

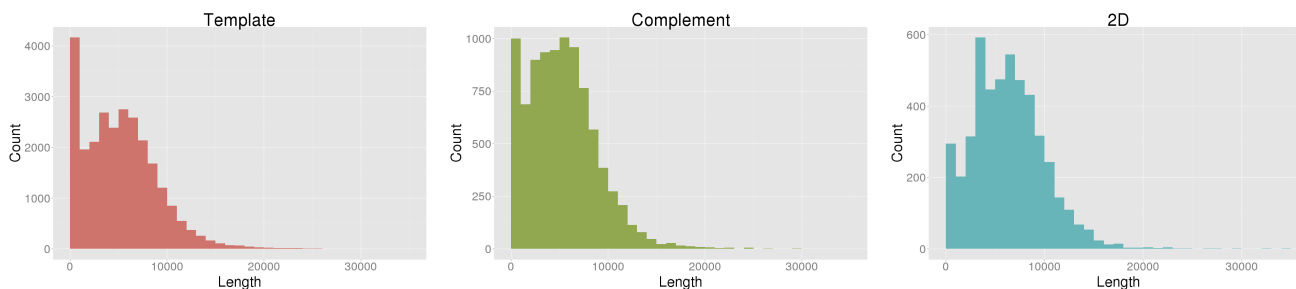
# NanoOK report for MARC\_EColi\_040115

## Pass and fail counts

| Type       | Pass | Fail  |
|------------|------|-------|
| Template   | 0    | 26382 |
| Complement | 0    | 8999  |
| 2D         | 0    | 4790  |

## Read lengths

| Type       | NumReads | TotalBases | Mean    | Longest | Shortest | N50  | N50Count | N90  | N90Count |
|------------|----------|------------|---------|---------|----------|------|----------|------|----------|
| Template   | 26382    | 139579249  | 5290.70 | 154207  | 7        | 7471 | 6628     | 3419 | 16871    |
| Complement | 8999     | 48453372   | 5384.31 | 388445  | 10       | 7204 | 2403     | 3299 | 6081     |
| 2D         | 4790     | 30370391   | 6340.37 | 34081   | 124      | 8036 | 1427     | 3825 | 3454     |



## Template alignments

|                                    |                |
|------------------------------------|----------------|
| Number of reads                    | 26382          |
| Number of reads with alignments    | 11046 (41.87%) |
| Number of reads without alignments | 15336 (58.13%) |

| ID               | Size    | Number of Reads | % of Reads | Mean read length | Aligned bases | Mean coverage | Longest Perf Kmer |
|------------------|---------|-----------------|------------|------------------|---------------|---------------|-------------------|
| Control sequence | 3560    | 939             | 3.56       | 2883.06          | 2464394       | 692.25        | 61                |
| Escherichia coli | 4641652 | 10107           | 38.31      | 7369.74          | 59159014      | 12.75         | 64                |

## Complement alignments

|                                    |               |
|------------------------------------|---------------|
| Number of reads                    | 8999          |
| Number of reads with alignments    | 4860 (54.01%) |
| Number of reads without alignments | 4139 (45.99%) |

| ID               | Size    | Number of Reads | % of Reads | Mean read length | Aligned bases | Mean coverage | Longest Perf Kmer |
|------------------|---------|-----------------|------------|------------------|---------------|---------------|-------------------|
| Control sequence | 3560    | 331             | 3.68       | 2695.99          | 665720        | 187.00        | 41                |
| Escherichia coli | 4641652 | 4529            | 50.33      | 6666.57          | 23065287      | 4.97          | 55                |

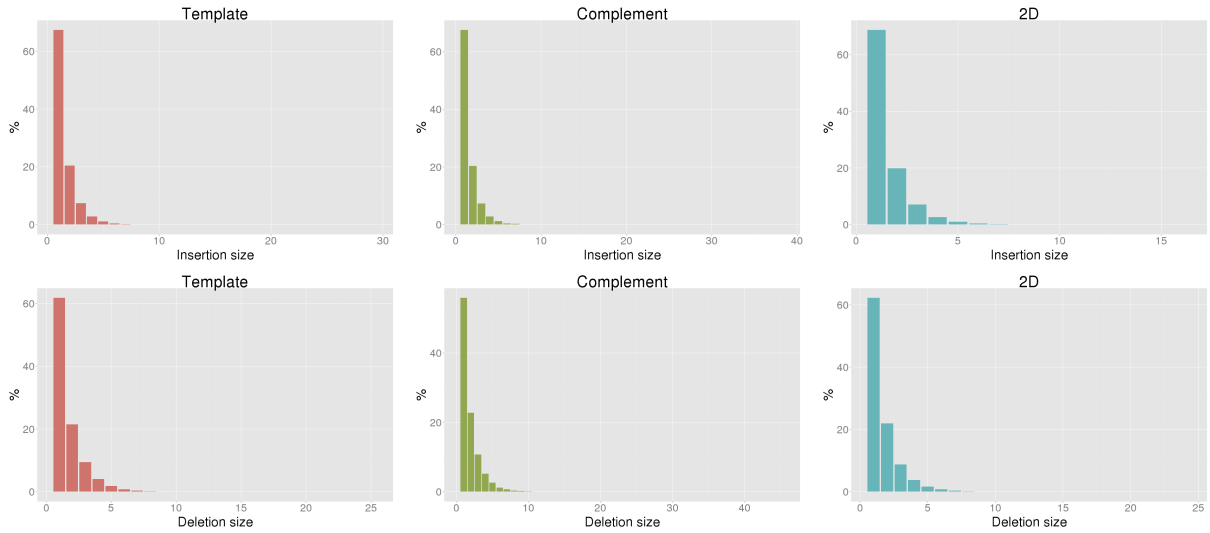
## 2D alignments

|                                    |               |
|------------------------------------|---------------|
| Number of reads                    | 4790          |
| Number of reads with alignments    | 3988 (83.26%) |
| Number of reads without alignments | 802 (16.74%)  |

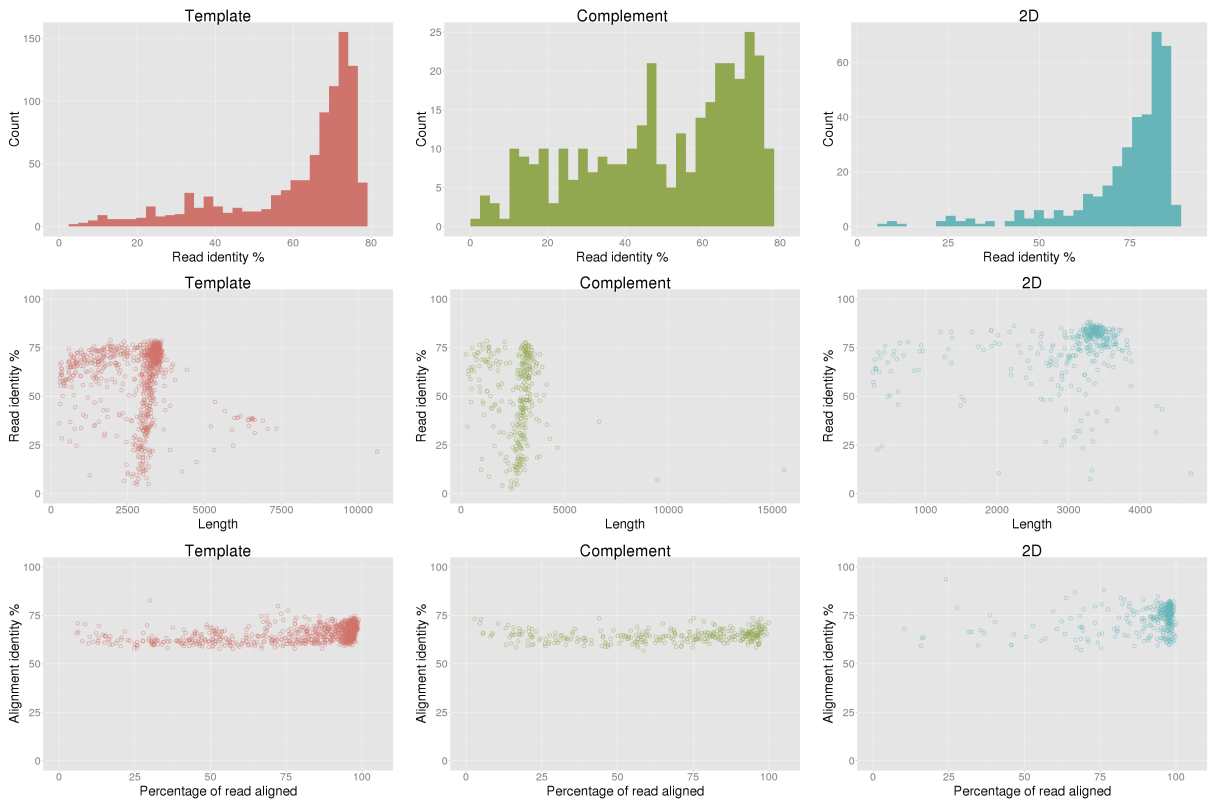
| ID               | Size    | Number of Reads | % of Reads | Mean read length | Aligned bases | Mean coverage | Longest Perf Kmer |
|------------------|---------|-----------------|------------|------------------|---------------|---------------|-------------------|
| Control sequence | 3560    | 369             | 7.70       | 2961.96          | 1100150       | 309.03        | 155               |
| Escherichia coli | 4641652 | 3619            | 75.55      | 7130.44          | 25680969      | 5.53          | 161               |

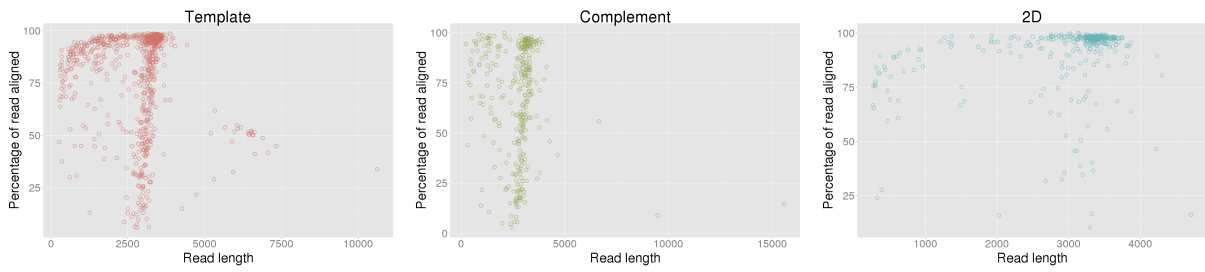
## Control sequence error analysis

|  | Template | Complement | 2D     |
|--|----------|------------|--------|
| Overall base identity (excluding indels)                 | 60.28%   | 47.95%     | 74.43% |
| Aligned base identity (excluding indels)                 | 79.01%   | 79.31%     | 85.70% |
| Identical bases per 100 aligned bases (including indels) | 66.22%   | 64.27%     | 73.95% |
| Inserted bases per 100 aligned bases (including indels)  | 4.71%    | 4.14%      | 4.44%  |
| Deleted bases per 100 aligned bases (including indels)   | 11.48%   | 14.82%     | 9.27%  |
| Substitutions per 100 aligned bases (including indels)   | 17.59%   | 16.77%     | 12.34% |
| Mean insertion size                                      | 1.54     | 1.53       | 1.50   |
| Mean deletion size                                       | 1.68     | 1.88       | 1.67   |

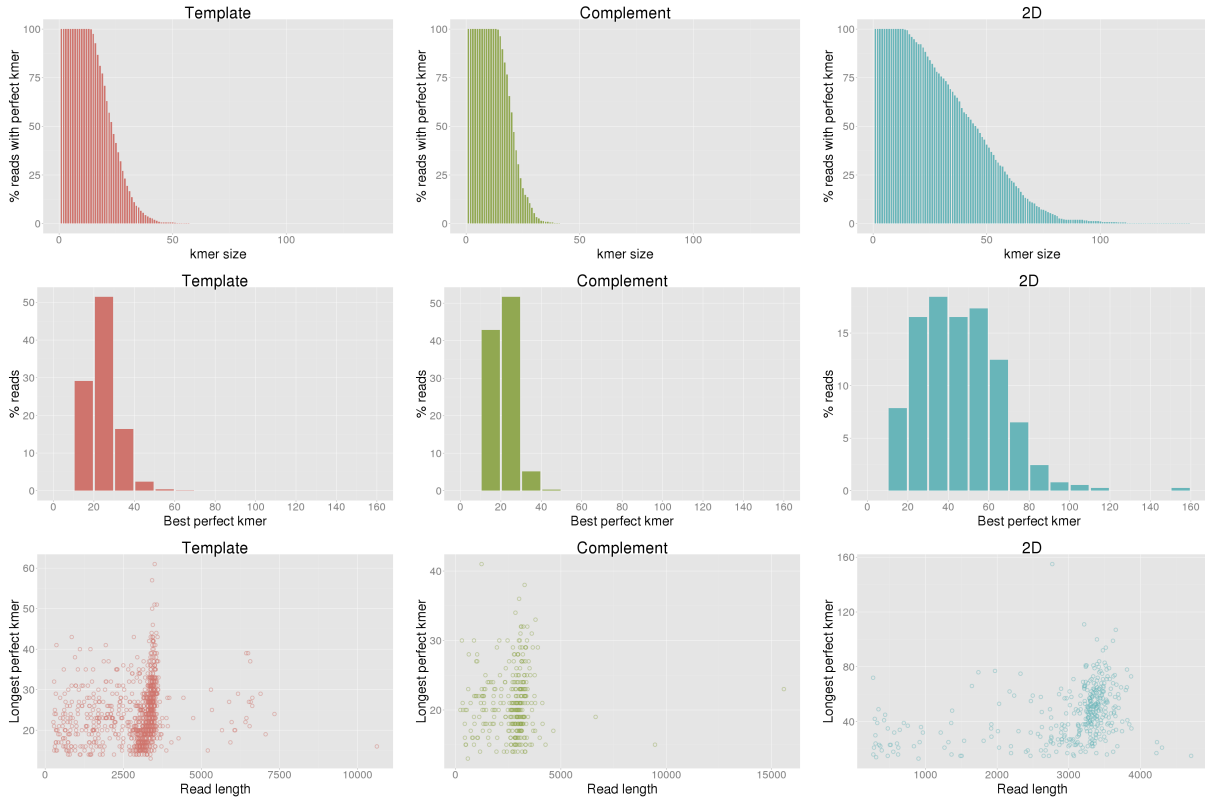


## Control sequence read identity

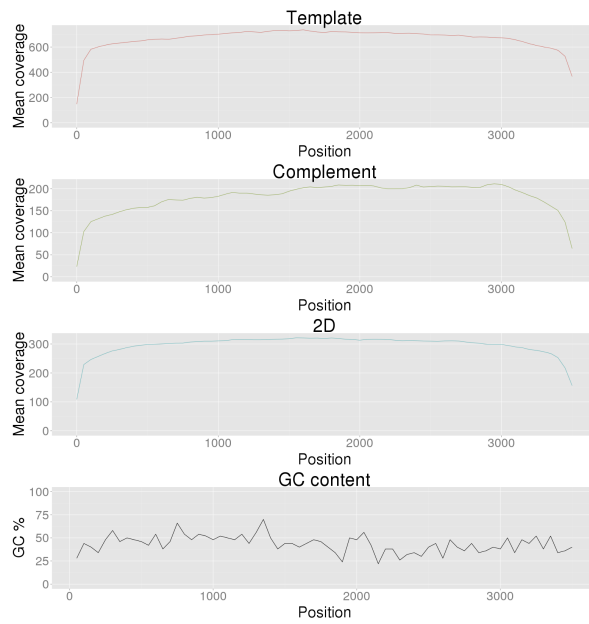




## Control sequence perfect kmers



## Control sequence coverage



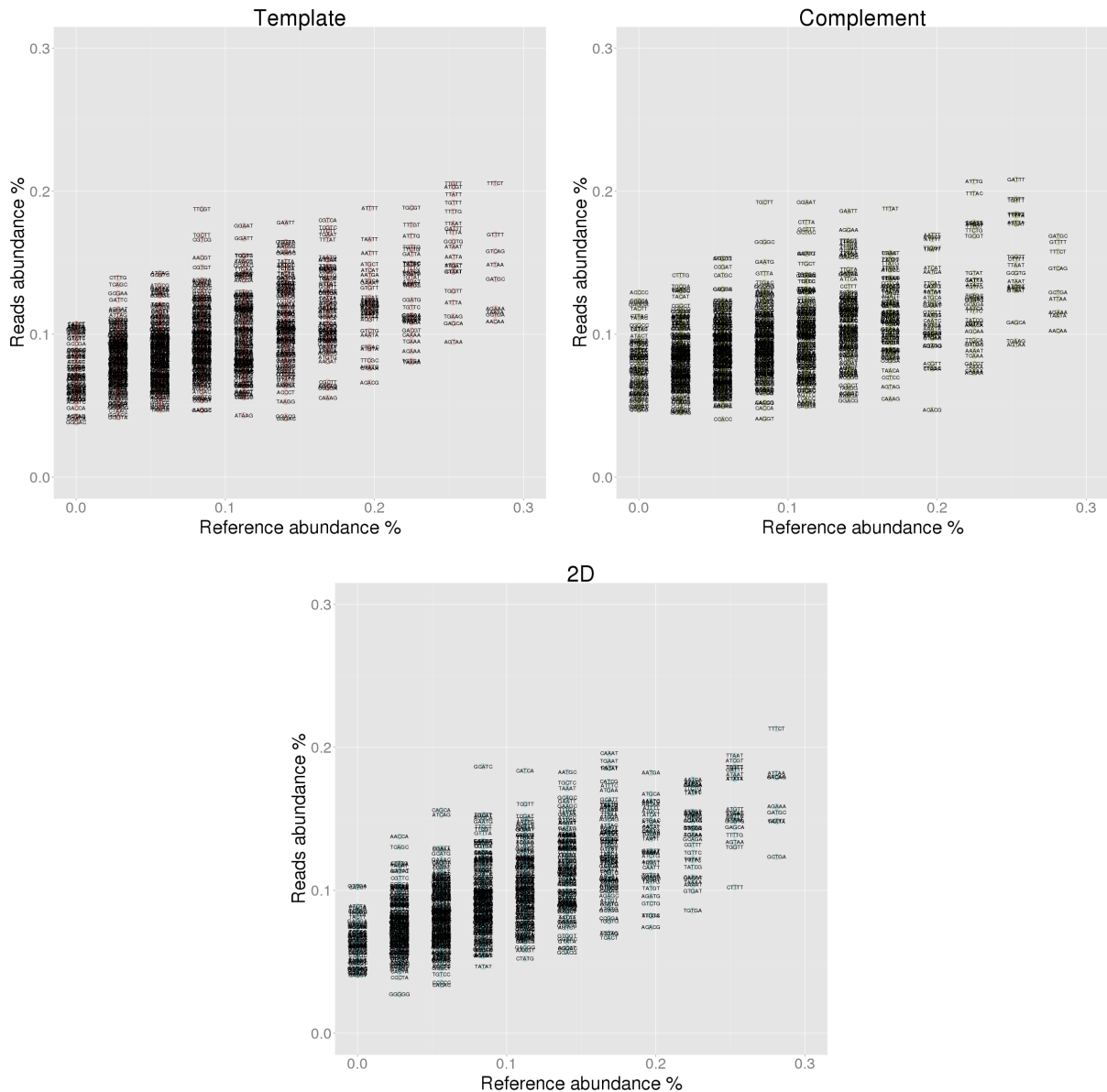
# Control sequence 5-mer analysis

## Under-represented 5-mers

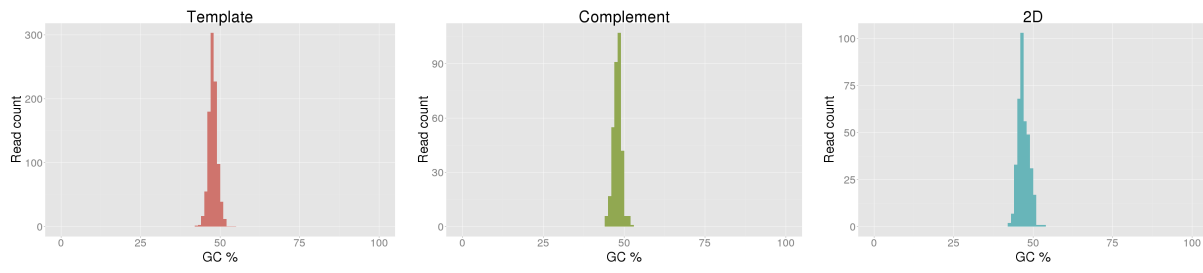
| Rank | Template |       |        |        | Complement |       |        |        | 2D    |       |        |        |
|------|----------|-------|--------|--------|------------|-------|--------|--------|-------|-------|--------|--------|
|      | kmer     | Ref % | Read % | Diff % | kmer       | Ref % | Read % | Diff % | kmer  | Ref % | Read % | Diff % |
| 1    | TTTTT    | 0.759 | 0.151  | -0.608 | TTTTT      | 0.759 | 0.113  | -0.645 | TTTTT | 0.759 | 0.058  | -0.700 |
| 2    | AAAAA    | 0.478 | 0.092  | -0.386 | AAAAA      | 0.478 | 0.061  | -0.416 | AAAAA | 0.478 | 0.055  | -0.422 |
| 3    | TGATG    | 0.393 | 0.131  | -0.263 | AAAAC      | 0.337 | 0.099  | -0.238 | TGATG | 0.393 | 0.135  | -0.259 |
| 4    | AAAAC    | 0.337 | 0.109  | -0.228 | TGATG      | 0.393 | 0.181  | -0.213 | AAAAC | 0.337 | 0.138  | -0.199 |
| 5    | GATGT    | 0.309 | 0.092  | -0.217 | GATGT      | 0.309 | 0.111  | -0.198 | GATGT | 0.309 | 0.114  | -0.195 |
| 6    | CTGAT    | 0.309 | 0.111  | -0.198 | GCAAT      | 0.309 | 0.121  | -0.188 | CTGAT | 0.309 | 0.120  | -0.189 |
| 7    | GCAAT    | 0.309 | 0.122  | -0.187 | AACAA      | 0.281 | 0.102  | -0.179 | TTATC | 0.309 | 0.151  | -0.158 |
| 8    | AACAA    | 0.281 | 0.109  | -0.172 | TAATA      | 0.281 | 0.113  | -0.168 | GCTGA | 0.281 | 0.123  | -0.158 |
| 9    | AATAT    | 0.309 | 0.137  | -0.172 | AGAAA      | 0.281 | 0.115  | -0.166 | CTTTT | 0.253 | 0.102  | -0.151 |
| 10   | GCTGA    | 0.281 | 0.114  | -0.167 | AATAT      | 0.309 | 0.145  | -0.164 | GCAAT | 0.309 | 0.166  | -0.143 |

## Over-represented 5-mers

| Rank | Template |       |        |        | Complement |       |        |        | 2D    |       |        |        |
|------|----------|-------|--------|--------|------------|-------|--------|--------|-------|-------|--------|--------|
|      | kmer     | Ref % | Read % | Diff % | kmer       | Ref % | Read % | Diff % | kmer  | Ref % | Read % | Diff % |
| 1    | CTTTG    | 0.028 | 0.140  | 0.111  | ACCCC      | 0.000 | 0.129  | 0.129  | AACCA | 0.028 | 0.138  | 0.109  |
| 2    | CATCT    | 0.000 | 0.107  | 0.107  | CCCCA      | 0.000 | 0.124  | 0.124  | GTCGA | 0.000 | 0.103  | 0.103  |
| 3    | TACTT    | 0.000 | 0.107  | 0.107  | GCCGA      | 0.000 | 0.123  | 0.123  | GCATC | 0.084 | 0.187  | 0.102  |
| 4    | TCAGC    | 0.028 | 0.135  | 0.107  | GAGGA      | 0.000 | 0.121  | 0.121  | TCAGC | 0.028 | 0.130  | 0.102  |
| 5    | ATCTA    | 0.000 | 0.105  | 0.105  | TACTT      | 0.000 | 0.118  | 0.118  | CATCT | 0.000 | 0.102  | 0.102  |
| 6    | TCTTA    | 0.000 | 0.105  | 0.105  | CTTTG      | 0.028 | 0.141  | 0.113  | CAGCA | 0.056 | 0.156  | 0.100  |
| 7    | ACTCT    | 0.000 | 0.105  | 0.105  | CTTAC      | 0.000 | 0.113  | 0.113  | ATCAG | 0.056 | 0.153  | 0.097  |
| 8    | CGAGA    | 0.000 | 0.104  | 0.104  | TATAC      | 0.000 | 0.111  | 0.111  | CTTTG | 0.028 | 0.119  | 0.091  |
| 9    | GCTCC    | 0.000 | 0.104  | 0.104  | TGCTT      | 0.084 | 0.192  | 0.108  | ACCAA | 0.028 | 0.118  | 0.090  |
| 10   | GAGGA    | 0.000 | 0.103  | 0.103  | CGGCC      | 0.000 | 0.106  | 0.106  | TACAT | 0.028 | 0.117  | 0.089  |

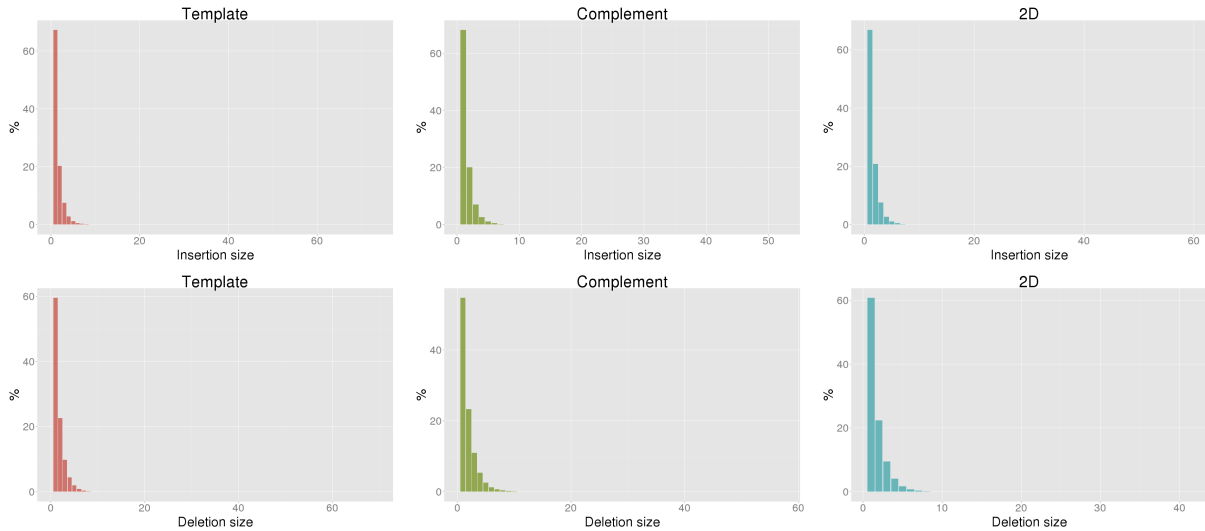


## Control sequence GC content

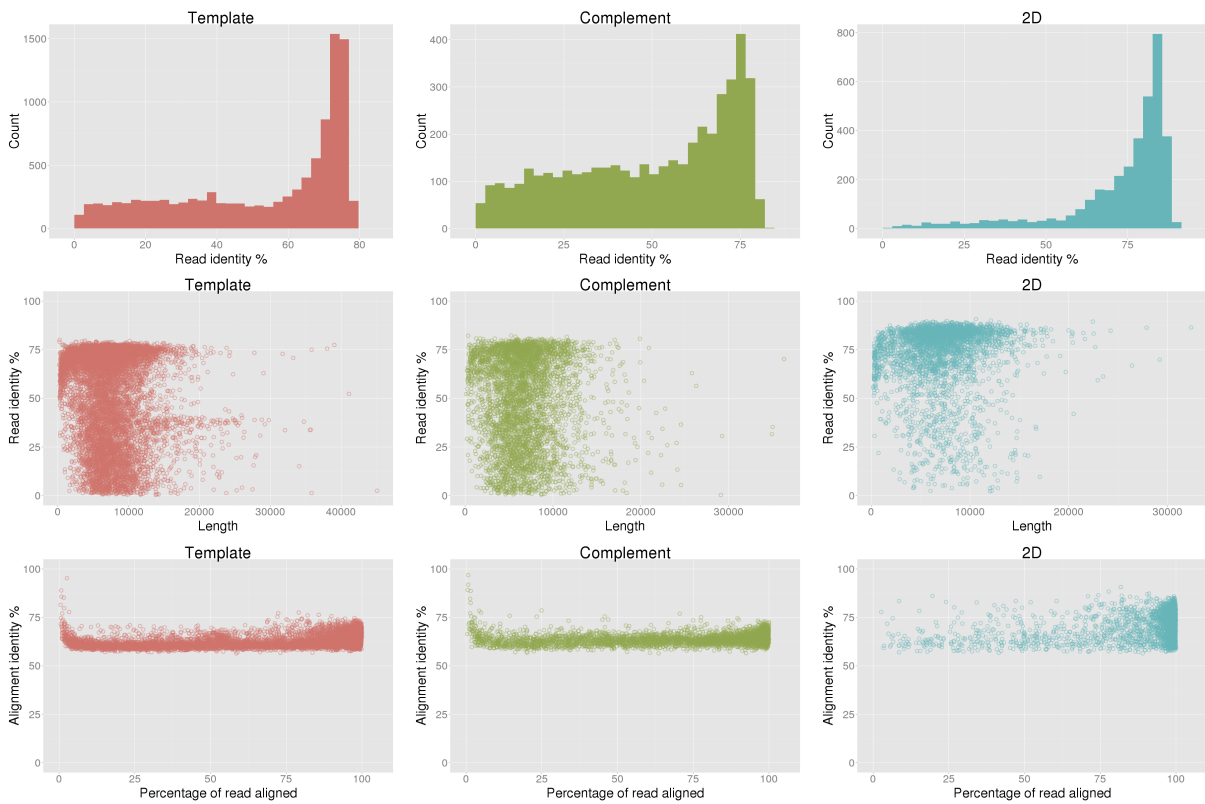


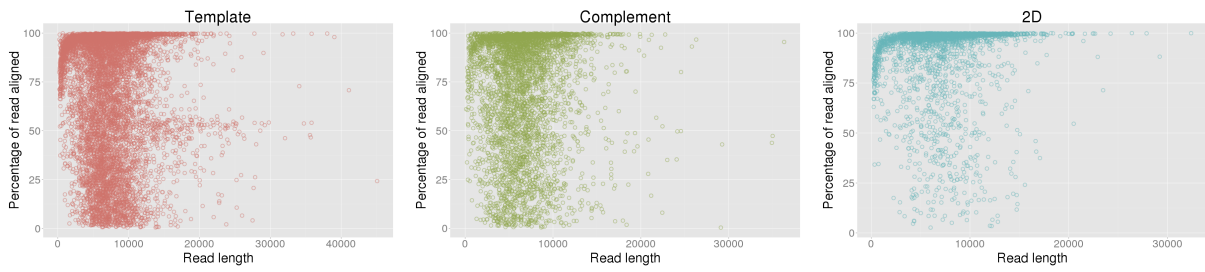
## Escherichia coli error analysis

|  | Template | Complement | 2D     |
|--|----------|------------|--------|
| Overall base identity (excluding indels)                 | 51.16%   | 48.86%     | 72.64% |
| Aligned base identity (excluding indels)                 | 77.83%   | 79.48%     | 85.09% |
| Identical bases per 100 aligned bases (including indels) | 64.42%   | 63.96%     | 72.99% |
| Inserted bases per 100 aligned bases (including indels)  | 4.87%    | 4.06%      | 4.66%  |
| Deleted bases per 100 aligned bases (including indels)   | 12.36%   | 15.47%     | 9.56%  |
| Substitutions per 100 aligned bases (including indels)   | 18.35%   | 16.51%     | 12.79% |
| Mean insertion size                                      | 1.55     | 1.52       | 1.54   |
| Mean deletion size                                       | 1.73     | 1.92       | 1.70   |

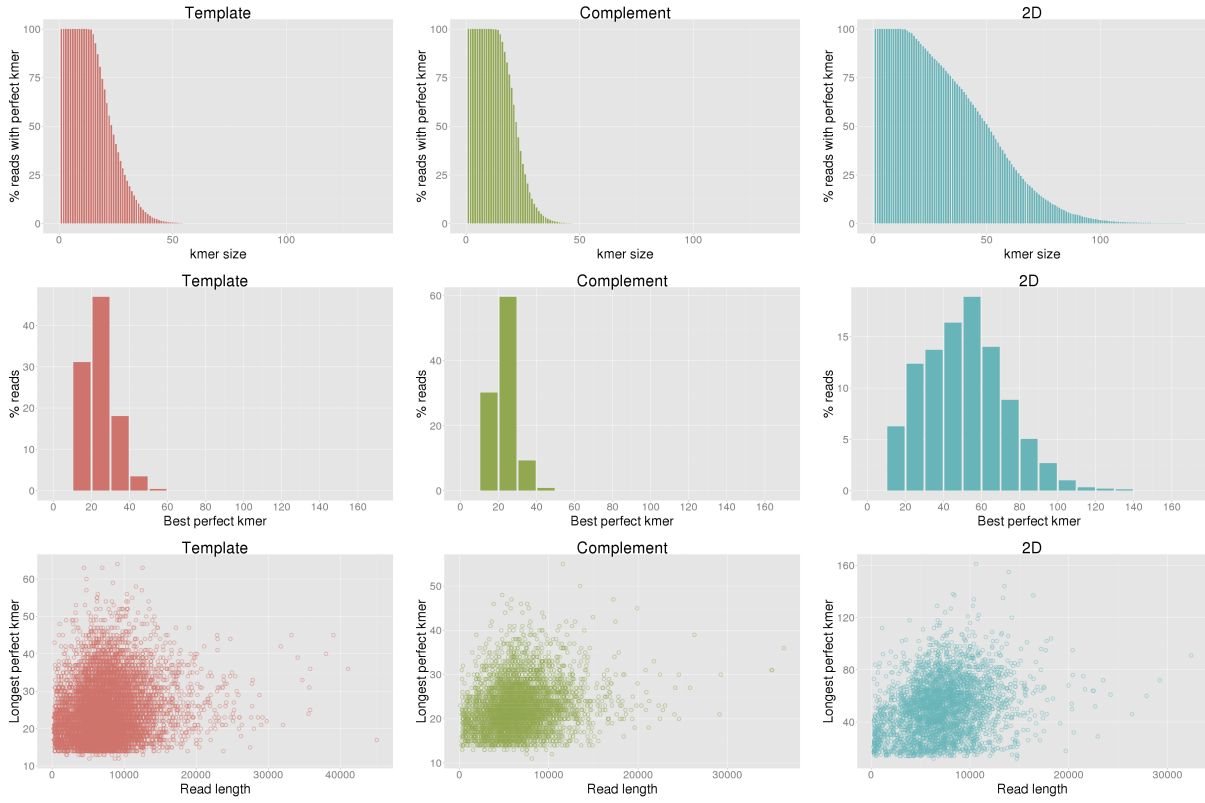


## Escherichia coli read identity

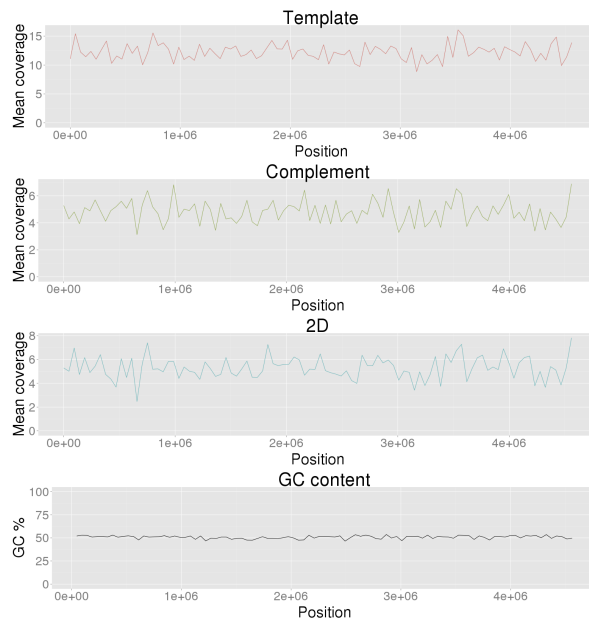




## Escherichia coli perfect kmers



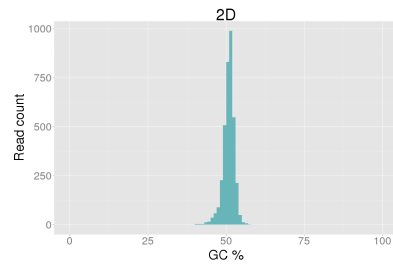
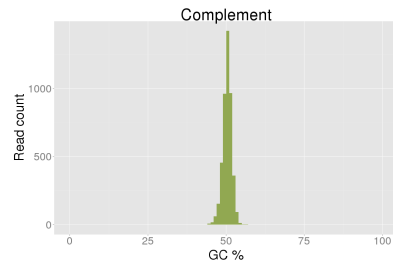
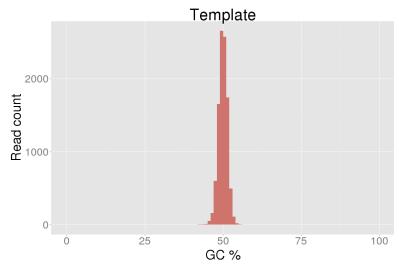
## Escherichia coli coverage



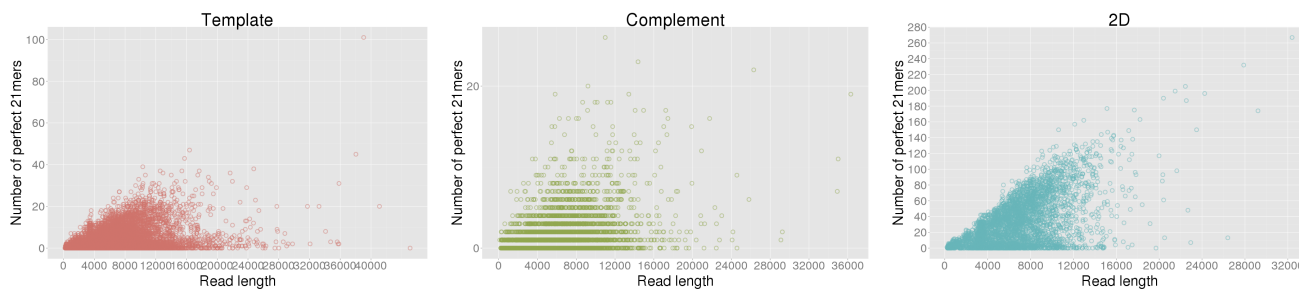




# Escherichia coli GC content



## 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                   | 8.73 | 8.90 | 5.09 | 0.00                     | 9.10 | 8.72 | 5.25 | 0.00             | 8.90  | 8.97  | 4.74 |
| C         | 8.63                   | 0.00 | 9.10 | 9.76 | 9.02                     | 0.00 | 8.80 | 9.45 | 8.76             | 0.00  | 10.18 | 8.75 |
| G         | 9.32                   | 9.13 | 0.00 | 8.42 | 8.92                     | 8.92 | 0.00 | 8.63 | 8.69             | 10.34 | 0.00  | 8.41 |
| T         | 5.35                   | 8.99 | 8.58 | 0.00 | 5.48                     | 8.71 | 9.00 | 0.00 | 4.78             | 8.82  | 8.66  | 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 (3.31%) | TTC (3.40%) | AAA (4.05%)  | TTC (2.85%) | TGC (2.87%) | AAA (3.98%)  | GCA (3.03%) | GCA (2.68%) | AAA (3.69%)  |
| 2    | AAA (2.85%) | TGC (2.95%) | TTC (3.67%)  | TGC (2.82%) | AAA (2.86%) | GCA (3.46%)  | TTC (2.99%) | AAA (2.64%) | GCA (3.43%)  |
| 3    | GCA (2.75%) | AAA (2.90%) | GCA (3.21%)  | AAA (2.67%) | GGC (2.80%) | GAA (3.19%)  | AAA (2.86%) | TGC (2.63%) | TTC (3.12%)  |
| 4    | TGC (2.65%) | GCA (2.75%) | GAA (2.92%)  | GCA (2.64%) | GCA (2.72%) | TTC (3.14%)  | TGC (2.58%) | TTC (2.62%) | GAA (2.99%)  |
| 5    | TCA (2.47%) | TCA (2.40%) | TTT (2.73%)  | CAG (2.44%) | TTC (2.64%) | TTT (2.74%)  | TCA (2.52%) | GCG (2.51%) | TTT (2.75%)  |
| 6    | ATC (2.41%) | GCC (2.40%) | TGC (2.61%)  | GAA (2.41%) | GAA (2.46%) | TGC (2.55%)  | GAA (2.50%) | GGC (2.51%) | AAT (2.48%)  |
| 7    | GAA (2.28%) | ATC (2.27%) | AAT (2.50%)  | GGC (2.29%) | GCC (2.31%) | TCA (2.32%)  | ATC (2.32%) | TCA (2.42%) | GTT (2.44%)  |
| 8    | TTT (2.27%) | GAA (2.27%) | TCA (2.21%)  | TCA (2.28%) | CAG (2.29%) | ATC (2.21%)  | TTT (2.22%) | GAA (2.28%) | TGC (2.30%)  |
| 9    | GCC (2.26%) | GGC (2.24%) | ATC (2.21%)  | GCC (2.24%) | TCA (2.21%) | CAA (2.16%)  | GCG (2.21%) | CAG (2.25%) | TCA (2.22%)  |
| 10   | CAG (2.14%) | AAT (2.23%) | GCC (2.14%)  | ATC (2.20%) | GCG (2.18%) | AAT (2.15%)  | CGC (2.16%) | CGC (2.19%) | GCC (2.21%)  |
|      |             |             |              |             |             |              |             |             |              |
| -10  | AGT (0.98%) | CCT (0.91%) | GGT (0.92%)  | AGA (1.02%) | AGA (0.96%) | CCC (0.95%)  | CTT (1.03%) | GTA (0.98%) | GTA (0.93%)  |
| -9   | CTC (0.94%) | GGG (0.91%) | TGT (0.90%)  | AGT (1.01%) | GGA (0.94%) | CTC (0.94%)  | CTC (0.99%) | ACT (0.96%) | GGG (0.93%)  |
| -8   | AGA (0.86%) | AGT (0.89%) | AGA (0.89%)  | GGA (0.89%) | AGT (0.94%) | AGT (0.91%)  | AGA (0.86%) | CCC (0.95%) | AGG (0.91%)  |
| -7   | CCC (0.85%) | CTT (0.89%) | CTT (0.83%)  | CTC (0.89%) | CCC (0.91%) | CCT (0.85%)  | GAG (0.85%) | CGA (0.90%) | CTT (0.84%)  |
| -6   | GGA (0.80%) | CGA (0.84%) | AGT (0.83%)  | CCC (0.84%) | CCT (0.89%) | ACT (0.82%)  | GGA (0.80%) | CTT (0.89%) | CGA (0.84%)  |
| -5   | GAG (0.76%) | GAG (0.82%) | AGG (0.82%)  | GAG (0.79%) | CTC (0.86%) | AGG (0.80%)  | CCC (0.77%) | CCT (0.87%) | GAG (0.74%)  |
| -4   | AGG (0.71%) | AGA (0.74%) | GGG (0.82%)  | AGG (0.70%) | GAG (0.82%) | GGG (0.74%)  | AGG (0.75%) | AGA (0.80%) | GGA (0.68%)  |
| -3   | GGG (0.65%) | GGA (0.69%) | GAG (0.64%)  | CTA (0.57%) | GGG (0.78%) | GAG (0.62%)  | GGG (0.66%) | GGA (0.71%) | AGA (0.68%)  |
| -2   | CTA (0.51%) | TAG (0.51%) | TAG (0.39%)  | GGG (0.56%) | TAG (0.51%) | CTA (0.44%)  | CTA (0.56%) | TAG (0.60%) | TAG (0.44%)  |
| -1   | TAG (0.46%) | CTA (0.47%) | CTA (0.35%)  | TAG (0.45%) | CTA (0.48%) | TAG (0.39%)  | TAG (0.49%) | CTA (0.58%) | CTA (0.39%)  |
|      |             |             |              |             |             |              |             |             |              |

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 (1.04%) | TTTC (1.09%) | AAAA (1.40%) | CTGC (0.91%) | TGGC (1.06%) | AAAA (1.17%) | ATCA (0.85%) | TGGC (0.94%) | GGCA (1.07%) |
| 2    | AAAA (0.93%) | AAAA (0.93%) | TTTT (1.24%) | CAGC (0.88%) | CAGC (1.01%) | CAAA (1.01%) | GGCA (0.84%) | CAGC (0.90%) | AAAA (1.04%) |
| 3    | GAAA (0.91%) | TTCA (0.92%) | GAAA (1.14%) | ATCA (0.87%) | CGGC (0.98%) | GAAA (1.00%) | TTCA (0.82%) | TTCA (0.83%) | GAAA (0.99%) |
| 4    | TTCA (0.86%) | TGCC (0.92%) | TTTT (0.96%) | TTGC (0.81%) | CTGC (0.91%) | TTTC (0.93%) | GAAA (0.82%) | CTGC (0.78%) | TTTC (0.94%) |
| 5    | ATCA (0.86%) | TTCC (0.90%) | GTTT (0.93%) | CCAG (0.80%) | TTGC (0.89%) | TGAA (0.92%) | TTTC (0.82%) | TTTC (0.77%) | TGAA (0.87%) |
| 6    | CAGC (0.82%) | TTGC (0.89%) | AAAT (0.90%) | TGGC (0.79%) | CAAA (0.82%) | GGCA (0.91%) | AACA (0.81%) | ATCA (0.76%) | CAAA (0.85%) |
| 7    | TGCC (0.81%) | CAGC (0.86%) | GGCA (0.89%) | CAAA (0.78%) | TTCC (0.82%) | AGCA (0.89%) | GCCA (0.79%) | TGCC (0.76%) | TGCA (0.85%) |
| 8    | CTTC (0.79%) | GAAA (0.86%) | CAAA (0.88%) | AAAA (0.77%) | AAAA (0.80%) | TTTT (0.88%) | CAAA (0.77%) | TTGC (0.75%) | TTTT (0.85%) |
| 9    | TTGC (0.79%) | TGGC (0.85%) | GCAA (0.85%) | TGCC (0.77%) | TGCC (0.80%) | ATCA (0.87%) | AAAA (0.76%) | GCCA (0.73%) | CGTT (0.85%) |
| 10   | TTTT (0.78%) | CTGC (0.84%) | TTCC (0.85%) | TTTT (0.77%) | ATCA (0.80%) | TGCA (0.84%) | AAAA (0.75%) | CGGC (0.72%) | GGAA (0.84%) |
|      |              |              |              |              |              |              |              |              |              |
| -10  | TCTA (0.12%) | TCTA (0.12%) | TTAG (0.11%) | TTAG (0.11%) | ACTA (0.12%) | TTAG (0.11%) | TCTA (0.14%) | ACTA (0.15%) | CTAT (0.11%) |
| -9   | GAGG (0.12%) | GGAC (0.12%) | ACTA (0.10%) | CTAA (0.11%) | CCTC (0.12%) | TCTA (0.11%) | GGAC (0.13%) | CGGA (0.14%) | CTAA (0.11%) |
| -8   | AGGG (0.12%) | CGGA (0.12%) | CGAG (0.10%) | CCCT (0.10%) | ACCT (0.12%) | GAGG (0.11%) | CTAT (0.13%) | CCCC (0.14%) | CCCT (0.11%) |
| -7   | TTAG (0.11%) | CCCT (0.12%) | TAGT (0.09%) | AGGG (0.10%) | CTAT (0.12%) | CTAT (0.10%) | GGGG (0.13%) | CTAT (0.14%) | CGGA (0.11%) |
| -6   | CTAA (0.11%) | TAGT (0.12%) | GGAC (0.09%) | TAGA (0.10%) | CTAA (0.11%) | GGAC (0.09%) | GAGG (0.13%) | CTAA (0.12%) | TCTA (0.10%) |
| -5   | GGAC (0.10%) | CTAA (0.10%) | TAGA (0.08%) | GGAC (0.10%) | TAGA (0.09%) | CGAG (0.08%) | CTAA (0.12%) | CCCT (0.11%) | ACTA (0.10%) |
| -4   | TAGA (0.07%) | TAGG (0.07%) | TCTA (0.07%) | GAGG (0.09%) | CCCT (0.08%) | CCCT (0.07%) | TAGA (0.08%) | TAGG (0.09%) | TAGG (0.07%) |
| -3   | CCTA (0.07%) | TAGA (0.05%) | TAGG (0.07%) | CCTA (0.07%) | TAGG (0.08%) | TAGG (0.07%) | TAGG (0.06%) | TAGA (0.08%) | TAGA (0.06%) |
| -2   | TAGG (0.06%) | CCTA (0.05%) | CCTA (0.05%) | TAGG (0.04%) | CCTA (0.06%) | CCTA (0.06%) | CCTA (0.06%) | CCTA (0.08%) | 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%) |
|      |              |              |              |              |              |              |              |              |              |

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    | CAGCA (0.34%) | CAGCA (0.35%) | GAAAA (0.42%) | CAGCA (0.38%) | CAGCA (0.43%) | CAGCA (0.52%) | CAGCA (0.36%) | CTGGC (0.39%) | CAGCA (0.42%) |
| 2    | TTATC (0.31%) | CTGGC (0.34%) | CAAAA (0.40%) | GCTGC (0.35%) | CTGGC (0.42%) | CGGCA (0.38%) | CGCCA (0.33%) | CAGCA (0.34%) | TGGCA (0.40%) |
| 3    | CTGGC (0.31%) | TTTGC (0.33%) | CAGCA (0.39%) | CATCA (0.32%) | CGGGC (0.34%) | GAAAA (0.37%) | TGGCA (0.31%) | CGCCA (0.29%) | CGGCA (0.36%) |
| 4    | CATCA (0.30%) | TTGCC (0.32%) | TGTTT (0.35%) | AATCA (0.31%) | GCTGC (0.33%) | GCAAA (0.33%) | CTGGC (0.30%) | CCAGC (0.28%) | GAAAA (0.35%) |
| 5    | CAAAA (0.29%) | GCTGC (0.30%) | CGTTT (0.34%) | CTGGC (0.31%) | CCAGC (0.33%) | CATCA (0.33%) | CGGCA (0.30%) | TGGCG (0.28%) | CAAAA (0.32%) |
| 6    | ATTTT (0.29%) | GCAGC (0.30%) | GCAAA (0.33%) | CCAGC (0.30%) | AATCA (0.32%) | CAGAA (0.32%) | GCAAA (0.29%) | GCCAG (0.28%) | GCGTT (0.31%) |
| 7    | GAAAA (0.29%) | TTTTT (0.30%) | AGAAA (0.32%) | GCAAA (0.28%) | TTTTT (0.31%) | TGGCA (0.31%) | TGGCG (0.28%) | GCAGC (0.28%) | GCAAA (0.30%) |
| 8    | GCAGC (0.28%) | TTTCA (0.30%) | TTTGC (0.31%) | CGGCA (0.28%) | CATCA (0.31%) | AATCA (0.30%) | CATCA (0.27%) | GCGGC (0.27%) | ATTTT (0.29%) |
| 9    | CGTTT (0.27%) | GAAAA (0.29%) | TCTTC (0.31%) | TTTTT (0.27%) | GCAGC (0.29%) | CTGCA (0.30%) | TGTTT (0.27%) | TTTCA (0.27%) | TGAAA (0.28%) |
| 10   | TTTTT (0.27%) | CGTTT (0.28%) | TGGCA (0.31%) | TTATC (0.27%) | GCAAA (0.28%) | ATAAA (0.30%) | GCCAG (0.26%) | CATCA (0.27%) | CTGCA (0.27%) |
|      |               |               |               |               |               |               |               |               |               |
| -10  | CCCTA (0.01%) | GGACC (0.01%) | TAGGT (0.01%) | CCCTA (0.01%) | CCCTA (0.01%) | GCCTT (0.01%) | TAGGA (0.01%) | TCCTA (0.01%) | ACCTA (0.01%) |
| -9   | GGACC (0.00%) | CCCTA (0.01%) | CCCTA (0.01%) | GGACC (0.01%) | CTAGC (0.01%) | CCCTA (0.01%) | GCTAG (0.01%) | TAGGA (0.01%) | CCCTA (0.01%) |
| -8   | CTAGC (0.00%) | GCTAG (0.00%) | ACTAG (0.00%) | CTAGT (0.01%) | GCTAG (0.00%) | CTAGC (0.01%) | CCCTA (0.01%) | GCTAG (0.01%) | ACTAG (0.01%) |
| -7   | ACTAG (0.00%) | ACTAG (0.00%) | GCTAG (0.00%) | ACTAG (0.01%) | ACCTA (0.00%) | GCTAG (0.00%) | ACTAG (0.01%) | CTAGC (0.01%) | CTAGT (0.00%) |
| -6   | CTAGT (0.00%) | CTAGC (0.00%) | CTAGT (0.00%) | GCTAG (0.00%) | ACTAG (0.00%) | CTAGT (0.00%) | CTAGC (0.00%) | ACTAG (0.00%) | GCTAG (0.00%) |
| -5   | GCTAG (0.00%) | CTAGT (0.00%) | CTAGA (0.00%) | CTAGC (0.00%) | CTAGT (0.00%) | ACTAG (0.00%) | CTAGT (0.00%) | CTAGT (0.00%) | CTAGC (0.00%) |
| -4   | CTAGA (0.00%) | TCTAG (0.00%) | CCTAG (0.00%) | CTAGA (0.00%) | CTAGA (0.00%) | CTAGA (0.00%) | CCTAG (0.00%) | TCTAG (0.00%) | CTAGG (0.00%) |
| -3   | TCTAG (0.00%) | CCTAG (0.00%) | CTAGG (0.00%) | CTAGG (0.00%) | TCTAG (0.00%) | TCTAG (0.00%) | TCTAG (0.00%) | CTAGG (0.00%) | CTAGA (0.00%) |
| -2   | CCTAG (0.00%) | CTAGG (0.00%) | CTAGC (0.00%) | CCTAG (0.00%) | CCTAG (0.00%) | CTAGG (0.00%) | CTAGG (0.00%) | CCTAG (0.00%) | TCTAG (0.00%) |
| -1   | CTAGG (0.00%) | CTAGA (0.00%) | TCTAG (0.00%) | TCTAG (0.00%) | CTAGG (0.00%) | CCTAG (0.00%) | CTAGA (0.00%) | CTAGA (0.00%) | CCTAG (0.00%) |
|      |               |               |               |               |               |               |               |               |               |

Most common

Least common

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