NanoOK report for JOGrady_MARC1b_run2

Pass and fail counts

Type	Pass	Fail
Template	14235	0
Complement	14235	0
2D	14235	0

Read lengths

Type	NumReads	TotalBases	Mean	Longest	Shortest	N50	N50Count	N90	N90Count
Template	14235	72180533	5070.64	47312	137	7318	3772	3450	9371
Complement	14235	69660350	4893.60	40899	147	7056	3762	3298	9336
2D	14235	72429202	5088.11	46475	141	7360	3758	3455	9338
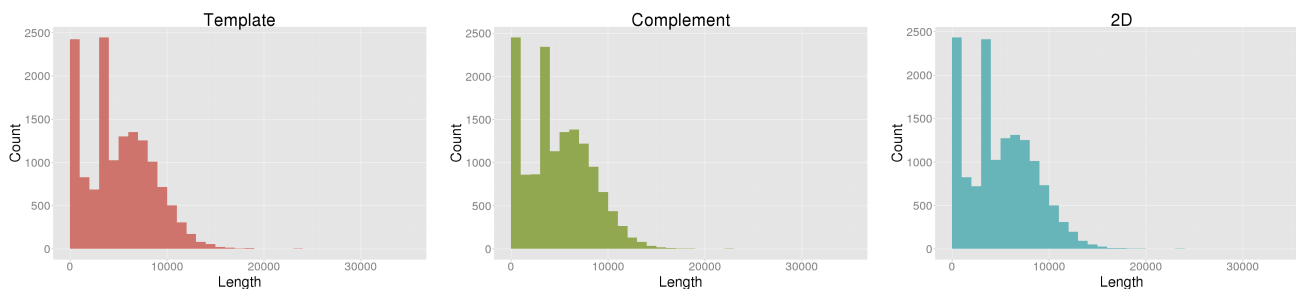
Template alignments

Number of reads	14235
Number of reads with alignments	11874 (83.41%)
Number of reads without alignments	2361 (16.59%)

ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	2044	14.36	3217.37	6981775	1961.17	64
Escherichia coli	4641652	9830	69.06	6503.17	68437244	14.74	93

Complement alignments

Number of reads	14235
Number of reads with alignments	11836 (83.15%)
Number of reads without alignments	2399 (16.85%)

| ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	2038	14.32	3071.46	6426091	1805.08	59
Escherichia coli	4641652	9798	68.83	6296.53	66103438	14.24	75

2D alignments

Number of reads	14235
Number of reads with alignments	12050 (84.65%)
Number of reads without alignments	2185 (15.35%)

ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	2060	14.47	3181.82	6821480	1916.15	218
Escherichia coli	4641652	9990	70.18	6447.83	67057690	14.45	264

Control sequence error analysis

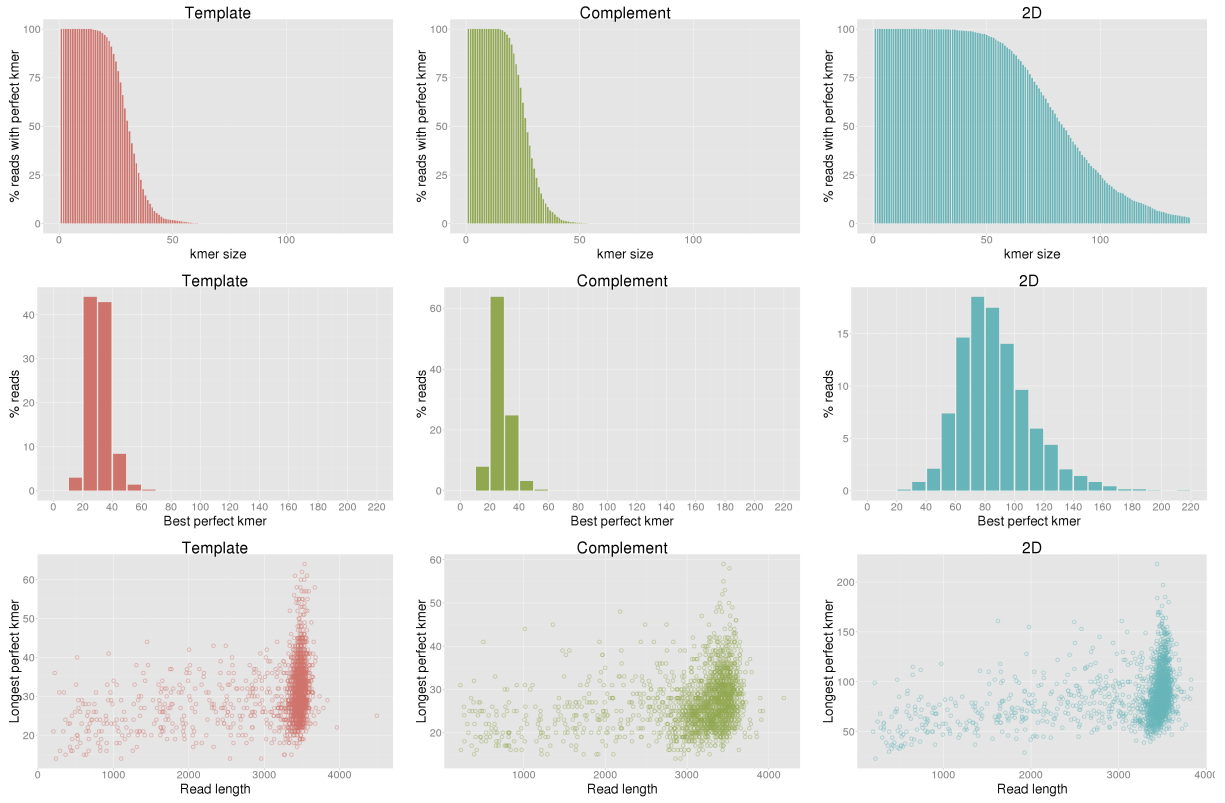|	Template	Complement	2D
Overall base identity (excluding indels)	74.64%	70.71%	90.24%
Aligned base identity (excluding indels)	81.81%	82.06%	94.59%
Identical bases per 100 aligned bases (including indels)	70.31%	68.88%	86.71%
Inserted bases per 100 aligned bases (including indels)	5.20%	4.75%	2.62%
Deleted bases per 100 aligned bases (including indels)	8.86%	11.31%	5.72%
Substitutions per 100 aligned bases (including indels)	15.63%	15.06%	4.96%
Mean insertion size	1.58	1.56	1.45
Mean deletion size	1.55	1.71	1.61
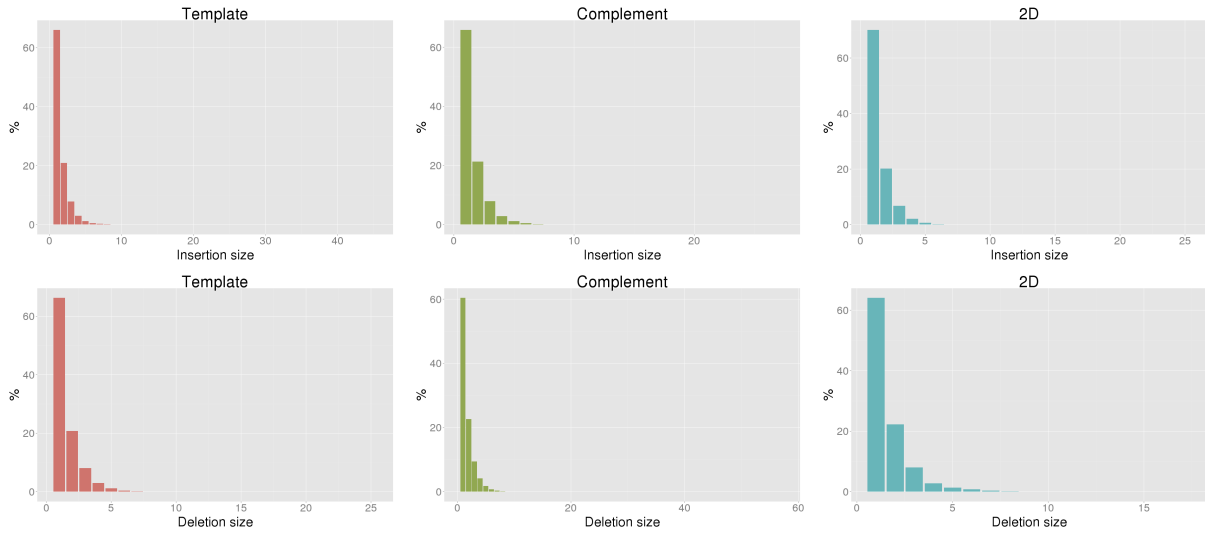
Control sequence read identity
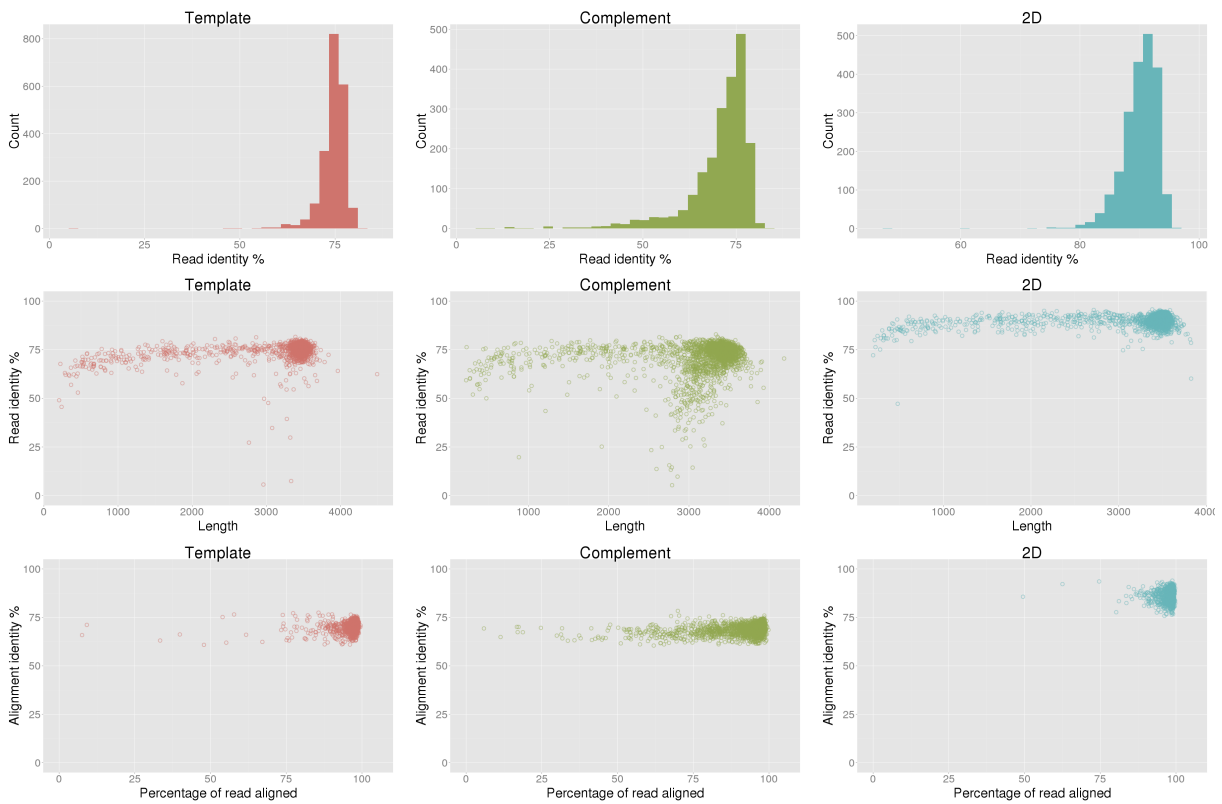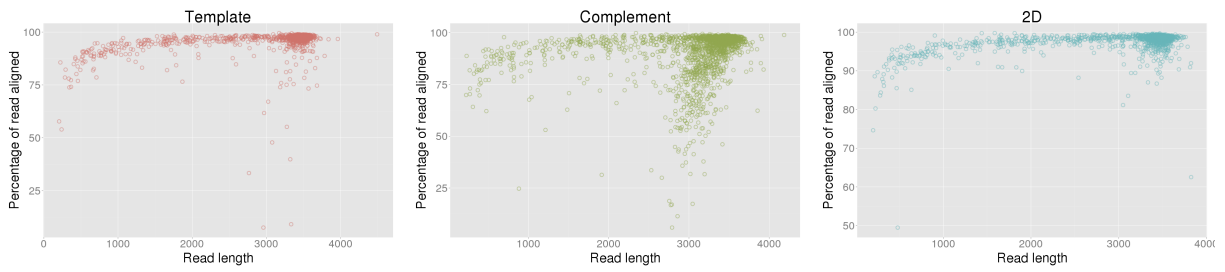
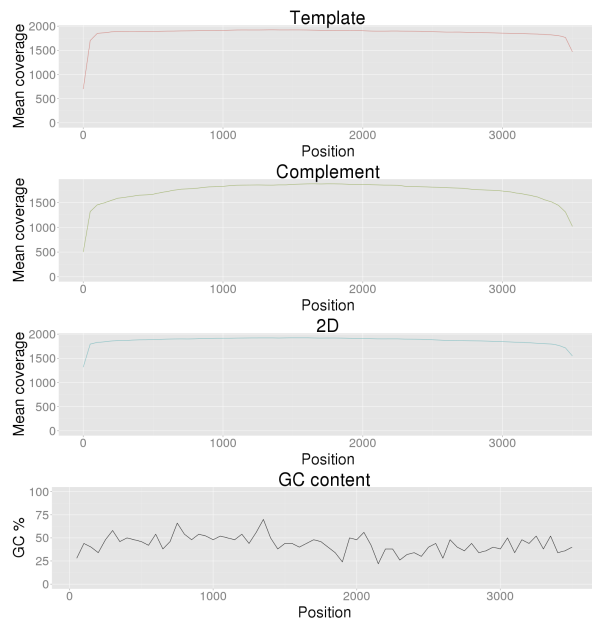## Control sequence perfect kmers

Control sequence coverage

3

Control sequence 5-mer analysis

Under-represented 5-mers

Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	TTTTT	0.759	0.162	-0.597	TTTTT	0.759	0.122	-0.637	TTTTT	0.759	0.069	-0.690
2	AAAAA	0.478	0.135	-0.343	AAAAA	0.478	0.086	-0.392	AAAAA	0.478	0.087	-0.391
3	TGATG	0.393	0.140	-0.253	AAAAC	0.337	0.147	-0.190	TGATG	0.393	0.191	-0.203
4	GATGT	0.309	0.105	-0.204	GATGT	0.309	0.138	-0.172	GATGT	0.309	0.155	-0.154
5	AAAAC	0.337	0.142	-0.195	GCAAT	0.309	0.139	-0.170	CTGAT	0.309	0.156	-0.153
6	CTGAT	0.309	0.125	-0.184	TGATG	0.393	0.226	-0.167	GCTGA	0.281	0.141	-0.140
7	GCTGA	0.281	0.120	-0.161	AACAA	0.281	0.120	-0.161	CTTTT	0.253	0.124	-0.129
8	GCAAT	0.309	0.148	-0.161	GTCAG	0.281	0.121	-0.160	TGTGA	0.225	0.100	-0.125
9	AGTAA	0.253	0.092	-0.161	TGAAG	0.253	0.097	-0.156	AAAAC	0.337	0.213	-0.124
10	AATAT	0.309	0.152	-0.157	ACACG	0.197	0.052	-0.145	TTATC	0.309	0.185	-0.124

Over-represented 5-mers

Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	CTTTG	0.028	0.146	0.118	ACCCC	0.000	0.127	0.127	TCAGC	0.028	0.170	0.142
2	CATCT	0.000	0.117	0.117	CCCCA	0.000	0.118	0.118	CATCA	0.112	0.250	0.138
3	TCTAC	0.000	0.111	0.111	GAGGA	0.000	0.113	0.113	ATCAG	0.056	0.190	0.134
4	CGAGA	0.000	0.110	0.110	TATAC	0.000	0.110	0.110	AACCA	0.028	0.148	0.120
5	TCAGC	0.028	0.131	0.103	TCAGC	0.028	0.130	0.102	GCATC	0.084	0.200	0.116
6	ACCCC	0.000	0.103	0.103	ACTCT	0.000	0.101	0.101	CATCT	0.000	0.112	0.112
7	CCCGC	0.000	0.102	0.102	CTTTG	0.028	0.128	0.100	CAGCA	0.056	0.168	0.112
8	TCTTA	0.000	0.102	0.102	GCCGA	0.000	0.099	0.099	ACCAA	0.028	0.132	0.103
9	GCGAA	0.028	0.130	0.102	TGCGA	0.028	0.125	0.097	TCACA	0.028	0.131	0.103
10	CCCCA	0.000	0.100	0.100	TACTT	0.000	0.096	0.096	TACAT	0.028	0.128	0.100
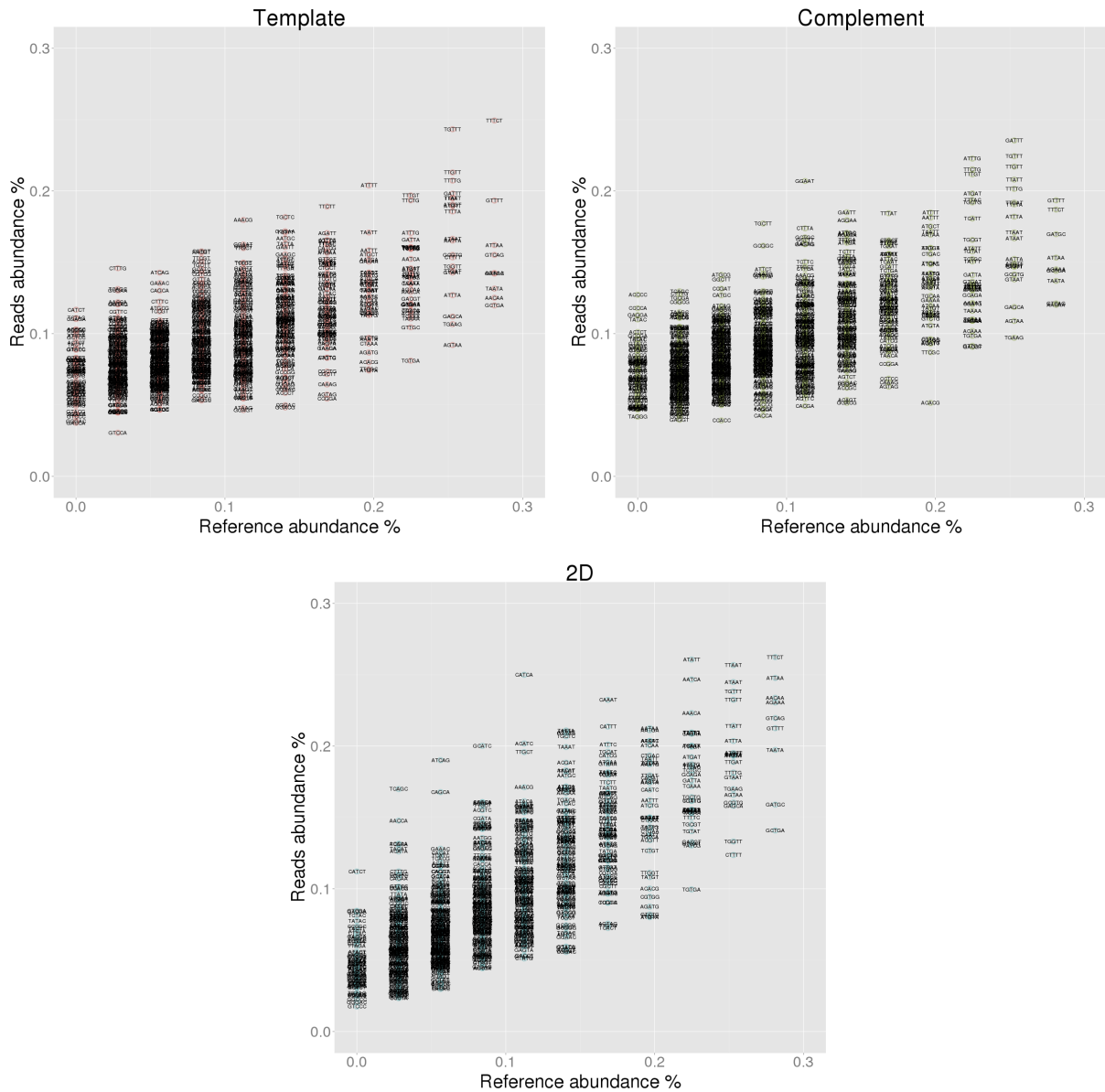
Control sequence GC content
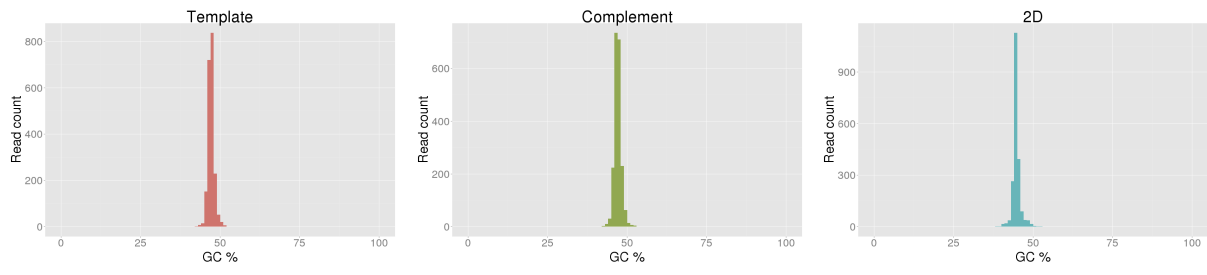
Escherichia coli error analysis

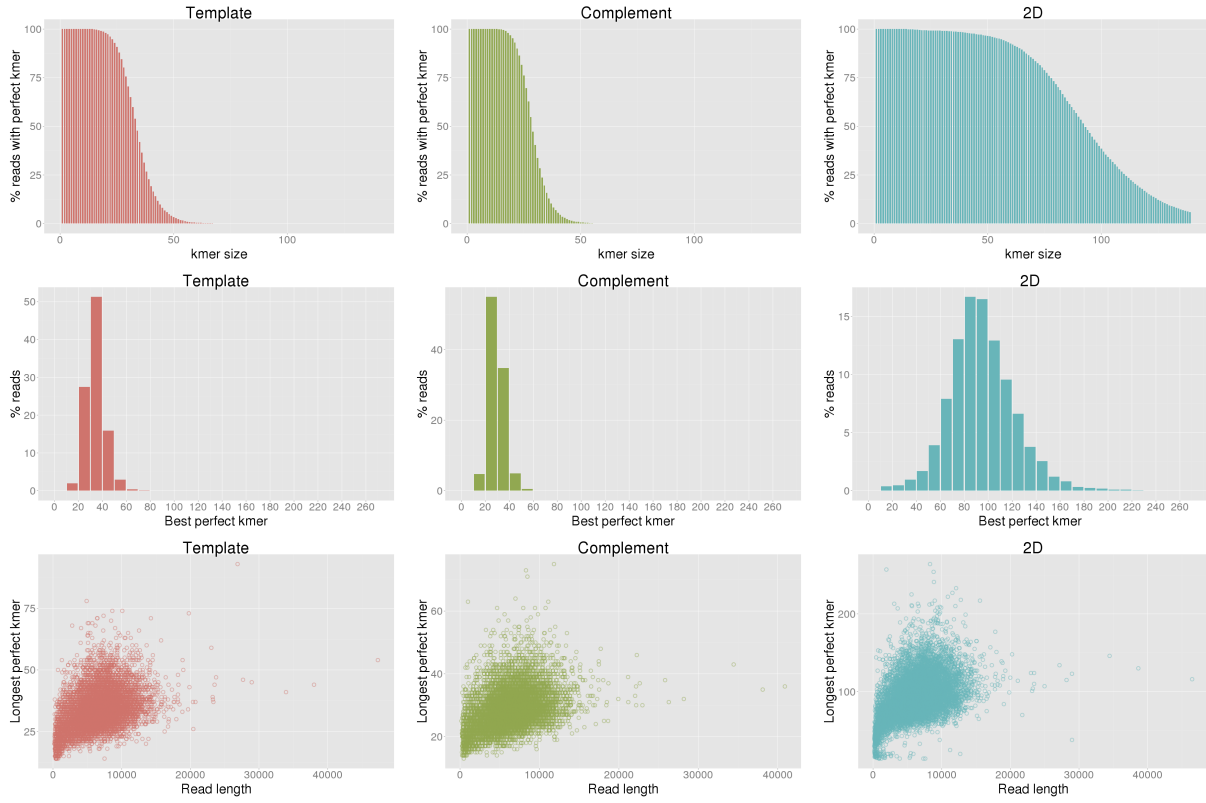|	Template	Complement	2D
Overall base identity (excluding indels)	74.58%	73.37%	90.55%
Aligned base identity (excluding indels)	81.47%	81.78%	94.60%
Identical bases per 100 aligned bases (including indels)	69.66%	68.48%	86.98%
Inserted bases per 100 aligned bases (including indels)	5.56%	5.10%	2.83%
Deleted bases per 100 aligned bases (including indels)	8.93%	11.16%	5.22%
Substitutions per 100 aligned bases (including indels)	15.85%	15.25%	4.97%
Mean insertion size	1.62	1.58	1.49
Mean deletion size	1.57	1.70	1.55
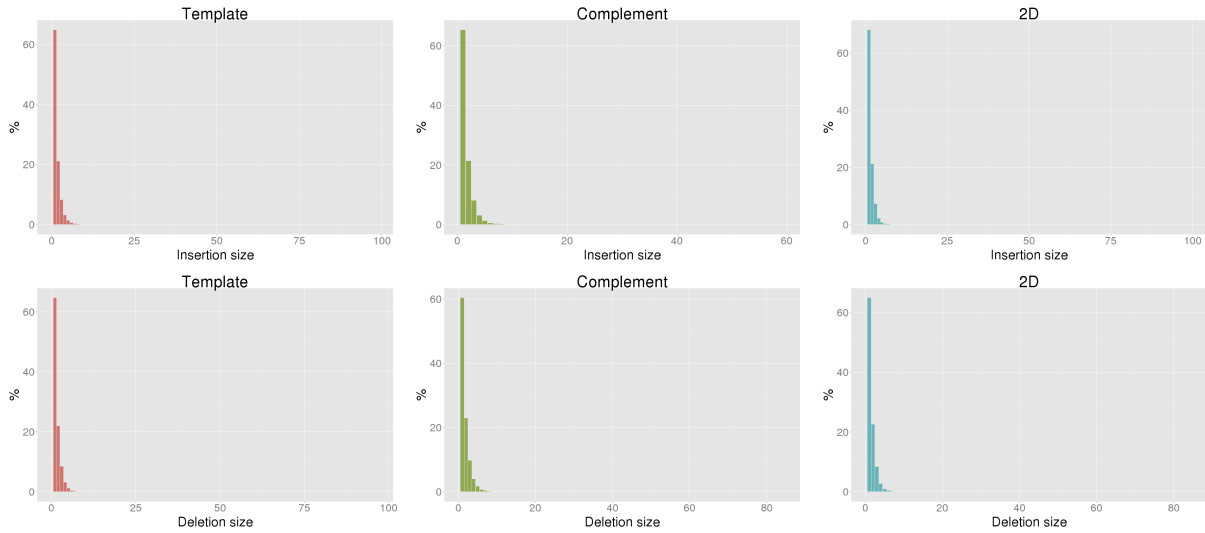
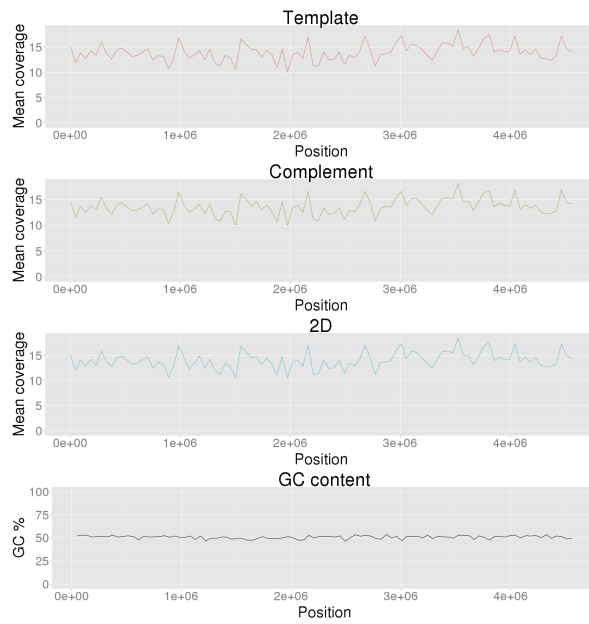# Escherichia coli read identity
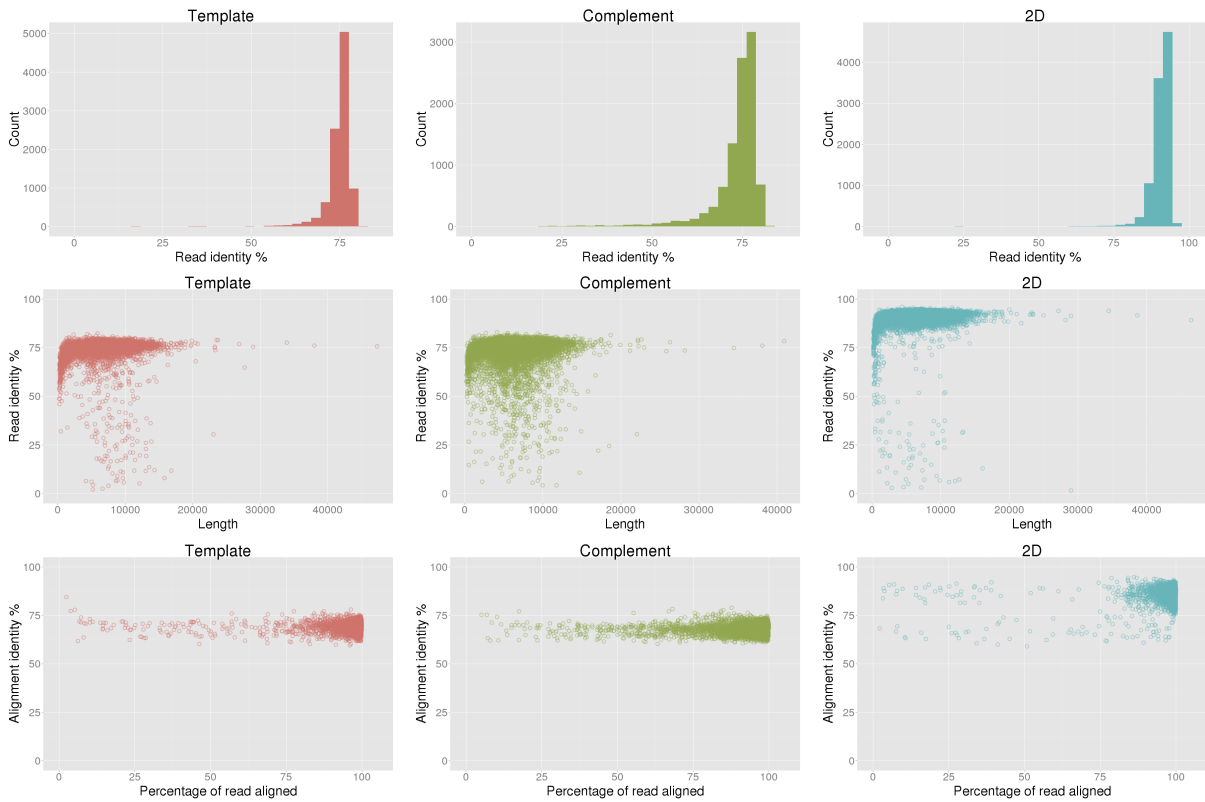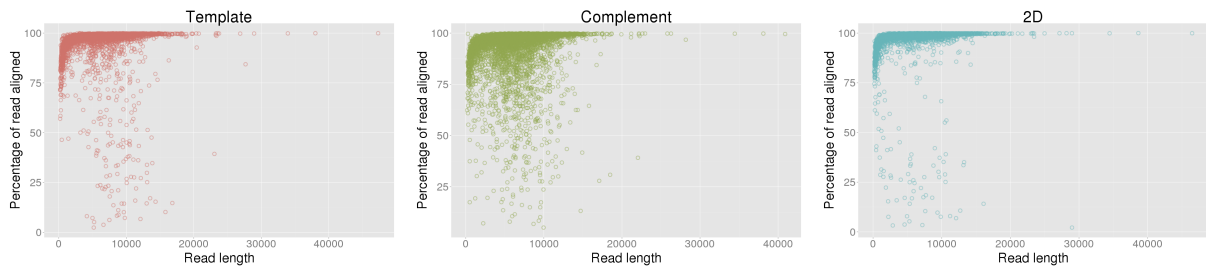
## Escherichia coli perfect kmers

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	0.104	-0.155	CGCCA	0.288	0.099	-0.189	TTTTT	0.251	0.043	-0.208
2	TTTTT	0.251	0.104	-0.146	AAAAA	0.247	0.071	-0.176	AAAAA	0.247	0.054	-0.194
3	AAAAA	0.247	0.101	-0.146	TTTTT	0.251	0.083	-0.168	CGCCA	0.288	0.219	-0.069
4	GCTGG	0.279	0.148	-0.131	CCAGC	0.289	0.158	-0.130	CAAAA	0.169	0.107	-0.063
5	CGCCA	0.288	0.167	-0.121	CACCA	0.184	0.054	-0.130	AAAAT	0.195	0.134	-0.061
6	CCAGC	0.289	0.170	-0.118	CGCTG	0.259	0.130	-0.129	GCCAG	0.280	0.219	-0.061
7	GCCAG	0.280	0.167	-0.112	CAGCA	0.261	0.134	-0.127	CGCTG	0.259	0.198	-0.061
8	CTGGC	0.278	0.175	-0.103	GCCAG	0.280	0.156	-0.123	GCTGG	0.279	0.223	-0.057
9	CAGCA	0.261	0.163	-0.098	CTGGC	0.278	0.158	-0.120	TGGTG	0.185	0.129	-0.056
10	TGGCG	0.275	0.180	-0.095	TGGCG	0.275	0.164	-0.112	AAAAG	0.132	0.076	-0.056

Over-represented 5-mers

Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	CCCCG	0.055	0.150	0.095	ACCCC	0.040	0.155	0.115	CAAAT	0.105	0.155	0.051
2	ACCCC	0.040	0.135	0.095	CCCCG	0.055	0.146	0.091	GGGGT	0.039	0.069	0.029
3	CCCCC	0.033	0.127	0.094	CCCCA	0.064	0.141	0.077	CCCAA	0.047	0.076	0.029
4	CCCCA	0.064	0.135	0.072	GAGAG	0.046	0.116	0.071	GGGTC	0.040	0.067	0.027
5	GCCCC	0.062	0.130	0.068	TACCC	0.073	0.142	0.069	TGAAT	0.121	0.148	0.027
6	CCTAG	0.003	0.071	0.068	TCCCC	0.056	0.123	0.067	CTCGT	0.042	0.068	0.026
7	CTCCC	0.040	0.104	0.064	CCTAA	0.026	0.091	0.065	CGTGA	0.102	0.127	0.026
8	GAGAG	0.046	0.110	0.064	CCTAG	0.003	0.067	0.065	CGGGG	0.054	0.080	0.025
9	TCTAC	0.048	0.110	0.063	CCCCT	0.039	0.104	0.064	AGGCA	0.093	0.118	0.025
10	TACCC	0.073	0.135	0.061	GACCC	0.040	0.101	0.061	ACCGT	0.123	0.148	0.025
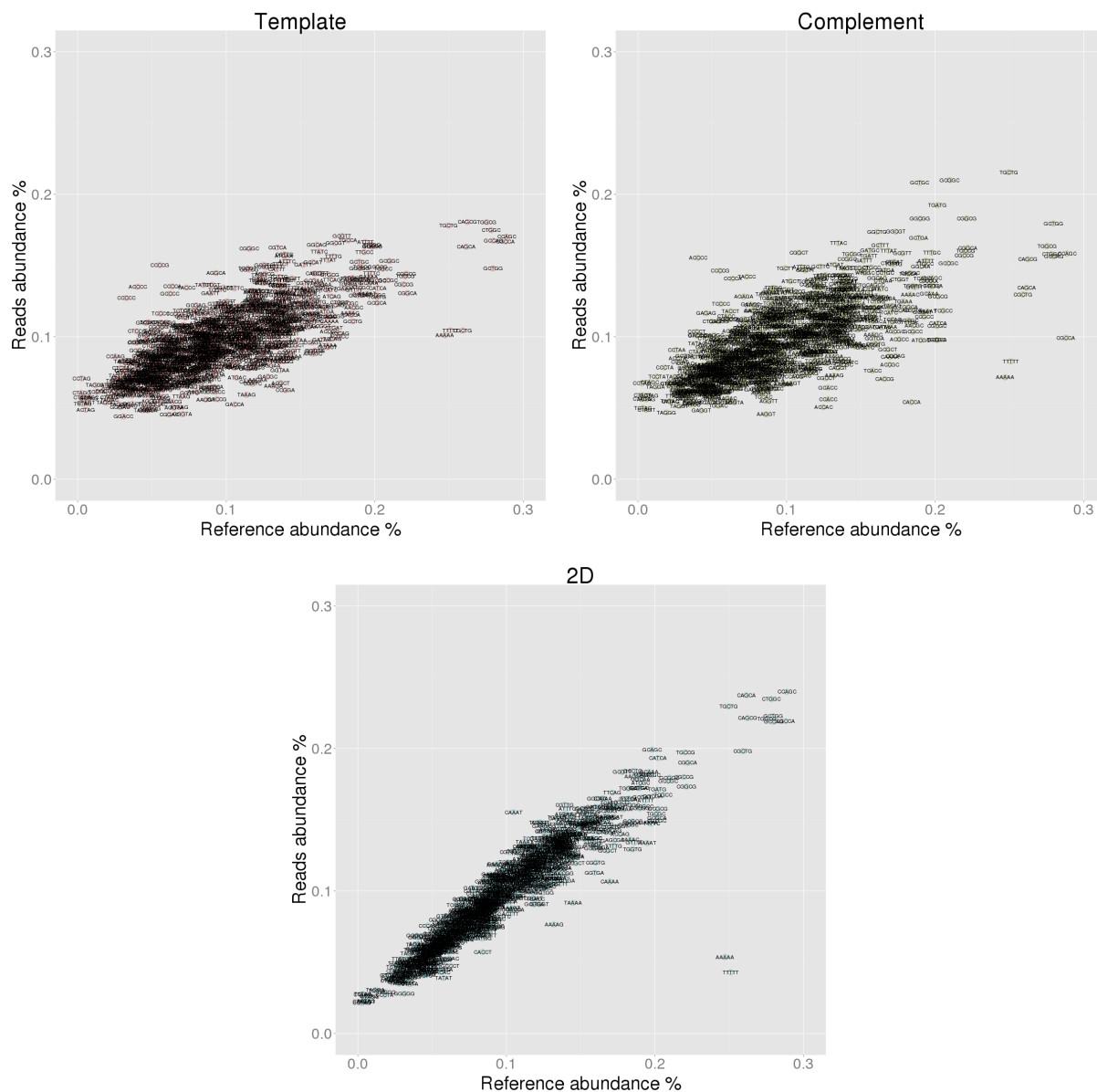
Escherichia coli GC content
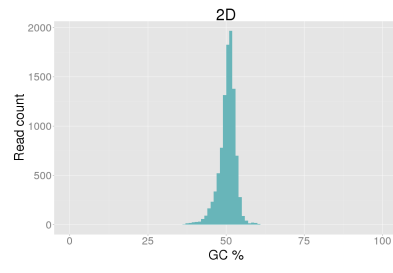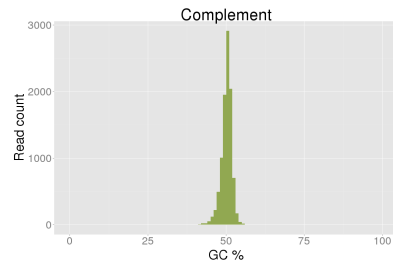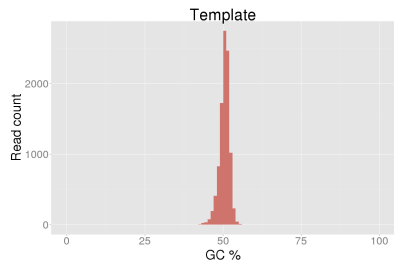
All reference 21mer analysis
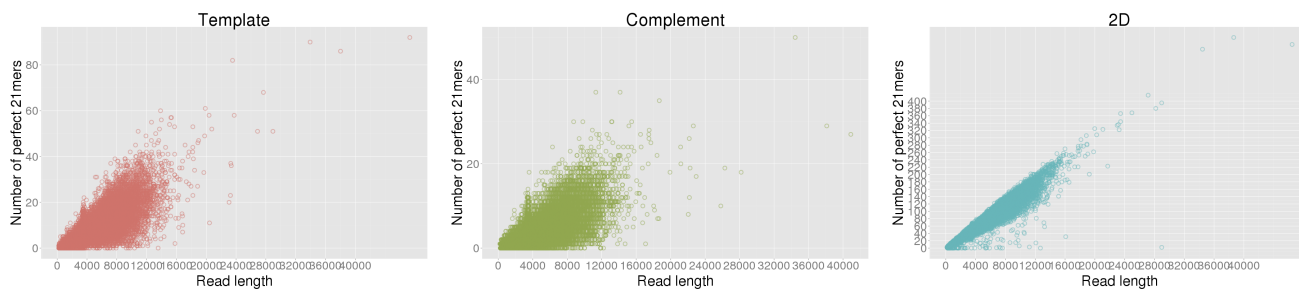
All reference substitutions

		Template substituted %				Complement substituted %				2D substituted %			
		a	c	g	t	a	c	g	t	a	c	g	t
Reference	A	0.00	8.90	9.50	4.67	0.00	9.19	8.90	4.92	0.00	8.40	8.79	3.96
	C	8.42	0.00	9.14	9.71	9.16	0.00	8.64	9.50	9.27	0.00	10.98	9.12
	G	9.20	9.11	0.00	8.02	8.83	8.78	0.00	8.77	8.95	11.09	0.00	8.62
	T	4.97	9.75	8.62	0.00	5.20	9.08	9.02	0.00	4.07	8.63	8.13	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 (3.26%)	TTC (3.53%)	AAA (4.17%)	TGC (2.81%)	TGC (2.91%)	AAA (4.24%)	GCA (3.05%)	TGC (2.60%)	AAA (3.98%)	
2	AAA (2.85%)	TGC (3.04%)	TTC (3.66%)	AAA (2.72%)	GGC (2.90%)	GCA (3.53%)	TTC (2.78%)	TCA (2.58%)	GCA (3.69%)	Most common
3	GCA (2.77%)	GCA (2.84%)	GCA (3.26%)	TTC (2.69%)	AAA (2.85%)	GAA (3.32%)	AAA (2.70%)	AAA (2.54%)	GAA (3.29%)	
4	TGC (2.71%)	AAA (2.69%)	GAA (3.04%)	GCA (2.69%)	TTC (2.61%)	TTC (3.10%)	TCA (2.65%)	GGC (2.50%)	TTC (2.90%)	
5	ATC (2.54%)	TCA (2.47%)	TTT (2.54%)	CAG (2.46%)	GCA (2.58%)	TTT (2.64%)	ATC (2.46%)	GCA (2.46%)	TTT (2.77%)	
6	TCA (2.54%)	GCC (2.43%)	TGC (2.49%)	GGC (2.41%)	GAA (2.34%)	TGC (2.48%)	TGC (2.40%)	GCG (2.45%)	AAT (2.44%)	
7	GCC (2.33%)	GGC (2.35%)	AAT (2.37%)	TCA (2.34%)	TCA (2.31%)	TCA (2.40%)	GAA (2.34%)	TTC (2.45%)	GCC (2.42%)	
8	GGC (2.24%)	ATC (2.28%)	TCA (2.30%)	ATC (2.31%)	GCC (2.27%)	ATC (2.32%)	CGC (2.24%)	ATC (2.24%)	TCA (2.37%)	
9	GAA (2.23%)	AAC (2.19%)	GCC (2.26%)	GAA (2.30%)	CAG (2.22%)	GCC (2.19%)	CAG (2.21%)	CGC (2.22%)	GCG (2.34%)	
10	TTT (2.21%)	GAA (2.16%)	CAA (2.22%)	TTT (2.15%)	AAT (2.17%)	CAA (2.10%)	GCG (2.17%)	CAG (2.13%)	ATC (2.27%)	
	TTC AAA	TTC AAA	TTC AAA	TTC AAA	TTC AAA	TTC AAA	TTC AAA	TTC AAA	TTT AAA	
-10	TGT (0.99%)	AGG (0.95%)	GGT (0.92%)	AGA (1.03%)	CTT (0.95%)	CTC (0.97%)	CTC (1.07%)	GTA (0.98%)	GTA (0.89%)	
-9	AGT (0.96%)	CCT (0.94%)	GGG (0.91%)	AGT (1.00%)	CCC (0.95%)	CTT (0.90%)	CTT (1.03%)	CCC (0.95%)	CCT (0.89%)	
-8	CCC (0.96%)	AGT (0.92%)	AGA (0.87%)	CTC (0.95%)	CGA (0.93%)	AGT (0.87%)	GAG (0.88%)	ACT (0.93%)	TGA (0.85%)	
-7	AGA (0.85%)	GAG (0.82%)	AGT (0.83%)	CCC (0.87%)	GGA (0.92%)	AGG (0.83%)	AGG (0.82%)	CGA (0.91%)	CTT (0.75%)	
-6	GGA (0.80%)	CTT (0.81%)	AGG (0.81%)	GAG (0.86%)	CTC (0.91%)	CCT (0.81%)	CCC (0.81%)	CTT (0.86%)	GAG (0.74%)	Least common
-5	GAG (0.75%)	CGA (0.74%)	TGT (0.78%)	GGA (0.83%)	CCT (0.87%)	GGG (0.81%)	AGA (0.81%)	CCT (0.84%)	CGA (0.71%)	
-4	AGG (0.69%)	AGA (0.67%)	CTT (0.73%)	AGG (0.71%)	GAG (0.86%)	ACT (0.76%)	GGA (0.77%)	AGA (0.79%)	AGA (0.52%)	
-3	GGG (0.68%)	GGA (0.64%)	GAG (0.64%)	CTA (0.66%)	GGG (0.79%)	GAG (0.64%)	GGG (0.73%)	GGA (0.73%)	GGA (0.50%)	
-2	CTA (0.53%)	TAG (0.55%)	TAG (0.41%)	GGG (0.60%)	TAG (0.56%)	CTA (0.47%)	CTA (0.64%)	CTA (0.71%)	TAG (0.49%)	
-1	TAG (0.44%)	CTA (0.53%)	CTA (0.37%)	TAG (0.48%)	CTA (0.54%)	TAG (0.39%)	TAG (0.52%)	TAG (0.65%)	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 Insertion	Template Deletion	Template Substitution	Complement Insertion	Complement Deletion	Complement Substitution	2D Insertion	2D Deletion	2D Substitution	
1	TTTC (1.01%)	TTTC (1.12%)	AAAA (1.46%)	CAGC (0.92%)	TGGC (1.05%)	AAAA (1.23%)	ATCA (0.93%)	TGGC (0.88%)	AAAA (1.17%)	Most common
2	GAAA (0.92%)	TGCC (0.95%)	TTTC (1.20%)	ATCA (0.89%)	CGGC (1.03%)	CAAA (1.08%)	GGCA (0.81%)	TTCA (0.85%)	GGCA (1.14%)	
3	AAAA (0.91%)	TTCA (0.92%)	GAAA (1.17%)	CTGC (0.88%)	CAGC (0.99%)	GAAA (1.01%)	CAGC (0.80%)	CAGC (0.85%)	GAAA (0.99%)	
4	ATCA (0.88%)	TTGC (0.87%)	GGCA (0.93%)	TGGC (0.83%)	CTGC (0.90%)	TGAA (0.97%)	GCCA (0.80%)	ATCA (0.85%)	TGAA (0.97%)	
5	TTCA (0.86%)	TGGC (0.87%)	GCAA (0.91%)	TTGC (0.81%)	TTGC (0.90%)	AGCA (0.93%)	AACA (0.78%)	CGGC (0.76%)	TTTT (0.92%)	
6	TGCC (0.84%)	CTGC (0.86%)	GTTC (0.90%)	CGGC (0.80%)	CAAA (0.82%)	ATCA (0.92%)	CGCC (0.78%)	TTGC (0.73%)	TGCA (0.91%)	
7	CAGC (0.83%)	CAGC (0.85%)	TTTT (0.89%)	AAAA (0.79%)	ATCA (0.82%)	TTTC (0.89%)	CAAA (0.77%)	CAAA (0.73%)	GGAA (0.89%)	
8	TTGC (0.80%)	AAAA (0.84%)	TGCC (0.88%)	CCAG (0.79%)	AAAA (0.80%)	AGAA (0.89%)	CGCA (0.77%)	CTGC (0.71%)	CGCC (0.88%)	
9	TTCT (0.79%)	GTTC (0.83%)	TGAA (0.85%)	CAAA (0.78%)	TGCC (0.76%)	GGCA (0.89%)	TTCA (0.76%)	GGCG (0.70%)	TGCC (0.86%)	
10	TGGC (0.77%)	TTCC (0.82%)	GGAA (0.85%)	ATGC (0.72%)	TTCA (0.76%)	TAAA (0.86%)	TGAA (0.75%)	GCCA (0.70%)	CAAA (0.86%)	
-10	TAGT (0.12%)	GGAC (0.14%)	CTAT (0.12%)	CGAG (0.12%)	CCCC (0.13%)	TTAG (0.11%)	GAGA (0.14%)	CCCC (0.15%)	TATA (0.11%)	Least common
-9	TCTA (0.12%)	TAGT (0.12%)	TAGT (0.10%)	AGGG (0.12%)	CCTC (0.13%)	ACTA (0.11%)	GGGA (0.14%)	CTAT (0.15%)	GAGA (0.11%)	
-8	AGGG (0.12%)	CTAT (0.12%)	ACTA (0.10%)	CTAA (0.11%)	CGAG (0.13%)	ACCT (0.11%)	CTAT (0.14%)	ACCT (0.14%)	TCTA (0.10%)	
-7	CTAA (0.12%)	TCTA (0.12%)	CGAG (0.10%)	TAGA (0.11%)	CTAA (0.12%)	GGAC (0.10%)	GAGG (0.14%)	GAGA (0.14%)	CGGA (0.10%)	
-6	TTAG (0.11%)	CGGA (0.11%)	GGAC (0.10%)	GAGG (0.10%)	ACCT (0.11%)	CTAT (0.10%)	CTAA (0.13%)	CTAA (0.13%)	CTAT (0.09%)	
-5	GGAC (0.11%)	CTAA (0.10%)	TAGA (0.08%)	GGAC (0.10%)	TAGA (0.11%)	CGAG (0.09%)	TCTA (0.13%)	TAGG (0.12%)	GGGA (0.09%)	
-4	TAGA (0.07%)	TAGG (0.08%)	TCTA (0.08%)	CCCT (0.10%)	CCCT (0.09%)	CCCT (0.07%)	TAGA (0.08%)	CCCT (0.11%)	TAGG (0.07%)	
-3	CCTA (0.07%)	CCTA (0.06%)	TAGG (0.06%)	CCTA (0.08%)	TAGG (0.09%)	CCTA (0.06%)	CCTA (0.08%)	CCTA (0.10%)	TAGA (0.06%)	
-2	TAGG (0.05%)	TAGA (0.06%)	CCTA (0.05%)	TAGG (0.06%)	CCTA (0.06%)	TAGG (0.06%)	TAGG (0.06%)	TAGA (0.09%)	CCTA (0.05%)	
-1	CTAG (0.01%)	CTAG (0.01%)	CTAG (0.01%)	CTAG (0.02%)	CTAG (0.02%)	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

Rank	Template Insertion	Template Deletion	Template Substitution	Complement Insertion	Complement Deletion	Complement Substitution	2D Insertion	2D Deletion	2D Substitution	
1	CTGGC (0.33%)	CAGCA (0.37%)	GAAAA (0.44%)	CAGCA (0.37%)	CTGGC (0.40%)	CAGCA (0.50%)	CAGCA (0.35%)	CTGGC (0.35%)	CAGCA (0.42%)	Most common
2	TTATC (0.33%)	CTGGC (0.36%)	CAAAA (0.39%)	GCTGC (0.32%)	CAGCA (0.38%)	GAAAA (0.38%)	CGCCA (0.34%)	CATCA (0.31%)	CGGCA (0.39%)	
3	CAGCA (0.32%)	TTGCC (0.32%)	CAGCA (0.39%)	CTGGC (0.31%)	GCGGC (0.35%)	CGGCA (0.35%)	CGGCA (0.30%)	CAGCA (0.30%)	TGGCA (0.39%)	
4	CGCCA (0.31%)	TTTGC (0.32%)	AGAAA (0.34%)	CCAGC (0.31%)	CCAGC (0.31%)	ATAAA (0.34%)	CATCA (0.29%)	CGCCA (0.29%)	GAAAA (0.38%)	
5	CAAAA (0.30%)	CATCA (0.31%)	TGTTC (0.33%)	CATCA (0.30%)	CATCA (0.30%)	CAGAA (0.34%)	CTGGC (0.28%)	GCGGC (0.29%)	CAAAA (0.34%)	
6	CATCA (0.29%)	ATTTC (0.30%)	TTGCC (0.32%)	GCGGC (0.29%)	TTTGC (0.30%)	AAGAA (0.33%)	GCAAA (0.28%)	GATGC (0.28%)	TTGCC (0.31%)	
7	ATTTC (0.28%)	TTTCA (0.30%)	TAAAA (0.32%)	TTATC (0.28%)	GCTGC (0.30%)	GCAAA (0.33%)	CAGGC (0.28%)	CCAGC (0.28%)	GCAAA (0.30%)	
8	TTTCT (0.28%)	CGCCA (0.30%)	GCAAA (0.32%)	GATGC (0.28%)	AATCA (0.29%)	AATCA (0.33%)	GCCAG (0.27%)	TATCG (0.28%)	CTGAA (0.29%)	
9	GCAGC (0.28%)	GCAGC (0.30%)	TGGCA (0.32%)	TTTGC (0.27%)	GATGC (0.29%)	AGAAA (0.32%)	TGGCA (0.27%)	TTTCA (0.28%)	CGCCA (0.28%)	
10	CCAGC (0.28%)	GCTGC (0.29%)	ATTTC (0.31%)	AATCA (0.27%)	TCAGC (0.29%)	CATCA (0.32%)	CCAGC (0.27%)	GCCAG (0.27%)	ATTTT (0.28%)	
-10	GGACC (0.01%)	GGACC (0.01%)	ACCTA (0.01%)	GGGAC (0.01%)	ACCTA (0.01%)	ACCTA (0.01%)	CCCTA (0.01%)	TCCTA (0.01%)	CCCTA (0.01%)	Least common
-9	CCCTA (0.01%)	CCCTA (0.01%)	CCCTA (0.01%)	GCTAG (0.01%)	GCTAG (0.01%)	CCCCT (0.01%)	TAGGA (0.01%)	TAGGA (0.01%)	TTAGA (0.01%)	
-8	GCTAG (0.00%)	GCTAG (0.00%)	ACTAG (0.00%)	CCCTA (0.01%)	CCCTA (0.01%)	CTAGC (0.01%)	CCTAG (0.01%)	GCTAG (0.01%)	CTAGC (0.01%)	
-7	CTAGT (0.00%)	ACTAG (0.00%)	GCTAG (0.00%)	CTAGC (0.00%)	CTAGC (0.01%)	GCTAG (0.01%)	GCTAG (0.01%)	ACTAG (0.01%)	ACTAG (0.00%)	
-6	CCTAG (0.00%)	CTAGC (0.00%)	CTAGC (0.00%)	CTAGA (0.00%)	CTAGA (0.00%)	CTAGA (0.00%)	CTAGC (0.00%)	CTAGC (0.01%)	GCTAG (0.00%)	
-5	ACTAG (0.00%)	CTAGT (0.00%)	CTAGT (0.00%)	CTAGT (0.00%)	CTAGT (0.00%)	CTAGT (0.00%)	CTAGA (0.00%)	CTAGT (0.00%)	CTAGT (0.00%)	
-4	CTAGC (0.00%)	CTAGA (0.00%)	CCTAG (0.00%)	ACTAG (0.00%)	ACTAG (0.00%)	CTAGG (0.00%)	CTAGT (0.00%)	CTAGA (0.00%)	CTAGG (0.00%)	
-3	CTAGG (0.00%)	TCTAG (0.00%)	TCTAG (0.00%)	CCTAG (0.00%)	CTAGG (0.00%)	ACTAG (0.00%)	ACTAG (0.00%)	TCTAG (0.00%)	CTAGA (0.00%)	
-2	CTAGA (0.00%)	CCTAG (0.00%)	CTAGG (0.00%)	TCTAG (0.00%)	CCTAG (0.00%)	TCTAG (0.00%)	TCTAG (0.00%)	CTAGG (0.00%)	TCTAG (0.00%)	
-1	TCTAG (0.00%)	CTAGG (0.00%)	CTAGA (0.00%)	CTAGG (0.00%)	TCTAG (0.00%)	CCTAG (0.00%)	CTAGG (0.00%)	CCTAG (0.00%)	CCTAG (0.00%)	

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