

## Algoritmi dinamici di allineamento di sequenze: come si trova il percorso finale?

(vedere Figure 3.5 e 3.7)

### Domanda

Abbiamo capito come si calcolano i punteggi con il metodo dinamico illustrato nelle figure 3.5 e 3.7, ma una volta trovato il punteggio massimo (ad esempio il valore 66 della figura 3.5), con che logica l'algoritmo trova il percorso "inverso", illustrato con le freccette nere nelle due figure del testo? Perché, ad esempio, per passare dalla casella con il valore 66 a quella con 53, si passa per la casella 56 e non per la casella 57 che ha un valore più alto? È chiaro che dal punto di vista dell'allineamento, si perderebbe l'appaiamento delle due treonine, ma chi glielo dice all'algoritmo?

### Risposta

Ci sono diversi metodi possibili. Il più semplice è quasi banale dal momento che non richiede un algoritmo per calcolare *ex novo* il percorso "inverso", ma memorizza il percorso direttamente nella fase di calcolo dinamico dei punteggi. In altre parole, ogni volta che si stabilisce la casella di provenienza tra le tre possibili (vedi spiegazione nel testo) il programma associa alla casella non solo il risultante punteggio totalizzato (scritto in nero nella figura), ma anche il percorso di provenienza (rappresentato dalla freccetta nella seguente tabella, modificata dalla figura 3.5).

	V	Q	T	Y
D	49	47	42	37
N	44	53	51	46
A	42	48	57	53
T	37	45	56	59
Y	32	40	51	66

Alla fine della procedura dinamica di calcolo dei punteggi, si dovrà individuare la casella con il punteggio massimo (66 nell'esempio illustrato), dalla quale si dovrà procedere a ritroso, seguendo il "filo di Arianna" lasciato precedentemente. Quindi il risultante percorso sarà dato dalle caselle con i seguenti punteggi: 66, 56, 48, 53, 49.

Come spiegato nel testo, è importante ricordare che un aminoacido (o una base nel caso di acidi nucleici) non si può appaiare in più di una posizione. Quindi, quando il percorso si muove in direzione verticale o orizzontale, allora si dovrà considerare solo una posizione per l'appaiamento. Ad esempio, il percorso passa per le caselle con i punteggi 48 e 53 (direzione verticale). Il programma dovrà pertanto valutare se la Q debba essere appaiata con la N oppure con la A. Si può vedere che la N presenta il punteggio più alto (4 punti per l'appaiamento Q-N, contro soli 3 punti per l'appaiamento Q-A), quindi la N sarà allineata alla Q. Il risultante allineamento sarà:

```

... V Q - T Y
   | |   | |
... D N A T Y
    
```