ARGON: fast, whole-genome simulation of the discrete time Wright-Fisher process

Pier Francesco Palamara

Abstract

Motivation: Simulation under the coalescent model is ubiquitous in the analysis of genetic data. The rapid growth of real data sets from multiple human populations led to increasing interest in simulating very large sample sizes at whole-chromosome scales. When the sample size is large, the coalescent model becomes an increasingly inaccurate approximation of the discrete time Wright-Fisher model (DTWF). Analytical and computational treatment of the DTWF, however, is generally harder.
Results: We present a simulator (ARGON) for the DTWF process that scales up to hundreds of thousands of samples and whole-chromosome lengths, with a time/memory performance comparable or superior to currently available methods for coalescent simulation. The simulator supports arbitrary demographic history, migration, variable mutation/recombination rates and gene conversion, and efficiently outputs pairwise identical-by-descent (IBD) sharing data.
Availability: ARGON (version 0.1) is written in Java, open source, and freely available at https://github.com/pierpal/ARGON.
Contact: ppalama@hsph.harvard.edu
Supplementary information: Supplementary data are available online.

1 Introduction

The coalescent process ([Kingman, 1982a, Kingman, 1982b, Kingman, 1982c]) can be constructed as an approximation of the discrete time Wright-Fisher process (DTWF, [Fisher et al., 1922, Wright, 1931]). It can be used to model recombination, population structure and migration ([Hudson et al., 1990]), and generally leads to a more tractable analytical and computational analysis compared to the DTWF. Simulators based on the coalescent process, such as the widely used MS program ([Hudson, 2002]), have been extensively adopted in computational methods and analyses of real data sets (e.g. [Beaumont et al., 2002, Li and Jakobsson, 2012, Excoffier et al., 2013]). The coalescent approximation, however, relies on the assumption that the sample size is small compared to the effective population size ($n \ll N_e$), and it has been shown that violations of this assumption may result in substantial distortions of key genealogical properties ([Wakeley and Takahashi, 2003, Bhaskar et al., 2014]).

Until very recently, coalescent simulators, did not scale up to long chromosomes and the very large sample sizes of modern day data sets, which now comprise hundreds of thousands of individuals (e.g. [Sudlow et al., 2015]). A Markovian approximation

([McVean and Cardin, 2005, Marjoram and Wall, 2006]) of the "spacial" algorithm described in [Wiuf and Hein, 2000], enabled substantial simplification of both analytical and computational treatment of the coalescent model, at the cost of slighly deflated correlation across markers. This approach was used in several efficient simulators (e.g. [Chen et al., 2009, Excoffier and Foll, 2011, Staab et al., 2015]). The recently developed simulators COSI2 ([Shlyakhter et al., 2014]) and SCRM ([Staab et al., 2015]) enable fast simulation under approximate coalescent models. The COSI algorithm proceeds backwards in time, while SCRM adopts a "spacial" approach. Both can be also used to simulate large sample sizes and chromosome-long regions under the "exact" coalescent process with reasonable time and memory requirements.

Here, we present ARGON, an efficient simulator of the DTWF process that scales up to very large chromosomes, and hundreds of thoudsands of samples. The simulator offers substantially improved performance compared to previous DTWF simulators (e.g. SIMCOAL, [Laval and Excoffier, 2004], Serial SIMCOAL [Anderson et al., 2005], GENOME [Liang et al., 2007]), and is comparable or superior to current coalescent simulators in terms of speed and memory usage. In its current version, ARGON supports variable recombination rates, non-crossover gene conversion, population structure, migration, and can efficiently output data on pairwise identical-by-descent (IBD) segment sharing. We show that ARGON remains accurate when the sample size approaches the effective population size, while MS's coalescent approach strongly deviates from the DTWF. We verified that, when $n \ll N_e$, ARGON's output is equivalent to that of MS.

2 Approach

ARGON proceeds backwards in time one generation at a time, occasionally sampling coalescent and recombination events subject to population structure and migration. Each individual is represented as a list of regions that are still being tracked at the current time (i.e. for which not all samples have found a common ancestor). Recombination events are sampled in genetic space from an exponential distribution, and rounded to the closest physiscal base pair position based on the desired recombination rate, which may be varying along the chromosome. A recombination event can be a crossover or a non-crossover event, based on user-specified rates. For each individual, a maximum of two parents are sampled, so that multiple recombination events result in alternating ancestry between two individuals from the previous generation. When two or more individuals choose the same parent, coalescence occurs if the individuals contain overlapping regions of genetic material. During coalescence, regions within individuals are annotated with links to descendant ancestral recombination graph (ARG [Griffiths, 1981]) nodes.

Compared to other DTWF implementations, ARGON offers substantially improved speed and memory usage. In the GENOME simulator, for instance, individuals are represented as arrays of blocks of genetic material of a fixed size, and individuals from all populations are explicitly represented in memory using arrays. In ARGON, large regions are represented as intervals with arbitrary boundary values, and hash map data structures are extensively used to take advantage of sparsity, avoiding explicit representation of all individuals. As in the GENOME simulator, ARGON can run in approximate mode, so that recombinations bleakpoints are rounded to blocks of

a user-specified genetic length. This reduces the granularity of the recombination process, improving speed and memory usage, at the cost of slightly inflated correlation of markers at a short genomic distance. We tested this approximation using non-recombinant blocks of size $10\mu M$ ($1\mu M = 0.000001$ Morgans). Additional tests using $50\mu M$ blocks are described in the Supplementary Note.

ARGON can efficiently output a list of pairwise IBD segments longer than a user-specified centimorgan length. Previous approaches to output simulated IBD sharing data (e.g. [Palamara et al., 2012, Palamara and Pe'er, 2013, Browning and Browning, 2013]) required comparing recent ancestry for all pairs of individual at each marginal tree in the sampled ARG, with computational cost quadratic in the sample size and linear in the number of analyzed marginal trees. IBD segments are delimited by the occurrence of recombination events that change the most recent common ancestor for pairs of samples. In ARGON, these events are detected by visiting internal ARG nodes, which requires less computational effort than the approach based on marginal trees.

3 Results

3.1 Accuracy for small sample sizes

When $n \ll N_e$, the coalescent is a good approximation of the DTWF, and we expect ARGON and MS to provide equivalent results. We performed extensive testing for several scenarios including population size variation, migration across multiple demes, and gene conversion. We report detailed results in the Supplementary Note. We find good agreement between ARGON and MS. We also tested the accuracy of COSI version 2.0, SCRM version 1.6.1, and MSprime version 0.1.6 ([Kelleher et al., 2015]), a new efficient simulator for which a preliminary version was released at the time of writing. MSprime matched MS results, although testing was limited due to several simulation functionalities not being available. SCRM also matched MS results. COSI2 presented small but significant deviations from MS and ARGON simulation results in multiple settings. The heterozygosity in a population of constant effective size $N_e = 20,000$, and mutation rate 2×10^{-8}, for instance, was observed to be 7.9983×10^{-4}, s.e. 3.5×10^{-7} for ARGON; 8.0039×10^{-4}, s.e. 3.5×10^{-7} for MS; 8.0051×10^{-4}, s.e. 3.5×10^{-7} for MSprime; 7.9854×10^{-4}, s.e. 3.4×10^{-7} for COSI2.

3.2 deviation of the coalescent from the DTWF

The coalescent becomes a poor approximation of the DTWF process when the sample size is not substantially smaller than the effective population size ([Wakeley and Takahashi, 2003, Bhaskar et al., 2014]). We verified that ARGON matches the theoretical prediction for the number of singletons and doubletons described in [Bhaskar et al., 2014] for the DTWF (see Table 1). We simulated populations of effective size $N_e = 1,000$ and $N_e = 20,000$ haploid individuals, and sampled all present-day individuals. While ARGON matches the prediction of [Bhaskar et al., 2014] in both exact and approximate mode, MS simulations substantially deviate from the DTWF model. Simulation using other coalescent approaches (e.g. COSI2, SCRM, MSprime), will similarly deviate in these scenarios.

	$n = N_e$	singletons %	doubletons %
MS	$20,000$	$\mathbf{-10.81 \pm 0.01}$	$\mathbf{+4.87 \pm 0.02}$
ARGON	$20,000$	$+0.01 \pm 0.01$	-0.02 ± 0.02
ARGON_{10}	$20,000$	$+0.00 \pm 0.01$	-0.02 ± 0.02
MS	$1,000$	$\mathbf{-10.80 \pm 0.05}$	$\mathbf{+4.80 \pm 0.09}$
ARGON	$1,000$	$+0.03 \pm 0.05$	-0.11 ± 0.09
ARGON_{10}	$1,000$	$+0.02 \pm 0.05$	$+0.17 \pm 0.09$

Table 1: Comparison of theoretical predictions and simulation results for the number of singleton and doubleton alleles when $n = N_e$. We performed $100,000$ independent simulations, for a region of 1 Mb, $\mu = 2 \times 10^{-8}$, and 1 cM/Mb. Reported errors are obtained as $100 \times (\theta_s - \theta_t)/\theta_s$, where θ_s is the average simulation result and θ_t is the theoretical expectation. The $\pm$ sign introduces a standard error. Statistically significant deviations from 0 are highlighted. ARGON_{10} results were obtained running ARGON with non-recombinant blocks of $10\mu M$.

3.3 Scalability to large sample size and whole-chromosome length

We tested the run time and memory usage of ARGON and two recently developed programs that enable simulating very large sample sizes and long chromosomes: COSI2 and SCRM (see Table 2). We find that for large parameter values SCRM generally performs worse than ARGON. COSI2 is generally faster, but scales poorly for memory usage as the size of the region and the sample size grows. Additional tests, including version 0.1.6 of MSprime, are detailed in the Supplementary Note. MSprime was observed to scale well compared to all other simulators, although the program crashed reporting insufficient memory (>60Gb) for large population expansion rates and long chromosomes.

We further tested the performance of approximate algorithms for the same set of simulation parameters (see Table 3). We compared ARGON with a minimum recombination block size of $10\mu M$ (AR_{10}), SCRM with the "-l" flag set to 0 (SC_0), and COSI2 with the "-u" flag set to 0 (CS_0). We find that ARGON's speed and memory usage is substantially improved, at the cost of slightly inflated correlation for neighboring markers (see Supplementary Note). For a constant population of size $N_e = 20,000$, for instance, squared correlation (r^2) of markers 0 to 50Kb apart (using 1 cM/Mb) was increased by ∼27% for AR_{10}, but remained unchanged for markers at a larger distance. For SC_0, r^2 was decreased by ∼1.9% between 0 and 50Kb, ∼5.8% between 50 and 100Kb, and ∼6.0% between 100 and 150Kb. For CS_0, r^2 for the same distance bins was decreased by ∼8.1%, ∼9.3%, and ∼9.0%. AR_{10} is faster than SC_0, and approximately as fast as CS_0, and uses less memory than both simulators for the large values of the test parameters. Additional simulations are detailed in the Supplementary Note, including comparison to the FASTSIMCOAL ([Excoffier and Foll, 2011]) and GENOME algorithms.

				Time			Memory		
Mb	$10^2\rho$	$n/10^3$	A	AR	SC	CS	AR	SC	CS
100	0.0	20	20	0.48	1.42	0.12	10.1	19.6	17.3
300	0.0	20	20	2.56	†	0.59	15.2	†	41.4
500	0.0	20	20	5.35	†	†	19.8	†	†
100	0.0	1	10	0.09	0.02	0.01	6.5	0.5	0.2
100	0.0	10	10	0.12	0.27	0.02	8.5	4.9	4.4
100	0.0	100	10	1.01	6.20	52.33	12.6	48.9	6.8
100	0.0	200	10	3.05	†	†	19.4	†	†
100	0.0	300	10	10.19	†	†	24.6	†	†
100	0.347	20	10	0.39	1.37	0.07	10.5	19.6	9.0
100	0.576	20	10	0.94	2.37	†	13.0	39.1	†
100	1.151	20	10	3.97	†	†	25.0	†	†

Table 2: Comparison of simulation algorithms (AR=ARGON, SC=SCRM, CS=COSI2). Simulation parameters were the chromosome length (Mb), exponential expansion rate (measured as $\rho = \log(A/C)/G$ where A = ancestral size, C = current size, $G = 500$ = generation of expansion start), number of haploid samples (in thousands), and haploid ancestral population size (A). Recombination and mutation rates were set to 1 cM/Mb and 2×10^{-8} mutations per bp per generation, respectively. We compare simulation time (in hours) and memory usage (in Gb). All tests were run on a 2.27GHz Intel Xeon L5640, using a single core and up to 60 Gb of memory. † represents runs terminated due to insufficient memory ($>$ 60Gb) or a memory error. Additional results are shown in the Supplementary Note.

				Time			Memory		
Mb	$10^2\rho$	$n/10^3$	A	AR_{10}	SC_0	CS_0	AR_{10}	SC_0	CS_0
100	0.0	20	20	0.12	1.23	0.15	8.2	19.6	17.3
300	0.0	20	20	0.46	†	0.89	9.7	†	41.4
500	0.0	20	20	0.79	†	†	11.0	†	†
100	0.0	1	10	0.02	0.02	0.01	3.3	0.5	0.3
100	0.0	10	10	0.04	0.28	0.03	7.9	4.9	4.4
100	0.0	100	10	0.34	6.36	⋆	8.0	48.9	⋆
100	0.0	200	10	0.66	†	⋆	8.4	†	⋆
100	0.0	300	10	1.59	†	†	17.6	†	†
100	0.347	20	10	0.09	2.51	0.07	8.2	39.1	8.9
100	0.576	20	10	0.15	†	0.13	8.6	†	9.4
100	1.151	20	10	0.22	†	†	12.0	†	†

Table 3: Comparison of approximate simulation algorithms (AR_{10}=ARGON with minimum recombination block of size $10\mu M$, SC_0=SCRM with "-l" set to 0, CS_0=COSI2 with "-u" set to 0). See Table 2 for a description of the simulation setup. † represents runs terminated due to insufficient memory ($>$ 60Gb) or a memory error. ⋆ indicates that the program took longer than 100 hours to complete. Additional results are shown in the Supplementary Note.

4 Discussion

We described ARGON, a fast and scalable simulator for the discrete time Wright-Fisher process. Accurate simulation of large genealogical and sequence data will facilitate analysis of modern day genomic data sets, which have now reached hundreds of thousands of samples ([Sudlow et al., 2015]). ARGON has similar or superior time and memory performance compared to current coalescent simulators when large samples and long chromosomes are simulated, and enables accurately simulating the DTWF process when the sample size approaches the effective population size, where the coalescent approximation makes current scalable simulators become imprecise. Version 0.1 of ARGON supports arbitrary demographic history with migration, variable mutation/recombination rate and gene conversion, and can output pairwise IBD segments for large sample sizes in both exact and approximate modes.

Acknowledgements

We thank Alkes L. Price for useful discussions and comments on an early draft, and John Wakeley for useful discussions.

Funding

This research was funded by NIH grant R01 MH101244.

References

[Anderson et al., 2005] Anderson, C. N., Ramakrishnan, U., Chan, Y. L., and Hadly, E. A. (2005). Serial simcoal: a population genetics model for data from multiple populations and points in time. *Bioinformatics*, 21(8):1733–1734.

[Beaumont et al., 2002] Beaumont, M. A., Zhang, W., and Balding, D. J. (2002). Approximate bayesian computation in population genetics. *Genetics*, 162(4):2025–2035.

[Bhaskar et al., 2014] Bhaskar, A., Clark, A. G., and Song, Y. S. (2014). Distortion of genealogical properties when the sample is very large. *Proceedings of the National Academy of Sciences*, 111(6):2385–2390.

[Browning and Browning, 2013] Browning, B. L. and Browning, S. R. (2013). Detecting identity by descent and estimating genotype error rates in sequence data. *The American Journal of Human Genetics*, 93(5):840–851.

[Chen et al., 2009] Chen, G. K., Marjoram, P., and Wall, J. D. (2009). Fast and flexible simulation of dna sequence data. *Genome research*, 19(1):136–142.

[Excoffier et al., 2013] Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., and Foll, M. (2013). Robust demographic inference from genomic and snp data.

[Excoffier and Foll, 2011] Excoffier, L. and Foll, M. (2011). Fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, 27(9):1332–1334.

[Fisher et al., 1922] Fisher, R. A. et al. (1922). On the dominance ratio. *Proceedings of the royal society of Edinburgh*, 42:321–341.

[Griffiths, 1981] Griffiths, R. C. (1981). Neutral two-locus multiple allele models with recombination. *Theoretical Population Biology*, 19(2):169–186.

[Hudson, 2002] Hudson, R. R. (2002). Generating samples under a wright–fisher neutral model of genetic variation. *Bioinformatics*, 18(2):337–338.

[Hudson et al., 1990] Hudson, R. R. et al. (1990). Gene genealogies and the coalescent process. *Oxford surveys in evolutionary biology*, 7(1):44.

[Kelleher et al., 2015] Kelleher, J., Etheridge, A. M., and McVean, G. (2015). Efficient coalescent simulation and genealogical analysis for large sample sizes. *bioRxiv*, page 033118.

[Kingman, 1982a] Kingman, J. (1982a). Exchangeability and the evolution of large populations.

[Kingman, 1982b] Kingman, J. F. (1982b). The coalescent. *Stochastic processes and their applications*, 13(3):235–248.

[Kingman, 1982c] Kingman, J. F. (1982c). On the genealogy of large populations. *Journal of Applied Probability*, pages 27–43.

[Laval and Excoffier, 2004] Laval, G. and Excoffier, L. (2004). Simcoal 2.0: a program to simulate genomic diversity over large recombining regions in a subdivided population with a complex history. *Bioinformatics*, 20(15):2485–2487.

[Li and Jakobsson, 2012] Li, S. and Jakobsson, M. (2012). Estimating demographic parameters from large-scale population genomic data using approximate bayesian computation. *BMC genetics*, 13(1):22.

[Liang et al., 2007] Liang, L., Zöllner, S., and Abecasis, G. R. (2007). Genome: a rapid coalescent-based whole genome simulator. *Bioinformatics*, 23(12):1565–1567.

[Marjoram and Wall, 2006] Marjoram, P. and Wall, J. D. (2006). Fast "coalescent" simulation. *BMC Genetics*, 7(1):16.

[McVean and Cardin, 2005] McVean, G. A. and Cardin, N. J. (2005). Approximating the coalescent with recombination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1459):1387–1393.

[Palamara et al., 2012] Palamara, P. F., Lencz, T., Darvasi, A., and Peer, I. (2012). Length distributions of identity by descent reveal fine-scale demographic history. *The American Journal of Human Genetics*, 91(5):809–822.

[Palamara and Pe'er, 2013] Palamara, P. F. and Pe'er, I. (2013). Inference of historical migration rates via haplotype sharing. *Bioinformatics*, 29(13):i180–i188.

[Shlyakhter et al., 2014] Shlyakhter, I., Sabeti, P. C., and Schaffner, S. F. (2014). Cosi2: An efficient simulator of exact and approximate coalescent with selection. *Bioinformatics*, 30(23):3427–3429.

[Staab et al., 2015] Staab, P. R., Zhu, S., Metzler, D., and Lunter, G. (2015). scrm: efficiently simulating long sequences using the approximated coalescent with recombination. *Bioinformatics*, 31(10):1680–1682.

[Sudlow et al., 2015] Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. (2015). Uk biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*, 12(3):1–10.

[Wakeley and Takahashi, 2003] Wakeley, J. and Takahashi, T. (2003). Gene genealogies when the sample size exceeds the effective size of the population. *Molecular biology and evolution*, 20(2):208–213.

[Wiuf and Hein, 2000] Wiuf, C. and Hein, J. (2000). The coalescent with gene conversion. *Genetics*, 155(1):451–462.

[Wright, 1931] Wright, S. (1931). Evolution in mendelian populations. *Genetics*, 16(2):97.

ARGON: fast, whole-genome simulation of the discrete time Wright-Fisher process

Supplementary Note

Pier Francesco Palamara

1 Time and memory performance

We report detailed time and memory performance evaluation results for several simulators. Table 1 shows results for ARGON version 0.1.160101, SCRM version 1.6.1 [Staab et al., 2015], COSI version 2.0 [Shlyakhter et al., 2014], and MSprime version 0.1.6 [Kelleher et al., 2015]. Table 2 reports performance evaluation for several algorithms run in approximate mode: ARGON version 0.1.160101 with 10 and 50 microMorgans, FASTSIMCOAL version 2.5.2.21 (implementing the SMC' approximation), SCRM version 1.6.1 with the "-l" flag set to 0, and COSI version 2.0 with the "-u" flag set to 0. Table 3 reports a comparison of ARGON 0.1.160101 and GENOME version 0.2 [Liang et al., 2007], both run with a non-recombinant block size of 10 and 50 microMorgans. GENOME could not be run in exact mode due to excessive time and memoty requirements.

				Time				Mem			
Mb	ρ	$n/10^3$	A	AR	SC	CS	MS'	AR	SC	CS	MS'
100	0.0	20	20	0.48	1.42	0.12	0.05	10.1	19.6	17.3	2.3
300	0.0	20	20	2.56	†	0.59	0.35	15.2	†	41.4	6.8
500	0.0	20	20	5.35	†	†	†	19.8	†	†	†
700	0.0	20	20	10.16	†	†	†	24.7	†	†	†
900	0.0	20	20	21.63	†	†	†	24.8	†	†	†
100	0.0	1	10	0.09	0.02	0.01	0.01	6.5	0.5	0.2	0.1
100	0.0	10	10	0.12	0.27	0.02	0.01	8.5	4.9	4.4	0.6
100	0.0	100	10	1.01	6.20	52.33	0.10	12.6	48.9	6.8	5.8
100	0.0	200	10	3.05	†	†	0.27	19.4	†	†	12.1
100	0.0	300	10	10.19	†	†	0.60	24.6	†	†	18.7
100	0.00173	20	10	0.28	0.97	0.06	0.03	9.4	19.6	8.8	1.7
100	0.00347	20	10	0.39	1.37	0.07	0.04	10.5	19.6	9.0	2.4
100	0.00448	20	10	0.60	1.90	0.08	0.06	11.5	39.1	9.1	3.1
100	0.00520	20	10	0.78	2.10	0.44	0.06	12.3	39.1	56.6	3.6
100	0.00576	20	10	0.94	2.37	†	†	13.0	39.1	†	†
100	0.01151	20	10	3.97	†	†	†	25.0	†	†	†

Table 1: Comparison of simulation algorithms (AR=ARGON, SC=SCRM, CS=COSI2, MS'=MSprime). Simulation parameters were the chromosome length (Mb), exponential expansion rate (measured as $\rho = \log(A/C)/G$ where A = ancestral size, C = current size, $G = 500$ = generation of expansion start), number of haploid samples (in thousands), and haploid ancestral population size (A). Recombination and mutation rates were set to 1 cM/Mb and 2×10^{-8} mutations per bp per generation, respectively. We compare simulation time (in hours) and memory usage (in Gb). All tests were run on a 2.27GHz Intel Xeon L5640, using a single core and up to 60 Gb of memory. † represents runs terminated due to insufficient memory ($>$ 60Gb) or a memory error.

				Time					Mem				
Mb	ρ	$n/10^3$	A	AR_{10}	AR_{50}	FS	SC_0	CS_0	AR_{10}	AR_{50}	FS	SC_0	CS_0
100	0.0	20	20	0.12	0.05	4.34	1.23	0.15	8.2	8.0	7.5	19.6	17.3
300	0.0	20	20	0.46	0.18	13.69*	†	0.89	9.7	8.7	7.5*	†	41.4
500	0.0	20	20	0.79	0.36	22.83*	†	†	11.0	9.7	7.5*	†	†
700	0.0	20	20	1.52	0.55	31.97*	†	†	12.4	10.2	7.5*	†	†
900	0.0	20	20	1.91	0.65	41.11*	†	†	13.5	10.9	7.5*	†	†
100	0.0	1	10	0.02	0.03	0.03	0.02	0.01	3.3	1.6	0.4	0.5	0.3
100	0.0	10	10	0.04	0.02	0.94	0.28	0.03	7.9	5.3	3.8	4.9	4.4
100	0.0	100	10	0.34	0.12	13.63	6.36	($\star$)	8.0	4.3	37.4	48.9	($\star$)
100	0.0	200	10	0.66	0.37	†	†	($\star$)	8.4	10.4	†	†	($\star$)
100	0.0	300	10	1.59	1.01	†	†	†	17.6	19.4	†	†	†
100	0.00173	20	10	0.06	0.04	2.90	1.31	0.10	8.1	6.7	7.5	19.6	8.8
100	0.00347	20	10	0.09	0.04	3.24	2.51	0.07	8.2	8.0	7.5	39.1	8.9
100	0.00448	20	10	0.17	0.09	4.68	†	0.09	8.7	8.1	7.5	†	9.2
100	0.00520	20	10	0.14	0.14	5.45	†	0.09	8.9	8.0	7.5	†	9.2
100	0.00576	20	10	0.15	0.10	5.51	†	0.13	8.6	8.0	7.6	†	9.4
100	0.01151	20	10	0.22	0.12	18.90	†	†	12.0	9.9	7.6	†	†

Table 2: Comparison of approximate algorithms (AR_{10}=ARGON with $10\mu M$ blocks, AR_{50}=ARGON with $50\mu M$ blocks, FS=FASTSIMCOAL, SC_0=SCRM with "-l" flag set to 0, CS_0=COSI2 with "-u" flag set to 0), time and memory performance. Simulation parameters were the chromosome length (Mb), exponential expansion rate (measured as $\rho = \log(A/C)/G$ where A = ancestral size, C = current size, $G = 500$ = generation of expansion start), number of haploid samples (in thousands), and haploid ancestral population size (A). Recombination and mutation rates were set to 1 cM/Mb and 2×10^{-8} mutations per bp per generation, respectively. We compare simulation time (in hours) and memory usage (in Gb). All tests were run on a 2.27GHz Intel Xeon L5640, using a single core and up to 60 Gb of memory. † represents runs terminated due to insufficient memory (> 60Gb) or a memory error. ($\star$) indicates that the program took longer than 100 hours to complete. * indicates that the result has been linearly extrapolated from previous runs with smaller parameter values.

				Time				Mem			
Mb	ρ	$n/10^3$	A	AR_{10}	AR_{50}	GEN_{10}	GEN_{50}	AR_{10}	AR_{50}	GEN_{10}	GEN_{50}
100	0.0	20	20	0.12	0.05	0.50	0.44	8.2	8.0	18.0	5.9
300	0.0	20	20	0.46	0.18	1.49	1.30	9.7	8.7	54.5	17.8
500	0.0	20	20	0.79	0.36	†	2.23	11.0	9.7	†	30.1
700	0.0	20	20	1.52	0.55	†	3.01	12.4	10.2	†	41.2
900	0.0	20	20	1.91	0.65	†	3.97	13.5	10.9	†	53.9
100	0.0	1	10	0.02	0.03	0.02	0.01	3.3	1.6	0.9	0.2
100	0.0	10	10	0.04	0.02	0.18	0.10	7.9	5.3	8.8	1.8
100	0.0	100	10	0.34	0.12	†	1.94	8.0	4.3	†	28.3
100	0.0	200	10	0.66	0.37	†	†	8.4	10.4	†	†
100	0.0	300	10	1.59	1.01	†	†	17.6	19.4	†	†
100	0.00173	20	10	0.06	0.04	0.36	0.38	8.1	6.7	17.5	5.5
100	0.00347	20	10	0.09	0.04	0.55	0.46	8.2	8.0	17.7	6.0
100	0.00448	20	10	0.17	0.09	0.71	0.62	8.7	8.1	17.8	7.5
100	0.00520	20	10	0.14	0.14	0.79	0.66	8.9	8.0	17.9	8.0
100	0.00576	20	10	0.15	0.10	0.86	0.83	8.6	8.0	18.8	9.2
100	0.01151	20	10	0.22	0.12	2.39	2.09	12.0	9.9	29.7	20.2

Table 3: Comparison of approximate algorithms (AR_{10}=ARGON with $10\mu M$ blocks, AR_{50}=ARGON with $50\mu M$ blocks, GEN_{10}=GENOME with $10\mu M$ blocks, GEN_{50}=GENOME with $50\mu M$ blocks), time and memory performance. Simulation parameters were the chromosome length (Mb), exponential expansion rate (measured as $\rho = \log(A/C)/G$ where A = ancestral size, C = current size, $G = 500$ = generation of expansion start), number of haploid samples (in thousands), and haploid ancestral population size (A). Recombination and mutation rates were set to 1 cM/Mb and 2×10^{-8}, respectively. We compare simulation time (in hours) and memory usage (in Gb). All tests were run on a 2.27GHz Intel Xeon L5640, using a single core and up to 60 Gb of memory. † represents runs terminated due to insufficient memory (> 60Gb) or a memory error.

2 Simulator accuracy

We compared ARGON, MS, COSI2, SCRM and MSprime to test accuracy for small sample sizes, when the coalescent process is a good approximation of the discrete time Wright-Fisher process. We tested a number of scenarios, for which we provide a description and corresponding MS command below. We did not test all scenarios for MSprime, because version 0.1.6 does not support some of the required features.

- Constant size: A population of constant size $N = 20,000$ haploid individuals; 50 haploid samples of 5 Mb, 10^{-8} recombinations per bp per generation, 2×10^{-8} mutations per bp per generation.

  ```
  ms 50 1 -t 4000 -r 2000 5000000
  ```

- Exponential expansion: A population of ancestral size $N = 20,000$ haploid individuals, expanding from generation 500 to a present day size of 200,000 haploid individuals; 50 haploid samples of 5 Mb, 10^{-8} recombinations per bp per generation, 2×10^{-8} mutations per bp per generation.

  ```
  ms 50 1 -t 40000 -r 20000 5000000 -G 1842.07 -eG 0.00125 0.0
  ```

- Island model: 4 demes of constant size 5,000 haploid individuals, symmetric pairwise migration rate of 0.001 individuals per generation; 20 samples of 5 Mb from each deme, 10^{-8} recombinations per bp per generation, 2×10^{-8} mutations per bp per generation.

  ```
  ms 80 1 -t 1000 -r 500 5000000 -I 4 20 20 20 20 300
  ```

- Island model, high migration: 4 demes of constant size 5000 haploid individuals, symmetric pairwise migration rate of 0.01 individuals per generation; 20 samples of 5 Mb from each deme, 10^{-8} recombinations per bp per generation, 2×10^{-8} mutations per bp per generation.

  ```
  ms 80 1 -t 1000 -r 500 5000000 -I 4 20 20 20 20 30
  ```

- Piecewise constant: A population of varying constant size. Size 100,000 haploid individuals from generation 0 to 100; 5,000 individuals from generation 100 to 200; 50,000 haploid individuals from generation 200 to 300; and ancestral size 10,000 haploid individuals from generation 300 on; 50 haploid samples of 5 Mb, 10^{-8} recombinations per bp per generation, 2×10^{-8} mutations per bp per generation.

  ```
  ms 50 1 -t 20000 -r 10000 5000000 -eN 0.0005 0.05
                   -eN 0.001 0.5 -eN 0.0015 0.1
  ```

The flag "-c 1 300" was added when simulating gene conversion with same rate as recombination, and a mean tract length of 300 bp. The flag "-p 10" was added in all cases to increse the number of reported significant digits, and "-seed" was used to generate independent random draws. We tested the following summary statistics:

- r_{50} correlation of markers at distance between 0 and 50 Kb.
- r_{100} correlation of markers at distance between 50 and 100 Kb.
- r_{150} correlation of markers at distance between 100 and 150 Kb.
- r^2_{50} squared correlation of markers at distance between 0 and 50 Kb.
- r^2_{100} squared correlation of markers at distance between 50 and 100 Kb.
- r^2_{150} squared correlation of markers at distance between 100 and 150 Kb.
- θ heterozygosity computed as $n/(n-1)\sum_i^l 2p_i(1-p_i)/l$ for n samples and l bases in the region, each with sample frequency p_i.
- S number of polymorphic sites in the sample.
- $S_{<.1}$ number of polymorphic sites with minor allele frequency below 10% in the sample.
- $S_{\in[.1,.4]}$ number of polymorphic sites with minor allele frequency between 10% and 40% in the sample.
- $S_{>.4}$ number of polymorphic sites with minor allele frequency above 40% in the sample.

r and r^2 statistics were obtained using Plink 1.9 using the flags "–r" and "–r2", for variants with MAF greater than 10%. The tables below report the value of these statistics for the considered pairs of simulators after $5,000$ independent draws. The column $\Delta\%$ reports differences of each feature as $100 \times (mean_1 - mean_2)/mean_1$. Z-scores corresponding to $p < 0.005$ are highlighted with an asterisk.

2.1 ARGON and MS

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.2178×10^{-2}	3.5×10^{-5}	9.2183×10^{-2}	3.5×10^{-5}	-0.0	-0.10
r_{100}	5.8527×10^{-2}	2.1×10^{-5}	5.8534×10^{-2}	2.1×10^{-5}	-0.0	-0.24
r_{150}	5.2552×10^{-2}	1.9×10^{-5}	5.2566×10^{-2}	1.9×10^{-5}	-0.0	-0.52
r^2_{50}	2.2052×10^{-2}	3.6×10^{-5}	2.2041×10^{-2}	3.6×10^{-5}	0.0	0.22
r^2_{100}	3.1454×10^{-3}	1.8×10^{-5}	3.1807×10^{-3}	1.8×10^{-5}	-1.1	-1.39
r^2_{150}	1.6384×10^{-3}	1.5×10^{-5}	1.6643×10^{-3}	1.6×10^{-5}	-1.6	-1.18
θ	7.9983×10^{-4}	3.5×10^{-7}	8.0039×10^{-4}	3.5×10^{-7}	-0.1	-1.13
S	1.7911×10^{4}	6.1	1.7920×10^{4}	6.2	-0.1	-1.03
$S_{<.1}$	9.5528×10^{3}	4.4	9.5597×10^{3}	4.5	-0.1	-1.10
$S_{\in[.1,.4]}$	6.5735×10^{3}	3.5	6.5698×10^{3}	3.5	0.1	0.75
$S_{>.4}$	1.7849×10^{3}	1.8	1.7910×10^{3}	1.8	-0.3	-2.40

Table 4: s_1 = ARGON, s_2 = MS; model: constant size

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	8.1168×10^{-2}	2.4×10^{-5}	8.1174×10^{-2}	2.4×10^{-5}	-0.0	-0.18
r_{100}	5.1466×10^{-2}	1.2×10^{-5}	5.1492×10^{-2}	1.2×10^{-5}	-0.1	-1.53
r_{150}	4.6663×10^{-2}	9.6×10^{-6}	4.6688×10^{-2}	9.6×10^{-6}	-0.1	-1.84
r_{50}^2	1.8070×10^{-2}	2.8×10^{-5}	1.8065×10^{-2}	2.8×10^{-5}	0.0	0.13
r_{100}^2	2.0247×10^{-3}	1.3×10^{-5}	2.0237×10^{-3}	1.3×10^{-5}	0.0	0.05
r_{150}^2	8.5814×10^{-4}	1.1×10^{-5}	8.5442×10^{-4}	1.1×10^{-5}	0.4	0.24
θ	8.1212×10^{-4}	3.4×10^{-7}	8.1153×10^{-4}	3.4×10^{-7}	0.1	1.23
S	1.9374×10^{4}	5.7	1.9361×10^{4}	5.8	0.1	1.60
$S_{<.1}$	1.1018×10^{4}	3.9	1.1009×10^{4}	4.0	0.1	1.61
$S_{\in[.1,.4]}$	6.5665×10^{3}	3.5	6.5662×10^{3}	3.4	0.0	0.06
$S_{>.4}$	1.7897×10^{3}	1.8	1.7861×10^{3}	1.8	0.2	1.41

Table 5: s_1 = ARGON, s_2 = MS; model: exponential expansion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	7.1071×10^{-2}	2.8×10^{-5}	7.1106×10^{-2}	2.8×10^{-5}	-0.0	-0.88
r_{100}	4.1573×10^{-2}	1.5×10^{-5}	4.1601×10^{-2}	1.5×10^{-5}	-0.1	-1.32
r_{150}	3.6115×10^{-2}	1.3×10^{-5}	3.6136×10^{-2}	1.3×10^{-5}	-0.1	-1.14
r_{50}^2	1.8373×10^{-2}	2.9×10^{-5}	1.8370×10^{-2}	2.9×10^{-5}	0.0	0.07
r_{100}^2	2.6738×10^{-3}	1.3×10^{-5}	2.6934×10^{-3}	1.3×10^{-5}	-0.7	-1.07
r_{150}^2	1.4057×10^{-3}	1.1×10^{-5}	1.4063×10^{-3}	1.1×10^{-5}	-0.0	-0.04
θ	8.0134×10^{-4}	3.4×10^{-7}	8.0134×10^{-4}	3.4×10^{-7}	0.0	0.00
S	1.9847×10^{4}	6.1	1.9844×10^{4}	6.3	0.0	0.34
$S_{<.1}$	1.1314×10^{4}	4.5	1.1313×10^{4}	4.6	0.0	0.16
$S_{\in[.1,.4]}$	6.8044×10^{3}	3.5	6.8034×10^{3}	3.6	0.0	0.20
$S_{>.4}$	1.7284×10^{3}	1.7	1.7283×10^{3}	1.7	0.0	0.04

Table 6: s_1 = ARGON, s_2 = MS; model: island model, high migration

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	7.1075×10^{-2}	2.7×10^{-5}	7.1083×10^{-2}	2.7×10^{-5}	-0.0	-0.21
r_{100}	4.1651×10^{-2}	1.5×10^{-5}	4.1649×10^{-2}	1.5×10^{-5}	0.0	0.09
r_{150}	3.6231×10^{-2}	1.3×10^{-5}	3.6217×10^{-2}	1.3×10^{-5}	0.0	0.76
r_{50}^2	1.8418×10^{-2}	2.8×10^{-5}	1.8444×10^{-2}	2.9×10^{-5}	-0.1	-0.64
r_{100}^2	2.7292×10^{-3}	1.3×10^{-5}	2.7299×10^{-3}	1.3×10^{-5}	-0.0	-0.04
r_{150}^2	1.4522×10^{-3}	1.1×10^{-5}	1.4542×10^{-3}	1.1×10^{-5}	-0.1	-0.13
θ	8.1465×10^{-4}	3.4×10^{-7}	8.1535×10^{-4}	3.4×10^{-7}	-0.1	-1.46
S	2.0168×10^{4}	6.3	2.0181×10^{4}	6.3	-0.1	-1.46
$S_{<.1}$	1.1496×10^{4}	4.6	1.1499×10^{4}	4.6	-0.0	-0.46
$S_{\in[.1,.4]}$	6.9133×10^{3}	3.6	6.9234×10^{3}	3.6	-0.1	-1.98
$S_{>.4}$	1.7585×10^{3}	1.7	1.7587×10^{3}	1.7	-0.0	-0.08

Table 7: s_1 = ARGON, s_2 = MS; model: island model

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	1.0771×10^{-1}	5.4×10^{-5}	1.0772×10^{-1}	5.4×10^{-5}	-0.0	-0.13
r_{100}	6.6321×10^{-2}	2.9×10^{-5}	6.6353×10^{-2}	3.0×10^{-5}	-0.0	-0.77
r_{150}	5.7374×10^{-2}	2.4×10^{-5}	5.7383×10^{-2}	2.4×10^{-5}	-0.0	-0.27
r_{50}^2	3.3727×10^{-2}	6.6×10^{-5}	3.3749×10^{-2}	6.6×10^{-5}	-0.1	-0.24
r_{100}^2	6.4249×10^{-3}	3.4×10^{-5}	6.4446×10^{-3}	3.4×10^{-5}	-0.3	-0.41
r_{150}^2	3.2411×10^{-3}	2.7×10^{-5}	3.1879×10^{-3}	2.7×10^{-5}	1.7	1.39
θ	4.0299×10^{-4}	2.4×10^{-7}	4.0235×10^{-4}	2.4×10^{-7}	0.2	1.89
S	9.2661×10^{3}	4.1	9.2569×10^{3}	4.1	0.1	1.59
$S_{<.1}$	5.0837×10^{3}	2.9	5.0822×10^{3}	2.9	0.0	0.37
$S_{\in[.1,.4]}$	3.2868×10^{3}	2.4	3.2811×10^{3}	2.4	0.2	1.68
$S_{>.4}$	8.9563×10^{2}	1.2	8.9358×10^{2}	1.3	0.2	1.16

Table 8: s_1 = ARGON, s_2 = MS; model: piecewise model

2.2 ARGON and COSI2

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.2178×10^{-2}	3.5×10^{-5}	9.2145×10^{-2}	3.5×10^{-5}	0.0	0.67
r_{100}	5.8527×10^{-2}	2.1×10^{-5}	5.8521×10^{-2}	2.1×10^{-5}	0.0	0.20
r_{150}	5.2552×10^{-2}	1.9×10^{-5}	5.2558×10^{-2}	1.9×10^{-5}	-0.0	-0.22
r_{50}^2	2.2052×10^{-2}	3.6×10^{-5}	2.2074×10^{-2}	3.6×10^{-5}	-0.1	-0.43
r_{100}^2	3.1454×10^{-3}	1.8×10^{-5}	3.1855×10^{-3}	1.8×10^{-5}	-1.3	-1.58
r_{150}^2	1.6384×10^{-3}	1.5×10^{-5}	1.6661×10^{-3}	1.5×10^{-5}	-1.7	-1.31
θ	7.9983×10^{-4}	3.5×10^{-7}	7.9854×10^{-4}	3.4×10^{-7}	0.2	2.64 *
S	1.7911×10^{4}	6.1	1.7888×10^{4}	6.0	0.1	2.69 *
$S_{<.1}$	9.5528×10^{3}	4.4	9.5462×10^{3}	4.4	0.1	1.06
$S_{\in[.1,.4]}$	6.5735×10^{3}	3.5	6.5568×10^{3}	3.4	0.3	3.42 *
$S_{>.4}$	1.7849×10^{3}	1.8	1.7855×10^{3}	1.8	-0.0	-0.24

Table 9: s_1 = ARGON, s_2 = COSI2; model: constant size

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	8.1168×10^{-2}	2.4×10^{-5}	8.1071×10^{-2}	2.4×10^{-5}	0.1	2.86 *
r_{100}	5.1466×10^{-2}	1.2×10^{-5}	5.1403×10^{-2}	1.2×10^{-5}	0.1	3.71 *
r_{150}	4.6663×10^{-2}	9.6×10^{-6}	4.6612×10^{-2}	9.6×10^{-6}	0.1	3.76 *
r_{50}^2	1.8070×10^{-2}	2.8×10^{-5}	1.8103×10^{-2}	2.9×10^{-5}	-0.2	-0.82
r_{100}^2	2.0247×10^{-3}	1.3×10^{-5}	2.0682×10^{-3}	1.3×10^{-5}	-2.1	-2.37
r_{150}^2	8.5814×10^{-4}	1.1×10^{-5}	8.6197×10^{-4}	1.2×10^{-5}	-0.4	-0.24
θ	8.1212×10^{-4}	3.4×10^{-7}	8.1033×10^{-4}	3.4×10^{-7}	0.2	3.72 *
S	1.9374×10^{4}	5.7	1.9343×10^{4}	5.7	0.2	3.85 *
$S_{<.1}$	1.1018×10^{4}	3.9	1.1004×10^{4}	3.9	0.1	2.54
$S_{\in[.1,.4]}$	6.5665×10^{3}	3.5	6.5548×10^{3}	3.4	0.2	2.40
$S_{>.4}$	1.7897×10^{3}	1.8	1.7840×10^{3}	1.8	0.3	2.24

Table 10: s_1 = ARGON, s_2 = COSI2; model: exponential expansion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	7.1071×10^{-2}	2.8×10^{-5}	7.1090×10^{-2}	2.7×10^{-5}	-0.0	-0.49
r_{100}	4.1573×10^{-2}	1.5×10^{-5}	4.1599×10^{-2}	1.5×10^{-5}	-0.1	-1.23
r_{150}	3.6115×10^{-2}	1.3×10^{-5}	3.6133×10^{-2}	1.3×10^{-5}	-0.0	-0.98
r^2_{50}	1.8373×10^{-2}	2.9×10^{-5}	1.8428×10^{-2}	3.5×10^{-5}	-0.3	-1.21
r^2_{100}	2.6738×10^{-3}	1.3×10^{-5}	2.7169×10^{-3}	2.0×10^{-5}	-1.6	-1.81
r^2_{150}	1.4057×10^{-3}	1.1×10^{-5}	1.4627×10^{-3}	1.8×10^{-5}	-3.9	-2.70 *
θ	8.0134×10^{-4}	3.4×10^{-7}	8.0011×10^{-4}	3.4×10^{-7}	0.2	2.56
S	1.9847×10^{4}	6.1	1.9814×10^{4}	6.2	0.2	3.79 *
$S_{<.1}$	1.1314×10^{4}	4.5	1.1296×10^{4}	4.6	0.2	2.80 *
$S_{\in[.1,.4]}$	6.8044×10^{3}	3.5	6.7908×10^{3}	3.6	0.2	2.71 *
$S_{>.4}$	1.7284×10^{3}	1.7	1.7277×10^{3}	1.7	0.0	0.29

Table 11: s_1 = ARGON, s_2 = COSI2; model: island model, high migration

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	7.1075×10^{-2}	2.7×10^{-5}	7.1084×10^{-2}	2.8×10^{-5}	-0.0	-0.23
r_{100}	4.1651×10^{-2}	1.5×10^{-5}	4.1671×10^{-2}	1.6×10^{-5}	-0.0	-0.91
r_{150}	3.6231×10^{-2}	1.3×10^{-5}	3.6239×10^{-2}	1.3×10^{-5}	-0.0	-0.44
r^2_{50}	1.8418×10^{-2}	2.8×10^{-5}	1.8519×10^{-2}	3.8×10^{-5}	-0.5	-2.14
r^2_{100}	2.7292×10^{-3}	1.3×10^{-5}	2.7876×10^{-3}	2.3×10^{-5}	-2.1	-2.21
r^2_{150}	1.4522×10^{-3}	1.1×10^{-5}	1.5044×10^{-3}	2.1×10^{-5}	-3.5	-2.20
θ	8.1465×10^{-4}	3.4×10^{-7}	8.1365×10^{-4}	3.5×10^{-7}	0.1	2.05
S	2.0168×10^{4}	6.3	2.0142×10^{4}	6.4	0.1	2.90 *
$S_{<.1}$	1.1496×10^{4}	4.6	1.1478×10^{4}	4.7	0.2	2.74 *
$S_{\in[.1,.4]}$	6.9133×10^{3}	3.6	6.9102×10^{3}	3.6	0.0	0.61
$S_{>.4}$	1.7585×10^{3}	1.7	1.7540×10^{3}	1.8	0.3	1.82

Table 12: s_1 = ARGON, s_2 = COSI2; model: island model

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	1.0771×10^{-1}	5.4×10^{-5}	1.0776×10^{-1}	5.4×10^{-5}	-0.0	-0.65
r_{100}	6.6321×10^{-2}	2.9×10^{-5}	6.6391×10^{-2}	3.0×10^{-5}	-0.1	-1.68
r_{150}	5.7374×10^{-2}	2.4×10^{-5}	5.7422×10^{-2}	2.4×10^{-5}	-0.1	-1.41
r^2_{50}	3.3727×10^{-2}	6.6×10^{-5}	3.3736×10^{-2}	6.5×10^{-5}	-0.0	-0.10
r^2_{100}	6.4249×10^{-3}	3.4×10^{-5}	6.4317×10^{-3}	3.5×10^{-5}	-0.1	-0.14
r^2_{150}	3.2411×10^{-3}	2.7×10^{-5}	3.2387×10^{-3}	2.7×10^{-5}	0.1	0.06
θ	4.0299×10^{-4}	2.4×10^{-7}	4.0218×10^{-4}	2.4×10^{-7}	0.2	2.39
S	9.2661×10^{3}	4.1	9.2493×10^{3}	4.2	0.2	2.86 *
$S_{<.1}$	5.0837×10^{3}	2.9	5.0745×10^{3}	2.9	0.2	2.24
$S_{\in[.1,.4]}$	3.2868×10^{3}	2.4	3.2823×10^{3}	2.5	0.1	1.30
$S_{>.4}$	8.9563×10^{2}	1.2	8.9249×10^{2}	1.2	0.4	1.85

Table 13: s_1 = ARGON, s_2 = COSI2; model: piecewise model

2.3 MS and COSI2

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.2183×10^{-2}	3.5×10^{-5}	9.2145×10^{-2}	3.5×10^{-5}	0.0	0.77
r_{100}	5.8534×10^{-2}	2.1×10^{-5}	5.8521×10^{-2}	2.1×10^{-5}	0.0	0.44
r_{150}	5.2566×10^{-2}	1.9×10^{-5}	5.2558×10^{-2}	1.9×10^{-5}	0.0	0.30
r_{50}^2	2.2041×10^{-2}	3.6×10^{-5}	2.2074×10^{-2}	3.6×10^{-5}	-0.1	-0.65
r_{100}^2	3.1807×10^{-3}	1.8×10^{-5}	3.1855×10^{-3}	1.8×10^{-5}	-0.2	-0.19
r_{150}^2	1.6643×10^{-3}	1.6×10^{-5}	1.6661×10^{-3}	1.5×10^{-5}	-0.1	-0.08
θ	8.0039×10^{-4}	3.5×10^{-7}	7.9854×10^{-4}	3.4×10^{-7}	0.2	3.79 *
S	1.7920×10^{4}	6.2	1.7888×10^{4}	6.0	0.2	3.71 *
$S_{<.1}$	9.5597×10^{3}	4.5	9.5462×10^{3}	4.4	0.1	2.15
$S_{\in[.1,.4]}$	6.5698×10^{3}	3.5	6.5568×10^{3}	3.4	0.2	2.66 *
$S_{>.4}$	1.7910×10^{3}	1.8	1.7855×10^{3}	1.8	0.3	2.16

Table 14: s_1 = MS, s_2 = COSI2; model: constant size

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	8.1174×10^{-2}	2.4×10^{-5}	8.1071×10^{-2}	2.4×10^{-5}	0.1	3.03 *
r_{100}	5.1492×10^{-2}	1.2×10^{-5}	5.1403×10^{-2}	1.2×10^{-5}	0.2	5.24 *
r_{150}	4.6688×10^{-2}	9.6×10^{-6}	4.6612×10^{-2}	9.6×10^{-6}	0.2	5.60 *
r_{50}^2	1.8065×10^{-2}	2.8×10^{-5}	1.8103×10^{-2}	2.9×10^{-5}	-0.2	-0.94
r_{100}^2	2.0237×10^{-3}	1.3×10^{-5}	2.0682×10^{-3}	1.3×10^{-5}	-2.2	-2.42
r_{150}^2	8.5442×10^{-4}	1.1×10^{-5}	8.6197×10^{-4}	1.2×10^{-5}	-0.9	-0.46
θ	8.1153×10^{-4}	3.4×10^{-7}	8.1033×10^{-4}	3.4×10^{-7}	0.1	2.50
S	1.9361×10^{4}	5.8	1.9343×10^{4}	5.7	0.1	2.21
$S_{<.1}$	1.1009×10^{4}	4.0	1.1004×10^{4}	3.9	0.0	0.90
$S_{\in[.1,.4]}$	6.5662×10^{3}	3.4	6.5548×10^{3}	3.4	0.2	2.37
$S_{>.4}$	1.7861×10^{3}	1.8	1.7840×10^{3}	1.8	0.1	0.82

Table 15: s_1 = MS, s_2 = COSI2; model: exponential expansion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	7.1106×10^{-2}	2.8×10^{-5}	7.1090×10^{-2}	2.7×10^{-5}	0.0	0.41
r_{100}	4.1601×10^{-2}	1.5×10^{-5}	4.1599×10^{-2}	1.5×10^{-5}	0.0	0.09
r_{150}	3.6136×10^{-2}	1.3×10^{-5}	3.6133×10^{-2}	1.3×10^{-5}	0.0	0.16
r_{50}^2	1.8370×10^{-2}	2.9×10^{-5}	1.8428×10^{-2}	3.5×10^{-5}	-0.3	-1.28
r_{100}^2	2.6934×10^{-3}	1.3×10^{-5}	2.7169×10^{-3}	2.0×10^{-5}	-0.9	-0.99
r_{150}^2	1.4063×10^{-3}	1.1×10^{-5}	1.4627×10^{-3}	1.8×10^{-5}	-3.9	-2.67 *
θ	8.0134×10^{-4}	3.4×10^{-7}	8.0011×10^{-4}	3.4×10^{-7}	0.2	2.56
S	1.9844×10^{4}	6.3	1.9814×10^{4}	6.2	0.2	3.39 *
$S_{<.1}$	1.1313×10^{4}	4.6	1.1296×10^{4}	4.6	0.2	2.61 *
$S_{\in[.1,.4]}$	6.8034×10^{3}	3.6	6.7908×10^{3}	3.6	0.2	2.47
$S_{>.4}$	1.7283×10^{3}	1.7	1.7277×10^{3}	1.7	0.0	0.25

Table 16: s_1 = MS, s_2 = COSI2; model: island model, high migration

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	7.1083×10^{-2}	2.7×10^{-5}	7.1084×10^{-2}	2.8×10^{-5}	-0.0	-0.03
r_{100}	4.1649×10^{-2}	1.5×10^{-5}	4.1671×10^{-2}	1.6×10^{-5}	-0.1	-1.00
r_{150}	3.6217×10^{-2}	1.3×10^{-5}	3.6239×10^{-2}	1.3×10^{-5}	-0.1	-1.20
r_{50}^2	1.8444×10^{-2}	2.9×10^{-5}	1.8519×10^{-2}	3.8×10^{-5}	-0.4	-1.57
r_{100}^2	2.7299×10^{-3}	1.3×10^{-5}	2.7876×10^{-3}	2.3×10^{-5}	-2.1	-2.18
r_{150}^2	1.4542×10^{-3}	1.1×10^{-5}	1.5044×10^{-3}	2.1×10^{-5}	-3.3	-2.12
θ	8.1535×10^{-4}	3.4×10^{-7}	8.1365×10^{-4}	3.5×10^{-7}	0.2	3.48 *
S	2.0181×10^{4}	6.3	2.0142×10^{4}	6.4	0.2	4.34 *
$S_{<.1}$	1.1499×10^{4}	4.6	1.1478×10^{4}	4.7	0.2	3.19 *
$S_{\in[.1,.4]}$	6.9234×10^{3}	3.6	6.9102×10^{3}	3.6	0.2	2.59 *
$S_{>.4}$	1.7587×10^{3}	1.7	1.7540×10^{3}	1.8	0.3	1.90

Table 17: s_1 = MS, s_2 = COSI2; model: island model

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	1.0772×10^{-1}	5.4×10^{-5}	1.0776×10^{-1}	5.4×10^{-5}	-0.0	-0.52
r_{100}	6.6353×10^{-2}	3.0×10^{-5}	6.6391×10^{-2}	3.0×10^{-5}	-0.1	-0.90
r_{150}	5.7383×10^{-2}	2.4×10^{-5}	5.7422×10^{-2}	2.4×10^{-5}	-0.1	-1.15
r_{50}^2	3.3749×10^{-2}	6.6×10^{-5}	3.3736×10^{-2}	6.5×10^{-5}	0.0	0.14
r_{100}^2	6.4446×10^{-3}	3.4×10^{-5}	6.4317×10^{-3}	3.5×10^{-5}	0.2	0.26
r_{150}^2	3.1879×10^{-3}	2.7×10^{-5}	3.2387×10^{-3}	2.7×10^{-5}	-1.6	-1.33
θ	4.0235×10^{-4}	2.4×10^{-7}	4.0218×10^{-4}	2.4×10^{-7}	0.0	0.50
S	9.2569×10^{3}	4.1	9.2493×10^{3}	4.2	0.1	1.29
$S_{<.1}$	5.0822×10^{3}	2.9	5.0745×10^{3}	2.9	0.2	1.88
$S_{\in[.1,.4]}$	3.2811×10^{3}	2.4	3.2823×10^{3}	2.5	-0.0	-0.35
$S_{>.4}$	8.9358×10^{2}	1.3	8.9249×10^{2}	1.2	0.1	0.62

Table 18: s_1 = MS, s_2 = COSI2; model: piecewise model

2.4 MSprime and COSI2

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.2136×10^{-2}	3.6×10^{-5}	9.2145×10^{-2}	3.5×10^{-5}	-0.0	-0.18
r_{100}	5.8503×10^{-2}	2.1×10^{-5}	5.8521×10^{-2}	2.1×10^{-5}	-0.0	-0.61
r_{150}	5.2525×10^{-2}	1.9×10^{-5}	5.2558×10^{-2}	1.9×10^{-5}	-0.1	-1.23
r^2_{50}	2.2021×10^{-2}	3.6×10^{-5}	2.2074×10^{-2}	3.6×10^{-5}	-0.2	-1.04
r^2_{100}	3.1598×10^{-3}	1.8×10^{-5}	3.1855×10^{-3}	1.8×10^{-5}	-0.8	-1.01
r^2_{150}	1.6740×10^{-3}	1.5×10^{-5}	1.6661×10^{-3}	1.5×10^{-5}	0.5	0.37
θ	8.0051×10^{-4}	3.5×10^{-7}	7.9854×10^{-4}	3.4×10^{-7}	0.2	4.04 *
S	1.7931×10^{4}	6.2	1.7888×10^{4}	6.0	0.2	4.98 *
$S_{<.1}$	9.5670×10^{3}	4.5	9.5462×10^{3}	4.4	0.2	3.30 *
$S_{\in[.1,.4]}$	6.5740×10^{3}	3.6	6.5568×10^{3}	3.4	0.3	3.47 *
$S_{>.4}$	1.7896×10^{3}	1.8	1.7855×10^{3}	1.8	0.2	1.61

Table 19: s_1 = MSprime, s_2 = COSI2; model: constant size

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	8.1239×10^{-2}	2.4×10^{-5}	8.1071×10^{-2}	2.4×10^{-5}	0.2	4.95 *
r_{100}	5.1493×10^{-2}	1.2×10^{-5}	5.1403×10^{-2}	1.2×10^{-5}	0.2	5.30 *
r_{150}	4.6679×10^{-2}	9.5×10^{-6}	4.6612×10^{-2}	9.6×10^{-6}	0.1	4.96 *
r^2_{50}	1.8168×10^{-2}	2.9×10^{-5}	1.8103×10^{-2}	2.9×10^{-5}	0.4	1.58
r^2_{100}	2.0534×10^{-3}	1.4×10^{-5}	2.0682×10^{-3}	1.3×10^{-5}	-0.7	-0.77
r^2_{150}	8.6613×10^{-4}	1.1×10^{-5}	8.6197×10^{-4}	1.2×10^{-5}	0.5	0.26
θ	8.1224×10^{-4}	3.5×10^{-7}	8.1033×10^{-4}	3.4×10^{-7}	0.2	3.91 *
S	1.9370×10^{4}	6.0	1.9343×10^{4}	5.7	0.1	3.26 *
$S_{<.1}$	1.1012×10^{4}	4.0	1.1004×10^{4}	3.9	0.1	1.43
$S_{\in[.1,.4]}$	6.5688×10^{3}	3.5	6.5548×10^{3}	3.4	0.2	2.87 *
$S_{>.4}$	1.7892×10^{3}	1.8	1.7840×10^{3}	1.8	0.3	2.04

Table 20: s_1 = MSprime, s_2 = COSI2; model: exponential expansion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	1.0786×10^{-1}	5.4×10^{-5}	1.0776×10^{-1}	5.4×10^{-5}	0.1	1.31
r_{100}	6.6421×10^{-2}	2.9×10^{-5}	6.6391×10^{-2}	3.0×10^{-5}	0.0	0.72
r_{150}	5.7458×10^{-2}	2.4×10^{-5}	5.7422×10^{-2}	2.4×10^{-5}	0.1	1.06
r_{50}^2	3.3852×10^{-2}	6.5×10^{-5}	3.3736×10^{-2}	6.5×10^{-5}	0.3	1.26
r_{100}^2	6.4862×10^{-3}	3.4×10^{-5}	6.4317×10^{-3}	3.5×10^{-5}	0.8	1.12
r_{150}^2	3.2866×10^{-3}	2.7×10^{-5}	3.2387×10^{-3}	2.7×10^{-5}	1.5	1.25
θ	4.0259×10^{-4}	2.4×10^{-7}	4.0218×10^{-4}	2.4×10^{-7}	0.1	1.21
S	9.2551×10^{3}	4.2	9.2493×10^{3}	4.2	0.1	0.98
$S_{<.1}$	5.0764×10^{3}	2.9	5.0745×10^{3}	2.9	0.0	0.46
$S_{\in[.1,.4]}$	3.2850×10^{3}	2.5	3.2823×10^{3}	2.5	0.1	0.76
$S_{>.4}$	8.9373×10^{2}	1.2	8.9249×10^{2}	1.2	0.1	0.73

Table 21: s_1 = MSprime, s_2 = COSI2; model: piecewise model

2.5 MSprime and ARGON

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.2136×10^{-2}	3.6×10^{-5}	9.2178×10^{-2}	3.5×10^{-5}	-0.0	-0.84
r_{100}	5.8503×10^{-2}	2.1×10^{-5}	5.8527×10^{-2}	2.1×10^{-5}	-0.0	-0.81
r_{150}	5.2525×10^{-2}	1.9×10^{-5}	5.2552×10^{-2}	1.9×10^{-5}	-0.1	-1.00
r_{50}^2	2.2021×10^{-2}	3.6×10^{-5}	2.2052×10^{-2}	3.6×10^{-5}	-0.1	-0.61
r_{100}^2	3.1598×10^{-3}	1.8×10^{-5}	3.1454×10^{-3}	1.8×10^{-5}	0.5	0.57
r_{150}^2	1.6740×10^{-3}	1.5×10^{-5}	1.6384×10^{-3}	1.5×10^{-5}	2.2	1.68
θ	8.0051×10^{-4}	3.5×10^{-7}	7.9983×10^{-4}	3.5×10^{-7}	0.1	1.37
S	1.7931×10^{4}	6.2	1.7911×10^{4}	6.1	0.1	2.30
$S_{<.1}$	9.5670×10^{3}	4.5	9.5528×10^{3}	4.4	0.1	2.26
$S_{\in[.1,.4]}$	6.5740×10^{3}	3.6	6.5735×10^{3}	3.5	0.0	0.10
$S_{>.4}$	1.7896×10^{3}	1.8	1.7849×10^{3}	1.8	0.3	1.85

Table 22: s_1 = MSprime, s_2 = ARGON; model: constant size

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	8.1239×10^{-2}	2.4×10^{-5}	8.1168×10^{-2}	2.4×10^{-5}	0.1	2.09
r_{100}	5.1493×10^{-2}	1.2×10^{-5}	5.1466×10^{-2}	1.2×10^{-5}	0.1	1.59
r_{150}	4.6679×10^{-2}	9.5×10^{-6}	4.6663×10^{-2}	9.6×10^{-6}	0.0	1.18
r_{50}^2	1.8168×10^{-2}	2.9×10^{-5}	1.8070×10^{-2}	2.8×10^{-5}	0.5	2.43
r_{100}^2	2.0534×10^{-3}	1.4×10^{-5}	2.0247×10^{-3}	1.3×10^{-5}	1.4	1.50
r_{150}^2	8.6613×10^{-4}	1.1×10^{-5}	8.5814×10^{-4}	1.1×10^{-5}	0.9	0.51
θ	8.1224×10^{-4}	3.5×10^{-7}	8.1212×10^{-4}	3.4×10^{-7}	0.0	0.25
S	1.9370×10^{4}	6.0	1.9374×10^{4}	5.7	-0.0	-0.48
$S_{<.1}$	1.1012×10^{4}	4.0	1.1018×10^{4}	3.9	-0.1	-1.07
$S_{\in[.1,.4]}$	6.5688×10^{3}	3.5	6.5665×10^{3}	3.5	0.0	0.46
$S_{>.4}$	1.7892×10^{3}	1.8	1.7897×10^{3}	1.8	-0.0	-0.20

Table 23: s_1 = MSprime, s_2 = ARGON; model: exponential expansion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	1.0786×10^{-1}	5.4×10^{-5}	1.0771×10^{-1}	5.4×10^{-5}	0.1	1.96
r_{100}	6.6421×10^{-2}	2.9×10^{-5}	6.6321×10^{-2}	2.9×10^{-5}	0.2	2.44
r_{150}	5.7458×10^{-2}	2.4×10^{-5}	5.7374×10^{-2}	2.4×10^{-5}	0.1	2.47
r_{50}^2	3.3852×10^{-2}	6.5×10^{-5}	3.3727×10^{-2}	6.6×10^{-5}	0.4	1.35
r_{100}^2	6.4862×10^{-3}	3.4×10^{-5}	6.4249×10^{-3}	3.4×10^{-5}	1.0	1.27
r_{150}^2	3.2866×10^{-3}	2.7×10^{-5}	3.2411×10^{-3}	2.7×10^{-5}	1.4	1.19
θ	4.0259×10^{-4}	2.4×10^{-7}	4.0299×10^{-4}	2.4×10^{-7}	-0.1	-1.18
S	9.2551×10^{3}	4.2	9.2661×10^{3}	4.1	-0.1	-1.87
$S_{<.1}$	5.0764×10^{3}	2.9	5.0837×10^{3}	2.9	-0.1	-1.78
$S_{\in[.1,.4]}$	3.2850×10^{3}	2.5	3.2868×10^{3}	2.4	-0.1	-0.52
$S_{>.4}$	8.9373×10^{2}	1.2	8.9563×10^{2}	1.2	-0.2	-1.12

Table 24: s_1 = MSprime, s_2 = ARGON; model: piecewise model

2.6 MSprime and MS

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.2136×10^{-2}	3.6×10^{-5}	9.2183×10^{-2}	3.5×10^{-5}	-0.1	-0.94
r_{100}	5.8503×10^{-2}	2.1×10^{-5}	5.8534×10^{-2}	2.1×10^{-5}	-0.1	-1.04
r_{150}	5.2525×10^{-2}	1.9×10^{-5}	5.2566×10^{-2}	1.9×10^{-5}	-0.1	-1.53
r_{50}^2	2.2021×10^{-2}	3.6×10^{-5}	2.2041×10^{-2}	3.6×10^{-5}	-0.1	-0.39
r_{100}^2	3.1598×10^{-3}	1.8×10^{-5}	3.1807×10^{-3}	1.8×10^{-5}	-0.7	-0.82
r_{150}^2	1.6740×10^{-3}	1.5×10^{-5}	1.6643×10^{-3}	1.6×10^{-5}	0.6	0.44
θ	8.0051×10^{-4}	3.5×10^{-7}	8.0039×10^{-4}	3.5×10^{-7}	0.0	0.24
S	1.7931×10^{4}	6.2	1.7920×10^{4}	6.2	0.1	1.25
$S_{<.1}$	9.5670×10^{3}	4.5	9.5597×10^{3}	4.5	0.1	1.15
$S_{\in[.1,.4]}$	6.5740×10^{3}	3.6	6.5698×10^{3}	3.5	0.1	0.84
$S_{>.4}$	1.7896×10^{3}	1.8	1.7910×10^{3}	1.8	-0.1	-0.55

Table 25: s_1 = MSprime, s_2 = MS; model: constant size

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	8.1239×10^{-2}	2.4×10^{-5}	8.1174×10^{-2}	2.4×10^{-5}	0.1	1.92
r_{100}	5.1493×10^{-2}	1.2×10^{-5}	5.1492×10^{-2}	1.2×10^{-5}	0.0	0.06
r_{150}	4.6679×10^{-2}	9.5×10^{-6}	4.6688×10^{-2}	9.6×10^{-6}	-0.0	-0.67
r_{50}^2	1.8168×10^{-2}	2.9×10^{-5}	1.8065×10^{-2}	2.8×10^{-5}	0.6	2.56
r_{100}^2	2.0534×10^{-3}	1.4×10^{-5}	2.0237×10^{-3}	1.3×10^{-5}	1.5	1.55
r_{150}^2	8.6613×10^{-4}	1.1×10^{-5}	8.5442×10^{-4}	1.1×10^{-5}	1.4	0.75
θ	8.1224×10^{-4}	3.5×10^{-7}	8.1153×10^{-4}	3.4×10^{-7}	0.1	1.46
S	1.9370×10^{4}	6.0	1.9361×10^{4}	5.8	0.0	1.08
$S_{<.1}$	1.1012×10^{4}	4.0	1.1009×10^{4}	4.0	0.0	0.53
$S_{\in[.1,.4]}$	6.5688×10^{3}	3.5	6.5662×10^{3}	3.4	0.0	0.53
$S_{>.4}$	1.7892×10^{3}	1.8	1.7861×10^{3}	1.8	0.2	1.22

Table 26: s_1 = MSprime, s_2 = MS; model: exponential expansion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	1.0786×10^{-1}	5.4×10^{-5}	1.0772×10^{-1}	5.4×10^{-5}	0.1	1.83
r_{100}	6.6421×10^{-2}	2.9×10^{-5}	6.6353×10^{-2}	3.0×10^{-5}	0.1	1.63
r_{150}	5.7458×10^{-2}	2.4×10^{-5}	5.7383×10^{-2}	2.4×10^{-5}	0.1	2.21
r_{50}^2	3.3852×10^{-2}	6.5×10^{-5}	3.3749×10^{-2}	6.6×10^{-5}	0.3	1.11
r_{100}^2	6.4862×10^{-3}	3.4×10^{-5}	6.4446×10^{-3}	3.4×10^{-5}	0.6	0.87
r_{150}^2	3.2866×10^{-3}	2.7×10^{-5}	3.1879×10^{-3}	2.7×10^{-5}	3.1	2.58 *
θ	4.0259×10^{-4}	2.4×10^{-7}	4.0235×10^{-4}	2.4×10^{-7}	0.1	0.71
S	9.2551×10^{3}	4.2	9.2569×10^{3}	4.1	-0.0	-0.31
$S_{<.1}$	5.0764×10^{3}	2.9	5.0822×10^{3}	2.9	-0.1	-1.41
$S_{\in[.1,.4]}$	3.2850×10^{3}	2.5	3.2811×10^{3}	2.4	0.1	1.13
$S_{>.4}$	8.9373×10^{2}	1.2	8.9358×10^{2}	1.3	0.0	0.08

Table 27: s_1 = MSprime, s_2 = MS; model: piecewise model

2.7 SCRM and MS

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.2183×10^{-2}	3.5×10^{-5}	9.2168×10^{-2}	3.5×10^{-5}	0.0	0.30
r_{100}	5.8534×10^{-2}	2.1×10^{-5}	5.8532×10^{-2}	2.1×10^{-5}	0.0	0.07
r_{150}	5.2566×10^{-2}	1.9×10^{-5}	5.2567×10^{-2}	1.9×10^{-5}	-0.0	-0.04
r_{50}^2	2.2041×10^{-2}	3.6×10^{-5}	2.2054×10^{-2}	3.6×10^{-5}	-0.1	-0.26
r_{100}^2	3.1807×10^{-3}	1.8×10^{-5}	3.1620×10^{-3}	1.8×10^{-5}	0.6	0.73
r_{150}^2	1.6643×10^{-3}	1.6×10^{-5}	1.6568×10^{-3}	1.5×10^{-5}	0.5	0.34
θ	8.0039×10^{-4}	3.5×10^{-7}	7.9989×10^{-4}	3.4×10^{-7}	0.1	1.02
S	1.7920×10^{4}	6.2	1.7914×10^{4}	6.1	0.0	0.69
$S_{<.1}$	9.5597×10^{3}	4.5	9.5571×10^{3}	4.4	0.0	0.41
$S_{\in[.1,.4]}$	6.5698×10^{3}	3.5	6.5698×10^{3}	3.5	0.0	0.00
$S_{>.4}$	1.7910×10^{3}	1.8	1.7875×10^{3}	1.7	0.2	1.41

Table 28: s_1 = MS, s_2 = SCRM; model: constant size

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	8.1174×10^{-2}	2.4×10^{-5}	8.1241×10^{-2}	2.4×10^{-5}	-0.1	-1.97
r_{100}	5.1492×10^{-2}	1.2×10^{-5}	5.1489×10^{-2}	1.2×10^{-5}	0.0	0.18
r_{150}	4.6688×10^{-2}	9.6×10^{-6}	4.6670×10^{-2}	9.7×10^{-6}	0.0	1.32
r_{50}^2	1.8065×10^{-2}	2.8×10^{-5}	1.8108×10^{-2}	2.8×10^{-5}	-0.2	-1.09
r_{100}^2	2.0237×10^{-3}	1.3×10^{-5}	2.0801×10^{-3}	1.3×10^{-5}	-2.7	-3.07 *
r_{150}^2	8.5442×10^{-4}	1.1×10^{-5}	8.8644×10^{-4}	1.2×10^{-5}	-3.6	-1.97
θ	8.1153×10^{-4}	3.4×10^{-7}	8.1265×10^{-4}	3.4×10^{-7}	-0.1	-2.33
S	1.9361×10^{4}	5.8	1.9370×10^{4}	5.7	-0.0	-1.11
$S_{<.1}$	1.1009×10^{4}	4.0	1.1005×10^{4}	3.9	0.0	0.72
$S_{\in[.1,.4]}$	6.5662×10^{3}	3.4	6.5709×10^{3}	3.5	-0.1	-0.96
$S_{>.4}$	1.7861×10^{3}	1.8	1.7939×10^{3}	1.8	-0.4	-3.06 *

Table 29: s_1 = MS, s_2 = SCRM; model: exponential expansion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	7.1106×10^{-2}	2.8×10^{-5}	7.1102×10^{-2}	2.7×10^{-5}	0.0	0.10
r_{100}	4.1601×10^{-2}	1.5×10^{-5}	4.1603×10^{-2}	1.5×10^{-5}	-0.0	-0.09
r_{150}	3.6136×10^{-2}	1.3×10^{-5}	3.6127×10^{-2}	1.3×10^{-5}	0.0	0.49
r_{50}^2	1.8370×10^{-2}	2.9×10^{-5}	1.8357×10^{-2}	2.9×10^{-5}	0.1	0.32
r_{100}^2	2.6934×10^{-3}	1.3×10^{-5}	2.6997×10^{-3}	1.3×10^{-5}	-0.2	-0.34
r_{150}^2	1.4063×10^{-3}	1.1×10^{-5}	1.4242×10^{-3}	1.1×10^{-5}	-1.3	-1.15
θ	8.0134×10^{-4}	3.4×10^{-7}	8.0214×10^{-4}	3.3×10^{-7}	-0.1	-1.69
S	1.9844×10^{4}	6.3	1.9859×10^{4}	6.1	-0.1	-1.71
$S_{<.1}$	1.1313×10^{4}	4.6	1.1316×10^{4}	4.5	-0.0	-0.47
$S_{\in[.1,.4]}$	6.8034×10^{3}	3.6	6.8134×10^{3}	3.5	-0.1	-1.99
$S_{>.4}$	1.7283×10^{3}	1.7	1.7288×10^{3}	1.7	-0.0	-0.21

Table 30: s_1 = MS, s_2 = SCRM; model: island model, high migration

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	7.1083×10^{-2}	2.7×10^{-5}	7.1093×10^{-2}	2.7×10^{-5}	-0.0	-0.26
r_{100}	4.1649×10^{-2}	1.5×10^{-5}	4.1669×10^{-2}	1.5×10^{-5}	-0.0	-0.94
r_{150}	3.6217×10^{-2}	1.3×10^{-5}	3.6241×10^{-2}	1.3×10^{-5}	-0.1	-1.31
r_{50}^2	1.8444×10^{-2}	2.9×10^{-5}	1.8427×10^{-2}	2.8×10^{-5}	0.1	0.42
r_{100}^2	2.7299×10^{-3}	1.3×10^{-5}	2.7573×10^{-3}	1.3×10^{-5}	-1.0	-1.49
r_{150}^2	1.4542×10^{-3}	1.1×10^{-5}	1.4797×10^{-3}	1.1×10^{-5}	-1.7	-1.64
θ	8.1535×10^{-4}	3.4×10^{-7}	8.1516×10^{-4}	3.4×10^{-7}	0.0	0.40
S	2.0181×10^{4}	6.3	2.0172×10^{4}	6.2	0.0	1.02
$S_{<.1}$	1.1499×10^{4}	4.6	1.1494×10^{4}	4.6	0.0	0.77
$S_{\in[.1,.4]}$	6.9234×10^{3}	3.6	6.9203×10^{3}	3.5	0.0	0.62
$S_{>.4}$	1.7587×10^{3}	1.7	1.7577×10^{3}	1.7	0.1	0.42

Table 31: s_1 = MS, s_2 = SCRM; model: island model

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	1.0772×10^{-1}	5.4×10^{-5}	1.0759×10^{-1}	5.3×10^{-5}	0.1	1.72
r_{100}	6.6353×10^{-2}	3.0×10^{-5}	6.6284×10^{-2}	2.9×10^{-5}	0.1	1.65
r_{150}	5.7383×10^{-2}	2.4×10^{-5}	5.7342×10^{-2}	2.3×10^{-5}	0.1	1.23
r_{50}^2	3.3749×10^{-2}	6.6×10^{-5}	3.3755×10^{-2}	6.4×10^{-5}	-0.0	-0.07
r_{100}^2	6.4446×10^{-3}	3.4×10^{-5}	6.4767×10^{-3}	3.3×10^{-5}	-0.5	-0.68
r_{150}^2	3.1879×10^{-3}	2.7×10^{-5}	3.2013×10^{-3}	2.7×10^{-5}	-0.4	-0.35
θ	4.0235×10^{-4}	2.4×10^{-7}	4.0283×10^{-4}	2.4×10^{-7}	-0.1	-1.41
S	9.2569×10^{3}	4.1	9.2710×10^{3}	4.0	-0.2	-2.46
$S_{<.1}$	5.0822×10^{3}	2.9	5.0914×10^{3}	2.7	-0.2	-2.32
$S_{\in[.1,.4]}$	3.2811×10^{3}	2.4	3.2862×10^{3}	2.4	-0.2	-1.50
$S_{>.4}$	8.9358×10^{2}	1.3	8.9341×10^{2}	1.3	0.0	0.09

Table 32: s_1 = MS, s_2 = SCRM; model: piecewise model

2.8 ARGON with approximation and MS

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.8023×10^{-2}	4.0×10^{-5}	9.2183×10^{-2}	3.5×10^{-5}	6.3	109.88 *
r_{100}	5.8530×10^{-2}	2.2×10^{-5}	5.8534×10^{-2}	2.1×10^{-5}	-0.0	-0.13
r_{150}	5.2551×10^{-2}	1.9×10^{-5}	5.2566×10^{-2}	1.9×10^{-5}	-0.0	-0.56
r_{50}^2	2.7992×10^{-2}	4.3×10^{-5}	2.2041×10^{-2}	3.6×10^{-5}	27.0	106.12 *
r_{100}^2	3.1324×10^{-3}	2.2×10^{-5}	3.1807×10^{-3}	1.8×10^{-5}	-1.5	-1.70
r_{150}^2	1.6134×10^{-3}	1.9×10^{-5}	1.6643×10^{-3}	1.6×10^{-5}	-3.1	-2.05
θ	8.0026×10^{-4}	3.7×10^{-7}	8.0039×10^{-4}	3.5×10^{-7}	-0.0	-0.26
S	1.7918×10^{4}	6.4	1.7920×10^{4}	6.2	-0.0	-0.22
$S_{<.1}$	9.5565×10^{3}	4.6	9.5597×10^{3}	4.5	-0.0	-0.50
$S_{\in[.1,.4]}$	6.5700×10^{3}	3.9	6.5698×10^{3}	3.5	0.0	0.04
$S_{>.4}$	1.7913×10^{3}	2.2	1.7910×10^{3}	1.8	0.0	0.11

Table 33: s_1 = ARGON with $10\mu M$ approximation, s_2 = MS; model: constant size

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	8.6346×10^{-2}	2.9×10^{-5}	8.1174×10^{-2}	2.4×10^{-5}	6.4	137.40 *
r_{100}	5.1525×10^{-2}	1.3×10^{-5}	5.1492×10^{-2}	1.2×10^{-5}	0.1	1.87
r_{150}	4.6686×10^{-2}	1.0×10^{-5}	4.6688×10^{-2}	9.6×10^{-6}	-0.0	-0.14
r_{50}^2	2.3278×10^{-2}	3.5×10^{-5}	1.8065×10^{-2}	2.8×10^{-5}	28.9	116.30 *
r_{100}^2	2.0818×10^{-3}	1.7×10^{-5}	2.0237×10^{-3}	1.3×10^{-5}	2.9	2.71 *
r_{150}^2	8.8709×10^{-4}	1.5×10^{-5}	8.5442×10^{-4}	1.1×10^{-5}	3.8	1.76
θ	8.1244×10^{-4}	3.7×10^{-7}	8.1153×10^{-4}	3.4×10^{-7}	0.1	1.81
S	1.9378×10^{4}	6.0	1.9361×10^{4}	5.8	0.1	2.04
$S_{<.1}$	1.1019×10^{4}	4.1	1.1009×10^{4}	4.0	0.1	1.75
$S_{\in[.1,.4]}$	6.5677×10^{3}	3.9	6.5662×10^{3}	3.4	0.0	0.29
$S_{>.4}$	1.7912×10^{3}	2.2	1.7861×10^{3}	1.8	0.3	1.79

Table 34: s_1 = ARGON with $10\mu M$ approximation, s_2 = MS; model: exponential expansion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	1.5175×10^{-1}	1.3×10^{-4}	9.2183×10^{-2}	3.5×10^{-5}	64.6	442.45 *
r_{100}	5.9464×10^{-2}	3.0×10^{-5}	5.8534×10^{-2}	2.1×10^{-5}	1.6	25.40 *
r_{150}	5.2784×10^{-2}	2.5×10^{-5}	5.2566×10^{-2}	1.9×10^{-5}	0.4	6.94 *
r_{50}^2	7.8270×10^{-2}	1.5×10^{-4}	2.2041×10^{-2}	3.6×10^{-5}	255.1	364.51 *
r_{100}^2	3.5439×10^{-3}	4.7×10^{-5}	3.1807×10^{-3}	1.8×10^{-5}	11.4	7.22 *
r_{150}^2	1.7672×10^{-3}	4.3×10^{-5}	1.6643×10^{-3}	1.6×10^{-5}	6.2	2.24
θ	8.0073×10^{-4}	5.8×10^{-7}	8.0039×10^{-4}	3.5×10^{-7}	0.0	0.50
S	1.7926×10^{4}	8.6	1.7920×10^{4}	6.2	0.0	0.57
$S_{<.1}$	9.5596×10^{3}	5.9	9.5597×10^{3}	4.5	-0.0	-0.01
$S_{\in[.1,.4]}$	6.5758×10^{3}	6.5	6.5698×10^{3}	3.5	0.1	0.81
$S_{>.4}$	1.7905×10^{3}	4.1	1.7910×10^{3}	1.8	-0.0	-0.11

Table 35: s_1 = ARGON with $50\mu M$ approximation, s_2 = MS; model: constant size

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	1.3411×10^{-1}	1.1×10^{-4}	8.1174×10^{-2}	2.4×10^{-5}	65.2	470.18 *
r_{100}	5.2318×10^{-2}	2.0×10^{-5}	5.1492×10^{-2}	1.2×10^{-5}	1.6	35.41 *
r_{150}	4.6871×10^{-2}	1.6×10^{-5}	4.6688×10^{-2}	9.6×10^{-6}	0.4	9.81 *
r_{50}^2	6.6789×10^{-2}	1.3×10^{-4}	1.8065×10^{-2}	2.8×10^{-5}	269.7	366.40 *
r_{100}^2	2.2923×10^{-3}	3.7×10^{-5}	2.0237×10^{-3}	1.3×10^{-5}	13.3	6.85 *
r_{150}^2	9.4097×10^{-4}	3.3×10^{-5}	8.5442×10^{-4}	1.1×10^{-5}	10.1	2.49
θ	8.1202×10^{-4}	5.8×10^{-7}	8.1153×10^{-4}	3.4×10^{-7}	0.1	0.73
S	1.9371×10^{4}	8.3	1.9361×10^{4}	5.8	0.1	0.99
$S_{<.1}$	1.1017×10^{4}	5.5	1.1009×10^{4}	4.0	0.1	1.18
$S_{\in[.1,.4]}$	6.5643×10^{3}	6.6	6.5662×10^{3}	3.4	-0.0	-0.26
$S_{>.4}$	1.7904×10^{3}	4.2	1.7861×10^{3}	1.8	0.2	0.94

Table 36: s_1 = ARGON with $50\mu M$ approximation, s_2 = MS; model: exponential expansion

2.9 SCRM with approximation and MS

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.2183×10^{-2}	3.5×10^{-5}	9.1822×10^{-2}	3.3×10^{-5}	0.4	7.50 *
r_{100}	5.8534×10^{-2}	2.1×10^{-5}	5.8417×10^{-2}	2.1×10^{-5}	0.2	3.94 *
r_{150}	5.2566×10^{-2}	1.9×10^{-5}	5.2498×10^{-2}	1.8×10^{-5}	0.1	2.60 *
r_{50}^2	2.2041×10^{-2}	3.6×10^{-5}	2.1633×10^{-2}	3.4×10^{-5}	1.9	8.24 *
r_{100}^2	3.1807×10^{-3}	1.8×10^{-5}	3.0075×10^{-3}	1.8×10^{-5}	5.8	6.80 *
r_{150}^2	1.6643×10^{-3}	1.6×10^{-5}	1.5699×10^{-3}	1.5×10^{-5}	6.0	4.30 *
θ	8.0039×10^{-4}	3.5×10^{-7}	8.0000×10^{-4}	2.9×10^{-7}	0.0	0.86
S	1.7920×10^{4}	6.2	1.7921×10^{4}	5.2	-0.0	-0.12
$S_{<.1}$	9.5597×10^{3}	4.5	9.5639×10^{3}	4.0	-0.0	-0.70
$S_{\in[.1,.4]}$	6.5698×10^{3}	3.5	6.5690×10^{3}	3.1	0.0	0.17
$S_{>.4}$	1.7910×10^{3}	1.8	1.7880×10^{3}	1.7	0.2	1.21

Table 37: s_1 = MS, s_2 = SCRM with "-l" set to 0; model: constant size

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	8.1174×10^{-2}	2.4×10^{-5}	8.0886×10^{-2}	2.1×10^{-5}	0.4	9.03 *
r_{100}	5.1492×10^{-2}	1.2×10^{-5}	5.1379×10^{-2}	1.1×10^{-5}	0.2	6.94 *
r_{150}	4.6688×10^{-2}	9.6×10^{-6}	4.6604×10^{-2}	9.3×10^{-6}	0.2	6.28 *
r_{50}^2	1.8065×10^{-2}	2.8×10^{-5}	1.7768×10^{-2}	2.7×10^{-5}	1.7	7.64 *
r_{100}^2	2.0237×10^{-3}	1.3×10^{-5}	1.9254×10^{-3}	1.3×10^{-5}	5.1	5.35 *
r_{150}^2	8.5442×10^{-4}	1.1×10^{-5}	7.8935×10^{-4}	1.1×10^{-5}	8.2	4.18 *
θ	8.1153×10^{-4}	3.4×10^{-7}	8.1218×10^{-4}	2.9×10^{-7}	-0.1	-1.45
S	1.9361×10^{4}	5.8	1.9377×10^{4}	4.8	-0.1	-2.13
$S_{<.1}$	1.1009×10^{4}	4.0	1.1021×10^{4}	3.6	-0.1	-2.23
$S_{\in[.1,.4]}$	6.5662×10^{3}	3.4	6.5683×10^{3}	3.1	-0.0	-0.46
$S_{>.4}$	1.7861×10^{3}	1.8	1.7882×10^{3}	1.7	-0.1	-0.85

Table 38: s_1 = MS, s_2 = SCRM with "-l" set to 0; model: exponential expansion

2.10 COSI2 with approximation and MS

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.2183×10^{-2}	3.5×10^{-5}	9.0614×10^{-2}	3.1×10^{-5}	1.7	33.56 *
r_{100}	5.8534×10^{-2}	2.1×10^{-5}	5.8279×10^{-2}	2.0×10^{-5}	0.4	8.79 *
r_{150}	5.2566×10^{-2}	1.9×10^{-5}	5.2433×10^{-2}	1.8×10^{-5}	0.3	5.08 *
r^2_{50}	2.2041×10^{-2}	3.6×10^{-5}	2.0384×10^{-2}	3.2×10^{-5}	8.1	34.40 *
r^2_{100}	3.1807×10^{-3}	1.8×10^{-5}	2.9096×10^{-3}	1.6×10^{-5}	9.3	11.26 *
r^2_{150}	1.6643×10^{-3}	1.6×10^{-5}	1.5267×10^{-3}	1.4×10^{-5}	9.0	6.47 *
θ	8.0039×10^{-4}	3.5×10^{-7}	7.9873×10^{-4}	2.7×10^{-7}	0.2	3.76 *
S	1.7920×10^{4}	6.2	1.7881×10^{4}	4.9	0.2	4.94 *
$S_{<.1}$	9.5597×10^{3}	4.5	9.5345×10^{3}	3.9	0.3	4.23 *
$S_{\in[.1,.4]}$	6.5698×10^{3}	3.5	6.5601×10^{3}	2.9	0.1	2.13
$S_{>.4}$	1.7910×10^{3}	1.8	1.7868×10^{3}	1.6	0.2	1.74

Table 39: s_1 = MS, s_2 = COSI2 with "-u" set to 0; model: constant size

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	8.1174×10^{-2}	2.4×10^{-5}	7.9725×10^{-2}	2.0×10^{-5}	1.8	46.38 *
r_{100}	5.1492×10^{-2}	1.2×10^{-5}	5.1201×10^{-2}	1.1×10^{-5}	0.6	17.88 *
r_{150}	4.6688×10^{-2}	9.6×10^{-6}	4.6510×10^{-2}	9.0×10^{-6}	0.4	13.53 *
r^2_{50}	1.8065×10^{-2}	2.8×10^{-5}	1.6704×10^{-2}	2.5×10^{-5}	8.1	36.26 *
r^2_{100}	2.0237×10^{-3}	1.3×10^{-5}	1.8652×10^{-3}	1.2×10^{-5}	8.5	8.96 *
r^2_{150}	8.5442×10^{-4}	1.1×10^{-5}	7.6492×10^{-4}	1.1×10^{-5}	11.7	5.75 *
θ	8.1153×10^{-4}	3.4×10^{-7}	8.1039×10^{-4}	2.7×10^{-7}	0.1	2.63 *
S	1.9361×10^{4}	5.8	1.9347×10^{4}	4.6	0.1	1.89
$S_{<.1}$	1.1009×10^{4}	4.0	1.1008×10^{4}	3.4	0.0	0.19
$S_{\in[.1,.4]}$	6.5662×10^{3}	3.4	6.5566×10^{3}	2.9	0.1	2.15
$S_{>.4}$	1.7861×10^{3}	1.8	1.7827×10^{3}	1.6	0.2	1.41

Table 40: s_1 = MS, s_2 = COSI2 with "-u" set to 0; model: exponential expansion

2.11 ARGON and MS, with gene conversion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.0235×10^{-2}	3.3×10^{-5}	9.0314×10^{-2}	3.4×10^{-5}	-0.1	-1.67
r_{100}	5.8409×10^{-2}	2.1×10^{-5}	5.8421×10^{-2}	2.1×10^{-5}	-0.0	-0.40
r_{150}	5.2502×10^{-2}	1.8×10^{-5}	5.2524×10^{-2}	1.9×10^{-5}	-0.0	-0.84
r_{50}^2	2.0442×10^{-2}	3.3×10^{-5}	2.0543×10^{-2}	3.4×10^{-5}	-0.5	-2.13
r_{100}^2	3.1009×10^{-3}	1.7×10^{-5}	3.1490×10^{-3}	1.7×10^{-5}	-1.5	-2.00
r_{150}^2	1.6348×10^{-3}	1.4×10^{-5}	1.6672×10^{-3}	1.5×10^{-5}	-1.9	-1.58
θ	7.9985×10^{-4}	3.2×10^{-7}	7.9986×10^{-4}	3.3×10^{-7}	-0.0	-0.02
S	1.7914×10^{4}	5.7	1.7914×10^{4}	6.0	0.0	0.00
$S_{<.1}$	9.5561×10^{3}	4.3	9.5581×10^{3}	4.3	-0.0	-0.33
$S_{\in[.1,.4]}$	6.5676×10^{3}	3.2	6.5663×10^{3}	3.4	0.0	0.28
$S_{>.4}$	1.7899×10^{3}	1.7	1.7893×10^{3}	1.7	0.0	0.25

Table 41: s_1 = ARGON, s_2 = MS; model: constant size, with gene conversion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	7.9434×10^{-2}	2.2×10^{-5}	7.9490×10^{-2}	2.3×10^{-5}	-0.1	-1.76
r_{100}	5.1360×10^{-2}	1.2×10^{-5}	5.1374×10^{-2}	1.2×10^{-5}	-0.0	-0.82
r_{150}	4.6614×10^{-2}	9.5×10^{-6}	4.6642×10^{-2}	9.6×10^{-6}	-0.1	-2.07
r_{50}^2	1.6686×10^{-2}	2.6×10^{-5}	1.6758×10^{-2}	2.7×10^{-5}	-0.4	-1.92
r_{100}^2	1.9921×10^{-3}	1.2×10^{-5}	2.0245×10^{-3}	1.3×10^{-5}	-1.6	-1.83
r_{150}^2	8.5838×10^{-4}	1.1×10^{-5}	8.7113×10^{-4}	1.1×10^{-5}	-1.5	-0.82
θ	8.1160×10^{-4}	3.2×10^{-7}	8.1173×10^{-4}	3.3×10^{-7}	-0.0	-0.28
S	1.9363×10^{4}	5.5	1.9369×10^{4}	5.8	-0.0	-0.75
$S_{<.1}$	1.1010×10^{4}	3.8	1.1016×10^{4}	4.0	-0.1	-1.09
$S_{\in[.1,.4]}$	6.5673×10^{3}	3.2	6.5676×10^{3}	3.4	-0.0	-0.06
$S_{>.4}$	1.7860×10^{3}	1.6	1.7857×10^{3}	1.7	0.0	0.13

Table 42: s_1 = ARGON, s_2 = MS; model: exponential expansion, with gene conversion

2.12 ARGON and COSI2, with gene conversion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.0235×10^{-2}	3.3×10^{-5}	9.0218×10^{-2}	3.4×10^{-5}	0.0	0.36
r_{100}	5.8409×10^{-2}	2.1×10^{-5}	5.8418×10^{-2}	2.1×10^{-5}	-0.0	-0.30
r_{150}	5.2502×10^{-2}	1.8×10^{-5}	5.2530×10^{-2}	1.9×10^{-5}	-0.1	-1.07
r^2_{50}	2.0442×10^{-2}	3.3×10^{-5}	2.0403×10^{-2}	3.3×10^{-5}	0.2	0.84
r^2_{100}	3.1009×10^{-3}	1.7×10^{-5}	3.1240×10^{-3}	1.7×10^{-5}	-0.7	-0.96
r^2_{150}	1.6348×10^{-3}	1.4×10^{-5}	1.6244×10^{-3}	1.4×10^{-5}	0.6	0.53
θ	7.9985×10^{-4}	3.2×10^{-7}	7.9789×10^{-4}	3.4×10^{-7}	0.2	4.20 *
S	1.7914×10^{4}	5.7	1.7876×10^{4}	6.0	0.2	4.59 *
$S_{<.1}$	9.5561×10^{3}	4.3	9.5407×10^{3}	4.4	0.2	2.50
$S_{\in[.1,.4]}$	6.5676×10^{3}	3.2	6.5519×10^{3}	3.4	0.2	3.36 *
$S_{>.4}$	1.7899×10^{3}	1.7	1.7830×10^{3}	1.7	0.4	2.87 *

Table 43: s_1 = ARGON, s_2 = COSI2; model: constant size, with gene conversion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	7.9434×10^{-2}	2.2×10^{-5}	7.9344×10^{-2}	2.3×10^{-5}	0.1	2.83 *
r_{100}	5.1360×10^{-2}	1.2×10^{-5}	5.1307×10^{-2}	1.2×10^{-5}	0.1	3.12 *
r_{150}	4.6614×10^{-2}	9.5×10^{-6}	4.6586×10^{-2}	9.3×10^{-6}	0.1	2.11
r^2_{50}	1.6686×10^{-2}	2.6×10^{-5}	1.6672×10^{-2}	2.6×10^{-5}	0.1	0.38
r^2_{100}	1.9921×10^{-3}	1.2×10^{-5}	2.0105×10^{-3}	1.2×10^{-5}	-0.9	-1.08
r^2_{150}	8.5838×10^{-4}	1.1×10^{-5}	8.7225×10^{-4}	1.1×10^{-5}	-1.6	-0.89
θ	8.1160×10^{-4}	3.2×10^{-7}	8.0976×10^{-4}	3.3×10^{-7}	0.2	4.00 *
S	1.9363×10^{4}	5.5	1.9340×10^{4}	5.7	0.1	2.90 *
$S_{<.1}$	1.1010×10^{4}	3.8	1.1010×10^{4}	4.0	0.0	0.00
$S_{\in[.1,.4]}$	6.5673×10^{3}	3.2	6.5467×10^{3}	3.3	0.3	4.48 *
$S_{>.4}$	1.7860×10^{3}	1.6	1.7832×10^{3}	1.7	0.2	1.20

Table 44: s_1 = ARGON, s_2 = COSI2; model: exponential expansion, with gene conversion

2.13 MS and COSi2, with gene conversion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	9.0314×10^{-2}	3.4×10^{-5}	9.0218×10^{-2}	3.4×10^{-5}	0.1	2.00
r_{100}	5.8421×10^{-2}	2.1×10^{-5}	5.8418×10^{-2}	2.1×10^{-5}	0.0	0.10
r_{150}	5.2524×10^{-2}	1.9×10^{-5}	5.2530×10^{-2}	1.9×10^{-5}	-0.0	-0.22
r^2_{50}	2.0543×10^{-2}	3.4×10^{-5}	2.0403×10^{-2}	3.3×10^{-5}	0.7	2.95 *
r^2_{100}	3.1490×10^{-3}	1.7×10^{-5}	3.1240×10^{-3}	1.7×10^{-5}	0.8	1.04
r^2_{150}	1.6672×10^{-3}	1.5×10^{-5}	1.6244×10^{-3}	1.4×10^{-5}	2.6	2.09
θ	7.9986×10^{-4}	3.3×10^{-7}	7.9789×10^{-4}	3.4×10^{-7}	0.2	4.16 *
S	1.7914×10^{4}	6.0	1.7876×10^{4}	6.0	0.2	4.48 *
$S_{<.1}$	9.5581×10^{3}	4.3	9.5407×10^{3}	4.4	0.2	2.83 *
$S_{\in[.1,.4]}$	6.5663×10^{3}	3.4	6.5519×10^{3}	3.4	0.2	2.99 *
$S_{>.4}$	1.7893×10^{3}	1.7	1.7830×10^{3}	1.7	0.4	2.62 *

Table 45: s_1 = MS, s_2 = COSI2; model: constant size, with gene conversion

feature	s_1 mean	s_1 s.e.	s_2 mean	s_2 s.e.	$\Delta\%$	z-score
r_{50}	7.9490×10^{-2}	2.3×10^{-5}	7.9344×10^{-2}	2.3×10^{-5}	0.2	4.49 *
r_{100}	5.1374×10^{-2}	1.2×10^{-5}	5.1307×10^{-2}	1.2×10^{-5}	0.1	3.95 *
r_{150}	4.6642×10^{-2}	9.6×10^{-6}	4.6586×10^{-2}	9.3×10^{-6}	0.1	4.19 *
r^2_{50}	1.6758×10^{-2}	2.7×10^{-5}	1.6672×10^{-2}	2.6×10^{-5}	0.5	2.29
r^2_{100}	2.0245×10^{-3}	1.3×10^{-5}	2.0105×10^{-3}	1.2×10^{-5}	0.7	0.79
r^2_{150}	8.7113×10^{-4}	1.1×10^{-5}	8.7225×10^{-4}	1.1×10^{-5}	-0.1	-0.07
θ	8.1173×10^{-4}	3.3×10^{-7}	8.0976×10^{-4}	3.3×10^{-7}	0.2	4.22 *
S	1.9369×10^{4}	5.8	1.9340×10^{4}	5.7	0.1	3.57 *
$S_{<.1}$	1.1016×10^{4}	4.0	1.1010×10^{4}	4.0	0.1	1.06
$S_{\in[.1,.4]}$	6.5676×10^{3}	3.4	6.5467×10^{3}	3.3	0.3	4.41 *
$S_{>.4}$	1.7857×10^{3}	1.7	1.7832×10^{3}	1.7	0.1	1.04

Table 46: s_1 = MS, s_2 = COSI2; model: exponential expansion, with gene conversion

References

[Kelleher et al., 2015] Kelleher, J., Etheridge, A. M., and McVean, G. (2015). Efficient coalescent simulation and genealogical analysis for large sample sizes. *bioRxiv*, page 033118.

[Liang et al., 2007] Liang, L., Zöllner, S., and Abecasis, G. R. (2007). Genome: a rapid coalescent-based whole genome simulator. *Bioinformatics*, 23(12):1565–1567.

[Shlyakhter et al., 2014] Shlyakhter, I., Sabeti, P. C., and Schaffner, S. F. (2014). Cosi2: An efficient simulator of exact and approximate coalescent with selection. *Bioinformatics*, 30(23):3427–3429.

[Staab et al., 2015] Staab, P. R., Zhu, S., Metzler, D., and Lunter, G. (2015). scrm: efficiently simulating long sequences using the approximated coalescent with recombination. *Bioinformatics*, 31(10):1680–1682.