file = original

goodFile = file – index hopping – “N” in index

betterFile = goodFile – low mean QS

bestFile = final

read = all 4 lines associated with fastq file

Open all files and be on the same line in each file

For line in files:

read1 = readReads(R1)

read2 = readReads(R2)

Index1 = ReadIndex(R2)

Index2 = ReadIndex(R3)

concatIndex (read1, index1)

concatIndex (read2, index2)

If matching(index1, index2) == FALSE

writeToBadFile(badFile1, read1, IH)

writeToBadFile(badFile2, read2, IH)

hoppingCounter += 1

break

elif containsN(index1) == TRUE

writeToBadFile(badFile1, read1, N)

writeToBadFile(badFile2, read2, N)

else

writeToGoodFile(goodFile1, read1)

writeToGoodFile(goodFile2, read2)

For each file (goodFile1, goodFile2)

read1 = readReads(goodFile1)

read2 = readReads(goodFile2, LN)

if compareMeans(meanQS(read1)) == FALSE

writeToBadFile(badFile1, read1, QS)

writeToBadFile(badFile2, read2, QS)

else

writeToGoodFile(betterFile1, read1)

writeToGoodFile(betterFile2, read2)

For each file (betterFile1, betterFile2)

read1 = readReads(betterFile1)

read2 = readReads(betterFile2)

index = getIndex(read1)

writeToGoodFile(index.f.txt, read1)

writeToGoodFile(index.r.txt, read2)