NanoOK report for 20150422_MN02862_FAA40596_BOWDEN05_MdC_MARC_Phase1b.1

Pass and fail counts

Type	Pass	Fail
Template	11126	0
Complement	11126	0
2D	11126	0

Read lengths

Type	NumReads	TotalBases	Mean	Longest	Shortest	N50	N50Count	N90	N90Count
Template	11126	72121069	6482.21	32110	148	7850	3511	3980	8302
Complement	11126	69903271	6282.88	30467	158	7608	3510	3884	8296
2D	11126	72877668	6550.21	32011	126	7931	3509	4036	8290
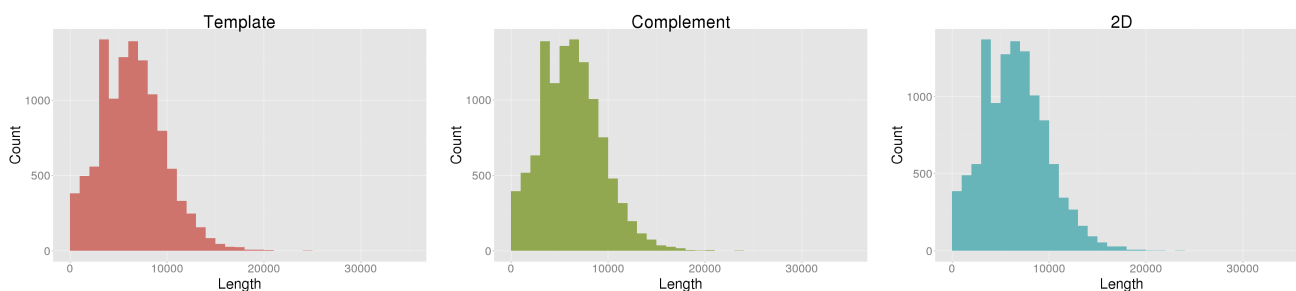
Template alignments

Number of reads	11126
Number of reads with alignments	11064 (99.44%)
Number of reads without alignments	62 (0.56%)

ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	875	7.86	3247.52	3017008	847.47	73
Escherichia coli	4641652	10189	91.58	6795.24	74650570	16.08	91

Complement alignments

Number of reads	11126
Number of reads with alignments	11057 (99.38%)
Number of reads without alignments	69 (0.62%)

| ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	873	7.85	3101.00	2815606	790.90	59
Escherichia coli	4641652	10184	91.53	6592.76	72820946	15.69	65

2D alignments

Number of reads	11126
Number of reads with alignments	11124 (99.98%)
Number of reads without alignments	2 (0.02%)

ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	882	7.93	3218.85	2956823	830.57	188
Escherichia coli	4641652	10242	92.05	6835.78	73026567	15.73	252

Control sequence error analysis

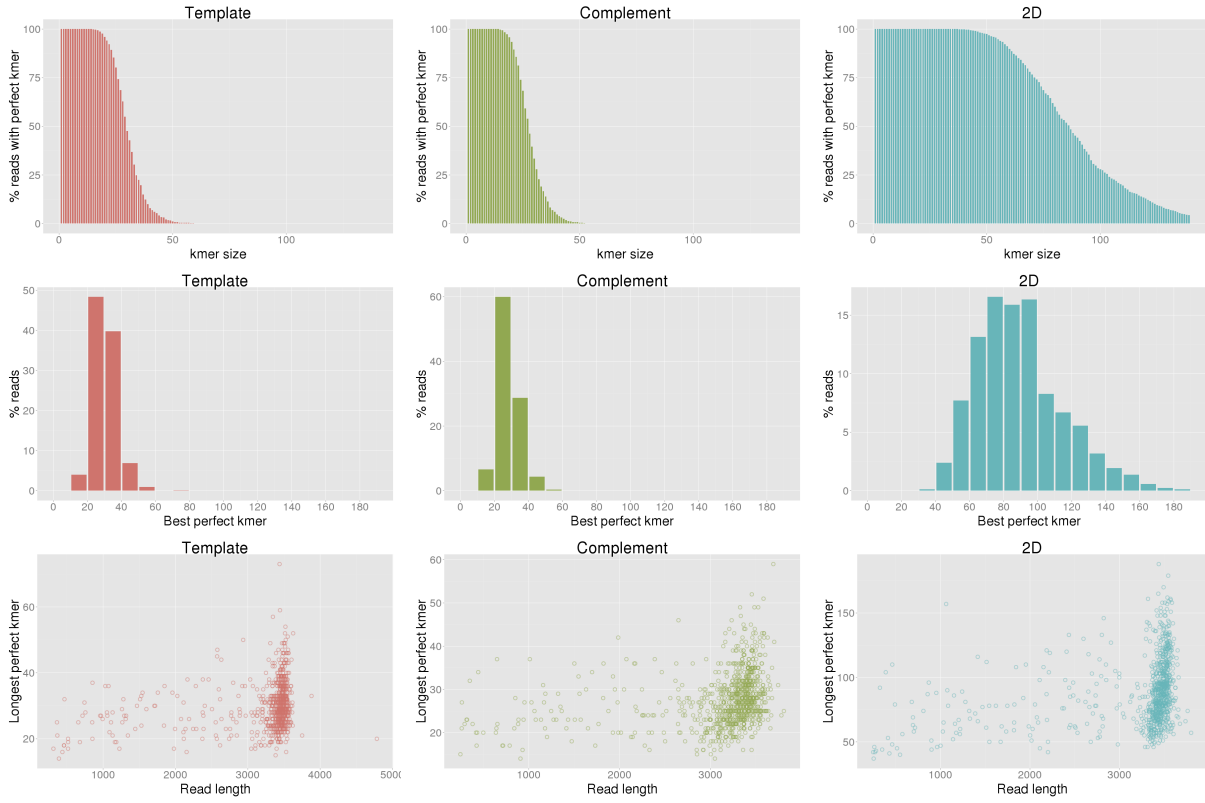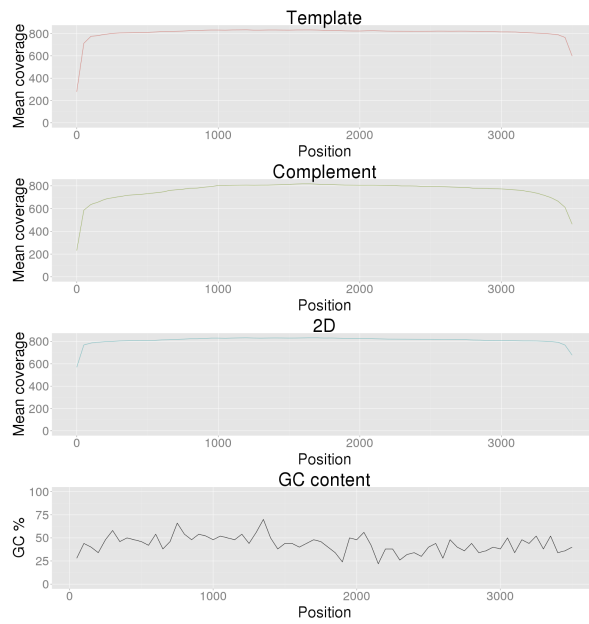|	Template	Complement	2D
Overall base identity (excluding indels)	74.40%	72.23%	90.86%
Aligned base identity (excluding indels)	81.63%	82.54%	94.99%
Identical bases per 100 aligned bases (including indels)	70.07%	69.44%	87.24%
Inserted bases per 100 aligned bases (including indels)	5.08%	4.46%	2.42%
Deleted bases per 100 aligned bases (including indels)	9.08%	11.41%	5.73%
Substitutions per 100 aligned bases (including indels)	15.77%	14.69%	4.60%
Mean insertion size	1.56	1.54	1.43
Mean deletion size	1.55	1.71	1.61
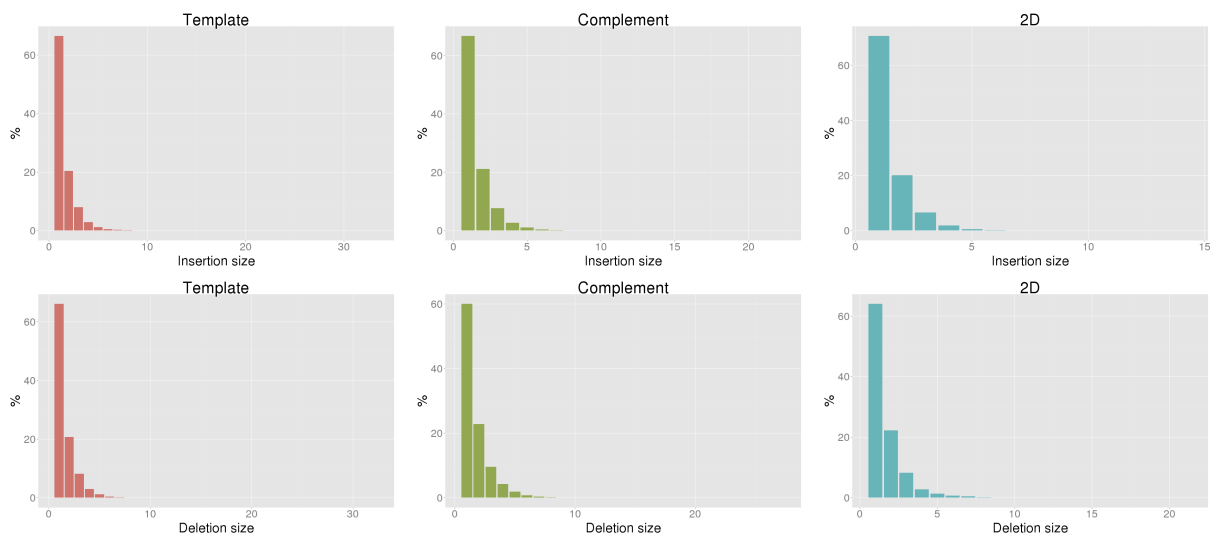
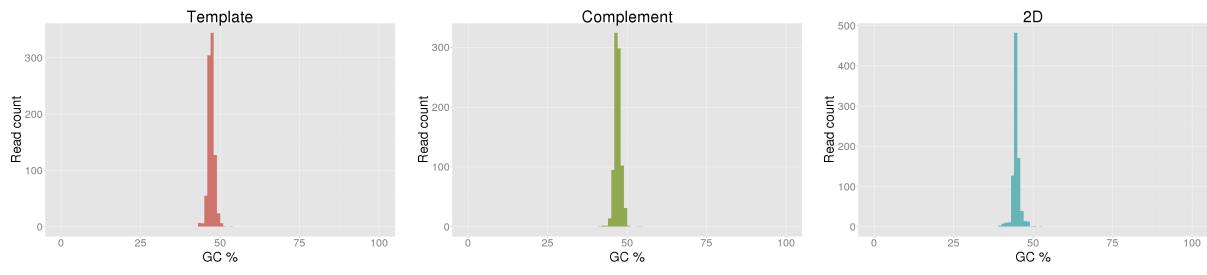# Control sequence read identity
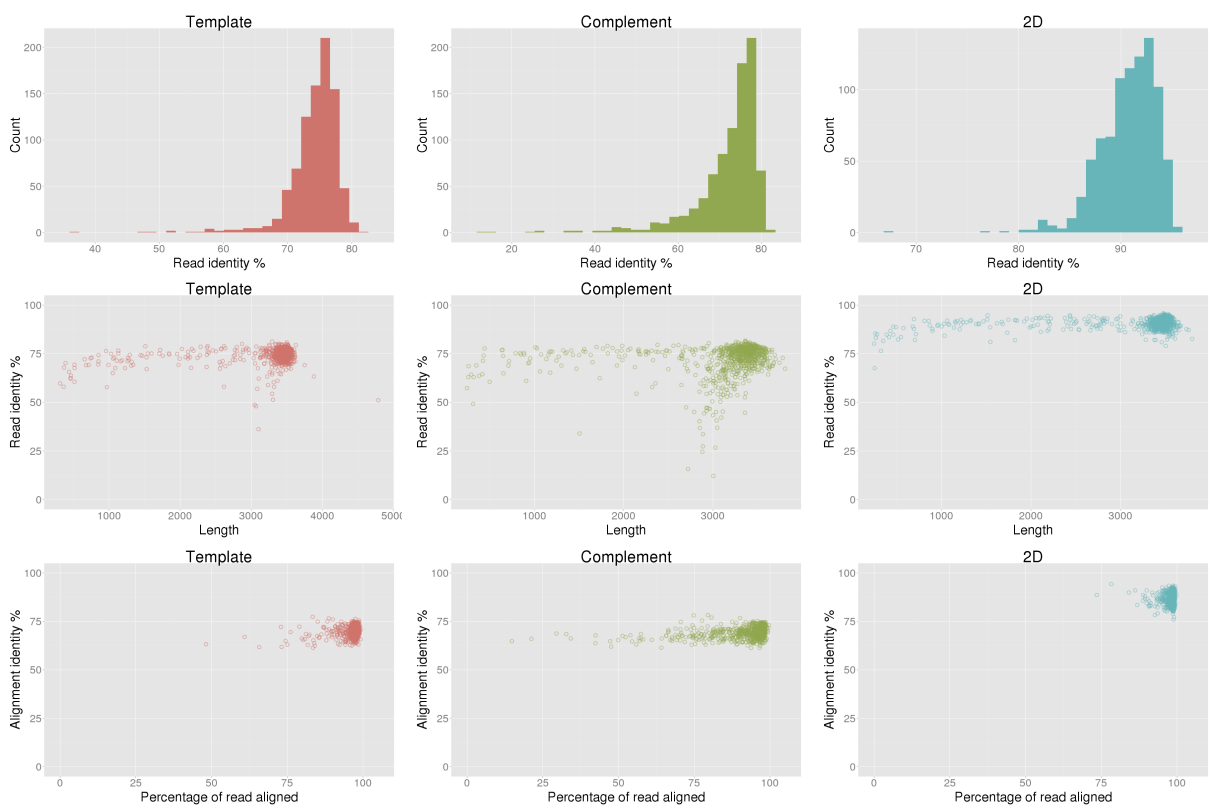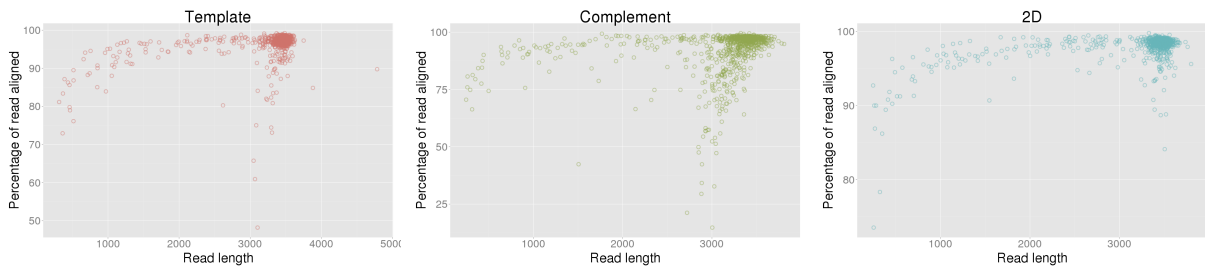
Control sequence perfect kmers

Control sequence coverage

3

Control sequence 5-mer analysis

Under-represented 5-mers

		Template				Complement				2D		
Rank	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	TTTTT	0.759	0.153	-0.606	TTTTT	0.759	0.102	-0.656	TTTTT	0.759	0.061	-0.698
2	AAAAA	0.478	0.127	-0.351	AAAAA	0.478	0.077	-0.401	AAAAA	0.478	0.083	-0.395
3	TGATG	0.393	0.146	-0.248	AAAAC	0.337	0.147	-0.190	TGATG	0.393	0.195	-0.198
4	GATGT	0.309	0.108	-0.201	GATGT	0.309	0.143	-0.166	CTGAT	0.309	0.156	-0.153
5	AAAAC	0.337	0.138	-0.199	GCAAT	0.309	0.144	-0.165	GATGT	0.309	0.161	-0.148
6	CTGAT	0.309	0.120	-0.189	TGATG	0.393	0.234	-0.160	GCTGA	0.281	0.145	-0.136
7	GCAAT	0.309	0.144	-0.165	GTCAG	0.281	0.122	-0.159	CTTTT	0.253	0.120	-0.133
8	AGTAA	0.253	0.090	-0.163	TGAAG	0.253	0.095	-0.158	TGTGA	0.225	0.100	-0.125
9	AATAT	0.309	0.147	-0.162	AACAA	0.281	0.125	-0.156	AAAAC	0.337	0.213	-0.124
10	AACAA	0.281	0.122	-0.159	ACACG	0.197	0.047	-0.150	TTATC	0.309	0.187	-0.122

Over-represented 5-mers

		Template				Complement				2D		
Rank	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	CTTTG	0.028	0.150	0.122	ACCCC	0.000	0.125	0.125	TCAGC	0.028	0.168	0.140
2	CATCT	0.000	0.117	0.117	CCCCA	0.000	0.114	0.114	CATCA	0.112	0.250	0.138
3	TCTAC	0.000	0.114	0.114	GAGGA	0.000	0.112	0.112	ATCAG	0.056	0.188	0.132
4	CGAGA	0.000	0.107	0.107	TATAC	0.000	0.111	0.111	AACCA	0.028	0.148	0.120
5	ATCTA	0.000	0.107	0.107	CTTTG	0.028	0.133	0.105	GCATC	0.084	0.199	0.115
6	CCCGC	0.000	0.106	0.106	TGCGA	0.028	0.129	0.101	CAGCA	0.056	0.168	0.111
7	ACCCC	0.000	0.104	0.104	TCAGC	0.028	0.128	0.100	CATCT	0.000	0.111	0.111
8	CCCCA	0.000	0.104	0.104	TACTT	0.000	0.099	0.099	TACAT	0.028	0.130	0.102
9	GCGAA	0.028	0.132	0.104	ACTCT	0.000	0.098	0.098	ACCAA	0.028	0.129	0.101
10	GCTCC	0.000	0.102	0.102	GGAAT	0.112	0.208	0.096	TCACA	0.028	0.129	0.101

Control sequence GC content

Escherichia coli error analysis

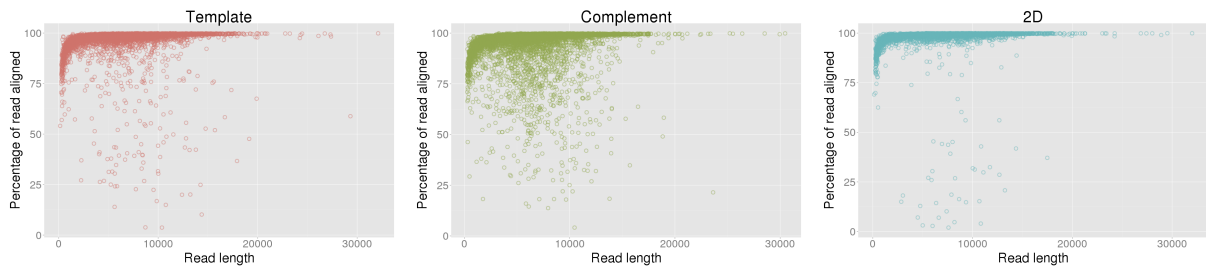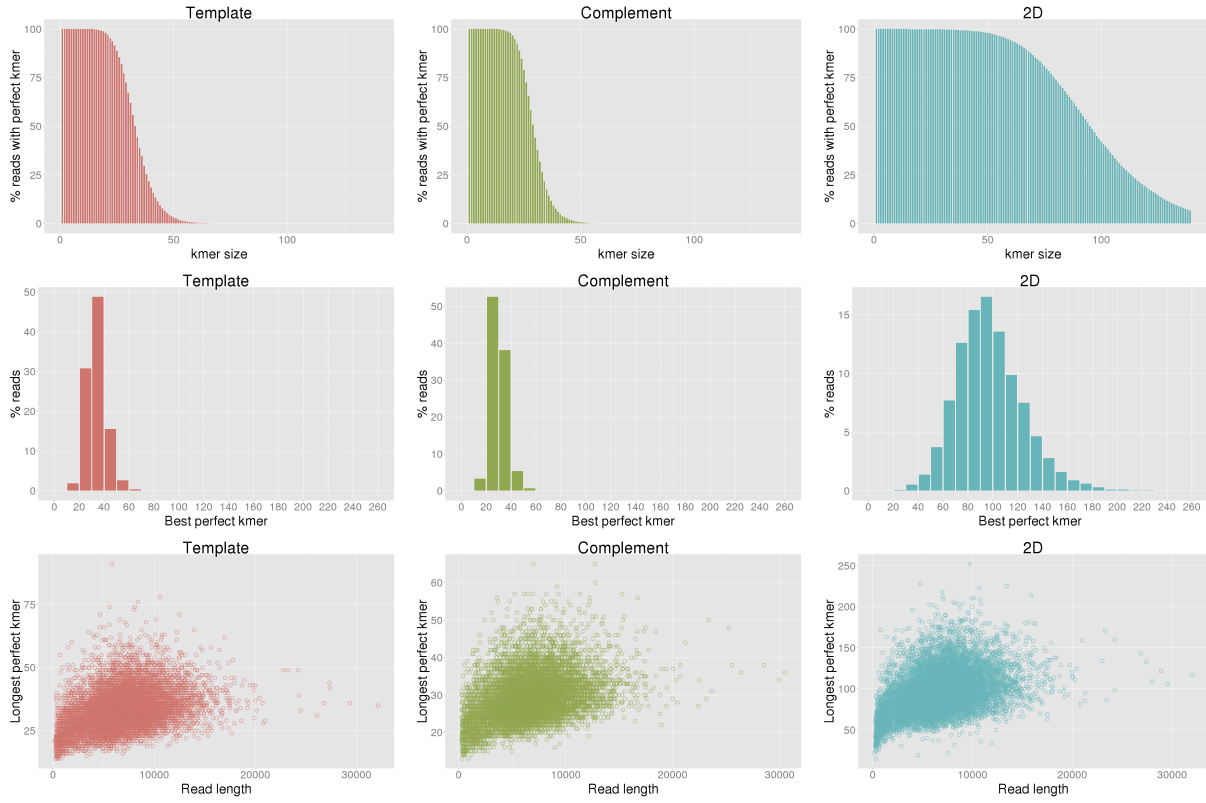|	Template	Complement	2D
Overall base identity (excluding indels)	74.78%	74.86%	91.51%
Aligned base identity (excluding indels)	81.20%	82.27%	95.10%
Identical bases per 100 aligned bases (including indels)	69.36%	69.02%	87.73%
Inserted bases per 100 aligned bases (including indels)	5.40%	4.75%	2.58%
Deleted bases per 100 aligned bases (including indels)	9.18%	11.35%	5.17%
Substitutions per 100 aligned bases (including indels)	16.06%	14.88%	4.52%
Mean insertion size	1.60	1.55	1.46
Mean deletion size	1.58	1.69	1.54
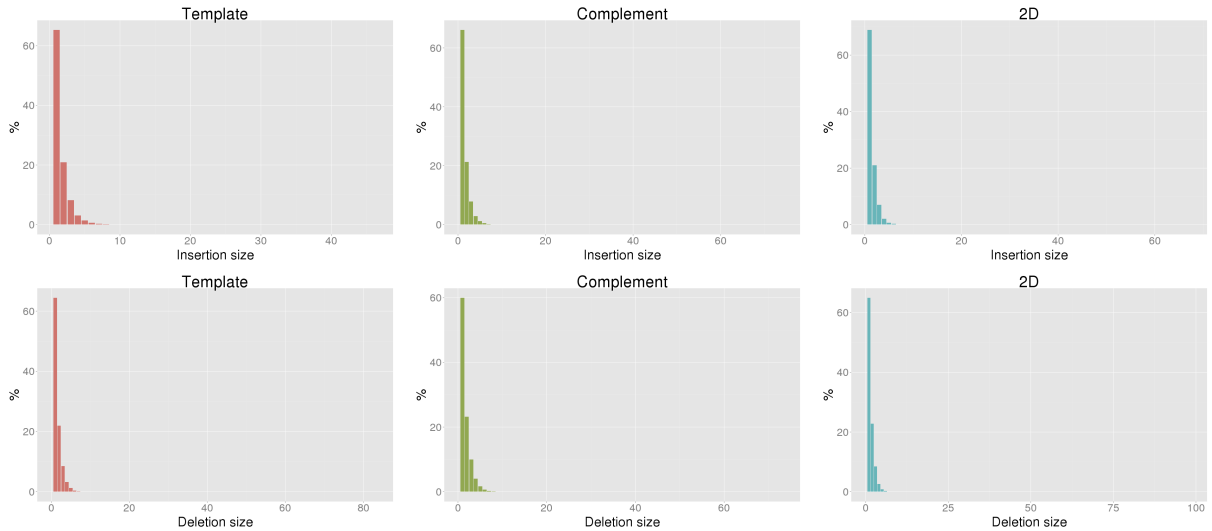
Escherichia coli read identity

6
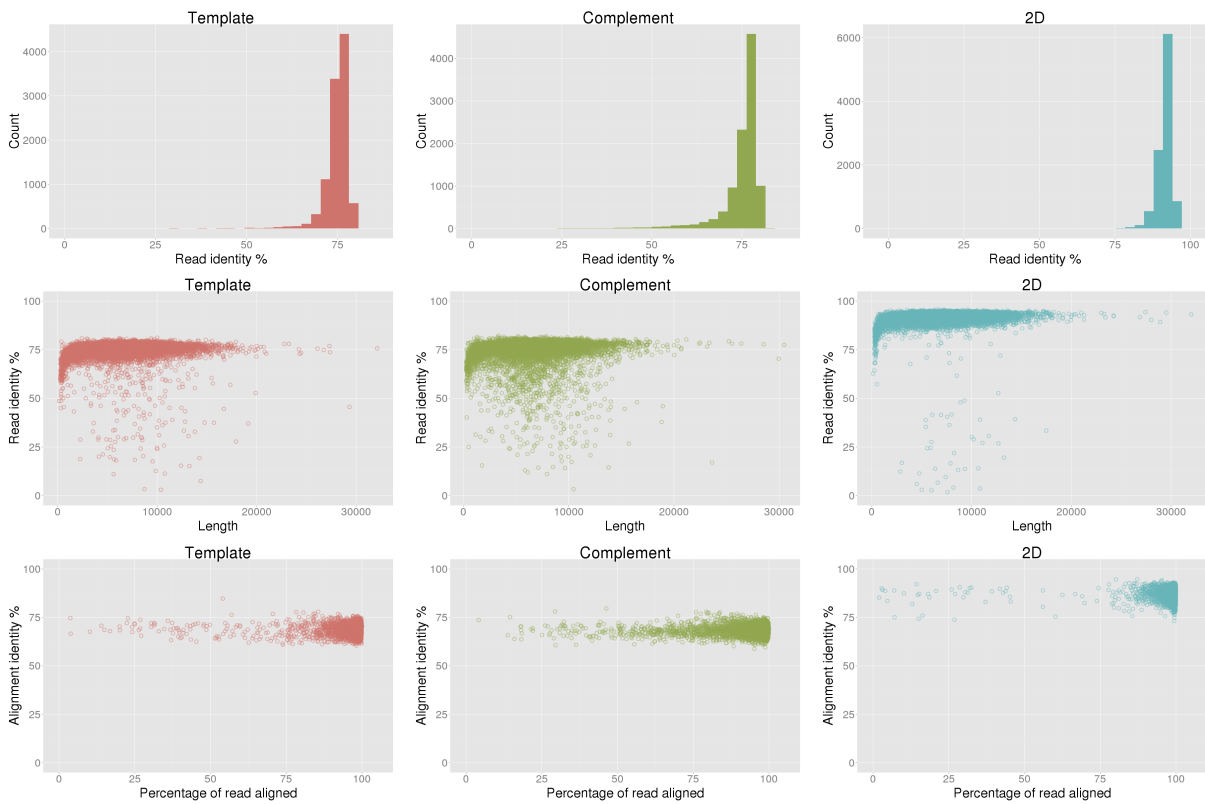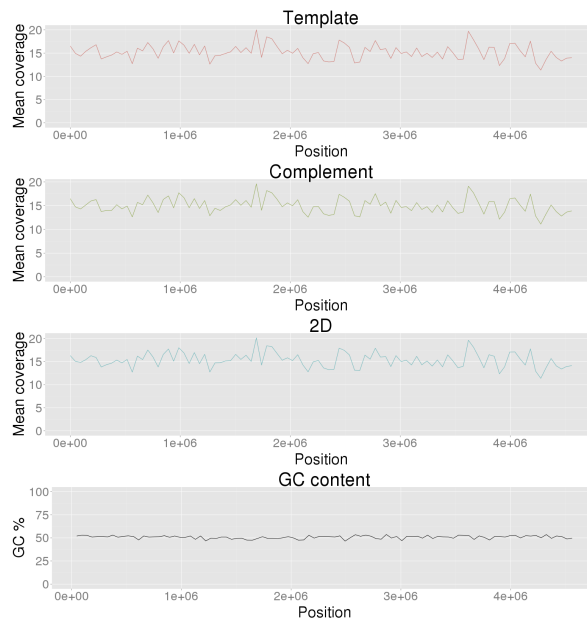
## Escherichia coli perfect kmers

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	0.106	-0.153	CGCCA	0.288	0.095	-0.193	TTTTT	0.251	0.041	-0.210
2	TTTTT	0.251	0.102	-0.149	AAAAA	0.247	0.064	-0.183	AAAAA	0.247	0.053	-0.194
3	AAAAA	0.247	0.099	-0.148	TTTTT	0.251	0.072	-0.179	CAAAA	0.169	0.102	-0.067
4	GCTGG	0.279	0.146	-0.134	CACCA	0.184	0.051	-0.133	CGCCA	0.288	0.222	-0.066
5	CGCCA	0.288	0.166	-0.122	CCAGC	0.289	0.163	-0.126	AAAAT	0.195	0.130	-0.065
6	CCAGC	0.289	0.176	-0.113	CAGCA	0.261	0.137	-0.124	TGGTG	0.185	0.127	-0.058
7	GCCAG	0.280	0.171	-0.109	CGCTG	0.259	0.136	-0.123	AAAAG	0.132	0.074	-0.058
8	CTGGC	0.278	0.177	-0.102	GCCAG	0.280	0.161	-0.119	TAAAA	0.145	0.089	-0.056
9	CAGCA	0.261	0.165	-0.097	CTGGC	0.278	0.161	-0.117	CGCTG	0.259	0.204	-0.056
10	CGGCA	0.222	0.125	-0.096	TGGCG	0.275	0.162	-0.114	TGGCG	0.275	0.221	-0.054

Over-represented 5-mers

| Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	CCCCG	0.055	0.153	0.098	ACCCC	0.040	0.152	0.112	CAAAT	0.105	0.156	0.051
2	ACCCC	0.040	0.137	0.096	CCCCG	0.055	0.143	0.088	GGGGT	0.039	0.070	0.030
3	CCCCC	0.033	0.129	0.096	CCCCA	0.064	0.137	0.073	CCCAA	0.047	0.077	0.030
4	CCCCA	0.064	0.141	0.077	TACCC	0.073	0.141	0.068	GAAGG	0.094	0.122	0.028
5	GCCCC	0.062	0.134	0.072	TCCCC	0.056	0.123	0.067	AGGCA	0.093	0.121	0.028
6	CTCCC	0.040	0.111	0.071	CTGAG	0.050	0.114	0.064	CTCGT	0.042	0.069	0.027
7	CCTAG	0.003	0.073	0.071	CCTAG	0.003	0.066	0.063	GGTAC	0.070	0.097	0.026
8	TCCCC	0.056	0.124	0.068	CCTAA	0.026	0.089	0.063	GGGTC	0.040	0.067	0.026
9	GAGAG	0.046	0.114	0.068	GACCC	0.040	0.101	0.061	CGGGG	0.054	0.080	0.026
10	TCTAC	0.048	0.114	0.067	CCCCT	0.039	0.100	0.061	GAATC	0.077	0.103	0.026

Template

Complement

2D
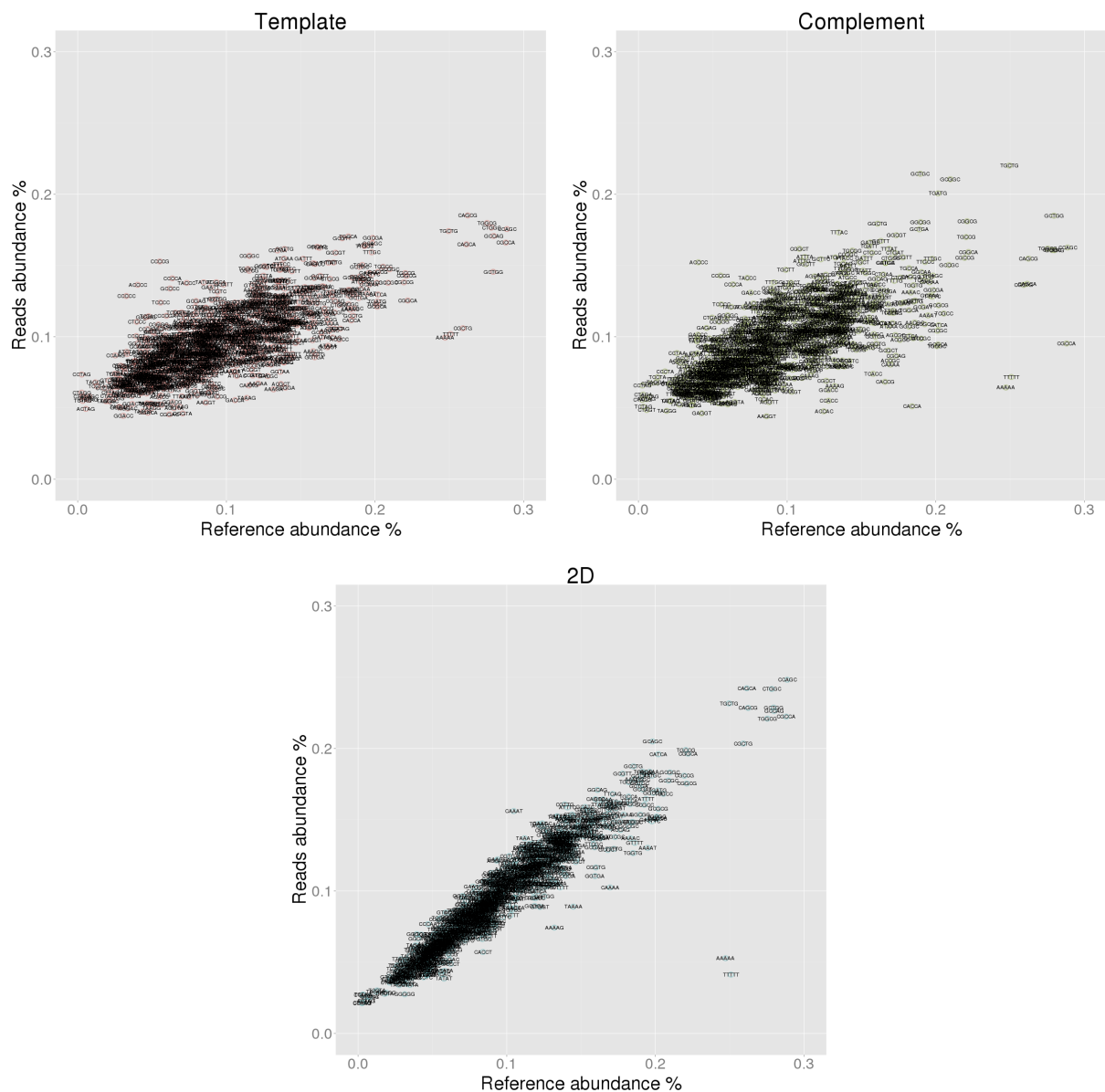
Escherichia coli GC content
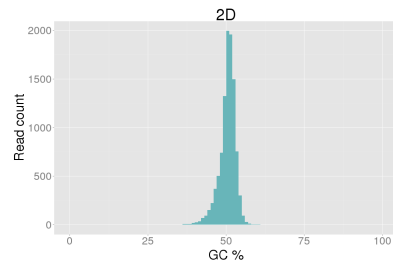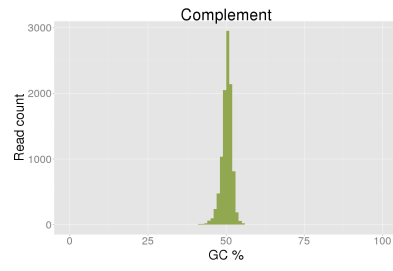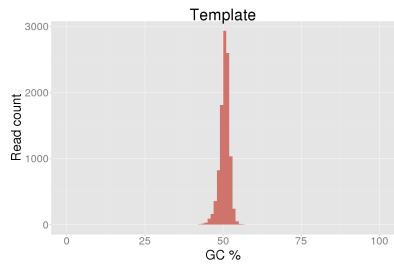
All reference 21mer analysis
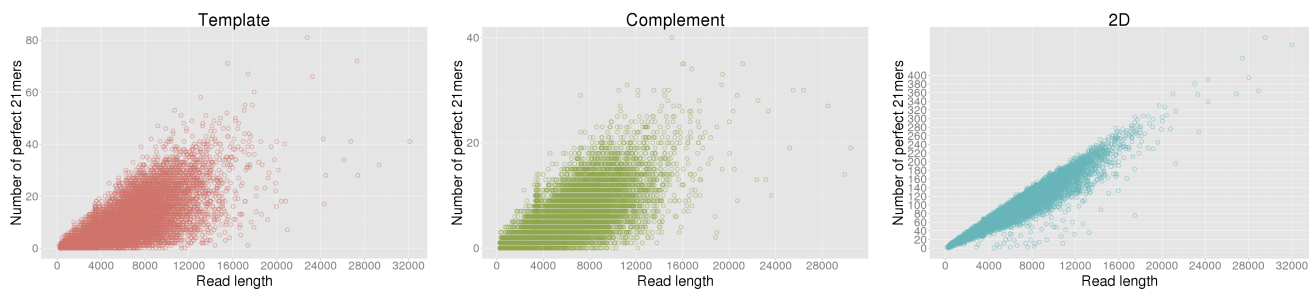
All reference substitutions

		Template substituted %				Complement substituted %				2D substituted %			
		a	c	g	t	a	c	g	t	a	c	g	t
Reference	A	0.00	8.93	9.33	4.64	0.00	9.06	8.74	4.86	0.00	8.37	8.60	3.78
	C	8.40	0.00	9.25	9.89	9.17	0.00	8.75	9.78	9.16	0.00	11.37	9.35
	G	9.45	9.18	0.00	7.99	8.97	8.89	0.00	8.72	9.11	11.50	0.00	8.43
	T	4.90	9.43	8.62	0.00	5.14	8.95	8.96	0.00	3.88	8.34	8.11	0.00

Kmer motifs before errors

3-mer error motif analysis

	Template			Complement			2D			
Rank	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution	
1	TTC (3.22%)	TTC (3.50%)	AAA (4.01%)	TGC (2.80%)	GGC (2.99%)	AAA (4.03%)	GCA (3.07%)	GGC (2.67%)	GCA (3.73%)	Most common
2	AAA (2.75%)	TGC (3.07%)	TTC (3.79%)	GCA (2.69%)	TGC (2.83%)	GCA (3.55%)	TTC (2.79%)	TGC (2.58%)	AAA (3.64%)	
3	GCA (2.74%)	GCA (2.79%)	GCA (3.22%)	TTC (2.64%)	AAA (2.78%)	GAA (3.32%)	TCA (2.66%)	TCA (2.53%)	GAA (3.27%)	
4	TGC (2.68%)	AAA (2.68%)	GAA (2.92%)	AAA (2.60%)	TTC (2.60%)	TTC (3.09%)	AAA (2.53%)	GCG (2.50%)	TTC (2.96%)	
5	ATC (2.60%)	GCC (2.52%)	TTT (2.65%)	CAG (2.55%)	GCA (2.58%)	TTT (2.65%)	ATC (2.48%)	AAA (2.43%)	TTT (2.84%)	
6	TCA (2.49%)	TCA (2.43%)	TGC (2.56%)	GGC (2.44%)	GAA (2.34%)	TGC (2.43%)	TGC (2.38%)	GCA (2.43%)	GCC (2.41%)	
7	GCC (2.31%)	GGC (2.42%)	AAT (2.33%)	GAA (2.32%)	CAG (2.31%)	ATC (2.38%)	GAA (2.27%)	TTC (2.42%)	GCG (2.33%)	
8	GGC (2.28%)	ATC (2.28%)	GCC (2.26%)	ATC (2.32%)	GCC (2.28%)	TCA (2.38%)	CAG (2.25%)	CGC (2.27%)	ATC (2.30%)	
9	TTT (2.21%)	AAC (2.22%)	TCA (2.22%)	TCA (2.29%)	TCA (2.25%)	GCC (2.14%)	GCG (2.24%)	CAG (2.24%)	GTT (2.30%)	
10	GAA (2.15%)	GCG (2.13%)	GGC (2.20%)	TTT (2.14%)	AAT (2.15%)	GGC (2.08%)	CGC (2.17%)	ATC (2.19%)	AAT (2.28%)	
	TTC	TTC	TTT	TTC	CCC	TTC	TTC	TTC	TTT	
	CC	CC	AAA	AA	AA	AA	CC	CC	AAA	
-10	TGT (0.97%)	GTA (0.95%)	GGT (0.95%)	AGA (1.02%)	CTT (0.95%)	CCC (0.95%)	CTT (1.06%)	GTA (0.99%)	TAT (0.89%)	Least common
-9	AGT (0.96%)	AGT (0.91%)	GGG (0.92%)	AGT (0.99%)	AGT (0.95%)	CTT (0.89%)	TAC (1.05%)	CCC (0.94%)	CCT (0.89%)	
-8	CCC (0.95%)	CCT (0.90%)	AGA (0.87%)	CTC (0.93%)	CCC (0.94%)	AGT (0.87%)	GAG (0.91%)	ACT (0.92%)	TGA (0.82%)	
-7	AGA (0.85%)	CTT (0.81%)	AGG (0.80%)	CCC (0.88%)	GGA (0.93%)	AGG (0.86%)	AGG (0.81%)	CGA (0.91%)	CTT (0.74%)	
-6	GGA (0.84%)	GAG (0.81%)	AGT (0.78%)	GGA (0.86%)	CTC (0.89%)	CCT (0.81%)	CCC (0.81%)	CTT (0.86%)	GAG (0.74%)	
-5	GAG (0.71%)	CGA (0.78%)	TGT (0.76%)	GAG (0.84%)	CCT (0.86%)	GGG (0.80%)	AGA (0.81%)	AGA (0.79%)	CGA (0.69%)	
-4	GGG (0.70%)	AGA (0.67%)	CTT (0.76%)	AGG (0.74%)	GAG (0.84%)	ACT (0.74%)	GGA (0.77%)	CCT (0.79%)	TAG (0.51%)	
-3	AGG (0.68%)	GGA (0.65%)	GAG (0.61%)	CTA (0.62%)	GGG (0.78%)	GAG (0.62%)	GGG (0.75%)	GGA (0.73%)	AGA (0.50%)	
-2	CTA (0.51%)	TAG (0.54%)	TAG (0.38%)	GGG (0.60%)	TAG (0.55%)	CTA (0.47%)	CTA (0.62%)	CTA (0.71%)	GGA (0.48%)	
-1	TAG (0.41%)	CTA (0.54%)	CTA (0.36%)	TAG (0.48%)	CTA (0.55%)	TAG (0.39%)	TAG (0.51%)	TAG (0.67%)	CTA (0.44%)	
	TTT	ATT	TTT	CTT	TTT	TTT	TTT	ATT	TTT	
	C C	CGA	C C	C C	CC	CC	CC	C A	C AA	
	AA	AAA	AAA	AAA	CC	CC	AAA			

Kmer space for 3-mers: 64   Random chance for any given 3-mer: 1.56%

4-mer error motif analysis

	Template			Complement			2D			
Rank	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution	
1	TTTC (1.04%)	TTTC (1.14%)	AAAA (1.36%)	CAGC (0.96%)	TGGC (1.09%)	AAAA (1.11%)	ATCA (0.99%)	TGGC (0.96%)	GGCA (1.21%)	Most common
2	ATCA (0.88%)	TGCC (0.97%)	TTTC (1.31%)	ATCA (0.92%)	CGGC (1.06%)	CAAA (1.03%)	GGCA (0.86%)	CAGC (0.90%)	AAAA (1.05%)	
3	AAAA (0.88%)	TGGC (0.91%)	GAAA (1.12%)	CTGC (0.88%)	CAGC (1.04%)	TGAA (0.97%)	GCCA (0.83%)	ATCA (0.83%)	TGAA (0.96%)	
4	GAAA (0.86%)	TTGC (0.90%)	GGCA (0.94%)	TGGC (0.85%)	CTGC (0.95%)	GAAA (0.95%)	CAGC (0.80%)	TTCA (0.83%)	TTTT (0.93%)	
5	CAGC (0.84%)	CAGC (0.89%)	TTTT (0.94%)	CGGC (0.85%)	TTGC (0.85%)	ATCA (0.94%)	AACA (0.77%)	CGGC (0.80%)	GAAA (0.92%)	
6	TTCA (0.82%)	TTCA (0.89%)	GTTC (0.91%)	CCAG (0.84%)	ATCA (0.81%)	GGCA (0.93%)	TTTC (0.77%)	CTGC (0.76%)	GGAA (0.92%)	
7	TGGC (0.82%)	CTGC (0.88%)	GCAA (0.90%)	TTGC (0.78%)	CAAA (0.80%)	AGCA (0.91%)	CGCA (0.76%)	TTGC (0.73%)	TTTC (0.91%)	
8	TGCC (0.81%)	AAAA (0.84%)	GGAA (0.85%)	CAAA (0.74%)	CTGG (0.78%)	AGAA (0.89%)	CCAG (0.76%)	GGCG (0.72%)	TGCA (0.87%)	
9	GGCA (0.79%)	TTCC (0.84%)	TGCC (0.85%)	ATGC (0.73%)	CCAG (0.78%)	TTTC (0.88%)	TTCA (0.75%)	TGCC (0.72%)	CGCC (0.86%)	
10	TTCT (0.79%)	CGCC (0.81%)	AAAT (0.82%)	AAAA (0.72%)	TTCC (0.76%)	TAAA (0.86%)	CGCC (0.75%)	GCCA (0.71%)	TGCC (0.85%)	
-10	GAGG (0.12%)	CCCT (0.13%)	TTAG (0.11%)	CGAG (0.12%)	CCTC (0.12%)	ACTA (0.11%)	GAGA (0.15%)	CCCC (0.15%)	ACTA (0.11%)	Least common
-9	TCTA (0.12%)	TCTA (0.13%)	ACTA (0.10%)	AGGG (0.12%)	CCCC (0.12%)	GTGT (0.10%)	GGGA (0.14%)	CTAT (0.15%)	TCGA (0.11%)	
-8	TAGT (0.12%)	CTAT (0.13%)	CGAG (0.10%)	CTAA (0.11%)	TAGA (0.12%)	GGAC (0.10%)	GAGG (0.14%)	ACCT (0.14%)	GAGA (0.10%)	
-7	CTAA (0.11%)	TAGT (0.12%)	TAGT (0.10%)	GGAC (0.11%)	GTGT (0.11%)	ACCT (0.10%)	CTAT (0.13%)	GAGA (0.14%)	CGGA (0.10%)	
-6	TTAG (0.11%)	CGGA (0.10%)	GGAC (0.09%)	TAGA (0.10%)	ACCT (0.11%)	CGAG (0.09%)	TCTA (0.13%)	CTAA (0.13%)	CTAT (0.10%)	
-5	GGAC (0.11%)	CTAA (0.09%)	TAGA (0.08%)	GAGG (0.10%)	CTAA (0.11%)	CTAT (0.09%)	CTAA (0.12%)	TAGG (0.11%)	GGGA (0.09%)	
-4	TAGA (0.07%)	TAGG (0.08%)	TCTA (0.08%)	CCCT (0.09%)	TAGG (0.08%)	TAGG (0.06%)	TAGA (0.09%)	CCCT (0.11%)	TAGG (0.07%)	
-3	CCTA (0.06%)	CCTA (0.06%)	TAGG (0.06%)	CCTA (0.08%)	CCCT (0.08%)	CCCT (0.06%)	CCTA (0.08%)	TAGA (0.09%)	TAGA (0.05%)	
-2	TAGG (0.05%)	TAGA (0.06%)	CCTA (0.04%)	TAGG (0.06%)	CCTA (0.06%)	CCTA (0.06%)	TAGG (0.06%)	CCTA (0.09%)	CCTA (0.04%)	
-1	CTAG (0.01%)	CTAG (0.01%)	CTAG (0.01%)	CTAG (0.01%)	CTAG (0.01%)	CTAG (0.01%)	CTAG (0.02%)	CTAG (0.02%)	CTAG (0.01%)	

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

5-mer error motif analysis

	Template			Complement			2D			
Rank	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution	Insertion	Deletion	Substitution	
1	CAGCA (0.34%)	CTGGC (0.38%)	GAAAA (0.42%)	CAGCA (0.40%)	CTGGC (0.40%)	CAGCA (0.52%)	CAGCA (0.37%)	CTGGC (0.39%)	CAGCA (0.44%)	Most common
2	CTGGC (0.34%)	CAGCA (0.36%)	CAGCA (0.41%)	CTGGC (0.33%)	CAGCA (0.39%)	CGGCA (0.36%)	CGCCA (0.34%)	CAGCA (0.30%)	CGGCA (0.41%)	
3	TTATC (0.32%)	TTGCC (0.33%)	CAAAA (0.37%)	CATCA (0.32%)	GCGGC (0.36%)	GAAAA (0.35%)	CGGCA (0.32%)	CATCA (0.30%)	TGGCA (0.41%)	
4	CGCCA (0.31%)	TTTGC (0.33%)	TGTTC (0.34%)	CCAGC (0.32%)	CCAGC (0.33%)	ATAAA (0.34%)	CTGGC (0.31%)	CGCCA (0.29%)	GAAAA (0.35%)	
5	CATCA (0.31%)	CGCCA (0.32%)	TGGCA (0.32%)	GCTGC (0.32%)	CATCA (0.32%)	CAGAA (0.34%)	CATCA (0.29%)	GCGGC (0.29%)	ATTTT (0.32%)	
6	ATTTC (0.29%)	GCTGC (0.31%)	GCCAG (0.32%)	GCGGC (0.30%)	GCTGC (0.31%)	CATCA (0.34%)	TGGCA (0.29%)	CCAGC (0.29%)	TTGCC (0.31%)	
7	CCAGC (0.28%)	GCCAG (0.30%)	CGTTC (0.32%)	AATCA (0.28%)	TCAGC (0.30%)	AATCA (0.33%)	GCCAG (0.29%)	TGGCG (0.29%)	GCGCA (0.29%)	
8	TTTCT (0.28%)	CATCA (0.30%)	CTGGC (0.32%)	GCCAG (0.27%)	AATCA (0.29%)	AAGAA (0.32%)	GATCA (0.28%)	GCCAG (0.28%)	TGGCG (0.29%)	
9	GAAAA (0.27%)	TTTCA (0.30%)	TTTTC (0.31%)	TTATC (0.27%)	ATAAA (0.29%)	TGGCA (0.32%)	CCAGC (0.27%)	TTTCA (0.27%)	GCAAA (0.28%)	
10	GCAGC (0.27%)	CCAGC (0.29%)	GCAAA (0.31%)	CGGCA (0.27%)	TTTGC (0.29%)	GCAAA (0.31%)	GCAAA (0.27%)	GCAGC (0.25%)	CAAAA (0.28%)	
-10	GGACC (0.01%)	CCCTA (0.01%)	ACCTA (0.01%)	GGGAC (0.01%)	ACCTA (0.01%)	CCCCT (0.01%)	CCCTA (0.01%)	TCCTA (0.01%)	TAGGA (0.01%)	Least common
-9	CCCTA (0.01%)	GGACC (0.01%)	CCCTA (0.01%)	CCCTA (0.01%)	CCCCC (0.01%)	ACCTA (0.01%)	TAGGA (0.01%)	TAGGA (0.01%)	ACTAG (0.01%)	
-8	CTAGC (0.00%)	ACTAG (0.00%)	ACTAG (0.00%)	CTAGC (0.01%)	GCTAG (0.01%)	GCTAG (0.01%)	GCTAG (0.01%)	GCTAG (0.01%)	CCCTA (0.01%)	
-7	GCTAG (0.00%)	CTAGC (0.00%)	GCTAG (0.00%)	GCTAG (0.01%)	CTAGC (0.00%)	CTAGC (0.00%)	CTAGT (0.01%)	ACTAG (0.01%)	GCTAG (0.01%)	
-6	CTAGT (0.00%)	GCTAG (0.00%)	CTAGC (0.00%)	CTAGT (0.01%)	CTAGT (0.00%)	ACTAG (0.00%)	CTAGC (0.00%)	CTAGC (0.01%)	CTAGT (0.00%)	
-5	ACTAG (0.00%)	CTAGT (0.00%)	CTAGT (0.00%)	ACTAG (0.00%)	ACTAG (0.00%)	CTAGA (0.00%)	CCTAG (0.00%)	CTAGT (0.01%)	CTAGC (0.00%)	
-4	CTAGA (0.00%)	TCTAG (0.00%)	CTAGG (0.00%)	CTAGA (0.00%)	CTAGA (0.00%)	CTAGT (0.00%)	ACTAG (0.00%)	TCTAG (0.00%)	CCTAG (0.00%)	
-3	CCTAG (0.00%)	CTAGG (0.00%)	CTAGA (0.00%)	TCTAG (0.00%)	TCTAG (0.00%)	CTAGG (0.00%)	TCTAG (0.00%)	CTAGG (0.00%)	TCTAG (0.00%)	
-2	CTAGG (0.00%)	CCTAG (0.00%)	CCTAG (0.00%)	CTAGG (0.00%)	CTAGG (0.00%)	CCTAG (0.00%)	CTAGG (0.00%)	CCTAG (0.00%)	CTAGG (0.00%)	
-1	TCTAG (0.00%)	CTAGA (0.00%)	TCTAG (0.00%)	CCTAG (0.00%)	CCTAG (0.00%)	TCTAG (0.00%)	CTAGA (0.00%)	CTAGA (0.00%)	CTAGA (0.00%)	

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