Demultiplex Assignment

David Degnan

October 15th, 2018

PART 1: DEMULTIPLEXER Script

Important Script Info

- This script and associated barcodes are specific to the "1294_S1_L008_R1_001.fastq.gz" data set (and associated R2, R3, and R4).

- Any barcodes that fail quality control metrics (not 8 nucleotides, average Phred score less than 25, and any N's) were placed with their associated sequences in an "Unknowns" file and designated with a "Index pair filed qc" in the Qname.

- Likewise, any sequences that failed quality control metrics (average Phred score less than 25) were placed in an "Unknowns" file with a "Sequence pair failed qc" in the Qname. Note that if one index/sequence failed in a pair, both reads counted as failed and were placed in Unknowns.

Script Validation

- All Read1 and Read2 files have the same number of lines (matched barcodes, Unknowns, mismatches).

- An expected 2,905,973,880 lines are expected (total size of the two paired end files) between all the generated files which is what we see here.

Table 1: Total lines of R1 and R2 files

FileName	Length	FileName.1	Length.1	FileName.2	Length.2
A1	61358416	A2	77298928	A3	121194736
A5	43352624	A7	27695432	A8	65844128
A10	88044384	A11	133161968	A12	29368128
B1	59977488	B2	55496736	B3	262814640
B4	67060024	B7	85055624	B8	34559192
B9	67879616	B10	86607696	C1	48753640
C2	51945624	C3	36721720	C4	574377936
C7	31995176	C9	80002064	C10	317899672
Mis	4736600	Unk	392771688	Total	2905973880

Raw Demultiplexer Script

```python
#!/usr/bin/python
#
import argparse
import gzip
#
# Associate the sequence and barcode name
barcodes = {"GTAGCGTA":"B1", "CGATCGAT":"A5", "GATCAAGG":"C1", "AACAGCGA":"B9",
"TAGCCATG":"C9", "CGGTAATC":"C3", "CTCTGGAT":"B3", "TACCGGAT":"C4",
```

```python
"CTAGCTCA":"A11", "CACTTCAC":"C7", "GCTACTCT":"B2", "ACGATCAG":"A1",
"TATGGCAC":"B7", "TGTTCCGT":"A3", "GTCCTAAG":"B4", "TCGACAAG":"A12",
"TCTTCGAC":"C10", "ATCATGCG":"A2", "ATCGTGGT":"C2", "TCGAGAGT":"A10",
"TCGGATTC":"B8", "GATCTTGC":"A7", "AGAGTCCA":"B10", "AGGATAGC":"A8"}
#
# Keep record of all the match file states
openFiles = {"B1":"closed", "A5":"closed", "C1":"closed", "B9":"closed",
"C9":"closed", "C3":"closed", "B3":"closed",
"C4":"closed", "A11":"closed", "C7":"closed", "B2":"closed", "A1":"closed",
"B7":"closed", "A3":"closed", "B4":"closed",
"A12":"closed", "C10":"closed", "A2":"closed", "C2":"closed",
"A10":"closed", "B8":"closed", "A7":"closed", "B10":"closed",
"A8":"closed"}
#
# Keep track of all opened matched files
matches = {}
#
#
def get_arguments():
    '''Takes 4 FASTQ file and demultiplexes. F1 is Read1, F2 is Index1, F3
    is Index2, and F4 is Read2.'''
    parser = argparse.ArgumentParser(description = "Demultiplexing File Inputs")
    parser.add_argument("-R", "--R1", help= "Designates the Read1 File", required=True, type=str)
    parser.add_argument("-I", "--I1", help= "Designates the Index1 File", required=True, type=str)
    parser.add_argument("-J", "--I2", help= "Designates the Index2 File", required=True, type=str)
    parser.add_argument("-S", "--R2", help= "Designates the Read2 File", required=True, type=str)
    return parser.parse_args()
#
#
def openPreliminaryFiles(R1, I1, I2, R2):
    '''Opens the original 4 FASTQ files with reads and indices, and
    generates 2 Unknown and 2 Mismatch files.'''
    File1 = R1
    File2 = I1
    File3 = I2
    File4 = R2
    Read1 = gzip.open(File1, mode="rt")
    Read2 = gzip.open(File2, mode="rt")
    Read3 = gzip.open(File3, mode="rt")
    Read4 = gzip.open(File4, mode="rt")
    Unk1 = open("R1_Unknowns.fastq", "a")
    Unk2 = open("R2_Unknowns.fastq", "a")
    Mis1 = open("R1_Mismatches.fastq", "a")
    Mis2 = open("R2_Mismatches.fastq", "a")
    return Read1, Read2, Read3, Read4, Unk1, Unk2, Mis1, Mis2
#
#
def Length_and_NCheck(Seq):
    '''Determines if the length is exactly 8 nucleotides and if there are
    N's in the barcode.'''
    if len(Seq) != 8 or "N" in Seq:
        TestPass = False
    else:
```

```python
        TestPass = True
    return TestPass
#
#
def AverageQualityScore(QS):
    '''Determines if the average quality score < 25'''
    TotalScore = 0
    for score in QS:
        TotalScore += ord(score) - 33
    averageScore = TotalScore / len(QS)
    if averageScore < 25:
        TestPass = False
    else:
        TestPass = True
    return TestPass
#
#
def QualityControl(fh):
    '''Reads in the next three lines in a FASTQ file after the header and
    tests for length, N's and average quality score > 25'''
    TestPass = True
    Seq = fh.readline().strip("\n")
    fh.readline()
    QS = fh.readline().strip("\n")
    if TestPass == True:
        TestPass = Length_and_NCheck(Seq)
    if TestPass == True:
        TestPass = AverageQualityScore(QS)
    return TestPass, Seq
#
#
def SeqRevComp(Seq):
    '''Creates the complement sequence'''
    # Reverses the sequence
    RevComp = ""
    Seq = Seq[::-1]
    for base in range(len(Seq)):
        if Seq[base] == "A":
            RevComp += "T"
        elif Seq[base] == "C":
            RevComp += "G"
        elif Seq[base] == "T":
            RevComp += "A"
        elif Seq[base] == "G":
            RevComp += "C"
    return RevComp
#
#
def grabLinesTestAndWrite(Read, NextRead, IndexPair, Filename1, Filename2, Unk1, Unk2):
    '''Takes the first four lines of a file and writes to its barcoded file,
    assuming the quality score is greater than 25. Otherwise, its goes to
    unknowns.'''
    Qname1 = Read.readline().strip("\n")
```

```python
        Qname1 = Qname1 + ":" + IndexPair + "\n"
        RSeq1 = Read.readline()
        Plus1 = Read.readline()
        Phred1 = Read.readline()

        Qname2 = NextRead.readline().strip("\n")
        Qname2 = Qname2 + ":" + IndexPair + "\n"
        RSeq2 = NextRead.readline()
        Plus2 = NextRead.readline()
        Phred2 = NextRead.readline()

        if AverageQualityScore(Phred1) == False or AverageQualityScore(Phred2) == False:
            Qname1 = Qname1.strip("\n") + ": Sequence pair failed qc\n"
            Qname2 = Qname2.strip("\n") + ": Sequence pair failed qc\n"
            Filename1 = Unk1
            Filename2 = Unk2

        ToWrite1 = Qname1 + RSeq1 + Plus1 + Phred1
        ToWrite2 = Qname2 + RSeq2 + Plus2 + Phred2
        Filename1.write(ToWrite1)
        Filename2.write(ToWrite2)

def NameMatchFiles(Bar1):
    '''Creates the name and opens the files for matches'''
    R1BName = "R1" + "_" + Bar1 + ".fastq"
    R2BName = "R2" + "_" + Bar1 + ".fastq"
    R1BFile = open(R1BName, "a")
    R2BFile = open(R2BName, "a")
    return R1BFile, R2BFile

def main():
    args = get_arguments()
    R1 = args.R1
    I1 = args.I1
    I2 = args.I2
    R2 = args.R2
    Read1, Read2, Read3, Read4, Unk1, Unk2, Mis1, Mis2 = openPreliminaryFiles(R1, I1, I2, R2)
    while True:
        Index1Header = Read2.readline()
        if not Index1Header:
            break
        Index1Pass, Index1Seq = QualityControl(Read2)
        Index2Header = Read3.readline()
        if not Index2Header:
            break
        Index2Pass, Index2Seq = QualityControl(Read3)
        Index2Seq = SeqRevComp(Index2Seq)
        if Index1Pass == True and Index2Pass == True \
        and Index1Seq in barcodes and Index2Seq in barcodes:
            Bar1 = barcodes[Index1Seq]
            Bar2 = barcodes[Index2Seq]
            BarPair = Bar1 + "_" + Bar2
            if Bar1 == Bar2:
```

```python
                if openFiles[Bar1] == "closed":
                    openFiles[Bar1] = "open"
                    R1BFile, R2BFile = NameMatchFiles(Bar1)
                    grabLinesTestAndWrite(Read1, Read4, BarPair, R1BFile, R2BFile, Unk1, Unk2)
                    matches[Bar1] = [R1BFile, R2BFile]
                elif openFiles[Bar1] == "open":
                    R1BFile = matches[Bar1][0]
                    R2BFile = matches[Bar1][1]
                    grabLinesTestAndWrite(Read1, Read4, BarPair, R1BFile, R2BFile, Unk1, Unk2)
            else:
                grabLinesTestAndWrite(Read1, Read4, BarPair, Mis1, Mis2, Unk1, Unk2)
        else:
            IndexPair = Index1Seq + "_" + Index2Seq + ": Index pair failed qc"
            grabLinesTestAndWrite(Read1, Read4, IndexPair, Unk1, Unk2, Unk1, Unk2)
    for key in matches.keys():
        File1 = matches[key][0]
        File2 = matches[key][1]
        File1.close()
        File2.close()

main()
```

Sbatch Script

```
Read1=/projects/bgmp/shared/2017_sequencing/1294_S1_L008_R1_001.fastq.gz
Read2=/projects/bgmp/shared/2017_sequencing/1294_S1_L008_R4_001.fastq.gz
Index1=/projects/bgmp/shared/2017_sequencing/1294_S1_L008_R2_001.fastq.gz
Index2=/projects/bgmp/shared/2017_sequencing/1294_S1_L008_R3_001.fastq.gz

cd /projects/bgmp/ddegnan/demultiplex/
./Demultiplexer.py -R $Read1 -S $Read2 -I $Index1 -J $Index2
```

- 20 cores were used and the program ran for 5 hours from 7pm-12am on Saturday, October 13th, 2018.

PART 2: DEMULTIPLEXING Statistics

1. The number of reads per file for Read1 was calculated and stored as Reads.txt and mismatched barcodes were stored (Bash)

```
grep -v "Mismatches" | grep -v "R2" | grep -v "Unknown" | cat *.fastq | \
grep -E "^+$" | wc -l >> Reads.txt

cat R1_Mismatches.fastq | grep -E "@K00337" | cut -d ":" -f11 | sort \
| uniq -c | sort -r > AllMismatches.txt
```

2. Generate Stats Table to determine the proportion of each read pair.

Step One: Combine the Reads and AllMismatches files (R)

```r
# Edit Reads file so that it can be merged with the matches file
ReadNums <- read.table("./Reads.txt")
matches <- c("A10_A10", "A11_A11", "A12_A12", "A1_A1", "A2_A2", "A3_A3",
             "A4_A4", "A7_A7", "A8_A8", "B10_B10", "B1_B1", "B2_B2",
             "B3_B3", "B4_B4", "B7_B7", "B8_B8", "B9_B9", "C10_C10",
             "C1_C1", "C2_C2", "C3_C3", "C4_C4", "C7_C7", "C9_C9")
ReadNums <- data.frame(ReadNums[1], matches)
colnames(ReadNums) <- c("Number", "MatchOrMismatch")

# Edit Mismatches file so that it can be merged with Reads file
Mismatches <- read.table("./AllMismatches.txt")
colnames(Mismatches) <- c("Number", "MatchOrMismatch")

# Merge files
Total <- rbind(ReadNums, Mismatches)

# Output the table
Total <- Total[,2:1]
write.table(Total, "Total.txt", quote=F, row.names=F, sep="\t")
```

Step Two: Add Barcode Sequences (Python)

```python
sequences = {"B1":"GTAGCGTA","B3":"CTCTGGAT","A3":"TGTTCCGT","B8":"TCGGATTC",
"A5": "CGATCGAT","C4": "TACCGGAT","B4": "GTCCTAAG","A7": "GATCTTGC",
"C1": "GATCAAGG","A11": "CTAGCTCA","A12": "TCGACAAG","B10": "AGAGTCCA",
"B9": "AACAGCGA","C7": "CACTTCAC","C10": "TCTTCGAC","A8": "AGGATAGC",
"C9": "TAGCCATG","B2": "GCTACTCT","A2": "ATCATGCG","C3": "CGGTAATC",
"A1": "ACGATCAG","C2": "ATCGTGGT","B7": "TATGGCAC","A10": "TCGAGAGT"}
toPrint = ["MatchOrMismatch\tNumber\tBar1Seq\tBar2Seq\n"]
with open("../demultiplexing-David-Degnan/Total.txt", "r") as fh:
        fh.readline()
        for line in fh:
            line = line.split()
            toWrite = line[0] + "\t" + line[1]
            barcodes = line[0].split("_")
            FirBar = barcodes[0]
            SecBar = barcodes[1]
            toWrite = toWrite + "\t" + sequences[FirBar] + "\t" + sequences[SecBar] + "\n"
            toPrint.append(toWrite)
with open("TotalwSeq.txt", "w") as ofh:
    for line in range(len(toPrint)):
        ofh.write(toPrint[line])
```

Step Three: Add Proportions (R). Unknowns are removed from data set.

```r
Total <- read.table("./TotalwSeq.txt", header = T)
Proportion <- format(round(Total$Number / sum(Total$Number), 8))
Total <- data.frame(Total[,1], Total[,3], Total[,4], Total[,2], Proportion)
colnames(Total) <- c("MatchOrMismatch", "Bar1Seq", "Bar2Seq", "Number",
                    "Proportion")
Total <- Total[order(Total$Number, decreasing=TRUE),]
```

- The total proportion of unknown reads (ones that failed the quality control) was 0.1351601 or 49,096,461 read pairs of the 363,246,735 read pairs.

- The total proportion of mismatches was 0.001629953 which means that index hopping occured 0.163%

of the time. The number of reads which had index hopping was in 592,075 read pairs of the total 363,246,735 read pairs.

- Index hopping is definitely an issue to take into consideration when processing Illumina data, as it occured in quite a few samples. Included below is the all the cases of index hopping in a table.

Table 2: Proportion of Matches/Mismatches

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
C4_C4	TACCGGAT	TACCGGAT	71797242	0.22854426
C10_C10	TCTTCGAC	TCTTCGAC	39737459	0.12649188
B3_B3	CTCTGGAT	CTCTGGAT	32851830	0.10457362
A11_A11	CTAGCTCA	CTAGCTCA	16645246	0.05298498
A3_A3	TGTTCCGT	TGTTCCGT	15149342	0.04822323
A10_A10	TCGAGAGT	TCGAGAGT	11005548	0.03503275
B10_B10	AGAGTCCA	AGAGTCCA	10825962	0.03446109
B7_B7	TATGGCAC	TATGGCAC	10631953	0.03384353
C9_C9	TAGCCATG	TAGCCATG	10000258	0.03183272
A2_A2	ATCATGCG	ATCATGCG	9662366	0.03075715
B9_B9	AACAGCGA	AACAGCGA	8484952	0.02700921
B4_B4	GTCCTAAG	GTCCTAAG	8382503	0.02668310
A8_A8	AGGATAGC	AGGATAGC	8230516	0.02619930
A1_A1	ACGATCAG	ACGATCAG	7669802	0.02441444
B1_B1	GTAGCGTA	GTAGCGTA	7497186	0.02386497
B2_B2	GCTACTCT	GCTACTCT	6937092	0.02208208
C2_C2	ATCGTGGT	ATCGTGGT	6493203	0.02066910
C1_C1	GATCAAGG	GATCAAGG	6094205	0.01939901
A5_A5	CGATCGAT	CGATCGAT	5419078	0.01724995
C3_C3	CGGTAATC	CGGTAATC	4590215	0.01461153
B8_B8	TCGGATTC	TCGGATTC	4319899	0.01375106
C7_C7	CACTTCAC	CACTTCAC	3999397	0.01273084
A12_A12	TCGACAAG	TCGACAAG	3671016	0.01168554
A7_A7	GATCTTGC	GATCTTGC	3461929	0.01101998
B7_A3	TATGGCAC	TGTTCCGT	85791	0.00027309
A3_B7	TGTTCCGT	TATGGCAC	82720	0.00026331
C1_C10	GATCAAGG	TCTTCGAC	17102	0.00005444
A11_A12	CTAGCTCA	TCGACAAG	14135	0.00004499
C4_B3	TACCGGAT	CTCTGGAT	12914	0.00004111
B3_C4	CTCTGGAT	TACCGGAT	11932	0.00003798
C4_C10	TACCGGAT	TCTTCGAC	10284	0.00003274
C10_C4	TCTTCGAC	TACCGGAT	9346	0.00002975
A12_A2	TCGACAAG	ATCATGCG	7546	0.00002402
B4_B7	GTCCTAAG	TATGGCAC	6702	0.00002133
C3_C4	CGGTAATC	TACCGGAT	6579	0.00002094
B3_C10	CTCTGGAT	TCTTCGAC	5002	0.00001592
C10_B3	TCTTCGAC	CTCTGGAT	4990	0.00001588
C4_A3	TACCGGAT	TGTTCCGT	4166	0.00001326
C10_C2	TCTTCGAC	ATCGTGGT	4102	0.00001306
B7_C4	TATGGCAC	TACCGGAT	4044	0.00001287
A3_C4	TGTTCCGT	TACCGGAT	3953	0.00001258
B7_C10	TATGGCAC	TCTTCGAC	3799	0.00001209
C4_A11	TACCGGAT	CTAGCTCA	3646	0.00001161
A11_C4	CTAGCTCA	TACCGGAT	3636	0.00001157
A11_B3	CTAGCTCA	CTCTGGAT	3326	0.00001059

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
C4_B7	TACCGGAT	TATGGCAC	3270	0.00001041
C4_C9	TACCGGAT	TAGCCATG	3147	0.00001002
A10_C4	TCGAGAGT	TACCGGAT	3117	0.00000992
C4_A10	TACCGGAT	TCGAGAGT	3079	0.00000980
C7_C9	CACTTCAC	TAGCCATG	3037	0.00000967
C10_A3	TCTTCGAC	TGTTCCGT	2608	0.00000830
B3_A11	CTCTGGAT	CTAGCTCA	2531	0.00000806
C10_B7	TCTTCGAC	TATGGCAC	2454	0.00000781
C4_C3	TACCGGAT	CGGTAATC	2319	0.00000738
A3_C10	TGTTCCGT	TCTTCGAC	2313	0.00000736
C9_C4	TAGCCATG	TACCGGAT	2257	0.00000718
C10_A11	TCTTCGAC	CTAGCTCA	2248	0.00000716
B3_B2	CTCTGGAT	GCTACTCT	2228	0.00000709
C10_A10	TCTTCGAC	TCGAGAGT	2140	0.00000681
B10_C4	AGAGTCCA	TACCGGAT	2132	0.00000679
C4_B10	TACCGGAT	AGAGTCCA	2099	0.00000668
A11_C10	CTAGCTCA	TCTTCGAC	2091	0.00000666
B9_C4	AACAGCGA	TACCGGAT	2071	0.00000659
C4_A1	TACCGGAT	ACGATCAG	2034	0.00000647
A2_C4	ATCATGCG	TACCGGAT	1992	0.00000634
C4_B4	TACCGGAT	GTCCTAAG	1973	0.00000628
B3_A3	CTCTGGAT	TGTTCCGT	1923	0.00000612
B4_C4	GTCCTAAG	TACCGGAT	1906	0.00000607
C4_A2	TACCGGAT	ATCATGCG	1888	0.00000601
B3_A10	CTCTGGAT	TCGAGAGT	1887	0.00000601
A1_C4	ACGATCAG	TACCGGAT	1884	0.00000600
C4_A8	TACCGGAT	AGGATAGC	1854	0.00000590
C4_C2	TACCGGAT	ATCGTGGT	1853	0.00000590
B9_C10	AACAGCGA	TCTTCGAC	1823	0.00000580
A5_C4	CGATCGAT	TACCGGAT	1799	0.00000573
B3_A5	CTCTGGAT	CGATCGAT	1790	0.00000570
C2_C4	ATCGTGGT	TACCGGAT	1766	0.00000562
A10_C10	TCGAGAGT	TCTTCGAC	1717	0.00000547
C4_B9	TACCGGAT	AACAGCGA	1711	0.00000545
C10_C1	TCTTCGAC	GATCAAGG	1708	0.00000544
C4_A5	TACCGGAT	CGATCGAT	1707	0.00000543
A11_B2	CTAGCTCA	GCTACTCT	1669	0.00000531
B3_B10	CTCTGGAT	AGAGTCCA	1626	0.00000518
A11_A2	CTAGCTCA	ATCATGCG	1618	0.00000515
A8_C4	AGGATAGC	TACCGGAT	1614	0.00000514
C10_A8	TCTTCGAC	AGGATAGC	1590	0.00000506
A3_B3	TGTTCCGT	CTCTGGAT	1589	0.00000506
A10_B3	TCGAGAGT	CTCTGGAT	1563	0.00000498
A8_C10	AGGATAGC	TCTTCGAC	1555	0.00000495
B2_C4	GCTACTCT	TACCGGAT	1522	0.00000484
C4_C1	TACCGGAT	GATCAAGG	1505	0.00000479
B1_C4	GTAGCGTA	TACCGGAT	1498	0.00000477
C4_B1	TACCGGAT	GTAGCGTA	1496	0.00000476
A2_C10	ATCATGCG	TCTTCGAC	1476	0.00000470
C4_B2	TACCGGAT	GCTACTCT	1438	0.00000458
B1_B2	GTAGCGTA	GCTACTCT	1435	0.00000457
B3_B7	CTCTGGAT	TATGGCAC	1400	0.00000446

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
C10_C9	TCTTCGAC	TAGCCATG	1395	0.00000444
C10_A1	TCTTCGAC	ACGATCAG	1378	0.00000439
B7_B3	TATGGCAC	CTCTGGAT	1356	0.00000432
C10_B8	TCTTCGAC	TCGGATTC	1277	0.00000406
C1_C4	GATCAAGG	TACCGGAT	1276	0.00000406
C10_B10	TCTTCGAC	AGAGTCCA	1272	0.00000405
A2_B3	ATCATGCG	CTCTGGAT	1206	0.00000384
C4_B8	TACCGGAT	TCGGATTC	1204	0.00000383
C9_C10	TAGCCATG	TCTTCGAC	1192	0.00000379
C2_C10	ATCGTGGT	TCTTCGAC	1191	0.00000379
C10_A2	TCTTCGAC	ATCATGCG	1180	0.00000376
C2_B3	ATCGTGGT	CTCTGGAT	1169	0.00000372
B10_C10	AGAGTCCA	TCTTCGAC	1162	0.00000370
C4_C7	TACCGGAT	CACTTCAC	1116	0.00000355
C10_B4	TCTTCGAC	GTCCTAAG	1103	0.00000351
A5_B3	CGATCGAT	CTCTGGAT	1079	0.00000343
B8_C10	TCGGATTC	TCTTCGAC	1076	0.00000343
C10_A5	TCTTCGAC	CGATCGAT	1070	0.00000341
A11_B1	CTAGCTCA	GTAGCGTA	1069	0.00000340
B3_A1	CTCTGGAT	ACGATCAG	1024	0.00000326
B4_C10	GTCCTAAG	TCTTCGAC	1020	0.00000325
C10_B1	TCTTCGAC	GTAGCGTA	1019	0.00000324
B3_B4	CTCTGGAT	GTCCTAAG	1017	0.00000324
A1_C10	ACGATCAG	TCTTCGAC	1015	0.00000323
B3_C9	CTCTGGAT	TAGCCATG	1011	0.00000322
B1_C10	GTAGCGTA	TCTTCGAC	1007	0.00000321
A5_C10	CGATCGAT	TCTTCGAC	1000	0.00000318
B3_A2	CTCTGGAT	ATCATGCG	997	0.00000317
B10_A11	AGAGTCCA	CTAGCTCA	995	0.00000317
C3_C10	CGGTAATC	TCTTCGAC	983	0.00000313
A11_B10	CTAGCTCA	AGAGTCCA	979	0.00000312
C10_B2	TCTTCGAC	GCTACTCT	977	0.00000311
B9_B3	AACAGCGA	CTCTGGAT	977	0.00000311
C4_A7	TACCGGAT	GATCTTGC	971	0.00000309
B8_B9	TCGGATTC	AACAGCGA	965	0.00000307
B2_B3	GCTACTCT	CTCTGGAT	952	0.00000303
A11_A10	CTAGCTCA	TCGAGAGT	947	0.00000301
B10_B1	AGAGTCCA	GTAGCGTA	945	0.00000301
C10_C3	TCTTCGAC	CGGTAATC	935	0.00000298
B2_C10	GCTACTCT	TCTTCGAC	929	0.00000296
B10_B3	AGAGTCCA	CTCTGGAT	929	0.00000296
B8_C4	TCGGATTC	TACCGGAT	919	0.00000293
B10_A10	AGAGTCCA	TCGAGAGT	901	0.00000287
B3_C2	CTCTGGAT	ATCGTGGT	898	0.00000286
C7_C4	CACTTCAC	TACCGGAT	885	0.00000282
B4_B3	GTCCTAAG	CTCTGGAT	876	0.00000279
C10_A7	TCTTCGAC	GATCTTGC	875	0.00000279
C4_A12	TACCGGAT	TCGACAAG	874	0.00000278
A8_B1	AGGATAGC	GTAGCGTA	870	0.00000277
C10_B9	TCTTCGAC	AACAGCGA	869	0.00000277
B7_B8	TATGGCAC	TCGGATTC	868	0.00000276
C10_C7	TCTTCGAC	CACTTCAC	860	0.00000274

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
A10_A3	TCGAGAGT	TGTTCCGT	859	0.00000273
B3_B1	CTCTGGAT	GTAGCGTA	856	0.00000272
C9_B3	TAGCCATG	CTCTGGAT	854	0.00000272
B3_A8	CTCTGGAT	AGGATAGC	853	0.00000272
A11_A3	CTAGCTCA	TGTTCCGT	847	0.00000270
A1_B3	ACGATCAG	CTCTGGAT	838	0.00000267
B4_A3	GTCCTAAG	TGTTCCGT	832	0.00000265
B3_B9	CTCTGGAT	AACAGCGA	817	0.00000260
C10_A12	TCTTCGAC	TCGACAAG	814	0.00000259
A3_A11	TGTTCCGT	CTAGCTCA	811	0.00000258
C7_C10	CACTTCAC	TCTTCGAC	804	0.00000256
A3_B8	TGTTCCGT	TCGGATTC	799	0.00000254
B7_A10	TATGGCAC	TCGAGAGT	797	0.00000254
A3_A10	TGTTCCGT	TCGAGAGT	787	0.00000251
A8_B3	AGGATAGC	CTCTGGAT	772	0.00000246
A12_C4	TCGACAAG	TACCGGAT	768	0.00000244
A2_C9	ATCATGCG	TAGCCATG	747	0.00000238
B1_B3	GTAGCGTA	CTCTGGAT	738	0.00000235
A11_B7	CTAGCTCA	TATGGCAC	736	0.00000234
B1_A11	GTAGCGTA	CTAGCTCA	730	0.00000232
B1_A1	GTAGCGTA	ACGATCAG	715	0.00000228
A3_B9	TGTTCCGT	AACAGCGA	706	0.00000225
B7_A11	TATGGCAC	CTAGCTCA	694	0.00000221
A2_B7	ATCATGCG	TATGGCAC	694	0.00000221
A11_C9	CTAGCTCA	TAGCCATG	691	0.00000220
B3_C3	CTCTGGAT	CGGTAATC	686	0.00000218
A2_A1	ATCATGCG	ACGATCAG	670	0.00000213
A3_B10	TGTTCCGT	AGAGTCCA	667	0.00000212
B7_B4	TATGGCAC	GTCCTAAG	654	0.00000208
B1_A10	GTAGCGTA	TCGAGAGT	646	0.00000206
A7_C4	GATCTTGC	TACCGGAT	646	0.00000206
A11_A1	CTAGCTCA	ACGATCAG	644	0.00000205
A10_A11	TCGAGAGT	CTAGCTCA	633	0.00000201
A7_C10	GATCTTGC	TCTTCGAC	625	0.00000199
B3_C1	CTCTGGAT	GATCAAGG	624	0.00000199
B3_B8	CTCTGGAT	TCGGATTC	621	0.00000198
B3_C7	CTCTGGAT	CACTTCAC	612	0.00000195
A10_A1	TCGAGAGT	ACGATCAG	595	0.00000189
C1_B3	GATCAAGG	CTCTGGAT	593	0.00000189
B7_B9	TATGGCAC	AACAGCGA	588	0.00000187
A2_A11	ATCATGCG	CTAGCTCA	587	0.00000187
B10_A3	AGAGTCCA	TGTTCCGT	583	0.00000186
A12_C10	TCGACAAG	TCTTCGAC	583	0.00000186
A3_A2	TGTTCCGT	ATCATGCG	570	0.00000181
C7_B3	CACTTCAC	CTCTGGAT	568	0.00000181
B9_A7	AACAGCGA	GATCTTGC	564	0.00000180
B2_B1	GCTACTCT	GTAGCGTA	564	0.00000180
A2_B4	ATCATGCG	GTCCTAAG	564	0.00000180
C9_A11	TAGCCATG	CTAGCTCA	555	0.00000177
A2_A3	ATCATGCG	TGTTCCGT	553	0.00000176
C2_C1	ATCGTGGT	GATCAAGG	550	0.00000175
A10_B2	TCGAGAGT	GCTACTCT	545	0.00000173

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
B2_A11	GCTACTCT	CTAGCTCA	540	0.00000172
A11_B4	CTAGCTCA	GTCCTAAG	529	0.00000168
C3_B3	CGGTAATC	CTCTGGAT	528	0.00000168
B7_A8	TATGGCAC	AGGATAGC	522	0.00000166
B9_A2	AACAGCGA	ATCATGCG	518	0.00000165
C9_A3	TAGCCATG	TGTTCCGT	508	0.00000162
A8_B7	AGGATAGC	TATGGCAC	501	0.00000159
B3_A7	CTCTGGAT	GATCTTGC	500	0.00000159
B9_A11	AACAGCGA	CTAGCTCA	498	0.00000159
B2_A3	GCTACTCT	TGTTCCGT	497	0.00000158
A3_C9	TGTTCCGT	TAGCCATG	494	0.00000157
B10_A1	AGAGTCCA	ACGATCAG	493	0.00000157
B8_B3	TCGGATTC	CTCTGGAT	492	0.00000157
A11_B9	CTAGCTCA	AACAGCGA	487	0.00000155
B7_C9	TATGGCAC	TAGCCATG	486	0.00000155
A11_A5	CTAGCTCA	CGATCGAT	486	0.00000155
A1_A11	ACGATCAG	CTAGCTCA	482	0.00000153
A11_A8	CTAGCTCA	AGGATAGC	478	0.00000152
C1_A1	GATCAAGG	ACGATCAG	474	0.00000151
A2_A8	ATCATGCG	AGGATAGC	473	0.00000151
A3_A5	TGTTCCGT	CGATCGAT	472	0.00000150
A3_A1	TGTTCCGT	ACGATCAG	471	0.00000150
A2_B10	ATCATGCG	AGAGTCCA	470	0.00000150
B3_A12	CTCTGGAT	TCGACAAG	461	0.00000147
A8_A11	AGGATAGC	CTAGCTCA	461	0.00000147
A8_A7	AGGATAGC	GATCTTGC	459	0.00000146
A1_A3	ACGATCAG	TGTTCCGT	455	0.00000145
B7_A1	TATGGCAC	ACGATCAG	453	0.00000144
B1_A3	GTAGCGTA	TGTTCCGT	452	0.00000144
A3_A8	TGTTCCGT	AGGATAGC	451	0.00000144
A2_B9	ATCATGCG	AACAGCGA	443	0.00000141
A8_A1	AGGATAGC	ACGATCAG	439	0.00000140
A3_B1	TGTTCCGT	GTAGCGTA	439	0.00000140
B10_A8	AGAGTCCA	AGGATAGC	435	0.00000138
A1_A10	ACGATCAG	TCGAGAGT	434	0.00000138
B9_B10	AACAGCGA	AGAGTCCA	433	0.00000138
B9_A3	AACAGCGA	TGTTCCGT	433	0.00000138
A8_A3	AGGATAGC	TGTTCCGT	432	0.00000138
C2_A10	ATCGTGGT	TCGAGAGT	431	0.00000137
A8_B10	AGGATAGC	AGAGTCCA	430	0.00000137
B9_A1	AACAGCGA	ACGATCAG	429	0.00000137
B4_A11	GTCCTAAG	CTAGCTCA	427	0.00000136
A2_C2	ATCATGCG	ATCGTGGT	427	0.00000136
A2_C1	ATCATGCG	GATCAAGG	426	0.00000136
A5_A3	CGATCGAT	TGTTCCGT	425	0.00000135
A8_A10	AGGATAGC	TCGAGAGT	424	0.00000135
A1_B10	ACGATCAG	AGAGTCCA	424	0.00000135
B7_B10	TATGGCAC	AGAGTCCA	422	0.00000134
B1_B10	GTAGCGTA	AGAGTCCA	420	0.00000134
B9_A8	AACAGCGA	AGGATAGC	419	0.00000133
A10_A8	TCGAGAGT	AGGATAGC	417	0.00000133
A8_A2	AGGATAGC	ATCATGCG	411	0.00000131

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
A11_C2	CTAGCTCA	ATCGTGGT	409	0.00000130
A10_B7	TCGAGAGT	TATGGCAC	409	0.00000130
B7_C7	TATGGCAC	CACTTCAC	404	0.00000129
B10_B7	AGAGTCCA	TATGGCAC	404	0.00000129
A1_B4	ACGATCAG	GTCCTAAG	402	0.00000128
A10_C2	TCGAGAGT	ATCGTGGT	400	0.00000127
B9_B7	AACAGCGA	TATGGCAC	399	0.00000127
A1_C9	ACGATCAG	TAGCCATG	399	0.00000127
A3_C2	TGTTCCGT	ATCGTGGT	396	0.00000126
A3_B2	TGTTCCGT	GCTACTCT	395	0.00000126
B10_B9	AGAGTCCA	AACAGCGA	391	0.00000124
A10_B10	TCGAGAGT	AGAGTCCA	385	0.00000123
C9_B7	TAGCCATG	TATGGCAC	381	0.00000121
B4_B9	GTCCTAAG	AACAGCGA	377	0.00000120
B2_B10	GCTACTCT	AGAGTCCA	374	0.00000119
B7_A2	TATGGCAC	ATCATGCG	373	0.00000119
A1_A2	ACGATCAG	ATCATGCG	373	0.00000119
A2_A10	ATCATGCG	TCGAGAGT	369	0.00000117
C2_A3	ATCGTGGT	TGTTCCGT	368	0.00000117
C1_B4	GATCAAGG	GTCCTAAG	367	0.00000117
B4_C9	GTCCTAAG	TAGCCATG	367	0.00000117
A11_C1	CTAGCTCA	GATCAAGG	367	0.00000117
C1_C2	GATCAAGG	ATCGTGGT	366	0.00000117
A5_A11	CGATCGAT	CTAGCTCA	366	0.00000117
B9_B1	AACAGCGA	GTAGCGTA	365	0.00000116
B10_A2	AGAGTCCA	ATCATGCG	365	0.00000116
A3_B4	TGTTCCGT	GTCCTAAG	365	0.00000116
A1_A8	ACGATCAG	AGGATAGC	362	0.00000115
B9_A10	AACAGCGA	TCGAGAGT	358	0.00000114
A10_C9	TCGAGAGT	TAGCCATG	358	0.00000114
C9_B4	TAGCCATG	GTCCTAAG	355	0.00000113
A1_B7	ACGATCAG	TATGGCAC	352	0.00000112
A12_B3	TCGACAAG	CTCTGGAT	352	0.00000112
C2_A11	ATCGTGGT	CTAGCTCA	349	0.00000111
C9_A10	TAGCCATG	TCGAGAGT	347	0.00000110
B2_A10	GCTACTCT	TCGAGAGT	347	0.00000110
A11_B8	CTAGCTCA	TCGGATTC	347	0.00000110
B8_B7	TCGGATTC	TATGGCAC	344	0.00000110
C9_A1	TAGCCATG	ACGATCAG	337	0.00000107
B4_C1	GTCCTAAG	GATCAAGG	336	0.00000107
A10_B8	TCGAGAGT	TCGGATTC	334	0.00000106
B4_A2	GTCCTAAG	ATCATGCG	331	0.00000105
B1_B4	GTAGCGTA	GTCCTAAG	331	0.00000105
A8_C7	AGGATAGC	CACTTCAC	329	0.00000105
C9_A2	TAGCCATG	ATCATGCG	328	0.00000104
A11_C3	CTAGCTCA	CGGTAATC	328	0.00000104
C7_A1	CACTTCAC	ACGATCAG	325	0.00000103
B4_A1	GTCCTAAG	ACGATCAG	325	0.00000103
A7_B3	GATCTTGC	CTCTGGAT	325	0.00000103
B9_C2	AACAGCGA	ATCGTGGT	322	0.00000102
C9_B10	TAGCCATG	AGAGTCCA	321	0.00000102
B9_B8	AACAGCGA	TCGGATTC	321	0.00000102

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
A1_A5	ACGATCAG	CGATCGAT	319	0.00000102
A10_B4	TCGAGAGT	GTCCTAAG	319	0.00000102
A3_C1	TGTTCCGT	GATCAAGG	318	0.00000101
B7_A7	TATGGCAC	GATCTTGC	317	0.00000101
C1_C9	GATCAAGG	TAGCCATG	316	0.00000101
C1_A3	GATCAAGG	TGTTCCGT	316	0.00000101
B10_C9	AGAGTCCA	TAGCCATG	316	0.00000101
A8_B2	AGGATAGC	GCTACTCT	315	0.00000100
C3_C2	CGGTAATC	ATCGTGGT	312	0.00000099
A10_A2	TCGAGAGT	ATCATGCG	310	0.00000099
B10_B4	AGAGTCCA	GTCCTAAG	309	0.00000098
A7_A8	GATCTTGC	AGGATAGC	309	0.00000098
A1_A12	ACGATCAG	TCGACAAG	306	0.00000097
C9_C1	TAGCCATG	GATCAAGG	304	0.00000097
B4_B1	GTCCTAAG	GTAGCGTA	303	0.00000096
A10_A5	TCGAGAGT	CGATCGAT	302	0.00000096
C7_B7	CACTTCAC	TATGGCAC	297	0.00000095
B4_A10	GTCCTAAG	TCGAGAGT	296	0.00000094
B7_C3	TATGGCAC	CGGTAATC	295	0.00000094
A10_B1	TCGAGAGT	GTAGCGTA	293	0.00000093
B2_A1	GCTACTCT	ACGATCAG	292	0.00000093
B1_B7	GTAGCGTA	TATGGCAC	292	0.00000093
A1_B2	ACGATCAG	GCTACTCT	292	0.00000093
A10_B9	TCGAGAGT	AACAGCGA	292	0.00000093
B8_A11	TCGGATTC	CTAGCTCA	289	0.00000092
B9_B4	AACAGCGA	GTCCTAAG	288	0.00000092
A1_C1	ACGATCAG	GATCAAGG	288	0.00000092
A8_B9	AGGATAGC	AACAGCGA	287	0.00000091
B4_B10	GTCCTAAG	AGAGTCCA	286	0.00000091
C7_A11	CACTTCAC	CTAGCTCA	285	0.00000091
A3_A7	TGTTCCGT	GATCTTGC	285	0.00000091
A11_A7	CTAGCTCA	GATCTTGC	285	0.00000091
A8_C9	AGGATAGC	TAGCCATG	284	0.00000090
B7_B2	TATGGCAC	GCTACTCT	283	0.00000090
B10_B2	AGAGTCCA	GCTACTCT	282	0.00000090
C2_B10	ATCGTGGT	AGAGTCCA	281	0.00000089
B7_B1	TATGGCAC	GTAGCGTA	281	0.00000089
B9_A5	AACAGCGA	CGATCGAT	280	0.00000089
A12_A10	TCGACAAG	TCGAGAGT	280	0.00000089
B9_C9	AACAGCGA	TAGCCATG	279	0.00000089
B7_C1	TATGGCAC	GATCAAGG	277	0.00000088
A11_C7	CTAGCTCA	CACTTCAC	275	0.00000088
C3_C9	CGGTAATC	TAGCCATG	273	0.00000087
B9_C1	AACAGCGA	GATCAAGG	273	0.00000087
B1_B9	GTAGCGTA	AACAGCGA	271	0.00000086
C2_A2	ATCGTGGT	ATCATGCG	270	0.00000086
A2_A12	ATCATGCG	TCGACAAG	269	0.00000086
A8_B8	AGGATAGC	TCGGATTC	268	0.00000085
A12_C9	TCGACAAG	TAGCCATG	268	0.00000085
B7_C2	TATGGCAC	ATCGTGGT	266	0.00000085
B8_A7	TCGGATTC	GATCTTGC	264	0.00000084
C2_A1	ATCGTGGT	ACGATCAG	263	0.00000084

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
B2_B7	GCTACTCT	TATGGCAC	261	0.00000083
A5_A8	CGATCGAT	AGGATAGC	261	0.00000083
C3_A11	CGGTAATC	CTAGCTCA	260	0.00000083
C1_A11	GATCAAGG	CTAGCTCA	260	0.00000083
A10_A12	TCGAGAGT	TCGACAAG	260	0.00000083
A2_B2	ATCATGCG	GCTACTCT	259	0.00000082
B10_C2	AGAGTCCA	ATCGTGGT	256	0.00000081
B1_C9	GTAGCGTA	TAGCCATG	255	0.00000081
B2_B4	GCTACTCT	GTCCTAAG	254	0.00000081
B1_A2	GTAGCGTA	ATCATGCG	253	0.00000081
A2_B1	ATCATGCG	GTAGCGTA	253	0.00000081
A10_C1	TCGAGAGT	GATCAAGG	252	0.00000080
A8_C3	AGGATAGC	CGGTAATC	251	0.00000080
C9_B1	TAGCCATG	GTAGCGTA	250	0.00000080
A5_A10	CGATCGAT	TCGAGAGT	250	0.00000080
C1_A10	GATCAAGG	TCGAGAGT	249	0.00000079
C3_A8	CGGTAATC	AGGATAGC	248	0.00000079
B8_A8	TCGGATTC	AGGATAGC	246	0.00000078
C2_B7	ATCGTGGT	TATGGCAC	245	0.00000078
A3_C7	TGTTCCGT	CACTTCAC	244	0.00000078
C9_A8	TAGCCATG	AGGATAGC	242	0.00000077
B2_A2	GCTACTCT	ATCATGCG	242	0.00000077
B4_B8	GTCCTAAG	TCGGATTC	240	0.00000076
A1_B9	ACGATCAG	AACAGCGA	239	0.00000076
A12_A1	TCGACAAG	ACGATCAG	239	0.00000076
B4_B2	GTCCTAAG	GCTACTCT	236	0.00000075
B8_A3	TCGGATTC	TGTTCCGT	235	0.00000075
B10_A5	AGAGTCCA	CGATCGAT	235	0.00000075
A1_B1	ACGATCAG	GTAGCGTA	234	0.00000074
B1_C2	GTAGCGTA	ATCGTGGT	233	0.00000074
A5_B10	CGATCGAT	AGAGTCCA	233	0.00000074
C3_A3	CGGTAATC	TGTTCCGT	231	0.00000074
C1_A2	GATCAAGG	ATCATGCG	231	0.00000074
B7_A5	TATGGCAC	CGATCGAT	231	0.00000074
B1_A8	GTAGCGTA	AGGATAGC	231	0.00000074
A8_A5	AGGATAGC	CGATCGAT	231	0.00000074
A3_C3	TGTTCCGT	CGGTAATC	230	0.00000073
C7_A3	CACTTCAC	TGTTCCGT	229	0.00000073
B8_A10	TCGGATTC	TCGAGAGT	229	0.00000073
C2_A8	ATCGTGGT	AGGATAGC	228	0.00000073
A8_B4	AGGATAGC	GTCCTAAG	228	0.00000073
A5_A1	CGATCGAT	ACGATCAG	228	0.00000073
B7_A12	TATGGCAC	TCGACAAG	227	0.00000072
B2_C9	GCTACTCT	TAGCCATG	226	0.00000072
C9_C7	TAGCCATG	CACTTCAC	223	0.00000071
A2_A7	ATCATGCG	GATCTTGC	223	0.00000071
B4_A8	GTCCTAAG	AGGATAGC	222	0.00000071
B8_A5	TCGGATTC	CGATCGAT	221	0.00000070
C2_B2	ATCGTGGT	GCTACTCT	220	0.00000070
C2_B1	ATCGTGGT	GTAGCGTA	215	0.00000068
C1_B7	GATCAAGG	TATGGCAC	215	0.00000068
A10_C3	TCGAGAGT	CGGTAATC	214	0.00000068

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
A2_A5	ATCATGCG	CGATCGAT	213	0.00000068
C3_B7	CGGTAATC	TATGGCAC	212	0.00000067
A8_C2	AGGATAGC	ATCGTGGT	210	0.00000067
C1_B1	GATCAAGG	GTAGCGTA	208	0.00000066
B4_A5	GTCCTAAG	CGATCGAT	208	0.00000066
A7_B7	GATCTTGC	TATGGCAC	207	0.00000066
C1_A8	GATCAAGG	AGGATAGC	206	0.00000066
C9_B9	TAGCCATG	AACAGCGA	204	0.00000065
C1_B10	GATCAAGG	AGAGTCCA	204	0.00000065
B2_B9	GCTACTCT	AACAGCGA	203	0.00000065
B2_A8	GCTACTCT	AGGATAGC	203	0.00000065
C7_C3	CACTTCAC	CGGTAATC	200	0.00000064
A3_A12	TGTTCCGT	TCGACAAG	200	0.00000064
C9_A12	TAGCCATG	TCGACAAG	199	0.00000063
C1_A7	GATCAAGG	GATCTTGC	198	0.00000063
B1_C1	GTAGCGTA	GATCAAGG	198	0.00000063
B10_C1	AGAGTCCA	GATCAAGG	197	0.00000063
A7_A2	GATCTTGC	ATCATGCG	195	0.00000062
A5_C9	CGATCGAT	TAGCCATG	194	0.00000062
A12_B4	TCGACAAG	GTCCTAAG	194	0.00000062
C1_B2	GATCAAGG	GCTACTCT	193	0.00000061
B4_C2	GTCCTAAG	ATCGTGGT	192	0.00000061
A12_A11	TCGACAAG	CTAGCTCA	191	0.00000061
C2_A5	ATCGTGGT	CGATCGAT	190	0.00000060
A1_C2	ACGATCAG	ATCGTGGT	190	0.00000060
A5_B4	CGATCGAT	GTCCTAAG	186	0.00000059
C9_B2	TAGCCATG	GCTACTCT	184	0.00000059
A7_A5	GATCTTGC	CGATCGAT	184	0.00000059
B1_A5	GTAGCGTA	CGATCGAT	182	0.00000058
A10_A7	TCGAGAGT	GATCTTGC	180	0.00000057
C7_A8	CACTTCAC	AGGATAGC	178	0.00000057
B9_B2	AACAGCGA	GCTACTCT	178	0.00000057
B8_A1	TCGGATTC	ACGATCAG	178	0.00000057
C9_C2	TAGCCATG	ATCGTGGT	177	0.00000056
C3_C7	CGGTAATC	CACTTCAC	177	0.00000056
A7_A11	GATCTTGC	CTAGCTCA	177	0.00000056
A12_A3	TCGACAAG	TGTTCCGT	177	0.00000056
C2_C9	ATCGTGGT	TAGCCATG	176	0.00000056
A5_C3	CGATCGAT	CGGTAATC	175	0.00000056
C3_A5	CGGTAATC	CGATCGAT	174	0.00000055
B2_C1	GCTACTCT	GATCAAGG	174	0.00000055
B10_C3	AGAGTCCA	CGGTAATC	174	0.00000055
C1_A5	GATCAAGG	CGATCGAT	172	0.00000055
B4_A7	GTCCTAAG	GATCTTGC	172	0.00000055
A5_B1	CGATCGAT	GTAGCGTA	172	0.00000055
C2_B9	ATCGTGGT	AACAGCGA	171	0.00000054
B8_C3	TCGGATTC	CGGTAATC	168	0.00000053
B2_A5	GCTACTCT	CGATCGAT	168	0.00000053
C9_B8	TAGCCATG	TCGGATTC	167	0.00000053
B10_B8	AGAGTCCA	TCGGATTC	166	0.00000053
C1_B9	GATCAAGG	AACAGCGA	165	0.00000053
B8_B10	TCGGATTC	AGAGTCCA	165	0.00000053

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
A5_B7	CGATCGAT	TATGGCAC	165	0.00000053
A1_C3	ACGATCAG	CGGTAATC	165	0.00000053
A8_C1	AGGATAGC	GATCAAGG	164	0.00000052
A5_A2	CGATCGAT	ATCATGCG	164	0.00000052
B10_C7	AGAGTCCA	CACTTCAC	163	0.00000052
C3_A10	CGGTAATC	TCGAGAGT	162	0.00000052
B2_C2	GCTACTCT	ATCGTGGT	162	0.00000052
A5_B2	CGATCGAT	GCTACTCT	161	0.00000051
A2_B8	ATCATGCG	TCGGATTC	161	0.00000051
A1_C7	ACGATCAG	CACTTCAC	160	0.00000051
A2_C3	ATCATGCG	CGGTAATC	159	0.00000051
C7_A5	CACTTCAC	CGATCGAT	158	0.00000050
A7_A3	GATCTTGC	TGTTCCGT	158	0.00000050
C2_B4	ATCGTGGT	GTCCTAAG	157	0.00000050
B4_A12	GTCCTAAG	TCGACAAG	156	0.00000050
A7_C1	GATCTTGC	GATCAAGG	156	0.00000050
A7_B10	GATCTTGC	AGAGTCCA	155	0.00000049
C7_B10	CACTTCAC	AGAGTCCA	153	0.00000049
B9_C7	AACAGCGA	CACTTCAC	151	0.00000048
B8_C9	TCGGATTC	TAGCCATG	149	0.00000047
A1_A7	ACGATCAG	GATCTTGC	149	0.00000047
A7_B9	GATCTTGC	AACAGCGA	147	0.00000047
A2_C7	ATCATGCG	CACTTCAC	146	0.00000046
A1_B8	ACGATCAG	TCGGATTC	146	0.00000046
C9_A7	TAGCCATG	GATCTTGC	145	0.00000046
A7_A1	GATCTTGC	ACGATCAG	145	0.00000046
C9_A5	TAGCCATG	CGATCGAT	143	0.00000046
C3_B10	CGGTAATC	AGAGTCCA	142	0.00000045
A7_B4	GATCTTGC	GTCCTAAG	142	0.00000045
C2_C3	ATCGTGGT	CGGTAATC	141	0.00000045
A5_C1	CGATCGAT	GATCAAGG	141	0.00000045
A12_B7	TCGACAAG	TATGGCAC	141	0.00000045
C7_A10	CACTTCAC	TCGAGAGT	140	0.00000045
B2_A7	GCTACTCT	GATCTTGC	140	0.00000045
A12_A8	TCGACAAG	AGGATAGC	139	0.00000044
A10_C7	TCGAGAGT	CACTTCAC	138	0.00000044
C9_C3	TAGCCATG	CGGTAATC	137	0.00000044
B10_A7	AGAGTCCA	GATCTTGC	137	0.00000044
A5_C2	CGATCGAT	ATCGTGGT	136	0.00000043
C3_C1	CGGTAATC	GATCAAGG	132	0.00000042
C3_B8	CGGTAATC	TCGGATTC	132	0.00000042
C3_A1	CGGTAATC	ACGATCAG	132	0.00000042
A5_B9	CGATCGAT	AACAGCGA	132	0.00000042
C3_A2	CGGTAATC	ATCATGCG	131	0.00000042
B9_C3	AACAGCGA	CGGTAATC	131	0.00000042
B8_B4	TCGGATTC	GTCCTAAG	129	0.00000041
B8_A2	TCGGATTC	ATCATGCG	129	0.00000041
A7_A10	GATCTTGC	TCGAGAGT	129	0.00000041
A7_B1	GATCTTGC	GTAGCGTA	128	0.00000041
B1_B8	GTAGCGTA	TCGGATTC	127	0.00000040
C7_A2	CACTTCAC	ATCATGCG	124	0.00000039
C2_B8	ATCGTGGT	TCGGATTC	124	0.00000039

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
C7_B8	CACTTCAC	TCGGATTC	118	0.00000038
B1_C3	GTAGCGTA	CGGTAATC	117	0.00000037
B1_A7	GTAGCGTA	GATCTTGC	117	0.00000037
A12_C1	TCGACAAG	GATCAAGG	117	0.00000037
C1_A12	GATCAAGG	TCGACAAG	116	0.00000037
A12_B10	TCGACAAG	AGAGTCCA	116	0.00000037
B4_C3	GTCCTAAG	CGGTAATC	115	0.00000037
B2_B8	GCTACTCT	TCGGATTC	115	0.00000037
B10_A12	AGAGTCCA	TCGACAAG	114	0.00000036
A7_C3	GATCTTGC	CGGTAATC	114	0.00000036
B2_C3	GCTACTCT	CGGTAATC	113	0.00000036
C7_B4	CACTTCAC	GTCCTAAG	112	0.00000036
C2_A7	ATCGTGGT	GATCTTGC	112	0.00000036
A8_A12	AGGATAGC	TCGACAAG	111	0.00000035
B9_A12	AACAGCGA	TCGACAAG	108	0.00000034
A5_C7	CGATCGAT	CACTTCAC	108	0.00000034
C3_B2	CGGTAATC	GCTACTCT	107	0.00000034
A7_B2	GATCTTGC	GCTACTCT	107	0.00000034
C3_B1	CGGTAATC	GTAGCGTA	106	0.00000034
B8_B1	TCGGATTC	GTAGCGTA	106	0.00000034
B8_C1	TCGGATTC	GATCAAGG	104	0.00000033
A5_A7	CGATCGAT	GATCTTGC	104	0.00000033
C1_C3	GATCAAGG	CGGTAATC	102	0.00000032
A5_B8	CGATCGAT	TCGGATTC	102	0.00000032
C3_B9	CGGTAATC	AACAGCGA	101	0.00000032
C3_B4	CGGTAATC	GTCCTAAG	101	0.00000032
A7_C9	GATCTTGC	TAGCCATG	101	0.00000032
B4_C7	GTCCTAAG	CACTTCAC	100	0.00000032
B8_C7	TCGGATTC	CACTTCAC	99	0.00000032
C7_B9	CACTTCAC	AACAGCGA	97	0.00000031
C7_B1	CACTTCAC	GTAGCGTA	97	0.00000031
B1_A12	GTAGCGTA	TCGACAAG	96	0.00000031
A7_C7	GATCTTGC	CACTTCAC	96	0.00000031
A5_A12	CGATCGAT	TCGACAAG	96	0.00000031
C3_A7	CGGTAATC	GATCTTGC	95	0.00000030
C2_A12	ATCGTGGT	TCGACAAG	94	0.00000030
B8_C2	TCGGATTC	ATCGTGGT	94	0.00000030
C7_C1	CACTTCAC	GATCAAGG	93	0.00000030
A12_B8	TCGACAAG	TCGGATTC	93	0.00000030
A12_A5	TCGACAAG	CGATCGAT	93	0.00000030
C1_B8	GATCAAGG	TCGGATTC	91	0.00000029
B8_B2	TCGGATTC	GCTACTCT	91	0.00000029
A12_B2	TCGACAAG	GCTACTCT	90	0.00000029
C7_C2	CACTTCAC	ATCGTGGT	89	0.00000028
C7_A7	CACTTCAC	GATCTTGC	85	0.00000027
B1_C7	GTAGCGTA	CACTTCAC	85	0.00000027
C7_B2	CACTTCAC	GCTACTCT	83	0.00000026
C1_C7	GATCAAGG	CACTTCAC	83	0.00000026
B2_A12	GCTACTCT	TCGACAAG	83	0.00000026
A12_B1	TCGACAAG	GTAGCGTA	82	0.00000026
B2_C7	GCTACTCT	CACTTCAC	81	0.00000026
A12_B9	TCGACAAG	AACAGCGA	81	0.00000026

MatchOrMismatch	Bar1Seq	Bar2Seq	Number	Proportion
C2_C7	ATCGTGGT	CACTTCAC	79	0.00000025
A12_C2	TCGACAAG	ATCGTGGT	79	0.00000025
B8_A12	TCGGATTC	TCGACAAG	75	0.00000024
A7_B8	GATCTTGC	TCGGATTC	75	0.00000024
A7_C2	GATCTTGC	ATCGTGGT	72	0.00000023
C3_A12	CGGTAATC	TCGACAAG	59	0.00000019
C7_A12	CACTTCAC	TCGACAAG	53	0.00000017
A12_A7	TCGACAAG	GATCTTGC	48	0.00000015
A12_C3	TCGACAAG	CGGTAATC	46	0.00000015
A7_A12	GATCTTGC	TCGACAAG	40	0.00000013
A12_C7	TCGACAAG	CACTTCAC	25	0.00000008