DNA Barcoding for Molecular Prospecting of Platyhelminthes

Logan Elkin, Will Butak, and Anindo Choudhury
Division of Natural Sciences, Saint Norbert College, De Pere, WI

Abstract:

DNA barcoding (Folmer et al. 1994), is an efficient method to distinguish species by short specific DNA sequences from a common region of their genome, such as the cytochrome *c* oxidase subunit 1 (CO1) gene. The early primers of Folmer et al. (1994) were not as universal as desired (Elias et al. 2007). Platyhelminthes (flatworms), being the fourth most speciose animal phylum, is a taxonomic group where universal barcoding primers are not very effective (Vanhove, et al. 2013). Recently, Van Steenkiste et al. (2014) and Elbrecht and Leese (2017) developed primers for parasitic and free-living flatworms that show promise as molecular barcodes. We tested the primers developed by Van Steenkiste et al. (2014), Dice 1F, Dice 11R, and Dice 14R, on a diverse collection of trematodes, a group with a significant number of undescribed taxa. A total of 120 amplifications were performed on 69 trematode samples from 27 genera. Dice 1F/11R and Dice 1F/14R primer sets were tested alongside the JB3/JB5 primer set (Bowles et al. 1992; Derycke et al. 2005). Overall amplification efficacy was notably larger for the Dice 1F/11R primer set. However, the JB3/JB5 primer pair led to a higher percentage of successful sequences, as compared to either of the Dice primers.

Introduction:

DNA barcoding (Folmer et al. 1994), is an efficient method to detect and distinguish species by short specific DNA sequences of a common region of their genome, such as the cytochrome *c* oxidase subunit 1 (CO1) gene. The early primers of Folmer et al. (1994) were not as universal as desired and were only moderately successful on certain groups of organisms (Elias et al. 2007). Platyhelminthes, being the fourth most speciose animal phylum, is a taxonomic group where universal primers are not very effective (Vanhove, et al. 2013). Recently, Van Steenkiste et al. (2014) and Elbrecht and Leese (2017) developed primers for parasitic and free-living flatworms that show promise in sequencing the CO1 barcoding region. The main objectives of this study were to (1) test the primers developed by Van Steenkiste et al., namely Dice 1F, Dice 11R, and Dice 14R for amplification efficacy on a variety of trematodes, and (2) use the resulting amplified barcodes to uncover cryptic species.

Materials and Methods:

In total, 69 extracts from 37 taxa belonging to 27 genera were used in this study (Table 1). Specimens had already been extracted prior to this study using Qiagen's DNEasy Blood and Tissue Kit (Qiagen Inc. CA). DNA concentrations were measured using ND-1000 Spectrophotometer V3.3 (NanoDrop Technologies, Inc., DE).

The CO1 gene was amplified by PCR, using the following primer sets: (1) Dice 1F/Dice 11R, (2) Dice 1F/14R, and (3) JB3/JB5. (Van Steenkiste et al. 2014; Bowles et al. 1992; Derycke et al. 2005). The JB3/JB5 primer set was the comparison, as they are frequently used for obtaining partial CO1 genes from platyhelminths (Pinto et al. 2018; Greiman et al. 2018). PCRs had a volume of 50 ul and contained; 25 uL of Phusion Hot Start Flex DNA polymerase in 1X Phusion HF buffer (New England Biolabs, Inc.), 0.5 uM of each primer from the primer set, and 5 ul of template DNA. The thermocycling conditions were slightly modified from the Van Steenkiste et al. protocol to accommodate the Phusion Hot Start Flex DNA polymerase. The thermocycling parameters were; 98 °C for 30 s, 94 °C for 90 s; 3 cycles of 94 °C for 40 s, 51 °C for 40 s, 72 °C for 1 min; 5 'touchdown' cycles of 94 °C for 40 s, 50 °C to 46 °C for 40 s (dropping 1 °C per cycle), 72 °C for 1 min; 35 cycles of 94 °C for 40 s, 45 °C for 40 s, 72 °C for 1 min; and a final extension at 72 °C for 7 min. The amplified products were run on a 1.5% TAE agarose gel stained with either ethidium bromide or 6X GelRed® Prestain Loading Buffer with Tracking Dye (Biotium, Fremont, CA). Samples were purified using ExoSAP-IT PCR Product Cleanup (Affymetrix, Inc. Santa Clara, CA) and sent for sequencing to MCLab (South San Francisco, CA). Sequences were assembled and aligned using Clustal W in Mega X (Kumar et al. 2018).

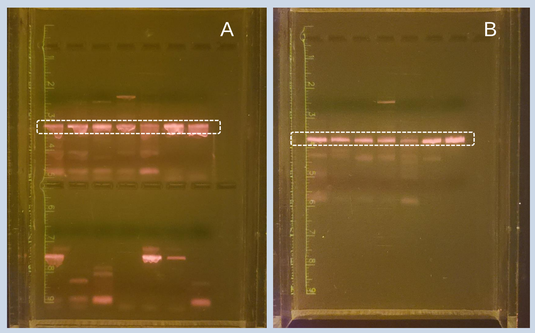

Figure 1. White box around target band. (A) Products of Dice 1F/11R amplification. (B) Purified products from Figure A.

Taxa	Extract ID #	DNA Concentration (ng/uL)	JB3/JB5 Primer	Dice 1F/11R Primer	Dice IF/14R Primer
Phyllodistomum sp.	X0049	8.3	-	+	+
Phyllodistomum sp.	X0051	20.5	-	+	+
Phyllodistomum sp.	X0055	6.2	+	+	-
Crepidostomum cooperi	X0092	5.1	+	N/A	N/A
Crepidostomum cooperi	X0095	9.6	+	N/A	N/A
Lissorchis sp.	X0104	14	+	-M	-
Lissorchis sp.	X0105	4.8	+	-M	-
Lissorchis sp.	X0107	8.2	+	+	-M
Lissorchis sp. (macropharynx)	X0110	6.3	N/A	+	+
Lissorchis sp. (gullaris)	X0111	4.1	N/A	+	+
Azygia longa	X0147	21.4	+	N/A	N/A
Azygia longa	X0148	4.1	N/A	+M	N/A
Azygia longa	X0149	3.9	+	N/A	N/A
Azygia sp.	X0150	0.4	N/A	+M	N/A
Hysterpmorpha sp.	X0208	25.8	N/A	+M	N/A
Crepidostomum sp.	X0209	2.7	N/A	-	N/A
Creptotrema sp.	X0216	3.2	+	+M	-
Paracreptotrematina limi	X0226	1.5	N/A	+M	+
Crepidostomum percopsisi	X0227	10.2	N/A	+	N/A
Achrolicanus auriculatum	X0238	6.5	N/A	+	N/A
Crepidostomum sp.	X0240	3.2	N/A	+	N/A
Crepidostomum sp.	X0244	4	N/A	-	N/A
Crepidostomum metoecus	X0247	3.1	N/A	+M	+
Crepidostomum farionis	X0248	6	N/A	+M	+
Plagiocirrus sp.	X0255	4.6	-	+M	-
Masenia sp.	X0338	11.2	-	+M	-
Bunodera saculata	X0345	2.3	+	N/A	N/A
Strigeidae gen. sp.	X0427	5	N/A	+M	+
Strigeidae gen. sp.	X0428	5.1	N/A	+M	+
Strigeidae gen. sp.	X0429	2.3	N/A	-	+
Cryptocotyle sp.	X0436	1.4	-	+	+
Opegaster sp.	X0504	2.8	-	+	-
Crepidostomum sp.	X0514	8.4	+	+	+
Plagioporus sp.	X0526	6.1	+	+M	-
Phyllodistomum sp.	X0561	3	+	N/A	N/A
Plagioporus sp.	X0566	1.6	N/A	+M	N/A
Nezpercella sp.	X0567	2.4	N/A	+	N/A
Hemiuroidea gen. sp.	X0570	8.1	N/A	+	N/A
Plagioporus sp.	X0574	9.8	N/A	+M	N/A
Plagioporus sp.	X0579	2.1	N/A	+M	N/A
Phyllodistomum sp.	X0581	15.4	N/A	+	N/A
Phyllodistomum sp.	X0593	8.1	N/A	+	N/A
Crepidostomum cooperi	X0599	6.4	N/A	+	N/A
Bunodera sp.	X0614	11.8	N/A	-	N/A
Bunodera eucaliae	X0627	18.3	N/A	-	N/A
Bunodera inconstans	X0629	2.1	+	N/A	N/A
Clinostomum sp.	X0638	22.1	N/A	-	N/A
Clinostomum sp.	X0669	6.8	N/A	+M	N/A
Haploporidae gen. sp.	X0685	1.1	N/A	+	N/A
Crassicuttis sp.	X0687	1.9	N/A	-	N/A
Haploporidae gen. sp.	X0688	1.5	N/A	+	N/A
Acanthostomum sp.	X0690	0	+	N/A	N/A
Haploporidae gen. sp.	X0692	0.8	N/A	+	N/A
Phyllodistomum sp.	X0698	0.6	N/A	+M	N/A
Haploporidae gen. sp.	X0703	2	N/A	+M	N/A
Gauhautiana sp.	X0709	1.7	N/A	+M	N/A
Masenia sp.	X0710	0.9	N/A	+M	N/A
Trematode gen. sp.	X0712	1.4	N/A	+M	N/A
Isoparorchis sp.	X0713	3	N/A	+M	N/A
Puntiotrema/Macrolecthius sp	X0716	4.3	N/A	+	N/A
Asymphyladora sp.	X0717	0.7	N/A	-	N/A
Masenia sp.	X0727	4	+	N/A	N/A
Cestrahelmins sp.	X0771	3.5	-	-	N/A
Cestrahelmins sp.	X0772	1.4	-	N/A	N/A
Cestrahelmins sp.	X0776	2	+	N/A	N/A
Crepidostomum sp.	X0781	3.3	-	+	N/A
Crepidostomum sp.	X0785	3	+	N/A	N/A
Acetodextra sp.	X0793	3.8	+	N/A	N/A
Acetodextra sp.	X0794	6.5	N/A	-	N/A

Table 1. Summative table of all extracts with corresponding taxa, DNA concentration, and amplification results.

Results:

Of the 120 amplifications, 86 successfully amplified regions of the CO1 gene; 37.2% of the amplifications showed mispriming (Table 2). We were unable to amplify the CO1 gene from the genera *Asymphyladora* and *Crassicuttis*, but the amplifications were only attempted with the Dice 1F/11R set. Amplification efficacy for JB3/JB5 and Dice 1F/11R primer sets were similar, 76.6% and 73.5%, respectively. The Dice 1F/14R set was less successful at 57.1% efficiency (Table 2). The overall sequencing efficiency (i.e., the number of readable sequences over total sequencing reactions ran) was 52.6% (Table 3). Sequencing efficiency varied greatly by primer set; the Dice primer sets were less effective compared to the JB3/JB5 set (Table 3).

Primer Set	# of Samples Amplified	Amplified Target Gene	Primer Dimer Complex	Heavier MP	Lighter MP
JB3/JB5	30	23	0	0	1
Dice 1F/11R	69	51	18	12	21
Dice 1F/14R	21	12	6	0	2

Table 2. Frequency table of all amplification reactions after correct thermocycling conditions were established. Four amplifications for Dice 1F/11R had both lighter and heavier mispriming.

Primer Set	# of Sequencing Attempts	# of Successful Sequences	Forward Primer Sequence Only	Reverse Primer Sequence Only
JB3/JB5	19	13	0	0
Dice 1F/11R	13	5	0	2
Dice 1F/14R	6	2	2	0

Table 3. Frequency table of all sequencing reactions.

Discussion:

Dice 1F/11R was the most effective amplification and sequencing primer set from VanSteenkiste et al. (2014) for trematodes in this study. It should be noted that Phusion Hot Start Flex DNA polymerase in 1X Phusion HF buffer protocol recommends 50-250ng of template DNA per 50uL reaction (New England Biolabs, Inc.). A majority of our amplifications contained less than 50 nanograms of template and were still effective (Table 1).

However, when the Dice 1F/11R primer set was compared to the JB3/JB5 primer set, the sequencing efficiency was notably lower without gel extraction (Table 3). Although the JB3/JB5 primer set produced successful sequences at a higher rate, the sequences generated using JB3/JB5 were 360-375 bp. Using the Dice 1F/11R primer set produced sequences that were 550-580 bp in length.

VanSteenkiste et al. (2014) were able to isolate sequences from non-misprimed bands without performing gel extractions. However, most of our amplifications resulted in variable levels of mispriming (Tables 1 and 2). Purifying these amplified products using ExoSAP-IT did not eliminate misprimed bands (Figure 1). Given the high frequency of mispriming with Dice primer sets requires the additional gel extraction step to acquire sequences effectively. Considering the need for gel extraction using the Dice primers, our study suggests the JB3/JB5 primer set may be the better primer set for obtaining a partial CO1 sequence, if used with the previously described thermocycling conditions.

In future studies, touchdown cycles will be selectively narrowed to produce more specific annealing. We also intend to gel extract successful but misprimed amplifications and obtain sequences from them.

Acknowledgments: We would like to thank the SNC Collaborative for a Student-Faculty Collaborative Research Grant.

References: **Bowles J., D. Blair and D.P. McManus**. 1992; *Mol. and Bioc. Parasit.* 54:165-174; **Derycke, S., T. Remerie, A. Vierstraete, T. Backeljau, J. Vanfleteren, M. Vincx and T. Moens.** 2005; *Mar. Ecol. Prog. Ser.* 300: 91–103; **Elbrecht, V. and F. Leese.** 2017; *Front. Environ. Sci.* 5:11; **Elias, M., R.I Hill, K.R. Willmott, K.K. Dasmahapatra, A.V.Z Brower**. 2007; *Proc. R. Soc. B*. 274: 2881-2889; **Folmer R.H.A , M. Nilges, P.J.M. Folkers, R.N.H. Konings and C.W. Hilbers.** 1994; *Mol. Mar. Biol. Tech*. 3: 294-299; **Greiman S.E., V.V. Tkach, and J.A. Cook**. 2013; *J. Parasitol.*, 99(6):1045-1049; **Kumar, S., G. Stecher, M. Li, C. Knyaz and K. Tamura.** 2018; Mol. Biol. Evol. 35(6): 1547-1549; **Pinto H.A., E.A. Pulido-Murillo, R.R. Braga, V.L.T. Mati, A.L. Melo, V.V. Tkach**. 2018; *Parasitol. Res*., 117:1975-1978; **Vanhove, M.P.M., B. Tessens, C. Schoelinck, U. Jondelius, D.T.J. Littlewood, T. Artois and T. Huyse.** 2013; *ZooKeys*, 365: 355-379; **Van Steenkiste, N.V., S.A. Locke, M. Castelin, D.J. Marcogliese and C.L. Abbott.** 2014; *Mol. Eco. Reso.* 15: 945-952.