Comparative genomic analysis of co-occurring hybrid zones of house mouse parasites Pneumocystis murina and Syphacia obvelata using genome polarisation

Jan Petružela[1], Beate Nürnberger[2], Alexis Ribas[3], Georgios Koutsovoulos[4], Dagmar Cizkova[5], A Fornuskova[2], Tatiana Aghová[5], Mark Blaxter[6], Joëlle Goüy de Bellocq[5], and Stuart Baird[5]

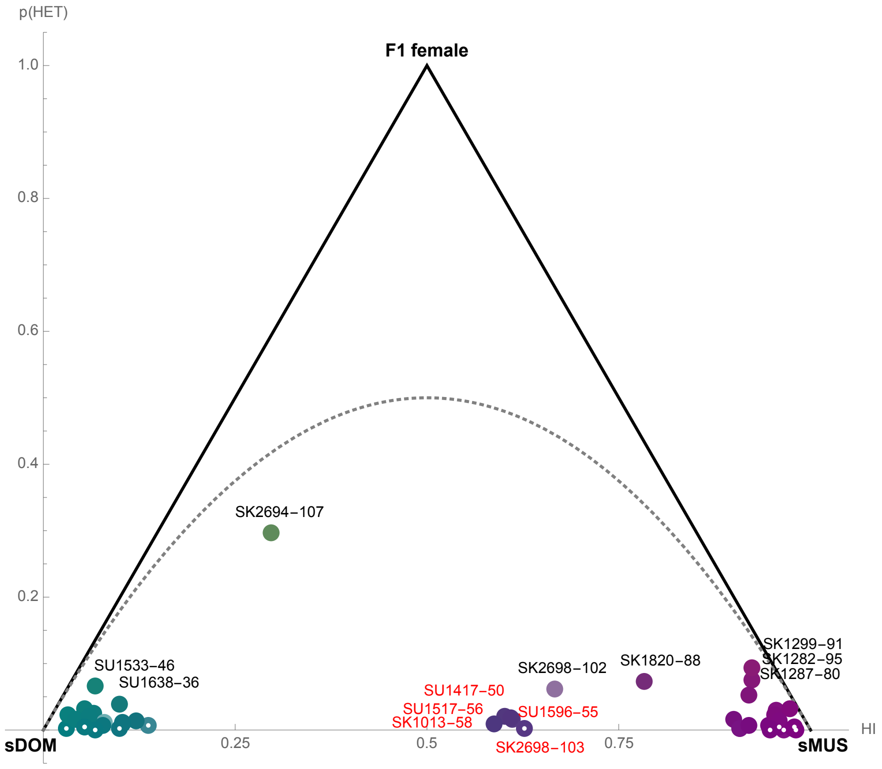
[1]Masaryk University Faculty of Science
[2]Institute of Vertebrate Biology CAS
[3]University of Barcelona
[4]Foundation of Research and Technology Hellas
[5]Institute of Vertebrate Biology Czech Academy of Sciences
[6]Wellcome Sanger Institute

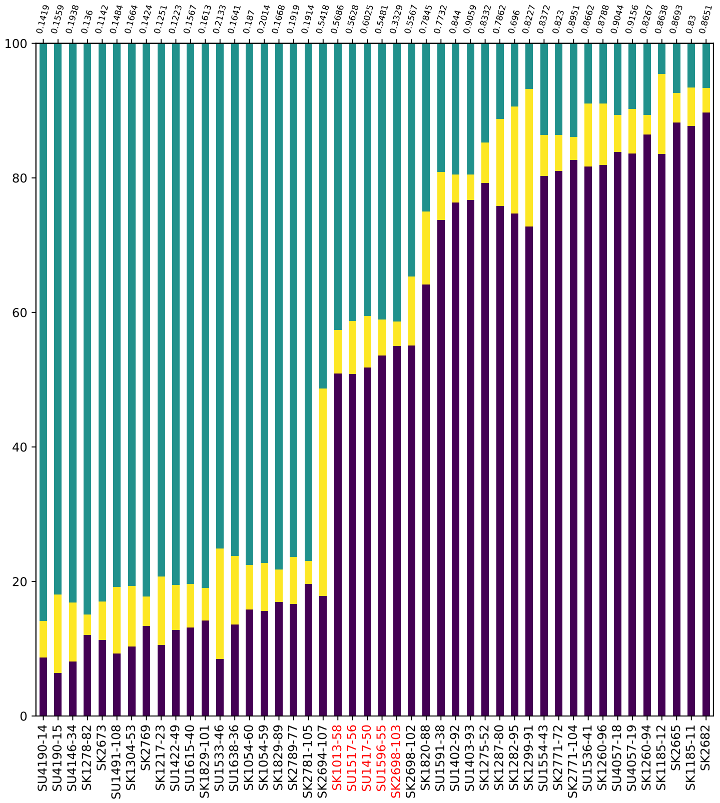
January 31, 2025

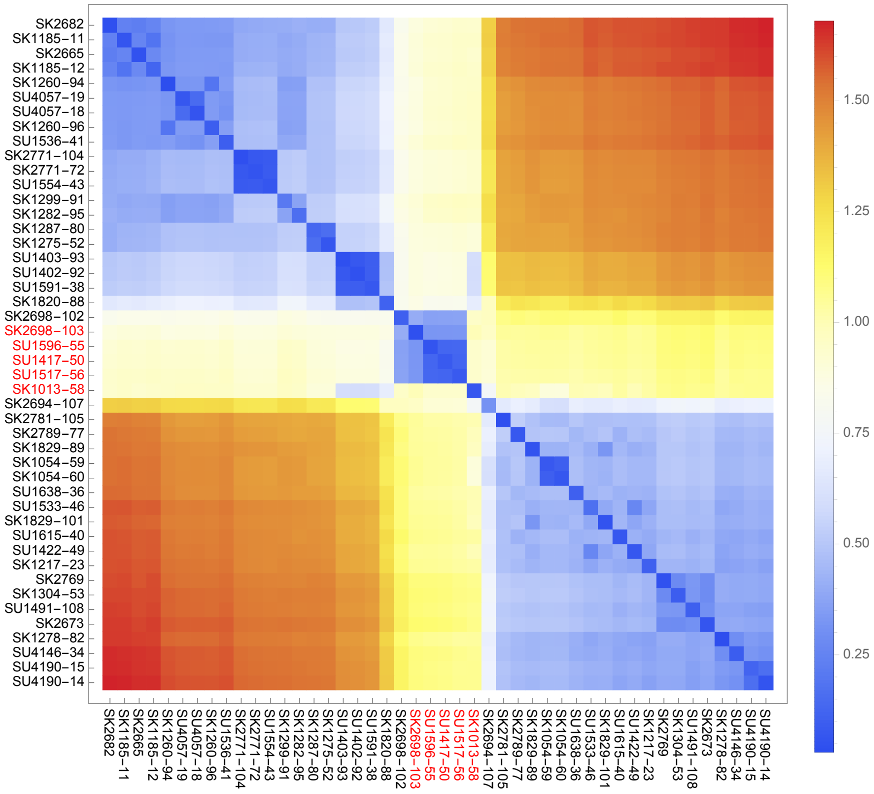
Abstract

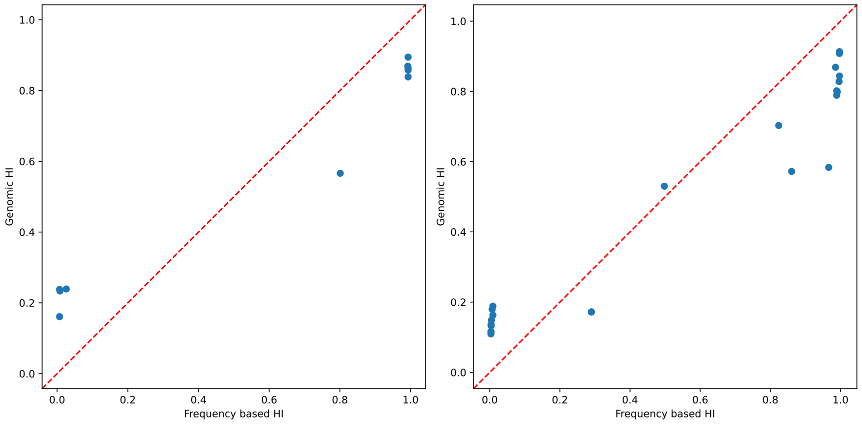
Parasites are expected to hybridise in similar ways to free living organisms, although this may be modified by their reduced genome architecture. Recombinant strains and taxa of hybrid origin can be studied in nature where hosts come into secondary contact. Here we apply genome-wide analysis to parasites from a contact previously characterised for many individuals using classical markers. The host contact is the European house mouse hybrid zone; the parasites are lung fungus Pneumocystis and gut pinworm Syphacia. The multilocus and multi-individual datasets are broadly consistent in scale and centering. Whole mitogenome comparisons confirm earlier suggestions that parasite divergence is low compared to their hosts, perhaps due to reduced genome stabilising selection. In the recombining genome we are able to show blocks of parasite genome of alternating host origin, including one Pneumocystis strain which appears to be an F3+ cross and one recombinant Syphacia strain found over multiple localities. Functional analyses of introgressing genes show enrichment for genes likely important for parasitic lifestyle.

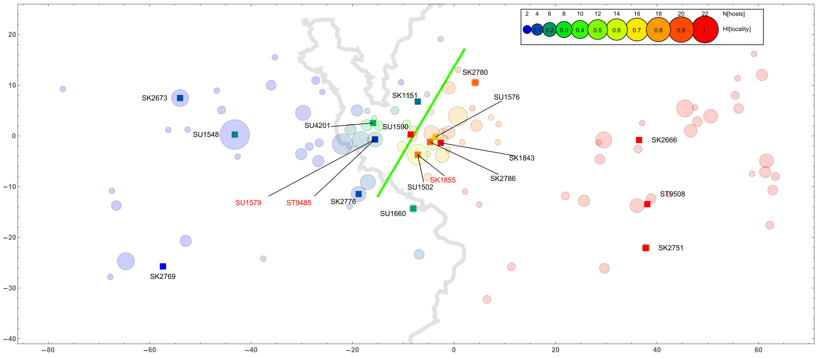
Hosted file

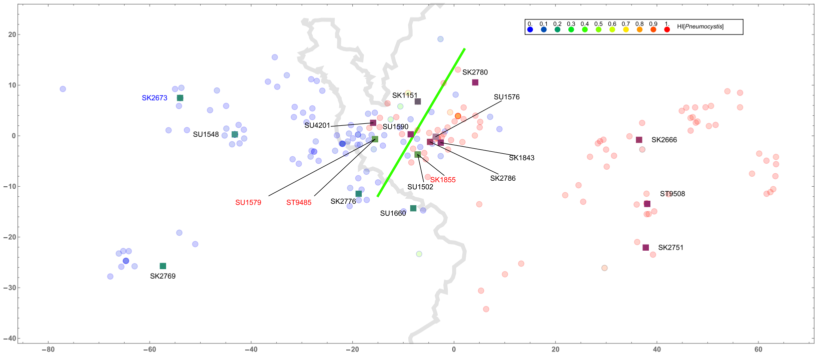

p(HET)
1.0
0.8
0.6
0.4
0.2
F1 female
SK2694−107
SU1533−46
SU1638−36
SK2698−102
SK1820−88
SK1299−91
SK1282−95
SK1287−80
SU1417−50
SU1517−56
SK1013−58
SU1596−55
SK2698−103
sDOM
0.25
0.5
0.75
sMUS
HI

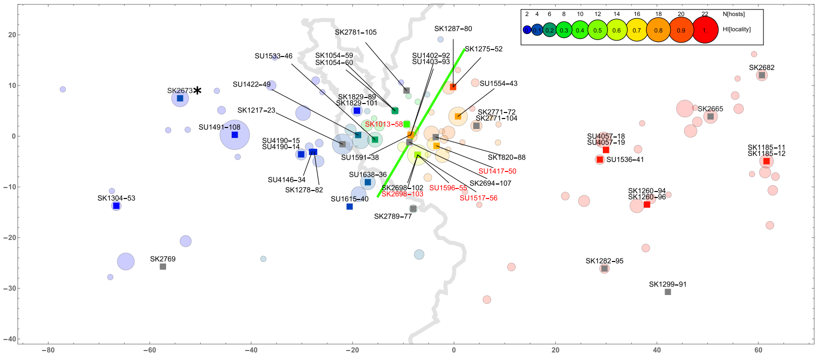

0.1419
0.1559
0.1938
0.136
0.1142
0.1484
0.1664
0.1424
0.1251
0.1223
0.1567
0.1613
0.2133
0.1641
0.187
0.2014
0.1668
0.1919
0.1914
0.5418
0.5686
0.5628
0.6025
0.5481
0.3329
0.5567
0.7845
0.7732
0.844
0.9059
0.8332
0.7862
0.696
0.8227
0.8372
0.823
0.8951
0.8662
0.8788
0.9044
0.9156
0.8267
0.8638
0.8693
0.83
0.8651
100
80
60
40
20
0
SU4190-14
SU4190-15
SU4146-34
SK1278-82
SK2673
SU1491-108
SK1304-53
SK2769
SK1217-23
SU1422-49
SU1615-40
SK1829-101
SU1533-46
SU1638-36
SK1054-60
SK1054-59
SK1829-89
SK2789-77
SK2781-105
SK2694-107
SK1013-58
SU1517-56
SU1417-50
SU1596-55
SK2698-103
SK2698-102
SK1820-88
SU1591-38
SU1402-92
SU1403-93
SK1275-52
SK1287-80
SK1282-95
SK1299-91
SU1554-43
SK2771-72
SK2771-104
SU1536-41
SK1260-96
SU4057-18
SU4057-19
SK1260-94
SK1185-12
SK2665
SK1185-11
SK2682

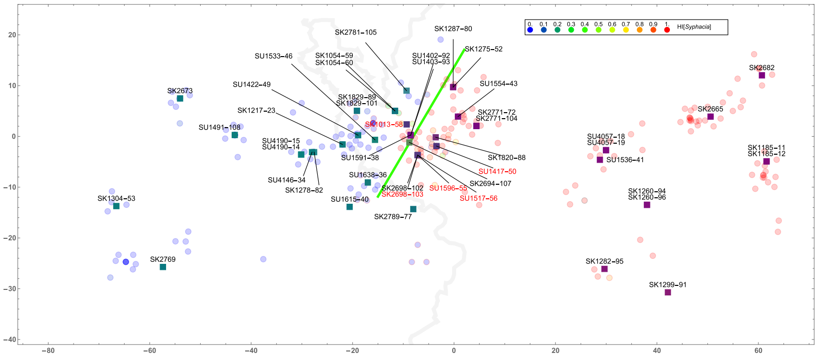

SK2682
SK1185−11
SK2665
SK1185−12
SK1260−94
SU4057−19
SU4057−18
SK1260−96
SU1536−41
SK2771−104
SK2771−72
SU1554−43
SK1299−91
SK1282−95
SK1287−80
SK1275−52
SU1403−93
SU1402−92
SU1591−38
SK1820−88
SK2698−102
SK2698−103
SU1596−55
SU1417−50
SU1517−56
SK1013−58
SK2694−107
SK2781−105
SK2789−77
SK1829−89
SK1054−59
SK1054−60
SU1638−36
SU1533−46
SK1829−101
SU1615−40
SU1422−49
SK1217−23
SK2769
SK1304−53
SU1491−108
SK2673
SK1278−82
SU4146−34
SU4190−15
SU4190−14
1.50
1.25
1.00
0.75
0.50
0.25


1.0
0.8
0.6
0.4
0.2
0.0
Genomic HI
0.0
0.2
0.4
0.6
0.8
1.0
Frequency based HI
1.0
0.8
0.6
0.4
0.2
0.0
Genomic HI
0.0
0.2
0.4
0.6
0.8
1.0
Frequency based HI

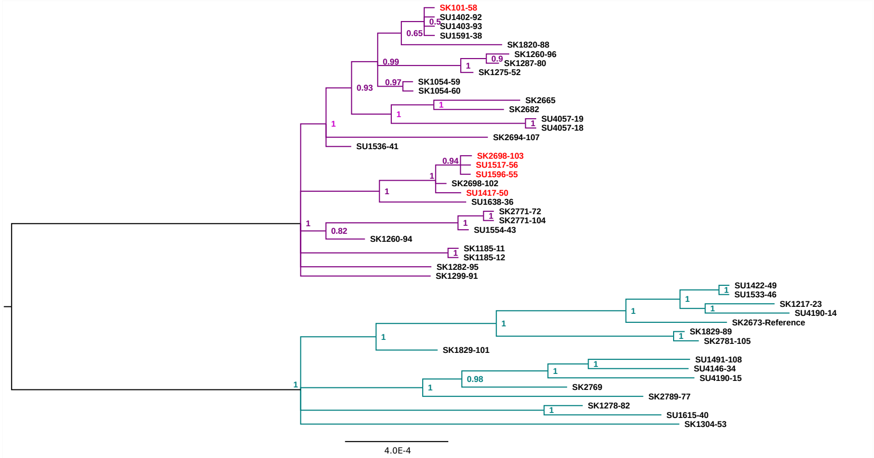

2 4 6 8 10 12 14 16 18 20 22 N[hosts]
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 HI[locality]
SK2780
SK1151
SU1576
SK2673
SU4201
SU1590
SU1548
SK2666
SK1843
SK1855
SK2786
SU1502
SU1579
ST9485
SK2776
ST9508
SU1660
SK2751
SK2769
20
10
0
−10
−20
−30
−40
−80
−60
−40
−20
0
20
40
60

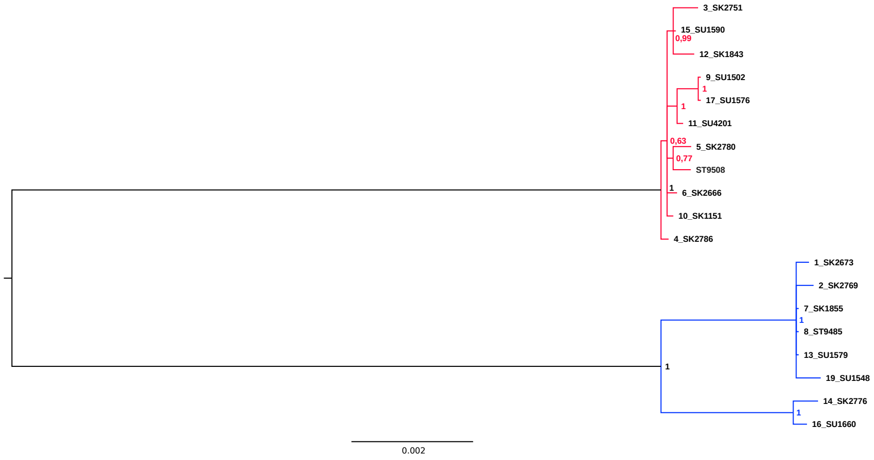

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 HI[Pneumocystis]
SK2780
SK1151
SU1576
SK2673
SU4201
SU1590
SU1548
SK2666
SK1843
SK1855
SK2786
SU1502
SU1579
ST9485
SK2776
ST9508
SU1660
SK2751
SK2769
20
10
0
−10
−20
−30
−40
−80
−60
−40
−20
0
20
40
60

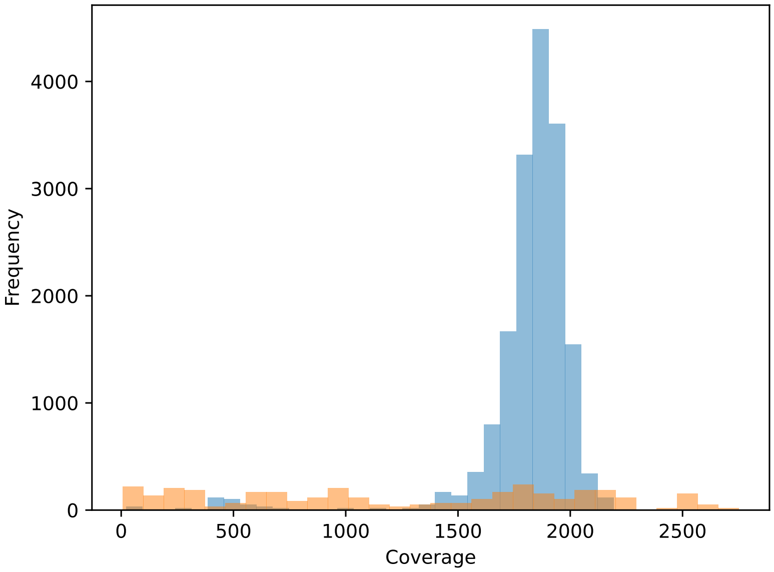

2 4 6 8 10 12 14 16 18 20 22 N[hosts]
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 HI[locality]
SK2781−105
SK1287−80
SK1275−52
SU1533−46
SK1054−59
SK1054−60
SU1402−92
SU1403−93
SK2682
SU1554−43
SK2673
SU1422−49
SK1829−89
SK1829−101
SK2665
SK1217−23
SK1013−58
SK2771−72
SK2771−104
SU1491−108
SU4057−18
SU4057−19
SK1185−11
SK1185−12
SU4190−15
SU4190−14
SK1820−88
SU1591−38
SU1536−41
SU1417−50
SK1638−36
SU4146−34
SK2694−107
SK1278−82
SK2698−102
SK2698−103
SU1596−55
SK1304−53
SU1615−40
SU1517−56
SK1260−94
SK1260−96
SK2789−77
SK2769
SK1282−95
SK1299−91
20
10
0
−10
−20
−30
−40
−80
−60
−40
−20
0
20
40
60

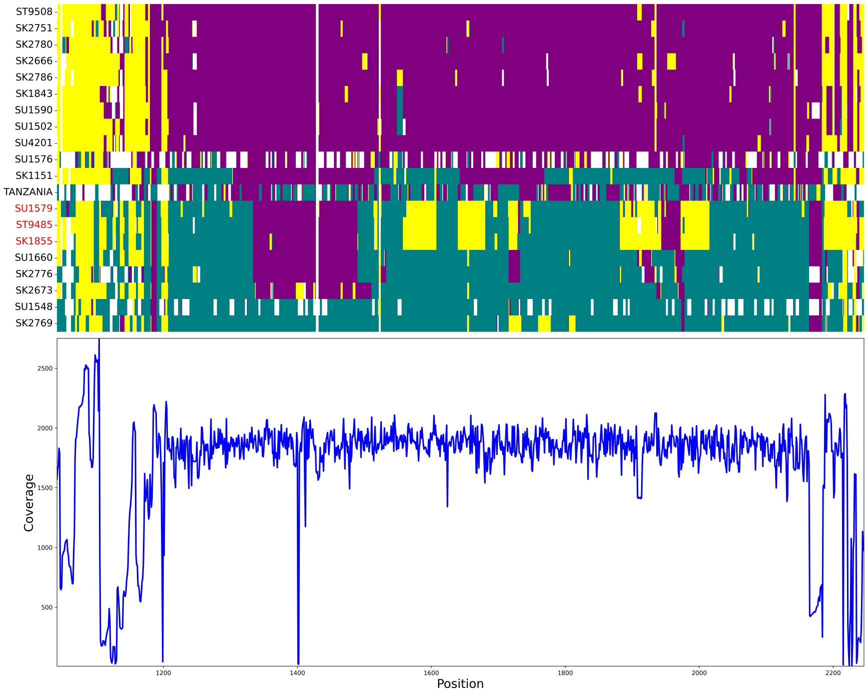

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 HI[Syphacia]
SK2781−105
SK1287−80
SK1275−52
SU1533−46
SK1054−59
SK1054−60
SU1402−92
SU1403−93
SK2682
SU1554−43
SK2673
SU1422−49
SK1829−89
SK1829−101
SK2665
SK1217−23
SK1013−58
SK2771−72
SK2771−104
SU1491−108
SU4057−18
SU4057−19
SK1185−11
SK1185−12
SU4190−15
SU4190−14
SK1820−88
SU1591−38
SU1536−41
SU1417−50
SK1638−36
SU4146−34
SK2694−107
SK1278−82
SK2698−102
SK2698−103
SU1596−55
SK1304−53
SU1615−40
SU1517−56
SK1260−94
SK1260−96
SK2789−77
SK2769
SK1282−95
SK1299−91
20
10
0
−10
−20
−30
−40
−80
−60
−40
−20
0
20
40
60

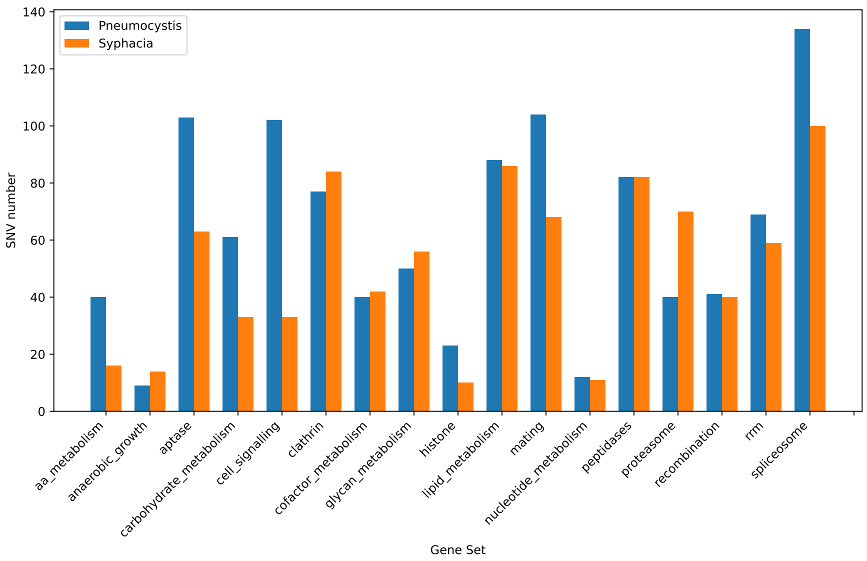

SK2751
1
SK2666
0.93
SK2780
SU4201
0.99
SK1151
SU1576
1
SU1590
SK1843
SU1502
SK2786
ST9508
SU1660
0.84
SK2776
1
SU1579
0.87
SK2673
ST9485
0.97
SK1855
0.98
SU1548
SK2769
NC_020332_reference
4.0E-4

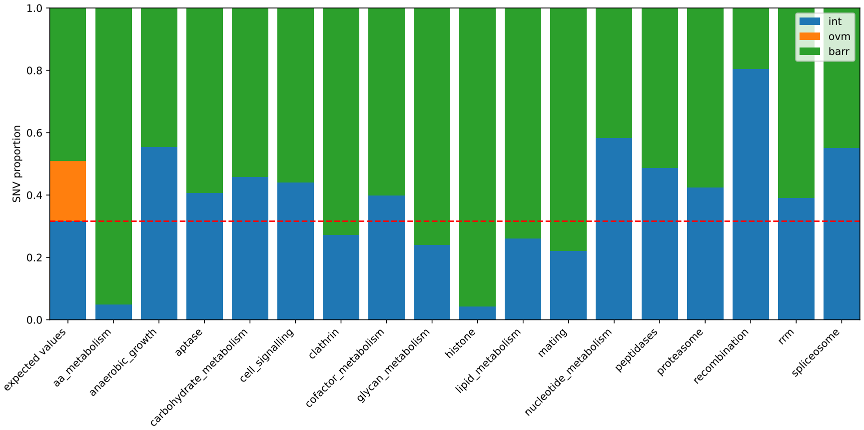

SK101-58
SU1402-92
0.5
SU1403-93
0.65
SU1591-38
SK1820-88
0.99
0.9
SK1260-96
SK1287-80
SK1275-52
0.97
SK1054-59
SK1054-60
0.93
SK2665
SK2682
SU4057-19
SU4057-18
SK2694-107
SU1536-41
0.94
SK2698-103
SU1517-56
SU1596-55
SK2698-102
SU1417-50
SU1638-36
SK2771-72
SK2771-104
SU1554-43
0.82
SK1260-94
SK1185-11
SK1185-12
SK1282-95
SK1299-91
SU1422-49
SU1533-46
SK1217-23
SU4190-14
SK2673-Reference
SK1829-89
SK2781-105
SK1829-101
SU1491-108
SU4146-34
SU4190-15
0.98
SK2769
SK2789-77
SK1278-82
SU1615-40
SK1304-53
4.0E-4

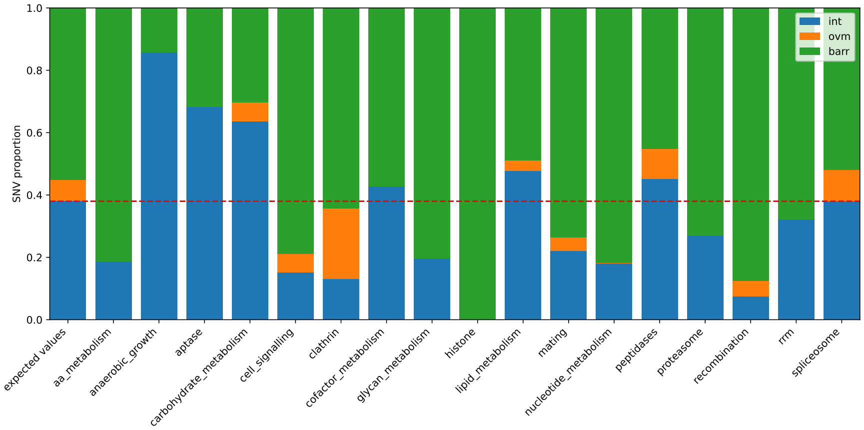

3_SK2751
15_SU1590
0,99
12_SK1843
9_SU1502
1
17_SU1576
1
11_SU4201
0,63
5_SK2780
0,77
ST9508
1
6_SK2666
10_SK1151
4_SK2786
1_SK2673
2_SK2769
7_SK1855
1
8_ST9485
13_SU1579
19_SU1548
1
14_SK2776
1
16_SU1660
0.002

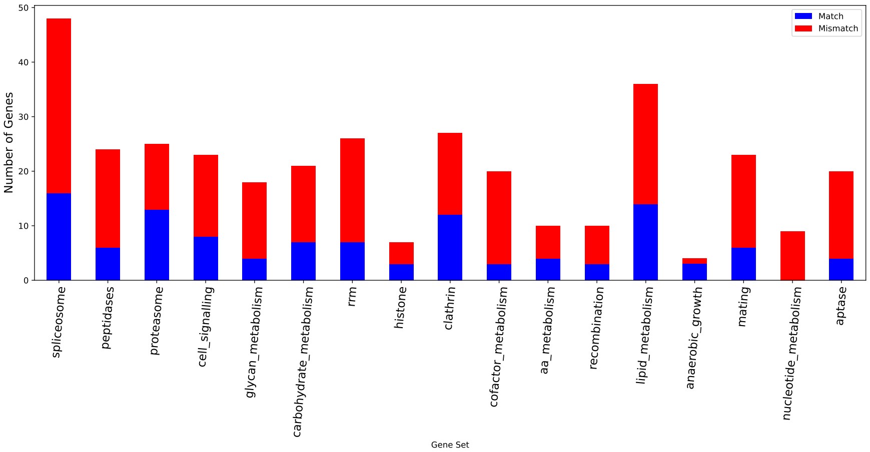

Frequency
4000
3000
2000
1000
0
0
500
1000
1500
2000
2500
Coverage


ST9508
SK2751
SK2780
SK2666
SK2786
SK1843
SU1590
SU1502
SU4201
SU1576
SK1151
TANZANIA
SU1579
ST9485
SK1855
SU1660
SK2776
SK2673
SU1548
SK2769
Coverage
2500
2000
1500
1000
500
1200
1400
1600
1800
2000
2200
Position


Pneumocystis
Syphacia
SNV number
0
20
40
60
80
100
120
140
aa_metabolism
anaerobic_growth
aptase
carbohydrate_metabolism
cell_signalling
clathrin
cofactor_metabolism
glycan_metabolism
histone
lipid_metabolism
mating
nucleotide_metabolism
peptidases
proteasome
recombination
rrm
spliceosome
Gene Set


int
ovm
barr
SNV proportion
0.0
0.2
0.4
0.6
0.8
1.0
expected values
aa_metabolism
anaerobic_growth
aptase
carbohydrate_metabolism
cell_signalling
clathrin
cofactor_metabolism
glycan_metabolism
histone
lipid_metabolism
mating
nucleotide_metabolism
peptidases
proteasome
recombination
rrm
spliceosome


int
ovm
barr
SNV proportion
0.0
0.2
0.4
0.6
0.8
1.0
expected values
aa_metabolism
anaerobic_growth
aptase
carbohydrate_metabolism
cell_signalling
clathrin
cofactor_metabolism
glycan_metabolism
histone
lipid_metabolism
mating
nucleotide_metabolism
peptidases
proteasome
recombination
rrm
spliceosome



Hosted file

Table1_assembly_stats.xlsx available at https://authorea.com/users/886532/articles/1264557-comparative-genomic-analysis-of-co-occurring-hybrid-zones-of-house-mouse-parasites-pneumocystis-murina-and-syphacia-obvelata-using-genome-polarisation

Hosted file

Table2_genic_content.xlsx available at https://authorea.com/users/886532/articles/1264557-comparative-genomic-analysis-of-co-occurring-hybrid-zones-of-house-mouse-parasites-pneumocystis-murina-and-syphacia-obvelata-using-genome-polarisation

Hosted file

Table3_category_content.xlsx available at https://authorea.com/users/886532/articles/1264557-comparative-genomic-analysis-of-co-occurring-hybrid-zones-of-house-mouse-parasites-pneumocystis-murina-and-syphacia-obvelata-using-genome-polarisation

Hosted file

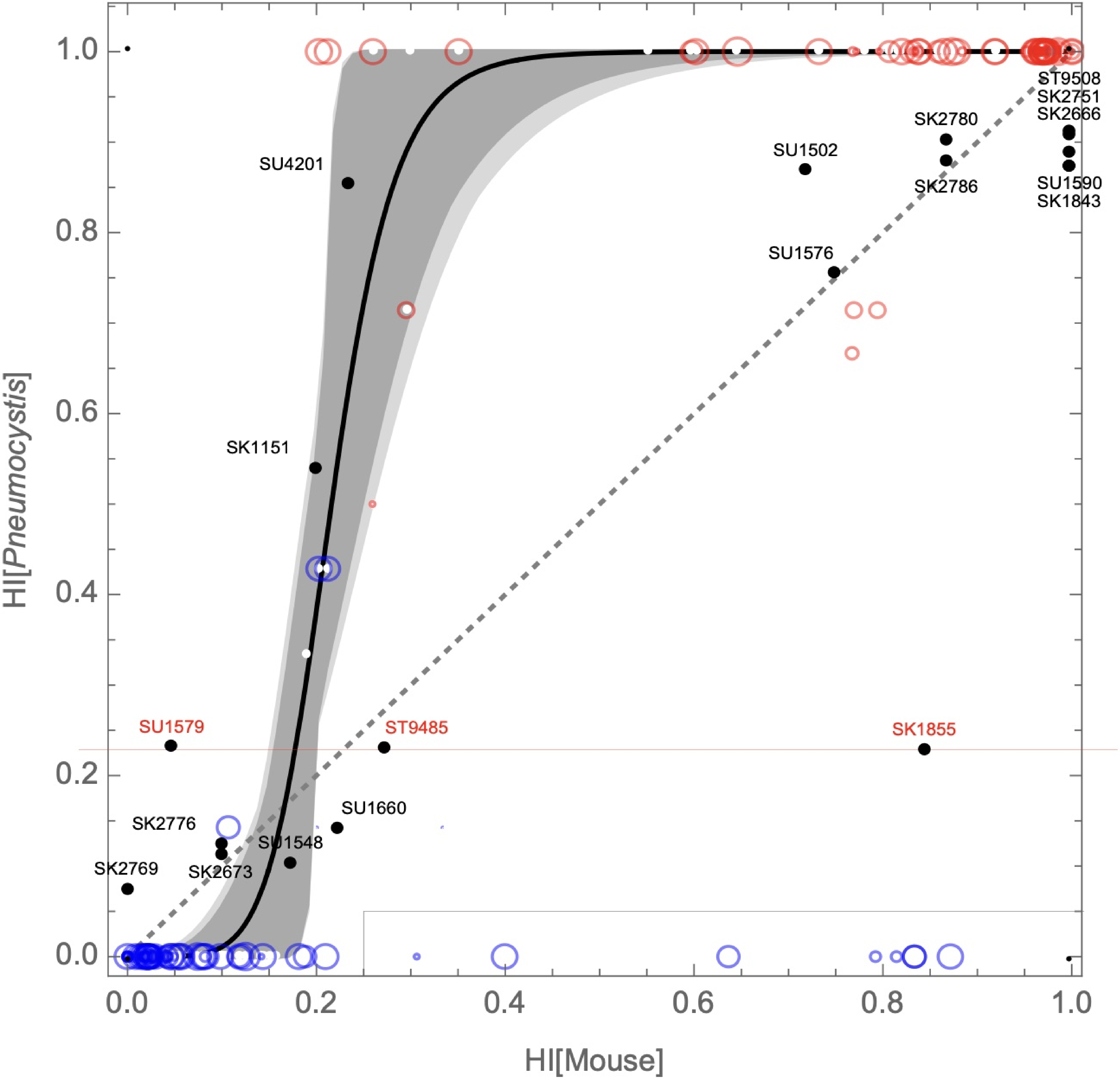
Table4_pneumocystis_contingency_table.xlsx available at https://authorea.com/users/886532/articles/1264557-comparative-genomic-analysis-of-co-occurring-hybrid-zones-of-house-mouse-parasites-pneumocystis-murina-and-syphacia-obvelata-using-genome-polarisation

Hosted file

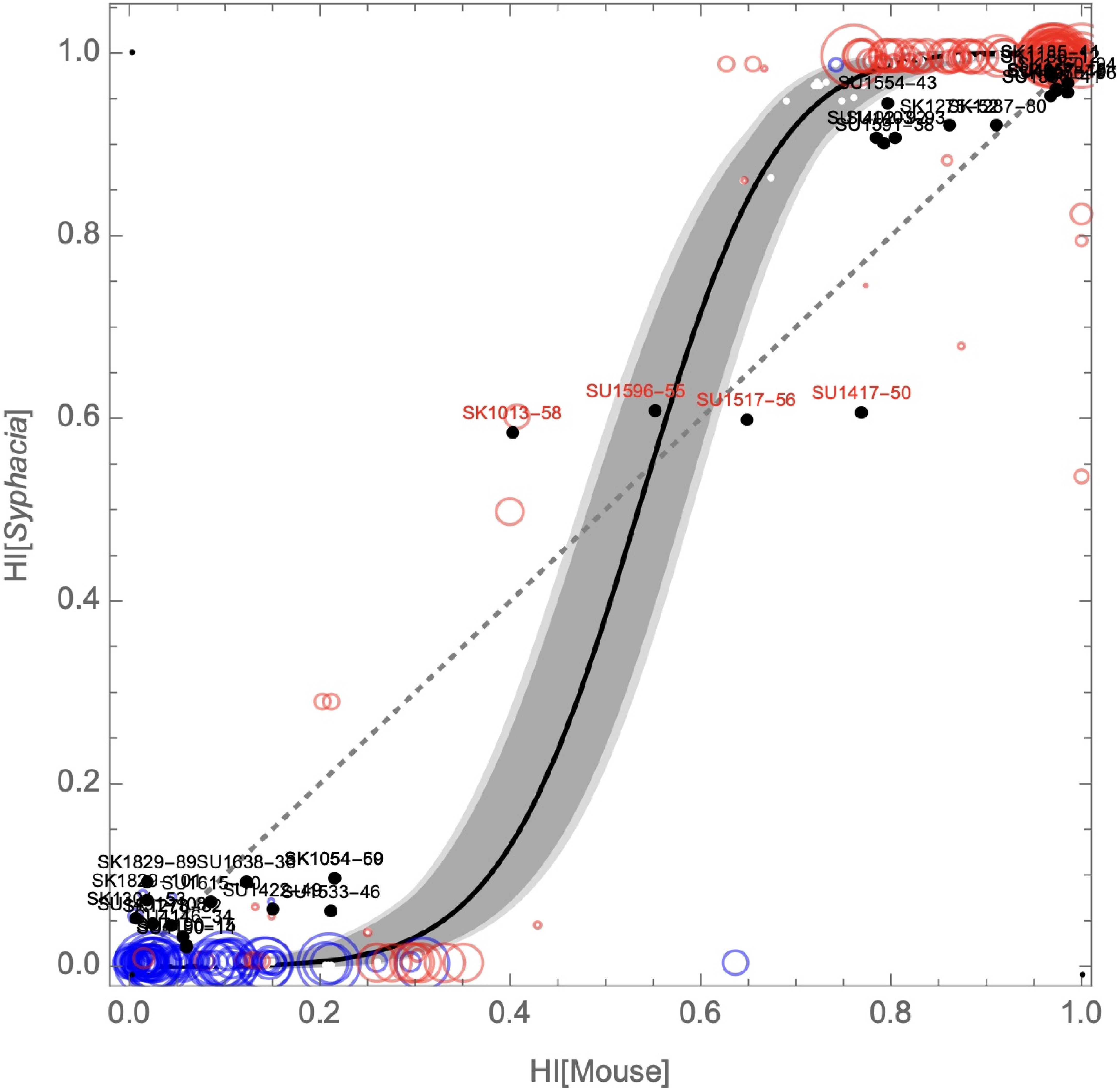
Table5_syphacia_contingency_table.xlsx available at https://authorea.com/users/886532/articles/1264557-comparative-genomic-analysis-of-co-occurring-hybrid-zones-of-house-mouse-parasites-pneumocystis-murina-and-syphacia-obvelata-using-genome-polarisation

Hosted file

Table6_intbarr_ratios.xlsx available at https://authorea.com/users/886532/articles/1264557-comparative-genomic-analysis-of-co-occurring-hybrid-zones-of-house-mouse-parasites-pneumocystis-murina-and-syphacia-obvelata-using-genome-polarisation



HI[*Pneumocystis*]
HI[Mouse]
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
ST9508
SK2751
SK2666
SK2780
SU1502
SU4201
SK2786
SU1590
SK1843
SU1576
SK1151
SU1579
ST9485
SK1855
SU1660
SK2776
SU1548
SK2769
SK2673


HI[*Syphacia*]
HI[Mouse]
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
SK1185-11
SU1554-43
SU1591-38
SK1287-80
SK1013-58
SU1596-55
SU1517-56
SU1417-50
SK1829-89
SU1638-36
SK1054-69
SK1829-101
SU1615-40
SU1533-46
SU1146-34