BỘ GIÁO DỤC VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ

Nguyễn Thị Thủy Trâm

NGHIÊN CỨU ĐẶC ĐIỂM BỘ GENE TY THỂ CỦA MỘT SỐ LOÀI LINH TRƯỞNG Ở TÂY NGUYÊN

LUẬN ÁN TIẾN SĨ SINH HỌC ỨNG DỤNG

Hà Nội - 2026

BỘ GIÁO DỤC VÀ ĐÀO TẠO **VIỆN HÀN LÂM KHOA HỌC VÀ CÔNG NGHỆ VIỆT NAM**

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ

Nguyễn Thị Thủy Trâm

NGHIÊN CỨU ĐẶC ĐIỂM BỘ GENE TY THỂ CỦA MỘT SỐ LOÀI LINH TRƯỞNG Ở TÂY NGUYÊN

LUẬN ÁN TIẾN SĨ SINH HỌC ỨNG DỤNG

Ngành: Công nghệ sinh học

Mã số: 9 42 02 01

Xác nhận của Học viện Khoa học và Công nghệ	Người hướng dẫn 1	Người hướng dẫn 2
	PGS.TS. Lê Thành Long	PGS.TS. Thái Kế Quân

LỜI CAM ĐOAN

Tôi xin cam đoan luận án: **"Nghiên cứu đặc điểm bộ gene ty thể của một số loài Linh trưởng ở Tây Nguyên"** là công trình nghiên cứu của chính mình dưới sự hướng dẫn khoa học của tập thể hướng dẫn. Luận án sử dụng thông tin trích dẫn từ nhiều nguồn tham khảo khác nhau và các thông tin trích dẫn được ghi rõ nguồn gốc. Các kết quả nghiên cứu của tôi được công bố chung với các tác giả khác đã được sự nhất trí của đồng tác giả khi đưa vào luận án. Các số liệu, kết quả được trình bày trong luận án là hoàn toàn trung thực và chưa từng được công bố trong bất kỳ một công trình nào khác ngoài các công trình công bố của tác giả. Luận án được hoàn thành trong thời gian tôi làm nghiên cứu sinh tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

Hà Nội, ngày 12 tháng 03 năm 2026

Nghiên cứu sinh

Nguyễn Thị Thủy Trâm

LỜI CẢM ƠN

Để hoàn thành luận án này, tôi đã nhận được rất nhiều sự quan tâm, hỗ trợ và tạo điều kiện từ các thầy cô, cơ quan, đồng nghiệp và gia đình.

Trước hết, tôi xin bày tỏ lòng biết ơn sâu sắc đến **PGS.TS. Lê Thành Long** – người thầy đã tận tình hướng dẫn, định hướng khoa học, đồng hành và khích lệ tôi trong suốt quá trình học tập và nghiên cứu. Tôi cũng xin trân trọng cảm ơn **PGS.TS. Thái Kế Quân** đã luôn quan tâm, hỗ trợ và giúp tôi tháo gỡ những khó khăn chuyên môn trong quá trình thực hiện đề tài.

Tôi xin chân thành cảm ơn **Ban Lãnh đạo Viện Khoa học Sự sống; Phòng Công nghệ sinh học động vật; Ban Lãnh đạo Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam**, **Phòng Đào tạo và các phòng chức năng của Học viện** đã tạo điều kiện thuận lợi để tôi tiến hành nghiên cứu và triển khai các thí nghiệm. Tôi cũng xin gửi lời cảm ơn đến các thầy cô, đồng nghiệp và học viên tại Viện Khoa học Sự sống, đặc biệt là cô Thảo, chị Thúy, các bạn Dung, Huy, Cường đã hỗ trợ và đồng hành trong quá trình tôi thực hiện đề tài.

Bên cạnh đó, tôi xin gửi lời tri ân sâu sắc đến gia đình – nguồn động viên tinh thần to lớn để tôi có thể hoàn thành công việc nghiên cứu. Tôi biết ơn ba mẹ, chồng và các con đã luôn thấu hiểu, chia sẻ và là chỗ dựa vững chắc trong suốt quá trình học tập.

Cuối cùng, tôi xin trân trọng cảm ơn tất cả những cá nhân và tập thể mà tôi không thể nêu hết tên trong lời cảm ơn này nhưng đã có những đóng góp quý báu, góp phần giúp tôi hoàn thành luận án.

Xin chân thành cảm ơn!

Hà Nội, ngày 12 tháng 3 năm 2026

Nghiên cứu sinh

Nguyễn Thị Thủy Trâm

MỤC LỤC

LỜI CAM ĐOAN ..... **I**
LỜI CẢM ƠN ..... **II**
DANH MỤC TỪ VIẾT TẮT ..... **V**
DANH MỤC BẢNG ..... **VI**
DANH MỤC HÌNH ..... **VIII**
MỞ ĐẦU **1**
CHƯƠNG 1. TỔNG QUAN VẤN ĐỀ NGHIÊN CỨU **5**
1.1. BỐI CẢNH VÀ SỰ CẦN THIẾT CỦA NGHIÊN CỨU BẢO TỒN LINH TRƯỞNG 5
1.1.1. Tầm quan của Linh trưởng và tình trạng đe dọa toàn cầu 5
1.1.2. Vị thế của Việt Nam trong khủng hoảng bảo tồn Linh trưởng 8
1.2. VAI TRÒ CỦA ADN TY THỂ TRONG NGHIÊN CỨU PHÁT SINH LOÀI 9
1.2.1. Đặc điểm cấu trúc và di truyền của hệ gene ty thể 9
1.2.2. Ứng dụng của ADN ty thể trong nghiên cứu Linh trưởng 10
1.3. TỔNG QUAN NGHIÊN CỨU ADN TY THỂ TRÊN LINH TRƯỞNG TRONG VÀ NGOÀI NƯỚC 11
1.3.1. Tổng quan nghiên cứu ADN ty thể linh trưởng ở Việt Nam 11
1.3.2. Tổng quan nghiên cứu ADN ty thể linh trưởng ở trên thế giới 15
1.4. MỤC TIÊU NGHIÊN CỨU 23
CHƯƠNG 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU **25**
2.1. ĐỐI TƯỢNG NGHIÊN CỨU 25
2.2. PHƯƠNG PHÁP NGHIÊN CỨU 25
2.2.1. Tách chiết ADN và giải trình tự 25
2.2.2. Phương pháp lắp ráp hệ gene ty thể 27
2.2.3. Phân tích tin sinh học 28
2.2.4. Cây phát sinh loài 30
CHƯƠNG 3. KẾT QUẢ VÀ THẢO LUẬN **31**
3.1. KẾT QUẢ GIẢI TRÌNH TỰ GENE 31
3.1.1. Kết quả chuẩn bị thư viện trình tự toàn bộ hệ gene Linh trưởng 31
3.1.2. Kết quả giải trình tự toàn bộ hệ gene Linh trưởng 32
3.2. ĐẶC ĐIỂM HỆ GENE TY THỂ HOÀN CHỈNH CỦA MỘT SỐ CÁ THỂ LINH TRƯỞNG 35
3.2.1 Đặc điểm hệ gene ty thể hoàn chỉnh của cá thể *Chà vá chân xám* (*Pygathrix cinerea*) Tây Nguyên Việt Nam 35
3.2.1.1. Thành phần nucleotide 37
3.2.1.2. Các gene mã hóa Protein 40
3.2.1.3. Vùng giàu A + T 41
3.2.1.4. *tRNA* và *rRNA* 42
3.2.1.5. Vùng lặp và vùng liên gene 43
3.2.1.6. Cây phân loài của *Chà vá chân xám* Tây Nguyên Việt Nam 46
3.2.1.7. Thảo luận 48

3.2.2. Đặc điểm hệ gene ty thể hoàn chỉnh của cá thể *Khỉ mặt đỏ* (*Macaca arctoides*) Tây Nguyên Việt Nam 50
3.2.2.1. Thành phần nucleotide 52
3.2.2.2. Các gene mã hóa Protein 56
3.2.2.3. Vùng giàu A + T 58
3.2.2.4. *tRNA* và *rRNA* 59
3.2.2.5. Vùng lặp và vùng liên gene 60
3.2.2.6. Thảo luận 62
3.2.3. Đặc điểm hệ gene ty thể hoàn chỉnh của cá thể *Chà vá chân nâu* (*Pygathrix nemaeus*) Tây Nguyên Việt Nam 63
3.2.3.2. Các gene mã hóa Protein 68
3.2.3.3. Vùng giàu A + T 70
3.2.3.4. *tRNA* và rRNA 70
3.2.3.5. Vùng lặp và vùng liên gene 73
3.2.3.6. Thảo luận 73
3.2.4. Đặc điểm hệ gene ty thể hoàn chỉnh của cá thể *Chà vá chân đe*n (*Pygathrix nigripes*) Tây Nguyên Việt Nam 75
3.2.4.1. Thành phần nucleotide 77
3.2.4.2. Các gene mã hóa Protein 78
3.2.4.3. Vùng giàu A + T 81
3.2.4.4. Thành phần acid amin 81
3.2.4.5. Vùng lặp và vùng liên gene 84
3.2.4.6. Thảo luận 84
3.2.5. Đặc điểm hệ gene ty thể hoàn chỉnh của cá thể *Khỉ vàng* (*Macaca mulatta*) Tây Nguyên Việt Nam 86
3.2.5.1. Thành phần nucleotide 89
3.2.5.2. Các gene mã hóa Protein 92
3.2.5.4. *tRNA* và *rRNA* 94
3.2.5.5. Vùng lặp và vùng liên gene 96
3.2.5.6. Cây phân loài của *Khỉ vàng* với các loài linh trưởng khác 97
3.2.5.7. Thảo luận 98
3.3. PHÂN TÍCH SAI KHÁC TRÌNH TỰ TY THỂ PHỤC VỤ NGHIÊN CỨU PHÁT SINH LOÀI 99
3.3.1. Phân tích sai khác trình tự gene *16S rRNA* và mối quan hệ phát sinh loài của các cá thể linh trưởng khu vực Tây Nguyên Việt Nam 100
3.3.2. Phân tích sai khác trình tự gene *cytb* và mối quan hệ phát sinh loài của các cá thể linh trưởng khu vực Tây Nguyên Việt Nam 107
3.3.3. Phân tích sai khác trình tự gene *D-loop* và mối quan hệ phát sinh loài của các cá thể linh trưởng khu vực Tây Nguyên Việt Nam 113
3.3.4. Mối quan hệ phát sinh loài của một số cá thể linh trưởng khu vực Tây Nguyên Việt Nam 116
3.3.5. Thảo luận 118
KẾT LUẬN VÀ KIẾN NGHỊ 120
KẾT LUẬN 120
KIẾN NGHỊ 120

DANH MỤC TỪ VIẾT TẮT

Allopatric speciation: sự phân hóa địa lý

AT-skew: mức độ mất cân bằng giữa adenine (A) và thymine (T) trong một trình tự DNA

Cytochrome b: Cytb

D-loop: Displacement loop, vùng kiểm soát (Control Region, CR),

GT-skew: mức độ mất cân bằng giữa guanine (G) và thymine (T) trong một trình tự DNA

PCG: Protein Coding Gene, vùng gene mã hóa protein

P. cinerea : *Pygathrix cinerea*

Phylogenetics: phát sinh loài

Phylogenetic Tree: cây phát sinh loài

P. nemaeus : *Pygathrix nemaeus*

P. nigripes : *Pygathrix nigripes*

M. arctoides : *Macaca arctoides*

Mitogenome: hệ gene ty thể

Marker: chỉ thị

tRNA: transfer RNA – RNA vận chuyển

12S rRNA: RNA ribosome 12S

16S rRNA: RNA ribosome 16S

DANH MỤC BẢNG

Bảng 3. 1. Nồng độ các mẫu ADN của Linh trưởng sau tách chiết. 31
Bảng 3. 2. Kết quả chuẩn bị thư viện trình tự toàn bộ hệ gene Linh trưởng (Whole geneome sequencing). 32
Bảng 3. 3. Vị trí và độ dài của 37 gene trong hệ gene ty thể hoàn chỉnh của cá thể *Chà vá chân xám* Tây Nguyên Việt Nam 36
Bảng 3. 4. Các chỉ số thành phần nucleotide trong các vùng khác nhau của 22 gene ty thể đại diện cho loài linh trưởng.....39
Bảng 3. 5. Giá trị GC-skew và AT-skew trong các gene mã hóa protein (PCGs). 41
Bảng 3. 6. Các vị trí bắt cặp không khớp của *t-RNAs*. AA—amino acid acceptor, T-arm, AC—anticodon.....43
Bảng 3.7. Vị trí và độ dài của 37 gene trong hệ gene ty thể hoàn chỉnh của cá thể *Khỉ mặt đỏ* Tây Nguyên Việt Nam 50
Bảng 3. 8. Các chỉ số thành phần nucleotide trong các vùng khác nhau của 23 gene ty thể đại diện cho loài linh trưởng.....55
Bảng 3. 9. Giá trị GC-skew và AT-skew trong các gene mã hóa protein (PCGs). 57
Bảng 3. 10. Các cặp base không bắt cặp của *t-RNAs*. AA—amino acid acceptor, T-arm, AC—anticodon.....60
Bảng 3. 11. Vị trí và độ dài của 37 gene trong hệ gene ty thể hoàn chỉnh của cá thể *Chà vá chân nâu* Tây Nguyên Việt Nam 63
Bảng 3. 12. Các chỉ số thành phần nucleotide trong các vùng khác nhau của 23 gene ty thể đại diện cho loài linh trưởng.....66
Bảng 3. 13. Giá trị GC-skew và AT-skew trong các gene mã hóa protein (PCGs). 69
Bảng 3. 14. Các cặp base không khớp của *t-RNAs*. AA—amino acid acceptor, T-arm, AC—anticodon. 71
Bảng 3. 15. Cấu trúc và đặc trưng của hệ gene ty thể hoàn chỉnh của *Chà vá chân đen* 76
Bảng 3. 16. Thành phần nucleotide của 13 gene mã hóa protein trong hệ gene ty thể của *Chà vá chân đen*.....78
Bảng 3. 17. AT-skew, và GC-skew của *Chà vá chân đen* Tây Nguyên Việt Nam.....80
Bảng 3. 18. Tỷ lệ phần trăm của thành phần axit amin trong toàn bộ hệ gene ty thể của các cá thể *chà vá: Chà vá chân đen* Tây Nguyên Việt Nam*, Chà vá chân đen* Trung Quốc*, và Chà vá chân đỏ* Trung Quốc. 83
Bảng 3. 19. Hệ gene ty thể hoàn chỉnh của *Khỉ vàng* chứa 37 gene, chi tiết về vị trí và chiều dài của chúng. 86
Bảng 3. 20. Tỷ lệ % AT trên các đoạn gene của các loài linh trưởng.....88
Bảng 3. 21. Thành phần các nucleotide giữa các loài linh trưởng. 89
Bảng 3. 22. Các cặp base không khớp được tìm thấy trong các gene *tRNA* củak*Khỉ vàng* Việt Nam.....94
Bảng 3. 23. Các vị trí khác biệt trên trình tự gene *16S rRNA* của cá thể *Chà vá chân đen* Tây Nguyên Việt Nam và cá thể *Chà vá chân đen* khác.....100
Bảng 3. 24. Các vị trí khác biệt trên trình tự gene *16S rRNA* của cá thể *Chà vá chân nâu* Tây Nguyên Việt Nam và cá thể *Chà vá chân nâu* khác 101
Bảng 3. 25. Các vị trí khác biệt trên trình tự gene *16S* của cá thể *Chà vá chân xám* Tây Nguyên Việt Nam và cá thể *Chà vá chân xám* khác.....102
Bảng 3. 26. Các vị trí khác biệt trên trình tự gene 16S của cá thể *Khỉ mặt đỏ* Tây Nguyên Việt Nam và cá thể *Khỉ mặt đỏ* khác 103

Bảng 3. 27. Các vị trí khác biệt trên trình tự gene *16S* của cá thể *Khỉ vàng* Tây Nguyên Việt Nam và cá thể *Khỉ vàng* khác 105
Bảng 3. 28. Các vị trí khác biệt trên trình tự gene *cytb* của cá thể *Chà vá chân đen* Tây Nguyên Việt Nam và cá thể *Chà vá chân đen* khác 107
Bảng 3. 29. Các vị trí khác biệt trên trình tự gene *cytb* của cá thể *Chà vá chân nâu* Tây Nguyên Việt Nam và cá thể *Chà vá chân nâu* khác 108
Bảng 3. 30. Các vị trí khác biệt trên trình tự gene *cytb* của cá thể *Chà vá chân xám* Tây Nguyên Việt Nam và cá thể *Chà vá chân xám* khác 109
Bảng 3. 31. Các vị trí khác biệt trên trình tự gene *cytb* của cá thể *Khỉ mặt đỏ* Tây Nguyên Việt Nam và cá thể *Khỉ mặt đỏ* khác 109
Bảng 3. 32. Các vị trí khác biệt trên trình tự gene *cytb* của cá thể *Khỉ vàng* Tây Nguyên Việt Nam và cá thể *Khỉ vàng* khác 110
Bảng 3. 33. Các vị trí khác biệt trên trình tự gene *D-loop* của cá thể *Chà vá chân xám* Tây Nguyên Việt Nam và cá thể *Chà vá chân xám* khác 113

DANH MỤC HÌNH

Hình 1. 1. Sự phân bố của các loài linh trưởng trên thế giới. Nguồn tham khảo từ *IUCN red list* và Estrada [1]. 5
Hình 1. 2. Sự phân bố của các loài linh trưởng trên thế giới với tỉ lệ % các loài đang giảm và bị đe doạ. 8
Hình 1. 3. Sự khác nhau trên trình tự *cytochrome b* giữa các loài linh trưởng ở các khu vực Châu Á (có cả Việt Nam) và cây phát sinh loài dựa vào sự khác biệt trình tự này. (Zhang và cs, 1998) 18
Hình 1. 4. ADN ty thể của *chà vá* 23
Hình 2. Loài thú linh trưởng được lưu trữ trong bảo tàng……18
Hình 3. 1. Kết quả chạy PCR các mẫu Linh trưởng; S001: *Chà vá chân xám*; S002: *Khỉ mặt đỏ*; S004: *Chà vá chân nâu*; S006: *Chà vá chân đen.* 32
Hình 3. 2. Cấu trúc gene ty thể hoàn chỉnh của *Chà vá chân xám* Tây Nguyên Việt Nam. 33
Hình 3. 3. Cấu trúc gene ty thể hoàn chỉnh của *Chà vá chân đen* Tây Nguyên Việt Nam. 33
Hình 3. 4. Cấu trúc gene ty thể hoàn chỉnh của *Khỉ vàng* Tây Nguyên Việt Nam. 34
Hình 3. 5. Cấu trúc gene ty thể hoàn chỉnh của *Chà vá chân nâu* Tây Nguyên Việt Nam. 34
Hình 3. 6. Cấu trúc gene ty thể hoàn chỉnh của *Khỉ mặt đỏ* Tây Nguyên Việt Nam. ...35
Hình 3. 7. Mức sử dụng codon đồng nghĩa tương đối (Relative Synonymous Codon Usage - RSCU) của các gene mã hóa protein trong ty thể *Chà vá chân xám*. 38
Hình 3. 8. Tần suất phần trăm thành phần axit amin trong toàn bộ hệ gene ty thể của các cá thể *chà vá*: *Chà vá chân xám* PP_623106, *Chà vá chân xám* JQ_821842, và *Chà vá chân đen* MH_064177. 38
Hình 3. 9. AT-skew và GC-skew trong 13 gene mã hóa protein gene ty thể *Chà vá chân xám* PP_623106. 40
Hình 3. 10. Cấu trúc thứ cấp của 22 gene *tRNA* của *Chà vá chân xám* Tây Nguyên Việt Nam. Màu đỏ đại diện cho các kết nối GC, trong khi màu xanh dương đại diện cho các kết nối AU. 45
Hình 3. 11. Cây phát sinh loài Maximum Likelihood của 55 hệ gene ty thể linh trưởng. PP633106 *Chà vá chân xám* màu xanh là *Chà vá chân xám* Tây Nguyên Việt Nam...47
Hình 3. 12. Mức sử dụng codon đồng nghĩa tương đối (Relative Synonymous Codon Usage - RSCU) của các gene mã hóa protein trong ty thể *Khỉ mặt đỏ*. 53
Hình 3. 13. Tần suất phần trăm thành phần axit amin trong hệ gene ty thể của các cá thể *khỉ*: *Khỉ mặt đỏ* Tây Nguyên Việt Nam, *Khỉ mặt đỏ* Trung Quốc (KM_360179) và *khỉ đuôi lợn* Trung Quốc (KP_330231).....54
Hình 3. 14. AT-skew và GC-skew trong 13 gene mã hóa protein gene ty thể *Khỉ mặt đỏ*. 57
Hình 3. 16. Cấu trúc thứ cấp của 22 gene *tRNA* của *Khỉ mặt đỏ* Tây Nguyên Việt Nam. Màu đỏ đại diện cho các kết nối GC, trong khi màu xanh dương đại diện cho các kết nối AU. 61
Hình 3. 16. Mức sử dụng codon đồng nghĩa tương đối (Relative Synonymous Codon Usage - RSCU) của các gene mã hóa protein trong ty thể *Chà vá chân nâu*. 66
Hình 3. 17. Tần suất phần trăm thành phần axit amin trong toàn bộ hệ gene ty thể của các cá thể *chà vá*: *Chà vá chân nâu* Tây Nguyên Việt Nam, *Chà vá chân nâu* Trung Quốc, và *Chà vá chân đen* Trung Quốc 66

Hình 3. 18. AT-skew và GC-skew trong 13 gene mã hóa protein gene ty thể *Chà vá chân nâu*. 69
Hình 3. 19. Cấu trúc thứ cấp của 22 gene *tRNA* của *Chà vá chân nâu* Tây Nguyên Việt Nam. Màu đỏ đại diện cho các kết nối GC, trong khi màu xanh dương đại diện cho các kết nối AU. 72
Hình 3. 20. AT-skew và GC- skew trên 13 gene mã hóa protein của hệ gene ty thể *Chà vá chân đen* Tây Nguyên Việt Nam. 79
Hình 3. 21. Thành phần axit amin và tỷ lệ tương đối (%) của chúng trong hệ gene ty thể hoàn chỉnh của *Chà vá chân đen* Tây Nguyên Việt Nam; *Chà vá chân đen* Trung Quốc và *Chà vá chân đỏ* Trung Quốc. 82
Hình 3. 22. Sử Dụng Codon Đồng Nghĩa Tương Đối (RSCU) của các gene mã hóa protein ty thể trong *Khỉ vàng* Việt Nam. 91
Hình 3. 23. Tỷ lệ phần trăm tần suất thành phần axit amin trong toàn bộ hệ gene ty thể của các cá thể khỉ của chi *macaca* : *Khỉ vàng* Việt Nam; *Khỉ vàng* Myanmar; *Khỉ vàng* Trung Quốc. 92
Hình 3. 24. Giá trị AT- và GC-skew của 13 gene mã hóa protein trong hệ gene ty thể của *Khỉ vàng* Tây Nguyên Việt Nam. 93
Hình 3. 25. Cấu trúc thứ cấp của 22 gene *tRNA* của *Khỉ vàng* Việt Nam. Màu đỏ là kết nối GC, màu xanh là kết nối AU. 95
Hình 3. 26. Cây phát sinh loài của *Khỉ vàng* với các loài linh trưởng khác.....97
Hình 3. 27. Cây phát sinh loài Linh trưởng Tây Nguyên từ trình tự *16S rRNA.* 106
Hình 3. 28. Cây phát sinh loài Linh trưởng Tây Nguyên từ trình tự gene *cytochrome b* 112
Hình 3. 29. Cây phát sinh loài Linh trưởng Tây Nguyên từ trình tự gene *D-loop.*.....115
Hình 3. 30. Cây phát sinh loài Linh trưởng Tây Nguyên từ trình tự của toàn hệ gene ty thể. 117

MỞ ĐẦU

Linh trưởng (Primates) là nhóm động vật có vú có mối quan hệ tiến hóa gần gũi với con người và đóng vai trò đặc biệt quan trọng trong các nghiên cứu về sinh học tiến hóa, sinh thái học, y sinh học và bảo tồn đa dạng sinh học. Bên cạnh giá trị khoa học, linh trưởng còn giữ vai trò then chốt trong hệ sinh thái rừng nhiệt đới, đặc biệt với chức năng phát tán hạt giống và duy trì cấu trúc quần xã thực vật. Do đó, sự suy giảm quần thể linh trưởng không chỉ ảnh hưởng đến bản thân các loài mà còn gây ra những hệ lụy nghiêm trọng đối với tính ổn định và khả năng phục hồi của hệ sinh thái.

Hiện nay, trên thế giới đã có hàng trăm hệ gene ty thể hoàn chỉnh của các loài linh trưởng được công bố và lưu trữ trên các cơ sở dữ liệu quốc tế như GenBank, đại diện cho phần lớn các họ và chi linh trưởng. Các dự án hệ gene quy mô lớn đã công bố dữ liệu hệ gene của hơn 200 loài linh trưởng nhằm nghiên cứu tiến hóa và đa dạng di truyền ở cấp độ toàn cầu (1).

Tuy nhiên, đối với linh trưởng Việt Nam, dữ liệu hệ gene ty thể hoàn chỉnh vẫn còn rất hạn chế. Việt Nam được ghi nhận là một trong những quốc gia có mức độ đa dạng loài linh trưởng cao trong khu vực Đông Nam Á, với khoảng 24–27 loài và phân loài thuộc ba họ chính, trong đó nhiều loài là đặc hữu và đang ở tình trạng nguy cấp hoặc cực kỳ nguy cấp theo phân loại của IUCN (1–3). Khu vực Tây Nguyên là một trong những vùng phân bố quan trọng của nhiều loài linh trưởng quý hiếm như *Chà vá chân xám* (*Pygathrix cinerea*), *Chà vá chân nâu* (*Pygathrix nemaeus), Chà vá chân đen* (*Pygathrix* nigripes), *Khỉ mặt đỏ* (*Macaca arctoides*) và *Khỉ vàng* (*Macaca mulatta*). Mặc dù vậy, số lượng hệ gene ty thể hoàn chỉnh được công bố cho các loài linh trưởng có nguồn gốc từ Việt Nam vẫn còn rất ít; nhiều loài linh trưởng tại Việt Nam chỉ được nghiên cứu thông qua một số marker ty thể riêng lẻ như *cytochrome b*, *16S rRNA* hoặc *D-loop*. Do sự thiếu hụt dữ liệu này, việc so sánh di truyền giữa các quần thể linh trưởng Việt Nam với các quần thể cùng loài ở các khu vực khác trong khu vực Đông Nam Á và trên thế giới còn gặp nhiều hạn chế. Điều này cũng làm giảm độ tin cậy của các phân tích phát sinh loài, đặc biệt trong việc xác định mối quan hệ tiến hóa giữa các loài hoặc đánh giá sự phân hóa di truyền giữa các quần thể địa lý.

Trong nhiều thập kỷ qua, các quần thể linh trưởng tại Tây Nguyên đang chịu sức ép ngày càng gia tăng do mất sinh cảnh, phân mảnh môi trường sống, săn bắt và buôn bán trái phép, cũng như tác động của biến đổi khí hậu. Trong bối cảnh đó, việc nghiên

cứu đặc điểm di truyền và mối quan hệ phát sinh loài của các loài linh trưởng có ý nghĩa đặc biệt quan trọng, không chỉ giúp làm rõ nguồn gốc tiến hóa và vị trí phân loại của các loài mà còn cung cấp cơ sở khoa học cho công tác bảo tồn. Trên thế giới, ADN ty thể đã được sử dụng rộng rãi như một công cụ hiệu quả trong nghiên cứu phát sinh loài nhờ các đặc điểm như di truyền theo dòng mẹ, không tái tổ hợp, tốc độ tiến hóa tương đối cao và kích thước hệ gene nhỏ. Các marker ty thể như *cytochrome b*, *16S rRNA* và vùng kiểm soát (*D-loop*), cũng như dữ liệu toàn bộ hệ gene ty thể (mitogenome), đã chứng minh khả năng phân giải tốt các mối quan hệ tiến hóa giữa các loài linh trưởng.

Mặc dù vậy, tại Việt Nam, các nghiên cứu di truyền linh trưởng vẫn còn hạn chế và phần lớn tập trung vào các đoạn gene đơn lẻ, với số lượng dữ liệu ADN ty thể hoàn chỉnh còn rất ít. Đặc biệt, đối với khu vực Tây Nguyên, cho đến nay vẫn thiếu các dữ liệu hệ thống về hệ gene ty thể hoàn chỉnh của các loài linh trưởng có nguồn gốc địa lý xác định. Khoảng trống dữ liệu này gây khó khăn cho việc so sánh tiến hóa, đánh giá quan hệ phát sinh loài cũng như đối chiếu các quần thể linh trưởng Việt Nam với các quần thể cùng loài đã được công bố trên thế giới.

Xuất phát từ thực tiễn đó, nghiên cứu này được thực hiện nhằm giải trình tự và phân tích hệ gene ty thể hoàn chỉnh của một số loài linh trưởng phân bố tại khu vực Tây Nguyên Việt Nam. Nghiên cứu tập trung vào việc mô tả đặc điểm cấu trúc, thành phần và tổ chức của hệ gene ty thể, so sánh sai khác trình tự với các dữ liệu đã công bố, đồng thời tái dựng mối quan hệ phát sinh loài của các loài nghiên cứu trong bối cảnh so sánh liên loài. Cách tiếp cận này không nhằm đánh giá đa dạng di truyền ở mức quần thể mà tập trung vào phân tích di truyền ở mức cá thể đại diện, phục vụ cho mục tiêu phát sinh loài và so sánh tiến hóa.

Việc bổ sung dữ liệu hệ gene ty thể hoàn chỉnh của các loài linh trưởng phân bố tại khu vực Tây Nguyên có ý nghĩa quan trọng. Dữ liệu này góp phần nâng cao độ phân giải của các phân tích phát sinh loài và bổ sung nguồn thông tin di truyền còn hạn chế của linh trưởng Việt Nam.

Kết quả của luận án cung cấp thêm dữ liệu ADN ty thể cho các loài linh trưởng Tây Nguyên, qua đó góp phần hoàn thiện cơ sở dữ liệu di truyền của linh trưởng Việt Nam. Đồng thời, các kết quả này tạo cơ sở khoa học cho các nghiên cứu về tiến hóa và hỗ trợ công tác bảo tồn các loài linh trưởng đang bị đe dọa.

Mục tiêu nghiên cứu

Mục tiêu tổng quát

Nghiên cứu này nhằm giải trình tự và phân tích đặc điểm hệ gene ty thể hoàn chỉnh của một số loài linh trưởng phân bố tại khu vực Tây Nguyên Việt Nam, qua đó so sánh sai khác trình tự di truyền và xác định mối quan hệ phát sinh loài của các loài nghiên cứu trong bối cảnh so sánh với các dữ liệu đã công bố trên thế giới.

Mục tiêu cụ thể

Phân tích đặc điểm cấu trúc, thành phần nucleotide và tổ chức gene của hệ gene ty thể hoàn chỉnh ở một số loài linh trưởng Tây Nguyên Việt Nam.

So sánh sai khác trình tự di truyền ty thể giữa các cá thể nghiên cứu và các trình tự tương đồng đã được công bố, phục vụ cho việc đánh giá quan hệ phát sinh loài.

Tái dựng cây phát sinh loài dựa trên các marker ty thể và toàn bộ hệ gene ty thể nhằm làm rõ vị trí phân loại và quan hệ tiến hóa của các loài linh trưởng nghiên cứu.

Câu hỏi nghiên cứu

Nghiên cứu này được thực hiện nhằm trả lời các câu hỏi khoa học sau:

Hệ gene ty thể hoàn chỉnh của các loài linh trưởng phân bố tại Tây Nguyên Việt Nam có đặc điểm cấu trúc và tổ chức gene như thế nào?

Các trình tự gene ty thể của các cá thể linh trưởng Tây Nguyên khác biệt ở mức độ nào so với các trình tự đã được công bố trên cơ sở dữ liệu quốc tế?

Các marker ty thể (*16S rRNA, cytochrome b, D-loop*) và dữ liệu toàn bộ hệ gene ty thể có cho kết quả nhất quán trong việc xác định mối quan hệ phát sinh loài hay không?

Các cá thể linh trưởng Tây Nguyên có biểu hiện các đặc điểm di truyền riêng biệt so với các quần thể cùng loài ở các khu vực khác hay không?

Nội dung nghiên cứu

Thu thập mẫu và tiến hành tách chiết ADN, xây dựng thư viện và giải trình tự thế hệ mới nhằm thu nhận dữ liệu hệ gene ty thể hoàn chỉnh của một số loài linh trưởng Tây Nguyên Việt Nam.

Lắp ráp, chú giải và phân tích đặc điểm cấu trúc hệ gene ty thể, bao gồm thành phần nucleotide, các gene mã hóa protein, gene *tRNA*, *rRNA*, vùng giàu A+T và các vùng lặp – liên gene.

Phân tích và so sánh sai khác trình tự di truyền ty thể giữa các cá thể nghiên cứu với các trình tự đã công bố trên cơ sở dữ liệu quốc tế.

Xây dựng và phân tích cây phát sinh loài dựa trên các vùng gene ty thể (*16S rRNA, cytochrome b*, *D-loop*) và toàn bộ hệ gene ty thể nhằm làm rõ mối quan hệ tiến hóa của các loài linh trưởng nghiên cứu.

Những đóng góp mới của luận án

Luận án đã cung cấp dữ liệu hệ gene ty thể hoàn chỉnh của một số loài linh trưởng phân bố tại khu vực Tây Nguyên Việt Nam, góp phần bổ sung cơ sở dữ liệu hệ gene ty thể của linh trưởng Việt Nam.

Làm rõ đặc điểm cấu trúc và tổ chức hệ gene ty thể của các loài linh trưởng nghiên cứu, đồng thời chỉ ra các sai khác trình tự di truyền có ý nghĩa trong so sánh phát sinh loài.

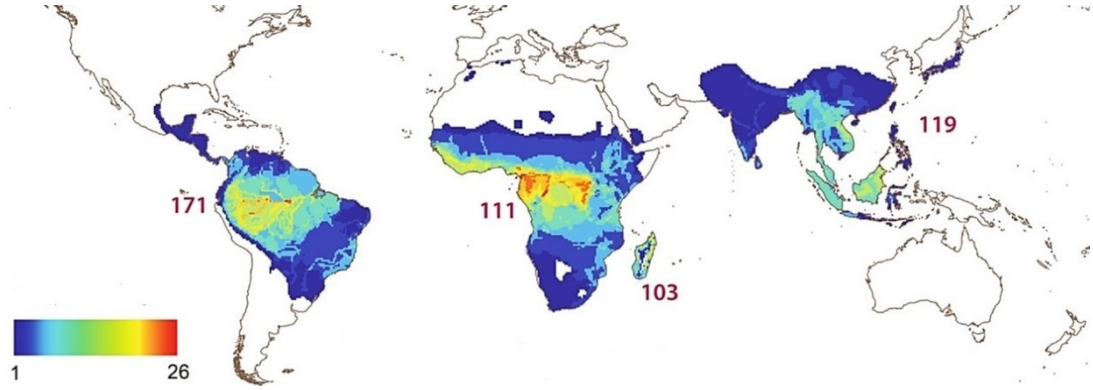
Xác định và củng cố mối quan hệ phát sinh loài của các loài linh trưởng Tây Nguyên trong bối cảnh so sánh với các loài cùng chi và cùng họ đã được công bố trên thế giới, cung cấp luận cứ khoa học phục vụ nghiên cứu tiến hóa và bảo tồn.

Chương 1. TỔNG QUAN VẤN ĐỀ NGHIÊN CỨU

1.1. Bối cảnh và sự cần thiết của nghiên cứu bảo tồn Linh trưởng

1.1.1. Tầm quan của Linh trưởng và tình trạng đe dọa toàn cầu

Bộ Linh trưởng (Primates) là một nhóm sinh vật đóng vai trò không thể thay thế trong hệ sinh thái toàn cầu, đóng vai trò trong việc duy trì đa dạng sinh học rừng nhiệt đới. Trên thế giới có khoảng 504 loài từ 79 chi và 16 họ (3) với trọng lượng từ 30g đến 200kg. Phần lớn các loài linh trưởng sống trong các khu rừng đất ẩm nhiệt đới, nhưng chúng cũng xuất hiện trong các khu vực rừng khô nhiệt đới, thảm thực vật ngập mặn trên mực nước triều cao, rừng ẩm trên núi, độ cao từ 1000m-4000m, rừng ôn đới lá rộng và lá rộng rụng lá, savan, đồng cỏ, đất ngập nước nội địa, vùng đá, và thậm chí cả sa mạc. Sự phân bố của linh trưởng trên thế giới tập trung ở 4 khu vực (Hình 1. 1) bao gồm: Châu Á (119 loài), Châu Phi (111 loài), Neotropics (Nam Mỹ) (171 loài), Madagascar (103 loài) và xuất hiện tự nhiên ở khoảng 90 quốc gia. 2/3 số loài linh trưởng tập trung ở 4 nước: Brazil, Madagascar, Indonesia và Công Gô.



Hình 1. 1. Sự phân bố của các loài linh trưởng trên thế giới. Nguồn tham khảo từ *IUCN red list* và Estrada (3).

Những biến động của Linh trưởng cũng sẽ trực tiếp hoặc gián tiếp ảnh hưởng đến cấu trúc, chức năng và khả năng phục hồi của hệ sinh thái. Sự tiến hóa, sự phân bố địa lý của chúng cũng liên quan chặc chẽ đến sự đa dạng hóa của thực vật hạt kín. Nhiều loài Linh trưởng đã được xác định hoặc nghi ngờ là loài thụ phấn quan trọng do chúng ăn các bộ phận của thực vật như hoa, mật hoa và một số loài ăn trái cây và hạt, cho phép chúng phân tán các hạt lớn và nhỏ trong khoảng cách xa, tăng cường khả năng tái sinh rừng. Vượn cáo là 1 loài Linh trưởng sinh sống ở Madagascar có mối quan hệ phức tạp với những cây sinh hạt lớn ở khu vực chúng sinh sống và sự tuyệt chủng của loài vượn cáo này là nguyên nhân trực tiếp hoặc gián tiếp đã làm cho một số loài cây Malagasy suy giảm khả năng tồn tại (4). Ở Thái Lan, hạt của loài cây

Lapsi (*Choerospondias axillaris*) được phân tán nhờ loài vượn ở trong khu vực này, tuy nhiên việc săn bắn vượn đã ảnh hưởng tiêu cực đến sự tăng sinh của loài cây này. Sự biến mất của quần thể vượn *Atelidae* và *Cebidae* ở rừng rậm Amazon do bị săn bắn cũng đã làm suy giảm nghiêm trọng nhiều loài cây gỗ cứng, loài cây có ý nghĩa đối với tiềm năng lưu trữ carbon của rừng. Trường hợp không có sự phát tán của hạt giống hợp tử bởi các cây nguyên sinh, các quần thể thực vật có thể bị giảm tính dị hợp tử di truyền và gia tăng sự phân hóa di truyền dưới quần thể, tăng sự phụ thuộc vào mật độ âm và giảm khả năng chọn lọc tự nhiên. Vì vậy, gián tiếp việc mất đi linh trưởng, động vật phát tán hạt giống, sẽ tác động nghiêm trọng đến con người. Với vai trò là những loài phát tán hạt giống quan trọng và là mắt xích trung tâm trong chuỗi thức ăn, sự suy giảm của quần thể linh trưởng có thể dẫn đến hiệu ứng domino, làm suy thoái cấu trúc và chức năng của hệ sinh thái mà chúng cư trú (3).

Về mặt tiến hóa, Linh trưởng là nhóm có mối quan hệ tiến hóa gần gũi với con người, cung cấp những hiểu biết vô giá về lịch sử phát triển, hành vi xã hội, và các quá trình sinh học phức tạp của loài người (5). Các quần thể linh trưởng hoang dã cũng có thể nắm giữ những manh mối có giá trị về nguồn gốc và sự tiến hóa của các mầm bệnh quan trọng và các quá trình lây truyền bệnh tật tự nhiên, do đó mang lại rất nhiều lợi ích cho con người (6) . Các mô hình động vật linh trưởng đã nâng cao hiểu biết của chúng ta rất nhiều về những căn bệnh trên con người. Các chứng bệnh như xơ vữa động mạch, bệnh hô hấp, HIV/AIDS, hay những rối loạn về sức khoẻ tâm thần như bệnh lý tâm thần, phản ứng điều trị với thuốc kích thích thần kinh, chức năng não, hay những căn bệnh có liên quan đến miễn dịch học, di truyền và gene, điều hòa nội tiết sinh sản, động lực của ký sinh trùng trên con người đã được nghiên cứu rất nhiều từ mô hình thí nghiệm linh trưởng (7). Khi các loài linh trưởng trở nên quý hiếm hoặc tuyệt chủng tại địa phương, chúng ta có nguy cơ mất đi các mối quan hệ sinh thái, xã hội và văn hóa phức tạp đã phát triển giữa con người và các loài linh trưởng trong nhiều thiên niên kỷ. Vì vậy, việc nghiên cứu và bảo tồn Linh trưởng không chỉ là trách nhiệm sinh thái mà còn là một nhiệm vụ nhân văn, nhằm bảo vệ nguồn gene và di sản tiến hóa chung của Trái Đất.

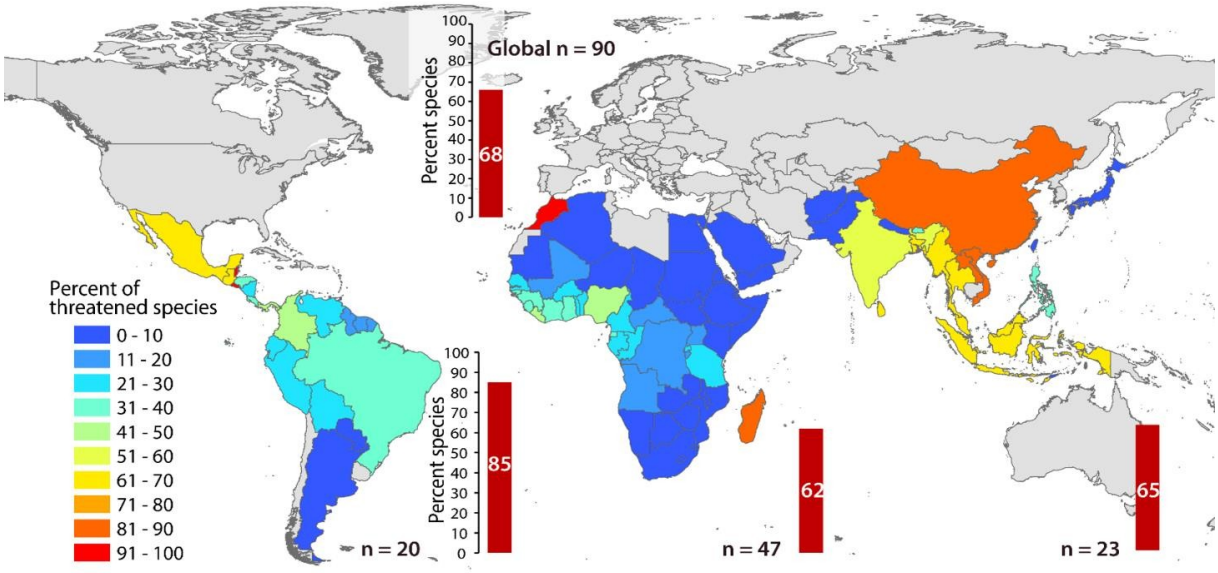
Trong hơn một thập kỉ qua, quần thể linh trưởng đang giảm dần trên thế giới bởi những tác động của con người như sự thu hẹp môi trường sống của linh trưởng do tốc độ đô thị hoá, khai thác rừng cho nông nghiệp, sự khai thác gỗ, nạn săn bắn, buôn bán trái phép, dịch bệnh. Trong 20 năm từ 1990-2010, diện tích đất nông nghiệp mở rộng ở các nước Châu Phi và Madagascar trong những khu vực có nhiều loài linh trưởng để phục vụ cho nông nghiệp ước chừng khoảng 1,5 triệu km^2 và gần 2 triệu

km^2 rừng mất độ che phủ. Hơn nữa, theo số liệu của Source of raw data FAOStat (*https://www.fao.org/faostat/en/#home*) , việc khai thác dầu cọ quá mức cũng đã làm ở Nam Mỹ và Châu Phi được dự báo sẽ dẫn đến những hậu quả tiêu cực nghiêm trọng đối với các quần thể linh trưởng ở những khu vực đó. Ở châu Á, việc mở rộng các đồn điền cao su ở khu vực Tây Nam Trung Quốc đã khiến loài vượn đen má trắng phương bắc (*Nomascus leucogenys*) và vượn Hải Nam (*Nomascus hainanus*) gần như tuyệt chủng (8).

Trái đất đang trải qua các giai đoạn biến đổi khí hậu, linh trưởng là loài động vật có nhiều đặc điểm rất dễ bị tổn thương bởi các điều kiện sinh thái thay đổi như có sự phân bố địa lý hạn chế, tuổi sinh sản lần đầu muộn và thời kỳ sinh đẻ kéo dài. Biến đổi khí hậu cũng gây ra sự mất môi trường sống cho động vật, buộc chúng phải di chuyển ra khỏi khu vực được bảo vệ khiến chúng dễ bị tổn thương hơn trước do mầm bệnh mới, đối thủ cạnh tranh mới, nguồn cung cấp thực phẩm thay đổi, hoặc chúng có thể trở thành thức ăn cho các động vật khác, thậm chí là các hoạt động săn bắn của con người. Ngoài ra, biến đổi khí hậu thường xuyên và nghiêm trọng hơn có thể gây ra lũ lụt, hạn hán, hỏa hoạn, bão và các sự kiện El Niño – Southern Oscillation có thể ảnh hưởng đến nguồn cung cấp thực phẩm sẵn có cho các quần thể linh trưởng, với các tác động tiêu cực đến sức khỏe, khả năng sinh sản và tỷ lệ tử vong.

Thế kỷ 21 chứng kiến sự gia tăng chưa từng có trong các mối đe dọa đối với sự sống còn của Linh trưởng. Theo Liên minh Bảo tồn Thiên nhiên Quốc tế (IUCN), Linh trưởng hiện là một trong những bộ động vật có xương sống bị đe dọa nghiêm trọng nhất. Năm 2017, theo số liệu của sách đỏ thế giới (the International Union for the Conservation of Nature Red List – IUCN red list) trên 60% loài linh trưởng, được đưa vào danh sách đang bị đe doạ tuyệt chủng (3) trong đó 87% số loài ở Madagascar đang bị đe dọa, 73% ở châu Á, 37% ở lục địa châu Phi và 36% ở Neotropics (Hình 1. 2).

Các nguyên nhân chính dẫn đến cuộc khủng hoảng này bao gồm: mất và suy thoái môi trường sống do phát triển nông nghiệp, khai thác gỗ, xây dựng cơ sở hạ tầng; săn bắt và buôn bán trái phép nhằm mục đích lấy thịt, vật nuôi cảnh, hoặc nguyên liệu cho y học cổ truyền; các mối đe dọa mới nổi như biến đổi khí hậu và nguy cơ lây truyền mầm bệnh (zoonotic diseases) giữa người và động vật hoang dã.



Hình 1. 2. Sự phân bố của các loài linh trưởng trên thế giới với tỉ lệ % các loài đang giảm và bị đe doạ.

Thanh màu biểu thị % số lượng linh trưởng đang bị đe dọa. Biểu đồ màu đỏ: quần thể linh trưởng đang bị đe dọa trên 4 khu vực trên thế giới. Nguồn tham khảo từ IUCN red list và Estrada và cs (3).

1.1.2. Vị thế của Việt Nam trong khủng hoảng bảo tồn Linh trưởng

Việt Nam được công nhận là một trong những quốc gia có tính đa dạng sinh học cao trên thế giới và là trung tâm đặc hữu của nhiều loài Linh trưởng là với nhiều loài cực kỳ nguy cấp (Critically Endangered - CR) cư trú như *Voọc cát bà*, *Voọc mũi hếch*, và một số loài Vượn. Nếu không kể đến 4 nước có tỉ lệ sinh sống của linh trưởng nhiều 1 cách khác biệt này thì Việt Nam được đánh giá là một trong những quốc gia có mức độ đa dạng sinh học về các loài linh trưởng tương đối cao. Với đặc điểm về khí hậu và địa hình của Việt Nam, nơi đây đã thu hút được 24-27 loài và phân loài linh trưởng thuộc 03 họ đó là: họ Cu li (*Loridae*), họ Khỉ (*Cercopithecidae*) và họ Vượn (*Hylobatidae*) chọn làm nơi cư trú trong tổng số hơn 500 loài và phân loài được Tổ chức Bảo tồn thiên nhiên quốc tế (IUCN) công nhận (9). Đặc biệt, trong đó có 6 loài ghi nhận là đặc hữu của Việt Nam: Khỉ đuôi dài côn đảo (*Macaca fascicularis condorensis*), vọoc Mông trắng (*Trachypithecus delacouri*), vọoc Cát Bà (*Trachypithecus poliocephalus*), vọoc Mũi hếch (*Voọc mũi hếch Bắc Bộ*), *Chà vá chân xám* (*Pygathrix cinerea*), Vượn đen Đông Bắc (*Nomascus nasutus).* Thú linh trưởng được xem là những loài chỉ thị cho sức khỏe hệ sinh thái, thước đo của mức độ đa dạng sinh học, đồng thời đóng vai trò quan trọng đối với đời sống văn hóa, tinh thần và khoa học phục vụ con người.

Tuy nhiên, do áp lực dân số, khai thác tài nguyên thiếu bền vững, và sự kiểm soát lỏng lẻo tại nhiều khu vực, hầu hết các loài Linh trưởng Việt Nam đều đang ở ngưỡng nguy cấp cao nhất với 90% linh trưởng đang đối mặt với nguy cơ đe dọa tuyệt chủng (10). Sự suy giảm nghiêm trọng về số lượng và sự phân mảnh của quần thể đã dẫn đến nguy cơ giao phối cận huyết và mất đa dạng di truyền, đe dọa nghiêm trọng đến khả năng tồn tại lâu dài của chúng.

Trước tình hình đó, công tác bảo tồn Linh trưởng tại Việt Nam không chỉ đòi hỏi các biện pháp bảo vệ tại chỗ (in-situ conservation) mà còn cần sự hỗ trợ đắc lực từ khoa học. Việc nghiên cứu toàn diện về mặt sinh học, sinh thái, và đặc biệt là di truyền học phân tử, trở thành yêu cầu tiên quyết để xác định các đơn vị bảo tồn ưu tiên (Conservation Units) và xây dựng chiến lược quản lý quần thể hiệu quả, chính xác, và bền vững (3).

1.2. Vai trò của ADN ty thể trong nghiên cứu phát sinh loài

1.2.1. Đặc điểm cấu trúc và di truyền của hệ gene ty thể

Ty thể là một bào quan có hệ gene riêng biệt và được cho là có nguồn gốc từ vi khuẩn hiếu khí cổ đại, kết hợp với tổ tiên của tế bào nhân thực thông qua quá trình nội cộng sinh. ADN ty thể (mtDNA) là một phân tử di truyền nằm trong ty thể – bào quan có chức năng hô hấp tế bào – và chỉ được truyền theo dòng mẹ giúp theo dõi phả hệ dòng mẹ và lịch sử tiến hóa của quần thể mà không bị ảnh hưởng bởi sự tái tổ hợp từ bố. Đã từ lâu, người ta cho rằng, hệ gene ty thể là nguồn dữ liệu phong phú và dễ dàng phân tích cho việc phân tích di truyền quần thể. Tỉ lệ đột biến cao trung bình ước chừng nhanh hơn gấp 5 đến 10 lần so với hệ gene trong nhân (11) giúp phân giải tốt các mối quan hệ phát sinh loài gần (ví dụ: giữa các phân loài hoặc quần thể). Hiện nay, ADN ty thể được ứng dụng rộng rãi trong nghiên cứu phát sinh loài và bảo tồn các loài linh trưởng, đặc biệt là các loài nguy cấp tại khu vực Đông Nam Á – trong đó có Việt Nam. ADN ty thể không tái tổ hợp, tốc độ đột biến cao, kích thước nhỏ (~16.500 cặp base ở động vật có xương sống) và có số lượng bản sao lớn trong mỗi tế bào đảm bảo sự ổn định của trình tự, đơn giản hóa việc xây dựng cây phát sinh loài (phylogenetic tree) do không có sự pha trộn giữa các dòng. Trừ một vài trường hợp ngoại lệ, tất cả các hệ gene ty thể của động vật đều chứa 37 gene: hai gene rRNA, 13 gene cho protein và 22 gene *tRNA*. Hệ gene của ty thể, cùng với RNA và protein từ tế bào chất, sẽ giúp ty thể có thể tự sao chép, phiên mã, xử lý mRNA và dịch mã protein. Vì vậy, cùng với sự di truyền từ dòng mẹ, với kích thước nhỏ và đầy đủ chức

năng của hệ gene ty thể, rất dễ giúp ADN ty thể trở thành công cụ lý tưởng để nghiên cứu phát sinh loài (phylogenetics) và lịch sử tiến hóa của sinh vật.

1.2.2. Ứng dụng của ADN ty thể trong nghiên cứu Linh trưởng

Phát sinh Chủng loại (Phylogeny): Do không chịu ảnh hưởng của quá trình tái tổ hợp (recombination) như ADN nhân, hệ gene ty thể cho phép các nhà nghiên cứu truy nguyên trực tiếp dòng dõi tổ tiên theo hệ mẫu (matrilineal lineages). Các vùng gen như *Cytochrome b (Cytb)* hay *COI* thường được sử dụng để phân giải mối quan hệ giữa các loài (interspecies), trong khi vùng kiểm soát (*D-loop*) hỗ trợ phân tích quan hệ trong cùng một loài (intraspecies). Một trong những ứng dụng nổi bật nhất là phát hiện "Eva ty thể" – tổ tiên chung theo dòng mẹ của toàn bộ nhân loại – được ước tính sống cách đây khoảng 150.000–200.000 năm ở châu Phi, dựa trên phân tích ADN ty thể của hàng trăm cá thể từ các châu lục khác nhau (12). Ngoài ra, ADN ty thể thường được sử dụng để xác định mối quan hệ giữa các loài, phân loài, quần thể, và trong giám định pháp y cũng như bảo tồn động vật hoang dã. Với sự phát triển của công nghệ giải trình tự gene thế hệ mới (NGS), ngày càng nhiều nghiên cứu sử dụng toàn bộ hệ gene ty thể thay vì chỉ một vài đoạn gene như *cytochrome b* hoặc COI, giúp nâng cao độ chính xác trong việc tìm ra mối quan hệ tiến hóa giữa các loài (13).

Ước tính thời gian phân tách (Molecular Clock): Tốc độ tích lũy đột biến của mtADN diễn ra nhanh và tương đối ổn định theo thời gian. Điều này cho phép các nhà khoa học sử dụng nó như một "đồng hồ phân tử" để ước tính thời điểm phân tách của các nhóm linh trưởng, chẳng hạn như thời điểm tách biệt giữa người và tinh tinh, hay giữa các phân loài *Voọc*.

Đánh giá khác biệt di truyền: Những đột biến xảy ra trên ADN ty thể không làm ảnh hưởng đến khả năng tồn tại của sinh vật, do đó, đột biến xảy ra trên ty thể không làm mất đi những kiểu gene của sinh vật đã được chọn lọc trong quá trình tiến hóa. Khi một loài nào đó có mối quan hệ tiến hóa càng xa thì số lượng sự khác biệt trở nên lớn hơn. Việc đánh giá mức độ khác biệt di truyền là cốt lõi để hiểu về "sức khỏe" di truyền của các quần thể linh trưởng, đặc biệt là các loài đang bị đe dọa. Trong những năm trước đây, sau các nghiên cứu đa dạng ADN ty thể ở người, trình tự ADN ty thể đã được sử dụng để kiểm tra sự khác biệt di truyền của loài vượn (14–16). Các chỉ số như độ đa dạng haplotype và độ đa dạng nucleotide phản ánh mức độ biến dị di truyền hiện tại. Quần thể có độ đa dạng thấp thường đối mặt với nguy cơ suy thoái cận huyết cao. Dữ liệu ADN ty thể giúp tái hiện các sự kiện trong quá khứ như "hiệu ứng nút thắt cổ chai" (bottleneck effect) – khi số lượng cá thể giảm sút

nghiêm trọng, hoặc sự bùng nổ dân số (population expansion*)*. Phân tích ADN ty thể giúp xác định sự di cư và sự phân bố của các dòng gen theo địa lý, từ đó hiểu được cách các rào cản tự nhiên (sông, núi) ảnh hưởng đến sự cách ly của các quần thể linh trưởng.

1.3. Tổng quan nghiên cứu ADN ty thể trên linh trưởng trong và ngoài nước

1.3.1. Tổng quan nghiên cứu ADN ty thể linh trưởng ở Việt Nam

Nghiên cứu sự phát sinh loài ở Việt Nam có lịch sử lâu đời, bắt đầu từ thế kỷ 19 khi các nhà khoa học Pháp tiến hành khảo sát hệ động thực vật phong phú tại đây. Giai đoạn này chủ yếu dựa vào phân loại hình thái học, với những đóng góp của các nhà khoa học như Henri Mouhot và Alexandre Yersin, đặt nền móng cho việc mô tả nhiều loài động thực vật mới. Trong những năm 60 của thế kỷ XX, thế giới đã bắt đầu quan tâm tới công tác nghiên cứu và bảo tồn đa dạng sinh vật. Sự phát sinh loài và phân loại của các loài động vật có vú ban đầu dựa trên các đặc điểm hình thái chung hoặc có nguồn gốc. Năm 1962; vườn quốc gia đầu tiên (Cúc Phương) được thành lập nhằm bảo tồn đa dạng sinh vật và loài linh trưởng quí hiếm và đặc hữu là loài *Voọc quần đùi trắng* (Trachypithecus delacouri). Bắt đầu từ thời gian này, những nghiên cứu về sinh học được triển khai sâu rộng hơn. Trung tâm Cứu hộ Linh trưởng Nguy cấp (EPRC) tại VQG Cúc Phương là một trong những cơ sở hàng đầu Đông Nam Á, hiện đang chăm sóc và phục hồi 170 cá thể thuộc 15 loài linh trưởng nguy cấp. Các dự án hợp tác với các tổ chức quốc tế như Tổ chức Bảo tồn Linh trưởng Quốc tế (IUCN Primate Specialist Group) và các viện nghiên cứu từ Đức, Nhật Bản, Mỹ đã thúc đẩy nghiên cứu di truyền, sinh thái và hành vi của các loài linh trưởng. Người Việt Nam đầu tiên có những công trình nghiên cứu về linh trưởng là Đào Văn Tiến. Năm 1957, ông công bố bài nghiên cứu "Một số ghi chép về động vật có vú ở miền Bắc Việt Nam", trong đó ghi nhận và phân loại các loài linh trưởng tại khu vực này, đánh dấu bước đi đầu tiên trong nghiên cứu linh trưởng tại Việt Nam (17–19). Các công trình của Đào Văn Tiến không chỉ giúp nâng cao hiểu biết về sự đa dạng và đặc điểm sinh thái của các loài linh trưởng mà còn tạo nền tảng khoa học cho các chương trình bảo tồn sau này, khiến ông được ghi nhận là người tiên phong trong nghiên cứu linh trưởng tại Việt Nam.

Giai đoạn năm 1975 đến nay các nghiên cứu và bảo tồn đa dạng sinh vật, được quan tâm và triển khai mạnh mẽ trên các vùng của cả nước và đạt được rất nhiều kết quả có giá trị. Đội ngũ cán bộ Việt Nam nghiên cứu về linh trưởng cũng phát triển và lớn mạnh không ngừng. Những kết quả nghiên cứu được thể hiện trong các công trình

của Đào Văn Tiến; Võ Quý nhận được Giải thưởng Môi trường Goldman (2003) và được công nhận là một nhà bảo tồn hàng đầu của Việt Nam; Nguyễn Cử đã đóng góp vào việc phát hiện và bảo tồn nhiều loài chim nguy cấp như gà lôi lam đuôi trắng (*Lophura hatinhensis*); Trần Triết được biết đến với các chương trình bảo tồn rùa và các dự án nâng cao nhận thức về bảo vệ động vật hoang dã; Lê Xuân Cảnh lãnh đạo nhiều dự án quốc gia và quốc tế nhằm bảo vệ các hệ sinh thái tự nhiên; Vũ Ngọc Long là người đi đầu trong các hoạt động bảo vệ rừng và nâng cao nhận thức cộng đồng về bảo tồn tài nguyên thiên nhiên; Hà Thăng Long đã vận động thành công việc mở rộng các khu bảo tồn và nâng cao nhận thức về bảo tồn *Voọc* trong cộng đồng; Nguyễn Thị Thu Hiền là người đi đầu trong việc bảo vệ các rạn san hô đang chịu ảnh hưởng của biến đổi khí hậu; Nguyễn Quảng Trường đóng góp lớn trong việc bảo tồn các loài bò sát quý hiếm như tắc kè và rắn hổ mang chúa. Các nhà bảo tồn trên đã và đang đóng góp to lớn trong việc bảo vệ đa dạng sinh học, từ rừng, núi, đến biển của Việt Nam. Họ không chỉ nghiên cứu và phát hiện các loài mới, mà còn tích cực bảo tồn các loài nguy cấp và nâng cao nhận thức cộng đồng về bảo vệ thiên nhiên.

Các nghiên cứu không chỉ bó hẹp do các cán bộ khoa học Việt Nam tiến hành, mà còn có sự hợp tác quốc tế sâu rộng với các chuyên gia linh trưởng và các tổ chức bảo tồn quốc tế. Kết quả của các điều tra, nghiên cứu về khu hệ linh trưởng của các địa phương, các vùng miền và các công trình nghiên cứu về sinh học, sinh thái của các loài linh trưởng ở Việt Nam đã được công bố trên các tạp chí khoa học trong nước và quốc tế. Nổi bật trong thời gian gần đây đã có rất nhiều công trình nghiên cứu chuyên sâu về linh trưởng Việt Nam của các tác giả trong và ngoài nước như: Nguyễn Cử (20), Nadler T (21). Năm 2020, một nghiên cứu tại dãy núi Kon Ka Kinh (Gia Lai) đã làm rõ tập tính kiếm ăn và sinh thái của *Chà vá chân xám* (*Pygathrix cinerea*), một trong những loài linh trưởng nguy cấp nhất thế giới (22) và đánh giá tác động của biến đổi khí hậu đến sự phân bố của cá thể *Chà vá chân xám Pygathrix cinerea* bằng mô hình hốc sinh thái (23). Trong khi đó, nghiên cứu tại Sơn Trà (Đà Nẵng) vào năm 2022 đã tập trung vào tập tính sinh sản và cấu trúc xã hội của *Chà vá chân nâu* (*Pygathrix nemaeus*), cho thấy loài này sở hữu cấu trúc bầy đàn phức tạp với hệ thống đa thê (24). Các nghiên cứu về sinh thái và tập tính của linh trưởng cũng đạt được nhiều tiến bộ lớn.

Tuy nhiên, các phân tích di truyền gần đây ngày càng đóng một vai trò quan trọng trong việc xác nhận hiện có hoặc thiết lập các nhóm thực vật và động vật có vú thường hoàn toàn khác nhau. Việc nhận diện các loài và xác định mối quan hệ giữa các loài cũng rất quan trọng trong sinh học, sinh thái học, tiến hóa, hệ thống học, quản

lý động vật hoang dã, bảo tồn và khoa học pháp y. Trên thế giới hiện đã có nhiều công trình nghiên cứu về sự đa dạng sinh học cũng như những nghiên cứu về quan hệ di truyền giữa các loài linh trưởng. Áp dụng các phương pháp sinh học phân tử vào đánh giá độ bảo tồn loài, xây dựng cây phát sinh loài và dự đoán được thời gian phân loài dựa trên trình tự bảo tồn trong trình tự nucleotide của ty thể. Năm 1982, Goodman M. đã tìm ra các bằng chứng amino acid về sự phát sinh ở bộ Linh trưởng và *Eutheria*. Năm 1984, Charles G.Sibley và cộng sự đã nghiên cứu sự phát sinh loài của Liên họ Người (*Hominoid)* bằng phương pháp lai ADN-ADN. Nghiên cứu về động lực tiến hóa phân tử gene *cytochrome b* ở linh trưởng mũi cong **(***stresirrhine primates)* của A.D.Yoder và cộng sự năm 1996. Năm 2001, R.Noda và cộng sự đã phân tích sự đa dạng trình tự *16S* rRNA ty thể ở Liên họ Người (*Hominoid).* Ana K.Z.Guillen và cộng sự năm 2004 cũng đã nghiên cứu sự đa dạng gene ở Vượn lớn châu Phi dựa trên trình tự ADN ty thể. Juan C.Opazo và cộng sự (2006) đã phân tích mối quan hệ di truyền và thời gian phân loài ở khỉ Tân thế giới (*platyrrhini)* trên nhiều trình tự bảo tồn khác nhau. Năm 2014, Luca Pozzi và cộng sự cũng đã thực hiện đánh giá quan hệ di truyền các loài linh trưởng và thời gian phân loài linh trưởng dựa trên toàn bộ trình tự gene ty thể (25).

Việt Nam là một trong những quốc gia có mức độ đa dạng loài linh trưởng cao nhất thế giới, với 24 đến 27 loài được ghi nhận, trong đó nhiều loài đặc hữu và nguy cấp như *Voọc cát bà* (*Trachypithecus poliocephalus*), *Voọc quần đùi trắng* (*Trachypithecus delacouri*), *Chà vá chân xám* (*Pygathrix cinerea*), và vượn đen má trắng phương Bắc (*Nomascus concolor*). Sự hiện diện của các loài linh trưởng này không chỉ góp phần tạo nên sự đặc sắc về sinh thái mà còn thu hút sự quan tâm của cộng đồng khoa học quốc tế trong việc nghiên cứu và bảo tồn.

Trong những năm gần đây, nhiều nghiên cứu đã tập trung vào việc điều tra và bảo tồn linh trưởng tại Việt Nam. Một trong những thành tựu đáng chú ý là nghiên cứu về đa dạng di truyền của các loài linh trưởng tại Khu Bảo tồn Thiên nhiên Xuân Liên vào năm 2020, do nhóm Lê Đức Minh thực hiện (26). Nghiên cứu này sử dụng phương pháp sinh học phân tử dựa trên trình tự gene *cytochrome b* và *16S* ty thể, giúp xác định sự đa dạng di truyền và cấu trúc quần thể của các loài linh trưởng. Kết quả nghiên cứu cung cấp dữ liệu quan trọng để bảo tồn các loài nguy cấp, đặc biệt là những loài có nguy cơ tuyệt chủng cao. Những kết quả này không chỉ làm sáng tỏ đặc điểm sinh thái của các loài mà còn hỗ trợ việc thiết lập các khu vực bảo vệ phù hợp.

Việt Nam là một trong những quốc gia có mức độ đa dạng sinh học cao nhất khu vực Đông Nam Á, đặc biệt là về thành phần các loài linh trưởng. Hiện nay, Việt

Nam có khoảng 24-27 loài linh trưởng, trong đó nhiều loài là đặc hữu và đang ở tình trạng nguy cấp theo phân loại của IUCN (1,27). Các phát hiện gần đây đã nhấn mạnh rằng tiềm năng sinh học ở Việt Nam chưa được khám phá hết. Trong đó, nổi bật là việc mô tả hai loài linh trưởng mới cho khoa học: cá thể *Chà vá chân xám* (*Pygathrix cinerea*) vào năm 2007 và loài vượn Trung bộ (*Nomascus annamensis*) vào năm 2010 (21). Những phát hiện này không chỉ có ý nghĩa về mặt khoa học mà còn góp phần thúc đẩy nỗ lực bảo tồn và khẳng định vai trò then chốt của Việt Nam trong việc bảo vệ đa dạng sinh học toàn cầu.

Trước áp lực từ quá trình phát triển kinh tế, đô thị hóa và khai thác tài nguyên, Việt Nam đã tăng cường đầu tư và triển khai các chính sách nhằm bảo vệ hệ sinh thái tự nhiên. Theo thống kê, cả nước hiện có khoảng 2,2 triệu ha rừng đặc dụng, với 30 vườn quốc gia và 114 khu bảo tồn loài và sinh cảnh được thành lập (28). Các khu vực này là nơi cư trú của phần lớn quần thể linh trưởng còn sót lại tại Việt Nam. Tuy nhiên, những nỗ lực này vẫn chưa đủ để ngăn chặn xu hướng suy giảm nghiêm trọng số lượng cá thể trong tự nhiên. Các mối đe dọa chính bao gồm mất sinh cảnh do chặt phá rừng, săn bắt và buôn bán trái phép động vật hoang dã, trong đó các loài linh trưởng thường là đối tượng bị nhắm đến.

Để bảo tồn và phục hồi các quần thể linh trưởng, cần thiết phải xây dựng một kế hoạch hành động quốc gia tổng thể. Kế hoạch này cần bao gồm: (1) Tăng cường khung pháp lý và thực thi các biện pháp xử lý nghiêm minh đối với các hành vi vi phạm (*Điều 244 Bộ luật hình sự năm 2015*); (2) Thành lập và vận hành hiệu quả các trung tâm cứu hộ động vật hoang dã, đặc biệt tập trung vào nhóm linh trưởng; (3) Quy hoạch và phát triển các khu bảo tồn ưu tiên cho sinh cảnh của các loài linh trưởng quý *hiếm (Luật Lâm nghiệp (2017) & Luật Đa dạng sinh học (2008))*; và (4) Tăng cường đầu tư cho nghiên cứu khoa học – đặc biệt là các nghiên cứu về phân bố, sinh thái, hành vi và di truyền học của các loài linh trưởng trong nước (29).

Trong bối cảnh đó, nghiên cứu về đa dạng di truyền có vai trò đặc biệt quan trọng. Thông qua các chỉ thị di truyền như gene ti thể (ví dụ: cytochrome b, *16S* rRNA, *D-loop*), các nhà khoa học có thể tái dựng quan hệ phát sinh loài, xác định cấu trúc quần thể, mức độ giao phối cận huyết và phát hiện các quần thể biệt lập có nguy cơ tuyệt chủng cao. Điều này giúp xác định đơn vị bảo tồn ưu tiên và thiết kế chiến lược bảo tồn phù hợp với thực tiễn từng vùng (30). Tại Việt Nam, các nghiên cứu về gene ty thể ở các loài như *Voọc mũi hếch* (*Voọc mũi hếch Bắc Bộ*), vượn (*Nomascus*

spp.) và khỉ mốc (*Macaca* spp.) đã chỉ ra mức độ đa dạng thấp ở nhiều quần thể, cho thấy sự cấp thiết trong công tác bảo tồn.

Bên cạnh ý nghĩa khoa học, nghiên cứu linh trưởng còn góp phần nâng cao nhận thức cộng đồng và thu hút sự hỗ trợ từ các tổ chức quốc tế. Các loài linh trưởng như *Chà vá chân nâu* hay *vượn đen má trắng* không chỉ là biểu tượng sinh thái mà còn có tiềm năng lớn trong phát triển du lịch sinh thái, đóng góp cho kinh tế địa phương nếu được quản lý bền vững (31).

Nghiên cứu về linh trưởng ở Việt Nam có vai trò không thể thay thế trong bảo tồn đa dạng sinh học. Đầu tư cho lĩnh vực này không chỉ là trách nhiệm bảo vệ di sản thiên nhiên của quốc gia, mà còn là một chiến lược lâu dài để đảm bảo sự phát triển bền vững trong tương lai.

Vai trò của nghiên cứu phát sinh loài ở Việt Nam vô cùng to lớn, không chỉ trong lĩnh vực khoa học mà còn ở các khía cạnh bảo tồn và phát triển bền vững. Các nghiên cứu này giúp làm sáng tỏ quá trình tiến hóa, xác định các loài nguy cấp và thiết lập các chiến lược bảo vệ hệ sinh thái hiệu quả, góp phần bảo tồn di sản thiên nhiên quý giá của đất nước. Đồng thời, nghiên cứu phát sinh loài còn mang lại giá trị kinh tế và xã hội, thúc đẩy du lịch sinh thái và khai thác bền vững tài nguyên thiên nhiên. Tuy đối mặt với các thách thức về suy thoái môi trường và hạn chế nguồn lực, lĩnh vực này vẫn có triển vọng lớn nhờ ứng dụng công nghệ hiện đại và hợp tác quốc tế, hứa hẹn tiếp tục đóng góp cho khoa học toàn cầu và bảo tồn đa dạng sinh học trong tương lai.

1.3.2. Tổng quan nghiên cứu ADN ty thể linh trưởng ở trên thế giới

Trên thế giới, ADN ty thể được phát hiện vào những năm 1960 bởi Margit Nass và Sylvan Nass, và được nghiên cứu rộng rãi từ thập niên 1980 khi các nhà khoa học nhận ra tiềm năng to lớn của nó trong truy vết tổ tiên di truyền. Sau khi "Eva ty thể" được công bố (12); ADN ty thể nhanh chóng trở thành công cụ trung tâm trong di truyền học tiến hóa và nhân học. Các nghiên cứu đã sử dụng ADN ty thể để xây dựng cây phát sinh loài cho các nhóm động vật có xương sống (32).

Việc so sánh các trình tự ADN ty thể đóng vai trò trụ cột trong nghiên cứu phát sinh chủng loài học (phylogenetics), trong đó nó cho phép các nhà sinh học làm sáng tỏ mối quan hệ tiến hóa giữa các loài. Nó cũng cho phép việc kiểm tra mối liên hệ của các quần thể, và như vậy nó rất quan trọng trong lĩnh vực nhân chủng học và

sinh học, thông qua việc dựng lại cây phả hệ theo dòng mẹ. Vì vậy, những trình tự của ty thể được sử dụng nhằm đánh giá phát sinh chủng loại trên nhiều đối tượng khác nhau. Trên thế giới, nghiên cứu về mối quan hệ phát sinh loài dựa trên trình tự các vùng gene ty thể cũng đã được thực hiện trên một số loài như Vượn lớn châu Phi (African great apes) và một số loài vượn khác (33,34). Đặt biệt các trình tự cytochrome b, *D-loop*, *16S* trên ADN ty thể là các trình tự được sử dụng nhiều để đánh giá sự đa dạng và mối quan hệ phát sinh loài của Linh trưởng như tinh tinh, bonobo, khỉ đột, đười ươi. Finstermeier đã xây dựng cây phát sinh loài của gần như toàn bộ các loài linh trưởng bằng cách sử dụng trình tự ADN ty thể hoàn chỉnh, cung cấp khung tham chiếu cho các nghiên cứu về tiến hóa, đa dạng và thích nghi (35).

ADN ty thể là công cụ di truyền quan trọng trong nghiên cứu tiến hóa do có đặc điểm di truyền theo dòng mẹ, không tái tổ hợp và có tốc độ tiến hóa nhanh. Các vùng khác nhau trong ADN ty thể tiến hóa với tốc độ không đồng đều, tạo điều kiện thuận lợi cho việc lựa chọn các chỉ thị (marker) di truyền phù hợp với từng mục tiêu nghiên cứu như phân tích đa dạng di truyền trong loài, xác định loài, hoặc khảo sát quan hệ phát sinh giữa các loài. Với đặc điểm di truyền theo dòng mẹ phổ biến ở hầu hết các loài động vật, ADN ty thể đặc biệt hữu ích trong việc truy vết phả hệ theo dòng mẹ. Ngoài ra, do không chịu ảnh hưởng của tái tổ hợp và nhiều đột biến xảy ra trên ADN ty thể là trung tính – không tác động đến khả năng sống sót của sinh vật – nên các kiểu gene được bảo tồn qua thời gian tiến hóa. Nhờ đó, ADN ty thể trở thành công cụ lý tưởng để phân tích mối quan hệ di truyền giữa các cá thể trong cùng loài hoặc giữa các loài khác nhau, từ đó ước tính được khoảng cách tiến hóa. Càng có quan hệ tiến hóa xa, mức độ khác biệt di truyền giữa các loài càng lớn. Năm 1998, Johns và cs đã thu thập khoảng 2000 trình tự gene *cytochrome b* của ty thể từ ngân hàng gene để tính toán và so sánh các mức độ khoảng cách di truyền giữa các loài cùng thế hệ, các chi họ và các lớp để phân loại động vật có xương sống (36). Kết quả cho thấy sự khác biệt về gene trong các nhóm động vật có xương sống, đặc biệt là lưỡng cư và bò sát. Sau đó, *cytochrome b* của ty thể cũng được ứng dụng có hiệu quả để khuếch đại và so sánh sự khác biệt giữa con người và các loài động vật gần với con người (37).

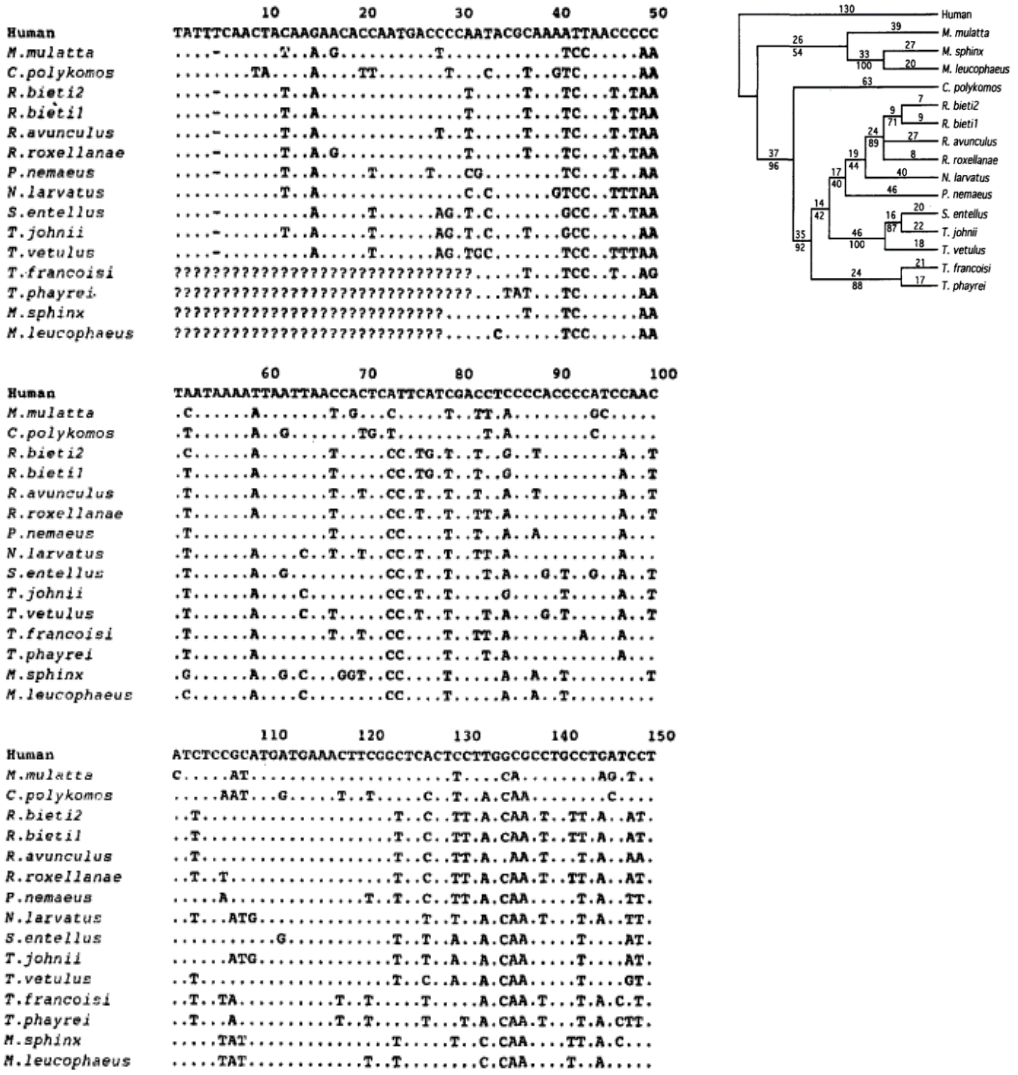
Gene Cytochrome b

ADN ty thể là trình tự gene được cho là ít biểu hiện sự biến đổi nhưng cho thấy đủ sự biến đổi giữa các điểm cụ thể để cho phép ước tính mức độ phát sinh loài. Các nghiên cứu đã sử dụng nhiều locus khác nhau trên hệ gene ty thể như *12S rRNA,*

16S rRNA, COII và khác. Tuy nhiên, locus chính được sử dụng để phân biệt loài cho đến gần đây là *cytochrome b* xuất hiện giữa các base 14,747 và 15,887 trong ADN ty thể của người. *Cytochrome b* là một thành viên trong họ cytochrome (a, b và c). Mỗi phân tử mang 2 nhân heme có khối lượng khoảng 40 kDa và hấp thu ánh sáng ở bước sóng 556-560 nm. *Cytochrome b* có vai trò vận chuyển điện tử trong quá trình hô hấp ở ty thể. Gene *cytochrome b* được sử dụng rộng rãi trong nghiên cứu tiến hóa vì quá trình đột biến xảy ra chậm chạp. Nhiều phần của gene *cytochrome b* được bảo tồn do chức năng của nó còn hạn chế. Mặc dù gene *cytochrome b* đột biến chậm nhưng tỷ lệ những đột biến im lặng lại xảy ra rất cao. Vì vậy, *cytochrome b* có tính ổn định và biến đổi phù hợp để nghiên cứu sâu hơn về những mối quan hệ tiến hóa. Phần lớn những vị trí hay xảy ra biến động nằm trong vùng mã hóa cho màng vận chuyển, các gốc amin và gốc cacboxyl ở vùng đầu tận cùng. *Cytochrome b* xuất hiện trong các sinh vật sống ở các cấp độ phân loại khác nhau; trong đó phần chính của cytochromes b có trong 95,9% ở động vật, trong đó 99,2% ở sinh vật nhân chuẩn và 79,7% ở động vật có vú (38). Protein *cytochrome b* được mã hóa từ 1048 phân tử trong số 1689 phân tử, số lượng gốc axit amin trong *cytochrome b* có thể thay đổi từ 300 đến 563 axit amin và được đặc trưng bởi nhiều chuỗi aa khác nhau *(dựa vào cơ sở dữ liệu Swiss-Prot, Zamyatnin và cs)*. Dù chúng ta biết rằng, các phân tử protein có cùng độ dài nhưng khác về trình tự aa, có thể có cùng chức năng, do đó, dù khác nhau về các thứ tự aa, thì chức năng của *cytochrome b* sẽ không khác nhau. *Cytochrome b* đóng vai trò tham gia vào quá trình truyền điện tử trong ty thể, có chức năng trong quá trình biến đổi thức ăn thành ATP, và chính đặc điểm cấu trúc đặc biệt của *cytochrome b* cho phép nó trở thành 1 protein xuyên màng (39).

Trình tự gene *cytochrome b* là một chỉ thị di truyền quan trọng, không chỉ giúp làm rõ các mối quan hệ phát sinh loài mà còn đóng góp đáng kể trong các nghiên cứu về cấu trúc quần thể, sự tiến hóa và bảo tồn linh trưởng. Nhờ tính linh hoạt, tốc độ tiến hóa trung bình, *cytb* vẫn tiếp tục là một công cụ thiết yếu trong di truyền học phân tử hiện đại. Trình tự gene *cytochrome b* (*cytb*) trong ADN ty thể đã được sử dụng rộng rãi trong nghiên cứu phát sinh loài ở các loài linh trưởng, nhờ khả năng cung cấp tín hiệu di truyền mạnh mẽ cho việc tái dựng quan hệ tiến hóa. Một nghiên cứu kinh điển của Yoder et al. (1996) đã sử dụng trình tự *cytb* để phân tích mối quan hệ phát sinh loài giữa các loài linh trưởng thuộc nhóm Strepsirrhini (gồm vượn cáo và cu li) (40) . Kết quả cho thấy dữ liệu *cytb* hỗ trợ giả thuyết rằng các loài vượn cáo Madagascar (Malagasy lemurs) và các loài lorisiforms Á–Phi (Asian–African lorises) là những nhóm có mối quan hệ chị em chặt chẽ, cung cấp bằng chứng di truyền cho sự hình thành đơn ngành trong các nhóm này. Năm 1998; sự khác nhau trên trình tự

cytochrome b giữa các loài linh trưởng ở các khu vực Châu Á (có cả Việt Nam) (Hình 1. 3) để phân tích mối quan hệ phát sinh loài, cho thấy khỉ Cựu Thế Giới chia thành hai phân họ, với Colobus (châu Phi) tách nhánh sớm nhất, trong khi *Rhinopithecus* và *Nasalis* tạo thành nhóm chị em với *Pygathrix*. Ngoài ra, ba loài *Voọc mũi hếch Bắc Bộ*, *R. roxellanae* và *R. bieti* có quan hệ gần gũi và nên được xếp vào cùng một phân chi, hỗ trợ giả thuyết rằng tổ tiên của khỉ colobine châu Á di cư từ châu Phi vào cuối Pliocene hoặc đầu Pleistocene (41).



Hình 1. 3. Sự khác nhau trên trình tự *cytochrome b* giữa các loài linh trưởng ở các khu vực Châu Á (có cả Việt Nam) và cây phát sinh loài dựa vào sự khác biệt trình tự này. (41)

Tiếp theo, Nascimento et al. (2005) đã tiến hành phân tích phát sinh loài và cấu trúc di truyền quần thể ở hai loài khỉ rú Nam Mỹ là Alouatta caraya và Alouatta belzebul thông qua dữ liệu trình tự gene *cytb*. Nghiên cứu cho thấy có sự chia sẻ các

haplotype tương đồng giữa các quần thể cách xa nhau về mặt địa lý. Điều này được lý giải là do sự tồn tại của các dạng đa hình tổ tiên vẫn được duy trì trong quần thể hiện tại, phản ánh những lịch sử tiến hóa lâu dài và phức tạp trong dòng giống của loài.

Trong vòng một thập kỷ gần đây, *cytb* tiếp tục là chỉ thị di truyền quan trọng trong nghiên cứu phát sinh loài linh trưởng. Pozzi và cs đã sử dụng *cytb* cùng với các chỉ thị khác để làm rõ hơn cây phát sinh loài của các loài linh trưởng Strepsirrhini, từ đó củng cố các giả thuyết về đa dạng sinh học và tiến hóa độc lập trên đảo Madagascar (25). Bên cạnh đó, Nater và cs đã kết hợp dữ liệu *cytb* với toàn bộ hệ gene để xác nhận sự phân tách loài giữa hai phân loài đười ươi Pongo tapanuliensis và Pongo abelii, đánh dấu sự phát hiện loài đười ươi thứ ba và mới nhất (42). Gần đây hơn, Reis đã sử dụng *cytb* cùng nhiều gene ti thể để hiệu chỉnh thời gian phân kỳ giữa các loài linh trưởng, giúp cải thiện độ chính xác trong phân tích niên đại phát sinh loài (43).

Bên cạnh vai trò trong việc xác định mối quan hệ phát sinh loài, gene *cytb* còn được sử dụng trong các nghiên cứu quần thể nhằm đánh giá mức độ đa dạng di truyền, cấu trúc quần thể và các mô hình lịch sử di cư. Ví dụ, các nghiên cứu ở khỉ mũ (*Sapajus*) và khỉ nhện (*Ateles*) đã sử dụng *cytb* để xác định các dòng di truyền khác biệt giữa các quần thể phân bố tại các khu vực sinh thái khác nhau của rừng Amazon và Trung Mỹ. Điều này giúp tái dựng lịch sử phân bố của loài và định hướng các chiến lược bảo tồn (44).

Gene *cytb* cũng rất hữu ích trong việc phát hiện loài mới hoặc loài có quan hệ gần. Nhiều trường hợp trước đây được xem là cùng một loài, sau khi phân tích *cytb* đã được phân tách thành các loài riêng biệt do có sự khác biệt rõ ràng về di truyền. Điều này đặc biệt quan trọng trong các vùng có mức độ đa dạng cao như Madagascar, Amazon, hay Đông Nam Á – nơi có nhiều loài linh trưởng chưa được nhận diện đầy đủ. *Cytb* đã đóng vai trò trung tâm trong việc làm sáng tỏ các phức hệ loài (species complexes) trong họ *Lemuridae*, *Callitrichidae*, và *Cercopithecidae*.

Gene *D-loop*

D-loop, hay còn gọi là vùng kiểm soát (control region), là phần không mã hóa dài nhất của hệ gene ty thể. Dù không trực tiếp mã hóa protein, vùng này có vai trò điều hòa trong sao chép và phiên mã ADN ty thể. *D-loop* có tốc độ đột biến cao, đặc

biệt trong các loài có tỷ lệ sinh sản nhanh, điều này làm cho nó trở thành công cụ lý tưởng trong nghiên cứu quần thể và mức độ cá thể.

Dựa trên giả thuyết được Clayton đề xuất từ năm 1982 và vẫn được công nhận đến nay, vùng *D-loop* trong ADN ty thể đóng vai trò trung gian trong quá trình sao chép sợi nặng, theo mô hình sao chép dịch chuyển sợi (strand-displacement model) (45). Trong mô hình này, sự tổng hợp của sợi ADN mới bắt đầu từ các đoạn RNA mồi có nguồn gốc từ promoter của sợi nhẹ (Light Strand Promoter – LSP), khởi phát tại vị trí OH trên sợi nặng (H) và thường kết thúc ở vùng gene Pro-*tRNA*. Tên gọi "*D-loop*" bắt nguồn từ cấu trúc đặc trưng của nó gồm ba sợi: một sợi con mới tổng hợp thay thế tạm thời sợi mẹ, tạo nên một vòng lặp chuyển vị (displacement loop). *D-loop* là vùng không mã hóa dài khoảng 1 kilobase trong tổng số khoảng 16 kilobase của hệ gene ty thể người, đóng vai trò là vùng điều hòa chính cho sự sao chép và phiên mã. Các nghiên cứu cho thấy trình tự *D-loop* có khả năng hình thành cấu trúc thứ cấp dạng thân–vòng giống như *tRNA*, và thường có sự chèn hoặc mất đoạn làm thay đổi kích thước của vùng này giữa các loài (46,47). Đặc biệt, *D-loop* chứa nhiều trình tự lặp lại có độ dài và số lượng biến đổi, góp phần tạo ra sự khác biệt về kích thước hệ gene ty thể giữa các loài mà không liên quan đến số lượng gene mã hóa. Do đó, *D-loop*, còn được gọi là vùng kiểm soát (control region), là một đặc điểm tiến hóa linh hoạt của hệ gene ty thể và thường được sử dụng trong các nghiên cứu về di truyền quần thể và tiến hóa loài (46–48).

Sự biến thiên di truyền cao trong *D-loop* phản ánh các quá trình lịch sử như phân tán, nút thắt cổ chai di truyền (bottleneck), hoặc tái tổ hợp di cư. Ví dụ, nghiên cứu của Vilà trên chó sói và chó nhà cho thấy rằng *D-loop* giúp làm rõ nguồn gốc đơn thuần của chó từ sói xám (49). Ngoài ra, vùng này cũng được ứng dụng rộng rãi trong nghiên cứu động vật hoang dã, giúp truy vết nguồn gốc cá thể như trong nghiên cứu của (50) về loài thằn lằn ở Úc. Tuy nhiên, tốc độ tiến hóa cao cũng gây ra sự tích tụ đột biến trung tính hoặc có thể gây nhầm lẫn trong phân tích các mối quan hệ sâu xa do hiện tượng bão hòa đột biến. Do đó, *D-loop* thường được sử dụng kết hợp với các gene mã hóa khác để tăng độ chính xác của phân tích phát sinh loài (51).

Trong vòng 10 năm trở lại đây, nhiều nghiên cứu đã khai thác vùng *D-loop* của ADN ty thể như một chỉ thị phân tử hiệu quả trong việc phân tích đa dạng di truyền và tiến hoá ở động vật. Ở dê Galla tại Kenya, phân tích 600 bp của *D-loop* đã ghi nhận 90 vị trí biến dị và 68 haplotype, cho thấy mức độ đa dạng di truyền rất cao, đồng thời AMOVA cũng chỉ ra rằng phần lớn biến dị tồn tại trong quần thể (52). Nghiên cứu ở bò bản địa Thổ Nhĩ Kỳ cũng xác định được 31 haplotype từ 62 mẫu cá

thể, phản ánh sự tồn tại của nhiều haplogroup và khả năng lai giữa các giống bò taurine và indicine (53) . Đối với bò Yak Bắc Á, phân tích toàn bộ ADN ty thể với hơn 500 mẫu đã xác định ba nhánh di truyền lớn với thời điểm tách nhánh khoảng 0,5 triệu năm trước, củng cố vai trò của *D-loop* trong phân tích phát sinh loài (54) . Về voi châu Á, tổng quan từ nhiều quốc gia cho thấy sự tồn tại của hai clade lớn (α và β), có khả năng phân hoá khoảng 1,5 triệu năm trước và không phân bố theo địa lý cụ thể, điều này cho thấy mức độ di cư và pha trộn gene mạnh (55). Ở Việt Nam, nghiên cứu trên năm giống vịt nội ghi nhận 17 haplotype với đa dạng nucleotide khá cao, phù hợp cho nghiên cứu phát sinh loài phân tử và bảo tồn nguồn gene (56). Những phát hiện này cho thấy vùng *D-loop* không chỉ giúp giải thích rõ hơn về cấu trúc di truyền và sự đa dạng trong quần thể mà còn góp phần làm sáng tỏ các mối quan hệ phát sinh loài và lịch sử tiến hoá của nhiều loài động vật.

Gene 16S rRNA

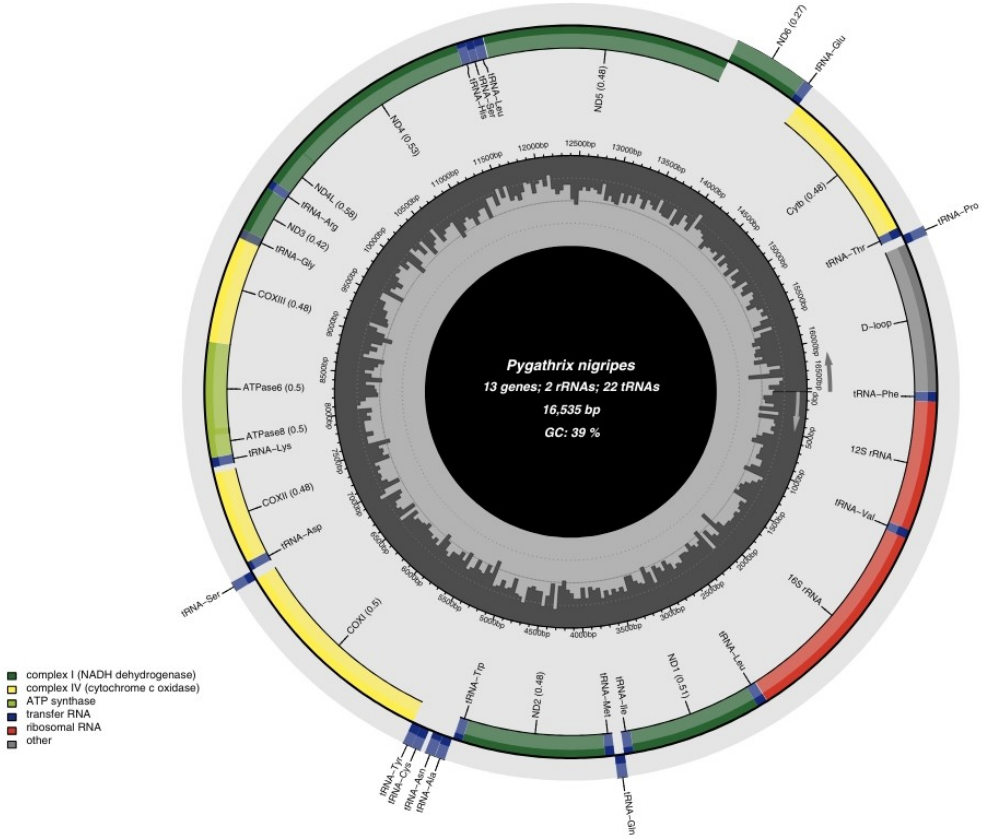
Một trong những trình tự gene trên ADN ty thể cũng thường được sử dụng trong nghiên cứu phát sinh loài là trình tự gene *16S* rRNA. Trình tự này thuộc tiểu đơn vị nhỏ của ribosome ty thể và có mặt ở hầu hết các sinh vật sống, do đó là chỉ thị phân tử hữu ích trong các phân tích tiến hóa. Trong một nghiên cứu tiêu biểu, Noda và cs (2001) đã giải trình tự gene *16S* rRNA của ADN ty thể từ 35 cá thể tinh tinh, 13 bonobo, 10 khỉ đột, 16 đười ươi và 23 cá thể vượn. Kết hợp dữ liệu thu được với các trình tự đã công bố, họ phát hiện rằng mức độ đa dạng di truyền ở các loài vượn người cao hơn ở người. Phân tích cây phát sinh loài từ dữ liệu này cũng cho thấy hai phân loài đười ươi là Sumatra và Bornean được phân tách rõ rệt, trong đó đười ươi Sumatra có đa dạng nucleotide cao hơn đáng kể. Một số loài vượn không hình thành các cụm đơn ngành, cho thấy sự biến thiên di truyền trong loài có thể tương đương với mức biến thiên giữa các loài trong chi Hylobates (34).

Phân tử *16S* rRNA có cấu trúc thứ cấp phức tạp gồm bốn domain, hình thành bởi các đoạn cặp base không hoàn toàn và tạo ra những đoạn phồng (loops) đặc trưng. Cấu trúc này cho phép phân tử ổn định và có thể tồn tại qua thời gian tiến hóa dài. Trình tự gene *16S* chứa các vùng bảo tồn cao (conserved) được duy trì ở nhiều loài và các vùng biến đổi (variable) phản ánh sự khác biệt về mặt tiến hóa giữa các loài hoặc dòng vi sinh vật. Có ít nhất tám vùng biến đổi chính (V1–V8), mỗi vùng dài khoảng 18–20 bp, xen kẽ với các vùng bán bảo tồn (semi-conserved) (57).

Gần đây, công nghệ giải trình tự thế hệ mới (NGS) đã cho phép khuếch đại và phân tích trình tự *16S* rRNA trên quy mô lớn. Các nghiên cứu sử dụng phương pháp

phân cụm theo đơn vị phân loại vận hành (OTU) hoặc các phương pháp giải trình tự amplicon theo đơn phân tử (ASV) đã làm rõ hơn sự đa dạng di truyền và phân loại vi sinh vật trong nhiều hệ sinh thái (58,59).

Tại Việt Nam, nghiên cứu về ADN ty thể bắt đầu phát triển mạnh từ đầu những năm 2000 với nhiều công trình giải trình tự các đoạn ADN ty thể ở người, động vật bản địa và các loài có nguy cơ tuyệt chủng. Ví dụ, nhiều nghiên cứu đã ứng dụng gene COI và *cytb* để nhận diện loài gà rừng, rắn, cá, và các động vật có vú nhỏ trong chương trình bảo tồn và phân loại học. Gần đây, các nhà khoa học cũng thực hiện các nghiên cứu ADN ty thể toàn bộ ở người Việt để hiểu rõ cấu trúc di truyền dân cư và mối quan hệ với các quần thể Đông Á khác (60). Trong lĩnh vực bảo tồn, ADN ty thể đóng vai trò quan trọng giúp xác định loài, tránh lai tạp và lập kế hoạch tái thả hiệu quả. ADN ty thể được ứng dụng rộng rãi trong nghiên cứu phát sinh loài và bảo tồn các loài linh trưởng, đặc biệt là các loài nguy cấp tại khu vực Đông Nam Á – trong đó có Việt Nam. Trong hơn hai thập kỷ qua, nghiên cứu về ADN ty thể ở linh trưởng Việt Nam đã đạt được nhiều tiến triển đáng kể. Năm 2007, Roos và cộng sự đã phân tích hệ thống phân loại của linh trưởng ở Việt Nam và Đông Nam Á dựa trên gene ty thể, mở đường cho các phân tích sau này (61). Hệ thống phân loại linh trưởng Việt Nam cũng được cập nhật nhiều lần giữa 2007–2014 dựa trên phân tích ADN ty thể của các chi như *Pygathrix*, *Trachypithecus* và *Nomascus* (62). Năm 2010, Thịnh cùng nhóm tại Viện Nghiên cứu và Cứu hộ Linh trưởng (EPRC) Cúc Phương chính thức mô tả loài mới *Nomascus annamensis* bằng cách sử dụng dữ liệu ADN ty thể *cytochrome b* kết hợp phân tích hình thái (63). Năm 2019, dựa trên dữ liệu ADN ty thể cứu hộ từ Việt Nam, các nghiên cứu đã chứng minh rằng loài *Trachypithecus crepusculus* có nguồn gốc lai (30). Nghiên cứu đáng chú ý khác là năm 2022, Roos và cộng sự đã giải trình tự toàn bộ ADN ty thể của loài *Nomascus siki*, tiết lộ sự không chính xác trong ghi nhận địa lý mẫu vật và nhấn mạnh vai trò của dữ liệu ADN ty thể trong phân loại chính xác hơn (64). Ngoài ra, nghiên cứu về *Chà vá chân nâu* ở Quảng Bình – mặc dù chưa công bố trên tạp chí quốc tế – đã thu thập mẫu *cytb* để đánh giá hiện trạng quần thể cho kế hoạch thiết lập khu bảo tồn.



Hình 1. 4. ADN ty thể của *Chà vá* (65)

1.4. Mục tiêu nghiên cứu

Hiện nay, dữ liệu về toàn bộ hệ gene ty thể của nhiều loài linh trưởng tại Việt Nam vẫn còn hạn chế, thậm chí thiếu hụt đối với một số nhóm. Điều này gây khó khăn trong việc xác định chính xác mối quan hệ giữa các phân loài hoặc quần thể địa phương, cũng như làm giảm độ tin cậy của các ước tính về thời gian phân tách trong lịch sử tiến hóa của chúng. Nghiên cứu của Yoder (40) đã chứng minh rằng trình tự *cytochrome b* có thể cung cấp tín hiệu phát sinh loài rõ ràng cho các nhóm linh trưởng Strepsirrhini, trong khi nghiên cứu của Noda (34) chỉ ra rằng gene *16S* rRNA có tính bảo tồn cao nhưng vẫn đủ biến đổi để phân biệt các loài trong chi Hylobates. *D-loop*, *16S rRNA* và *cytochrome b* là những chỉ thị phân tử quan trọng giúp đánh giá mức độ tương đồng – khác biệt di truyền giữa các loài, qua đó hỗ trợ tái dựng cây phát sinh loài một cách tin cậy. Tuy nhiên, cần lưu ý rằng việc phân tích đơn lẻ từng chỉ thị vẫn có những giới hạn nhất định. Trong những năm gần đây, các nghiên cứu dựa trên toàn bộ Hệ gene Ty thể được xem là có độ phân giải cao hơn và phản ánh chính xác hơn mối quan hệ tiến hóa. Dù vậy, số lượng nghiên cứu công bố theo hướng này vẫn còn hạn chế, nên việc mở rộng dữ liệu Hệ gene Ty thể sẽ đặc biệt quan trọng cho các phân

tích phát sinh loài trong tương lai. Ngoài ra, những nghiên cứu trên toàn bộ Hệ gene Ty thể sẽ là minh chứng quan trọng trong

Mục tiêu tổng quát:

Giải trình tự và phân tích đặc điểm hệ gene ty thể hoàn chỉnh của một số loài linh trưởng phân bố tại khu vực Tây Nguyên Việt Nam, từ đó so sánh sai khác trình tự và xác định mối quan hệ phát sinh loài của các loài nghiên cứu trong bối cảnh so sánh với các dữ liệu đã công bố trên thế giới.

Mục tiêu cụ thể:

Phân tích đặc điểm cấu trúc, thành phần nucleotide và tổ chức gene của hệ gene ty thể hoàn chỉnh ở một số loài linh trưởng Tây Nguyên Việt Nam.

So sánh sai khác trình tự di truyền ty thể giữa các cá thể nghiên cứu và các trình tự tương đồng đã được công bố, phục vụ cho việc đánh giá quan hệ phát sinh loài.

Tái dựng cây phát sinh loài dựa trên các marker ty thể và toàn bộ hệ gene ty thể nhằm làm rõ vị trí phân loại và quan hệ tiến hóa của các loài linh trưởng nghiên cứu.

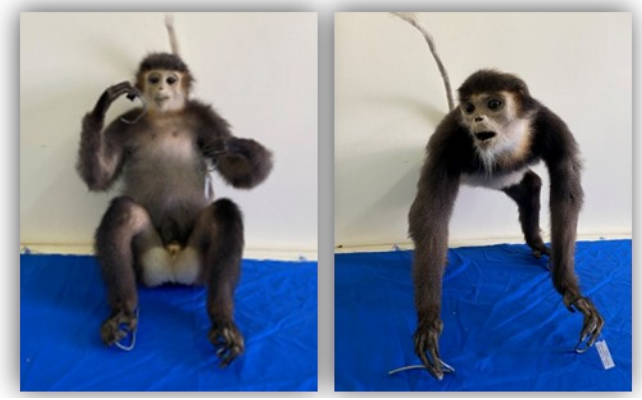
Chương 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng nghiên cứu

Đối tượng nghiên cứu là các mẫu tươi của các loài thú Linh trưởng hiện đang lưu giữ tại Bảo tàng Sinh học – Viện Nghiên cứu Khoa học Tây Nguyên. Các loài này đã được định danh về mặt hình thái (Hình 2) bởi GS. Đặng Huy Huỳnh từ những năm 1990 lưu trữ ở dạng tiêu bản tại bảo tàng và 1 phần mẫu mô tươi được cắt và lưu trữ trong nito lỏng -80ºC bao gồm các tiêu bản như (Bảng 2. 1).

Bảng 2. 1. Các loài thú linh trưởng được sử dụng trong nghiên cứu.

STT	Tên thường gọi	Tên khoa học	Số lượng
1	*Chà vá chân đen*	*Pygathrix nigripes*	1
2	*Khỉ mặt đỏ*	*Macaca arctoides*	1
3	*Chà vá chân nâu*	*Pygathrix nemaeus*	1
4	*Khỉ vàng*	*Macaca mulatta*	1
5	*Chà vá chân xám*	*Pygathrix cinerea*	1
Tổng cộng			**5**



Hình 2. Loài thú linh trưởng được lưu trữ trong bảo tàng.

2.2. Phương pháp nghiên cứu

2.2.1. Tách chiết ADN và giải trình tự

Mẫu mô các cá thể Linh trưởng khu vực Tây Nguyên Việt Nam lưu trữ trong bảo tàng lưu trữ mẫu tươi tại Bảo tàng Sinh học – Viện Nghiên cứu Khoa học Tây Nguyên, sau khi được sự cho phép của cơ quan có thẩm quyền đã được thu nhận được rửa sạch bằng dung dịch muối sinh lý 0,9%. Sau đó, mẫu mô được trữ ở -8ºC trong eppendorf 1,5ml và chuyển nhanh về phòng thí nghiệm để thực hiện thí nghiệm.

Đầu tiên, ADN tổng số được chiết xuất từ mô của các cá thể Linh trưởng bằng bộ kit QIAamp ADN Mini (Qiagene) theo hướng dẫn của nhà sản xuất. Bộ kit hoạt động dựa trên nguyên lý công nghệ màng silica đã được thương mại hóa rộng rãi. Quy trình chuẩn của bộ kit không yêu cầu các bước đặc biệt để tách riêng ADN ty thể, do đó, tổng lượng ADN (cả hai loại ADN nhân và ADN ty thể) được thu nhận.

Thu nhận ADN tổng bao gồm các bước:

Ly giải tế bào: Các đệm ly giải mạnh (như Buffer AL) và enzyme Proteinase K được sử dụng để phá vỡ màng tế bào và màng nhân, giải phóng tất cả các thành phần bên trong, bao gồm cả ADN hệ gene và ADN ty thể.

Gắn kết ADN: Trong điều kiện đệm có nồng độ muối và pH thích hợp, tất cả các phân tử ADN (kích thước lớn hay nhỏ) sẽ gắn kết một cách chọn lọc và hiệu quả với màng silica trong cột ly tâm.

Rửa: Các bước rửa với Buffer AW1 và AW2 loại bỏ tạp chất, protein và các chất ức chế PCR khác (ví dụ như cation hóa trị hai) mà không làm ảnh hưởng đến sự gắn kết của ADN trên màng.

Rửa giải: ADN tinh sạch được rửa giải khỏi màng silica bằng đệm elution có nồng độ muối thấp (Buffer AE) hoặc nước cất.

Thu nhận ADN ty thể (sử dụng REPLI-g Mitochondrial ADN Kit) bao gồm các bước

Chuẩn bị và biến tính ADN: Cho 1–10 µl ADN tổng đã tinh sạch (thu được từ kit QIAamp ADN Mini) vào ống. Điều chỉnh thể tích lên 20 µl bằng Nước không có RNase.

Chuẩn bị hỗn hợp khuếch đại (Master Mix): cho 27 µl REPLI-g mt Reaction Buffer vào 2 µl bộ mt Primer Mix (cho linh trưởng); sau đó trộn đều hỗn hợp này.

Trộn mẫu và ủ nhiệt: Thêm 29 µl hỗn hợp khuếch đại vào ống chứa 20 µl ADN mẫu.

Biến tính: Ủ mẫu trong 5 phút ở 75ºC.

Làm nguội: Nhanh chóng làm nguội về nhiệt độ phòng (15–25ºC).

Phản ứng khuếch đại (Amplification)

Thêm Enzyme: Thêm 1 µl REPLI-g Midi ADN Polymerase (giữ lạnh trên đá) vào ống phản ứng. Trộn đều.

Khuếch đại: Ủ mẫu trong 8 giờ ở 33ºC. Phản ứng đẳng nhiệt này đã làm giàu ADN ty thể.

Bất hoạt Enzyme: Bất hoạt enzyme bằng cách ủ mẫu trong 3 phút ở 65ºC.

Kit này sử dụng công nghệ khuếch đại gen tổng thể bằng dịch chuyển chuỗi đa điểm (MDA) có chọn lọc, sử dụng các cặp mồi chuyên biệt chỉ bám vào ADN ty thể, khuếch đại nó lên tới 40 triệu lần.

Chất lượng mẫu ADN được đánh giá bằng một số phương pháp: đo nồng độ bằng phương pháp huỳnh quang (Qubit), đo tỷ lệ OD260/OD280 bằng phương pháp đo độ hấp thụ, và điện di trên gel agarose để xác định kích thước ADN. Chỉ những mẫu đáp ứng các tiêu chí về chất lượng (nồng độ ≥2 ng/μL, lượng ≥90 ng, và OD260/OD280 ≥ 1.70) mới được xem là phù hợp cho các bước thí nghiệm giải trình tự toàn hệ gene bằng bộ kit NEBNext Ultra II DNA Library Prep Kit for Illumina. Các mẫu có kích thước DNA <1000 bp bị loại bỏ.

Sau đó, thư viện giải trình tự toàn hệ gene được chuẩn bị bằng bộ kit NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs) theo hướng dẫn của nhà sản xuất. Nồng độ thư viện được xác định bằng phương pháp huỳnh quang và kích thước trung bình của thư viện được xác định bằng máy Bioanalyzer (Agilent) theo hướng dẫn đánh giá thư viện của Illumina. Các mẫu được xem là phù hợp để giải trình tự khi nồng độ ≥0.50 ng/μL (đối với kích thước gene <1 Gb) hoặc ≥2 ng/μL (đối với kích thước gene >1 Gb). Các thư viện sau đó được giải trình tự bằng phương pháp giải trình tự thế hệ mới 150 PE trên máy NovaSeq (Illumina) (66).

2.2.2. Phương pháp lắp ráp hệ gene ty thể

Hệ gene ty thể được lắp ráp lại *de novo* sử dụng công cụ GetOrganelle với các thông số tối ưu để lắp ráp *de novo* trình tự hệ gene mẫu và chất lượng được đánh giá bằng Quast. (Lắp ráp *de novo* là phương pháp xây dựng trình tự hệ gene từ các đoạn đọc ngắn hoặc dài mà không cần dựa trên trình tự hệ gene tham chiếu). Kết quả lắp ráp cho thấy chiều dài contig phù hợp với kích thước của hệ gene ty thể, khoảng 16 kb. Dữ liệu trình tự thô ban đầu được làm sạch bằng công cụ fastp v0.23.1 trước khi lắp ráp. Chất lượng lắp ráp được đánh giá bằng công cụ Quast v5.2.0 và sắp xếp các đoạn đọc khớp vào các contig đã lắp ráp. Việc này cho phép phát hiện các vùng có độ phủ thấp bất thường so với các vùng lân cận. Hệ gene ty thể được chú thích sử dụng hệ thống chú thích MITOS (phiên bản 6b33f95) với cơ sở dữ liệu ty thể chuyên dụng, cũng như Mitos WebServer và MitoFish. Mito Annotator được sử dụng để tạo bản đồ di truyền của toàn bộ hệ gene ty thể Linh Trưởng.

2.2.3. Phân tích tin sinh học

Phân tích sinh tin học bắt đầu bằng việc làm sạch dữ liệu trình tự thô bằng công cụ fastp v0.23.1 (67), loại bỏ các nucleotide có chất lượng kém, không đáng tin cậy hoặc không xác định (loại N) dựa trên điểm Phred (68). Sau khi làm sạch, các đoạn đọc được lắp ráp *de novo* bằng công cụ GetOrganelle (v1.7.7.0) (69). Dữ liệu trình tự thô, sau khi được tối ưu hóa, lắp ráp *de novo* và chú thích hệ gene, được lưu trữ ở định dạng tệp FASTQ, chứa thông tin đọc và trình tự cùng với điểm chất lượng tương ứng. Chất lượng lắp ráp *de novo* được đánh giá bằng công cụ Quast v5.2.0 và sắp xếp cục bộ các đoạn đọc vào các contig đã lắp ráp (68) và căn chỉnh cục bộ của các đoạn đọc lên các contig đã được lắp ráp. Điều này cho phép phát hiện các vùng có giá trị độ phủ sâu bất thường thấp so với các vùng lân cận, điều này đã được ghi nhận trong kết quả lắp ráp. Lắp ráp được chú thích bằng hệ thống chú thích MITOS (phiên bản 6b33f95) (70) với một cơ sở dữ liệu ty thể chuyên biệt.

Gene ty thể được chú thích bằng Mitos WebServer phiên bản 2.1.9 (70) và Mito Fish phiên bản 4.03 (71). Mito Annotator phiên bản 4.03 (71) đã được áp dụng để tạo ra một bản đồ gene của toàn hệ gene ty thể Linh Trưởng. Thành phần axit amin và nucleotide đã được đánh giá và so sánh cho tất cả 55 hệ gene ty thể của loài linh trưởng bằng cách sử dụng MEGA 11: AT-skew = (A − T)/(A + T) và GC-skew = (G − C)/(G + C). Các chuỗi *t-RNA* đã được căn chỉnh với các gene của các loài tương tự. Bằng cách so sánh với các gene ty thể khác của *Pygathrix*, các gene mã hóa protein (PCGs) giàu A + T và *r-RNA* đã được xác định. Giá trị Sử dụng Codon Đồng nghĩa Tương đối (RSCU) của toàn hệ gene ty thể Linh Trưởng đã được tính toán bằng MEGA 11. RSCU cho biết tần suất tương đối của một codon nhất định so với các codon khác có thể mã hóa cho cùng một axit amin. Chỉ số này giúp nghiên cứu các yếu tố ảnh hưởng đến quá trình dịch mã, sự chọn lọc tự nhiên, và các đặc điểm tiến hóa của gene. Cấu trúc thứ cấp của các dự đoán *t-RNA* đã được xác định bằng phần mềm *tRNA*scan-SE phiên bản 2.0 (72) và Mitos WebServer(70).

Phân tích các gene mã hóa protein và chức năng của chúng

Các gene mã hóa protein trong hệ gene ty thể rất cần thiết cho nhiều chức năng tế bào, chủ yếu là sản xuất năng lượng thông qua con đường phosphoryl hóa oxy hóa. Ở linh trưởng, những gene này bao gồm các gene mã hóa cho các tiểu đơn vị của NADH dehydrogenease (ND), cytochrome c oxidase (COX) và ATP synthase, tất cả đều quan trọng cho quá trình tổng hợp ATP. Những biến thể trong các gene này có thể ảnh hưởng đến hiệu quả chuyển hóa và sức khỏe tổng thể, làm cho chúng trở nên

quan trọng trong việc hiểu các thích nghi của loài, đặc biệt là đối với các áp lực môi trường (73).

Phân tích tỷ lệ A-T trong hệ gene ty thể

Sự phong phú tương đối của adenine và thymine (A-T) so với guanine và cytosine (G-C) trong các chuỗi gene đóng vai trò quan trọng trong các quá trình tiến hóa, ảnh hưởng đến các yếu tố như tỷ lệ đột biến và thiên lệch sử dụng codon. Những biến đổi trong nội dung A-T và G-C có thể ảnh hưởng đến hiệu suất dịch mã protein và khả năng thích ứng, định hình các quỹ đạo tiến hóa của loài theo các áp lực môi trường. Tỷ lệ A-T, đo lường sự phong phú tương đối của adenine (A) và thymine (T) trong hệ gene ty thể, là một chỉ số quan trọng về các áp lực tiến hóa. Các biến thể trong tỷ lệ A-T có thể phản ánh động lực quần thể lịch sử và sự thích ứng với các sinh cảnh sinh thái. Nghiên cứu cho thấy các loài có nội dung A-T cao có thể trải qua các áp lực chọn lọc khác nhau, ảnh hưởng đến quỹ đạo tiến hóa của chúng (Błażej *et al.*, 2017).

Phân tích Codon Usage Bias và ý nghĩa tiến hóa

Codon usage bias, đề cập đến việc ưu tiên sử dụng một số codon hơn những codon khác trong các gene mã hóa protein, bị ảnh hưởng bởi các yếu tố như áp lực đột biến và chọn lọc tự nhiên. Các nghiên cứu cho thấy sự thay thế nucleotide ưu tiên có thể có những tác động đáng kể đến hiệu suất dịch mã protein và sức khỏe tổng thể (Anderson *et al.*, 1981). Trong *Chà vá chân đen*, việc hiểu ưu tiên sử dụng codon là điều cần thiết để làm rõ cách mà thành phần gene ảnh hưởng đến sự thích ứng và thành công tiến hóa.

Nghiên cứu so sánh giữa các loài linh trưởng

Các nghiên cứu hệ gene so sánh giữa các loài *Chà vá* khác nhau sẽ giúp tiết lộ các mẫu hình khác biệt trong cấu trúc gene ty thể. Ví dụ, nghiên cứu của Liedigk (74) cho thấy sự biến đổi trong AT-skew giữa các loài *Chà vá* tương quan với các sinh cảnh sinh thái và sự thích ứng tiến hóa. Những phân tích so sánh như vậy có thể cung cấp những cái nhìn quý giá về mối quan hệ tiến hóa giữa các loài trong chi này và phản ứng của chúng đối với các thách thức môi trường.

Hiểu biết về sự đa dạng gene được thể hiện qua các gene mã hóa protein ty thể là điều thiết yếu để phát triển các chiến lược bảo tồn có mục tiêu. Tran nhấn mạnh rằng các quần thể có sự đa dạng gene lớn hơn có khả năng thích ứng tốt hơn với các môi trường đang thay đổi, do đó giảm nguy cơ tuyệt chủng (75). Việc giám sát di truyền,

kết hợp với dữ liệu hệ gene ty thể, nên trở thành một phần tích cực trong kế hoạch bảo tồn cho các loài đang bị đe dọa nghiêm trọng.

2.2.4. Cây phát sinh loài

Để xác định vị trí phân tử trong cây tiến hóa của các loài linh trưởng khác, 13 gene mã hóa protein (PCGs) từ 55 loài đã được căn chỉnh và phân tích bằng phương pháp Tối đa hóa khả năng (Maximum Likelihood) sử dụng IQ-tree phiên bản 2.2.2.6 (76). Mô hình thay thế phù hợp cho tập dữ liệu đã được chọn bằng ModelFinder phiên bản 2.2.0 (77) tích hợp trong IQ-TREE, dựa trên tiêu chí Bayesian Information Criterion (BIC). Hỗ trợ nhánh đã được thu được bằng cách sử dụng phương pháp bootstrap siêu nhanh (78). Số gia nhập GeneBank của tất cả các chuỗi được sử dụng được hiển thị trong Tài liệu bổ sung.

Chương 3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Kết quả giải trình tự gene

3.1.1. Kết quả chuẩn bị thư viện trình tự toàn bộ hệ gene Linh trưởng

Nồng độ ADN của các mẫu linh trưởng sau tách chiết dao động từ 2,67–9,94 ng/µL (≥2 ng/µL), với tổng lượng ADN thu được từ 133,5–497 ng (Bảng 3. 1). Tỷ lệ hấp thụ quang học OD260/OD280 nằm trong khoảng 1,70–1,96 (≥ 1.70), phù hợp với ngưỡng đánh giá ADN có độ tinh sạch tốt (≈1,7–2,0) cho thấy tất cả các mẫu đều đạt tiêu chuẩn về chất lượng và số lượng để sử dụng cho cho xây dựng thư viện giải trình tự thế hệ mới (NGS). Trong số các mẫu phân tích, mẫu *Chà vá chân đen* có nồng độ và tổng lượng ADN cao nhất (9,94 ng/µL; 497 ng), trong khi mẫu *Khỉ vàng* có giá trị thấp nhất (2,67 ng/µL; 133,5 ng); tuy nhiên, tất cả các mẫu đều vượt ngưỡng tối thiểu cần thiết cho xây dựng thư viện giải trình tự thế hệ mới.

Bảng 3. 1. Nồng độ các mẫu ADN của Linh trưởng sau tách chiết.

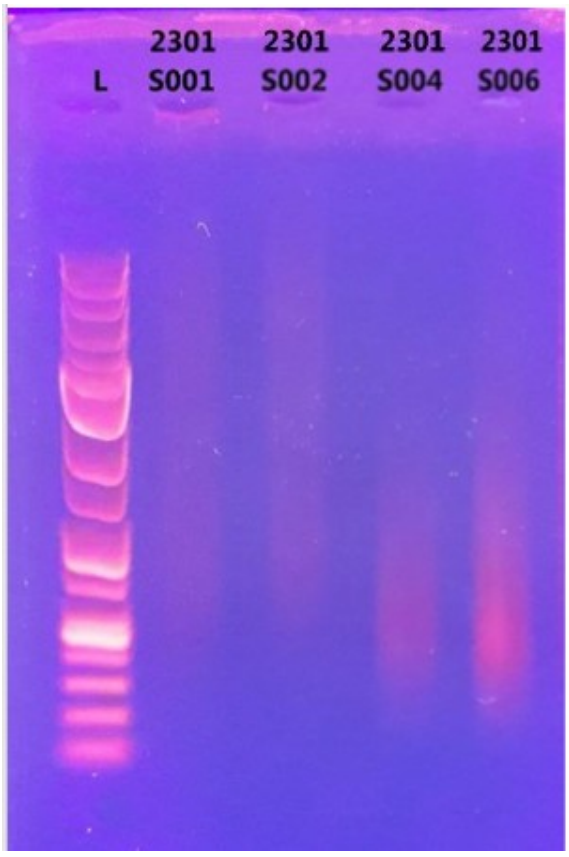
STT	Tên mẫu	Nồng độ ng/ul	OD260/280	Lượng (ng)
1	*Chà vá chân đen*	9.94	1.81	497.00
2	*Khỉ mặt đỏ*	5.82	1.72	291.00
3	*Chà vá chân nâu*	5.2	1.7	260.00
4	*Khỉ vàng*	2.67	1.96	133.50
5	*Chà vá chân xám*	6.42	1.72	321.00

Kết quả chuẩn bị thư viện giải trình tự toàn bộ hệ gene cho thấy nồng độ thư viện của các mẫu dao động trong khoảng từ 2,17 đến 57,00 ng/µL (Bảng 3. 2). Trong đó, các mẫu *Chà vá chân nâu*, *Khỉ mặt đỏ* và *Chà vá chân xám* có nồng độ thư viện cao, lần lượt đạt 57,00 ng/µL, 55,00 ng/µL và 51,00 ng/µL. Ngược lại, mẫu *Khỉ vàng* có nồng độ thư viện thấp nhất (2,17 ng/µL). Kích thước trung bình của các đoạn ADN trong thư viện nằm trong khoảng 412–554 bp. Mẫu *Khỉ mặt đỏ* có kích thước trung bình lớn nhất (554 bp), trong khi mẫu *Chà vá chân nâu* có kích thước trung bình nhỏ nhất (412 bp). Các giá trị này phù hợp với yêu cầu kỹ thuật của các nền tảng giải trình tự thế hệ mới, đảm bảo cho quá trình giải trình tự và phân tích dữ liệu ở các bước tiếp theo.

Bảng 3. 2. Kết quả chuẩn bị thư viện trình tự toàn bộ hệ gene Linh trưởng (Whole geneome sequencing).

STT	Tên mẫu	Nồng độ thư viện (ng/ul)	Kích thước trung bình (bp)
1	*Chà vá chân đen*	30.2	460
2	*Khỉ mặt đỏ*	55.00	554
3	*Chà vá chân nâu*	57.00	412
4	*Khỉ vàng*	2.17	522
5	*Chà vá chân xám*	51.00	513

Kết quả điện di sản phẩm PCR (Hình 3. 1) cho thấy các mẫu *Chà vá chân xám*, *Khỉ mặt đỏ*, *Chà vá chân nâu* và *Chà vá chân đen* đều xuất hiện các băng ADN rõ ràng đảm bảo cho các bước xử lý tiếp theo trong quy trình xây dựng thư viện

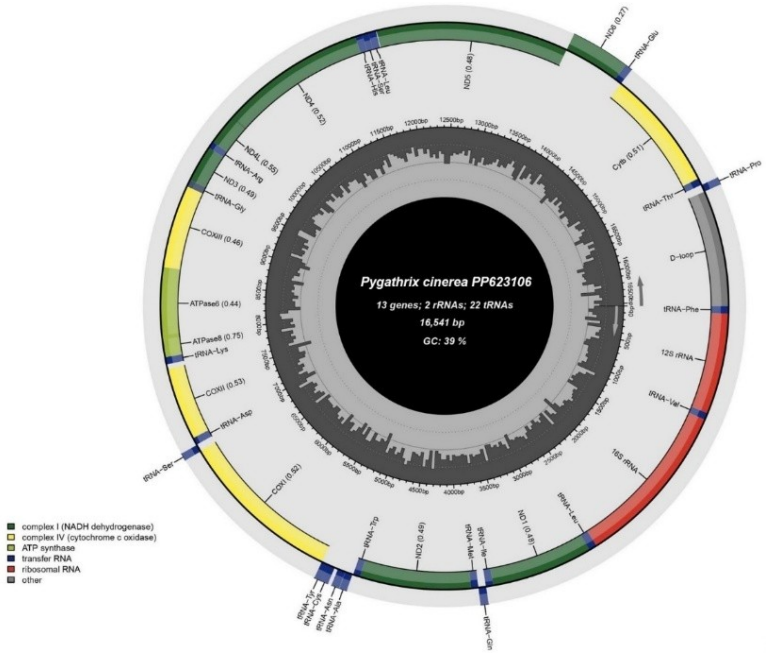


Hình 3. 1. Kết quả chạy PCR các mẫu Linh trưởng; S001: *Chà vá chân xám*; S002: *Khỉ mặt đỏ*; S004: *Chà vá chân nâu*; S006: *Chà vá chân đen.*

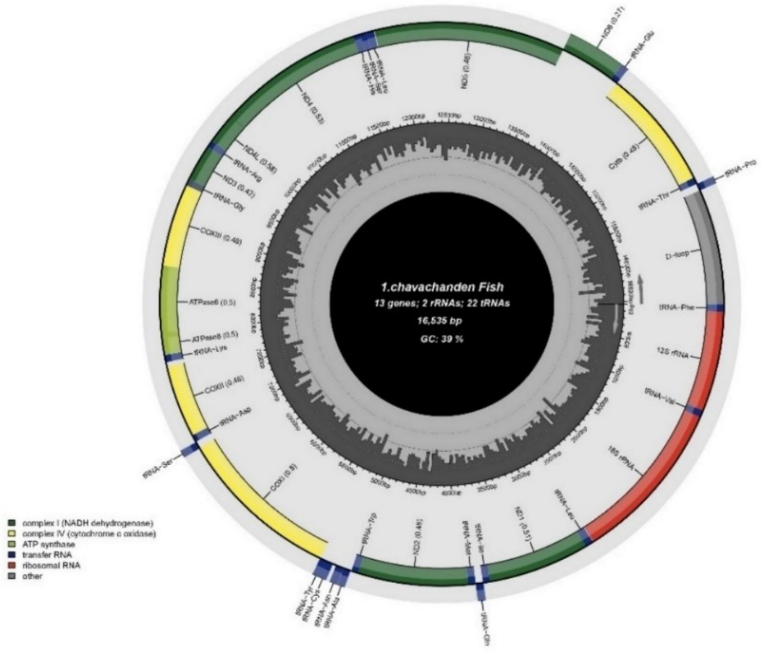
3.1.2. Kết quả giải trình tự toàn bộ hệ gene Linh trưởng

Giải mã được 5 trình tự toàn bộ hệ ty thể của 5 cá thể linh trưởng ở Tây Nguyên (Bảng. S1, S2, S3, S4, S5). Gene ty thể dạng vòng được mô tả với các gene được biểu

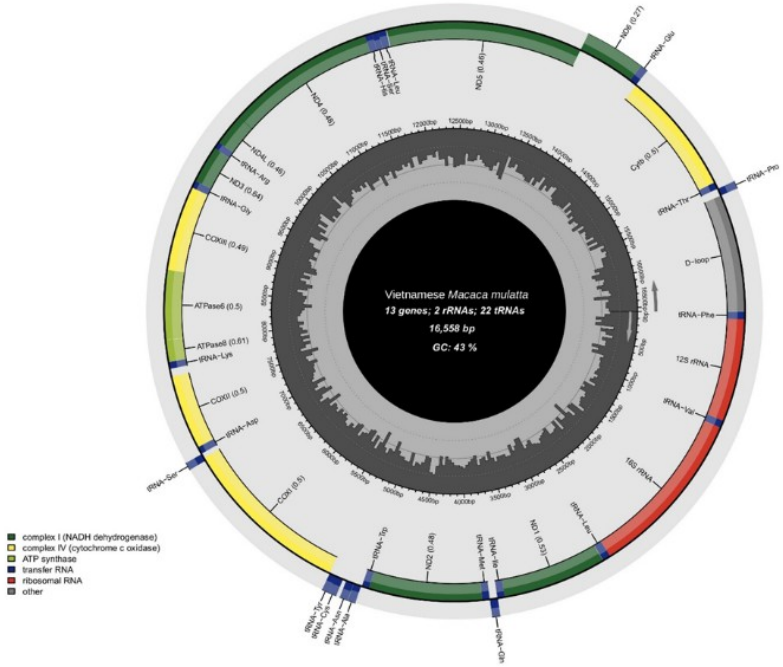
diễn bằng các khối màu. Vùng điều khiển được hiển thị bằng màu xám, các gene 16S (*rRNA* lớn) và 12S (*rRNA* nhỏ) được hiển thị bằng màu đỏ, 22 gene transfer RNA (*tRNA*) được ghi nhãn bằng màu xanh đậm, và 13 gene mã hóa protein (PCGs) được hiển thị bằng màu xanh lá cây và màu vàng. Các gene nằm trên sợi H được hiển thị trên vòng ngoài, trong khi các gene nằm trên sợi L được hiển thị trên vòng trong.



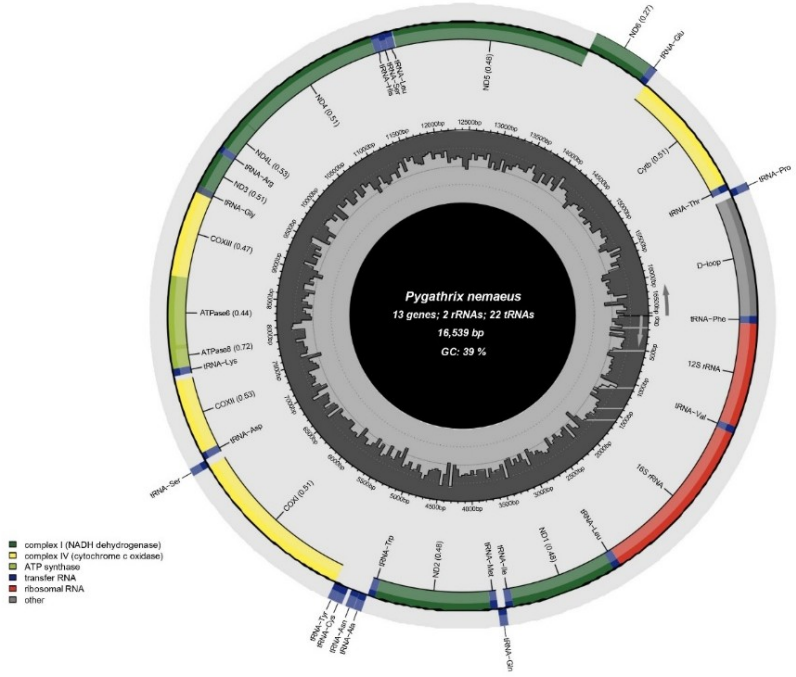
Hình 3. 2. Cấu trúc gene ty thể hoàn chỉnh của *Chà vá chân xám* Tây Nguyên Việt Nam.



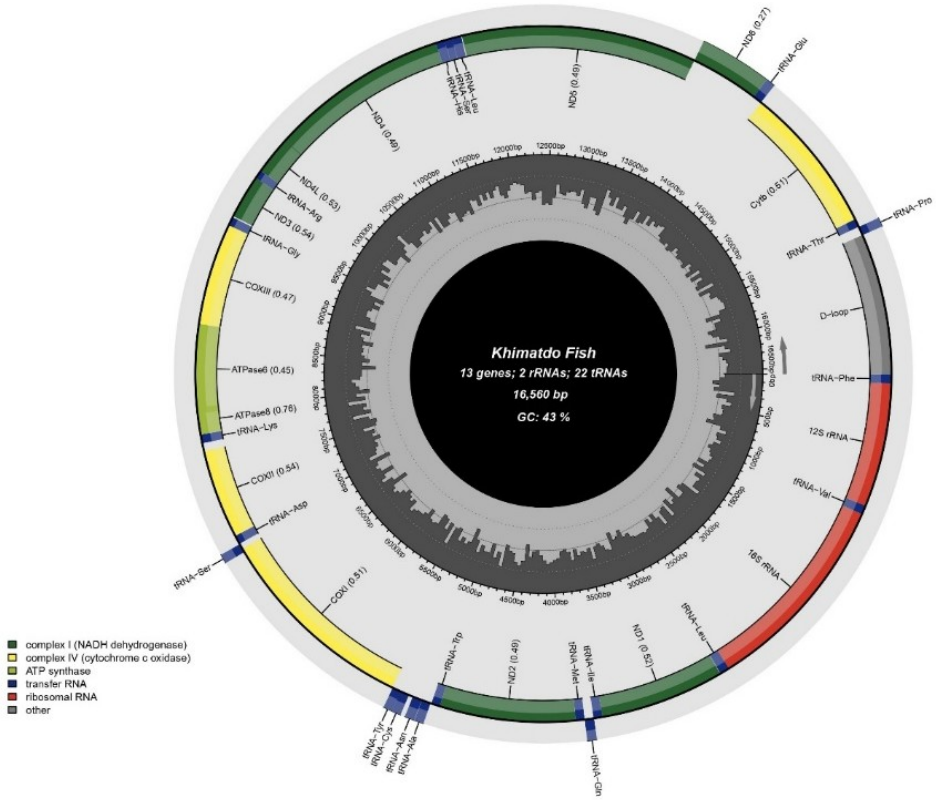
Hình 3. 3. Cấu trúc gene ty thể hoàn chỉnh của *Chà vá chân đen* Tây Nguyên Việt Nam.



Hình 3. 4. Cấu trúc gene ty thể hoàn chỉnh của *Khỉ vàng* Tây Nguyên Việt Nam.



Hình 3. 5. Cấu trúc gene ty thể hoàn chỉnh của *Chà vá chân nâu* Tây Nguyên Việt Nam.



Hình 3. 6. Cấu trúc gene ty thể hoàn chỉnh của *Khỉ mặt đỏ* Tây Nguyên Việt Nam.

3.2. Đặc điểm hệ gene ty thể hoàn chỉnh của một số cá thể Linh trưởng

3.2.1 Đặc điểm hệ gene ty thể hoàn chỉnh của cá thể *Chà vá chân xám* (*Pygathrix cinerea*) Tây Nguyên Việt Nam

Gene ty thể hoàn chỉnh của *Chà vá chân xám* bao gồm 16,541 cặp base (bp) và chứa 37 gene điển hình có trong ty thể của loài linh trưởng (Hình 3. 2; Bảng 3. 3). Những gene này bao gồm một vùng điều khiển (*D-loop*); hai gene ribosomal RNA (rRNA: 12S rRNA và *16S* rRNA); 22 gene transfer RNA (*tRNA*); và 13 gene mã hóa protein (PCGs). Phần lớn các gene nằm trên sợi H, ngoại trừ gene ND6 và tám gene *tRNA* (*tRNA*-Gln, *tRNA*-Ala, *tRNA*-Asn, *tRNA*-Cys, *tRNA*-Tyr, *tRNA*-Ser, *tRNA*-Glu) nằm trên sợi L. Gene ty thể của *Chà vá chân xám* thể hiện các trình tự tổ chức gene tương tự như các loài *Chà vá* khác. Tuy nhiên, đáng chú ý rằng chuỗi ADN ty thể của *Chà vá chân xám* Tây Nguyên Việt Nam được phân tích trong nghiên cứu này có kích thước lớn hơn một chút (16,541 bp) so với các chuỗi đã được công bố trước đó, bao gồm *Chà vá chân xám* (JQ_821842, 16,535 bp) và *Chà vá chân đen* (MH_064177, 16,536 bp) (79).

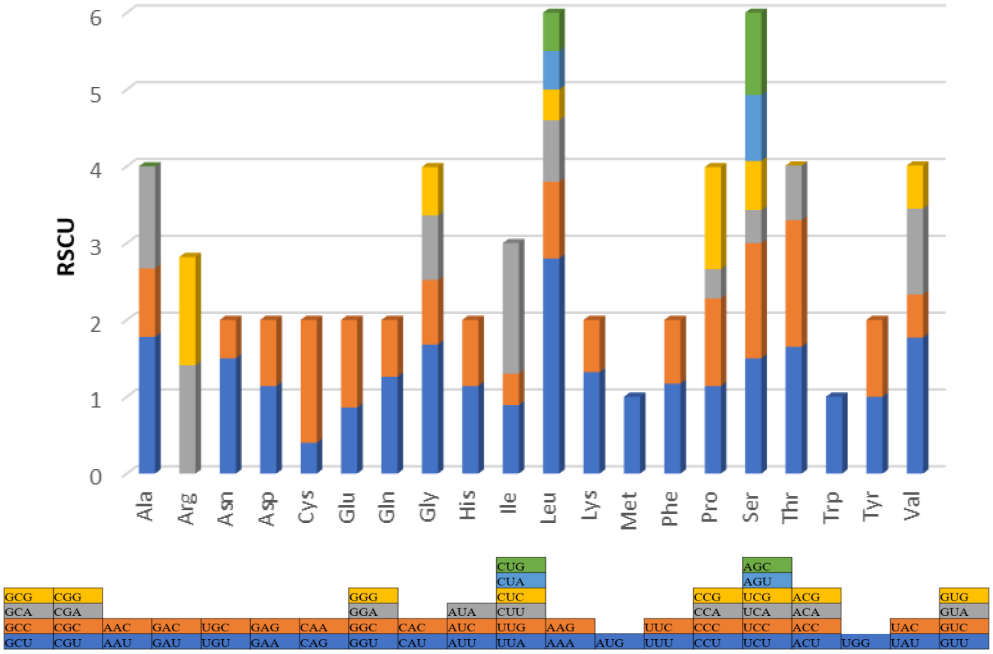
Bảng 3. 3. Vị trí và độ dài của 37 gene trong hệ gene ty thể hoàn chỉnh của cá thể *Chà vá chân xám* Tây Nguyên Việt Nam .

Tên	Bắt đầu	Kết thúc	Mạch	Chiều dài	Độ che phủ
trnF(ttc)	137	207	+	71	0
rrnS	208	1156	+	949	0
trnV(gta)	1157	1222	+	66	−2
rrnL	1221	2785	+	1565	0
trnL2(tta)	2786	2860	+	75	2
nad1	2863	3813	+	951	4
trnI(atc)	3818	3887	+	70	−3
trnQ(caa)	3885	3956	-	72	0
trnM(atg)	3957	4024	+	68	0
nad2	4025	5062	+	1038	4
trnW(tga)	5067	5133	+	67	7
tRNA(gca)	5141	5209	-	69	1
trnN(aac)	5211	5283	-	73	32
trnC(tgc)	5316	5381	-	66	0
trnY(tac)	5382	5446	-	65	1
cox1	5448	6989	+	1542	0
trnS2(tca)	6990	7058	-	69	3
trnD(gac)	7062	7130	+	69	1
cox2	7132	7791	+	660	68
trnK(aaa)	7860	7927	+	68	1
atp8	7929	8111	+	183	−22
atp6	8090	8764	+	675	5
cox3	8770	9552	+	783	1
trnG(gga)	9554	9620	+	67	0

nad3	9621	9965	+	345	1
trnR(cga)	9967	10,031	+	65	0
nad4l	10,032	10,325	+	294	−4
nad4	10,322	11,689	+	1368	10
trnH(cac)	11,700	11,768	+	69	0
trnS1(agc)	11,769	11,827	+	59	0
trnL1(cta)	11,828	11,898	+	71	3
nad5	11,902	13,701	+	1800	15
nad6	13,717	14,235	-	519	0
trnE(gaa)	14,236	14,304	-	69	4
Cytb	14,309	15,442	+	1134	7
trnT(aca)	15,450	15,514	+	65	2
trnP(cca)	15,517	15,583	-	67	

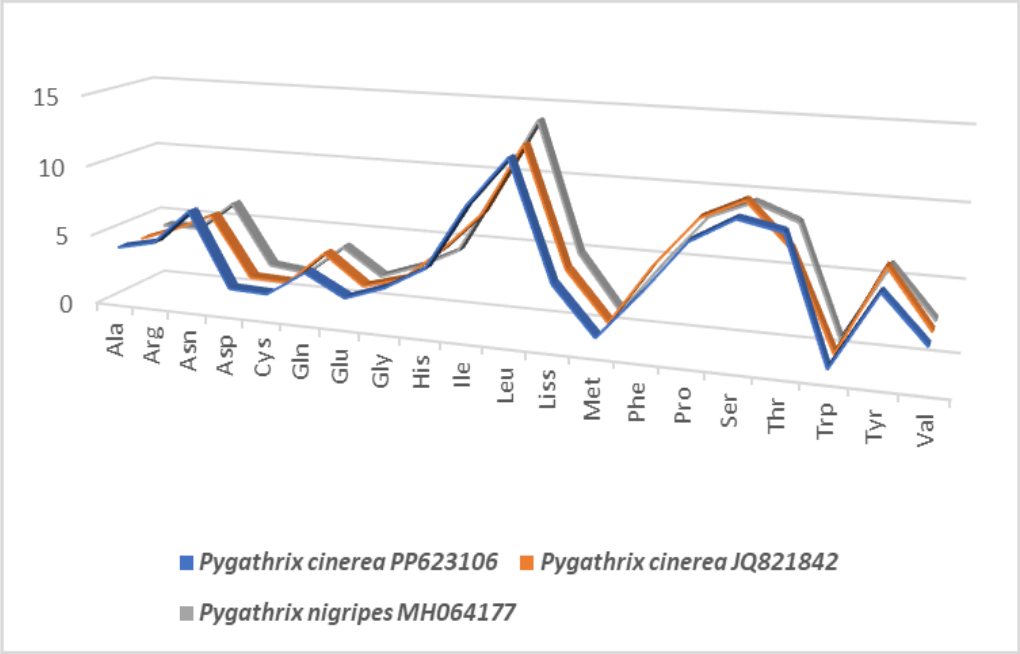
3.2.1.1. Thành phần nucleotide

Gene ty thể của *Chà vá chân xám* Tây Nguyên Việt Nam thể hiện hàm lượng A + T cao, chiếm 61,1%, tương tự như *Chà vá chân xám khác* (61,1%) nhưng khác một chút so với các mẫu *Chà vá chân nâu* (61,3%) và *Chà vá chân đen* (61,4%) (Bảng 3. 4). Tương tự như các loài linh trưởng khác, *hệ gene ty thể* của *Chà vá chân xám* thể hiện sự ưu tiên mạnh mẽ về A/T trong việc sử dụng codon (Hình 3. 7). Phân tích so sánh tần suất axit amin giữa ba cá thể *Chà vá* cho thấy phân bố rất tương đồng, như được thể hiện qua hàm lượng A + T, AT-skew và GC-skew tương đương (Hình 3. 8). Các chỉ số này, thường được sử dụng để đánh giá mô hình thành phần nucleotide trong các gene ty thể, nhấn mạnh tính bảo tồn của tần suất axit amin trong chi *chà vá*. Các axit amin phổ biến nhất ở *Chà vá chân xám* Tây Nguyên Việt Nam và các mẫu *Chà vá chân xám* khác bao gồm *Ala* (3.9–4%), *Asp* (1.8%), *Cys* (1.7%), *Asn* (6.1–7.1%), *Gln* (3.5–4.1%), *Ile* (7.9–9.1%), *Phe* (4.4–5.4%), và *Thr* (7.3–9.1%).



Hình 3. 7. Mức sử dụng codon đồng nghĩa tương đối (Relative Synonymous Codon Usage - RSCU) của các gene mã hóa protein trong ty thể *Chà vá chân xám*.

Với giá trị RSCU lớn hơn 1 cho thấy codon đó được sử dụng nhiều hơn so với kỳ vọng, trong khi giá trị nhỏ hơn 1 cho thấy mức sử dụng thấp hơn.



Hình 3. 8. Tần suất phần trăm thành phần axit amin trong toàn bộ hệ gene ty thể của các cá thể *chà vá*: *Chà vá chân xám* PP_623106, *Chà vá chân xám* JQ_821842, và *Chà vá chân đen* MH_064177.

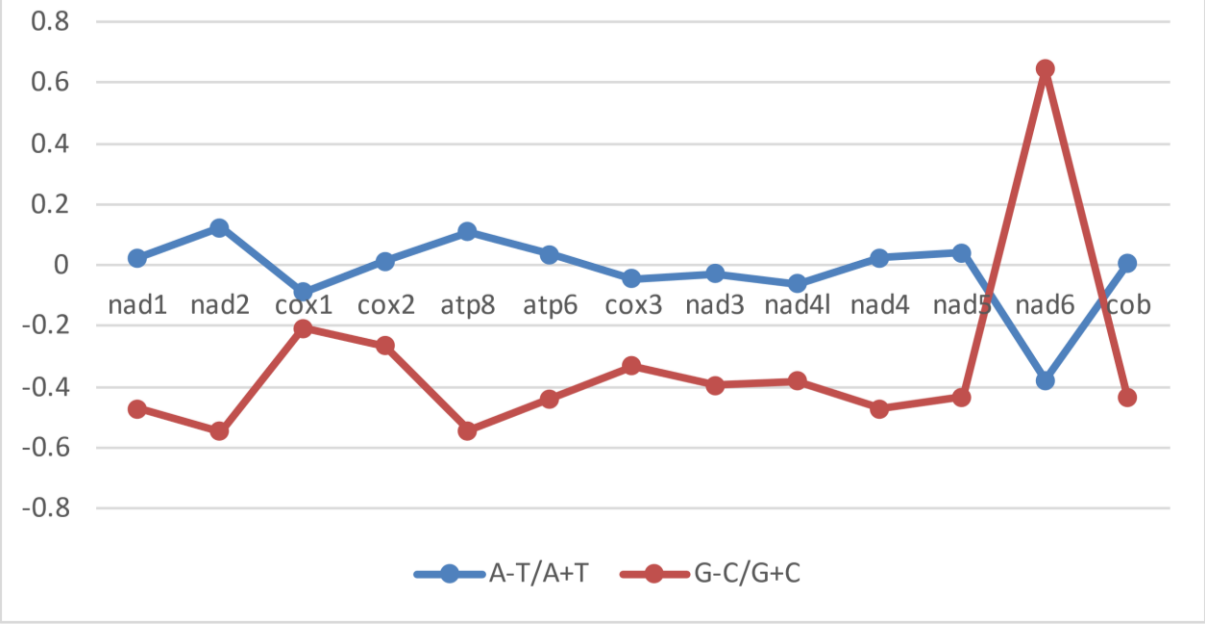
Bảng 3. 4. Các chỉ số thành phần nucleotide trong các vùng khác nhau của 22 gene ty thể đại diện cho loài linh trưởng.

	Số gia nhập			**Gene mã hóa protein (PCGs)**		**16S (rrnL)**		**12S (rrnS)**	
		Độ dài	**AT (%)**	**Độ dài (bp)**	**AT (%)**	**Độ dài (bp)**	**AT (%)**	**Độ dài (bp)**	**AT (%)**
Chà vá chân xám Tây Nguyên	PP_623106	16,541	61.1	11,292	61	1565	60	949	59.4
Chà vá chân xám	JQ_821842	16,535	61.1	11,292	61	1562	59.9	948	59.4
Chà vá chân nâu	JF_293096	16559	61.3	11,292	61	1564	59.7	948	59.4
Chà vá chân đen	MH_064177	16,534	61.4	11,304	62	1563	59.7	949	58.8
Khỉ đuôi lợn	KP_330231	17,050	57.2	11,310	57	1562	57.8	947	54.6
Voọc sọc đầu đen	KU_899140	16,548	61.9	11,277	62	1563	61	935	58.3
Khỉ vòi	DQ_355298	16,570	60.9	11,298	61	1568	59.9	949	57.5
Voọc mũi hếch vàng	KM_504390	16,552	61.5	11,292	62	722	59.6	949	58.2
Voọc mũi hếch xám	JQ_821836	16,553	61.6	11,295	62	1571	60	949	58.3
Voọc mũi hếch đen	JQ_821839	16,550	61.5	11,298	62	1570	59.7	949	58
Voọc mũi hếch Bắc Bộ	JF_293093	16,552	61.6	11,289	62	1572	60.2	949	58.5
Khỉ cộc	KM_360179	16,559	56.7	11,313	56	1566	57.3	951	55.8
Khỉ vàng	JQ_821843	16,564	56.8	11,310	56	1560	57.4	947	55.6
Khỉ đầu chó	NC_001992	16,521	56.3	11,292	56	1572	57.7	947	55.1
Khỉ tai mào trắng	KM_588314	16,499	59.7	11,280	59	1555	60.5	953	56.6
Homo sapiens neanderthalensis	NC_011137	16,565	55.6	11,283	55	1560	57.2	954	54.4
Homo sapiens neanderthalens	OM_062614	16,565	55.7	11,283	55	1560	57.3	954	54.4
Khỉ mũ trán trắng	NC_002763	16,554	60.9	11,268	61	1556	60.2	958	57.6

3.2.1.2. Các gene mã hóa Protein

Gene ty thể của *Chà vá chân xám* chứa 13 gene mã hóa protein điển hình (PCGs), tổng chiều dài lên tới 11,292 bp và chiếm 68,27% tổng thể của gene ty thể. Mười hai trong số các gene này nằm trên sợi H (mạnh nặng), trong khi gene nad6 nằm trên sợi L (mạch nhẹ). Tổng hàm lượng A + T của các PCGs là 61,4%, dao động từ 58,7% (cox3) đến 68,9% (atp8). Tỉ lệ A + T này phù hợp với thành phần nucleotide tổng thể của gene ty thể *Chà vá chân xám*.

Các chỉ số AT-skew và GC-skew dùng để đánh giá độ biến thiên thành phần nucleotide trên các PCGs; phần lớn các PCGs của *Chà vá chân xám* cho thấy AT-skew dương (Bảng 3. 5), nghĩa là tần suất xuất hiện của axit amin adenine cao hơn thymine. Mô hình này cũng được quan sát ở các loài linh trưởng khác, trong đó phần lớn các PCGs có GC-skew âm, dao động từ −0.35 đến −0.42, cho thấy thành phần nucleotide thiên về C. Tuy nhiên, nad6 lại khác biệt với xu hướng này, thể hiện GC-skew dương (0.644) và AT-skew âm (−0.378). Mô hình này tương tự như những gì được quan sát ở các mẫu *Chà vá chân xám khác* (−0.387 và 0.648) và *Chà vá chân đen* (−0.341 và 0.596) (Hình 3. 9).



Hình 3. 9. AT-skew và GC-skew trong 13 gene mã hóa protein gene ty thể *Chà vá chân xám* PP_623106.

Bảng 3. 5. Giá trị GC-skew và AT-skew trong các gene mã hóa protein (PCGs).

	Số gia nhập	T(U)	C	A	G	Tổng	AT-Skew	GC-Skew
Chà vá chân xám Tây Nguyên Việt Nam	PP_623106	30.9	26.1	30.5	12.5	868.6	−0.00605	−0.35278
Chà vá chân xám	JQ_821842	30.9	26.1	30.5	12.5	868.6	−0.0062	−0.35277
Chà vá chân nâu	JF_293096	30.9	26.1	30.5	12.5	868.6	−0.00678	−0.35155
Chà vá chân đen	MH_064177	31.2	25.9	30.6	12.3	807.4	−0.01002	−0.35463
Khỉ đuôi lợn	KP_330231	26.5	30.9	30.2	12.3	870	0.065732	−0.42945
Voọc sọc đầu đen	KU_899140	31.5	25.4	30.9	12.2	867.5	−0.00896	−0.34998
Khỉ vòi	DQ_355298	30.5	26.7	30.6	12.2	869.1	0.001593	−0.37082
Voọc mũi hếch vàng	KM_504390	31.4	25.6	30.8	12.2	868.6	−0.01054	−0.35253
Voọc mũi hếch xám	JQ_821836	31.6	25.5	30.6	12.4	868.8	−0.01638	−0.34612
Voọc mũi hếch đen	JQ_821839	31.5	25.6	30.7	12.2	869.1	−0.01181	−0.35612
Voọc mũi hếch Bắc Bộ	JF_293093	31.5	25.5	30.8	12.3	868.4	−0.01138	−0.34898
Khỉ cộc	KM_360179	26.4	31	30	12.5	870.2	0.063577	−0.4244
Khỉ vàng	JQ_821843	26.6	31	29.8	12.6	870	0.056444	−0.42052
Khỉ đầu chó	NC_001992	26.2	31.2	29.7	12.9	868.6	0.061768	−0.41663
Khỉ tai mào trắng	KM_588314	28.3	27.8	31.1	12.9	867.7	0.046296	−0.36736
Homo sapiens neanderthalensis	OM_062614	26.2	31.9	28.8	13.1	867.8	0.046422	−0.41631
Khỉ mũ trán trắng	NC_002763	29.9	26.7	31.3	12.1	866.8	0.02306	−0.37663

3.2.1.3. Vùng giàu A + T

Vùng điều khiển (*D-loop*), nằm giữa gene *trnP* và gene *trnF*, có chiều dài 1094 cặp base (vị trí 15,584–16,541 và tiếp tục từ 1–136). Kích thước này là điển hình cho các gene ty thể của động vật có xương sống (47).

Kích thước của vùng này ở *Chà vá chân xám* Tây Nguyên Việt Nam lớn hơn so với các mẫu *Chà vá chân xám* khác JQ_821842 (1092 bp; 15,444–16,535 nt) và *Chà vá chân đen* (1092 bp; 15,443–16,534 nt), *Khỉ vàng* (1085 bp; 15,480–16,564), *Khỉ đầu chó* (1076 bp; 15,446–16,521 nt), và *Khỉ tai mào trắng* (1079 bp; 15,421–16,499 nt), nhưng ngắn hơn so với Homo sapiens (1121 bp; 16,022–16,567 nt, 1–575 nt). Sự biến đổi kích thước này chủ yếu được cho là do sự khác biệt về số lượng bản sao và các lặp lại nối tiếp trong vùng (80). Vùng điều khiển thể hiện mức độ biến đổi cao so với các khu vực khác của gene ty thể, phản ánh sự hiện diện của nhiều lặp lại nối tiếp (TRs) và sự biến đổi trong số lượng bản sao của chúng (80). Tổng GC-skew trong vùng điều khiển là −0.349, AT-skew là 58%, và hàm lượng A + T là −0.013. Không có sự thiếu hụt lặp lại rõ ràng nào được quan sát trong vùng điều khiển của *Chà vá chân xám* Tây Nguyên Việt Nam.

3.2.1.4. *tRNA* và *rRNA*

Tám gene *tRNA* nằm trên sợi L, trong khi các gene còn lại nằm trên sợi H (Hình 3. 10) (Bảng 3. 3). Tất cả 22 gene *tRNA* được xác định trong hệ gene ty thể của *Chà vá chân xám* đều có anticodon tương thích với những gì tìm thấy ở các loài linh trưởng khác. Các gene *tRNA* có kích thước dao động từ 59 bp (cho *trnS1*) đến 75 bp (cho *trnL2*). Thành phần nucleotide trung bình của các gene *tRNA* là A: 32,7%, T: 30,4%, G: 19,3%, và C: 17,5%. Gene *trnT* có hàm lượng GC cao nhất (50,8%), trong khi *trnD* và *trnH* có hàm lượng thấp nhất (23,2%). 7 gene *tRNA* trong *Chà vá chân xám* chứa tổng cộng 11 cặp base không khớp nhau. Những sự không khớp này được tìm thấy trong thân chấp nhận axit amin (AA) (amino acid acceptor stem), thân pseudouridine (TΨC), và thân anticodon (AC) (Bảng 3. 6).

Gene ty thể của *Chà vá chân xám* cũng chứa hai gene ribosomal RNA (rRNA): đơn vị ribosome lớn (*16S* rRNA) và đơn vị ribosome nhỏ (12S rRNA). Gene *16S* rRNA nằm giữa các gene *trnV* và *trnL2*, trong khi gene *12S* rRNA nằm giữa các gene *trnF* và *trnV.* Sắp xếp này là điển hình cho hầu hết động vật có xương sống, trong đó gene *trnV* tách biệt các đơn vị ribosome lớn và nhỏ (81). Gene *12S* rRNA dài 949 bp, và gene *16S* rRNA dài 1565 bp. Những kích thước này tương tự như những gì tìm thấy ở các mẫu *Chà vá chân xám* khác (948 bp và 1562 bp, tương ứng). Hàm lượng A + T của gene *16S* rRNA là 60%, và hàm lượng A + T của gene *12S* rRNA là 59,4%. Các giá trị này phù hợp với những gì quan sát được ở các loài linh trưởng khác, bao gồm các mẫu *Chà vá chân xám* JQ_821842 (59,9% cho *16S* và 59,4% cho 12S) và *Chà vá chân đen* MH_064177 (60% cho *16S* và 58,8% cho 12S) (Bảng 3. 4).

Bảng 3. 6. Các vị trí bắt cặp không khớp của *t-RNAs*. AA—amino acid acceptor, T-arm, AC—anticodon.

tRNA	**Cặp base không khớp**	**thân**	**Tần số**
Phenylalanine GAA	A-C	AC	1
Valine TAC	A-C	AC	1
Methionine CAT	A-G	AA	1
	U-U	T-arm	2
	C-U	T-arm	1
	A-U	T-arm	1
Glycine TCC	A-C	AC	1
Serine GCT	A-A	AA	1
	A-A	AC	1
Threonine TGT	U-C	AA	1

3.2.1.5. Vùng lặp và vùng liên gene

Phân tích so sánh các gene ty thể tiết lộ một số trường hợp chồng chéo gene:

atp8 và *atp6:* Một sự chồng chéo 22 bp được quan sát nhất quán giữa các gene atp8 và atp6 ở *Chà vá chân xám* Tây Nguyên Việt Nam, các mẫu *Chà vá chân xám* khác và *Chà vá chân đen*.

trnV và *rrnL*: Một sự chồng chéo 2 bp được tìm thấy ở *Chà vá chân xám* Tây Nguyên Việt Nam, các cá thể *Chà vá chân xám* khác, và *Chà vá chân đen*.

trnI và *trnQ*: Một sự chồng chéo 3 bp có mặt ở *Chà vá chân xám* Tây Nguyên Việt Nam, các cá thể *Chà vá chân xám* khác, và *Chà vá chân đen*.

nad4L và *nad4*: Một sự chồng chéo 4 bp được quan sát ở *Chà vá chân xám* Tây Nguyên Việt Nam và các mẫu *Chà vá chân xám* khác.

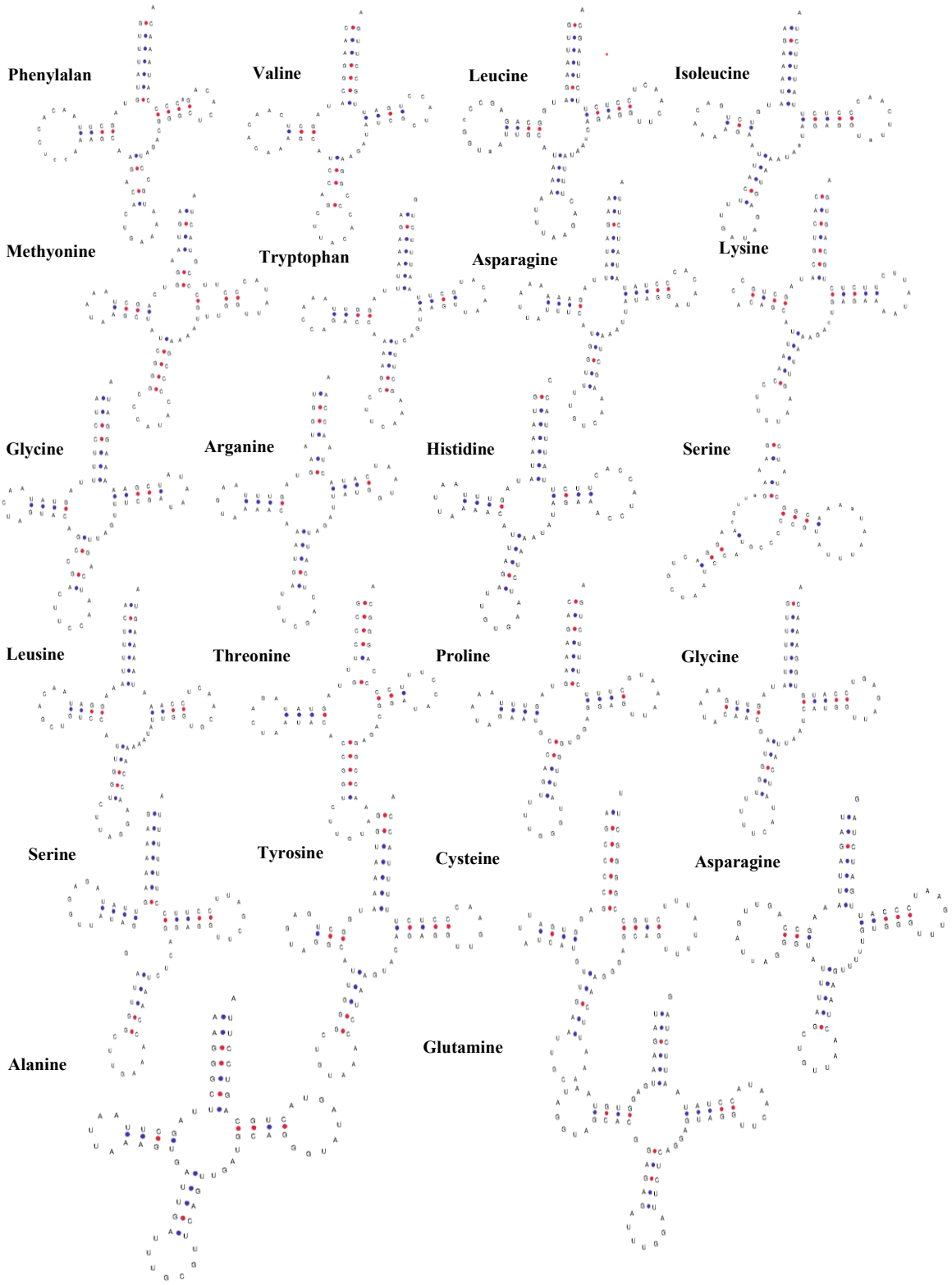
cytb-0 và *cytb-1*: Một sự chồng chéo 8 bp chỉ được tìm thấy ở *Chà vá chân đen*.

Những sự chồng chéo này, đặc biệt là sự chồng chéo 22 bp giữa atp8 và atp6, cùng với 2 bp giữa *trnV* và *rrnL*, là phổ biến trong các loài vượn và nằm trên sợi H.

Gene ty thể của *Chà vá chân xám* chứa khoảng hai mươi vùng khoảng cách liên gene không mã hóa, có kích thước dao động từ 1 bp đến 68 bp. Trong khi những vùng này trước đây được coi là không chức năng, các nghiên cứu gần đây cho thấy một số có thể đóng vai trò điều chỉnh trong việc biểu hiện gene.

Vùng khoảng cách liên gene dài nhất trong *Chà vá chân xám* dài 68 bp, nằm giữa *trnK* và *cox2*. Vùng này ngắn hơn đáng kể so với vùng 172 bp tìm thấy ở các mẫu *Chà vá chân xám* JQ_821842 vùng 167 bp trong *Chà vá chân đen* MH_064177.

Chiều dài của các vùng khoảng cách liên gene thay đổi giữa các loài vượn, với *Chà vá chân xám* có các vùng khoảng cách ngắn hơn so với *Chà vá chân nâu* (194 bp với 20 vùng) và *Khỉ đuôi lợn* (182 bp với 22 vùng), nhưng dài hơn so với *Chà vá chân đen* (167 bp với 20 vùng). Các vùng khoảng cách liên gene giữa *Chà vá chân xám* và *Chà vá chân đen* nhìn chung khá tương đồng, ngoại trừ vùng giữa *nad5* và *trnL1,* dài 15 bp ở *Chà vá chân xám* và 37 bp ở *Chà vá chân đen*.



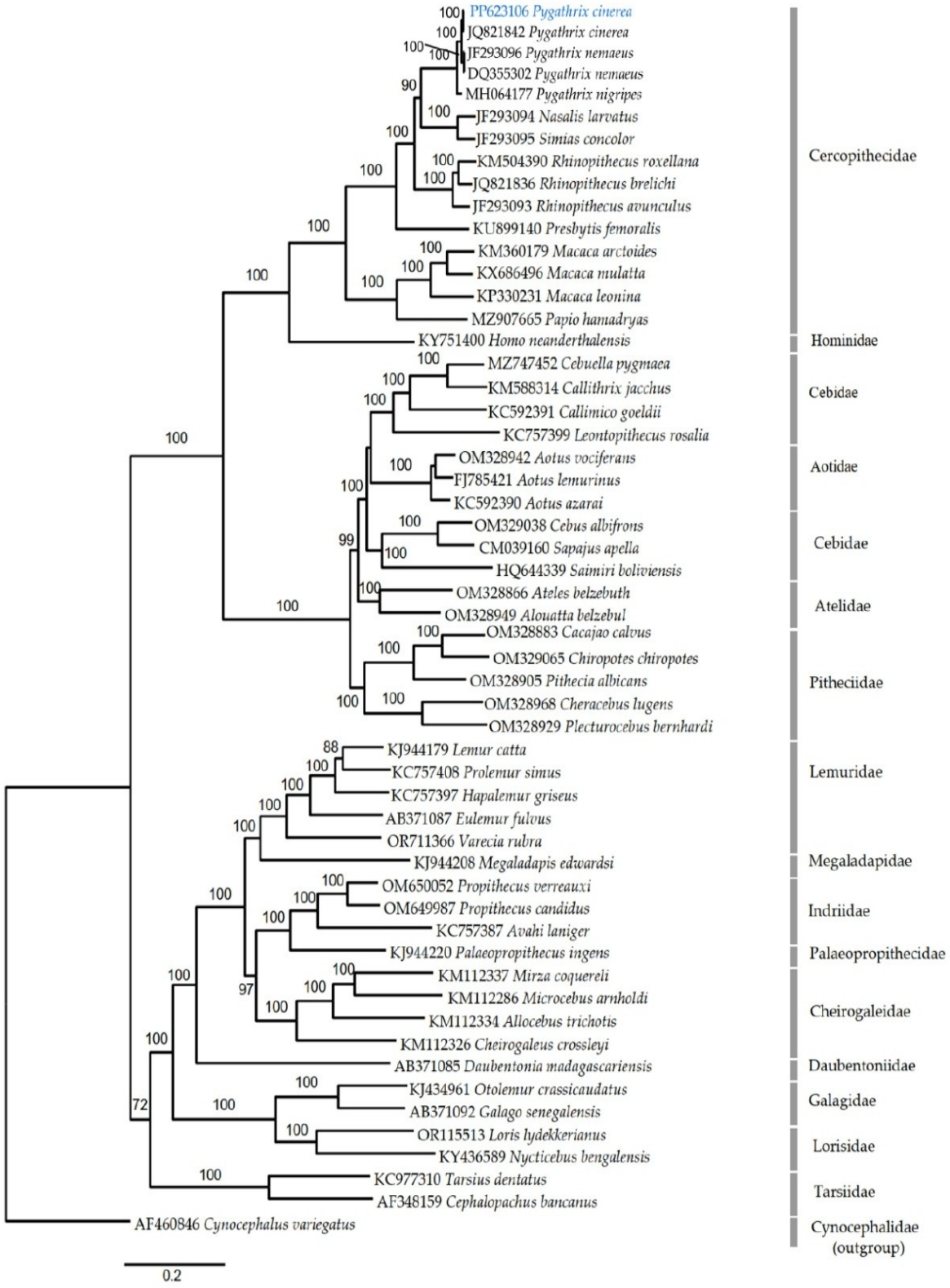
Hình 3. 10. Cấu trúc thứ cấp của 22 gene *tRNA* của *Chà vá chân xám* Tây Nguyên Việt Nam. Màu đỏ đại diện cho các kết nối GC, trong khi màu xanh dương đại diện cho các kết nối AU.

3.2.1.6. Cây phát sinh loài của *Chà vá chân xám* Tây Nguyên Việt Nam

Mô hình Tamura–Nei và phương pháp Maximum Likelihood đã được sử dụng để phân tích lịch sử tiến hóa (82). Các nhánh trong cây cho thấy tỷ lệ cây mà trong đó các taxa liên quan được nhóm lại với nhau. Cây kết quả có log likelihood cao nhất là (−433,849.22). Bằng cách áp dụng các thuật toán Neighbor-Join và BioNJ trên ma trận khoảng cách cặp tính toán bằng mô hình Tamura–Nei, các cây ban đầu cho tìm kiếm heuristic đã được tạo tự động. Topology với giá trị log likelihood tốt nhất sau đó được chọn. Phân tích này bao gồm 55 hệ gene ty thể. Các vị trí codon được đưa vào gồm có 1st, 2nd, 3rd và +Noncoding. Các phân tích tiến hóa đã được thực hiện trong MEGA11 (81).

Cấu trúc cây của các loài linh trưởng nhất quán theo phương pháp Maximum Likelihood và phân tích với hỗ trợ bootstrap (>70). Các họ hàng gần nhất còn sống của linh trưởng là Hominidae. Cây được gốc từ nhóm ngoài Cynocephalus variegatus dựa trên nghiên cứu của Mason và cs (83). Toàn bộ cây linh trưởng có nguồn gốc từ loài *Dermoptera*, cụ thể là *Cynocephalus variegatus*, và cấu trúc kết quả nhất quán với cấu trúc thu được từ các nghiên cứu trước đây (32,84–87).

Cây phát sinh loài được xây dựng bằng cách sử dụng toàn hệ gene ty thể của hơn 50 loài linh trưởng cho thấy rõ rằng các loài *Pygathrix* (*Chà vá chân xám*, *Chà vá chân nâu*, và *Chà vá chân đen*) tạo thành một nhánh đơn ngành được hỗ trợ tốt (Hình 3. 11). Trong các nghiên cứu trước, vị trí tương đối của *Pygathrix* và Nasalis gần hơn so với Rhinopithecus, và nghiên cứu hiện tại đã xác nhận điều này (35,79). Một nguồn gốc chung của nhóm này hiện nay được chấp nhận rộng rãi, mặc dù các mối quan hệ phát sinh loài giữa các chi và loài của nó vẫn phần lớn chưa được biết đến (79). Nghiên cứu này cho thấy mối quan hệ này giữa các loài linh trưởng và xác nhận rằng các loài *Chà vá* (chi *Pygathrix*) là một dòng tiến hóa riêng biệt trong nhóm khỉ mũi lạ. Nhánh *Pygathrix* có mối quan hệ gần nhất với các chi *Nasalis* (khỉ mũi) và *Simias* (khỉ *Mentawai*). Cùng nhau, ba chi này tạo thành một nhóm đơn ngành lớn hơn của các cá thể khỉ mũi lạ, điều này nhất quán với các nghiên cứu hình thái và phân tử trước đây.



Hình 3. 11. Cây phát sinh loài Maximum Likelihood của 55 hệ gene ty thể linh trưởng. PP633106 *Chà vá chân xám* màu xanh là *Chà vá chân xám* Tây Nguyên Việt Nam.

Chuỗi gene ty thể của *Chà vá chân xám* được tạo ra trong nghiên cứu này (mã số gia nhập GeneBank PP_623106) khác biệt đáng kể so với một gene ty thể *Chà vá chân xám* đã được công bố trước đó (mã số gia nhập GeneBank JQ_821842), với 43 sự khác biệt nucleotide. Điều này gợi ý sự hiện diện của các dòng gene khác nhau trong loài này. Sự phân kỳ giữa các gene ty thể cho thấy chuỗi đã công bố trước đó (JQ_821842) có thể đến từ một quần thể địa lý hoặc phân loài khác của *Chà vá chân xám*, điều này có thể giải thích cho những khác biệt gene quan sát được, hoặc khả năng tồn tại các phân loài bí mật hoặc các đơn vị tiến hóa quan trọng trong quần thể *Chà vá chân xám*.

Chuỗi gene ty thể được tạo ra trong nghiên cứu này được xác nhận chắc chắn trong nhánh *Pygathrix*, khẳng định sự phân loại loài của mẫu là *Chà vá chân xám*. Tuy nhiên, những khác biệt gene lớn so với gene ty thể *Chà vá chân xám* đã được công bố trước đó gợi ý rằng cần có thêm nhiều nghiên cứu về quần thể để tìm hiểu lịch sử tiến hóa và cấu trúc quần thể của loài linh trưởng này đang bị đe dọa nghiêm trọng.

3.2.1.7. Thảo luận

Cấu trúc và tổ chức hệ gen ty thể của cá thể *Chà vá chân xám* Tây Nguyên Việt Nam (PP_623106) tuân theo khuôn mẫu điển hình của động vật có vú, bao gồm 37 gen với trật tự sắp xếp bảo tồn cao. Kích thước hệ gen (16,541 bp) và hàm lượng A+T (61,1%) nằm trong khoảng giá trị đã được báo cáo cho chi *Pygathrix* (Bảng 3. 4), phản ánh tính ổn định ở cấp bậc chi.

Vùng điều khiển của *Chà vá chân xám* Tây Nguyên Việt Nam có kích thước lớn hơn một chút (1094 bp) so với các mẫu *Chà vá chân xám* khác (1092 bp); những đột biến này có thể liên quan đến ảnh hưởng của chọn lọc và trôi gene, hỗ trợ quan điểm về một dòng tiến hóa riêng biệt (88); tuy nhiên, sự khác biệt hai nucleotide này nằm trong vùng các chuỗi lặp lại và do đó không có ý nghĩa sinh học, vì mẫu hình này cũng được quan sát thấy ở các loài linh trưởng khác không trải qua hiện tượng thắt cổ chai cá thể hoặc phân kỳ tiến hóa nhanh chóng (89). Do đó, sự khác biệt kích thước nhỏ này không đủ để rút ra kết luận về các dòng tiến hóa khác biệt hoặc sự đa dạng ở cấp độ quần thể.

Các lặp lại liên tiếp và GC-skew được quan sát trong vùng điều khiển tương thích với các loài linh trưởng khác, nhấn mạnh tính bảo tồn của những yếu tố điều tiết này. Cấu trúc và thành phần cụ thể của những yếu tố này có thể ảnh hưởng đến việc khởi đầu sao chép hoặc điều chỉnh phiên mã (90).

Sự bảo tồn đáng kể về kích thước gene *rRNA* và hàm lượng A + T giữa các loài *Chà vá* nhấn mạnh vai trò cơ bản của chúng trong tổng hợp protein ty thể (91) làm nổi bật giá trị của những gene này như những dấu ấn đáng tin cậy cho phân tích phát sinh, đặc biệt là trong việc giải quyết các mối quan hệ giữa các taxa gần gũi (92,93). Mặc dù các gene rRNA thường được bảo tồn cao, có bốn vị trí chèn nucleotide được tìm thấy trong *rrnS* và *rrnL* có thể dẫn đến việc điều khiển biểu hiện gene ở cấp độ dịch mã, và mẫu hình sửa đổi rRNA này có thể xảy ra để đáp ứng với sự thay đổi môi trường, giai đoạn phát triển và trạng thái bệnh lý (94) hoặc chỉ ra một quỹ đạo tiến hóa độc đáo của *Chà vá chân xám* Tây Nguyên Việt Nam.

Sự hiện diện của tất cả 22 gene *tRNA* với các anticodon trong *Chà vá chân xám* là cần thiết cho việc dịch chính xác các protein ty thể. Tính bảo tồn của những gene này nhấn mạnh vai trò cơ bản của chúng trong việc duy trì chức năng ty thể. Sự biến đổi trong kích thước gene *tRNA* và hàm lượng GC, được quan sát ở *Chà vá chân xám* Tây Nguyên Việt Nam, có thể phản ánh sự khác biệt trong biểu hiện và điều chỉnh gene (95).

Việc xác định các cặp base không khớp trong một số gene *tRNA*, đặc biệt là ở các thân AA, TΨC và AC, là điều thú vị. Những sự không khớp này có thể ảnh hưởng đến cấu trúc và chức năng của *tRNA*.

Các sự chồng chéo gene được quan sát giữa một cặp gene (*atp8* và *atp6*; *trnV* và *rrnL*) trong các loài *Chà vá* gợi ý một ý nghĩa chức năng tiềm năng vượt ra ngoài những sự trùng hợp ngẫu nhiên. Những sự chồng chéo này có thể đóng vai trò điều tiết trong biểu hiện gene, có khả năng ảnh hưởng đến phiên mã hoặc dịch mã của những gene này (96).

Các vùng khoảng cách liên gene, mặc dù trước đây được coi là không chức năng, hiện đang được công nhận như những yếu tố điều khiển tiềm năng (97). Sự biến đổi trong chiều dài các khoảng cách giữa các gene khác nhau giữa các loài *Pygathrix*, với các khoảng cách dài hơn ở *Chà vá chân nâu* và *Macaca leonnina* so với *Chà vá chân xám*, gợi ý một vai trò tiềm năng trong việc điều khiển biểu hiện gene (98).

Các đặc điểm gene độc đáo được xác định trong *Chà vá chân xám* Tây Nguyên Việt Nam, đặc biệt là trong vùng điều khiển và các gene *tRNA*, nhấn mạnh tầm quan trọng của việc bảo tồn sự đa dạng gene của nó. Sự đa dạng này là rất quan trọng cho khả năng thích ứng và sức mạnh trong bối cảnh thay đổi và đe dọa môi trường. Hiểu biết về sự khác biệt gene giữa *Chà vá chân xám* Tây Nguyên và các mẫu *Chà vá chân xám* khác có thể thông báo cho các nỗ lực bảo tồn bằng cách xác định các quần thể khác biệt có thể cần các chiến lược quản lý cụ thể.

Các sai khác trình tự quan sát được giữa cá thể *Chà vá chân xám* Tây Nguyên và các mẫu công bố trước đây cho thấy khả năng tồn tại nhiều dòng dõi ty thể trong phạm vi phân bố của loài tại Việt Nam. Tuy nhiên, nghiên cứu này không nhằm đánh giá cấu trúc quần thể, mà cung cấp dữ liệu nền phục vụ so sánh phát sinh loài và định hướng cho các nghiên cứu di truyền quần thể sau này đối với một loài linh trưởng cực kỳ nguy cấp.

3.2.2. Đặc điểm hệ gene ty thể hoàn chỉnh của cá thể *Khỉ mặt đỏ* (*Macaca arctoides*) Tây Nguyên Việt Nam

Gene ty thể hoàn chỉnh của *Khỉ mặt đỏ* bao gồm 16,675 cặp base (bp) và chứa 37 gene điển hình có trong ty thể của loài linh trưởng (Hình 3. 6; Bảng 3.7), phần lớn các gene nằm trên sợi H, ngoại trừ gene ND6 và tám gene *tRNA* (*tRNA-Glu*, *tRNA-Pro*, *tRNA-Gln*, *tRNA-Ala*, *tRNA-Asn*, *tRNA-Cys*, *tRNA-Tyr, tRNA-Ser*) nằm trên sợi L. Gene ty thể của *Khỉ mặt đỏ* thể hiện các trình tự tổ chức gene tương tự như các loài thú linh trưởng khác như *Chà vá chân xám, Chà vá chân đen, Chà vá chân nâu, Khỉ vàng*. Tuy nhiên, đáng chú ý rằng chuỗi ADN ty thể của *Khỉ mặt đỏ* Tây Nguyên Việt Nam được phân tích trong nghiên cứu này có kích thước lớn hơn chút (16,675 bp) so với *Khỉ mặt đỏ* Trung Quốc (KM_360179, 16,559 bp) là 116 cặp nucleotide và nhỏ hơn *khỉ đuôi lợn Trung Quốc* (KP_330231, 17,050 bp) với 375 cặp nucleotide.

Bảng 3.7. Vị trí và độ dài của 37 gene trong hệ gene ty thể hoàn chỉnh của cá thể *Khỉ mặt đỏ* Tây Nguyên Việt Nam .

Tên	Bắt đầu	Kết thúc	Mạch	Chiều dài	Độ che phủ
*cytb*_b	7	558	+	552	3
trnT(aca)	566	629	+	64	1
trnP(cca)	631	698	-	68	1,084
trnF(ttc)	1782	1853	+	72	0
rrnS	1854	2804	+	951	0
trnV(gta)	2805	2873	+	69	-2
rrnL	2872	4438	+	1,567	−2
trnL2(tta)	4439	4513	+	75	2
nad1	4516	5466	+	951	4
trnI(atc)	5471	5539	+	69	-3

trnQ(caa)	5537	5607	-	71	0
trnM(atg)	5608	5675	+	68	0
nad2	5676	6707	+	1032	10
trnW(tga)	6718	6784	+	67	3
tRNA(gca)	6792	6860	-	69	1
trnN(aac)	6862	6934	-	73	32
trnC(tgc)	6967	7035	-	69	1
trnY(tac)	7036	7100	-	65	31
cox2	7132	7791	+	660	68
trnK(aaa)	7860	7927	+	68	1
atp8	7929	8111	+	183	−22
atp6	8090	8764	+	675	5
cox3	8770	9552	+	783	1
trnG(gga)	9554	9620	+	67	0
nad3	9621	9965	+	345	1
trnR(cga)	9967	10,031	+	65	0
nad4l	10,032	10,325	+	294	−4
nad4	10,322	11,689	+	1368	0
trnH(cac)	11,700	11,768	+	69	0
trnS1(agc)	11,769	11,827	+	59	0
trnL1(cta)	11,828	11,898	+	71	3
nad5	11,902	13,701	+	1800	15
nad6	13,717	14,235	-	519	0
trnE(gaa)	14,236	14,304	-	69	4
Cytb	14,309	15,442	+	1134	7
trnT(aca)	15,450	15,514	+	65	2
trnP(cca)	15,517	15,583	-	67	

3.2.2.1. Thành phần nucleotide

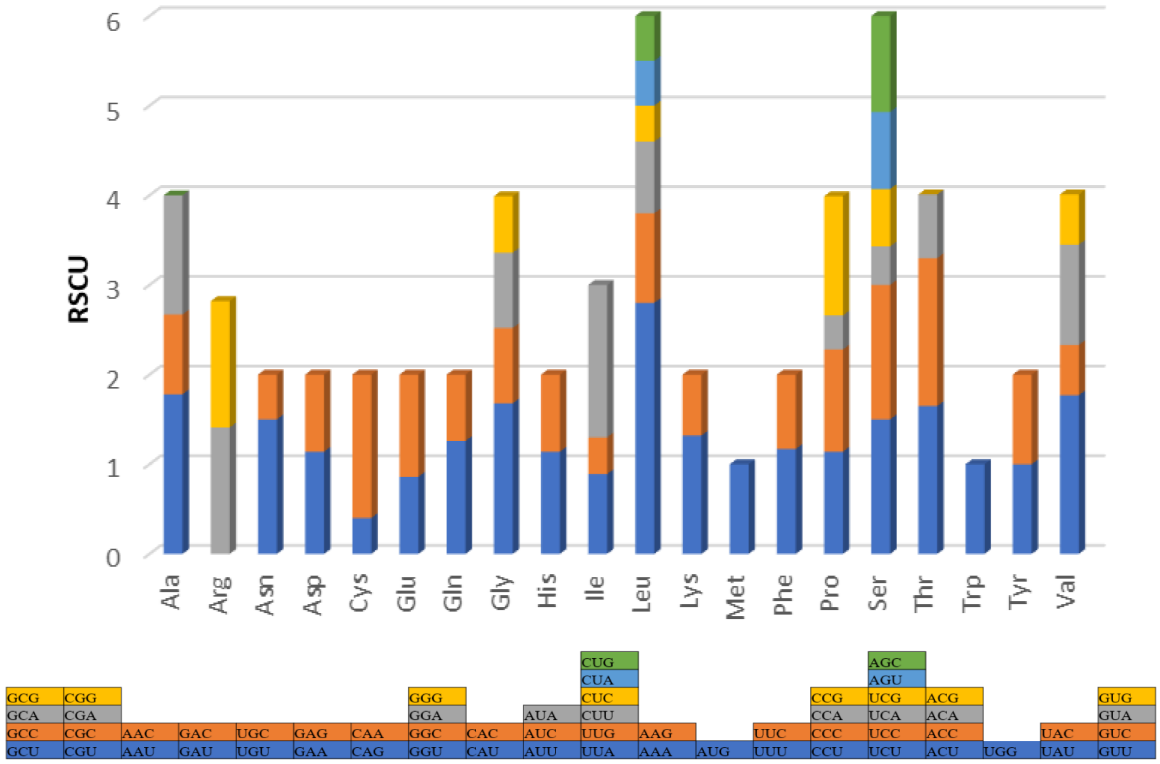
Tỷ lệ A/T của toàn bộ hệ gene ty thể dao động từ 55,6% đến 61,9%. Tỷ lệ A/T toàn bộ hệ gene ty thể của *Khỉ mặt đỏ* dao động từ 56,7–57%, thấp hơn rõ rệt so với các loài *Chà vá* (trên 61%) nhưng tương đồng với các loài *khỉ* khác như *Khỉ vàng*, *khỉ đầu chó* và *khỉ đuôi lợn (*Bảng 3. 8*).*

Phân bố tần suất axit amin trong toàn bộ hệ gene ty thể của ba cá thể *khỉ*: *Khỉ mặt đỏ* Tây Nguyên Việt Nam, *Khỉ mặt đỏ* Trung Quốc (KM_360179) và *khỉ đuôi lợn* Trung Quốc (KP_330231) nhìn chung tương đối đồng đều và có xu hướng tương đồng cao. Các axit amin chiếm tỷ lệ cao nhất ở cả ba cá thể bao gồm *Leucine (Leu), Serine (Ser), Isoleucine (Ile), Phenylalanine (Phe)* và *Valine (Val)* (Hình 3. 2). Đây là những axit amin thường gặp trong các protein mã hóa bởi hệ gene ty thể, phản ánh đặc điểm bảo tồn của các gene liên quan đến chuỗi hô hấp tế bào. So sánh giữa hai cá thể *Khỉ mặt đỏ* (Việt Nam và Trung Quốc) cho thấy sự tương đồng rất rõ rệt về tần suất của hầu hết các axit amin, với chênh lệch phần trăm không đáng kể. Trong khi đó, *khỉ đuôi lợn* Trung Quốc vẫn có sự phân bố axit amin tương tự về tổng thể, nhưng xuất hiện một số sai khác nhỏ ở một vài axit amin như *Leu, Ser* hoặc *Ala.* Các axit amin có tần suất thấp như *Cysteine (Cys), Tryptophan (Trp) và Histidine (His)* đều duy trì ở mức thấp và ổn định giữa ba cá thể, phản ánh tính bảo tồn chức năng của các protein ty thể và áp lực chọn lọc mạnh lên các vị trí mã hóa quan trọng.

Sự khác biệt về cấu trúc protein ty thể của *Khỉ mặt đỏ* giữa hai khu vực địa lý không đáng kể, phù hợp với mối quan hệ di truyền gần gũi và mức độ sai khác thấp. Đối với khỉ *đuôi lợn*, sự khác biệt về thành phần axit amin chỉ ở mức phần trăm nhỏ, phản ánh mức độ khác biệt di truyền nhất định nhưng chưa đủ lớn để làm thay đổi xu hướng chung của thành phần axit amin trong các gene ty thể.

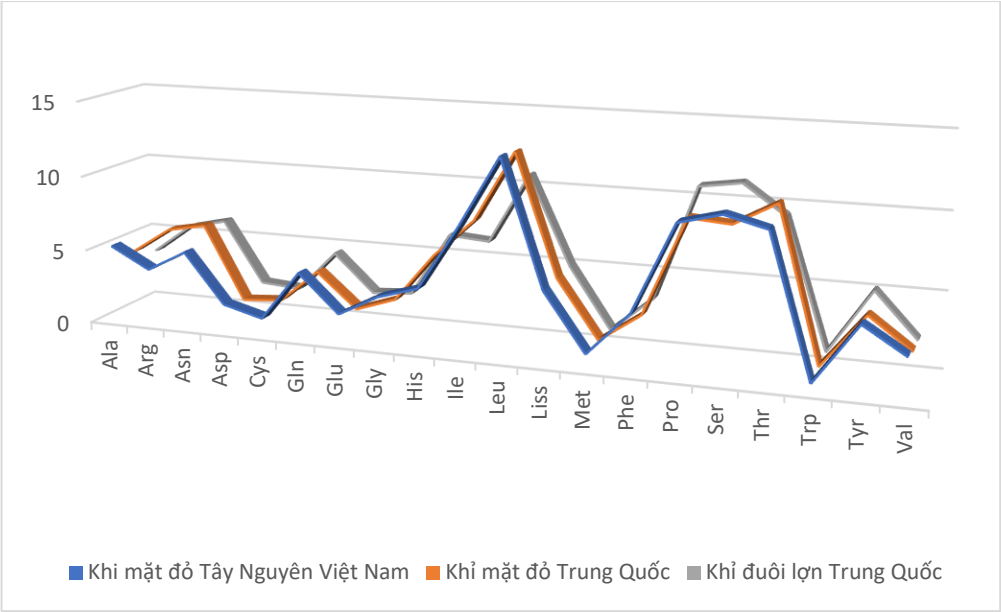
Gene ty thể của *Khỉ mặt đỏ* Tây Nguyên Việt Nam và mẫu tham chiếu (KM_360179) đều có độ dài hệ gene ty thể tương đương nhau (16.559–16.675 bp), cho thấy tính ổn định cao về mặt cấu trúc. Vùng PCGs của *Khỉ mặt đỏ* đạt chiều dài khoảng 11.301–11.313 bp, tương đồng với các loài *khỉ* khác trong chi *Macaca*. Điều này cho thấy *Khỉ mặt đỏ* mang đặc trưng thành phần nucleotide điển hình của chi *Macaca*, phản ánh mối quan hệ phát sinh chủng loại gần gũi trong nhóm này. Xét riêng từng vùng, PCGs của *Khỉ mặt đỏ* có tỷ lệ A/T khoảng 56–56,7%, trong khi *rrnL* đạt khoảng 57,3–57,6% và *rrnS* khoảng 55,6–55,8%. Sự chênh lệch không lớn giữa các vùng cho thấy áp lực chọn lọc tương đối đồng đều trên hệ gene ty thể của loài này. Đặc biệt, tỷ lệ A/T thấp hơn ở vùng *rrnS* phản ánh mức độ bảo tồn cao của gene *ribosome 12S*, phù hợp với chức năng thiết yếu trong bộ máy dịch mã ty thể. Nhìn chung, thành phần nucleotide ty thể của *Khỉ mặt đỏ* thể hiện tính ổn định cao, ít biến

động giữa các cá thể và mang đặc trưng trung gian giữa nhóm *khỉ* và nhóm *chà vá*, góp phần củng cố vị trí phân loại và mối quan hệ tiến hóa của loài trong bộ Linh trưởng.



Hình 3. 12. Mức sử dụng codon đồng nghĩa tương đối (Relative Synonymous Codon Usage - RSCU) của các gene mã hóa protein trong ty thể *Khỉ mặt đỏ*.

Phân tích chỉ số RSCU của các gen mã hóa protein trong hệ gene ty thể của *Khỉ mặt đỏ* cho thấy sự đặc điểm giàu A/T phổ biến của hệ gene ty thể ở các loài linh trưởng (). Các codon kết thúc bằng A hoặc U có giá trị RSCU lớn hơn 1 và được sử dụng ưu tiên, trong khi nhiều codon kết thúc bằng G hoặc C có RSCU thấp hoặc không xuất hiện trong hệ gene ty thể. Xu hướng này thể hiện rõ ở các amino acid có nhiều codon đồng nghĩa như leucine, serine, threonine và alanine. RSCU của *Khỉ mặt đỏ* nhìn chung tương đồng với các loài thú linh trưởng khác. Với giá trị RSCU lớn hơn 1 cho thấy codon đó được sử dụng nhiều hơn so với kỳ vọng, trong khi giá trị nhỏ hơn 1 cho thấy mức sử dụng thấp hơn.



Hình 3. 13. Tần suất phần trăm thành phần axit amin trong hệ gene ty thể của các cá thể *khỉ*: *Khỉ mặt đỏ* Tây Nguyên Việt Nam, *Khỉ mặt đỏ* Trung Quốc (KM_360179) và *khỉ đuôi lợn* Trung Quốc (KP_330231).

Phân bố tần suất axit amin trong toàn bộ hệ gene ty thể của ba cá thể *khỉ*: *Khỉ mặt đỏ* Tây Nguyên Việt Nam, *Khỉ mặt đỏ* Trung Quốc (KM_360179) và *khỉ đuôi lợn* Trung Quốc (KP_330231) nhìn chung tương đối đồng đều và có xu hướng tương đồng cao. Các axit amin chiếm tỷ lệ cao nhất ở cả ba cá thể bao gồm *Leucine (Leu), Serine (Ser), Isoleucine (Ile), Phenylalanine (Phe)* và *Valine (Val)* (Hình 3. 13). Đây là những axit amin thường gặp trong các protein mã hóa bởi hệ gene ty thể, phản ánh đặc điểm bảo tồn của các gene liên quan đến chuỗi hô hấp tế bào. So sánh giữa hai cá thể *Khỉ mặt đỏ* (Việt Nam và Trung Quốc) cho thấy sự tương đồng rất lớn về tần suất của hầu hết các axit amin, với chênh lệch phần trăm không đáng kể. Trong khi đó, *khỉ đuôi lợn* Trung Quốc vẫn có sự phân bố axit amin tương tự về tổng thể, nhưng xuất hiện một số sai khác nhỏ ở một vài axit amin như *Leu, Ser* hoặc *Ala.* Các axit amin có tần suất thấp như *Cysteine (Cys), Tryptophan (Trp) và Histidine (His)* đều duy trì ở mức thấp và ổn định giữa ba cá thể, phản ánh tính bảo tồn chức năng của các protein ty thể và áp lực chọn lọc mạnh lên các vị trí mã hóa quan trọng.

Điều này cho thấy cấu trúc protein ty thể của *Khỉ mặt đỏ* giữa hai khu vực địa lý vẫn được duy trì ổn định, phù hợp với mối quan hệ gần về mặt di truyền và sự khác biệt ở mức độ thấp, đối với phân loài *khỉ đuôi lợn* thể hiện khác biệt ở mức độ phần trăm thấp phản ánh sự sai khác di truyền ở mức thấp, nhưng chưa đủ lớn để làm thay đổi xu hướng chung của thành phần axit amin trong các gene ty thể.

Bảng 3. 8. Các chỉ số thành phần nucleotide trong các vùng khác nhau của 23 gene ty thể đại diện cho loài linh trưởng.

	Số gia nhập	**Gene mã hóa protein (PCGs)**				***16S* (*rrnL*)**		**12S (*rrnS*)**	
		Độ dài	**AT (%)**	**Độ dài (bp)**	**AT (%)**	**Độ dài (bp)**	**AT (%)**	**Độ dài (bp)**	**AT (%)**
Khỉ mặt đỏ Tây Nguyên Việt Nam		16,675	57	11301	56.7	1,567	57.6	951	55.6
Khỉ mặt đỏ Trung Quốc	KM_360179	16,559	56.7	11,313	56	1566	57.3	951	55.8
Khỉ đuôi lợn Trung Quốc	KP_330231	17,050	57.2	11,310	57	1562	57.8	947	54.6
Chà vá chân xám Tây Nguyên	PP_623106	16,541	61.1	11,292	61	1565	60	949	59.4
Chà vá chân xám	JQ_821842	16,535	61.1	11,292	61	1562	59.9	948	59.4
Chà vá chân nâu	JF_293096	16559	61.3	11,292	61	1564	59.7	948	59.4
Chà vá chân đen	MH_064177	16,534	61.4	11,304	62	1563	59.7	949	58.8
Khỉ đuôi lợn	KP_330231	17,050	57.2	11,310	57	1562	57.8	947	54.6
Voọc sọc đầu đen	KU_899140	16,548	61.9	11,277	62	1563	61	935	58.3
Khỉ vòi	DQ_355298	16,570	60.9	11,298	61	1568	59.9	949	57.5
Voọc mũi hếch vàng	KM_504390	16,552	61.5	11,292	62	722	59.6	949	58.2
Voọc mũi hếch xám	JQ_821836	16,553	61.6	11,295	62	1571	60	949	58.3
Voọc mũi hếch đen	JQ_821839	16,550	61.5	11,298	62	1570	59.7	949	58
Voọc mũi hếch Bắc Bộ	JF_293093	16,552	61.6	11,289	62	1572	60.2	949	58.5
Khỉ vàng	JQ_821843	16,564	56.8	11,310	56	1560	57.4	947	55.6
Khỉ đầu chó	NC_001992	16,521	56.3	11,292	56	1572	57.7	947	55.1
Khỉ tai mào trắng	KM_588314	16,499	59.7	11,280	59	1555	60.5	953	56.6

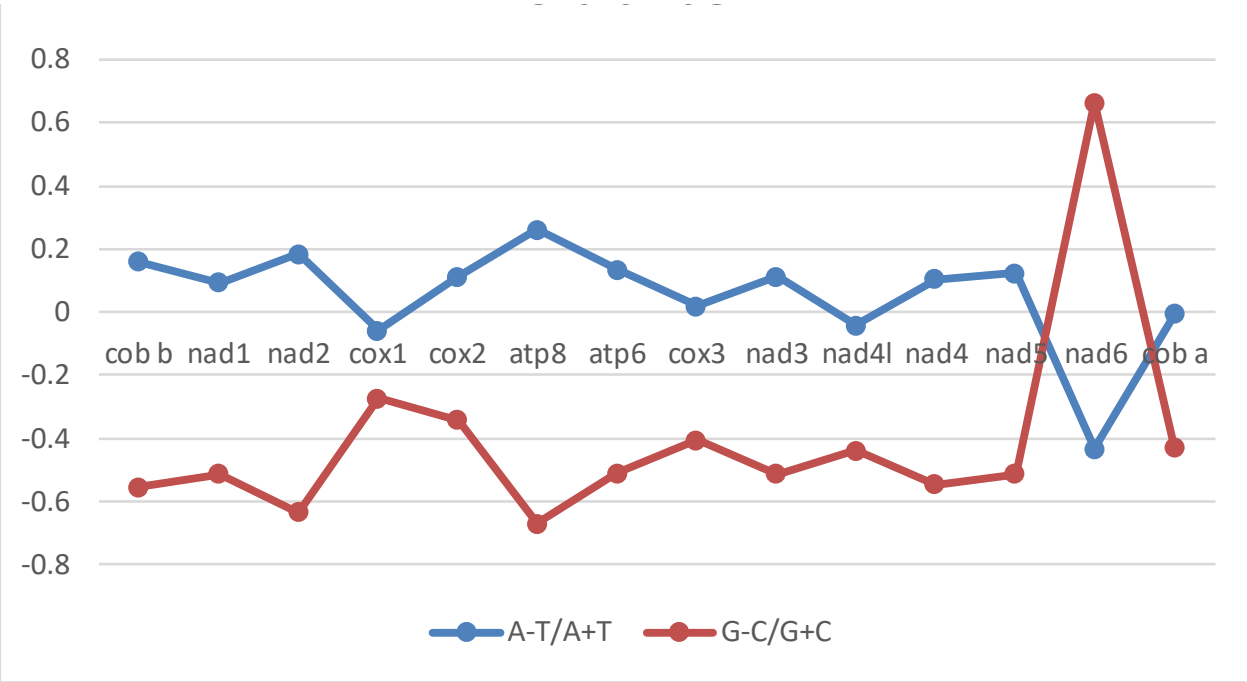
Homo sapiens neanderthalensis	NC_011137	16,565	55.6	11,283	55	1560	57.2	954	54.4
Homo sapiens neanderthalens	OM_062614	16,565	55.7	11,283	55	1560	57.3	954	54.4
Khỉ mũ trán trắng	NC_002763	16,554	60.9	11,268	61	1556	60.2	958	57.6
Cheo cheo đốm Ấn Độ	NC_037993	16,444	61.4	11,310	61	1576	61.6	958	56.4

3.2.2.2. Các gene mã hóa Protein

Gene ty thể của *Khỉ mặt đỏ* chứa 13 gene mã hóa protein điển hình (PCGs), tổng chiều dài lên tới 11,301 bp và chiếm 67,78% tổng thể của gene ty thể. Mười hai trong số các gene này nằm trên sợi H (mạnh nặng), trong khi gene *nad6* nằm trên sợi L (mạch nhẹ). Các chỉ số AT-skew và GC-skew dùng để đánh giá độ biến thiên thành phần nucleotide trên các PCGs.

Giá trị AT-skew của 13 gene mã hóa protein trong hệ gene ty thể *Khỉ mặt đỏ* nhìn chung đều mang giá trị dương (Bảng 3. 9), cho thấy A xuất hiện với tần suất cao hơn T trên mạch được phân tích. Mức độ AT-skew khác nhau giữa các gene, trong đó một số gene như *ND1, ND2, COX1* và *COX2* có AT-skew cao hơn so với các gene còn lại, phản ánh sự thiên lệch rõ rệt về thành phần adenine trong các vùng này (Hình 3. 14). Ngược lại, GC-skew của hầu hết các gene đều mang giá trị âm, cho thấy C chiếm ưu thế so với G trong các gene mã hóa protein ty thể. Mức độ GC-skew giữa các gene có sự biến động nhất định, với các gene thuộc nhóm *ND* (đặc biệt *ND4* và *ND5*) thể hiện GC-skew âm mạnh hơn, trong khi các gene *ATP6, ATP8* và *COX* có giá trị GC-skew âm thấp hơn. Sự khác biệt về biên độ AT-skew và GC-skew giữa các gene cho thấy tính không đồng nhất về thành phần nucleotide trong các gene mã hóa protein của gene ty thể *Khỉ mặt đỏ*. Tuy nhiên, xu hướng chung vẫn được duy trì ổn định, với thiên lệch A và C là đặc điểm nổi bật xuyên suốt 13 gene được phân tích.

Khỉ mặt đỏ Tây Nguyên Việt Nam có AT-skew = 0,0598 và GC-skew = −0,4212, trong khi *Khỉ mặt đỏ* Trung Quốc (KM_360179) có AT-skew = 0,0653 và GC-skew = −0,4253 (Bảng 3. 9). Như vậy, cá thể Trung Quốc có AT-skew cao hơn nhẹ (chênh lệch ~0,0055), cho thấy mức ưu thế của A so với T nhỉnh hơn so với cá thể Việt Nam. Đồng thời, GC-skew âm mạnh hơn ở cá thể Trung Quốc (−0,4253 so với −0,4212), phản ánh C chiếm ưu thế so với G rõ hơn. Tuy nhiên, cả hai cá thể đều duy trì cùng xu hướng AT-skew dương và GC-skew âm, cho thấy mức độ thiên lệch nucleotide tương đồng và ổn định giữa hai cá thể ở 2 vùng khác nhau.



Hình 3. 14. AT-skew và GC-skew trong 13 gene mã hóa protein gene ty thể *Khỉ mặt đỏ*.

Bảng 3. 9. Giá trị GC-skew và AT-skew trong các gene mã hóa protein (PCGs).

	Số gia nhập	**T(U)**	**C**	**A**	**G**	**Tổng**	**AT-Skew**	**GC-Skew**
Khỉ mặt đỏ Tây Nguyên Việt Nam		26.6	30.8	30.0	12.5	807.2	0.059797	-0.42116
Khỉ mặt đỏ Trung Quốc	KM_360179	26.4	31.0	30.0	12.5	870.2	0.065256	-0.42529
Khỉ đuôi lợn	KP_330231	26.5	30.9	30.2	12.3	870	0.065732	−0.42945
Chà vá chân xám	JQ_821842	30.9	26.1	30.5	12.5	868.6	−0.0062	−0.35277
Chà vá chân nâu	JF_293096	30.9	26.1	30.5	12.5	868.6	−0.00678	−0.35155
Chà vá chân đen	MH_064177	31.2	25.9	30.6	12.3	807.4	−0.01002	−0.35463
Voọc sọc đầu đen	KU_899140	31.5	25.4	30.9	12.2	867.5	−0.00896	−0.34998
Khỉ vòi	DQ_355298	30.5	26.7	30.6	12.2	869.1	0.001593	−0.37082
Voọc mũi hếch vàng	KM_504390	31.4	25.6	30.8	12.2	868.6	−0.01054	−0.35253
Voọc mũi hếch xám	JQ_821836	31.6	25.5	30.6	12.4	868.8	−0.01638	−0.34612
Voọc mũi hếch đen	JQ_821839	31.5	25.6	30.7	12.2	869.1	−0.01181	−0.35612

Voọc mũi hếch Bắc Bộ	JF_293093	31.5	25.5	30.8	12.3	868.4	−0.01138−0.34898
Khỉ vàng	JQ_821843	26.6	31	29.8	12.6	870	0.056444−0.42052
Khỉ đầu chó	NC_001992	26.2	31.2	29.7	12.9	868.6	0.061768−0.41663
Khỉ tai mào trắng	KM_588314	28.3	27.8	31.1	12.9	867.7	0.046296−0.36736
Homo sapiens neanderthalensis	OM_062614	26.2	31.9	28.8	13.1	867.8	0.046422−0.41631
Khỉ mũ trán trắng	NC_002763	29.9	26.7	31.3	12.1	866.8	0.02306 −0.37663

3.2.2.3. Vùng giàu A + T

Vùng điều khiển (*D-loop*) của *Khỉ mặt đỏ* nằm giữa hai gene *trnP* và *trnF*, có chiều dài 1083 cặp base, tương ứng với các vị trí 699–1781 trên hệ gene ty thể. Kích thước này nằm trong khoảng điển hình của vùng *D-loop* ở động vật có xương sống, vốn thường dao động quanh 1,0–1,2 kb (48).

So sánh với các loài linh trưởng khác cho thấy kích thước vùng *D-loop* của *Khỉ mặt đỏ* Tây Nguyên Việt Nam tương đương với cá thể *Khỉ mặt đỏ* Trung Quốc (KM_360179) (1083 bp; 16,006–16,559 và tiếp nối từ 1–529). Kích thước này cũng gần tương đồng với các loài thuộc họ Cercopithecidae và Colobinae như *Chà vá chân đen* (1092 bp; 15,443–16,534 nt), *Khỉ vàng* (1085 bp; 15,480–16,564 nt), *khỉ đầu chó* (1076 bp; 15,446–16,521 nt) và *khỉ tai mào trắng* (1079 bp; 15,421–16,499 nt). Ngược lại, vùng *D-loop* của *Khỉ mặt đỏ* có kích thưởng nhỏ hơn nhiều so với *khỉ đuôi lợn* Trung Quốc (1592 bp; 16,500–17,050 và tiếp nối từ 1–1041), trong khi ở *Homo sapiens* vùng này dài 1121 bp (16,022–16,567 nt và 1–575 nt).

Sự khác biệt về chiều dài vùng *D-loop* giữa các loài được cho là chủ yếu bắt nguồn từ sự biến đổi về số lượng và độ dài của các lặp lại nối tiếp (tandem repeats, TRs), cũng như sự khác nhau trong số lượng bản sao của các motif lặp này (80,83). Đây là đặc điểm phổ biến của vùng *D-loop* – khu vực không mã hóa nhưng đóng vai trò trung tâm trong việc

giữa khởi đầu sao chép và phiên mã của hệ gene ty thể.

Do chịu áp lực chọn lọc thấp hơn so với các gene mã hóa protein, vùng *D-loop* thường thể hiện mức độ biến đổi trình tự cao nhất trong toàn bộ hệ gene ty thể. Sự hiện diện của nhiều lặp lại nối tiếp và sự dao động về số lượng bản sao của chúng làm cho vùng này dễ phát sinh đột biến chèn/xóa (indels), từ đó dẫn đến sự khác biệt đáng kể về chiều dài giữa các loài và thậm chí giữa các quần thể trong cùng một loài.

3.2.2.4. *tRNA* và *rRNA*

Cấu trúc của 22 gene *tRNA* trong gene ty thể của *Khỉ mặt đỏ* Tây Nguyên Việt Nam có cấu trúc lá chẻ ba (cloverleaf structure) (Hình 3. 15). Trong số đó, tám gene *tRNA* (*trn^{Gln}, tRNA^{la}, tRNA^{sn}, trn^{Cys}, trn^{Tyr}, trn^{Ser}, trn^{Glu}* và *trn^{Pro}*) nằm trên sợi L, trong khi 14 gene *tRNA* còn lại nằm trên sợi H, phù hợp với mô hình tổ chức gene ty thể đã được ghi nhận ở các loài linh trưởng khác. Các gene *tRNA* của *Khỉ mặt đỏ* có chiều dài dao động từ 59 bp đến 75 bp, trong đó *trnS1* (agc) là gene ngắn nhất (59 bp) và *trnL2* (tta) là gene dài nhất (75 bp). Thành phần nucleotide của các gene *tRNA* cho thấy sự ưu thế của A và T, phản ánh đặc trưng giàu A+T của hệ gene ty thể. Mặc dù có sự khác nhau nhất định về chiều dài và hình thái các vòng các *tRNA*, cấu trúc tổng thể dạng lá clover vẫn được duy trì ở tất cả các gene, cho thấy mức độ bảo tồn cao của hệ thống *tRNA* ty thể ở loài này.

Về anticodon, tất cả 22 gene *tRNA* đều mang bộ ba anticodon đặc trưng, được xác định tại vòng đối mã, phù hợp với bảng mã di truyền ty thể của động vật có xương sống. Các anticodon được ghi nhận bao gồm UGU (Thr), GAA (Phe), UAC (Val), UAA và UAG (Leu), GAU (Ile và Met), UCA (Trp), GUC (Asp), UUU (Lys), UCC (Gly), UGG (Arg và Pro), GCU (Ser), UUC (Glu), GUA (Tyr), GCA (Cys), GUU (Asn), GUG (His), UGC (Ala) và UUG (Gln). Đáng chú ý, một số axit amin như Leucine được mã hóa bởi hai gene *tRNA* khác nhau, trong khi trnI và trnM cùng mang anticodon GAU, phản ánh đặc điểm phổ biến của hệ gene ty thể và cơ chế giải mã linh hoạt trong quá trình dịch mã. Nhìn chung, sự hiện diện đầy đủ và nhất quán của các anticodon cho thấy hệ thống *tRNA* ty thể của *Khỉ mặt đỏ* được tổ chức hoàn chỉnh, đảm bảo chức năng nhận diện codon và tham gia hiệu quả vào quá trình tổng hợp protein ty thể.

Ngoài cấu trúc lá chẻ ba điển hình, 2 gene *tRNA* trong *ty thể Khỉ mặt đỏ* chứa tổng cộng 6 cặp base không khớp (Bảng 3. 10). Cụ thể, *tRNA-Methionine* ghi nhận ba dạng cặp base không bắt cặp, bao gồm A–G, A–U và A–C, với tổng tần số là 5 lần. Trong đó, một cặp A–G xuất hiện tại thân nhận axit amin (AA stem), trong khi bốn cặp không bắt cặp còn lại (A–U với tần số 3 và A–C với tần số 1) được ghi nhận tại thân TΨC (T-arm). Điều này cho thấy vùng TΨC của trnM có mức độ biến đổi cao hơn so với thân nhận axit amin. Đối với *tRNA*-Arginine (*trnR*, anticodon TCG), chỉ ghi nhận một cặp base không bắt cặp dạng A–A, xuất hiện tại thân nhận axit amin (AA stem). So với trnM, số lượng cặp base không bắt cặp ở trnR thấp hơn đáng kể. Nhìn chung, các cặp base không bắt cặp trong các gene *tRNA* của *Khỉ mặt đỏ* chủ yếu tập trung ở thân nhận axit amin và thân TΨC, trong khi không ghi nhận sự không bắt

cặp tại thân đối mã (anticodon stem), cho thấy vùng liên quan trực tiếp đến nhận diện codon được bảo tồn cao.

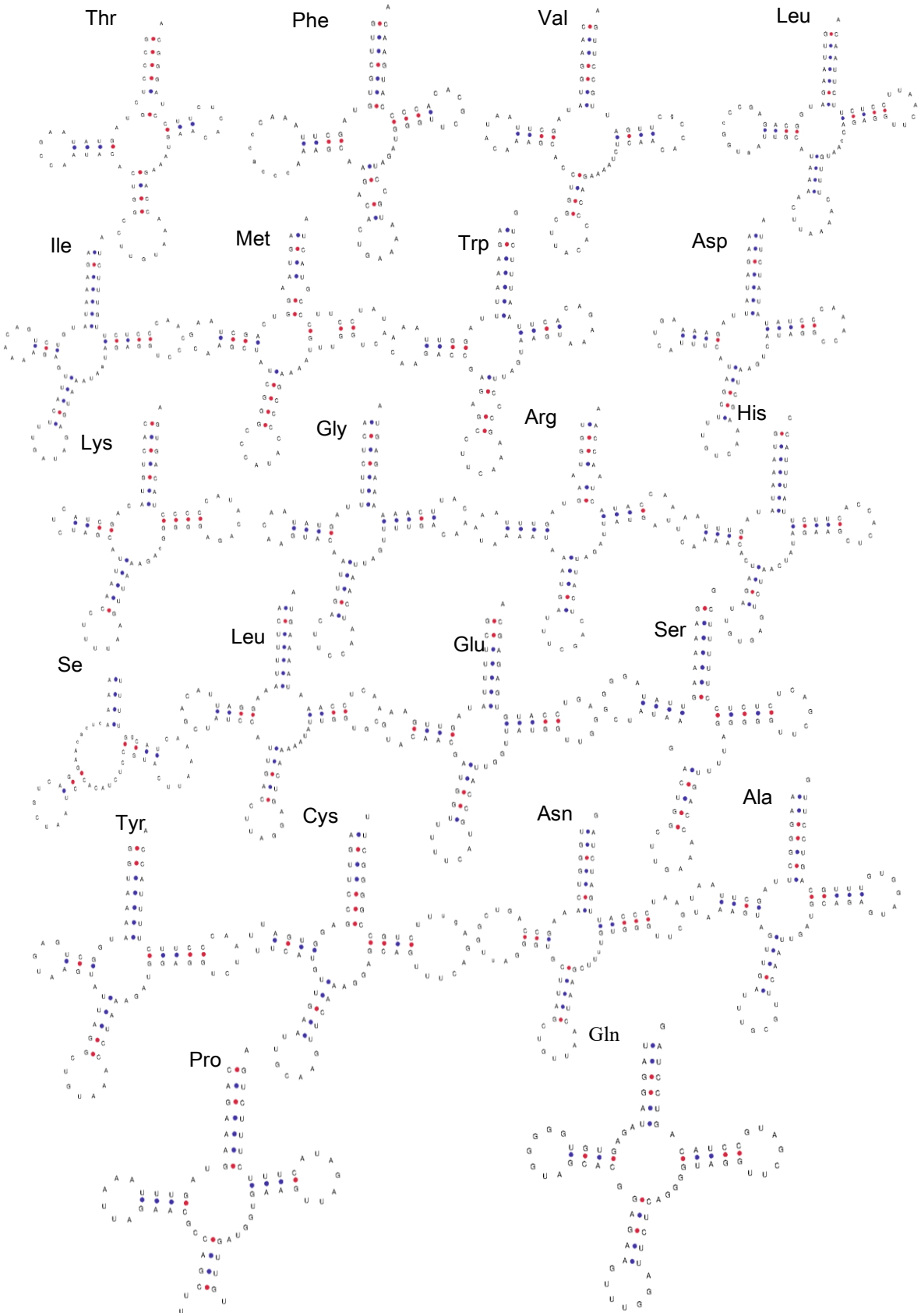
Bảng 3. 10. Các cặp base không bắt cặp của *t-RNAs*. AA—amino acid acceptor, T-arm, AC—anticodon.

tRNA	**Cặp base không khớp**	**thân**	**Tần số**
Methyonine CAT	A-G	AA	1
	A-U	T-arm	3
	A-C	T-arm	1
Arginine TCG	A-A	AA	1

3.2.2.5. Vùng lặp và vùng liên gene

Phân tích vùng liên gene và vùng lặp trên hệ gene ty thể của *Khỉ mặt đỏ* cho thấy sự hiện diện của một số trường hợp chồng chéo gene (vùng lặp), phản ánh kiểu tổ chức gene điển hình của hệ gene ty thể ở linh trưởng (Bảng 3.7). Cụ thể, một vùng chồng chéo 22 bp được ghi nhận giữa hai gene atp8 và atp6, trong đó *atp8* và *atp6* nằm chồng lên nhau. Đây là vùng liên gene phổ biến và được bảo tồn cao trong hệ gene ty thể của động vật có xương sống. Ngoài ra, sự chồng chéo 2 bp được quan sát giữa *trnV* và *rrnL*, sự chồng chéo 3 bp giữa *trnI* và *trnQ*, và sự chồng chéo 4 bp giữa *nad4L* và *nad4*. Các vùng lặp này đều có giá trị độ che phủ âm và chủ yếu nằm trên sợi H, tương tự như mô hình đã được ghi nhận ở nhiều loài linh trưởng khác.

Bên cạnh các vùng lặp, gene ty thể của *Khỉ mặt đỏ* Tây Nguyên Việt Nam còn chứa nhiều vùng khoảng cách liên gene không mã hóa, thể hiện bằng các giá trị độ che phủ dương. Các vùng liên gene này có chiều dài dao động từ 1 bp đến 68 bp, phân bố rải rác giữa các gene mã hóa protein, *tRNA* và *rRNA*. Vùng liên gene dài nhất (68 bp) được ghi nhận giữa *trnY* và *cox2*, trong khi các vùng liên gene đáng kể khác bao gồm 32 bp giữa *trnN* và *trnC*, 31 bp giữa *trnY* và gene kế cận, và 15 bp giữa *nad5* và *nad6*. Ngoài ra, nhiều vùng liên gene ngắn (1–4 bp) cũng được ghi nhận, cho thấy cấu trúc hệ gene ty thể tương đối chặt chẽ.



Hình 3. 15. Cấu trúc thứ cấp của 22 gene tRNA của *Khỉ mặt đỏ* Tây Nguyên Việt Nam. Màu đỏ đại diện cho các kết nối GC, trong khi màu xanh dương đại diện cho các kết nối AU.

3.2.2.6. Thảo luận

Trong nghiên cứu này, hệ gene ty thể hoàn chỉnh của *Khỉ mặt đỏ* Tây Nguyên Việt Nam đã được xác định và phân tích chi tiết, cung cấp thêm dữ liệu di truyền quan trọng cho loài linh trưởng này. Hệ gene ty thể của *Khỉ mặt đỏ* Tây Nguyên Việt Nam thể hiện mức độ tương đồng cao với cá thể *Khỉ mặt đỏ* Trung Quốc, phản ánh tính ổn định di truyền của loài này trong phạm vi phân bố rộng ở khu vực châu Á. Sự sai khác về kích thước tổng thể chủ yếu xuất phát từ vùng không mã hóa, đặc biệt là vùng điều khiển (*D-loop*), trong khi số lượng và vị trí 37 gene ty thể vẫn được bảo tồn. Điều này cho thấy cấu trúc hệ gene ty thể của *Khỉ mặt đỏ* có tính ổn định cao, phù hợp với đặc điểm chung của các loài trong chi *Macaca* (47,99).

Thành phần nucleotide của hệ gene ty thể *Khỉ mặt đỏ* thể hiện xu hướng giàu A+T trung bình (~56–57%), thấp hơn nhiều so với các loài thuộc họ *Colobinae* (>61%), nhưng tương đồng với các loài khỉ Cựu Thế giới khác như *Khỉ vàng, khỉ đầu chó* và *khỉ đuôi lợn*. Sự khác biệt này phản ánh áp lực chọn lọc và lịch sử tiến hóa khác nhau giữa hai nhóm Cercopithecinae và Colobinae. Việc các vùng PCGs, *rrnL* và *rrnS* có tỷ lệ A+T tương đối gần nhau cho thấy áp lực chọn lọc tương đối đồng đều trên các vùng chức năng chính của ty thể, trong đó *rrnS* có A+T thấp hơn nhẹ, phù hợp với vai trò bảo tồn cao của gene ribosomal *12S*.

Phân tích AT-skew và GC-skew của các gene mã hóa protein cho thấy xu hướng AT-skew dương và GC-skew âm ổn định, phản ánh sự ưu thế của A so với T và C so với G trên mạch được phân tích. Mô hình thiên lệch này tương đồng giữa cá thể *Khỉ mặt đỏ* Việt Nam và Trung Quốc, cho thấy tính bảo tồn về cơ chế sao chép và phiên mã ty thể giữa các khu vực địa lý. Sự khác biệt nhỏ về biên độ skew giữa các gene nhiều khả năng liên quan đến vị trí gene trên mạch và áp lực chức năng khác nhau trong chuỗi hô hấp ty thể.

Kết quả phân tích RSCU và tần suất axit amin cho thấy hệ gene ty thể của *Khỉ mặt đỏ* có mức độ sử dụng codon ưu tiên kết thúc bằng A hoặc U, phù hợp với nền nucleotide giàu A+T. Sự tương đồng rất cao về thành phần axit amin giữa hai cá thể *Khỉ mặt đỏ* (Việt Nam và Trung Quốc) cho thấy mức độ phân hóa di truyền thấp, trong khi các sai khác nhỏ so với khỉ đuôi lợn cho thấy sự phân hóa loài nhưng không làm thay đổi xu hướng chức năng tổng thể của protein ty thể (100).

Vùng điều khiển (*D-loop*) của *Khỉ mặt đỏ* có chiều dài 1083 bp, nằm trong khoảng điển hình của động vật có xương sống và tương đồng với nhiều loài linh trưởng khác (101). Sự khác biệt lớn về chiều dài *D-loop* giữa *Khỉ mặt đỏ* và *khỉ đuôi*

lợn Trung Quốc cho thấy vai trò của các lặp lại nối tiếp (tandem repeats) trong việc tạo nên biến thiên kích thước vùng này. Do chịu áp lực chọn lọc thấp, *D-loop* trở thành vùng biến đổi cao, có giá trị lớn trong các nghiên cứu về đa dạng di truyền, phân hóa quần thể và lịch sử tiến hóa của loài.

Những sai khác nhỏ quan sát được giữa *Khỉ mặt đỏ* và các loài *macaca* khác có ý nghĩa chủ yếu ở cấp độ phân loại loài, không làm thay đổi xu hướng chức năng chung của protein ty thể. Điều này củng cố giá trị của dữ liệu ADN ty thể trong việc làm rõ mối quan hệ phát sinh loài trong chi *macaca*.

3.2.3. Đặc điểm hệ gene ty thể hoàn chỉnh của cá thể *Chà vá chân nâu* (*Pygathrix nemaeus*) Tây Nguyên Việt Nam

Gene ty thể hoàn chỉnh của *Chà vá chân nâu* bao gồm 16,539 cặp base (bp) và chứa 37 gene điển hình có trong ty thể của loài linh trưởng (Hình 3. 5; Bảng 3. 11). ADN ty thể của *Khỉ mặt đỏ* Tây Nguyên Việt Nam (16,539 bp) được phân tích trong nghiên cứu này có kích thước nhỏ hơn *Chà vá chân nâu (*JF_293096) 16,559 bp là 20 bp và lớn hơn so với và *Chà vá chân nâu* (NC008220) 15,467bp là 1,072 cặp nucleotide. Tuy nhiên kích thước của *Chà vá chân nâu* (NC008220) được công bố không có số liệu của vùng *D-loop*, do đó sự chênh lệch về kích thước này phù hợp với độ dài điển hình của khoảng 1092 bp của vùng này.

Bảng 3. 11. Vị trí và độ dài của 37 gene trong hệ gene ty thể hoàn chỉnh của cá thể *Chà vá chân nâu* Tây Nguyên Việt Nam

Tên	Bắt đầu	Kết thúc	Mạch	Chiều dài	Độ che phủ
*cytb*_b	7	558	+	552	7
trnT(aca)	566	629	+	64	1
trnP(cca)	631	698	-	68	1,083
trnF(ttc)	1782	1853	+	72	0
rrnS	1854	2804	+	951	0
trnV(gta)	2805	2873	+	69	-2
rrnL	2872	4438	+	1567	−2
trnL2(tta)	4439	4513	+	75	2
nad1	4516	5466	+	951	4

trnI(atc)	5471	5539	+	69	-3
trnQ(caa)	5537	5607	-	71	0
trnM(atg)	5608	5675	+	68	0
nad2	5676	6707	+	1032	10
trnW(tga)	6718	6784	+	67	3
tRNA(gca)	6792	6860	-	69	1
trnN(aac)	6862	6934	-	73	32
trnC(tgc)	6967	7035	-	69	0
trnY(tac)	7036	7100	-	65	4
cox1	7105	8643	+	1539	2
trnS2(tca)	8646	8714	-	69	3
trnD(gac)	8718	8785	+	68	1
cox2	8787	9446	+	660	86
trnK(aaa)	9533	9599	+	67	1
atp8	9601	9801	+	201	-40
atp6	9762	10436	+	675	5
cox3	10442	11224	+	783	1
trnG(gga)	11226	11293	+	68	0
nad3	11294	11638	+	345	1
trnR(cga)	11640	11704	+	65	0
nad4l	11705	11998	+	294	-4
nad4	11995	13365	+	1371	7
trnH(cac)	13373	13441	+	69	0
trnS1(agc)	13442	13500	+	59	0
trnL1(cta)	13501	13571	+	71	6
nad5	13578	15377	+	1800	2
nad6	15390	15911	-	522	0

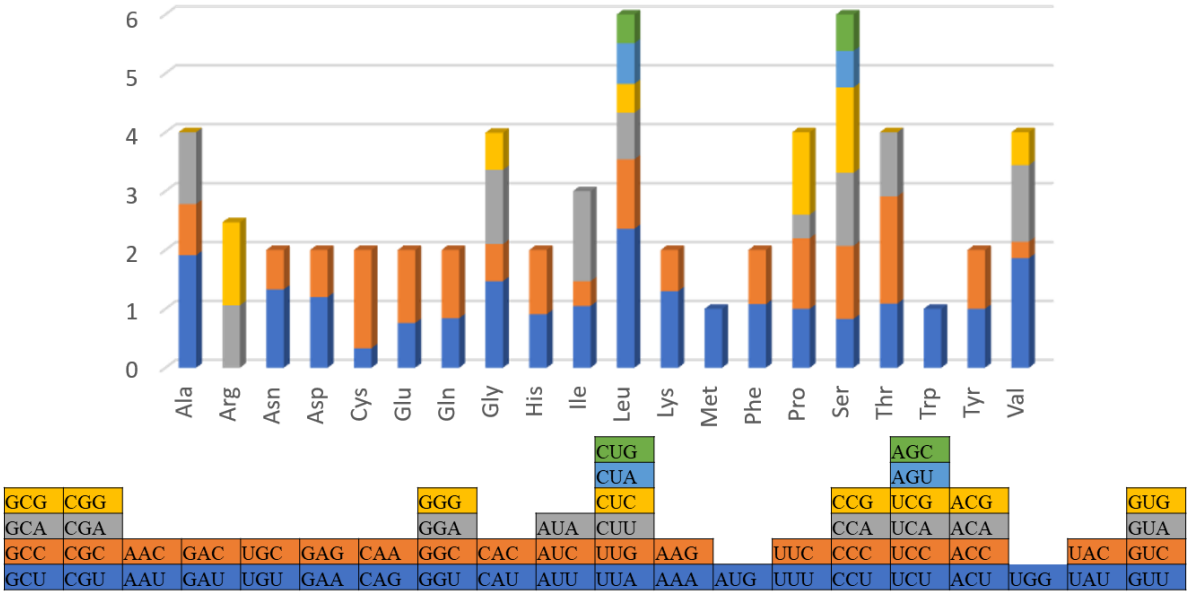
trnE(gaa)	15912	15980	-	69	4
*cytb*_a	15985	16560	+	576	

3.2.3.1. Thành phần nucleotide

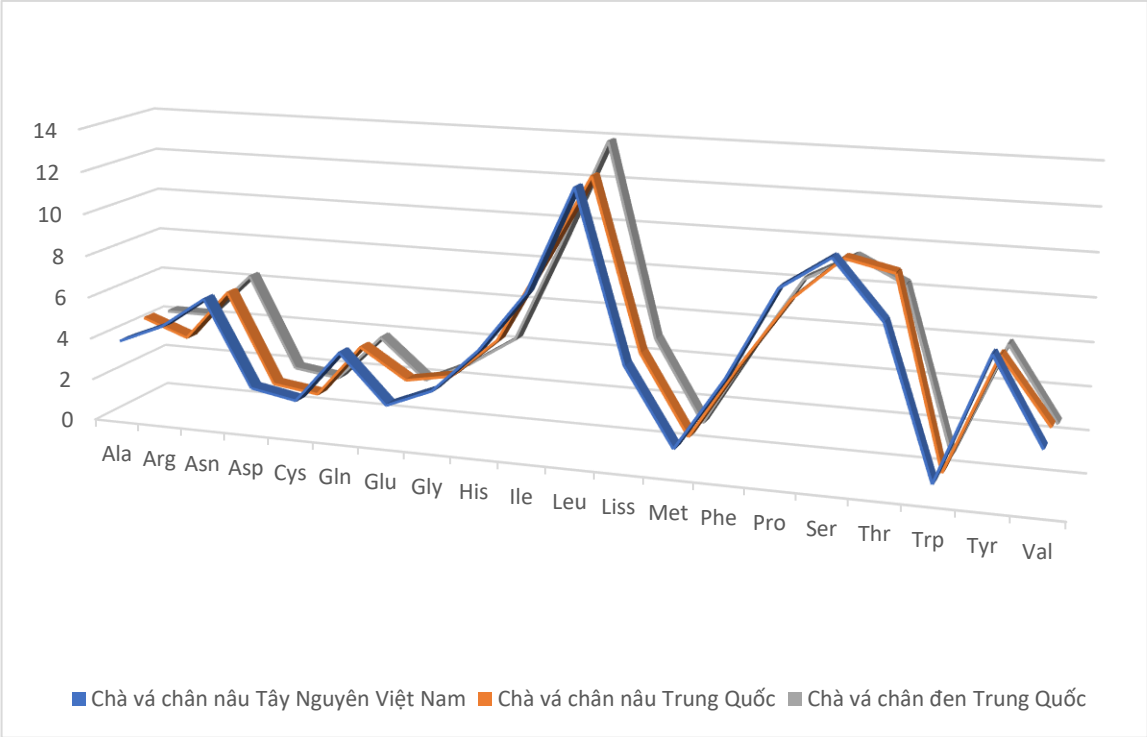
Gene ty thể của *Chà vá chân nâu* Tây Nguyên Việt Nam thể hiện hàm lượng A + T 61.1% nhỏ hơn cá thể *Chà vá chân nâu* ở khu vực khác NC008220 và *Chà vá chân nâu* JF_293096 cùng có hàm lượng A + T chiếm 61.3% (Bảng 3. 12). Tương tự như các loài linh trưởng khác, hàm lượng ADN ty thể của *Chà vá chân nâu* thể hiện sự thiên lệch sử dụng codon rõ rệt. Phần lớn các codon kết thúc bằng A hoặc U có giá trị RSCU > 1, trong khi nhiều codon kết thúc bằng G hoặc C có RSCU < 1 hoặc không xuất hiện trong hệ gene ty thể. Xu hướng này được ghi nhận rõ ở các amino acid có nhiều codon đồng nghĩa như leucine, serine, alanine và proline, phản ánh ưu thế của các codon giàu A/U trong hệ gene ty thể của loài (Bảng 3. 12).

Sự phân bố tần suất (%) của 20 axit amin trong toàn bộ các gene mã hóa protein ty thể của ba cá thể linh trưởng gồm *Chà vá chân nâu* Tây Nguyên Việt Nam, *Chà vá chân nâu* Trung Quốc) và *Chà vá chân đen* Trung Quốc rất tương đồng (Hình 3. 17). Trong số đó, *Leu* là axit amin chiếm tần suất cao nhất ở cả ba cá thể, dao động khoảng 12–14%. Tiếp theo là các axit amin có tần suất tương đối cao như *Ser* và *Thr*, đạt khoảng 7–9%. Các axit amin khác như *Ala*, *Ile*, *Pro* và *Val* có tần suất trung bình, phân bố khá đồng đều giữa các cá thể. Ngược lại, *Trp* và *Met* là những axit amin có tần suất thấp nhất, chỉ chiếm khoảng dưới 1–2%. Nhìn chung, sự khác biệt về tần suất axit amin giữa ba cá thể là rất nhỏ, không ghi nhận sự thay đổi rõ rệt về thứ hạng hay xu hướng phân bố của các axit amin.

RSCU của các codon trong 13 gene mã hóa protein ty thể của cá thể *Chà vá chân nâu* (Hình 3. 16), các codon được sử dụng ưu tiên chủ yếu là những codon giàu A/T, đặc biệt là các codon kết thúc bằng A hoặc T. Ngược lại, phần lớn các codon giàu G/C, nhất là các codon kết thúc bằng G hoặc C, có giá trị RSCU thấp. Xu hướng này được ghi nhận đồng nhất ở hầu hết các axit amin có nhiều codon đồng nghĩa như *Leu*, *Ser*, *Ala*, *Thr* và *Pro*. Đối với các axit amin chỉ được mã hóa bởi một codon duy nhất trong hệ gene ty thể như *Met* (AUG) và *Trp* (UGG), giá trị RSCU bằng 1, phản ánh sự không tồn tại của thiên lệch sử dụng codon đồng nghĩa đối với các axit amin này.



Hình 3. 16. Mức sử dụng codon đồng nghĩa tương đối (Relative Synonymous Codon Usage - RSCU) của các gene mã hóa protein trong ty thể *Chà vá chân nâu*.



Hình 3. 17. Tần suất phần trăm thành phần axit amin trong toàn bộ hệ gene ty thể của các cá thể *chà vá*: *Chà vá chân nâu* Tây Nguyên Việt Nam, *Chà vá chân nâu* Trung Quốc, và *Chà vá chân đen* Trung Quốc

Bảng 3. 12. Các chỉ số thành phần nucleotide trong các vùng khác nhau của 23 gene ty thể đại diện cho loài linh trưởng.

	Số gia nhập	Gene mã hóa protein (PCGs)				*16S* (*rrnL*)		12S (*rrnS*)	
		Độ dài	AT (%)	Độ dài (bp)	AT (%)	Độ dài (bp)	AT (%)	Độ dài (bp)	AT (%)
Chà vá chân nâu Tây Nguyên		16,539	61.1	11292	61.5	1564	59.5	948	59.4
Chà vá chân nâu	NC008220	15,467	61.3	11292	61.4	1564	59.7	948	59.4
Chà vá chân nâu	JF_293096	16559	61.3	11,292	61	1564	59.7	948	59.4
Khỉ mặt đỏ	KM_360179	16,559	56.7	11,313	56	1566	57.3	951	55.8
Khỉ đuôi lợn	KP_330231	17,050	57.2	11,310	57	1562	57.8	947	54.6
Chà vá chân xám	JQ_821842	16,535	61.1	11,292	61	1562	59.9	948	59.4
Chà vá chân đen	MH_064177	16,536	61.4	11,304	62	1563	59.7	949	58.8
Khỉ đuôi lợn	KP_330231	17,050	57.2	11,310	57	1562	57.8	947	54.6
Voọc sọc đầu đen	KU_899140	16,548	61.9	11,277	62	1563	61	935	58.3
Khỉ vòi	DQ_355298	16,570	60.9	11,298	61	1568	59.9	949	57.5
Voọc mũi hếch vàng	KM_504390	16,552	61.5	11,292	62	722	59.6	949	58.2
Voọc mũi hếch xám	JQ_821836	16,553	61.6	11,295	62	1571	60	949	58.3
Voọc mũi hếch đen	JQ_821839	16,550	61.5	11,298	62	1570	59.7	949	58
Voọc mũi hếch Bắc Bộ	JF_293093	16,552	61.6	11,289	62	1572	60.2	949	58.5
Khỉ mặt đỏ	KM_360179	16,559	56.7	11,313	56	1566	57.3	951	55.8
Khỉ vàng	JQ_821843	16,564	56.8	11,310	56	1560	57.4	947	55.6
Khỉ đầu chó	NC_001992	16,521	56.3	11,292	56	1572	57.7	947	55.1
Khỉ tai mào trắng	KM_588314	16,499	59.7	11,280	59	1555	60.5	953	56.6
Homo sapiens neanderthalensis	NC_011137	16,565	55.6	11,283	55	1560	57.2	954	54.4
Homo sapiens neanderthalens	OM_062614	16,565	55.7	11,283	55	1560	57.3	954	54.4
Khỉ mũ trán trắng	NC_002763	16,554	60.9	11,268	61	1556	60.2	958	57.6

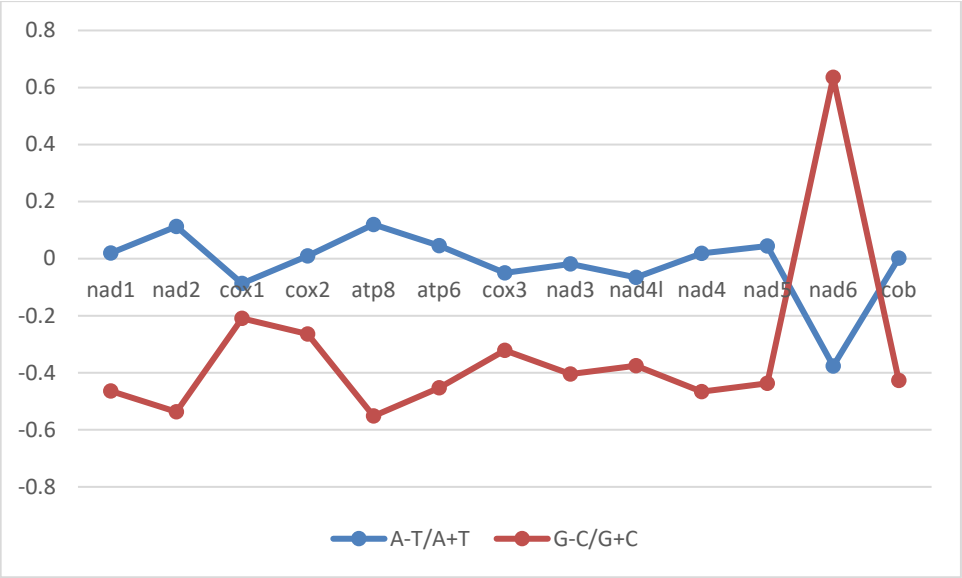
Cheo cheo đốm Ấn Độ	NC_037993	16,444	61.4	11,310	61	1576	61.6	958	56.4

3.2.3.2. Các gene mã hóa Protein

Gene ty thể của *Chà vá chân nâu* Tây Nguyên Việt Nam chứa 13 gene mã hóa protein điển hình (PCGs), với tổng chiều dài khoảng 11,3 kb, chiếm phần lớn chiều dài hệ gene ty thể hoàn chỉnh. Trong số này, 12 gene nằm trên mạch nặng (H-strand), trong khi gene *nad6* nằm trên mạch nhẹ (L-strand), phù hợp với mô hình tổ chức gene ty thể đặc trưng của linh trưởng. Các chỉ số AT-skew và GC-skew cho thấy giá trị AT-skew của 13 gene mã hóa protein trong hệ gene ty thể *Chà vá chân nâu* nhìn chung đều mang giá trị âm, phản ánh T xuất hiện với tần suất cao hơn A trên mạch được phân tích (3. 13). Mức độ AT-skew có sự khác biệt giữa các gene, trong đó một số gene như *nad1, nad2, cox1* và *cox2* thể hiện AT-skew âm rõ rệt hơn so với các gene còn lại, cho thấy mức độ thiên lệch mạnh về thành phần thymine trong các vùng này (Hình 3. 18). Ngược lại, GC-skew của hầu hết các gene mã hóa protein đều mang giá trị âm, cho thấy C chiếm ưu thế so với G trong các gene ty thể của *Chà vá* chân nâu. Mức độ GC-skew giữa các gene có sự biến động nhất định, trong đó các gene thuộc nhóm NAD, đặc biệt là *nad4* và *nad5*, có GC-skew âm mạnh hơn, trong khi các gene *atp6, atp8* và các gene *cox* thường có giá trị GC-skew âm thấp hơn. Sự khác biệt về biên độ AT-skew và GC-skew giữa các gene phản ánh tính không đồng nhất về thành phần nucleotide trong các gene mã hóa protein của hệ gene ty thể *Chà vá* chân nâu.

Tuy nhiên, xu hướng chung vẫn được duy trì ổn định, với AT-skew âm và GC-skew âm là đặc điểm nổi bật xuyên suốt 13 gene mã hóa protein được phân tích, cho thấy sự bất đối xứng nucleotide nhất quán trong hệ gene ty thể của loài này.

Ở cấp độ toàn bộ các gene mã hóa protein, cá thể *Chà vá chân nâu* Tây Nguyên Việt Nam có AT-skew âm và GC-skew âm, phản ánh ưu thế của T so với A và của C so với G. So sánh với các cá thể *Chà vá chân nâu* khác đã công bố cho thấy mức độ skew giữa các cá thể chỉ dao động nhẹ, không làm thay đổi xu hướng tổng thể. Do đó, độ thiên lệch nucleotide trong các gene mã hóa protein ty thể của *Chà vá chân nâu* là tương đối ổn định, bất chấp sự khác biệt về nguồn gốc địa lý.



Hình 3. 18. AT-skew và GC-skew trong 13 gene mã hóa protein gene ty thể *Chà vá chân nâu*.

Bảng 3. 13. Giá trị GC-skew và AT-skew trong các gene mã hóa protein (PCGs).

	Số gia nhập	**T(U)**	**C**	**A**	**G**	**Tổng**	**AT-Skew**	**GC-Skew**
Chà vá chân nâu Tây Nguyên Việt Nam		31.0	26	30.5	12.5	868.6	-0.00749	-0.35205
Chà vá chân nâu	JF_293096	30.9	26.1	30.5	12.5	868.6	−0.00678	−0.35155
Chà vá chân nâu	NC008220	30.9	26.1	30.5	12.5	868.6	−0.00678	−0.35155
Chà vá chân đen	MH_064177	31.2	25.9	30.6	12.3	807.4	−0.01002	−0.35463
Chà vá chân xám	PP_623106	30.9	26.1	30.5	12.5	868.6	−0.00605	−0.35278
Chà vá chân xám	JQ_821842	30.9	26.1	30.5	12.5	868.6	−0.0062	−0.35277
Khỉ đuôi lợn	KP_330231	26.5	30.9	30.2	12.3	870	0.065732	−0.42945
Voọc sọc đầu đen	KU_899140	31.5	25.4	30.9	12.2	867.5	−0.00896	−0.34998
Khỉ vòi	DQ_355298	30.5	26.7	30.6	12.2	869.1	0.001593	−0.37082
Voọc mũi hếch vàng	KM_504390	31.4	25.6	30.8	12.2	868.6	−0.01054	−0.35253
Voọc mũi hếch xám	JQ_821836	31.6	25.5	30.6	12.4	868.8	−0.01638	−0.34612
Voọc mũi hếch đen	JQ_821839	31.5	25.6	30.7	12.2	869.1	−0.01181	−0.35612
Voọc mũi hếch Bắc Bộ	JF_293093	31.5	25.5	30.8	12.3	868.4	−0.01138	−0.34898

Khỉ cộc	KM_360179	26.4	31	30	12.5	870.2	0.063577	−0.4244
Khỉ vàng	JQ_821843	26.6	31	29.8	12.6	870	0.056444	−0.42052
Khỉ đầu chó	NC_001992	26.2	31.2	29.7	12.9	868.6	0.061768	−0.41663
Khỉ tai mào trắng	KM_588314	28.3	27.8	31.1	12.9	867.7	0.046296	−0.36736
Homo sapiens neanderthalensis	OM_062614	26.2	31.9	28.8	13.1	867.8	0.046422	−0.41631
Khỉ mũ trán trắng	NC_002763	29.9	26.7	31.3	12.1	866.8	0.02306	−0.37663

3.2.3.3. Vùng giàu A + T

Vùng điều khiển (*D-loop*), nằm giữa gene *trnP* và gene *trnF*, có chiều dài 1083 cặp base (vị trí 699–1,871). Kích thước này là điển hình cho các gene ty thể của động vật có xương sống (47). Kích thước của vùng này ở *Chà vá chân nâu* Tây Nguyên Việt Nam tương tự so với các mẫu *Chà vá chân nâu* khác JF_293096 (1092 bp; vị trí 15468-16559) cũng như *Chà vá chân đen* (1092 bp; 15,443–16,534 nt), *Khỉ vàng* (1085 bp; 15,480–16,564), *Khỉ đầu chó* (1076 bp; 15,446–16,521 nt), và *Khỉ tai mào trắng* (1079 bp; 15,421–16,499 nt), và ngắn hơn so với Homo sapiens (1121 bp; 16,022–16,567 nt, 1–575 nt). Sự biến đổi kích thước này chủ yếu được cho là do sự khác biệt về số lượng bản sao và các lặp lại nối tiếp trong vùng (80). Vùng điều khiển thể hiện mức độ biến đổi cao so với các khu vực khác của gene ty thể, phản ánh sự hiện diện của nhiều lặp lại nối tiếp (TRs) và sự biến đổi trong số lượng bản sao của chúng (80).

3.2.3.4. *tRNA* và rRNA

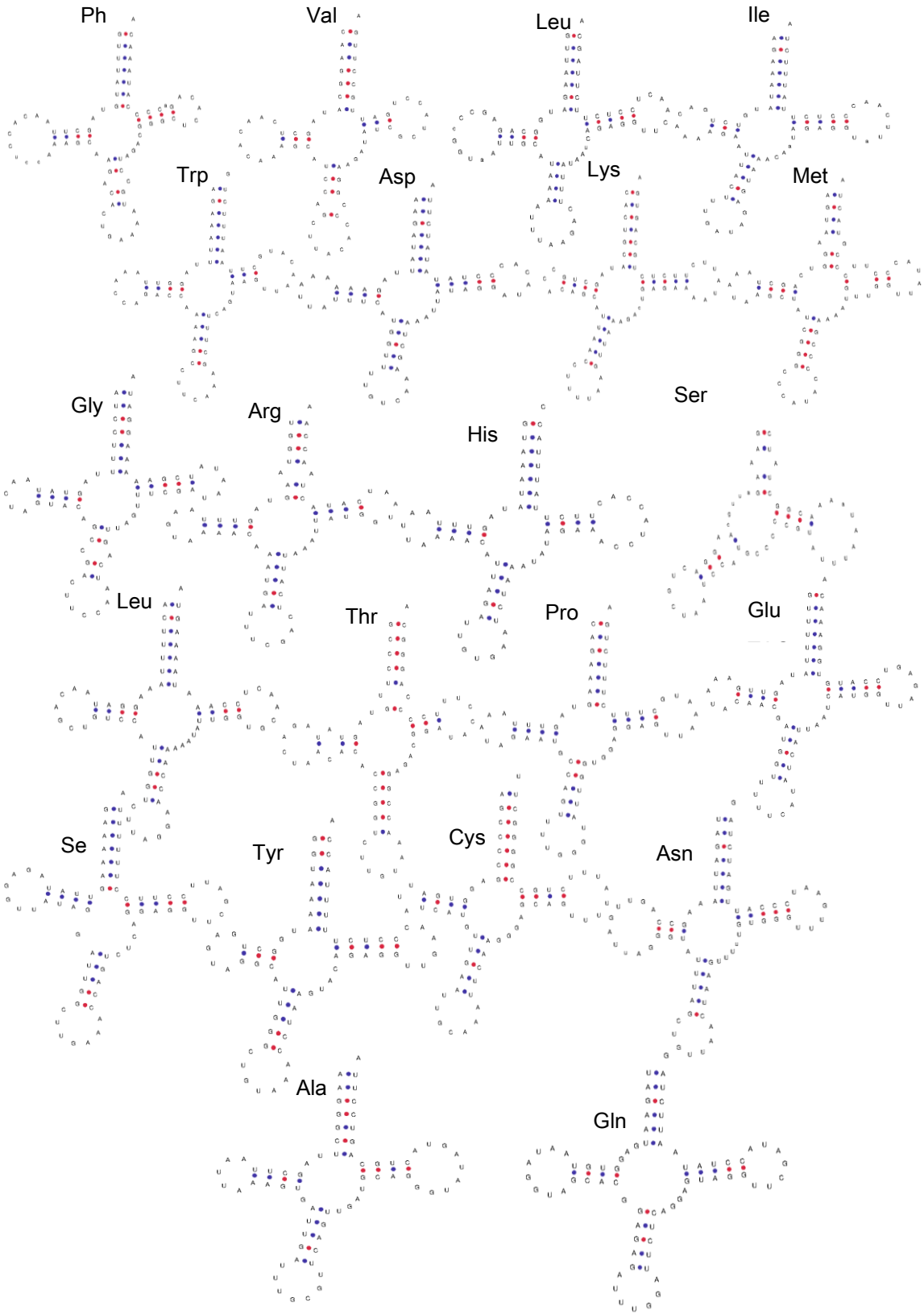
Toàn bộ 22 gene transfer RNA (*tRNA*) trong hệ gene ty thể của *Chà vá chân nâu* Tây Nguyên Việt Nam (Hình 3. 19) có tám gene *tRNA* nằm trên mạch nhẹ (L-strand), trong khi các gene *tRNA* còn lại nằm trên mạch nặng (H-strand) (Bảng 3.4). Tất cả 22 gene *tRNA* được xác định trong hệ gene ty thể của *Chà vá chân nâu* đều mang anticodon tương thích với những gì đã được ghi nhận ở các loài linh trưởng khác, cho thấy mức độ bảo tồn cao về chức năng dịch mã. Chiều dài các gene *tRNA* dao động từ 59 bp đến 75 bp, trong đó *trnS1* là gene có kích thước nhỏ nhất (59 bp) và *trnL2* là gene có kích thước lớn nhất (75 bp). Các gene *tRNA* nhìn chung có kích thước ngắn và tương đối đồng đều, phù hợp với đặc điểm của *tRNA* ty thể ở linh trưởng. Một số gene *tRNA* trong hệ gene ty thể *Chà vá chân nâu* được ghi nhận chứa các cặp base không khớp, với tổng cộng 9 vị trí không bắt cặp được phát hiện ở nhiều gene *tRNA* khác nhau (Bảng 3. 14). Cụ thể, *tRNA*-Methionine (*trnM*) là gene ghi nhận nhiều cặp base không khớp nhất, với bốn dạng không bắt cặp, bao gồm A–G tại thân nhận axit

amin (AA stem) (1 lần), cùng với U–U (2 lần), C–U (1 lần) và A–U (1 lần) tại thân TΨC (T-arm). Điều này cho thấy thân TΨC của *trnM* có mức độ biến đổi cao hơn so với thân nhận axit amin, phản ánh tính linh hoạt cấu trúc của vùng này. Đối với *tRNA*-Arginine (*trnR*, anticodon CGA), chỉ ghi nhận một cặp base không khớp dạng A–A, xuất hiện tại thân nhận axit amin, cho thấy mức độ biến đổi cấu trúc thấp hơn đáng kể so với *trnM*. Tương tự, *tRNA*-Serine (*trnS*) ghi nhận hai cặp base không khớp, bao gồm một cặp A–A tại thân nhận axit amin và một cặp A–A tại thân anticodon (AC stem). Ngoài ra, *tRNA*-Phenylalanine (*trnF*) và *tRNA*-Valine (*trnV*) mỗi gene ghi nhận một cặp base không khớp dạng A–C, đều xuất hiện tại thân anticodon. *tRNA*-Glycine (*trnG*) ghi nhận một cặp không khớp C–A tại thân anticodon, trong khi *tRNA*-Threonine (*trnT*) có hai cặp base không khớp, bao gồm U–C tại thân nhận axit amin và A–C tại nhánh D (D-arm).

Nhìn chung, các cặp base không bắt cặp trong các gene *tRNA* của *Chà vá chân nâu* phân bố chủ yếu tại thân nhận axit amin, thân TΨC và thân anticodon, trong khi số lượng không bắt cặp tại các vùng này là thấp và phân tán, không tập trung quá mức vào một gene duy nhất (ngoại trừ *trnM*). Điều này cho thấy mặc dù tồn tại một số biến đổi cấu trúc cục bộ, cấu trúc chẻ ba tổng thể và chức năng dịch mã của các *tRNA* ty thể vẫn được bảo tồn.

Bảng 3. 14. Các cặp base không khớp của *t-RNAs*. AA—amino acid acceptor, T-arm, AC—anticodon.

tRNA	**Cặp base không khớp**	**thân**	**Tần số**
Phenylalanine	A-C	AC	1
Valine	A-C	AC	1
Methylonine	A-G	AA	1
	U-U	T-arm	2
	C-U	T-arm	1
	A-U	T-arm	1
Glycine	C-A	AC	1
Arginine TCG	A-A	AA	1
Serine	A-A	AA	1
	A-A	AC	1
Threonine	U-C	AA	1
	A-C	D arm	1



Hình 3. 19. Cấu trúc thứ cấp của 22 gene *tRNA* của *Chà vá chân nâu* Tây Nguyên Việt Nam. Màu đỏ đại diện cho các kết nối GC, trong khi màu xanh dương đại diện cho các kết nối AU.

3.2.3.5. Vùng lặp và vùng liên gene

Phân tích so sánh các gene ty thể của *Chà vá chân nâu* Tây Nguyên Việt Nam cho thấy một số vùng lặp (overlapping genes) và vùng liên gene, phản ánh đặc điểm sắp xếp chặt chẽ của hệ gene ty thể. Cụ thể, gene *atp8* và *atp6* có một vùng lặp dài 40 bp, được xác định tại ranh giới giữa hai gene này. Đây là dạng chồng chéo phổ biến trong hệ gene ty thể của động vật có xương sống và đã được ghi nhận rộng rãi ở các loài linh trưởng, cho thấy mức độ bảo tồn cao của cấu trúc gene trong vùng này. Ngoài ra, gene *trnV* và *rrnL* ghi nhận một sự chồng chéo 2 bp, tương tự như mô hình đã được quan sát ở nhiều loài linh trưởng khác. Gene trnI và trnQ cũng có sự chồng chéo 3 bp, trong khi gene *nad4L* và *nad4* có 4 bp chồng lên nhau. Các trường hợp chồng chéo này đều xảy ra trên mạch nặng (H-strand) và phù hợp với các đặc điểm đã được mô tả ở hệ gene ty thể linh trưởng. Nhìn chung, các vùng liên gene trong hệ gene ty thể *Chà vá* chân nâu chủ yếu xuất hiện giữa các gene mã hóa protein hoặc giữa gene *tRNA* và rRNA, với độ dài chồng chéo ngắn (từ 2–40 bp), cho thấy áp lực chọn lọc nhằm duy trì kích thước hệ gene ty thể nhỏ gọn.

Bên cạnh các vùng liên gene, hệ gene ty thể của *Chà vá chân nâu* cũng chứa nhiều vùng khoảng cách liên gene không mã hóa. Các vùng liên gene này có kích thước dao động từ 1 bp đến 86 bp, phân bố rải rác giữa các gene ty thể. Trong số đó, vùng liên gene dài nhất (86 bp) nằm giữa *cox2* và *trnK*, tiếp theo là một số vùng có kích thước trung bình như 32 bp giữa *trnN* và *trnC*.

Tổng thể, số lượng và chiều dài các vùng liên gene ở *Chà vá chân nâu* Tây Nguyên Việt Nam cho thấy hệ gene ty thể có mức độ nén cao, với phần lớn các vùng liên gene có kích thước ngắn. Mặc dù trước đây các vùng này thường được xem là không mang chức năng, nhiều nghiên cứu gần đây cho rằng một số vùng liên gene có thể tham gia vào điều hòa sao chép và phiên mã gene ty thể.

So sánh với các loài linh trưởng khác cho thấy *Chà vá chân nâu* có hệ thống vùng lặp và vùng liên gene tương đồng với các loài *Chà vá* và linh trưởng, phản ánh tính bảo tồn cao trong tổ chức hệ gene ty thể, đồng thời vẫn duy trì một mức độ biến thiên nhất định ở các vùng không mã hóa.

3.2.3.6. Thảo luận

Hệ gene ty thể của *Chà vá chân nâu* Tây Nguyên Việt Nam có chiều dài và tổ chức trình tự tương đồng với các mẫu đã công bố, trong đó sự khác biệt về tổng kích thước chủ yếu bắt nguồn từ mức độ hoàn chỉnh của vùng điều khiển trong các dữ liệu tham chiếu. Điều này cho thấy các sai khác về chiều dài ADN ty thể không nhất thiết

phản ánh biến thiên sinh học thực sự. So sánh kích thước hệ gene cho thấy cá thể *Chà vá chân nâu* Tây Nguyên Việt Nam có chiều dài nhỏ hơn mẫu JF_293096 (16.559 bp) nhưng lớn hơn mẫu NC008220 (15.467 bp). Sự chênh lệch đáng kể với NC008220 chủ yếu bắt nguồn từ việc thiếu số liệu vùng *D-loop* trong hệ gene này. Trên thực tế, chiều dài vùng *D-loop* của *Chà vá chân nâu* Tây Nguyên Việt Nam đạt 1.083 bp, tương đương với các mẫu *Chà vá chân nâu* và *Chà vá* khác (~1.085–1.092 bp), cho thấy sự khác biệt về tổng chiều dài hệ gene không phản ánh sai khác sinh học thực sự, mà chủ yếu do khác biệt trong chú giải và mức độ hoàn chỉnh của trình tự công bố (47,99).

Hệ gene ty thể của *Chà vá chân nâu* Tây Nguyên Việt Nam có hàm lượng A+T đạt 61,1%, thấp hơn nhẹ so với các mẫu *Chà vá chân nâu* khác (61,3%) nhưng vẫn nằm trong khoảng đặc trưng của linh trưởng. Xu hướng giàu A/T này là đặc điểm phổ biến của ADN ty thể động vật, đặc biệt ở các loài linh trưởng, và có liên quan mật thiết đến cơ chế sao chép, phiên mã cũng như áp lực chọn lọc ở cấp độ phân tử ((48,102)

Sự thiên lệch A/T cao đóng vai trò quan trọng trong việc hình thành thiên lệch sử dụng codon đồng nghĩa, khi các codon kết thúc bằng A hoặc U được sử dụng ưu tiên hơn so với các codon kết thúc bằng G hoặc C. Điều này đã được quan sát rõ rệt trong phân tích RSCU của *Chà vá chân nâu*, đặc biệt ở các amino acid có nhiều codon đồng nghĩa như *Leu*, *Ser*, *Ala* và *Pro*. Xu hướng này phù hợp với các nghiên cứu trước đây trên hệ gene ty thể linh trưởng và động vật có xương sống, cho thấy mối liên hệ chặt chẽ giữa thành phần nucleotide và chiến lược sử dụng codon ((100,103).

Phân tích AT-skew và GC-skew trong 13 gene mã hóa protein của *Chà vá* chân nâu Tây Nguyên Việt Nam cho thấy AT-skew âm và GC-skew âm, phản ánh ưu thế của T so với A và của C so với G trên mạch được phân tích. Kiểu bất đối xứng nucleotide này là đặc điểm phổ biến của gene ty thể động vật và được cho là hệ quả của cơ chế sao chép bất đối xứng của ADN ty thể, trong đó hai mạch chịu các áp lực đột biến khác nhau trong quá trình sao chép (48,104). Sự khác biệt về biên độ AT-skew và GC-skew giữa các gene, đặc biệt là các gene thuộc nhóm *NAD* (như *nad4* và *nad5*), cho thấy tính không đồng nhất về thành phần nucleotide giữa các PCGs. Tuy nhiên, xu hướng skew tổng thể vẫn ổn định và tương đồng giữa các cá thể *Chà vá* chân nâu có nguồn gốc địa lý khác nhau, cho thấy độ thiên lệch nucleotide của hệ gene ty thể mang tính bảo tồn cao, ít chịu ảnh hưởng bởi sự phân hóa quần thể (102).

Việc chỉ ghi nhận 9 cặp base không khớp, trong toàn bộ 22 gene *tRNA* của *Chà vá chân nâu* Tây Nguyên Việt Nam có cấu trúc thứ cấp dạng chẻ ba điển hình phân

bố chủ yếu ở thân nhận axit amin, thân TΨC và thân anticodon, cho thấy mặc dù tồn tại biến đổi cấu trúc cục bộ, cấu trúc tổng thể và chức năng dịch mã của *tRNA* vẫn được duy trì ổn định. Đặc biệt, sự tập trung mismatch ở *trnM* phản ánh tính linh hoạt cao hơn của một số *tRNA* nhất định, điều đã được ghi nhận ở nhiều loài động vật khác (105). Hai gene *rRNA* (*12S rRNA* và *16S rRNA*) của *Chà vá chân nâu* có chiều dài và hàm lượng A+T (59,4–59,5%) tương đương với các loài *Chà vá* và linh trưởng khác. Do vai trò thiết yếu trong cấu trúc ribosome ty thể, các gene rRNA thường chịu áp lực chọn lọc mạnh, dẫn đến mức độ bảo tồn cao cả về kích thước lẫn thành phần nucleotide (Cann et al., 1987).

Vùng điều khiển (*D-loop*) của *Chà vá chân nâu* Tây Nguyên Việt Nam có chiều dài 1.083 bp, phù hợp với khoảng kích thước điển hình của động vật có xương sống. Sự biến thiên chiều dài *D-loop* giữa các loài và cá thể chủ yếu liên quan đến số lượng và kiểu lặp nối tiếp (tandem repeats), vốn là nguyên nhân chính tạo ra tính đa hình cao của vùng này (47,99).

Các vùng chồng chéo gene như *atp8–atp*6 (40 bp), *trnV–rrnL* (2 bp), *trnI–trnQ* (3 bp) và *nad4L–nad4* (4 bp) đều được ghi nhận ở *Chà vá chân nâu*, tương đồng với các loài *Chà vá* khác. Những vùng chồng chéo này phản ánh áp lực tiến hóa nhằm tối ưu hóa không gian hệ gene ty thể, một đặc điểm điển hình của ADN ty thể động vật (106). Ngược lại, các vùng liên gene không mã hóa thể hiện sự biến thiên lớn hơn về chiều dài, góp phần tạo ra khác biệt nhẹ về tổng kích thước hệ gene giữa các loài và cá thể.

Những biến thiên được ghi nhận chủ yếu tập trung ở vùng không mã hóa và các khoảng liên gene, gợi ý vai trò tiềm năng của các vùng này trong điều hòa biểu hiện gene ty thể. Tuy nhiên, các sai khác này cần được xem xét thận trọng do nghiên cứu chỉ sử dụng một cá thể đại diện.

3.2.4. Đặc điểm hệ gene ty thể hoàn chỉnh của cá thể *Chà vá chân đen* (*Pygathrix nigripes*) Tây Nguyên Việt Nam

Nghiên cứu này cho kết quả hệ gene ty thể hoàn chỉnh đầu tiên của *Chà vá chân đen* Tây Nguyên Việt Nam, gồm 16,535 bp (Hình 3. 3) với cấu trúc đặc trưng, có kích thước lớn hơn hệ gene ty thể của hai mẫu *Chà vá chân đen* khác được thu thập ở Nanning, Trung Quốc (16,536 bp) (Su *et al.*, 2018).

Bảng 3. 15. Cấu trúc và đặc trưng của hệ gene ty thể hoàn chỉnh của *Chà vá chân đen*

Tên	**Bắt đầu**	**Kết thúc**	**Mạch**	**Chiều dài**
nad6_a	22	276	L	255
trnE(gaa)	277	345	L	69
cytb	350	1483	H	1134
trnT(aca)	1491	1554	H	64
trnP(cca)	1557	1623	L	67
A + T rich region	1624	2716	-	1093
trnF(ttc)	2717	2787	H	71
rrnS	2788	3736	H	949
trnV(gta)	3737	3802	H	66
rrnL	3801	5362	H	1562
trnL2(tta)	5363	5437	H	75
nad1	5440	6390	H	951
trnI(atc)	6395	6464	H	70
trnQ(caa)	6462	6533	L	72
trnM(atg)	6534	6601	H	68
nad2	6602	7639	H	1038
trnW(tga)	7644	7710	H	67
tRNA(gca)	7719	7787	L	69
trnN(aac)	7789	7861	L	73
trnC(tgc)	7894	7959	L	66
trnY(tac)	7960	8024	L	65
cox1	8026	9567	H	1542
trnS2(tca)	9568	9636	L	69
trnD(gac)	9640	9707	H	68
cox2	9709	10368	H	660
trnK(aaa)	10436	10503	H	68
atp8	10505	10687	H	183

atp6	10666	11340	H	675
cox3	11346	12128	H	783
trnG(gga)	12130	12196	H	67
nad3	12197	12541	H	345
trnR(cga)	12543	12607	H	65
nad4l	12608	12901	H	294
nad4	12898	14265	H	1368
trnH(cac)	14276	14344	H	69
trnS1(agc)	14345	14403	H	59
trnL1(cta)	14404	14474	H	71
nad5	14478	16277	H	1800
nad6_b	16290	16535	L	246

Hệ gene ty thể của *Chà vá chân đen* Tây Nguyên Việt Nam đã được giải trình tự, cho thấy sự đa dạng của các gene mã hóa protein với các độ dài khác nhau. Đáng chú ý, gene dài nhất được xác định là *cytochrome b* (*cytb*) với kích thước 1,134 bp, trong khi gene ngắn nhất là ATP8 với kích thước 183 bp (Bảng 3. 15). Thành phần nucleotide tổng thể cho thấy sự phong phú cao hơn của adenine (A) và thymine (T) so với guanine (G) và cytosine (C). Sự giàu có A-T này là đặc trưng của nhiều hệ gene ty thể của linh trưởng và gợi ý về các chiến lược tiến hóa thích nghi.

3.2.4.1. Thành phần nucleotide

Với tỷ lệ phần trăm của mỗi nucleotide cho từng gene, cùng với tổng số nucleotide cho gene đó. Các giá trị trung bình trên tất cả các gene là: Thymine (T): 31.24%; Guanine (G): 12.37% (Bảng 3. 16) cho thấy sự giàu có A-T chung, điều này là đặc trưng của nhiều hệ gene ty thể. *Nad6a* (22-276): Gene này cho thấy tỷ lệ T rất cao (43.92%) và tỷ lệ C thấp (9.41%). Tỷ lệ T cao như vậy có thể gợi ý một vai trò trong các chức năng điều hòa hoặc một sự thích ứng cụ thể làm tăng hiệu quả phiên mã của nó. *Nad1* (5440-6390): Hiển thị thành phần nucleotide cân bằng hơn, với T chiếm 29.13%, A chiếm 30.70% và C chiếm 29.65%. *Nad6b* (16290-16535): thể hiện tỷ lệ C rất thấp (5.28%) và tỷ lệ G cao (32.52%). *Atp8* (10505-10687): Có tỷ lệ A cao nhất (38.8%), đóng vai trò trong các quá trình chuyển hóa, có khả năng tăng cường tổng hợp ATP trong ty thể.

Các giá trị trung bình cho thấy các xu hướng chung trong việc sử dụng nucleotide: Có sự ưu tiên cho A và T (giàu A-T) so với C và G (thiếu G-C) và tỷ lệ G thấp ở tất cả các gene có thể ảnh hưởng đến tính ổn định và chức năng của hệ gene. Thành phần nucleotide có thể phản ánh các vai trò chức năng của các gene này. Các gene có tỷ lệ T và A cao hơn có thể được biểu hiện tích cực hơn hoặc tham gia vào các con đường chuyển hóa quan trọng. Sự biến đổi trong thành phần nucleotide cũng có thể cung cấp cái nhìn về các thích ứng tiến hóa, chẳng hạn như phản ứng với các áp lực môi trường hoặc thay đổi môi trường sống.

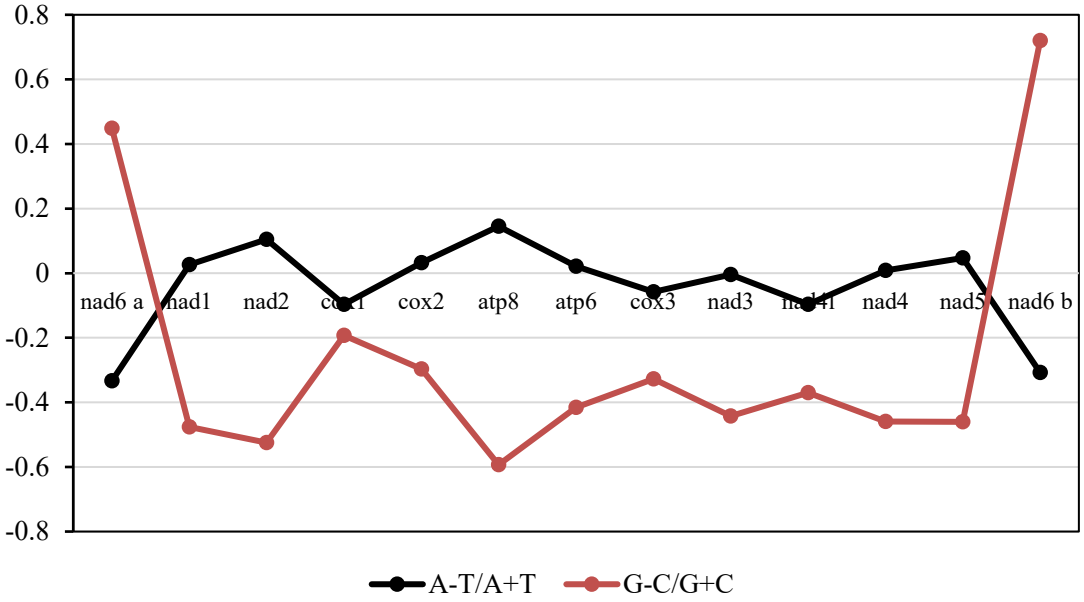
Bảng 3. 16. Thành phần nucleotide của 13 gene mã hóa protein trong hệ gene ty thể của *Chà vá chân đen*.

	T(U)	C	A	G	Tổng
22-276 - nad6 a	43.92157	9.411765	21.96078	24.70588	255
5440-6390 + nad1	29.12723	29.653	30.70452	10.51525	951
6602-7639 + nad2	28.42004	27.842	35.06744	8.67052	1038
8026-9567 + cox1	32.36057	24.44877	26.6537	16.53696	1542
9709-10368 + cox2	29.84848	24.84848	31.81818	13.48485	660
10505-10687 + atp8	28.96175	25.68306	38.79781	6.557377	183
10666-11340 + atp6	30.07407	27.25926	31.40741	11.25926	675
11346-12128 + cox3	31.41762	26.94764	27.96935	13.66539	783
12197-12541 + nad3	32.46377	25.50725	32.17391	9.855072	345
12608-12901 + nad4l	34.69388	25.17007	28.57143	11.56463	294
12898-14265 + nad4	31.21345	27.04678	31.72515	10.01462	1368
14478-16277 + nad5	30.27778	26.61111	33.27778	9.833333	1800
16290-16535 - nad6 b	40.65041	5.284553	21.54472	32.52033	246
Trung bình	31.2426	25.66075	30.72978	12.36686	780

3.2.4.2. Các gene mã hóa Protein

Hầu hết các cá thể *chà vá*, bao gồm *Chà vá chân đen, Chà vá chân xám*, và *Chà vá chân nâu*, cho thấy thành phần nucleotide tương tự với A và T có phần phong phú hơn C và G (Bảng 3. 16). Các cá thể khỉ của chi *macaca* thể hiện sự giảm đáng kể

trong A và T, kèm theo sự gia tăng tương ứng trong C, điều này gợi ý về các áp lực tiến hóa khác nhau hoặc các thích ứng chức năng. Nội dung GC trung bình dường như cao hơn ở các cá thể khỉ của chi *macaca*, được phản ánh qua GC-skew dương của chúng, chỉ ra một thành phần nucleotide cân bằng hơn so với các cá thể *chà vá*.



Hình 3. 20. AT và GC skew trên 13 gene mã hóa protein của hệ gene ty thể *Chà vá chân đen* Tây Nguyên Việt Nam.

Các giá trị AT-skew và GC-skew của 13 gene mã hóa protein trong hệ gene ty thể của *Chà vá chân đen* Tây Nguyên Việt Nam (Hình 3. 20) cho thấy các gene như *nad1*, *nad2*, *cox1*, *cox2*, *atp6*, *cox3*, *nad3*, *nad4l*, *nad4* và *nad5* đều thể hiện xu hướng này với mức độ khác nhau. Ngược lại, hai đoạn gene *nad6* (*nad6a* và *nad6b*) có giá trị AT-skew âm rõ rệt, phản ánh sự ưu thế của T so với A, cho thấy đặc điểm phân bố nucleotide khác biệt của gene này so với các gene mã hóa protein còn lại. Đối với GC-skew, hầu hết các gene có giá trị âm, thể hiện sự ưu thế của C so với G trong thành phần nucleotide. Xu hướng này được ghi nhận nhất quán ở đa số các gene thuộc cả nhóm NADH dehydrogenase, cytochrome c oxidase và ATP synthase. Một số gene như *cox1* và *cox3* cho thấy GC-skew ít âm hơn, phản ánh mức độ chênh lệch giữa G và C thấp hơn so với các gene khác.

Phân tích thành phần nucleotide và các chỉ số skew cho thấy các cá thể nghiên cứu có xu hướng phân bố nucleotide tương đối đồng nhất (Bảng 3. 17). Kết quả phân tích AT-skew và GC-skew của hệ gene ty thể cho thấy *Chà vá chân đen* Tây Nguyên Việt Nam và cá thể *Chà vá chân đen* MH_064177 có thành phần nucleotide và xu hướng thiên lệch base gần như tương đồng. Cụ thể, cá thể *Chà vá chân đen* Tây Nguyên Việt Nam có tỷ lệ T = 31,2%, A = 30,7%, C = 25,7% và G = 12,4%, trong

khi cá thể MH_064177 có các giá trị tương ứng là T = 31,2%, A = 30,6%, C = 25,9% và G = 12,3%. Cả hai cá thể đều thể hiện rõ ưu thế của A+T so với C+G trong hệ gene ty thể

Giá trị AT-skew của hai cá thể *Chà vá chân đen* đều âm nhẹ, lần lượt là –0,00827 (Tây Nguyên Việt Nam) và –0,01002 (MH_064177), cho thấy hàm lượng T nhỉnh hơn A (Bảng 3. 17). Đồng thời, GC-skew đều mang giá trị âm rõ (–0,34959 và –0,35463), phản ánh sự ưu thế của C so với G trong thành phần nucleotide. Mức độ sai khác giữa hai cá thể là rất nhỏ, cho thấy mô hình thiên lệch nucleotide được bảo tồn cao trong nội bộ loài *Chà vá chân đen*, bất kể sự khác biệt về nguồn gốc địa lý. So sánh nội bộ giữa các cá thể *Chà vá chân đen* cho thấy các chỉ số skew ổn định và nhất quán, không ghi nhận sự đảo chiều AT-skew hay GC-skew. Kết quả phân tích cho thấy cấu trúc phân bố nucleotide trong hệ gene ty thể của *Chà vá chân đen* ít biến đổi. Những kết quả này cung cấp cơ sở dữ liệu định lượng quan trọng cho các phân tích tiếp theo về so sánh nội loài và đánh giá mức độ biến thiên di truyền trong hệ gene ty thể của *Chà vá chân đen*.

Bảng 3. 17. AT-skew, và GC-skew của *Chà vá chân đen* Tây Nguyên Việt Nam

	Số gia nhập	**T(U)**	**C**	**A**	**G**	**Tổng**	**AT-Skew**	**GC-Skew**
Chà vá chân đen Tây Nguyên Việt Nam		31.2	25.7	30.7	12.4	780.0	-0.00827	-0.34959
Chà vá chân đen	MH_064177	31.2	25.9	30.6	12.3	807.4	−0.01002	−0.35463
Chà vá chân xám	JQ_821842	30.9	26.1	30.5	12.5	868.6	−0.0062	−0.35277
Chà vá chân nâu	JF_293096	30.9	26.1	30.5	12.5	868.6	−0.00678	−0.35155
Macaca leonina	KP_330231	26.5	30.9	30.2	12.3	870	0.065732	−0.42945
Presbytis femoralis	KU_899140	31.5	25.4	30.9	12.2	867.5	−0.00896	−0.34998
Nasalis larvatus	DQ_355298	30.5	26.7	30.6	12.2	869.1	0.001593	−0.37082
Rhinopithecus roxellana	KM_504390	31.4	25.6	30.8	12.2	868.6	−0.01054	−0.35253
Rhinopithecus brelichi	JQ_821836	31.6	25.5	30.6	12.4	868.8	−0.01638	−0.34612
Rhinopithecus bieti	JQ_821839	31.5	25.6	30.7	12.2	869.1	−0.01181	−0.35612

Rhinopithecus avunculus	JF_293093	31.5	25.5	30.8	12.3	868.4	−0.01138	−0.34898
Macaca arctoides	KM_360179	26.4	31	30	12.5	870.2	0.063577	−0.4244
Macaca mulatta	JQ_821843	26.6	31	29.8	12.6	870	0.056444	−0.42052
Papio hamadryas	NC_001992	26.2	31.2	29.7	12.9	868.6	0.061768	−0.41663
Callithrix jacchus	KM_588314	28.3	27.8	31.1	12.9	867.7	0.046296	−0.36736
Homo sapiens neanderthalensis	OM_062614	26.2	31.9	28.8	13.1	867.8	0.046422	−0.41631
Cebus albifrons	NC_002763	29.9	26.7	31.3	12.1	866.8	0.02306	−0.37663

3.2.4.3. Vùng giàu A + T

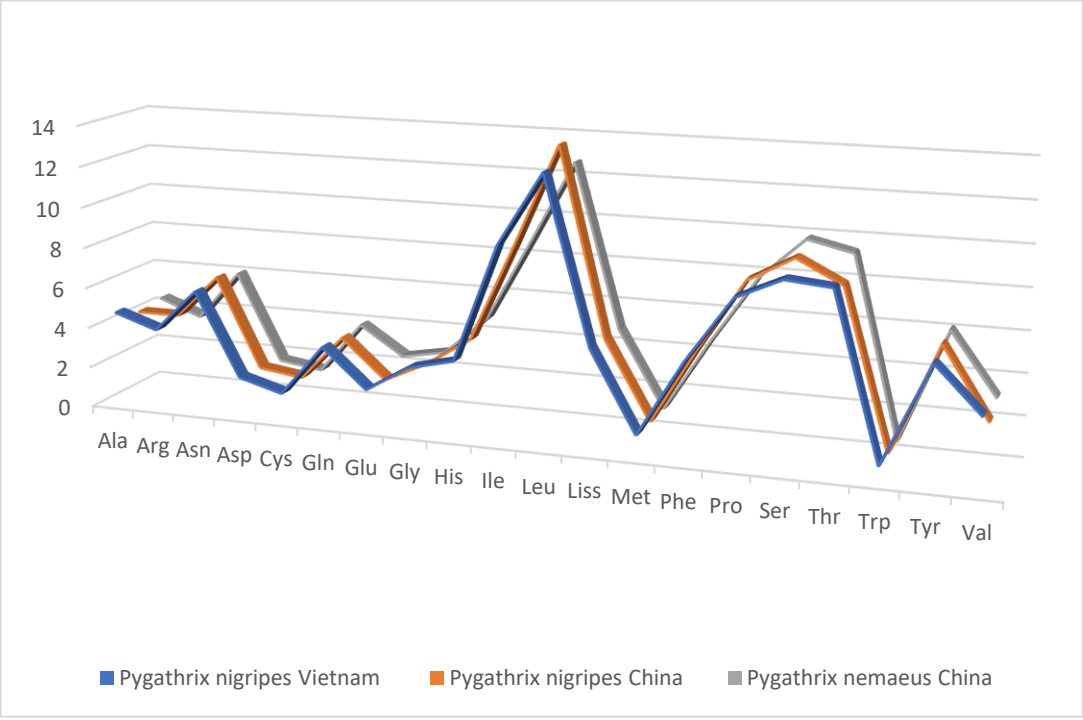
Vùng *D-loop* trong hệ gene ty thể của *Chà vá chân đen* Tây Nguyên Việt Nam có chiều dài 1.093 bp, kéo dài từ vị trí 1624 đến 2716, nằm giữa hai gene *trnP* và *trnF* (Bảng 3. 15). Đây là vùng không mã hóa dài nhất trong toàn bộ hệ gene ty thể của loài, chiếm tỷ lệ đáng kể trong tổng chiều dài genome. Vị trí và kích thước của vùng *D-loop* phù hợp với cấu trúc điển hình của hệ gene ty thể ở linh trưởng, nơi vùng này thường nằm giữa các gene *tRNA* và đóng vai trò trung tâm trong điều hòa sao chép và phiên mã ty thể. Sự tồn tại của vùng không mã hóa dài, giàu A+T cho thấy đây là khu vực có mức độ biến thiên nucleotide cao hơn so với các gene mã hóa protein và *rRNA*, đồng thời có tiềm năng chứa các yếu tố điều hòa như điểm khởi đầu sao chép mạch nặng và mạch nhẹ, cũng như các trình tự liên quan đến kiểm soát phiên mã. Kết quả xác định vùng *D-loop* trong nghiên cứu này cung cấp cơ sở dữ liệu quan trọng cho các phân tích tiếp theo về biến dị di truyền và tiến hóa phân tử của *Chà vá chân đen* ở Việt Nam

3.2.4.4. Thành phần acid amin

Thành phần axit amin và tỷ lệ tương đối (%) của ba cá thể *chà vá*, bao gồm *Chà vá chân đen* Tây Nguyên Việt Nam, *Chà vá chân đen* Trung Quốc và *Chà vá chân nâu* Trung Quốc cho thấy phổ axit amin của ba cá thể có cấu trúc tương đối tương đồng, phản ánh tính bảo tồn cao của các gene mã hóa protein ty thể trong nội bộ nhóm *Chà vá* (Hình 3. 21). Ở cả ba cá thể, leucine (*Leu*) là axit amin chiếm tỷ lệ cao nhất, dao động từ 12,7% đến 13,9%, tiếp theo là isoleucine (*Ile*) (8,7–9,5%), serine (*Ser*) (8,9–9,8%), threonine (*Thr*) (8,3–9,3%) và proline (*Pro*) (7,8–8,2%). Nhóm các axit amin này chiếm tỷ trọng lớn trong tổng thành phần axit amin, phản ánh vai trò chủ đạo của chúng trong cấu trúc và chức năng của các protein ty thể.

So sánh giữa các cá thể, *Chà vá chân đen* Tây Nguyên Việt Nam có tỷ lệ *Ile* (9,5%) và Val (3,5%) cao hơn nhẹ so với hai cá thể còn lại, trong khi *Chà vá chân đen* Trung Quốc thể hiện tỷ lệ *Leu* (13,9%) và *Ser* (9,4%) cao nhất trong ba cá thể. Đối với *Chà vá* chân nâu Trung Quốc, tỷ lệ *His* (4,8%), *Ser* (9,8%) và *Thr* (9,3%) cao hơn so với hai cá thể *Chà vá* chân đen, cho thấy sự khác biệt nhẹ về phân bố axit amin giữa các loài trong cùng chi (Bảng 3. 18). Các axit amin có tỷ lệ thấp như methionine (*Met*) (1,1% ở cả ba cá thể), tryptophan (*Trp*) (0,6–0,8%) và cysteine (*Cys*) (1,4–1,6%) thể hiện mức độ biến thiên rất nhỏ giữa các cá thể. Điều này cho thấy các axit amin hiếm trong protein ty thể được duy trì ổn định, bất kể nguồn gốc địa lý hay loài.

Tổng hợp các kết quả trên cho thấy thành phần và tỷ lệ axit amin trong hệ gene ty thể của các cá thể *Chà vá* có mức độ tương đồng cao, đồng thời tồn tại một số khác biệt nhỏ về tỷ lệ tương đối của từng axit amin giữa các cá thể. Những khác biệt này phản ánh sự biến thiên ở mức độ loài, đồng thời cung cấp cơ sở dữ liệu quan trọng cho các phân tích tiếp theo về tiến hóa protein ty thể và so sánh liên loài trong chi *Pygathrix*.



Hình 3. 21. Thành phần axit amin và tỷ lệ tương đối (%) của chúng trong hệ gene ty thể hoàn chỉnh của *Chà vá chân đen* Tây Nguyên Việt Nam; *Chà vá chân đen* Trung Quốc và *Chà vá chân đỏ* Trung Quốc.

Bảng 3. 18. Tỷ lệ phần trăm của thành phần axit amin trong toàn bộ hệ gene ty thể của các cá thể *chà vá: Chà vá chân đen* Tây Nguyên Việt Nam*, Chà vá chân đen* Trung Quốc*, và Chà vá chân đỏ* Trung Quốc.

	Chà vá chân đen **Tây Nguyên Việt Nam**	***Chà vá chân đen*** **Trung Quốc**	***Chà vá chân đỏ*** **Trung Quốc**
Ala	4.7	4.2	4.4
Arg	4	4.2	3.6
Asn	6.1	6.3	6
Asp	2	1.9	1.7
Cys	1.4	1.6	1.4
Gln	3.9	3.8	3.9
Glu	2	1.9	2.5
Gly	3.3	2.9	2.9
His	3.8	4.3	4.8
Ile	9.5	9	8.7
Leu	13	13.9	12.7
Liss	5	4.9	4.8
Met	1.1	1.1	1.1
Phe	4.8	4.8	4.6
Pro	7.9	8.2	7.8
Ser	8.9	9.4	9.8
Thr	8.7	8.3	9.3
Trp	0.8	0.7	0.6
Tyr	5.7	6	6.1
Val	3.5	2.6	3.1

3.2.4.5. Vùng lặp và vùng liên gene

Các vùng liên gene trong hệ gene ty thể của *Chà vá chân đen* chủ yếu có chiều dài ngắn; vùng liên gene giữa *nad6_a* (kết thúc tại vị trí 276) và *trnE(gaa)* (bắt đầu tại vị trí 277) không ghi nhận khoảng cách, cho thấy hai gene này được sắp xếp liền kề. Một vùng liên gene nhỏ được ghi nhận giữa *trnE(gaa)* (345) và *cytb* (350), với khoảng cách 4 bp, cho thấy sự tồn tại của một đoạn không mã hóa ngắn. Giữa *cytb* và *trnT(aca)* tồn tại khoảng cách 7 bp (1483–1491), phản ánh một vùng liên gene điển hình có kích thước nhỏ. Các vùng liên gene ngắn (từ 2–10 bp) tiếp tục được ghi nhận giữa nhiều cặp gene *tRNA* liền kề, như *trnT(aca)*/*trnP(cca)*, *trnF(ttc)*/*rrnS*, *rrnS*/*trnV(gta)* và *trnV(gta)*/*rrnL*. Đặc biệt, vùng giàu A+T (A+T-rich region) là vùng liên gene dài nhất trong hệ gene ty thể, với chiều dài 1.093 bp, nằm giữa *trnP(cca)* và *trnF(ttc)*. Đây là vùng không mã hóa có vai trò điều hòa, được xem là trung tâm khởi đầu sao chép và phiên mã hệ gene ty thể. Sự hiện diện của vùng liên gene lớn này là đặc điểm phổ biến và mang tính bảo tồn ở hệ gene ty thể linh trưởng.

Bên cạnh các vùng liên gene, một số vùng lặp gene (chồng chéo) cũng được ghi nhận trong hệ gene ty thể của *Chà vá* chân đen Tây Nguyên Việt Nam. Cặp gene *atp8* (10505–10687) và *atp6* (10666–11340) có sự chồng chéo 22 bp, trong đó đoạn cuối của *atp8* trùng với đoạn đầu của *atp6*. Đây là vùng lặp gene điển hình, đã được ghi nhận phổ biến trong hệ gene ty thể của nhiều loài động vật có vú và linh trưởng. Cặp gene *trnI(atc)* (6395–6464) và *trnQ(caa)* (6462–6533) có sự chồng chéo 3 bp, phản ánh mức độ nén cao trong nhóm gene *tRNA*. Một vùng lặp nhỏ 2 bp được ghi nhận giữa *trnV(gta)* và *rrnL*, cho thấy sự sắp xếp chặt chẽ giữa gene *tRNA* và gene rRNA.

Các vùng lặp gene này có kích thước nhỏ (từ 2–22 bp), không làm gián đoạn khung đọc mở của các gene mã hóa protein, đồng thời phản ánh cơ chế tổ chức genome ty thể tối ưu, trong đó các gene có thể chia sẻ một phần trình tự nucleotide nhằm giảm thiểu độ dài tổng thể của hệ gene.

3.2.4.6. Thảo luận

Nghiên cứu này lần đầu tiên công bố hệ gene ty thể hoàn chỉnh của *Chà vá chân đen* Tây Nguyên Việt Nam, một loài linh trưởng đang bị đe dọa nghiêm trọng và đặc hữu của Việt Nam (107). Loài này đối mặt với những mối đe dọa đáng kể từ sự phá hủy môi trường sống, săn trộm và phân mảnh, chủ yếu là do nạn phá rừng và mở rộng nông nghiệp (21,27). Nghiên cứu gần đây dự đoán rằng phân bố của *Chà vá chân đen* có thể giảm đáng kể, với ước tính mất 66,19% diện tích hiện tại vào năm 2070 do biến đổi khí hậu(108). Hiểu biết về cấu trúc gene của *Chà vá chân đen* là

điều cần thiết để phát triển các chiến lược bảo tồn hiệu quả nhằm bảo vệ loài đặc biệt này (109).

Kích thước này tương đương và chỉ sai khác rất nhỏ so với các cá thể *Chà vá chân đen* được công bố trước đó, cho thấy tính bảo tồn cao về tổ chức hệ gene ty thể trong nội bộ loài. Các nghiên cứu ADN ty thể gần đây trên linh trưởng đều ghi nhận rằng kích thước và vị trí gene của ADN ty thể rất ổn định, trong khi sự khác biệt chủ yếu tập trung ở mức độ nucleotide và vùng không mã hóa (35,110). Do đó, kết quả của nghiên cứu này củng cố quan điểm rằng ADN ty thể của *Chà vá chân đen* duy trì cấu trúc bảo tồn mạnh, phản ánh áp lực chọn lọc cao lên các thành phần chức năng cốt lõi của hệ gene ty thể.

Thành phần nucleotide của 13 gene mã hóa protein cho thấy hệ gene ty thể của *Chà vá chân đen* Tây Nguyên Việt Nam có xu hướng giàu A+T và nghèo G+C, với giá trị trung bình A+T chiếm trên 60%. Sự giàu A+T trong ADN ty thể thường được giải thích bởi cơ chế đột biến không đối xứng trong quá trình sao chép, trong đó mạch ADN tồn tại ở trạng thái đơn sợi trong thời gian dài dễ tích lũy các đột biến chuyển vị theo hướng tăng A và T. Đồng thời, nhiều nghiên cứu gần đây nhấn mạnh rằng áp lực chọn lọc chức năng cũng đóng vai trò quan trọng trong việc duy trì thiên lệch này, do các codon giàu A/T có thể làm tăng hiệu quả phiên mã và dịch mã của protein ty thể.

Đáng chú ý, sự biến thiên thành phần nucleotide giữa các gene trong hệ gene ty thể *Chà vá chân đen* không đồng đều. Các gene như n*ad6* và *atp8* thể hiện mức độ thiên lệch A/T cao hơn đáng kể so với các gene khác, trong khi các gene có kích thước lớn như c*ox1*, *nad4* và *nad5* có thành phần nucleotide cân bằng hơn. Hiện tượng này đã được ghi nhận trong nhiều nghiên cứu gần đây và được cho là phản ánh sự khác biệt về áp lực chọn lọc chức năng giữa các gene ty thể ((111,112). Kết quả phân tích AT-skew và GC-skew cho thấy *Chà vá chân đen* Tây Nguyên Việt Nam và các cá thể *Chà vá chân đen* khác có mô hình skew gần như tương đồng, với AT-skew âm nhẹ và GC-skew âm rõ rệt, sự nhất quán này phản ánh tính ổn định cao của mô hình phân bố nucleotide trong nội bộ loài, bất chấp sự khác biệt về nguồn gốc địa lý (113).

Các nghiên cứu gần đây chỉ ra rằng skew nucleotide là đặc trưng mang tính "chữ ký tiến hóa" (evolutionary signature) của ADN ty thể, ít biến động hơn so với trình tự nucleotide riêng lẻ (114) (115). Do đó, sự tương đồng về AT/GC-skew giữa

các cá thể *Chà vá chân đen* cho thấy ADN ty thể của loài này chịu áp lực tiến hóa tương đối đồng nhất.

Phân tích thành phần axit amin cho thấy phổ axit amin trong các protein ty thể của *Chà vá chân đen* Tây Nguyên Việt Nam rất tương đồng với các cá thể *Chà vá chân đen* khác. Các axit amin như leucine, isoleucine, serine, threonine và proline chiếm tỷ lệ cao nhất, trong khi methionine, tryptophan và cysteine luôn ở mức thấp. Mô hình này phù hợp với các nghiên cứu gần đây cho thấy protein ty thể có xu hướng giàu các axit amin kỵ nước, đặc biệt là leucine và isoleucine, nhằm duy trì cấu trúc không gian và chức năng của các phức hợp hô hấp nằm trong màng trong ty thể. Sự khác biệt nhỏ về tỷ lệ tương đối axit amin giữa các cá thể *Chà vá* chân đen được xem là biến thiên mức thấp, không đủ để gây ảnh hưởng đáng kể đến chức năng protein.

3.2.5. Đặc điểm hệ gene ty thể hoàn chỉnh của cá thể *Khỉ vàng* (*Macaca mulatta*) Tây Nguyên Việt Nam

Cấu trúc hệ gene ty thể hoàn chỉnh của *Khỉ vàng* cho thấy một cấu trúc tròn gồm 16,566 bp với các 37 gene được biểu diễn bằng các khối màu (Hình 3. 4). Vùng điều khiển xuất hiện màu xám, trong khi các gene *16S* (rRNA lớn) và *12S* (rRNA nhỏ) được chỉ định bằng màu đỏ. 22 gene RNA chuyển vị (*tRNA*) được đánh dấu bằng màu xanh đậm, và 13 gene mã hóa protein (PCGs) được hiển thị bằng màu xanh lá và vàng; chi tiết vị trí và độ dài của từng gene được thể hiện trong (Bảng 3. 19).

Bảng 3. 19. Hệ gene ty thể hoàn chỉnh của *Khỉ vàng* chứa 37 gene, chi tiết về vị trí và chiều dài của chúng.

Tên	**Bắt đầu**	**Kết thúc**	**Mạch**	**Chiều dài**	**Vùng lặp**
trnF(ttc)	248	319	H	72	0
rrnS	320	1266	H	947	0
trnV(gta)	1267	1335	H	69	0
rrnL	1334	2894	H	1561	0
trnL2(tta)	2895	2969	H	75	2
nad1	2972	3922	H	951	4
trnI(atc)	3927	3995	H	69	0
trnQ(caa)	3993	4064	L	72	1

trnM(atg)	4066	4133	H	68	0
nad2	4134	5165	H	1032	10
trnW(tga)	5176	5242	H	67	7
tRNA(gca)	5250	5318	L	69	1
trnN(aac)	5320	5392	L	73	32
trnC(tgc)	5425	5493	L	69	0
trnY(tac)	5494	5556	L	63	4
cox1	5561	7099	H	1539	2
trnS2(tca)	7102	7170	L	69	3
trnD(gac)	7174	7241	H	68	1
cox2	7243	7902	H	660	97
trnK(aaa)	8000	8063	H	64	1
atp8	8065	8259	H	195	0
atp6	8226	8900	H	675	5
cox3	8906	9688	H	783	1
trnG(gga)	9690	9757	H	68	0
nad3	9758	10102	H	345	1
trnR(cga)	10104	10168	H	65	0
nad4l	10169	10462	H	294	0
nad4	10459	11829	H	1371	7
trnH(cac)	11837	11905	H	69	0
trnS1(agc)	11906	11964	H	59	0
trnL1(cta)	11965	12035	H	71	6
nad5	12042	13841	H	1800	12
nad6	13854	14375	L	522	0
trnE(gaa)	14376	14444	L	69	4
Cytb	14449	15585	H	1137	4
trnT(aca)	15590	15653	H	64	1
trnP(cca)	15655	15722	L	68	206

Bảng 3. 20. Tỷ lệ % AT trên các đoạn gene của các loài linh trưởng

	Số gia nhập	Gene mã hóa protein (PCGs)		*16S* (*rrnL*)		12S (*rrnS*)	
		Chiều dài	AT (%)	Chiều dài (bp)	AT (%)	Chiều dài (bp)	AT (%)	Chiều dài (bp)	AT (%)
Khỉ vàng Tây Nguyên Việt Nam		16566	56.9	11304	57	1561	57.4	947	55.2
Khỉ vàng Myanamar	JQ_821843	16,564	56.8	11,310	56	1560	57.4	947	55.6
Khỉ đuôi lợn	KP_330231	17,050	57.2	11,310	57	1562	57.8	947	54.6
Khỉ vòi	DQ_355298	16,570	60.9	11,298	61	1568	59.9	949	57.5
Chà vá chân đen	MH_064177	16,536	61.4	11,304	62	1563	59.7	949	58.8
Voọc sọc đầu đen	KU_899140	16,548	61.9	11,277	62	1563	61	935	58.3
Voọc mũi hếch vàng	KM_504390	16,552	61.5	11,292	62	722	59.6	949	58.2
Voọc mũi hếch xám	JQ_821836	16,553	61.6	11,295	62	1571	60	949	58.3
Voọc mũi hếch đen	JQ_821839	16,550	61.5	11,298	62	1570	59.7	949	58
Voọc mũi hếch Bắc Bộ	JF_293093	16,552	61.6	11,289	62	1572	60.2	949	58.5
Khỉ vàng	JQ_821843	16,564	56.8	11,310	56	1560	57.4	947	55.6
Khỉ đầu chó	NC_001992	16,521	56.3	11,292	56	1572	57.7	947	55.1
Khỉ tai mào trắng	KM_588314	16,499	59.7	11,280	59	1555	60.5	953	56.6
Homo sapiens neanderthalensis	NC_011137	16,565	55.6	11,283	55	1560	57.2	954	54.4
Homo sapiens neanderthalens	OM_062614	16,565	55.7	11,283	55	1560	57.3	954	54.4
Khỉ mũ trán trắng	NC_002763	16,554	60.9	11,268	61	1556	60.2	958	57.6
Chà vá chân đen	MH_064177	16,536	61.4	11,304	62	1563	59.7	949	58.8
Khỉ đuôi lợn	KP_330231	17,050	57.2	11,310	57	1562	57.8	947	54.6
Voọc sọc đầu đen	KU_899140	16,548	61.9	11,277	62	1563	61	935	58.3

Sự sắp xếp này là đặc trưng cho các hệ gene ty thể và phản ánh một tổ chức gọn gàng, nâng cao hiệu quả phiên mã và sao chép. Các gene chủ yếu nằm trên sợi nặng (H strand), với chiều dài khác nhau; ví dụ, gene *nad5* là dài nhất với 1800 cặp bazơ, trong khi *trnK* là ngắn nhất với 64 cặp bazơ. Sự biến đổi trong chiều dài gene

này gợi ý về các vai trò chức năng và áp lực tiến hóa khác nhau, đặc biệt trong chuyển hóa năng lượng và tổng hợp protein, điều này rất quan trọng cho khả năng thích nghi của loài với các môi trường đa dạng (112).

Phân tích các vùng giữa các gene trong hệ gene ty thể hoàn chỉnh của *Khỉ vàng* cho thấy tổng cộng 366 nucleotide giữa các gene qua 36 khoảng cách giữa gene. Đáng chú ý, vùng giữa gene *trnF* và *trnP* dài nhất là 206 nucleotide, trong khi vùng ngắn nhất là giữa *trnL2* và *nad1*, chỉ có 2 nucleotide. Những phát hiện này gợi ý về sự biến đổi trong tiềm năng điều hòa của các vùng giữa gene, với các khoảng cách dài hơn có thể chỉ ra các khu vực cho các yếu tố điều hòa, trong khi các khoảng cách ngắn hơn phản ánh một tổ chức hệ gene gọn gàng hơn. Sự hiện diện của các nucleotide giữa các gene là rất quan trọng để hiểu biết về điều hòa gene và chức năng ty thể, cung cấp cái nhìn sâu sắc về các thích ứng tiến hóa của loài này (116).

3.2.5.1. Thành phần nucleotide

Bảng 3. 21. Thành phần các nucleotide giữa các loài linh trưởng.

	Số gia nhập	T(U)	C	A	G	Tổng	AT-Skew	GC-Skew
Khỉ vàng Tây Nguyên Việt Nam		30.9	26.1	30.5	12.5	868.6	−0.00605	−0.35278
Khỉ vàng Myanmar	JQ_821843	26.6	31	29.8	12.6	870	0.056444	−0.42052
Khỉ đuôi lợn	KM_360179	26.4	31	30	12.5	870.2	0.063577	−0.4244
Papio hamadryas	NC_001992	26.2	31.2	29.7	12.9	868.6	0.061768	−0.41663
Callithrix jacchus	KM_588314	28.3	27.8	31.1	12.9	867.7	0.046296	−0.36736
Homo sapiens neanderthalensis	OM_062614	26.2	31.9	28.8	13.1	867.8	0.046422	−0.41631
Cebus albifrons	NC_002763	29.9	26.7	31.3	12.1	866.8	0.02306	−0.37663

Thành phần nucleotide của hệ gene ty thể của *Khỉ vàng* Tây Nguyên Việt Nam cho thấy tỷ lệ AT là 56.9%. Hệ gene ty thể hoàn chỉnh dài 16,566 cặp base, trong đó vùng mã hóa protein chiếm 11,304 cặp base. Trong hệ gene này, tỷ lệ AT trong các gene mã hóa protein (PCGs) là 57%, trong khi RNA ribosome lớn (*rrnL*) và RNA ribosome nhỏ (*rrnS*) có tỷ lệ AT lần lượt là 57.4% và 55.2% (Bảng 3. 21).

Đáng chú ý, *Khỉ vàng* Việt Nam và *Khỉ vàng* Myanmar có các đặc điểm hệ gene ty thể rất giống nhau, với chỉ một số khác biệt nhỏ. *Khỉ vàng* Việt Nam có chiều dài hệ gene là 16,566 bp và tỷ lệ AT là 56.9%, trong khi *Khỉ vàng* Myanmar ngắn hơn một chút với 16,564 bp và tỷ lệ AT là 56.8%. Các cá thể khỉ của chi *macaca* khác như *Macaca arctoides* cũng có chiều dài hơi ngắn hơn với tỷ lệ AT là 56.7%. Tỷ lệ

AT rất quan trọng để hiểu các động lực tiến hóa vì nó có thể ảnh hưởng đến sự ổn định và chức năng của ADN ty thể.

Phân tích thành phần nucleotide cho thấy các gene mã hóa protein (PCGs) trong *Khỉ vàng* có tỷ lệ AT hơi cao hơn là 57%. Ngược lại, các gene RNA ribosome lớn (*rrnL*) và RNA ribosome nhỏ (*rrnS*) cho thấy tỷ lệ AT lần lượt là 57.4% và 55.2%. Sự biến đổi trong tỷ lệ AT giữa các vùng hệ gene khác nhau nhấn mạnh tầm quan trọng chức năng của thành phần nucleotide liên quan đến biểu hiện gene và các quá trình chuyển hóa. Các nghiên cứu so sánh với các loài linh trưởng khác, đặc biệt là *Pygathrix* và Presbytis, cho thấy sự gia tăng đáng kể trong tỷ lệ AT, làm nổi bật sự khác biệt di truyền giữa các loài linh trưởng và cung cấp cái nhìn về các thích ứng tiến hóa hình thành các hệ gene ty thể qua các dòng họ khác nhau (Bergey et al. 2020) (101).

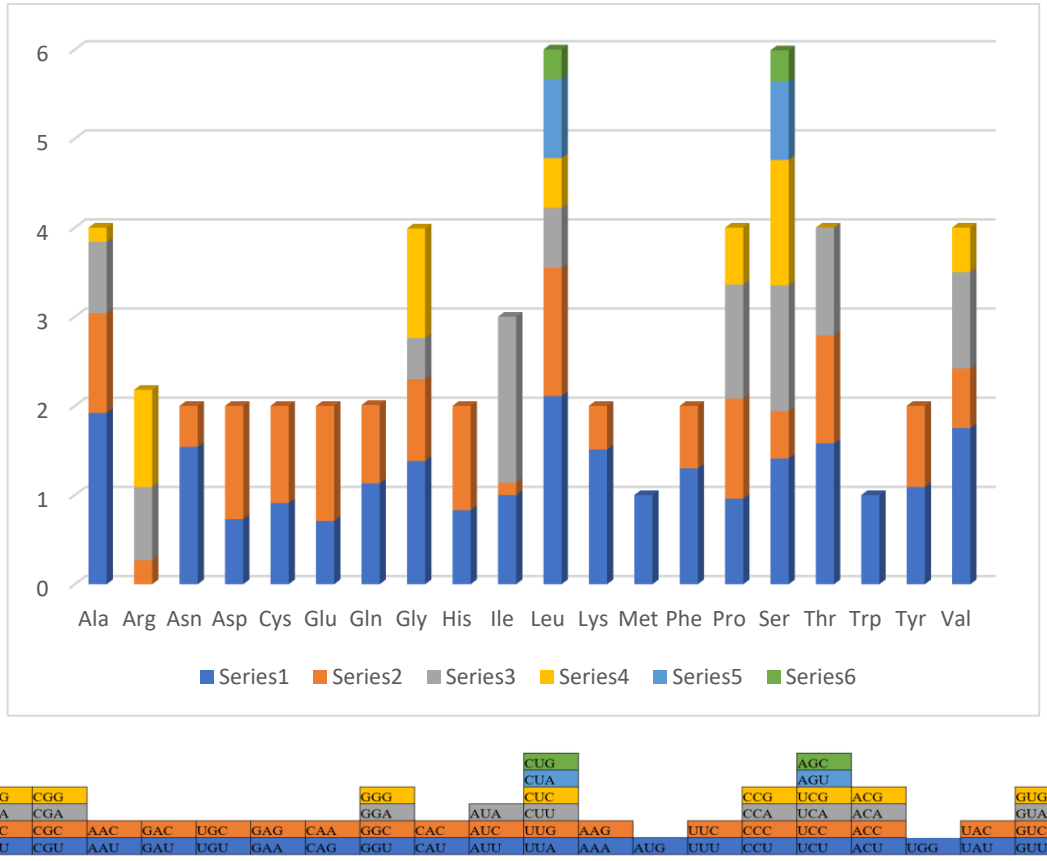
Cả *Khỉ vàng* Việt Nam và *Khỉ vàng* Myanmar đều có chiều dài chuỗi mã hóa protein tương đương, với biến thể Việt Nam đo được 11,304 bp và biến thể Myanmar là 11,310 bp. Chiều dài của các gene RNA ribosomal là giống nhau ở cả hai cá thể, với *rrnL* là 1,561 bp và *rrnS* là 947 bp. Những sự tương đồng này gợi ý một mối quan hệ di truyền gần gũi giữa hai cá thể, có thể chỉ ra tổ tiên chung gần đây và các áp lực tiến hóa tương tự.

Các giá trị GC-skew và AT-skew của các gene mã hóa protein (PCGs) ở các loài linh trưởng khác nhau, bao gồm *Khỉ vàng* Việt Nam cung cấp cái nhìn sâu sắc về thành phần nucleotide và các động lực tiến hóa tiềm năng của hệ gene ty thể; GC-skew được tính là (G - C) / (G + C), trong khi AT-skew được tính là (A - T) / (A + T). Các chỉ số này

Đối với *Khỉ vàng*, GC-skew có giá trị là -0.35278, cho thấy sự chiếm ưu thế của A và T so với G và C trong hệ gene ty thể. GC-skew âm này gợi ý một thiên lệch có thể ảnh hưởng đến các quá trình sao chép và sửa chữa ADN ty thể, cũng như sự ổn định tổng thể của nó (Sahyoun et al. 2014). Ngược lại, AT-skew là -0.00605, gần bằng 0, cho thấy việc sử dụng nucleotide A và T khá cân bằng, điều này có thể phản ánh một thích ứng chức năng trong bối cảnh chuyển hóa năng lượng và hiệu quả của ty thể.

Dữ liệu so sánh từ các cá thể khỉ của chi *macaca* khác, chẳng hạn như *KKhỉ vàng* Myanmar và *khỉ đuôi lợn*, cho thấy các xu hướng tương tự trong các giá trị thiên lệch nucleotide, củng cố ý tưởng về các mô hình tiến hóa bảo tồn trong chi này. Các phân tích thiên lệch này rất quan trọng để hiểu các áp lực tiến hóa tác động lên các hệ gene ty thể, vì chúng có thể tiết lộ những cái nhìn về chức năng gene, động lực sao chép và các chiến lược thích ứng tiềm năng trong các môi trường khác nhau.

3.2.15.2. Các gene mã hóa protein



Hình 3. 22. Sử Dụng Codon Đồng Nghĩa Tương Đối (RSCU) của các gene mã hóa protein ty thể trong *Khỉ vàng* Việt Nam.

Trong bối cảnh này, các axit amin leucine (*Leu*) và serine (*Ser*) thường thể hiện các giá trị RSCU cao, gợi ý rằng các codon tương ứng được ưa chuộng trong các gene ty thể của loài này. Sự ưa chuộng cho một số codon nhất định có thể được quy cho nhiều yếu tố, bao gồm sự sẵn có của các *tRNA* tương ứng và hiệu quả tổng thể của cơ chế dịch mã, điều này rất quan trọng cho việc tổng hợp hiệu quả các protein ty thể.

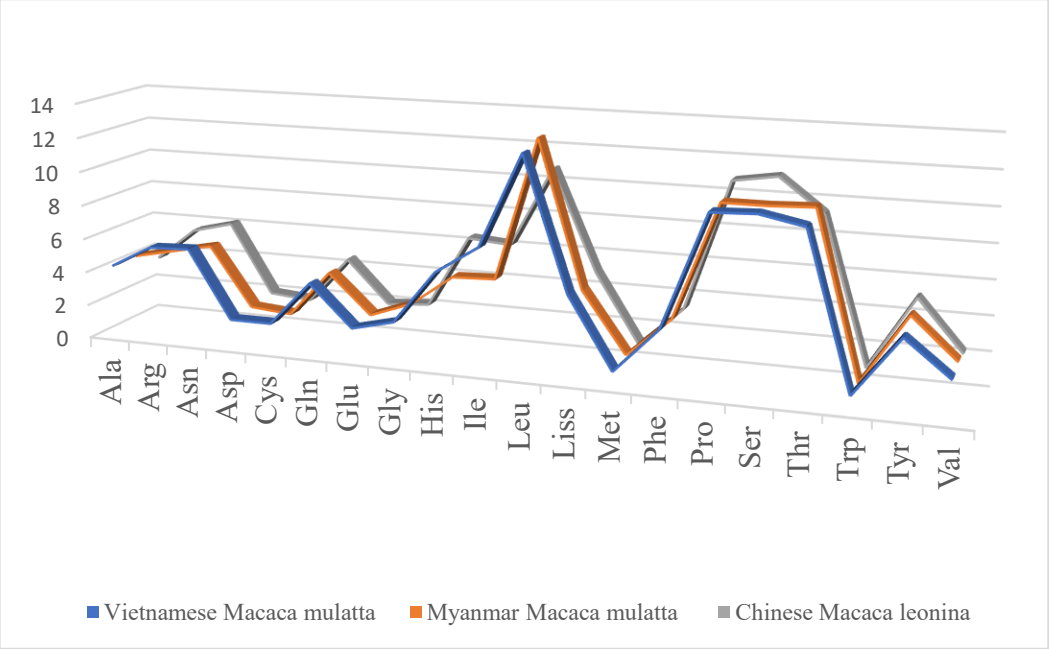
Sự ưu tiên sử dụng codon quan sát được là quan trọng vì nó có thể ảnh hưởng đến hiệu quả của biểu hiện gene và tổng hợp protein, từ đó tác động đến chức năng ty thể và chuyển hóa năng lượng. Hiểu biết về những mô hình sử dụng codon này là rất cần thiết, vì chúng không chỉ phản ánh các áp lực tiến hóa đã hình thành hệ gene ty thể mà còn cung cấp cái nhìn về cách mà những linh trưởng này đã thích ứng với các sinh cảnh sinh thái của chúng (117).

Các phân tích so sánh về các mô hình RSCU giữa các cá thể khỉ của chi *macaca* khác nhau có thể làm rõ hơn các mối quan hệ tiến hóa và các thích ứng chức năng. Bằng cách kiểm tra cách sử dụng codon thay đổi giữa các loài liên quan, các nhà nghiên cứu có thể xác định các xu hướng tiến hóa bảo tồn và phân kỳ, nhấn mạnh tác

động của các áp lực môi trường cụ thể lên cấu trúc và chức năng của các gene ty (118). Những so sánh này có thể cho thấy rằng một số codon được ưa chuộng phổ biến giữa các loài, cho thấy áp lực chọn lọc mạnh mẽ, trong khi những codon khác có thể thể hiện sự biến đổi, cho thấy các thích ứng với các bối cảnh sinh thái cụ thể.

Hơn nữa, những kết quả này có những ý nghĩa đối với các chiến lược bảo tồn. Việc xác định các mô hình sử dụng codon trong hệ gene ty thể nhấn mạnh tầm quan trọng của việc duy trì sự đa dạng di truyền trong các cá thể *Khỉ vàng*. Sự khác biệt di truyền là rất quan trọng cho khả năng thích nghi và sự bền bỉ của loài, đặc biệt là trước những thay đổi môi trường và mất mát môi trường sống (119,120). Bằng cách hiểu biết về cấu trúc di truyền và các thích ứng tiến hóa của *Macaca mulatta*, các nỗ lực bảo tồn có thể được điều chỉnh tốt hơn để bảo vệ không chỉ loài mà còn cả sự biến đổi di truyền của nó, điều này là cần thiết cho sự sống còn lâu dài của nó.

3.2.5.2. Các gene mã hóa Protein

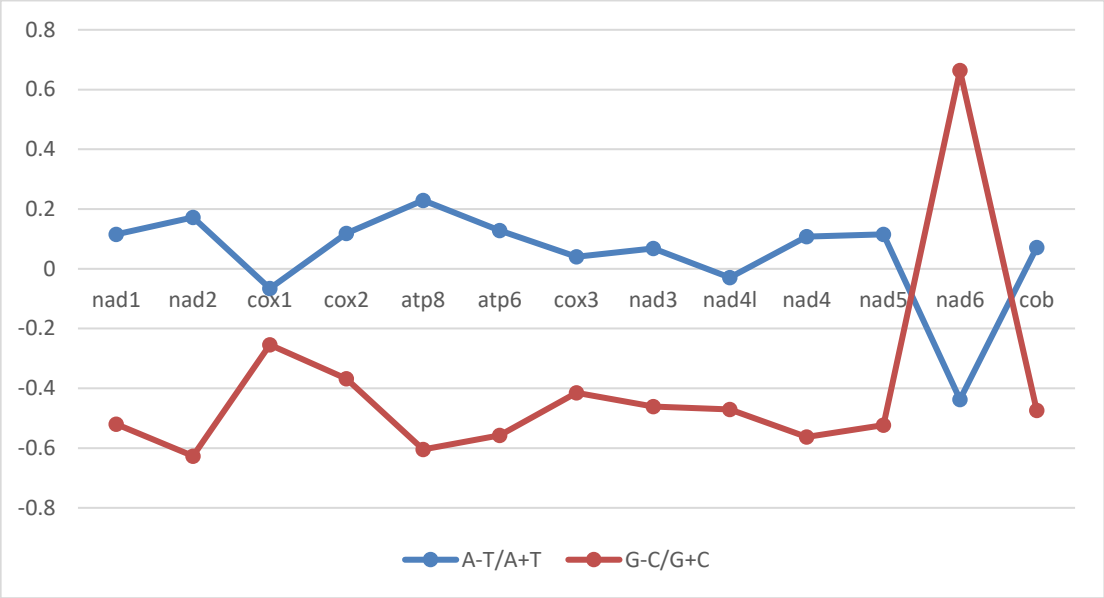


Hình 3. 23. Tỷ lệ phần trăm tần suất thành phần axit amin trong toàn bộ hệ gene ty thể của các cá thể khỉ của chi *macaca* : *Khỉ vàng* Việt Nam; *Khỉ vàng* Myanmar; *Khỉ vàng* Trung Quốc.

Tỷ lệ phần trăm tần suất thành phần axit amin trong hệ gene ty thể của *Khỉ vàng* Việt Nam (Hình 3. 23) cung cấp cái nhìn sâu sắc về sự phân bố của các axit amin trong các protein ty thể, điều này rất quan trọng cho nhiều chức năng chuyển hóa và sản xuất năng lượng. Các axit amin như methionine và tryptophan được đại diện với tần suất thấp, điều này có thể ảnh hưởng đến các động lực chức năng tổng thể của các protein ty thể. Ngược lại, kết quả cho thấy rằng một số axit amin chiếm ưu thế trong thành phần của các protein ty thể được mã hóa trong hệ gene. Đặc biệt,

leucine và serine nổi bật như những axit amin phong phú nhất, phản ánh vai trò thiết yếu của chúng trong chức năng ty thể và tổng hợp protein. Những phát hiện này phù hợp với nghiên cứu trước đây nhấn mạnh sự phổ biến của các axit amin này trong các protein ty thể ở nhiều loài khác nhau (65,121)

Phân tích thành phần axit amin không chỉ cung cấp cái nhìn tổng quan về khả năng mã hóa protein của hệ gene ty thể trong *Khỉ vàng* mà còn cho phép so sánh với các loài linh trưởng khác. Những so sánh như vậy có thể làm sáng tỏ các thích ứng tiến hóa và tầm quan trọng chức năng của các protein ty thể qua các dòng khác nhau (Spuhler 1988). Sự kiểm tra chi tiết này nâng cao hiểu biết của chúng ta về các động lực tiến hóa và vai trò sinh hóa của các protein ty thể trong *Khỉ vàng* Việt Nam.



Hình 3. 24. Giá trị AT- và GC-skew của 13 gene mã hóa protein trong hệ gene ty thể của *Khỉ vàng* Tây Nguyên Việt Nam.

Giá trị AT- và GC-skew cho 13 gene mã hóa protein (PCGs) trong hệ gene ty thể của *Khỉ vàng* Tây Nguyên Việt Nam như *nad1*, nad2, cox1, và atp6, có giá trị dương, cho thấy sự phong phú tương đối của adenine (A) so với thymine (T). Xu hướng này gợi ý về việc tối ưu hóa hiệu suất phiên mã, điều này rất quan trọng cho các yêu cầu chuyển hóa cao của chức năng ty thể. Ngược lại, GC-skew âm nhất quán trong các gene này cho thấy sự chiếm ưu thế của cytosine (C) so với guanine (G), điều này có thể nâng cao sự ổn định của ADN ty thể và ảnh hưởng đến động lực biểu hiện gene.

Các giá trị skew quan sát được (Hình 3. 24) phản ánh các thích ứng tiến hóa của những gene này đối với vai trò của chúng trong chuyển hóa năng lượng và hiệu suất hô hấp. Ví dụ, các gene như nad5 và cox2 cũng cho thấy AT-skew dương, củng cố tầm quan trọng chức năng của chúng trong chuỗi vận chuyển điện tử và tổng hợp

ATP. Sự hiện diện của GC-skew âm trong các PCGs gợi ý một cấu trúc hệ gene bảo tồn hỗ trợ chức năng ty thể hiệu quả. Tổng thể, những mô hình này không chỉ làm nổi bật các động lực tiến hóa phức tạp trong *Khỉ vàng* mà còn nhấn mạnh tầm quan trọng của các gene này trong việc duy trì khả năng thích nghi chuyển hóa trong các môi trường đa dạng.

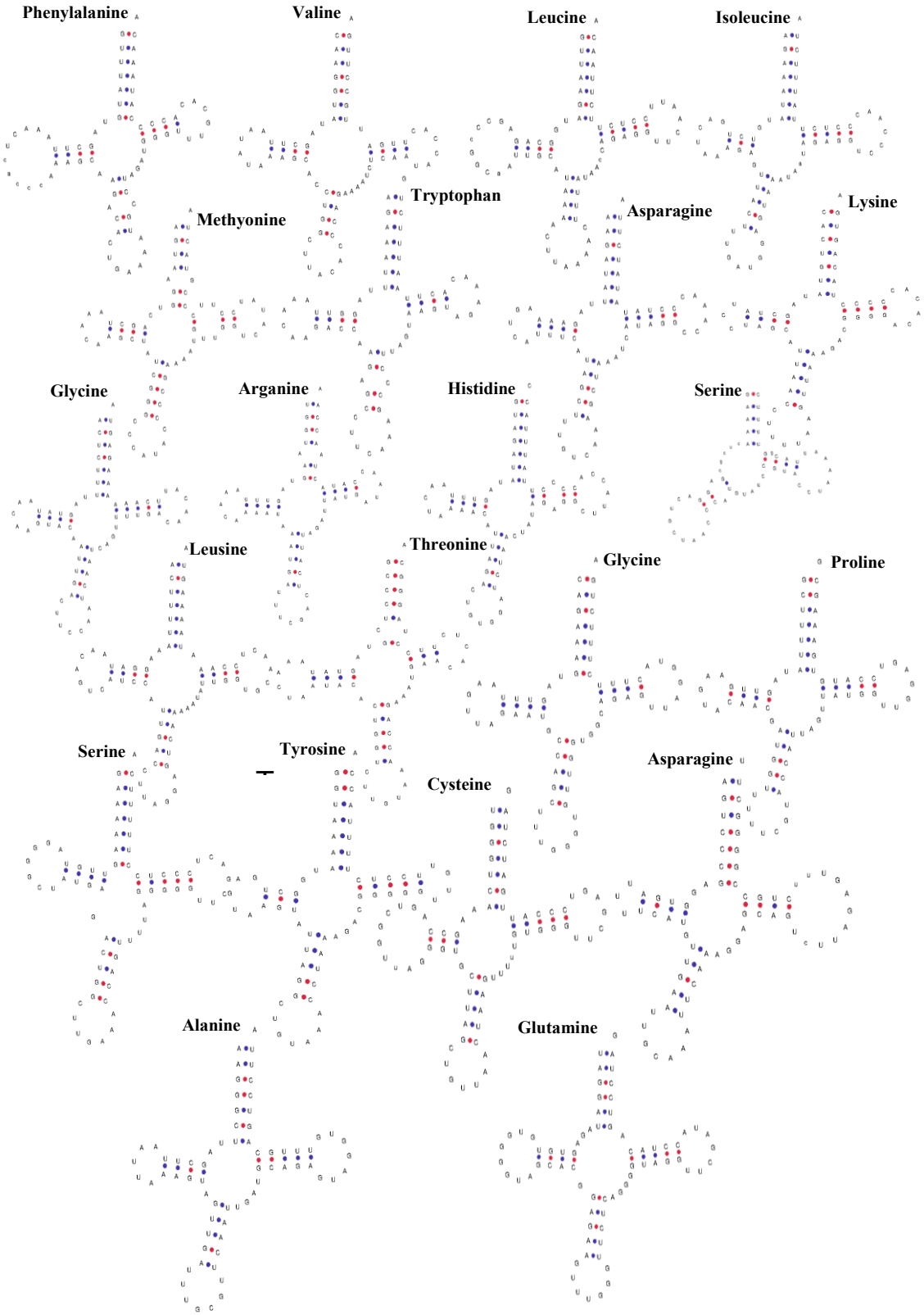
3.2.5.4. *tRNA* và *rRNA*

Bảng 3. 22. Các cặp base không khớp được tìm thấy trong các gene *tRNA* củak*Khỉ vàng* Việt Nam.

tRNA	**Các Cặp Base Không Khớp**	**Thân**	**Tần suất**
Phenylalanine GAA	A-C	AC	1
Tryptophan TCA	A-C	AC	1
Methionine CAT	A-G	AA	1
	U-U	T-arm	2
Arginine TCG	A-A	AA	1
Histidine	A-C	T-arm	1
Threonine TGT	C-U	AA	1
Threonine TGT	C-U	AA	1

Cặp base không khớp được tìm thấy trong các gene *tRNA* của *Khỉ vàng* Việt Nam (Bảng 3. 22) rất quan trọng để hiểu biết về tính toàn vẹn cấu trúc và hiệu suất chức năng của các phân tử *tRNA*. Ví dụ, các cặp không khớp như A-C trong *tRNA* cho phenylalanine và tryptophan, cũng như A-G trong methionine, chỉ ra sự sai lệch so với cặp base Watson-Crick tiêu chuẩn. Những sai lệch này có thể ảnh hưởng đến sự ổn định và gấp nếp của *tRNA*, có khả năng ảnh hưởng đến khả năng của nó trong việc vận chuyển chính xác các axit amin trong quá trình tổng hợp protein.

Sự hiện diện của các cặp base không khớp trong các vùng thân của *tRNA*, đặc biệt là trong vùng tiếp nhận axit amin (AA) và T-arm, gợi ý về các thích ứng tiến hóa có thể nâng cao tính linh hoạt hoặc khả năng chức năng của những *tRNA* này trong bối cảnh ty thể. Hiểu biết về những biến thể này cung cấp cái nhìn sâu sắc về các áp lực tiến hóa đã hình thành hệ gene ty thể của *Macaca mulatta*, nhấn mạnh cách mà những linh trưởng này đã điều chỉnh cấu trúc *tRNA* của họ để tối ưu hóa chức năng ty thể và hiệu suất tổng hợp protein.



Hình 3. 25. Cấu trúc thứ cấp của 22 gene *tRNA* của *Khỉ vàng* Việt Nam. Màu đỏ là kết nối GC, màu xanh là kết nối AU.

Các phân tử *tRNA* được minh họa với các thân và vòng tương ứng trong ấu trúc thứ cấp của 22 gene *tRNA* trong hệ gene ty thể của *Khỉ vàng* Việt Nam (Hình 3. 25) cho thấy sự gấp nếp tinh vi cần thiết cho chức năng của chúng. Trong hình, các kết nối được mã hóa màu: các cặp GC được chỉ định bằng màu đỏ, trong khi các cặp AU được hiển thị bằng màu xanh. Mã màu này làm nổi bật các tương tác cặp base quan trọng cho tính ổn định và chức năng của từng *tRNA*.

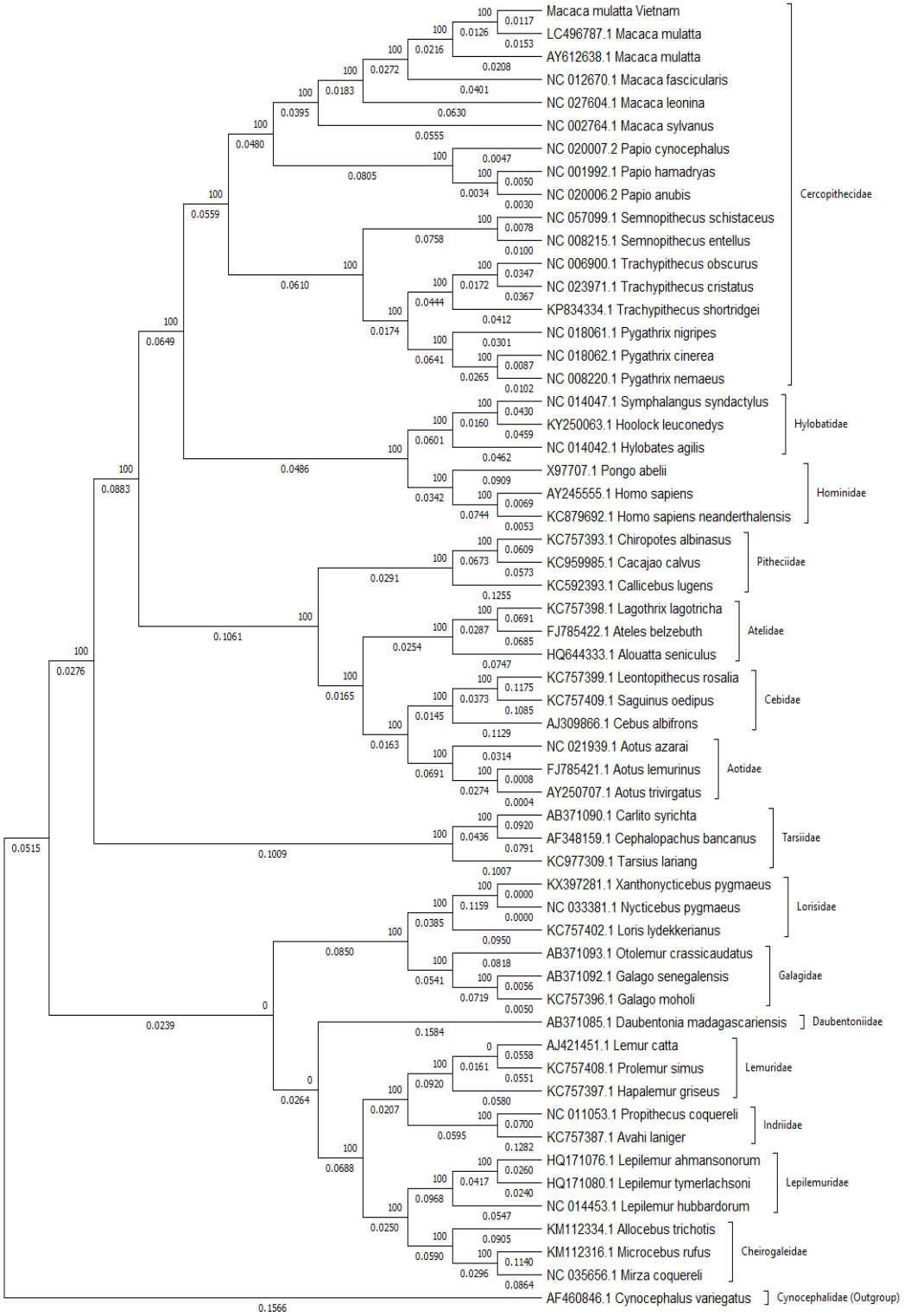
Cấu trúc thứ cấp của các gene *tRNA* rất quan trọng cho vai trò của chúng trong dịch mã, vì những phân tử này chịu trách nhiệm mang các axit amin đến ribosome trong quá trình tổng hợp protein. Sự hiện diện của các cặp base GC và AU ảnh hưởng đến sự ổn định của cấu trúc *tRNA*, với các cặp GC thường cung cấp liên kết mạnh hơn nhờ vào ba liên kết hydro so với hai liên kết hydro được hình thành bởi các cặp AU (Hình 3. 25).

3.2.5.5. Vùng lặp và vùng liên gene

Phân tích cấu trúc hệ gene ty thể hoàn chỉnh của *Khỉ vàng* cho thấy sự tồn tại phổ biến của các vùng liên gene ngắn xen kẽ giữa các gene mã hóa protein, *tRNA* và rRNA, phản ánh mức độ nén cao và tổ chức chặt chẽ của ADN ty thể (Bảng 3. 19). Phần lớn các gene được sắp xếp liền kề hoặc chỉ cách nhau từ 1–10 bp, điển hình như giữa *trnL2(tta)* và *nad1* (2 bp), *nad1* và *trnI(atc)* (4 bp), *nad2* và *trnW(tga)* (10 bp), cũng như giữa nhiều cặp gene *tRNA* liền kề. Ngoài ra, một số vùng liên gene có kích thước trung bình hơn cũng được ghi nhận, đáng chú ý là khoảng cách 32 bp giữa *trnN(aac)* và *trnC(tgc)*, 97 bp giữa *cox2* và *trnK(aaa)*, và đặc biệt là vùng liên gene lớn 206 bp giữa *trnP(cca)* và gene kế tiếp. Sự hiện diện của các vùng liên gene với độ dài khác nhau cho thấy hệ gene ty thể của *Khỉ vàng*, mặc dù mang đặc trưng nén cao, vẫn duy trì một số đoạn không mã hóa có thể liên quan đến điều hòa phiên mã *RNA* ty thể.

Bên cạnh các vùng liên gene, các vùng lặp gene (chồng chéo) cũng được ghi nhận với tần suất đáng kể trong hệ gene ty thể của *Khỉ vàng*. Một số cặp gene có sự chồng chéo ngắn từ 1–7 bp, điển hình như *trnQ(caa)*/*trnI(atc)* (1 bp), *trnW(tga)*/*nad2* (7 bp), *cox1*/*trnS2(tca)* (2 bp), *atp6*/*atp8* (5 bp), và *nad4*/*nad4l* (7 bp). Đáng chú ý, gene *nad5* có vùng chồng chéo 12 bp với gene lân cận, phản ánh mức độ nén cao của các gene mã hóa protein trong hệ gene ty thể. Các vùng lặp này đều có kích thước nhỏ và không làm gián đoạn khung đọc của các gene, cho thấy sự thích nghi của ADN ty thể theo hướng tối ưu hóa không gian di truyền trong khi vẫn bảo đảm tính toàn vẹn chức năng của các gene thiết yếu.

3.2.5.6. Cây phân loài của *Khỉ vàng* với các loài linh trưởng khác



Hình 3. 26. Cây phát sinh loài của *Khỉ vàng* với các loài linh trưởng khác

Cây phân loài cho thấy *Khỉ vàng* Việt Nam có mối quan hệ gần gũi với các loài khỉ khác, như *Macaca fascicularis (khỉ đuôi dài)* và *Khỉ vàng* từ các khu vực khác nhau. Sự tập hợp gần gũi này gợi ý rằng các loài này chia sẻ một tổ tiên chung gần đây, cho thấy rằng chúng đã phân kỳ tương đối gần đây trong các khía cạnh tiến hóa. Những sự tương đồng di truyền này phản ánh các thích ứng của chúng với các môi trường và lịch sử sống tương tự, nhấn mạnh các kết nối tiến hóa trong chi *Macaca* .

Ngược lại, khi so sánh *Khỉ vàng* Việt Nam với các linh trưởng có quan hệ xa hơn, chẳng hạn như các thành viên trong họ Hominidae (ví dụ, con người và tinh tinh), sự khác biệt di truyền đáng kể rõ ràng. Các nhánh dài hơn dẫn đến những loài này cho thấy một lịch sử tiến hóa lâu dài và sự phân hóa di truyền lớn hơn. Sự khác biệt này có thể được quy cho các yếu tố như môi trường sống, hành vi và các thích ứng sinh thái đã phát triển qua hàng triệu năm. Mối quan hệ này làm nổi bật các con đường tiến hóa đã dẫn đến những đặc điểm riêng biệt quan sát được trong chi Macaca so với các họ linh trưởng khác.

3.2.5.7. Thảo luận

Hệ gene ty thể của *Khỉ vàng* Việt Nam cho thấy mức độ tương đồng cao với cá thể *Khỉ vàng* khu vực Myanmar, phản ánh mối quan hệ tiến hóa gần và lịch sử phân bố liên tục của loài trong khu vực Đông Nam Á lục địa. Các sai khác nucleotide quan sát được, đặc biệt trong một số gene mã hóa protein, có thể liên quan đến thích ứng sinh thái địa phương (122).

Nghiên cứu làm nổi bật các mối quan hệ phả hệ trong chi *macaca* . Các nghiên cứu trước đây đã chỉ ra sự phân kỳ gene di truyền đáng kể giữa các cá thể khỉ của chi *macaca* , cho thấy các động lực tiến hóa phức tạp được hình thành bởi các áp lực môi trường và thích ứng với môi trường sống. Mối quan hệ di truyền gần gũi giữa cá thể Việt Nam và Myanmar, với chỉ 704 khác biệt nucleotide, gợi ý về tổ tiên chung gần đây và các áp lực tiến hóa tương tự. Những khác biệt quan sát được ở các gene mã hóa protein, đặc biệt là trong gene *nad1*, đóng vai trò quan trọng trong chuỗi vận chuyển điện tử, có thể phản ánh các thích ứng với các sinh cảnh sinh thái cụ thể (123). Những biến thể này rất quan trọng để hiểu cách mà những linh trưởng này đã tiến hóa để phát triển trong các môi trường đa dạng, bao gồm cả các khu vực đô thị.

Phân tích thành phần nucleotide cho thấy tỷ lệ AT là 56.9% trong hệ gene ty thể, với các vùng mã hóa protein có tỷ lệ AT hơi cao hơn (57%). Sự ưu tiên sử dụngnucleotide này có thể ảnh hưởng đến sự ổn định và chức năng của ADN ty thể, tác động đến quá trình sao chép và phiên mã (124). Phân tích Sử Dụng Codon Tương Đương Tương Đối (RSCU) cho thấy sự ưa thích đối với một số codon, đặc biệt là đối với leucine và serine, có thể tăng cường hiệu suất dịch mã. Sự ưa thích này rất quan

trọng để tối ưu hóa quá trình tổng hợp protein nhằm đáp ứng các nhu cầu chuyển hóa đa dạng, đặc biệt trong các môi trường đô thị có biến động (117).

Phân tích các vùng giữa gene cho thấy sự biến đổi trong tiềm năng điều chỉnh, với các vùng giữa gene dài hơn có thể chỉ ra các khu vực cho các yếu tố điều chỉnh. Sự quan sát 366 nucleotide giữa gene nhấn mạnh tầm quan trọng của những vùng này trong việc điều chỉnh gene và chức năng ty thể. Khả năng điều chỉnh này đặc biệt quan trọng để hiểu cách mà *Khỉ vàng* thích ứng với các thách thức môi trường, vì các yếu tố điều chỉnh đóng vai trò thiết yếu trong biểu hiện gene ty thể và các phản ứng chuyển hóa (116).

Việc xác định các cặp base không khớp trong các gene *tRNA* cho thấy những thích ứng cấu trúc có thể nâng cao sự linh hoạt và khả năng chức năng của *tRNA* trong bối cảnh ty thể. Các cặp không khớp, như A-C và A-G, cho thấy sự sai lệch so với cặp base tiêu chuẩn, điều này có thể ảnh hưởng đến sự ổn định của các phân tử *tRNA* và khả năng của chúng trong việc vận chuyển chính xác các axit amin trong quá trình tổng hợp protein. Những biến thể này gợi ý về các áp lực tiến hóa đã hình thành cấu trúc *tRNA* để tối ưu hóa chức năng ty thể (65).

Nghiên cứu này đã thành công trong việc mô tả hệ gene ty thể hoàn chỉnh của *Khỉ vàng* Việt Nam, lấp đầy một khoảng trống quan trọng trong hiểu biết của chúng ta về cấu trúc di truyền và các thích ứng tiến hóa của loài này. Những hiểu biết thu được từ thành phần nucleotide, mối quan hệ phả hệ và các cơ chế điều chỉnh đóng góp vào kiến thức rộng hơn về di truyền và tiến hóa của linh trưởng. Hơn nữa, nghiên cứu này nhấn mạnh tầm quan trọng của đa dạng di truyền trong các chiến lược bảo tồn, vì việc duy trì sự biến đổi di truyền là rất cần thiết cho khả năng thích ứng và sự bền bỉ của các cá thể *Khỉ vàng* trước những thay đổi môi trường. Các nghiên cứu trong tương lai nên tập trung vào việc điều tra các tác động chức năng của những biến thể di truyền quan sát được và khám phá tiềm năng của các nỗ lực bảo tồn nhằm tăng cường đa dạng di truyền trong loài này.

3.3. Phân tích sai khác trình tự ty thể phục vụ nghiên cứu phát sinh loài

Trong nghiên cứu tiến hóa phân tử và bảo tồn sinh học, việc sử dụng các trình tự gene ty thể như *16S rRNA*, *cytb*, *D-loop* và toàn bộ hệ gene ty thể đã trở thành công cụ then chốt để xác định mối quan hệ phát sinh loài giữa các nhóm động vật, đặc biệt là linh trưởng (32). Việt Nam là nơi có đa dạng sinh học cao với nhiều loài linh trưởng đặc hữu có nguy cơ tuyệt chủng cao (125). Việc tiến hành xây dựng và phân tích cây phát sinh loài từ bốn vùng gene khác nhau của ADN ty thể – bao gồm *16S rRNA, cytb*, vùng *D-loop* và toàn bộ hệ gene ty thể – có thể đánh giá tính nhất quán và giá trị bảo tồn của từng cây. Mỗi vùng gene phản ánh một khía cạnh khác nhau của tiến

hóa và cấu trúc di truyền. Gene *16S rRNA* cung cấp độ phân giải ổn định ở cấp độ loài hoặc chi, với ưu điểm là độ bảo tồn cao và khả năng định danh rõ ràng các nhánh lớn. *Cytb* thể hiện tốc độ tiến hóa tốt hơn cho các nhánh gần, phù hợp với việc xác định mối quan hệ giữa các cá thể trong cùng một loài. Các mẫu linh trưởng tại Tây Nguyên từ các cá thể khác nhau có sự phân tách đáng kể trên cây *cytb*. Vùng *D-loop* có tốc độ tiến hóa nhanh, giúp nhận diện sự phân hóa nội bộ rõ rệt. Tuy nhiên, chính vì biến động cao nên dữ liệu dễ bị nhiễu, đòi hỏi xử lý thận trọng trong quá trình căn chỉnh và phân tích. Toàn bộ hệ gene ty thể kết hợp cả ba loại thông tin trên, mang lại độ phân giải cao nhất và nhất quán với cả ba cây đơn lẻ. Đây là công cụ mạnh mẽ để tái tạo lịch sử phát sinh loài toàn diện nhất.

3.3.1. Phân tích sai khác trình tự gene *16S rRNA* và mối quan hệ phát sinh loài của các cá thể linh trưởng khu vực Tây Nguyên Việt Nam

Bảng 3. 23. Các vị trí khác biệt trên trình tự gene *16S rRNA* của cá thể *Chà vá chân đen* Tây Nguyên Việt Nam và cá thể *Chà vá chân đen* khác

	2	3	6	10	13	14	16	17	18	19	20	21	22	25	26	27	30	23	33	39	41	45	46	47	48	52
Chà vá chân đen Tây Nguyên Việt Nam	T	T	G	G	A	G	G	C	C	C	G	G	T	T	T	G	T	A	A	A	G	T	T	A	T	C
Chà vá chân đen NC018061	C	.	T	A	T	C	A	A	A	A	C	C	C	C	C	A	C	C	T	T	G	G	A	T	A	A
Chà vá chân đen MH_064177	C	C	T	A	T	C	A	-	-	A	C	C	C	C	C	A	C	C	T	T	A	G	A	T	A	A
Chà vá chân đen JQ821840	C	.	T	A	T	C	A	A	A	A	C	C	C	C	C	A	C	C	T	T	G	G	A	T	A	A

	53	54	55	56	57	58	59	61	63	66	67	68	69	70	71	72	75	76	78	82	83	84	85	86	87	88
Chà vá chân đen Tây Nguyên Việt Nam	A	G	A	G	G	T	T	A	C	C	T	C	T	T	C	T	G	C	C	G	T	T	C	A	C	A
Chà vá chân đen NC018061	T	A	T	T	A	A	A	C	T	A	C	A	A	G	T	A	A	A	G	A	G	G	A	G	A	T
Chà vá chân đen MH_064177	T	A	T	T	A	A	A	C	T	A	C	A	A	A	T	A	A	A	G	A	G	G	A	G	A	T
Chà vá chân đen JQ821840	T	A	T	T	A	A	A	C	T	A	C	A	A	G	T	A	A	A	G	A	G	G	A	G	A	T

	89	90	95	97	100	102	103	110	111	112	114	115	117	118	119	120	122	125	151	159	157	171	133	138	144	167
Chà vá chân đen Tây Nguyên Việt Nam	G	T	C	T	A	T	A	T	A	T	C	T	C	C	C	A	C	A	G	C	C	G	G	C	G	T
Chà vá chân đen NC018061	A	G	T	A	C	G	G	A	G	A	A	C	G	T	A	C	T	G	A	T	T	A	G	T	G	T
Chà vá chân đen MH_064177	A	G	T	A	C	G	G	A	G	A	A	C	G	T	A	C	C	A	G	C	C	G	A	C	A	G
Chà vá chân đen JQ821840	A	G	T	A	C	G	G	A	G	A	A	C	G	T	A	C	T	G	A	T	T	A	G	T	G	T

	6	7	7	7	9	1	1	2
	8	5	5	7	8	2	8	2
	4	6	8	6	2	1	9	2
Chà vá chân đen Tây Nguyên Việt Nam	G	A	C	C	T	C	C	T
Chà vá chân đen NC018061	G	G	T	T	T	C	G	C
Chà vá chân đen MH_064177	A	A	C	C	C	T	C	T
Chà vá chân đen JQ821840	G	G	T	T	T	C	G	C

Cá thể *Chà vá chân đen* Tây Nguyên thể hiện nhiều sai khác nucleotide so với các mẫu đối chiếu (NC018061, MH_064177, JQ821840) với gần 60 vị trí biến đổi (Bảng 3. 23). Tại vị trí 1236, cá thể *Chà vá chân đen* Tây Nguyên mang nucleotide T, trong khi cả ba mẫu đối chiếu đều mang C. Ở cụm vị trí 1278–1280, trình tự của cá thể *Chà vá chân đen* Tây Nguyên là G–G–T, khác với kiểu gen đồng nhất C–C–C của các mẫu còn lại. Tại vị trí 1339, cá thể Tây Nguyên mang A, trong khi các mẫu đối chiếu lần lượt mang T (NC018061) và G (MH_064177, JQ821840). Ngoài các sai khác đơn lẻ, nhiều chuỗi đột biến liên tiếp được ghi nhận. Trong đoạn 1508–1513, cá thể Tây Nguyên có trình tự C–C–C–C–C–C, trong khi cả ba mẫu đối chiếu đều mang T–T–T–T–T–T. Tương tự, tại đoạn 1607–1610, cá thể *Chà vá chân đen* Tây Nguyên mang A–T–T–T, khác biệt so với kiểu gen G–C–C–C của các mẫu tham chiếu. Một số vị trí như 1510, 1512 và 1721 cho thấy sự không đồng nhất giữa các mẫu đối chiếu

Sự khác biệt này thể hiện tính khác biệt di truyền giữa cá thể *Chà vá chân đen* Tây Nguyên và các cá thể *Chà vá chân đen* từ các vùng địa lý khác.

Trên cây phát sinh loài dựa trên trình tự gene *16S* (Hình 3. 27), cá thể *Chà vá chân đen* Tây Nguyên nằm ở một nhánh riêng biệt rõ rệt, tách biệt với các mẫu tham chiếu. Việc cá thể nghiên cứu nằm trong một nhánh riêng biệt trên cây phát sinh loài đã khẳng định giá trị của bộ dữ liệu hệ gen ty thể hoàn chỉnh, đóng góp làm nền tảng di truyền cho loài linh trưởng đặc hữu, nguy cấp này

Bảng 3. 24. Các vị trí khác biệt trên trình tự gene *16S rRNA* của cá thể *Chà vá chân nâu* Tây Nguyên Việt Nam và cá thể *Chà vá chân nâu* khác

	4	4	4	6	6	7	8	8	8
	5	0	0	9	1	2	1	0	5	5
	3	1	2	0	1	4	0	4	1	4
Chà vá chân nâu Tây Nguyên Việt Nam	T	C	C	T	C	C	T	C	C	C
Chà vá chân nâu NC008220	C	T	T	C	T	T	C	C	T	C
Chà vá chân nâu JF_293096	C	T	C	C	T	T	C	T	T	T
Chà vá chân nâu DQ355302	C	T	T	C	T	T	C	C	T	C

10 vị trí nucleotide khác biệt trong trình tự gene *16S rRNA* giữa cá thể *Chà vá chân nâu* Tây Nguyên Việt Nam và ba mẫu tham chiếu trên thế giới (NC008220, JF_293096, DQ355302) (Bảng 3. 24), các biến dị chủ yếu là thay thế base T↔C hoặc C↔T. Cá thể *Chà vá chân nâu* Tây Nguyên thể hiện kiểu gene T-C-C-T-C-C-T-C-C-C ở 10 vị trí này, trong khi các mẫu tham chiếu có xu hướng ngược lại (C-T-T-C-T-T-C-T-T-C). Ở vị trí 4450–4452; cá thể *Chà vá chân nâu* Tây Nguyên mang bộ ba T-C-C so với C-T-T (NC008220), C-T-C (JF_293096). Các vị trí 4460–4461; cá thể *Chà vá chân nâu* Tây Nguyên C-C, trong khi mẫu đối chiếu chủ yếu là T-T → thể hiện sự thay đổi đồng loạt; vị trí 4515; cá thể *Chà vá chân nâu* Tây Nguyên mang C trong khi mẫu JF_293096 là T.

Bảng 3. 25. Các vị trí khác biệt trên trình tự gene *16S* của cá thể *Chà vá chân xám* Tây Nguyên Việt Nam và cá thể *Chà vá chân xám* khác

	1	2	3	3	3	3	4	4	4	4	4	5	5	5	5	6	7	8	1 0	1 0	1 2	1 3	1 3	1 3	1 4	1 4
	2	5	0	1	2	9	1	4	6	8	9	0	4	5	6	6	5	9	3	5	8	1	9	9	7	7
	0	6	7	2	9	0	0	0	8	8	9	5	0	0	7	5	0	1	5	1	3	1	3	5	4	5
	5	1	5	9	1	1	8	2	4	2	2	9	8	1	2	7	3	1	4	3	1	9	6	1	3	3
Chà vá chân xám Tây Nguyên Việt Nam	G	A	T	T	G	T	A	C	C	T	C	C	C	G	C	G	T	C	A	T	A	G	T	T	A	T
Chà vá chân xám NC088556	G	A	T	T	G	T	A	C	C	T	C	C	C	G	C	G	T	C	A	T	A	G	T	T	A	T
Chà vá chân xám NC018063	A	G	C	C	A	C	C	T	T	C	T	T	T	A	T	A	C	T	G	C	G	A	C	C	G	C
Chà vá chân xám JQ_821842	A	G	C	C	A	C	C	T	T	C	T	T	T	A	T	A	C	T	G	C	G	A	C	C	G	C

	1	1	1	1	1	1	1	1	1	1	1	1
	4	5	5	5	5	5	5	5	6	6	6	6
	9	0	3	8	8	8	8	9	0	3	4	4
	9	2	7	0	1	1	2	7	6	3	2	6
	2	1	5	9	3	5	3	5	0	0	7	1
Chà vá chân xám Tây Nguyên Việt Nam	T	C	T	C	T	C	G	G	C	T	A	T
Chà vá chân xám NC088556	T	C	T	C	T	C	G	G	C	T	A	T
Chà vá chân xám NC018063	C	T	C	T	C	T	A	A	T	C	G	C
Chà vá chân xám JQ_821842	C	T	C	T	C	T	A	A	T	C	G	C

Cá thể *Chà vá chân xám* Tây Nguyên có các vị trí nucleotide giống với mẫu NC088556, nhưng khác biệt so với hai mẫu khác (NC018063, JQ_821842) (Bảng 3. 25). Các điểm khác biệt lên tới hơn 30 vị trí; sự khác biệt xảy ra tại nhiều vị trí phân bố đều trên toàn trình tự, với các dạng đột biến phổ biến là A↔G, T↔C.

Đáng chú ý, cá thể *Chà vá chân xám* Tây Nguyên có kiểu nucleotide hoàn toàn giống với mẫu NC088556 tại tất cả các vị trí được ghi nhận → cho thấy khả năng hai mẫu này cùng khu vực địa lý hoặc là cùng dòng di truyền.

So với 2 mẫu tham chiếu NC018063 và JQ_821842, cá thể *Chà vá chân xám* Tây Nguyên khác biệt tại nhiều vị trí như vị trí 1103–1105: *Chà vá chân xám* Tây Nguyên có T-G-A, trong khi NC018063 có C-C-C. Vị trí 1510–1513: *Chà vá chân xám* Tây Nguyên mang C-G-T-T, trong khi mẫu khác có T-A-G-C. Các vị trí như 1709, 1721, 1733, 1823... thể hiện sự thay thế nucleotide đồng loạt giữa các mẫu.

Trên cây phát sinh loài (Hình 3. 27) cá thể *Chà vá chân xám* Tây Nguyên phân bố gần với mẫu NC088556, nhưng tách xa hẳn với NC018063 và JQ_821842. Điều này chỉ ra rằng *Chà vá chân xám* Tây Nguyên đã biệt lập so với các dòng ở khu vực khác.

Bảng 3. 26. Các vị trí khác biệt trên trình tự gene 16S của cá thể *Khỉ mặt đỏ* Tây Nguyên Việt Nam và cá thể *Khỉ mặt đỏ* khác

	3	2 8	5 5	5 6	1 5	1 2	3 1	4 3	5 8	6 3	6 5	7 3	7 4	8 0	8 3	9 9	1 0 7	1 1 3	1 2 5	1 3 0	1 3 5	1 3 5	1 3 6	1 3 2	1 3 6	1 3 7
Khỉ mặt đỏ Tây Nguyên Việt Nam	T	C	T	T	T	T	T	G	T	C	T	A	C	G	T	G	A	C	A	T	A	C	C	A	C	C
Khỉ mặt đỏ KJ567055	C	T	C	C	C	C	C	A	C	T	T	G	C	A	C	A	G	C	G	C	C	G	T	C	G	T
Khỉ mặt đỏ KM350179	C	T	C	T	C	C	C	G	C	C	C	G	T	A	C	A	A	T	A	T	C	G	T	C	G	T
Khỉ mặt đỏ 16S NC025201	C	T	C	C	C	C	C	A	C	T	T	G	C	A	C	A	G	C	G	C	C	G	T	C	G	T

	1 3 7 2	1 3 7 4	1 3 7 5	1 3 7 6	1 3 7 7	1 3 7 8	1 3 7 9	1 3 8 0	1 3 8 1	1 3 8 2	1 3 8 3	1 3 8 6	1 3 8 7	1 3 8 8	1 3 8 9	1 3 9 0	1 3 9 1	1 3 9 2	1 3 9 5	1 3 9 6	1 3 9 8	1 3 9 9	1 4 0 1	1 4 0 2	1 4 0 3	1 4 0 6
Khỉ mặt đỏ Tây Nguyên Việt Nam	G	A	C	A	C	A	C	C	G	C	C	T	C	A	C	T	C	C	T	C	A	A	A	T	A	T
Khỉ mặt đỏ KJ567055	T	G	T	T	T	G	T	T	C	A	A	A	T	T	A	A	A	G	C	T	C	G	G	A	T	G
Khỉ mặt đỏ KM350179	T	G	T	T	T	G	T	T	C	A	A	A	T	T	A	A	A	G	C	T	C	G	G	A	T	G
Khỉ mặt đỏ 16S NC025201	T	G	T	T	T	G	T	T	C	A	A	A	T	T	A	A	A	G	C	T	C	G	G	A	T	G

	1 4 0 8	1 4 0 9	1 4 1 0	1 4 1 1	1 4 1 3	1 4 1 5	1 4 1 7	1 4 1 8	1 4 2 0	1 4 2 2	1 4 2 4	1 4 2 5	1 4 2 8	1 4 2 9	1 4 3 0	1 4 3 3	1 4 3 5	1 4 3 7	1 4 4 1	1 4 4 4	1 4 4 5	1 4 4 7	1 4 4 8	1 4 4 9
Khỉ mặt đỏ Tây Nguyên Việt Nam	A	G	G	A	T	T	T	T	A	T	A	A	C	C	C	T	A	A	A	T	A	G	A	A	G	G
Khỉ mặt đỏ KJ567055	G	T	T	C	G	C	G	G	G	A	T	C	G	G	T	C	G	G	T	C	C	T	T	C	A	C

Khỉ mặt đỏ KM350179	G	T	T	C	G	C	G	G	G	A	T	C	G	G	T	C	G	G	T	C	C	T	T	C	A	C
Khỉ mặt đỏ 16S NC025201	G	T	T	C	G	C	G	G	G	A	T	C	G	G	T	C	G	G	T	C	C	T	T	C	A	C

	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
	5	5	5	5	5	5	6	6	6	6	6	6	7	7	7	7	7	8	8	8	8	8	8	8	8	9
	2	3	4	7	8	9	0	1	3	5	6	8	0	1	2	7	8	0	1	3	4	6	7	8	9	0
Khỉ mặt đỏ Tây Nguyên Việt Nam	A	A	G	G	T	A	A	C	T	G	T	A	C	G	T	G	A	G	G	G	C	C	T	T	G	G
Khỉ mặt đỏ KJ567055	T	T	C	C	C	T	G	T	C	A	A	G	A	C	A	A	G	A	A	A	A	G	C	C	C	A
Khỉ mặt đỏ KM350179	T	T	C	C	C	T	G	T	C	A	A	G	A	C	A	A	G	A	A	A	A	G	C	C	C	A
Khỉ mặt đỏ 16S NC025201	T	T	C	C	C	T	G	T	C	A	A	G	A	C	A	A	G	A	A	A	A	G	C	C	C	A

	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	9	9	9	9	9	0	0	0	0	0	0	0	0	1	1	1	1	2	2	2	2	3	3	3	3	4
	1	3	4	5	7	1	2	3	4	5	6	8	9	0	4	8	9	4	6	7	9	0	2	3	4	3
Khỉ mặt đỏ Tây Nguyên Việt Nam	A	A	A	A	C	G	T	A	T	A	G	T	T	A	C	G	C	G	C	T	A	C	C	C	C	C
Khỉ mặt đỏ KJ567055	C	T	C	T	A	C	G	C	C	T	T	A	C	T	T	T	G	A	T	C	T	A	T	T	T	A
Khỉ mặt đỏ KM350179	C	T	C	A	A	C	G	C	C	T	T	A	T	T	T	T	G	A	T	C	T	A	T	T	T	A
Khỉ mặt đỏ 16S NC025201	C	T	C	T	A	C	G	C	C	T	T	A	C	T	T	T	G	A	T	C	T	A	T	T	T	A

	1	1	1	1	1	1	1	1	1	1	1
	5	5	5	5	5	5	5	5	5	5	5
	4	5	5	5	5	6	6	6	6	6	6
	4	2	5	7	8	1	2	5	6	7	9
Khỉ mặt đỏ Tây Nguyên Việt Nam	T	A	G	T	T	T	T	C	T	T	A
Khỉ mặt đỏ KJ567055	C	T	A	C	C	A	A	A	-	-	-
Khỉ mặt đỏ KM350179	C	T	A	C	C	A	A	A	G	G	T
Khỉ mặt đỏ 16S NC025201	C	T	A	C	C	A	A	A	-	-	-

Cá thể *Khỉ mặt đỏ* Tây Nguyên có hơn 70 vị trí biến đổi nucleotide so với các mẫu tham chiếu khác (KJ567055, KM350179, NC025201) (Bảng 3. 26). Các biến đổi xuất hiện trên toàn trình tự, chủ yếu là các thay thế T↔C, A↔G, và C↔G. Một số chuỗi liên tục thể hiện các biến đổi đặc trưng, như từ vị trí 1508–1513, mẫu Tây Nguyên mang C–C–C–C–C–C, trong khi các mẫu đối chiếu mang T–T–T–T–T–T. Đặc biệt, chuỗi từ 1607–1610: *Khỉ mặt đỏ* Tây Nguyên mang A–T–T–T, trong khi các mẫu khác mang G–C–C–C. So sánh chi tiết: vị trí 1124: Tây Nguyên có T, trong khi mẫu đối chiếu có C; vị trí 1361: G so với A → phản ánh sự chuyển G→A phổ biến ở các dòng tiến hóa khác biệt; vị trí 1721–1725: A–C–A–T–C ở mẫu Tây Nguyên khác biệt rõ rệt với các mẫu khác mang G–G–T–G–T; nhiều vị trí khác như

1865, 1885, 1903... thể hiện xu hướng biến đổi trùng nhau giữa các mẫu đối chiếu, trong khi cá thể linh trưởng ở Tây Nguyên lại mang dạng duy nhất.

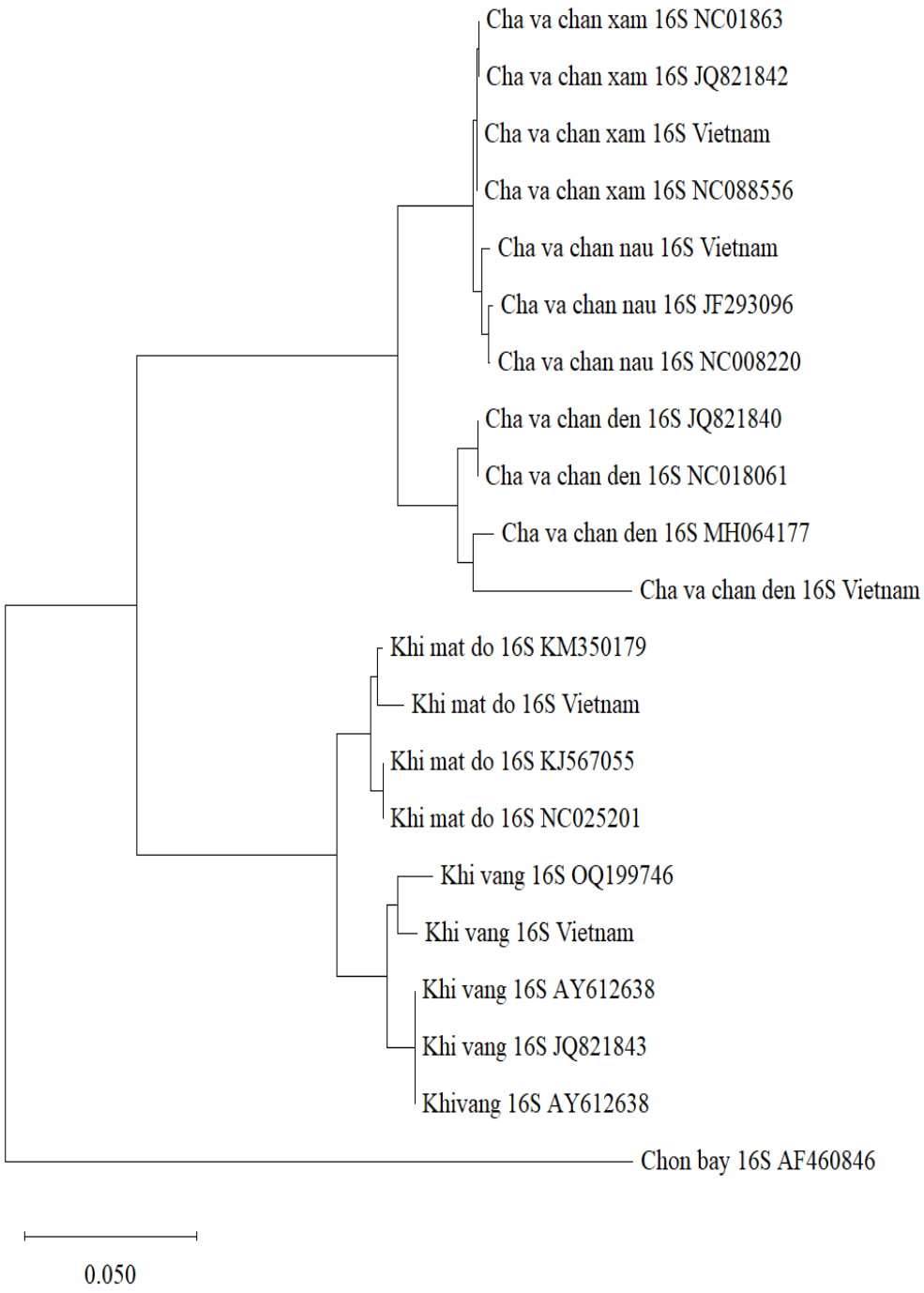
Mức độ biến dị nucleotide giữa *Khỉ mặt đỏ* Tây Nguyên và các mẫu khác tương đối cao so với một loài thông thường. *Khỉ mặt đỏ* Tây Nguyên cũng thể hiện rõ sự phân hóa trên cây phát sinh loài (Hình 3. 27), xuất hiện riêng biệt và không hòa lẫn với các nhánh mẫu tham chiếu.

Bảng 3. 27. Các vị trí khác biệt trên trình tự gene *16S* của cá thể *Khỉ vàng* Tây Nguyên Việt Nam và cá thể *Khỉ vàng* khác

	2 7	4 4	4 9	1 3 4	1 6 8	2 1 7	2 6 8	3 3 1	3 7 5	4 0 7	4 8 5	4 8 8	4 9 3	5 0 6	5 7 6	6 1 0	6 2 3	8 1 0	1 2 9	1 3 4	1 3 5	1 3 9	1 8 7	1 1 2	1 2 1	1 2 6	1 2 5
Khỉ vàng Tây Nguyên Việt Nam	T	C	T	A	T	G	G	T	A	C	C	C	C	C	A	T	T	T	G	A	A	C	C	C	T	A
Khỉ vàng NC005943	C	T	C	G	C	A	A	C	G	T	T	T	T	T	G	C	C	C	A	G	T	T	T	T	C	G
Khỉ vàng JQ_821843	C	T	C	G	C	A	A	C	G	T	T	T	T	T	G	C	C	C	A	G	T	T	T	T	C	G
Khỉ vàng AY612638	C	T	C	G	C	A	A	C	G	T	T	T	T	T	G	C	C	C	A	G	T	T	T	T	C	G

Cá thể *Khỉ vàng* Tây Nguyên được so sánh với ba trình tự tham chiếu khác (NC005943, JQ_821843, AY612638) cho thấy 26 vị trí nucleotide khác biệt trên trình tự gene *16S rRNA* (Bảng 3. 27). Trong tất cả các vị trí kể trên, cá thể *Khỉ vàng* Tây Nguyên đều có nucleotide khác biệt hoàn toàn so với ba cá thể đối chứng còn lại (đa số thuộc khu vực Đông Á, Ấn Độ). ADN ty thể của *Khỉ vàng* Tây Nguyên có đặc điểm riêng biệt như: vị trí 1124 là T thay vì C (ở 3 mẫu còn lại), vị trí 1244: T so với C, vị trí 1361: G so với A, ị trí 1508–1513: một loạt các vị trí đồng nhất với C (Tây Nguyên) so với T (các mẫu khác), vị trí 1607: A so với G, vị trí 1721, 1723: A – A so với T – T.

Trình tự gene *16S rRNA* là một trong những vùng gene phổ biến nhất được sử dụng trong nghiên cứu phát sinh loài, đặc biệt ở động vật có xương sống và linh trưởng. Do có tính bảo tồn cao nhưng vẫn chứa các vùng biến đổi vừa phải, gene này cung cấp một công cụ hiệu quả để phân tích mối quan hệ tiến hóa giữa các loài và cá thể khác nhau. Trong nghiên cứu này, cây phát sinh loài được xây dựng từ trình tự *16S rRNA* của các loài linh trưởng tại Tây Nguyên đã cho thấy nhiều đặc điểm đáng chú ý về sự phân hóa di truyền, lịch sử tiến hóa và cấu trúc cá thể. Do tốc độ tiến hóa tương đối chậm của gene này, nó đặc biệt hữu ích trong việc xây dựng cây phát sinh loài giữa các loài hoặc nhóm loài có thời gian tách biệt tiến hóa từ vài triệu năm trước.



Hình 3. 27. Cây phát sinh loài Linh trưởng Tây Nguyên từ trình tự *16S rRNA*.

Cây phát sinh loài Linh trưởng Tây Nguyên từ trình tự *16S* rRNA gợi ý mối quan hệ di truyền giữa các loài linh trưởng, đặc biệt là sự khác biệt giữa các loài ở Việt Nam so với các loài từ khu vực khác. Các loài linh trưởng ở Tây Nguyên Việt Nam, như *Chà vá chân xám* và *Chà vá chân đen* xuất hiện ở nhiều nhánh khác nhau, cho thấy sự khác biệt di truyền đáng kể trong các cá thể tại Việt Nam. Cá thể *Chà vá chân nâu* từ Việt Nam thể hiện sự khác biệt di truyền so với các loài linh trưởng khác trong khu vực khi nằm ở 1 nhánh biệt so với 2 cá thể khác. Đối với *Khỉ mặt đỏ* Việt

Nam cũng cho thấy sự phân hóa di truyền. Mặc dù các loài tập trung chặt chẽ trong các nhánh nhỏ, các mẫu từ Việt Nam lại xuất hiện ở vị trí riêng biệt, thể hiện sự khác biệt về mặt di truyền của cá thể này. Điều này đặc biệt quan trọng trong bối cảnh Việt Nam là một khu vực có sự đa dạng môi trường sống lớn, từ vùng núi cao đến các khu rừng nhiệt đới, tạo điều kiện cho các loài linh trưởng phát triển theo những hướng tiến hóa riêng biệt.

Kết quả từ cây phát sinh này phù hợp với lý thuyết về sự phân hóa địa lý (allopatric speciation), điều kiện khí hậu và địa hình phức tạp tại Tây Nguyên – với cao nguyên, dãy núi, rừng sâu và sông ngòi – đóng vai trò là yếu tố chính tạo nên mô hình phân bố rời rạc của các loài linh trưởng. Điều này cũng được ghi nhận trong các nghiên cứu tương tự tại Đông Dương và Nam Trung Hoa (126). Tuy nhiên cần có các phân tích tiếp theo kết hợp ADN trong nhân (nuclear ADN), dữ liệu hình thái, và dữ liệu môi trường để xây dựng mô hình phát sinh loài toàn diện và cũng cần tiến hành phân tích thời gian phân ly (molecular dating) để xác định lịch sử tiến hóa chính xác hơn.

3.3.2. Phân tích sai khác trình tự gene *cytb* và mối quan hệ phát sinh loài của các cá thể linh trưởng khu vực Tây Nguyên Việt Nam

Bảng 3. 28. Các vị trí khác biệt trên trình tự gene *cytb* của cá thể *Chà vá chân đen* Tây Nguyên Việt Nam và cá thể *Chà vá chân đen* khác

	2 1	4 6	7 2	1 7 4	2 2 4	2 6 4	3 0 3	3 0 4	3 0 9	3 3 0	4 3 5	4 9 8	5 0 1	5 1 3	5 2 5	5 2 6	5 8 2	5 8 5	6 5 4	7 3 5	7 7 1	7 8 9	8 1 6	8 9 1	1 0 3	1 0 6
Chà vá chân đen Tây Nguyên Việt Nam	C	C	G	C	T	C	A	C	C	T	A	G	A	C	T	A	T	A	T	T	A	C	A	T	G	T
Chà vá chân đen NC018061	T	C	A	T	C	T	G	T	T	C	G	A	G	T	C	G	C	T	C	C	T	T	G	C	A	C
Chà vá chân đen MH_064177	C	T	G	C	T	C	A	C	C	T	A	G	A	C	T	A	T	A	T	T	A	C	A	T	G	C
Chà vá chân đen JQ821840	T	C	A	T	C	T	G	T	T	C	G	A	G	T	C	G	C	T	C	C	T	T	G	C	A	C

26 vị trí nucleotide khác biệt được tìm thấy trong vùng gene *cytb* giữa cá thể *Chà vá chân đen* Tây Nguyên Việt Nam và các mẫu tham chiếu (NC018061, MH_064177, JQ821840) (Bảng 3. 28). Các biến dị bao gồm: Các vị trí đáng chú ý như: 1247 (G vs. A), 1303 (A vs. G), 1363 (C vs. T), 1458 (T vs. C), 1522 (A vs. G), 1602 (T vs. C), 1785 (T vs. C). Biến dị thường thấy là các đột biến chuyển (transition) như A↔G và T↔C, phổ biến trong tiến hóa ty thể. *Chà vá chân đen* Tây Nguyên có một chuỗi liên tục A–T–T–A–C–A tại các vị trí cuối, trong khi các mẫu khác thường mang G–C–C–T–T–G. Các khác biệt này cho thấy cá thể Tây Nguyên mang nhiều biến trên gene *cytb* so với các mẫu từ các khu vực khác.

So sánh với dữ liệu từ 52 vị trí khác biệt trong gene *16S rRNA* giữa cá thể *Chà vá chân đen* Tây Nguyên và các mẫu đối chứng, các vị trí biến dị liên tục tại vùng 1508–1513 và 1607–1610 (C–C–C–C–C–C và A–T–T–T ở Tây Nguyên) có thể thấy sự khác biệt giữa cá thể *Chà vá chân đen* Tây Nguyên và các mẫu tham chiếu xuất hiện đồng thời trên cả hai vùng gene – *16S rRNA* (bảo tồn cao) và *cytb* (mã hóa protein).

Bảng 3. 29. Các vị trí khác biệt trên trình tự gene *cytb* của cá thể *Chà vá chân nâu* Tây Nguyên Việt Nam và cá thể *Chà vá chân nâu* khác

	1	1	2	3	4	4	5	6	7	8	9	9	9	1 0	1 0	1 0	1 0	1 0	1 0	1 0	1 0	1 0	1 0	1 0	1 0	1 0
	1 0	1 6	9 4	0 3	0 5	7 7	1 6	2 1	5 3	1 6	0 7	3 9	9 4	2 0	2 4	2 6	2 7	2 8	2 9	3 3	3 6	3 7	3 8	4 0	4 1	4 2
Chà vá chân nâu Tây Nguyên Việt Nam	G	C	A	G	A	T	T	C	T	A	G	A	T	C	A	C	C	A	T	T	T	A	A	A	C	G
Chà vá chân nâu NC008220	A	T	G	A	G	C	C	T	C	G	A	G	C	T	C	A	G	T	A	C	C	C	C	T	T	A
Chà vá chân nâu JF_293096	A	T	G	A	G	C	C	T	C	G	A	G	C	T	C	A	G	T	A	C	C	C	C	T	T	A
Chà vá chân nâu DQ355302	A	T	G	A	G	C	C	T	C	G	A	G	C	T	C	A	G	T	A	C	C	C	C	T	T	A

	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
	4	5	5	5	6	6	6	6	7	7	7	8	8	8	8	8	9	9	0	0	0	0	0	1	1	1
	3	4	6	7	0	1	4	8	0	3	8	2	3	4	5	7	1	3	4	5	6	8	9	1	2	3
Chà vá chân nâu Tây Nguyên Việt Nam	C	A	A	A	A	A	A	G	C	T	C	A	T	G	G	C	A	C	G	A	T	T	G	A	A	A
Chà vá chân nâu NC008220	T	C	G	G	G	C	C	A	T	A	T	C	A	C	A	T	G	T	A	C	C	C	T	G	C	C
Chà vá chân nâu JF_293096	T	C	G	G	G	C	C	A	T	A	T	C	A	C	A	T	G	T	A	C	C	C	T	G	C	C
Chà vá chân nâu DQ355302	T	C	G	G	G	C	C	A	T	A	T	C	A	C	A	T	G	T	A	C	C	C	T	G	C	C

	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	4
	4	6	9	0	2	4	6	7	8	9	1	4	5	6	7	1
Chà vá chân nâu Tây Nguyên Việt Nam	A	C	C	G	T	T	T	T	T	C	A	A	T	A	A	C
Chà vá chân nâu NC008220	T	T	A	A	C	A	A	A	C	A	T	C	C	T	C	T
Chà vá chân nâu JF_293096	T	T	A	A	C	A	A	A	C	A	T	C	C	T	C	T
Chà vá chân nâu DQ355302	T	T	A	A	C	A	A	A	C	A	T	C	C	T	C	T

Có khoảng 60 nucleotide khác biệt trong vùng gene *cytb* giữa cá thể *Chà vá chân nâu* Tây Nguyên Việt Nam và ba mẫu tham chiếu từ GeneBank (NC008220, JF_293096, DQ355302) (Bảng 3. 29) Dạng biến đổi phổ biến: A↔G, T↔C, G↔C → đa phần là đột biến chuyển (transition), nhưng có một số đột biến đảo (transversion). Đặc biệt, cá thể *Chà vá chân nâu* Tây Nguyên chia sẻ rất ít hoặc không chia sẻ alen với cả ba mẫu tham chiếu.

Đối chiếu với gene *16S* (Bảng 3. 24) mặc dù số lượng biến dị ít hơn (chỉ 10 vị trí) so với *cytb*, nhưng các vị trí thay thế có tính ổn định. Các biến dị T↔C xảy ra liên tục tại các vị trí 4450–4484. Kết hợp dữ liệu từ cả hai bảng chỉ ra rằng biến dị tồn tại đồng thời ở cả vùng bảo tồn (*16S*) lẫn vùng mã hóa protein (*cytb*), từ đó có gợi ý sự tồn tại của một dòng di truyền riêng biệt tại khu vực Tây Nguyên.

Bảng 3. 30. Các vị trí khác biệt trên trình tự gene *cytb* của cá thể *Chà vá chân xám* Tây Nguyên Việt Nam và cá thể *Chà vá chân xám* khác

	435	445	684	713	1067
Chà vá chân xám Tây Nguyên Việt Nam	A	T	T	C	T
Chà vá chân xám NC 088556	A	T	T	C	T
Chà vá chân xám NC018063	G	C	C	T	C
Chà vá chân xám cytb JQ_821842	G	C	C	T	C

Bảng 3. 30 ghi nhận 5 vị trí nucleotide khác biệt trong vùng gene *cytb* giữa cá thể *Chà vá chân xám* Tây Nguyên Việt Nam và các mẫu *Chà vá chân xám* khác (NC088556, NC018063, JQ_821842). Các vị trí biến dị bao gồm: 1446, 1447, 1454, 1458, 1463. Cá thể linh trưởng ở Tây Nguyên mang cặp base A–T tại hai vị trí đầu (1446–1447), trong khi các mẫu đối chiếu có G–C. Vị trí 1454: *Chà vá chân xám* Tây Nguyên có T, trong khi các mẫu khác có C →T. Hai vị trí còn lại (1458, 1463) có sự thay đổi đồng nhất giữa cá thể *Chà vá chân xám* Tây Nguyên và các mẫu từ các khu vực khác.

Bảng 3. 31. Các vị trí khác biệt trên trình tự gene *cytb* của cá thể *Khỉ mặt đỏ* Tây Nguyên Việt Nam và cá thể *Khỉ mặt đỏ* khác

	45	121	127	151	267	333	432	475	486	507	515	519	531	552	654	741	789	850	960	969	996	997	1038	1104
Khỉ mặt đỏ Tây Nguyên Việt Nam	T	C	T	T	A	A	A	G	G	T	G	C	A	C	C	C	C	G	C	T	G	C	T	G
Khỉ mặt đỏ NC025201	C	C	C	T	G	G	G	A	G	C	G	T	G	T	C	T	T	A	G	C	A	T	C	A

Khỉ mặt đỏ KM_360179	T	T	C	C	A	G	G	A	A	C	A	C	A	C	T	T	T	A	C	T	A	T	C	G
Khỉ mặt đỏ KJ567055	C	C	C	T	G	G	G	A	G	C	G	T	G	T	C	T	T	A	G	C	A	T	C	A

Khi so sánh với cá thể *Khỉ mặt đỏ* Tây Nguyên với các cá thể *Khỉ mặt đỏ* tham chiếu (NC025201, KM_360179, KJ567055) đã cho thấy 25 vị trí nucleotide khác biệt trong vùng gene *cytb* (Bảng 3. 31). Các biến dị chủ yếu là dạng T↔C, A↔G, phân bố rải rác trên toàn trình tự. Một số vị trí tiêu biểu như: tại vị trí 1245, 1256, 1285: ADN ty thể của cá thể *Khỉ mặt đỏ* Tây Nguyên mang T–C–T, trong khi các mẫu khác có C–T–C; vị trí 1456: có C thay vì A (hoặc G); vị trí 1496–1499: chuỗi C–T–G–C ở mẫu *Khỉ mặt đỏ* Tây Nguyên khác biệt hoàn toàn so với các mẫu còn lại.

Sự hiện diện của nhiều đột biến trên cả gene *cytb* và *16S rRNA* củng cố khả năng hình thành một dòng di truyền biệt lập tại Tây Nguyên.

Bảng 3. 32. Các vị trí khác biệt trên trình tự gene *cytb* của cá thể *Khỉ vàng* Tây Nguyên Việt Nam và cá thể *Khỉ vàng* khác

	5	3 1	4 5	5 7	6 7	8 7	9 0	1 2 4	1 3 6	1 9 8	2 2 8	2 7 3	2 7 6	3 0 0	3 0 6	3 2 1	3 4 8	3 6 9	3 7 8	3 8 9	3 9 4	4 0 0	4 1 5	4 2 4	4 3 6	4 4 8	4 4 7
Khỉ vàng Tây Nguyên Việt Nam	T	A	C	T	A	G	A	G	G	T	C	T	C	A	C	T	C	C	A	T	C	G	G	G	T	A
Khỉ vàng OQ199746	T	A	T	T	A	G	A	A	A	C	T	T	C	A	T	T	C	C	A	C	T	A	G	G	C	A
Khỉ vàng KJ567053	C	C	T	C	G	A	G	A	A	T	C	C	T	G	T	C	T	T	G	C	T	A	A	A	C	G
Khỉ vàng JQ_821843	C	C	T	C	G	A	G	A	A	T	C	C	T	G	T	C	T	T	G	C	T	A	A	A	C	G
Khỉ vàng AY612638	C	C	T	C	G	A	G	A	A	T	C	C	T	G	T	C	T	T	G	C	T	A	A	A	C	G

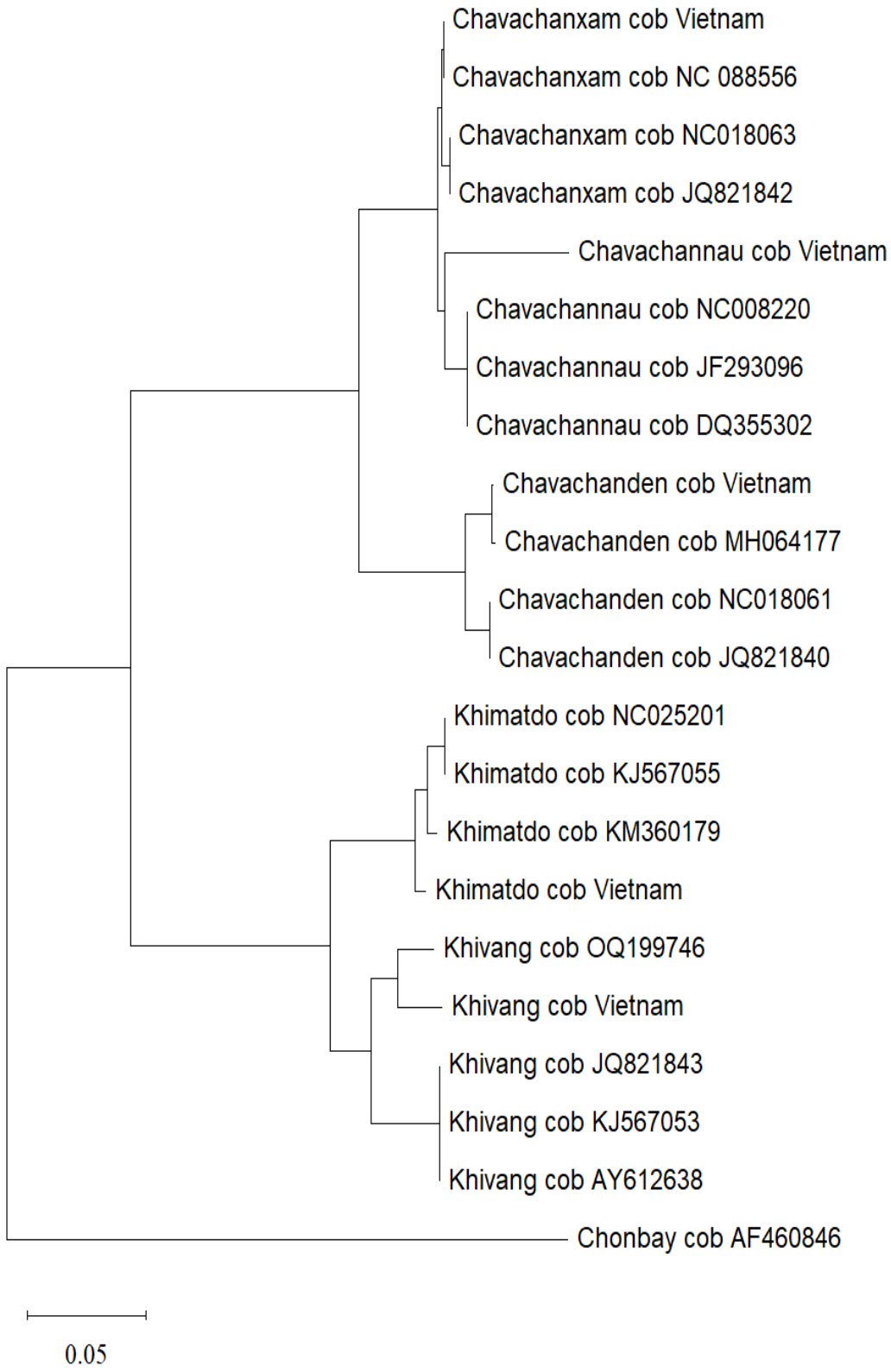
	4 5 0	4 5 9	4 6 5	4 6 8	4 8 0	4 9 2	5 1 2	5 1 9	5 5 2	5 6 4	5 8 0	5 8 2	5 8 5	5 9 1	6 0 0	6 2 4	6 4 5	6 5 6	6 6 7	6 7 0	7 1 1	7 1 3	7 4 5	7 7 1	7 7 4	7 8 4
Khỉ vàng Tây Nguyên Việt Nam	A	C	T	T	A	T	A	C	T	T	G	T	A	A	A	C	G	T	T	A	T	T	T	C	A	C
Khỉ vàng OQ199746	G	T	C	T	A	C	G	C	C	T	A	T	A	G	A	T	G	C	T	A	T	C	C	C	G	C
Khỉ vàng KJ567053	A	C	T	C	C	C	A	T	C	C	A	C	G	A	G	T	A	C	C	G	C	T	C	T	A	T
Khỉ vàng JQ_821843	A	C	T	C	C	C	A	T	C	C	A	C	G	A	G	T	A	C	C	G	C	T	C	T	A	T
Khỉ vàng AY612638	A	C	T	C	C	C	A	T	C	C	A	C	G	A	G	T	A	C	C	G	C	T	C	T	A	T

	7 8 9	7 9 2	8 0 1	8 0 7	8 1 3	8 1 6	8 4 0	8 5 0	8 8 5	8 8 6	8 9 1	8 9 5	9 0 0	9 1 5	9 2 0	9 6 5	9 7 1	9 8 9	9 9 7	9 9 1	9 9 7	1 0 0 9	1 0 0 2	1 0 0 0	1 0 0 0	1 0 0 1	1 0 1 2
Khỉ vàng Tây Nguyên Việt Nam	A	T	T	C	C	C	A	T	C	C	A	C	G	A	G	T	A	C	C	G	C	T	C	T	A	T
Khỉ vàng OQ199746	A	T	T	C	C	C	A	T	C	C	A	C	G	A	G	T	A	C	C	G	C	T	C	T	A	T
Khỉ vàng KJ567053	G	C	C	T	T	T	G	C	T	T	G	T	A	G	A	C	G	T	T	A	T	C	T	C	G	C
Khỉ vàng JQ_821843	G	C	C	T	T	T	G	C	T	T	G	T	A	G	A	C	G	T	T	A	T	C	T	C	G	C
Khỉ vàng AY612638	G	C	C	T	T	T	G	C	T	T	G	T	A	G	A	C	G	T	T	A	T	C	T	C	G	C

Khỉ vàng Tây Nguyên Việt Nam	T	C	T	G	G	G	T	G	C	C	G	T	C	C	C	C	A	A	G	T	A	G	T	A	C	T	A
Khỉ vàng OQ199746	C	T	C	G	A	A	C	A	C	T	A	C	C	C	C	C	G	A	A	C	A	A	C	A	C	T	A
Khỉ vàng KJ567053	C	T	T	A	G	A	C	A	T	C	G	C	T	T	T	T	A	C	A	T	G	A	C	G	T	C	G
Khỉ vàng JQ_821843	C	T	T	A	G	A	C	A	T	C	G	C	T	T	T	T	A	C	A	T	G	A	C	G	T	C	G
Khỉ vàng AY612638	C	T	T	A	G	A	C	A	T	C	G	C	T	T	T	T	A	C	A	T	G	A	C	G	T	C	G

Bảng 3. 32 ghi nhận hơn 30 vị trí nucleotide khác biệt trong vùng gene *cytb* giữa cá thể *Khỉ vàng* Tây Nguyên Việt Nam và các mẫu tham chiếu từ GeneBank (OQ199746, KJ567053, JQ_821843, AY612638). Các vị trí biến dị phân bố từ đầu đến cuối gene, nổi bật là các chuỗi liên tục từ vị trí 1267–1270, 1360–1363, 1497–1500. Cá thể *Khỉ vàng* Tây Nguyên mang nhiều dạng alen không phổ biến, như vị trí 1285: *Khỉ vàng* Tây Nguyên có G, trong khi các mẫu khác có A hoặc T; vị trí 1499–1500: C–A ở *Khỉ vàng* Tây Nguyên so với G–C hoặc T–C ở mẫu khác.

Tổng thể, *Khỉ vàng* Tây Nguyên thể hiện một mẫu hình biến dị phân tán và độc lập, phản ánh khả năng tiến hóa biệt lập tương đối lâu dài.



Hình 3. 28. Cây phát sinh loài Linh trưởng Tây Nguyên từ trình tự gene *cytb*

Gene *cytb* là một gene mã hóa protein nằm trong hệ thống chuỗi chuyền điện tử ty thể, *cytb* có tốc độ tiến hóa trung bình và cung cấp lượng thông tin di truyền lớn, rất phù hợp cho việc phân tích ở cấp độ loài và phân loài. Trong nghiên cứu này, cây

phát sinh loài được xây dựng từ trình tự gene *cytb* của các loài linh trưởng ở Tây Nguyên đã mang lại nhiều kết quả quan trọng giúp hiểu rõ mối quan hệ tiến hóa và sự phân hóa di truyền giữa các cá thể trong khu vực. *Cytb* là một trong 13 gene mã hóa protein nằm trong hệ gene ty thể và đã được chuẩn hóa để sử dụng trong các nghiên cứu phân loại học, tiến hóa học và xác định loài nhờ vào khả năng phản ánh tốt sự biến đổi di truyền ở cấp độ giữa các loài.
Với đặc tính trung gian về tốc độ tiến hóa, *cytb* có thể phát hiện được sự phân hóa gần và sâu, phù hợp cho các nghiên cứu về các nhóm linh trưởng có lịch sử tiến hóa phức tạp như ở Đông Dương.

Trình tự *cytb* của cá thể *Chà vá chân nâu* của Việt thể hiện sự khác biệt di truyền đáng kể so với 2 loài đến từ Trung Quốc và Myanma. Bên cạnh đó *Khỉ mặt đỏ* ở Tây Nguyên cũng thể hiện sự khác biệt di truyền trên vùng trình tự *D-loop* với các cá thể *Khỉ mặt đỏ* khác trong khu vực. Cụ thể, trong khi các loài từ các khu vực khác thường tập trung trong các nhánh riêng biệt, các mẫu từ Việt Nam có xu hướng phân bố đa dạng hơn, cho thấy khả năng tồn tại các phân loài có lịch sử tiến hóa riêng biệt. Điều này gợi ý rằng Việt Nam có thể là một khu vực quan trọng trong sự phát triển và đa dạng hóa của các loài linh trưởng thuộc nhóm này.

3.3.3. Phân tích sai khác trình tự gene *D-loop* và mối quan hệ phát sinh loài của các cá thể linh trưởng khu vực Tây Nguyên Việt Nam

Bảng 3. 33. Các vị trí khác biệt trên trình tự gene *D-loop* của cá thể *Chà vá chân xám* Tây Nguyên Việt Nam và cá thể *Chà vá chân xám* khác

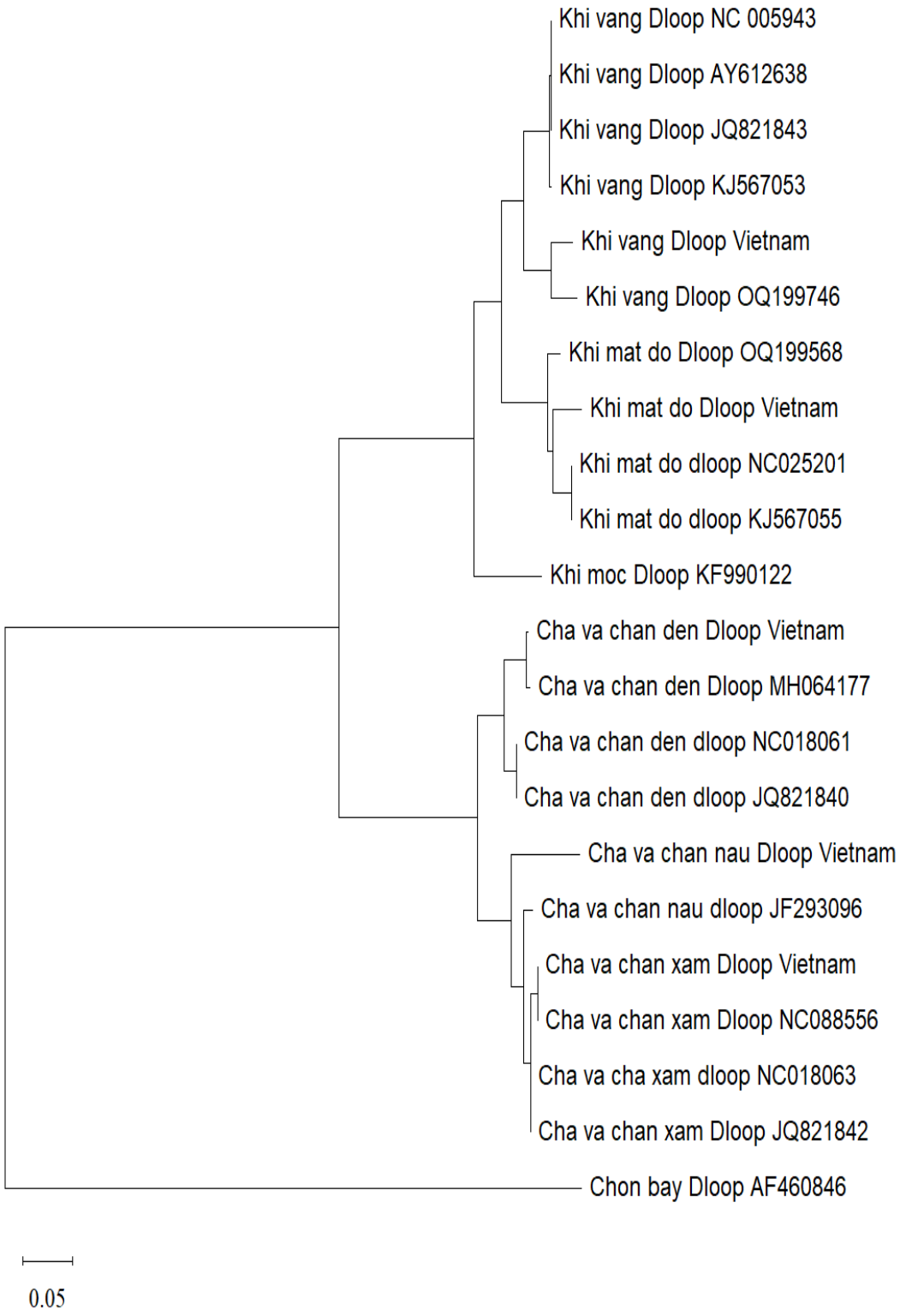
	2 2 7	2 3 1	2 3 3	2 4 1	3 9 3	4 7 8	7 4 8
Chà vá chân xám Tây Nguyên Việt Nam	C	T	C	G	G	C	T
Chà vá chân xám NC088556	C	T	C	G	G	C	T
Chà vá chân xám NC018063	T	C	T	A	A	T	C
Chà vá chân xám JQ_821842	T	C	T	A	A	T	C

Vùng *D-loop* là một vùng không mã hóa nhưng có tốc độ tiến hóa rất nhanh, do đó đặc biệt phù hợp để phát hiện đa dạng di truyền nội loài. Tuy nhiên so với các cá thể linh trưởng khác có trên 100 vị trí biến đổi nucleotide thì cá thể *Chà vá chân xám* Tây Nguyên Việt Nam chỉ có 7 vị trí nucleotide biến đổi trong vùng *D-loop*. Cá

thể *Chà vá chân xám* Tây Nguyên và NC088556 giống nhau hoàn toàn tại các vị trí khảo sát ⇒ Có thể cùng nguồn gốc địa lý hoặc di truyền gần. Cá thể *Chà vá chân xám* Tây Nguyên khác hoàn toàn với hai mẫu còn lại (NC018063, JQ_821842) tại tất cả các vị trí, điều này cho thấy mức độ phân hóa di truyền lớn. Các đột biến chủ yếu là thay thế base (transition hoặc transversion), không có chèn hay mất đoạn. Sự khác biệt hoàn toàn giữa nhóm Tây Nguyên – NC088556 và nhóm NC018063 – JQ_821842 gợi ý rằng có thể tồn tại 2 dòng tiến hóa chính trong loài *Chà vá chân xám.*

Vùng *D-loop* (Displacement loop) là một trong những vùng có tốc độ tiến hóa nhanh nhất trong hệ gene ty thể và thường được ứng dụng rộng rãi trong nghiên cứu cấu trúc quần thể, phát sinh loài gần và lịch sử tiến hóa của các loài động vật có xương sống. Với đặc tính không mã hóa và giàu trình tự biến thiên, *D-loop* đóng vai trò như một chỉ dấu mạnh mẽ để đánh giá sự khác biệt di truyền nội bộ loài và sự phân hóa giữa các quần thể cận lâm sàng. Trong nghiên cứu này, việc xây dựng cây phát sinh loài từ vùng *D-loop* đã làm sáng tỏ thêm về mối quan hệ di truyền giữa các loài linh trưởng ở Tây Nguyên cũng như cấu trúc tiến hóa nội bộ của từng loài.

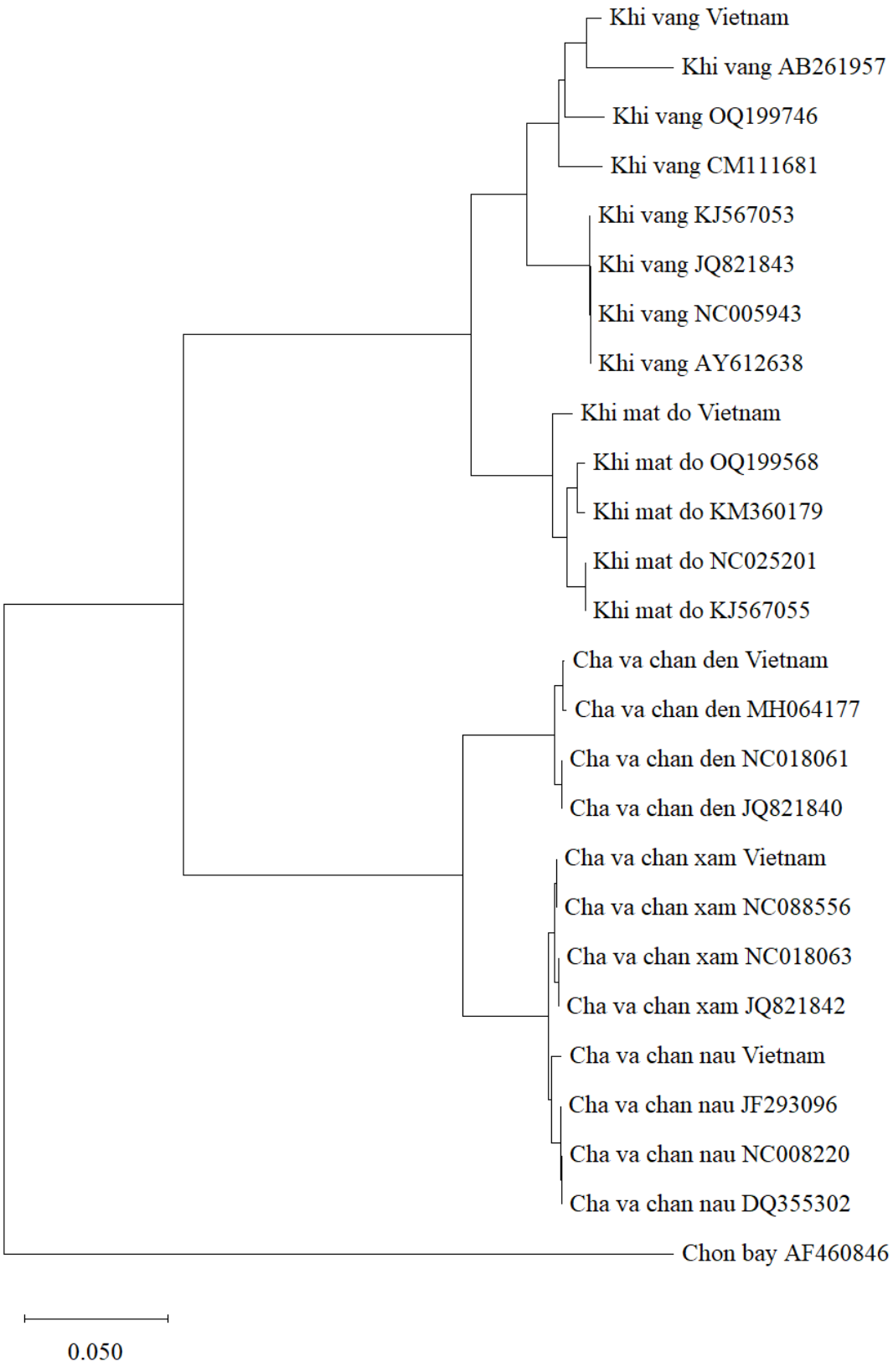
So sánh với các loài từ khu vực khác như *Khỉ vàng* và *Khỉ mặt đỏ*, các mẫu từ Việt Nam cho thấy sự khác biệt di truyền đáng kể. Điều đặc biệt ở đây là cá thể *Chà vá chân nâu* vẫn nằm ở 1 nhánh riêng biệt giống như so sánh trên trình tự gene *16S*. Điều này gợi ý về sự phong phú di truyền và khả năng tiến hóa riêng biệt của *Chà vá chân nâu* ở Tây Nguyên Việt Nam. Sự khác biệt này có thể cho thấy có thể Việt Nam là nơi tồn tại các điều kiện sinh thái độc đáo, tạo nên các dòng di truyền đặc trưng cho cá thể *Chà vá chân nâu*, khác biệt so với các loài ở khu vực lân cận.



Hình 3. 29. Cây phát sinh loài Linh trưởng Tây Nguyên từ trình tự gene *D-loop*.

3.3.4. Mối quan hệ phát sinh loài của một số cá thể linh trưởng khu vực Tây Nguyên Việt Nam

Cây phát sinh loài được xây dựng dựa trên trình tự toàn bộ hệ gene ty thể cho thấy mối quan hệ di truyền rõ ràng giữa các loài linh trưởng phân bố tại khu vực Tây Nguyên Việt Nam, đồng thời phản ánh sự phân hóa phù hợp với hệ thống phân loại hình thái – phân loại học đã được công nhận. Các cá thể *Khỉ vàng* Tây Nguyên Việt Nam nằm trong cùng một nhánh với các trình tự tham chiếu đã công bố trên GenBank (AB261957, OQ199746, CM111681, KJ567053, JQ_821843, NC005943, AY612638). Sự tập hợp chặt chẽ này cho thấy: mức độ tương đồng di truyền cao giữa cá thể *Khỉ vàng* Việt Nam và các cá thể *Khỉ vàng* khác trong khu vực và trên thế giới; cấu trúc hệ gene ty thể của loài này được bảo tồn tương đối ổn định, phản ánh mối quan hệ phát sinh loài gần gũi trong nội bộ loài. Cá thể *Khỉ mặt đỏ* Tây Nguyên Việt Nam tạo thành một nhánh riêng biệt nhưng vẫn nằm trong cụm *macaca*, cùng nhóm với các trình tự OQ199568, KM_360179, NC025201 và KJ567055. Điều này cho thấy *Khỉ mặt đỏ* có sự phân hóa di truyền rõ rệt so với *Khỉ vàng*. Tuy nhiên, mối quan hệ họ hàng gần giữa hai loài trong cùng chi *Macaca* vẫn được duy trì, phù hợp với nguồn gốc tiến hóa chung. Các cá thể *Chà vá chân đen* Tây Nguyên Việt Nam (MH_064177, NC018061, JQ821840) tạo thành một nhánh riêng biệt, tách rõ khỏi nhóm *Macaca*. Điều này phản ánh sự khác biệt di truyền lớn giữa hai chi *Macaca* và *Pygathrix* và tính ổn định di truyền trong nội bộ loài *Chà vá* chân đen. Cá thể *Chà vá chân xám* Tây Nguyên Việt Nam nhóm chặt với các trình tự tham chiếu NC088556, NC018063 và JQ_821842. Sự phân nhóm này cho thấy *Chà vá chân xám* Tây Nguyên Việt Nam có tính đặc trưng di truyền cao và phân biệt rõ ràng với các loài *Chà vá* khác. Kết quả củng cố vị trí phân loại độc lập của *Chà vá* chân xám trong chi *chà vá*. Cá thể *Chà vá chân nâu* Tây Nguyên Việt Nam nằm cùng nhánh với các trình tự JF_293096, NC008220 và DQ355302, cho thấy mức độ tương đồng di truyền cao giữa cá thể nghiên cứu và các cá thể đã công bố và mối quan hệ phát sinh loài gần gũi giữa *Chà vá chân nâu* với các loài *Chà vá* khác, đặc biệt là *Chà vá chân xám* và *Chà vá* chân đen, nhưng vẫn thể hiện sự phân hóa rõ ràng ở cấp loài.



Hình 3. 30. Cây phát sinh loài Linh trưởng Tây Nguyên từ trình tự của toàn hệ gene ty thể.

3.3.5. Thảo luận

Luận án đã công bố lần đầu tiên hệ gen ty thể hoàn chỉnh của 5 loài linh trưởng bao gồm (*Chà vá chân xám, Chà vá chân nâu, Chà vá chân đen, Khỉ mặt đỏ và Khỉ vàng*) từ các cá thể có nguồn gốc địa lý xác định tại khu vực Tây Nguyên, Việt Nam. Các kết quả thu được đã góp phần củng cố cơ sở dữ liệu di truyền cho các loài này, vốn còn rất hạn chế trong các ngân hàng gen quốc tế.

Về cấu trúc và tổ chức gen: Các hệ gen ty thể của cả 5 loài đều thể hiện cấu trúc điển hình của động vật có xương sống, với 37 gen được sắp xếp theo trật tự bảo tồn cao. Các chỉ số thành phần nucleotide như hàm lượng A+T và AT/GC-skew cho thấy sự khác biệt rõ rệt giữa hai phân họ Colobinae (giàu AT) và Cercopithecinae (giàu GC hơn). Sự khác biệt có hệ thống này nhất quán với các nghiên cứu trước đây và được giải thích là do lịch sử tiến hóa và cơ chế sao chép ADN ty thể khác biệt giữa hai nhóm. Các đặc điểm cấu trúc khác như sự xuất hiện của mismatch trong tRNA hay các vùng chồng chéo gen ngắn cũng là những đặc điểm chung, phản ánh tính bảo tồn chức năng của bộ máy dịch mã ty thể.

Kết quả cho thấy, mặc dù trên các cây phát sinh loài xây dựng từ từng vùng gene riêng lẻ và từ toàn bộ hệ gene ty thể, các loài *Chà vá* (*Chà vá chân nâu, Chà vá chân xám, Chà vá chân đen*) vẫn duy trì cấu trúc phân nhánh ổn định và không xuất hiện sự phân hóa sâu ở cấp độ loài, phân tích chi tiết sai khác nucleotide lại ghi nhận số lượng lớn các biến dị tại nhiều vị trí trên trình tự. Hiện tượng này phản ánh hai cấp độ tiến hóa khác nhau: ở cấp độ phát sinh loài, cấu trúc di truyền tổng thể và mối quan hệ họ hàng chung vẫn được bảo tồn; trong khi đó, ở cấp độ vi mô, các đột biến trung tính vẫn liên tục tích lũy theo thời gian, đặc biệt tại các vùng gene có tốc độ tiến hóa cao hơn như *cytb* và *D-loop*.

Gene *16S rRNA* – với đặc tính bảo tồn cao – cho thấy số lượng vị trí biến dị tương đối hạn chế ở hầu hết các loài, song các biến dị này thường có tính ổn định và xuất hiện đồng thời tại nhiều vị trí liên tiếp. Ngược lại, gene *cytb* và vùng *D-loop* thể hiện mức độ biến dị cao hơn rõ rệt, phù hợp với vai trò của các vùng này trong việc phản ánh sự phân hóa gần, sự khác biệt giữa các cá thể và các dòng di truyền địa phương. Việc các biến dị xuất hiện đồng thời trên cả vùng bảo tồn (*16S rRNA*) và vùng mã hóa protein (*cytb*) ở nhiều loài cho thấy đây không phải là các sai khác ngẫu nhiên mà là dấu hiệu của sự phân hóa di truyền có hệ thống.

Đối với các loài *chà vá*, cá thể linh trưởng ở Tây Nguyên thường mang kiểu nucleotide riêng biệt và chia sẻ rất ít alen với các mẫu tham chiếu từ các khu vực khác. Tuy nhiên, mức độ sai khác này chưa đủ lớn để làm thay đổi vị trí phân nhánh

trên cây phát sinh loài, cho thấy các loài này vẫn duy trì tính thống nhất về mặt phát sinh loài, đồng thời tồn tại sự đa dạng di truyền nội loài đáng kể.

Đối với các loài thuộc chi *Macaca*, đặc biệt là *Khỉ mặt đỏ* và *Khỉ vàng*, kết quả phân tích cũng cho thấy xu hướng tương tự: các cá thể linh trưởng ở Tây Nguyên xuất hiện ở nhánh riêng biệt trên cây phát sinh loài, phản ánh sự khác biệt di truyền đáng kể so với các cá thể ở khu vực khác, trong khi vẫn duy trì mối quan hệ họ hàng gần gũi trong nội bộ chi.

Xét tổng thể, sự nhất quán giữa các cây phát sinh loài xây dựng từ từng vùng gene riêng lẻ và từ toàn bộ hệ gene ty thể khẳng định độ tin cậy của dữ liệu ty thể trong nghiên cứu phát sinh loài linh trưởng. Những kết quả này phù hợp với các nghiên cứu gần đây về tiến hóa và đa dạng di truyền linh trưởng ở khu vực Đông Dương và Đông Nam Á, nơi các cao nguyên và dãy núi đóng vai trò như các "đảo sinh học" thúc đẩy quá trình phân hóa di truyền nội loài (8,64,74)

KẾT LUẬN VÀ KIẾN NGHỊ

KẾT LUẬN

Nghiên cứu này đã giải trình tự và phân tích hệ gene ty thể hoàn chỉnh của năm loài linh trưởng đại diện phân bố tại khu vực Tây Nguyên Việt Nam. Cách tiếp cận sử dụng một cá thể đại diện cho mỗi loài được thực hiện nhằm mô tả đặc điểm ADN ty thể và phân tích mối quan hệ phát sinh loài trong bối cảnh so sánh liên loài, phù hợp với mục tiêu nghiên cứu của luận án.

Kết quả cho thấy hệ gene ty thể của các loài nghiên cứu có cấu trúc điển hình của linh trưởng, với sự bảo tồn cao về thành phần gene, vị trí các gene và đặc điểm chức năng. Tuy nhiên, vẫn tồn tại những sai khác trình tự đặc trưng giữa các loài và giữa các cá thể đại diện so với các trình tự đã công bố, phản ánh tín hiệu phân hóa di truyền ở mức liên loài và liên chi, phục vụ phân tích phát sinh loài.

Phân tích phát sinh loài dựa trên từng vùng gene riêng lẻ và toàn bộ hệ gene ty thể cho thấy các cây phát sinh thu được có tính nhất quán cao, củng cố vị trí phân loại của các loài linh trưởng Tây Nguyên Việt Nam trong mối quan hệ với các loài họ hàng gần. Các kết quả này khẳng định giá trị của dữ liệu ty thể trong việc làm rõ quan hệ tiến hóa giữa các loài linh trưởng, đặc biệt trong bối cảnh khu vực Đông Nam Á có mức độ đa dạng và đặc hữu cao.

Mặc dù số lượng mẫu còn hạn chế và chưa cho phép đánh giá đa dạng di truyền ở mức quần thể, nghiên cứu này đã cung cấp bộ dữ liệu di truyền nền có giá trị, góp phần bổ sung cơ sở dữ liệu ADN ty thê cho các loài linh trưởng Việt Nam. Những kết quả thu được là cơ sở khoa học quan trọng cho các nghiên cứu tiếp theo với cỡ mẫu lớn hơn, phục vụ phân tích di truyền quần thể và hỗ trợ công tác bảo tồn các loài linh trưởng nguy cấp tại Việt Nam.

KIẾN NGHỊ

1. Tăng cường nghiên cứu di truyền học phân tử trong bảo tồn linh trưởng; các kết quả từ luận án cho thấy việc ứng dụng các trình tự gene ty thể như *cytb*, *16S rRNA* và *D-loop* mang lại dữ liệu chính xác về quan hệ phát sinh loài và mức độ khác biệt di truyền của linh trưởng tại Tây Nguyên. Tuy nhiên các nghiên cứu chỉ dừng lại ở một cá thể riêng lẻ. Do đó, cần mở rộng nghiên cứu trên đơn vị quần thể và mở rộng phạm vi nghiên cứu sang các vùng khác như Đông Bắc, Bắc Trung Bộ và Tây Nam Bộ để có cái nhìn toàn diện hơn về quần thể linh trưởng Việt Nam.
2. Bổ sung dữ liệu di truyền vào hệ thống quản lý bảo tồn quốc gia:

Kiến nghị xây dựng một ngân hàng dữ liệu gene linh trưởng Việt Nam tích hợp dữ liệu hình thái, sinh thái và di truyền – nhằm phục vụ việc nhận diện nhanh các loài, cá thể trong cứu hộ, tái thả và giám sát quần thể trong tự nhiên.

3. Đẩy mạnh hợp tác quốc tế và đào tạo nguồn nhân lực:

Kiến nghị mở rộng hợp tác với các tổ chức quốc tế như IUCN Primate Specialist Group, Viện Max Planck (Đức), Đại học Kyoto (Nhật Bản) để cập nhật phương pháp, thiết bị, và đào tạo chuyên gia trẻ trong lĩnh vực di truyền học bảo tồn.

DANH MỤC CÔNG TRÌNH CÔNG BỐ LIÊN QUAN ĐẾN LUẬN ÁN

* Danh mục tạp chí SCI-E

1. Mai Thi Phuong Nguyen[†]& **Tram Thi Thuy Nguyen**[†], Tung Thanh Ha, Chi Nguyen Quynh Ho, Cuong Phan Minh Le, Huy Nghia Quang Hoang, Quynh Thi Nhu Nguyen, Tao Thien Nguyen, Dung Tri Luu, Khoa Dang Dang, Quan Ke Thai and Long Thanh Le*. (2024). Characterization of the Complete Mitochondrial Geneome of the Central Highland Grey-Shanked Douc Langur (*Pygathrix cinerea*), a Critically Endangered Species Endemic to Vietnam (Mammalia: Primates). *Curr Issues Mol Biol* 46(9): 9928-9947.https://doi.org/10.3390/cimb46090592

2. Tram Thi Thuy Nguyen, Chi Nguyen Quynh Ho, Cuong Phan Minh Le, Nhan Lu Chinh Phan, Loan Tung Dang, Quan Ke Thai, Mai Thi Phuong Nguyen, Phuc Bui Tran, Long Thanh Le, Vu Quang Manh. (2025). The complete mitochondrial geneome of the Vietnamese *macaca mulatta. Natural history* 126(3): 2-20. https://doi.org/10.59879/1Ia6i

* Danh mục tạp chí Quốc gia

3. **Nguyen Thi Thuy Tram**, Le Thanh Long, Ho Nguyen Quynh Chi, Phan Lu Chinh Nhan, Ha Thanh Tung, Dang Thi Tung Loan, Nguyen Thi Phuong Mai (2025) Analysis of the complete mitochondrial geneome of Vietnamese *Pygathrix nigripes*. *Vietnam Journal of Biotechnology* 23(1): 25-39. DOI: 10.15625/vjbt-21785

DANH MỤC TÀI LIỆU THAM KHẢO

1. Mittermeier RA, Rylands AB, Schwitzer C, 2022, IUCN SSC Primate Specialist Group: Report 2018–2021, Rep Int Primatol Soc IPS Quito Ecuad.
2. Roos C, Boonratana R, Supriatna J, Fellowes JR, Groves CP, Nash SD, et al, 2014, An updated taxonomy and conservation status review of Asian primates, *Asian Primates J*.
3. Estrada A, Garber PA, Rylands AB, Roos C, Fernandez-Duque E, Di Fiore A, et al, 2017, Impending extinction crisis of the world's primates: Why primates matter, *Sci Adv*, 3(1):e1600946.
4. Federman S, Dornburg A, Daly DC, Downie A, Perry GH, Yoder AD, et al, 2016, Implications of lemuriform extinctions for the Malagasy flora. Proc Natl Acad Sci, 113(18):5041–6.
5. Evolutionary Human Sciences, 2025, Insights into human evolution from 60 years of research on chimpanzees at Gombe, *Cambridge Core*.
6. Wolfe ND, Escalante AA, Karesh WB, Kilbourn A, Spielman A, Lal AA, 1998, Wild primate populations in emerging infectious disease research: the missing link? *Emerg Infect Dis,* 4(2):149.
7. Phillips KA, Bales KL, Capitanio JP, Conley A, Czoty PW, 'T Hart BA, et al, 2014, Why primate models matter. Am J Primatol, 76(9):801–27.
8. Fan PF, Fei HL, Luo AD, 2014, Ecological extinction of the critically endangered northern white-cheeked gibbon Nomascus leucogenys in China. *Oryx*. 48(1):52–5.
9. Blair ME, Sterling EJ, Hurley MM, 2011, Taxonomy and conservation of Vietnam's primates: a review. *Am J Primatol*. 73(11):1093–106.
10. Covert HH, Duc HM, Quyet LK, Ang A, Harrison-Levine A, Van Bang T, 2017, Primates of Vietnam: Conservation in a Rapidly Developing Country. *Anthropol Now,* 9(2):27–44.
11. Nicholls DG, 1979, Brown adipose tissue mitochondria. *Biochim Biophys Acta BBA-Rev Bioenerg*, 549(1):1–29.
12. Cann RL, Stoneking M, Wilson AC, 1987, Mitochondrial DNA and human evolution, *Nature*, 325(6099):31–6.
13. Galtier N, Nabholz B, Glémin S, Hurst GDD, 2009, Mitochondrial DNA as a marker of molecular diversity: a reappraisal, *Mol Ecol*, 18(22):4541–50.
14. Gagneux P, Wills C, Gerloff U, Tautz D, Morin PA, Boesch C, et al, 1999, Mitochondrial sequences show diverse evolutionary histories of African hominoids, *Proc Natl Acad Sci*, 96(9):5077–82.

15. Muir CC, Galdikas BMF, Beckenbach AT, 2000, *mtDNA* Sequence Diversity of Orangutans from the Islands of Borneo and Sumatra, *J Mol Evol*, 51(5):471–80.
16. Warren BH, Bermingham E, Bowie RC, Prys-Jones RP, Thébaud C, 2003, Molecular phylogeography reveals island colonization history and diversification of western Indian Ocean sunbirds (Nectarinia: Nectariniidae), *Mol Phylogenet Evol*, 29(1):67–85.
17. Van Tien D, 1960, Recherches zoologiques dans la region de Vinh-Linh, *Zool Anz*, 164:221–39.
18. Van Tien D, 1991, Recherches zoologiques dans la région de Thai-Nguyen (Nord Vietnam), *Zool Anz*, 166:298–308.
19. Van Tien D, 1970, Sur les Formes de L'écureuil de MacClelland, Callosciurus (Tamiops) MacClellandi Horsfield (Sciuridae, Rodentia) au Vietnam. *Mitteilungen Aus Dem Mus Für Naturkunde Berl Zool Mus Inst Für Spez Zool Berl*, 46(1):53–60.
20. Eames JC, Lambert FR, Cu N, 1994, A survey of the Annamese Lowlands, Vietnam, and its implications for the conservation of Vietnamese and Imperial Pheasants Lophura hatinhensis and L. imperialis, *Bird Conserv Int*, 4(4):343–82.
21. Nadler T, 2010, Status of Vietnamese primates-complements and revisions. *Conserv Primates Indoch*, 3–16.
22. Long HT, 2020, Feeding behaviour and diet of grey-shanked douc langurs (Pygathrix cinerea) in Kon ka Kinh National Park, Vietnam, *Vietnam J Primatol*, 3(2):59–83.
23. Vu TT, Tran DV, Tran HTP, Nguyen MD, Do TA, Ta NT, et al, 2020, An assessment of the impact of climate change on the distribution of the grey-shanked douc *Pygathrix cinerea* using an ecological niche model. *Primates*, 61(2):267–75.
24. Lippold LK, Vu NT, Tran DN, Nguyen AD, 2022, Feeding ecology of red-shanked douc langurs at Son Tra nature reserve, Da Nang, Vietnam, *Primate Conserv*, 36:141–54.
25. Pozzi L, Hodgson JA, Burrell AS, Sterner KN, Raaum RL, Disotell TR, 2014, Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes, *Mol Phylogenet Evol*, 75:165–83.
26. Giang CTH, Van Thanh N, Ha NM, Van NTH, Hai ND, Huong DT, 2016, Assessing mammal diversity in Xuan Lien Nature Reserve using molecular approaches, *Acad J Biol*, 38(2):171–8.
27. Donnelly D, 2024, *Endangered Species Conservation in Vietnam*.

28. Báo cáo kinh nghiệm quốc tế trong quản lý môi trường, 2020, *Bộ tài nguyên và môi trường.*
29. Nghị định số 84/2021/NĐ-CP của Chính phủ: Sửa đổi, bổ sung một số điều của Nghị định số 06/2019/NĐ-CP ngày 22 tháng 01 năm 2019 của Chính phủ về quản lý thực vật rừng, động vật rừng nguy cấp, quý, hiếm và thực thi Công ước về buôn bán quốc tế các loài động vật, thực vật hoang dã nguy cấp
30. Roos C, Helgen KM, Miguez RP, Thant NML, Lwin N, Lin AK, et al, 2020, Mitogenomic phylogeny of the Asian colobine genus *Trachypithecus* with special focus on *Trachypithecus phayrei* (Blyth, 1847) and description of a new species, *Zool Res*, 41(6):656.
31. Maheshwari A, Snowden E, Van Hoang L, 2023, Vietnam's vital role in primate conservation, *Science*, 382(6672):777–777.
32. Avise JC, 2000, Phylogeography: the history and formation of species, *Harvard university press*.
33. Guillén AKZ, Barrett GM, Takenaka O, 2005, Genetic diversity among African great apes based on mitochondrial DNA sequences, *Biodivers Conserv*, 14(9):2221–33.
34. Noda R, Kim CG, Takenaka O, Ferrell RE, Tanoue T, Hayasaka I, et al, 2001, Mitochondrial 16S rRNA sequence diversity of *hominoids*, *J Hered*, 92(6):490–6.
35. Finstermeier K, Zinner D, Brameier M, Meyer M, Kreuz E, Hofreiter M, et al, 2013, A mitogenomic phylogeny of living primates, *PloS One*, 8(7):e69504.
36. Johns GC, Avise JC, 1998, A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome b gene, *Mol Biol Evol*, 15(11):1481–90.
37. Bataille M, Crainic K, Leterreux M, Durigon M, de Mazancourt P, 1999, Multiplex amplification of mitochondrial DNA for human and species identification in forensic evaluation, *Forensic Sci Int*, 99(3):165–70.
38. Zamyatnin AA, Belozerskaya TA, Zamyatnin AA, 2021, Taxonomy of Mitochondrial Cytochrome B Proteins of the Same Amino Acid Sequence Length, *Sci World J*, 2021:1–9.
39. Degli Esposti M, De Vries S, Crimi M, Ghelli A, Patarnello T, Meyer A, 1993, Mitochondrial cytochrome b: evolution and structure of the protein. *Biochim Biophys Acta BBA-Bioenerg*, 1143(3):243–71.

40. Yoder AD, Cartmill M, Ruvolo M, Smith K, Vilgalys R, 1996, Ancient single origin for Malagasy primates, *Proc Natl Acad Sc*i, 93(10):5122–6. doi:10.1073/pnas.93.10.5122
41. Zhang Y ping, Ryder OA, 1998, Mitochondrial cytochrome b gene sequences of old world monkeys: With special reference on evolution of Asian colobines, *Primates*, 39(1):39–49.
42. Nater A, Mattle-Greminger MP, Nurcahyo A, Nowak MG, De Manuel M, Desai T, et al, 2017, Morphometric, behavioral, and genomic evidence for a new orangutan species, *Curr Biol*, 27(22):3487–98.
43. Reis MD, Gunnell GF, Barba-Montoya J, Wilkins A, Yang Z, Yoder AD, 2018, Using phylogenomic data to explore the effects of relaxed clocks and calibration strategies on divergence time estimation: primates as a test case. *Syst Biol,* 67(4):594–615.
44. Ruiz-Garcia M, Cerón Á, Sánchez-Castillo S, Rueda-Zozaya P, Pinedo-Castro M, Gutierrez-Espeleta G, et al, 2017, Phylogeography of the mantled howler monkey (Alouatta palliata; Atelidae, Primates) across its geographical range by means of mitochondrial genetic analyses and new insights about the phylogeny of Alouatta, *Folia Primatol (Basel),* 88(5):421–54.
45. Clayton DA, 1982, Replication of animal mitochondrial DNA, *Cell*, 28(4):693–705.
46. Sbisà E, Tanzariello F, Reyes A, Pesole G, Saccone C, 1997, Mammalian mitochondrial *D-loop* region structural analysis: identification of new conserved sequences and their functional and evolutionary implications, *Gene,* 205(1–2):125–40.
47. Shadel GS, Clayton DA, 1997, Mitochondrial dna maintenance in vertebrates, *Annu Rev Biochem*, 66(1):409–35.
48. Saccone C, De Giorgi C, Gissi C, Pesole G, Reyes A, 1999, Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system, *Gene*, 238(1):195–209.
49. Vilà C, Maldonado JE, Wayne RK, 1999, Phylogenetic relationships, evolution, and genetic diversity of the domestic dog, *J Hered*, 90(1):71–7.
50. Moritz C, 1994, Defining ‘evolutionarily significant units’ for conservation, *Trends Ecol Evol,* 9(10):373–5.
51. Wan Q, Wu H, Fujihara T, Fang S, 2004, Which genetic marker for which conservation genetics issue?, *Electrophoresis*, 25(14):2165–76.

52. Masila EM, Ogada SO, Ogali IN, Kennedy GM, Too EK, Ommeh CS, 2024, Mitochondrial DNA D-Loop Polymorphisms among the Galla Goats Reveals Multiple Maternal Origins with Implication on the Functional Diversity of the HSP70 Gene, *Genet Res*, 2024:1–15.
53. Demir E, Moravčíková N, Argun Karsli B, Kasarda R, Aytekin I, Bilginer U, et al, 2023, Mitochondrial DNA diversity of D-loop region in three native Turkish cattle breeds, *Arch Anim Breed*, 66(1):31–40.
54. Wang X, Pei J, Xiong L, Bao P, Chu M, Ma X, et al, 2024, Genetic diversity, phylogeography, and maternal origin of yak (Bos grunniens), *BMC Genomics*, 15;25(1).
55. Srikulnath K, Ariyaraphong N, Singchat W, Panthum T, Lisachov A, Ahmad SF, et al, 2022, Asian elephant evolutionary relationships: New perspectives from mitochondrial D-loop haplotype diversity, *Sustainability,* 15(1):720.
56. Vinh NT, DaNg PK, PhuoNg TB, Giang NTP, Do L, Bo HX, 2024, Genetic diversity of Vietnamese native ducks based on sequences of the mitochondrial DNA d-loop region, *Adv Anim Vet Sci*,13(1):73–80.
57. Clarridge JE, 2004, Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases, *Clin Microbiol Rev*,17(4):840–62.
58. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP, 2016, DADA2: High-resolution sample inference from Illumina amplicon data, *Nat Methods*, 13(7):581–3.
59. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al, 2012, The SILVA ribosomal RNA gene database project: improved data processing and web-based tools, *Nucleic Acids Res,* 41(D1):D590–6.
60. Macholdt E, Arias L, Duong NT, Ton ND, Van Phong N, Schröder R, et al, 2020, The paternal and maternal genetic history of Vietnamese populations, *Eur J Hum Genet*, 28(5):636–45.
61. Roos C, Thanh VN, Walter L, Nadler T, 2007, Molecular systematics of Indochinese primates, *Vietn J Primatol*, 1(1):41–53.
62. Nadler T, Thanh VN, Streicher U, 2007, Conservation status of Vietnamese primates, *Vietnam J Primatol,* 1(1):7–26.
63. Thinh VN, Rawson B, Hallam C, Kenyon M, Nadler T, Walter L, et al, 2010, Phylogeny and distribution of crested gibbons (genus *Nomascus*) based on mitochondrial cytochrome b gene sequence data, *Am J Primatol,* 72(12):1047–54.

64. Roos C, Miguez RP, Zinner D, Nadler T, Thinh VN, 2022, Importance of genetic data from type specimens: The questionable type locality of southern white-cheeked gibbon, Nomascus siki (Delacour, 1951), *Zool Res,* 43(4):666.
65. Nguyen MTP, Nguyen TTT, Ha TT, Ho CNQ, Le CPM, Hoang HNQ, et al, 2024, Characterization of the Complete Mitochondrial Genome of the Central Highland Grey-Shanked Douc Langur (*Pygathrix cinerea*), a Critically Endangered Species Endemic to Vietnam (Mammalia: Primates), *Curr Issues Mol Biol*, 46(9):9928–47.
66. Illumina I, 2011, Quality scores for next-generation sequencing, *Tech Note Inform*, 31.
67. Chen S, Zhou Y, Chen Y, Gu J, 2018, fastp: an ultra-fast all-in-one FASTQ preprocessor, *Bioinformatics*, 34(17):i884–90.
68. Gurevich A, Saveliev V, Vyahhi N, Tesler G, 2013, QUAST: quality assessment tool for genome assemblies, *Bioinformatics*, 29(8):1072–5.
69. Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, et al, 2020, GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes, *Genome Biol*, 21(1):241.
70. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, et al, 2013, MITOS: Improved de novo metazoan mitochondrial genome annotation, *Mol Phylogenet Evol*, 69(2):313–9.
71. Iwasaki W, Fukunaga T, Isagozawa R, Yamada K, Maeda Y, Satoh TP, et al, 2013, MitoFish and MitoAnnotator: A Mitochondrial Genome Database of Fish with an Accurate and Automatic Annotation Pipeline, *Mol Biol Evol,* 30(11):2531–40.
72. Lowe TM, Eddy SR, 1997, tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence, *Nucleic Acids Res,* 25(5):955–64.
73. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, et al, 1981, Sequence and organization of the human mitochondrial genome, *Nature*, 290(5806):457–65.
74. Liedigk R, Roos C, Brameier M, Zinner D, 2014, Mitogenomics of the Old World monkey tribe Papionini, *BMC Evol Biol*, 14(1):176.
75. Tran DV, Vu TT, Tran BQ, Nguyen MD, Vu PT, Tran TH, et al, 2024, Modelling the change in the distribution of the black-shanked douc, *Pygathrix nigripes* (Milne-Edwards) in the context of climate change: Implications for conservation, *Raffles Bull Zool*.

76. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al, 2020, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era, *Mol Biol Evol*, 37(5):1530–4.
77. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS, 2017, ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods,14(6):587–9.
78. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS, 2018, UFBoot2: Improving the Ultrafast Bootstrap Approximation, *Mol Biol Evol*, 35(2):518–22.
79. Liedigk R, Yang M, Jablonski NG, Momberg F, Geissmann T, Lwin N, et al, 2012, Evolutionary History of the Odd-Nosed Monkeys and the Phylogenetic Position of the Newly Described Myanmar Snub-Nosed Monkey *Rhinopithecus strykeri*, *PLoS ONE*, 7(5):e37418.
80. Wang ZL, Li C, Fang WY, Yu XP, 2016, The Complete Mitochondrial Genome of two *Tetragnatha* Spiders (Araneae: Tetragnathidae): Severe Truncation of tRNAs and Novel Gene Rearrangements in Araneae, *Int J Biol Sci*, 12(1):109–19.
81. Tamura K, Stecher G, Kumar S, 2021, MEGA11: Molecular Evolutionary Genetics Analysis Version 11, *Mol Biol Evol*, 38(7):3022–7.
82. Tamura K, Nei M, 1993, Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees, *Mol Biol Evol.*
83. Mason VC, Li G, Minx P, Schmitz J, Churakov G, Doronina L, et al, 2016, Genomic analysis reveals hidden biodiversity within colugos, the sister group to primates, *Sci Adv*, 2(8):e1600633.
84. Brown WM, Prager EM, Wang A, Wilson AC, 1982, Mitochondrial DNA sequences of primates: Tempo and mode of evolution, *J Mol Evol,* 18(4):225–39.
85. Melnick DJ, Hoelzer GA, 1993, What is mtDNA good for in the study of primate evolution?, *Evol Anthropol Issues News Rev,* 2(1):2–10.
86. Vilstrup JT, Ho SY, Foote AD, Morin PA, Kreb D, Krützen M, et al, 2011, Mitogenomic phylogenetic analyses of the Delphinidae with an emphasis on the Globicephalinae, *BMC Evol Biol*, 11(1):65.
87. Avise JC, 1994, Molecular Markers, Natural History and Evolution, *Boston, MA: Springer US*
88. Sigurðardóttir S, Helgason A, Gulcher JR, Stefansson K, Donnelly P, 2000, The Mutation Rate in the Human *mtDNA* Control Region, *Am J Hum Genet*, 66(5):1599–609.

89. Prado-Martinez J, Sudmant PH, Kidd JM, Li H, Kelley JL, Lorente-Galdos B, et al, 2013, Great ape genetic diversity and population history, *Nature*, 499(7459):471–5.
90. Cuppari A, Fernández-Millán P, Battistini F, Tarrés-Solé A, Lyonnais S, Iruela G, et al, 2019, DNA specificities modulate the binding of human transcription factor A to mitochondrial DNA control region, *Nucleic Acids Res*, 47(12):6519–37.
91. Brar GA, Weissman JS, 2015, Ribosome profiling reveals the what, when, where and how of protein synthesis, *Nat Rev Mol Cell Biol*, 16(11):651–64.
92. Yang B, Wang Y, Qian PY, 2016, Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis, *BMC Bioinformatics*, 17(1):135.
93. Kim M, Morrison M, Yu Z, 2016, Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes, *J Microbiol Methods*,84(1):81–7.
94. Sloan KE, Warda AS, Sharma S, Entian KD, Lafontaine DLJ, Bohnsack MT, 2017, Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function, *RNA Biol*, 14(9):1138–52.
95. Seligmann H, Krishnan NM, Rao BJ, 2006, Possible multiple origins of replication in primate mitochondria: Alternative role of tRNA sequences, *J Theor Biol*, 241(2):321–32.
96. Sabath N, Graur D, 2010, Detection of Functional Overlapping Genes: Simulation and Case Studies, *J Mol Evol*, 71(4):308–16.
97. Yadav ML, Mohapatra B, 2018, Encyclopedia of Animal Cognition and Behavior, *Cham: Springer International Publishing*, p. 1–5.
98. Sardana R, O'Dell M, Flavell R, 1993, Correlation between the size of the intergenic regulatory region, the status of cytosine methylation of rRNA genes and nucleolar expression in wheat, *Mol Gen Genet MGG,* 236–236(2–3):155–62.
99. Wang W, Liu JY, Wang HF, Yang MY, Liu QY, Ding MX, 2016, The complete mitochondrial genome of white-tufted-ear marmoset, *Callithrix jacchus* (Primates: Callitrichinae), *Mitochondrial DNA Part A*, 27(3):1920–1.
100. Xia W, Jiang H, Fan W, Li X, Liu Y, Xu L, et al, 2023, Complete mitochondrial genome of *Mastigias papua* (Scyphozoa: Rhizostomeae: Mastigiidae) based on next-generation sequencing and phylogenetic analysis, *Mitochondrial DNA Part B*, 8(11):1294–7.

101. Bergey CM, Trujillo AE, Disotell TR, Tosi AJ, Burrell AS, 2020, Genomics of adaptation, convergence, and evolutionary conservation in haplorrhine primates, *Am J Biol Anthropol*.

102. Faith JJ, Pollock DD, 2003, Likelihood Analysis of Asymmetrical Mutation Bias Gradients in Vertebrate Mitochondrial Genomes, *Genetics*, 165(2):735–45.

103. Min XJ, Hickey DA, 2007, DNA Barcodes Provide a Quick Preview of Mitochondrial Genome Composition, *Plos One*, 2(3):e325.

104. Tanaka M, Ozawa T, 1994, Strand Asymmetry in Human Mitochondrial DNA Mutations, *Genomics*, 22(2):327–35.

105. Lavrov DV, Brown WM, Boore JL, 2000, A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. *Proc Natl Acad Sci*, 97(25):13738–42.

106. Boore JL, 1999, Animal mitochondrial genomes, *Nucleic Acids Res*, 127(8):1767–80.

107. Ha NH, Dam VM, Van Dung T, 2023, The population status of the black-shanked douc langurs (Pygathrix nigripes) in Chua Chan Mountain, Dong Nai Province, Vietnam, *Vietnam J Primatol*, 3(4):69–78.

108. Tran DV, Vu TT, Tran BQ, Nguyen MD, Vu PT, Tran TH, et al, 2024, Modelling the change in the distribution of the black-shanked douc, Pygathrix nigripes (Milne-Edwards) in the context of climate change: Implications for conservation, *Raffles Bull Zool,* 13: 68-77.

109. Freund D, Signs M, Yoganand K, 2021, Primates of the Greater Mekong: Status, Threats and Conservation Efforts, *WWF-Greater Mekong*, p. 74.

110. Perna NicoleT, Kocher ThomasD, 1995, Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes, *J Mol Evol,* 41(3).

111. Medina NS, Moreno-Carmona M, Canal NA, Prada-Quiroga CF, Medina NS, Moreno-Carmona M, et al, 2025, Mitochondrial Genome Variations and Possible Adaptive Implications in Some Tephritid Flies (Diptera, Tephritidae), *Int J Mol Sci*, 26(12).

112. Nabholz B, Ellegren H, Wolf JB, 2012, High levels of gene expression explain the strong evolutionary constraint of mitochondrial protein-coding genes, *Mol Biol Evol*, 30(2):272–84.

113. Pataradool T, Siriyasatien P, Sukhumavasi W, Tiwananthagorn S, Suankratay C, Preativatanyou K, 2025, First complete mitochondrial genome of Armillifer moniliformis (Pentastomida: Porocephalida) isolated from a human case in

Northern Thailand: comparative and phylogenetic analyses, *Parasitol Res,* 124(6):69.

114. Li Y, Thomas GWC, Richards S, Waterhouse RM, Zhou X, Pfrender ME, 2024, Rapid evolution of mitochondrion-related genes in haplodiploid arthropods, *BMC Biol*, 22(1):229.

115. Zhang RQ, Bai YW, Hu JT, Wang PY, Qi Y, Zhang TX, et al, 2024, Complete mitochondrial genomes and phylogenetic analysis of native and non-native fishes in a national key wetland of China, *Front Environ Sci*, 12.

116. Zardoya R, 2020, Recent advances in understanding mitochondrial genome diversity, *F1000Research*.

117. Roller M, Lucić V, Nagy I, Perica T, Vlahoviček K, 2013, Environmental shaping of codon usage and functional adaptation across microbial communities, *Nucleic Acids Res*. 41(19):8842–52.

118. Parvathy ST, Udayasuriyan V, Bhadana V, 2022, Codon usage bias, *Mol Biol Rep*, 49(1):539–65.

119. Da Silva MJF, Minhos T, Sá RM, Bruford M, 2012, Using genetics as a tool in primate conservation, *Nat Educ Knowl*, 3(10):89.

120. Wang W, Zheng Y, Zhao J, Yao M, 2019, Low genetic diversity in a critically endangered primate: shallow evolutionary history or recent population bottleneck?, *BMC Evol Biol*, 19(1):134.

121. Smith DG, McDonough J, 2005, Mitochondrial DNA variation in Chinese and Indian rhesus macaques (Macaca mulatta), *Am J Primatol*, 65(1):1–25.

122. Spuhler JN, 1988, Evolution of mitochondrial DNA in monkeys, apes, and humans, *Am J Phys Anthropol*, 31(S9):15–48.

123. Zhang QL, Zhang L, Zhao TX, Wang J, Zhu QH, Chen JY, et al, 2017, Gene sequence variations and expression patterns of mitochondrial genes are associated with the adaptive evolution of two Gynaephora species (Lepidoptera: Lymantriinae) living in different high-elevation environments, *Gene*, 610:148–55.

124. Sahyoun AH, Bernt M, Stadler PF, Tout K, 2014, GC skew and mitochondrial origins of replication, *Mitochondrion*, 17:56–66.

125. Bett NN, Blair ME, Sterling EJ, 2012, Ecological Niche Conservatism in Doucs (Genus Pygathrix), *Int J Primatol*, 33(4):972–88.

126. terling EJ, Hurley MM, 2025, Vietnam: a natural history, *Yale University Press.*

PHỤ LỤC

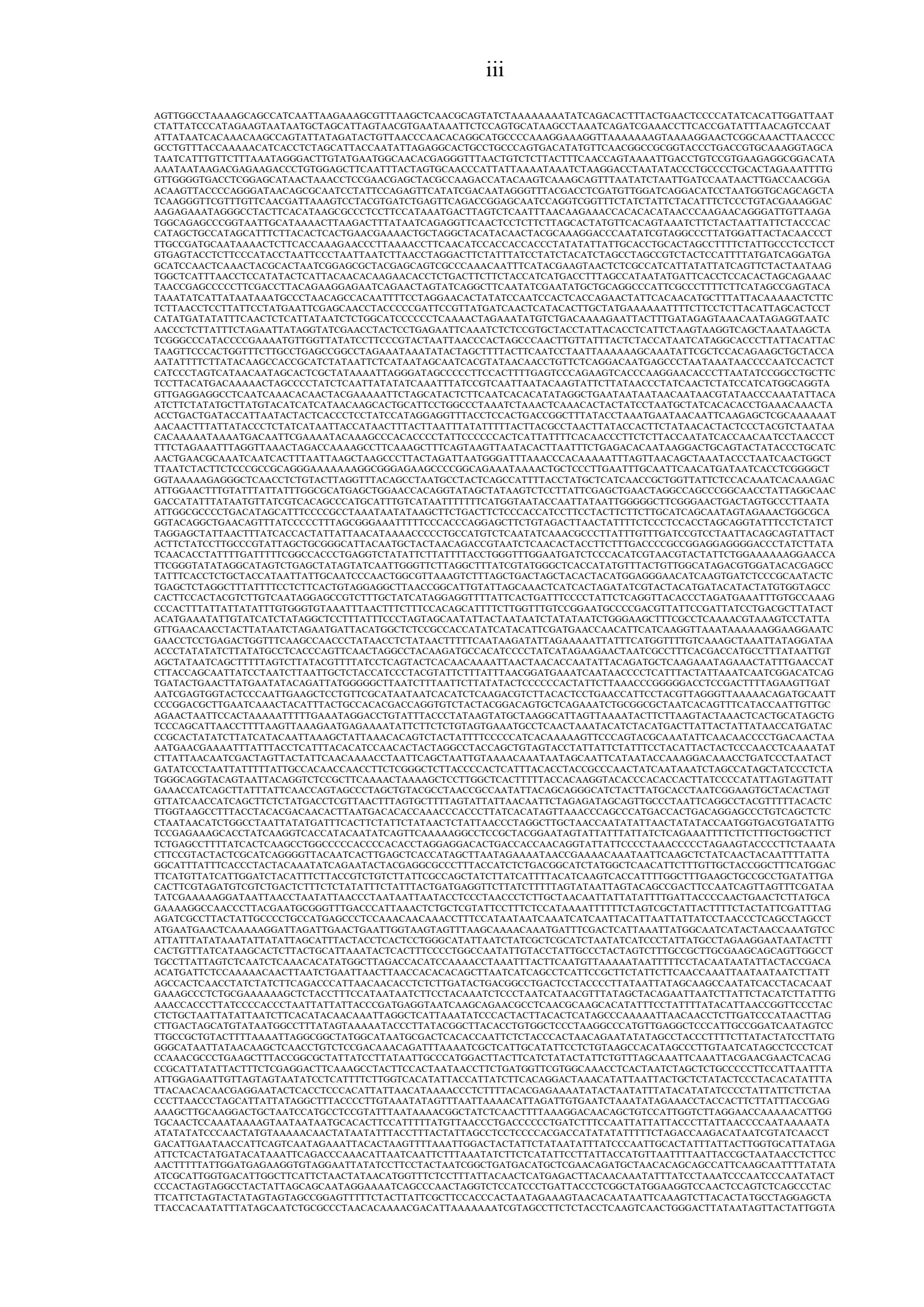
Bảng S1. Trình tự nucleotide trên ADN ty thể *Chà vá chân xám* Tây Nguyên Việt Nam

TATCTTTTAGGTGTGCATAACTCAACTGTCGTTTTCTCAACTAATAAAAATTTATTTCACACACCACCCTTGACACCTCCACCCACCTCTTCTTTCCCATAATCCCCAA
AGATGGTATCATAATACCCATTATTCAGTTTATGTAGCTTAACACCTCCAAAGCAAGACACTGAAAATGCCTAGACGGGCTCACAGACCCCATAAACAAACAGGTTT
GGTCCCGGCCTTTCTATTAGCTTTTTAGCAAAATTACACATGCAAGCATCCCCGCTCCAGTGAAGATACCCTATAAATCACGACGATTAAGAGGAGTAAGTATCAAG
CACGCTTAATGCAGCTCAAGACACTTTGCTTAGCCACACCCCCACGGGAAACAGCAGTGACTAATATTTAGCAATAAACGAAAGTTTAACTAAGCTATGCTACTAAG
GGTTGGTCAATTTCGTGCCAGCCACCGCGGCCATACGATTGACCCTAGTTAATAGACCTCGGCGTAAAGGGTGTTTTAGATAAATTTAATTAAATAAAGCTAAACTC
TTTCTAAACTGTAAAATCCTAGCTAATGTAAAATAGACTACGAAAGTGGCTTTAAAGCTTCTGAATACACAACAGCTAAGACTCAAACTGGGATTAGATACCCCACT
ATGCTTAGCCCTAAACTTTAATAGTTAAAAATAACAAAACTATTCGCCAGAACACTACAAGCAACAGCTTAAAATTTAAAGGACTTGGCGGTGCTCCATATCCCCCT
AGAGGAGCCTGTTCTATAATCGATAAACCCCGATTCACCTCACCACCTCTTGCTAAGCCTATATACCGCCATCTTCAGCAAACCTCAACAAGAGATATAAAGTAAGC
ACAAACGCCCACGCAAAAACGTTAGGTCAAGGTGTAGCTTATGAGCTGGAAGAAATGGGCTACATTTTCTACCCTAGAAAACCCCACGATAGCTCTTATGAAACTTA
AGAGTCCAAGGAGGATTTAGTAGTAAACTAAGAATAGAGTGCTTAGTTGAACTAGGCCATGAAGCACGCACACACCGCCCGTCGCTCTCCTCAAATATAATTAGAA
CTAACTTAACTAAATATTTCTGCATTCATACAGAGGAGATAAGTCGTAACATGGTAAGTGTACTGGAAAGTGCACTTGGATAAATCAAGGCATAGCTCAACACAAA
GCATCCAGCTTACACCCGGAAGATATCGCTACCTGATTGCCTTGAGCCAATACTAGCTCAACAAACCTACCAATACTACTATTAAATTAATATATTATTAAACCATTC
ACAAGTATAAAAGTATAGGAGATAGAAATTTACTCTGACGCTATAGATGCAGTACCGCAAGGGAAAGATGAAAAAAGCAACCAAGCATAAAACAGCAAGGACTCA
CCCCTATACCTTTTGCATAATGAACTAGCTAGAGACCACTTCGCAAAGAGAACTAAAGCCAAGTACCCCGAAACCAGACGAGCTACCCAAAAACAGCTAAAAGAGC
ACACCCGTCTATGTAGCAAAATAGTGGGAAGATTTCTGGGTAGAGGTGATACGCCTACCGAGCCTGGAGATAGCTGGTTATCCAAGATAGAATTTTAGTTCAACCTT
AAGTTTACTTACAGAACTACCAATCCTCCTGTAAACTTAACTGTTAGTCTAAAGAGGGACAGCTCTTTAGACATCAGGAAAAAACCTTATGTAGAGAGTAAAAATT
TCCAAACTCATAGTTGGCCTAAAAGCAGCCATCAATTAAGAAAGCGTTTAAGCTCAACGCAATGTCTAAAAAAATATCAAACACTTTACTGAACTCCTCATATCACA
TTGGATTAATCTATTATTTCATAGAAGTAATAATGCTAGCATTAGTAACGTGAATAAATTCTCCAGTGCATAAGCCTAAATCGGATCGAAACCTTCACCGATATTTGA
CAGTCTAATGTTATAACCACAAACAAGCCAATATTATATACTGTTAACCCAACACAGGCATGCCCTAAAGGAAAGGTTAAAAAAAGTAAAAAGGAACTCGGCAAA
CTCAACCCCGCCTGTTTACCAAAAACATCACCTCTAGCATTACCAGTATTAGAGGCACTGCCTGCCCAGTGACATATGTTTAACGGCCGCGGTACCCTGACCGTGCA
AAGGTAGCATAATCATTTGTTCTTTAAATAGGGACTTGTATGAATGGCAACACGAGGGTTTAACTGTCTCTTACTTTCAACCAGTAAAATTGACCTGTCCGTGAAGAG
GCGGACATAAAATAATAAGACGAGAAGACCCTGTGGAGCTTCAATTTACTAGTACAACCCATTATTAAAATAAATCTAAGGACCTAACATACCCTGCCCCTGTACTA
GAAATTTTGGTTGGGGTGACCTCGGAGCATAATTAAACCTCCGAACGAACTACGCCAAGGCCATACAAGTCAAAGCAGTCTAATATCTAAATTGATCCAATAACTTG
ACCAACGGAACAAGTTACCCCAGGGATAACAGCGCAATCCTATTCCAGAGTCCATATCGACAATAGGGTTTACGACCTCGATGTTGGATCAGGACATCCTAATGGTG
CAGCAGCTATCAAGGGTTCGTTTGTTCAACGATTAAAGTCCTACGTGATCTGAGTTCAGACCGGAGCAATCCAGGTCGGTTTCTATCTATTCTATATTTCTCCCTGTA
CGAAAGGACAAGAGAAATAGGGCCTACTTCACATAAGCGCCCTCCTCCCATAAATGACCTAGTCTCAATTTAGCAAGAAATTACACACACACAACCCAAGAACAGG
GATTGTTAAGATGGCAGAGCCCGGTAATTGCATAAAATTTAAGACTTTATAATCAGAGGTTCAACTCCTCTTCTTAGCACTATGTTCACAATAAATCTTCTACTCATT
ATCCTACCCACTATAGCTGCCATAGCATTCTTACACTTACTGAACGAAAACTATTAGGCTATATACAACTACGCAAAGGACCCAACATCGTGGGTCCTTATGGACT
ACTACAACCCTTTGCCGATGCAATAAAACTCTTCACCAAAGAACCCTTAAAACCCTCAACATCCACTACCACCCTATATATTATTGCACCCGCCCTAGCCTTTTCTAT
TGCCCTTCTCCTGTGAGTACCTCTTCCCATACCCAATTCCCTAATTAATCTTAATCTAGGACTTCTATTTATCCTAGCTACATCTAGTCTAGCTGTTTACTCCATTTTAT
GATCAGGATGAGCATCCAACTCAAACTACGCATTAATCGGAGCGCTACGAGCAGTCGCCCAAACAATTTCATACGAGGTAACTCTCGCCATCATTATACTATCAGTT
CTACTAATAAGTGGCTCATTCAACCTCCACGCACTCATTACAACACAAGAACACCTCTGACTTCTCCTACCATCATGACCTTTAACCATAATATGATTCACCTCCACA
CTAGCAGAAACCAATCGAGCCCCCTTCGACCTCACAGAAGGAGAATCAGAACTAGTATCAGGCTTCAATATCGAGTACGCCGCAGGTCCATTCGCTCTTTTCTTCAT
AGCCGAATATATAAATATTATTATAATAAATGCCCTAACAGCCACAATTTTTCTAGGGACACTATACCCAATCCACTCACCAGAACTATTCACAACATGCTTTGTTAC
AAAAACTCTTCTCCTAACCTCCTTATTCCTATGGATTCGAGCAACCTACCCCCGATTCCGTTATGATCAACTCATACACTTACTATGAAAAAATTTCCTTCCTCTCACA
TTAGCACTCCTCATATGATATATCTCAACTCTTATTATAACCTCTGGCATCCCCCCTCAAAGCTAGAAATATGTCTGACAAAAGAATTACTTTGATAGAGTAAATAAT
AGAGGTATTCAACCCTCTTATTTCTAGAATTATAGGTATTGAACCTACTCCTGAGAATCCAAATCTCTCCGTGCTACCTATTACACCTCATTCTAAGTAAGGTCAGCT
AAATAAGCTATCGGGCCCATACCCCGAAAATGTTGGTTATACCCTTCCCGTACTAATTAATCCACTAGCCCAACTTGTTATCTACTCTACCATAATCATAGGCACCCT
TATTACATCACTAAGTTCCCACTGATTTCTAGCCTGAACCGGCCTAGAGATAAATATACTAGCTTTTACCTCAATCCTAATTAAAAAAGCAAACATTCGCTCCACAGA
AGCTGCTACCAAGTATTTCCTTACACAAGCCACCGCATCTATAATTCTCATAATAGCAATCATGTATAACAACCTACTCTCAGGACAATGAACCCTGATAAACAATC
CCAATCAACTCTCATCCTTAATCATAACAATAGCACTCGCTATAAAATTAGGAATAACCCCCTTCCACTTTTGAGTTCCAGAAGTCACCCAAGGAACACCCTTAATAT
CCGGCCTGCTTCTCCTCACATGACAAAAACTAGCCCCTATCTCAATTATATATCAAATTTATCCATCAATTAACACAAGTATTCTTATAACCCTATCAACCCTATCCAT
CATAGCAGGCAGTTGAGGGGGTCTCAATCAAACACAACTACGAAAAATTCTAGCATACTCTTCAATCACACACATAGGCTGAATAATAATAACAATAACGTATAAC
CCAAATATTACAATCTTCTATCTGCTCATGTACATCATTATAACAAGCACTGCATTCCTAGCCCTGAACCTAAACTCAAATACCACCACCCTAATACTATCACGCACC
TGAAACAAACTAACCTGACTAATACCATTAATACTACTTACCCTCCTATCCATAGGAGGTCTACCTCCACTGACCGGCTTTCTACCTAAATGGATAACAATTCAAGAA
CTTACAAAAAATAGCAACTTTATCATACCCTCTATCATAATCACCATAACTTTACTCAATCTATATTTCTATTTACGTTTAACTTATATTACTTCCATAACACTACTCC
CTACGTCTAATAACACAAAAATAAAATGACAGTTCGAAAATACAAAGCCTACACCCCTTTTCCCCCCACTAATTATTTTTACAACCCTTCTCTTACCAATATCGCCAA
CAGTTCTAACCCTATTCTAGAAATTTAGGTTAAACCAGACCAAAAGCCTTCAAAGCTTTCAGTAAGTTAACATGCTTAATTCTTCAAACACTAAGGACTGCAGTAC
TATACCCTGCATCAACTGAACGCAAATCAATCACTTTAATTAAGCTAAGCCCTTACTAGATCAATGGGATTCAAACCCACAAAAACTTAGTTAACAGCTAAATACCC
TAATCAACTGGCTTTAATCTACTTCTCCCGCCGCAGGGAAAAAAAGGCGGGAGAAGCCCCGGCAGAAATAAAACTGCTCCCTTGAATTTGCAATTCAACATGATAAT
CACCTCGGGGCTGGTAAAAAGAGGGTTCAACCTCTGTACTTAGGTTTACAGCCCAATGCCTACTCAGCCATTTTACCTATGCTCATCAACCGCTGGTTATTCTCTACA
AATCACAAGGACATTGGAACTTTGTATTTATTATTTGGTGCATGAGCTGGAACCACAGGTATAGCTATAAGTCTCCTTATTCGAGCTGAACTAGGCCAGCCCGGCAA
CCTACTAGGCAACGACCATATTTATAATGTTATTGTTACAGCCCATGCATTTGTTATAATTTTCTTCATGGTTATACCAATTATAATCGGGGGCTTCGGAAACTGACTA
GTTCCCTTAATAATTGGCGCTCCTGACATAGCATTCCCCCGCCTAAATAATATAAGCTTCTGAACTCTCCCGCCATCTTTCCTACTTCTTCTCGCATCAGCAATAGTAG
AGGCTGGCGCAGGAACAGGCTGAACAGTCTATCCTCCTCTAGCAGGAAATTTTTCCCACCCAGGAGCTTCTGTAGACTTAACTATTTTTTCACTCCACCTAGCAGGTA
TTTCCTCTATCTTAGGAGCTATTAATTTTATTACTACTATTATTAACATAAAACCCCCTGCCATGTCTCAGTATCAGACACCCCTATTTGTTTGATCCGTCCTAATTACA
GCAGTCTTACTGCTTCTATCCTTACCTGTATTAGCTGCGGGCATTACAATGCTATTAATCAGACCGTAATCTCAACACTACCTTCTTTGACCCCGCCGGAGGAGGAGAC
CCAATCTTATATCAACACTTATTCTGATTTTTCGGTCACCCTGAGGTTTATATTCTTATCTTACCTGGGTTTGGAATGATCTCCCACATTGTAACATATTATTCCGGAA
AAAAGGAACCATTCGGATATATAGGCATAGTCTGAGCTATAGTATCAATTGGGTTTTTAGGCTTTATCGTATGAGCTCACCATATATTTACTGTTGGCATAGACGTAG
ACACACGAGCCTATTTCACCTCCGCCACCATAATTATTGCAATCCCAACTGGCGTTAAAGTCTTTAGCTGACTAGCTACACTACATGGAGGAAATATCAAATGATCC
CCTGCAATACTCTGAGCCCTAGGCTTTATTTTCCTCTTCACCGTAGGAGGCTTAACTGGTATTGTACTAGCAAACTCATCACTAGATATCGTACTACATGATACATAC
TATGTGGTAGCCCACTTCCACTACGTTTTGTCAATAGGAGCCGTCTTTGCTATCATAGGGGGTTTTATCCACTGATTTCCACTATTCTCAGGATACGCCCTAGACGAA
ATTTGTGCCAAAGCCCACTTTATTATTATATTTGTAGGCGTAAATTTAACTTTCTTTCCGCAGCATTTTCTTGGTTTGTCCGGAATACCTCGACGTTATTCCGATTATCC
TGATGCTTACACCACATGAAATGTTGTATCATCTATAGGCTCCTTCATCTCCCTAGTAGCTATATTACTAATAATCTATATAATCTGAGAAGCTTTCGCCTCAAAACGT
AAAGTCCTATTTATTGAACAACCTACCTACAACCTAGAATGATTGCACGGCTCTCCACCACCATATCATACATTCGACGAACCAACATTCATTAAGGTTAAATAAAA
AAGGAAGGAATCGAACCTCCTGAGATTGGTTTCAAGCCAATCCTATAACCTCTATAACTTTTTCAATAAGATATTAGAAAAAATTATTTCATGGTTTTGTCAAAGCTA
AATTATAGGATACACCCTATATATCTTATATGCCTCACCCAGTTCAACTAGGCCTACAAGATGCCACATCCCCTATCATAGAAGAATTAATTGCCTTCCACGACCATG
CCTTTATAATTGTAGCTATGATCAGCTTTTTAGTCTTATACGTCTTATCTTCAGTACTCACAACAAAATTAACTAACACCAATATTACAGATGCTCAAGAAATAGAAA
CTATTTGGACCATCTTACCAGCAATTATCTTAATCTTAATTGCTCTGCCATCCCTACGTATTCTCTACTTAACGGATGAGGTTAATAATCCTTCATTTACTATTAAATC
AATCGGACATCAGTGATACTGAACTTATGAGTATACAGATTATGGGGGCTTAATCTTTAATTCCTATATACTCCCTCCACTATTCTTAAACCCAGGGGACCTTCGACT
TCTGGAAGTTGATAATCGAGTGGTACTCCCAATTGAAGCTCCTGTTCGCATAATAATTACATCTCAAGACGTCTTACACTCCTGAACTATTCCTACGCTAGGATTAAA
AACAGATGCAATTCCCGGACGCTTAAATCAAACCACATTCACTGCCACACGACCAGGTGTCTACTACGGACAGTGCTCAGAAATTTGCGGCGCTAATCACAGTTTTA
TACCAATTGTTGCAGAACTAATTCCACTAAAAATTTTTGAAATAGGGCCTGTATTCACCCTGTAGATATACTAAAGCGTTTAATTAAAGTATTTCTTAAGTACTAAAC
TCACTGCATAGCTGCCACAGCATTAACCTTTTAAGTTAAAGACTGAGAAAATATTCTTCTCTGCAGTGAAATGCCTCAACTAAATACATCTACATGGCTAATTACTAT
TATAACCATGCTACCCGCACTATATCTTATTATACAACTAAAACTATTAAATATAGTCTACTATTTTCCCCCGTCACAAAAAGTCTCTAGTACGCAAATATTTAACAA
CCCCTGACAACTAAAATGAACGAAAATTTATTTACCTCATTTACACACCCAACACTCCTAGGGCTACCAGCCGTAGTACCTATTATTCTATTTCCTACATTACTACTTC
CAACCTCAAAATATCTTATTAATAATCGACTAATTACTATTCAACAAAATCTAATTCAGCTAATCGTAAAACAAATAATAATTCATAATATTAAAGGACAAACC
TGGTCTCTAATACTAATATCCCTAATCATTTTTATTGCCACAACTAACCTTCTCGGACTCTTACCCCACTCATTTACACCTACCGCCCAACTATCAATAAATCTGGCTA
TAGCAATCCCTCTATGAGCAGGTACAGTAATCACAGGCCTTCGCTTCAAAACTAAAAGCTCCTTAGCGCACTTTTTACCACAAGGTACACCCACACTACTTATCCCCA
TATTAGTAGTTATTGAGACCATCAGTTTATTTATTCAACCAGTAGCCCTTGCTGTACGCCTAACCGCCAATATTACAGCAGGACATCTACTTATGCACCTAATTGGAA
GTGCTACGCTAGTACTATCAACCATCAGCCTCTCTATAACCTCACTAACTTTAGTGCTTCTAGTATTATTAACAATTCTAGAAATAGCAGTCGCCCTAATTCAAGCCT
ACGTTTTCACACTATTGGTAAGCCTTTATCTACATGACAACACTTAATGACACACCAAACCCACCCCTATCACATAGTTAAACCCAGTCCATGACCACTGACAGGAG
CCCTGTCAGCTCTCCTAATAACATCTGGCCTAATTATATGATTCCACTTTATTCTACAACCCTATTAACCCTAGGTCTACTAACCAATATGTTAACCATATACCAATG
ATGGCGTGATATCATCCGAGAAAGCACCTATCAAGGTCACCATACAACACCAGTTCAAAAAGGCCTCCGCTACGGAATAGTATTATTTATCATCTCAGAGGTTTTCT
TCTTTGCTGGCTTCTTCTGAGCTTTTTATCACTCAAGCCTGGCCCCGACTCCACACCTAGGAGGACACTGACCACCTACAGGTATTATTCCCCTAAACCCCCTAGAAG
TACCCCTTCTAAATACCTCCGTACTACTTGCATCAGGGGTTACAATCACTTGAGCTCACCACAGCTTAATAGAAAATAACCGAAAACAAATAATTCAAGCCCTATCA
ACTACAATTTTATTAGGCATTTATTTCACCCTACTACAAATATCAGAGTACTATGAAGCACCCTTTACCATCTCTGATGGAATCTATGGTTCAACATTCTTTGTTGCTA
CCGGCTTTCATGGACTTCATGTTATCATTGGATCTACATTTCTTACTACCTGTCTTATTCGTCAGCTATCGTACCATTTTACATCAAGTCACCATTTTGGCTTTGAAGCT
GCCGCCTGATATTGACATTTCGTAGATGTCGTCTGACTCTTTCTCTATATCTCCATTTACTGATGAGGTTCCTATCCTTTTAGTATAACTAGTACAGCCGACTTCCAAT
CAGTTAGTTTCGATAATATCGAAAAAGGATAATTAATCTAATACTAACTCTAATAATTAATACCTCCTTAACCATCTTGCTAATAATTATTATATTTTGGCTACCCCA
ACTCAACTCTTATGTAGAAAAAGCCAACCCCTACGAATGCGGGTTTGACCCACTAAACTCTGCTCGCATTCCCTTTTCCATAAAATTCTTCCTAGTCGCTATTACCTTC
CTGCTGTTTGATTTAGAAATCGCCTTGCTATTGCCCTTGCCATGAGCCCTCCAAACAACAAACCTTCCCGTAATGATCAAGTCATCAATTACATTAATTATTATCTTAA
CCCTCAGCCTGGCCTATGAATGAACTCAAAAAGGTTTAGATTGAACTGAATTGGTAAGTAGTTTAAGTAAAACAAATGATTTCGACTCATTAGATTATGGTAATCAT
ACTAACCAAATGTCCATTATTTATATAAACATTATATTAGCATTCACTACCTCACTCCTGGGCATATTAATCTATCGCTCACATCTAATGTCATCCCTACTATGCCTAG
AAGGAATAATACTTTCACTGTTTATCATAAGCACTCTTACTGCATTAAACACTCACTTTCCCCTCGCCAATATAGTACCTATTGCCTTACTAGTATTTGCCGCTTGCGA
GGCAGCAGTGGGCCTTGCCCTATTAATTTCAATCTCAAACACATACGGCTTAGACCATATCCAGAACCTAAACTTACTTCAATGTTAAAAATAATTTTCCCTACAATG
ATACTATTACCAACAACATGATTTTCCAAAAACAACCTAGTCTGAATTAACCTAACTACACATAGCTTAATAATCAGCCTCATTCCCCTTATATTCTTCAATCAAATC

AATAATAACCTTATTAGCCACTCGACCTATTTATCTTCAGATCCATTAACAACACCTCTTCTAATACTGACAGCCTGACTCCTGCCCCTCATAATTATAGCAAGCCAA
TATCACCTACACAATGAAAGTCCCCTGCGAAAAAAACTTTACCTCTCCATAATAATCTTCCTACAAATCTCCCTAATCATAACATTTATAGCCACAGAGTTAATCTTA
TTCTACATCCTATTTGAAACAACCCTTATCCCCACCCTAATTATTATCACCCGATGAGGTAACCAAGCAGAACGCCTCAACGCAAGCACATATTTCTTATTTTATACA
TTAGCTGGCTCCCTACCTCTGCTAATTATATTAATTTATACACATAACAAATTAGGCTCATTAAATATTCCACTACTAACACTCATAGCCCAAAAACTAACAACCTCT
TGATCCCATAATTTAGCTTGACTAGCATGCATAATGGCTTTTATAGTAAAAATACCCTTATATGGCCTACACCTATGACTCCCTAAAGCCCATGTTGAGGCTCCCATT
GCCGGGTCAATAGTTCTTGCCGCCGTGCTCTTAAAATTAGGCGGCTATGGCATAATACGACTTACACCAATTCTCAACCCACTAACAGAACATATAGCCTACCCTTTT
CTTATACTATCCTTATGGGGCATAATTATAACTAGCTCAACCTGTCTCCGACAAACAGATTTAAAATCGCTCATTGCATACTCCTCTGTAAGCCACATAGCCCTTGTA
ATCATAGCCTCCCTCATTCAAACCCCCTGAAGCTTTACTGGCGCTATCATCCTTATAATTGCCCACGGACTTACCTCATCTATATTATTCTGCTTAGCAAACTCAAATT
ACGAACGAACCCACAGCCGCATTATATTACTTTCTCGAGGACTTCAAAGTCTACTTCCCACTAATAGCCTTCTGATGATTTGTAGCAAACCTCACCAATCTAGCTCTAC
CTCCCTCCATTAACTTAATCGGGGAGTTATTAGTAGTGACATCCTCATTTTCTTGGTCACATATTACCATTATCTTCACAGGACTAAACATATTAATTACTGCTCTATA
TTCCCTACACATATTCATTACAACACAACGAGGAACACTCACCTCCCATATTATTAACATAAAACCCTCTTTTACACGAGAAAACATGCTAATATTTATACATATATC
TCCTATTATTCTTCTAACCCTTAATCCTAGCATTATTATAGGCTTCACCCCTTGTAAATATAGTTTAATTAAAACATTAGATTGTGAATCTAAATATAGAAACCTACCA
CTTCTTATTTACCGAGAAAGATTGCAAGGACTGCTAATCCATGCCCCCGTATTTAATAAAACGGCTATCTCAACTTTTAAAGGATAACAGCTGTCCATTGGTCTTAGG
AACCAAAAATATTGGTGCAACTCCAAATAAAAGTAATAATAATGCACACCTCCATTTTTATACTAGCCCTAACCCCCTTAATCTTTCCGATTATTATTACCCTTATTA
GCCCCAATAAAAATAATATATACCCCAACTATGTAAAAACAACTATGATATTTACCTTTACTATTAGTCTCATCCCCACAACCATATATACTTTCCTAGGTCAAGATA
CAATCATATCAACCTGACATTGAATAACTATTCAATCACTAGAAATTACACTAAGTTTTAAATTGGACTATTACTCCGTAATATTTACCCCAATTGCACTATTTATTAC
TTGGTGCATTATAGAATTCTCACTATGATACATAGACTCAGACCCAAACATTAACCAATTCTTCAAATATCTTCTCATCTTCCTTATTACCATGCTAATTTTAGTTACC
GCTAACAACCTCTTCCAGCTCTTTATTGGGTGAGAAGGTGTAGGAATTATATCATTTCTACTAATCGGCTGATGATACGCTCGAACAGACGCTAACACAGCAGCCAT
TCAAGCAATTCTGTATAACCGCATTGGTGATATTGGTTTCATTCTAGCCATAATATGGTTTCTCCTCCATTATAACTCATGAGACTTACAACAAATATTTATCCTAGAT
CCCAACCCCGATCTACTTCCATTAGTGGGTCTACTATTAGCAGCAACAGGAAAATCAGCCCAACTCGGCCTCCATCCCTGATTACCCTCGGCTATAGAAGGCCCAAC
TCCAGTATCAGCCCTACTTCACTCCAGTACTATAGTAGTAGCTGGGGTTTTCTTACTTATTCGCTTCCACCCACCAATAGAAAATAATACAACAATTCAAAGTCTTAC
ACTATGCCTAGGAGCTATTACTACCATATTCATAGCAATCTGCGCCCTAACACAAAATGACATTAAAAAAATTGTAGCCTTCTCTACCTCAAGTCAACTGGGACTTAT
AATAGTTACTATTGGTATTAATCAACCGCACCTAGCATTTCTACATATCTGTACTCATGCCTTTTTCAAGGCTATACTATTTATCTGCTCTGGGTCTATAATTCATAAC
CTAAATAATGAACAAGACATCCGAAAAATAGGAGGACTATTTAAAACAATGCCCCTCACCTCAACTTCCCTGATAATCGGTAGCCTAGCACTCACAGGCATACCTTT
TCTTACAGGTTATTACTCCAAAGACCTCATCATCGAAACCGCAAACACATCATACACCAACGCCTGGGCCCTGTGTATTACTCTTATCGCTACCTCTATAACAAGCGC
CTACAGCACCCGAACTATTATTCTCACACTAACAGGATCACCTCGTTTTTCAACTTCCGTATATATTAATGAGAACAACCCAACCCTACTAAACCCAATAAAACGCCT
AGCAGCAGGTAGTCTACTCGCAGGATTTTCATCGTCAACAACATCTCTCCGACTACAGTTCCTCAATTAACAATACCTTATCACCTGAAACTCCTAGCCTTATGCGT
AACCACCCTAGGCTTCTTAACAGCCCTAGATCTGACTCTTATAACTAACAGTCTCAAAATAAATACCCCATCGCACATATTCAAATTCTCCAATATACTAGGATATTT
TCCCATTACAATTCACCGAACAGTTCCCTACCAAAACCTAACCATAAGCCAAAACCTAGCCTTCCTATTACTAGACTTACTCTGACTAGAGAAATCAATACCTAAAA
CAATTTCACACACCCATATTATTACAGCCATCACCTCAACCACCCAAAAAGGCATAATCAAGCTATATTTCCTCTCTTTTCTTATTTCCCTCACACTAATCCCACTTTT
AATTATATAATCTATTACCCCGAGTAATTTCAATAACAATATAAACACCAACAAATAATGTTCAACCAACAACTACGACCAACCAACGCCCATAATCATACAAAGCA
CCCGCACCAATAGAATCCTCACGAATCAACCCCGACCCCTCCCCCTCAAAAATCACCCAACTCCCTATATTATCCAAATTAATTATCACTACCAACTCATTATAATCT
ATAACCCACAAAACTAAATATACCTCCATTGCCAATCCAACCAAAAAACTCCCCAAAACCTCAAATCCTGAAACCCATGCTTCAGGATATTCTTCAATAGCCATCGC
AGTAGTATAACCAAAGACAACCATTATACCCCCCAGATAAATCAAAAACATTATTAAACCTATATAAGTACCCCCATAACTTAAAATAATAGCACAACCAATCACA
CCACTAACAACCAATGCTAAACCCCCATAAATAGGAGAAGGCTTAGAAGAAAAGCCCACAAAACCCATAACTAATAATACGCTTAATAAAAATAAAATATACGAC
ATTGTTTCCACATGGACTCTAACCATGATTAATGATATGAAAAACCATCGTTGTATTTCAACTATAAAAACACTAATGATCCCCGTACGCAAATCTAATCCAATTATA
AAATTAATTAATCACTCCCTCATTGATCTACCAACCCCATCAAACATCTCAGCATGATGAAACTTTGGTTCCCTTTTAGCAACCTGCTTAATTTTACAAATCATCACA
GGTCTATTCCTAGCAATACATTACTCACCTAGCACCTCCTCAGCCTTCTCCTCAATCGCCCATATCACTCGGAGCGTAAACTACGGCTGAATTATCCGCTACCTTCAT
GCCAATGGTGCCTCCATATTCTTTATCTGCCTATTCCTACACGTAGGTCGAGGGTTATATTATGGCTCATTCCTTCTTCTTGAAACTTGAAACATTGGCATTGCACTAT
TACTTATAGTTATAGCAACGGCCTTTATAGGCTATGTACTCCCATGAGGACAAATATCATTTTGAGGTGCTACAGTAATTACAAATTTATTATCCGCAATCCCATACA
TCGGAACAAATCTCGTTCAATGAGTATGAGGTGGATATTCTATCGATAATCCAACCCTTACCCGATTCTTCACTCTCCACTTTACCCTGCCTTTCATTATCACAGCCTT
TACAGTTCTACATCTACTTTTCCTACACGAAACAGGATCTAATAATCCATGCGGAATTCCCTCAAACTCCGACAAAATCCCCTTCCACCCCTACTACACAATCAAAGA
TATGCTAGGCCTAGTCCTCCTTATTCTTCCCCTAATAACTCTAGTATTATTTTCACCTGACCTTTTAGGTGACCCAGACAACTATACACCAGCTAATCCACTAAACACC
CCACCACACATCAAACCAGAATGGTACTTCTTATTCGCATATGCAATCCTACGATCTGTACCTAACAAATTGGGAGGCGTACTAGCACTTCTTATATCCATTCTTATC
CTAATAATTATCCCCATACTTCACAAATCCAAACAACAAAGCATAATATTCCGCCCATTCAGCCAATTTACACTATGATTATTAATCACAGTTCTATTAACCCTAACC
TGAATTGGAAGCCAACCAGTAAACCAACCCTTTATTATAATTGGACAGGTAGCATCTATAATATATTTCACTACAATTCTGATCTTAATACCACTAGCCTCTATAATC
GAAAACAATCTCCTCAAATGAACCTGCCCTTGTAGTATAGACTAATACACCGGTCTTGTAAACCGGAGACGGATACCTTTCCCCAGGGCAACTCAGAAAGAAAGCA
TTTAACTCCTCCACCAATACCCAAAACTGGCATTCTATTTAAACTACTTTCTGCATTCTAAGGAGGTATAACCTTTAGAGAATAATTTAGTACAATCTAATTTTATATG
TCCTTATGTAATTCGTGCATTACTGCTAGTCACCATGGATATTATATAGTACTATAAATGTTTTACCGTCCATAGGACATAAAATTACATATTTACTGGCAGTTATATC
TAGGACATGCTTATAAGCAAGTACTCTAATAGAATATCTGACTATAACACATACCATTCACATATCCAAATCCCGTGGTCACACCCATTGGAATACCAACCAATAGC
GATTAATCCATTATCGTACATAGTACATAATATTCTTTACCGGACATAGCACATCGCAGTCGAGCATCCCTAAACCAATCCTCGTCAACACGGATATTCCCGTCAGTT
AGGTGTCCCTTGATCACCATCCTCCGTGAAATCAATATCCCGCACAAGAGTGCTACTCTCCTCGCCCCGGGCCCATAACCTGTGGGGGTAGCTACACACGACGTCAA
TGGACATCTGGTTCTTACCTCAGGGCCATACCGCCAAGATCGCCCACACGTTCCCCTTAAATAAGACATCTCGATGGATCACGGGTCTATCACCCTATTAACCAGTCA
CGGGAGCTCTCCATGCATTTGGTATCTTTTATCTCTGGTCTGCACGCGACCCCATTGCAGTATGCTGGTCTCGCCACAATCAGTCCCGCAGCGCCTGTCTTTGATTCCT
AGCACATACCATTATTAACCGCACCTACGTTCCATATTTTAATCCCGCATGAACCTTATCATGGTGTTATTTAATCCATGCTTGTAAGACATACATATAATTAACCGT
ATACCCCGATCGCACCTTGATCCAATTTACAACTATCTCTACTCAAACCCCCCCCCCCCCCCATATTCGACCGCACCCTAATAATCCGCCTTTGCCAAACCCCAAAAAC
AAAAACTTGCAATCTGGTCGAATTTTACATTTTTTTTT

Bảng S2. Trình tự nucleotide trên ADN ty thể *Chà vá chân đen* Tây Nguyên Việt Nam

AAGACTACCCAAAACCTCAAACCCTGAGACTCATGCCTCAGGGTATTCTTCAATAGCTATCGCAGTAGTATAACCAAAGACAACCATTATACCTCCCAGATAAATCA
AAAATATCATTAAACCCATATAAGTACCCCCATAATTTAAAATAATAGCACAACCAACTATTCCACTAACAACCAATGCTAAGCCCCCATAAATAGGAGAAGGCTTA
GAAGAAAACCCCACAAAACCCATAACTAATAATACACTTAATAAAAATAAAATATATGACATTGTTTCCACATGGACTCTAACCATGATTAATGATATGAAAAACC
ATCGTTGTATTTCAACTATAAAAACACTAATGACCTCCCTACGTAAATCCAATCCAATTATAAAATTAATTAATCACTCTCTTATCGATCTACCCACCCCGTCAAACA
TTTCCGCGTGATGAAACTTTGGTTCTCTTTTAGCAACCTGCTTAATTTTACAAATCATCACCGGTCTATTCCTAGCAATACACTACTCACCCAACACCTCCTCAGCCTT
CTCCTCAATCGCTCATATTACCCGAGATGTAAATTATGGCTGAATTATCCGCTATCTTCATGCCAATGGTGCCTCCATATTCTTTATTTGCCTATTCCTACATGTAGGC
CGAGGACTATACTATGGCTCATTTCTTCTTCTTGAAACTTGAAACATTGGCATTGCACTATTACTCATAGTTATAGCAACGGCCTTCATAGGCTACGTACTCCCATGA
GGACAAATATCATTTTGGGGCGCTACAGTAATTACAAATTTACTATCCGCAATCCCATATATCGGAACGAATCTTGTTCAATGGGTATGAGGGGGATATTCTATCGA
CAATCCAACCCTTACCCGATTCTTCACCCTCCACTTTACCCTACCTTTCATTATTACAGCCTTTACAATTCTACACCTACTTTTTCTACACGAAACAGGGTCTAATAAT
CCCTGTGGAATCCCCTCAAACTCTGACAAAATTCCCTTCCACCCCTACTACACAGTCAAAGATATACTAGGCCTAGTCCTCCTTATTTTTCCCCTAATAATTCTAGTAT
TATTTTCACCTGACCTTTTAGGTGACCCAGACAACTATACACCAGCTAATCCACTAAACACCCCACCACACATCAAACCAGAGTGATATTTCCTATTCGCATATGCAA
TCTTACGATCCGTACCTAACAAATTAGGGGGCGTGCTAGCACTCCTTATATCTATCCTCATCCTAATAATTATCCCCATACTTCACAAATCCAAGCAACAAAGTATAA
TATTCCGCCCACTTAGCCAATTCACACTATGACTACTAACCACAGTTCTACTAACCCTAACCTGAATTGGAAGCCAACCAGTTAGCCAGCCCTTTATTATAATTGGAC
AGATGGCATCTATATTATACTTCATTACAATTCTAATCTTAATACCGCTAGCCTCTATAATGCAAAACAATCTCCTCAAATGAACCTGCCCTTGTAATATAGACTAAT
ATACCGGTCTTGTAAACCGGAGACGGATACCTTCCCCAGGGCAACTCAGAAAGAAAGCATTTAACTCTACCACCAATACCCAAAACTGGCATTCTATTTAAACTACC
TTCTGCATTCTAAGGGGGTATAACCTCTAAAGAATAACTTAGTACAATTTAACTTTATATGTCCTTATGTAATTCGTGCATTACTGCTAGTCACCATGGATATTATATA
GTACTATAAGTGTTTTATCGTCCATAGGACATAAAATTACATATTTACCAGCAATTCTATCTAGGACATGCTTACAAGCAAGTACCTTACTGAAACCTCTAACTGTAA
CACATACCACCCAAGTTACCAAATTCTATGGTTCCGCCCACCGGAATACCAACCAATAGCGATTAATTCATTATCGTACATAGTACATAATATTCTTTATCGTACATA
GCACATTACAGTTGAGCATCCCTAAACCAATCCCCATCAACACGGATATACCTGTCAGTTAGGTGTCCCTTGATCACCATCCTCCGTGAAATCAATATCCCGCACAA
GAGTGCTACTCTCCTCGCCCCGGGCCCATAACCTGTGGGGGTAGCTATATATGACATCAATGGACATCTGGTTCTTACCTCAGGGCCATGCCGTCAAGATCGCCCAC
ACGTTCCCCTTAAATAAGACATCTCGATGGATCACGGGTCTATCACCCTATTAACCAGTCACGGGAGCTCTCCATGCATTTGGTATCTTTTATCTCTGGTCTGCACGC
GACCCCATTGCAGTATGCTGGTCTCGCCACAATCAATCCCGCAGCGCCTGTCTTTGATTCCTAGCACATACCATTATTAACCGCACCTACGTTCCATATTTTAATCCC
GCATGAACCTTATCATGGTGTTATTTAATCCATGCTTGTAAGACATATACATAATCATCCCATATACCCCGACCACACCTTGATTCAATTTACAACTATCTCTACTCAA
ACCCCCCCTCCCCCCACATCCGACCGCACCCTTATAATCCGCCTTTGCCAAACCCCAAAAACAAAAACTTACAATCTGGTCGAATTTTACATTTTTTTTTATCTTTTAG
GTATGCACAACTCAACTGTCGTTTTCTCAACTAATAAAAATTTATTTCACACACCACCCTTTACATCTCCACTCACCTCCTTTTTTTCCCATAACCCTAAAGATACACC
ACAACATCCACTATTCCGTTTATGTAGCTTAACATGTCTAAAGCAAGACACTGAAAATGCCTAGACGGGCTCACAAACCCCATAAACAAACAGGTTTGGTCCCGGCC
TTTCTATTAGCTTTTAGCAAAATTACACATGCAAGCATCCCCGTCCCAGTGAAGATACCCTATAAATCACAACGATTAAAAGGAGTGAGTATCAAGCACGCTTAATG
CAGCTCAAGACACTTTGCTTAGCCACACCCCCACGGGAAACAGCAGTGACTAATATTTAGCAATAAACGAAAGTTTAACTAAGCTATGTTATTAGGGGTTGGTCAAT
TTCGTGCCAGCCACCGCGGTCATACGATTGACCCTAGTTAATAGACCTCGGCGTAAAGGGTGTTTTAGATAAATTTAACTAAATAAAGCTGAACTCTTTCTAAACTGT
AAAATCCTAGCTAATGTAAAATAGACTACGAAAGTGGCTTTAAAGCTTCTGAATACACAACAGCTAAGACTCAAACTGGGATTAGACACCCCACTATGCTTAGCCCT
AAACCTTAATAGTTAAAAATAACAAAACTATTCGCCAGAACACTACAAGCAACAGCTTAAAATTTAAAGGACTTGGCGGTGCTCCATATCCCCCTAGAGGAGCCTGT
TCTATAATCGATAAACCCCGATCCACCTCACCACCTCTCGCTCAGCCTATATACCGCCATCTTCAGCAAACCTCAACAAGAGATATAAAGTAAGCACAAACGCCCAC
GCAAAAACGTTAGGTCAAGGTGTAGCCTATGAGCTGGGAGAAATGGGCTACATTTTCTATTCTAGAAAACCCCACGATAACCCTTATGAAACTTAAGGGTCCAAGG
AGGATTTAGTAGTAAACTAAGAATAGAGTGCTTAGTTGAACCAGGCCATGAAGCACGCACACACCGCCCGTCGCTCTCCTCAAATATAATTAGAACTAATTTAACT
AAATATTTCTGCATTTATACAGAGGAGATAAGTCGTAACATGGTAAGTGTACTGGAAGGTGCACTTGGATAAATCAAGGCATAGCTCAACACAAAGCATCCAGCTT
ACACCCGGAAGATATCGCCACCCGATTGCCTTGAGCCAATACTAGCTCAAACCCAACCAACACTACTATTAAACTGATAAATATATTAAACCATTCACAAATATAAA
AGTATAGGAGATAGAAATTTACTCCGGCGCTATAGATACAGTACCGCAAGGGAAAGATGAAAAAAAGCAATTAAGCACAAAACAGCAAGGACTCACCCCTATACC
TTTTGCATAATGAGCTAACTAGAGACCACTTCACGAAGAGAACTAAAGCCAAGTACCCCGAAACCAGACGAGCTACCCAAAAACAGCTGAAAGAGCGCACCCATCT
ATGTAGCAAAATAGTGGGAAGATTTCTGGGTAGAGGTGATACGCCTACCGAGCCTGGAGATAGCTGGTTATCCAAGATAGAATTTTAGTTCAACCTTAAGCTTACCT
ACAGAACTATAAATCTCCCTGTAAGCTTAACTGTTAGTCTAAAGAGGGACAGCTCTTTAGATACCAGGAAAAAACCTTATGTAGAGAGTAAAAAACTCCAAACTCAT



TTAATCAACCACACCTAGCATTTCTACACATCTGTACCCACGCCTTTTTCAAGGCTATACTATTTATCTGCTCTGGGTCTATAATTCATAACCTAAATAATGAACAAG
ACATCCGAAAAATAGGAGGACTATTTAAAACAATACCCCTTACCTCAACTTCCCTAATAATCGGTAGCCTAGCACTCACAGGCATACCTTTTCTTACAGGCTACTACT
CCAAAGATCTCATCATCGAAACCGCAAACATATCATACACCAACGCCTGAGCCCTGTCCATTACTCTTATTGCTACCTCCATAACGAGCGCCTACAGTACCCGAACT
ATTATTCTCACACTAACAGGACCACCTCGTTTTTCAACTTCAGTATACATTAATGAAAATAACCTAGCCTTACTAAACCCAATAAAACGTCTAGCAACAGGTAGCCTA
CTCGCAGGATTTTTCATTGTCAATAACATCTCTCCTACTACATTTCCTCAACTAACAATACCCTATTACCTAAAACTCCTAGCCTTATGCGTAACCACTCTAGGCTTCT
TAATAGCACTAGATCTGACCCTTATAACTAACAATCTCAAAATAAACATCCCATCACACATATTCAAATTCTCCAACATACTAGGATTCTTTCCTATAACAATCCACC
GAACAGTTCCCTACCAAAACCTAACCATAAGCCAAAACCTAGCCTTCCTATTACTAGACTTATTCTGACTAGAGAAATCAATACCTAAAACAATTTCACATACCCAT
ATTAATACAGCCATCACCTCAACTACCCAAAAAGGCATAATCAAGCTCTATTTCCTCTCTTTCCTTATTCCCCTCACACTAATCCCACTTTTAATAATATAATCTATTA
CCCCGAGTAATTTCAATAACAATATAAACACCAACAAATAATGTTCAACCAACAACTACGACTAGCCAACGCCCATAATCGTATAAAGCACCAGCACCAATAGAAT
CCTCACGAATCAACCCTGACCCCTCCCCCTCAAAAATCACCCAACTTCCCATATTATTCAAATTAACTATCACTACCAACTCATTATAATCTATAACCCATAAAACCA
AATACACTTCCATTGCCAACCCAACTAA

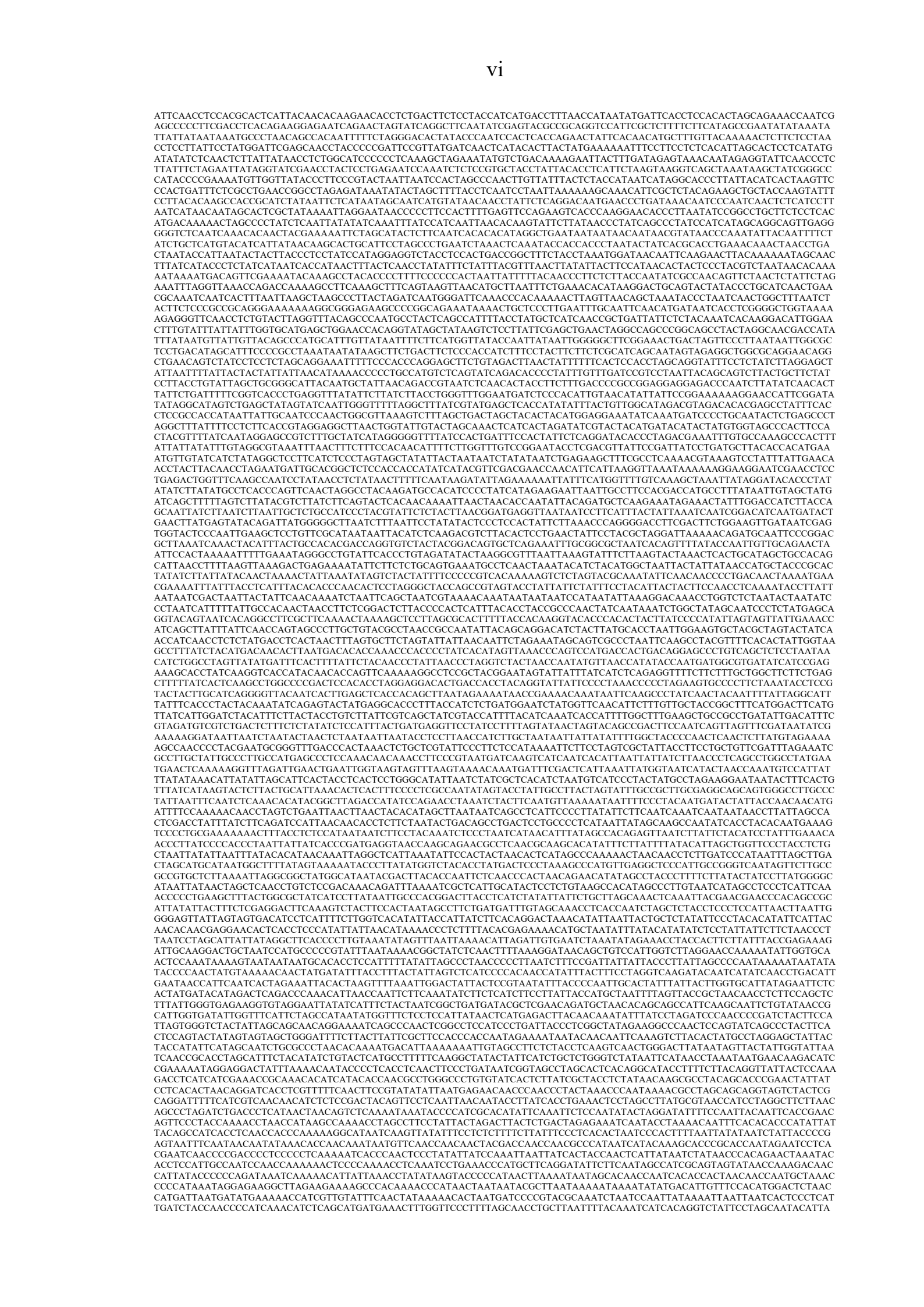
Bảng S3. Trình tự nucleotide trên ADN ty thể *Khỉ mặt đỏ* Tây Nguyên Việt Nam

ACAGCCGTACACCTACTATTCCTACACGAAACAGGATCAAACAACCCCTGCGGAATCTCCTCCAACTCGGACAAAATCGCCTTTCACCCCTACTATACAATCAAAGA
CATCCTAGGTCTAGTCCTCCTTCTCTTTATCCTAGCAACACTAACACTACTCTCACCCAACCTCCTAAACGACCCAGACAACTACATTCCAGCCGACCCACTAAACAC
TCCCCCACATATCAAGCCAGAGTGATACTTTCTATTCGCATACACAATCCTACGATCTGTCCCCAACAAACTGGGAGGCGTACTAGCACTCTTCCTATCAATCCTCAT
CCTAGCAGCCATCCCCATACTTCACAAATCCAAACAACAAAGCATAATATTCCGCCCACTCAGCCAATTTCTATTCTGACTCCTGATCACAATCCTGCTGACCCTCAC
CTGAATTGGAAGCGAACCAGTAACCCAACCTCTTACCACCATCGGCCAAGCAGCATCCATAATATACTTCATTACAATTCTAATTCTAATACCACTGGCCTCCCTAAT
CGAAAACAGCCTACTCAAATGAACCTGCCCTCGTAGTATAAGCCAATACACTGGCCTTGTAAACCAGAAATGAACACTCTTCCTAGGGCAATCAGAAAGAAAGTAT
CTAACTTCACCATCAACACCCAAAGCTGGCGTTCTAATTTAAACTACTTTCTGCATTCTTATGTTGTACAAGTCACATAATGTACTTCTAGTACTAACCCACTCACAAT
ACTACTATGTAATTCGTGCATTACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATCGTACATAATACATATTATTGTACACCAACTTAACAATTCC
AGAAAGCATGCTTACAAGCAAGTACACTAATATAAACTTCAACAGTAGCGCATGGAACAGTTCCCTCGTACCCAAACCTGTCCCTCCCCACGGATATCAACCAAACC
AGTTCATGCCAGTCGTCCATAAGTACATTATATCGTTCATCGGACATAGTACATACCTATTAAACAATCCCTTTCACCACGGATGCCCCCCCTCACTTAGGAATCCCT
TACTCACCATCCTCCGTGAAATCAATATCCCGCACAAGAGTGCTACTCTCCTCGCTCCGGGCCCATAACTCGTGGGGGTAGCTATACTTGAGCTGTATCCGGCATCTG
GTTCTTACCTCAGGGCCATGACAATCAAGATCGCCCACACGTTCCCCTTAAATAAGACATCTCGATGGATCACAGGTCTATCACCCTATTAACCAGTCACGGGAGAT
TTCCATGCATTTGGTATCTTTTATCTCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATGCGCCTGTCTTTGATTCCTAGTACA
TGCAGTTATTAATCGCACCTACGTTCAATATTCTAGCTCCACGCAAACCTTAGCAAGGTGTTATTTAATTCATGCTTGTAGGACATACCAATAATCACTCCAGCCAAC
GCCCCCATTACGTTACAAACCATATCTCATCAAACCCCCCTACCCCCATCTCCGATCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAAGTCTTAACACA
TCCAATCAGAGCCTATGTTTTCATCTTTTAGGTGCACACAACTCTAACTGCCATTCTCTCAACTAACGAACATTTACTTCACTAAACACCCTCTACACCAACCCACAA
CAAACCCCTTTCATACAACCCGAAAGAAATCGCCTCATAACCATGTTAGTCCCTCTGTTCGTGTAGCTTAAACCCACCCAAAGCAAGACACTGAAAATGCCTAGATG
GGTTCGCACACCCCATGAACAAATAGGCTTGGTCCTGGCCTTTCTATTAGCTCTTAGCAGGATTACACATGCAAGCATCCCCGCTCCGGTGAAGACGCCCTACAAAT
CATTACGACCAAGAGGAGCAAGCATCAAGCACGCACACGCAGCTCAAAACGCTTTGCCTAGCCACACCCCCACGGGAGACAGCAGTGACAAATATTTAGCAATAAA
CGAAAGTTCAACTAAGCCATGCTATATTTAGGGTTGGTCAATCTCGTGCCAGCCACCGCGGTTACACGATTAACCCTAGCTAATAGAGATCGGCGTAAAGGGTGTTT
AAGATCTGACATAATAAAGCTAAACCCCATCTAAACTGTAAAACTCTAGCCAACGTAAAATAAACTACGAAAGTGGCTTTAAAGCTTCTGAACACACAATAGCCAA
GACCCAAACTGGGATTAGACACCCTACTATGCTTGGCCCTAAACCTCAGTAGTTAGATAACAAAACTACTCGCCAGAATACTACAAGCAACAGCTTGAAACTCAAA
GGACTTGACGGTGCTTTACATCCCCCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATCCACCCCACCCTCTCTTGCTCAGCCTATATACCGCCATCTTCAGCAA
ACCCTAATGAAGGTCACAAAGTGAGCGCAAACGCCGCCACCGCCGCGAATACGTTAGGTCAAGGTGTAGCCTATGAGACGGTAAAAAATGGGCTACATTTTCTATC
TCAGAAAATTCCACGAAAACCCTTATGAAATTTAAAGGTCTAAGGAGGATTTAGCAGTAAATTAAGAATAGAGTGCTTAATTGAACCAGGCCATAAAGCGCGCACA
CACCGCCCGTCACTCCTCTCAAATATACTTAAGGAATATCTTAACTAAACACCCTAATATTTATATAGAGAGGATAAGTCGTAACATGGTAAGCGTACTGGAAGGTG
CGCTTGGATAAATCAAGGTATAGCTTAATACAAAGCACCTGGCTTACACCCAGAAAATCTCAACACCGCTTGATTGCCTTGAGCTAATACTAGCCCTAAGCACAACC
AACACTAATACCAAACCAACACGCACACATTAAACCATTCACCTACACAAAGTATAGGCGATAGAAATTTTAACCTGGCGCTATAGATATAGTACCGTAAGGGAAA
GATAGAAAACTACCCAAGCACAAAACAGCAAGGACTAACTCCTGTACCTTTCGCATAATGAATTAGCTAGAAACAATTTCACAAAGAGGACTACAAGCCAAATTCC
CCGAAACCAGACGAGCTACCCAAAAACAGCTAAAAGAGCACCACCCATCTATGTGGCAAAATAGTGGGAAGATTTCTGGGTAGAGGTGACAAGCCTACCGAGCCTG
GTGATAGCTGGTTATCCAAGACAGAATCTTAGTTCAACTTTAAGCTTACCTACAGAACTACCTAATCCCCCTGTAAACTTAATTGCTAGTCTAAAGAGGGACAGCTCT
TTAGACACTAGGAAAAAACCTTGTAGAGAGAGTAAAACCCTCAATTATCATAGTTGGCTTAAAAGCAGCCATCAATTAAGAAAGCGTTCAAGCTCAACACCAATCA
CTCCAAAAATACCAAACACACAACTGAACTCCTCACACCACATTGGATTAATCTATCACCATATAGAAGCAATAATGCTAGTATAAGTAACATGAACATTTTCTCCA
CTGCATAAGCCTAAATCGGACCGAAAGCCTCACCGATACCTAACAGCCAGGCATAGCAAACACAAACAAGCCCATGCTACCTCCACTGTTAACCCAACACAGGCAT
GCCTCAAGGAAAGGTTAAAAAAAGTAAAAGGAACTCGGCAAACTCAAACCCCGCCTGTTTACCAAAAACATCACCTCTAGCATTACTAGTATTAGAGGCACTGCCT
GCCCGGTGACACACGTTTAACGGCCGCGGTACCCTGACCGTGCAAAGGTAGCATAATCACTTGTTCTTTAAATGGGGACTCGCATGAAAGGCACCACGAGGGTTTAA
CTGTCTCTTACTTTCAACCAGTGAAATTGACCTGCCCGCGAAGAGGCGGGCATAAAATAATAAGACGAGAAGACCCTATGGAGCTTTAATCTATTAATGCAATCAAA
AATCATACAAACCCACGGATCTTTAAACCACCGACACCTGCATTAAAGATTTTGGTTGGGGCGACCTCGGAGCACAACAAAACCTCCGAACAACACATGCCAAGAC
CACACAAGTCAAAGCAAATTAACACTCATAATTGATCCAATAATTTGATCAACGGAACAAGTTACCCTAGGGATAACAGCGCAATTCTATTCTAGAGTCCATATCGA
CAATAGAGCTTACGACCTCGATGTTGGATCAGGACATCCTAATGGTGCAGCAGCTATCAAGGGTTCGTTTGTTCAACGATTAAAGTCCTACGTGATCTGAGTTCAGA
CCGGAGCAATCCAGGTCGGTTTCTATCTATTGCACATTTCTCCCTGTACGAAAGGACAAGAGAAATAAGGCCCACTTCATATAAGCGCCTTCATTATATAAATGACC
TAGTCTTAATTTAGCAAAATACCACATATTCCACCCGAAAGCAGGGTTTGTTAAGATGGCAGAGCCCGGTAATTGCATAAAACTTAAAACTTTGTAACCAGAGGTTC
AATTCCTCTTCTTAACATCATGATTACAGCAAACCTCTTACTTATAATCATATCCACATTAGCTGCCATGGCATTTCTCACACTCATTGAACGAAAATTATTAGGCTAC
ATGCAGCTACGAAAAGGACCCAACGTTGTAGGCCCATATGGACTGTTACAACCCTTCGCTGACGCTGAAAACTTTTCACCAAAGAACCCCTAAAACCCGCAACATC
TACCACCATCCTCTATATTATCGCACCTGCTCTAGCCTTCTTCATCACCCTTCTCCTATGAACCCCCCTCCCCATACCCAACGCTCTAATCAACTTCAACCTAGGACTC
CTATTCATCCTAGCCATACTCAGCCTAATCGTCTACTCCATCCTATGATCAGGATGAGCATCAAACTCAAACTACGCCCTAATCGGAGCCCTGCGAGCTGTCGCCCAA
ACAATTTCATACGAAGTCACCCTTACTATTATCCTACTATCAGTCCTACTGATAAGCGGCTCATTTAACCTCAACACACTTATTACAACACAAGAACACCTATGACTC
CTCCTACCATCATGACCCCTAGCCATAATATGATTTACCTCTACACTAGCAGAAACAAACCGAACCCCCTTTGATCTCATAGAGGGCGAATCAGAATTAGTATCAGG
CTTTAACATTGAATACGCCGCAGGCCCATTCGCCCTATTCTTCATAGCTGAATATATAAACATCATCATGATAAACGCCCTAACAACTACAATCTTCCTAGGCACATT
CTACCCCACCCACTCACCAGAACTATTCACAATATGCTTCACCACCAAAACACTCTTCCTAACCTCCCTATTCCTATGAATCCGAGCAACATATCCCCGATTCCGTTA
CGACCAACTTATACACTTACTATGAAAAAACTTCCTCCCCCTCACACTAGCACTACTTATATGATCTACCTCAACACCTATCGCAATCTCCAGCATCCCCCCTCAAAC
CTAGAAATATGTCTGACAAAAGAGTTACTTTGATAGAGTAAATAATAGAGGTCCCAACCCTCTTGTTTCTAGGACTGTAGGCATCGAACCTACCCCTGAGAATCCAA
ACTTCTCCGTGCTACCCCCACACTCTATCCTAAGTAAGGTCAGCTAAGCAAGCTATCGGGCCCATACCCCGAAAATGTTGGTCATATCCTTCCCGTACTAATTAACCC
ACCAGCTCAACTCACTATCTATACCACAATAATCACAGGCACATTTATCACAATACTAAGCTCACACTGATTTCTCACCTGAACAGGTCTAGAAATAAACATACTAG
CCTTCATTCCAATCCTAATAAAAAAAACAAGCCTACGCTCTACAGAAGCTGCTATCAAATACTTCTTAATACAATCTACCGCATCCATAATCCTCATAATAGCAATCA
TCTCAAACAACCTACTATCAGGATGCTGAACAACAACGAGTTACATCAATCAACCTCCATCCCTAATAATAACAACCGCCCTCGCCATAAAACCTAGGAATAGCCCCC
TTCCACTTCTGAGTCCCAGAAGTCACCCAAGGAACACCCCTAATCTCCGGCCTACTCCTCATTACATGACAAAAACTAGCCCCCATCTCAATCATATACCAAATTCAC
CCATCAATTAACACCCACATCCTCCTAACCCTCTCCACCCTATCCATTGCAGTAGGCAGTTGAGGGGGTCTAAACCAAACACAATTACGCAAAATCCTAGGATACTC
TTCAATCACACACATAGGCTGAATAATAATAACACTAGTATACAACCCAACAACTACAATTCTTTATTTAACTATCTACATTATCCTGACAACTACCATATTCCTGAC
CCTTAACTTAAGCTCAAGCACCACAACCCTCATACTATCTCGCACCTGAAACAAATCAACCCACCTAGCACCACTAATAGCACCCATCCTCATATCTCTAGGAGGCC
TACCCCCCCTAACCGGTTTCCTACCTAAATGGATCATCATCCAAGAACTCACAATAAACAATAATTTCCTATCCCCTTTATTATAACTACCATAACCCTACTTAACCT
ATACTTTTACATACGCTTAATCTACACCACCTCCCTGACACTATTCCCTACATCCAACAACACAAAAATAACATGACAACTCGAAAATACAAAACCCATACTCCTCAT
CCCCCCACTCATCATCACCTCCACTCTTCTACTACCAATCTCTCCCCTACTCTTAACCTTCCACTAGAAATTTAGGTTAAACAAGACCGAGAGCCTTCAAAGCCCTTA
GTAAGTAAAGACACTTAATTTCTGAAACACACAAGGACTGCAAACACCTACTCTGCATCAATTGAACGCAAATCAACTACTTTAATTAAGCTAAGCCCTTACTAGAT
CAATGGGATTCGAACCCACAAAAGCTTAGTTAACAGCTAAGCACCCTAACCGACTGGCTTTGACCCACTTCTCCCGCCGCAGGAAAAAAAAGGCGGGAGAAGCCCC
GGCAGAAACTCTGAAACTGCTCCTTTGAACTTGCAATTCAACATGAAAATCACCTCGGAACTGGTAAAAGAAGGGTTAGACCTCCATCTTTAGGTTTACAGCCTAAT
ACTTACTCAGCCATTTTACCGCCTATGCTCACCAATCGTTGACTCTTTTCAACAAACCACAAAGATATTGGAACCCTATATTTACTATTTGGTGCGTGAGCTGGAATC
ATAGGCACTGCCCTAAGCCTTCTCATTCGAGCTGAACTAGGCCAACCCGGCAACCTACTAGGCAACGACCACATCTACAACGTCATTGTAACGACCCATGCATTTGT
TATAATTTTCTTTACAGTTATACCCATTATAATTGGGGGGTTCGGAAACTGACTAGTACCCTTAATAATCGGCGCTCCCGACATAGCATTTCCCCGTCTAAATAATAT
AAGCTTCTGACTCCTCCCTCCTTCTTTCCTGCTACTAATAGCATCAGCCGTAGTAGAAGCTGGCGCCGGAACAGGTTGAACGGTATATCCCCCCCTAGCAGGAAACTT
CTCCCACCCAGGAGCTTCTGTGGATTTAGTCATCTTCTCCCTCCACCTAGCAGGTATTTCCTCTATCTTAGGAGCCATCAACTTTATTACTACTATTATCAACATAAAA
CCCCCCGCGATATCCCAATACCAAACCCCTCTATTTGTCTGATCGATCTTAATCACAGCAGTTCTTCTACTCCTCTCTCTGCCAGTCCTAGCCGCTGGCATTACCATGC
TACTAACTGATCGCAACCTCAATACTACTTTCTTTGATCCTGTTGGAGGAGGAGATCCTATCCTATATCAACACCTGTTTTGATTCTTTGGTCACCCTGAAGTTTACAT
CCTTATTCTTCCTGGCTTTGGAATAGTCTCTCACATCGTAGCTTACTACTCTGGAAAAAAGGAACCGTTTGGGTATATAGGTATAGTTTGAGCCATAATATCAATTGG
GTTCTTAGGTTTTATTGTATGAGCCCACCACATATTTACAGTTGGCATAGACGTAGATACACGAGCCTATTTCACTTCCGCCACTATAATCATTGCAATCCCCACCGG
TGTGAAAGTTTTTAGCTGACTTGCTACACTCCATGGAGGCAACATTAAATGATCCCCAGCAATACTTTGAGCTCTGGGCTTTATCTTCCTATTTACTGTAGGGGGCCT
AACCGGCATCATCCTGGCAAACTCATCCCTAGATATCGTACTACACGACACATACTACGTTGTCGCTCACTTCCATTATGTTCTATCAATAGGAGCTGTATTTGCCAT
TATAGGAGGCTTTATACATTGATTTCCCTTATTCTCAGGCTATACACTGAACCAAACCTATGCCAAAGCCCACTTCATCATTATATTCATCGGCGTAAATTTAACCTTC
TTCCCACAACACTTCCTTGGTCTGTCCGGGATACCCCGACGCTACTCCGATTATCCCGATGCCTACACTACATGAAATATCCCGGACTCATCTATAGGCTCCTTCATCTCAC
TAGTAGCAATAATTTTAATAATCTACATAATCTGGGAAGCCTTTGCTTCAAAACGTAAAGTATTACTAATCGAACAACCCCTTACCAACTTAGAATGACTAAATGGC
TGCCCCCCCACCTTATCACACATTCGAAGAACCAGCCTATATTAAACTAAACGAAAAAGGAGAGAGTCGAACCCCCTAAAACTGGTTTCAAGCCAGTCTTATAGCCCC
TATAACTTTTTCAACAAGATATTAGAAAAGTCATTTCATAGCTTTGTCAAAGCTAAGTCATAGGCCAAACCCTATATATCTTCATGGCCCACCCAGTTCAACTAAGC
CTACAAGACGCCACATCCCCTATTATAGAGGAATTAATTACTTTCCATGACCATGCTTTTATAGCCATATCTCTAATCAGCTTCCTAGTGTTATATGCCCTACTCTCAA
CACTCACAACAAAGCTAACCAACACTAACATCACAGACGCCCAAGAAATAGAGACTATCTGAACTATCTTACCCGCAATTATCTTAATCCTAATCGCTCTCCCATCC
CTGCGCATCCTATACCTAACAGATGAAGTCAACGACCCGTCCTTCACCATCAAATCAATTGGGCACCAGTGATACTGAACCTACGGACTACGGAGGCCT
AATTTTCAACTCCTACATGCTACCCCCACTATTCTTAAATCCGGGGGACCTTCGACTCCTAGAAGTTGATAATCGAGTAGTCCTTCCAATCGAAGCCCCTGTACGTAT
AATAATTACATCCCAAGACGTCTTGCATTCATGAACTATCCCCACACTAGGCCTAAAGACAGACGCCGTACCTGGACGTCTAAATCAAACCGTTTTCACAGCCACAC
GACCGGGCGTCTATTACGGACAATGCTCAGAGATCTGCGGCGCAAATCATAGCTTCATACCAATTGTTGCGGAACTAATCCCACTAAAAATCTTTGAAATAGGACCC

GTATTTACTCTGTAAATAATCCTATTTACCCCCCCCAACACGCCCCCACCTCATAAACTCTCAAAATACTCTAAACTCACTGTACAGCTACTCTAGCATTAACCTTTTA
AGTTAAAGACGGGGGAGACATACCCCCTACAGTGAAATGCCCCAGTTAGACACATCGACATGATTCACCACTATTATAACAATACTTCCTACACTATACCTTATCAT
ACAATTAAAACTACTAAACACAAACCACTACCAACCGCCTCTTACAAATAAACCCCAGCCTTCAATCTCACAATAACCACTGACAACCAAAATGAACGAAAACTTGT
TTGCCCCTTTCTCAGCCCCAACAATCCTAAACCAACCCGCCACAATCCCCATTATCCTATTGCCCACCCTATTAGTCTTAACCTCAAAACACCCCATCAATAATCGAC
TAACCACCATTCAACAAAGTCTAACTCAATTCACCCTAAAACAAATAATAAAGACTCACAACGCTAAAGGACAAACCTGGTCTTTAATACTAATATCCCTAATCACT
TTTATTGCCATAACAAACCTCCTAGGACTTCTGCCTTACTCGTTCACACCAACCACCCAACTTTCCATAAACCTAGCTATAGCAATTCCCTTATGAGCGGGCACGGTA
GTAGTAGGGCTTCGCTTTAAAACCAAAAATTCCCTAGCCCACCTCCTACCACAAGGCACACCCACACCTCTCATCCCCATATTGGTAGCAATCGAAACTATCAGCCT
ACTCATTCAACCAGTAGCTCTAGCCGTACGACTAACCGCAAACATCACAGCCGGCCATCTGCTAATTCACCTGACCGGAAACGCCGTACTAGCACTATCAACCATCA
ACCTTTCCATAACTCTACTAACTTCCACACTTCTAGTACTACTAACTGTGCTAGAAATCGCAGTAGCCCTGATTCAAGCCTATGTCTTCACACTCTTAATTAGCCTCTA
TCTGCGCAACAACACCTAATGACCCACCAAACCCATGCTTATCACATAGTTAAACCCAGTCCCTGGCCACTGACAGGAGCCCTATCAGCCTTCTTAATAACATCTGG
CTTAGTCACGTGATTTCACTTCTACTCTACCACCCTACTAATACTAGGCCTACTAACCAACATACTAACCTTGTATCAATGATGACGTGACATTGTACGAGAAAGCAC
ATACCAAGGCCACCACACAGTATCCGTCCAAAAAGCCTTCGATACGGAATAACACTATTCATCATCTCAGAAGTATTCTTTTTTATTGGCTTCTTCTGAGCGTTTTA
CCACTCAAGTCTCGCCCCAACACCCTGCTTAGGAAGCCATTGACCACCAACAGGCATCACCCCCCTAAATCCTATAGAAGTACCCCTCCTAAACACTTCTGTATTACT
TGCATCTGGAGTTACAATCACCTGGGCTCATCATAGCCTCATAAATAATGACCGAAAACAAACAATCCAAGCCCTACTCATCACAATCCTACTAGGCACCTATTTCA
CCCTATTACAAATCTCAGAATACTTTGAAGCACCCTTTACCATCTCCGATGGCATCTATGGTTCAACATTCTTCATCACCACGGGCTTCCATGGACTCCACGTTATTAT
CGGATCCACATTCCTCTTTATTTGCCTCATCCGCCAACTACTATACCACTTCACATCAAACCACCACTTCGGCTTCGAAGCTGCCGCTTGGTACTGACACTTTGTAGAT
GTAATCTGACTACTCCTCTATATTTCTATCTATTGATGAGGATCCTACTCTTTTAGTATAACAAGTACAATTGACTTCCAATCAATTAGTTTTGACAACATTCAAAAAA
GAGTAATCAACCTAGTGCTAGCCTTAACAATCAACACCCTATTAACCTCACTACTGATAATAATCATATTCTGACTACCCCAACTCAACTTCTATACGAGAAAAACTA
GCCCCTACGAATGCGGATTTGACCCCCTAAGCCCTGCTCGCATCCCATTCTCAATAAAATTCTTCCTAATCGCCATTACTTTCCTACTATTTGACCTAGAAATCGCCCT
ACTACTATCCCTACCGTGAGCCATTCAGACAACAAACCTTCCAATAATAATCAAATCAACCACCACCTTTATTATCATCCTAATTCTTAGCTTAGCCTATGAATGAAC
CCAAAAAGGGCTAGACTGAGCCGAATTGGTAAGTAGTTTAAATAAAATAAATGATTTCGACTCATTAGATTATGATAACCATACTAACCAAATGACCCCTACCTACA
TAAACATTATACTAGCATTTACTATCTCTCTTCTAGGCATACTAATCTACCGCTCACACCTAGTAGCTTCTCTCCTCTGCCTAGAGGGGATAATAATATCACTCTTTAT
CATAACTGCCCTTATTGCTTCAAACACACACTCCCCCCTAATCAACATTATACCCATCATTTTGCTAGTATTTGCCGCCTGTGAGGCAGCAGTAGGTCTTGCCTTACTA
ATCTCAATCTCCAATACATACGGACTAGATCACGTCCACAACCTAAGCCTGCTTCAATGCTAAAAATAGTTATTCCCACAACCATACTACTACCTACCTATAACATGACCT
CTAAAAATAATACAATTTGAATCAACTTAACTATACACAGCTTAGCTATCAGCCTCATTCCCTTACTATTCTTTAACCAAACTAACACTAACCTCCTCAATCACTCAA
CCCACCTATCTTCCGACCCACTAACAACACCCCTCCTAACACTAACTATCTGACTCCTGCCCCTCATGATTATAGCAAGCCAGTATCATCTATGCAATGAACCCCCCT
CAAAAAAGAAACTCTTCCTCCTCATAACGATCCTCCTACAAATCACCCTAATTCTAACATTCATGGCCACAGAACTAATTATATTCTACATTCTATTCGAAACCACCC
TTATCCCCACCCTAGTTATCATTACTAAGTGAGGCAACCAAGCAAAACGCCTCAACGCAAGCACATACTTTCTATTCTACACACTAACCGGCTCTTTACCCCTACTCA
TCATACTAAGCCACACCTACAACAACACAGGCTCATTAAATATTACATTACTAACACTCACAGACCAAAAACTAACAACCACCTGATCCCACGGCCTTGCCTGACTA
GCATGCATAATAGCCTTTATAACAAAAATACCCCTATATGGCTTACACCTGTGACTTCCCAAAGCTCATGTCGAAGCCCCCATTGCAGGATCAATAGTTCTTGCCGCA
GTACTCCTAAAACTAGGCGGCTACGGCATGATACGATTTACTCCCATCCTTAACCCTCTAACAGAATACATAGCCTACCCATTCCTCATACTATCTTTATGGGGCATA
ATTATAACAAGCCTAACATGCCTCCGACAAACAGATCTAAAATCACTTATCGCATACTCTTCTGTAAGCCACATAGCCCTTGTAATCGTAGCCTCCCTCATCCAAACC
CCCTGAAGCTTCTCTGGCGCAACCATCCTCATAATCGCCCACGGGCTCACCTCCTCTATGTATTTCTGCTTAGCCAATTCAAACTATGAACGCACCCATAGCCGTACC
ATACTACTATCCCGAGGACTTCAAATCCTGCTTCCACTAACAGCCTTCTGATGATTCGCAGCAAGCCTTACTAACCTTGCTCTACCCCCCACCATCAATCTACTAGGC
GAACTCTTTGTAATCGCGACCTCATTTTCCTGATCCTATATCACCATCATATTAACAGGACTTAACATACTAATCACAGCCCTCTACTCACTCCACATATTCATTACAA
CACAACGAGGAGCACTTACACACCACATAATCAACATAAAACCCCCTTCACACGAGAAAACATATTAATATTCATACACCTCGCTCCAATTATCCTCCTATCTCTC
AACCCCAACATCATCCTAGGGTTTACCTCCTGTAAATATAGTTTAATCAAAACACTAGATTGTGAGTCTAACTATAGAAGCTCACCACTTCTTATTTACCGAGAAAAC
TCACAAGGACTGCTAATCCACATCTCCGTACTTAAAACTACGGTTTTCTCAACTTTTAAAGGATTACAGCTATCCATTGACCTTAGGAGTCAAAAATATTGGTGCAAC
TCCAAATAAAAGTAATAATCATGCACACCACCATCATAATAACAACCCTTATCTCCCTGACTCTCCCAATTTTGGCCACTCTCACCAACCCTTACAAAAAACGCCCAT
ACCCAGATTACGTAAAAACAACCGTAATATATGCCTTCATCATCAGCCTCTCCTCAACAACTTTATTTATCTTCTCAAACCAAGAAACAACCATTTGGAGTTGACATT
GAATAATGACCCAAACATTAAACCTAACACTAAGCTTTAAATTAGATTACTTCTCCATAATATTTACTCCAATCGCACTATTCATCACCTGATCCATTATAGAATTCT
CACTGTGATACATAACCCCAGACCCAAATATTAATCAATTCTTCAAATACCTCCTTATTTTCCTTACCGCTATACTAACTCTGGTCACTGCCAACAACCTCTTCCAATT
CTTTATCGGCTGAGAAGGCGTAGGAATCATATCTTTTCTCCTAATTAGCTGATGACACGCTCGAACAGACGCCAACACAGCAGCCGTCCAAGCAATCTTGTATAACC
GCATTGGCGATATCGGTCTTATCCTAGCCATAACGTGATTCCTCCTACATTATAACTCATGGGACCTCCAACAAATATTGACCCTAAATTCCAACCCAAGCCCCCTAC
CACTAGCAGGCCTCCTCCTGGCAGCAGCAGGGAAATCAGCTCAATTCGGCCTTCACCCCTGACTACCCTCTGCCATGGAAGGCCCAACTCCAGTCTCAGCCCTACTT
CACTCCAGCACCATAGTCGTCGCCGGAGTATTCTTGCTCATCCGCTTCCACCCTCTAATAGAAACCAACACAACAATTCAAAACTTTACGCTATGTCTAGGGGCTATT
ACTACCTTATTTACAGCCGTCTGCGCCCTCACACAAACGACATTAAAAAAATTGTAGCCTTCTCCACTTCAAGCCAACTAGGCCTAATAAATAGTCACCATCGGCATC
AACCAACCATACTTAGCATTCCTACACATCTGCACCCATGCTTTCTTTAAAGCCATACTTTTCATATGCTCCGGATCCATCATCCACAACCTAAATAACGAACAAGAC
ATTCGAAAAATAGGAGGCCTATTCAAAACAATACCCCTCACCTCAACTTCCATAATCATTGGTAACCTAGCACTCACAGGAATACCCTTCCTCACAGGCTTCTACTCC
AAAGACCTTATTATTGAAGCCACAAATACGTCATATACCAACGCCTGAGCCCTATTCACTACTATCATCGCTACCTCCCTAACAAGCGCCTACAGCACCCGAACTATT
CTCCTCACCCTAACAGGACAACCCCGCTTTCCAGCCCTAACAAACATTAATGAAAACAACCCCACCCTACCTAAACCCAATTAAACGCCTCACAATAGGCAGCATAGT
TGCAGGATTTCTTATCACCAACAACGTCCCCCCCACCTCACTCCCCCAACCAGCAATACCCCTCTACTTAAAACTCTCAGCCCTATGCGCAACTGCCCTGGGCTTTCT
AGCAGCCTCAGACCTCATCCTTATAACAAACAAACTGAAAACAAAAAACCCATCACAAATATTCAAATTCTCCAACATACTAGGATATTTTCCCACCACAATCTACC
GCATAATCCCCTATCAAAACCTACTCATAGGCCAAAACTTAGCTCTCCTCCTATTAGACTCGACATGACTAGAAAAATCTATACCTAAAATAATTTCACAAGCACAC
ATCGCTGCTTCCACAGCCGTAACCACCCAAAAAGGTATAATCAAACTCTATTCCCTCTCTTTCCTCATCCCCCTTACCCTAGCCCTAATCCTAATAATATAACCTATTA
CCCCGAACAATCTCAATCACAACATACACCAACAAATAGCGTTCAACCAGCAACCACCACCAATCAACGCCCATAATCATACAAGGCACCCGCGCCAATAGAAT
CACCTCGAACCAATCCTGACCCCTCCCCCTCAAAAATCACCCAATCACCAGTATCACTAAAACCAACCACCACTACCACCCACCTCATCCAAACCTAAAACCCATAAA
ACTATTACCACCTCCATTATCAACCCTACTAAAAGACCCCCTAAAACCTCAATTCCCGAACCCCATGTCTCAGGATACTCTTCAATGGCCATTGCTGTGGTATAACCA
AAAACAACCATCATGCCCCCCAGATAAACCAAAAATATCATTAAACCCATATAAGCCCCCCACAATTTAAAACGATCACACAACCAACCACACCACTAACAATTA
ACGCTAAACCCCCATAAATAGGAGAAGGTTTAGAAGAAAACCCCACAAACCCCATAACCAATAAAACACTCAATACAAACAGAATATACGCCATTGCTCTCACATG
GACTCCAACCATAACCAATGGCATGAAAAACCATCGTTGTACTTCAACTATAAAAGCACTAATGACTCCAATACGCAAATCCAACCCAATCATAAAAATAATCAATC
GCTCCCTCATTGATCTACCCACTCCATCCAACCTTTCCATGTGATGAAACTTCGGCTCACTTCTCACAGCCTGCCTAATTTTACAAATCGTCACAGGCCTCCTCTTAGC
AATGCACTACTCACCAGATACCTCCTCCGCCTTCTCCTCAATTGCACACATTACCCGAGATGTAAAATACGGCTGAACTACTCGCTACCTTCACGCCAACGGTGCCTC
TATACTCTTCATCTGCCTTTTCCTACACATCGGCCGAGGCCTCTACTATGGCTCATACCTCCTCCTAGAAACCTGAAACATTGGCATCATACTCCTTCTCCTAACCATA
ACAACAGCTTTCATGGGTTACGTTCTCCCATGGGGCCAAATATCATTCTGAGGGGCAACAGTAATCACAAACCTACTATCAGCAGTCCCATACATCGGAACCGATCT
CGTTCAGTGAGTTTGAGGAGGATACGCTATCGATAGTCCCACTCTCACACGATTCTTCACCTTGCACTTTATCCTACCCTTCATTACCATCGCCCTC

Bảng S4. Trình tự nucleotide trên ADN ty thể *Chà vá chân nâu* Tây Nguyên Việt Nam

AACTCAACTGTCGTTTTCTCAACTAATAAAAATTTATTTCATACACCACCCTTGACACCTCCACCCACCTCTTCTTTCCCATAATCCCCAAAGATGGTATCATAATACC
CATTATTCAGTTTATGTAGCTTAACACCTCCAAAGCAAGACACTGAAAATGCCTAGACGGGCTCACAGACCCCATAAACAAACAGGTTTGGTCCCGGCCTTTCTATT
AGCTTTTAGCAAAATTACACATGCAAGCATCCCCGCTCCAGTGAAGATACCCTATAAATCACGACGATTAAAAGGAGTAAGTATCAAGCACGCTTAATGCAGCTCA
AGACACTTTGCTTAGCCACACCCCCACGGGAAACAGCAGTGACTAATATTTAGCAATAAACGAAAGTTTAACTAAGCTATGCTACTAAGGGTTGGTCAATTTCGTGC
CAGCCACCGCGGCCATACGATTGACCCTAGTTAATGACCTCGGCGTAAAGGGTGTTTTAGATAAATTTAATTAAATAAAGCTAAACTCTTTCTAAACTGTAAAATC
CTAGCTAATGTAAAATAGACTACGAAAGTGGCTTTAAAGCTTCTGAATACACAACAGCTAAGACTCAAACTGGGATTAGATACCCCACTATGCTCAGCCCTAAACTT
TAATAGTTAAAAATAACAAAACTATTCGCCAGAACACTACAAGCAACAGCTTAAAATTTAAAGGACTTGGCGGTGCTCCATATCCCCCTAGAGGAGCCTGTTCTATA
ATCGATAAACCCCGATTCACCTCACCACCTCTTGCTAAGCCTATATACCGCCATCTTCAGCAAACCTCAACAAGAGATATAAAGTAAGCACAAACGCCCACGCAAAA
ACGTTAGGTCAAGGTGTAGCTTATGAGCTGGAAGAAATGGGCTACATTTTCTACCCTAGAAAACCCCACGATAGCTCTTATGAAACTTAAGAGTCCAAGGAGGATTT
AGTAGTAAACTAAGAATAGAGTGCTTAGTTGAACTAGGCCATGAAGCACGCACACACCGCCCGTCGCTCTCCTCAAATATAATTAGAACTAACTTAACTAAATATTT
CTGCATTCATACAGAGGAGATAAGTCGTAACATGGTAAGTGTACTGGAAAGTGCACTTGGATAAATCAAGGCATAGCTCAACACAAAGCATCCAGCTTACACCCGG
AAGATATCGCTACCTGATTGCCTTGAGCCAATACTAGCTCAACAAACCTACCAATACTACTATTAAATTAATATATTATTAAACCATTCACAAGTATAAAAGTATAG
GAGATAGAAATTTACTCTGACGCTATAGATGCAGTACCGCAAGGGAAAGATGAAAAAAGCAACCAAGCATAAAACAGCAAGGACTCACCCCTATACCTTTTGCATA
ATGAGCTAGCTAGAGACCACTTCGCAAAGAGAACTAAAGCCAAGTACCCCGAAACCAGACGAGCTACCCAAAAACAGCTAAAAGAGCACACCCGTCTATGTAGCA
AAATAGTGGGAAGATTTCTGGGTAGAGGTGATACGCCTACCGAGCCTGGAGATAGCTGGTTATCCAAGATAGAATTTTAGTTCAACCTTAAGTTTACCTACAGAACT
ACCAATCCTCCTGTAAACTTAACTGTTAGTCTAAAGAGGGACAGCTCTTTAGACATCAGGAAAAAACCTTATGTAGAGAGTAAAAAATTTTCAAACTCATAGTTGGC
CTAAAAGCAGCCATCAATTAAGAAAGCGTTTAAGCTCAACGCAACGTCTAAAAAAAATATCAAACACTTTACTGAACTCCTCATATCACATTGGATTAATCTATCAT
TTCATAGAAGCAATAATGCTAGCATTAGTAACGTGAATAAATTCTCCAGTGCATAAGCCTAAATCGGATCGAAACCTTCACCGATATTTGACAGTCTAATGTTATAA
CCACAAACAAGCCAATATTATATATACTGTTAACCCAACACAGGCATGCCCTAAAGGAAAGGTTAAAAAAAGTAAAAGGAACTCGGCAAACTCAACCCCGCCTGTT
TACCAAAAACATCACCTCTAGCATCACCAGTATTAGAGGCACTGCCTGCCCAGTGACATATGTTTAACGGCCGCGGTACCCTGACCGTGCAAAGGTAGCATAATCAT
TTGTTCTTTAAATAGGGACTTGTATGAATGGCAACACGAGGGTTTAACTGTCTCTTACTTTCAACCAGTAAAATTGACCTGTCCGTGAAGAGGCGGACATAAAATAA
TAAGACGAGAAGACCCTGTGGAGCTTCAATTTACTAGTACAGCCCATTATTAAAATAAATCTAAGGACCTAACATACCCTGCCCCTGTACTAGAAATTTTGGTTGGG
GTGACCTCGGAGCATAATTAAACCTCCGAACGAACTACGCCAAGGCCATACAAGTCAAAGCAGTCTAATATCTAAATTGATCCAATAACTTGACCAACGGAACAAG
TTACCCCAGGGATAACAGCGCAATCCTATTCCAGAGTCCATATCGACAATAGGGTTTACGACCTCGATGTTGGATCAGGACATCCTAATGGTGCAGCAGCTATCAAG
GGTTCGTTTGTTCAACGATTAAAGTCCTACGTGATCTGAGTTCAGACCGGAGCAATCCAGGTCGGTTTCTATCTATTCTATATTTCTCCCTGTACGAAAGGACAAGAG
AAATAGGGCCCACTTCACACAAGCGCCCTCCTCCCATAAATGACCTAGTCTCAATTTAGCAAGAAATTACACACACAACCCAAGAACAGGGATTGTTAAGATGGCA
GAGCCCGGTAATTGCATAAAATTTAAGACTTTATAATCAGAGGTTCAACTCCTCTTCTTAGCACTATGTTCACAATAAATCTTCTACTCATTATCCTACCCACTATAGC
TGCCATAGCATTTCTTACACTTACTGAACGAAAACTATTAGGCTATATACAACTACGCAAAGGACCCAACATCGTGGGTCCTTATGGACTACTACAACCCTTTGCCG
ATGCAATAAAACTCTTCACCAAAGAACCCTTAAAACCCTCAACATCCTCTACCACCCTGTATATTATTGCACCCGCCCTAGCCTTTTCTATTGCCCTTCTCCTGTGAGT
ACCTCTTCCCATACCCAATTCCCTAATTAATCTTAATCTAGGACTTCTATTTATCCTAGCTACATCTAGCCTAGCTGTTTACTCCATTTTATGATCAGGATGAGCATCC
AACTCAAACTACGCATTAATCGGAGCGCTGCGAGCAGTCGCCCAAACAATTTCATACGAGGTAACTCTCGCCATCATTATACTATCAGTTCTACTAATAAGTGGCTC



CTCACCTAGCACCTCCTCAGCCTTCTCCTCAATCGCCCATATCACTCGAGACGTAAACTACGGCTGAATTATCCGCTACCTTCATGCCAATGGTGCCTCCATATTCTTT
ATCTGCCTATTCCTACATGTAGGTCGAGGGTTATATTATGGCTCATTCCTTCTTCTTGAAACTTGAAACATTGGCATTGCACTATTACTTATAGTTATAGCAACGGCCT
TTATAGGCTATGTACTCCCATGAGGCAAATATCATTTTGAGGTGCTACAGTAATTACAAATTTATTATCCGCAATCCCATACATCGGAACAAATCTCGTTCAATGAG
TATGAGGTGGGTATTCTATCGATAATCCAACCCTTACCCGATTCTTCACTCTCCACTTTACCCTGCCTTTCATTATCACAGCCTTTACAGTTCTACATCTACTTTTCCTA
CACGAAACAGGATCTAATAACCCATGCGGAATTCCCTCAAACTCCGACAAAATCCCCTTCCACCCCTACTACACAATCAAAGATATGCTAGGCCTAGTCCTCCTTAT
TCTTCTCCTAATAACTCTAGTATTATTTTCACCTGACCTTTTAGGTGACCCAGACAACTATACACCAGCTAATCCACTAAACACCCCACCACACATCAAACCAGAATG
ATACTTCTTATTCGCATATGCAATCCTACGATCTGTACCTAACAAATTGGGAGGCGTACTAGCACTTCTTATATCCATTCTTATCCTAATAGTTATCCCCATACTTCAC
AAATCCAAACAACAAAGCATAATATTCCGCCCATTCAGCCAATTTACACTATGATTACTAATCACAGTTTTATTAACCCTAACCTGAATTGGAAGCCAACCAGTAAA
CCAACCCTTTATTATAATTGGACAGGTAGCATCTATAATATATTTCACTACAATTCTGATCTTAATACCACTAGCCTCTATAATCGAAAACAATCTCCTCAAATGAAC
CTGCCCTTGTAGTATAGACTAACACACCGGTCTTGTAAACCGGAGACGGATACCTTTCCCCAGGGCAACTCAGAAAGAAAGCATTTAACTCCTCCACCAATACCCAA
AACTGGCATTCTATTTAAACTACTTTCTGCATTCTAAGGAGGTATAATCTTTAGAGAATAATTTAGTACAATCTAATTTTATATGTCCTTATGTAATTCGTGCATTACT
GCTAGTCACCATGGATATTATATAGTACTATAAATGTTTTACCGTCCATAGAACATAAAATTACATATTTACTAGCAGTTATATCCAGAACATGCTTATAAGCAAGCA
CCTTAATAGAATATTTAACTATAACACATACCATTCATACATTCAAATCCCATGGTCACACCCACTGGAATACCAACCAATAGCGATTAATCCATTATCGTACATAGT
ACATAATATCCTTTACCGGGCATAGCACATCACAGTTGAGCATCCCTAAACCAATTCTTGCCAACACGGATATTCCCGCCAGTTAGGTGTCCCTTGATCACCATCCTC
CGTGAAATCAATATCCCGCACAAGAGTGCTACTCTCCTCGCCCCGGGCCCATAACCTGTGGGGGTAGCTATACATGACGTCAACGGACATCTGGTTCTTACCTCAGG
GCCATACCGCCAAGATCGCCCACACGTTCCCCTTAAATAAGACATCTCGATGGATCACGGGTCTATCACCCTATTAACCAGTCACGGGAGCTCTCCATGCATTTGGT
ATCTTTTATCTCTGGTCTGCACGCGACCCCATTGCAGTATGCTGGTCTCGCCACAATCAGTCCCGCAGCGCCTGTCTTTGATTCCTAGCACATGCCATTATTTACCGCA
CCTACGTTCCATATTTTAACCCCGCATGAACCTTATCATGGTGTTATTTAATCCATGCTTGTAAGACATACATATAATTAACCGTATACCCCGATCGCACCTTGATCCA
ATTTACAACTATCTCTACTCAAACCCCCCCCCCCCCCCCATATCCGACCGCACCCTAATAATACGCCTTTGCCAAACCCCAAAAACAAAAACTTGCAATCTGGTCGAAT
TTTACATTTTTTTTTTATCTTTTAGGTGTGCAT

Bảng S5. Trình tự nucleotide trên ADN ty thể *Khỉ vàng* Tây Nguyên Việt Nam

TCTCATCAAACCCCCCCACCCCCATCTCCGACCTTCATCCAAACCCACTCTTGCCAAACCCCAAAAACAAAAGTCTTAACATATCCGATCAGAGCCCGTATTTT
TATCTTTTGGGTGTACACAACTCCAACTGCCATTCCCTCAACTAACAAACATTTACTTCACCAAACACCCTCTTCACCAACCCACGACAAATCCTTCTCATACAACCC
GAAAGAAACTGCCTCACAATTATACTGAGACCACCGTTTATGTAGCTTAAACTCACCCAAAGCAAGACACTGAAAATGCCTAGATGGGTTTGCACACCCCATAAAC
AAATAGGCTTGGTCCTGGCCTTTCTATTAGCTCTTGGCAGGATTACACATGCAAGCATCCTCGCTCCGGTGAAGACGCCCTATAAATCATCATGACCAATGGGAGCA
AGCATCAAGCACGCACGCGCAGCTCAAAACGCTTTGCCTAGCCACACCCCCACGGGAGACAGCAGTGACAAGTATTTAGCAATAAACGAAAGTTCAACTAAGCCAT
GCTATATTTAGGGTTGGTCAATTTCGTGCCAGCCACCGCGGTTACACGATTAACCCTAGCTAATAGAGACCGGCGTAGAGGGTGTTTAAGATCTGACATAATAAAGC
TAAACTCCATCTAAACTGTAAAACTCTAGCTGATGTAAAATAAACTACGAAAGTGGCTTTAAAGCTTCTGAACACACAATAGCCAGGACCCAAACTGGGATTAGAC
ACCCTACTATGCTTGGCCCTAAACCTCAGTAGTTAAATAACAAAACTACTCGCCAGAATACTACGAGCTATAGCTTAAAACTCAAAGGACTTGACGGTGCTTTACAC
CCCCTAGAGGAGCCTGTTCCGTAATCGATAAACCCCGATCCACCCCACCCTCTCTTGCTCAGCCTATATACCGCCATCTTCAGCAAACCCTAATGAAGGTCACAAAGT
GAGCGCAAACGCCGTCACCGCGAGTACGTTAGGTCAAGGTGTAGCCTATGAGATGGTAAAAAATGGGCTACATTTTCTACCTCAGAAAATTCCACGAAAACCCTTAT
GAAATTTAAGGGCCCAAGGAGGATTTAGCAGTAAATTAAGAATAGAGTGCTTAATTGAACCAGGCCATAAAGCGCGCACACACCGCCGTCACTCCTCTCAAATAT
ATTTAAGGAACATCTTAACTAAACGCCCTAATATTTATATAGAGAGGATAAGTCGTAACATGGTAAGCGTACTGGAAGGTGCGCTTGGATAAATCAAGGTATAGCTT
AATATAAAGCACCTGGCTTACACCCAGAAAATCTCAATACCACTTGATTGCCTTGAGCCAATACTAGCCCTAAGCATAACTAACACTAATACCAAACCAACATGCAC
TAAACCATTCACCTACACAAAGTATAGGCGATAGAAATTTTAATCTGGCGCTATAGACGCAGTACCGTAAGGGAAAGATAAAAACCACCCAAGCACAAAACAGCAA
GGACTAACTCCTGTACCTTTCGCATAATGAATTAGCTAGAAACAATTTCACAAAGAGGACTACAAGCCAAATTCCCCGAAACCAGACGAGCTACCCAAAAACAGCT
AAGAGAGCGCACCCATCTATGTGGCAAAATAGTGGGAAGATTTCTGGGTAGAGGTGACAAGCCTATCGAGCCTGGTGATAGCTGGTTATCCAAGACAGAATCTTAG
TTCAACCTTAAGCTTACCTATAGAACCACCTAATCCCCCTGTAAGCTTAATTGCTAGTCTAAAGAGGGACAGCTCTTTAGACACTAGGAAAAAACCTTATAGAGAGA
GTAAAACCCCCAATCACCATAGTTGGCCTAAAAGCAGCCATCAATTAAGAAAGCGTTCAAGCTCAACACTAACCACCCCAAAAATACCAAACACATAACTGAACTC
CTCACACCACATTGGATTAATCTATTACCATATAGAAGTAATAATGCTAGTATAAGTAACATGAATATTTTCTCCACTGCATAAGCCCAAATCGGATCGAAAACCTC
ACCGATACCTAACAGCCAGGCATAGCAAGCACAAACAAGCCCATGCTACCTTCACTGTTAACCCAACACAGGCATGCCTCAAGGAAAGGTTAAAAAAAGTAAAAG
GAACTCGGCAAACTTAAACCCCGCCTGTTTACCAAAAACATCACCTCTAGCATTACTAGTATTAGAGGCACTGCCTGCCCAGTGACACACGTTTAACGGCCGCGGTA
CCCTGACCGTGCAAAGGTAGCATAATCACTTGTTCTTTAAATGGGGACTCGCATGAAAGGCACCACGAGGGTTTAACTGTCTCTTACTTTTAACCAGTGAAATTGAC
CTGCCCGCGAAGAGGCGGGCATAAAATAATAAGACGAGAAGACCCTATGGAGCTTTAATGTATTAATGCAATCAAAAACTACACAAACCCACGGACCTTTAAACTA
CCGACACCTGCATTAAAAATTTTGGTTGGGGCGACCTCGGAGCACAACAAAACCTCCGAATAACACATGCTAAGACCACACAAGTCAAAGCGAACTAACACTTACA
ATTGATCCAATAATTTGATCAACGGAACAAGTTACCCTAGGGATAACAGCGCAATTCTATTCTAGAGTCCATATCGACAATAGAGCTTACGACCTCGATGTTGGATC
AGGACATCCTAATGGTGCAGCAGCTATCAAGGGTTCGTTTGTTCAACGATTAAAGTCCTACGTGATCTGAGTTCAGACCGGAGCAATCCAGGTCGGTTTCTATCTATT
GCACATTTCTCCCTGTACGAAAGGACAAGAGAAATAAGGCCCACTTCACAAAGCGCCTTCACTACATAAATGACCTAATCTTAATTTAGCAAAATACCACACATTCC
ACCCGAAAACAGGGTTTGTTAAGATGGCAGAGCCCGGCAATTGCATAAAACTTAAAACTTTATAACCAGAGGTTCAATTCCTCTTCTTAACACCATGACCACCATAA
ACCTCCTACTTATAATCATACCCACACTAGCCGCCATAGCATTCCTCACACTCGTTGAACGAAAATTACTAGGCCACATGCAACTACGAAAAGGGCCCAATATTGTA
GGTCCATGTGGACTATTACAACCCTTCGCTGATGCCGTAAAACTTTTCACTAAAGAACCCCTAAAACCCGCAACATCTACCACCATCCTCTACATCGCCGCACCCGCT
CTAGCCTTCTTCATTGCCCTCCTCCTATGAACCCCCCTCCCCATACCCAACGCCCTAATCAACCTCAACCTAGGGCTCCTATTCACCCTAGCCATACTCAGCCTAACTG
TCTACTCCATCCTATGATCAGGATGAGCATCAAACTCAAACTACGCCCTAATCCGGAGCCCTACGAGCCGTCGCCCAAACAATTTCTTACGAAGTCACTCTTACTATTA
TCCTCCTATCAGTCCTACTAATAAGCGGTTCATTTAACCTCAACACACTCATTACAACGCAAGAACATCTATGACTCCTCCTACCATCATGACCCCTAGCCATAATAT
GATTCACCTCCACACTAGCAGAAACAAACCGAGCCCCTTTTGATCTCATAGAGGGCGAATCAGAATTAGTATCAGGCTTTAACATTGAATACGCCGCAGGTCCATTC
GCCTTATTCTTCATAGCCGAATACACAAACATCATCATAATAAACGCCCTAACAACCACAATCTTCCTAGGCACATTCCACTCCATTCACCAGAACTATTCACA
ACATGCTTCGCCACCAAAACACTCTTCCTAACCTCCCTGTTCCTATGAATCCGGGCAGCACACCCCCGATTCCGTTACGACCAACTTATACATCTACTATGAAAAAAC
TTCCTCCCCCTCACACTAGCACTACTCATATGATCCACTTCAACACCCATCGCAATCTCCAGCATTCCCCCCCAAACCTAGAAATATGTCTGACTAAAGAGTTACTTT
GATGGAGTAAATAATAGAGGTCCCAACCCTCTTATTTCTAGGACTGTAGGTATCGAACCTACCCCTGAGAACCCAAACTTCTCCGTGCTACCCCACACACTCTATCCT
AAAGTAAGGTCAGCTAAACAAGCTATCGGGCCCATACCCCGAAAATGTTGGTCATATCCTTCCCGTACTAATTAACCCACTAGCTCAACCCATTATCTACACCACAA
TAATCACAGGCACACTTATCACAACACTAAGCTCACACTGATTTCTCGCCTGAACAGGTCTAGAAATAAATATACTAGCCTTCATTCCAATCTTAATAAAAAAAACA
AGTCTCCGCTCTACAGAAGCCGCTACCAAATACTTCCTAGTACAATCTACCGCATCCATAATCCTCATAATAGCAATCATCTCCAACAACCTACTATCAGGATGCTGA
ACAACAACAAACTACATCAGCCAACCCCCATCCCTAATAATAACAACTGCTCTCGCCATAAAACTAGGAATAGCCCCCTTCCACTTCTGAATCCCAGAAGTTACCCA
AGGAACACCCCTAACCTCCGGTCTACTCCTTCTTACATGACAAAAACTAGCCCCCATCTCAATCATATACCAAATTCACCCATCAATCAACACCCACATCCTCCTAAC
CCTCTCCACCTTGTCCATTGCAGTAGGCAGTTGAGGAGGTCTAAACCAAACACAATTACGCAAAATCCTAGGGTACTCTTCAATCACACATATGGGTTGAATAATGA
TAACACTAGTATACAACCCAACAACTACAATTCTTTATTTAACTATATACATTATCCTAACAACTACCATATTCCTGACTCTTAACTTAAGCTCAAGCACCACAACCC
TTATACTATCCCGCACCTGAAACAAATCACCCACCTAGCACCACTAATAGCACCCATTCTCATATCTCTAGGAGGCCTACCCCCCTTAACCGGTTTCCTACCTAAAT
GAATTATCATCCAAGAACTCACAATGAACAACAACTTCCTCATCCCCTCTATTATAACTATCATAACCCTACTTAACCTATACTTTTATATACGCCTAATCTACACCA
CCTCCCTAACGCTATTCCCTACATCCAACAACACAAAAATAACATGACAACTTGAAAATACAAAACCCACACTCCTTATCCCCCCACTCATCATCATCTCCACTCTTC
TACTACCAATCTCCCCCCTACTCTTAACTTTTCACTAGAAATTTAGGTTAAACAAGACCAAGAGCCTTCAAAGCCCTTAGTAAGTAGAAACACTTAATTTCTGAAACA
CATAAGGACTGCAAACACCTACTCTGCATCAATTGAACGCAAATCAACTACTTTAATTAAGCTAAGCCCTTACTAGATCAATGGGACTCGAACCCACAAAAACTTAG
TTAACAGCTAAGCACCCTAACCGACTGGCTTTGACCCACTTCTCCCGCCGCAGGAAAAAAAAGGCGGGAGAAGCCCCGGCAGAAACTCTAAGACTGCTCCTTTGAA
TTTGCAATTCAACATGAAAATCACCTCGGGACTGGTAAAAGAGGAACAACCCCGTCTTTAGGTTTACAGCCTAATACTTACTCAGCCATTTTACCACTTATGCTCAT
CAATCGCTGACTCTTTTCAACAAATCACAAAGACATTGGAACCCTGTATTTACTGTTTGGTGCATGAGCTGGAGTCATAGGCACTGCCCTAAGCCTCCTCATTCGAGC
TGAACTGGGCCAACCCGGCAGCCTACTAGGCAACGACCACATCTATAACGTTATTGTAACGGCCCATGCATTTGTTATAATTTTCTTTACAGTCATACCCATTATAAT
TGGAGGGTTCGGGAACTGACTAGTACCTCTAATAATCGGCGCCCCCGACATAGCATTTCCCCGTCTAAATAATATGAGCTTCTGACTCCTACCCCCTTCTTCCTGCT
ACTAATAGCATCCGCCGTGGTAGAAGCTGGCGCCGGAACAGGATGAACAGTATACCCCCCTTTGGCAGGAAACTTCTCCCACCCAGGAGCTTCTGTAGACTTAGTCA
TCTTCTCTCTCCACCTAGCAGGTATTTCCTCTATCTTAGGAGCCATCAACTTCATTACCACCATCATCAACATAAAACCCCCCGCAGTATCCCAATACCAAACCCCTTT
ATTTGTCTGATCAATCTTAATCACAGCAATCCTTCTACTCCTCTCTCTACCAGTCCTAGCCGCTGGCATTACCATGCTACTAACAGATCGCAACCTCAATACTACTTTC
TTTGACCCTGTCGGAGGGGGAGATCCTATCCTATATCAACATCTATTTGATTCTTTGGCCACCCCGAAGTCTACATCCTTATTCTTCCCGGCTTTGGAATGATCTCTC
ACATTGTGGCTTACTACTCTGGAAAAAAAGAACCATTTGGGTATATGGGCATAGTTTGGGCCATAATATCAATTGGGTTTTTAGGTTTTATTGTATGAGCCCACCACA
TGTTTACAGTTGGCATAGACGTAGACACACGAGCCTATTTCACTTCCGCCACTATAATCATTGCAATCCCCACCGGTGTGAAAGTTTTTAGTTGGCTTGCTACACTTC
ATGGGGGTAACATTAAATGGTCCCCCGCAATACTTTGAGCCCTAGGCTTTATCTTCCTATTTACTGTAGGTGGTCTAACCGGCATCATCTTGGCAAACTCATCCCTAG
ATATCGTACTACACGATACATACTACGTCGTCGCCCACTTTCATTATGTTCTGTCAATAGGAGCTGTATTCGCCATTATAGGGGGCTTTATACACTGATTTCCTTTATT
CTCAGGCTACACACTGAACCAAACCTGTGCCAAAGCCCACTTCATCATTATATTCGTGGGCGTAAATTTAACCTTCTTTCCACAGCACTTCCTTGGCCTGTCCGGAAT
GCCCCGACGCTACTCTGATTATCCCGATGCCTACACCCACATGAAATACCCTATCATCTATAGGCTCCTTCATCTCACTAGTAGCAGTAATTTTAATAATCTACATGAT
CTGAGAAGCCTTTGCTTCAAAACGTAAAGTACTGCTAATTGAACAACCCCTTACTAACCTAGAATGACTAAATGGCTGCCCCCCGCCTTATCACACATTCGAAGAAC
CAGCCTACATTAAGCTAGTTGAAAAAGGAGGGAGTCGAACCCCCTAAAACTGGTTTCAAGCCAGTCTCATAGCCCCTACAACTTTTTCAACAAGATATTAGAAAAGT
CATTTCATAGCTTTGTCAAAGCTAAATCATAGGCCAAACCCTATATATCTTACATGGCCCACCCAGTTCAACTAAGCCTGCAAGACGCCACATCTCCTGTTATAGAGG
AATTAATTACTTTCCACGACCATGCTTTTATAGCCATATCTCTTATCAGCTTCCTAGTATTATATGCCCTGCTCTCAACACTCACAACAAAACTAACCAATACTAACAT
CACAGACGCCCAAGAAATAGAAACCATCTGAACTATCTTGCCCGCAATTATCTTAATCCTAATTGCCCTCCCATCTCTACGCATCCTATACCTAACAGATGAAGTCAA
CGACCCATCCTTCACCATCAAATCAATCGGACATCAGTGATACTGAACCTACGAATACACGGACTACGGGGGCCTAATTTTCAACTCCTACATACTACCCCCACTATT
CCTAAACCCAGGAGACCTCCGACTCCTAGAAGTTGACAATCGAGTAGTTCTTCCAATTGAAGCTCCTGTACGTATAATAATTACATCTCAAGACGTCTTACACTCATG
AACTATCCCCACACTAGGCCTAAAAACAGACGCCGTACCTGGACGTCTAAATCAAACCGTTTTCACGGCCACACGACCAGGCGTCTATTACGGACAATGCTCAGAG
ATCTGTGGCGCAAATCATAGCTTCATACCAATTGTTGCAGAACTGATCCCACTAAAAATCTTTGAAATAGGACCTGTGTTTACTCTGTAAATAACCCTGCCCACTCTC
TTCCCTCCTAGCACTCCTCTCCCACCCCATAAACTCTCAAAACACCCTACAAACTCACTGTATAGCTATCCTAGCATTAACCTTTTAAGTTAAAGACGGGGACACACC
CCCTACAGTGAAATGCCCCAGTTAGACACATCGACATGATTCACCACTATTGTGCGATGCTCCCTACACTATATCTTATCATACAGCTAAAGCTATACAGCTAAACACGAA
CCACTACCAACCACCCCTTACAAAAAATCCCACCCCTCAATCCCACAACATCCGCTGACAACCAAAATGAACGAAAACTTGTTTACCCCTTTCTCAACCCCAATAAT
TTTAAATCAACCCGCTACAATCCCCATTATCCTATTCCCTACCCTACTAATCTTAACCCCCAAGCACCCCATTAATAATCGACTAGCTACCATTCAACAAGACCTGAT
TCGATTCACCCTAAAACAAATAATAATAACTCACAACGCTAAAGGACAAACCTGATCCTTAATACTAACATCCCTAATCACCTTTATTGCCATAACAAACCTCCTAG

GACTTATGCCTTACTCATTCACACCAACTACCCAACTTTCCATAAACCTAGCTATAGCAATCCCTCTATGATCGGCCACAGTAATAGTGGGGCTCCGCTTTAAAACCA
AAAACTCCCTAGCCCACCTCCTACCACAAGGCACACCCACACCTCTCATTCCCATGTTAGTAGCAATCGAAACTATCAGCCTACTTATTCAACCAGTAGCTCTAGCCG
TACGATTAACCGCAAACATCACAGCCGGCCATCTGCTAATACACCTAATCGGAAACGCTGTACTAGCACTATCAACCACCAACCTCTCCATAACTCTGCTAACCTCC
ACACTTCTAGTATTACTAACCACTCTAGAAATTGCAGTAGCCCTAATTCAAGCCTATGTCTTCGCACTCTTAATTAGCCTTTATCTACATAACAACACCTAATGACCC
ACCAGACCCACGCTTACCACATAGTTAAACCCAGTCCCTGACCGCTGACAGGAGCTATGTCAGCCTTCTTAACAACATCCGGCTTAGTCATGTGATTTCACTTCTACT
CTACCACCCTGCTAATACTAGGCCTACTAACCAATACACTAACCTTATATCAATGATGACGTGACATTGTGCGAGAAAGTACATACCAAGGCCACCACACAACACCC
GTCCAAAAAAGCCTTCGATACGGAATAACACTATTCATCATCTCAGAAGTATTCTTCTTCATCGGCTTCTTCTGGGCATTTTACCACTCAAGCCTTGCCCCAACACCC
TGCTTAGGAGGCCATTGACCACCAACAGGCATCACCCCCCTAAACCCCATAGAAGTACCCCTCCTAAACACTTCTGTATTGCTTGCATCCGGAGTTACAATCACCTG
AGCTCATCACAGCCTCATAAACAACGACCGAAAACAAACAATCCAAGCCCTACTCATCACAATCCTACTCGGCACCTATTTCACCCTACTACAAATCTCAGAATACT
TCGAAGCACCCTTCACCATCTCCGATGGCATCTACGGTTCAACATTCTTCCTAACCACAGGCTTCCATGGACTCCATGTTATTATCGGATCTACATTCCTCTTTATCTG
CTTCATCCGCCAACTACTACACCACTTCACATCAAACCATCACTTCGGCTTCGAGGCTGCCGCTTGATATTGACACTTCGTAGATGTTGTCTGACTACTCCTCTACATT
TCTGTCTATTGATGAGGATCCTACTCTTTTAGTATAACAAGTACAATTGACTTCCAATCAATCAGTTTTGACAACATTCAAAAAAGAGTAATTAATCTAGTGCTAGCC
TTAACAATCAACACCCTATTAACCTCGCTATTGATAGTTATCATATTCTGACTACCCCAGCTCAACCCCTATATAGAAAAAACCAGCCCCTACGAGTGCGGATTTGAC
CCCTTAAACCCCGCTCGCATCCCATTCTCAATAAAATTCTTCCTAGTTGCCATTACTTTCCTACTATTTGACCTAGAAATCGCCCTACTACTATCCCTACCGTGAGCCA
TTCAAACAACAAATCTCCCAATAATAATCAAATCGACCATCGCCTTTATTATTATCCTAATTCTCAGTCTAACCTATGAATGAACCCAAAAAGGGCTAGACTGAGCC
GAATTGGTAAGTAGTTTAAACAAAATAAATGATTTCGACTCATTAGATTATGATAACCATACTAACCAAATGACTCTTACCTACATAAATATTATATTAGCATTCACT
ATCTCTCTTCTAGGCATACTAACCTACCGCTCACACCTAGTAGCCTCTCTCCTCTGCCTAGAAGGAATAATGATATCACTCTTCATCATAGCCGCCCTTATTGCCTCAA
ACATACACTCCCCTCTAATCAACATTATACCCATCATCTTGTTAGTATTCGCCGCCTGTGAAACAGCAGTAGGTCTTGCTTTACTAATCTCAATCTCCAACACATACG
GACTAGATTACGTTCACAACCTAAACCTGCTTCAATGCTAAAAATAATTATCCCCACAACCATACTACTGCCCATAACATGACTCTCTAAAAATAATACAATTTGAAT
CAACTTAACTATACATAGCTTAACTATCAGCCTCATTCCCTTACTATTCTTCAACCAAACCAACACCAACCTCCTCAATCACTCAACCTACCTATCCTCCGACCCACTA
ACAACACCCCTCCTAACACTAACAATCTGACTTCTACCCCTCATGATTATAGCGAGCCAGTATCCACCTATGCAATGAACCCCTCTCACAAAAGAAACTCTTCCTCTTC
ATAACAATTCTCCTACAAATCACCCTAATCCTAACATTCATGGCCACAGAACTAATTATATTTTACATTCTATTTGAAACCACCCTCATTCCCACCCTAATTATCATTA
CCAAATGGGGCAACCAAGCAAAACGCTTCAACGCAAGTACTTACTTTCTATTCTACACACTAACCGGCTCTCTACCCCTACTCATCATACTAAGCCACACCTACAAC
AACACAGGCTCATTAAATATTACATTACTAACACTCACAGACCAAAAACTAACAACCACCTGATCTCACAACTTTACCTGACTAGCATGCATAATAGCCTTTATAAC
AAAAATACCCCTATATGGCCTACACCTGTGACTTCCTAAAGCCCATGTCGAAGCCCCCATTGCGGGATCAATAGTTCTTGCCGCAGTACTCCTAAAACTAGGTGGCT
ATGGCATAATACGACTTACTCCCATCCTTAACCCCCTAACAGAATACATAGCTTACCCATTCCTCATATTATCCTTGTGGGGCATAATCATAACAAGCCTAACATGCC
TCCGACAAACAGATCTAAAATCACTTATCGCATACTCTTCTGTAAGCCACATAGCCCTTGTAATTGTAGCCTCCCTCATCCAAACCCCCTGAAGCTTCTCCGGCGCAA
CCATCCTCATAATTGCTCACGGACTCACCTCTTCCATATATTTCTGTCTAGCCAATTCAAACTATGAACGCACTCACAACCGTACTATACTACTGTCCCGAGGACTTC
AAATCCTACTTCCACTAATAGCCTTCTGATGATTAACAGCAAGCCTCACTAACCTTGCCCTACCCCCCACTATCAACCTACTAGGCGAACTCTTTGTAATCGCGACCT
CATTCTCCTGGTCCCATATCACCATTATATTAACAGGACTTAACATACTAATTACGGCCCTCTACTCCCTCCACATATTCACCACAACACAACGAGGAGCACTCACAC
ATCACATAATCAACATAAAACCCCCCTTCACACGAGAAAACATATTAATATTCATACACCTCGCTCCAATTATCCTCCTATCTCTCAACCCCAACATCATCCTGGGGT
TTACTTCCTGTAGATATAGTTTAACTAAAACATTAGACTGTGGATCTAACTATAGAGGCTTACCACCCCTTATTTACCGAGAAAACTCGCGAGGACTGCTAACCCATG
TATCCGTACCTAAAATTACGGTTTTCTCAACTTTTAAAGGATAACAGTCATCCATTGACCTTAGGAGTCAAAAATATTGGTGCAACTCCAAATAAAAGTAATAATCAT
GCACACCCCTATCATAATAACAACCCTTATCTCCCTGACTCTCCCGATTTTTGCCACCCTCATCAACCCTTACAAAAAACGTCCATACCCAGATTACGTAAAAACAAC
CGTAATATATGCTTTCATCATCAGCCTCCCCTCAACAACTTTATTCATCTTCTCAAACCAAGAAACAACCATTTGGAGCTGACATTGAATAATAACCCAAACACTAGA
CCTAACACTAAGCTTCAAATTAGATTACTTCTCCATAATATTCATTCCAATCGCACTATTCATCACCTGGTCCATCATAGAATTCTCACTATGATACATAAACTCAGAT
CCAAATATCAACCAATTCTTCAAATACCTCCTTATTTTCCTCACTGCCATACTAATTCTAGTCACTGCCAACAACCTCTTCCAATTCTTTATCGGCTGAGAGGGCGTAG
GAATCATATCTTTTCTCCTAATTAGCTGATGATACGCTCGAACAGACGCCAATACAGCAGCCATCCAGGCAATCTTATATAATCGCATCGGCGACATTGGTCTTATCC
TAGCTATAGCATGATTTCTCCTGCATTGTAACTCATGAGACCTCCAACAAATACTAGCTCTAAATCTCAACTCAAACCCCCTCCCACTAATAGGCCTCCTCCTAGCAG
CAGCAGGAAAATCAGCTCAATTCGGCCTTCACCCCTGACTACCCTCTGCCATAGAAGGCCCAACTCCAGTCTCAGCCCTACTTCACTCCAGTACCATAGTTGTTGCCG
GAGTGTTCTTACTCATCCGCTTCCACCCTCTAATAGAAACCAACACAATGATTCAAAATCTTACATTATGCCTGGGGGCTATTACTACCCTATTTACAGCCATCTGCG
CCCTCACACAAAACGACATTAAAAAAATTGTAGCCTTCTCCACCTCAAGCCAACTGGGCCTAATAATAATCACCATTGGTATCAACCAACCATACTTAGCATTCCTA
CACATCTGCACCCATGCTTTCTTTAAAGCCATACTTTTTATATGCTCCGGATCCATTATCCATAACCTAAACAACGAACAAGACATTCGAAAAATAGGAGGCCTATTC
AAAACAATACCTCTCACCTCAACTTCCATAATCATTGGCAACCTAGCACTTACAGGAATACCCTTCCTCACAGGCTTCTACTCCAAAGACCTTATCATCGAAGCCACA
AACACGTCGTACACCAACGCCTGAGCCCTATTCATTACTCTCATCACTACCTCTCTAACAAGCGCCTACAGCACTCGAACTATCCTCCTCACCCTAACAGGACAACCC
CGTTTTCCAGCTCTAACAAACATCAACGAAAACAACCCCGCCCTACTAAACCCAATTAAACGCCTCACAATAGGCAGTATAATCACAGGATTTCTTATCACCAACAA
CATCCCCCCCACTTCACTCCCCCAACCAGCAATACCCCTCTACTTAAAGCTCTCAGCCCTATGCGCAACTGCCCTAGGCTTCCTAGCAGCCCTAGACCTCACCCTTAT
AACAAACAAACTGAAAATAAAAAACCCATCACAAATATTCAAATTCTCCAATATACTAGGATATTTTCCCACCACAATTCACCGCACAATCCCCTATCAAAACCTAC
TCATAAGCCAAAATTTAGCCCTTCTCCTATTAGACTCGACATGACTAGAAAAATCTATACCTAAAATAATCTCACAAGCACACATCACTGCCTCCATAGCTGTAACC
ACCCAAAAAGGCATAATCAAACTCTATTCCCTTTCTTTCCTCATCCCCCTTACCCTGGCCCTAATTCTAATAATATAACCTATTACCCCGAACAATCTCAATCACAAC
ATACACACCAACAAACAACGTTCAACCAGCAACTACCACTAATCAACGCCCATAATCATATAAGGCACCCGCACCAATAGAATCACCTCGAATCAATCCTGACCCCT
CCCCCTCAAAAATTACCCAGTCACCGGTATCACTAAAACTAACCACCACTACTACCGCCTCATCTAAACTTAAAACCCATAAAACCAATATCACCTCCATCATCAAC
CCCACTAAAAGACCCCCTAAAACCTCAAACCCAGAAACCCCATGTCTCAGGATACTCTTCAATAGCCATCGCAGTAGTATAACCAAAAACAACCATCATACCTCCCAG
ATAAACTAAAAACATCATTAAACCTATATAAGCTCCCCCACAACTTAAAACAACCACACAACCAACCACACCCACTAACAATTAACGCCAGACCCCCATAAATAGGA
GAGGGCTTAGAAGAAAACCCCACAAACCCCATAACTAATAAAACACTCAATGCAAACAAAACATACGTCATCGCTTTCACATGGACTCCAACCATAACTAATGGTA
CGAAAAACCATCGTTGTACTTCAACTATAAAAGCACCAATGATTCCAATACGCAAATCCAACCCAATCATAAAAATAATTAACCGCTCCTTCATTGATTTACCCACC
CCACCCAACCTCTCCATGTGATGAAACTTTGGCTCACTTCTTGCAGCCTGCCTAGTTTTACAAATCGTCACAGGCCTACTCCTAGCAATACACTACTCACCAGACACC
TCCTCCGCCTTCTCCTCAATTGCACATATCACCCGAGATGTAAAGTACGGCTGAATCACTCGCTACCTCCACGCCAATGGTGCCTCTATACTCTTTATCTGCCTTTTCC
TACACATCGGTCGAGGCCTCTACTACGGCTCATATCTCCTCCTAGAAACCTGAAACATTGGCATTATACTCCTTCTTATAACCATAACAACAGCTTTTATGGGCTATG
TTCTCCCATGGGGCCAAATGTCATTCTGGGGGGCAACAGTAATTACAAACCTACTATCAGCAATCCCGTATATTGGAACCAATCTAGTCCAATGAATTTGAGGAGGA
TACGCCATCGACAGCCCCACTCTCACACGATTCTTCACCTTACACTTTATTCTACCCTTCATTATCATCGCCCTCACAGCTGTACACCTACTATTCCTACACGAAACAG
GATCAAACAACCCCTGCGGAATCTCCTCCGACTCGGACAAAATCGTCTTCCACCCCTACTACACAATCAAAGACATCCTAGGCCTAGTCCTCCTTCTCTTTATCCTAG
CAACACTAACACTACTCTCACCCAACTTCCTAAACGACCCAGACAACTACATCCCAGCCGACCCACTAAATACCCCCCCACATATCAAGCCAGAGTGGTACTTCCTA
TTTGCATACACAATTCTACGATCCGTCCCCAACAAACTGGGAGGCGTACTAGCACTCTTCCTATCGATCTTCATCCTAGCAGCCATCCCCATACTCCACAAATCCAAA
CAACAAAGCATAATATTCCGCCCACTCAGCCAATTCCTATTCTGACTCCTAATAACAGTCCTATTAACCCTTACCTGAGTTGGAAGTGAACCAGTAACTCAAACCCTT
ACCACTATCGGCCAAGTAGCATCCATAATATACTTCATCACAATTCTAATCCTAATACCACTGGCCTCCCTAATCGAAAACAACCTACTCAAATGAACTTGCCCTCGT
AGTATAAACTAATACACTGGTCTTGTAAACCAGAAATGAACACTCTTCCTAGGGCAATCAGAAAGAAAGTACCTAACTCTGCCACCAACACCCAAAGCTGGCATTCT
AACTTAAACTACTTTCTGCATTCTTATGTTGTATAAACTTCATAAAATAACCCTAACACCAACCTACCCATAATATTACTATGTAATTCGTGCATTACTGCTAGCCAAC
ATGTATAATATATAGTACTATATATGCTTAACTGTACATAACACATATCATTACATATCAACTTAACATCCTAGACAACATGCTTACAAGCAAGTATCCTAATACAAG
TCTCAACAGCAATACATATCATGGCCCTTTCAAACTTAACTTGTCCCCCCTCATGAATATCAACTGAACCAATTCATGCTAGTCGTCCATAGTACATTAAGTCGTTCA
TCGGACATAGTACATATTCATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTAGGAATCCCTTGTTCACCATCCTCCGTGAAATCAATATCCCGCACAAGA
GTGCTACTCTCCTCGCTCCGGGCCCATAACTCGTGGGGGTAGCTATACCTGAACTGTATCCGGCATCTGGTTCTTACCTCAGGGCCATAACAATCAAGATCGCCCACA
CGTTCCCCTTAAATAAGACATCTCGATGGATCACGGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATCTCTGGTCTGCACGCA
ACCCCATTGCAGAATGCTGACTCCCACCACATCTCGTCCTGTATGCGCCTGTCTTTGATTCCTAGTACATGCAGTTGTTGATCGCACCTACGTTCAATATTCTAGCTCC
ACGCAAACCTTAGCAAGGTGTTATTTAATTCATGCTTGTAGGACATATTAATAACCATTCCAACCAATATCACCCACACCACTAACCACAACCGTG