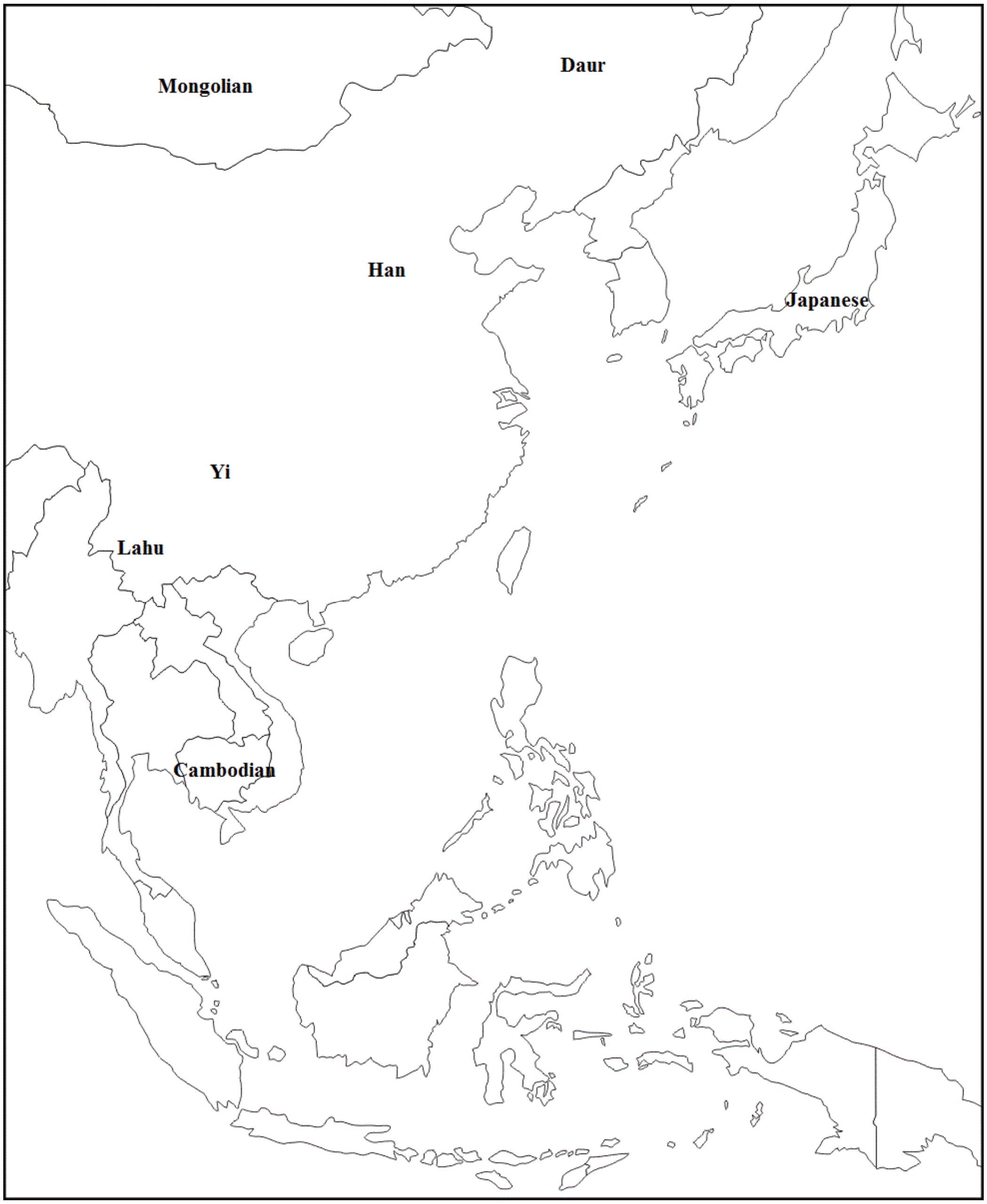
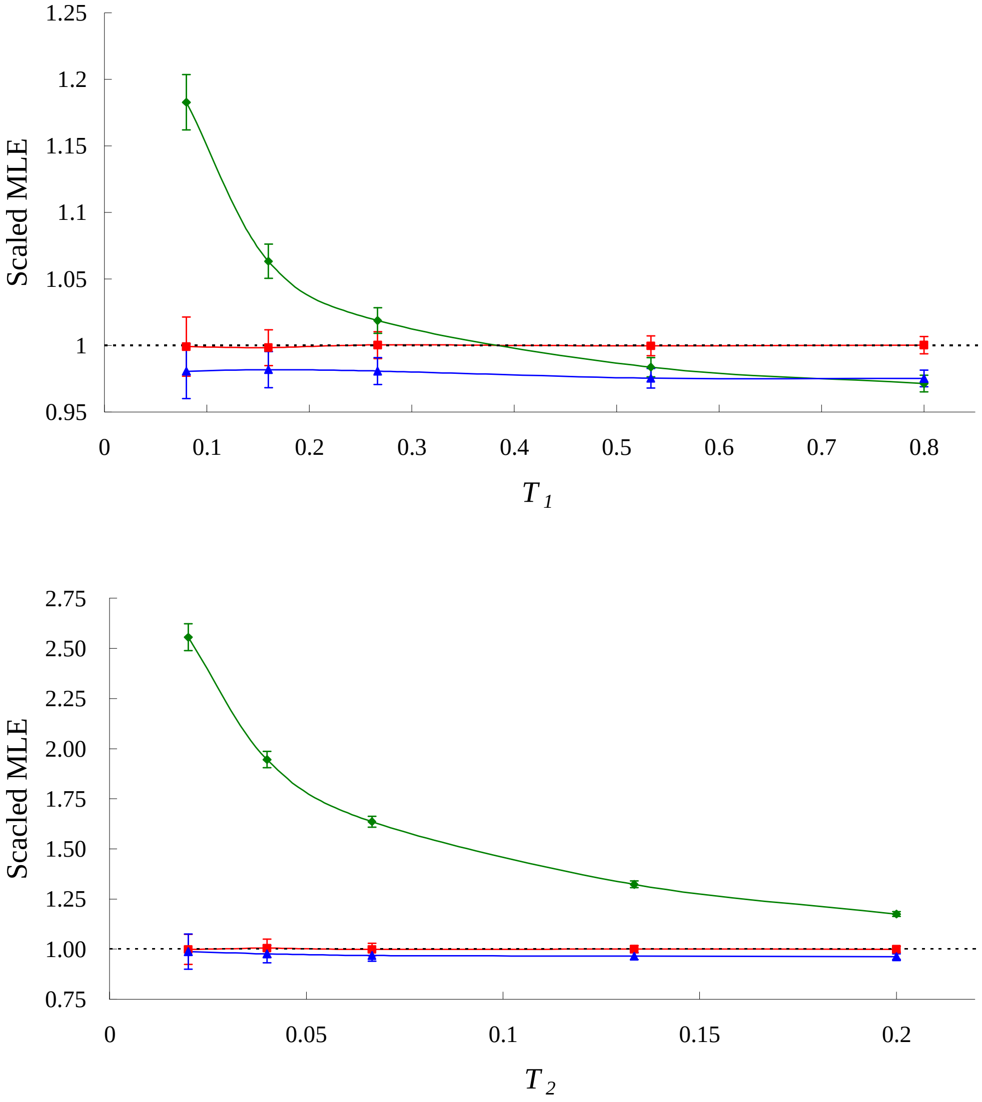
Supplemental Figures

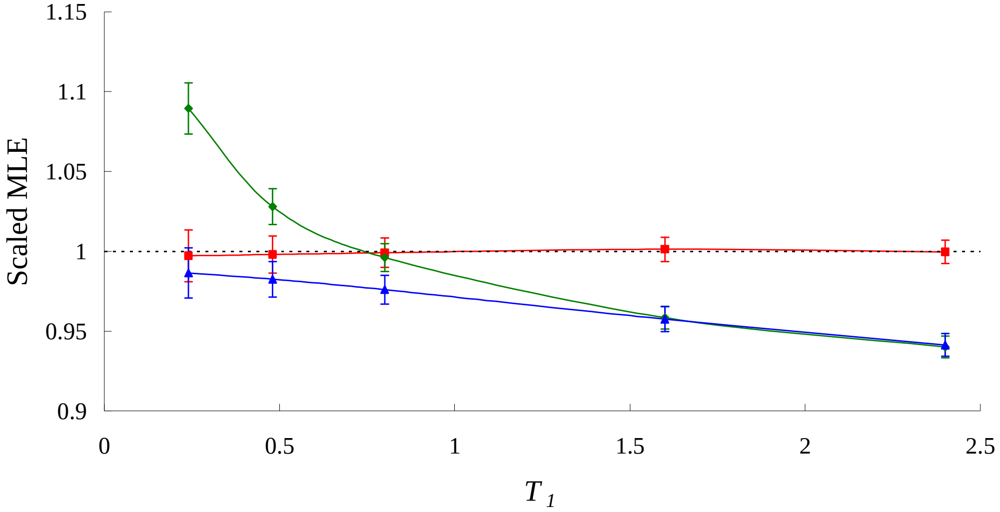


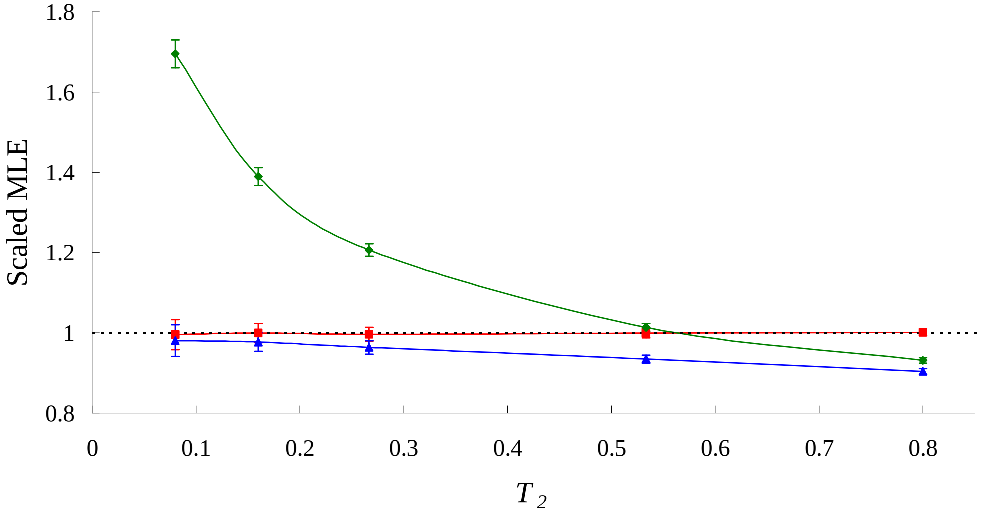
Sup. Fig. 1 Geographic locations of the seven East Asian Populations

A


Scaled MLE
1.25
1.2
1.15
1.1
1.05
1
0.95
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
T_1
Scaled MLE
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
0.75
0
0.05
0.1
0.15
0.2
T_2

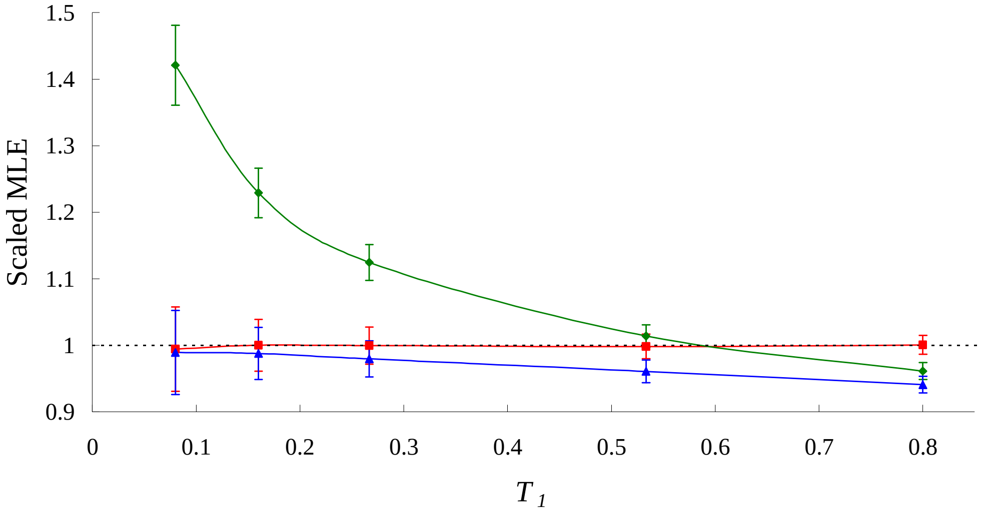
B



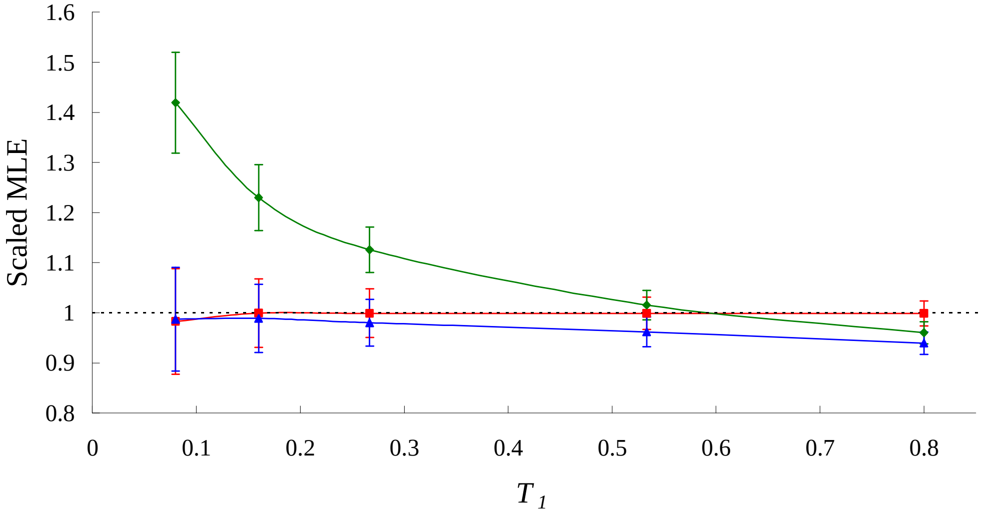


Sup. Fig. 2 Scaled maximum likelihood estimate of the divergence time plotted against the true value for different effective population sizes

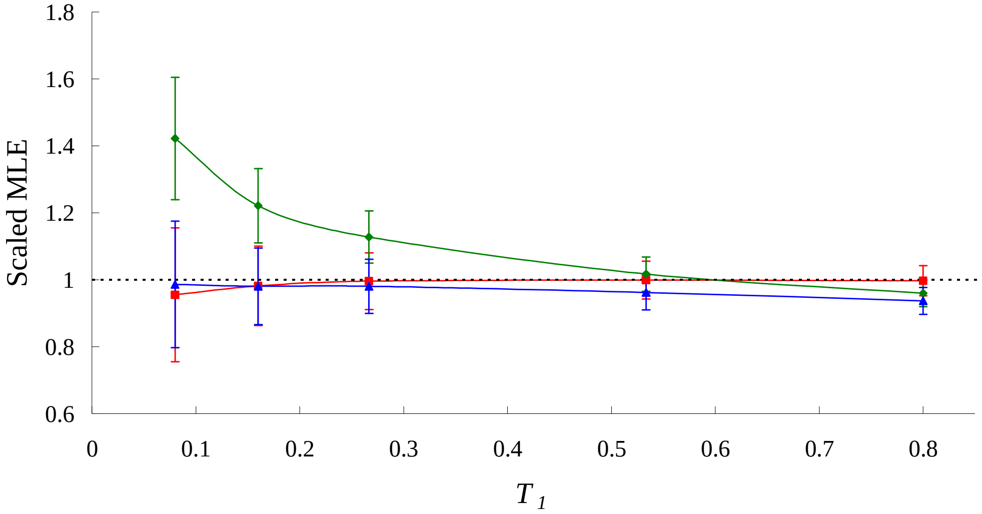
A


1.5
1.4
1.3
1.2
1.1
1
0.9
Scaled MLE
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
T_1

B


1.6
1.5
1.4
1.3
1.2
1.1
1
0.9
0.8
Scaled MLE
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
T_1

C



D

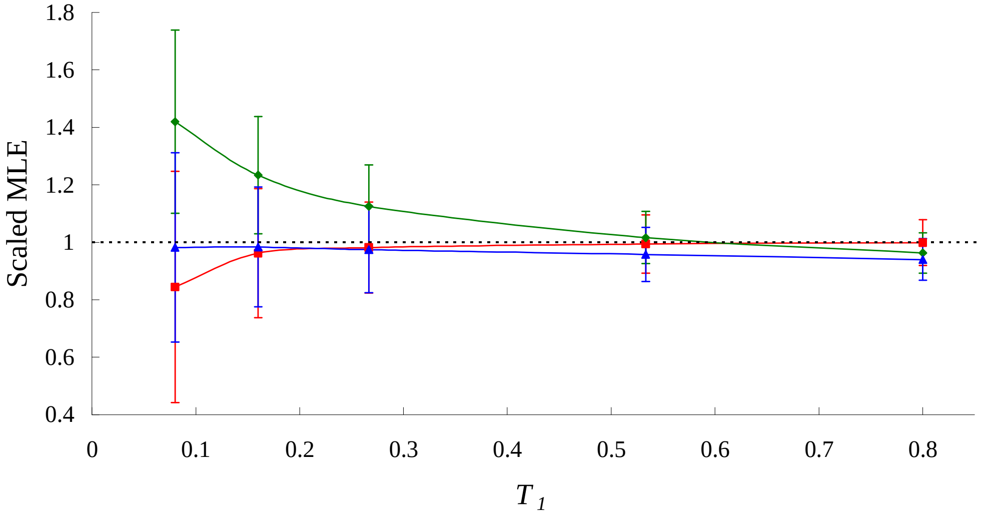
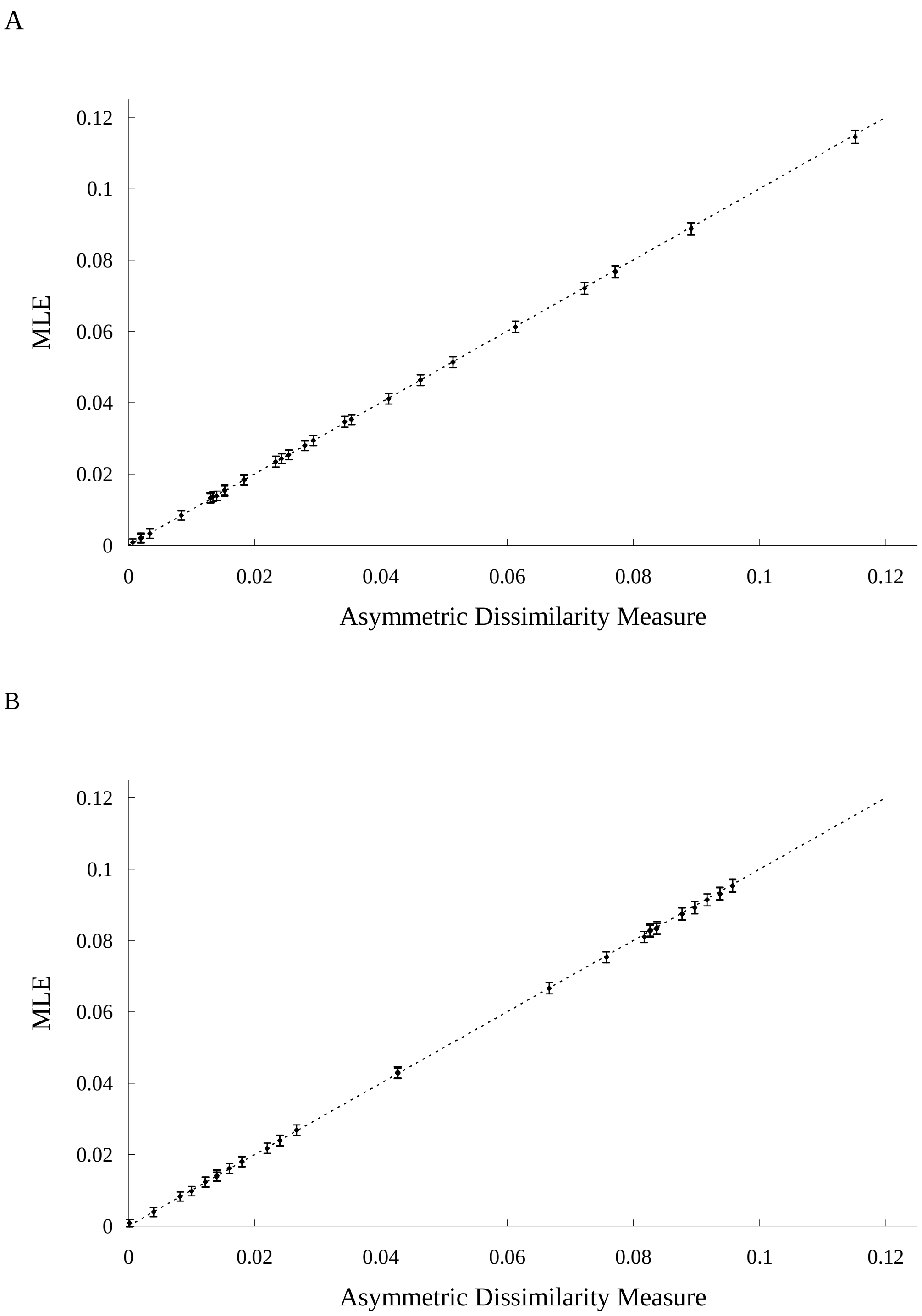


Fig. 3 Scaled maximum likelihood estimate of divergence time plotted against the true value for different number of SNPs



Sup. Fig. 4 Maximum likelihood estimate of asymmetric dissimilarity measure plotted against the expected value

Supplemental Figure Legends

Sup. Fig. 1 Geographic locations of the seven East Asian Populations

Sup. Fig. 2 Data were simulated from population structure shown in Fig. 1B with population size A) N_1=937.5 and N_2=3750 and B) N_1=937.5 and N_2=312.5, and different divergence times ranging from 3,000 years to 30,000 years. For each combination of population sizes and divergence time, 1,000 data sets were simulated. Average value of 1,000 maximum likelihood estimates was scaled by and plotted against the true value of divergence time, with bars representing scaled standard deviations. Red line represents the result from fully-parameterized method. Green line represents the result from IS-based method. Blue line represents the result from Beta-based method.

Sup. Fig. 3 Data were simulated from population structure shown in Fig. 1B with population size N_1=N_2=937.5, and different divergence times ranging from 3,000 years to 30,000 years. For each combination of population sizes and divergence time, 1,000 data sets were simulated, each including A) 30,000, B) 10,000, C) 3,000 and D) 1,000 SNPs, respectively. Average value of 1,000 maximum likelihood estimates was scaled by and plotted against the true value of divergence time, with bars representing scaled standard deviations. Red line represents the result from fully-parameterized method. Green line represents the result from IS-based method. blue line represents the result from Beta-based method.

Sup. Fig. 4 Data were simulated using population tree shown in A) Fig. 4A. and B) Fig. 4B. For each population tree, 1,000 data sets were simulated. Asymmetric dissimilarity measures were estimated from each data set. The average of 1,000 maximum likelihood estimates was plotted against the expected value of the corresponding asymmetric dissimilarity measure, with bars representing standard deviations.