

Index Input File

↓ pull last 2 columns

Create bucket  
dict

key = Index c - last column

value = Bin c - 2nd last column  
↳ 26 bins total

↓  
Input files

R1                      I1                      I2                      R2  
↓                      ↓                      ↓                      ↓

• Header	• R2	• I2
• R2	• I1	• quality

↓  
reverse complement I2

↓  
Compare I1 + I2 by Hamming function

If meets  
threshold

↓  
get Bin <sup>From Bucket Dict</sup>

Fastq 2 (r1)      Fastq 2 (r2)

If miss threshold

↓  
get Swap Bin <sup>From Bucket Dict</sup>

Fastq 2 (r1)      Fastq 2 (r2)

If bad

↓  
get bad Bin <sup>From Bucket Dict</sup>

Fastq 2 (r1)      Fastq 2 (r2)