

CẤU TRÚC VÀ ĐẶC ĐIỂM GEN CỦA ĐƠN VỊ SAO CHÉP RIBOSOME CỦA SÁN LÁ GAN NHỎ *OPISTHORCHIS VIVERRINI*, *O. FELINEUS* VÀ *CLONORCHIS SINENSIS*

Lê Thanh Hoà^{1,2,✉}, Nguyễn Thị Bích Nga¹, Đoàn Thị Thanh Hương^{1,2}, Lê Thị Kim Xuyên¹, Nguyễn Thị Khuê¹

¹Viện Công nghệ sinh học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

²Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

✉Người chịu trách nhiệm liên lạc. E-mail: imibtvn@gmail.com

Ngày nhận bài: 08.01.2019

Ngày nhận đăng: 25.03.2019

TÓM TẮT

Opisthorchiasis là bệnh ký sinh trùng do các loài sán lá gan nhỏ *Opisthorchis viverrini*, *O. felineus* và *Clonorchis sinensis* thuộc họ Opisthorchiidae gây nên. Việt Nam có phân bố của hai loài, *C. sinensis* ở các tỉnh phía Bắc và *O. viverrini* ở các tỉnh miền Trung. Bên cạnh hệ gen ty thể, trình tự DNA của ribosome (rDNA) của các loài này rất cần thu nhận để cung cấp các chỉ thị phân tử trong nghiên cứu xác định loài, phân loại, phá hệ và tiến hoá. Trong nghiên cứu này, chuỗi nucleotide hoàn chỉnh của đơn vị sao chép ribosome (rTU) từ các loài *O. viverrini* (mẫu Việt Nam), *O. felineus* (mẫu Nga) và *C. sinensis* (mẫu Việt Nam) đã được phân tích cấu trúc và đặc điểm gen. Các rTU mỗi loài có độ dài được xác định là 7.839 bp cho *O. viverrini*, 6.948 bp cho *O. felineus* và 7.296 bp cho *C. sinensis*, bao gồm các cấu trúc 18S, ITS1, 5.8S, ITS2 và 28S. Riêng vùng IGS chưa thu được hoàn toàn cho cả 3 loài. Trong vùng giao gen ITS1 của cả ba loài đều có 2 chuỗi cấu trúc lặp liên kế (tandem repeat) kích thước 47-48 bp/chuỗi, nhưng không có trong ITS2. Chuỗi nucleotide của 18S, ITS1, ITS2 và 28S là các chỉ thị phân tử có giá trị mà nghiên cứu này cung cấp sử dụng cho chẩn đoán, giám định, phân loại và di truyền quần thể. Như vậy, trình tự rTU của 3 loài trong họ Opisthorchiidae đã được xác định và cung cấp chỉ thị phân tử cho việc sử dụng nghiên cứu phân tích phá hệ để giải quyết tình trạng phân loại loài/họ trong phân lớp Opisthorchioidea và lớp sán lá Trematoda.

Từ khoá: Cấu trúc; *Clonorchis sinensis*; Đơn vị sao chép ribosome (rTU); *Opisthorchis viverrini*; *Opisthorchis felineus*; phá hệ; rDNA.

ĐẶT VẤN ĐỀ

Họ Opisthorchiidae bao gồm các thành viên sán lá gan nhỏ ký sinh ở túi mật và gan, chủ yếu là các loài *Clonorchis sinensis*, *Opisthorchis viverrini*, *O. felineus*, ký sinh và gây bệnh ở người (Sripa *et al.*, 2012; Petney *et al.*, 2013). *O. felineus* phân bố kéo dài từ Viễn đông (Liên bang Nga) đến châu Âu, đó là phần địa lý châu Á và châu Âu của Nga, Kazakstan, Ukraine, Belarussia, một số nước Đông Âu, Đức, Italy và Hy Lạp (Fedorova *et al.*, 2018). *C. sinensis* phân bố ở Trung Quốc, Đài Loan, Hàn Quốc, Nhật Bản, Việt Nam và phía Đông nước Nga. *O. viverrini* phân bố ở Đông Nam Châu Á, gây bệnh trên người ở Thái Lan, Lào, Malaysia, Việt Nam, Campuchia và Trung Quốc. *O. viverrini* và *C. sinensis* được coi là các yếu tố tham gia tích cực trong cơ chế tiến triển gây ung thư biểu mô túi mật tiên phát

(cholangiocarcinoma) ở người, với gần như 100% tử vong (Sripa *et al.*, 2012; Kim *et al.*, 2016). Việt Nam có phân bố của cả hai loài, trong đó *C. sinensis* ở các tỉnh phía Bắc, đến Nghệ An; *O. viverrini* phát hiện ở các tỉnh miền Trung từ Thừa Thiên Huế vào đến Nam Trung Bộ (Le *et al.*, 2006; 2012; Doanh, Nawa, 2016).

Một trong những dữ liệu gen học quan trọng cần xác lập, khám phá và khai thác là DNA ribosome (ribosomal DNA, rDNA) hay còn gọi là đơn vị sao chép ribosome (ribosomal transcription unit, rTU) có trong hệ gen nhân của tế bào (Blair, 2006). Chuỗi gen của rTU được tận dụng làm chỉ thị phân tử có giá trị trong giám định, phân biệt loài và xác định dòng bố (paternity) đối với các loài độc lập, hoặc loài lai ngoại loài (hybrid hoặc introgressive hybridization) hoặc lai chéo trong quần thể (Blair,

2006; Blair *et al.*, 2016; Nguyen *et al.*, 2018). Trong hệ gen nhân ở tế bào động vật, rDNA sắp xếp hàng trăm đơn vị nối tiếp nhau thành dãy. Mỗi một đơn vị, ở ký sinh trùng, có độ dài khoảng 7 - 9 kb, chứa 3 gen mã hóa (18S, 5.8S và 28S) tách biệt nhau bởi 2 vùng giao gen, lần lượt là ITS-1 và ITS-2 (internal transcribed spacer 1 and 2). Các đơn vị được nối với nhau bằng một chuỗi nucleotide, chứa nhiều cấu trúc lặp, gọi là vùng bản lề IGS (non-transcribed intergenic spacer). Khung gen đặc trưng có sắp xếp là IGS-ETS-18S-ITS1-5.8S-ITS2-28S-IGS, tạo nên các cấu trúc nối tiếp nhau (Blair, 2006). Bất kỳ gen ribosome nào (18S, 28S) hay vùng giao gen (ITS-1, ITS-2) đều được sử dụng trong phân tích phân loại và quan hệ tiến hóa phân tử và truy xuất nguồn gốc phả hệ của các loài (Weider *et al.*, 2005; Blair, 2006).

Đến nay, chưa có bất kỳ dữ liệu gen của toàn bộ đơn vị sao chép gen (rDNA/rTU) của các loài này, kể cả từ mẫu Việt Nam và mẫu thể giới, mặc dù một số vùng gen ribosome riêng biệt (18S, 28S hoặc vùng giao gen ITS) đã được thu nhận và sử dụng trong nghiên cứu chẩn đoán và dịch tễ học (Thaenkham *et al.*, 2011; Tkach *et al.*, 2016; Dao *et al.*, 2017; Le *et al.*, 2017). Trong khuôn khổ bài báo này, chúng tôi giới thiệu quá trình thu nhận toàn bộ rDNA của 3 loài thuộc họ Opisthorchiidae, giải trình tự và phân tích cấu trúc, sắp xếp, đặc điểm từng gen và vùng giao gen ở mỗi loài.

NGUYÊN VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Mẫu nghiên cứu thuộc họ Opisthorchiidae

Mẫu nghiên cứu là sán lá gan nhỏ trưởng thành do các đối tác trong nước (Viện Sốt rét-Ký sinh trùng và Côn trùng trung ương); và mẫu của loài *O. felineus* được các đối tác quốc tế cung cấp. Các mẫu đã được phân biệt loài bằng phương pháp hình thái học và thẩm định bằng sinh học phân tử. Mẫu ở dạng tươi đông lạnh hoặc trong cồn 70% hoặc được cung cấp DNA tổng số, bảo quản mẫu ở -20°C .

Các mẫu bao gồm: *C. sinensis* (mẫu Việt Nam, ký hiệu chúng là NH, danh pháp là Csin-CsNH-Vietnam), *O. viverrini* (mẫu Việt Nam, ký hiệu chúng là PY2, danh pháp là Oviv-OvPY2-Vietnam) và *O. felineus* (mẫu của CHLB Nga, ký hiệu chúng là UstTula, danh pháp là Ofel-UstTula-Russia).

Tách chiết DNA tổng số

DNA tổng số của từng mẫu riêng biệt được tách

chiết bằng bộ sinh phẩm QIAamp DNA Mini kit (QIAGEN Inc, Mỹ) hoặc bộ sinh phẩm GeneJET™ Genomic DNA Purification Kit (Thermo Scientific Inc., MA, USA) theo hướng dẫn của nhà sản xuất. DNA tổng số được thu nhận với dung tích 100 μL /mẫu sán, kiểm tra nồng độ bằng cách đo OD (optical density) trên máy đo quang phổ kế Nanodrop 1000. DNA tổng số được pha thành nồng độ 50 ng/ μL , sử dụng 2 μL cho mỗi phản ứng PCR có dung tích 50 μL .

Môi sử dụng để thu nhận toàn bộ rDNA

Toàn phần mỗi đơn vị rDNA (5' 18S-ITS1-5.8S-ITS2-28S-IGS 3') có độ dài khoảng từ 7-8 cho đến 9-10 kb, tùy thuộc mỗi loài sán lá (trematode). Môi được thiết kế dựa trên một số công bố về rDNA đã có trên Ngân hàng gen và dữ liệu của chúng tôi (Le *et al.*, 2017). PCR được áp dụng với phối hợp các cặp mồi khác nhau để thu các phần DNA (phân đoạn) dài ngắn khác nhau (Bảng 1).

Thực hiện PCR, kiểm tra sản phẩm và giải trình tự

Phản ứng PCR có tổng thể tích 50 μL , gồm: 25 μL PCR mastermix (Thermo Fisher Scientific, MA, USA), 2 μL mỗi loại mồi (10 pmol/ μL), 3 μL khuôn DNA, 2 μL DMSO (dimethyl sulfoxide) và 16 μL nước khử ion DEPC, được thực hiện trên máy MJ PTC-100 (USA). Chu trình nhiệt gồm: 1 chu kỳ ở $95^{\circ}\text{C}/5$ phút, 35 chu kỳ ở [$94^{\circ}\text{C}/1$ phút, $52^{\circ}\text{C}/3$ phút; $72^{\circ}\text{C}/2$ phút], chu kỳ cuối ở $72^{\circ}\text{C}/10$ phút. Sản phẩm PCR được điện di kiểm tra trên thạch agarose 1%, với chỉ thị DNA λ (HindIII). Sản phẩm PCR đơn băng được tinh sạch bằng GeneJET PCR Purification Kit hoặc phân lập băng DNA đúng kích thước dự tính (nếu nhiều băng) bằng bộ kit GeneJET Gel Extraction Kit (Thermo Fisher Scientific Inc.) và gửi đi giải trình tự tại Macrogen (Hàn Quốc).

Xử lý số liệu, xác định đặc điểm rDNA của các loài

Phần lớn sản phẩm PCR được giải trình tự trực tiếp. Chuỗi nucleotide của từng phân đoạn được nối liên tiếp với nhau để thu nhận hoàn toàn trình tự rDNA của mỗi loài. Sử dụng chuỗi rDNA để phân tích cấu trúc, sắp xếp gen, vùng biên, nội gen và vùng giao gen. Các chương trình ChromasPro, BioEdit, GENEDOC2.7 đã được sử dụng để thu nhận, xử lý, phân tích các chuỗi nucleotide, đặc điểm các gen ribosome (18S, 5.8S, 28S) và vùng giao gen (internal transcribed spacer, ITS) gồm ITS1 và ITS2

và vùng bản lề IGS. Độ dài từng gen được xác định dựa vào các công bố trước đây (Zheng *et al.*, 2014; Briscoe *et al.*, 2016; Le *et al.*, 2017; Dao *et al.*,

2017). Các cấu trúc chuỗi nucleotide lặp được xác định bằng phần mềm Tandem Repeats Finder (Benson, 1999).

Bảng 1. Mồi sử dụng cho PCR và giải trình tự đơn vị sao chép gen các loài trong họ Opisthorchiidae.

Tên mồi	Chuỗi nucleotide (5'–3')	Độ dài (bp)	Nhiệt độ (°C)	Gen/Vùng đích	Tài liệu tham khảo
UD18SF	AACCTGGTTGATCCTGCCAG	20	59	18S (F)	Olson <i>et al.</i> , 2003
NS1F	GTAGTCATATGCTTGCTC	19	48	18S (F)	Le <i>et al.</i> , 2017
U18SF	GCGAATGGCTCATTAATCAGC	22	57	18S (F)	Le <i>et al.</i> , 2017
U18S2R*	GGTTCTGTTCTAATAATCCAC	22	50	18S (R)	Le <i>et al.</i> , 2017
U18SAR*	CCGTCGCCGACACAAGGCCGAC	22	67	18S (R)	Le <i>et al.</i> , 2017
NS2F*	GCAAGTCTGGTGCCAGCAGCC	21	66	18S (F)	Le <i>et al.</i> , 2017
NS2R*	GGCCTGCTTTGAGCACTC	18	59	18S (R)	Le <i>et al.</i> , 2017
U18S2F	TCGTGACTGGGATCGGGGC	19	64	18S (F)	Le <i>et al.</i> , 2017
NS5F*	TGAATGGTTTAGCAAGTCCTCGG	24	61	18S (F)	Le <i>et al.</i> , 2017
U18SR*	GGAACCAATCCGAGGACCTTGC	22	63	18S (R)	Le <i>et al.</i> , 2017
NS8R*	CACCTACGGAAACCTTGTTACGACTT	26	60	18S (R)	Le <i>et al.</i> , 2017
U3SF	GGTACCGGTGGATCACTCGGCTCGTG	26	67	5.8S (F)	Le <i>et al.</i> , 2017
U3SR	CGACCCTCGGACAGGCG	17	64	5.8S (R)	Le <i>et al.</i> , 2017
U28SF*	CTAACAAGGATCCCTTAGTAAC	23	52	28S (F)	Le <i>et al.</i> , 2017
U28S2R*	ACAACCCGACTCCAAGGTC	19	59	28S (R)	Le <i>et al.</i> , 2017
U28F*	TCGGAGACGGCGGCTTG	17	63	28S (F)	Le <i>et al.</i> , 2017
U28S2F*	ATCACCGGCCCGTCCCATG	19	65	28S (F)	Le <i>et al.</i> , 2017
U28SR	GTCTTTCGCCCTATACTCAC	21	57	28S (R)	Le <i>et al.</i> , 2017
1500R	GCTATCCTGAGGGAAACTTCG	21	57	28S (R)	Olson <i>et al.</i> , 2003
U28SF1	GTGTAACAACCTCACCTGCCGAATC	24	58	28S (F)	Nghiên cứu này
U28SF2	AGGCCGTACCCATATCCGC	19	56	28S (F)	Nghiên cứu này
U28SF3	GACCCGAAAGATGGTGAACATATGC	24	57	28S (F)	Nghiên cứu này
28ENDR	TTCTGACTTAGAGGCGTTCAGTC	23	55	28S (R)	Nghiên cứu này
28ENDF	TTCTGACTTAGAGGCGTTCAGTC	23	56	28S (F)	Nghiên cứu này
U28TUF	CGACGTCGCTTTTTGATCCTTCG	23	61	28S (F)	Nghiên cứu này
U18TUR	CGGGTCAGGGCATAGTGGC	19	63	18S (R)	Nghiên cứu này

Viết tắt: F, xuôi; R, ngược; Tm: Nhiệt độ nóng chảy của mồi. *Các mồi chủ yếu dùng giải trình tự.

KẾT QUẢ NGHIÊN CỨU

Thực hiện PCR thu nhận toàn bộ DNA ribosome

Với các cặp mồi khác nhau, chúng tôi thực hiện PCR thu nhận các phân đoạn có độ dài khoảng 3 - 5 kb từ DNA khuôn của mỗi loài. Một số sản phẩm đại diện được trình bày ở Hình 1. Kết quả điện di kiểm tra cho thấy, sản phẩm PCR thu được của từng phân đoạn là các băng đơn nhất, sáng rõ, chứng tỏ các cặp mồi đã được thiết kế đặc hiệu và chu trình nhiệt sử dụng trong nghiên cứu là phù hợp.

Phân tích cấu trúc đơn vị sao chép ribosome của các loài *C. sinensis* (Việt Nam), *O. viverrini* (Việt Nam) và *O. felineus* (Nga)

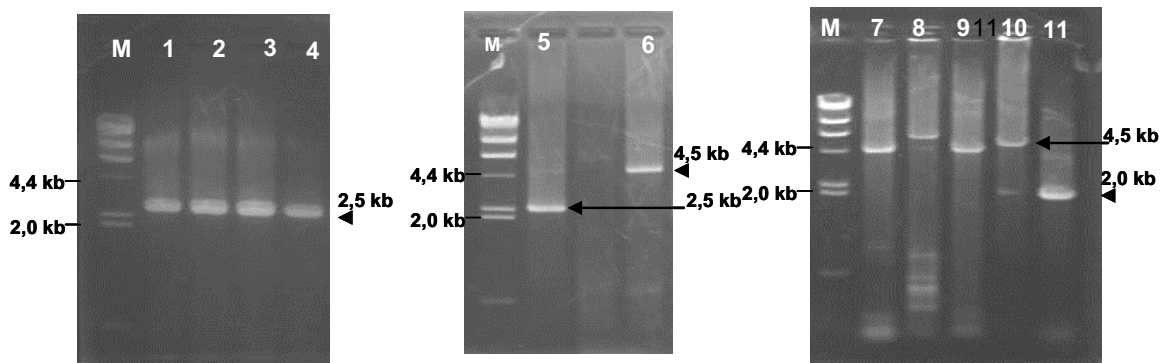
Các mồi (Bảng 1) được sử dụng để thực hiện PCR thu nhận sản phẩm có độ dài khác nhau của rDNA từng loài và giải trình tự. Sau khi xử lý chuỗi, chúng tôi thu được: i) Chuỗi nucleotide rDNA toàn phần của loài *O. viverrini* (mẫu OvPY2, Việt Nam) có độ dài là 7.839 bp; chứa toàn phần vùng IGS; ii) Chuỗi nucleotide rDNA toàn phần của loài *C. sinensis* (mẫu CsNH, Việt Nam) có độ dài là 7.601 bp; chứa một

phần vùng IGS; Chuỗi nucleotide rDNA của loài *O. felineus* (mẫu UstTula, Nga) có độ dài là 7.261 bp, thiếu vùng IGS (Bảng 2).

Trong cấu trúc của rDNA của từng loài, 6 phân đoạn DNA của 6 gen/vùng gen đã được xác định bao gồm: 18S, ITS1, 5.8S, ITS2, 28S, IGS. Đơn vị rDNA của 3 loài trong đó loài *O. viverrini* đã được xác định toàn bộ chuỗi nucleotide chứa toàn phần vùng IGS; của loài *C. sinensis* chỉ có một phần IGS được giải trình tự; còn của loài *O. felineus* còn thiếu tất cả vùng IGS này.

Độ dài gen và vùng giao gen của các loài *O. viverrini*, *O. felineus* và *C. sinensis*

Dựa trên một số công bố trước đây về rDNA của một số loài sán dẹt (Briscoe *et al.*, 2016; Le *et al.*, 2017) và dữ liệu rDNA đã được phân tích công bố trong Ngân hàng gen, chúng tôi đã xác định được kích thước và vị trí của từng gen/vùng giao gen trong rTU của từng loài bao gồm gen 18S, vùng giao gen ITS1, gen 5.8S, vùng giao gen ITS2, gen 28S và vùng ngoại biên IGS (Bảng 2).



Hình 1. Điện di kiểm tra một số sản phẩm PCR đại diện từ các mẫu nghiên cứu trên agarose 1%. M: chỉ thị DNA λ(*Hind*III) 1, 4, 7, 9: *C. sinensis*; 2, 5, 8: *O. felineus*; 3, 6, 10, 11: *O. viverrini*.

Bảng 2 cho thấy, các gen ribosome của cả 3 loài đều có độ dài hoàn toàn giống nhau, cụ thể: 1.991 bp ở gen 18S, 157 bp ở gen 5.8S và 4.190 bp ở gen 28S.

Tuy nhiên, độ dài vùng giao gen ITS ở mỗi loài là khác nhau:

- Vùng giao gen ITS1 ở *O. viverrini* là 643 bp, ở *O. felineus* là 632 bp và ở *C. sinensis* là 654 bp. Vùng giao gen ITS1 của *Opisthorchis* spp. (*O. viverrini* và *O. felineus*) chứa 2 cấu trúc lặp liên kề (tandem repeat) có độ dài 48 bp và của loài *C. sinensis* chứa 47 bp của mỗi một cấu trúc lặp (Bảng 2).

- Vùng giao gen ITS2 ở 2 loài, *O. viverrini* và *O. felineus* thuộc chi *Opisthorchis* đều có độ dài 294 bp; ở loài *C. sinensis* (chi *Clonorchis*) là 300 bp. Tất cả chuỗi nucleotide ITS2 ở cả 3 loài đều không có cấu trúc lặp.

- Vùng IGS nằm ở đầu cuối 3' của rDNA đều chưa xác định chính xác, ở loài *O. viverrini* là khoảng 564 bp, ở loài *C. sinensis* là khoảng 309 bp, mới được một phần, nhưng chưa phải là chuỗi cuối cùng, do sản phẩm PCR của vùng này có nhiều độ dài khác nhau. Vùng IGS của loài *O. felineus* chưa

thu nhận được hoàn toàn.

BÀN LUẬN

Trong nghiên cứu này, cấu trúc, sắp xếp gen và vùng giao gen trong đơn vị sao chép gen rDNA hay rTU của 3 loài *O. viverrini*, *O. felineus* và *C. sinensis*, là các loài sán lá gan nhỏ gây bệnh trên người, đã được trình bày. Chuỗi nucleotide đã thu nhận có độ dài bao phủ gần hết rDNA và các gen/vùng gen chính đã được xác định. Độ dài gen 18S là 1.991 bp và 28S là 4.190 bp, khoảng 4,1 – 4,5 kb như công bố ở các loài sán dẹt khác (Blair *et al.*, 2006; Zheng *et al.*, 2014; Briscoe *et al.*, 2016; Le *et al.*, 2017). Gen 5.8S và các vùng giao gen (ITS1 và ITS2), cũng như một phần IGS cũng đã được xác định. Chuỗi nucleotide của các gen ribosome chính, 18S và 28S được sử dụng trong so sánh chuỗi thực hiện nghiên cứu mối quan hệ về loài trong một họ và trong một lớp (class) phân định rõ ràng và tin tưởng vị trí phân loại của các loài với nhau (Olson *et al.*, 2003; Lokyer *et al.*, 2003; Thaenklam *et al.*, 2011; Dao *et al.*, 2017; Le *et al.*, 2017). Việc sử dụng đơn phương chuỗi gen ribosome hoặc cùng kết hợp với phân tích gen ty thể đã giúp việc phân loại, chẩn

đoán, nghiên cứu phả hệ và tiến hóa, cũng như dịch tễ học phân tử và di truyền quần thể của các loài nói chung và ký sinh trùng nói riêng, chính xác hơn, rõ ràng hơn, đặc biệt là những loài mới hoặc loài lai.

Bảng 2. Vị trí của gen ribosome và vùng giao gen trong đơn vị sao chép gen (rTU) của các loài *O. viverrini* (7.839 bp), *C. sinensis* (7.601 bp) và *O. felineus* (7.261 bp).

Gen/ Vùng giao gen	Vị trí (5'→3')	Cấu trúc lặp	Độ dài (bp)	Giao gen (bp)	Ghi chú
<i>O. viverrini</i> (7.839 bp)					
					GenBank: (đang đăng kí)
18S	1–1991		1991	0	Gen ribosome
ITS1	1992–2634		643	+65	65 bp đến ITS1-RP1
	2057–2104	ITS1-RP1	48	+2	Cấu trúc lặp
	2107–2154	ITS1-RP2	48	+480	Cấu trúc lặp
5.8S	2635–2791		157	0	Gen ribosome
ITS2	2792–3085		294	0	Không có lặp
28S	3086–7275		4190	0	Gen ribosome
IGS	7276–7839		564		Kết thúc
<i>O. felineus</i> (7.261 bp)					
					GenBank: (đang đăng kí)
18S	1–1991		1991	0	Gen ribosome
ITS1	1992–2623		632	+65	65 bp đến ITS1-RP1
	2057–2104	ITS1-RP1	48	+2	Cấu trúc lặp
	2107–2154	ITS1-RP2	48	+469	Cấu trúc lặp
5.8S	2624–2780		157	0	Gen ribosome
ITS2	2781–3071		294	0	Không có lặp
28S	3072–7261		4190	0	Gen ribosome
IGS	---/---				Còn thiếu
<i>C. sinensis</i> (7.601 bp)					
					GenBank: (đang đăng kí)
18S	1–1991		1991	0	Gen ribosome
ITS1	1992–2645		654	+66	66 bp đến ITS1-RP1
	2058–2104	ITS1-RP1	47	+3	Cấu trúc lặp
	2108–2154	ITS1-RP2	47	+491	Cấu trúc lặp
5.8S	2646–2802		157	0	Gen ribosome
ITS2	2803–3102		300	0	Không có lặp
28S	3103–7292		4190	0	Gen ribosome
IGS	7293–7601//		309		Một phần

RP: repeat (chuỗi lặp)

Có hai cấu trúc lặp trong ITS1 (mỗi đơn vị là 47 bp hoặc 48 bp) ở cả 3 loài *O. viverrini*, *O. felineus* và *C. sinensis*, đã được tìm thấy. Điều này cho thấy, cũng như một số loài ở các họ khác, vùng giao gen ITS của sán dẹt có chứa trình tự nucleotide có tính

chất đa hình nội loài và các cấu trúc lặp có số lượng và độ dài khác nhau, như ở các họ Schistosomatidae, Opisthorchiidae, Heterophyidae, Paramphistomatidae và nhiều họ khác (Lockyer *et al.*, 2003; Van *et al.*, 2009; Tatonova *et al.*, 2012;

Zheng *et al.*, 2014; Le *et al.*, 2017). Điều này khuyến cáo về sự thận trọng khi sử dụng chuỗi ITS để xác định nội và ngoại loài, vì chúng thực sự không thích hợp khi phân tích phả hệ ở mức độ tiến hóa cao, do sự có mặt của các cấu trúc lặp làm thay đổi so sánh và độ tin tưởng (Blair *et al.*, 2006; Tatonova *et al.*, 2012; Le *et al.*, 2017). Tuy nhiên, chuỗi gen 18S và 28S (phần đầu 5') là có thể xem xét phân tích phả hệ, do trong chuỗi 18S và 28S không có cấu trúc lặp và chỉ có các sai khác có tính chất chỉ thị phân từ phân biệt loài (Olson *et al.*, 2003; Blair *et al.*, 2006; Thaenkham *et al.*, 2011; Le *et al.*, 2017; Dao *et al.*, 2017).

Mặc dù vậy, chuỗi nucleotide của các vùng giao gen ITS1 và ITS2 đã có những đóng góp nhất định trong phân loại làm sáng tỏ di truyền quần thể bên cạnh kết hợp phân tích hình thái học và nuôi cấy (Weider *et al.*, 2005; Blair, 2006; Tatonova *et al.*, 2012; Blair *et al.*, 2016).

Chỉ thị phân từ có nguồn gốc từ hệ gen ty thể và nhân tế bào, trong đó có các chuỗi gen và vùng giao gen của đơn vị sao chép gen đang được thu nhận ngày càng nhiều và đang được sử dụng ngày càng rộng rãi và tin tưởng, kể cả ở động vật bậc cao và thấp. Đối với ngành ký sinh trùng, do tính cấp thiết của kết hợp xác định hình thái học và phân từ, có rất nhiều nghiên cứu xác lập dữ liệu và tận dụng chỉ thị phân từ để có những nghiên cứu tiến hóa và phân loại chính xác.

KẾT LUẬN

Đơn vị sao chép ribosome của các loài *C. sinensis* (chủng Csin-NH-VN, Việt Nam), *O. viverrini* (chủng Oviv-PY2-VN, Việt Nam) và *O. felineus* (Ofel-UstTula-RU, Nga) đã được thu nhận, xác định cấu trúc, kích thước gen/vùng gen. Độ dài của từng gen (18S, 5.8S, 28S) ở cả 3 loài là như nhau; tuy nhiên vùng giao gen (ITS1 và ITS2) của các loài *O. viverrini*, *O. felineus* và *C. sinensis* là khác nhau; trong đó ITS1 của tất cả 3 loài đều chứa 2 cấu trúc lặp (47-48 bp/mỗi loại) và ITS2 không có cấu trúc lặp. Tính đa hình trong ITS1 do có chứa số lượng và kích thước các cấu trúc lặp khác nhau, khuyến cáo cần thận trọng khi sử dụng phân tích phả hệ ở mức độ tiến hóa cao do chúng có thể làm thay đổi so sánh chuỗi và độ tin tưởng quan hệ loài. Xác định và phân tích chuỗi nucleotide DNA ribosome ở 3 loài sán lá gan nhỏ và các loài ký sinh trùng khác là cần thiết cho nghiên cứu chẩn đoán, phả hệ, phân loại và dịch tễ bệnh.

Lời cảm ơn: Nghiên cứu này được tài trợ bởi Quỹ Phát triển khoa học và công nghệ Quốc gia (NAFOSTED) trong đề tài mã số 108.02-2017.09.

TÀI LIỆU THAM KHẢO

- Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27: 573–580.
- Blair D (2006) Ribosomal DNA variation in parasitic flatworms. In: Maule A, editor. *Parasitic Flatworms: Molecular Biology, Biochemistry, Immunology and Control*. CAB International pp. 96–123.
- Blair D, Nawa Y, Mitreva M, Doanh PN (2016) Gene diversity and genetic variation in lung flukes (genus *Paragonimus*). *Trans R Soc Trop Med Hyg* 110(1): 6–12.
- Briscoe AG, Bray RA, Brabec J, Littlewood DTJ (2016) The mitochondrial genome and ribosomal operon of *Brachycladium goliath* (Digenea: Brachycladiidae) recovered from a stranded minke whale. *Parasitol Int* 65(3): 271–275.
- Dao TTH, Nguyen TTG, Gabriël S, Bui KL, Dorny P, Le TH (2017) Updated molecular phylogenetic data for *Opisthorchis* spp. (Trematoda: Opisthorchioidea) from ducks in Vietnam. *Parasit Vectors* 10(1): 575.
- Doanh PN, Nawa Y (2016) *Clonorchis sinensis* and *Opisthorchis* spp. in Vietnam: current status and prospects. *Trans R Soc Trop Med Hyg* 110(1): 13–20.
- Fedorova OS, Fedotova MM, Sokolova TS, Golovach EA, Kovshirina YV, Ageeva TS, Kovshirina AE, Kobayakova OS, Ogorodova LM, Odermatt P (2018) *Opisthorchis felinus* infection prevalence in Western Siberia: A review of Russian literature. *Acta Trop* 178: 196–204.
- Kim TS, Pak JH, Kim JB, Bahk YY (2016) *Clonorchis sinensis*, an oriental liver fluke, as a human biological agent of cholangiocarcinoma: a brief review. *BMB Rep* 49(11): 590–597. Review.
- Le TH, De NV, Blair D, Sithithaworn P and McManus DP (2006) *Clonorchis sinensis* and *Opisthorchis viverrini*: development of a mitochondrial-based multiplex PCR for their identification and discrimination. *Exp Parasitol* 112(2): 109–114.
- Le TH, Nguyen KT, Nguyen NT, Doan HT, Dung DT, Blair D (2017) The ribosomal transcription units of *Haplorchis pumilio* and *H. taichui* and the use of 28S sequences for phylogenetic identification of common heterophyids in Vietnam. *Parasit Vectors* 10: 17.
- Lockyer AE, Olson PD, Ostergaard P, Rollinson D, Johnston DA, Attwood SW, et al. (2003) The phylogeny of the Schistosomatidae based on three genes with emphasis on the interrelationships of *Schistosoma* Weinland, 1858. *Parasitology* 126(3): 203–224.

- Nguyen TBN, De NV, Nguyen TKL, Quang HH, Doan HTT, Agatsuma T, Le TH (2018) Distribution status of hybrid types in large liver flukes, *Fasciola* species (Digenea: Fasciolidae), from ruminants and humans in Vietnam. *Korean J Parasitol* 56(5): 453–461.
- Olson PD, Cribb TH, Tkach VV, Bray RA, Littlewood DTJ (2003) Phylogeny and classification of the Digenea (Platyhelminthes: Trematoda). *Int J Parasitol* 33: 733–755.
- Petney TN, Andrews RH, Saijuntha W, Wenz-Mücke A, Sithithaworn P (2013) The zoonotic, fish-borne liver flukes *Clonorchis sinensis*, *Opisthorchis felineus* and *Opisthorchis viverrini*. *Int J Parasitol* 43(12–13): 1031–1046.
- Sripa B, Brindley PJ, Mulvenna J, Laha T, Smout MJ, Mairiang E, Bethony JM, Loukas A (2012) The tumorigenic liver fluke *Opisthorchis viverrini*—multiple pathways to cancer. *Trends Parasitol* 28(10): 395–407.
- Tatonova YV, Chelomina GN, Besprosvannykh VV (2012) Genetic diversity of nuclear ITS1-5.8S-ITS2 rDNA sequence in *Clonorchis sinensis* Cobbold, 1875 (Trematoda: Opisthorchidae) from the Russian Far East. *Parasitol Int* 61(4): 664–674.
- Thaenkham U, Nawa Y, Blair D, Pakdee W (2011). Confirmation of the paraphyletic relationship between families Opisthorchiidae and Heterophyidae using small and large subunit ribosomal DNA sequences. *Parasitol Int* 60(4): 521–523.
- Tkach VV, Kudlai O, Kostadinova A (2016) Molecular phylogeny and systematics of the Echinostomatoidea Looss, 1899 (Platyhelminthes: Digenea). *Int J Parasitol* 46: 171–185.
- Van KV, Dalsgaard A, Blair D, Le TH (2009) *Haplorchis pumilio* and *H. taichui* in Vietnam discriminated using ITS-2 DNA sequence data from adults and larvae. *Exp Parasitol* 123: 146–151.
- Weider LJ, Elser JJ, Crease TJ, Mateos M, Cotner JB, Markow TA (2005) The functional significance of ribosomal rDNA variation: Impacts on the evolutionary ecology of organisms. *Annu Rev Ecol Syst* 36: 219–242.
- Zheng X, Chang QC, Zhang Y, Tian SQ, Lou Y, Duan H, Guo HG, Wang CR, Zhu XQ (2014) Characterization of the complete nuclear ribosomal DNA sequences of *Paramphistomum cervi*. *Scientific World J* 2014: 751907.

STRUCTURE AND CHARACTERIZATION OF THE RIBOSOMAL TRANSCRIPTION UNITS OF SMALL LIVER FLUKES, *OPISTHORCHIS VIVERRINI*, *O. FELINEUS* AND *CLONORCHIS SINENSIS*

Le Thanh Hoa^{1,2}, Nguyen Thi Bich Nga¹, Doan Thi Thanh Huong^{1,2}, Le Thi Kim Xuyen¹, Nguyen Thi Khue¹

¹Institute of Biotechnology, Vietnam Academy of Science and Technology

²Graduate University of Science and Technology, Vietnam Academy of Science and Technology

SUMMARY

Opisthorchiasis is a zoonotic parasitic infection caused by small liver fluke species, *Opisthorchis viverrini*, *O. felineus* and *Clonorchis sinensis*, in the family Opisthorchiidae. Vietnam has both species, of which *C. sinensis* is distributed in the northern and *O. viverrini* in the central provinces. In addition to the mitochondrial genomes, the ribosomal DNA sequences (rDNA) of these species are highly needed to obtain for providing molecular markers in species identification, classification, phylogeny and evolutionary studies. In this study, the near/complete nucleotide sequences of ribosomal transcription units (rTU) from *O. viverrini* (Vietnamese sample), *O. felineus* (Russian sample) and *C. sinensis* (Vietnamese sample) were analyzed. All rTUs for three species were determined, which is 7,839 bp for *O. viverrini*, 6,948 bp for *O. felineus* and 7,296 bp for *C. sinensis* containing structures of 18S, ITS1, 5.8S, ITS2 and 28S. The IGS region was not obtained for all three species. In all three species, sequence analysis revealed 2 tandem repetitive elements of 47–48 bp/each in ITS1 but not in ITS2. The nucleotide sequences of 18S, ITS1, ITS2 and 28S are valuable ribosomal markers that this study provides for diagnosis, identification, taxonomic classification and population genetics. In conclusion, the rTU sequences for the three species of the family Opisthorchiidae have been identified and provides molecular markers for the use of phylogenetic analysis for species/family classification in the superfamily Opisthorchioidea and the class Trematoda.

Keywords: Ribosomal transcription unit (rTU); *Clonorchis sinensis*; *Opisthorchis viverrini*; *O. felineus*; rDNA sequence; phylogeny