Supporting Information

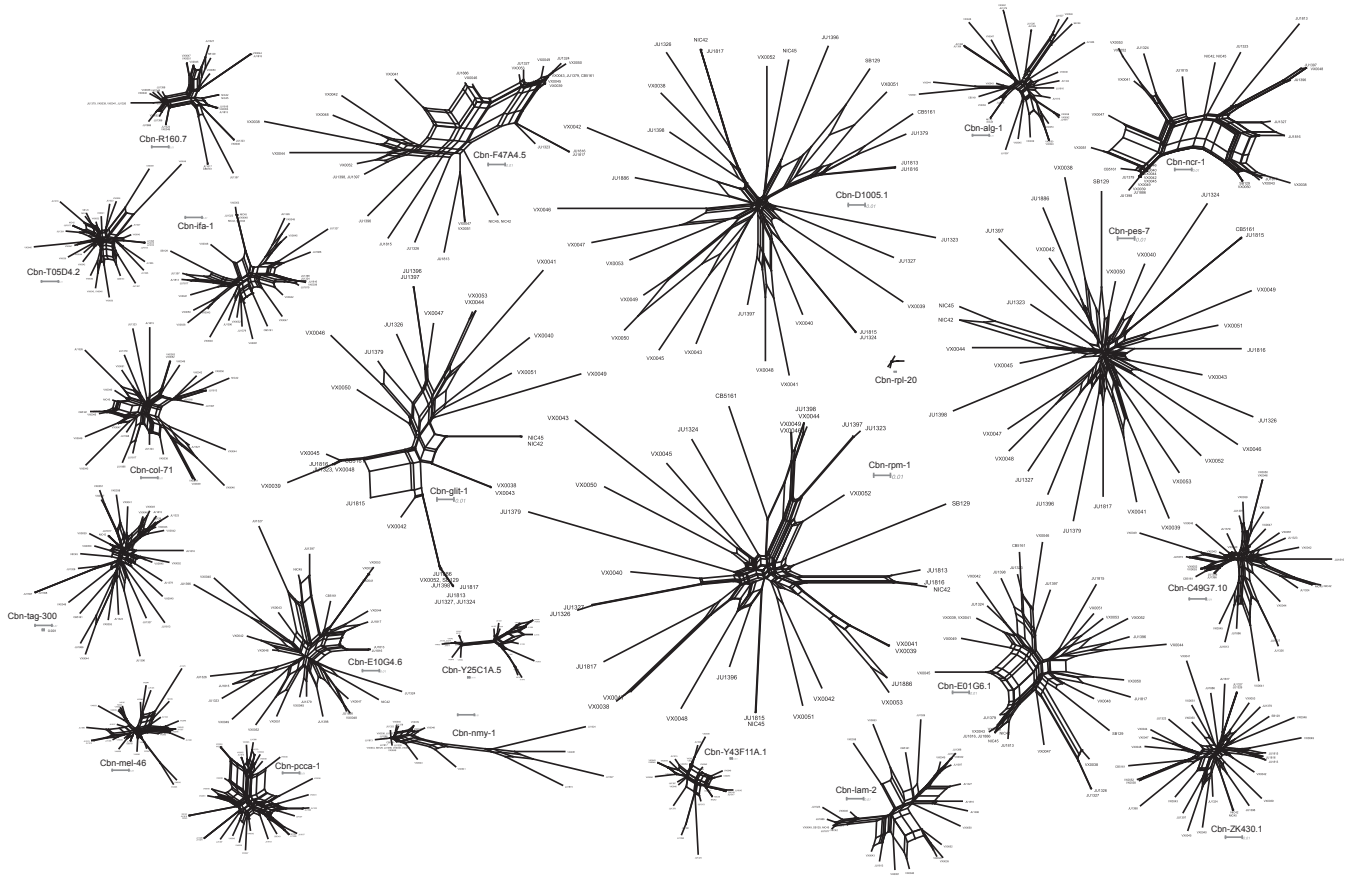## Dey et al. 10.1073/pnas.1303057110

Fig. S1. Per-locus genealogy networks from SplitsTree4 using Jukes-Cantor distances for third codon positions (gaps excluded). Most loci exhibit extensive reticulation owing to recombination and star-shaped topologies indicative of little population substructure within a single species. All genealogical networks are shown at approximately equal scales.
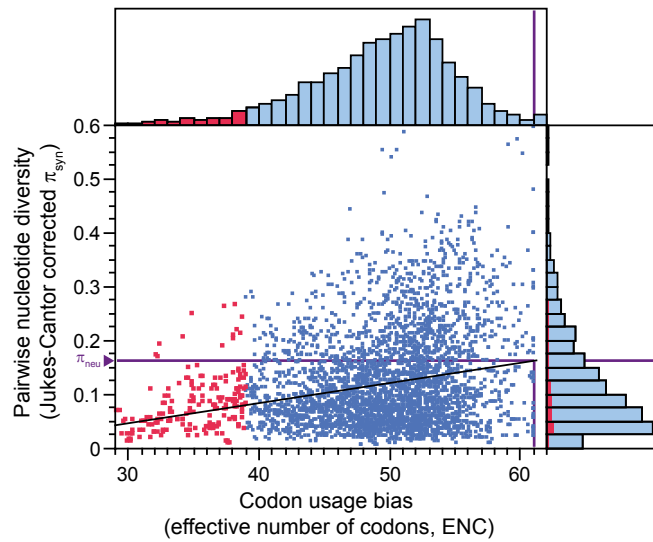
Fig. S2. Nucleotide diversity ($\pi_{syn\text{-}JC}$) is reduced in genes with strong codon bias. Correcting for this effect provides an estimate of neutral polymorphism in the absence of translational selection on synonymous sites. The equation for the least-squares regression line (black line) was used to estimate π_{neu} for effective number of codons (ENC) = 61 (intersection with purple lines). Strongly biased genes (ENC $\leq$ 39) shown in red. Histograms for observed $\pi_{syn\text{-}JC}$ and π_{neu} are indicated adjacent to the biplot.

Table S1. Collection information for *Caenorhabditis brenneri* isofemale strains

Strain	Location	Latitude	Longitude	Substrate	Lab source/collected by
VX0038	Shyamnagar, India (East)	22.819N	88.402E	Rotting hibiscus	Alivia Dey
VX0039	Shyamnagar, India (East)	22.819N	88.402E	Rotting apple	Alivia Dey
VX0040	Shyamnagar, India (East)	22.819N	88.402E	Unidentified rotten fruit	Alivia Dey
VX0041	Shyamnagar, India (East)	22.819N	88.402E	Unidentified rotten fruit	Alivia Dey
VX0042	Shyamnagar, India (East)	22.819N	88.402E	Rotting banana	Alivia Dey
VX0043	Shyamnagar, India (East)	22.819N	88.402E	Rotting carambola fruit	Alivia Dey
VX0044	Shyamnagar, India (East)	22.819N	88.402E	Rotting marigold flower	Alivia Dey
VX0045	Shyamnagar, India (East)	22.819N	88.402E	Rotting marigold flower	Alivia Dey
VX0046	Garia, Calcutta, India (East)	22.466N	88.374E	Rotting sapodilla fruit	Alivia Dey
VX0047	Haridwar, India (North)	29.951N	78.168E	Rotting hibiscus flower	Alivia Dey
VX0048	Haridwar, India (North)	29.951N	78.168E	Rotting hibiscus flower	Alivia Dey
VX0049	Garia, Calcutta, India (East)	22.466N	88.374E	Rotting banana	Alivia Dey
VX0050	Garia, Calcutta, India (East)	22.466N	88.374E	Rotting banana	Alivia Dey
VX0051	Haridwar, India (North)	29.951N	78.168E	Rotting hibiscus flower	Alivia Dey
VX0052	Shyamnagar, India (East)	22.819N	88.402E	Rotting guava	Alivia Dey
VX0053	Shyamnagar, India (East)	22.819N	88.402E	Rotting apple	Alivia Dey
JU1323	Kerala, India (South)	8.309N	77.081E	Rotting coconut and wood	(1)
JU1324	Kerala, India (South)	8.309N	77.081E	Rotting pineapple	(1)
JU1326	Kerala, India (South)	9.6N	76.36E	Rotting banana leaves and stem	(1)
JU1327	Kerala, India (South)	9.55N	76.8E	Rotting pineapples	(1)
JU1813	Montabo, French Guiana	4.949N	52.313W	Small fruits (most likely figs)	(2)
JU1815	Comte river, French Guiana	4.6593N	52.3534W	Rotten seeds/fruits	(2)
JU1816	Cacao, French Guiana	4.52813N	52.5009W	Rotten lemons	(2)
JU1817	Cacao, French Guiana	4.53155N	52.4888W	Rotten kumquat	(2)
JU1379	LaReunion	21.0476S	55.6905E	Rotting *Alpinia purpurata* (red ginger)	(1)
JU1396	Medellin, Colombia	5.9N	75.9W	Rotting noni fruit	(1)
JU1397	Medellin, Colombia	5.9N	75.9W	Rotting orange fruit	(1)
JU1398	Medellin, Colombia	5.9N	75.9W	Rotting orange fruit	(1)
JU1886	Thailand	8.03N	98.93E	Unidentified rotting fruit	Isolated by Marie-Anne Felix
NIC 42	Nouragues Forest Station, French Guiana	4.08788N	52.68W	Figs	(1)
NIC 45	Nouragues Forest Station, French Guiana	4.08788N	52.68W	Figs	(1)
CB5161	Trinidad			Sugar cane	CGC
SB129	Bohrok, Sumatra			Banana plants	CGC

CGC, Caenorhabditis Genetics Center.

1. Kiontke KC, et al. (2011) A phylogeny and molecular barcodes for Caenorhabditis with numerous new species from rotting fruits. *BMC Evol Biol* 11(1):339.
2. Felix M-A, et al. (2013) Species richness, distribution and genetic diversity of Caenorhabditis nematodes in a remote tropical rainforest. *BMC Evol Biol* 13(1):10.

Table S2. Primer sequences used for each locus

Gene name	Primer ID	Forward primer	Reverse primer	Length amplified (bp)	Length retained (bp)	Strains sequenced
Cbn-tag-300	p11	ACGGTATGTGGAGTGAACGG	TCGGCGGCTGTGAGAGCAAT	698	663	33
Cbn-pes-7	p12	GCTCATCGTGACAAGACCGT	GCAAACATCGCCACCGTCTG	719	645	32
Cbn-Y25C1A.5	p15	CAGACTGCCGACACTCTTCA	CTCCAAGGGTTAGACACATGC	856	546	31
Cbn-E01G4.6	p17	GAGTTCCATTCCTTCGTCCA	TATCTTCCATGGCGACCAAT	867	635	30
Cbn-R160.7	p18	CTGACGAGGAGAAGGCAAAG	GGCGCCTTGACACTTTCTAC	666	456	33
Cbn-Y43F11A.1	p21	TGGAATCCAAAGAGTTCAAAAGA	ACCGATCCATTTGGAGATGA	803	497	29
Cbn-D1005.1exon5	p22*	CGAAGATGACACCAAAGCAA	AACGGGAGACGTAGGCAAC	748	1,113	32
Cbn-D1005.1exon6	p86*	GAATCAAGCCAGGAGCATTC	ATCCCAAAACTCCTCCGATT	741	1,113	32
Cbn-col-71	p22a	GTAACTGCTGATACCGAGGG	GGCTCTTCCTTGCTTTCCAT	564	529	32
Cbn-ZK430.1	p24†	TATCCCGTTTTCCTCAATCG	TCGTCTGACGTCTCCACAAG	709	585	33
Cbn-ZK430.1	p24a†	GATTGCTCCTCGGACTCTGC	AGACCCTCCTGCTCACCTTC	721	585	33
Cbn-ZK430.1	p24b†	TTGCTCCTCGGACTCTGCGT	CCTACGCTCCGAAACTCCTA	624	585	33
Cbn-T05D4.2	p31	AAAAGTTGAAGAGCTTCGGAAA	ATGGGGAGTCCACACCAAT	658	475	31
Cbn-lam-2	p35	GGAACCGAGGAAGATTGTCA	GAAGTGGTATGGAGCGCATT	624	590	30
Cbn-ncr-1	p37	CAGCCGATGTCACAGAAAAC	CCAGCCATAACCATTCCAAC	650	614	32
Cbn-rpl-20	p41	ATGCCAACCAAGGCTCTCGG	ACGATGTGGGAGTGGGAAGC	528	493	33
Cbn-pcca-1	p42	CTACGTCATCCGTGGAGTCA	CAGCTTGCTCTGGAAGAACC	545	522	33
Cbn-mel-46	p42a	GCCAACACGGGAGAAAGCCA	TTCCTTCCGCCTCATCTTGC	643	606	31
Cbn-nmy-1	p44	AGTTAGCCACCGCAAGAAGA	CGTTGGCTCTCTTCTCATCC	743	706	27
Cbn-glit-1	p45	GTCCAGTGCCGATGCTTATT	CAGGGCTGGAATGTACTCGT	722	691	33
Cbn-C49G7.10	p51	CTGGGATACCGATTTTCACG	AGTGCGATGAGTGCAACAAC	701	477	31
Cbn-rpm-1	p52	ACGGTGAAGCCCTTAACGGT	GCGTACTCCTCTGGTTGGTT	652	622	33
Cbn-ifa-1	p81	AAGGGTCGTGAACAAGATCG	CGACAAGTTGCTTGAGTCCA	799	754	33
Cbn-F47A4.5	p89	ATCATGGCTGGTTTGGCTAC	TTGTCTTGATCTGGCGACTG	630	593	32
Cbn-alg-1	p92	CACCACTCCACAAACCCTTT	GAGCCTTCTCCAGAGTCGTG	822	782	33
Cbn-E01G6.1	p94	CAAGAGTGCAACTTCGTCCA	GGTCAATGTTCTTGGCTGGT	637	601	33

*Primers p22 and p86 were used for the fragment of *Cbn-D1005.1*.
†Primers p24, p24a, and p24b were used for the fragment of *Cbn-ZK430.1*

Table S3. Nucleotide diversity and variant frequency spectrum summaries for pooled, scattered, and local samples of *C. brenneri*

Locus	Pooled sampling scheme				Scattered sampling scheme				Local sampling scheme (Shyamnagar)			
	π_{syn-JC}	π_{rep-JC}	$D_{Taj-syn}$	$D_{Taj-syn}/D_{min}$	π_{syn-JC}	π_{rep-JC}	$D_{Taj-syn}$	$D_{Taj-syn}/D_{min}$	π_{syn-JC}	π_{rep-JC}	$D_{Taj-syn}$	$D_{Taj-syn}/D_{min}$
Cbn-tag-300	0.1047	0.00081	−1.349	−0.421	0.1053	0.00099	−1.185	−0.403	0.0973	0.00041	−0.732	−0.230
Cbn-pes-7	0.2841	0.00293	−1.138	−0.213	0.2959	0.00359	−0.857	−0.043	0.2543	0.00284	−0.622	0.083
Cbn-Y25C1A.5	0.0424	0.00000	−0.733	−0.256	0.0471	0.00000	−0.313	−0.087	0.0314	0.00000	−0.021	0.043
Cbn-E01G4.6	0.1488	0.00996	−0.360	−0.003	0.1495	0.01039	−0.262	0.074	0.1273	0.00730	−0.175	0.137
Cbn-R160.7	0.0762	0.00759	−0.340	−0.065	0.0851	0.00832	−0.296	−0.031	0.0659	0.00507	−0.574	−0.207
Cbn-Y43F11A.1	0.0486	0.01214	−0.879	−0.306	0.0467	0.01316	−0.679	−0.264	0.0360	0.00803	−1.059	−0.530
Cbn-D1005.1	0.2933	0.00170	−0.573	0.039	0.2852	0.00200	−0.462	0.146	0.3007	0.00204	−0.293	0.452
Cbn-col-71	0.0830	0.00027	−1.080	−0.338	0.0905	0.00016	−0.728	−0.218	0.0859	0.00031	−0.103	0.086
Cbn-ZK430.1	0.1793	0.00529	−0.629	−0.085	0.1716	0.00539	−0.560	−0.046	0.1902	0.00503	0.328	0.523
Cbn-T05D4.2	0.0998	0.02050	−0.930	−0.276	0.0938	0.01711	−0.508	−0.118	0.0809	0.02233	−0.827	−0.315
Cbn-lam-2	0.1469	0.00075	−0.427	−0.032	0.1462	0.00076	−0.076	0.163	0.1561	0.00144	−0.041	0.275
Cbn-ncr-1	0.1345	0.00071	−0.226	0.036	0.1399	0.00085	−0.378	0.005	0.1047	0.00075	0.168	0.263
Cbn-rpl-20	0.0060	0.00033	−1.516	−0.736	0.0041	0.00109	−1.353	−0.839	0.0109	0.00163	−0.326	−0.181
Cbn-pcca-1	0.1046	0.00469	−0.210	0.014	0.1020	0.00436	0.091	0.167	0.1017	0.00514	−0.144	0.090
Cbn-mel-46	0.0999	0.00552	−1.227	−0.382	0.1076	0.00593	−0.890	−0.276	0.0939	0.00423	−0.736	−0.239
Cbn-nmy-1	0.0733	0.00034	−1.633	−0.561	0.0990	0.00015	−1.129	−0.404	0.0346	0.00067	−1.620	−0.879
Cbn-glit-1	0.2504	0.00254	0.054	0.259	0.2217	0.00252	0.188	0.378	0.2689	0.00237	−0.024	0.485
Cbn-C49G7.10	0.1190	0.01339	−1.174	−0.353	0.1246	0.01241	−0.641	−0.142	0.1066	0.01315	−0.722	−0.216
Cbn-rpm-1	0.4333	0.00336	−0.155	0.355	0.4522	0.00330	−0.068	0.602	0.4096	0.00367	0.040	0.870
Cbn-ifa-1	0.1051	0.00077	−0.667	−0.160	0.1104	0.00073	−0.626	−0.146	0.1125	0.00081	−0.189	0.086
Cbn-F47A4.5	0.1657	0.00071	−0.314	0.029	0.1621	0.00058	−0.318	0.061	0.1766	0.00044	0.202	0.422
Cbn-alg-1	0.0972	0.00000	−0.953	−0.274	0.0991	0.00000	−0.965	−0.308	0.0900	0.00000	−0.114	0.088
Cbn-E01G6.1	0.1554	0.00144	−0.602	−0.095	0.1538	0.00135	−0.317	0.046	0.1555	0.00140	−0.559	−0.048
Mean	0.1414	0.0042	−0.742	−0.166	0.1432	0.00414	−0.536	−0.073	0.1344	0.00387	−0.354	0.046

Table S4. Numbers of monomorphic and polymorphic sites for each locus in the globally pooled sample of *C. brenneri*

Locus	Monomorphic sites*	Dimorphic SNPs*	Trimorphic SNPs*	Tetramorphic SNPs*
Cbn-tag-300	588	80	13	0
Cbn-pes-7	487	154	33	3
Cbn-Y25C1A.5	526	27	1	0
Cbn-E01G4.6	545	99	8	1
Cbn-R160.7	413	44	1	0
Cbn-Y43F11A.1	452	60	2	0
Cbn-D1005.1	853	198	36	4
Cbn-col-71	476	63	5	1
Cbn-ZK430.1	496	94	13	2
Cbn-T05D4.2	385	101	4	0
Cbn-lam-2	517	71	4	0
Cbn-ncr-1	549	58	7	1
Cbn-rpl-20	487	11	0	0
Cbn-pcca-1	467	55	2	0
Cbn-mel-46	537	86	12	0
Cbn-nmy-1	639	67	1	0
Cbn-glit-1	573	116	4	1
Cbn-C49G7.10	401	87	4	0
Cbn-rpm-1	473	127	33	5
Cbn-ifa-1	681	70	3	0
Cbn-F47A4.5	515	67	11	0
Cbn-alg-1	699	76	7	0
Cbn-E01G6.1	516	78	7	0
Mean	12,275	1,889	211	18

*Maximum observed across 1,000 randomized haplotype samples per locus.

Table S5. Per-locus recombination estimates ($\rho = 4N_ec$, R_{min}) for scattered and local samples of *C. brenneri*

Locus	Scattered sampling scheme		Local sampling scheme (Shyamnagar)	
	ρ	R_{min}	ρ	R_{min}
Cbn-tag-300	0.136	10	0.108	15
Cbn-pes-7	0.155	30	0.155	43
Cbn-Y25C1A.5	0.018	2	0.031	3
Cbn-E01G4.6	0.157	18	0.084	25
Cbn-R160.7	0.058	6	0.011	9
Cbn-Y43F11A.1	0.098	4	0.055	8
Cbn-D1005.1	0.089	51	0.090	65
Cbn-col-71	0.127	9	0.108	16
Cbn-ZK430.1	0.161	19	0.171	24
Cbn-T05D4.2	0.155	11	0.167	13
Cbn-lam-2	0.036	10	0.116	14
Cbn-ncr-1	0.055	13	0.003	19
Cbn-rpl-20	0.105	0	0.045	0
Cbn-pcca-1	0.083	11	0.123	15
Cbn-mel-46	0.071	9	0.066	14
Cbn-nmy-1	0.003	7	0.000	13
Cbn-glit-1	0.023	20	0.048	35
Cbn-C49G7.10	0.142	12	0.155	22
Cbn-rpm-1	0.149	29	0.125	37
Cbn-ifa-1	0.074	13	0.082	17
Cbn-F47A4.5	0.070	11	0.007	17
Cbn-alg-1	0.104	10	0.128	19
Cbn-E01G6.1	0.072	11	0.073	17
Mean	0.093	13.7	0.085	20.0

Median observed across 1,000 randomized haplotype samples per locus.

Table S6. Genetic differentiation per locus between strains from Shyamnagar (eastern India) and French Guiana

Locus	N_{st}	F_{st}	D_{xy}	Fixed differences	Polymorphic in French Guiana, monomorphic in Shyamnagar	Polymorphic in Shyamnagar, monomorphic in French Guiana	Shared variants	Significance (K_{st}*)
Cbn-tag-300	0.008	0.008	0.019	0	12	24	15	NS
Cbn-pes-7	0.073	0.071	0.053	0	29	70	35	$0.01 < P < 0.05$
Cbn-Y25C1A.5	0.194	0.194	0.011	0	4	2	6	$0.01 < P < 0.05$
Cbn-E01G4.6	0.117	0.116	0.035	0	16	26	27	$0.01 < P < 0.05$
Cbn-R160.7	0.189	0.188	0.020	0	4	8	15	$0.01 < P < 0.05$
Cbn-Y43F11A.1	0.232	0.231	0.017	0	10	12	6	$0.001 < P < 0.01$
Cbn-D1005.1	0.065	0.065	0.055	0	23	100	85	$0.01 < P < 0.05$
Cbn-col-71	0.038	0.038	0.022	0	10	18	12	NS
Cbn-ZK430.1	0.056	0.057	0.036	0	6	35	31	$0.01 < P < 0.05$
Cbn-T05D4.2	0.159	0.158	0.028	0	6	37	10	$P < 0.001$
Cbn-lam-2	0.156	0.156	0.029	0	8	33	15	$0.01 < P < 0.05$
Cbn-ncr-1	0.180	0.179	0.031	0	21	15	21	$0.001 < P < 0.01$
Cbn-rpl-20	0.175	0.175	0.003	0	0	5	1	$0.01 < P < 0.05$
Cbn-pcca-1	0.157	0.156	0.030	0	5	17	21	$0.01 < P < 0.05$
Cbn-mel-46	0.100	0.100	0.020	0	8	34	9	$0.01 < P < 0.05$
Cbn-nmy-1	−0.082	−0.082	0.013	0	17	10	11	NS
Cbn-glit-1	0.043	0.042	0.045	0	9	48	44	NS
Cbn-C49G7.10	0.092	0.091	0.036	0	14	27	18	$0.01 < P < 0.05$
Cbn-rpm-1	0.084	0.086	0.067	0	15	61	56	$0.01 < P < 0.05$
Cbn-ifa-1	−0.047	−0.046	0.019	0	2	22	24	NS
Cbn-F47A4.5	0.016	0.020	0.037	0	17	27	31	$0.01 < P < 0.05$
Cbn-alg-1	0.038	0.039	0.017	0	5	26	17	NS
Cbn-E01G6.1	0.083	0.083	0.028	0	10	41	17	NS
Mean	0.092	0.092	0.029	0	251	698	527	

Significance (K_{st}*) was not adjusted for multiple tests. NS, not significant.

Table S7. Codon use bias metrics and estimates of selection on preferred codons for each locus

Locus	F_{op}*	ENC	Selection intensity for preferred codons[†]		
			$4N_es$	−2Ln	+2Ln
Cbn-tag-300	0.530	48.84	1.0250	0.26	1.86
Cbn-pes-7	0.391	54.18	0.7500	0.09	1.45
Cbn-Y25C1A.5	0.454	47.01	0.8500	−0.78	2.76
Cbn-E01G4.6	0.528	35.91	0.2375	−0.56	1.05
Cbn-R160.7	0.349	59.54	0.6500	−0.57	2.03
Cbn-Y43F11A.1	0.415	55.30	0.4625	−1.09	2.16
Cbn-D1005.1	0.518	40.42	1.7250	0.76	2.89
Cbn-col-71	0.809	30.39	0.1250	−0.30	0.55
Cbn-ZK430.1	0.443	44.28	−0.4625	−1.34	0.36
Cbn-T05D4.2	0.268	54.30	0.5125	−0.78	1.90
Cbn-lam-2	0.559	39.30	0.2750	−0.46	1.03
Cbn-ncr-1	0.473	47.25	0.7000	−0.10	1.55
Cbn-rpl-20	0.742	38.70	−0.1875	−4.28	3.59
Cbn-pcca-1	0.720	38.49	0.1625	−0.84	1.19
Cbn-mel-46	0.434	54.52	0.6125	−0.27	1.56
Cbn-nmy-1	0.501	45.46	0.5500	−0.32	1.48
Cbn-glit-1	0.478	55.00	0.6000	−0.15	1.39
Cbn-C49G7.10	0.384	35.76	−0.0250	−0.96	0.90
Cbn-rpm-1	0.431	56.93	0.0250	−0.59	0.63
Cbn-ifa-1	0.703	33.22	0.9500	0.18	1.79
Cbn-F47A4.5	0.445	50.77	0.4750	−0.36	1.35
Cbn-alg-1	0.606	43.01	0.0125	−0.74	0.78
Cbn-E01G6.1	0.531	40.38	0.3125	−0.38	1.01
Mean	0.509	45.608			
MLE for all loci			0.4125	0.25	0.575

MLE, Maximum likelihood estimate.
*Caenorhabditis elegans codon table.
[†]Based on analysis of a scattered sample.