

Relazione sul tool mVista: uso e funzionamento

VISTA è una suite completa di programmi e database per l'analisi comparativa delle sequenze genomiche.

Panoramica su mVISTA

mVISTA è un insieme di programmi per il confronto delle sequenze di DNA provenienti da due o più specie, lunghe anche varie migliaia di basi e visualizzati con varie informazioni. mVista permette una facile identificazione delle somiglianze e delle differenze delle varie sequenze, anche visivamente permettendo di visualizzare le varie lunghezze a diversi livelli di risoluzione. E' implementato come un server on-line che fornisce l'accesso a livello mondiale di molteplici strumenti di allineamento. Il modulo di mVISTA è progettato per visualizzare allineamenti globali di sequenze genomiche di specie diverse e determinerà l'identità percentuale tra due sequenze utilizzando una finestra scorrevole di lunghezza predefinita, visualizzandola come una curva continua. Il programma individua anche i colori e le regioni di alta conservazione.

mVista utilizza vari programmi per l'allineamento i principali sono:

- AVID è un programma per l'allineamento globale delle sequenze di DNA di lunghezza arbitraria. Le caratteristiche principali dell'algoritmo sono che può allineare centinaia di kilobasi rapidamente, la sua precisione e abilità per rilevare le omologie è però debole. Il programma funziona in modo ricorsivo per trovare via via gli allineamenti migliori.
- LAGAN è un programma per l'allineamento globale a coppie e per allineamenti multipli di sequenze. LAGAN esegue progressivo allineamento a coppie, diretto da un albero filogenetico.

Utilizzo di mVISTA

L'utilizzo di mVISTA per l'analisi comparativa di due o più sequenze avviene introducendo tali sequenze in formato FASTA (solo testo), insieme ad un file di annotazione del gene e inoltrarle al server Web.

Funzionamento nel dettaglio

In ingresso come prima cosa ci chiede il numero di specie che vogliamo comparare, il server mVista è in grado di elaborare da 2 a 100.

Segue la pagina di presentazione principale che contiene i campi dove inserire le sequenze da analizzare, che possiamo poi caratterizzare con un nome (per la specie scelta, che sarà poi mostrata nella legenda), annotazioni varie (se l'annotazione di un gene è disponibile si può presentare in un semplice formato testo da visualizzare sul grafico), e fare anche un'analisi delle sequenze ripetute e anche un'analisi con rVISTA (programma che cerca potenziali elementi regolatori nelle regioni non codificanti del genoma umano); le dimensioni totali delle sequenze non devono superare le 10 megabasi e nessuna sequenza deve essere più lunga di 2.

A seconda delle specifiche della vostra analisi è possibile scegliere uno dei programmi di allineamento seguenti:

-AVID, allineamento globale a coppie. Per tutte le sequenze nel set, AVID genera allineamenti a coppie "tutti-contro-tutti", che possono essere visualizzate utilizzando qualsiasi sequenza come una base (di riferimento). Se alcune delle sequenze sono in un formato bozza, Avid generato il loro allineamento con la sequenza completa che sarà utilizzata come base.

-LAGAN, allineamento globale a coppie e multiplo di sequenze. Se alcune delle sequenze sono in un formato bozza verranno prima reindirizzate su AVID per ottenere il migliore allineamento a coppie. Mentre l'allineamento multiplo sarà visualizzata da Vista, che calcola e visualizza la conservazione delle sequenze su una qualsiasi sequenza.

-Shuffle-LAGAN - allineamenti globali e locali delle sequenze. Rileva riarrangiamenti e inversioni in sequenze.

Alcuni minuti dopo aver inoltrato le sequenze al server è possibile ricevere una mail con un link web personale da cui si può accedere ai risultati delle analisi.

In questo link sono riportati i risultati della ricerca in tre opzioni di visualizzazione. Queste tre opzioni sono: Text Browser , che fornisce tutte le informazioni dettagliate - sequenze, allineamenti, le statistiche per le sequenze conservate, ecc; VISTA Browser, che è uno strumento di visualizzazione interattiva che può essere utilizzato per navigare in modo dinamico sugli allineamenti risultanti e regolare la curva VISTA e in ultimo, un file PDF, che è un incrocio statico degli alliniamenti VISTA.