web reference: Regier, J.C. 2005. Protocols, Concepts, and Reagents for preparing DNA sequencing templates. Version 3/14/06. www.umbi.umd.edu/users/jcrlab/PCR_primers.pdf

[The latest version of this document can be found at www.umbi.umd.edu/users/jcrlab/PCR_primers.pdf]

Jerome C. Regier Voice: 301-405-7679 E-mail: regier@umd.edu
Center for Biosystems Research, University of Maryland Biotechnology Institute,
Plant Sciences Building -- Rooms 5140 and 5142, College Park, MD 20742 USA

Protocols, Concepts, and Reagents for preparing DNA sequencing templates

---

1. Introduction and acknowledgments. This document describes laboratory protocols and PCR primers that are used in the Regier lab to amplify and then to sequence 25 genes across Lepidoptera and 62 genes across Arthropoda and near relatives. While successful amplification and sequencing may be taxon specific, the overall rate of success should generally be very high, based on our experience testing hundreds of diverse taxa. The procedures have been adapted from widely available protocols. The primers have been developed either together with our collaborators or else adopted from others. To summarize our approach, gene-specific PCR amplicons are generated starting from total mRNA that has been extracted from a wide taxonomic diversity of Lepidoptera or arthropods generally. To facilitate direct amplicon sequencing, we have aimed to identify orthologous genes with little or no intra-specific polymorphism and with expression patterns that are temporally and spatially broad. As systematists, we are interested in sequences that show taxon-specific variation. The genes described in this document display a wide range of variability and have been used over a

correspondingly wide taxonomic range. At present, this document contains little rate information on individual genes, although genes which are considered useful for resolving Mesozoic and more recent divergences (*i.e.,* within Lepidoptera) are identified. Rate information can be found in past and upcoming publications, and may eventually be incorporated into an updated version of this document.

Protocol development mostly occurred through collaborations with Charles Mitter (Dept. of Entomology, University of Maryland) and Jeffrey Shultz (Dept. of Entomology, University of Maryland), beginning in 1990. This document represents our first systematic effort to present a practical guide, separate from those summaries presented in our publications.

Primer development in the Regier lab aimed at identifying genes informative of lepidopteran phylogeny began in collaboration with Charles Mitter.

> [See Friedlander et al. 1992. Nuclear gene sequences for higher level phylogenetic analysis: 14 promising candidates. Syst. Biol. 41: 483-490.]
> [Recent publication: Regier et al. 2005. Phylogenetics of eggshell morphogenesis in *Anthereaea* (Lepidoptera: Saturniidae): Unique origin and repeated reduction of the aeropyle crown. Syst. Biol. 54: 254-267.]

Subsequently, primers that amplified across Arthropoda were developed for elongation factor-1, RNA polymerase II (subunit 1), and elongation factor-2 in collaboration with Jeffrey Shultz.

> [Starting with: Regier and Shultz 1997. Molecular phylogeny of the major arthropod groups indicates polyphyly of crustaceans and a new hypothesis for the origin of hexapods. Mol. Biol. Evol. 14: 902-913.]
> [Recent publication: Regier, Shultz, and Kambic. 2005. Pancrustacean phylogeny: Hexapods are terrestrial crustaceans and maxillopods are not monophyletic. Proc. R. Soc. B 272: 395-401.]

Then, in 2001 the laboratories of Regier and Cliff Cunningham (Duke University) began collaborating on arthropod phylogeny. To date, the very encouraging outcome of that collaboration has been that primers were developed that amplified 59 genes across Arthropoda and near outgroups, and that at least 20 of these genes also appeared useful within Lepidoptera. These primers are presented in this document and will also appear in upcoming joint publications.

Since 1992, primer development in the Regier lab has been dependent on generous funding from the National Science Foundation (Proposal numbers DEB-9212669, DEB-9509174, DEB-9629791, DEB-9981970, DEB-0075605, DEB-0120635, and DEB-0212910) and from the U.S. Department of Agriculture (Proposal numbers 93-37302-8981 and 9502086).

This undertaking also would not have been possible without major assistance from graduate students, postdoctoral fellows, and technicians -- both past and present. Former graduate students and postdocs (all joint with C. Mitter) include Brian Wiegmann (Department of Entomology, North Carolina State University), Quentin Fang (Dept. of Biology, Georgia Southern University), Timothy Friedlander (Gaithersburg, MD), Soowon Cho (Dept. of Plant Biology, Chungbuk National University, South Korea), Andrew Mitchell (Agricultural Scientific Collections Unit, New South Wales Department of Primary Industries, Orange, Australia), and Chris Desjardins (The Institute for Genomic Research). Austen Ganley, a postdoc with Cliff

Cunningham at the time, also played a vital role. The technical assistance of Zaile Du, Diane Shi, and Bernie Ball (Cunningham lab) is gratefully acknowledged. Other laboratories have independently pursued new genes and primers for studies of insect phylogeny, and they have greatly facilitated our own efforts. Brian Wiegmann (Dept. of Entomology, North Carolina State University), John Moulton (Dept. of Entomology and Plant Pathology, University of Tennessee), Andrew Brower (Oregon State University), Rob DeSalle (Dept. of Entomology, American Museum of Natural History), and Brian D. Farrell (Dept. of Organismic and Evolutionary Biology, Harvard University) deserve special acknowledgment.

2. <u>Specimen storage conditions.</u> Operationally, a successful specimen storage strategy is one that subsequently allows amplification of a specimen's nucleic acid sequences by minimizing their degradation during storage. Given the sensitivity of PCR, elimination of all degradation is unlikely to be necessary. The extent of degradation that is tolerable will depend on 1) amount of specimen, 2) sequence concentration (*e.g.*, mitochondrial DNA is generally more abundant than nuclear DNA), 3) sequence stability (*e.g.*, DNA is more stable than RNA), and 4) degree of primer matching (*e.g.*, nondegenerate, fully matching primers would be more effective than degenerate and/or partially mismatched primers). As some of these are unlikely to be known, our aim is always to minimize degradation as much as practical.

Upon extraction, specimens that have been live-frozen at -85°C provide satisfactory yield and quality of template even after storage for many (>10) years and, at least in some cases, are easier to extract and may provide higher yields of nucleic acids than those stored in 100% ethanol at -85°C. However, there are at least two problems: 1) Many specimens cannot be delivered live to a -85° freezer and, 2) if the freezer unexpectedly fails, the specimens' nucleic acids become highly susceptible to rapid degradation. Therefore, our routine approach is to place live specimens in 100% ethanol and, then, as quickly as possible, transfer them to -85°C. Sometimes, access to -85° may take several weeks. In such cases temporary storage at -20°C (*e.g.*, a "home" freezer) or at 0°C (*e.g.*, an ice cooler) is helpful, particularly at the beginning. Indeed, long-term specimen storage in 100% ethanol at -20°C may be as good as at -85°C. Other potentially useful techniques are 1) to rip open the specimen in order to enhance ethanol penetration of tissues and 2) to replace the 100% ethanol with fresh ethanol after a relatively short time (*e.g.*, a few hours). The former suggestion raises the potential problems of 1) cross-specimen contamination if multiple specimens are dissected without first cleaning the cutting tools, and 2) subsequent difficulty with species identification, so the particularities of the situation ought to be taken into account in deciding how best to proceed. As a rule-of-thumb, we add at least 15-volumes of ethanol relative to the estimated specimen volume and then replace the ethanol with fresh ethanol after a few weeks or when the specimen is first thawed for extraction of nucleic acids. When collecting in the field, access to any refrigeration may be impractical for hours or even days. At least for small specimens, storage in 100% ethanol at room temperature seems satisfactory (even up to a year), as may be the case for small parts (*e.g.*, legs) of larger specimens. Generally, we have not had success amplifying sequences from dead specimens kept at room temperature outside of ethanol, although rapid desiccation may be an acceptable alternative when ethanol and refrigeration are not available.

3. <u>Nucleic acid extraction protocol.</u> We use the SV Total RNA Isolation System (Promega, cat. # Z3100), which is designed to yield total RNA devoid of DNA. However, we have slightly

modified the manufacturer's protocol in order to recover some DNA along with the RNA. We do this to keep available the option of direct genomic amplification, even though we haven't done this for years. The few times we've directly compared DNA-PCR and RT-PCR, we've gotten better results with RT-PCR, so we stick with that. (See also section 5.) Our modified extraction also saves some time, and we haven't noticed a consistent difference in RT-PCR performance with and without DNA. The Promega kit comes with an abbreviated protocol sheet to follow during the extraction. We usually extract with up to six specimens at a time, and the protocol takes 30 minutes, excluding the time it takes for specimen preparation.

Supplies in addition to those in the kit (all sterile and RNase free): 1.5-ml polypropylene centrifuge tubes, pestles to fit centrifuge tubes (available from Kontes and Fisher Scientific), aerosol-barrier pipet tips, disposable Petri dishes, scalpel and disposable blades, 100% ethanol, Kimwipes, water bath at 70°C, timer

(Slightly) modified protocol:

1. For each extraction to be performed, remove 175 µl Lysis Buffer from the stock tube and add to a centrifuge tube. To the tube, add 3.5 µl of 2-mercaptoethanol (included with the kit) for each 175 µl Lysis Buffer. Mix. Quick spin. Place 175 µl of this mixture into individual extraction tubes.

2. Take one specimen at a time from the -85° freezer and place on ice. Using forceps, we usually transfer the specimen to a sterile plastic Petri dish containing 100% ethanol at room temperature. Working under a dissecting microscope (if necessary), we excise (with a scalpel) a portion approximately equal to one adult Drosophila. (*N.B.*: The final yield of nucleic acids can be quite low if too much tissue is extracted.) Since most of the genes we study are largely ubiquitously expressed, the particular tissue selected probably doesn't matter too much. We usually choose to extract the head and/or thorax, and, if possible, avoid the abdomen, where gut contamination might present problems. We've also had success with legs. Pick up the tissue with forceps, absorb the excess ethanol into a Kimwipe, and place the tissue into the centrifuge tube containing the lysis buffer + 2-mercaptoethanol. Using a pestle, homogenize / sheer / macerate the specimen until it is a fine suspension. You cannot over-grind. Some people have reported problems extracting specimens stored in ethanol. We haven't, but we grind long (*e.g.*, up to 5 minutes) and hard. The RNA is completely stable at room temperature in the lysis buffer, which is guanidinium-based, so once you're done, the homogenate can be set aside and you can start the extraction of another specimen. Shearing may also fragment genomic DNA, increasing its yield.

3. Add 350 µl of the RNA Dilution Buffer and mix by vortexing. (The instructions say to mix gently, but again, we're trying to fragment the DNA.) Place in a water bath for 3 minutes.

4. Centrifuge for 10 minutes at 12,000-14,000 × *g*. Transfer the cleared lysate to a fresh tube, avoiding the pelleted material against the wall of the tube.

5. Add 200 µl 100% ethanol to the cleared lysate and mix by vortexing quickly. Transfer the mixture to the Spin Basket Assembly and centrifuge at 12,000-14,000 × *g* for one minute. Discard the eluate.

6. Add 600 µl RNA Wash Solution (with ethanol added, per the kit's instructions) to the basket. Centrifuge at 12,000-14,000 × *g* for one minute. Discard the eluate.

7. Repeat step 6 but with 250 µl RNA Wash Solution. (Note that the DNase incubation step and the subsequent multiple washes have been eliminated.)

8. Add 100 µl Nuclease-Free Water to the Spin Basket membrane. Centrifuge at 12,000-14,000 × *g* for one minute. The nucleic acid is in the eluate. Store it at -85°C.

4. <u>Guidelines for PCR primer design.</u> Our primers vary from 17-32 nucleotides in length, excluding the 5' *M13* tails (see next paragraph), but most are 17-21 nucleotides in length. Longer lengths are preferred, but benefits may be counterbalanced by the effects of increasing degeneracy. Primers are designed by examining aligned protein sequences and are almost always made completely degenerate, regardless of whether or not a particular synonymous substitution has (yet) been observed. For 17-mers, we oftentimes have little success in PCR using primers greater than 128-fold degenerate. As a result, in scanning across aligned protein sequences in search of suitable primer sites, we typically avoid arginine and leucine residues (except perhaps at the 3' end, where the third codon position can be excluded) because of their 8-fold degeneracy. Similarly, fourfold degenerate codons (*e.g.*, glycine and valine) are kept to a minimum, leaving twofold (*e.g.*, glutamic acid and lysine) and nondegenerate (*e.g.*, methionine and tryptophan) codons as the preferred targets. According to lore, primer mismatches are least well tolerated at the 3' end, so we try to keep degeneracy particularly low near the 3' end. If fortunate enough to identify a useful 9-residue or longer stretch, we may design two partially overlapping primers and use the inner primer in a hemi-nested reamplification strategy (see section 6). Alternatively, a longer single primer may function more effectively, possibly avoiding a second round of PCR, so these alternatives need to be tested. However, identifying suitably long stretches is uncommon, so much so that we never bother worrying about theoretically complicating secondary structures, such as hairpins, and biased base compositions, but rather synthesize and test the primers directly. Deoxyinosine is occasionally inserted into primer sequences on the hope that it binds relatively nonselectively, thereby reducing the complexity of the primer mixture. We have never systematically tested this.

To the 5' ends of all primers, we add 18-nt long *M13* sequences, one (called *M13REV*) for all forward primers and another (called *M13(-21)*) for all reverse primers. (The *M13REV* and *M13(-21)* sequences are listed in section 10.) Initially, we did this so that PCR amplicons could be directly sequenced by dye-terminator technology using the same M13 primers, rather than gene-specific degenerate PCR primers. This strategy has been effective. More recently, we have discovered (Regier & Shi 2005. BioTechniques 38: 34-38) that addition of 5' tails can dramatically increase the yield of PCR amplicons, and that this effect is specific to degenerate primers, so addition of 5' tails seems to be a doubly effective strategy.

Currently, our primers are synthesized by Integrated DNA Technologies, Inc. (www.idtdna.com). Originally, we used primers made by Oligos etc. (www.oligosetc.com). These were consistently good, but Oligos' prices became too high. Twice (out of many dozens of orders), IDT has sent us bad primers, which we discovered only after much frustration. As a result, we tried primers made by GeneProbe Technologies (www.geneprobetech.com), but they

were terrible, so we're back with IDT. As the coupling efficiencies are lower at IDT than at Oligos, we order PAGE-purified primers, which adds to their expense and greatly reduces their yield. For example, we just received 5 nmoles of a PAGE-purified primer that was synthesized at the 250 nmole scale. We are certain that PAGE purification is not necessary for successful PCR and sequencing in most cases, but we're uncertain whether or not it might help in some borderline circumstances, so we continue ordering PAGE-purified primers.

5. RT-PCR protocol. For all primers, we use employ RT-PCR rather than direct gene amplification in order to have a good estimate of the amplicon length (thereby facilitating gel isolation) and to avoid having to sequence introns, which are too variable for most of our phylogenetic studies. Prior knowledge of orthologs from other taxa allows us to predict, usually with great accuracy, the length of an RT-PCR fragment from a novel species. The sensitivity of RT-PCR also compares favorably with DNA-PCR, at least when amplifying mRNAs of widely- and moderately-to-abundantly expressed, single-copy, protein-coding genes.

We will start by describing our basic RT-PCR reaction and then discuss various modifications.

Part 1: The RT reaction

reaction ingredients	10 μl reaction	20 × stock
25 mM $MgCl_2$	2 μl (5 mM)	40 μl
10X PCR buffer II*	1 μl (1 ×)	20 μl
10 mM each of dATP, dCTP, dGTP, & TTP	2 μl (2 mM each)	40 μl
H_2O	1.75 μl	35 μl
RNase inhibitor (20 u/μl)	0.5 μl (10 units)	10 μl
reverse transcriptase (50 u/μl)	0.5 μl (25 units)	10 μl
reverse primer (20-135 pmoles/μl)	1.25 μl (25-168.75 pmoles)**	25 μl
template	1.0 μl	

thermocycler conditions: Run RT program.
* , 10 × PCR buffer II: 100 mM Tris-HCl, pH 8.3 (at 23°C), 500 mM KCl
** , see comment 4 below

Comments

1. The RNase inhibitor (N808-0119) and MuLV reverse transcriptase (N808-0018) are purchased from Applied Biosystems. The $MgCl_2$ and 10X PCR buffer II (N808-0010) are supplied with the AmpliTaq DNA polymerase (see below), also from Applied Biosystems. The RNase inhibitor and RT (+ Taq and Taq antibody, see below) are stored in a freezer block at -20°C. RNase inhibitor and RT (+ Taq and Taq antibody, see below) are supplied in 50% glycerol, and their high viscosity makes for inaccurate pipetting. To minimize this, we prewet the pipet tip with the protein solution (by pipeting in and out one time) and then dispense 5-10% larger volume than indicated in the recipe. As least in one case, we have observed that the activity of the RT noticeably decreased (perhaps twofold) after ~9 months in cold storage; the

stated expiration date is at least 12 months. As a result, we now store RT only six months before purchasing more.

2. The dNTPs are purchased from Promega. We used to use dNTPs from Sigma, but a later batch proved inferior over a range of concentrations, and we switched. We discovered that the RT reaction is sensitive to the supplier, the batch, and the concentration of dNTPs, so we titrate each batch of dNTPs (0.5 mM - 2 mM final concentration for each dNTP) and compare with the previous batch. The standard RT protocol in the literature recommends a final concentration of 1 mM for each dNTP, which we have found optimal with some batches, but our current batch works best at 2 mM. Either more or less gives inferior results. This effect of dNTP concentration on the RT reaction does not appear to be common knowledge, and, as the result can be quite dramatic, it is worth the effort to standardize reaction conditions each time a component is changed.

3. The H_2O is salt-free, organic-free "Millipore" water that has been further treated with diethylpyrocarbonate and then autoclaved, according to standard procedures.

4. In principle, the amount of primer required might be exptected to be proportional to the degree of primer degeneracy. Other than cost, we have not observed deleterious effects from adding excess primer. Primer stock solutions are 20 pmole/μl for *EF-1α, EF-2*, dopa decarboxylase, and enolase; 135 pmole/μl for period, and 40 pmole/μl for RNA polymerase II (largest subunit), *CAD*, wingless, and the other 59 genes at the bottom of the list in section 10. Recently, we have started adding 1.875 μl of the *DDC* primers (20 pmoles / μl stock) to the RT and PCR (see below) reactions, rather than 1.25 μl. In the future, we will start making our *DDC* primer stocks at 40 pm/μl.

5. The template concentration in our total nucleic acid preparations is largely unknown but can affect the amplicon yield. If the template is not limiting, we will use 2.0 μl (and also eliminate the additional H_2O from the recipe). If it is very limiting, we have used as little as 0.1 μl (and added correspondingly more H_2O). Larger amounts of template sometimes result in reduced, not higher, amplicon yields.

6. When amplifying many templates with the same primer, we make a stock solution of everything except the template, and then aliquot that stock into the separate PCR tubes before adding the templates individually. In the above recipe using a 20× stock, 9 μl of stock would first be added to a PCR tube, followed by 1 μl of template. The contents would be gently mixed and quick spun, and then placed in the thermocycler (described below). If we are using multiple primers in single-day's experiment, then we may make separate stocks for each primer, or else make one large stock without primer, and then individually add that plus the template to the PCR tubes.

7. RT-PCR reactions are carried out in a laminar-flow hood, when available. Otherwise, components are combined on top of a clean lab bench that has been further cleaned with RNase AWAY (cat. # 7005, Molecular BioProducts).

8. PCR tubes -- both 8-tube strips + caps and individual 200-□l tubes -- are purchased from Applied Biosystems. Pipet tips are all sterile, RNase-free, and contain aerosol barriers.

9. Before assembling the reaction components, the components themselves are thawed, thoroughly mixed, quick-spun, and stored on ice. Remember that the RTase, RNase inhibitor, Taq DNA polymerase, and Taq antibody (the latter two described below) should be mixed by tube inversion or gentle "flicking" with the finger but should not be vortexed; this could denature them. Prior to dispensing components into individual PCR tubes, the tubes are placed in aluminum 96-well storage racks (cat. # 420094, Stratagene) that have been prechilled to 0°C.

10. We currently use thermocyclers from MJ Research (model PTC-200). The four steps in the *RT* program are as follows:

42□C, 35 min;
99□C, 5 min;
4□C, for ever;
End

Part 2: The PCR Reaction

This section will describe the second part of the RT-PCR reaction. (It is **not** a guide to a stand-alone PCR reaction, which is described below.)

reaction ingredients	50□l reaction	20□ stock
RT reaction (from Part 1)	10 □l	
25 mM $MgCl_2$	3 □l (□ 2.5 mM)	60 □l
10X PCR buffer II	4 □l (1□)	80 □l
H_2O	31.25 □l	625 □l
AmpliTaq (5 u/□l) + Taq antibody	0.5 □l (1.25 units)	10 □l
forward primer (20-135 pmoles/□l)	1.25 □l (25-168.75 pmoles)	25 □l

thermocycler conditions: Run *TOUCHDN* program.

Comments

1. Combining Taq polymerase with a commercially available antibody provides a convenient (but not cheap) means of doing hot-start reactions. All of our RT-PCR and PCR reactions (described below) use this hot start approach. To prepare this solution, AmpliTaq DNA polymerase (5 u/□l, cat. # N808-0156) is gently but thoroughly mixed with an equal volume of TaqStart antibody (7 □M, cat. # 5400-2, Clontech), incubated at 23□C for 10 minutes, and then stored at -20□C. This inactivated form of Taq is stable for many months.

2. Over the years, we have tested RTase and Taq in combined RT-PCR reactions from various companies but have not found consistent improvement over the Applied Biosystems' reagents.

3. As for the RT reaction, a stock solution -- this time consisting of $MgCl_2$, 10X PCR buffer II, additional H_2O, AmpliTaq + Ab, and forward primer (optional, depending on number of forward primers, see Part 1, comment 6) -- is prepared. This can be done while the RT reaction is in the

thermocycler. In the above recipe using a 20X stock, 40 µl of the stock would be added to each PCR tube containing the RT reaction. The tubes are then placed in the thermocycler (described below). We usually do not worry about thorough mixing or quick spinning because of thermal mixing during the PCR reaction itself.

4. The amount of forward primer added in the PCR part should equal the amount of reverse primer added in the RT part.

5. We use a Touchdown protocol because the specific primers in our degenerate mixture anneal over an unknown range of T_m's, particularly during the first round of amplification. The 11 steps in the *TOUCHDN* program (38 cycles total) are as follows:

94°C, 30 sec;
55°C, 30 sec, -.4°/cycle
72°C, 1 min + 2 sec/cycle
Goto 1, 24X
94°C, 30 sec
45°C, 30 sec
72°C, 2 min + 3 sec/cycle
Goto 5, 12×
72°C, 10 min
4°C, for ever
End

Modifications

1. When an RT-PCR reaction fails and we're confident that it's not due to bad reagents or human error, we usually assign tentative blame to a poorly matched primer site. If alternative sites are available, we will try them. Another approach that sometimes works is to prime the RT reaction with $dT_{12\text{-}18}$ (20 pmole/µl stock) instead of the reverse gene-specific primer, and to place both forward and reverse gene-specific primers in the PCR reaction. While $dT_{12\text{-}18}$ priming of the RT reaction sometimes yields better results (and can be somewhat simpler to set up, since all reactions are identical except for template), gene-specific priming yields better results more consistently, so that is our standard approach. The modified thermocycler program (called *RT(DT)*) for using $dT_{12\text{-}18}$ is described below:

23°C, 10 min
42°C, 45 min
99°C, 5 min
4°C, for ever
End

2. We occasionally employ other approaches after encountering failed amplifications. If no amplicon band of the desired size is visible on agarose gels, we 1) increase the template concentration, 2) increase the Mg^{+2} concentration in the PCR part of the reaction (up to 4 mM), or 3) decrease the annealing temperature range in the *TOUCHDN* thermocycler program from 55× 45°C to 50× 40°C. Conversely, if too many bands are present near the expected length, we decrease the Mg^{+2} concentration in the PCR part of the reaction (to 2 mM) or increase the

annealing temperature range in the *TOUCHDN* thermocycler program from 55°– 45°C to 60°– 50°C or 65°– 55°C.

6. <u>Hemi-nested and M13 PCR protocols.</u> Not infrequently, RT-PCR bands will be insufficiently abundant or pure to sequence directly after gel purification (described in sections 7, 8). In this case, we proceed to a nested (or, more often, hemi-nested) PCR strategy to increase purity and yield.

<u>reaction ingredients</u>	<u>50µl reaction</u>	<u>20× stock</u>
25 mM $MgCl_2$	4 µl (2 mM)	80 µl
10X PCR buffer II	5 µl (1×)	100 µl
H_2O	36.75 µl	735 µl
10 mM each of dATP, dCTP, dGTP, & TTP	1 µl (0.2 mM each)	20 µl
AmpliTaq (5 u/µl) + antibody	0.5 µl (1.25 units)	10 µl
forward primer (20-40 pmoles/µl)	1.25 µl (25-50 pmoles)	25 µl
reverse primer (20-40 pmoles/µl)	1.25 µl (25-50 pmoles)	25 µl
template	0.25 µl	

thermocycler conditions: Run *3STEP* program.

<u>Comments</u>

1. A dNTP concentration of 0.2mM each consistently gives good amplicon yields, unlike that described for the RT-PCR reaction (see comment 2, section 5).

2. The template concentration doesn't seem to matter (*e.g.*, 0.1-1 µl) if one is reamping a visible, gel-isolated band. If the band is invisible, it might be safer to increase the template volume to 1-2 µl.

3. The concentration of the *M13* primer stocks is 20 pmoles/µl. Period primers, which during the RT-PCR reaction, are 135 pmoles/µl, work well in reamps at 40 pmoles/µl.

4. When performing a hemi-nested reamp -- in which only one new primer is used -- the "old" primer is oftentimes replaced with its *M13* sequence alone, both for convenience and cost savings.

5. The thermocycler program labeled *3STEP* is described below:

94°C, 30 sec
50°C, 30 sec
72°C, 1 min, +2 sec/cycle
Goto 1, 21×
72°C, 10 min
4°C, for ever
End

7. <u>Preparative gel electrophoresis and DNA band excision.</u> PCR amplicons are excised from low-melting point agarose gels of concentration 1.0-1.1% (w/v). On a typical 20-cm wide, 25-cm long, 450-ml (~0.9 cm thick) gel, we insert three combs. Individual combs either have 9 wells (max. vol. / well, ~32 l; comb cut from a 21-well comb with 14 teeth removed) or, less commonly, 7 wells (max. vol. / well, >100 l; comb cut from a 36-well comb with 27 teeth removed). Reducing the number of wells per comb is designed to minimize cross-sample contamination, but has the added benefit of making the combs easier to remove without tearing the fragile prep gels.

Gel electrophoresis buffer is the standard 1X TAE buffer with 0.5 g/ml ethidium bromide. This same buffer is used to dissolve the agarose when pouring the gel. Many labs appear to leave EtBr out of the reservoir buffer (but not the gel!), but our experience is that this leads -- after electrophoresis -- to zones of different fluorescence levels in the gel, and may complicate band identification. Gels (prep and analytical) can be stored for days in reservoir buffer. If only a portion of the gel is used, the remainder of the gel is stored in fresh reservoir buffer. Reusing reservoir buffer can result in fainter bands.

PCR bands are visualized after transferring the prep gel from the gel tray to plastic wrap resting directly on a uv transilluminator (λ_{mean} = 302 nm). Individual bands are excised using a sterile spatula, one per band. Prolonged exposure of the gel to ultraviolet light can seriously degrade the quality of DNA sequencing, and should be kept to less than 2-3 minutes total. (Of course, this depends on the strength of the uv bulbs.) We routinely expose only a few gel lanes-worth of gel at a time, keeping the overall exposure of each band < 2 minutes. The phenotype of a sequencer chromatogram displaying overexposed DNA is a rapid reduction in fluorescent signal strength with length of sequence, *i.e.*, bad tailing rather than uniform intensity of signal across the chromatogram, thereby, reducing the amount of readable sequence.

8. <u>DNA extraction from low-MP agarose and other means of primer removal, and semi-quantification of gel-isolated template.</u> Excised gel pieces are placed in a 1.5-ml microcentrifuge tube, quick spun, and stored at 4C up to 16 hr or at -20C for longer periods. Each gel piece should be 300 l. To begin the extraction, the tubes with gel pieces are incubated at 70C for 7 minutes to melt the gel. Then, 1-ml of resin (Wizard PCR preps DNA purification resin, Promega cat. # A7181) is pipetted into each tube, the contents are thoroughly mixed by inversion over 20 seconds, and further incubated at 28-30C for 2 minutes. The contents are then pipetted into the barrels of 3-ml syringes (*e.g.,* BD 3-ml syringe, Ref. # 309585, available from Fisher Scientific) attached to minicolumns (Wizard minicolumns, Promega cat. # A7211), which in turn are attached to a vacuum manifold (Vac-Man Laboratory vacuum manifold, Promega cat. # A7231). A moderate vacuum (*e.g.*, aspirator is fine, but we use an old vacuum pump placed inside a chemical fume hood) is applied until the liquid contents are gone. The vacuum is released. Then, 2 ml 80% isopropanol is added to each barrel + minicolumn and a vacuum reapplied. The vacuum is released and the syringe barrels are discarded. The minicolumns (still attached to the manifold) are washed once more with a few hundred microliters of 80% isopropanol. The minicolumn contents should not be overdried during the vacuum steps. After this, the minicolumns are removed from the manifold, inserted into empty, capless microcentrifuge tubes and spun at 10,000 *g* for 2 minutes to remove the last bits of isopropanol. The minicolumns are then transferred to a clean, labeled tube (with cap still

attached). 50 µl nuclease-free distilled water is pipetted into each minicolumn. After 60 seconds, the minicolumn / tube assemblage is spun at 10,000 × *g* for 20 seconds. The DNA has now been eluted into the tube and should be stored at -20°C. Normal recoveries are 50-75%, with more agarose at the start yielding less DNA at the end.

After gel isolation, we combine 10 µl of the gel-isolated template with 5 µl of (3× PCR buffer II + 7.5 mM $MgCl_2$ + 30% glycerol + bromphenol blue) and run this on an analytical gel, along with an appropriate DNA ladder standard (*e.g.*, "100 bp DNA Ladder," New England BioLabs, cat. no. N3231L). The relative fluorescent intensities of the template and DNA standard bands permit semi-quantification of template, which is necessary for planning the future DNA sequencing reactions.

Promega sells a kit (Wizard SV Gel and PCR Clean-UP System, cat. # A9281) that contains most of the components listed above, but we've been buying selected components for years. The technical bulletin for the entire protocol is cat. # TB308. Promega items are sold through Fisher Scientific. The manifold mentioned above works well for scaling up to 20 minicolumns. Doing more than that at one time is probably not too wise, as the low MP agarose can resolidify over time, and this lowers the DNA recovery. Vacuum adapters (Promega cat. # A1331) are useful for those with big fingers. Also, the stopcocks seem to develop leaks moderately often, and their replacement (stopcock one-way Luer Lock, Promega cat. # A7261) will keep the vacuum at acceptably high levels.

Sometimes, we skip gel isolation (see section 9 for rationale) and, instead, simply remove salts and primers from the preceding (RT-) PCR reaction; we call this *direct purification*. To do direct purification, we have found the QuickStep 2 PCR purification kit from Edge BioSystems (cat. # 33617) most convenient. We now buy the component cartridges (Performa DTR gel filtration cartridges, Edge BioSystems cat. # 42453) and resin (QuickStep 2 SOPE resin, Edge BioSystems cat. # 72418) rather than the kit. Here's the general protocol:

1. Add 1/5 volume (relative to the volume of the PCR reaction) of SOPE resin directly to the PCR reaction mixture and mix well. Let the suspension stand at room temperature.

2. Centrifuge a gel filtration cartridge for 3 minutes at 750 × *g*. Transfer the cartridge to a clean microcentrifuge tube.

3. Carefully pipet the COPE/PCR reaction mixture onto the center of the gel bed surface of the prepared gel filtration cartridge.

4. Centrifuge for 2 minutes at 750 × *g*. The eluate contains the purified DNA.

9. <u>General guidelines for dealing with difficult-to-amplify templates.</u> As described in sections 3 and 5, we always begin the preparation of a new sequencing template with an RT-PCR reaction. If its yield is sufficient and its length < 900 bp, the amplicon is directly sequenced after gel purification. However, this best-case scenario is infrequent. Typically, the amplicon yield is too low (perhaps even invisible) or there are other bands closely positioned to the desired one or the desired band is too long to sequence in its entirety from both ends.

If no band is visible after gel analysis of an aliquot (we visualize 1/6th of the total reaction on an analytical gel) of the RT-PCR reaction and no hemi-nested primer site is available, then we either try to improve the reaction or we're finished.

If a faint amplicon band is visible after the RT-PCR reaction and no hemi-nested primer site is available, then we gel isolate the desired band, reamp it with *M13* primers, re-gel-isolate the band, and sequence it.

If no band is visible after the RT-PCR reaction and a hemi-nested primer site is available, we gel isolate the invisible "band," using as a length marker the same PCR fragment from another taxon that yielded a visible band. If no markers are available, then we use direct purification, but we consider this a second-best option relative to gel isolation. Following gel isolation of the invisible band, we do a PCR reaction using a hemi-nested primer and re-gel-isolate the band.

If a faint band is visible after the RT-PCR reaction and a hemi-nested primer site is available, we gel isolate the band. Following gel isolation, we either reamp it with *M13* primers and re-gel-isolate the band OR we reamp it with a hemi-nested primer and re-gel-isolate the band. The choice depends on the amplicon's length and purity.

If a major band of the expected length results after a hemi-nested or *M13* reamp, the band is gel isolated and sequenced.

If a faint band of the expected length results after hemi-nested PCR, the band is gel isolated, reamped with M13 primers, re-gel isolated, and sequenced.

If no band results from the hemi-nested reamp, and no band was visible in the prior RT-PCR reaction, then we're finished, unless there are other nested primer sites or other PCR conditions to try.

If no band is visible in the hemi-nested reamp, but a band was visible in the RT-PCR reaction, then the gel-isolated RT-PCR band is reamped with M13 primers, re-gel isolated, and sequenced.

If no band or a weak band results after an *M13* reamp on a visible -- either faint for bright -- RT-PCR band, then we usually discover that the RT-PCR band is not the correct gene.

Our general rule-of-thumb is that invisible "bands" are PCRed one time only. An informed guess is that approximately half of the visible bands that result from PCRing of invisible bands twice are false positives. If you're really desperate, it may be worth the try, but you should expect to spend time at the computer convincing yourself that you have the right sequence. In our experience, one-time PCRing of invisible bands seldom (*i.e.*, perhaps <1%) results in false positives, but of course great care to avoid contamination is assumed.

The reader may question whether all the gel isolations we routinely do (up to three for each sequencing template) are really necessary. In most cases, the answer is, "We don't know." However, we do know that gel isolation never yields less than the best results, so when we decide that we are unable to amplify a particular gene from a particular taxon, we at least feel that we have done our best.

10. <u>List of PCR primer sequences (+ acknowledgments) used to amplify lepidopteran and (more generally) arthropod mRNAs and cDNAs, plus commentary.</u>

<u>General comments (READ ME 1st)</u>

1. Identifying useful protein-coding nuclear genes for phylogenetic inference in Lepidoptera and across Arthropoda has been a goal of Regier and Charles Mitter since 1990, Regier and Jeff Shultz since 1995, and Regier and Cliff Cunningham since 2001. Of course, there are researchers in other laboratories who have developed genes for arthropod phylogeny. Some of these genes are in use in our lab, and their initial developers are acknowledged below (see also section 1).

2. Forward primer sequences (labeled *F* in our nomenclature) bind to the anti-sense strand of a ds-cDNA and, thus, generate a sense sequence. Reverse primers (labeled *R* in our nomenclature) bind to the sense strand of a ds-cDNA and, thus, generate an anti-sense sequence. **All of the PCR primers designed and/or used in our lab are bipartite, containing a specific *M13* sequence (18 nt in length) at the 5' end and a gene-specific, degenerate sequence (17-32 nt in length) at the 3' end (see section 4). For simplicity and clarity of presentation, *M13* sequences are removed from all primer sequences described below. We emphasize that their inclusion is very useful in practice (see section 4).**

The *M13* sequence in all *F* primers is called *M13REV* and its sequence is:
`CAG GAA ACA GCT ATG ACC`.

The *M13* sequence in all *R* primers is called *M13(-21)* and its sequence is:
`TGT AAA ACG ACG GCC AGT`.

3. A few primers contain the modified base deoxyinosine (codename: *I*).

4. Amplification schemes are described below by the primer pairs involved, with nested primer pairs (if any) separated from the RT-PCR primer pairs by a forward slash, *e.g.,* 30F_41.21R / 45.71F_53.5R means that the primer pair 30F and 41.21R was used in an RT-PCR reaction and that this was followed (after gel isolation) by a nested reamp using 45.71F and 53.5R. Not all taxa will necessarily require the nested reamp (see section 9).

5. Primers for the first seven genes listed below -- *DDC*, enolase, period, *PEPCK, CAD*, wingless, and *EF-1*□ -- are used in our lab for lepidopteran phylogenetics and have been developed by ourselves and others. The first six genes have a useful number of non-synonymous, as well as synonymous, changes for higher-level phylogenetic inference in Lepidoptera. *EF-1*□ is used at lower levels within Lepidoptera, where almost all of the changes are synonymous. Primers for the next two genes -- *EF-2* and *Pol II* -- plus *EF-1*□ were developed in our lab for use in higher-level arthropod phylogeny. We have largely restricted our analyses of these three genes to their non-synonymous changes.

More recently, and in collaboration with Cliff Cunningham and co-workers at Duke University, we have developed primers for 59 protein-coding nuclear genes that amplify across arthropods. They are listed immediately after *Pol II*, beginning with *acc*. We are currently testing the utility of these genes for higher-level arthropod phylogeny. In addition, and based on

sequencing of these 59 genes for three diverse Lepidoptera and subsequent assessment of their higher-than-average inter-species variability, we suggest that the following 20 "arthropod" genes may also prove useful for lepidopteran phylogenetics: *acc, aspec* (11F_15R / 11F_12R)*, 109fin, 113fin, 192fin, 197fin, 262fin, 267fin, 270fin, 3006fin, 3007fin, 3017fin, 3070fin, 36fin, 40fin, 42fin, 44fin, 69fin, 8053fin,* and *8091fin*. (*267fin* encodes *CAD* and completely overlaps the region covered by our "lepidopteran" *CAD* primers listed above, although one of the two primers is different from any of those listed for the "lepidopteran" *CAD* primers.) Regier, Mitter, and colleagues plan to test their phylogenetic utility across Lepidoptera in the future.

6. For the first eight genes, primers locations are identified relative to the 5'-most primer site on the sense strand of the ds-cDNA. For the remaining genes, their PCR-amplified fragments are listed by approximate total length in nucleotides rather than by the relative locations of the primers.

dopa decarboxylase (*DDC*)

We have used DDC primers exclusively in Lepidoptera and near outgroups. The 1.7sF_4sR fragment has been amplified across Lepidoptera. The outermost fragments are largely untested outside Macrolepidoptera. Our first *DDC* primers were published in Fang et al. 1997. Syst. Biol. 46: 269-283.

location	name	primer sequence (-M13 tail)
1	1.1vF	GAY TAY ATY RCR GAR TA
28	1.2F	GAR AAY ATY AGA GAY AGR CAR GT
259	1.7sF	GCH TGY ATY GGI TTY WCN TGG AT
271	1.8R	CAT NAC NAC YTC IAR YTC IGT RCA
292	1.9sR	CAT YTG RCC BAR CCA RTC IAD CAT
317	1.9sF	ATG HTI GAY TGG YTV GGY CAR ATG
766	3.2sF	TGG YTI CAY GTI GAY GCN GCN TAY GC
829	3.3sR	CCA YTT RTG NGG RTT RAA RTT RAA
854	3.3sF	TTY AAY TTY AAY CCN CAY AAR TGG
967	4sR	GGD ATY TGC CAR TGH CKR TAR TC
1309	7.5sR	TCC CAN GAN ACR TGV ATR TC

The typical strategy for amplifying *DDC* is 1.1vF_1.9sR / 1.1vF_1.8R; 1.7sF_4 sR/ 1.7sF_3.3sR; and 3.2sF_7.5sR / 3.3sF_7.5sR.

enolase

Enolase primers have been previously used in beetle phylogenetics (see Farrell *et al*. 2001. Evolution 55: 2011-2027). Based on available GenBank sequences, we developed other primers (not shown) that amplified enolase across arthropods, but we oftentimes found highly divergent paralogs co-amplifying. Now, with new primers defined (shown below), we use enolase within Lepidoptera only, where no paralogs have (yet) been detected.

location	name	primer sequence (-M13 tail)
1	28LF	AAC CCC ACH GTK GAG GTN GAY YTG GTD ACH GA
385	156LF	TTY AAY GTB ATY AAY GG
643	241LR	GCD ACR TCC ATD CCR AT

661	248LF	ATY GGH ATG GAY GTH GC
949	344LR	CCR ATY TGR TTV ACY TT
1090	391LR	GTY TTR ATC TGR CCR GT
1135	406LR	ATC TGG TTR TAY TTD GC

The typical strategy for amplifying enolase is 28LF_406LR followed by nested reamps.

period

We have used period only in Lepidoptera, specifically, Macrolepidoptera. Period evolves faster than any others listed here, and it is the hardest to amplify successfully. Our period primers were first published in Regier et al. 1998. Mol. Biol. Evol. 15: 1172-1182.

location	name	primer sequence (-M13 tail)
1	177sF	GTN ATH TCN ATG CAY GAY GG
62	197sF	GGN TTY CCN AAR GAY ATG TGG
493	341sF	TAY YTN GGN TAY YTN CCN CAR GA
658	397nR	GAC CAN GGR TTD AYR AA
676	402nF	TTY RTH AAY CCN TGG TC
1018	514nR	TAN TCR TGR TGN GGN GAD RT
1021	515sR	TRN TCR TGR TGN GGN GA
1072	532sR	TCR TTR TAR TTN ARY TGR TTR TA

The typical strategy for amplifying period is to generate by RT-PCR 177sF_514nR, 177sF_515sR, 177sF_532sR, 197sF_514nR, 197sF_515sR, and 197sF_532sR, and then to use the most promising ones for nested reamps. It is a tedious approach, but, as we said above, period is difficult to amplify and perseverance is required.

Phosphoenolpyruvate carboxykinase (*PEPCK*)

We have used *PEPCK* only in Lepidoptera, primarily for resolving basal clades. Our *PEPCK* primers were first published in Friedlander *et al*. 1996. Mol. Biol. Evol. 13: 594-604.

location	name	primer sequence (-M13 tail)
1	284dF	GAG GGC TGG CTR GCM GAR CAY ATG
91	18.5dF	TGT GGN AAR ACC AAY YTG GCC ATG
151	19.5dF	GGN GAY GAY ATI GCB TGG ATG
326	20.5dF	GGI GTI TGG TGG GAR GGI ATG G
405	21dNR	CAI AAY CTI GAR TTI GGR TGN GC
632	511dR	GGM CGC ATT GCR AAY GGR TCR TGC AT
730	22.5R	GAA CCA RTT RAC RTG RAA GAT C

Successful amplification has been more variable with this gene than others. The largest consistently amplifiable *PEPCK* fragment was obtained with 284dF_511dR.

CAD (pyrimidine biosynthesis)

CAD primers and their utility for dipteran phylogenetics were first described by Moulton and Wiegmann, 2003 (Mol. Phylog. Evol. 31: 363-378). Based on available GenBank sequences, we

have developed *CAD* primers that work well within Lepidoptera. Serendipitously, it turned out that our 46F primer (see below) exactly matched Moulton and Wiegmann's 54F. Also serendipitously, *CAD* turned out to be one of the 59 genes selected in our screen of genes useful for deep-level arthropod phylogeny (see gene *267fin* below), and primer *267fin*1F exactly matched our 356F. The other two *CAD* primers identified in the arthropod gene screen are *267fin2F* and *267 fin3R*.

location	name	primer sequence (-M13 tail)
1	46F	GTN GTN TTY CAR ACN GGN ATG GT
709	287nR	TTR TGN CCN CKR TTR CCR TA
736	295nF	TAY GGY AAY MGN GGN CAY AA
781	309R	TCN ACN GCR AAN CCR TGR TTY TG
904	350R	RTG YTC NGG RTG RAA YTG
922	356F (=267fin1F)	CAR TTY CAY CCN GAR CA
1126	267fin2F	GCN GGN GAR TTY GAY TA
1339	496F	CAR ACN GCN YTN AAY TGY GG
1582	576R	TCN TCY TCR TTR TTN GCR AA
1603	582F	TTY GCN AAY AAY GAR GAN GA
1672	606nR	ACN ACY TCR TAY TCN ACY TCY TTC CA
1696	613F	TGG AAR GAR GTN GAR TAY GAR GT
1726	267fin3R	TTY TCC ATR TTR CAN AC
1873	673R	GCR TAY TGN AYR TTR CAY TC
1894	681F	GAR TGY AAY RTN CAR TAY GC
2080	743nF	GGN GTN ACN ACN GCN TGY TTY GAR CC
2200	782R	GCY TTY TGR AAN GCY TCY TCR AA
2224	791F	TTY GAR GAR GCN TTY CAR AAR GC
2734	963R	GCR CAC CAR TCR AAY TC
2752	970F	GAR TTY GAY TGG TGY GC
2929	1028R	TTR TTN GGN ARY TGN CCN CCC AT

The typical strategy for amplifying *CAD* is to generate the following by RT-PCR: 46F_350R, 295nF_267fin3R, 496F_782R, and 743F_1028R. Then, after gel isolation a series of nested reamps are performed.

wingless

Wingless primers (Wg1aF, Wg2aR) and their utility for insect phylogenetics were first described by Brower and DeSalle, 1998 (Insect Mol. Biol. 7: 73-82). To these, we added M13 tails. More recently, we have designed modified primers (wg1sF, wg2sR / wg2nR) to amplify taxa that do not work with the original ones.

location	name	primer sequence (-M13 tail)
1	wg1sF	GAR TGY AAR TGY CAY GG
10	Wg1aF	GAR TGY AAR TGY CAY GGY ATG TCT GG
412	Wg2aR	ACT ICG CAR CAC CAR TGG AAT GTR CA
418	wg2sR	ACY TCR CAR CAC CAR TGR AA
421	wg2nR	ACY TCR CAR CAC CAR TG

The strategy for amplifying wingless by RT-PCR is Wg1aF_Wg2aR. If this doesn't work, we originally tried wg1sF_wg2sR, and this was generally successful. However, when using wgs2R, we have now observed that poor-quality reads may result from the R-primer end (but not the F-primer end) when the nucleotide that is immediately 5' to the wg2sR primer sequence is "T." Therefore, we will now shift to use of wg1sF_wg2nR, although the effectiveness of this new primer pair has yet to be tested.

elongation factor-1□ (*EF-1□*)

Our original *EF-1□* primers (described in Cho et al. 1995. Mol. Biol. Evol. 12: 650-656) were used specifically in Lepidoptera and with DNA-PCR rather than RT-PCR, since Lepidoptera (but not some other orders of insects) lack introns over the length encompassed by these primers. Since then, we have been sequencing *EF-1□* from a wide range of arthropods and, thus, have switched to RT-PCR and to use of more degenerate primers (shown below), both within Lepidoptera and beyond. We have tested the 30F primer only in Lepidoptera. The others work consistently well across Arthropoda, including Lepidoptera.

location	name	primer sequence (-M13 tail)
1	30F	CAY ATY AAY ATH GTS GTI ATH GG
97	40.6F	ATY GAR AAR TTY GAR AAR GAR GC
136	40.71F	TCN TTY AAR TAY GCN TGG GT
247	42.8R	ATC ATR TTY TTD ATR AAR TC
268	42.8F	GAY TTY ATH AAR AAY ATG AT
406	45.71sR	TCC ATY TTR TTN ACN WCI AC
427	45.71F	GTN GSN GTI AAY AAR ATG GA
709	52R	CCD ATY TTR TAN ACR TCY TG
730	52F	CAR GAY GTN TAY AAR ATH GG
832	52.5R	TCR TGR TGC ATY TCN AC
849	52.4F	TCN GTN GAR ATG CAY CAY G
1042	53.5R	ATR TGV GMI GTR TGR CAR TC
1228	41.21R	TGY CTC ATR TCD CGV ACR GCR AA

Our typical strategy for amplifying *EF-1□* is 30F_41.21R (for Lepidoptera) or 40.6F_41.21R (for other arthropods) followed by various nested reamps.

elongation factor-2 (*EF-2*)

We have used *EF-2* across arthropods. Our *EF-2* primers were first published in Regier and Shultz 2001. Mol. Phylog. Evol. 20: 136-148.

location	name	primer sequence (-M13 tail)
1	20F	ATG GTN AAY TTY ACN GTI GA
76	95F	GCN CAY GTN GAY CAY GGI AA
382	431F	GTN GTN GTI GAY TGY GT
382	431TF	GTN GTN GTI GAY GCY GT
385	436R	TCN GTY TGN ACR CAN ACI CC
403	452F	GGN GTN TGY GTI CAR AC
406	455F	GGN GTN TGY GTN CAR ACI GA
637	691R	GTR AAN GCC CAN CCR TG
655	707sF	CAY GGN TGG GCI TTY WC

1111	1216sR	TAC ATC ATN ARI GGN WC
1123	1228R	ACC ATY TTI GAN AYR TAC AT
1129	1232sF	GSN CCN YTI ATG ATG TA
1144	1228sF	ATG ATG TAY RTH TCN AAR ATG GT
1285	1390R	CCC ATC ATN ARI ATN GT
1303	1390F	ACN ATN YTN ATG ATG GG
1528	1640R	CAY TGN ACC ATN GGR TC
1546	1655F	GAY CCN ATG GTN CAR TG
1561	1672R	CCN GCD ATD ATR TCY TCI CC
1969	2080sF	CAR TAY YTI AAY GAR ATI AAR GA
1978	2092R	GCC CAY TGR AAN CCI GC
2179	2293R	TCN GGR CAY TGD ATY TC

431TF works better than 431F for some tardigrades. The typical strategy for amplifying EF-2 is to generate the following by RT-PCR: 20F_1390R or 95F_1390R, 707sF_1672R, and 1655F_2092R or 1655F_2293R. Then, after gel isolation a series of nested reamps are performed.

largest subunit of RNA polymerase II (*Pol II*)

We have used *Pol II* across arthropods. Our first *Pol II* primers were published in Shultz and Regier 1999. Proc. R. Soc. Lond. B 267: 1011-1019.

location	name	primer sequence (-M13 tail)
1	35F	GGN AAR MGI GTI GAY TT
422	37F	GAY TTY GAY GGI GAY GAR ATG
538	39F	GTN ATG GGN ATH GTI CAR GA
625	46R	GGM ATY TTN CCR TCC CA
742	48R	TCR TCY TCI TCR TCN GGR TG
760	47sF	CAY CCN GAY GAI GAR GA
1141	49sF	CAR ACI TTY GAR AAY MAI GT
1342	41R	CCR TAR TCR TCY TTI ATR AAR TG
1354	42R	YTC NGG ICC RTA RTC RTC
1360	5F	CCI CAY TTY ATH AAR GAY GA
1372	6F	GAY GAY TAY GGI CCN GA
1429	43R	CCN CCC ATI GCR TGR AA
1447	7F	TTY CAY GCN ATG GGN GG
1483	44R	YTG DAT RTA ICC NGT YTC
1501	15F	ACW GCH GAR ACH GGK TAY ATY CA
1525	14F	YTK ATH AAR GCT ATG GA
1873	17R	TTY TGN GCR TTC CAD ATC AT
1894	18F	ATG ATH TGG AAY GYN CAR AA
2167	27R	GCN CCA ACC ATY TCN CC
2218	28R	ART GRA AIG TRT TNA RIG TCA TYT G
2243	30F	CAR ATG ACI YTN AAY ACI TTY CAY T
2437	22R	TCN GGR TCR TAR TAD ATN GC
2662	23R	TTN TCI GCR TTR TCR TC
2800	32R	CCY TGN ARI GTC ATR TC

The typical strategy for amplifying *Pol II* is to generate the following by RT-PCR: 39F_43R, 5F_17R, 18F_28R, 30F_32R. Then, after gel isolation, a series of nested reamps are performed.

acc (acetyl-coA carboxylase)

acc amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GTN TGG GCN GGN TGG GG
2F:  TGG GCN GGN TGG GGN CAY GC
4R:  GCY TCY TCD ATD ATY TTY TG
```

amplification strategy: 1F_4R (507 nt) / 2F_4R (501 nt)

aspec (β-spectrin)

aspec amplifies across arthropods and fragment 2 (see below) is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  CAY CAR GAR CAY AAR GG
4F:  CAY AAR GGN GAR ATH GAY GC
6R:  YTC RAA RTC YTC RTG YTT
7R:  ATY TTY TCY TCY TGN GC
11F: TGG ATH MGN GAR AAR GA
12R: ACY TCN ACY TTC CAC CAR TC
15R: AAR TCR TCR AAY TTY TTY TG
19F: TAY ACN TGG TTY ACN ATG GA
21R: RTC RTT RTG YTT RAA CAT CAT
22R: TCY TTR TCR AAR TGY TTR AA
```

amplification strategies (fragment 1): 2F_6R (297 nt) and 2F_7R (327 nt) / 4F_6R (285 nt)
amplification strategy (fragment 2): 11F_15R (915 nt) / 11F_12R (594 nt)
amplification strategy (fragment 3): 19F_22R (456 nt) / 19F_21R (450 nt)

2fin (pre-mRNA splicing factor)

2fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
3F:  GAR AAY ATG CCN ATG CCN TGG GAR CA
4R:  RTT RAA YTC RTT CCA RTC YTC
5R:  ATR TCR TTR AAY TCR TTC CA
6F:  GAY TAY GGN AAR AAR AAY AAY GT
7F:  GAR AAR CAR ACN AAR GAR CA
8R:  GGR TTR TTN CCY TTR TC
```

amplification strategy (fragment 1): 3F_5R (537 nt) / 3F_4R (531 nt)
amplification strategy (fragment 2): 6F_8R (735 nt) / 7F_8R (618 nt)

6fin (casein kinase)

6fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  TTY GAR CAY GTN AAY AA
3R:  RTA RTT RTC RTG NCC RTG RAA
```

amplification strategy: 2F_ 3R (342 nt)

25fin (signal recognition particle)

25fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  AAR CAR AAY GCN ACN AAR GC
2F:  CAR GCN TGY GAR GCN CAR GC
4R:  CAT YTT CAT DAT RTT YTG RAA YTG
```

amplication strategy: 1F_4R (537 nt) / 2F_4R (303 nt)

26fin (spliceosome-associated protein)

26fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  ATG GAY GCN GAR TAY GC
3F:  GAY GAR GAR ATG AAR AAR ATH GT
4R:  RTC DAT RCA RTT YTC YTG
5R:  GCR TAD ATR TAR TCY TT
```

amplification strategy: 2F_5R (1200 nt) / 3F_4R (729 nt)

36fin (syntaxin)

36fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  AAY GTN GAR GAR GTN AA
2R:  GCC ATR TCC ATR AAC ATR TCR TG
3R:  GCC ATR TCC ATR AAC AT
```

amplification strategy: 1F_3R (477 nt) / 1F_2R (471 nt)

40fin (phosphogluconate dehydrogenase)

40fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  ATH GAR TAY GGN GAY ATG CA
3R:  GTR TGI GCN CCR AAR TAR TC
4R:  CCN GTC CAR TTN GTR TG
```

amplification strategy: 2F_4R (801 nt) / 2F_3R (747 nt)

42fin (putative GTP-binding protein)

42fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GCN GAR AAY TTY CCI TTY TG
2R:  GCC ATD ATR AAN CCY TTY TCR AAR TC
```

amplification strategy: 1F_2R (846 nt)

44fin (glucosamine phosphate isomerase)

44fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  TTY AAY ATG GAY GAR TAY GT
3R:  CCY TTR AAR TAY TTN AC
```

amplification strategy: 2F_3R (528 nt)

58fin (clathrin heavy chain)

58fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
3F:  AAY TAY GAY GCN CCN GAY AT
5F:  TAY GAY AAR TAY GAR GAR TA
6R:  TAY TCY TCR TAY TTR TCR TA
7F:  TAY GAR GAR TAY GAY AAY GC
8R:  GCR TTR TCR TAY TCY TCR TA
9R:  ATR TTR TCR AAR TTR TC
```

amplification strategy (fragment 1): 3F_8R (936 nt) / 3F_6R (927 nt)
amplification strategy (fragment 2): 5F_9R (348 nt) / 7F_ 9R (339 nt)

62fin (protein phosphatase)

62fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  CAY GAR CCN GAR TTY GAY TA
3R:  RCA YTC RAA YTT RTC RAA DAT
4R:  RCA YTC RAA YTT RTC RAA
```

amplification strategy: 2F_4R (768 nt) / 2F_3R (765 nt)

63fin (□-adaptin)

63fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  GAY ATH GAY TTY GGN CAY ATG
3R:  ACR AAR TAR TAN GTR TAR TCY TG
```

amplification strategy: 2F_3R (501 nt)

69fin (clathrin coat assembly protein)

69fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  AAY GAR GCN TAY TTY GAY GT
3R:  TTR TAY TTY TCN CCR TAC ATR TC
4R:  CCY TTR AAN GGY TTR TAY TTY TC
```

amplification strategy: 2F_4R (642 nt) / 2F_3R (630 nt)

73fin (acetylglucosaminyltransferase)

73fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  TGY GTN TAY TAY GAR CAR GG
3R:  GGR TGR TTN CCR AAR TC
4R:  GTD ATD ATR TAR TCC ATR AA
```

amplification strategy: 2F_4R (1194 nt) / 2F_3R (867 nt)

96fin (ATP synthase)

96fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GAR GAR ATG ATH CAR AC
3R:  TCN GTR TAC ATR TAN CC
```

amplification strategy: 1F_3R (459 nt)

109fin (gelsolin)

109fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GTN TAY TTY TGG CAR GG
2R:  TCR TCY TGR CAR AAR TC
3R:  DAT CAT DAT RTC RTC RTC
```

amplification strategy: 1F_3R (627 nt) / 1F_2R (603 nt)

113fin (glycogen synthase)

113fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GCN CAY TTY CAY GAR TGG
2R:  CCC CAN GGY TCR TAR TA
3R:  CCC ATN ACN GTR CAY TC
```

amplification strategy: 1F_3R (1005 nt) / 1F_2R (975 nt)

127fin (methylmalonate semialdehyde dehydrogenase)

127fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GAY GCN GAR GGN GAY GT
2R:  GTR TTY TCY TTR TTN GCR TC
3R:  GGN CCR AAD ATY TCY TC
```

amplification strategy: 1F_3R (831 nt) / 1F_2R (468 nt)

149fin (protein kinase)

149fin amplifies across arthropods. Sequences from *Mastigoproctus* (Chelicerata: Arachnida) and *Limulus* (Chelicerata: Xiphosura) were polymorphic enough to make sequence assembly challenging, but not impossible.

```
1F:  CAY TAY GCN TGY TTY TGG
2F:  CAY TAY GCN TGY TTY TGG G
3R:  GGR TCY TCR TTC ATR CAD AT
4R:  GGR TCY TCR TTC ATR CA
```

amplification strategy: 1F_4R (942 nt) / 2F_3R (936 nt)

166fin (CDC 5-related protein / cell division)

166fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  ATG AAR TAY GGN AAR AAY CAR TGG
3R:  TCR TCY TCR TCC ATR TC
```

amplification strategy: 2F_3R (333 nt)

192fin (glutamyl- & prolyl-tRNA synthetase)

192fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GAR AAR GAR AAR TTY GC
2R:  TCC CAR TGR TTR AAY TTC CA
```

amplification strategy: 1F_2R (405 nt)

197fin (triosephosphate isomerase)

197fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  AAY TGG AAR ATG AAY GG
2R:  GCC CAN ACN GGY TCR TA
```

amplification strategy: 1F_2R (441 nt)

220fin (F-box protein)

220fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GAY TGY GAR AAY GTN GG
2R:  CCY TTY TCR TGN ACR TAD AT
4R:  GCC ATN GCR TTR TCR AAD ATY TCR TT
```

amplification strategy: 1F_4R (885 nt) / 1F_2R (570 nt)

226fin (glutamine amidotransferase)

226fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  CAR TTY CAY CCN GAR GT
2R:  TCN GTR TCR TTR TGR TG
4R:  ACY TCR TCY TTR TGR AAR TC
```

amplification strategy: 1F_4R (648 nt) / 1F_2R (582 nt)

247fin (leucyl-tRNA synthetase)

247fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  CAY TGY ACN GGN ATG CC
2R:  TCR TGR TCC ATR CAN GGY TG
```

amplification strategy: 1F_2R (447 nt)

262fin (proteasome subunit)

262fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  TGG GGN CAY GAR TAY GT
2R:  TTN GGY TCC ATD ATR TC
```

amplification strategy: 1F_2R (525 nt)

265fin (histidyl-tRNA synthetase)

265fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GAR TTY TAY CAR TGY GA
2F:  GAR TTY TAY CAR TGY GAY TTY GAY ATH GC
3R:  ACN CCN GTR TAR TAR TC
```

amplification strategy: 1F_3R (459 nt) / 2F_3R (447 nt)

267fin (*CAD*, pyrimidine biosynthesis)

267fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera; however, *267fin* is the *CAD* gene, for which

lepidopteran primers are already available (see above). These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  CAR TTY CAY CCN GAR CA
2F:  GCN GGN GAR TTY GAY TA
3R:  TTY TCC ATR TTR CAN AC
```

amplification strategy: 1F_3R (804 nt) / 2F_3R (600 nt)

268fin (AMP deaminase)

268fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  ATG AAY CAR AAR CAY YT
2R:  CCN GCY TCN CCR CAR TG
3R:  ARY TCR CAC ATR TCR CA
```

amplification strategy: 1F_3R (1083 nt) / 1F_2R (768 nt)

270fin ("hypothetical protein")

270fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  TGY GTN GAY AAY TTY GAR GC
3R:  TCD ATN CCC CAY TCR TT
```

amplification strategy: 2F_3R (450 nt)

274fin (methionine aminopeptidase)

274fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  CCN GGN ATG ACN ATG AT
2R:  TCC ATR TCR TCR TGN AC
```

amplification strategy: 1F_2R (537 nt)

3006fin (dynamin)

3006fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  CCN GAY ATG GCN TTY GA
2R:  TCY TCR TGR TTN GTR TTC ATR TA
```

amplification strategy: 1F_2R (222 nt)

3007fin (glucose phosphate dehydrogenase)

3007fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  AAR AAR AAR ATH TAY CC
2R:  ARR TGR TTY TGC ATN ACR TC
3R:  ACY TTN ACY TTY TCR TC
```

amplification strategy: 1F_3R (705 nt) / 1F_2R (621 nt)

3009fin (G protein-coupled receptor kinase)

3009fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  AAR ATG TAY GCN ATG AAR TG
3R:  GCR TGN GGY TTY TTY TT
```

amplification strategy: 2F_3R (369 nt)

3012fin (DNA replication licensing factor)

3012fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  TGY AAR CCN GGN GAY GA
2R:  CAT YTT RTC RAA YTC RTC
3R:  TGY TGY TCC ATN GCY TCR TG
```

amplification strategy: 1F_3R (567 nt) / 1F_2R (525 nt)

3017fin (tetrahydrofolate synthase)

3017fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  ATG GAR GAR TTY AAY YT
2R:  CCR TGN GCD ATR TTN GCR AA
3R:  AAY TTY TCC ATN CCD ATR TC
```

amplification strategy: 1F_3R (699 nt) / 1F_2R (594 nt)

3031fin (myosin)

3031fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  TTY GGN AAR TAY ATH GA
2F:  AAR ATH GAR CAR TAY YT
3R:  AAR TTY TCR AAN CCR AAD ATR TC
4F:  GAY GAR ATH TAY TGY CA
5R:  GCR TCC ATN ACR TGR TC
6R:  CCY TGY TCY TTN GCR TAY TGY TCR CAY TG
```

amplification strategy (fragment 1): 1F_3R (642 nt) / 2F_3R (588 nt)

amplification strategy (fragment 2): 4F_6R (450 nt) / 4F_5R (426 nt)

3044fin (prohormone convertase)

3044fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  TAY ACN GAY GAY TGG TTY AA
2R:  CCR TCR CAR TTR CAR TCR TC
```

amplification strategy: 1F_2R (324 nt)

3055fin (protein kinase)

3055fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GAN GAY ATH CAR CAR GA
2F:  TGG ATH ATH ATG GAR TA
3R:  TCN GGN GCC ATC CAR AA
```

amplification strategy: 1F_3R (339 nt) / 2F_3R (246 nt)

3059fin (arginine methyltransferase)

3059fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera.. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GGN ATH CAY GAR GAR ATG
3R:  TAR AAN ACN GTY TGY TTC CAR TG
```

amplification strategy: 1F_3R (738 nt)

3064fin ("transmembrane protein")

3064fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
6F:  CAR GAR GAR TTY GGN TGG AA
7R:  AAN CCR AAC ATR TAR TA
```

amplification strategy: 6F_7R (606 nt)

3066fin (RNA helicase)

3066fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  CAR CAR ATH GAR GAR GAR AC
3R:  TTN GCC ATR TCR TAR TT
```

amplification strategy: 1F_3R (744 nt)

3070fin (alanyl-tRNA synthetase)

3070fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
2F:  TAY CAY CAY ACN TTY TTY GAR ATG
4F:  ATG AAR GAY AAY TTY TGG GAR ATG GG
5R:  GGR AAN CCR TAN GTR TCR TA
```

amplification strategy: 2F_5R (942 nt) / 4F_5R (705 nt)

3089fin (acetyltransferase)

3089fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GAR AAY TAY CAR ATG AAR TAY TA
2F:  CAR ATG AAR TAY TAY TTY TAY CA
3R:  GCR TAR TAY TTN GGY TC
```

amplification strategy: 1F_3R (312 nt) / 2F_3R (303 nt)

3094fin (ATPase)

3094fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GAR ATH ATG GAR TTY GT
2F:  TTY ATH GAY GAR ATH GAY GC
3R:  ATN GCY TGY TCR AAR TG
```

amplification strategy: 1F_3R (651 nt) / 2F_3R (390 nt)

3114fin (glutaminyl-tRNA synthetase)

3114fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  AAY CCN GAR AAR GAR GA
2R:  CAR TGN GTR TAR TCR TA
```

amplification strategy: 1F_2R (378 nt)

3121fin (protein kinase)

3121fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GAY GCN ATG TTY CCN GT
2F:  GAR GGN GAY AAY TTY TAY GT
3R:  ATY TCN CCR AAR TAR TC
```

amplification strategy: 1F_3R (498 nt) / 2F_3R (438 nt)

3136fin (histone deacetylase)

3136fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  AAY TAY TAY TAY GGN CAR GG
2R:  GTY TCR TAN GTC CAR CA
3R:  CCR AAR TAY TCR AAR TAR TCR TTR TA
```

amplification strategy: 1F_3R (906 nt) / 1F_2R (849 nt)

3152fin (protein kinase)

3152fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  CAY TGY ATH CAR CAR AT
2R:  TGY TGR TCY TCR TCC CAR AA
```

amplification strategy: 1F_2R (279 nt)

3153fin (RNA helicase)

3153fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GAY GCN GGN AAR TTY CA
2R:  GTR TTR TCY TGC ATY TC
3R:  GGR TCC ATR AAR TCR AAR TG
```

amplification strategy: 1F_3R (666 nt) / 1F_2R (573 nt)

3196fin (RNA polymerase II, subunit 2)

3196fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  ATG ATG CCN AAY GAR GC
3R:  ACC ATY TCC ATC ATY TCN GG
4R:  GCN ACR TTY TGY TCY TG
5F:  AAY TGG GGN GAY CAR AAR AAR GC
6R:  GGR TGD ATY TCR CAR TG
7R:  TGR TTR TGR TCN GGR AA
```

amplification strategy (fragment 1): 1F_4R (594 nt) / 1F_3R (543 nt)
amplification strategy (fragment 2): 5F_7R (762 nt) / 5F_6R (708 nt)

3202fin (ATP synthase)

3202fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  TTY GAR CAR ATG GAR GC
3R:  ACN CCR AAR TGR AAY TG
4R:  ATR TAR TTR TCD ATY TT
```

amplification strategy: 1F_4R (606 nt) / 1F_3R (504 nt)

8018fin (proteasome)

8018fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  TAY CAR GAY GAY GCN GT
2R:  ACR TCR TCR TTN ACR TC
```

amplification strategy: 1F_2R (303 nt)

8028fin ("nucleolar cystein-rich protein")

8028fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  ATG GAR TGY GAY AAR TG
2R:  TCR AAY TGR TCR TCY TCR CA
```

amplification strategy: 1F_2R (324 nt)

8029fin (neurofibromin)

8029fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
4F:  CAR GCN AAR ATH TGG GG
6F:  ATG TGG GAY GAY ATH GC
7R:  CAT RCA NGC YTC CAT DAT YTC
```

amplification strategy: 4F_7R (663 nt) / 6F_7R (420 nt)

8053fin (phosphatidylinositol kinase)

8053fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  AAY CCN AAR TGG ACN AA
2F:  TGY TGY CCN TGY TGY TT
3R:  GGR TGY TTR AAN GGR AA
```

amplification strategy: 1F_3R (435 nt) / 2F_3R (402 nt)

8070fin (SH2 domain binding protein)

8070fin amplifies across arthropods. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  GAR GGN GAY AAR ATG GAY CA
2R:  TGR AAN GCY TGR TCR TAR TC
3R:  TGR TAR TAR TAY TGR AA
```

amplification strategy: 1F_3R (534 nt) / 1F_2R (519 nt)

8091fin (glucose phosphate isomerase)

8091fin amplifies across arthropods and is sufficiently variable at the level of its protein sequence to be a candidate for testing across Lepidoptera. These primers were developed through the NSF-funded collaboration of Regier and Cliff Cunningham.

```
1F:  TTY TGG GAY TGG GTN GG
2R:  AAD ATY TTR TGY TCR TA
3R:  ACN CCC CAY TGR TCR AA
```

amplification strategy: 1F_3R (714 nt) / 1F_2R (666 nt)

11. Updates to earlier versions.

Please check www.umbi.umd.edu/users/jcrlab/PCR_primers.pdf for the latest version of this document. I (JCR) would be grateful to be informed of any errors in this document or any general comments that might improve it.

Version 4/01/05:	The first release.
Version 4/06/05:	Primer 2R and a nested reamp strategy were added to gene *8070fin*.
Version 4/15/05:	An error in the amplification strategy for gene *3202fin*, which was inadvertantly introduced in the previous version, was corrected. The URL for downloading this document was corrected.
Version 4/25/05:	The introduction and acknowledgments were revised. Primers for the *PEPCK* gene were added.
Version 5/23/05:	An error in the amplification strategy for *CAD* was corrected. Other small inconsistencies in primer name usage were corrected.
Version 6/7/05:	Primer *581F* of the *CAD* gene was replaced by *582F*. Also, the three primers for *267fin* (also encoding *CAD*) were duplicated in the *CAD* gene section, and the "typical strategy for amplifying *CAD*" was modified.
Version 8/30/05:	Primer 743F of the *CAD* gene was replaced by 743nF. The original primer unexplicably contained errors, although it still worked moderately well. Also, two modified *wg* primers were added.
Version 3/14/06:	A new primer for amplifying *wingless* was added and a problem with an old primer was noted. The list of genes that were being used / tested in Lepidoptera was slightly modified.