

Appendix

Online resources, dataset, additional histograms

A Online resources

The following resources are available online:

- The **interactive webtool MoD Map**, that includes the 2-dimensional Molecular Distance Maps in Figures 2, 3, 5, 6, 7, 8, is available for use directly at

<http://www.csd.uwo.ca/MoDMap/interintra/>

In these interactive Molecular Distance Maps it is possible to zoom in, select a single point in order to see information about the corresponding DNA sequence and the FCGR of that DNA sequence, and it is possible to display the distance between any two points.

- The **distance matrices**, that were generated using each of the six distances (DSSIM, descriptor, Euclidean, Manhattan, Pearson, approximate information) and the datasets used (First experiment - a single complete chromosome from each of the six organisms; Second experiment - fragments randomly sampled from all chromosomes of the six organisms; Preview experiment - fragments randomly sampled from *H. sapiens* and *M. musculus*) are available as plain-text, comma-separated values (csv) at

https://github.com/rallis/intraSupplemental_Material.

In the first experiment distance matrix, the row and column numbers 1–234 correspond to *H. sapiens* fragments, 235–264 correspond to *E. coli* fragments, 265–274 correspond to *S. cerevisiae* fragments, 275–475 correspond to *A. thaliana* fragments, 476–496 correspond to *P. falciparum* fragments, and 497–508 correspond to *P. furiosus* fragments.

In the second experiment distance matrix, the row and column numbers 1–240 correspond to *H. sapiens* fragments, 241–270 correspond to *E. coli* fragments, 271–343 correspond to *S. cerevisiae* fragments, 344–393 correspond to *A. thaliana* fragments, 394–514 correspond to *P. falciparum* fragments, and 515–526 correspond to *P. furiosus* fragments.

In the preview experiment distance matrix, the row and column numbers 1–240 correspond to *H. sapiens* fragments, 241–450 correspond to *M. musculus* fragments.

- The **code**, that was used to generate the Molecular Distance Maps, histograms, and CGR images, is available at

https://github.com/rallis/intraSupplemental_Material.

The code was written and executed using Wolfram Mathematica 9.

B Dataset

The datasets for all the experiments can be obtained from

<http://www.ncbi.nlm.nih.gov/>.

Tables 1, 2, and 3 in this Appendix list the NCBI accession numbers, and the names of the chromosomes (or genomes) for the three experiments, respectively. In order to download one of the sequences, the following link can be used, where <NCBI Accession> has to be replaced by the corresponding NCBI accession number listed in the tables:

<http://www.ncbi.nlm.nih.gov/nuccore/<NCBI Accession>>.

TABLE 1: The dataset for the first experiment

NCBI Accession	Name
NC_000021.8	<i>H. sapiens</i> chrom. 21, GRCh37.p13 Primary Assembly
NC_000913.3	<i>E. coli</i> str. K-12 substr. MG1655, complete genome
NC_001136.10	<i>S. cerevisiae</i> S288c chrom. IV
NC_003070.9	<i>A. thaliana</i> chrom. 1
NC_004317.2	<i>P. falciparum</i> 3D7 chrom. 14
NC_018092.1	<i>P. furiosus</i> COM1 chrom., complete genome

TABLE 2: The dataset for the second experiment

NCBI Accession	Name	Sampled fragments	Total fragment nr.
NC_000001.10	<i>H. sapiens</i> chrom. 1, GRCh37.p13	{249,640,856,188,243,1310,916,197,869,936}	1501
NC_000002.11	<i>H. sapiens</i> chrom. 2, GRCh37.p13	{221,505,458,916,84,1464,838,627,1336,955}	1588
NC_000003.11	<i>H. sapiens</i> chrom. 3, GRCh37.p13	{1051,725,1231,771,496,336,297,1202,660,836}	1298
NC_000004.11	<i>H. sapiens</i> chrom. 4, GRCh37.p13	{785,116,684,272,98,805,661,612,75,716}	1251
NC_000005.9	<i>H. sapiens</i> chrom. 5, GRCh37.p13	{623,783,563,717,747,156,47,6,739,366}	1184
NC_000006.11	<i>H. sapiens</i> chrom. 6, GRCh37.p13	{815,166,192,147,646,452,419,460,1066,143}	1115
NC_000007.13	<i>H. sapiens</i> chrom. 7, GRCh37.p13	{897,611,847,498,234,910,166,793,438,873}	1035
NC_000008.10	<i>H. sapiens</i> chrom. 8, GRCh37.p13	{740,822,532,432,96,482,316,280,173,861}	952
NC_000009.11	<i>H. sapiens</i> chrom. 9, GRCh37.p13	{296,315,584,574,752,108,66,280,265,554}	800
NC_000010.10	<i>H. sapiens</i> chrom. 10, GRCh37.p13	{398,271,708,478,388,428,762,429,215,568}	875
NC_000011.9	<i>H. sapiens</i> chrom. 11, GRCh37.p13	{201,561,535,559,699,593,641,801,168,470}	874
NC_000012.11	<i>H. sapiens</i> chrom. 12, GRCh37.p13	{623,586,163,627,415,522,245,685,244,471}	869
NC_000013.10	<i>H. sapiens</i> chrom. 13, GRCh37.p13	{580,576,457,310,13,349,373,549,116,132}	637
NC_000014.8	<i>H. sapiens</i> chrom. 14, GRCh37.p13	{579,434,219,212,383,286,508,322,572,237}	588
NC_000015.9	<i>H. sapiens</i> chrom. 15, GRCh37.p13	{522,244,446,504,440,467,181,271,306,326}	544
NC_000016.9	<i>H. sapiens</i> chrom. 16, GRCh37.p13	{225,101,276,232,316,513,183,40,3,81}	525
NC_000017.10	<i>H. sapiens</i> chrom. 17, GRCh37.p13	{264,155,384,507,214,152,484,273,141,412}	518
NC_000018.9	<i>H. sapiens</i> chrom. 18, GRCh37.p13	{434,48,374,447,433,281,81,117,268,134}	497
NC_000019.9	<i>H. sapiens</i> chrom. 19, GRCh37.p13	{302,1,9,203,79,126,49,158,156,329}	372
NC_000020.10	<i>H. sapiens</i> chrom. 20, GRCh37.p13	{112,319,75,345,351,118,54,8,157,1}	396
NC_000021.8	<i>H. sapiens</i> chrom. 21, GRCh37.p13	{165,68,110,57,131,50,72,8,188,157}	234
NC_000022.10	<i>H. sapiens</i> chrom. 22, GRCh37.p13	{57,62,189,99,211,168,196,154,172,45}	232
NC_000023.10	<i>H. sapiens</i> chrom. X, GRCh37.p13	{250,303,21,816,693,170,581,744,621,145}	1007
NC_000024.9	<i>H. sapiens</i> chrom. Y, GRCh37.p13	{89,167,136,19,16,93,163,29,147,65}	171
NC_000913.3	<i>Escherichia coli</i> str. K-12 substr. MG1655	1 through 30	30
NC_001133.9	<i>S. cerevisiae</i> S288c chrom. I	{1}	1
NC_001134.8	<i>S. cerevisiae</i> S288c chrom. II	{3,2,5,4,1}	5
NC_001135.5	<i>S. cerevisiae</i> S288c chrom. III	{2,1}	2
NC_001136.10	<i>S. cerevisiae</i> S288c chrom. IV	{10,4,2,9,3,5,6,7,8,1}	10
NC_001137.3	<i>S. cerevisiae</i> S288c chrom. V	{2,3,1}	3
NC_001138.5	<i>S. cerevisiae</i> S288c chrom. VI	{1}	1
NC_001139.9	<i>S. cerevisiae</i> S288c chrom. VII	{3,2,5,1,7,6,4}	7
NC_001140.6	<i>S. cerevisiae</i> S288c chrom. VIII	{2,1,3}	3
NC_001141.2	<i>S. cerevisiae</i> S288c chrom. IX	{2,1}	2
NC_001142.9	<i>S. cerevisiae</i> S288c chrom. X	{4,3,2,1}	4
NC_001143.9	<i>S. cerevisiae</i> S288c chrom. XI	{3,4,1,2}	4
NC_001144.5	<i>S. cerevisiae</i> S288c chrom. XII	{1,4,6,2,7,5,3}	7
NC_001145.3	<i>S. cerevisiae</i> S288c chrom. XIII	{3,4,1,6,2,5}	6
NC_001146.8	<i>S. cerevisiae</i> S288c chrom. XIV	{3,5,4,2,1}	5
NC_001147.6	<i>S. cerevisiae</i> S288c chrom. XV	{2,3,4,1,5,7,6}	7
NC_001148.4	<i>S. cerevisiae</i> S288c chrom. XVI	{2,6,5,4,1,3}	6
NC_003070.9	<i>A. thaliana</i> chrom. 1	{190,108,156,198,132,62,199,194,187,76}	201
NC_003071.7	<i>A. thaliana</i> chrom. 2	{63,124,58,11,101,34,95,56,51,62}	131
NC_003074.8	<i>A. thaliana</i> chrom. 3	{72,122,25,100,70,49,66,21,133,88}	156
NC_003075.7	<i>A. thaliana</i> chrom. 4	{34,122,8,108,35,15,87,20,11,1}	123
NC_003076.8	<i>A. thaliana</i> chrom. 5	{125,1,50,113,127,45,74,106,18,76}	179
NC_004325.1	<i>P. falciparum</i> strain 3D7, chrom. 1	{4,1,3,2}	4
NC_000910.2	<i>P. falciparum</i> 3D7 chrom. 2	{2,6,5,4,1,3}	6
NC_000521.3	<i>P. falciparum</i> 3D7 chrom. 3	{1,7,4,2,6,3,5}	7
NC_004318.1	<i>P. falciparum</i> 3D7 chrom. 4	{1,4,3,7,2,8,5,6}	8
NC_004326.1	<i>P. falciparum</i> 3D7 chrom. 5	{7,6,1,3,8,2,4,5}	8
NC_004327.2	<i>P. falciparum</i> 3D7 chrom. 6	{5,1,8,6,2,9,7,4,3}	9
NC_004328.2	<i>P. falciparum</i> 3D7 chrom. 7	{3,9,4,1,6,5,2,10,8,7}	10
NC_004329.2	<i>P. falciparum</i> 3D7 chrom. 8	{9,8,1,2,6,3,5,4,7}	9
NC_004330.1	<i>P. falciparum</i> 3D7 chrom. 9	{2,8,7,1,10,3,6,5,9,4}	10
NC_004314.2	<i>P. falciparum</i> 3D7 chrom. 10	{9,3,4,7,5,8,10,11,1,2}	11
NC_004315.2	<i>P. falciparum</i> 3D7 chrom. 11	{4,9,13,3,12,10,5,1,11,2}	13
NC_004316.3	<i>P. falciparum</i> 3D7 chrom. 12	{7,11,14,8,12,9,6,5,13,1}	15
NC_004331.2	<i>P. falciparum</i> 3D7 chrom. 13	{2,17,12,16,4,14,19,15,13,7}	19
NC_004317.2	<i>P. falciparum</i> 3D7 chrom. 14	{12,18,1,3,16,17,2,5,10,4}	21
NC_018092.1	<i>P. furiosus</i> COM1 chrom.	1 through 12	12

TABLE 3: The dataset for the preview experiment

NCBI Accession	Name
NC_000001.11	<i>H. sapiens</i> chrom. 1, GRCh38 Primary Assembly
NC_000002.12	<i>H. sapiens</i> chrom. 2, GRCh38 Primary Assembly
NC_000003.12	<i>H. sapiens</i> chrom. 3, GRCh38 Primary Assembly
NC_000004.12	<i>H. sapiens</i> chrom. 4, GRCh38 Primary Assembly
NC_000005.10	<i>H. sapiens</i> chrom. 5, GRCh38 Primary Assembly
NC_000006.12	<i>H. sapiens</i> chrom. 6, GRCh38 Primary Assembly
NC_000007.14	<i>H. sapiens</i> chrom. 7, GRCh38 Primary Assembly
NC_000008.11	<i>H. sapiens</i> chrom. 8, GRCh38 Primary Assembly
NC_000009.12	<i>H. sapiens</i> chrom. 9, GRCh38 Primary Assembly
NC_000010.11	<i>H. sapiens</i> chrom. 10, GRCh38 Primary Assembly
NC_000011.10	<i>H. sapiens</i> chrom. 11, GRCh38 Primary Assembly
NC_000012.12	<i>H. sapiens</i> chrom. 12, GRCh38 Primary Assembly
NC_000013.11	<i>H. sapiens</i> chrom. 13, GRCh38 Primary Assembly
NC_000014.9	<i>H. sapiens</i> chrom. 14, GRCh38 Primary Assembly
NC_000015.10	<i>H. sapiens</i> chrom. 15, GRCh38 Primary Assembly
NC_000016.10	<i>H. sapiens</i> chrom. 16, GRCh38 Primary Assembly
NC_000017.11	<i>H. sapiens</i> chrom. 17, GRCh38 Primary Assembly
NC_000018.10	<i>H. sapiens</i> chrom. 18, GRCh38 Primary Assembly
NC_000019.10	<i>H. sapiens</i> chrom. 19, GRCh38 Primary Assembly
NC_000020.11	<i>H. sapiens</i> chrom. 20, GRCh38 Primary Assembly
NC_000021.9	<i>H. sapiens</i> chrom. 21, GRCh38 Primary Assembly
NC_000022.11	<i>H. sapiens</i> chrom. 22, GRCh38 Primary Assembly
NC_000023.11	<i>H. sapiens</i> chrom. X, GRCh38 Primary Assembly
NC_000024.10	<i>H. sapiens</i> chrom. Y, GRCh38 Primary Assembly
NC_000067.6	<i>Mus musculus</i> strain C57BL/6J chrom. 1, GRCm38.p3 C57BL/6J
NC_000068.7	<i>Mus musculus</i> strain C57BL/6J chrom. 2, GRCm38.p3 C57BL/6J
NC_000069.6	<i>Mus musculus</i> strain C57BL/6J chrom. 3, GRCm38.p3 C57BL/6J
NC_000070.6	<i>Mus musculus</i> strain C57BL/6J chrom. 4, GRCm38.p3 C57BL/6J
NC_000071.6	<i>Mus musculus</i> strain C57BL/6J chrom. 5, GRCm38.p3 C57BL/6J
NC_000072.6	<i>Mus musculus</i> strain C57BL/6J chrom. 6, GRCm38.p3 C57BL/6J
NC_000073.6	<i>Mus musculus</i> strain C57BL/6J chrom. 7, GRCm38.p3 C57BL/6J
NC_000074.6	<i>Mus musculus</i> strain C57BL/6J chrom. 8, GRCm38.p3 C57BL/6J
NC_000075.6	<i>Mus musculus</i> strain C57BL/6J chrom. 9, GRCm38.p3 C57BL/6J
NC_000076.6	<i>Mus musculus</i> strain C57BL/6J chrom. 10, GRCm38.p3 C57BL/6J
NC_000077.6	<i>Mus musculus</i> strain C57BL/6J chrom. 11, GRCm38.p3 C57BL/6J
NC_000078.6	<i>Mus musculus</i> strain C57BL/6J chrom. 12, GRCm38.p3 C57BL/6J
NC_000079.6	<i>Mus musculus</i> strain C57BL/6J chrom. 13, GRCm38.p3 C57BL/6J
NC_000080.6	<i>Mus musculus</i> strain C57BL/6J chrom. 14, GRCm38.p3 C57BL/6J
NC_000081.6	<i>Mus musculus</i> strain C57BL/6J chrom. 15, GRCm38.p3 C57BL/6J
NC_000082.6	<i>Mus musculus</i> strain C57BL/6J chrom. 16, GRCm38.p3 C57BL/6J
NC_000083.6	<i>Mus musculus</i> strain C57BL/6J chrom. 17, GRCm38.p3 C57BL/6J
NC_000084.6	<i>Mus musculus</i> strain C57BL/6J chrom. 18, GRCm38.p3 C57BL/6J
NC_000085.6	<i>Mus musculus</i> strain C57BL/6J chrom. 19, GRCm38.p3 C57BL/6J
NC_000086.7	<i>Mus musculus</i> strain C57BL/6J chrom. X, GRCm38.p3 C57BL/6J
NC_000087.7	<i>Mus musculus</i> strain C57BL/6J chrom. Y, GRCm38.p3 C57BL/6J

C Additional histograms for the first experiment (one complete chromosome per organism)

In this section we display all histograms depicting the distance distributions for all pairs of organisms from the first experiment dataset (similar to the pair *H. sapiens* and *A. thaliana* in Figure 4 in the main paper). The histograms are shown in Figures S1 through S15. These histograms are computed in correspondence with the computation of the *relative overlap between the intragenomic and intergenomic distance histograms*. For a distance measure d_α , the range from 0 to the maximum distance $\max(d_\alpha)$ is divided into 100 bins (intervals) of equal size (this is done independently of the two species which are represented in the histogram). Then, we count the number of pairwise distances which fall into each bin and represent the tally of each bin as a bar. When plotting the histograms, only the smallest interval that contains all non-empty bins is shown. In some cases, for example Figure S6(c), only few bins are shown because all distances lie within a small range relative to the overall range $[0, \max(d_\alpha)]$ of the distance.

Note that the number of used fragments for the species *H. sapiens* and *A. thaliana* is higher than that for the other four species, see Table 2 in the main paper. Therefore, *H. sapiens* and *A. thaliana* will also contribute a higher number of distances to the histograms. This means, for example, that the number of pairwise distances between any two fragments from *H. sapiens* is significantly larger than the number pairwise distances between any two fragments from *E. coli*. In Figure S1, this results in a small area of the bars representing *E. coli* as compared to the area of the bars representing *H. sapiens*.

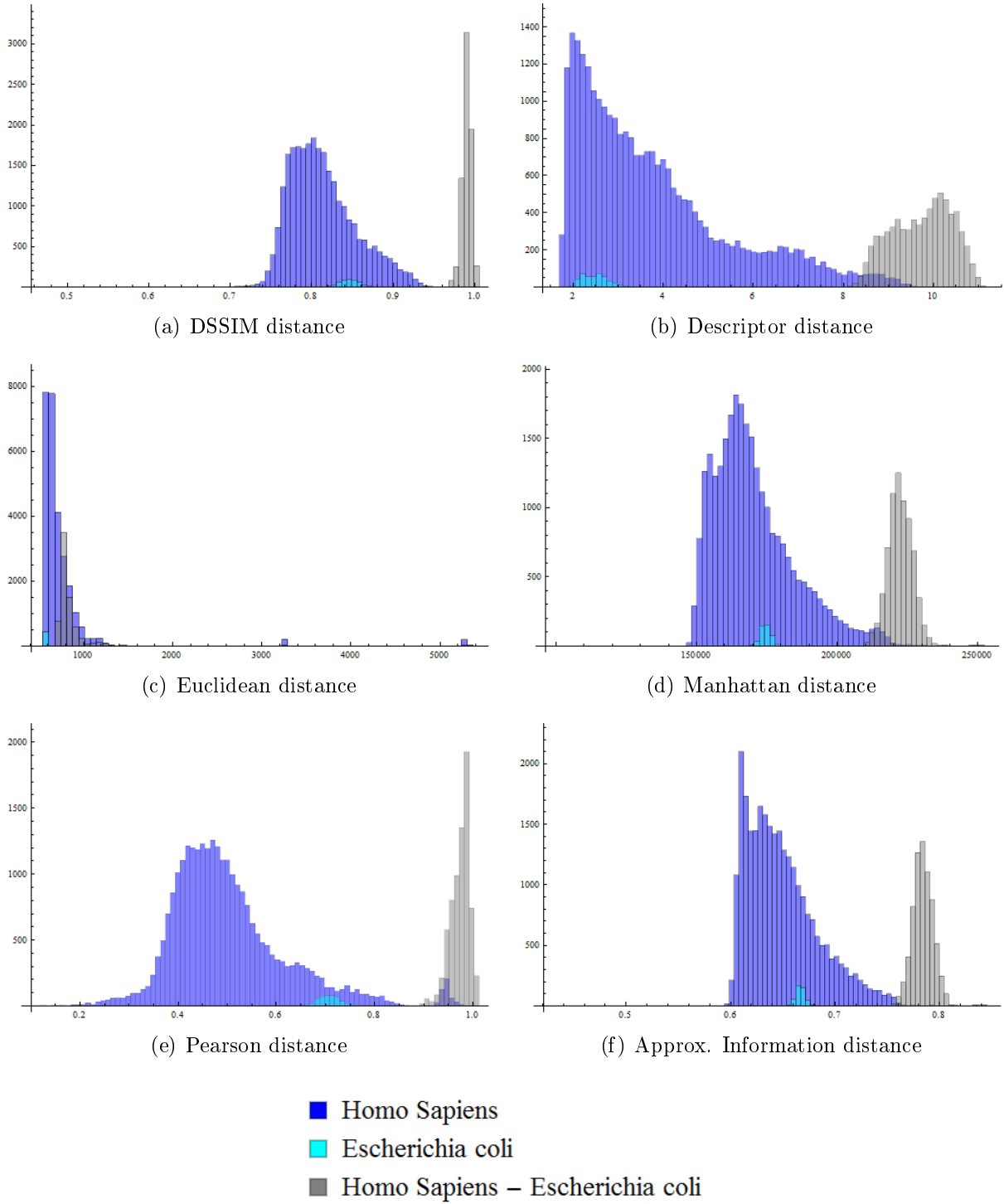


Figure S1: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *H. sapiens* and *E. coli*.

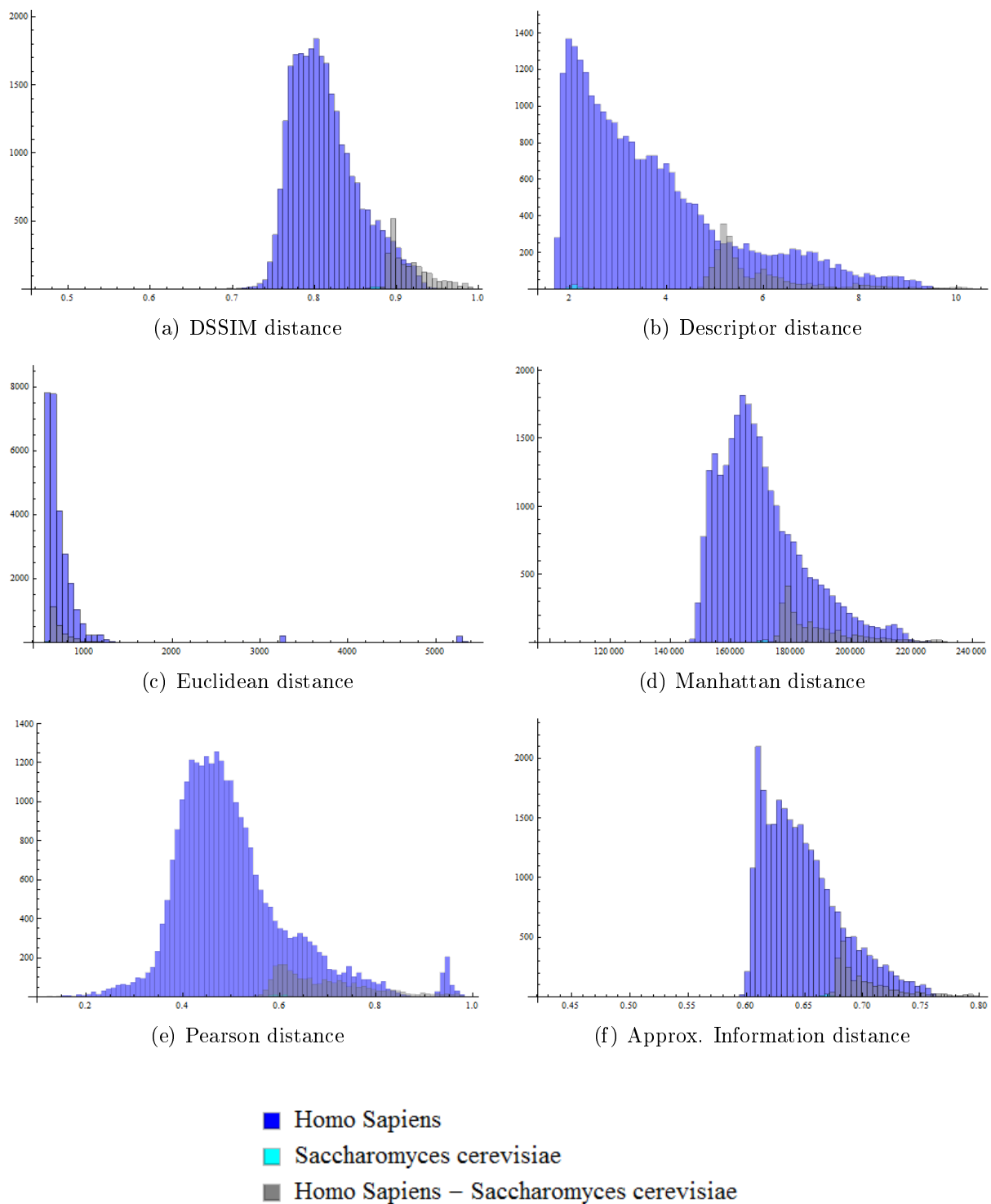


Figure S2: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *H. sapiens* and *S. cerevisiae*.

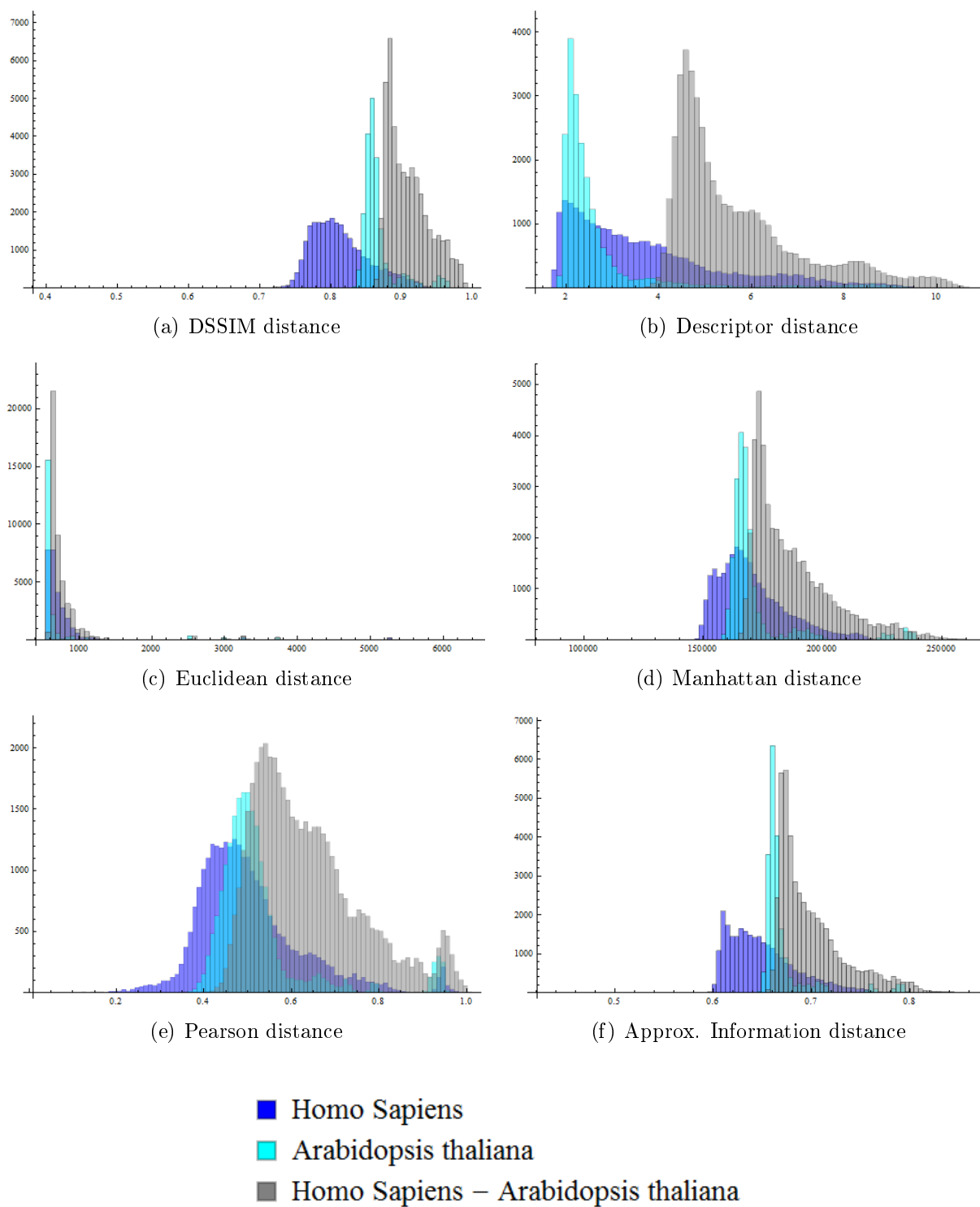


Figure S3: First experiment (one complete chromosome per organism) : Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *H. sapiens* and *A. thaliana*.

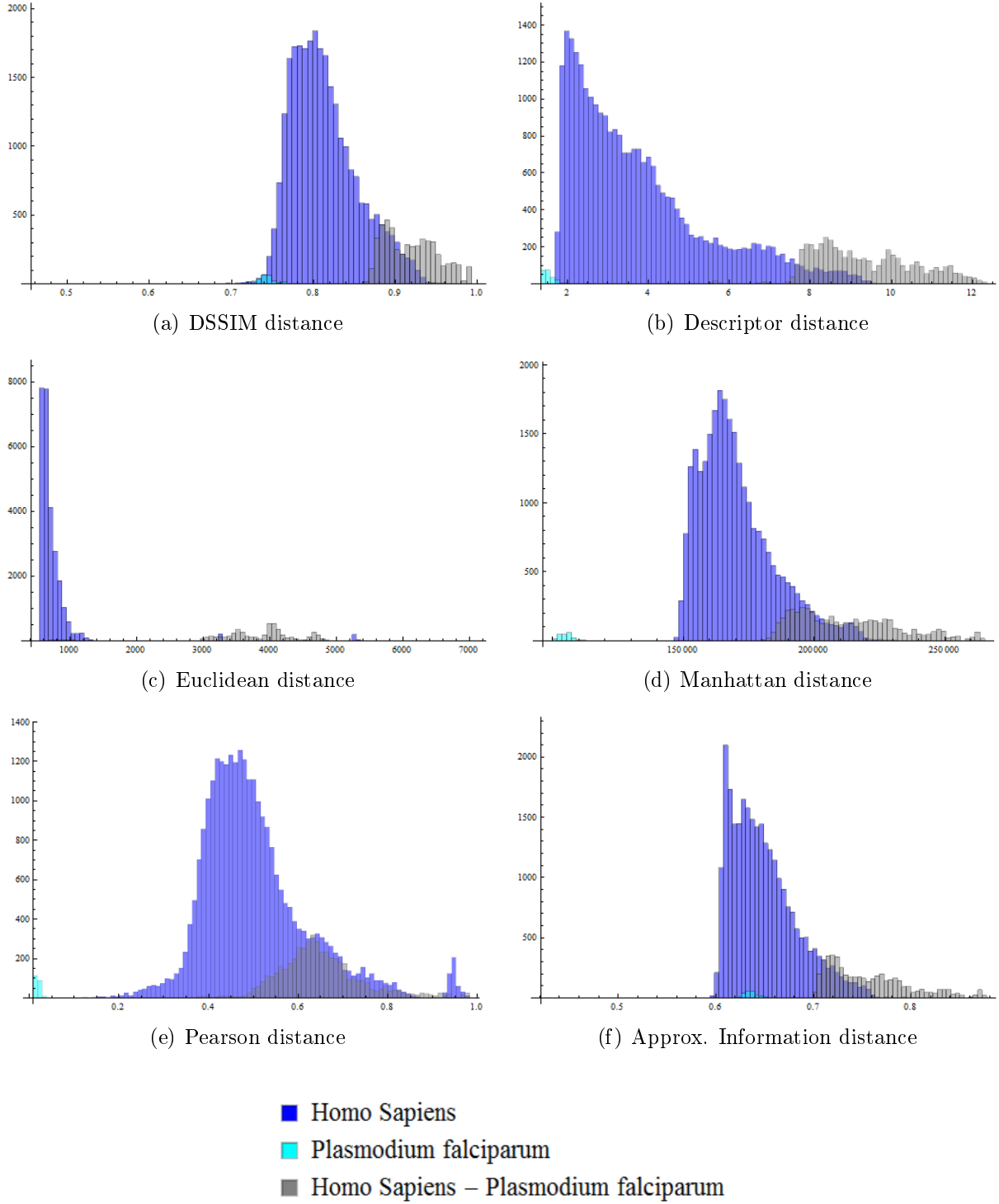


Figure S4: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *H. sapiens* and *P. falciparum*.

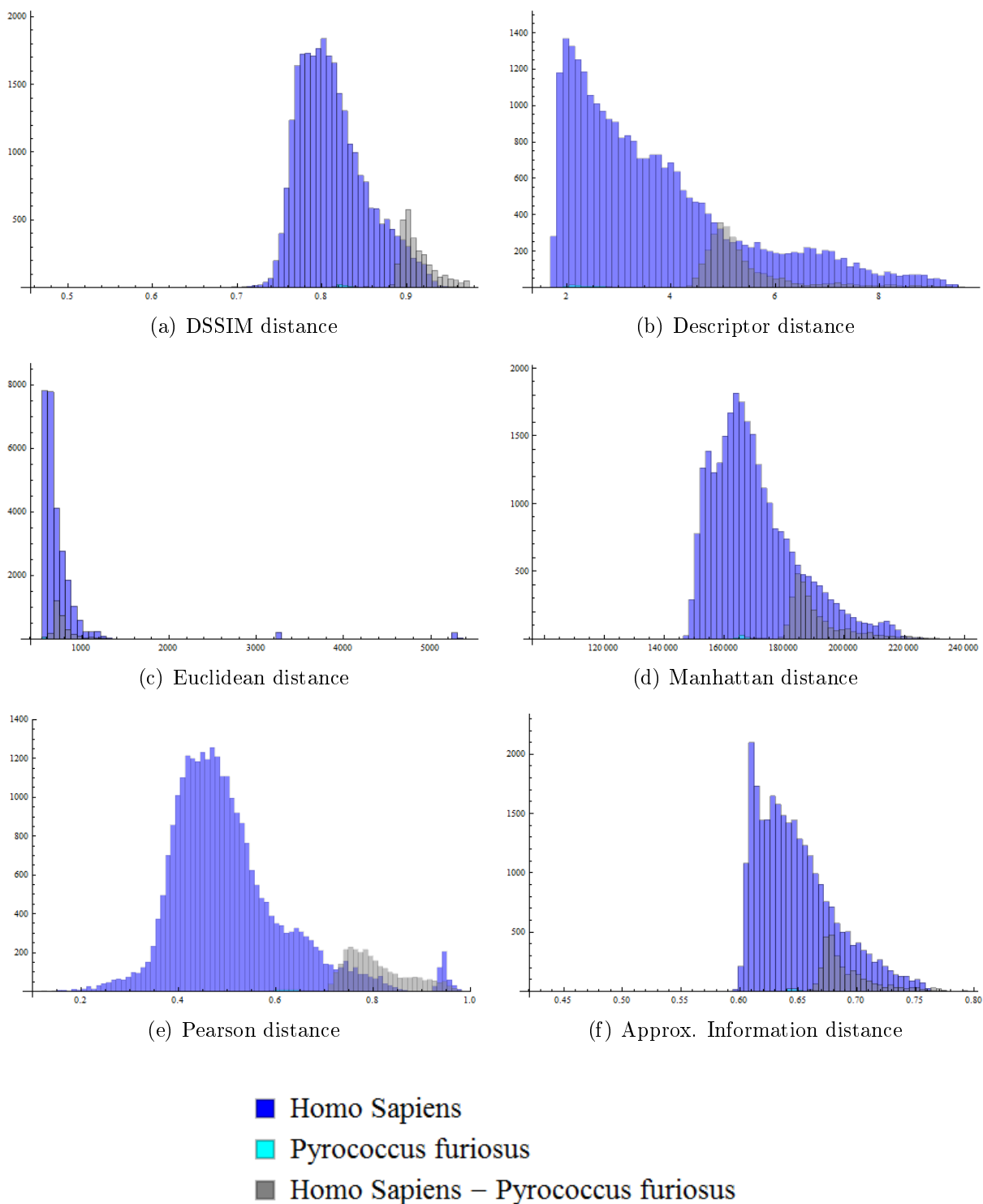


Figure S5: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *H. sapiens* and *P. furiosus*.

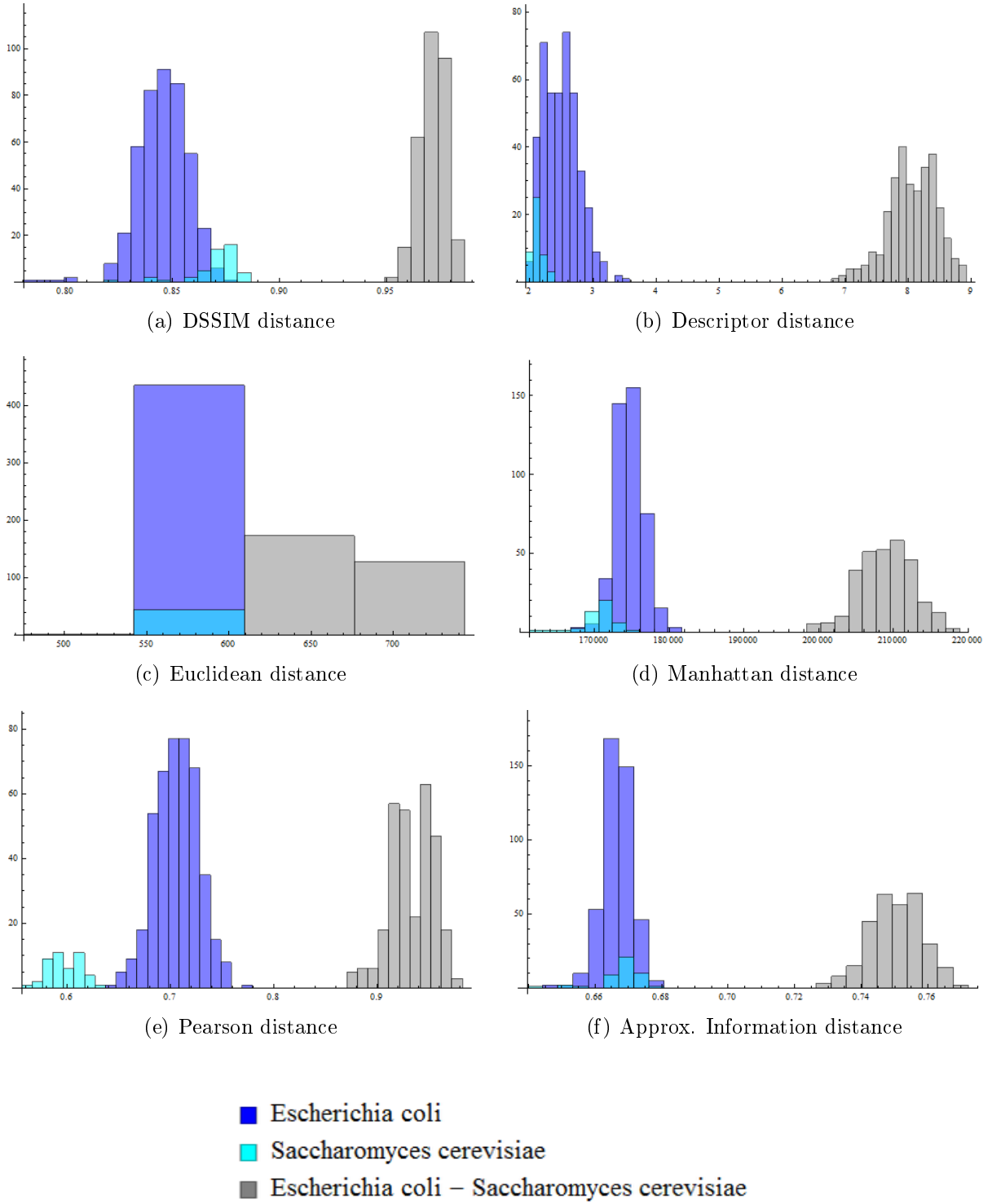


Figure S6: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *E. coli* and *S. cerevisiae*.

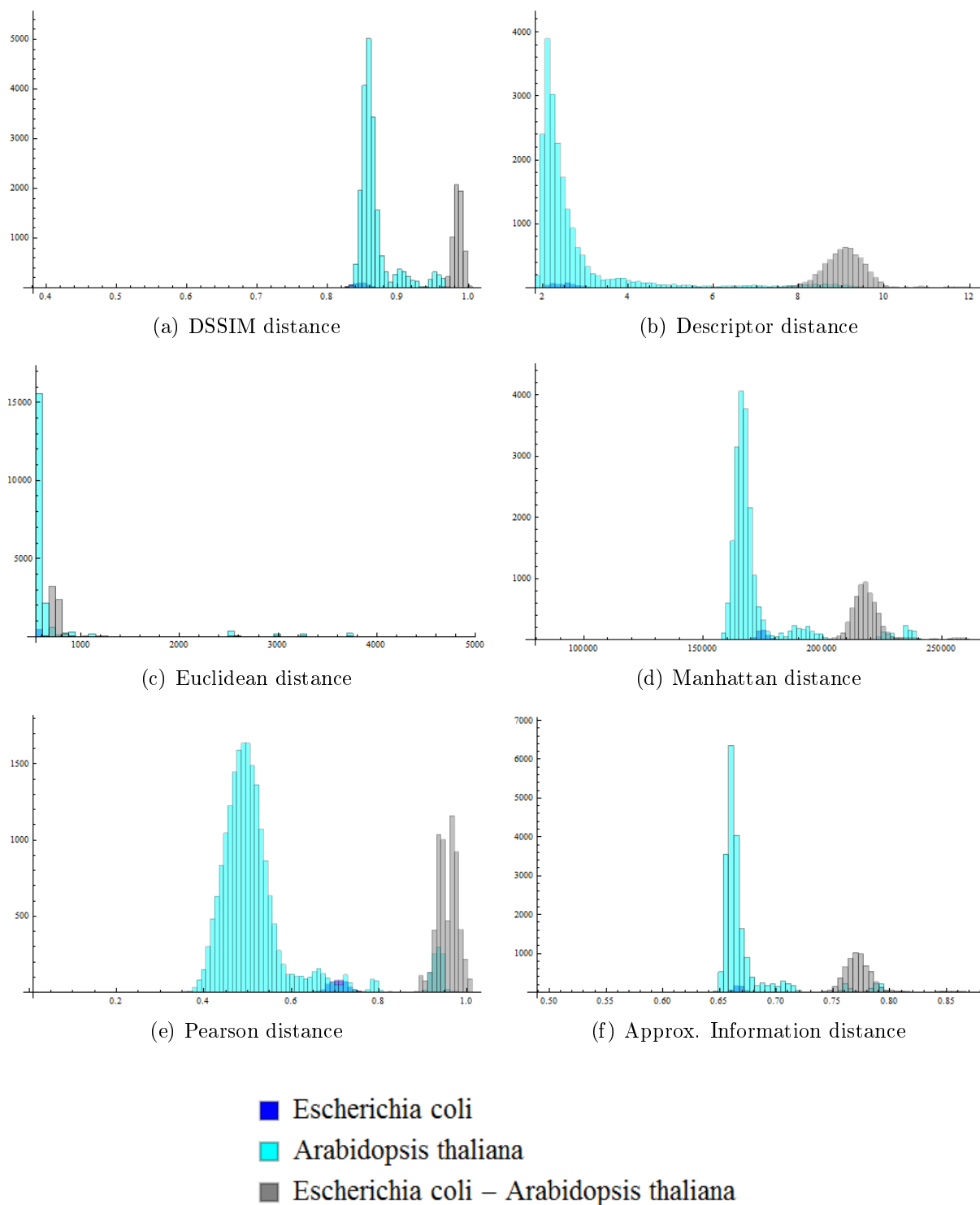


Figure S7: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *E. coli* and *A. thaliana*.

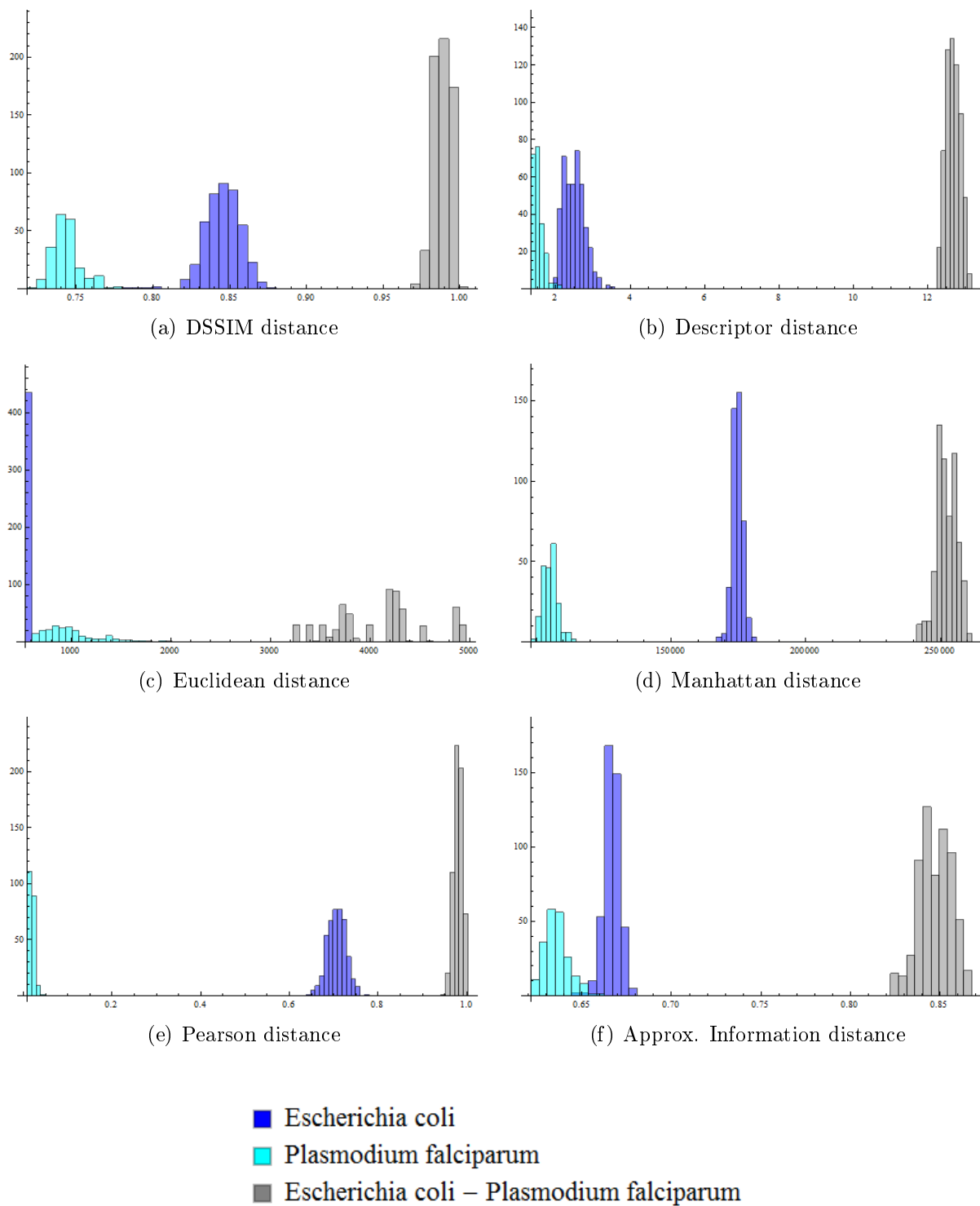


Figure S8: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *E. coli* and *P. falciparum*.

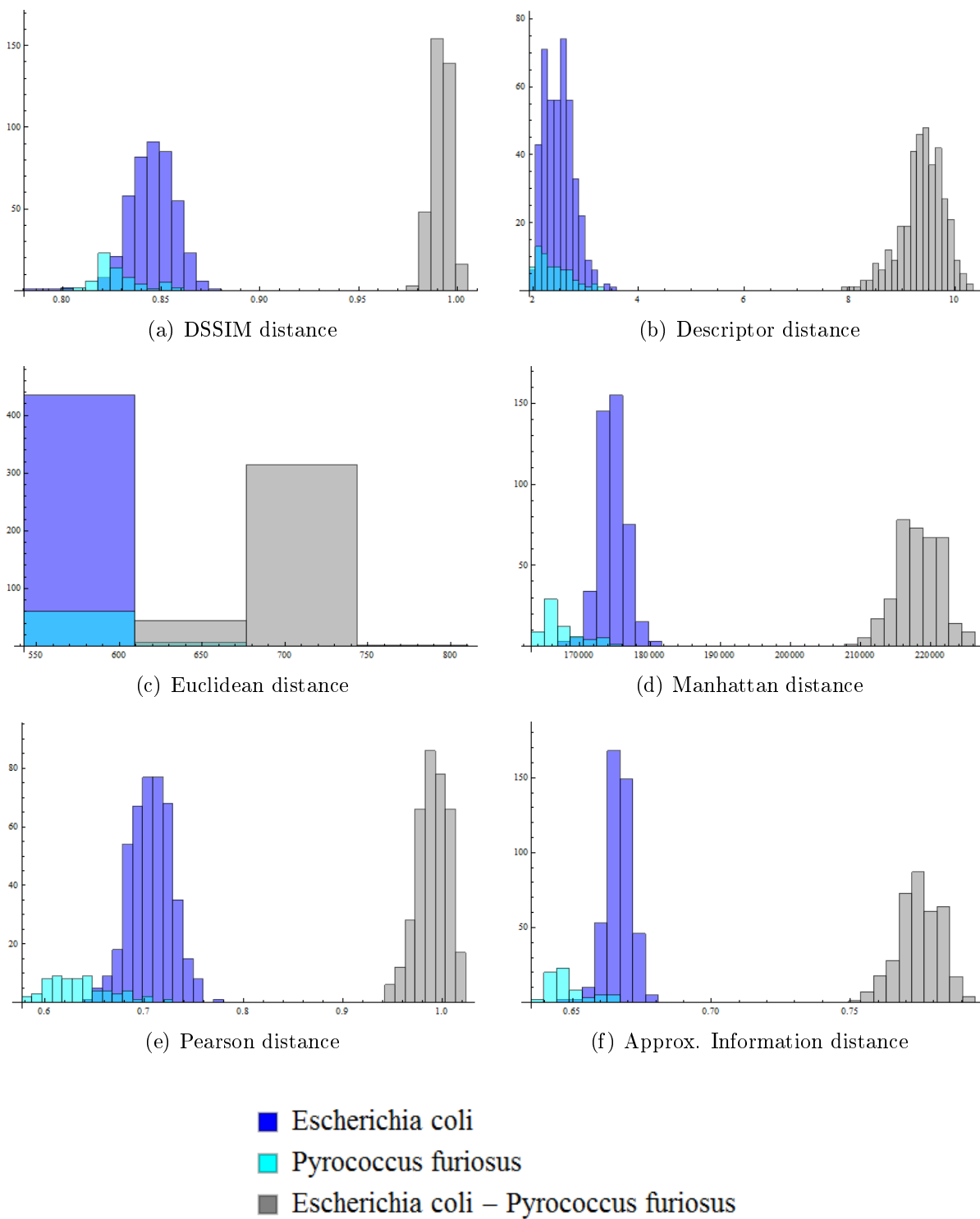


Figure S9: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *E. coli* and *P. furiosus*.

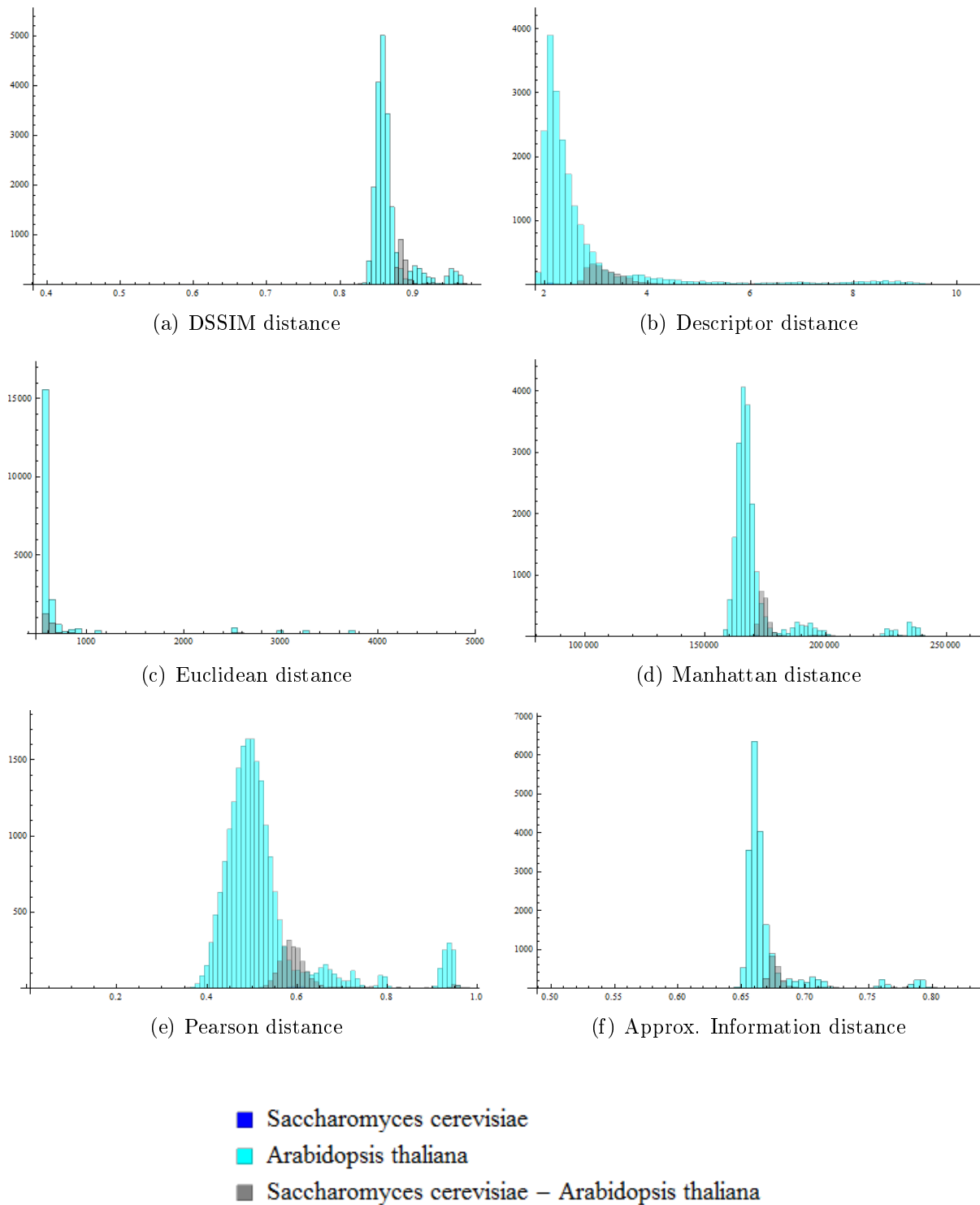


Figure S10: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *S. cerevisiae* and *A. thaliana*.

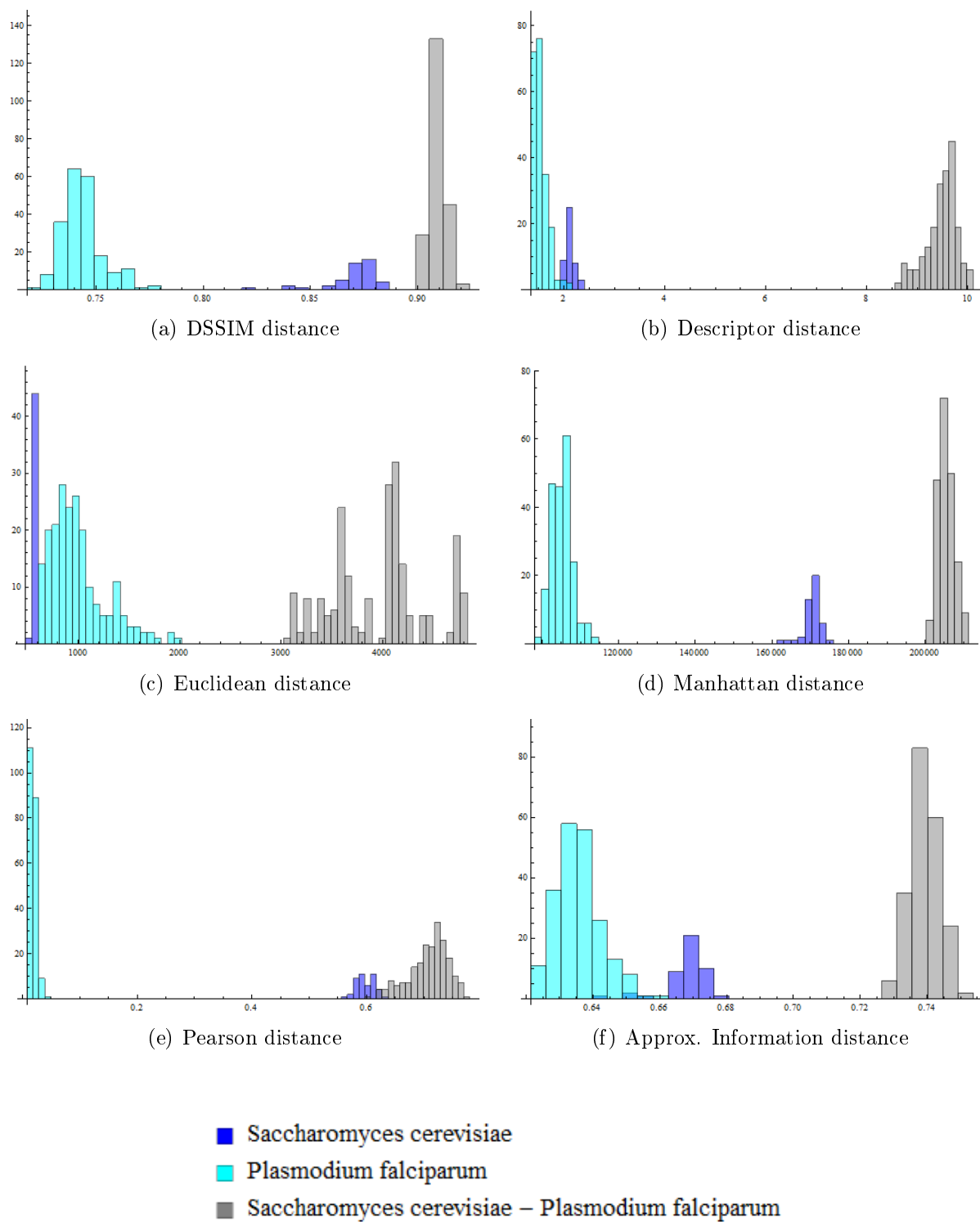


Figure S11: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *S. cerevisiae* and *P. falciparum*.

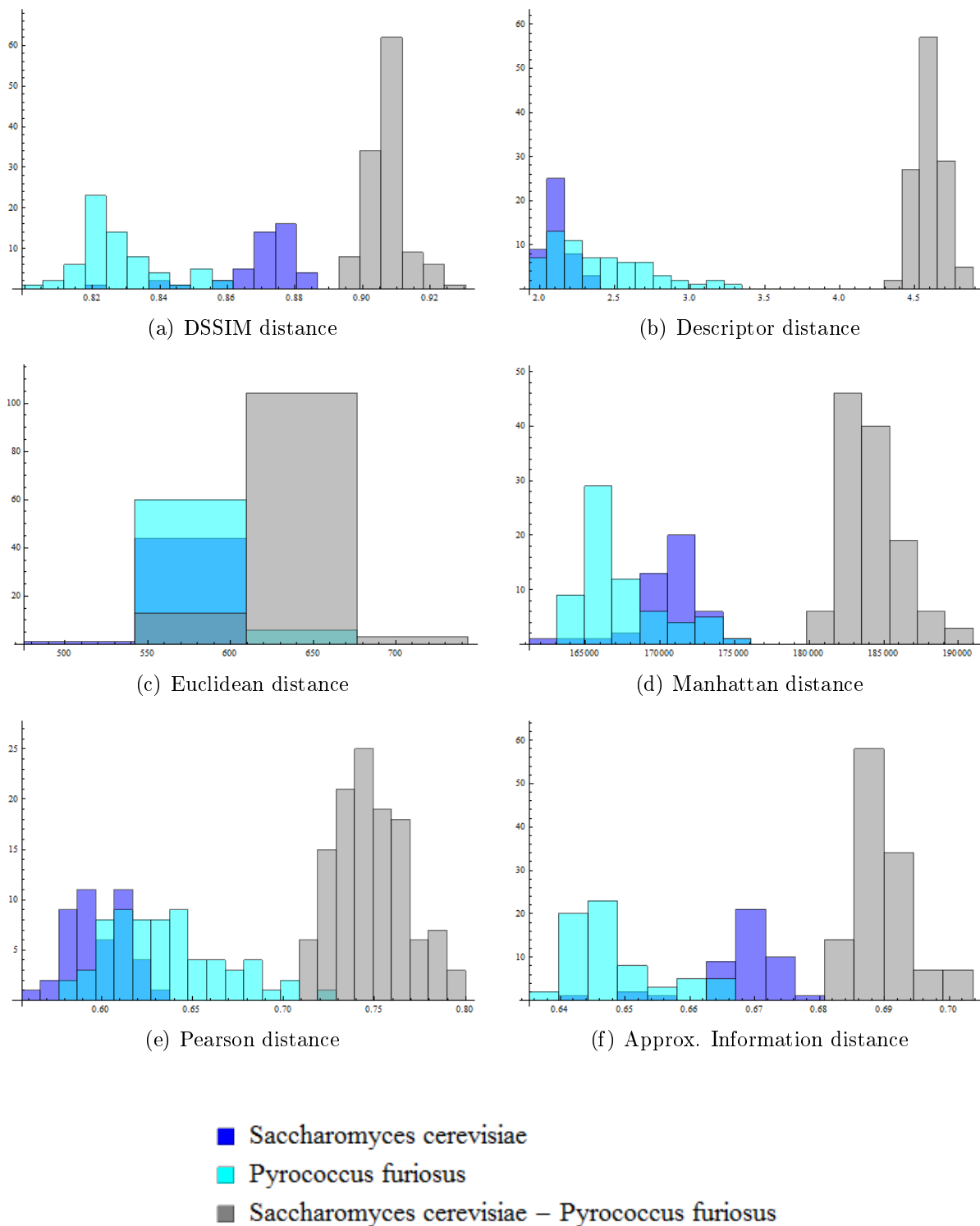


Figure S12: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *S. cerevisiae* and *P. furiosus*.

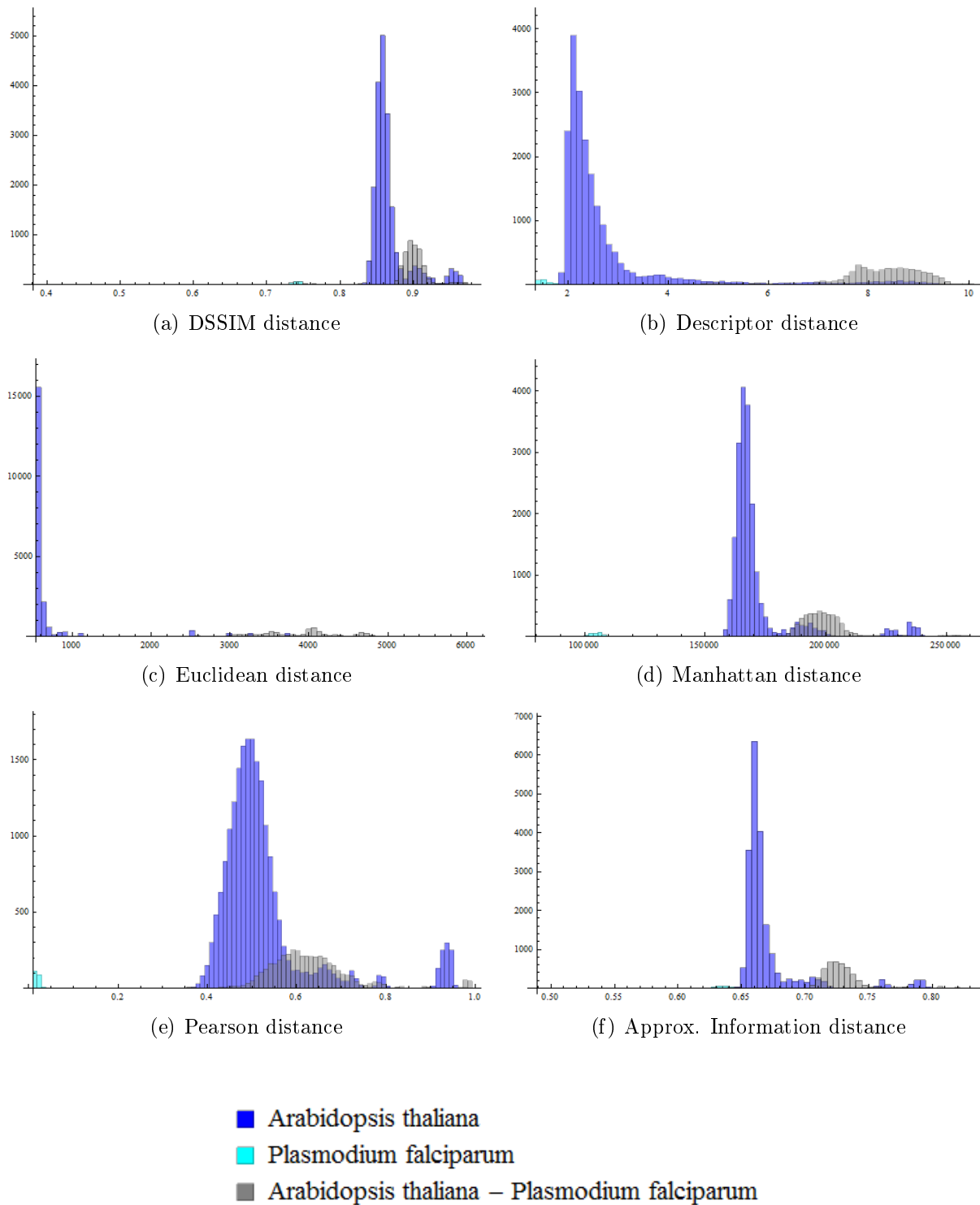


Figure S13: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *A. thaliana* and *P. falciparum*.

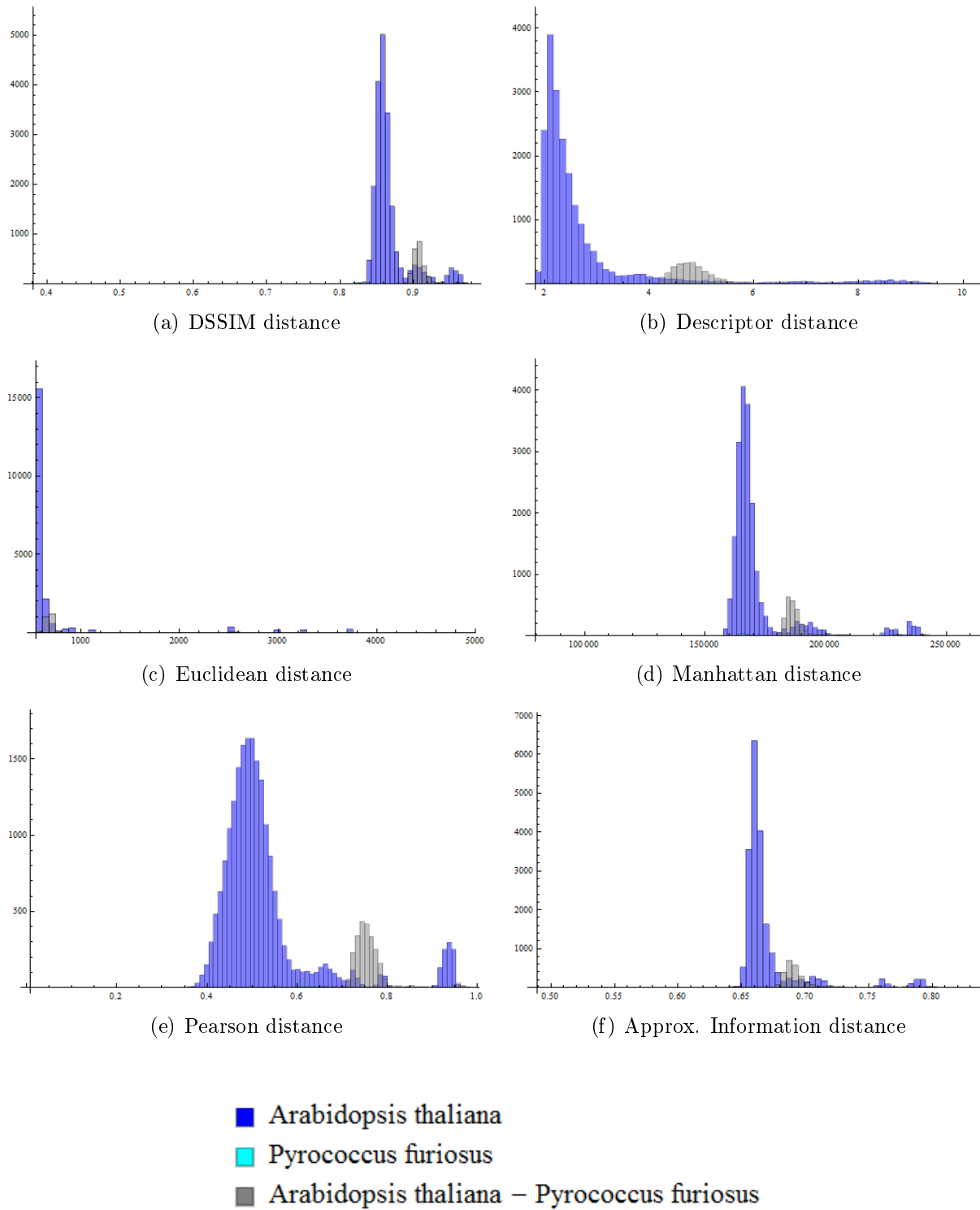


Figure S14: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *A. thaliana* and *P. furiosus*.

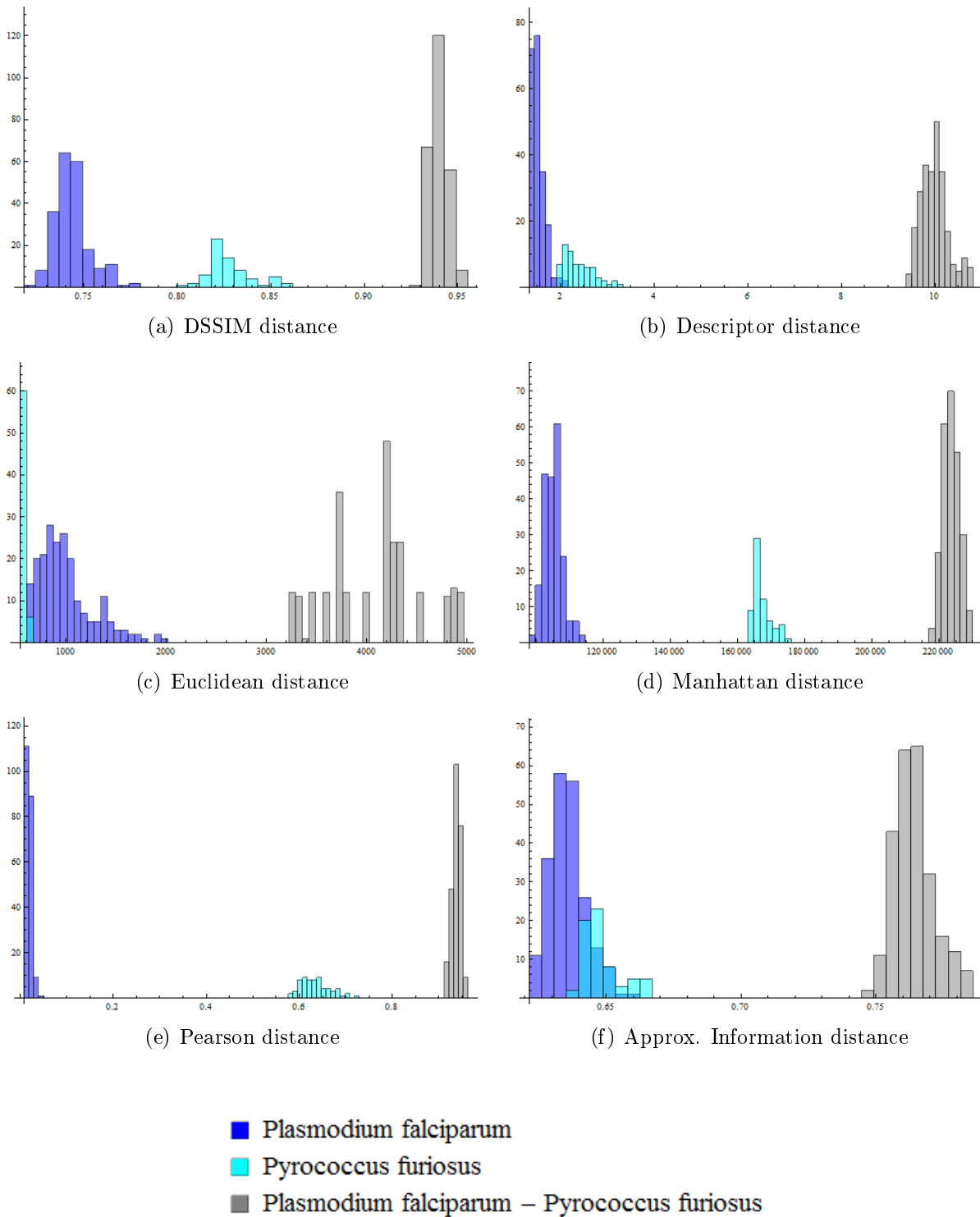


Figure S15: First experiment (one complete chromosome per organism): Histograms of pairwise intragenomic and intergenomic distances among 150,000 bp sequences from *P. falciparum* and *P. furiosus*.