

Demultiplexing Overview

My demultiplexing script took two hours and thirty-seven minutes to run after leaving all the files open. This was a significant reduction from my initial 8 hours when I opened and closed each file every time I used it. I used a very strict quality score cutoff of 32 which resulted in a large number of reads being placed into the unknown file. I chose to report my percentages as a percentage of the reads that correctly ended up in a barcoded file. The barcode, followed by the count, followed by the percentage is reported below. Also included is a heatmap of all barcode combinations, and a bar graph of all barcode counts, unknown counts and index hopped counts. These are found in the graphs folder.

| Barcode | Count | Percent |
|----------|----------|---------|
| CTCTGGAT | 28273074 | 11% |
| TACCGGAT | 59803610 | 22% |
| AGAGTCCA | 9424161 | 3.5% |
| GTAGCGTA | 6425781 | 2.4% |
| ATCATGCG | 8385008 | 3.1% |
| AACAGCGA | 7418886 | 2.8% |
| TCGACAAG | 3174898 | 1.2% |
| TCGAGAGT | 9284990 | 3.5% |
| CGGTAATC | 3730418 | 1.4% |
| TAGCCATG | 8657718 | 3.2% |
| TCTTCGAC | 34156815 | 13% |
| CTAGCTCA | 14820156 | 5.5% |
| TATGGCAC | 9086426 | 3.4% |
| ACGATCAG | 6841163 | 2.5% |
| GATCTTGC | 3054876 | 1.1% |
| AGGATAGC | 7217701 | 2.7% |
| TGTTCCGT | 13452889 | 5.0% |
| GCTACTCT | 5780573 | 2.1% |
| CGATCGAT | 4834036 | 1.8% |
| GTCCTAAG | 7309039 | 2.7% |
| GATCAAGG | 5204218 | 1.9% |
| TCGGATTC | 3638673 | 1.4% |
| CACTTCAC | 3369605 | 1.3% |
| ATCGTGGT | 5606619 | 2.1% |