

CÁC PHƯƠNG PHÁP XÂY DỰNG MA TRẬN BIẾN ĐỔI AXÍT AMIN

1. Họ và tên nghiên cứu sinh: ĐẶNG CAO CƯỜNG
2. Giới tính: Nam
3. Ngày sinh: 25/10/1983
4. Nơi sinh: Hải Phòng
5. Quyết định công nhận nghiên cứu sinh số: 3613/SĐH ngày 22 tháng 10 năm 2009 của Giám đốc Đại học Quốc gia Hà Nội.
6. Các thay đổi trong quá trình đào tạo: Không
7. Tên đề tài luận án: Các phương pháp xây dựng ma trận biến đổi axit amin
8. Chuyên ngành: Khoa học Máy tính
9. Mã số: 62 48 01 01
10. Cán bộ hướng dẫn khoa học: TS. Lê Sỹ Vinh, TS. Lê Sĩ Quang

11. Tóm tắt các **kết quả mới** của luận án:

- 1) Đề xuất cải tiến cho phương pháp ước lượng mô hình bằng ML để giảm thời gian chạy và tăng độ chính xác bằng cách chia tách nhỏ dữ liệu đầu vào bằng hai phương pháp ngẫu nhiên và dựa theo cấu trúc cây phân loài. Ưu điểm của thuật toán cải tiến được kiểm định bằng thực nghiệm trên các bộ dữ liệu chuẩn của Pfam và virút cúm.
- 2) Đề xuất một cải tiến khác để giảm thời gian ước lượng bằng cách giảm thiểu các bước xây dựng cây và tối ưu tham số. Kết quả thực nghiệm cho thấy thời gian giảm được 2 lần so với phương pháp gốc.
- 3) Dựa trên các thuật toán đã cải tiến để xây dựng một hệ thống ước lượng tự động cung cấp cho người dùng nhiều tùy chọn và kết quả. Thực nghiệm cho thấy hiệu quả đáp ứng rất tốt của hệ thống với lượng người dùng lớn.
- 4) Cuối cùng, luận án đưa ra mô hình FLU áp dụng cho virút cúm. Thực nghiệm cho thấy FLU tốt hơn hẳn tất cả các mô hình hiện nay khi ứng dụng vào nghiên cứu các chuỗi prôtêin của virút cúm.

12. Khả năng ứng dụng trong thực tiễn:

Các mô hình biến đổi axit amin có nhiều ứng dụng trong thực tiễn giúp tìm kiếm các chuỗi prôtêin trong cơ sở dữ liệu, xây dựng các cây phân loài, ...

13. Những hướng nghiên cứu tiếp theo:

Xây dựng mô hình biến đổi axit amin cho các loài sinh vật khác.

14. Các công trình đã công bố có liên quan đến luận án:

1.Cuong DC, Quang LS, Gascuel O, and Vinh LS (2010), “FLU, an amino acid substitution model for influenza proteins”, *BMC Evolutionary Biology* vol. 10 (1), p. 99-110.

2.Cuong DC, Lefort V, Vinh LS, Quang LS and Gascuel O (2011), “ReplacementMatrix: a web server for maximum-likelihood estimation of amino acid replacement rate matrices”, *Bioinformatics* vol. 27 (19), pp. 2758–2760.

3.Dat LV, Cuong DC, Quang LS and Vinh LS (2011), “A Fast and Efficient Method for Estimating Amino Acid Substitution Models”, *Proc. of the 2011 Third International Conference on Knowledge and Systems Engineering*, pp. 85 –91.

4.Sau NV, Cuong DC, Quang LS and Vinh LS (2011), “Protein Type Specific Amino Acid Substitution Models for Influenza Viruses”, *Proc. of the 2011 Third International Conference on Knowledge and Systems Engineering*, pp. 98 –103.

5.Quang LS, Cuong DC, and Gascuel O (2012), “Modeling Protein Evolution with Several Amino Acid Replacement Matrices Depending on Site Rates”, *Mol Biol Evol* vol. 29(10), pp. 2921–2936.