Chapter 3

DNA Barcodes for Insects

John James Wilson

Abstract

DNA barcoding refers to the technique of sequencing a short fragment of the mitochondrial cytochrome c oxidase subunit I (COI) gene, the "DNA barcode," from a taxonomically unknown specimen and performing comparisons with a reference library of barcodes of known species origin to establish a species-level identification. The library barcodes gain their value due to an intimate association—through the vouchered specimens from where they came—with other data; particularly Linnaean names, collection localities, and morphology in the form of digital images. Consequently, this chapter details means of efficiently obtaining barcodes along two general streams: rapid barcode assembly to populate the library and retrieval of barcodes from highly prized specimens, but also emphasizes organization and collection of the barcode collaterals.

Key words: BOLD, Databasing, Tissue subsampling, DNA extraction, High-throughput, DNA amplification, Sequencing, Sequence editing, Sequence aligning, Cytochrome oxidase, DNA barcoding

1. Introduction

The first task of DNA barcoding is the association of sequences with species names (1). These sequences, from "correctly" identified individuals, delineated with external information and classical morphological methods, are then incorporated into the reference barcode library—the Barcode of Life Datasystem (BOLD) (2). Campaigns charged with the goal of populating BOLD have embraced different approaches: large-scale regional inventories using freshly caught specimens (3, 4) or barcode "blitzes" of national parks (5) or museum collections (6). Sequence acquisition is the primary driver and consequently specimens may have "interim" taxonomic names (4). Alternative campaigns focus on

W. John Kress and David L. Erickson (eds.), *DNA Barcodes: Methods and Protocols*, Methods in Molecular Biology, vol. 858,
DOI 10.1007/978-1-61779-591-6_3,

providing "accurate" taxonomic names for sequences, e.g., Lepidoptera Barcode of Life: Sphingidae (7). These approaches do not truly represent a dichotomy, all have the same goal and often are combined (e.g., refs. 8–10); however, the diversity demonstrates a requirement for rapid, efficient, high-throughput laboratory methods to get from "field to fasta," together with methods for important specimens such as types which may deserve more individualized analysis.

Although the molecular methods in this chapter focus on generating library sequences, they can also be used by those wishing to use the library as an identification key. Through sequence comparisons, the reference library can be employed to identify unknown newly collected individuals to a species-level taxon. While much deliberation has surrounded the choice of particular criteria used during sequence comparison (11), DNA barcoding for identification to species is now relatively uncontroversial.

This vision of DNA barcoding still requires the protracted process of traditional species description (12, 13) and growth of the library through generating barcodes for "known" species is ultimately restricted by what is already "known." Another option is for barcoding to involve curation of the reference library though reciprocal illumination. Many studies document cases where species taxonomy was considerably and rapidly improved through the combination of DNA barcodes, morphological and ecological analyses and then swiftly incorporated back into an improved barcode reference library (e.g., ref. 9). We see the incorporation of "collateral" information in the library as vital to the success of the barcoding initiative. In fact, the connection of sequences (digital data easily transferred and analyzed) with other data (images, collection data, and historical taxonomy) is exactly what makes barcodes valuable communicators of biodiversity. Therefore our chapter provides protocols for the generation of the collaterals alongside generation of the sequences.

2. Materials

2.1. Specimen Collection

1. 99.9% Ethyl alcohol (Commercial Alcohols). Store in a flammable liquids cabinet.

2.2. Tissue Subsampling

1. ELIMINase® (Decon Labs Inc.™).
2. KimWipes (Kimberly–Clark Corporation).
3. Forceps (Fine Science Tools).
4. Microplate (Eppendorf).
5. Cap-strips (ABgene).

2.3. DNA Extraction and Lysis Buffers

1. 0.5 M EDTA pH 8.0: 186.1 g EDTA (Fisher Scientific®), ~20.0 g NaOH (Fisher Scientific®), made up to 1,000 ml with ddH_2O. Vigorously mix on magnetic stirrer with heater. The disodium salt of EDTA will not go into solution until the pH of the solution is adjusted to ~8.0 by the addition of NaOH. Give a brief rinse to NaOH granules with ddH_2O in a separate glass before dissolving them.
2. 1 M Tris–HCl pH 8.0: 26.5 g Trizma® base (Sigma®), 44.4 g Trizma® HCl (Sigma®), made up to 500 ml with ddH_2O.
3. Proteinase K 20 mg/ml: Add 5 ml of ddH_2O to a 100 mg package of Proteinase K (Promega®). Store in 0.5 ml aliquots at –20°C.
4. 0.1 M Tris–HCl pH 6.4: 6.06 g Trizma® base made up to 500 ml with ddH_2O. Adjust pH with HCl to 6.4–6.5.
5. 1 M NaCl: 29.22 g of NaCl (Fisher Scientific®) made up to 500 ml with ddH_2O.
6. 1 M Tris–HCl pH 7.4: 9.7 g Trizma® base, 66.1 g Trizma® HCl, made up to 500 ml with ddH_2O.
7. Insect Lysis Mix: 16.5 g of GuSCN (Sigma®), 12 ml of 0.5 M EDTA pH 8.0, 6 ml of 1 M Tris–HCl pH 8.0, 1 ml Triton X-100 (Sigma®), 10 ml Tween-20 (Fluka®), made up to a final volume of 200 ml with ddH_2O.
8. Cap-strips.

2.4. High-Throughput DNA Extraction

1. AcroPrep™ 96 1 ml filter plate with 3.0 µm Glass Fiber media over 0.2 µm Bio-Inert membrane, natural housing (PALL®).
2. Axyseal™ sealing film (Axygen Scientific®).
3. PP MASTERBLOCK®, 96-well, 2 ml (Greiner Bio-One®).
4. Binding Buffer: 354.6 g of GuSCN, 20 ml of 0.5 M EDTA pH 8.0, 50 ml of 0.1 M Tris–HCl pH 6.4, 20 ml of Triton X-100, made up to final volume of 500 ml with ddH_2O. Vigorously mix on magnetic stirrer with heater. If any recrystallization occurs, pre-warm at 56°C to dissolve before use.
5. Protein Wash Buffer: 26 ml of Binding Buffer, 70 ml of EtOH 96%, made up to 100 ml with ddH_2O. Stable at room temperature for ~1 week, discard if any crystallization occurs.
6. Wash Buffer: 300 ml of EtOH 96%, 23.75 ml of 1 M NaCl, 4.75 ml of 1 M Tris–HCl pH 7.4, 0.475 ml 0.5 M EDTA pH 8.0, made up to 475 ml with ddH_2O. Mix well, store at –20°C.
7. Microplate.

2.5. Archival Specimen DNA Extraction

1. DNeasy 96 Blood & Tissue Kit single columns (Qiagen): Buffers AL, AW1, AW2, and AE are included in the kit.
2. EtOH 96%.
3. Fisherbrand Premium Flat Top Microcentrifuge Tubes 1.5 ml (Fisher Scientific).

2.6. PCR Amplification

1. ELIMINase®.
2. KimWipes.
3. D-(+)-trehalose dehydrate (Sigma-Aldrich).
4. 10× PCR Buffer supplied with enzyme (Invitrogen).
5. 50 mM $MgCl_2$ (Invitrogen).
6. 10 mM dNTP mix (New England Biolabs).
7. 100 μM Primer Stock: Dissolve desiccated primer (Integrated DNA Technologies) in ($\$ \times 10$) μl ultrapure ddH_2O. $ is different for every primer and is the number measured in "nmol" that can be found on the tube which the desiccated primer arrives in. Store at –20°C.
8. Taq Polymerase (Invitrogen).
9. Microplate.
10. Cap-strips.

2.7. High-Throughput PCR Check

1. 2% Agarose E-gel® 96 gel (Invitrogen).
2. Mother E-Base™ (Invitrogen).

2.8. Single Specimen PCR Check

1. 50× TAE buffer: Dissolve 242 g Tris(hydroxymethyl)aminomethane in 500 ml ddH_2O, add 57.1 ml glacial acetic acid and 100 ml 0.5 M EDTA and make up to 1,000 ml with more ddH_2O.
2. Ultrapure Agarose (Fisher Scientific).
3. Parafilm (Fisher Scientific).
4. 6× Loading dye (Fermentas, Thermo Fisher Scientific).
5. 100-bp DNA ladder (Invitrogen).
6. GelRed 10,000× in water (Biotium).

2.9. Cycle Sequencing

1. ELIMINase®.
2. KimWipes.
3. Microplate.
4. Dye terminator mix v3.1 (Applied Biosystems).
5. 5× ABI Sequencing buffer (Applied Biosystems).
6. D-(+)-trehalose dehydrate.
7. Primer.

2.10. Sequencing Clean-Up

1. Sephadex® G-50 (Sigma-Aldrich).
2. MultiScreen® Column Loader (Millipore).
3. Acroprep™ 96 Filter plate with 0.45-μm GHP membrane (PALL).

4. MicroAmp® Optical 96-well Reaction Plate (Applied BioSystems®, Cat. No. N801-0560).
5. 0.1 mM EDTA pH 8.0 (Fisher Scientific).
6. Septa (Applied Biosystems).

3. Methods

3.1. Barcode of Life Datasystem

1. BOLD (2) is the recommended place to manage your barcoding efforts. We advise initiating your online "project" as the first step in any new study.
2. To create a user account on BOLD, navigate to the homepage http://www.boldsystems.org.
3. Under *MANAGEMENT & ANALYSIS* click on **Request a new user account.*
4. Fill in the required personal details, invent and reconfirm a password, and *Submit Request*. You will receive an e-mail with your username and password.
5. To create a new project, log in and under *Project Options* in the left-hand column click *Create New Project* (see Note 1).
6. The *Project Title* should be meaningful to facilitate easy recall later (e.g., "Moths of the Olympic Peninsula").
7. The *Project Code* is a short form of the title (3–5 letters, e.g., "LOP") and forms the basis of *Process ID*s (see Note 2).
8. Select *COI-5P—Cytochrome Oxidase subunit 1 5′ Region* as *Primary Marker* (see Note 3).
9. At the bottom of the form you can *Assign Users* of the project and select the type of access they will have.
10. *Save* the new project (see Note 4).
11. Each project can initially hold 999 records. If you will be submitting more than 999 specimens, repeat steps 1–5 above to create additional projects. Later the data can be temporarily merged online for analysis (see Note 5).

3.2. Specimens

1. For beginners to insect collecting, *Basic Techniques for Observing and Studying Moths and Butterflies* (14) is a good resource regardless of taxon focus.
2. For those eager to expand to large-scale regional inventories, Janzen et al. (4) provide details of a mammoth "guinea pig" initiative in Area de Conservacion Guanacaste, Costa Rica.
3. When collecting insects for DNA barcoding collectors should consider the data required by BOLD and GenBank to make sequence records "BARCODE" compliant (15).

Table 1
Specimen collection information

Field ID	Collectors	Collection date	Location (GPS)	Notes	Storage conditions
QUI001	Billy Black	13/9/2010	47.9053, -124.6261	Light trap	Collected into an individual tube of 99.9% EtOH which was transferred to the −20°C freezer on return to lab

4. The collector should note the way in which the insect was collected, e.g., light trap (see Table 1).
5. The storage conditions (see Table 1) should be recorded—Was the specimen stored in EtOH? Was it oven-dried? This information is crucial for understanding incidents of low DNA yield (see Note 6).
6. Care should be taken to prevent DNA degradation in your collected specimens. We advise specimens be collected into an individual tube of 99.9% EtOH (see Note 7) which is later deposited into a freezer (−20°C). If this is not possible, longer-term storage under ethanol at room temperature should suffice (16).
7. The ethanol should be frequently changed to ensure it remains at high concentration (as water diffuses out of the specimen, the ethanol concentration decreases) (16).
8. Museum collections can also be mined for barcoding (6). The most recently collected specimens and those preserved in ways which minimize DNA degradation are preferred, as are those with a good "record" (i.e., with information available to complete Table 1, and especially types).

3.3. Submitting Specimen Data to BOLD

1. Data can be submitted by typing directly on the webpage, but using a spreadsheet will most likely save time. Most data can be copied directly from the field records (e.g., Table 1).
2. Download a blank data template from BOLD (from http://www.boldsystems.org click *Documentation* > *Data management* > *Data submission protocol* > *download blank data submission template*).
3. Open the template in Microsoft Excel.
4. On the *Voucher Info* sheet, enter *Sample ID*s (see Note 8).
5. *Museum ID* and *Collection Code* are optional (see Note 9).
6. *Institution Storing* must be completed for each sample. This is the location of the voucher specimen, not subsampled pieces of tissue (can be a private collection, see Table 2, museum or university).

Table 2
Minimum data required for submission of specimen records to BOLD

Voucher info sheet			Taxonomy sheet	Collection data sheet
Sample ID	**Field ID**	**Institution Storing**	**Phylum**	**Country**
QUI001	QUI001	Research collection of Carlisle Cullen	Arthropoda	USA

7. Proceed to the *Taxonomy* sheet and complete as fully as possible.
8. Proceed to the *Specimen Details* sheet and enter all known information (see Note 10).
9. Proceed to the *Collection Data* sheet and enter all known information (see Note 11).
10. Save the file (*File* > *Save*).
11. Spreadsheets can be uploaded by sending them through e-mail to submissions@boldsystems.org and must contain data in the *Sample ID*, *Field ID*, *Institution Storing*, *Phylum*, and *Country* columns (Table 2).

3.4. Specimen Imaging

1. Make a pedestal by mounting a 15 cm piece of drinking straw vertically in a wooden base. Plug the top of the straw with modeling clay to allow single-specimens to be pinned in the top (14).
2. Take pictures using the high-quality mode on your camera. If a fairly wide aperture (for shallow depth of field) is employed, background shadows will be negligible (14).
3. The specimen should be centered in the image frame.
4. Photos should be taken as close-up to the specimen as possible, leaving very little gap around the edges.
5. Use Landscape orientation.
6. Use 2 × 3 aspect ratio if possible. This will ensure that the images are not skewed when viewed in the BOLD image library.
7. If desired, a measurement scale may be included in the image to provide a size reference.
8. Use a standardized orientation (Table 3) as this makes it much easier to compare specimens within a project.

3.5. Submitting Images to BOLD

1. Create a folder on your desktop called *Images* and place in it all the image files (in .jpg format) you would like to upload.
2. To create a list of the files in the *Images* folder open a terminal window (*Start* > *Run* and type "cmd" into the black box that appeared in Windows), navigate to the *Images* folder (see Note 12), and then run one of the following commands: Windows: *dir* > *list.txt* ; MacOS: *ls* > *list.txt*.

Table 3
Common standardized animal orientations for specimen imaging

Orientation	Explanation
Dorsal	The anterior of the specimen should be facing the top of the image frame The specimen should be face-down, with the dorsal aspect of the head visible
Lateral	The anterior of the specimen should be facing the left side of the image frame The specimen should be oriented with the feet towards the bottom of the image
Ventral	The anterior of the specimen should be facing the top of the image frame The specimen should be face-up, with the ventral aspect of the head visible

3. Download a blank image submission template from BOLD (from www.boldsystems.org click *Documentation* > *Data management* > *Image submission protocol* > *please click here to download a blank image submission template*). Save the file (*ImageData.xls*) in the *Images* folder on the desktop.
4. Open *ImageData.xls* in Microsoft Excel.
5. Next open *list.txt* (see Note 13) and move the data into the *Image File* column in *ImageData.xls*. The cells in this column should contain the name of an image file including the extension *(.jpg*).
6. In the *Original Specimen* column type *yes* for original or *no* for not original.
7. In the *View Metadata* column choose one of the standard options from Table 3.
8. In the *Caption* column type any information you wish to appear by the image on BOLD.
9. Obtain the *Sample ID*s and *Process ID*s from BOLD by clicking on *Data Spreadsheets* under the *Downloads* menu on the left side of your *Project Console* (see Note 4). Choose to download the *Progress Report*, open the file *bold.xls* and copy the data from the *Sample ID* and *Process ID* columns into the appropriate columns in *ImageData.xls*.
10. Once you have filled in all the mandatory columns (see Note 14), save the file (*File* > *Save*).
11. The folder *Images* needs to be zipped before submission to BOLD. Most modern operating systems have built-in functionality for zipping (see Note 15) so this simply requires right-clicking on the folder and selecting *Compress "Images"* or something similar.
12. Navigate to your BOLD project's *Project Console* and under the *Uploads* menu on the left click *Specimen Images*. Browse through to *Images.zip* and click *Submit* (see Note 16).

3.6. Tissue Subsampling

1. Make sure that your specimens are correctly organized. This protocol is appropriate with pinned specimens stored in a drawer, specimens stored in ethanol tubes, and specimens stored in glassine envelopes. Print off a list of specimens and *Specimen ID*s so you can double-check they are going into the correct well as you go (e.g., QUI001 goes in A1, QUI002 goes in B1, …).
2. Use clean gloves.
3. Clean workspace and wipe bench with Eliminase.
4. Work on top of a KimWipe.
5. Get a microplate. These instructions are for a 96-well microplate; however, the procedure is essentially the same when working with single tubes. This is going to be the microplate in which the lysis takes place. Make sure the microplate is in the correct orientation, e.g., A1 is in the top right corner of the plate (Fig. 1).
6. It is worth assigning two wells as control wells (Fig. 1) at this point.
7. Put cap-strips on all rows.
8. Turn on the gas slowly and light the Bunsen burner, so that there is a small blue flame. When it is not possible or dangerous to use a Bunsen or gas burner an alternative Eliminase dip protocol can be just as effective (see Note 17) and can be easily adapted for use in the field.

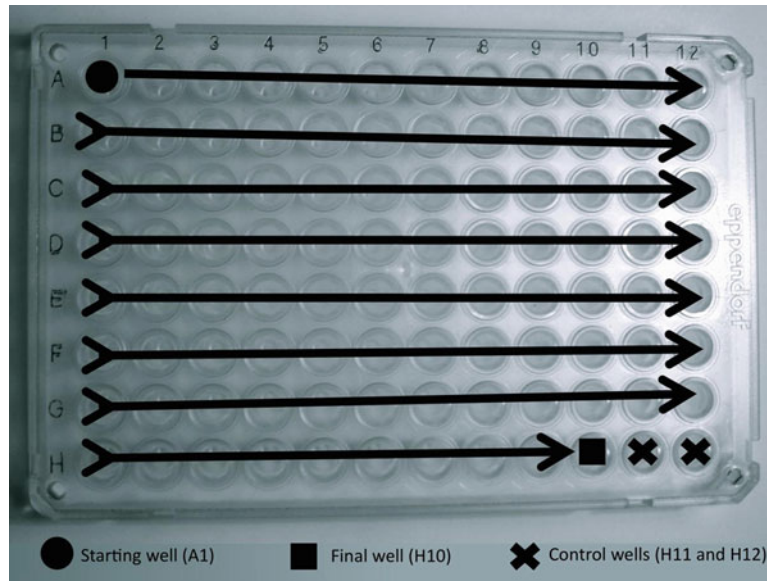

Fig. 1. Diagram detailing how to organize and fill a microplate with tissue, including designation of control wells.

9. Remove the first cap-strip (row A) from the microplate, and place it on a KimWipe. Take out the first specimen.
10. Take the forceps and dip them in ethanol (carefully shaking off any excess, not near the flame) and put them in the flame for a few seconds to burn off the ethanol.
11. Remove a small piece of tissue from the specimen (about a 2–3-mm-long piece of insect leg) and place it in the first well (A1) in the lysis plate.
12. It is also possible to use whole specimens in the lysis (17) or whole abdomens, in the case of a combined lysis/genitalia dissection (18) (see Note 18).
13. Continue on to the next specimen and well, making sure to sterilize the forceps between each sample with ethanol and flame. Place each specimen back in its drawer/tube/envelope before moving onto the next one.
14. Put the cap-strip back on as you finish a row and then carefully take the cap-strip off the next row.
15. Return your specimens to the freezer or cabinet.
16. For specimens in ethanol, the subsampled tissue needs to be completely dried before moving on to the next stage. Incubate at 56°C for 30 min, with the cap-strips slightly loosened, to evaporate residual ethanol.

3.7. Tissue Lysis

1. For one plate mix 5 ml of Insect Lysis Buffer and 0.5 ml of Proteinase K, 20 mg/ml in a sterile container (see Note 19 for single tubes).
2. Carefully remove all the cap-strips from the microplate (prepared as above). These instructions are for a microplate of tissue containing 96 wells (Fig. 1) but the procedure is the same for single tubes.
3. Add 50 µl of Lysis Mix to each well using a multichannel pipette. If you are careful not to touch the microplate with the tips, you can use the same tips right across the microplate.
4. Cover microplate with cap-strips.
5. Incubate at 56°C for a minimum of 6 h or overnight to allow digestion. It is not necessary to shake the microplate during incubation.
6. Centrifuge at 1,500 × *g* for 15 s to remove any condensate from the cap-strips (see Note 20).

3.8. High-Throughput DNA Extraction (Ivanova et al. (19))

1. Retrieve your microplate from the lysis stage above and remove cap-strips. Add 100 µl of Binding Mix to each sample using multichannel pipette. Cover plate with new cap-strips. Shake vigorously for 10–15 s and centrifuge at 1,000 × *g* for 20 s to remove any sample from the cap-strips.

2. Remove cap-strips and transfer the lysate (about 150 µl) from the wells into the wells of a GF plate placed on top of a square-well block using multichannel pipette. Seal the plate with self-adhering cover.
3. Centrifuge at 5,000×*g* for 5 min to bind DNA to the GF membrane.
4. First wash step: Add 180 µl of Protein Wash Buffer to each well of GF plate (see Note 21). Seal with a new cover and centrifuge at 5,000×*g* for 2 min.
5. Second wash step: Add 750 µl of Wash Buffer to each well of the GF plate (see Note 22). Seal with a new self-adhering cover and centrifuge at 5,000×*g* for 5 min.
6. To avoid incomplete Wash Buffer removal open the sealing cover, close it, and centrifuge the GF plates again for 5 min at 6,000×*g*.
7. Remove the self-adhering cover. Place GF plate on the lid of a tip box (see Note 23). Incubate at 56°C for 30 min to evaporate residual ethanol.
8. Position a PALL collar on a new microplate and place the GF plate on top. Dispense 30–60 µl of ddH_2O (pre-warmed to 56°C) directly onto the membrane in each well of GF plate and incubate at room temperature for 1 min. Seal plate.
9. Place the assembled plates on a clean square-well block to prevent cracking of the collection plate and centrifuge at 5,000×*g* for 5 min to collect the DNA eluate. Remove the GF plate and discard it.
10. Cover DNA microplate with cap-strips or aluminum PCR foil. This is your DNA and it can be temporarily stored at 4°C or at –20°C for long-term storage. Label it well.

3.9. Archival Specimen DNA Extraction (See Note 24)

1. Vortex the sample from lysis stage (about 150 µl) for 15 s (see Note 25).
2. Add 200 µl Buffer AL and vortex it (a white precipitate will most likely form). Add 200 µl EtOH 96% and vortex until it is homogeneous (there should be a lot less white precipitate).
3. Pipette the liquid (set the pipette to 650 µl) into a spin column, make sure to label the cap, and centrifuge it at 6,000×*g* for 1 min.
4. Discard the collection tube (see Note 26) and put the spin column into a new tube. Add 500 µl Buffer AW1 and centrifuge at 6,000×*g* for 1 min.
5. Discard the collection tube and put the spin column into a new tube. Add 500 µl Buffer AW2 and centrifuge at 20,000×*g* for 5 min.

6. Discard the collection tube and remove the spin column carefully so as not to let it touch the liquid. Put the spin column into a 1.5 or 1.7-ml microcentrifuge tube (see Note 27) and label it.
7. Add 100 µl Buffer AE (elution buffer) and let it sit at room temperature for 1 min. Centrifuge at 6,000 × *g* for 1 min.
8. Pipette the DNA out of the bottom of the microcentrifuge tube (about 100 µl), place it back into the spin column, and place the spin column back in the microcentrifuge tube. Centrifuge for an additional 1 min at 6,000 × *g*.
9. The liquid in the bottom of the microcentrifuge tube is your DNA. DNA can be temporarily stored at 4°C or at −20°C for long-term storage. Label it well.

3.10. Designing PCR Primers

1. Primers should be between 20 and 30 nt in length.
2. Avoid complementarity within and between primers.
3. The GC content should be approximately 50%.
4. Avoid mono- or dinucleotide repetition within primers.
5. The primer should end on a G or a C.
6. Primers should end on the second (or first if necessary) position of a codon.
7. The melting temperatures of primer pairs should be within 5°C of one another.
8. To design COI primers for a particular taxonomic group, try aligning as many COI genes from closely related taxa as possible (try surfing GenBank http://www.ncbi.nlm.nih.gov/genbank/) for the desired species group. Design primers that are situated in regions that are conserved across all taxa.
9. Remember to target the "barcode" region [i.e., overlapping with the region targeted by the Folmer primers (20) (Table 4)].
10. Primers can be tailed with M13 tails (Table 4) to improve amplification success (16) and facilitate high-throughput sequencing protocols (21). However, some tailed versions can form strong primer dimers, reducing PCR efficiency (e.g., LepF1 and LepR1).

3.11. PCR Set-Up

1. Prepare PCR master mix either for a single tube or 96-well microplate following the recipe in Table 5 where details on ingredient preparation are also provided.
2. We use LepF1 and LepR1 (Table 4) as the primer pair in a first amplification attempt (see Note 28).
3. Remember as above to wear clean gloves, clean benches with Eliminase and work on top of a KimWipe. Also work in a cold block if possible.

Table 4
Common primers used for DNA barcoding insects

Name	Sequence	Use with	Direction	References
LCO1490	GGTCAACAAATCATAAAGATATTGG	HCO2198	F	(20)
HCO2198	TAAACTTCAGGGTGACCAAAAAATCA	LCO1490	R	(20)
LepF1	ATTCAACCAATCATAAAGATATTGG	LepR1	F	(24)
LepR1	TAAACTTCTGGATGTCCAAAAAATCA	LepF1	R	(24)
MLepF1	GCTTTCCCACGAATAAATAATA	LepR1	F	(25)
MLepR1	CCTGTTCCAGCTCCATTTTC	LepF1	R	(25)
M13F (-21)	TGTAAAACGACGGCCAGT		F	(26)
M13R (-27)	CAGGAAACAGCTATGAC		R	(26)

Table 5
Basic recipe for PCR

	Amount of ingredient (µl)		
Ingredient	**Single tube**	**96-well microplate**	**Ingredient preparation**
10% Trehalose	6.25	625	Dissolve 5 g D-(+)-trehalose dehydrate in 50 ml of total volume of molecular grade ddH_2O. Store at –20 C
ddH_2O	2	200	Store at 4°C
10× Buffer	1.25	125	10× PCR Buffer for Platinun Taq. Store at –20°C
50 mM $MgCl_2$	0.625	62.5	50 mM $MgCl_2$. Store at –20°C
10 mM dNTPs	0.0625	6.25	10 mM dNTPs mix. Store at –20°C in 100 µl aliquots
10 µM F Primer working solution	0.125	12.5	Add 20 µl of 100 µM primer stock to 180 µl ultrapure ddH_2O. Store at –20°C
10 µM R Primer Working Solution	0.125	12.5	Add 20 µl of 100 µM primer stock to 180 µl ultrapure ddH_2O. Store at –20°C
Taq (5 U/µl)	0.06	6	Platinum Taq polymerase. Store at –20°C in 50 µl aliquots
Total	10.5	1,050	

4. Label your mix tube and microplate (see Note 29).
5. Return PCR ingredients to the freezer.
6. Mixes in tubes can be stored at –20°C for up to 3 months (1–3 freeze–thaw cycles do not affect performance). The content of a tube should be mixed by pipetting before use.
7. For microplate (see Note 30): Aliquot 1/8 of total mix volume to each of the tubes in an 8-tube PCR strip (see Note 31). Dispense desired volume (10.5 µl for 12.5 µl reactions) into each well of the 96-well plate using multichannel pipette.
8. Retrieve your DNA plate/tube from the fridge. Add 1–2 µl of DNA extract (see Note 32) to each tube/well (see Note 33). Seal and return DNA. Seal microplate with self-adhering aluminum foil (for PCR) or close the tube.
9. Centrifuge the microplate/tube at 1,000 × *g* for 20 s and start thermocycling.

3.12. PCR Thermocycle Program

1. Typical conditions for COI amplification include the initial denaturation at 94°C for 1 min; five cycles of 94°C for 30 s, annealing at 45–50°C for 40 s, and extension at 72°C for 1 min; followed by 30–35 cycles of 94°C for 30 s, 51–54°C for 40 s, and 72°C for 1 min; with a final extension at 72°C for 10 min, followed by indefinite hold at 4°C.
2. Centrifuge the microplate/tube at 1,000 × *g* for 20 s.

3.13. High-Throughput PCR Check

1. Precast agarose gels (E-gels) and docks (E-bases) to use them on are available from Invitrogen™. This system is bufferless, so exposure to Ethidium Bromide is minimized. However, gloves should be worn when handling and loading the gel.
2. The recommended program for 2% Agarose E-gel® 96 gel is EG and the run time is 6 min.
3. Plug the Mother E-Base™ into an electrical outlet. Press and release the pwr/prg (power/program) button on the base to select program EG.
4. Remove gel from the package and remove plastic comb from the gel.
5. Slide gel into the two electrode connections on the Mother or Daughter E-Base™.
6. Load 16 µl of ddH_2O into wells with 12-multichannel pipettor.
7. Load 4 µl of sample from your PCR microplate.
8. To begin electrophoresis, press and release the pwr/prg button on the E-Base™. The red light changes to green.
9. At the end of run (signaled with a flashing red light and rapid beeping), press and release the pwr/prg button to stop the beeping.

10. Remove gel cassette from the base and capture a digital image of the gel on UV transilluminator equipped with digital camera.
11. As a rough guide, set the filter to two for Ethidium Bromide and the exposure time to 2 s.
12. Analyze the image and align or arrange lanes in the image using the E-editor™ 2.0 software available at: http://tools.invitrogen.com/egels/.
13. White bands indicate product; square slots are the loading wells.

3.14. PCR Check: Important and Old Specimens

1. This protocol requires you make the gel yourself which is more time consuming, but cheaper on materials and produces gels that are more sensitive to product.
2. Gel should be ~5 mm thick: measure the size of the gel tray and determine the volume of liquid you will need to make a 5-mm thick gel (e.g., for a gel tray measuring 10×20 cm you would need to start with—20×10×0.5—100 ml of 1× TAE buffer, see Note 34).
3. Make sure that your tray is on a flat surface with tape securely on the sides.
4. Tape the edges of the tray so that it will hold liquid.
5. Measure out the agarose powder onto a piece of weigh paper using a metal spatula. The amount of agarose powder that you need depends on the percentage of the gel. Generally, 2% gels are best (e.g., to make a 2% agarose gel for the tray that takes 100 ml of buffer you need 1 g of agarose powder).
6. Add the agarose to a large beaker or Erlenmeyer flask. Add the 1× TAE buffer to the agarose.
7. Place the flask in the microwave on high power for 30 s. Gently swirl the flask using heat resistant gloves. Heat for another few seconds until the agarose has dissolved.
8. Be very careful because the agarose could burn you.
9. Let the flask sit for 5 min on the lab bench at room temperature to cool.
10. Place a comb of desired well width into the tray. Pour the hot liquid into the middle of the tray, trying to avoid creating bubbles. Push any large bubbles to the edges of the tray using a clean pipette tip.
11. Allow the gel to cool for 30 min. Remove the tape from the edges.
12. Set up the gel rig.
13. The liquid in the base should be 1× TAE buffer (see Note 35).
14. The gel should now be set and you can remove the tape from the edges of the gel tray.
15. Slowly lower the tray into place in the gel rig.

16. Add more 1× TAE buffer to the gel rig until the gel is completely submerged.
17. Carefully remove the comb from the gel by rocking it back and forth while pulling up slowly.
18. Cut a piece of parafilm and place flat on the lab bench. For every PCR product you will be adding to the gel place a 1 μl drop of loading dye (see Note 36) onto the parafilm. Take your PCR product and add 6 μl to one of the droplets of dye. Using the same pipette tip draw up the PCR product/dye droplet.
19. With a steady hand, add this to a well in the gel. With multiple samples be sure to keep track of which well is holding which sample.
20. The loading dye makes the product heavy so it will sink to the bottom of the well. You can hold the tip directly above the well without entering the gel. When you add the samples to the wells be careful not to poke a hole in the gel.
21. Always run a DNA ladder in a well beside your samples. A ladder of 100 bp would be appropriate and you should add 1 μl of ladder for every 5 mm of well width.
22. Close the top of the gel rig.
23. Remember to have the black electrode near the wells and the red electrode at the opposite end of the gel. DNA runs towards the positive electrodes.
24. Run the gel with the rig set to 150 V.
25. The loading dye forms two bands that you can see—wait until they have moved close to the bottom of the gel then turn off the rig (approximately 20 min).
26. Carefully transfer the gel from the tray into a plastic container for staining. Pour in diluted GelRed (see Note 37) until it covers the gel. Let it sit with moderate manual mixing for 20 min.
27. Pour the GelRed back into the bottle carefully using the funnel.
28. Capture a digital image of a gel on a UV transilluminator, equipped with digital camera, usually located in your institution's dark room.
29. As a rough guide set the filter to two for Ethidium Bromide and the exposure time to 2 s.
30. Print and save the image.

3.15. Cycle Sequencing Set-Up

1. When sequencing PCR product, you sequence in both forward and reverse directions. This is done with two different reactions and each reaction mix should include only a forward or reverse primer, not both. For example, for each microplate of PCR product, two microplates must be set up for sequencing, one with the forward primer and one with the reverse primer (see Note 38).

Table 6
Basic recipe for cycle sequencing

Ingredient	Amount of ingredient (µl)	
	Single tube	96-well microplate
Dye terminator mix v3.1	0.25	26
5× ABI sequencing buffer	1.875	195
10% Trehalose	5	520
10 µM Primer working solution	1	104
ddH_2O	0.875	91
Total	9	936

2. Prepare cycle sequencing master mix either for a single tube or 96-well microplate following the recipe in Table 6 and details on ingredient preparation in Table 5.
3. Remember as above to wear clean gloves, clean benches with Eliminase and work on top of a KimWipe. Also work in a cold block if possible.
4. Label your mix tube and any microplates (see Note 29).
5. Return cycle sequencing ingredients to the freezer.
6. Mixes in tubes (or pre-made plates, see Note 39) can be stored at –20°C for up to 3 months (see Note 40). The content of a tube should be mixed by pipetting before use.
7. For microplate : Aliquot 1/8 of total mix volume (115 µl) into each of the tubes in an 8-tube PCR strip (see Note 31). Dispense desired volume (9 µl) into each well of the 96-well plate using multichannel pipette. Changing tips after every row (see Note 41).
8. Retrieve your PCR product from the fridge. Add 1.5 µl of PCR product (see Note 42) to each tube/well (see Note 33). Seal and return PCR plate. Seal cycle sequencing microplate with self-adhering aluminum foil (for PCR) or close the tube.
9. Centrifuge the microplate/tube at 1,000 × *g* for 20 s and start thermocycling.

3.16. Cycle Sequencing Thermocycle Program

1. Denaturation at 96°C for 2 min.
2. Thirty cycles of 96°C for 30 s, annealing at 55°C for 15 s.
3. Additional extension at 60°C for 4 min.
4. Indefinite hold at 4°C (see Note 43).

3.17. Sequencing Clean-Up and Analysis (Ivanova and Grainger (22))

1. Sequencers should be operated by specially trained technicians, many facilities exist and will often require the Cycle Sequencing microplate, a supply of sequencing primer and a plate record (Table 7).
2. Measure dry Sephadex® G-50 with the MultiScreen® Column Loader into the Acroprep™ 96 Filter plate with 0.45 μm GHP membrane.
3. Hydrate the wells with 300 μl of ddH_2O.
4. Let the Sephadex® hydrate overnight in the fridge or for 3–4 h at room temperature before use.
5. Put Acroprep™ plate together with MicroAmp® Optical 96-well Reaction Plate and secure with at least two rubber bands.
6. Make sure the two sets weigh the same (adjust weight by using different rubber bands).
7. Centrifuge at 750 × *g* for 3 min—this is to drain the water from the wells. Discard water from MicroAmp® plates (these plates could be reused for the same procedure without autoclaving).
8. Add the entire volume of the cycle sequencing reaction to the center of Sephadex® columns.
9. Add 25 μl of 0.1 mM EDTA pH 8.0 to each well of the new (or autoclaved) MicroAmp® plate.

Table 7
Example of a plate record for a 3730*xl* DNA Analyzer (Applied Biosystems)

Container name	Plate ID	Description	Container type
LOP Plate1	LOP Plate1	COI-Barcodes	96-Well
AppServer	AppInstance		
Well	Sample name	Comment	Results group 1
A01	LOP001-11	LepF1	CC
B01	LOP013-11	LepF1	CC

App type	Owner	Operator	Plate sealing	Schedule pref
Regular	CCDB	CCDB	Septa	1234
Instrument Protocol 1	Analysis Protocol 1			
FolA700	3730BDTv3-KB-DeNovo_v5.1			
FolA700	3730BDTv3-KB-DeNovo_v5.1			

10. To elute DNA attach MicroAmp® plate to the bottom of the Acroprep™ plate—secure them with tape and with rubber bands.
11. Make sure the sets weigh the same (adjust weight by using different rubber bands).
12. Centrifuge at 750 × *g* for 3 min.
13. Remove MicroAmp® plate and cover its top with Septa.
14. Place MicroAmp® plate into the black plate base and attach the white plate retainer.
15. Stack assembled plate in 3730*xl* DNA Analyzer (Applied Biosystems)—do not forget barcode and plate record.
16. Discard Sephadex® from Acroprep™ plate.
17. Using the Plate Manager of the Data Collection software (Applied Biosystems), import the plate record(s) for the plate being run.
18. Begin the run within Run Scheduler.

3.18. Uploading Raw Sequences to BOLD

1. The sequencing outputs a folder of files. The files you are interested in have an extension *.ab1*, e.g., *LOP001-11_F.ab1*. These raw files (traces) can be edited into the form we are use to seeing DNA sequences represented in, i.e., a string of letters. However, as editing can be a subjective task, BOLD also requires the raw files (traces) be uploaded as part of a barcode's collateral data.
2. To add trace files of your new sequences to the appropriate records in BOLD create a folder on your desktop called *Traces* and place in it all the *.ab1* files that you would like to upload.
3. To create the list of files in the *Traces* folder, open a terminal window (*Start* > *Run* and type "cmd" into the black box that appeared in Windows), navigate to the *Traces* folder (see Note 44), and then run one of the following commands: Windows: *dir* > *list.txt*; MacOS: *ls* > *list.txt*.
4. Download a blank trace submission template from BOLD (from www.boldsystems.org click *Documentation* > *Data management* > *Trace submission protocol* > *please click here to download a blank trace submission template*). Save the file (*data.xls*) in the *Traces* folder on the desktop.
5. Open *data.xls* in Microsoft Excel.
6. You can then open *list.txt* and move the data into the *Filename* (*.ab1*) column in *data.xls* (see Note 45). The cells in this column should contain the name of a trace file including the extension (*.ab1*) (see Note 46).
7. In the *FORWARD PCR PRIMER* column enter the registered name of the forward primer used during the PCR. Copy it down through the entire column to the end of your file list.

Table 8
Example of file data.xls

A	B	C	D	E	F	G	J
Filename (.ab1)	Score file (.phd1)	Forward PCR Primer	Reverse PCR Primer	Sequencing Primer	Read direction	Process ID	Marker
LOP001-11_F.ab1		LepF	LepR	LepF	Forward	LOP001-11[a]	COI-5P
LOP001-11_R.ab1		LepF	LepR	LepR	Reverse	LOP001-11[b]	COI-5P

[a]Formula typed into this cell is = left(A2,9)
[b]Formula in this cell is = left(A3,9)

8. In the *REVERSE PCR PRIMER* column enter the registered name of the reverse primer used during the PCR. Copy it down through the entire column to the end of your file list.
9. In the column *SEQUENCING PRIMER* enter the registered name of your sequencing primer. It will alternate between forward and reverse. For example, *LepF1* should line up with the read direction Forward (Table 8).
10. In the *Read Direction* column enter *Forward* or *Reverse* depending on the direction of the *.ab1* files it refers to (see Note 47).
11. In the *Process ID* column you need to type in the formula "=left (A2, $)" where A2 is the column with your first *.ab1* file and $ is the number of characters in the *Process ID*. For example LOP001-11 has nine characters ($ = 9). Therefore, $ may be more or less depending on the number of letters in the *Project Code*.
12. Press *Control* and *A* to select the entire page. Press *Control* and *C* to copy the page, and then go *Edit* > *Paste Special* and chose *values* and press *OK*. This removes the formulas from your sheet.
13. Save this file under the name *data.xls* and save it in your folder *Traces*.
14. Delete *list.txt* from your *Traces* folder.
15. The folder *Traces* needs to be zipped before submission to BOLD. See Subheading 3.5 for details on how to do this. Save as *Traces.zip*.
16. Navigate to your BOLD project's *Project Console* and under the *Uploads* menu on the left click *Trace Files*. *Browse* through to *Traces.zip* and click *Submit* (see Note 48).

3.19. Sequence Editing

1. Open CodonCode (http://www.codoncode.com) and choose *Create a new project* and press *OK*.
2. Go to *File*> *Import*> *Add Folder*> *Traces* then press *Import*.
3. To see the files you just imported press ▶ beside the *Unassembled Samples* folder.
4. Your .ab1 files should be of the form "LOP001-11_F" where the first part "LOP001-11" refers to the *Process ID* and the second part "F" refers to the sequence direction, i.e., Forward.
5. Sort files by quality by double-clicking on *Quality*. Any sequences that are of very poor quality or of short length highlight them and click the trashcan to delete them.
6. Next select the *Contig* menu and move the cursor over *Advanced Assembly*. From the options that appear select *Assemble in Groups*.
7. A window will appear asking if you would like to *Define sample name parts?* Choose *Define names…* to bring up another small window.
8. There are two parts to our filenames. The first will be the *Process ID* and for your purposes the option in the *Meaning* menu can be left as *Clone*. Since the *Process ID* is followed by underscore choose _ *(underscore)* in the *Delimiter* menu for *Clone*.
9. For the second part choose *Direction* in the *Meaning* menu. We can ignore the *Delimiter* for the *Direction* part because there is no actual delimiter following the direction.
10. Delete all the additional parts that may appear on this window.
11. Next click *Preview…* to check how aligner is interpreting the sample names. Click *Close* to exit the preview.
12. Click *OK* to return to the *Assemble in Groups* window.
13. In order to assemble our files according to direction you should choose *Direction* in the *Name Part* section. Then click *Assemble*.
14. You should now have two folders, one called *F* with the forward sequences and one called *R* with the reverse sequences.
15. Next you need to cut the primers from your sequences. Highlight the *R* folder and reverse and compliment the sequences using the button with three black arrows on it.
16. Double click the *R* folder to open it. For the reverse sequences, you need to find the forward primer motif (e.g., LepF1) and delete it from the beginning of the consensus sequence. You will find the primer around 50 nucleotides from the end of the raw sequence. For example, in Fig. 2a, you would need to delete the sequence marked in bold and everything to the left of it.

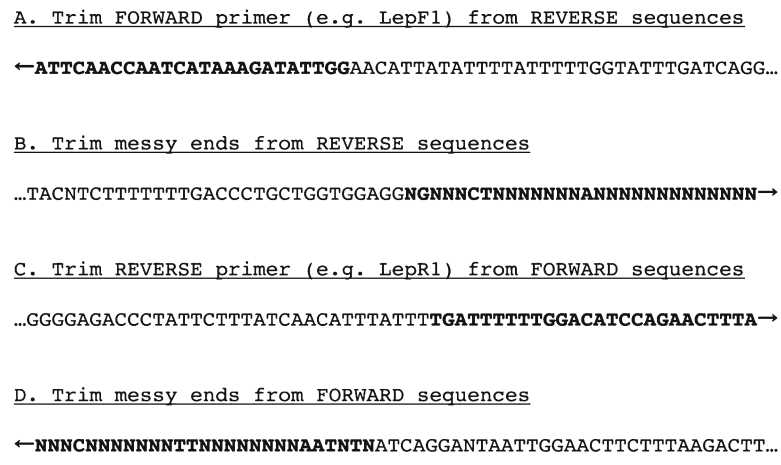

Fig. 2. An example of sequence editing.

17. When you have located the primer, highlight it on the consensus sequence at the bottom of the window and press the *Delete* key.
18. Next go to the opposite end of the consensus, the far right, and delete the consensus sequence from the point where many *N*s appear all the way to its very right-hand edge. For example, in Fig. 2b, you would delete the sequence marked in bold and everything to the right of it. Close the window.
19. Double click the *F* folder to open it. Go to the far right of the consensus sequence and find the reverse and complement of the reverse primer (e.g., LepR1) at the very end. This means that at the right end of the forward sequences, you will find the complement of the reverse primer backwards (e.g., if the reverse primer is ATGC then you will find GCAT at the end of your forward sequence). This should be around position 690–700 bp on the consensus sequence. For example, in Fig. 2c, you would need to delete the sequence marked in bold and everything to the right of it.
20. When you have located the primer, highlight it on the consensus sequence at the bottom of the window and press the *Delete* key.
21. Next go to the opposite end of the consensus, the far left, and delete the consensus sequence from the point where many *N*s appear all the way to its very left-hand edge. For example, in Fig. 2d, you would delete the sequence marked in bold and everything to the left of it.
22. Dissolve both folders by clicking on the button marker with a red X.
23. Highlight all sequences and press the button marked with a black N. This time in order to assemble our files according to

Process ID choose *Clone* in the *Name Part* menu. Then click *Assemble*.

24. Specimens which only sequenced successfully in one direction will have files which remain in the *Unassembled Samples* folder (see Note 49).
25. Open each folder (contig) by double-clicking and make sure that forward and reverse sequences have the correct orientation, i.e., forward sequence is in black with the arrow pointing to the right and reverse in red with the arrow pointing to the left. If they are backwards, reverse-complement the two files in the folder by closing the window, highlighting the folder and clicking the button with three black arrows.
26. Correct ambiguous positions ("N"s) and gaps ("-"s) in consensus sequences by checking the original trace chromatograms, which are present in the CodonCode project. This is done by double-clicking on the consensus sequence. Always open both trace files (forward and reverse) and compare them.
27. Generally if reads conflict (i.e., different colored peaks appear in the same location on the forward and reverse chromatograms) you can decide which is more reliable based on sequence quality (e.g., less background noise, cleaner peaks, taller peaks).
28. Correct bases in contigs first, and then check the single sequences in the *Unassembled samples* folder. This is a good idea because not all contigs will be kept, some will be dissolved or deleted.
29. Make sure single sequences are also in the correct orientation before uploading to BOLD.
30. To export the consensus sequences select all the folders using shift click, go *File*>*Export*>*Consensus sequences…*, choose *Current selection*. Open the *Options* and check *Include gaps in FASTA* but uncheck all other options by clicking. Press *Export*. Save the file to the desktop as *sequences.fas* (see Note 49).

3.20. Sequence Aligning

1. Open the file *sequences.fas* in BioEdit (see Note 50).
2. Make sure *Mode:* is set to *Edit* using the drop-down menu.
3. Another drop-down menu will become visible to the right of the *Edit* drop-down. Make sure this is set to *Insert*.
4. Sequences that have ended up in the FASTA file in the wrong orientation may be corrected by highlighting the sequence name by clicking the cursor on it, clicking the *Sequences* menu at the top of the screen. Moving the cursor down the drop-down to *Nucleic Acid* and clicking, *Reverse complement*.
5. Sequences all need to be 658 bp and aligned to each other before being uploaded to BOLD. This can be done by typing additional Ns at the beginning and end of your sequences in

Unaligned FASTA

```
LOP001-11     AACTTTATATTTTATTTTTGGAATTTGAGCAGGAATAGTAGGAACCTCTT
LOP002-11     ATATTTTATTTTTGGAATTTGAGCTGGATTAATTGGAACTTCATT
LOP003-11     AACTCTATATTTTATTTTTGGAATTTGAGCAGGATTACTAGGAACT
LOP004-11     TCTATATTTTATTTTTGGAATTTGAGCAGGTTTAGTTGGAACTTCATT
LOP005-11_F   AATGATGTTCCTAACATACCTGCTCAAATACCAAAAATAAAATATAAAGT
```

Aligned FASTA

```
LOP001-11     AACTTTATATTTTATTTTTGGAATTTGAGCAGGAATAGTAGGAACCTCTT
LOP002-11     NNNNNNATATTTTATTTTTGGAATTTGAGCTGGATTAATTGGAACTTCAT
LOP003-11     AACTCTATATTTTATTTTTGGAATTTGAGCAGGATTACTAGGAACTNNNN
LOP004-11     NNNTCTATATTTTATTTTTGGAATTTGAGCAGGTTTAGTTGGAACTTCAT
LOP005-11     AACTTTATATTTTATTTTTGGTATTTGAGCAGGTATGTTAGGAACATCAT
MHAHC824-05   AACTTTATATTTTATTTTTGGAATTTGAGCAGGAATAGTAGGAACCTCTT
```

Fig. 3. An example of an unaligned and aligned FASTA file.

the BioEdit *Edit* mode. Be sure to check across the whole alignment of your sequences that you have added the correct number of Ns.

6. In Fig. 3 featuring a 50-bp barcode for simplification, LOP001-11 is of full length. LOP002-11 needs 6 Ns adding to the left side of the sequence to become aligned, while LOP003-11 needs 4 Ns adding to the right side of the sequence to be 50 bps long. LOP005-11_F needs to be reverse complemented to be in the same orientation as the other sequences.
7. Sequences which were not part of a consensus (i.e., when one direction failed but the single sequence is of sufficient length and quality for submission to BOLD) may appear in the FASTA still tagged with the direction. This needs to be deleted, e.g., the sequence named LOP005-11_F should be renamed as LOP005-11 (Fig. 3).
8. If you are having trouble with the alignment, a good quality (i.e., 658[0n]) sequence can be downloaded from BOLD and imported into your BioEdit file as a guide, e.g., MHAHC824-05 (Fig. 3). Be sure to delete this sequence before saving the file.
9. Save the file (*File* > *Save*).

3.21. Sequence Upload and Publication

1. Open the file *sequences.fas* in a text editor. This can usually be done by double-clicking on the file icon on the desktop.
2. Under the *Edit* menu, click *Select All* then under *Edit*, click *Copy*.
3. Navigate to your BOLD project's *Project Console* and under the *Uploads* menu on the left click *Sequences*.
4. Right click on the box *Paste sequences in fasta format:* and click *Paste*.
5. Select the *Markers:* as *COI-5P* and select or register your *Run site*.

6. Click *Submit*.
7. Once you are happy with all your data make your project public on BOLD by adding *Public* as a user (see Subheading 3.1).

4. Notes

1. The user who creates the project becomes the *Project Manager*. The *Project Manager* is the only user who can add new users. The *Project Manager* can be changed by contacting the BOLD team at support@boldsystems.org.
2. We recommend avoiding the letters A, C, G, T, U, R, Y, M, K, W, S, B, D, H, V, and N. These are IUPAC nucleotide codes including the ambiguity codes which will appear in your sequences when a sequence-editing program is unable to make a base call. If you use these letters you may have difficulties later on during sequence editing and manipulation.
3. BOLD can store data for other regions besides COI-5P. The extracted DNA obtained by following these methods can be stored in a freezer and subsequently used as a template for other gene regions. For information about amplification of other regions frequently sequenced for insects see refs. 16 and 21.
4. To return to the *Console* of your new project simply log into BOLD and click on the project name in your list of projects.
5. If you will be creating multiple projects it may be worth launching a *Campaign*. To start a *Campaign* contact support@boldsystems.org. Your *Campaign* can then be selected on the *Create New Project* page.
6. We've had incidences where specimens that should have sequenced successfully failed unexpectedly. After consultation with the collector we were able to attribute the failure to storage conditions.
7. DNA leaks into the storage fluid (23). For large specimens which may be damaged by ethanol, whole specimen can be stored dry and legs can be removed into ethanol.
8. If you enter *Sample ID* on any other sheet it will overwrite important macros. When copying and pasting information into the spreadsheet, use *Paste Special* and select *Values* to avoid overwriting formulas. Avoid using the *Project Code* in *Sample ID*s if possible, because the *Project Code* will form the basis of *Process ID*s. We advise *Sample ID*= *Field ID* where possible. Also keep in mind that each project in BOLD can have a maximum of 999 samples. Alternatively *Sample ID* can refer directly to the catalog number of a museum collection voucher.

9. Museum abbreviations should follow standard registers for biorepositories, e.g., http://www.biorepositories.org/.
10. In the *Sex* column use M for male, F for female, or H for hermaphrodite. *Reproduction* refers to the type of lifecycle (use S for sexual, A for asexual, or P for parthenogenic). For *Life Stage* use either I for immature or A for adult. *Extra info* will show up on taxonomic identification trees generated by BOLD. *Notes* will not be seen on the tree, but they will appear on BOLD on the *Specimen Page*.
11. *Latitude* (North–South) and *Longitude* (East–West) must be in decimal format. A useful website for this conversion is http://www.calculatorcat.com/latitude_longitude.phtml. *Elevation* must be in meters, but it is not necessary to put "m" for meters beside the number.
12. To navigate to the folder, type *cd desktop*, press enter, type *cd Images*, press enter.
13. In Excel, go to *File* > *Open…* (change *Enable:* to *All files*) navigate to *Desktop* > *Images* > *list.txt* and click *Open*. A window will open called *Text Import Wizard*. Select *Fixed width* and click *Next* >. Scroll down until you can see your first file named .jpg and move the arrowed line beside it over to the right a little so that it touches the file name click *Next* > and *Finish*.
14. For information on the optional columns: *Measurement*, *Copyright* etc, please refer to the BOLD handbook (http://www.boldsystems.org > *Documentation* > *Image Submission Protocol*).
15. If your machine does not have built in zipping functionality try downloading a free program (WinZip: http://www.winzip.com; WinRar: http://www.rarsoft.com; or MacZipIt: http://www.maczipit.com).
16. BOLD will give a message advising if the submission was successful. If problems have occurred, refer to the *Image Submission—Tips and Troubleshooting* section of the BOLD handbook (http://www.boldsystems.org > *Documentation* > *Image Submission Protocol*).
17. Eliminase protocol:
    (a) Get four clean jars. Label the first one "Eliminase" and add about 10 cm^3 of Eliminase.
    (b) Label the next three jars "Wash 1," "Wash 2," and "Wash 3" and fill them all with 10 ml or more of ddH_2O.
    (c) Dip your forceps into the Eliminase, shake them a bit, remove them and wipe off excess liquid using a clean KimWipe.
    (d) Dip your forceps and shake them slightly in Wash 1, then Wash 2, and finally in Wash 3. Dry them off using a clean KimWipe.

18. In these cases the tissue must be retrieved from the lysis buffer prior to undertaking the DNA extraction.
19. For single tube lysis add 50 µl of lysis buffer and 5 µl of ProK.
20. This is high speed so place the lysis plate on a clean square-well block to prevent cracking.
21. You will need to put about 18 ml of Wash Buffer in the reservoir for a full plate.
22. You will need to put about 75 ml of Wash Buffer in the reservoir for a full plate.
23. Square-well blocks can be washed with ELIMINase® (or with any other DNA removing detergent), autoclaved, and reused.
24. The Qiagen kits can be stored at room temperature (15–25°C). Some ingredients may need storage at lower temperatures.
25. The Qiagen protocol suggests that you use a mortar and pestle to crush part of the insect into a powder, but we have found that this is unnecessary.
26. The following buffers contain toxic components that must not be thrown in the regular garbage: AL, ATL, AW1. Be sure that all liquid waste from DNA extraction is kept in a well-labelled glass jar.
27. This tube is not from the kit. It is one with a cap (see Subheading 2.5).
28. We find this has a very high success rate (99%) with recently collected specimens (<5 years old) and most likely these will be the only primers you will need. If you do not have success with these primers (see Subheading 3.14) the next strategy we recommend to try would be two amplification reactions: (1) using primer combination MLepF1 and LepR1 and (2) using primer combination MLepR1 and LepF1 (see ref. 25 and Table 4). These products are then sequenced in a single direction using LepR1 for the first product and LepF1 for the second product. The sequences can then be combined into a single contig close in length to the target 658-bp full-length barcode. If you have a very old and very important specimen but the above strategies have been unsuccessful, a further (time-consuming and expensive) option is amplification using "micro" primers, consult (8) and (9).
29. Put a star in the corner A1.
30. If you plan to fill several microplates include extra volume to allow for pipetting mistakes and dead volume in the digital multichannel pipettor (e.g., for making ten microplates with 12.5 µl reactions each, include about 40 extra reactions).
31. Once you become proficient can use the final row of the microplate instead of 8-tube PCR strip.

32. Can use less DNA template if you find it is too concentrated.
33. Here it is essential to ensure your plates are in the same orientation.
34. TAE buffer is stored at 50× TAE. You will need to add ddH_2O to the concentrated TAE buffer to get a 1× dilution, e.g., add 5 ml of 50× TAE to 250 ml of ddH_2O to make 1× TAE.
35. Buffer in the rig is good for ~15 runs.
36. We use 6× loading dye, so this means 1 μl of dye for 6 μl of product.
37. GelRed should be diluted 10,000:1 with ddH_2O. This means that you need to add 25 μl of GelRed to 250 ml water. GelRed can be used several times. Make sure to keep track of the number of uses.
38. An exception is when using MLepF1 and MLepR1 (see Note 29).
39. Addition of trehalose makes possible the freezing of aliquoted sequencing mixes. Currently, the Canadian Centre for DNA Barcoding uses a batch strategy for making sequencing plates. Mixes are aliquoted directly into 96-well plates, using a Biomek® FX robot; plates are covered with PCR film and stored at –20°C for up to 3 months.
40. The mixes are light sensitive so should be stored in a light-proof box or wrapped in aluminum foil.
41. If you are a skillful pipettor and set up your forward and reverse sequencing plates at the same time you can use the same tips to put PCR product into each plate. You will only need to use one box of tips if you only touch the very edge of the wells near the top, and do not contaminate the tips with primer in the process.
42. The volume of PCR product added to the sequencing reaction can be varied between 0.5 and 1.2 μl depending on the strength of the gel band.
43. The annealing temperature can be varied according to the primer specificity, but 55°C works well for most COI sequencing reactions.
44. To navigate to the folder, type *cd desktop*, press enter, type *cd Traces*, press enter.
45. In Excel, go to *File*> *Open*… (change *Enable:* to *All files*) navigate to *Desktop*> *Traces*> *list.txt* and click *Open*. A window called *Text Import Wizard* will open. Select *Fixed width* and click *Next*>. Scroll down until you can see your first file named .ab1 and move the arrowed line beside it over to the right a little so that it touches the file name click *Next*> and *Finish*.

46. Score files are not compulsory and the column can be left blank; however, for more information on these and how you can include them see the BOLD handbook.
47. Usually the files names should be of the form (e.g., LOP001-11_F.ab1) *Process ID* (LOP001-11) and direction (F), which should make this easier.
48. BOLD will give a message advising if the submission was successful. If problems have occurred, refer to the *Trace Submission—Tips and Troubleshooting* section of the BOLD handbook (http://www.boldsystems.org > *Documentation* > *Trace Submission Protocol*).
49. To export the single direction sequences, go *File* > *Export* > *Samples…*, choose *Current selection*. Open the *Options* and select *Include gaps in FASTA* but deselect all other options. Press *Export*. Save the file to the desktop as *sequences.fas*.
50. BioEdit can be downloaded for free from http://www.mbio.ncsu.edu/bioedit/bioedit.html.

Acknowledgment

Heather Braid compiled the "Barcoding in the Hanner Lab" protocols (http://barcoding.wikia.com/wiki/Barcoding_in_the_Hanner_Lab_Wiki), which greatly aided with the structuring and content of this chapter, and also provided Fig. 1.

References

1. Floyd R, Wilson JJ, Hebert PDN (2009) DNA barcodes and insect biodiversity. In: Footit RG, Adler PH (eds) Insect biodiversity: science and society. Blackwell Publishing, Oxford, pp 417–431
2. Ratnasingham S, Hebert PDN (2007) BOLD: The barcode of life data system (www.barcodinglife.org). Mol Ecol Notes 7:355–364
3. Janzen DH, Hajibabaei M, Burns JM et al (2005) Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. Phil Trans R Soc Lond B 360:1835–1845
4. Janzen DH, Hallwachs W, Blandin P, Burns JM, Cadiou J-M, Chacon I et al (2009) Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity. Mol Ecol Res 9:1–25
5. Xhou X, Robinson JL, Geraci CJ, Parker CR et al (2011) Accelerated construction of a regional DNA-barcode reference library: caddisflies (Trichoptera) in the Great Smoky Mountains National Park. J Nor Amer Benth Soc 30:131–162
6. iBOL (2010) Barcoding blitz targets Australian Lepidoptera. Barcode Bulletin 1(4):5
7. Vaglia T, Haxaire J, Kitching IJ et al (2008) Morphology and DNA barcoding reveal three cryptic species within the *Xylophanes neoptolemus* and *loelia* species-group (Lepidoptera: Sphingidae). Zootaxa 1923:18–36
8. Hausmann A, Hebert PDN, Mitchell A et al (2009) Revision of the Australian *Oenochroma vinaria* Guenée, 1858 species-complex (*Lepidoptera*, *Geometridae, Oenochrominae*): DNA barcoding reveals cryptic diversity and assesses status of type specimen without dissection. Zootaxa 2239:1–21
9. Wilson JJ, Landry JF, Janzen DH et al (2010) Identity of the ailanthus webworm moth, a

complex of two species: evidence from DNA barcoding, morphology and ecology. Zookeys 46: 41–60

10. Dincă V, Zakharov EV, Hebert PDN, Vila R (2011) Complete DNA barcode reference library for a country's butterfly fauna reveals high performance for temperate Europe. Proc R Soc Lond B 278:347–355
11. Virgilio M, Backeljau T, Nevado B, de Meyer M (2010) Comparative performances of DNA barcoding across insect orders. BMC Bioinformatics 11:4567–4573
12. Vogler AP (2006) Will DNA barcoding advance efforts to conserve biodiversity more efficiently than traditional taxonomic methods? Front Ecol Environ 4:270–272
13. Pons J, Barraclough TG, Gomez-Zurita J et al (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Sys Biol 55:595–609
14. Winter WD Jr (2000) Basic techniques for observing and studying moths & butterflies (Memoir No 5). The Lepidopterist's Society, Cambridge
15. Hanner R. (2005) Proposed standards for BARCODE records in INSDC (BRIs). http://barcoding.si.edu/PDF/DWG_data_standards-Final.pdf
16. Regier JC (2008) Protocols, concepts, and reagents for preparing DNA sequencing templates. Version 12/4/08. http://www.umbi.umd.edu/users/jcrlab/PCR_primers.pdf
17. Porco D, Rougerie R, Deharveng L, Hebert PDN (2010) Coupling non-destructive DNA extraction and voucher retrieval for small soft-bodied Arthropods in a high-throughput context: the example of Collembola. Mol Ecol Res 10: 942–945
18. Knölke S, Erlacher S, Hausmann A et al (2005) A procedure for combined genitalia extraction and DNA extraction in Lepidoptera. Insect Syst Evol 35:401–409
19. Ivanova NV, deWaard J, Hebert PDN (2006) An inexpensive, automation-friendly protocol for recovering high-quality DNA. Mol Ecol Notes 6:998–1002
20. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotech 3: 294–299
21. Wilson JJ (2010) Assessing the value of DNA barcodes and other priority gene regions for molecular phylogenetics of Lepidoptera. PLoS ONE 5:e10525
22. Ivanova N, and Grainger C (2006) Protocols: Sequencing. Canadian Centre for DNA Barcoding CCDB Protocols. http://www.dnabarcoding.ca
23. Shokralla S, Singer GAC, Hajibabaei M (2010) Direct PCR amplification and sequencing of specimens' DNA from preservative ethanol. BioTechniques 48:232–234
24. Hebert PDN, Penton EH, Burns J, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly, *Astraptes fulgerator*. Proc Nat Acad Sci USA 101: 14812–14817
25. Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN (2006) DNA barcodes distinguish species of tropical Lepidoptera. Proc Nat Acad Sci USA 103:968–971
26. Messing J (1983) New M13 vectors for cloning. Meth Enzymol 101:20–78