Supplementary figures

Figure S1: Boxplots of percentage of total sequence reads mapping to the house sparrow reference genome.

Figure S2: PCA of allelic frequencies (as in Fig 1B) but showing within-species population structuring instead of among species.

Figure S3: Cross-validation of ADMIXTURE runs shows highest support for K = 2 & 3.

Figure S4: Violin plots show the distribution of pairwise genome-wide F_{ST} estimates from 100 Kb sliding windows with a 25 Kb step among and within species.

Figure S5: Genome-wide F_{ST} patterns across all chromosomes from a 100 Kb sliding windows with a 25 Kb step. Black lines represent house vs. Bactrianus, blue lines denote Spanish vs. Bactrianus and red lines show House vs. Spanish.

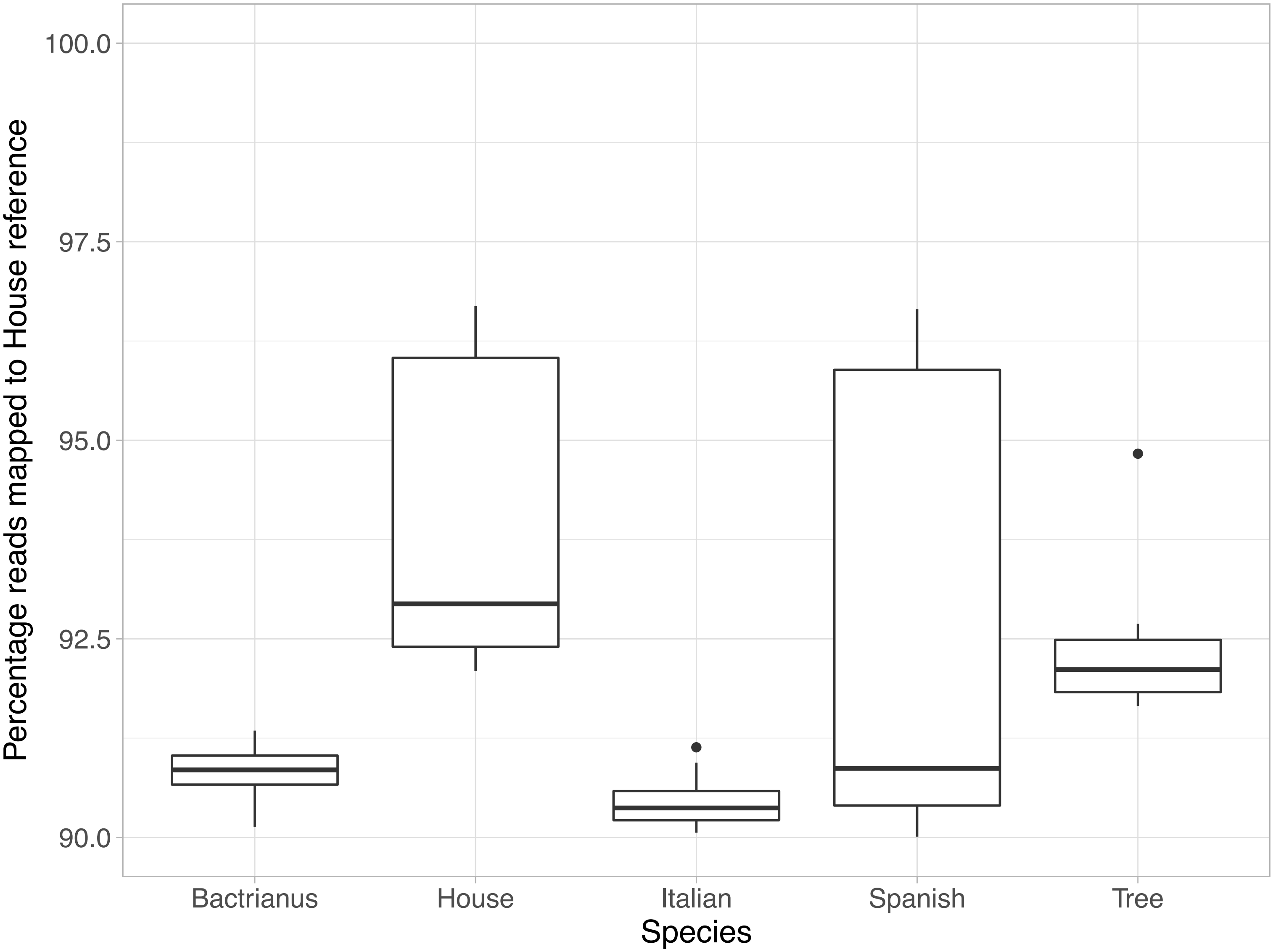
Figure S6: Violin plots showing the distribution of pairwise genome-wide d_{XY} (a) and f_d (b) estimates from 100 Kb sliding windows with a 25 Kb step among and within species.

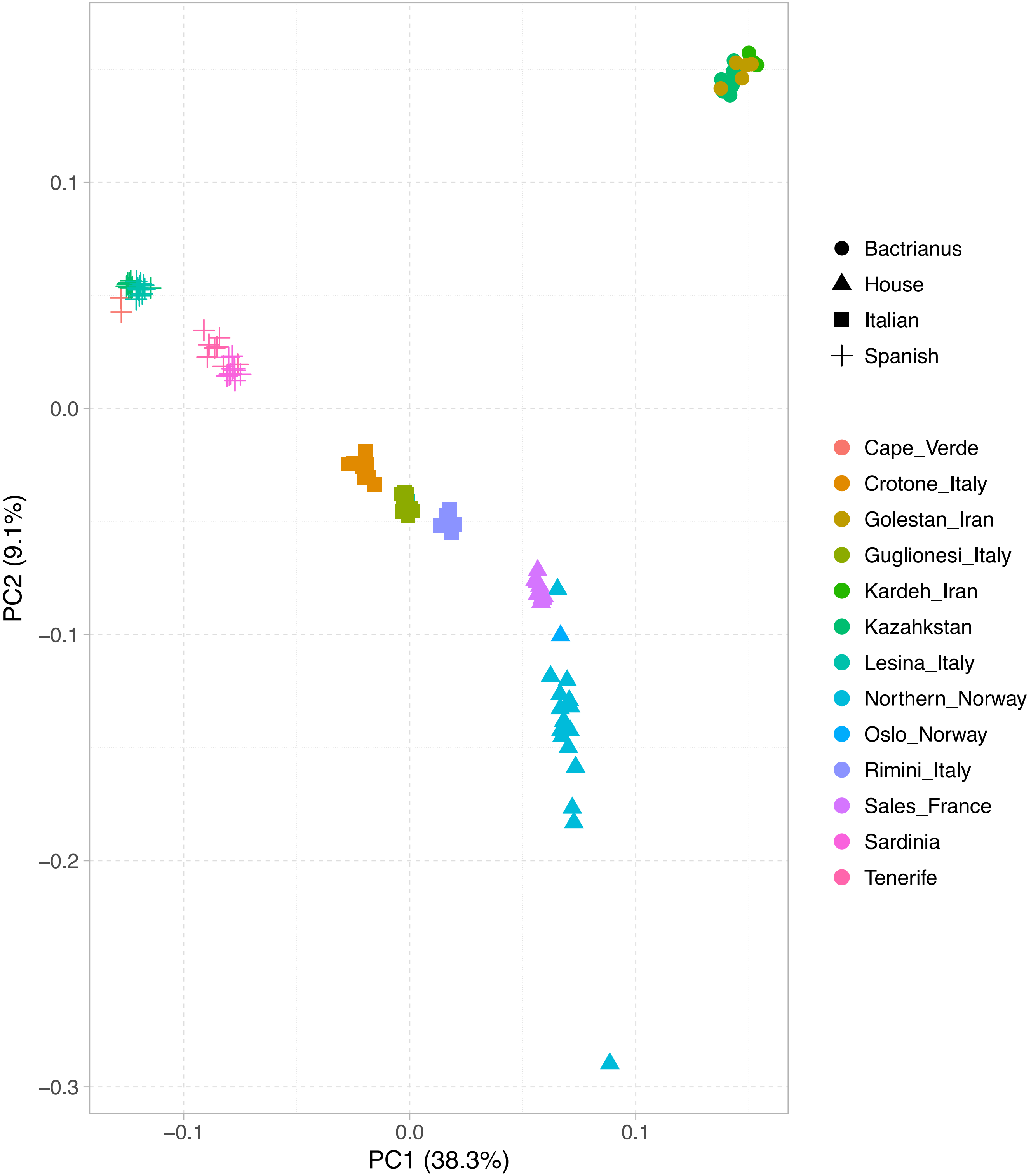
Figure S7: Demographic models of isolation, migration and admixture tested using site-frequency spectrum methods. Bottom panel shows migration parameters included in the models when intraspecific gene flow is allowed.

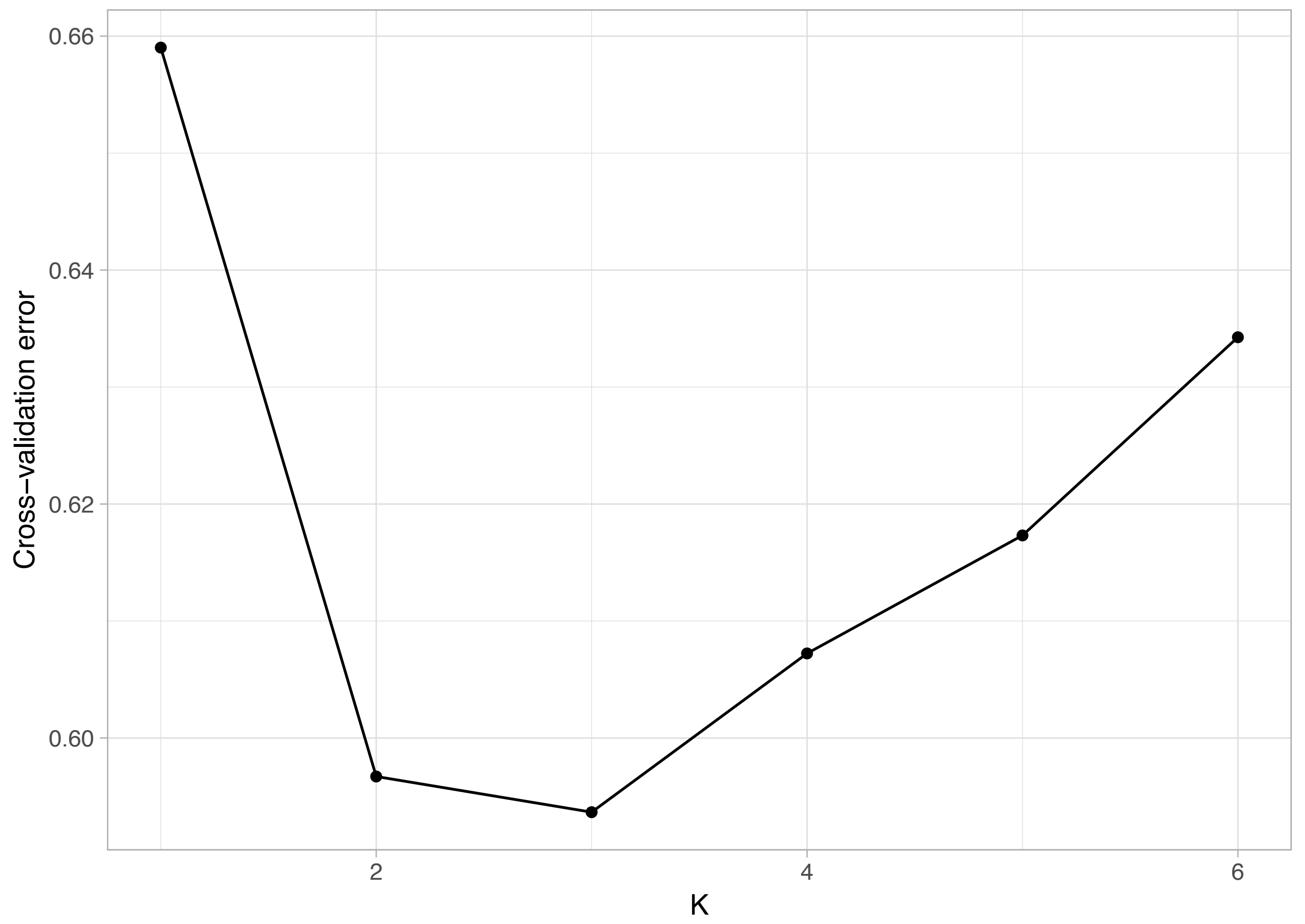
Figure S8: Log-likelihood distributions derived from 1,000,000 coalescent simulations for each of the demographic models. These distributions provide a means of distinguishing different scenarios independent of AIC and standard model selection.

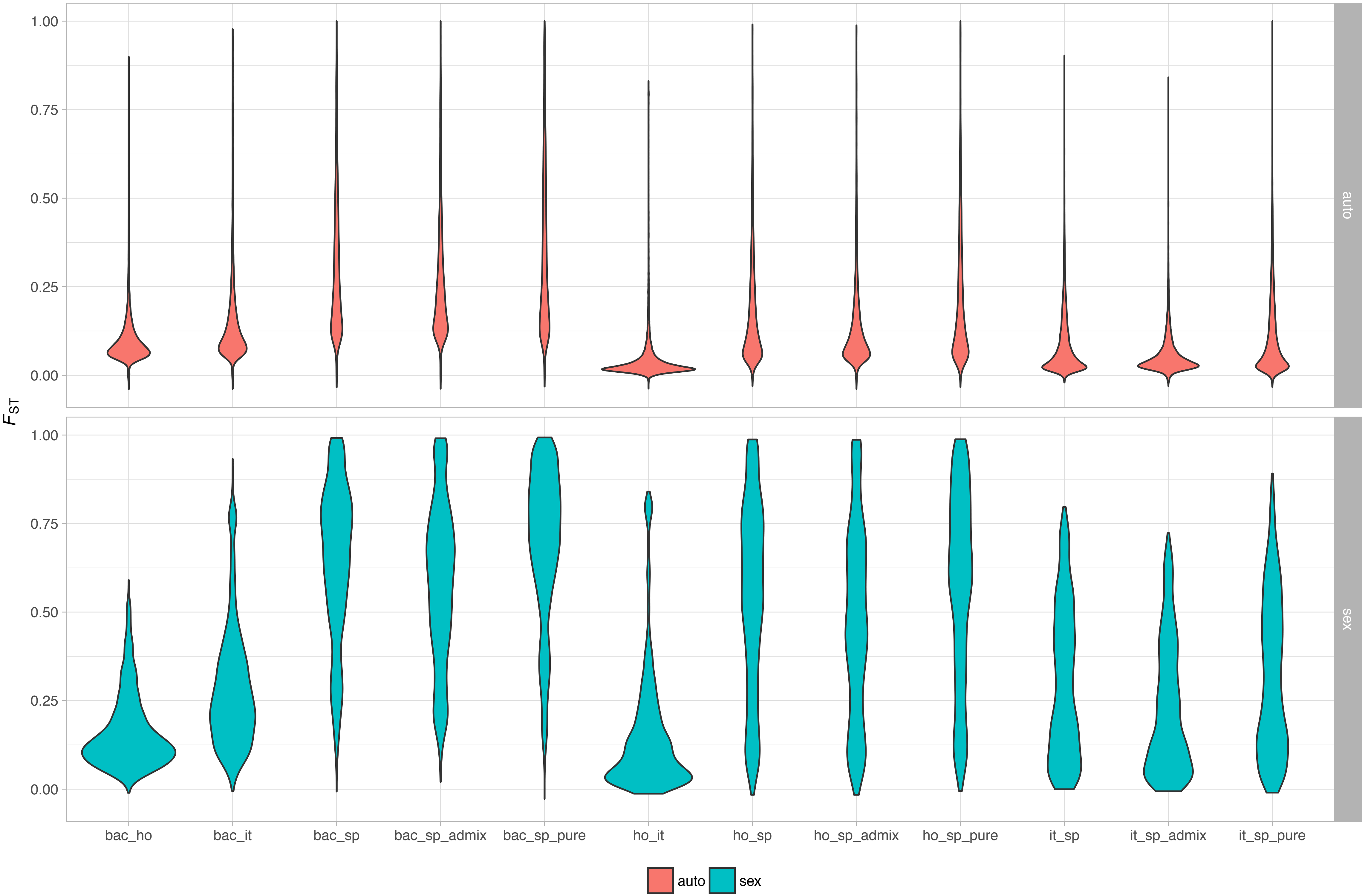
Figure S9: Genome-wide plot of iHS across the house sparrow genome. The full dataset has been downsampled to 20% (176, 701 SNPs) in order to aid visualisation. Dashed horizontal line represents significance at the level of $P = 0.001$.

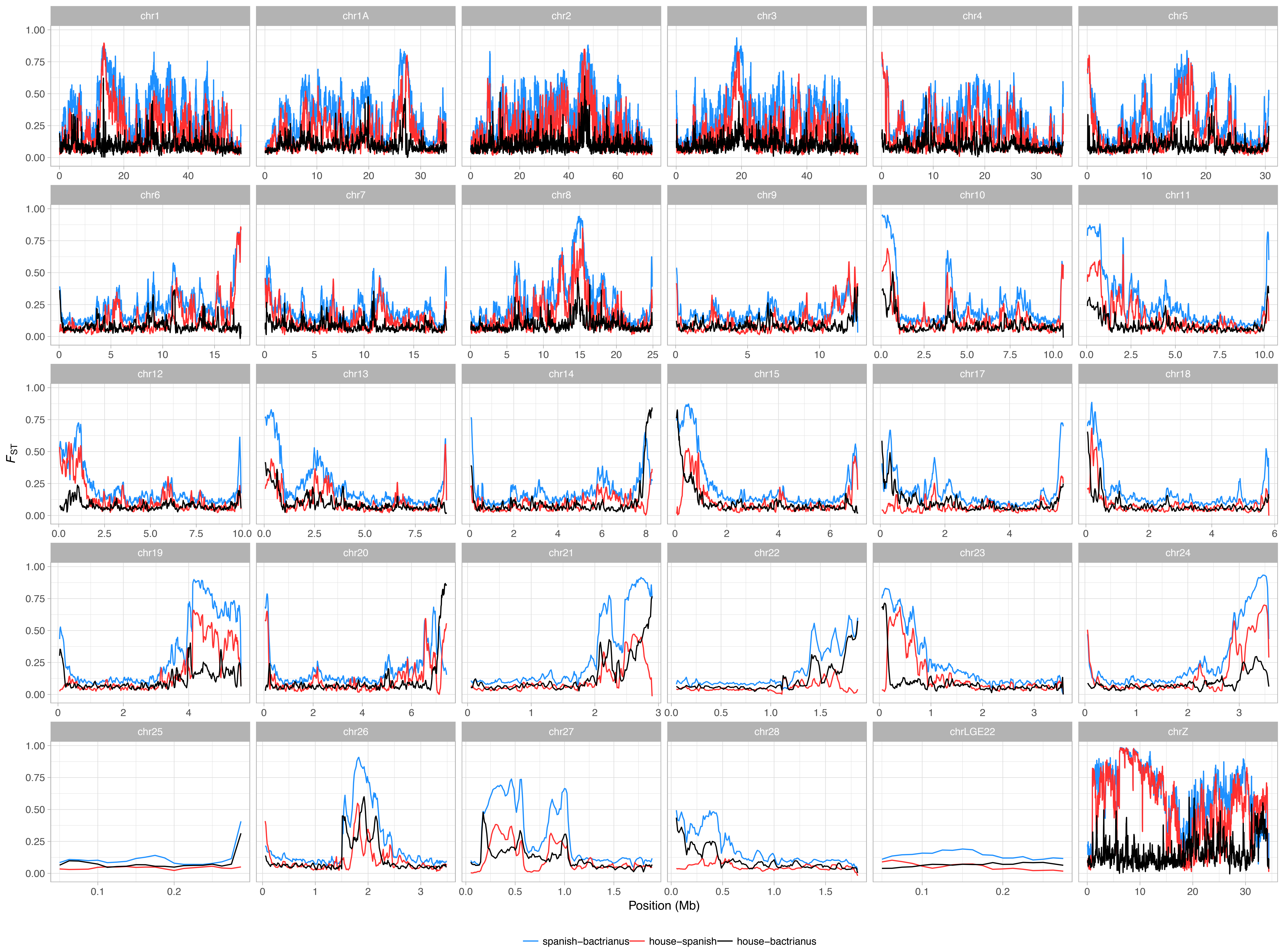
Figure S10: Closer examination of genomic divergence between house and Bactrianus sparrows for chromosomes 1, 2, 3 & 8 where the strongest signatures of selective sweeps were observed. For each chromosome, top panel – log10 (p) xpEHH (blue = background, red = outliers where $P < 0.0001$); second panel – mean absolute nucleotide divergence (d_{XY}); third panel – relative differentiation (F_{ST}); fourth panel – proportion of putatively introgressed sites per 100 Kb window (f_d) between the house and Spanish sparrow.

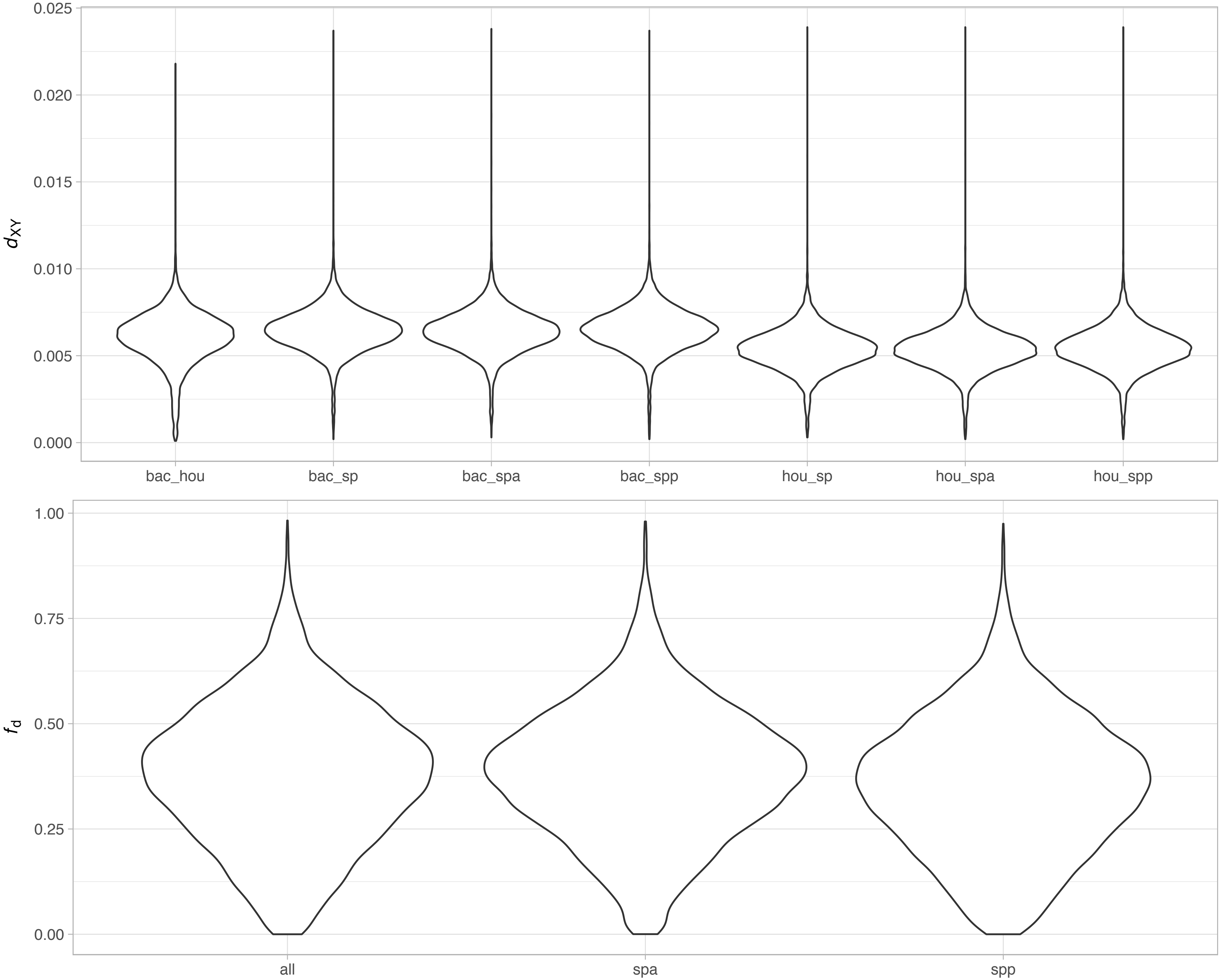

Percentage reads mapped to House reference
100.0
97.5
95.0
92.5
90.0
Bactrianus
House
Italian
Spanish
Tree
Species

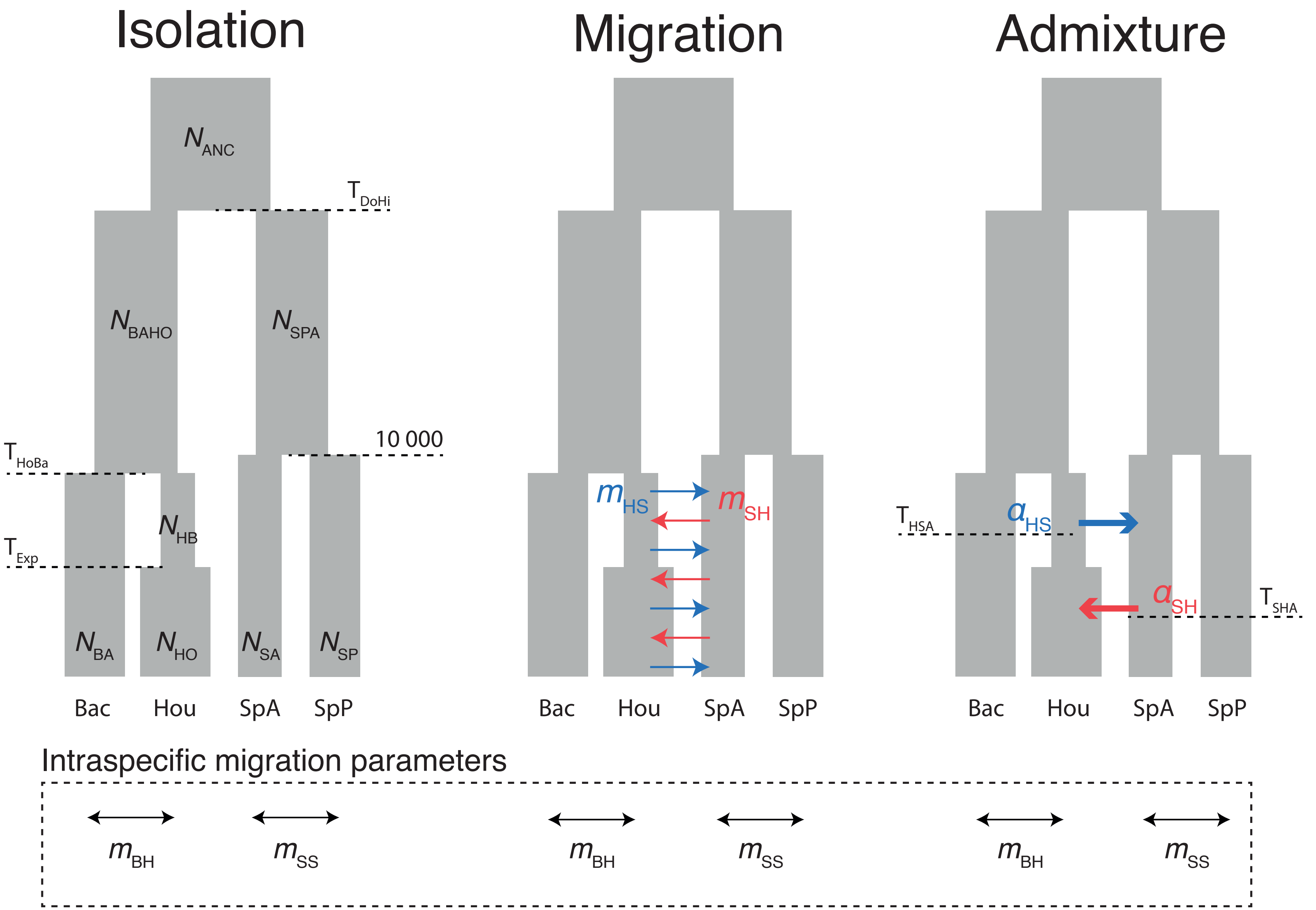

0.1
0.0
−0.1
−0.2
−0.3
PC2 (9.1%)
−0.1
0.0
0.1
PC1 (38.3%)
● Bactrianus
▲ House
■ Italian
+ Spanish
● Cape_Verde
● Crotone_Italy
● Golestan_Iran
● Guglionesi_Italy
● Kardeh_Iran
● Kazahkstan
● Lesina_Italy
● Northern_Norway
● Oslo_Norway
● Rimini_Italy
● Sales_France
● Sardinia
● Tenerife

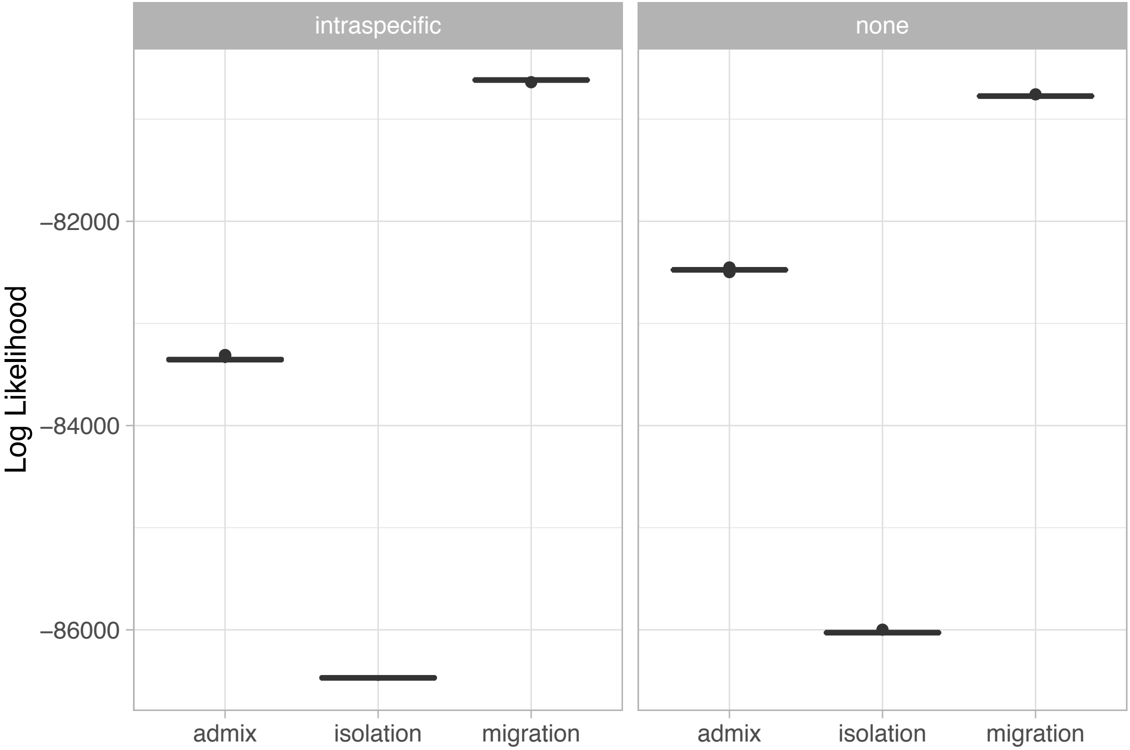

0.66
0.64
0.62
0.60
Cross−validation error
2
4
6
K

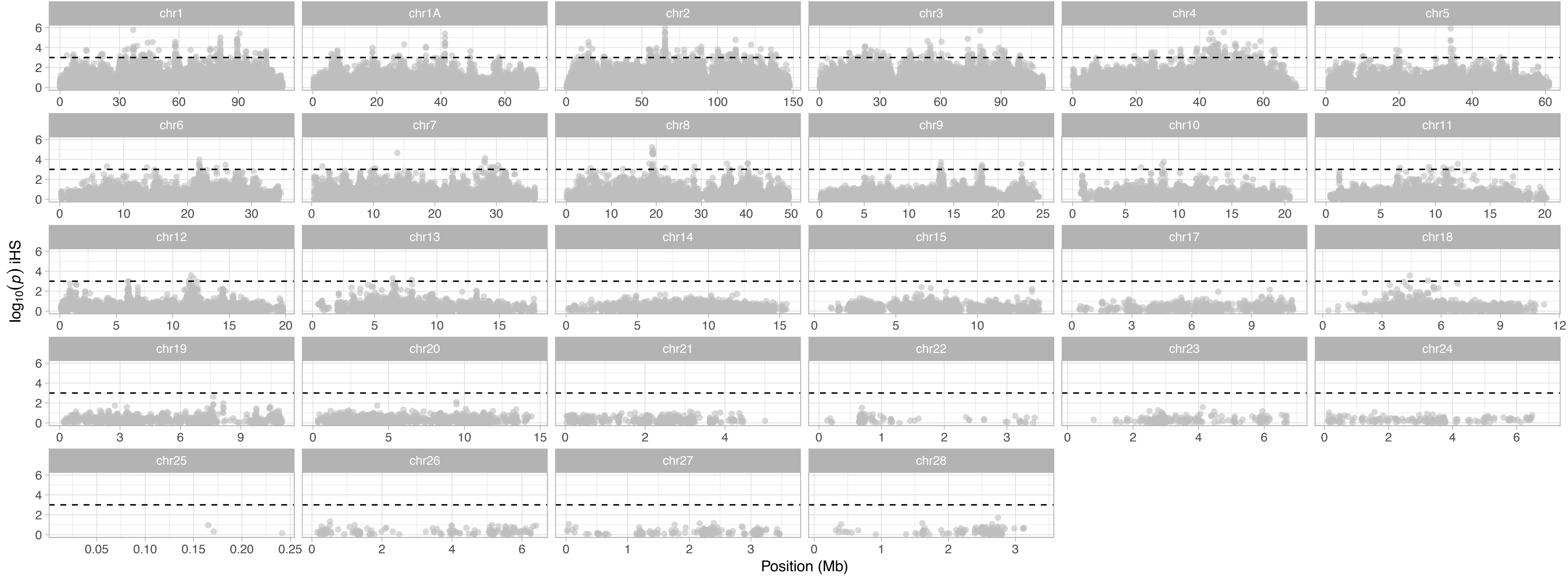

1.00
0.75
0.50
0.25
0.00
F_{ST}
auto
sex
bac_ho
bac_it
bac_sp
bac_sp_admix
bac_sp_pure
ho_it
ho_sp
ho_sp_admix
ho_sp_pure
it_sp
it_sp_admix
it_sp_pure
auto
sex

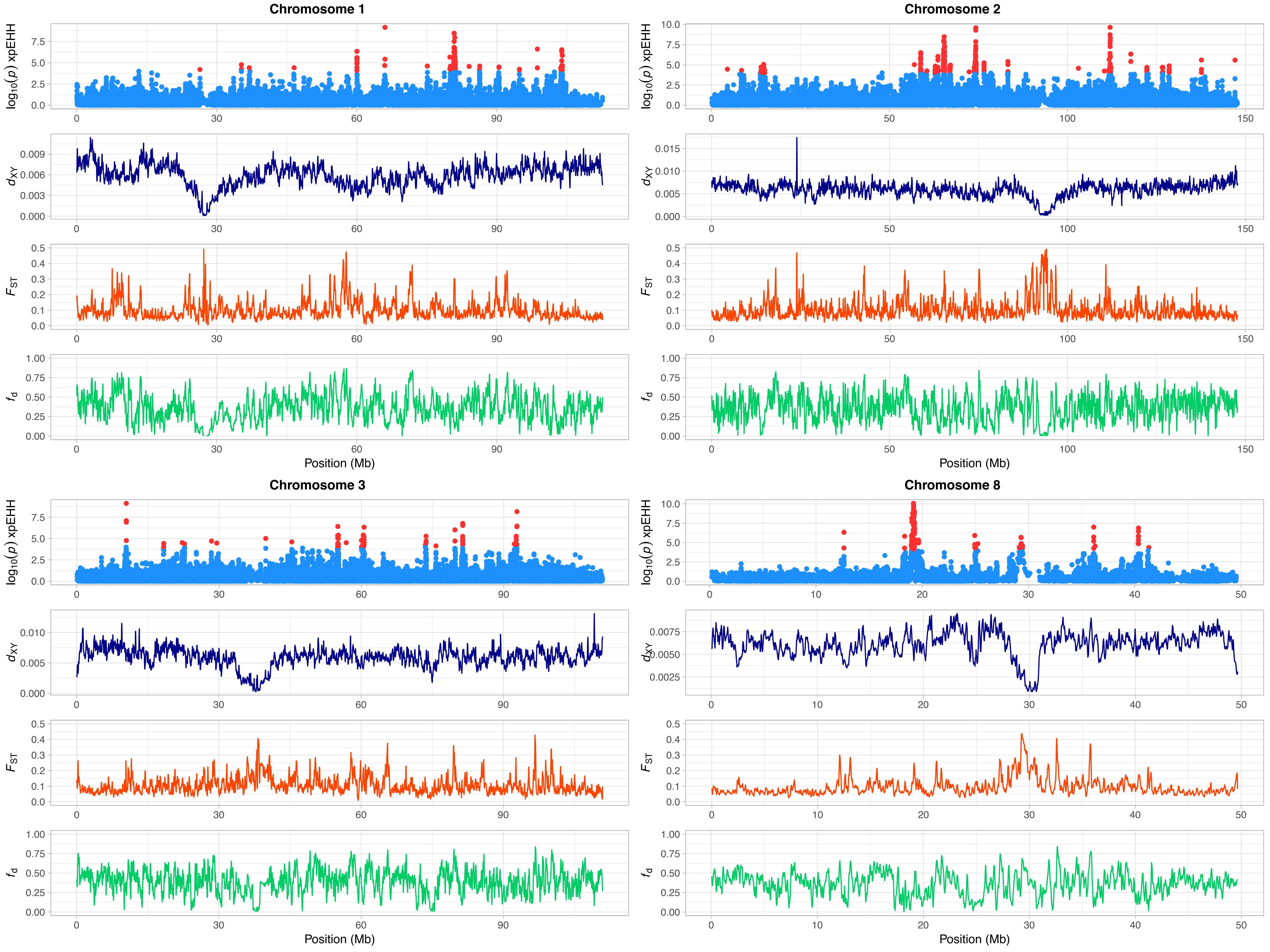

chr1
chr1A
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr17
chr18
chr19
chr20
chr21
chr22
chr23
chr24
chr25
chr26
chr27
chr28
chrLGE22
chrZ
1.00
0.75
0.50
0.25
0.00
F_{ST}
Position (Mb)
spanish−bactrianus
house−spanish
house−bactrianus


d_{XY}
0.025
0.020
0.015
0.010
0.005
0.000
bac_hou
bac_sp
bac_spa
bac_spp
hou_sp
hou_spa
hou_spp
f_d
1.00
0.75
0.50
0.25
0.00
all
spa
spp


Isolation
Migration
Admixture
N_{ANC}
T_{DoHi}
N_{BAHO}
N_{SPA}
10 000
T_{HoBa}
N_{HB}
T_{Exp}
N_{BA}
N_{HO}
N_{SA}
N_{SP}
Bac
Hou
SpA
SpP
m_{HS}
m_{SH}
Bac
Hou
SpA
SpP
T_{HSA}
a_{HS}
a_{SH}
T_{SHA}
Bac
Hou
SpA
SpP
Intraspecific migration parameters
m_{BH}
m_{SS}
m_{BH}
m_{SS}
m_{BH}
m_{SS}


intraspecific
none
Log Likelihood
−82000
−84000
−86000
admix
isolation
migration
admix
isolation
migration


chr1
chr1A
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr17
chr18
chr19
chr20
chr21
chr22
chr23
chr24
chr25
chr26
chr27
chr28
$\log_{10}(p)$ iHS
Position (Mb)
0
2
4
6
0
30
60
90
0
20
40
60
0
50
100
150
0
30
60
90
0
20
40
60
0
20
40
60
0
10
20
30
0
10
20
30
0
10
20
30
40
50
0
5
10
15
20
25
0
5
10
15
20
0
5
10
15
20
0
5
10
15
20
0
5
10
15
0
5
10
15
0
5
10
0
3
6
9
0
3
6
9
12
0
3
6
9
0
5
10
15
0
2
4
0
1
2
3
0
2
4
6
0
2
4
6
0.05
0.10
0.15
0.20
0.25
0
2
4
6
0
1
2
3
0
1
2
3


Chromosome 1
Chromosome 2
Chromosome 3
Chromosome 8
$\log_{10}(p)$ xpEHH
d_{XY}
F_{ST}
f_d
Position (Mb)
0
30
60
90
50
100
150
10
20
40
0.0
2.5
5.0
7.5
10.0
0.000
0.003
0.006
0.009
0.005
0.010
0.015
0.0025
0.0050
0.0075
0.1
0.2
0.3
0.4
0.5
0.00
0.25
0.50
0.75
1.00

Supplementary tables

Table S1: Breakdown of samples used.

Table S2: Priors for demographic inference.

Table S3: Model selection for SFS based demographic analyses using scenarios with and without intraspecific gene flow.

Table S4: Number of xpEHH outlier peaks per chromosome.

Table S5: Table of highest xpEHH outlier peaks per chromosome.

Table S6: Enriched GO terms for genes within 250 Kb of xpEHH peaks between the house and Bactrianus sparrow.