

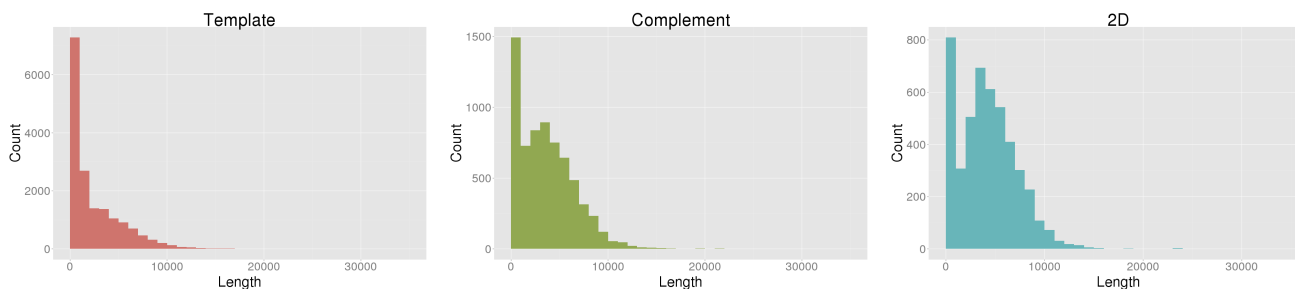
NanoOK report for ZF-screens_MARC_phase_1b_e_coli_run1

Pass and fail counts

Type	Pass	Fail
Template	0	16817
Complement	0	6679
2D	0	4673

Read lengths

Type	NumReads	TotalBases	Mean	Longest	Shortest	N50	N50Count	N90	N90Count
Template	16817	44906135	2670.28	139593	10	5406	2617	1104	8899
Complement	6679	24878534	3724.89	51769	11	5569	1586	2473	4143
2D	4673	20043200	4289.15	32601	123	5913	1246	2941	3102



Template alignments

Number of reads	16817
Number of reads with alignments	4115 (24.47%)
Number of reads without alignments	12702 (75.53%)

ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	477	2.84	2573.91	1099343	308.80	38
Escherichia coli	4641652	3638	21.63	4943.90	14983900	3.23	58

Complement alignments

Number of reads	6679
Number of reads with alignments	2783 (41.67%)
Number of reads without alignments	3896 (58.33%)

ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	286	4.28	2764.88	634218	178.15	38
Escherichia coli	4641652	2497	37.39	5398.28	11244118	2.42	48

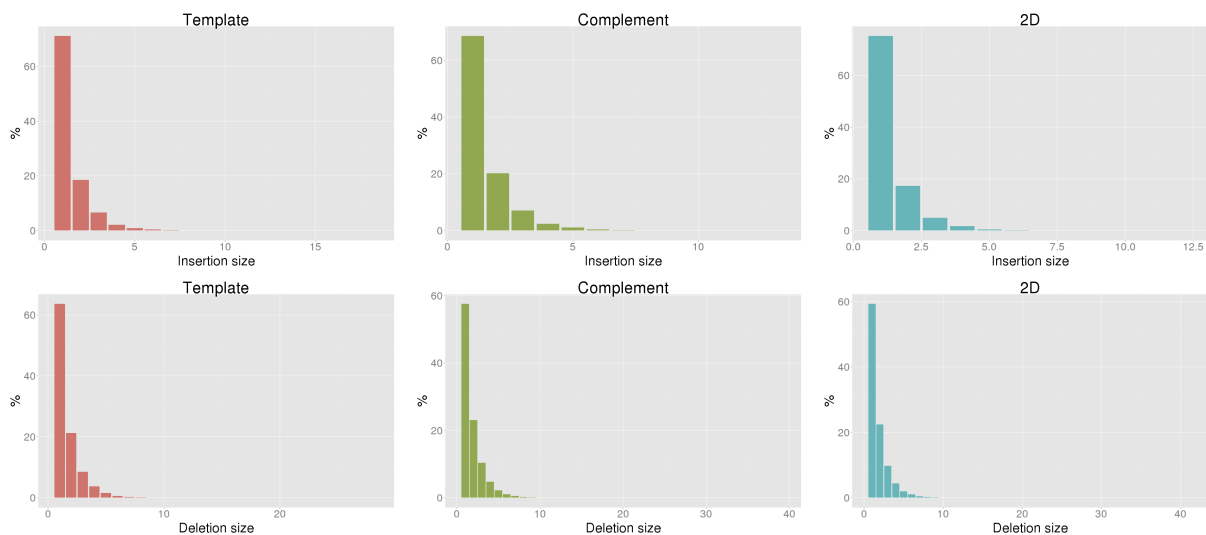
2D alignments

Number of reads	4673
Number of reads with alignments	3435 (73.51%)
Number of reads without alignments	1238 (26.49%)

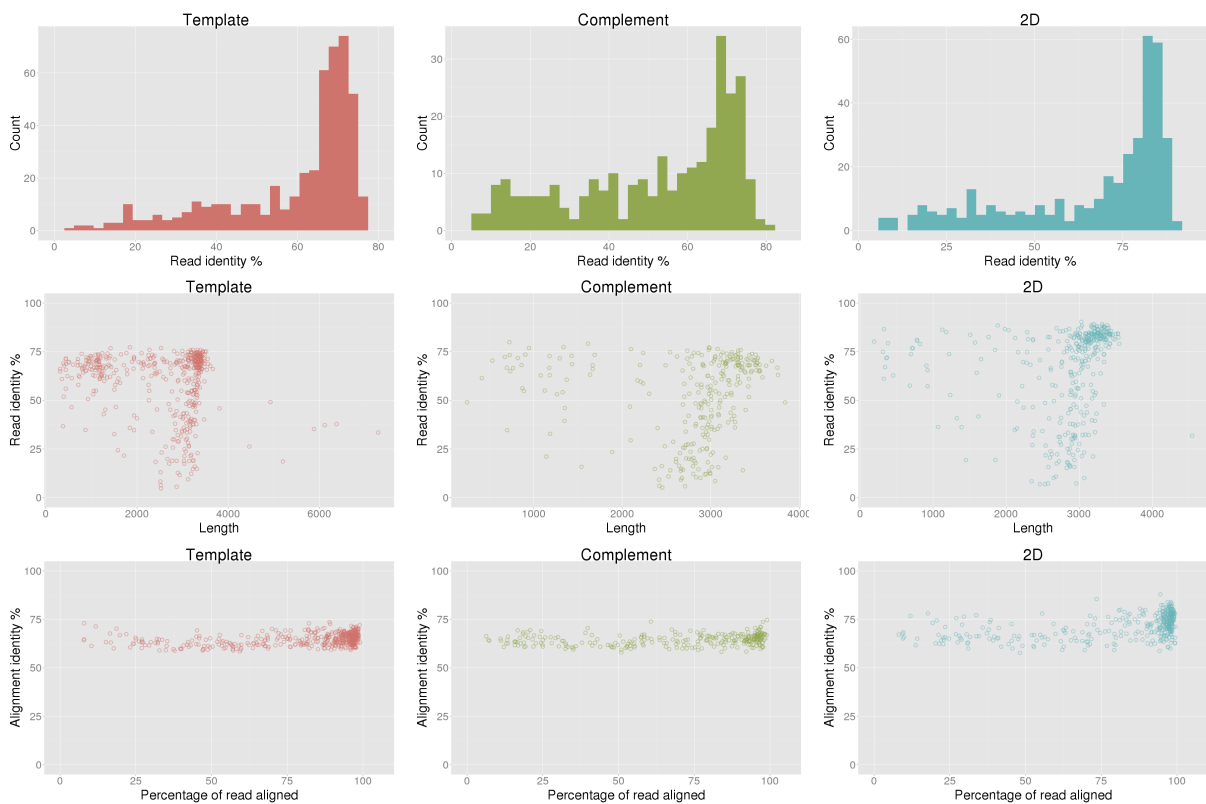
ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	379	8.11	2826.13	985598	276.85	124
Escherichia coli	4641652	3056	65.40	5179.72	14945880	3.22	140

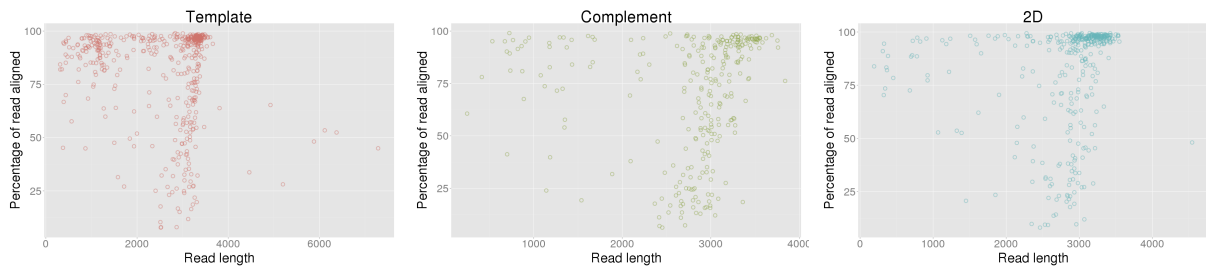
Control sequence error analysis

	Template	Complement	2D
Overall base identity (excluding indels)	57.97%	51.74%	66.95%
Aligned base identity (excluding indels)	77.20%	78.84%	85.95%
Identical bases per 100 aligned bases (including indels)	64.74%	64.51%	72.76%
Inserted bases per 100 aligned bases (including indels)	4.17%	4.09%	2.51%
Deleted bases per 100 aligned bases (including indels)	11.97%	14.09%	12.84%
Substitutions per 100 aligned bases (including indels)	19.12%	17.31%	11.90%
Mean insertion size	1.46	1.50	1.36
Mean deletion size	1.63	1.80	1.77

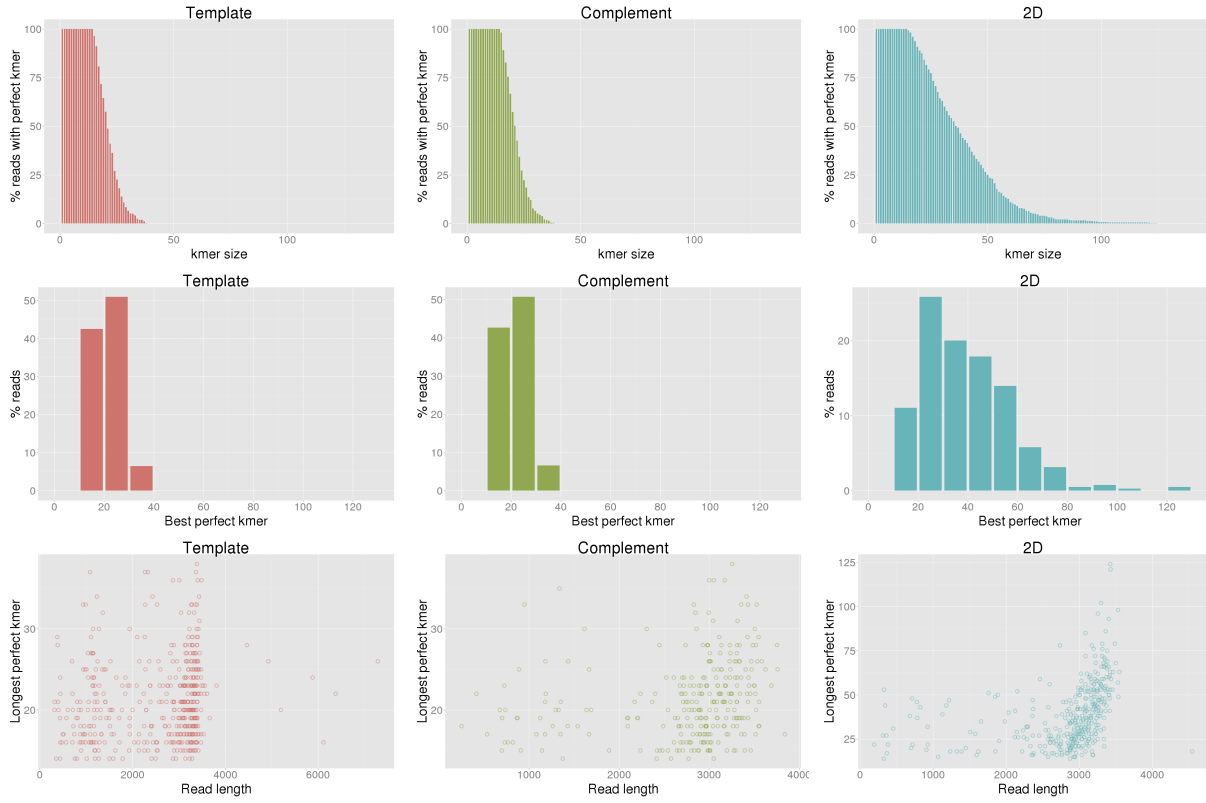


Control sequence read identity

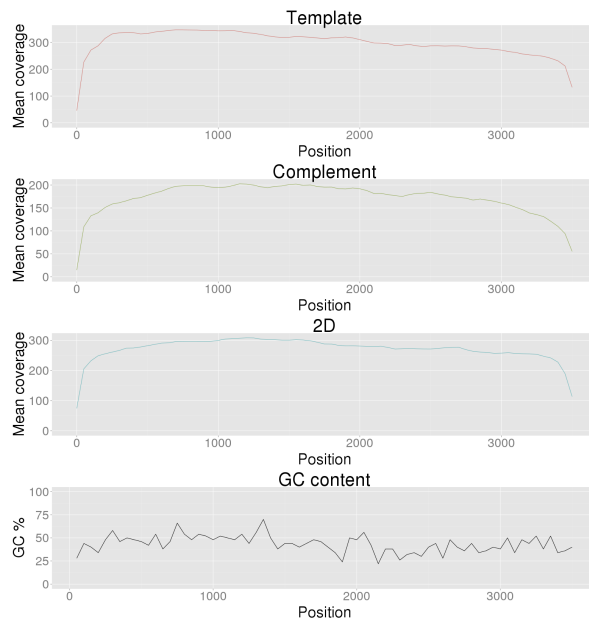




Control sequence perfect kmers



Control sequence coverage



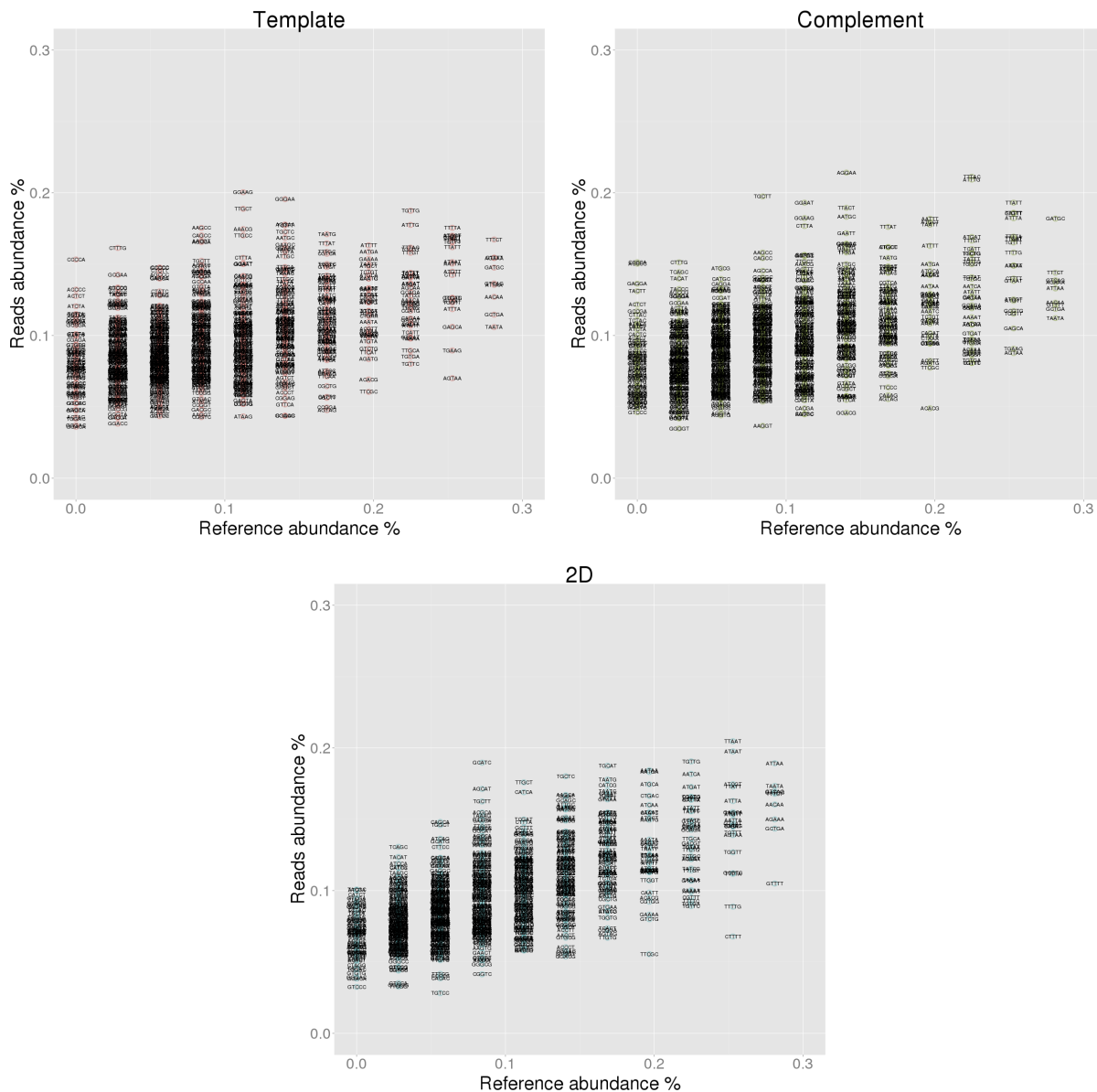
Control sequence 5-mer analysis

Under-represented 5-mers

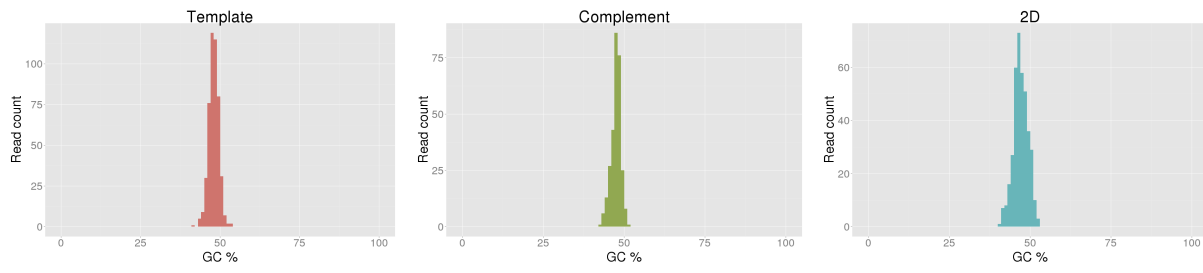
Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	TTTTT	0.759	0.122	-0.636	TTTTT	0.759	0.067	-0.691	TTTTT	0.759	0.027	-0.731
2	AAAAA	0.478	0.114	-0.364	AAAAA	0.478	0.062	-0.416	AAAAA	0.478	0.034	-0.444
3	TGATG	0.393	0.129	-0.264	AAAAC	0.337	0.120	-0.217	TGATG	0.393	0.167	-0.227
4	GATGT	0.309	0.093	-0.216	TGATG	0.393	0.180	-0.214	AAAAC	0.337	0.117	-0.220
5	AAAAC	0.337	0.132	-0.205	GATGT	0.309	0.100	-0.209	CTTTT	0.253	0.068	-0.185
6	CTGAT	0.309	0.109	-0.201	GCAAT	0.309	0.129	-0.180	GATGT	0.309	0.133	-0.176
7	AATAT	0.309	0.123	-0.186	AATAT	0.309	0.131	-0.178	GTTTT	0.281	0.105	-0.176
8	AGTAA	0.253	0.070	-0.183	TAATA	0.281	0.112	-0.169	TTATC	0.309	0.134	-0.175
9	GCAAT	0.309	0.128	-0.182	CTGAT	0.309	0.143	-0.166	GCAAT	0.309	0.140	-0.169
10	TAATA	0.281	0.106	-0.175	AGTAA	0.253	0.088	-0.165	CTGAT	0.309	0.143	-0.166

Over-represented 5-mers

Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	CCCCA	0.000	0.153	0.153	CCCCA	0.000	0.151	0.151	GCATC	0.084	0.190	0.105
2	CTTTG	0.028	0.161	0.133	ACCCC	0.000	0.150	0.150	TCAGC	0.028	0.131	0.103
3	ACCCC	0.000	0.132	0.132	GAGGA	0.000	0.136	0.136	TACGC	0.000	0.101	0.101
4	ACTCT	0.000	0.128	0.128	TACTT	0.000	0.131	0.131	ATCTA	0.000	0.100	0.100
5	ATCTA	0.000	0.120	0.120	CTTTG	0.028	0.151	0.123	CATCT	0.000	0.097	0.097
6	TCTTA	0.000	0.115	0.115	ACTCT	0.000	0.122	0.122	TACAT	0.028	0.124	0.096
7	GCTCC	0.000	0.115	0.115	GCCGA	0.000	0.117	0.117	GTCGA	0.000	0.095	0.095
8	GCGAA	0.028	0.143	0.114	TCAGC	0.028	0.144	0.116	TTAGA	0.000	0.094	0.094
9	TCTAC	0.000	0.114	0.114	CTTAC	0.000	0.115	0.115	CAGCA	0.056	0.148	0.092
10	CCCGC	0.000	0.112	0.112	TGCTT	0.084	0.198	0.113	CCCGC	0.000	0.091	0.091

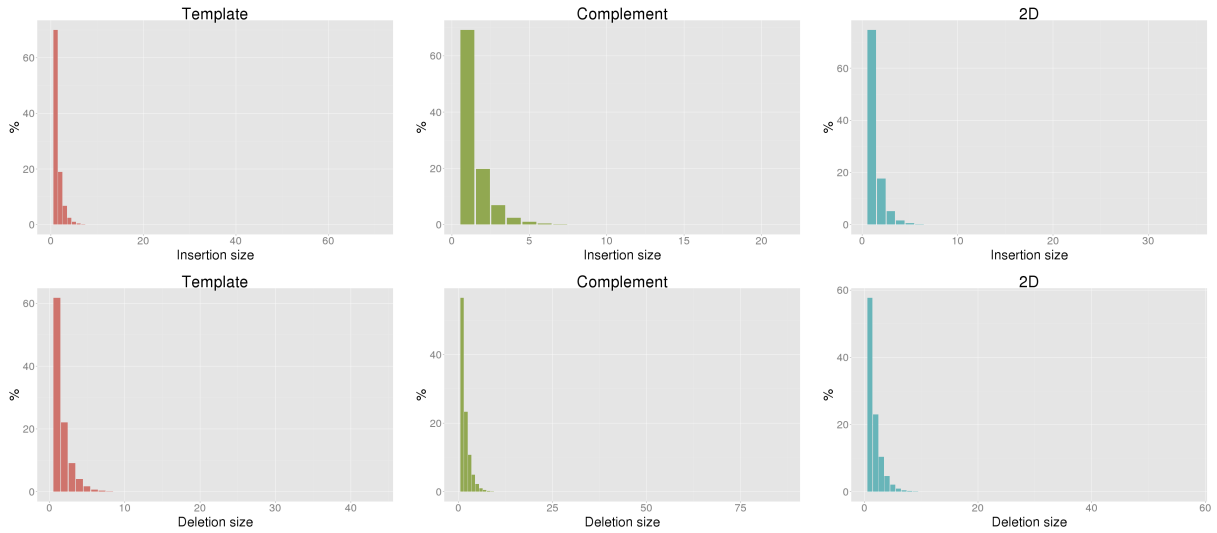


Control sequence GC content

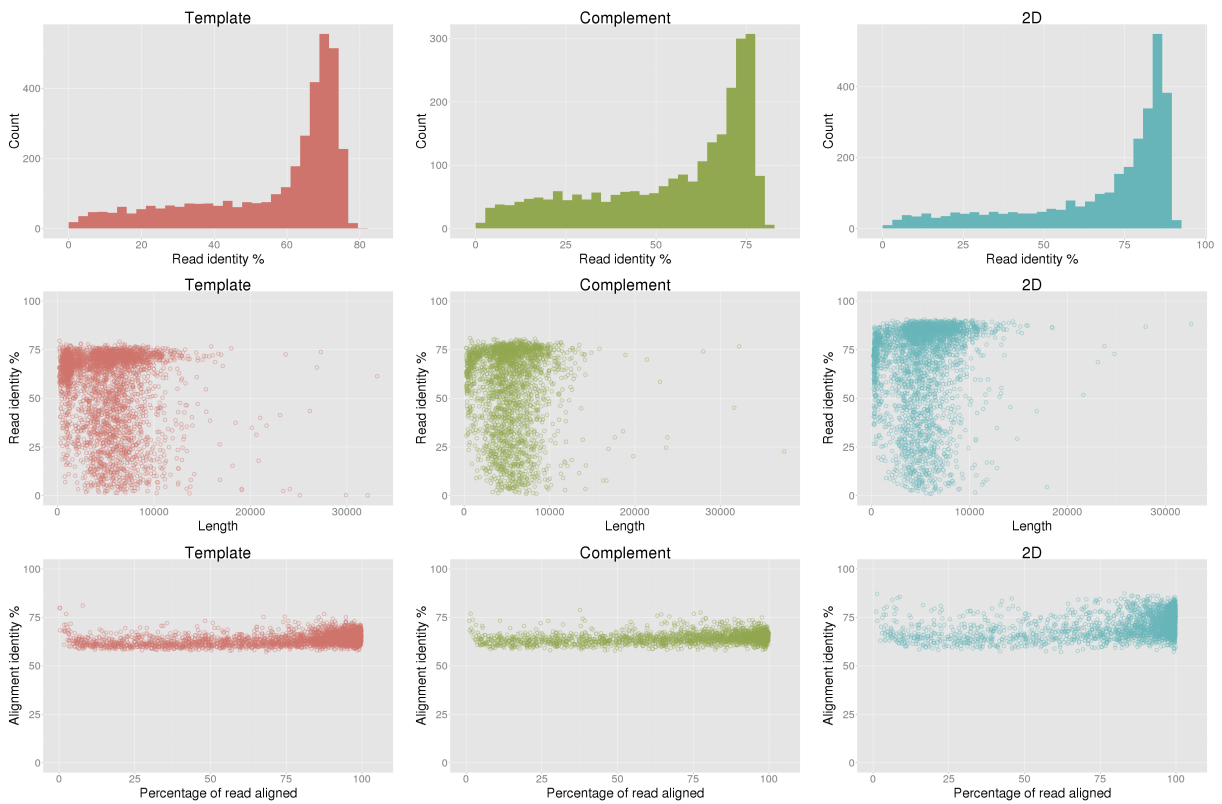


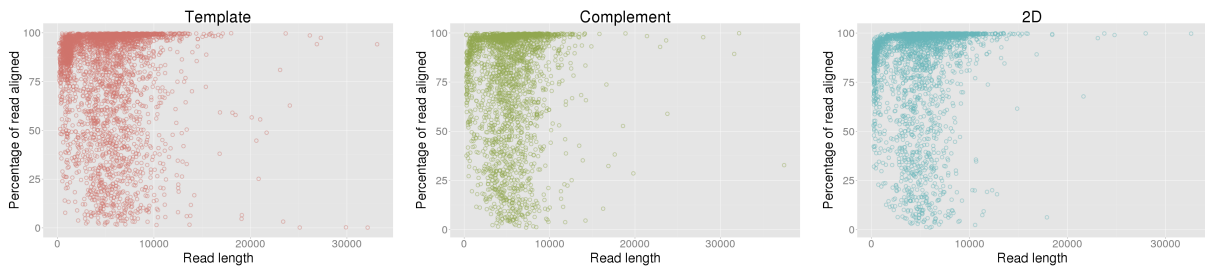
Escherichia coli error analysis

	Template	Complement	2D
Overall base identity (excluding indels)	53.14%	53.77%	68.76%
Aligned base identity (excluding indels)	76.75%	79.35%	86.31%
Identical bases per 100 aligned bases (including indels)	63.79%	64.46%	72.83%
Inserted bases per 100 aligned bases (including indels)	4.43%	3.90%	2.50%
Deleted bases per 100 aligned bases (including indels)	12.46%	14.87%	13.12%
Substitutions per 100 aligned bases (including indels)	19.32%	16.78%	11.55%
Mean insertion size	1.50	1.49	1.37
Mean deletion size	1.68	1.83	1.80

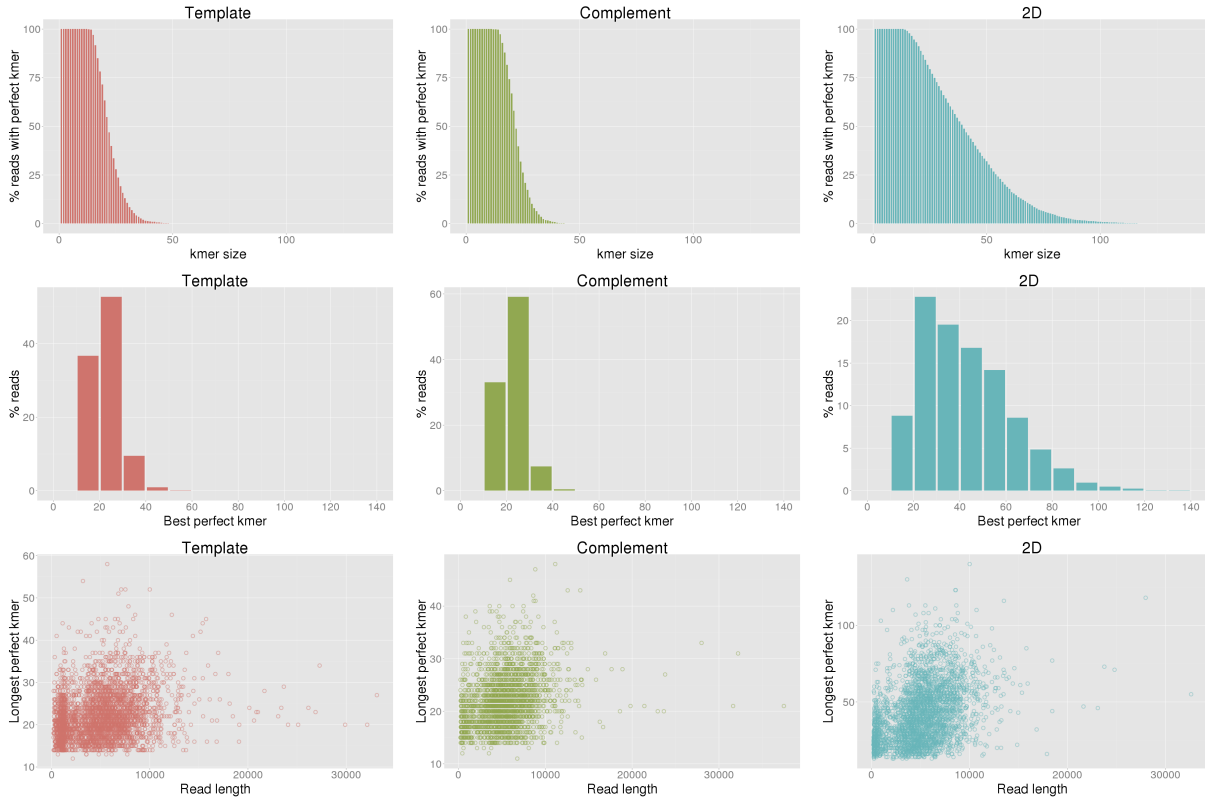


Escherichia coli read identity

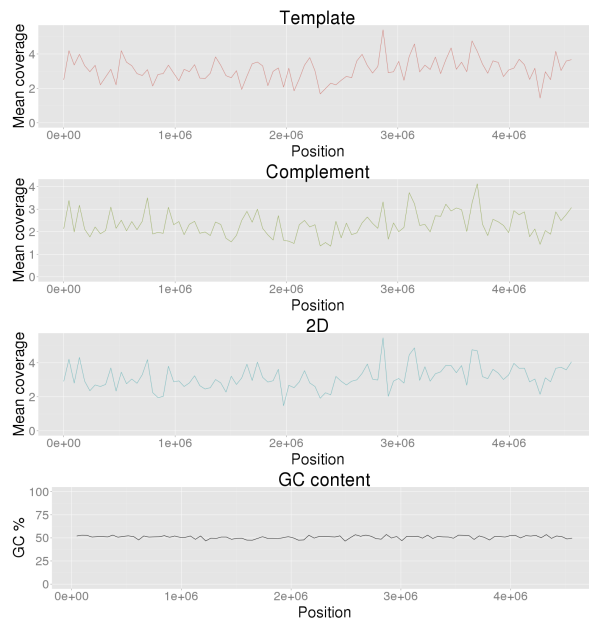




Escherichia coli perfect kmers



Escherichia coli coverage



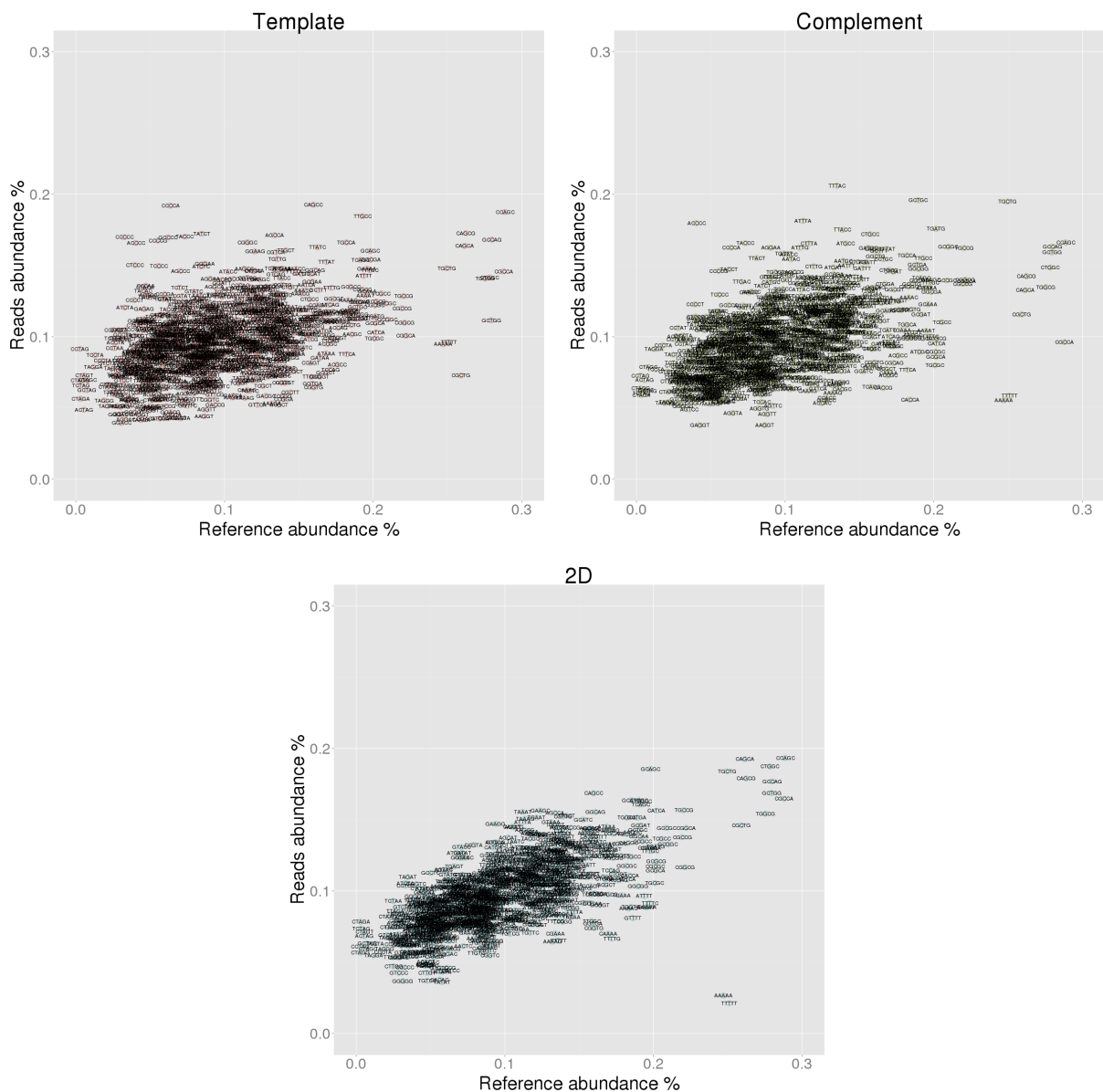
Escherichia coli 5-mer analysis

Under-represented 5-mers

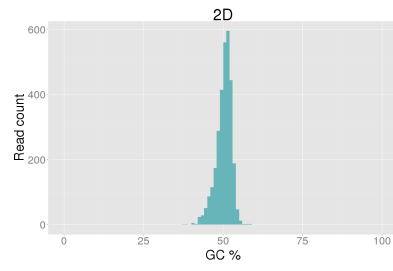
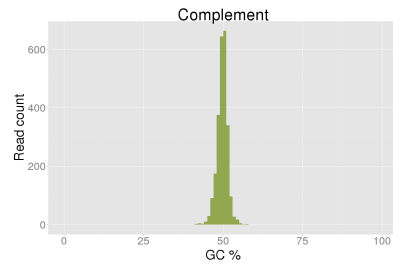
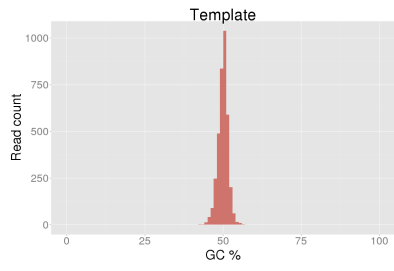
Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	CGCTG	0.259	0.073	-0.186	TTTTT	0.251	0.059	-0.192	TTTTT	0.251	0.021	-0.230
2	GCTGG	0.279	0.112	-0.168	CGCCA	0.288	0.096	-0.192	AAAAA	0.247	0.027	-0.221
3	TTTTT	0.251	0.096	-0.155	AAAAA	0.247	0.055	-0.192	CGCCA	0.288	0.165	-0.123
4	AAAAA	0.247	0.095	-0.152	CGCTG	0.259	0.116	-0.143	TGGCG	0.275	0.154	-0.121
5	CGCCA	0.288	0.146	-0.142	TGGCG	0.275	0.135	-0.141	CGCTG	0.259	0.146	-0.113
6	CTGGC	0.278	0.142	-0.137	CTGGC	0.278	0.149	-0.130	GCTGG	0.279	0.169	-0.111
7	TGGCG	0.275	0.141	-0.134	CAGCA	0.261	0.133	-0.129	GAAAA	0.196	0.089	-0.107
8	CGGCA	0.222	0.101	-0.121	CACCA	0.184	0.056	-0.128	AAAAT	0.195	0.089	-0.106
9	GCCAG	0.280	0.168	-0.112	CCAGC	0.289	0.166	-0.122	TTTTC	0.198	0.091	-0.106
10	CGGCG	0.221	0.110	-0.111	TGGCG	0.201	0.080	-0.121	TTTTG	0.172	0.067	-0.106

Over-represented 5-mers

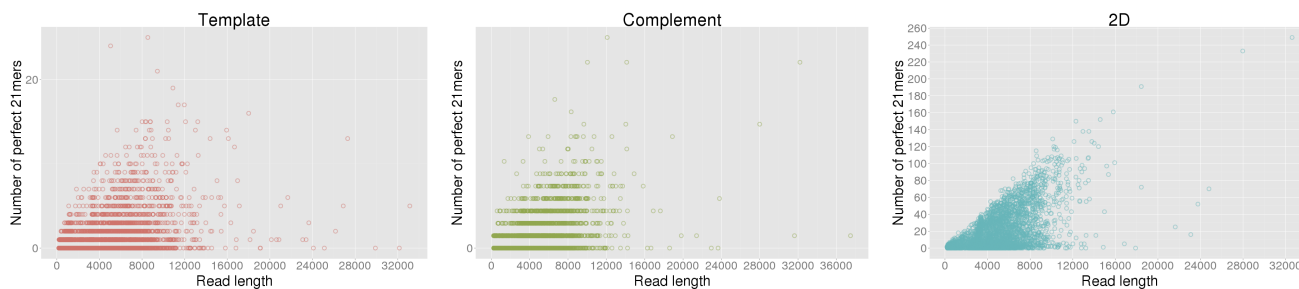
Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	CCCCC	0.033	0.170	0.137	ACCCC	0.040	0.180	0.140	CTAGA	0.003	0.079	0.075
2	CCCCA	0.064	0.192	0.128	CCCCA	0.064	0.163	0.099	TAGAT	0.035	0.110	0.075
3	ACCCC	0.040	0.166	0.126	TACCC	0.073	0.166	0.092	ATCTA	0.033	0.105	0.073
4	CCCCG	0.055	0.168	0.112	CCCCG	0.055	0.146	0.091	CCTCC	0.033	0.104	0.071
5	CTCCC	0.040	0.150	0.111	TACCT	0.062	0.147	0.085	TCTAG	0.003	0.074	0.070
6	GCCCC	0.062	0.170	0.108	CCCCT	0.039	0.123	0.083	TCTAA	0.025	0.093	0.068
7	TACCC	0.073	0.170	0.097	TAGGA	0.012	0.091	0.080	CTAGT	0.006	0.071	0.066
8	TCCCC	0.056	0.150	0.094	CCTAT	0.028	0.106	0.078	GGGTC	0.040	0.105	0.065
9	CCCAA	0.047	0.136	0.089	CCCTA	0.018	0.094	0.076	ACTAG	0.006	0.068	0.063
10	CCTAG	0.003	0.091	0.089	CTCCC	0.040	0.114	0.074	GCCTC	0.050	0.113	0.062



Escherichia coli GC content



All reference 21mer analysis



All reference substitutions

Reference	Template substituted %				Complement substituted %				2D substituted %			
	a	c	g	t	a	c	g	t	a	c	g	t
A	0.00	9.38	8.44	4.97	0.00	9.46	8.30	5.24	0.00	9.85	8.44	4.55
C	8.41	0.00	9.02	9.90	9.03	0.00	8.56	9.65	8.66	0.00	9.95	8.82
G	9.41	9.07	0.00	8.31	9.08	8.85	0.00	8.76	8.51	10.37	0.00	8.32
T	5.35	8.52	9.23	0.00	5.55	8.21	9.31	0.00	4.81	8.10	9.63	0.00

Kmer motifs before errors

3-mer error motif analysis

Rank	Template			Complement			2D		
	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution
1	GCA (2.89%)	AAA (3.13%)	AAA (4.20%)	GCA (2.81%)	AAA (3.11%)	AAA (4.08%)	GCA (3.32%)	AAA (2.74%)	GCA (3.75%)
2	TTC (2.84%)	TTC (3.02%)	TTC (3.29%)	AAA (2.66%)	GCA (2.83%)	GCA (3.57%)	TTC (2.60%)	GCA (2.68%)	AAA (3.33%)
3	AAA (2.79%)	TGC (2.97%)	GCA (3.14%)	TTC (2.65%)	GGC (2.76%)	GAA (3.21%)	TCA (2.58%)	GGC (2.54%)	GAA (3.11%)
4	ATC (2.41%)	GCA (2.78%)	TTT (2.92%)	TGC (2.62%)	TGC (2.75%)	TTT (2.91%)	AAA (2.48%)	TGC (2.53%)	GTT (2.72%)
5	TGC (2.36%)	GCC (2.56%)	GAA (2.81%)	CAG (2.50%)	GAA (2.65%)	TTT (2.89%)	ATC (2.44%)	CAG (2.39%)	TTT (2.66%)
6	TTT (2.36%)	GGC (2.25%)	TGC (2.49%)	GGC (2.45%)	TTT (2.48%)	TCA (2.38%)	GAA (2.37%)	TCA (2.38%)	TTC (2.52%)
7	TCA (2.30%)	TCA (2.19%)	AAT (2.38%)	ATC (2.23%)	CAG (2.39%)	TGC (2.38%)	AAT (2.27%)	GAA (2.31%)	AAT (2.49%)
8	GCC (2.24%)	GAA (2.18%)	CAA (2.28%)	TCA (2.21%)	GCC (2.34%)	CAA (2.27%)	GCG (2.26%)	GCC (2.25%)	TCA (2.27%)
9	GGC (2.17%)	AAC (2.15%)	TCA (2.20%)	GAA (2.20%)	TCA (2.25%)	AAT (2.16%)	TGC (2.13%)	GCG (2.23%)	GCC (2.15%)
10	AAT (2.17%)	CAG (2.13%)	GCC (2.18%)	AAT (2.19%)	TTT (2.20%)	GCC (2.11%)	CAG (2.12%)	CGC (2.20%)	CAG (2.05%)
-10	AGA (1.01%)	AGG (1.00%)	GGG (0.96%)	CTT (1.02%)	ACT (0.95%)	CCC (0.96%)	ACT (1.03%)	CTT (1.00%)	CTC (0.98%)
-9	CTC (0.97%)	GAG (0.93%)	CTC (0.93%)	AGT (0.97%)	CGA (0.94%)	CTC (0.86%)	CTC (1.00%)	AGA (0.96%)	ACT (0.96%)
-8	AGT (0.94%)	CCT (0.91%)	ACT (0.93%)	GGA (0.91%)	GTG (0.94%)	AGG (0.86%)	AGA (0.97%)	CGA (0.95%)	CCC (0.92%)
-7	GGA (0.91%)	CTT (0.90%)	CTT (0.90%)	CTC (0.86%)	AGG (0.92%)	GGG (0.84%)	GGA (0.93%)	CTC (0.95%)	CTT (0.91%)
-6	CCC (0.85%)	AGT (0.89%)	AGG (0.90%)	GAG (0.84%)	CCT (0.89%)	AGT (0.84%)	GAG (0.89%)	CCC (0.92%)	GAG (0.89%)
-5	GAG (0.77%)	CGA (0.83%)	GGT (0.89%)	AGG (0.81%)	GAG (0.87%)	CCT (0.83%)	AGG (0.81%)	ACT (0.91%)	CGA (0.84%)
-4	AGG (0.75%)	AGA (0.83%)	AGT (0.74%)	CCC (0.75%)	GGG (0.82%)	ACT (0.74%)	GGG (0.74%)	CCT (0.86%)	AGA (0.77%)
-3	GGG (0.72%)	GGA (0.70%)	GAG (0.70%)	GGG (0.64%)	CTC (0.77%)	GAG (0.68%)	CCC (0.70%)	GGA (0.80%)	GGA (0.70%)
-2	CTA (0.48%)	TAG (0.57%)	TAG (0.42%)	CTA (0.54%)	CTA (0.51%)	TAG (0.41%)	CTA (0.61%)	TAG (0.67%)	TAG (0.53%)
-1	TAG (0.38%)	CTA (0.49%)	CTA (0.33%)	TAG (0.45%)	TAG (0.51%)	CTA (0.41%)	TAG (0.51%)	CTA (0.67%)	CTA (0.42%)

Kmer space for 3-mers: 64

Random chance for any given 3-mer: 1.56%

4-mer error motif analysis

Rank	Template			Complement			2D		
	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution
1	T TTC (1.00%)	T TTC (1.08%)	A AAA (1.40%)	C AGC (0.97%)	T GGC (1.07%)	A AAA (1.15%)	G GCA (0.94%)	T GC (0.98%)	G GCA (1.22%)
2	A AAA (0.91%)	T GCC (1.00%)	T TTC (1.29%)	T GGC (0.92%)	C AGC (1.04%)	C AAA (1.03%)	A TCA (0.88%)	C AGC (0.97%)	G GAA (0.95%)
3	T GGC (0.88%)	T TG (0.99%)	G AAA (1.14%)	C TG (0.82%)	C GC (0.96%)	G AAA (1.01%)	G CCA (0.84%)	T TCA (0.83%)	T GTT (0.92%)
4	G AAA (0.84%)	A AA (0.95%)	T TTT (1.06%)	A TCA (0.82%)	C TG (0.92%)	G GCA (0.97%)	T TCA (0.80%)	C AAA (0.81%)	G AAA (0.91%)
5	C AGC (0.83%)	G AAA (0.91%)	C AAA (0.98%)	C CAG (0.81%)	C AAA (0.87%)	A GCA (0.92%)	T GCA (0.78%)	C TG (0.81%)	T GCA (0.91%)
6	G CAA (0.82%)	C AGC (0.90%)	G CAA (0.97%)	G GCA (0.80%)	T GC (0.82%)	T TTT (0.91%)	C GCA (0.78%)	C GC (0.80%)	C GTT (0.89%)
7	G CAA (0.82%)	C TG (0.89%)	G GCA (0.96%)	T GCC (0.77%)	C CAG (0.82%)	T TTT (0.88%)	A GCA (0.77%)	T GC (0.77%)	A GCA (0.88%)
8	T TTT (0.79%)	T GGC (0.89%)	G GAA (0.90%)	C GC (0.76%)	A AAA (0.81%)	T GAA (0.88%)	C TTC (0.76%)	T AAA (0.76%)	T GAA (0.84%)
9	G CCA (0.78%)	C AAA (0.83%)	A AT (0.88%)	C AAA (0.76%)	T TG (0.79%)	G GAA (0.87%)	A AC (0.76%)	C GC (0.75%)	A GAA (0.83%)
10	A TCA (0.78%)	C GC (0.82%)	T GC (0.86%)	T TTC (0.75%)	C TG (0.79%)	A GAA (0.87%)	C AGC (0.76%)	C CAG (0.74%)	C AAA (0.83%)
-10	C GAG (0.12%)	C TAT (0.12%)	G GG (0.11%)	T AGA (0.12%)	A CTC (0.11%)	T AGT (0.11%)	T ATA (0.14%)	C CGA (0.15%)	G GAC (0.12%)
-9	A TAG (0.11%)	T AGT (0.11%)	T AGA (0.10%)	G TCC (0.12%)	A TCA (0.11%)	C TAT (0.11%)	A TCA (0.13%)	G GAC (0.15%)	T CGA (0.12%)
-8	C TAA (0.11%)	T CTA (0.11%)	C GAG (0.09%)	A CTA (0.11%)	G GAC (0.10%)	T CTA (0.11%)	G TCC (0.13%)	C GGA (0.15%)	C TTG (0.11%)
-7	T CTA (0.11%)	G GAC (0.11%)	T AGT (0.09%)	T AGT (0.10%)	A CCT (0.10%)	A CTA (0.10%)	C CCT (0.11%)	C TAA (0.14%)	C GGA (0.11%)
-6	T TAG (0.09%)	C TAA (0.10%)	G GAC (0.08%)	C GAG (0.10%)	C CTC (0.10%)	A CCT (0.09%)	C CCT (0.11%)	A CTA (0.13%)	T CTA (0.09%)
-5	G GAC (0.08%)	C GGA (0.10%)	T CTA (0.07%)	G GAC (0.08%)	T AGA (0.10%)	G GAC (0.08%)	G GAC (0.11%)	C CCT (0.11%)	A CTA (0.09%)
-4	T AGA (0.08%)	T AGG (0.08%)	A CTA (0.07%)	C CCT (0.08%)	C CCT (0.09%)	T AGG (0.07%)	T AGA (0.09%)	T AGG (0.09%)	T AGG (0.07%)
-3	C CTA (0.06%)	C CTA (0.07%)	T AGG (0.06%)	C CTA (0.08%)	T AGG (0.07%)	C CCT (0.06%)	C CTA (0.08%)	T AGA (0.09%)	T AGA (0.06%)
-2	T AGG (0.06%)	T AGA (0.06%)	C CTA (0.04%)	T AGG (0.05%)	C CTA (0.06%)	C CTA (0.05%)	T AGG (0.05%)	C CTA (0.09%)	C CTA (0.06%)
-1	C TAG (0.01%)	C TAG (0.01%)	C TAG (0.01%)	C TAG (0.01%)	C TAG (0.01%)	C TAG (0.01%)	C TAG (0.02%)	C TAG (0.01%)	C TAG (0.01%)

Kmer space for 4-mers: 256 Random chance for any given 4-mer: 0.39%

5-mer error motif analysis

Rank	Template			Complement			2D		
	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution
1	C AGC (0.38%)	T TGCC (0.37%)	G AAAA (0.43%)	C AGC (0.36%)	C AGC (0.46%)	C AGC (0.57%)	C AGC (0.45%)	C TGCC (0.42%)	C AGC (0.49%)
2	C TGCC (0.35%)	C AGC (0.37%)	T GCA (0.40%)	C CAGC (0.35%)	C TGCC (0.44%)	C GCA (0.42%)	C GCA (0.38%)	C AGC (0.36%)	T GCA (0.47%)
3	C GCA (0.33%)	T TGGC (0.36%)	C AGC (0.39%)	C TGCC (0.35%)	G GGC (0.35%)	G AAAA (0.36%)	T GCA (0.37%)	C CAGC (0.33%)	C GCA (0.44%)
4	C AAAA (0.32%)	C TGCC (0.35%)	C AAAA (0.38%)	C ATCA (0.32%)	C CAGC (0.35%)	C TCA (0.34%)	C GCA (0.35%)	G CAG (0.31%)	G CTT (0.36%)
5	T GCA (0.32%)	G CTGC (0.34%)	G CAAA (0.38%)	G CAGC (0.32%)	G CTGC (0.34%)	C ATCA (0.34%)	C TGCC (0.32%)	G CAGC (0.31%)	G AAAA (0.33%)
6	T TATC (0.32%)	G CAGC (0.34%)	C TGCC (0.36%)	C GCA (0.31%)	G AAAA (0.33%)	T TACC (0.33%)	C ATCA (0.30%)	G CAGC (0.30%)	A GAA (0.31%)
7	G CAAA (0.31%)	C GCA (0.33%)	A GAAA (0.35%)	G CTGC (0.30%)	G CAGC (0.32%)	A ATCA (0.33%)	C TCA (0.29%)	C GCA (0.29%)	A GAA (0.31%)
8	G CAGC (0.31%)	G AAAA (0.32%)	T TTTT (0.34%)	T GCA (0.29%)	A ATCA (0.31%)	T GCA (0.33%)	G CAAA (0.28%)	T GGC (0.28%)	T GCA (0.31%)
9	T TTTT (0.30%)	G CAGC (0.31%)	T TTTT (0.33%)	C GCA (0.29%)	C ATCA (0.30%)	A GAAA (0.33%)	T GGC (0.27%)	G CAAA (0.28%)	C GTT (0.29%)
10	C GTTG (0.30%)	G CAAA (0.30%)	T TGCC (0.32%)	G CAGC (0.28%)	C GCA (0.30%)	C AGA (0.32%)	T TTT (0.26%)	A TAAA (0.28%)	T TTT (0.28%)
-10	C TAGC (0.00%)	A GGTC (0.01%)	A CCTA (0.01%)	C TAGA (0.01%)	T CCTA (0.01%)	A CCTT (0.01%)	T AGGG (0.01%)	T AGGA (0.01%)	C CCTA (0.01%)
-9	G CTAG (0.00%)	T AGGG (0.01%)	C CCTA (0.01%)	C TTAG (0.01%)	T AGGG (0.00%)	T AGGG (0.01%)	C TAGT (0.01%)	T CCTA (0.01%)	T AGGA (0.01%)
-8	C TAGT (0.00%)	G CTAG (0.00%)	T TAGA (0.00%)	C CCTA (0.01%)	A CTA (0.00%)	G CCTT (0.01%)	C CTAG (0.01%)	C TAGT (0.01%)	T TAGA (0.01%)
-7	T CTAG (0.00%)	G GACC (0.00%)	T CTAG (0.00%)	G ACCT (0.01%)	C TAGC (0.00%)	C TAGC (0.00%)	G TAGG (0.01%)	G CTAG (0.01%)	A CTA (0.01%)
-6	C CCTA (0.00%)	A CTAG (0.00%)	G CTAG (0.00%)	G CTAG (0.01%)	G CTAG (0.00%)	G CTAG (0.00%)	T AGGA (0.01%)	C TAGC (0.00%)	G CTAG (0.01%)
-5	G GACC (0.00%)	C TAGC (0.00%)	C CTAG (0.00%)	G GTCC (0.01%)	C TAGA (0.00%)	C CCTA (0.00%)	C CCTA (0.01%)	T CTAG (0.00%)	C TAGG (0.00%)
-4	C TAGG (0.00%)	C TAGT (0.00%)	A CTAG (0.00%)	A CTAG (0.00%)	T CTAG (0.00%)	C TAGT (0.00%)	A CTAG (0.01%)	A CTAG (0.00%)	C TAG (0.00%)
-3	C TAGA (0.00%)	C TAGG (0.00%)	C TAGG (0.00%)	C CTAG (0.00%)	C TAGT (0.00%)	C TAGA (0.00%)	C TAGG (0.00%)	C TAGA (0.00%)	C TAGC (0.00%)
-2	C CTAG (0.00%)	C TAGA (0.00%)	C TAGT (0.00%)	C TAGT (0.00%)	C TAGC (0.00%)	C CTAG (0.00%)	G CTAG (0.00%)	C TAGC (0.00%)	C TAGA (0.00%)
-1	A CTAG (0.00%)	C CTAG (0.00%)	C TAGC (0.00%)	C TAGC (0.00%)	C CTAG (0.00%)	A CTAG (0.00%)	C TAGC (0.00%)	C CTAG (0.00%)	T CTAG (0.00%)

Kmer space for 5-mers: 1024 Random chance for any given 5-mer: 0.10%