

Count sequences around a specific target (e.g., AGT) in a CGR plot

Source code: count_CGR_v6.cpp

Example input file: CGR_coord.txt

Example output file: AGT.txt

Instructions:

1) Compile source code to app called CGR using the line command:

```
g++ count_CGR_v6.cpp -o count_CGR
```

2) Run the program by invoking the following line command:

```
./count_CGR 16 AGT CGR_coord.txt AGT.txt
```

Note: '16' is the resolution around the x- and y- coordinate of the target (in this case AGT). The lower the resolution (e.g., 4), the more sequences will be found around the target. The higher the resolution (e.g., 512), the fewer sequences will be found. Using a resolution of 16 yielded 20 similar sequences (out of the 4297 possible) (see Terminal output below). The outfile (AGT.txt) contains the source file, the target, the x- and y-coordinates of the target, the number of 'similar' sequences around the target, and the total number of nucleotides in the sequence.

Terminal output.

```
CGR_coord.txt  AGT  x=0.125 y=0.375
TGAGCGCTGTGCGAACGTCTCACAGCAGTTCGGATGTTGCTCTTTCCTCAAGT
GTTGCAGGTCTATGGAAATGCAGGATCTAACCAGCCCAGCATAGCCGACTGAGT
GGTCTATGGAAATGCAGGATCTAACCAGCCCAGCATAGCCGACTGAGTGGTAGT
CAATGAGCAGCAGTGAAATAGCTTCAACAGCAGCAGACGGGTCTTTAGACAGT
TATTCTCCCTACCCCTTCTCACAAACTATGGCTGCATATGGGCCAAACACAGT
CTCCTATGGTGCATTGTGGGCAGGCATCAAGACGGAAAGTGGATTGTCACAGT
CAGGATTATATTCAGGAAATAATTCACCAACTCCTCCGGATTCAACAGT
ACAGCAGGACTATCCGTCTTATCCCGGCTTTGGCCAGGGTCAGTACGCACAGT
CGTCCACCAATGCCACTTACCAACTCCAGGAACCACCTTCTGGCGTCACAAGT
CACCTTCTGGCGTCACAAGTCAGGCGGTACAGACCCACAGCAGAGTACAGT
TCACAAGTCAGGCGGTACAGACCCACAGCAGAGTACAGTACAATCCACAGT
AACTTGGCAGACACACATCTATTTTTCAATGACCTAGAAGAGTGTGACCAAGT
AAGTCCTGCTATATGGATTAGGAATTGTGTTTCCAATAGAAAATATTTACAGT
CTCTAGTACTGTGAATCCAGTGAAAGTAAGCCATGAGAATGTTCTCACACAGT
TTGAAGGGACAATAGCATAAGAAATAATCAATAAAAAGGCCTTGGCCTGACAGT
CAATAGCATAAGAAATAATCAATAAAAAGGCCTTGGCCTGACAGTACAGCAAGT
TTGGCCTGACAGTACAGCAAGTACTTTGAATTTTAGCACATTGCAAAAAGTAGT
ATAATGGATCTTTTTGATAATGTTATCTAAGACCAAAAAGCATGAATGTCAAGT
TTTATTTTCTTAGTCTTTTATCCCATTCATTTTCTAATTTGTGTACGTGAGT
GTGTGTGTGTGTGTATTGGTTAGTTAGAAGGGACTTGATAAAAATAACAGT
```