Supplementary Materials: Identification and quantification of chimeric sequencing reads in a highly multiplexed RAD-seq protocol

Maria Luisa Martin Cerezo\textsuperscript{1,2*}, Rohan Raval \textsuperscript{1}, Bernardo de Haro Reyes \textsuperscript{1}, Marek Kucka\textsuperscript{3}, Frank Yingguang Chan\textsuperscript{3} and Jarosław Bryk \textsuperscript{1*}
| Adapter name     | Sequence                                                                 |
|------------------|---------------------------------------------------------------------------|
| i5-top_#01       | CGCTCTCCGATCTVBBNAAAGACTGAGGTCA/3Phos/                                   |
| i5-top_#02       | CGCTCTCCGATCTVBBNATGTGTTGGCTGCA/3Phos/                                  |
| i5-top_#04       | CGCTCTCCGATCTVBBNCTCTCATTCTTGCA/3Phos/                                 |
| i5-top_#05       | CGCTCTCCGATCTVBBNGACCTTGAGTGCA/3Phos/                                  |
| i5-top_#06       | CGCTCTCCGATCTVBBNCAAGGTGATGCA/3Phos/                                  |
| i5-top_#07       | CGCTCTCCGATCTVBBNCGGAAGGTGCA/3Phos/                                   |
| i5-top_#10       | CGCTCTCCGATCTVBBNTCTTCTCAGTGCA/3Phos/                                  |
| i5-bottom_#01    | CGCTCTCCGATCTVBBNCGTAAGAGCGTCGTGTAGGGAAAGAGTGT                         |
| i5-bottom_#02    | CGCTCTCCGATCTVBBNCGTAAGAGCGTCGTGTAGGGAAAGAGTGT                         |
| i5-bottom_#04    | CGCTCTCCGATCTVBBNCGTAAGAGCGTCGTGTAGGGAAAGAGTGT                         |
| i5-bottom_#05    | CGCTCTCCGATCTVBBNCGTAAGAGCGTCGTGTAGGGAAAGAGTGT                         |
| i5-bottom_#06    | CGCTCTCCGATCTVBBNCGTAAGAGCGTCGTGTAGGGAAAGAGTGT                         |
| i5-bottom_#07    | CGCTCTCCGATCTVBBNCGTAAGAGCGTCGTGTAGGGAAAGAGTGT                         |
| i5-bottom_#10    | CGCTCTCCGATCTVBBNCGTAAGAGCGTCGTGTAGGGAAAGAGTGT                         |
| i5-bottom_#11    | CGCTCTCCGATCTVBBNCGTAAGAGCGTCGTGTAGGGAAAGAGTGT                         |
| i5-bottom_#12    | CGCTCTCCGATCTVBBNCGTAAGAGCGTCGTGTAGGGAAAGAGTGT                         |

Table 1: i5 adapter’s sequences
Table 2: i7 adapter’s sequences

| Adapter name       | Sequence                                                                 |
|--------------------|--------------------------------------------------------------------------|
| i7-top_#01_AGAGTTCG| GTGACTGGAGTTCAAGACGTGGTCTCTCCGATCTVBBNAGAGTTCG                           |
| i7-top_#02_ACCTGTTG| GTGACTGGAGTTCAGACGTGGTCTCTCCGATCTVBBNACCTGTTG                           |
| i7-top_#03_CTGGTTCA| GTGACTGGAGTTCAAGACGTGGTCTCTCCGATCTVBBNCTGGTTCA                         |
| i7-top_#04_CGACAAGA| GTGACTGGAGTTCAGACGTGGTCTCTCCGATCTVBBNACGACAGA                          |
| i7-top_#05_CAGTCGAA| GTGACTGGAGTTCAAGACGTGGTCTCTCCGATCTVBBNACAGTCGAA                        |
| i7-top_#06_GTCAGAAC| GTGACTGGAGTTCAGACGTGGTCTCTCCGATCTVBBNCTAGACAC                           |
| i7-top_#07_TTCGTTCCG| GTGACTGGAGTTCAGACGTGGTCTCTCCGATCTVBBNTTGTCCG                           |
| i7-top_#08_TCGCATTTC| GTGACTGGAGTTCAAGACGTGGTCTCTCCGATCTVBBNCTCAGATTTC                      |
| i7-top_#09_TCGAACCA| GTGACTGGAGTTCAGACGTGGTCTCTCCGATCTVBBNCTCAGACCA                        |
| i7-bottom_#01_AGAGTTCG| TACGAACTCTVBBAGATCGGAAGAGCA                                            |
| i7-bottom_#02_ACCTGTTG| TACCAACGTTNVBAGATCGGAAGAGCA                                             |
| i7-bottom_#03_CTGGTTCA| TATGAAACGNNVBBAGATCGGAAGAGCA                                           |
| i7-bottom_#04_CGACAAGA| TATTTTGTNCVBBAGATCGGAAGAGCA                                             |
| i7-bottom_#05_CAGTCGAA| TATTGGTGCNVAAGAGCA                                                      |
| i7-bottom_#06_GTCAGAAC| TATGTTCTGACNNVBBAGATCGGAAGAGCA                                         |
| i7-bottom_#07_TTCGTTCCG| TACGGAACANVBBAGATCGGAAGAGCA                                             |
| i7-bottom_#08_TCGCATTTC| TACGGAACANVBBAGATCGGAAGAGCA                                             |
| i7-bottom_#09_TCGAACCA| TATGTTCTGANVBBAGATCGGAAGAGCA                                           |
## Table 3: Combinatorial outer adapter sequences

| Adapter name | Sequence |
|--------------|----------|
| i501_AGCACTGGA | AATGATACGGCGACCACCGAGATCTACAC{AGCATGGA}ACACTCTTTCCCTACACGAC*G |
| i502_CCTGGAAT | AATGATACGGCGACCACCGAGATCTACAC{CCTGGAAT}ACACTCTTTCCCTACACGAC*G |
| i503_GCAAGCAA | AATGATACGGCGACCACCGAGATCTACAC{GCAAGCAA}ACACTCTTTCCCTACACGAC*G |
| i504_TGAGGATG | AATGATACGGCGACCACCGAGATCTACAC{TGAGGATG}ACACTCTTTCCCTACACGAC*G |
| i701_ACACCTCAG | CAAGCAGAAGACGCGCATACGAGAT{CTGAGTGT}GTGACTGGAGTTCAGACGTGTGC*T |
| i702_CAGTCGAA | CAAGCAGAAGACGCGCATACGAGAT{TTCGACTG}GTGACTGGAGTTCAGACGTGTGC*T |
| i703_GGCTCAAT | CAAGCAGAAGACGCGCATACGAGAT{ATTGAGCC}GTGACTGGAGTTCAGACGTGTGC*T |
| i704_TTCCGCTT | CAAGCAGAAGACGCGCATACGAGAT{AAGCGGAA}GTGACTGGAGTTCAGACGTGTGC*T |
Table 4: Proportion of chimeric sequences per plate. Median (+/- standard deviation), mean, maximum and minimum values are shown.

| Plate     | Mismatches | Median | Stdev | Mean | Max  | Min  |
|-----------|------------|--------|-------|------|------|------|
| PlateA-1  | 0.00       | 0.54   | 0.14  | 0.55 | 0.87 | 0.36 |
| PlateA-2  | 0.00       | 0.59   | 0.17  | 0.64 | 0.91 | 0.38 |
| PlateA-3  | 0.00       | 0.49   | 0.12  | 0.53 | 0.69 | 0.33 |
| PlateA-4  | 0.00       | 0.90   | 0.23  | 0.89 | 1.21 | 0.44 |
| PlateB-1  | 0.00       | 0.83   | 0.19  | 0.89 | 1.39 | 0.68 |
| PlateB-2  | 0.00       | 1.32   | 0.16  | 1.30 | 1.51 | 1.03 |
| PlateB-3  | 0.00       | 1.25   | 0.67  | 1.28 | 2.33 | 0.31 |

Analysis were performed using the four type A libraries. Demultiplexed chimeras, classified by the combination of barcodes, demultiplexing group and plate, were used as samples in Stacks (ustacks, cstacks, sstacks, tsv2bam, gstacks). SNPs were called without filtering, with -r set up to 10. Analysis were run for 0, 2 and 4 mismatches.