

Tóm tắt Khóa luận tốt nghiệp

XÂY DỰNG CƠ SỞ DỮ LIỆU HAI GENE *HSP-70* VÀ *REVERSE TRANSCRIPTASE-RNaseH* Ở MỘT SỐ LOÀI VIRUS THỰC VẬT

Sinh viên: Nguyễn Văn Thái

Khóa: 2001 – 2005

Với sự phát triển của kỹ thuật giải trình tự, một số lượng lớn các gene *hsp-70* và *RT-RNaseH* đã được giải trình tự. Những trình tự gene này được lưu trữ trong CSDL sinh học lớn như NCBI, EMBL, DDBj,... Vì các CSDL này quá lớn và chứa rất nhiều thông tin khác nhau, không tập trung thành từng gene cụ thể nên khó có thể thực hiện việc truy xuất các thông tin phục vụ trực tiếp cho một nghiên cứu chuyên biệt. Do vậy, mục tiêu của chúng tôi là tiến hành xây dựng cơ sở dữ liệu hai gene *hsp-70* và *reverse transcriptase-RNaseH* ở một số loài virus thực vật.

Để đạt được mục tiêu trên, khóa luận cần đảm bảo thực hiện nội dung như sau:

- Dùng Perl script để thu nhận trình tự các nucleotide và protein của hai gene từ trang CSDL GenBank (NCBI cơ sở dữ liệu nucleotide).
- Xác định gene và protein của hai gene *hsp-70* và *Reverse transcriptase-RNaseH* (*RT-RNaseH*) trong genome hay ORF (Open Reading Frame) của virus.
- Tìm hiểu về mô hình dữ liệu quan hệ, sử dụng mô hình này vào việc lưu trữ dữ liệu các trình tự nucleotide và protein của hai gene, tạo CSDL hai gene này. Dùng Perl script để chuyển tự động các dữ liệu vào CSDL.
- Sử dụng giao thức CGI kết hợp với ngôn ngữ lập trình Perl, để thiết kế trang web CSDL về hai gene *hsp-70* và *RT-RNaseH* ở trên hai họ virus *Closteroviridae* và *Caulimoviridae*.

Sau khi thực hiện các nội dung trên chúng tôi đạt được những kết quả như sau:

- Chúng tôi đã tải được 325 trình tự gene *hsp-70* và *RT-RNaseH* từ cơ sở dữ liệu NCBI.
- Thông qua việc tìm hiểu về hai họ virus, trình tự gene tương đồng, trình tự protein bảo tồn và kết hợp với ClustalW. Chúng tôi đã xác định được vị trí gene *hsp-70* và *RT-RNaseH* trong ORF hay nằm trong genome của chúng.
- CSDL có 325 trình tự được tích hợp với Web.
- Trang Web CSDL gene *hsp-70* và *RT-RNaseH* gồm có 6 trang chính, đó là HOME, SEARCH, TOOL, TAXONOMY, LINK, ABOUT PAGE. Ngoài ra, từ những trang web chính này còn có thể kết nối đến những trang phụ khác để cung cấp những tiện ích cho người dùng. Từ các trang web này, người sử dụng có thể truy xuất thông tin, so sánh một trình tự quan tâm với các trình tự trong cơ sở dữ liệu gene *hsp-70* và *RT-RNaseH*, tìm kiếm trình tự, các đặc tính của loài,...

