CHALMERS UNIVERSITY OF TECHNOLOGY

Examination in Bioinformatics, MVE360

Monday 10 March 2014, 08:30-12:30

Solutions

Updated 2017-03-08

Question 1.
5 p

a)
- i) 3
- ii) 2

See lecture notes for definitions of these two distances.

b) Solutions showing the full score matrices can be obtained using these on-line programs:

- i) http://www.cse.chalmers.se/∼kemp/cgi-bin/global_alignment.html
- ii) http://www.cse.chalmers.se/∼kemp/cgi-bin/local_alignment.html

Question 3.
12 p

a) 28

b)

```
@sequences = <DATA>;

$numSequences = $#sequences + 1;

for ( $i = 0 ; $i < $numSequences ; ++$i ) {
    chomp($sequences[$i]);
}

$numColumns = length($sequences[0]);

$score = 0;

for ( $column = 0 ; $column < $numColumns ; ++$column ) {
    for ( $i = 0 ; $i < $numSequences - 1 ; ++$i ) {
        for ( $j = $i+1 ; $j < $numSequences ; ++$j ) {
            $pair = substr($sequences[$i], $column, 1) .
                    substr($sequences[$j], $column, 1);
            $score += $blosum62{$pair};
        }
    }
}

print "Alignment score is $score\n";
```

Question 4.
3 p

a) `VVVAA`
b) `VLVGVVV`
c) `V, A`
d) `XXXXATLSA`
e) `XGXAATXSA`
f) `ALAGAAATTTLST`

Question 5. a) `[AC].V.{4}[^ED]`
9 p

`^A.[ST]{2}.{0,1}V`

b)
```
#!/usr/bin/perl

print "Type in a PROSITE pattern: ";
$prosite_pattern = <STDIN>;
chomp($prosite_pattern);
$prosite_pattern =~ s/\.$//;

@pattern_elements = split("-", $prosite_pattern);

foreach $e (@pattern_elements) {
    $e =~ s/\{(.*)\}/[^\1]/;
    $e =~ s/\((.*)\)/{\1}/;
    $e =~ s/x/./;
    $e =~ s/^</^/;
    $e =~ s/>$/\$/;
    push(@perl_pattern_elements, $e);
}

$perl_pattern = join("", @perl_pattern_elements);

if ( $sequence =~ /$perl_pattern/ ) {
    print "Sequence matches\n";
} else {
    print "Sequence does not match\n";
}
```