

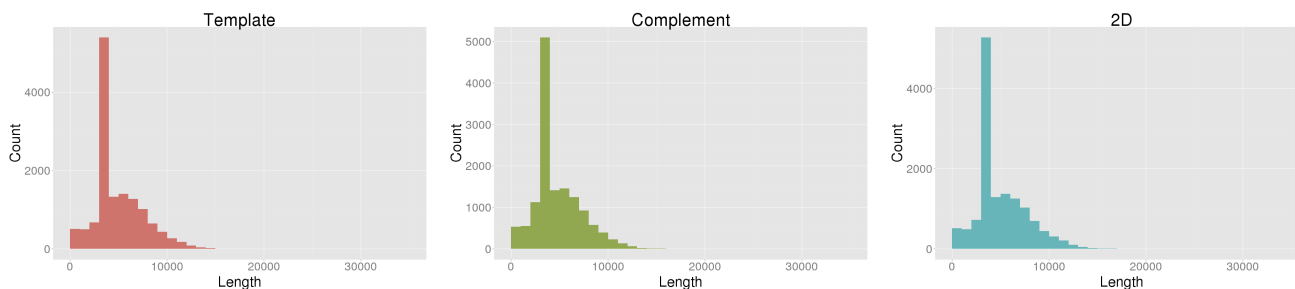
# NanoOK report for ZF-screens\_MARC\_phase\_1a\_e\_coli\_run1

## Pass and fail counts

Type	Pass	Fail
Template	13836	0
Complement	13836	0
2D	13836	0

## Read lengths

Type	NumReads	TotalBases	Mean	Longest	Shortest	N50	N50Count	N90	N90Count
Template	13836	68250103	4932.79	28699	243	5857	4211	3406	10914
Complement	13836	65002154	4698.05	27216	213	5619	4175	3132	10843
2D	13836	69426261	5017.80	29179	238	5980	4184	3444	10884



## Template alignments

Number of reads	13836
Number of reads with alignments	13722 (99.18%)
Number of reads without alignments	114 (0.82%)

ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	5014	36.24	3296.94	17451979	4902.24	81
Escherichia coli	4641652	8708	62.94	5928.92	55499569	11.96	78

## Complement alignments

Number of reads	13836
Number of reads with alignments	13716 (99.13%)
Number of reads without alignments	120 (0.87%)

ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	5001	36.14	3090.12	15651991	4396.63	62
Escherichia coli	4641652	8715	62.99	5674.42	53593333	11.55	67

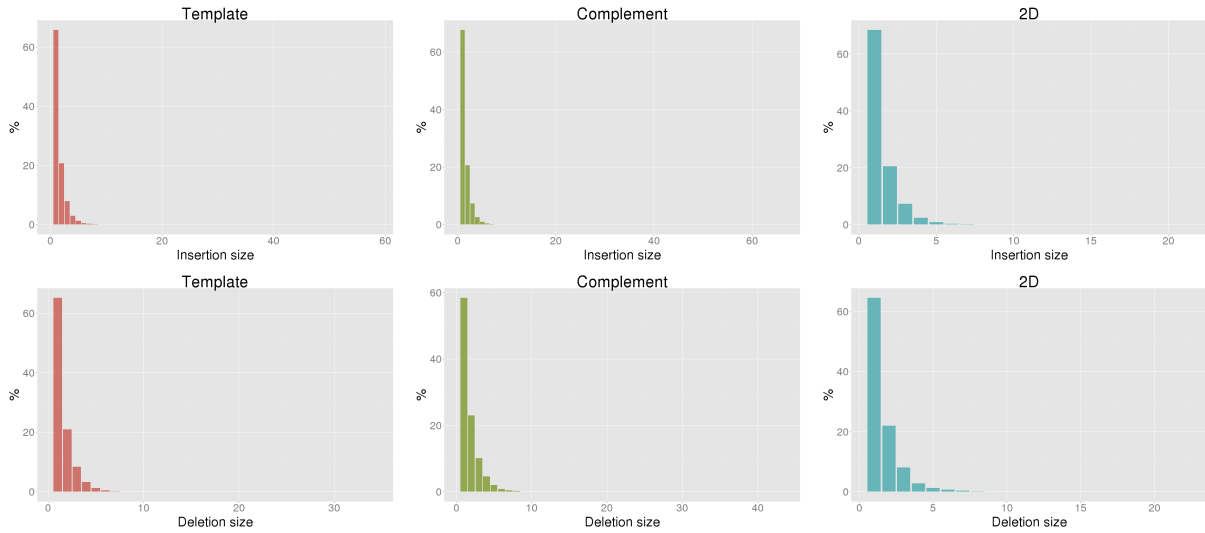
## 2D alignments

Number of reads	13836
Number of reads with alignments	13832 (99.97%)
Number of reads without alignments	4 (0.03%)

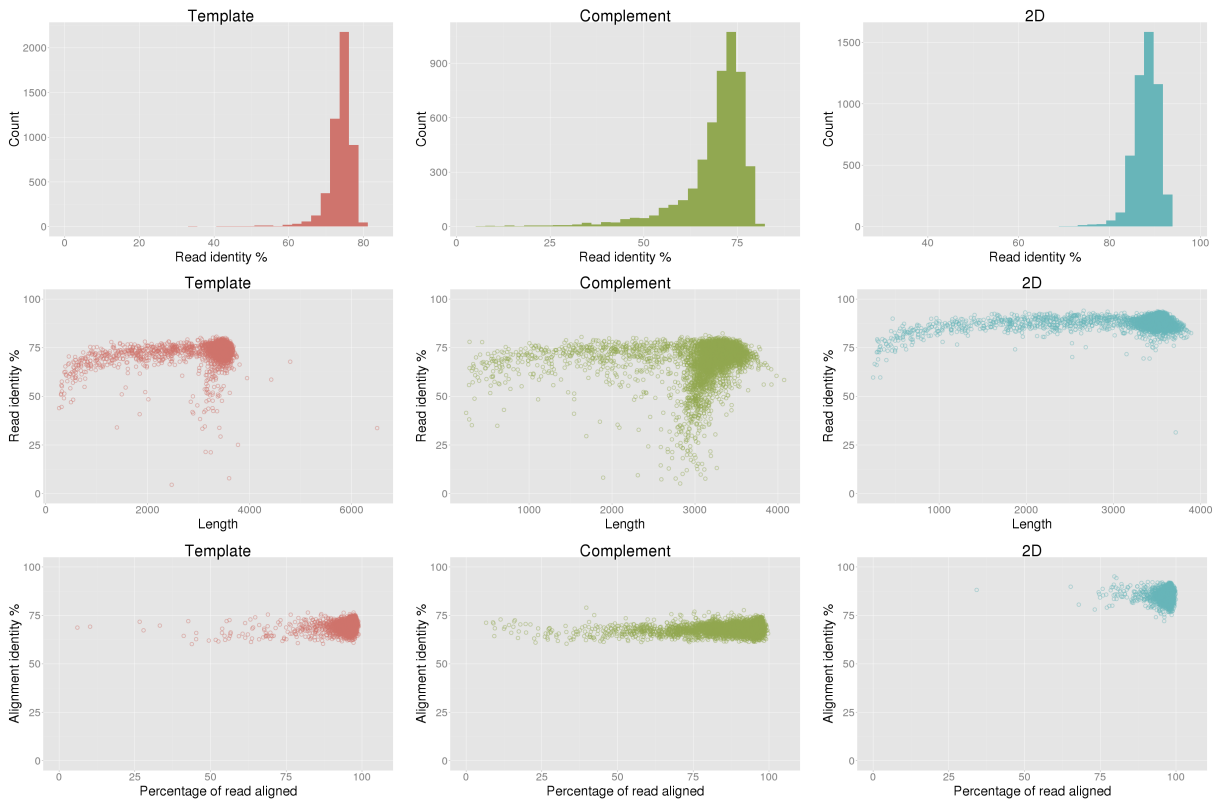
ID	Size	Number of Reads	% of Reads	Mean read length	Aligned bases	Mean coverage	Longest Perf Kmer
Control sequence	3560	5043	36.45	3305.98	17228290	4839.41	194
Escherichia coli	4641652	8789	63.52	6000.11	54767306	11.80	224

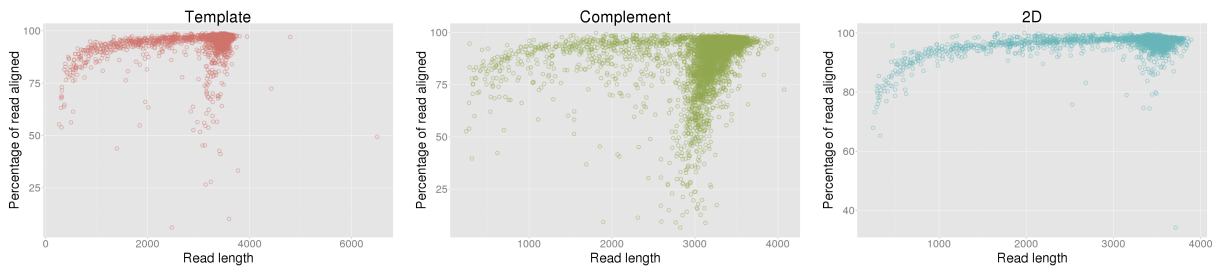
## Control sequence error analysis

	Template	Complement	2D
Overall base identity (excluding indels)	73.80%	69.03%	88.10%
Aligned base identity (excluding indels)	81.67%	81.84%	93.64%
Identical bases per 100 aligned bases (including indels)	69.91%	68.16%	85.26%
Inserted bases per 100 aligned bases (including indels)	5.24%	4.14%	3.22%
Deleted bases per 100 aligned bases (including indels)	9.16%	12.58%	5.73%
Substitutions per 100 aligned bases (including indels)	15.69%	15.12%	5.79%
Mean insertion size	1.59	1.52	1.49
Mean deletion size	1.58	1.76	1.59

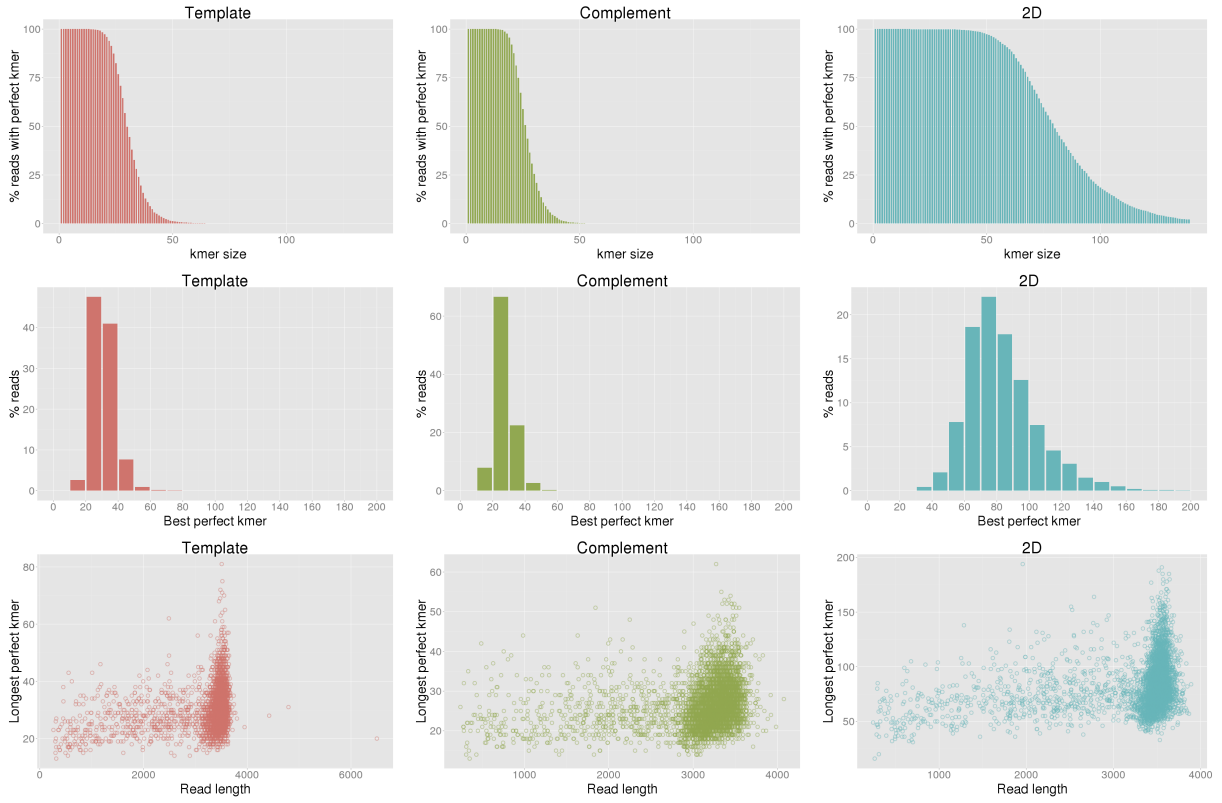


## Control sequence read identity

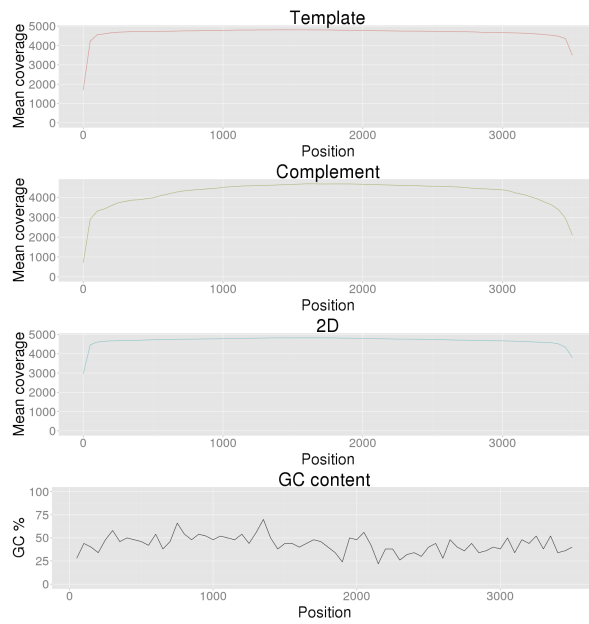




## Control sequence perfect kmers



## Control sequence coverage



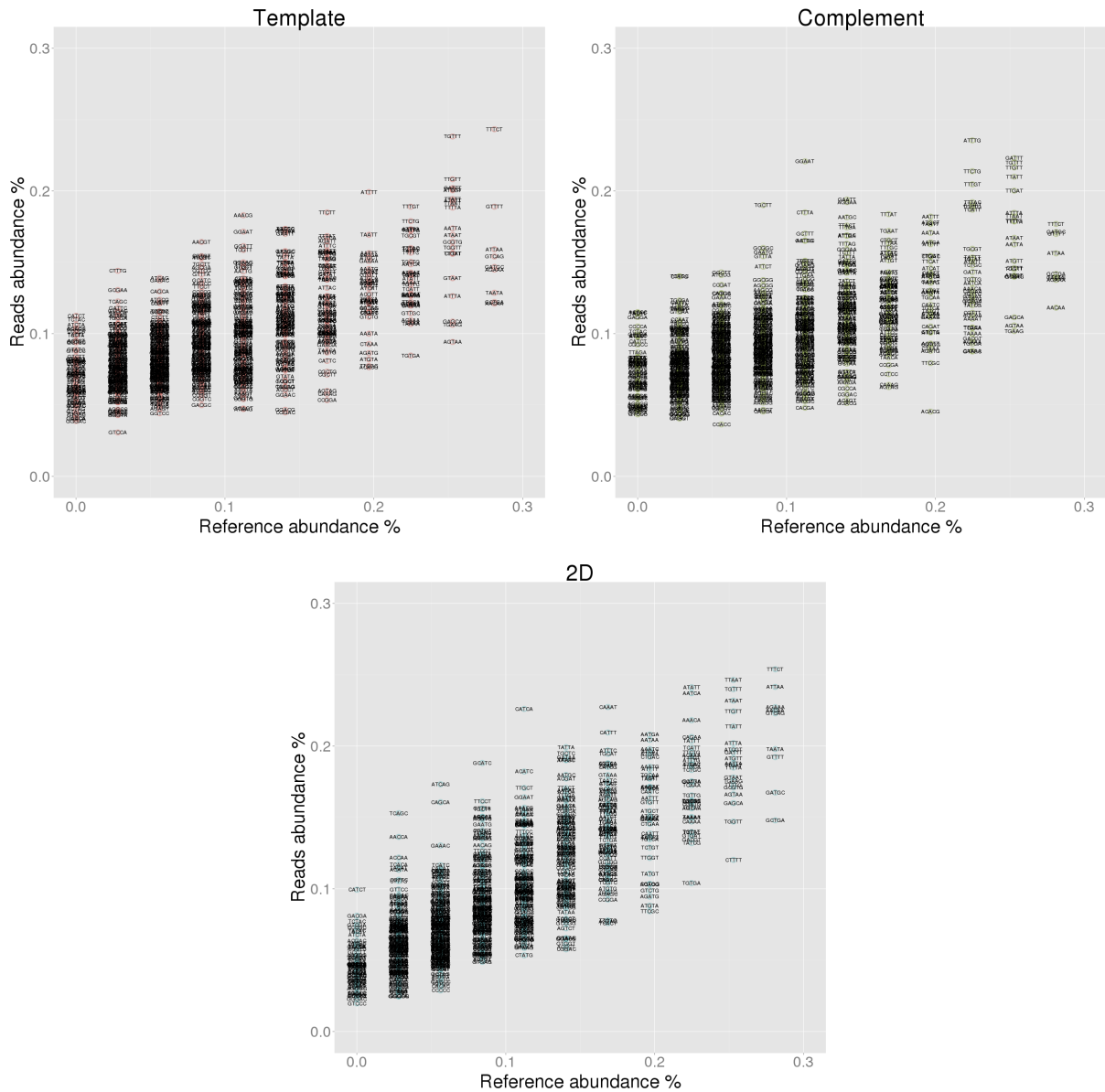
# Control sequence 5-mer analysis

## Under-represented 5-mers

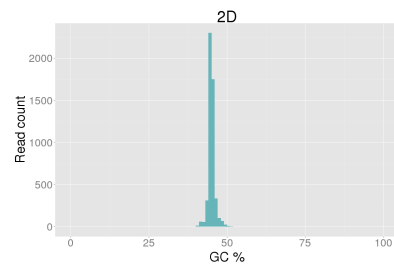
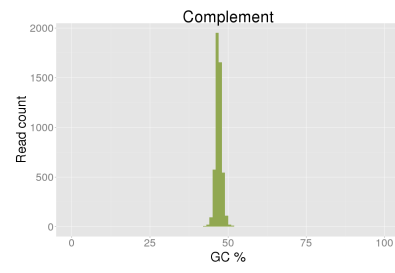
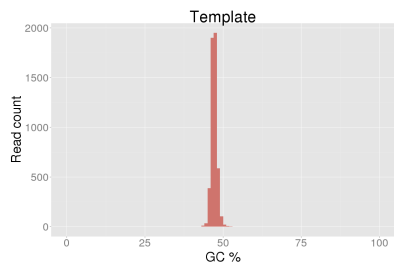
Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	TTTTT	0.759	0.152	-0.607	TTTTT	0.759	0.092	-0.667	TTTTT	0.759	0.060	-0.698
2	AAAAA	0.478	0.123	-0.355	AAAAA	0.478	0.061	-0.417	AAAAA	0.478	0.079	-0.399
3	TGATG	0.393	0.148	-0.246	AA AAC	0.337	0.127	-0.210	TGATG	0.393	0.197	-0.196
4	AAAAC	0.337	0.136	-0.201	TGATG	0.393	0.210	-0.183	CTGAT	0.309	0.153	-0.156
5	GATGT	0.309	0.112	-0.197	GATGT	0.309	0.126	-0.183	GATGT	0.309	0.157	-0.152
6	CTGAT	0.309	0.125	-0.184	GCAAT	0.309	0.144	-0.165	GCTGA	0.281	0.148	-0.133
7	AACAA	0.281	0.121	-0.160	AACAA	0.281	0.118	-0.163	CTTTT	0.253	0.120	-0.133
8	GCTGA	0.281	0.122	-0.159	TTATC	0.309	0.152	-0.157	AAAAC	0.337	0.205	-0.132
9	AGTAA	0.253	0.094	-0.159	ACACG	0.197	0.045	-0.151	TGTGA	0.225	0.104	-0.121
10	GCAAT	0.309	0.150	-0.159	TGAAG	0.253	0.102	-0.151	TTATC	0.309	0.188	-0.121

## Over-represented 5-mers

Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	CTTTG	0.028	0.144	0.116	ACCCC	0.000	0.115	0.115	TCAGC	0.028	0.153	0.125
2	CATCT	0.000	0.113	0.113	TATAC	0.000	0.115	0.115	ATCAG	0.056	0.173	0.117
3	TCTAC	0.000	0.110	0.110	TACTT	0.000	0.114	0.114	CATCA	0.112	0.226	0.114
4	ATCTA	0.000	0.106	0.106	TCAGC	0.028	0.140	0.112	AACCA	0.028	0.136	0.108
5	ACCCC	0.000	0.104	0.104	CTTTG	0.028	0.140	0.112	CAGCA	0.056	0.161	0.104
6	CGAGA	0.000	0.103	0.103	GAGGA	0.000	0.112	0.112	GCATC	0.084	0.188	0.104
7	GCGAA	0.028	0.131	0.102	GGAAT	0.112	0.221	0.108	CATCT	0.000	0.099	0.099
8	CCCCA	0.000	0.102	0.102	TGCTT	0.084	0.190	0.106	ACCAA	0.028	0.122	0.094
9	TACTT	0.000	0.099	0.099	CCCCA	0.000	0.105	0.105	TCACA	0.028	0.117	0.089
10	TCTTA	0.000	0.099	0.099	TCTAC	0.000	0.101	0.101	TACAT	0.028	0.115	0.086

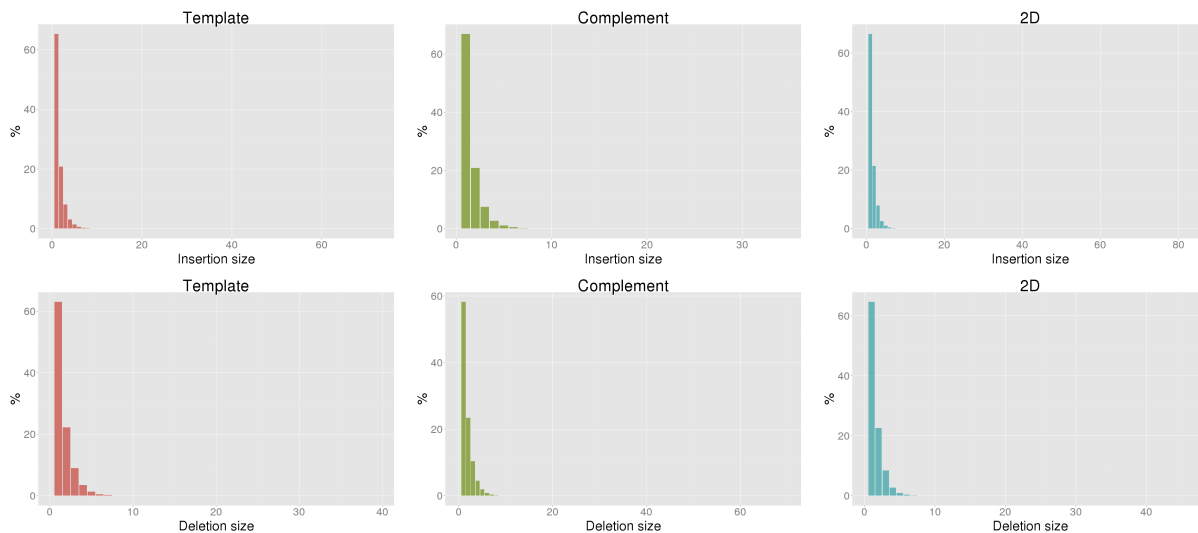


## Control sequence GC content

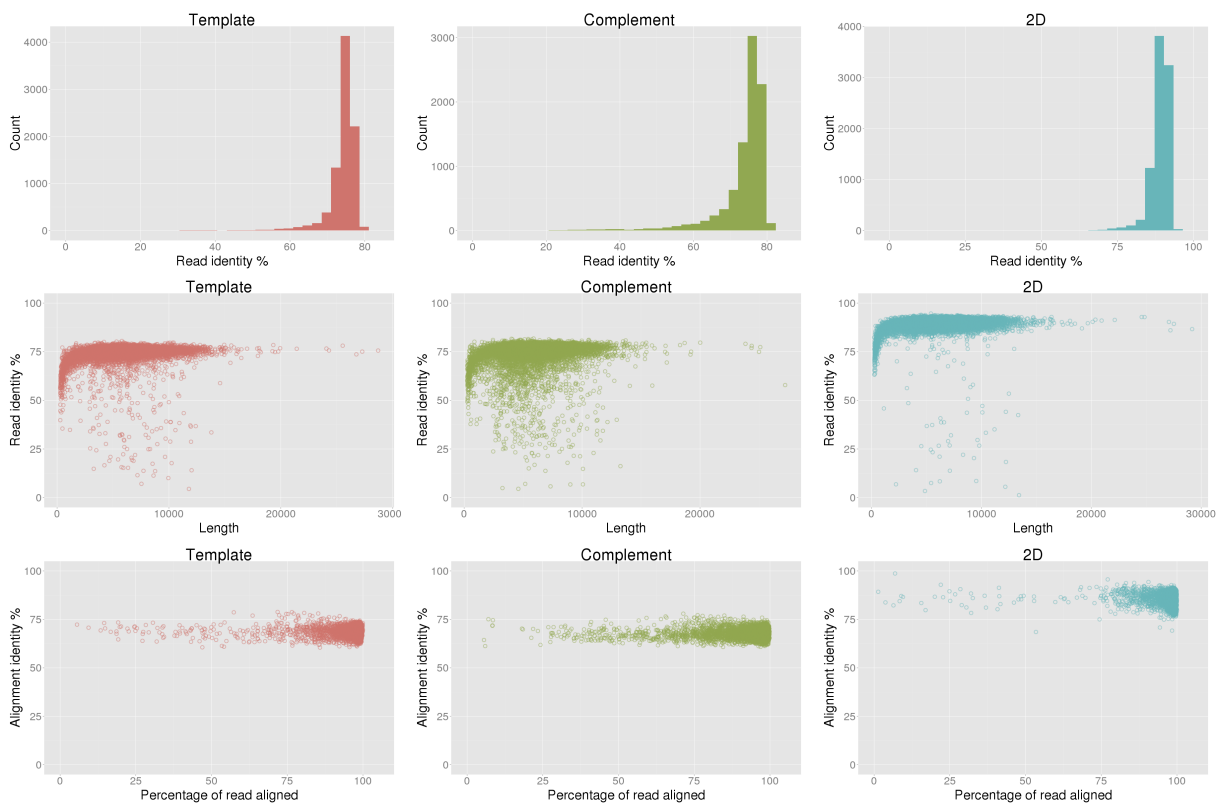


## Escherichia coli error analysis

	Template	Complement	2D
Overall base identity (excluding indels)	74.20%	73.86%	89.11%
Aligned base identity (excluding indels)	81.16%	81.92%	93.90%
Identical bases per 100 aligned bases (including indels)	69.03%	68.15%	85.80%
Inserted bases per 100 aligned bases (including indels)	5.40%	4.46%	3.35%
Deleted bases per 100 aligned bases (including indels)	9.55%	12.34%	5.27%
Substitutions per 100 aligned bases (including indels)	16.02%	15.04%	5.57%
Mean insertion size	1.61	1.54	1.53
Mean deletion size	1.61	1.75	1.55

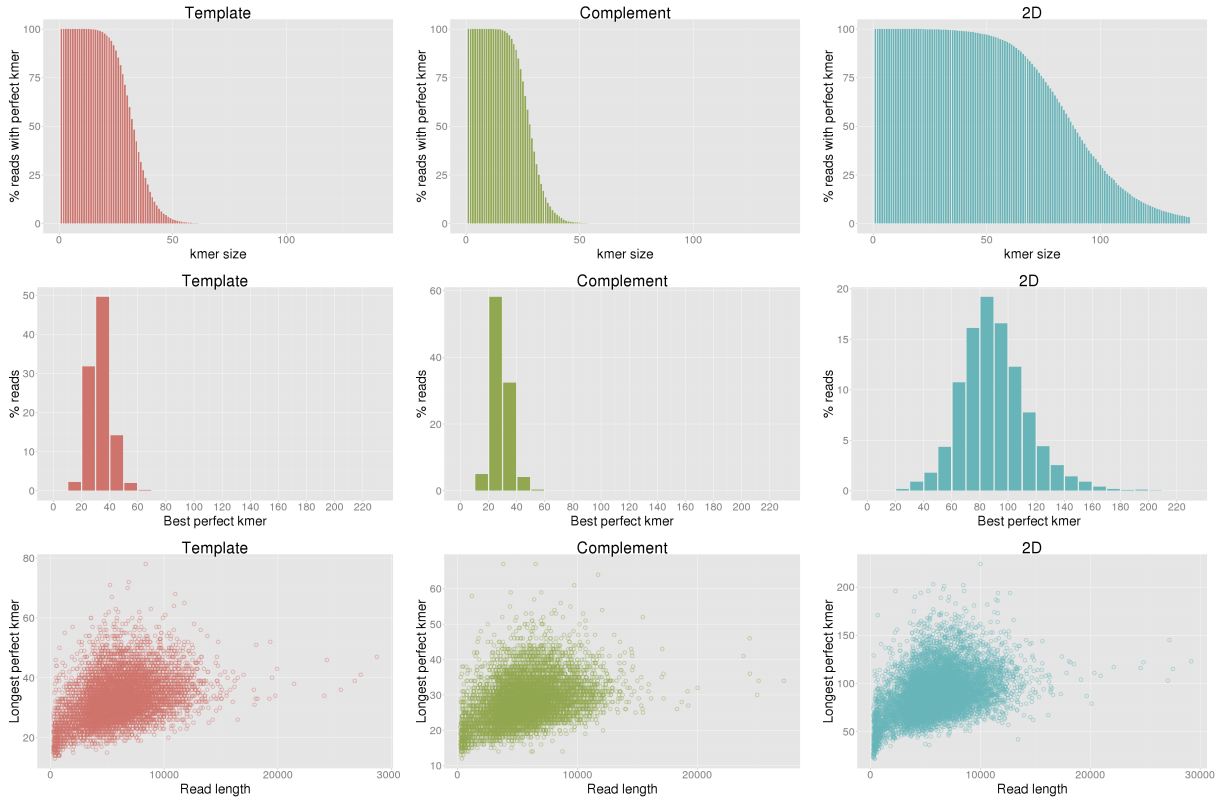


## Escherichia coli read identity

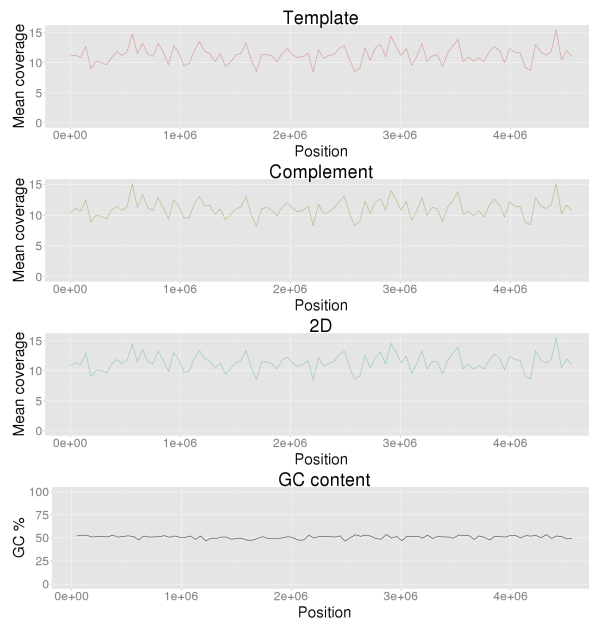




## Escherichia coli perfect kmers



## Escherichia coli coverage



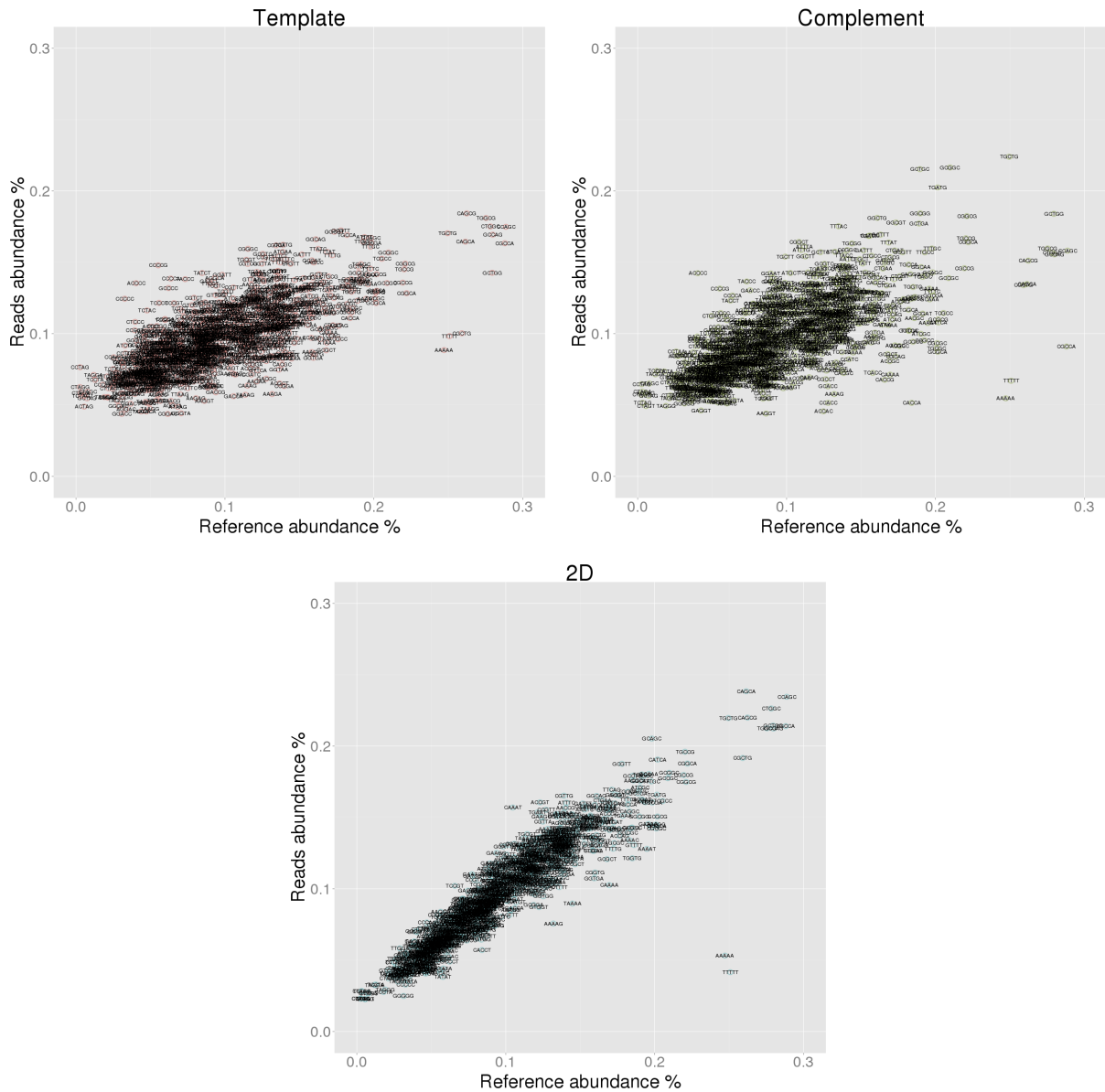
# Escherichia coli 5-mer analysis

## Under-represented 5-mers

Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	CGCTG	0.259	0.100	-0.159	CGCCA	0.288	0.091	-0.197	TTTTT	0.251	0.042	-0.209
2	AAAAA	0.247	0.089	-0.159	AAAAA	0.247	0.055	-0.192	AAAAA	0.247	0.053	-0.194
3	TTTTT	0.251	0.098	-0.153	TTTTT	0.251	0.067	-0.184	CGCCA	0.288	0.214	-0.074
4	GCTGG	0.279	0.143	-0.137	CACCA	0.184	0.052	-0.132	GCCAG	0.280	0.212	-0.068
5	CGCCA	0.288	0.163	-0.125	CCAGC	0.289	0.158	-0.131	CGCTG	0.259	0.192	-0.067
6	CCAGC	0.289	0.175	-0.114	CAGCA	0.261	0.135	-0.127	CAAAA	0.169	0.103	-0.067
7	GCCAG	0.280	0.169	-0.110	CGCTG	0.259	0.134	-0.125	AAAAT	0.195	0.128	-0.067
8	CTGGC	0.278	0.175	-0.103	GCCAG	0.280	0.156	-0.124	GCTGG	0.279	0.214	-0.065
9	CAGCA	0.261	0.165	-0.097	CTGGC	0.278	0.157	-0.122	TGGTG	0.185	0.121	-0.064
10	TGGCG	0.275	0.181	-0.094	TGGCG	0.275	0.160	-0.116	TGGCG	0.275	0.212	-0.063

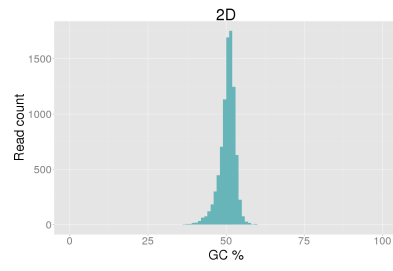
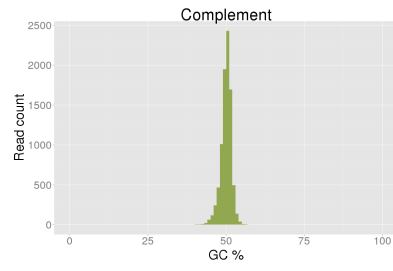
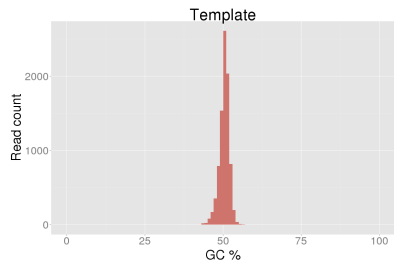
## Over-represented 5-mers

Rank	Template				Complement				2D			
	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %	kmer	Ref %	Read %	Diff %
1	ACCCC	0.040	0.135	0.095	ACCCC	0.040	0.142	0.102	CAAAT	0.105	0.157	0.052
2	CCCCG	0.055	0.148	0.093	CCCCG	0.055	0.132	0.076	ACCGT	0.123	0.161	0.038
3	CCCCC	0.033	0.125	0.092	TACCC	0.073	0.137	0.063	TCCGT	0.066	0.102	0.036
4	CCCCA	0.064	0.139	0.075	CCCCA	0.064	0.127	0.063	GAATC	0.077	0.110	0.033
5	CCTAG	0.003	0.076	0.074	CTGAG	0.050	0.112	0.063	TGAAT	0.121	0.154	0.032
6	GCCCC	0.062	0.132	0.070	CCTAG	0.003	0.065	0.062	GATTC	0.078	0.109	0.032
7	CTCCC	0.040	0.108	0.069	CCTAA	0.026	0.087	0.062	GGATT	0.098	0.130	0.031
8	TCTAC	0.048	0.116	0.069	TCCTA	0.013	0.074	0.061	GAAGG	0.094	0.125	0.031
9	TCCCC	0.056	0.122	0.066	TACCT	0.062	0.123	0.061	GGGGT	0.039	0.070	0.031
10	TACCC	0.073	0.138	0.065	TAGGA	0.012	0.071	0.060	TTGGA	0.029	0.058	0.029

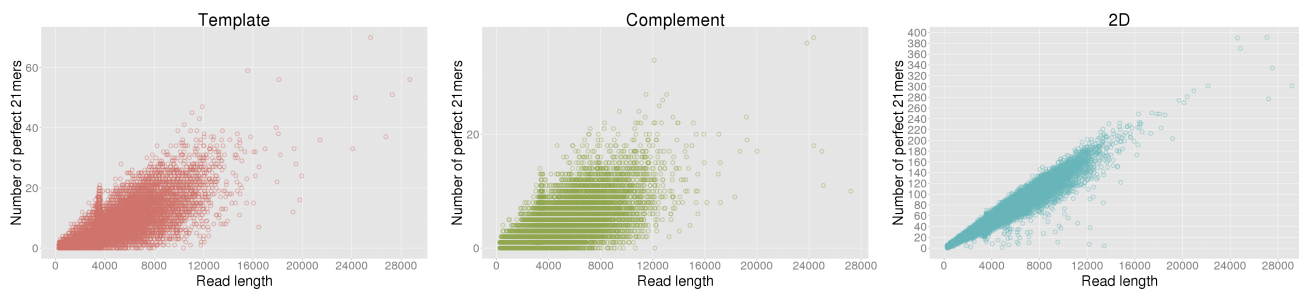




# Escherichia coli GC content



## 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	9.39	9.55	4.79	0.00	9.52	9.12	5.06	0.00	8.89	8.89	4.21
C	8.04	0.00	8.79	9.49	8.83	0.00	8.34	9.19	8.81	0.00	10.55	8.74
G	9.06	8.83	0.00	7.72	8.64	8.61	0.00	8.52	8.73	10.72	0.00	8.29
T	5.23	10.05	9.06	0.00	5.40	9.36	9.41	0.00	4.36	9.06	8.75	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	TTC (3.19%)	TTC (3.63%)	AAA (4.43%)	TGC (2.84%)	TGC (3.00%)	AAA (4.30%)	GCA (3.08%)	TGC (2.75%)	AAA (4.23%)
2	AAA (2.94%)	TGC (3.13%)	TTC (3.63%)	AAA (2.76%)	AAA (2.97%)	GCA (3.49%)	AAA (2.98%)	AAA (2.61%)	GCA (3.64%)
3	GCA (2.79%)	AAA (2.82%)	GCA (3.22%)	TTC (2.73%)	GGC (2.74%)	GAA (3.44%)	TTC (2.83%)	TCA (2.54%)	GAA (3.42%)
4	TGC (2.67%)	GCA (2.81%)	GAA (3.09%)	GCA (2.68%)	TTC (2.69%)	TTC (3.15%)	TCA (2.73%)	TTC (2.45%)	TTT (3.13%)
5	TCA (2.62%)	TCA (2.37%)	TTT (2.83%)	GAA (2.40%)	GCA (2.50%)	TTT (2.72%)	GAA (2.53%)	GCA (2.41%)	TTC (2.94%)
6	ATC (2.61%)	GCC (2.27%)	AAT (2.55%)	TCA (2.39%)	GAA (2.42%)	TCA (2.47%)	TGC (2.41%)	GCC (2.30%)	AAT (2.65%)
7	TTT (2.46%)	GAA (2.23%)	TCA (2.41%)	ATC (2.35%)	AAT (2.31%)	TGC (2.44%)	ATC (2.35%)	GAA (2.29%)	TCA (2.45%)
8	GAA (2.32%)	ATC (2.23%)	TGC (2.39%)	TTT (2.31%)	TCA (2.29%)	ATC (2.27%)	AAT (2.31%)	ATC (2.27%)	GTT (2.43%)
9	AAT (2.22%)	GGC (2.22%)	CAA (2.38%)	CAG (2.30%)	CAG (2.22%)	AAT (2.25%)	TTT (2.14%)	GGC (2.27%)	GCC (2.26%)
10	GCC (2.09%)	AAT (2.18%)	GCC (2.17%)	GGC (2.26%)	ATC (2.16%)	CAA (2.18%)	CAG (2.08%)	CAG (2.19%)	ATC (2.14%)
-10	AGT (0.96%)	GGG (0.98%)	GGG (0.88%)	CGA (0.99%)	CCC (0.95%)	CTC (0.93%)	CTT (1.03%)	GGG (1.06%)	CCT (0.90%)
-9	CGA (0.94%)	GTA (0.93%)	GGT (0.86%)	AGA (0.99%)	AGG (0.95%)	AGT (0.89%)	TAC (1.02%)	CCT (0.96%)	GTA (0.89%)
-8	CCC (0.91%)	AGT (0.89%)	AGG (0.85%)	CTC (0.89%)	CTC (0.93%)	CTT (0.87%)	GAG (0.92%)	GTA (0.94%)	ACT (0.85%)
-7	GGA (0.85%)	GAG (0.82%)	AGA (0.85%)	GGA (0.85%)	ACT (0.91%)	AGG (0.82%)	AGA (0.83%)	ACT (0.88%)	GAG (0.78%)
-6	AGA (0.84%)	CTT (0.79%)	TGT (0.84%)	GAG (0.85%)	CGA (0.90%)	CCT (0.81%)	CCC (0.79%)	CGA (0.83%)	CTT (0.76%)
-5	GAG (0.78%)	CGA (0.72%)	AGT (0.79%)	CCC (0.84%)	CCT (0.89%)	GGG (0.75%)	AGG (0.78%)	AGA (0.82%)	CGA (0.64%)
-4	GGG (0.72%)	AGA (0.71%)	CTT (0.74%)	AGG (0.72%)	GAG (0.86%)	ACT (0.73%)	GGA (0.78%)	CTT (0.80%)	TAG (0.54%)
-3	AGG (0.67%)	GGA (0.67%)	GAG (0.70%)	CTA (0.67%)	GGG (0.74%)	GAG (0.65%)	GGG (0.74%)	GGA (0.77%)	GGA (0.53%)
-2	CTA (0.55%)	TAG (0.63%)	TAG (0.43%)	GGG (0.54%)	TAG (0.61%)	CTA (0.46%)	CTA (0.63%)	CTA (0.68%)	AGA (0.53%)
-1	TAG (0.45%)	CTA (0.55%)	CTA (0.38%)	TAG (0.50%)	CTA (0.51%)	TAG (0.41%)	TAG (0.60%)	TAG (0.67%)	CTA (0.43%)

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	TTTC (0.97%)	TTTC (1.24%)	AAAA (1.62%)	CTGC (0.90%)	TGGC (1.01%)	CAAA (1.20%)	CAAA (0.96%)	ATCA (0.88%)	AAAA (1.27%)
2	GAAA (0.95%)	TTGC (0.90%)	TTTT (1.24%)	ATCA (0.88%)	CGGC (0.95%)	AAAA (1.19%)	ATCA (0.91%)	GGAA (0.85%)	GGCA (1.09%)
3	AAAA (0.92%)	TGCC (0.88%)	GAAA (1.20%)	CAGC (0.84%)	CAAA (0.95%)	TGAA (1.03%)	TGAA (0.82%)	TGGC (0.83%)	TTTT (1.06%)
4	TTTT (0.91%)	CTGC (0.87%)	TTTT (1.05%)	TTGC (0.82%)	TTGC (0.95%)	AGAA (0.98%)	GGCA (0.80%)	ATGC (0.83%)	GAAA (1.03%)
5	TTCA (0.88%)	TTCA (0.86%)	GCAA (0.97%)	CAAA (0.80%)	CTGC (0.94%)	GGAA (0.98%)	TTCA (0.78%)	TTCA (0.82%)	TGAA (0.99%)
6	ATCA (0.88%)	TGCC (0.86%)	GTTC (0.89%)	TGGC (0.79%)	CAGC (0.92%)	AGCA (0.97%)	AACA (0.78%)	CAAA (0.80%)	TGCA (0.98%)
7	TTGC (0.80%)	AAAA (0.85%)	TGCA (0.88%)	CGGC (0.75%)	ATCA (0.79%)	ATCA (0.93%)	TGCA (0.77%)	AGCA (0.77%)	GGAA (0.96%)
8	TTCT (0.79%)	GTTC (0.83%)	TGAA (0.88%)	AAAA (0.74%)	ATGC (0.79%)	TAAA (0.89%)	GAAA (0.75%)	CAGC (0.77%)	ATCA (0.91%)
9	CTGC (0.77%)	GAAA (0.83%)	GTTT (0.88%)	CCAG (0.74%)	AGCA (0.77%)	TTTC (0.89%)	CGCA (0.74%)	TTGC (0.76%)	CAAA (0.90%)
10	GTTT (0.76%)	TTCC (0.82%)	GGCA (0.87%)	ATTC (0.73%)	AAAA (0.76%)	GGAA (0.86%)	TTTC (0.74%)	ATGC (0.76%)	TTTC (0.89%)
-10	AGGG (0.13%)	TAGT (0.12%)	TAGA (0.12%)	TAGA (0.12%)	CCTC (0.12%)	GTGT (0.11%)	CGAG (0.14%)	CCGA (0.15%)	ACTA (0.10%)
-9	TCTA (0.13%)	CTAT (0.11%)	CTAT (0.11%)	GGGG (0.12%)	TAGA (0.12%)	GGAC (0.11%)	GAGG (0.14%)	ACCT (0.14%)	TCTA (0.10%)
-8	CGAG (0.13%)	TAGG (0.11%)	TAGT (0.11%)	CCTA (0.11%)	CCCT (0.12%)	ACCT (0.10%)	CTTG (0.14%)	GAGA (0.14%)	TCGA (0.10%)
-7	TTAG (0.12%)	CTAA (0.11%)	ACTA (0.09%)	GGAC (0.11%)	ACTA (0.11%)	ACTA (0.10%)	CCCC (0.14%)	CTAA (0.13%)	GAGA (0.10%)
-6	GAGG (0.12%)	GGAC (0.11%)	CGAG (0.09%)	GAGG (0.11%)	TAGG (0.11%)	CGAG (0.09%)	CTAT (0.13%)	CCCC (0.13%)	CGGA (0.10%)
-5	GGAC (0.10%)	TCTA (0.11%)	TCTA (0.09%)	CGAG (0.11%)	CCCC (0.10%)	CTAT (0.08%)	TCTA (0.12%)	CTAT (0.13%)	CTAT (0.09%)
-4	CCTA (0.08%)	CGGA (0.11%)	GGAC (0.09%)	AGGG (0.10%)	ACCT (0.10%)	CCTA (0.07%)	TAGA (0.10%)	CCTA (0.13%)	TAGG (0.07%)
-3	TAGA (0.07%)	CCTA (0.08%)	TAGG (0.06%)	CCCT (0.10%)	CGAG (0.10%)	CCCT (0.06%)	CCTA (0.09%)	CCCT (0.13%)	TAGA (0.06%)
-2	TAGG (0.05%)	TAGA (0.06%)	CCTA (0.04%)	TAGG (0.07%)	CCTA (0.06%)	TAGG (0.06%)	TAGG (0.05%)	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 (0.04%)	CTAG (0.02%)	CTAG (0.01%)

Most common

Least common

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	TTATC (0.36%)	GATGC (0.38%)	GAAAA (0.52%)	CAGCA (0.35%)	CTGGC (0.40%)	CAGCA (0.46%)	TCAAA (0.34%)	TATCG (0.48%)	TGGCA (0.39%)
2	CGTTT (0.34%)	CTGGC (0.38%)	CAAAA (0.40%)	GCTGC (0.31%)	ACAAA (0.38%)	ACAAA (0.40%)	CAGCA (0.31%)	GATGC (0.43%)	GAAAA (0.38%)
3	TTTCT (0.31%)	ATTTT (0.37%)	AGAAA (0.39%)	GATGC (0.30%)	GATGC (0.37%)	AAGAA (0.40%)	CGCCA (0.30%)	ACAAA (0.36%)	CGGCA (0.38%)
4	CTGGC (0.31%)	CAGCA (0.36%)	TAAAA (0.36%)	AATCA (0.30%)	CAGCA (0.36%)	GAAAA (0.40%)	GCAAA (0.30%)	CTGGC (0.36%)	CAGCA (0.37%)
5	CAAAA (0.30%)	GTTTC (0.35%)	CAGCA (0.35%)	TTATC (0.29%)	GCGGC (0.33%)	CAGAA (0.38%)	TGTTC (0.30%)	TGGAA (0.35%)	CAAAA (0.36%)
6	GTTGC (0.30%)	TATCG (0.35%)	CGTTT (0.34%)	CTGGC (0.28%)	TGATG (0.33%)	AATCA (0.36%)	GCAAT (0.29%)	CATCA (0.34%)	GCAAT (0.34%)
7	AGAAA (0.30%)	TGTTT (0.33%)	TGTTT (0.33%)	CATCA (0.28%)	ATTGC (0.32%)	AGAAA (0.34%)	CGGCA (0.28%)	GCGAC (0.31%)	AATGG (0.34%)
8	CAGCA (0.30%)	GCCAG (0.33%)	CTGAA (0.33%)	CCAGC (0.27%)	GCAGC (0.30%)	TTATC (0.33%)	CGTCA (0.28%)	TTATT (0.31%)	CTGAA (0.33%)
9	GAAAA (0.29%)	GCAGC (0.32%)	TTTTT (0.33%)	GCGGC (0.27%)	TTATC (0.30%)	ATGAA (0.32%)	AATCA (0.28%)	GCCAG (0.30%)	AATCA (0.32%)
10	GTTTT (0.28%)	TTTGC (0.32%)	ATTTT (0.32%)	ACAAA (0.26%)	AATCA (0.30%)	ATAAA (0.32%)	TTGCA (0.27%)	TTTCA (0.30%)	CAGAA (0.31%)
-10	CCCTA (0.01%)	CTTAG (0.01%)	CCCTA (0.01%)	CCCTA (0.01%)	ACCTA (0.01%)	CTTAG (0.01%)	GGACC (0.01%)	TCCTA (0.01%)	GCTAG (0.01%)
-9	GCTAG (0.00%)	TCCTA (0.01%)	GCTAG (0.01%)	CTAGA (0.01%)	CCCCC (0.01%)	TAGGG (0.01%)	CTTAG (0.01%)	GCTAG (0.01%)	CCCTA (0.01%)
-8	CTAGT (0.00%)	GCTAG (0.01%)	CTTAG (0.00%)	TAGGG (0.01%)	CTAGA (0.01%)	GCTAG (0.01%)	TAGGG (0.01%)	TAGGA (0.01%)	TTAGA (0.00%)
-7	GGACC (0.00%)	ACTAG (0.00%)	ACTAG (0.00%)	CTAGC (0.01%)	CTAGC (0.01%)	CTAGC (0.01%)	CCCTA (0.01%)	CTAGA (0.01%)	CCTAG (0.00%)
-6	CCTAG (0.00%)	CTAGA (0.00%)	CTAGG (0.00%)	GGGGG (0.01%)	CCCTA (0.01%)	CCCTT (0.01%)	CTAGA (0.01%)	CTAGC (0.01%)	CTAGA (0.00%)
-5	CTAGC (0.00%)	CTAGC (0.00%)	CTAGA (0.00%)	CTAGT (0.01%)	CTAGT (0.00%)	CTAGT (0.00%)	CTAGT (0.00%)	ACTAG (0.00%)	CTAGC (0.00%)
-4	CTAGA (0.00%)	CCTAG (0.00%)	CCTAG (0.00%)	CCTAG (0.00%)	ACTAG (0.00%)	ACTAG (0.00%)	CTAGC (0.00%)	CTAGT (0.00%)	CTAGT (0.00%)
-3	ACTAG (0.00%)	CTAGT (0.00%)	CTAGT (0.00%)	ACTAG (0.00%)	CCTAG (0.00%)	CTAGG (0.00%)	ACTAG (0.00%)	TCTAG (0.00%)	CTAGG (0.00%)
-2	CTAGG (0.00%)	TCTAG (0.00%)	CTAGC (0.00%)	TCTAG (0.00%)	CTAGG (0.00%)	TCTAG (0.00%)	TCTAG (0.00%)	CTAGG (0.00%)	TCTAG (0.00%)
-1	TCTAG (0.00%)	CTAGG (0.00%)	TCTAG (0.00%)	CTAGG (0.00%)	TCTAG (0.00%)	CCTAG (0.00%)	CTAGG (0.00%)	CCTAG (0.00%)	ACTAG (0.00%)

Most common

Least common

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