via PCR
Evolutionary Analyses

CHAPTER 13
polymerase chain reaction and direct sequencing overcome the limitations of traditional DNA methods for evolutionary studies.¹

**UNIVERSAL PRIMERS**

At first glance, it would appear that the lack of sequence knowledge for most species would limit the application of the PCR, since some knowledge of the sequence is required to design primers for PCR. Fortunately, there is a rapid means of obtaining primer sequences for new species. By choosing sequences that are conserved among widely divergent species, "universal" primers can be designed that will amplify a particular nuclear or organelle gene fragment from nearly all members of a major taxonomic group, e.g., plants or fungi. This extends the phylogenetic range of comparative sequence analyses to the taxonomic level of class or phylum. But the method also has important uses for identifying specimens and scoring types in population work. For those purposes, mitochondrial DNA (mtDNA) sequences are often appropriate. Below, we describe "universal" primers for both nuclear and mitochondrial sequences. Mitochondrial genes are a rigorous test of the method; since mtDNA evolves so quickly, it should be difficult to design primers for this molecule.

**Primers for Nuclear Genes**

The concept of "universal" primers has been used for a number of years in the direct sequencing of abundant RNAs.¹ Those same primers are easily paired to amplify ribosomal DNA sequences via PCR. Amplification of the rDNA sequences offers several advantages over direct rRNA sequencing. First, DNA can be used as the starting material, and this is usually easier to prepare from tissues than is RNA. Second, much smaller samples can be used.¹ Finally, DNA sequencing methods can be used — including a variety of sequencing enzymes and techniques — to obtain data from both strands and to sequence through the complex secondary structures found in ribosomal RNAs.

As an example, we have designed primers that amplify a region of approximately 515 bp in size of 18S rDNA from many fungi, protists, algae, plants and animals (the size of the amplified region plus the primers is approximately 555 bp). The primers NS1 and NS2 were based on conserved nucleotide sequences among the 18S rDNA from Saccharomyces cerevisiae, Dicyostelium discoideum, and Styloclisia pusilla and have been described elsewhere.²

\[ \text{NS1: 5'}-\text{GTAGTCATATGCTTGTCTC-3'} \]
\[ \text{NS2: 5'}-\text{GGCTGCTGCGCACCAGCTTGC-3'} \]

NS1 and NS2 do not amplify bacterial or mitochondrial rRNA genes using the conditions described below. Sequence differences in the region amplified by NS1 and NS2 (and for some organisms, length differences) may be useful for initial estimates of molecular diversity, i.e., determining the number of different species in a natural population of various organisms. These primers have also found use in detecting the presence of fungal DNA contamination in ancient DNA isolated from plant and animal tissues. Medlin and coworkers³ describe other rDNA primers that amplify the entire 18S gene from lower eukaryotes; these primers will be more useful than NS1 and NS2 for phylogenetic studies of these organisms.

**Primers for Mitochondrial DNA**

Studies of mitochondrial gene sequences are appropriate for many problems in evolutionary and population biology.⁴ Because of its clonal inheritance in vertebrates, it is ideal for reconstructing maternal phylogenies. Its high rate of point-mutation evolution makes it ideal for high-resolution population studies within species and the rapid fixation of mutations within species makes this molecule ideal for species identification, especially in small organisms.

Because vertebrate mtDNA evolves so quickly (roughly 10 times as fast as nuclear genes), it should be difficult to find conserved sequences to use as priming sites for the PCR. The location of two primers that amplify a 307-bp region (the amplified fragment is 376 bp) of the cytochrome b gene from most vertebrates tested (mammals, birds, amphibians, reptiles, and fishes) is shown in Figure 1. Primers L14841 and H15149 were based on conserved regions of published nucleotide sequences. The letters L and H refer to the light and heavy strands of mtDNA and the number refers to the position of the 3' base of the primer in the complete mtDNA sequence reported for a human.

\[ L14841: 5’-\text{AAAAGCTTCCATCCACATCTCAGCATGATGAAAA-3'} \]
\[ H15149: 5’-\text{AAACTGACGCCCTCTCAGAATGATATTGCTGCCTCACA-3'} \]

We have also had success using truncated versions of these primers to amplify mammal, bird, and amphibian DNA (C. Orrego, pers. communication). Primers MVZ3 and MVZ4 have the sequence underlined above and amplify a 311-bp region (the amplified fragment is 365 bp). The sequences from the mitochondrial cytochrome b gene contain phylogenetic information of high resolving power and great taxonomic range. Alignment of sequences is facilitated because of the overall conservation of the protein structure and function. Close relationships can be assessed through changes due to transitions at "silent" sites in codons; more distant relationships can be examined by analyzing transversion differences or amino acid replacements.⁴

**METHODS**

**DNA Preparation**

DNA was extracted from tissues by digestion in 100 mM Tris pH 8.0, 10 mM EDTA, 100 mM NaCl, 0.1% SDS, 50 mM DTT, 0.5 μg/ml Proteinase K
of genomic DNA and 2.5 units of Taq polymerase (Perkin Elmer/Cetus).
Each cycle of the polymerase chain reaction consisted of denaturation for 1
min at 93°C, hybridization for 1 min at 50°C, and extension for 2-5 min
at 72°C. This cycle was repeated 25-40 times depending on the initial
concentration of template DNA in the sample. Electrophoresis of 5 µl of
the amplified mixture was done in a 2% agarose gel (NuSieve, FMC Corp.)
in 40 mM Tris-acetate (pH 8.0) and the DNA stained with ethidium
bromide.

The gel fragment containing the amplified product was cut from the gel
and melted in 1 ml of distilled water, and 1 µl of this mixture was used as
the template in a second chain reaction to generate single stranded DNA for
sequencing. In this second reaction, the concentration of one or the other
primer was reduced 100-fold. After 40 cycles of amplification, free
nucleotides and salts were removed by 2-4 cycles of centrifugal dialysis
(Centricon 30, Amicon). The DNA was sequenced with a commercial kit
(Sequenase, U.S. Biochemical) using the primer that had been limiting in
the second chain reaction.

**Technical Notes**

It is possible to improve the specificity and yield of the amplified product
by raising the temperature of the annealing step for the DNA from some
organisms. This increased specificity permits single-stranded DNA template
to be produced directly from the original template in a single PCR
amplification of 35 cycles using a primer ratio of 50:1. Modified reaction
mixtures use reduced concentrations of each dNTP (32 µM) and cycling
parameters are: an initial denaturation at 93°C for 3 min followed by 35
cycles consisting of denaturation at 93°C for 25 sec, annealing at 55°C
for 25 sec and extension at 72°C for 2 min. Using longer extension times at
72°C during the linear phase may increase the yield of single-stranded
template for sequencing. Those cycling parameters were used with primers
NS1 and NS2 for fungal DNA and eliminated the additional bands observed
in lane 1 in Figure 2.

For studies where the same gene is being amplified and sequenced from
many different individuals or organisms, it is imperative to scrupulously
avoid cross-contamination of DNA samples during DNA isolation and
manipulation of amplified products. Only positive displacement pipets
which have disposable tips and pistons should be used in setting up PCR
reactions. Pipets that have been used with amplified DNA should never
subsequently be used for DNA isolation from tissues or for DNA dilutions
prior to another round of amplification. Controls that contain no DNA and
utilize all other reagents, diluents, etc., should be included in every
experiment to check for contamination. In extreme circumstances, it may
be necessary to switch to primers which amplify a different segment of the
target gene and to obtain new pipets that are dedicated to DNA isolation or
setting up PCR reactions.
RESULTS

Figure 2 shows the results of amplifying an rDNA fragment using 5 ng of total DNA from various organisms with primers NS1 and NS2. A major amplified DNA product of about 500 bp is observed from all the organisms tested and small variations in length are detected. Because these primers have such broad range, they may be useful for surveys of symbiotic organisms. For example, once sequence data have been obtained for a group of mycorrhizal fungi and host plants, species-specific DNA probes could be used to survey and identify individual symbiont partners in a natural population.

The primers for mtDNA amplify the cytochrome b region from a wide range of species. Figure 3 shows an alignment of mtDNA sequences from the cytochrome b obtained from five vertebrates. The sequences illustrate both the broad utility of the primers and the collinearity of the alignment. The primers work on human, mouse and cow mtDNA despite considerable sequence mismatch (Figure 4).

---

Figure 2. Amplification of a region of the nuclear small subunit rRNA gene using primers NS1 and NS2. Conditions for amplification, electrophoresis, and detection by visualization of bands stained with ethidium bromide were as described in Methods except that an annealing temperature of 45°C was used. Lane 1: Laccaria bicolor, a mycorrhizal basidiomycete; lane 2: Alder glauces; lane 3: Drosophila melanogaster; lane 4: Anadromus aximus, a spiker; lane 5: plasmid DNA from a clone of the rDNA repeat unit from Tyromyces unicolor; lane 6: a basidiomycete; lane 7: negative control, no DNA; lane 8: blank; lane 9: molecular weight standards, PhixX174RF HaeIII digest.
AMPLIFICATION WITH MISMATCHED PRIMER

PRIMER
Fish 5'—CCATCCAAACATCTCAGCATGATGAAA—3'

TEMPLATE
Human ..............................C......................
Cow ..............................T........................
Mouse ..............................T........................

Figure 4. Amplification with mismatched primers. The sequence of part of the cytochrome b primer L14841 is shown. Despite mismatches in the middle of the primer, it can be used to amplify a variety of mammalian templates. It is most important that the 3' end of the primer be well-matched to the template.

DISCUSSION
The ability to rapidly amplify sequences from virtually any species from just a few molecules of DNA has many applications in the study of evolution and ecology. Here we discuss some promising uses of the technique, and try to stimulate the invention of other applications in these fields.

Molecular Systematics
While the advantages of nucleotide sequence data for systematic studies have long been recognized, the accumulation of comparative sequence data has until now been tedious. The rapid sequencing which PCR/universal primer technology provides will facilitate the extension of molecular systematics to many more groups. The homologous data that are provided by sequences provide a common phylogenetic framework in various taxonomic groups. Sequence data also provide a degree of resolution that has not been possible with previous methods. The PCR amplification and direct sequencing of mitochondrial DNA has already been useful in obtaining sequences that confirm the identity of an African root for the tree relating human mtDNAs. In another study, a phylogenetic analysis of primate class II histocompatibility genes demonstrated that the origin of many of the alleles of the HLA-DQa loci predate the speciation event that led to humans.

Because amplification can be accomplished from only a few molecules of the target sequence, many samples previously refractory to molecular analysis can be utilized. The ability to work from museum and herbarium specimens should promote the use of molecular data in a variety of systematic studies. Also exciting is the application of these techniques to ancient specimens DNA from a 7,000-year-old brain preserved in a Florida peat bog was found to contain a novel type of mtDNA not yet observed among living Native Americans.

Population Biology
The rapid amplification of sequences allows the serious consideration of population studies at the DNA sequence level. Single-stranded amplifications allow rapid sequencing of DNA from tens or hundreds of individuals without the tedious cloning steps required previously. Once a representative set of sequence data is available, more convenient and simple analytical methods, such as allele specific oligonucleotide probes (see Chapter 16), can be used to obtain allele frequency data.

The ability to work from museum specimens will facilitate study of gene frequencies over time. Thomas and coworkers have studied a series of rodent populations through 70 years by amplifying mtDNA sequences from the skins of preserved specimens. Variation in gene frequency over such long periods has not been accessible to modern molecular techniques until now.

Ecology and Marine Biology
Traditional molecular techniques require relatively large tissue samples for analysis. In particular, many invertebrates are too small for the manipulations normally required for molecular investigations. For this reason, direct analysis of individual small organisms, organisms that are symbiotic, or those that are not easily grown in pure culture has been difficult. The polymerase chain reaction can be used to extend the range of organisms that is accessible to molecular investigations. It should now be possible to study directly the genetic structure of natural populations of single-celled organisms such as protozoa and algae. The method should find wide application in analyses of dispersal and recruitment, especially in marine environments. Finally, the sensitivity of the polymerase chain reaction should allow detection and identification of small numbers of single-celled organisms, symbionts and parasites, even from complex mixtures of DNA isolated from their hosts.

Non-Invasive Sampling
The collection of samples for genetic analysis has frequently required sampling of blood or tissues. This has limited the use of genetic analyses
in behavioral and ecological studies where disturbance of the subjects must be minimized. The ability to amplify sequences from forensic samples such as single hairs offers new opportunities for behavioral scientists and conservation biologists. Non-invasive sampling should also facilitate collection of samples for studies of endangered species.

Synergism of Molecular and Organismal Studies

By opening up a new range of species to molecular analysis, we hope a productive interaction between molecular and population biologists can arise. The comparative sequences gathered for phylogenetic reconstruction, for example, may shed light on protein structure and function. Molecular study of organisms adapted to unique environments and not usually studied in the laboratory may reveal unusual molecular adaptations. Conversely, knowledge of the molecular structure of genetic variants may contribute to our understanding of how organisms are evolving.

ACKNOWLEDGMENTS

We thank Mary Ann Brow and Randy Saiki for many helpful suggestions on DNA sequencing and PCR, Monique Garde for providing DNA from mycorrhizal fungi, Dave Irwin for sheep DNA, Leticia Aviles for spider DNA, and Allan Wilson, Norm Arneheim, Cristian Orrego, Tom Bruins, Ellen Prager and John Taylor for helpful advice and comments on the manuscript.

REFERENCES