	Branch length																			
k-mer	**0.00001**				**0.0001**				**0.001**				**0.01**				**0.1**			
	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr
4	2.94	2	2.24	0.22	3.78	4	2.42	0.24	3.92	4	2.52	0.25	3.88	4	2.44	0.24	3.88	4	2.34	0.23
5	1.64	2	1.73	0.17	2.16	2	2.1	0.21	2.34	2	2.23	0.22	2.32	2	2.28	0.22	2.64	2	2.14	0.21
6	0.94	0	1.51	0.15	1.52	2	1.73	0.17	1.56	2	1.59	0.15	1.52	2	1.8	0.18	1.68	2	1.83	0.18
7	9.3	10	1.28	0.12	9.26	10	1.16	0.11	9.3	10	1.21	0.12	9.2	10	1.23	0.12	8.86	10	1.68	0.16
8	8.84	10	1.62	0.16	8.98	10	1.4	0.14	8.84	10	1.5	0.15	8.8	10	1.44	0.14	8.6	8	1.62	0.16
9	0.64	0	1.29	0.12	1.22	0	1.55	0.15	1.16	0	1.48	0.14	0.9	0	1.34	0.13	1.14	0	1.48	0.14
10	0.54	0	1.23	0.12	1.14	0	1.51	0.15	1.02	0	1.4	0.14	0.9	0	1.34	0.13	1.12	0	1.45	0.14
11	0.54	0	1.23	0.12	1.06	0	1.43	0.14	0.92	0	1.31	0.13	0.86	0	1.31	0.13	1.08	0	1.43	0.14
12	0.52	0	1.22	0.12	0.96	0	1.4	0.14	0.86	0	1.27	0.12	0.82	0	1.24	0.12	0.98	0	1.28	0.12
16	0.44	0	1.15	0.15	0.74	0	1.22	0.12	0.74	0	1.19	0.11	0.72	0	1.15	0.11	0.82	0	1.2	0.12

Figure A1: Euclidean distance RF-scores for k-mer 4-16 and for a range of substitution per site of 10E-5 to 10-E1 incremented by orders of magnitude, assembled genomes.

	Branch length																			
k-mer	**0.00001**				**0.0001**				**0.001**				**0.01**				**0.1**			
	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr
4	3.02	2	2.15	0.21	4.2	4	2.55	0.25	4.4	4	2.37	0.23	4.22	4	2.41	0.24	4.2	4	2.43	0.24
5	1.38	2	1.54	0.15	2.02	2	2.02	0.2	2.14	2	2.07	0.2	2.26	2	2.19	0.21	2.66	2	2.03	0.2
6	0.64	0	1.13	0.11	1.02	0	1.49	0.14	1.16	0	1.5	0.15	1.06	0	1.48	0.14	1.44	2	1.58	0.15
7	9.28	10	1.22	0.12	9.14	10	1.34	0.13	9.04	10	1.37	0.13	9.22	10	1.23	0.12	9.3	10	1.34	0.13
8	9.12	10	1.4	0.14	8.82	10	1.73	0.17	8.96	10	1.34	0.13	8.56	9	1.77	0.17	9.04	10	1.31	0.13
9	0.1	0	0.43	0.04	0.14	0	0.51	0.05	0.26	0	0.78	0.07	0.28	0	0.85	0.08	0.26	0	0.83	0.08
10	0.12	0	0.47	0.04	0.1	0	0.52	0.05	0.24	0	0.71	0.07	0.22	0	0.62	0.06	0.18	0	0.7	0.07
11	0.08	0	0.39	0.03	0.08	0	0.48	0.04	0.18	0	0.64	0.06	0.22	0	0.62	0.06	0.2	0	0.66	0.06
12	0.12	0	0.47	0.04	0.08	0	0.39	0.03	0.18	0	0.57	0.05	0.2	0	0.6	0.06	0.12	0	0.47	0.04
16	0.26	0	0.67	0.06	0.22	0	0.69	0.06	0.26	0	0.78	0.07	0.22	0	0.62	0.06	0.22	0	0.62	0.06

Figure A2: Cosine distance RF-scores for k-mer 4-16 and for a range of substitution per site of 10E-5 to 10E-1 incremented by orders of magnitude, assembled genomes.

	Branch length																			
k-mer	**0.00001**				**0.0001**				**0.001**				**0.01**				**0.1**			
	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr
4	3.04	2	2.17	0.21	4.2	4	2.64	0.26	4.36	4	2.36	0.23	4.26	4	2.33	0.23	4.22	4	2.46	0.24
5	1.42	2	1.56	0.15	1.98	2	1.97	0.19	2.12	2	2.08	0.2	2.3	2	2.2	0.22	2.68	2	2.01	0.2
6	0.64	0	1.13	0.11	1.02	0	1.49	0.14	1.12	0	1.48	0.14	1.04	0	1.49	0.14	1.42	2	1.56	0.15
7	9.16	10	1.24	0.12	9.26	10	1.38	0.13	9.3	10	1.07	0.1	9.24	10	1.38	0.13	9.22	10	1.33	0.13
8	9.08	10	1.22	0.12	9.16	10	1.3	0.13	8.96	10	1.51	0.15	9.22	10	1.26	0.12	9.16	10	1.24	0.12
9	8.92	10	1.4	0.14	8.46	8	1.47	8	8.56	1.73	0.17	8	8.42	1.91	0.19	0.08	8.66	9	1.73	0.17
10	0.1	0	0.43	0.04	0.1	0	0.52	0.05	0.24	0	0.71	0.07	0.22	0	0.62	0.06	0.18	0	0.7	0.07
11	0.1	0	0.43	0.04	0.8	0	0.48	0.04	0.16	0	0.54	0.05	0.22	0	0.62	0.06	0.16	0	0.61	0.06
12	0.12	0	0.47	0.04	0.06	0	0.34	0.03	0.16	0	0.54	0.05	0.2	0	0.6	0.06	0.1	0	0.43	0.04
16	0.26	0	0.67	0.06	0.3	0	1.18	0.11	0.24	0	0.76	0.07	0.22	0	0.62	0.06	0.22	0	0.62	0.06
21	0.26	0	0.67	0.06	0.18	0	0.57	0.05	0.36	0	0.91	0.09	0.22	0	0.62	0.06	0.24	0	0.76	0.07
22	0.28	0	0.69	0.06	0.14	0	0.51	0.05	0.32	0	0.83	0.08	0.2	0	0.6	0.06	0.28	0	0.8	0.08
23	0.24	0	0.65	0.06	0.14	0	0.51	0.05	0.34	0	0.85	0.08	0.22	0	0.62	0.06	0.22	0	0.74	0.07

Figure A3: Cosine distance RF-scores for k-mer 4-16 and for a range of substitution per site of 10E-5 to 10E-1 incremented by orders of magnitude, raw reads at 10X uniform coverage

	Branch length																			
k-mer	**0.00001**				**0.0001**				**0.001**				**0.01**				**0.1**			
	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr
10	8.78	8	0.98	0.09	8.68	8	0.95	0.09	8.84	8	0.99	0.09	8.86	8	0.99	0.09	8.78	8	0.98	0.09
11	10	10	0	0	10	10	0	0	9.96	10	0.28	0.02	10	10	0	0	10	10	0	0
12	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0
16	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0
21	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0
22	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0
23	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0

Figure A4: Cosine distance RF-scores for k-mer 10-23 and for a range of substitution per site of 10E-5 to 10E-1 incremented by orders of magnitude, Illumina raw reads 10X coverage

	Branch length																			
k-mer	**0.00001**				**0.0001**				**0.001**				**0.01**				**0.1**			
	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr
10	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0
11	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0
12	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0
16	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0
21	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0
22	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0
23	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0	10	10	0	0

Figure A5: Euclidean distance RF-scores for k-mer 10-23 and for a range of substitution per site of 10E-5 to 10E-1 incremented by orders of magnitude, Illumina raw reads 10X coverage

	Branch length																			
k-mer	**0.00001**				**0.0001**				**0.001**				**0.01**				**0.1**			
	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr	mean	median	stdv	stderr
10	9.8	10	0.63	0.19	9.6	10	0.84	0.26	10	10	0	0	10	10	0	0	10	10	0	0
11	8	8	0	0	8.4	8	0.84	0.26	7.8	8	0.63	0.19	8	8	0	0	8	8	0.94	0.29
12	10	10	0	0	9.2	10	1.39	0.44	9.6	10	0.84	0.26	9.6	10	1.26	0.39	9.2	10	1.68	0.53
16	9.2	10	1.03	0.32	9.2	10	1.39	0.44	9.4	10	0.96	0.3	9.4	10	1.34	0.42	9.6	10	0.84	0.26
21	9	9	1.05	0.33	8.8	9	1.39	0.44	8.4	8	1.57	0.49	8.6	10	2.11	0.66	7.8	8	2.2	0.69
22	9.4	10	0.96	0.3	8.6	8	1.34	0.42	7.6	8	1.83	0.58	8.4	10	2.27	0.71	7.8	8	2.2	0.69
23	9.4	10	0.96	0.3	8.8	9	1.39	0.44	7.6	8	1.83	0.58	8.4	10	2.27	0.71	7.8	8	2.2	0.69

Figure A6: Cosine distance RF-scores for k-mer 10-23 and for a range of substitution per site of 10E-5 to 10E-1 incremented by orders of magnitude, Illumina raw reads 100X coverage