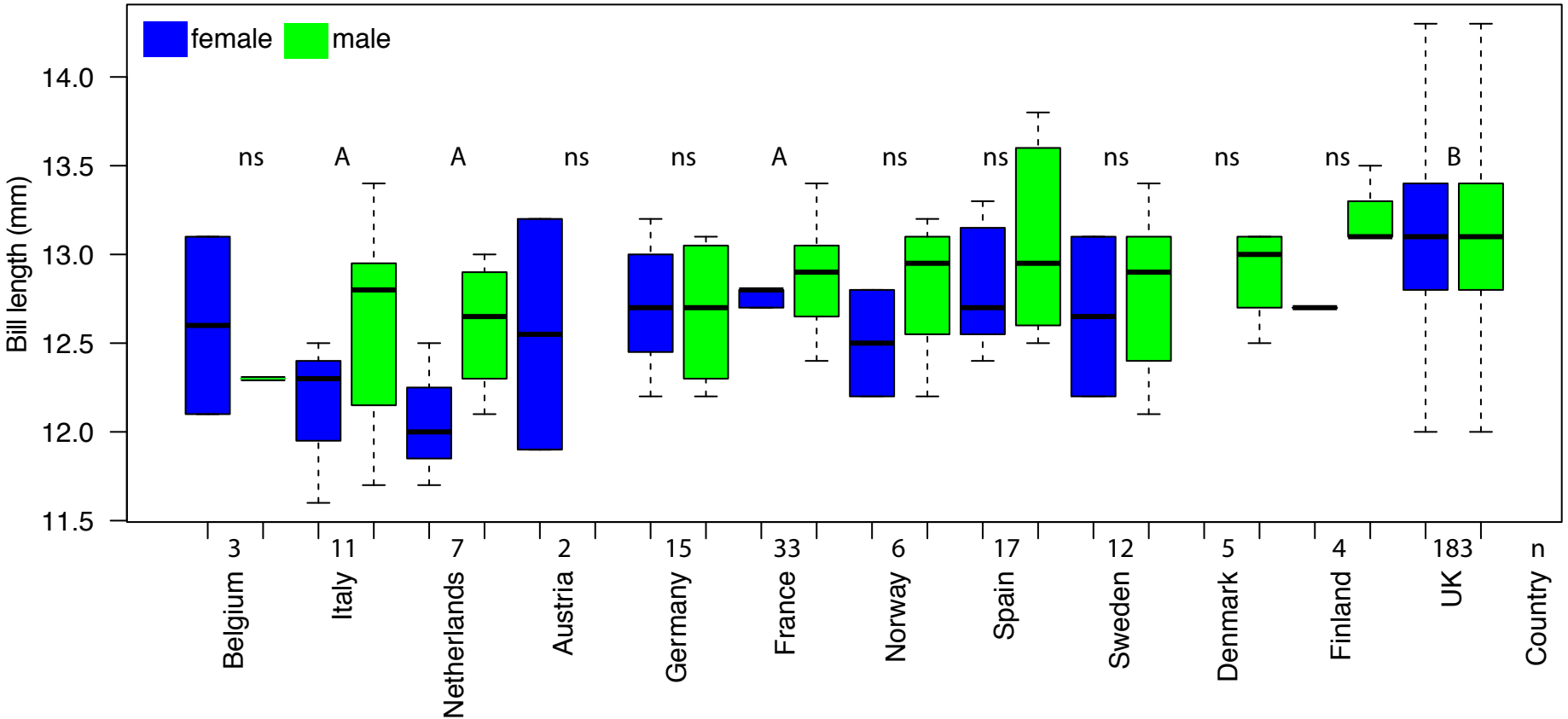
Supplementary material 2. Simulation of a chromosome similar in size to the chromosome 4A in great tit, and containing a region of reduced recombination of approximately the same size as the COL4A5 region. We used SLiM (Haller et al. 2017) to simulate a very simple scenario of reduced recombination in a 1.6Mb region within a 20Mb long chromosome. Recombination rate was equal to 1e-5 from 0 to 11.5Kb and from 13.1Kb to 20Mb. Recombination rate was equal to 1e-100 11.5Kb to 13.1Kb. Population size was equal to 1000. Simulation lasted 2000 generations. Mutation rate was 1e-7. We implemented 3 types of mutations: i) neutral, 50% of the mutations, ii) slightly deleterious (fitness effect of -0.1), 25% of the mutations, iii) severely deleterious (fitness effect of -0.9), 25% of the mutations. We exported a VCF dataset containing about 8500 SNPs from a randomly chosen simulation. We used vcftools to estimate LD across the chromosomes. To represent long distance LD, we kept LD values between each pairs of markers distant from 20kb to 200kb. We performed eigenGWAS tests (Chen et al. 2017). We

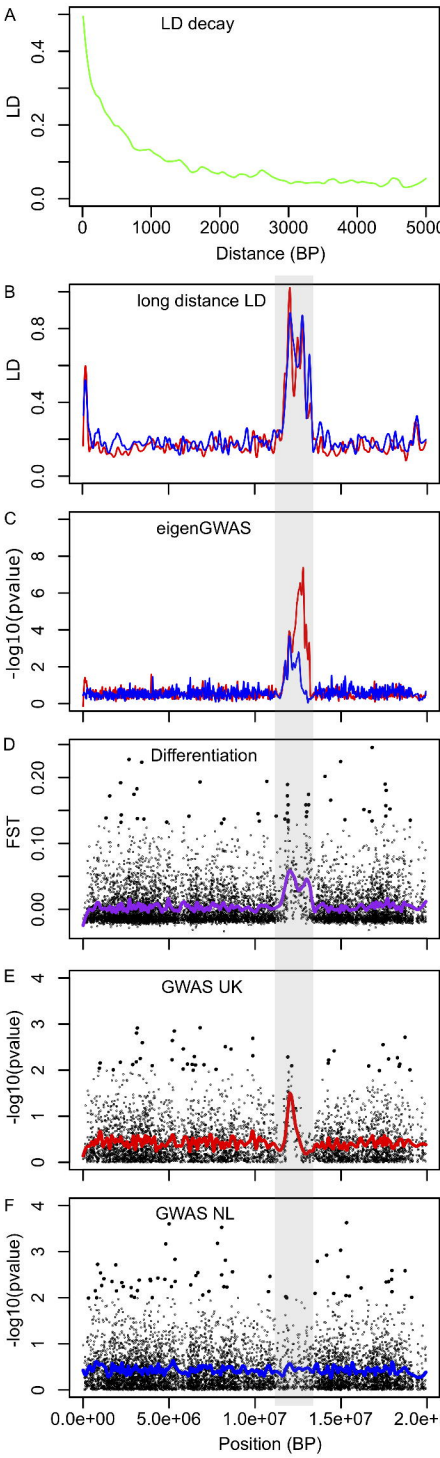
estimated iES (Sabesti et al. 2007) using the R package rehh (Gautier et al. 2016). We smoothed statistics using the loess function native to R.

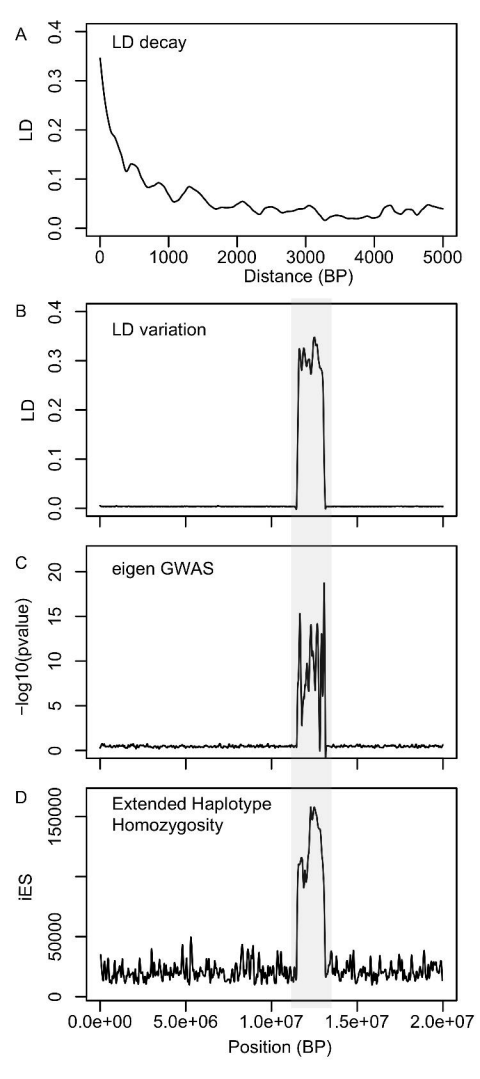

A
Bill length (mm)
13.8
13.7
13.6
13.5
13.4
1982
1986
1990
1994
1998
2002
2006
B
Tarsus length (mm)
23.6
23.4
23.2
23.0
1982
1986
1990
1994
1998
2002
2006
C
Bill length / Tarsus length
0.59
0.58
0.57
1982
1986
1990
1994
1998
2002
2006
Year


*including every possible combinations
of 10 to 25 consecutive years of data
all the models
significant models
Slopes of linear models* of bill length evolution
0.015
0.000
−0.015
n=56
n=67
n=29
n=47
Not included
Included
Not included
Included
Data from the 1982−1986 period


14
13
12
Bill length (mm)
1850
1875
1900
1925
1950
1975
2000
Year


female
male
Bill length (mm)
14.0
13.5
13.0
12.5
12.0
11.5
ns
A
A
ns
ns
A
ns
ns
ns
ns
ns
B
3
11
7
2
15
33
6
17
12
5
4
183
n
Belgium
Italy
Netherlands
Austria
Germany
France
Norway
Spain
Sweden
Denmark
Finland
UK
Country


A
LD decay
LD
0.0
0.2
0.4
0
1000
2000
3000
4000
5000
Distance (BP)
B
long distance LD
LD
0.0
0.4
0.8
C
eigenGWAS
-log10(pvalue)
0
2
4
6
8
D
Differentiation
FST
0.00
0.10
0.20
E
GWAS UK
-log10(pvalue)
0
1
2
3
4
F
GWAS NL
-log10(pvalue)
0
1
2
3
4
0.0e+00
5.0e+06
1.0e+07
1.5e+07
2.0e+07
Position (BP)


A
LD decay
LD
0.0
0.1
0.2
0.3
0.4
0
1000
2000
3000
4000
5000
Distance (BP)
B
LD variation
LD
0.0
0.1
0.2
0.3
0.4
C
eigen GWAS
−log10(pvalue)
0
5
10
15
20
D
Extended Haplotype
Homozygosity
iES
0
50000
150000
0.0e+00
5.0e+06
1.0e+07
1.5e+07
2.0e+07
Position (BP)