

Bioinformatics – Global Max Similarity Matrix Paths

Conor Heffron

Overview:

- Compute the global max similarity matrix *and* best global upmost and downmost alignments between the two sequences:
 - A A G T G C C T C A A G A T A
 - A C C G T C T C A G C A A T A

Manual workings / Notes

	-	A	A	G	T	G	C	C	T	C	A	A	G	A	T	A
-	0	-2	-4	-6	-8	-10	-12	-14	-16	-18	-20	-22	-24	-26	-28	-30
A		A	A								A	A		A		A
C							C	C		C						
C							C	C		C						
G				G		G							G			
T																
C							C	C		C						
T					T				T						T	
C							C	C		C						
A		A	A								A	A		A		A
G				G		G							G			
C							C	C		C						
A		A	A								A	A		A		A
A		A	A								A	A		A		A
T					T				T						T	
A		A	A								A	A		A		A

- A A G T G C C T C A A G A T A
- A C C G T C T C A G C A A T A
- + - - - - + - - - - - + + +
- 5 matches, 10 mismatches, zero gaps
- score = $5x1 + 10x(-1) + 0x(-2) = -5$

- A - A G T G C C T C A A G A T A
- A C C G T C T C A G C A - A T A
- + = - + + - - + - - - + = + + +
- 8 matches, 5 mismatches, 2 gaps
- score = $8x1 + 6x(-1) + 2x(-2) = -2$

- - A A G T G C C T C A A G A T A
- A C C G T C T C A G C A - A T A
- = - - + + - - + - - - + = + + +
- 7 matches, 7 mismatches, 2 gaps
- score = $7x1 + 7x(-1) + 2x(-2) = -4$

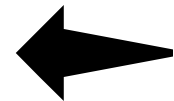
Global Maximum Similarity Matrix From R Code Output:

- Note: Paths Added Manually

Path 1

| | | | | | | | | | | | | | | | | |
|---|---|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | - | A | A | G | T | G | C | C | T | C | A | A | G | A | T | A |
| - | 0 | -2 | -4 | -6 | -8 | -10 | -12 | -14 | -16 | -18 | -20 | -22 | -24 | -26 | -28 | -30 |

| | | | | | | | | | | | | | | | | |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| A | -2 | 1 | -1 | -3 | -5 | -7 | -9 | -11 | -13 | -15 | -17 | -19 | -21 | -23 | -25 | -27 |
| C | -4 | -1 | 0 | -2 | -4 | -6 | -8 | -10 | -12 | -12 | -14 | -16 | -18 | -20 | -22 | -24 |
| C | -6 | -3 | -2 | -1 | -1 | -3 | -5 | -7 | -9 | -11 | -11 | -13 | -15 | -17 | -19 | -21 |
| G | -8 | -5 | -4 | -3 | -2 | 0 | -2 | -4 | -6 | -8 | -10 | -12 | -14 | -16 | -16 | -18 |
| T | -10 | -7 | -6 | -5 | -2 | -2 | -1 | -3 | -5 | -7 | -7 | -9 | -11 | -13 | -15 | -17 |
| C | -12 | -9 | -6 | -5 | -4 | -3 | -1 | -2 | -2 | -4 | -6 | -6 | -8 | -10 | -12 | -14 |
| T | -14 | -11 | -8 | -5 | -6 | -5 | -2 | -2 | -1 | -3 | -5 | -5 | -7 | -9 | -11 | -13 |
| C | -16 | -13 | -10 | -7 | -6 | -5 | -4 | -1 | -3 | -2 | -4 | -6 | -6 | -8 | -8 | -10 |
| A | -18 | -15 | -12 | -9 | -8 | -7 | -4 | -3 | 0 | -2 | -3 | -3 | -5 | -7 | -9 | -9 |
| G | -20 | -17 | -14 | -11 | -10 | -9 | -6 | -5 | -2 | 1 | -1 | -3 | -2 | -4 | -6 | -8 |
| C | -22 | -19 | -16 | -13 | -12 | -11 | -8 | -7 | -4 | -1 | 0 | -2 | -2 | -1 | -3 | -5 |
| A | -24 | -21 | -18 | -15 | -12 | -13 | -10 | -9 | -6 | -3 | 0 | -1 | -3 | -3 | -2 | -4 |
| A | -26 | -23 | -20 | -17 | -14 | -13 | -12 | -11 | -8 | -5 | -2 | -1 | 0 | -2 | -4 | -1 |
| T | -28 | -25 | -22 | -19 | -16 | -13 | -14 | -11 | -10 | -7 | -4 | -3 | -2 | -1 | -1 | -3 |
| A | -30 | -27 | -24 | -21 | -18 | -15 | -14 | -13 | -12 | -9 | -6 | -5 | -2 | -1 | -2 | 0 |



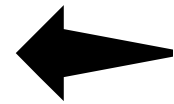
0 -> -1 -> -2 -> -3 -> -2 -> -1 -> 1 -> 0 -> -1 -> -2 -> -3 -> -2 -> -3 -> -2 -> -1 -> -2 -> 0

Score = 0

Alternative Path 2

| | | | | | | | | | | | | | | | | |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| - | 0 | A | A | G | T | G | C | C | T | C | A | A | G | A | T | A |
| A | -2 | 1 | -1 | -3 | -5 | -7 | -9 | -11 | -13 | -15 | -17 | -19 | -21 | -23 | -25 | -27 |
| C | -4 | -1 | 0 | -2 | -4 | -6 | -8 | -10 | -12 | -12 | -14 | -16 | -18 | -20 | -22 | -24 |
| C | -6 | -3 | -2 | -1 | -1 | -3 | -5 | -7 | -9 | -11 | -11 | -13 | -15 | -17 | -19 | -21 |
| G | -8 | -5 | -4 | -3 | -2 | 0 | -2 | -4 | -6 | -8 | -10 | -12 | -14 | -16 | -16 | -18 |
| T | -10 | -7 | -6 | -5 | -2 | -2 | -1 | -3 | -5 | -7 | -7 | -9 | -11 | -13 | -15 | -17 |
| C | -12 | -9 | -6 | -5 | -4 | -3 | -1 | -2 | -2 | -4 | -6 | -6 | -8 | -10 | -12 | -14 |
| T | -14 | -11 | -8 | -5 | -6 | -5 | -2 | -2 | -1 | -3 | -5 | -5 | -7 | -9 | -11 | -13 |
| C | -16 | -13 | -10 | -7 | -6 | -5 | -4 | -1 | -3 | -2 | -4 | -6 | -6 | -8 | -8 | -10 |
| A | -18 | -15 | -12 | -9 | -8 | -7 | -4 | -3 | 0 | -2 | -3 | -3 | -5 | -7 | -9 | -9 |
| G | -20 | -17 | -14 | -11 | -10 | -9 | -6 | -5 | -2 | 1 | -1 | -3 | -2 | -4 | -6 | -8 |
| C | -22 | -19 | -16 | -13 | -12 | -11 | -8 | -7 | -4 | -1 | 0 | -2 | -2 | -1 | -3 | -5 |
| A | -24 | -21 | -18 | -15 | -12 | -13 | -10 | -9 | -6 | -3 | 0 | -1 | -3 | -3 | -2 | -4 |
| A | -26 | -23 | -20 | -17 | -14 | -13 | -12 | -11 | -8 | -5 | -2 | -1 | 0 | -2 | -4 | -1 |
| T | -28 | -25 | -22 | -19 | -16 | -13 | -14 | -11 | -10 | -7 | -4 | -3 | -2 | -1 | -1 | -3 |

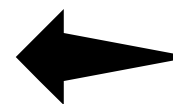
| | | | | | | | | | | | | | | | | |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|----|----|----|----|---|
| A | -30 | -27 | -24 | -21 | -18 | -15 | -14 | -13 | -12 | -9 | -6 | -5 | -2 | -1 | -2 | 0 |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|----|----|----|----|---|



0 -> -1 -> -1 -> 0 -> -1 -> 0 -> 1 -> 0 -> -1 -> -2 -> -1 -> -2 -> -2 -> -1 -> 0 -> 1 -> 0

Alternative Path 3

| | - | A | A | G | T | G | C | C | T | C | A | A | G | A | T | A |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| - | 0 | -2 | -4 | -6 | -8 | -10 | -12 | -14 | -16 | -18 | -20 | -22 | -24 | -26 | -28 | -30 |
| A | -2 | 1 | -1 | -3 | -5 | -7 | -9 | -11 | -13 | -15 | -17 | -19 | -21 | -23 | -25 | -27 |
| C | -4 | -1 | 0 | -2 | -4 | -6 | -8 | -10 | -12 | -12 | -14 | -16 | -18 | -20 | -22 | -24 |
| C | -6 | -3 | -2 | -1 | -1 | -3 | -5 | -7 | -9 | -11 | -11 | -13 | -15 | -17 | -19 | -21 |
| G | -8 | -5 | -4 | -3 | -2 | 0 | -2 | -4 | -6 | -8 | -10 | -12 | -14 | -16 | -16 | -18 |
| T | -10 | -7 | -6 | -5 | -2 | -2 | -1 | -3 | -5 | -7 | -7 | -9 | -11 | -13 | -15 | -17 |
| C | -12 | -9 | -6 | -5 | -4 | -3 | -1 | -2 | -2 | -4 | -6 | -6 | -8 | -10 | -12 | -14 |
| T | -14 | -11 | -8 | -5 | -6 | -5 | -2 | -2 | -1 | -3 | -5 | -5 | -7 | -9 | -11 | -13 |
| C | -16 | -13 | -10 | -7 | -6 | -5 | -4 | -1 | -3 | -2 | -4 | -6 | -6 | -8 | -8 | -10 |
| A | -18 | -15 | -12 | -9 | -8 | -7 | -4 | -3 | 0 | -2 | -3 | -3 | -5 | -7 | -9 | -9 |
| G | -20 | -17 | -14 | -11 | -10 | -9 | -6 | -5 | -2 | 1 | -1 | -3 | -2 | -4 | -6 | -8 |
| C | -22 | -19 | -16 | -13 | -12 | -11 | -8 | -7 | -4 | -1 | 0 | -2 | -2 | -1 | -3 | -5 |
| A | -24 | -21 | -18 | -15 | -12 | -13 | -10 | -9 | -6 | -3 | 0 | -1 | -3 | -3 | -2 | -4 |
| A | -26 | -23 | -20 | -17 | -14 | -13 | -12 | -11 | -8 | -5 | -2 | -1 | 0 | -2 | -4 | -1 |
| T | -28 | -25 | -22 | -19 | -16 | -13 | -14 | -11 | -10 | -7 | -4 | -3 | -2 | -1 | -1 | -3 |
| A | -30 | -27 | -24 | -21 | -18 | -15 | -14 | -13 | -12 | -9 | -6 | -5 | -2 | -1 | -2 | 0 |



0 -> -1 -> -2 -> 0 -> -1 -> 0 -> 1 -> 0 -> -1 -> -2 -> -3 -> -2 -> -3 -> -1 -> 0 -> 1 -> 0