Table 1. Number Of Pairs Discovered. Each integer denotes the number of pairs for each setting. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level). Without confidence pruning (confidence level = 0.0), the longer associated pattern pairs are the supersets of the shorter pairs according to the downward closure property. For instance, if the pair AAACA-HNLSL is found in the set of TFBS5mer-TF5mer pairs, then the two sub-pairs AAACA-HNLS and AAACA-NLSL must also be found in the set of TFBS5mer-TF4mer pairs. Note that such property is only applied when there is no confidence pruning. In other words, it is only applied to the second row. With confidence pruning (confidence level > 0.0), it can be observed that the degree of overlapping between different sets of pairs is decreased when the confidence level is increased. For instance, at the confidence level=0.1, 86.3% of the TFBS5mer-TF5mer pairs have all of their sub-pairs in the set of TFBS4mer-TF4mer pairs. At the confidence level=0.3, only 11.1% of the TFBS5mer-TF5mer pairs have all of their sub-pairs in the set of TFBS4mer-TF4mer pairs. Besides, it is interesting that the degree of overlapping is significantly decreased when TFBS k-mers are elongated. For example, at the confidence level=0.1, 92.6% of the TFBS4mer-TF5mer pairs have all of their sub-pairs in the set of TFBS4mer-TF4mer pairs, whereas only 9.09% of the TFBS5mer-TF4mer pairs have all of their sub-pairs in the set of TFBS4mer-TF4mer pairs.

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0	4229	1334	871	420	262	200	100	74	59
0.1	3407	1155	761	414	262	200	100	74	59
0.2	710	435	294	354	240	184	95	74	59
0.3	260	163	115	264	180	134	82	72	59
0.4	28	15	4	170	126	92	52	44	36
0.5	2	0	0	94	70	48	22	20	16
0.6	0	0	0	34	24	10	0	12	14
0.7	0	0	0	6	6	2	0	0	0
0.8	0	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0

Table 2. Number Of Pairs (duplicated pairs removed). Each integer denotes the number of pairs for each setting. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level). Since some DNA k-mers are the reverse complements of themselves, the number of pairs are not necessary to be half of the corresponding number in Table 1

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	2153	689	452	210	131	100	56	42	34
0.1	1739	599	397	207	131	100	56	42	34
0.2	374	232	159	177	120	92	53	42	34
0.3	141	91	65	132	90	67	46	41	34
0.4	15	8	2	85	63	46	26	22	18
0.5	1	0	0	47	35	24	11	10	8
0.6	0	0	0	17	12	5	0	6	7
0.7	0	0	0	3	3	1	0	0	0
0.8	0	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0	0
1.0	0	0	0	0	0	0	0	0	0

Table 3. Number Of Merged Pairs. Each integer denotes the number of merged pairs for each setting. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level). The number is significantly reduced as compared to that in Table 2.

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0	1002	192	110	69	29	21	14	8	4
0.1	803	171	96	67	29	21	14	8	4
0.2	137	69	44	50	24	17	14	8	4
0.3	51	23	19	40	23	14	14	8	4
0.4	6	5	2	33	21	12	7	4	2
0.5	1	0	0	18	11	7	3	3	2
0.6	0	0	0	8	8	3	0	2	2
0.7	0	0	0	2	2	1	0	0	0
0.8	0	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0

Table 4. Mean and SD of Support. Each block denotes the mean and SD of support of the pairs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	8.53±2.65	8.97±3.20	8.85±2.98	9.75±3.67	9.88±3.68	9.62±3.35	9.80±2.79	9.64±2.81	9.39±2.55
0.1	8.75±2.87	9.23±3.36	9.08±3.11	9.79±3.68	9.88±3.68	9.62±3.35	9.80±2.79	9.64±2.81	9.39±2.55
0.2	10.71±4.80	10.63±4.65	10.25±4.23	10.23±3.81	10.14±3.73	9.85±3.40	9.95±2.78	9.64±2.81	9.39±2.55
0.3	11.80±6.41	11.71±6.26	11.15±5.62	10.82±4.16	10.63±4.11	10.25±3.80	10.28±2.82	9.71±2.81	9.39±2.55
0.4	22.96±8.57	23.67±7.76	24.50±10.50	11.47±4.81	11.40±4.59	11.09±4.24	11.38±2.77	11.18±2.66	10.78±2.37
0.5	40.00±0.00	N/A	N/A	13.55±5.39	13.63±5.05	13.63±4.48	13.64±2.23	13.50±1.69	13.38±0.48
0.6	N/A	N/A	N/A	16.00±5.62	15.08±5.28	13.40±3.56	N/A	13.50±0.76	13.29±0.45
0.7	N/A	N/A	N/A	8.67±2.36	10.33±2.36	12.00±0.00	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 5. Mean and SD of ϕ-coefficient. Each block denotes the mean and SD of ϕ-coefficient of the pairs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.18±0.14	0.29±0.14	0.32±0.13	0.44±0.14	0.49±0.11	0.50±0.09	0.46±0.11	0.50±0.10	0.51±0.09
0.1	0.21±0.14	0.31±0.14	0.33±0.13	0.45±0.13	0.49±0.11	0.50±0.09	0.46±0.11	0.50±0.10	0.51±0.09
0.2	0.38±0.12	0.41±0.12	0.43±0.11	0.48±0.11	0.51±0.11	0.50±0.09	0.47±0.10	0.50±0.10	0.51±0.09
0.3	0.48±0.09	0.51±0.08	0.52±0.07	0.52±0.10	0.54±0.10	0.53±0.09	0.49±0.10	0.51±0.10	0.51±0.09
0.4	0.57±0.03	0.58±0.04	0.58±0.04	0.57±0.09	0.58±0.09	0.57±0.08	0.53±0.08	0.56±0.08	0.55±0.07
0.5	0.57±0.00	N/A	N/A	0.62±0.08	0.64±0.07	0.63±0.05	0.60±0.04	0.63±0.03	0.62±0.02
0.6	N/A	N/A	N/A	0.70±0.06	0.71±0.06	0.69±0.04	N/A	0.62±0.02	0.62±0.02
0.7	N/A	N/A	N/A	0.80±0.04	0.79±0.03	0.77±0.00	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 6. Mean and SD of Lift. Each block denotes the mean and SD of lift of the pairs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	5.09±4.80	8.67±6.13	9.56±6.16	15.29±8.23	17.92±7.34	18.14±5.80	14.98±4.45	17.96±3.60	18.54±2.96
0.1	5.71±5.09	9.28±6.31	10.13±6.37	15.50±8.11	17.92±7.34	18.14±5.80	14.98±4.45	17.96±3.60	18.54±2.96
0.2	11.71±6.96	13.49±7.64	14.37±7.66	16.60±8.11	18.32±7.46	18.43±5.87	15.42±4.07	17.96±3.60	18.54±2.96
0.3	16.25±7.55	18.67±8.12	19.40±7.92	18.33±8.53	19.81±7.79	19.72±6.04	15.89±3.90	18.00±3.65	18.54±2.96
0.4	11.56±5.37	10.96±4.22	11.97±5.84	20.97±9.30	21.41±8.53	20.87±6.67	16.67±3.31	18.75±3.28	19.12±2.39
0.5	5.94±0.00	N/A	N/A	21.97±11.37	22.57±10.46	21.41±8.05	18.15±3.50	19.37±3.23	18.95±0.66
0.6	N/A	N/A	N/A	24.62±15.14	25.80±13.76	25.49±8.90	N/A	18.75±0.63	19.02±0.67
0.7	N/A	N/A	N/A	51.93±15.80	42.07±14.87	31.55±0.00	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 7. Mean and SD of Forward Conviction. Each block denotes the mean and SD of forward conviction of the pairs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	1.15±0.14	1.21±0.16	1.21±0.15	1.89±0.71	1.89±0.67	1.77±0.53	2.39±1.15	2.33±1.02	2.12±0.58
0.1	1.17±0.14	1.23±0.16	1.23±0.15	1.90±0.71	1.89±0.67	1.77±0.53	2.39±1.15	2.33±1.02	2.12±0.58
0.2	1.38±0.16	1.38±0.16	1.38±0.16	1.97±0.70	1.94±0.68	1.81±0.53	2.42±1.17	2.33±1.02	2.12±0.58
0.3	1.52±0.11	1.53±0.09	1.52±0.09	2.08±0.68	2.02±0.64	1.90±0.47	2.44±1.25	2.32±1.03	2.12±0.58
0.4	1.76±0.10	1.75±0.06	1.78±0.01	2.33±0.72	2.23±0.66	2.06±0.47	2.53±1.25	2.40±1.05	2.18±0.59
0.5	2.04±0.00	N/A	N/A	2.58±0.73	2.49±0.70	2.26±0.51	3.33±1.49	3.08±1.16	2.79±0.33
0.6	N/A	N/A	N/A	3.34±0.62	3.33±0.57	3.17±0.40	N/A	2.79±0.36	2.73±0.31
0.7	N/A	N/A	N/A	3.50±0.28	3.70±0.29	3.90±0.00	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 8. Mean and SD of Backward Conviction. Each block denotes the mean and SD of backward conviction of the pairs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	2.04±1.73	3.44±2.62	4.01±2.70	2.61±1.61	3.50±2.29	3.71±2.15	1.96±0.86	2.46±1.55	2.45±0.83
0.1	2.14±1.83	3.52±2.69	4.07±2.79	2.64±1.61	3.50±2.29	3.71±2.15	1.96±0.86	2.46±1.55	2.45±0.83
0.2	3.17±2.33	4.02±3.05	4.30±3.05	2.72±1.64	3.51±2.30	3.70±2.14	2.00±0.86	2.46±1.55	2.45±0.83
0.3	4.28±2.55	5.99±3.15	6.44±3.14	2.80±1.58	3.46±2.31	3.44±2.02	2.10±0.88	2.49±1.57	2.45±0.83
0.4	4.90±1.99	5.87±3.17	5.28±2.33	2.87±1.30	3.61±2.40	3.68±2.35	2.29±0.89	2.90±1.79	2.74±0.64
0.5	2.75±0.00	N/A	N/A	3.13±1.25	4.35±2.65	4.85±2.72	2.64±1.17	3.25±2.47	2.54±0.10
0.6	N/A	N/A	N/A	3.71±1.63	4.21±2.55	3.50±0.76	N/A	2.54±0.07	2.57±0.06
0.7	N/A	N/A	N/A	3.90±0.00	4.87±0.00	4.87±0.00	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 9. Mean and SD of TFBS Prediction Score. Each block denotes the mean and SD of TFBS Prediction Score of the pairs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.29±0.35	0.49±0.40	0.52±0.42	0.36±0.36	0.41±0.41	0.40±0.43	0.21±0.33	0.18±0.33	0.13±0.30
0.1	0.31±0.36	0.47±0.41	0.50±0.42	0.37±0.36	0.41±0.41	0.40±0.43	0.21±0.33	0.18±0.33	0.13±0.30
0.2	0.40±0.36	0.46±0.41	0.44±0.43	0.39±0.36	0.40±0.41	0.38±0.42	0.22±0.34	0.18±0.33	0.13±0.30
0.3	0.45±0.39	0.49±0.43	0.50±0.44	0.41±0.35	0.38±0.39	0.36±0.40	0.24±0.35	0.18±0.34	0.13±0.30
0.4	0.56±0.37	0.63±0.38	0.88±0.13	0.45±0.35	0.38±0.40	0.33±0.40	0.38±0.39	0.33±0.40	0.25±0.37
0.5	0.73±0.00	N/A	N/A	0.39±0.37	0.42±0.40	0.42±0.42	0.34±0.40	0.34±0.42	0.31±0.41
0.6	N/A	N/A	N/A	0.53±0.38	0.62±0.37	0.83±0.09	N/A	0.14±0.31	0.24±0.38
0.7	N/A	N/A	N/A	0.30±0.43	0.63±0.45	0.95±0.00	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 10. Mean and SD of TFBS Binding Prediction Score. Each block denotes the mean and SD of TFBS Binding Prediction Score of the pairs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.06±0.16	0.20±0.29	0.26±0.32	0.15±0.23	0.24±0.30	0.25±0.33	0.11±0.20	0.12±0.24	0.12±0.26
0.1	0.07±0.16	0.20±0.28	0.24±0.31	0.15±0.23	0.24±0.30	0.25±0.33	0.11±0.20	0.12±0.24	0.12±0.26
0.2	0.14±0.23	0.22±0.29	0.23±0.32	0.15±0.21	0.23±0.29	0.23±0.32	0.12±0.21	0.12±0.24	0.12±0.26
0.3	0.16±0.25	0.22±0.31	0.25±0.34	0.18±0.23	0.25±0.29	0.27±0.34	0.13±0.22	0.12±0.24	0.12±0.26
0.4	0.38±0.27	0.47±0.34	0.75±0.25	0.21±0.25	0.25±0.31	0.24±0.33	0.22±0.25	0.23±0.29	0.23±0.32
0.5	0.55±0.00	N/A	N/A	0.23±0.27	0.30±0.33	0.33±0.37	0.23±0.28	0.27±0.33	0.27±0.35
0.6	N/A	N/A	N/A	0.39±0.29	0.50±0.31	0.68±0.11	N/A	0.13±0.28	0.22±0.35
0.7	N/A	N/A	N/A	0.24±0.34	0.50±0.35	0.75±0.00	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 11. Mean and SD of Binding Prediction Score. Each block denotes the mean and SD of Binding Prediction Score of the pairs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.13±0.28	0.30±0.37	0.35±0.39	0.28±0.36	0.36±0.39	0.36±0.41	0.16±0.27	0.15±0.29	0.16±0.34
0.1	0.14±0.29	0.31±0.37	0.34±0.39	0.29±0.36	0.36±0.39	0.36±0.41	0.16±0.27	0.15±0.29	0.16±0.34
0.2	0.26±0.34	0.31±0.37	0.32±0.38	0.30±0.36	0.35±0.39	0.35±0.41	0.16±0.28	0.15±0.29	0.16±0.34
0.3	0.26±0.35	0.29±0.37	0.33±0.39	0.31±0.35	0.35±0.39	0.37±0.42	0.17±0.28	0.15±0.30	0.16±0.34
0.4	0.50±0.33	0.56±0.36	0.83±0.17	0.34±0.36	0.34±0.40	0.31±0.40	0.30±0.32	0.28±0.36	0.30±0.42
0.5	0.75±0.00	N/A	N/A	0.33±0.38	0.39±0.42	0.39±0.43	0.30±0.34	0.31±0.40	0.32±0.42
0.6	N/A	N/A	N/A	0.51±0.36	0.61±0.37	0.82±0.08	N/A	0.15±0.34	0.26±0.42
0.7	N/A	N/A	N/A	0.26±0.37	0.53±0.37	0.79±0.00	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 12. Mean and SD of TFBS Prediction Score (Merged Pairs). Each block denotes the mean and SD of TFBS Prediction Score of the merged pairs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.20±0.34	0.38±0.40	0.44±0.42	0.34±0.38	0.47±0.43	0.50±0.45	0.29±0.35	0.34±0.40	0.14±0.24
0.1	0.22±0.35	0.40±0.40	0.41±0.42	0.35±0.38	0.47±0.43	0.50±0.45	0.29±0.35	0.34±0.40	0.14±0.24
0.2	0.38±0.38	0.47±0.40	0.46±0.42	0.35±0.35	0.39±0.40	0.43±0.43	0.29±0.35	0.34±0.40	0.14±0.24
0.3	0.39±0.39	0.40±0.41	0.39±0.42	0.41±0.35	0.43±0.40	0.38±0.41	0.27±0.36	0.34±0.40	0.14±0.24
0.4	0.72±0.33	0.67±0.35	0.88±0.13	0.42±0.34	0.36±0.37	0.33±0.38	0.44±0.37	0.61±0.39	0.27±0.27
0.5	0.73±0.00	N/A	N/A	0.36±0.30	0.40±0.34	0.44±0.39	0.60±0.43	0.60±0.43	0.42±0.42
0.6	N/A	N/A	N/A	0.45±0.40	0.51±0.40	0.82±0.12	N/A	0.42±0.42	0.42±0.42
0.7	N/A	N/A	N/A	0.45±0.45	0.48±0.48	0.95±0.00	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 13. Mean and SD of TFBS Binding Prediction Score (Merged Pairs). Each block denotes the mean and SD of TFBS Binding Prediction Score of the merged pairs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.04±0.15	0.17±0.29	0.24±0.34	0.17±0.30	0.31±0.37	0.37±0.39	0.14±0.20	0.18±0.24	0.14±0.24
0.1	0.05±0.16	0.19±0.30	0.24±0.34	0.18±0.30	0.31±0.37	0.37±0.39	0.14±0.20	0.18±0.24	0.14±0.24
0.2	0.16±0.26	0.24±0.31	0.24±0.33	0.17±0.26	0.25±0.32	0.29±0.35	0.14±0.20	0.18±0.24	0.14±0.24
0.3	0.19±0.30	0.24±0.35	0.26±0.38	0.22±0.27	0.31±0.34	0.34±0.39	0.13±0.20	0.18±0.24	0.14±0.24
0.4	0.54±0.32	0.52±0.35	0.75±0.25	0.23±0.26	0.26±0.28	0.26±0.32	0.25±0.23	0.37±0.22	0.27±0.27
0.5	0.55±0.00	N/A	N/A	0.21±0.28	0.30±0.31	0.30±0.36	0.44±0.33	0.44±0.33	0.39±0.39
0.6	N/A	N/A	N/A	0.40±0.36	0.44±0.35	0.68±0.13	N/A	0.38±0.38	0.39±0.39
0.7	N/A	N/A	N/A	0.36±0.36	0.38±0.38	0.75±0.00	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 14. Mean and SD of Binding Prediction Score (Merged Pairs). Each block denotes the mean and SD of Binding Prediction Score of the merged pairs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.07±0.22	0.25±0.38	0.33±0.41	0.26±0.38	0.44±0.44	0.46±0.43	0.25±0.34	0.25±0.35	0.25±0.43
0.1	0.08±0.24	0.27±0.38	0.34±0.41	0.27±0.39	0.44±0.44	0.46±0.43	0.25±0.34	0.25±0.35	0.25±0.43
0.2	0.25±0.37	0.33±0.39	0.32±0.40	0.32±0.39	0.40±0.42	0.42±0.43	0.25±0.34	0.25±0.35	0.25±0.43
0.3	0.30±0.40	0.35±0.42	0.40±0.44	0.42±0.41	0.47±0.45	0.49±0.45	0.19±0.31	0.25±0.35	0.25±0.43
0.4	0.61±0.34	0.61±0.37	0.83±0.17	0.39±0.39	0.41±0.43	0.39±0.42	0.37±0.36	0.50±0.36	0.50±0.50
0.5	0.75±0.00	N/A	N/A	0.35±0.41	0.46±0.44	0.37±0.43	0.49±0.38	0.49±0.38	0.47±0.47
0.6	N/A	N/A	N/A	0.54±0.42	0.54±0.42	0.82±0.08	N/A	0.45±0.45	0.47±0.47
0.7	N/A	N/A	N/A	0.39±0.39	0.39±0.39	0.79±0.00	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 15. Number of Pairs with at least one related PDB chain. Each integer denotes the number of pairs with at least one related PDB chain for each setting. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level)

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	1347	500	323	151	80	55	19	10	6
0.1	1140	427	274	150	80	55	19	10	6
0.2	283	162	101	131	71	48	18	10	6
0.3	102	62	44	95	50	33	16	10	6
0.4	11	6	2	62	32	20	14	9	6
0.5	1	N/A	N/A	29	19	12	5	4	3
0.6	N/A	N/A	N/A	12	9	5	N/A	1	2
0.7	N/A	N/A	N/A	1	2	1	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 16. Number of Pairs with at least one related PDB chain (Merged Pairs). Each integer denotes the number of merged pairs with at least one related PDB chain for each setting. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level)

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	379	112	69	46	19	14	7	4	1
0.1	337	106	58	46	19	14	7	4	1
0.2	90	49	30	35	15	10	7	4	1
0.3	34	15	13	30	15	8	6	4	1
0.4	5	4	2	24	12	6	5	3	1
0.5	1	N/A	N/A	14	7	4	2	2	1
0.6	N/A	N/A	N/A	5	5	3	N/A	1	1
0.7	N/A	N/A	N/A	1	1	1	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 17. Fraction of Verified Pairs. Each fraction denotes the fraction of verified pairs for each setting. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level)

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.3942094	0.684	0.7585139	0.6225166	0.8125	0.8363636	0.7894737	0.9	1
0.1	0.4184211	0.7189696	0.7810219	0.6266667	0.8125	0.8363636	0.7894737	0.9	1
0.2	0.6466431	0.7530864	0.7920792	0.6641221	0.8309859	0.8333333	0.8333333	0.9	1
0.3	0.6568627	0.7258065	0.75	0.7263158	0.88	0.9393939	0.875	0.9	1
0.4	1	1	1	0.7419355	0.875	0.9	0.9285714	1	1
0.5	1	N/A	N/A	0.7586207	0.8947368	0.9166667	1	1	1
0.6	N/A	N/A	N/A	1	1	1	N/A	1	1
0.7	N/A	N/A	N/A	1	1	1	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 18. Fraction of Verified Pairs (Merged Pairs). Each fraction denotes the fraction of merged pairs verified for each setting. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level)

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.3271768	0.6696429	0.7391304	0.5652174	0.8421053	0.8571429	0.8571429	0.75	1
0.1	0.3560831	0.6981132	0.7931034	0.5652174	0.8421053	0.8571429	0.8571429	0.75	1
0.2	0.6444444	0.7755102	0.7333333	0.6857143	0.8666667	0.9	0.8571429	0.75	1
0.3	0.6470588	0.8	0.7692308	0.8	0.8666667	1	0.8333333	0.75	1
0.4	1	1	1	0.7916667	0.9166667	1	0.8	1	1
0.5	1	N/A	N/A	0.5714286	0.8571429	0.75	1	1	1
0.6	N/A	N/A	N/A	1	1	1	N/A	1	1
0.7	N/A	N/A	N/A	1	1	1	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 19. Maximal TFBS Prediction Score over 50 runs. Each block denotes the maximal performance of TFBS Prediction Score of the pairs over 50 random runs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFzBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.14469±0.3095	0.09185±0.25819	0.04296±0.17842	0.05925±0.2266	0.03767±0.18401	0.02667±0.15762	0.04292±0.19751	0.04678±0.19528	0.01695±0.12908
0.1	0.15028±0.3172	0.09067±0.25405	0.04943±0.19171	0.0558±0.22102	0.04018±0.19093	0.02375±0.14916	0.05356±0.20521	0.02815±0.16226	0.02018±0.13098
0.2	0.16204±0.32886	0.10662±0.27429	0.06548±0.22314	0.06866±0.23892	0.03588±0.1824	0.02222±0.14515	0.03775±0.18377	0.02703±0.16216	0.01695±0.12908
0.3	0.17702±0.34032	0.10803±0.28737	0.07302±0.23572	0.08541±0.26954	0.04931±0.2151	0.02456±0.14834	0.05358±0.20581	0.02381±0.14286	0.02119±0.13251
0.4	0.28267±0.38034	0.25946±0.37558	0.25±0.43301	0.06072±0.22946	0.06365±0.23856	0.02998±0.16459	0.05769±0.23316	0.03409±0.16507	0.02778±0.16434
0.5	0.87037±0.12963	N/A	N/A	0.078±0.25463	0.05713±0.20329	0.04167±0.19983	0.06818±0.22841	0.05±0.21794	0.0625±0.24206
0.6	N/A	N/A	N/A	0.14163±0.30753	0.06483±0.22311	0.1±0.3	N/A	0.08333±0.27639	0.07143±0.25754
0.7	N/A	N/A	N/A	0.23333±0.37268	0.16667±0.37268	0±0	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 20. Maximal TFBS Binding Prediction Score over 50 runs. Each block denotes the maximal performance of TFBS Binding Prediction Score of the pairs over 50 random runs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.00332±0.04768	0.00222±0.03928	0.00115±0.03386	0.00516±0.06929	0.00382±0.06166	0.005±0.07053	0.01549±0.11083	0.01351±0.11546	0.01695±0.12908
0.1	0.00321±0.04947	0.00247±0.04357	0.00131±0.03623	0.00631±0.07446	0.00763±0.08704	0±0	0.01634±0.11728	0.0045±0.03849	0±0
0.2	0.00496±0.05402	0.00374±0.04772	0.00612±0.07444	0.00857±0.09168	0.00729±0.08036	0.00543±0.07352	0.02456±0.14703	0.01351±0.11546	0.01695±0.12908
0.3	0.00751±0.07192	0.00241±0.02201	0.005±0.04689	0.00821±0.08726	0.00556±0.07433	0±0	0.01829±0.1221	0.01389±0.11703	0.00254±0.01936
0.4	0.03571±0.18558	0.00667±0.02494	0±0	0.00588±0.07647	0.00794±0.08873	0.00036±0.00342	0.02404±0.14091	0.01136±0.07452	0.00185±0.01096
0.5	0.5±0.5	N/A	N/A	0.01064±0.10259	0.01429±0.11867	0.0026±0.01785	0.02273±0.10415	0.03333±0.1453	0±0
0.6	N/A	N/A	N/A	0.02941±0.16896	0.02083±0.09991	0±0	N/A	0±0	0±0
0.7	N/A	N/A	N/A	0.11111±0.24845	0±0	0±0	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 21. Maximal Binding Prediction Score over 50 runs. Each block denotes the maximal performance of Binding Prediction Score of the pairs over 50 random runs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.0073±0.0802	0.00354±0.05666	0.00237±0.04362	0.00774±0.07773	0.00509±0.06492	0.005±0.07053	0.025±0.1479	0.02703±0.16216	0.01695±0.12908
0.1	0.00741±0.07814	0.00394±0.06054	0.00197±0.04048	0.00966±0.09144	0.00763±0.08704	0±0	0.01812±0.12757	0.01351±0.11546	0±0
0.2	0.01156±0.09813	0.00736±0.07301	0.0068±0.0822	0.0096±0.09399	0.0099±0.09392	0.00543±0.07352	0.03158±0.17488	0.01351±0.11546	0.01695±0.12908
0.3	0.01517±0.11226	0.01361±0.10318	0.01304±0.10344	0.00947±0.09183	0.00926±0.0891	0±0	0.02439±0.15426	0.02083±0.1301	0.00847±0.06454
0.4	0.07143±0.25754	0.06667±0.24944	0±0	0.01471±0.1141	0.00794±0.08873	0.00047±0.00451	0.03846±0.19231	0.02273±0.14903	0.02778±0.16434
0.5	0.5±0.5	N/A	N/A	0.01463±0.10918	0.01429±0.11867	0.01042±0.07141	0.04545±0.2083	0.05±0.21794	0±0
0.6	N/A	N/A	N/A	0.03088±0.14039	0.04167±0.19983	0±0	N/A	0±0	0±0
0.7	N/A	N/A	N/A	0.11111±0.24845	0±0	0±0	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 22. Maximal TFBS Prediction Score over 50 runs (Merged Pairs). Each block denotes the maximal performance of TFBS Prediction Score of the merged pairs over 50 random runs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

Confidence	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.14387±0.3094	0.09185±0.25819	0.04301±0.17852	0.05925±0.2266	0.03767±0.18401	0.02667±0.15762	0.04292±0.19751	0.04678±0.19528	0.01695±0.12908
0.1	0.15047±0.31761	0.09067±0.25405	0.04943±0.19171	0.0558±0.22102	0.04018±0.19093	0.02375±0.14916	0.05356±0.20521	0.02815±0.16226	0.02018±0.13098
0.2	0.16227±0.32903	0.10662±0.27429	0.06548±0.22314	0.06866±0.23892	0.03588±0.1824	0.02222±0.14515	0.03775±0.18377	0.02703±0.16216	0.01695±0.12908
0.3	0.17702±0.34032	0.10803±0.28737	0.07302±0.23572	0.08541±0.26954	0.04931±0.2151	0.02456±0.14834	0.05358±0.20581	0.02381±0.14286	0.02119±0.13251
0.4	0.28267±0.38034	0.25946±0.37558	0.25±0.43301	0.06072±0.22946	0.06365±0.23856	0.02998±0.16459	0.05769±0.23316	0.03409±0.16507	0.02778±0.16434
0.5	0.87037±0.12963	N/A	N/A	0.078±0.25463	0.05713±0.20329	0.04167±0.19983	0.06818±0.22841	0.05±0.21794	0.0625±0.24206
0.6	N/A	N/A	N/A	0.14163±0.30753	0.06483±0.22311	0.1±0.3	N/A	0.08333±0.27639	0.07143±0.25754
0.7	N/A	N/A	N/A	0.23333±0.37268	0.16667±0.37268	0±0	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 23. Maximal TFBS Binding Prediction Score over 50 runs (Merged Pairs). Each block denotes the maximal performance of TFBS Binding Prediction Score of the merged pairs over 50 random runs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.00334±0.04777	0.00222±0.03928	0.00115±0.03386	0.00516±0.06929	0.00382±0.06166	0.005±0.07053	0.01564±0.11138	0.01351±0.11546	0.01695±0.12908
0.1	0.00322±0.04956	0.00247±0.04357	0.00131±0.03623	0.00631±0.07446	0.00763±0.08704	0±0	0.01634±0.11728	0.0045±0.03849	0±0
0.2	0.00496±0.05402	0.00374±0.04772	0.00612±0.07444	0.00857±0.09168	0.00729±0.08036	0.00543±0.07352	0.02456±0.14703	0.01351±0.11546	0.01695±0.12908
0.3	0.00751±0.07192	0.00241±0.02201	0.005±0.04689	0.00821±0.08726	0.00556±0.07433	0±0	0.01829±0.1221	0.01389±0.11703	0.00254±0.01936
0.4	0.03571±0.18558	0.00667±0.02494	0±0	0.00588±0.07647	0.00794±0.08873	0.00036±0.00342	0.02404±0.14091	0.01136±0.07452	0.00185±0.01096
0.5	0.5±0.5	N/A	N/A	0.01064±0.10259	0.01429±0.11867	0.0026±0.01785	0.02273±0.10415	0.03333±0.1453	0±0
0.6	N/A	N/A	N/A	0.02941±0.16896	0.02083±0.09991	0±0	N/A	0±0	0±0
0.7	N/A	N/A	N/A	0.11111±0.24845	0±0	0±0	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 24. Maximal Binding Prediction Score over 50 runs (Merged Pairs). Each block denotes the maximal performance of Binding Prediction Score of the merged pairs over 50 random runs under different settings. The first row denotes the different settings for the value of k, whereas the first column denotes the minimal confidence (confidence level).

	TFBS4mer-TF4mer	TFBS4mer-TF5mer	TFBS4mer-TF6mer	TFBS5mer-TF4mer	TFBS5mer-TF5mer	TFBS5mer-TF6mer	TFBS6mer-TF4mer	TFBS6mer-TF5mer	TFBS6mer-TF6mer
0.0	0.00709±0.07892	0.00354±0.05666	0.00237±0.04365	0.00774±0.07773	0.00509±0.06492	0.005±0.07053	0.02525±0.14863	0.02703±0.16216	0.01695±0.12908
0.1	0.00743±0.07824	0.00394±0.06054	0.00197±0.04048	0.00969±0.09155	0.00763±0.08704	0±0	0.01812±0.12757	0.01351±0.11546	0±0
0.2	0.01156±0.09813	0.00736±0.07301	0.0068±0.0822	0.0096±0.09399	0.0099±0.09392	0.00543±0.07352	0.03158±0.17488	0.01351±0.11546	0.01695±0.12908
0.3	0.01517±0.11226	0.01361±0.10318	0.01304±0.10344	0.00947±0.09183	0.00926±0.0891	0±0	0.02439±0.15426	0.02083±0.1301	0.00847±0.06454
0.4	0.07143±0.25754	0.06667±0.24944	0±0	0.01471±0.1141	0.00794±0.08873	0.00047±0.00451	0.03846±0.19231	0.02273±0.14903	0.02778±0.16434
0.5	0.5±0.5	N/A	N/A	0.01463±0.10918	0.01429±0.11867	0.01042±0.07141	0.04545±0.2083	0.05±0.21794	0±0
0.6	N/A	N/A	N/A	0.03088±0.14039	0.04167±0.19983	0±0	N/A	0±0	0±0
0.7	N/A	N/A	N/A	0.11111±0.24845	0±0	0±0	N/A	N/A	N/A
0.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.9	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A