Table 1: Average runtime, Pearson's correlation coefficient and Fowlkes-Mallows index for $k = 4$ and $k = 8$ and the neighbor-joining algorithm. The 'reference' method calculates distances from the original sequences. We show only assembly algorithms that gave the highest and the lowest correlation. From d-type measures, the one with the highest correlation is selected.

| Dataset	method	finished	$\frac{\text{assem.}}{\text{ms}}$	$\frac{\text{distances}}{\text{ms}}$	corr.	B_4	B_8	FM/upgmak4	FM/upgmak8
	reference	112/112	0	4,129	1	1	1	1	1
	$\max(\lvert R_A\rvert, \lvert R_B\rvert)$	112/112	0	343	.801	.66	.32	0.67	0.32
	$\text{Dist}_{\text{MESSG}}(R_A, R_B)$	101/112	0	983,094	.981	1	.99	0.99	1
	$\text{Dist}_{\text{MESSGq}}(R_A, R_B)$	112/112	0	63,425	.98	1	.99	1	0.99
	$\text{Dist}_{\text{C}}(C_A, C_B)$ ABySS	87/112	21,628	17,469	.951	.98	.8	0.98	0.87
	$\text{Dist}_{\text{C}}(C_A, C_B)$ Edena	72/112	285	18,483	.96	1	.85	1	0.88
	$\text{Dist}_{\text{C}}(C_A, C_B)$ SSAKE	68/112	2,079	17,735	.944	.97	.84	0.97	0.87
	$\text{Dist}_{\text{C}}(C_A, C_B)$ Velvet	110/112	385	23,567	.958	.99	.91	0.99	0.91
	$\text{Dist}_{\text{C}}(C_A, C_B)$ SPAdes	43/112	13,529	22,661	.973	.99	.93	0.99	0.95
	$\text{Dist}(RC_A, RC_B)$ ABySS	112/112	4,067	826,926	.976	1	.99	1	1
	$\text{Dist}(RC_A, RC_B)$ Edena	112/112	156	732,929	.978	1	.99	1	1
	$\text{Dist}(RC_A, RC_B)$ SSAKE	111/112	1,562	872,481	.978	1	.99	1	1
	$\text{Dist}(RC_A, RC_B)$ Velvet	107/112	239	903,229	.967	1	.98	1	0.99
	$\text{Dist}(RC_A, RC_B)$ SPAdes	111/112	11,550	935,862	.982	1	.99	1	1
	$\text{Dist}_q(RC_A, RC_B)$ ABySS	112/112	4,067	24,660	.965	.99	.92	1	0.94
	$\text{Dist}_q(RC_A, RC_B)$ Edena	112/112	156	16,540	.968	.99	.95	1	0.96
	$\text{Dist}_q(RC_A, RC_B)$ SSAKE	112/112	1,557	28,155	.969	1	.96	1	0.98
Influenza	$\text{Dist}_q(RC_A, RC_B)$ Velvet	112/112	245	22,581	.955	1	.98	1	0.98
	$\text{Dist}_q(RC_A, RC_B)$ SPAdes	112/112	11,494	28,409	.972	1	.95	1	0.98
	Mash	84/112	0	187	.906	.48	.72	0.51	0.73
	d_2	109/112	0	379	.709	.43	.99	0.3	0.67
	d_2^*	107/112	0	379	.837	.4	.9	0.38	0.71
	d_2^{q*}	109/112	0	365	.631	.32	.27	0.37	0.11
	D_2	110/112	0	374	.443	.32	$1.31 \cdot 10^{-3}$	0.68	0.32
	D_2^*	110/112	0	376	$-.102$	.44	$2.24 \cdot 10^{-3}$	0.5	$3.99 \cdot 10^{-3}$
	D_2^{q*}	111/112	0	370	0	.32	.27	0.37	0.12
	d_2^q	112/112	0	371	.631	.32	.27	0.37	0.11
	D_2^q	112/112	0	364	0	.32	.27	0.37	0.12
	longest contig ABySS	87/112	4,150	2,675	.668	.62	.43	0.63	0.43
	longest contig Edena	72/112	186	2,866	.673	.63	.47	0.64	0.48
	longest contig SSAKE	65/112	1,876	2,504	.66	.59	.37	0.61	0.35
	longest contig Velvet	110/112	245	222	.567	.47	.22	0.45	0.23
	longest contig SPAdes	43/112	12,970	4,157	.75	.71	.55	0.74	0.56
	reference	9/9	0	1,748,984	1	1	1	1	1
	$\max(\lvert R_A\rvert, \lvert R_B\rvert)$	9/9	0	29,340	.181	.72	.83	0.55	0.37
	$\text{Dist}_{\text{MESSG}}(R_A, R_B)$	9/9	0	42,332,682	.965	1	.9	0.99	0.95
	$\text{Dist}_{\text{MESSGq}\alpha}(R_A, R_B)$	9/9	0	1,118,585	.897	1	.94	0.91	0.94
	$\text{Dist}_{\text{C}}(C_A, C_B)$ ABySS	9/9	35,145	48,256,963	.949	1	.91	0.95	0.94
	$\text{Dist}_{\text{C}}(C_A, C_B)$ Edena	9/9	7,038	44,548,818	.892	1	.84	0.95	0.91
	$\text{Dist}_{\text{C}}(C_A, C_B)$ SSAKE	5/9	69,156	55,880,178	.901	1	.95	0.98	0.95
	$\text{Dist}_{\text{C}}(C_A, C_B)$ Velvet	4/9	11,090	59,898,794	.98	1	.99	1	0.97
	$\text{Dist}_{\text{C}}(C_A, C_B)$ SPAdes	2/9	76,514	31,517,537	.869	1	.89	0.87	0.93
	$\text{Dist}(RC_A, RC_B)$ ABySS	0/9	NaN	NaN	NaN	NaN	NaN	NaN	NaN
	$\text{Dist}(RC_A, RC_B)$ Edena	2/9	2,326	255,000,727	.995	1	.91	1	0.95
	$\text{Dist}(RC_A, RC_B)$ SSAKE	0/9	NaN	NaN	NaN	NaN	NaN	NaN	NaN
	$\text{Dist}(RC_A, RC_B)$ Velvet	0/9	NaN	NaN	NaN	NaN	NaN	NaN	NaN
	$\text{Dist}(RC_A, RC_B)$ SPAdes	0/9	NaN	NaN	NaN	NaN	NaN	NaN	NaN
	$\text{Dist}_{q\alpha}(RC_A, RC_B)$ ABySS	9/9	40,453	530,594	.849	1	.86	0.89	0.93
	$\text{Dist}_{q\alpha}(RC_A, RC_B)$ Edena	9/9	9,462	531,265	.894	1	.91	0.87	0.86
Hepatitis	$\text{Dist}_{q\alpha}(RC_A, RC_B)$ SSAKE	9/9	78,036	550,453	.881	1	.89	0.92	0.93
	$\text{Dist}_{q\alpha}(RC_A, RC_B)$ Velvet	9/9	21,409	430,359	.976	1	.96	0.95	0.92
	$\text{Dist}_{q\alpha}(RC_A, RC_B)$ SPAdes	9/9	6,012	343,114	.982	1	.98	0.98	0.95
	Mash	9/9	0	2,350	.967	1	.92	0.96	0.97
	d_2	7/9	0	27,145	.973	1	.86	0.98	0.96
	d_2^*	7/9	0	28,189	.972	1	.9	0.89	0.97
	d_2^{q*}	9/9	0	29,296	.972	1	.9	0.89	0.97
	D_2	8/9	0	30,458	$-.181$	.93	.43	0.66	0.54
	D_2^*	7/9	0	27,718	$-.783$	.93	.41	0.66	0.54
	D_2^{q*}	9/9	0	27,151	$-.782$	.92	.4	0.66	0.54
	d_2^q	9/9	0	27,885	.973	1	.87	0.98	0.96
	D_2^q	7/9	0	31,481	$-.187$	.93	.41	0.66	0.54
	longest contig ABySS	9/9	35,145	2,493,455	.53	.95	.69	0.65	0.69
	longest contig Edena	9/9	7,038	1,581,613	.515	.92	.76	0.78	0.8
	longest contig SSAKE	5/9	69,156	764,321	.334	.86	.62	0.71	0.66
	longest contig Velvet	4/9	11,090	515	.296	.92	.47	0.64	0.58
	longest contig SPAdes	2/9	76,514	3,242,365	.395	.87	.78	0.82	0.77
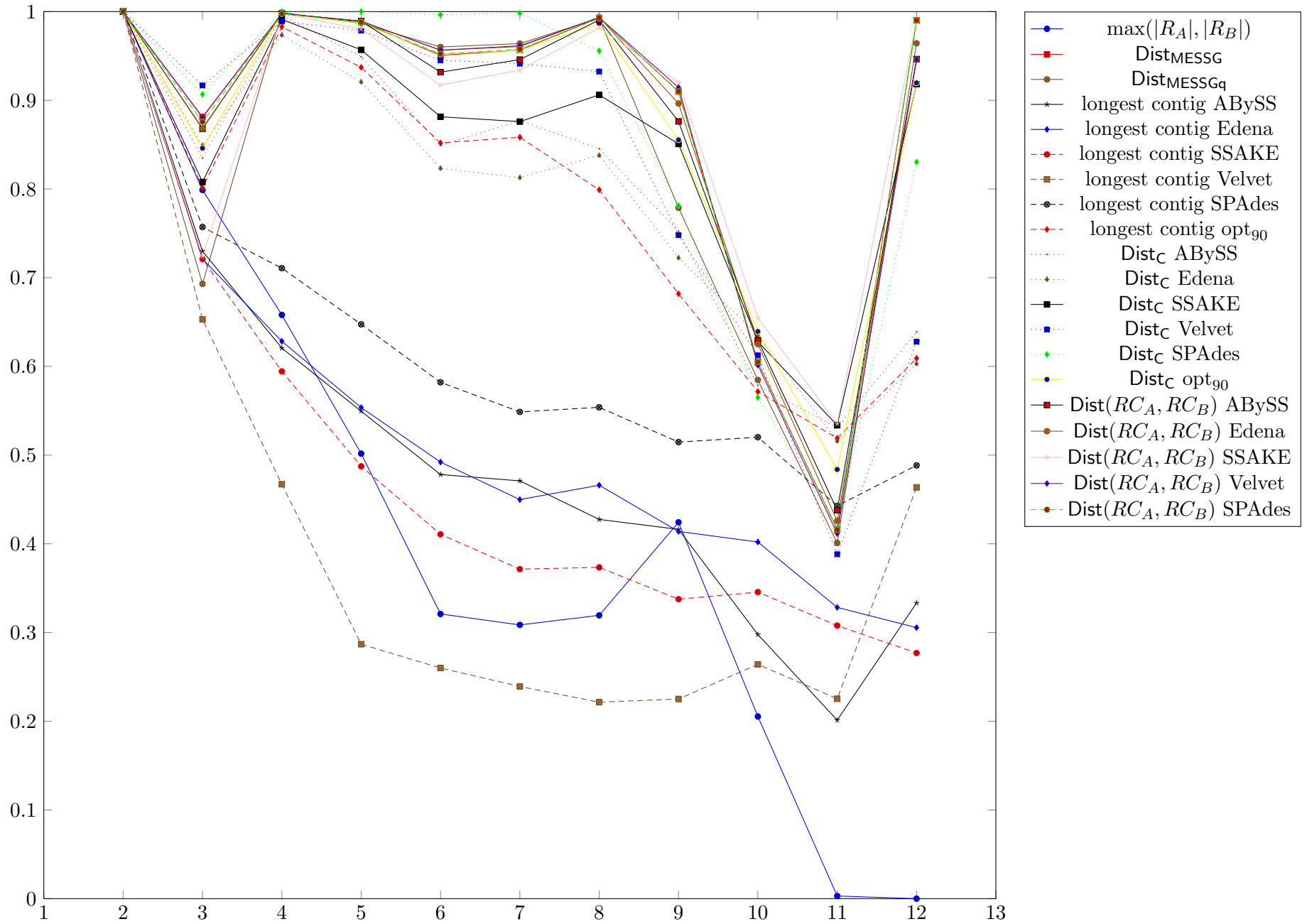
Figure 1: Plot of Fowlkes-Mallows index B_k versus k on influenza dataset. The index compares trees generated by the neighbor-joining algorithm. The tree is compared with the tree generated from the original sequences. If all values are equal to 1, the structures of the trees are the same.
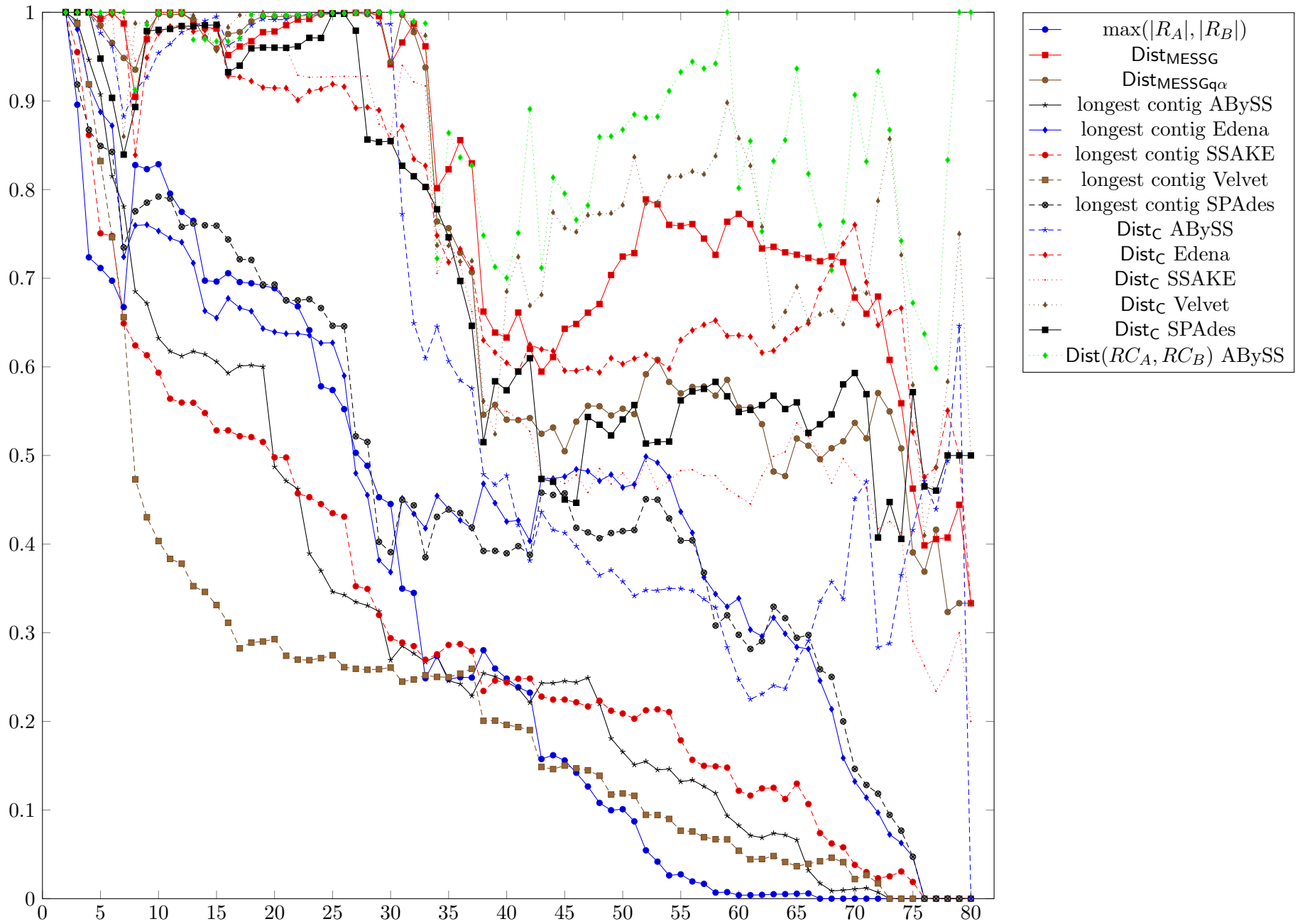
Figure 2: Plot of Fowlkes-Mallows index B_k versus k on hepatitis dataset. The index compares trees generated by the neighbor-joining algorithm.

Legend:
- $\max(|R_A|, |R_B|)$
- $\text{Dist}_{\text{MESSG}}$
- $\text{Dist}_{\text{MESSGq}\alpha}$
- longest contig ABySS
- longest contig Edena
- longest contig SSAKE
- longest contig Velvet
- longest contig SPAdes
- Dist_C ABySS
- Dist_C Edena
- Dist_C SSAKE
- Dist_C Velvet
- Dist_C SPAdes
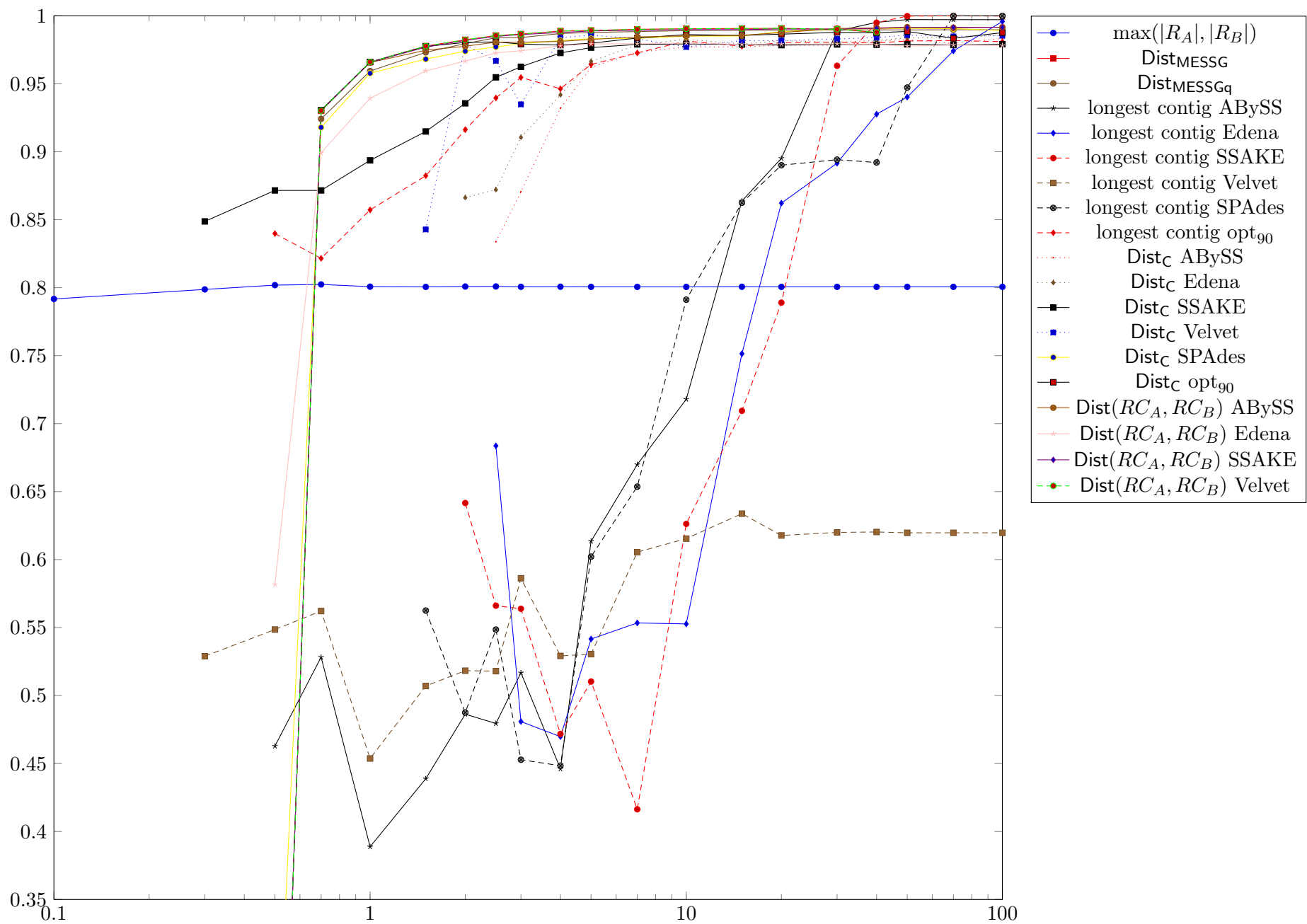- $\text{Dist}(RC_A, RC_B)$ ABySS

Figure 3: Plot of average Pearson's correlation coefficient for several choices of coverage values.
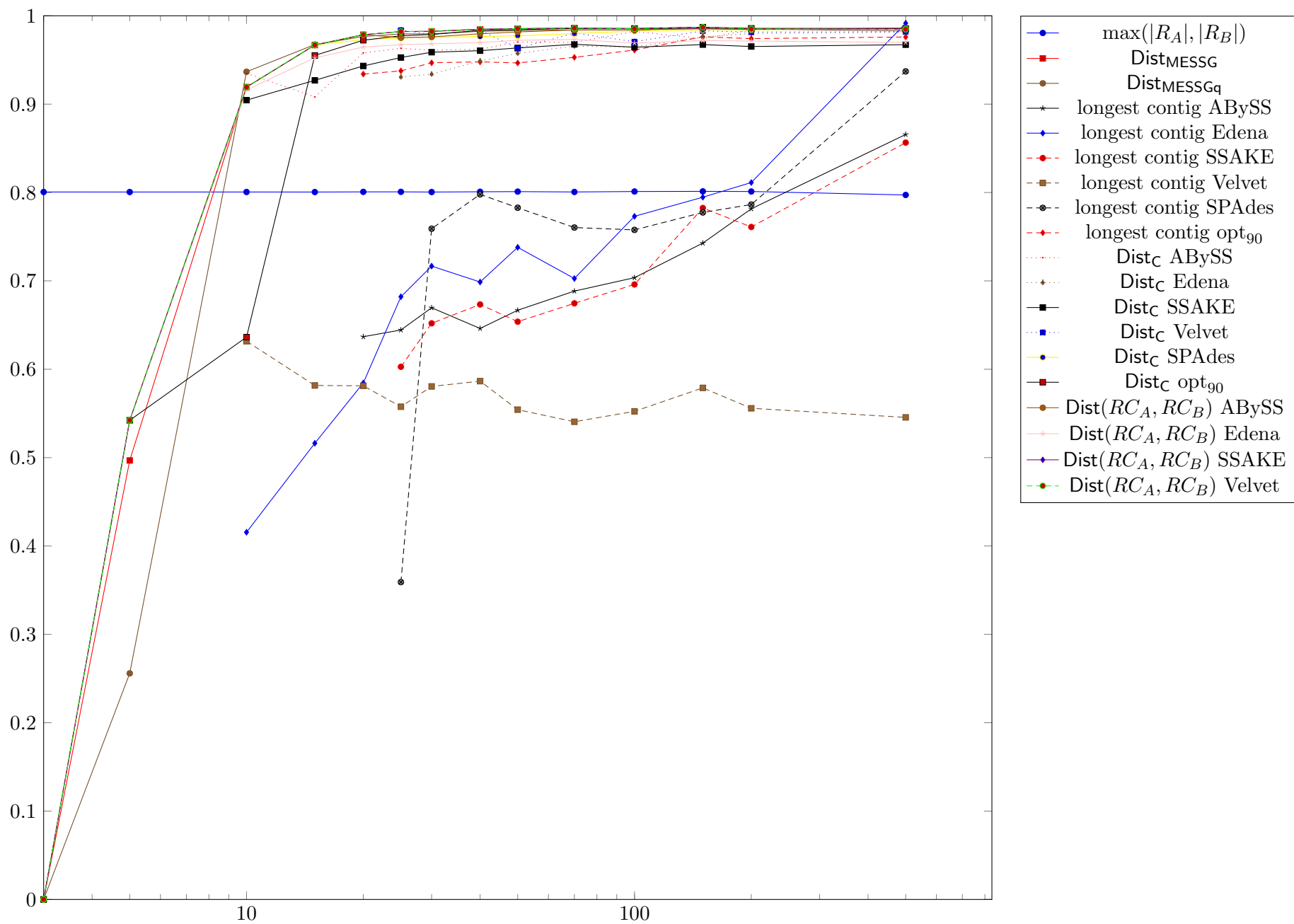
Figure 4: Plot of average Pearson's correlation coefficient for several choices of read length.

The legend entries, top to bottom:
- $\max(|R_A|, |R_B|)$
- $\text{Dist}_{\text{MESSG}}$
- $\text{Dist}_{\text{MESSGq}}$
- longest contig ABySS
- longest contig Edena
- longest contig SSAKE
- longest contig Velvet
- longest contig SPAdes
- longest contig opt_{90}
- Dist_C ABySS
- Dist_C Edena
- Dist_C SSAKE
- Dist_C Velvet
- Dist_C SPAdes
- Dist_C opt_{90}
- $\text{Dist}(RC_A, RC_B)$ ABySS
- $\text{Dist}(RC_A, RC_B)$ Edena
- $\text{Dist}(RC_A, RC_B)$ SSAKE
- $\text{Dist}(RC_A, RC_B)$ Velvet