CE 573
Lab 3
Due: 3/30/08

Lab 3- Identification of Bacteria in an Environmental Sample

Introduction

The main purpose of this laboratory was to identify unknown bacterial organisms found in a given environmental sample taken from a lab-scale bioreactor. Our methods followed a culture-dependent method of identification. In contrast to this laboratory, a majority of scientific experiments involving the identification of organisms use culture independent methods. The lab also demonstrates the ability of the polymerase chain reaction (PCR) to efficiently amplify a small amount of DNA for the gene encoding 16s ribosomal RNA. This amplified DNA could then be sequenced and the organism could be identified.

Methods

Single bacteria colonies were isolated using streak plating on nutrient agar of an environmental sample provided from a lab-scale bioreactor. The selected colony was introduced to 1.0 mL of nutrient broth and then added to 9.0 additional mL of nutrient broth to form a 10 mL pure liquid culture. Each lab group isolated a colony from 5 different plates to prepare 5 pure cultures. DNA was extracted from these samples using a DNA extraction kit (Quiagen DNA Stool Mini Kit Catalog #51504). Next, two separate Mastermix solutions were prepared using Quiagen HotStart Taq mastermix reagents with a 341 forward primer and two different reverse DNA primers, 518 and 907. PCR products were obtained from the Thermocycler and purified using Quiagen QUIAquick PCR Purification Kit. The Thermocycler was run at several different temperatures to obtain ideal results. Gel Electrophoresis was then performed as a verification of the PCR results using a 1% pre-poured agarose gel, 10,000xSYBR Gold Stain, 1xTAE buffer, and MBI Fermentas GeneRuler DNA ladder. Loading dye was added to the sample drops so that image analysis could be performed. Electrophoresis was run for approximately 50 minutes at 90mA and 100V. The quality DNA was sent for sequencing by the KU DNA Sequencing Facility and the results were used to identify organisms based on the Genebank nucleotide-nucleotide BLAST.

Results

The table below summarizes the results for our laboratory group. In the experiment there were twelve samples, two of which were water control samples and should not have contained any DNA. Three of the remaining samples produced usable sequences and were able to be identified. In general, the samples that produced sequences had a significantly higher DNA concentration than the samples that did not. This could account for some of the reason the other samples did not give us good sequences but sample two had the highest concentration of DNA and still was not usable. The sequences that were not usable were most likely samples in which the PCR was unsuccessful. When you look at the DNA sequencing files in Chromas Lite, the

sequences that were not good produced overlapping curves and did not show clear smooth peaks. This indicates that the sequence represents the free nucleotides that were added before attempting PCR. If the PCR did not work correctly the free nucleotides would still be present and would produce this type of crowded image.

In addition to the bad sequences there was also a problem with the results of our gel electrophoresis. The electrophoresis image analysis did not produce bands where expected for any of our samples. It did, however, show the image of the DNA ladder. Because all of the samples were found to contain some amount of DNA, there is some question as to why we saw no image. It is possible that we ran the electrophoresis for too long or another technical aspect was performed incorrectly.

Because our samples did contain DNA, they could still be sent for sequencing and we did identify organisms. All three of our good samples were found to be DNA sequences of organisms in the Aeromonas genus.

Group 1 DNA Lab data

Sample	*DNA Sequence*	*Organism*	*Concentration, ng/µL*
1	Sequence not good		13.1
2	Sequence not good		29.1
3	AATATTGCACAATGGGGGAAACCCTGATGCAGC CATGCCGCGTGTGTGAAGAAGGCCTTCGGGTT GTAAAGCACTTTCAGCGAGGAGGAAAGGTTGGT AGCGAATAACTGCCAGCTGTGACGTTACTCGCA GAAGAAGCACCGGCTAACTCCGTGCCAGCAGC CGCGGTAA	Aeromonas	23.5
4	AATATTTGCACAATGGGGGAAACCCTGATGCAG CCATGCCGCGTGTGTGAAGAAGGCCTTCGGGT TGTAAAGCACTTTCAGCGAGGAGGAAAGGTTGA TGCCTAATACGTATCAGCTGTGACGTTACTCGC AGAAGAAGCACCGGCTAACTCCGTGCCAGCAG CCGCGGTAA	Aeromonas	23.4
5	AATATTGCACAATGGGGGAAACCCTGATGCAGC CATGCCGCGTGTGTGAAGAAGGCCTTCGGGTT GTAAAGCACTTTCAGCCAGGAGGAAAGGGTGG TGGCTAATAACTGTCAACTGTGACGTTACTCAC AGAAGAAGCACCGGCTAACTCCGTGCCAGCAG CCGCGGTAA	Aeromonas	24.2
7	Sequence not good		3.2
8	Sequence not good		2.4
9	Sequence not good		4.4
10	Sequence not good		2.8
11	Sequence not good		1.7

Discussion

The procedures of this experiment follow what is considered to be culture-dependent methods because we attempted to isolate an individual colony of bacteria before performing PCR or any DNA analysis. Usually it is not practical to follow this method because you cannot always isolate the exact organism from a sample. This becomes especially relevant in environmental samples that contain many different organisms. In the environment, approximately 20% of naturally occurring bacteria have

actually been isolated so far (Muyzer et al., 695). It is clear that there must be a way to characterize those microorganisms of interest that are not able to be extracted from complex environments. Once DNA containing a community of microorganisms is extracted, individual organisms can be isolated using denaturing gradient gel electrophoresis (DGGE). In electrophoresis of natural samples, a single set of 16s rRNA primers would produce what appears to be a single band. This band actually contains many related genes that are not identical (Madigan et al., 605). When you want to sequence these results, you must separate the genes. Performing DGGE will separate these genes based on the melting points of different DNA fragments coding for 16s rRNA (Muyzer et al., 695). Sequence variation of the genes accounts for the differences in melting points. Once the molecule reaches is melting point, migration in the DGGE gel will stop, leaving base pairs with different melting temperatures separated (Muyzer et al., 695). PCR performed in combination with DGGE was found to be successful in producing products and examining the diversity of oxygenic phototrophic microorganisms in both cultures and complex microbial communities (Nubel et al., 1997).

A significant portion of this experiment was dependent on the functioning of the polymerase chain reaction (PCR). The goal of this procedure was to amplify the sequence of DNA that encodes for the 16s rRNA gene. We want to isolate this gene because it is known to be a very accurate evolutionary chronometer. It can be used to identify organisms and also relate them to each other based on the amount of variation in sequences (Madigan et al., 309). Because PCR is very powerful at amplifying this DNA, it has allowed for the detection of organisms existing in the environment in very small concentrations as well as very small sample volumes (Wintzingerode et al., 214). PCR requires the use of a PCR primer, a single strand of DNA that provides the region we want to amplify. The PCR primer needed to produce accurate results depends strongly on the expected number of members in the population (Wintzingerode et al., 215). In our experiment, we used two different primer pairs that contained the same forward primer, and different locations of the reverse primer. The primer location tells the DNA where to begin forward and reverse transcription. The first primer covered a distance of 180 base pairs and the second covered 560 base pairs of the DNA encoding for 16s rRNA. A larger set of base pairs may narrow down the number of possible organism matches and produce more accurate results because it contains more information and is long enough to be specific to the 16s rRNA region.

Despite the important scientific role PCR plays, there can be some significant problems with relying on it for identification of organisms. If great care is not taken during laboratory procedures, samples could be at risk for contaminating DNA. Contamination can lead to DNA amplification in negative controls as well as co-amplification in experimental reactions (Wintzingerode et al., 221). This may produce either ambiguous results when sequencing, or completely false results in gel electrophoresis (Wintzingerode et al., 221). UV treatment and pre-PCR uracil DNA glycosylase digestion are two decontamination systems that have shown to be reliable in PCR experiments (Wintzingerode et al., 221). Another problem with PCR as 16S rRNA sequence analysis is that often environmental samples show a low sequence similarity to known sequences. (Wintzingerode et al., 223). Often environmental samples will

represent uncultured or novel organisms. Many cultivated microorganisms also do not have high quality 16S rRNA sequences available (Wintzingerode et al. 223-224).

The ultimate goal of performing PCR for this experiment was to produce sequences that would provide information of the microorganisms present in our bioreactor sample. It is often necessary to identify unknown organisms in environmental communities and our results demonstrate how difficult it can be to do this accurately. Of the three samples that produced good sequences, all matched closely with species of organisms in the genus Aeromonas. These organisms are facultative anaerobic bacteria. Most species in the aeromonas genus are considered to be pathogens, and infection can result from coming in contact with contaminated water. It is clear that the results of our sequencing are not representative of the total microbial population in the bioreactor sample because we only found one type of organism. It is logical that many environmental microorganisms are not culturable which may explain the lack in variety. Our results were most likely accurate since three separate samples represented organisms in the same genus, but they do not represent all the other organisms that would have originally been present. We checked our sequences based off of the Genbank database. This database contains a vast amount of sequence data but cannot always be considered to be completely accurate. The imperfections of the database have led to the development of an error rate (ER). If one assumes that only entries to the database that have reported discrepancies were verified, the ER associated with the Genbank database is equal to 2.887 errors per 1000 bases (Krawetz, 3951). This error could occasionally lead to inaccurate results when comparing sequences.

Conclusion

Based on this laboratory, it was found that PCR of the 16S rRNA gene can be effectively used to identify an unknown organism in an environmental sample. However, errors in PCR, bad sequences, and small errors in the Genbank database can impact results and produce either no or false identification. Because only species from one specific genus were found in a sample known to contain multiple types of microorganisms, it is clear that there are some limitations to this otherwise useful method of identification.

Questions

1. *For each reactant added to the PCR mastermix, what is its purpose?*

10xPCR buffer- serves as a buffer to maintain pH

25 mM $MgCl_2$ – affects enzyme activity

dNTP mix – free nucleotides used to match with single stranded separated DNA

Primer F- forward transcription primer

Primer R- reverse transcription primer, two different ones provide different sizes

HotStarTaq DNA Polymerase- creates phosphodiester bonds, doesn't degrade at high temperatures

Template DNA- DNA from our pure culture obtained from an environmental sample. DNA that we want to perform PCR

2. *Some researchers use a different annealing temperature in their PCR program. Why?*

The melting temperature of the PCR primer determines the needed annealing temperature. Different PCR primers produce different lengths of nucleotide bases. Because G and C bases contain more hydrogen bonds than A or T bases, a section of DNA with lots of G and C bases would need a higher temperature than one with few G and C bases.

3. *After electrophoresis, the following results were obtained. Explain what might have happened:*

A: The lane loaded with the negative control has a small, faint band.

This could result from either contamination of the negative control or a reverse running of the gel: from positive to negative.

B: The lane loaded with the ladder does not have anything and the entire gel is blank.

This would probably result from a problem with the way electrophoresis was performed. Since you know there is DNA information in the ladder, there was something wrong with the procedure.

C: The lane loaded with the ladder does not have anything, but the lanes loaded with samples display one band for each lane.

You know electrophoresis was performed correctly since you have results for your samples, this means that there was a problem with the ladder, possibly the wrong concentration of ladder was added.

4. *After PCR amplification of total community DNA using a specific primer set, why is it usually necessary to either clone or run DGGE on the products before sequencing them?*
When amplifying a total community's DNA, what appears to be one line in electrophoresis is actually a group of similar but non-identical genes. DGGE and cloning are two methods that successfully separate these genes so that sequencing will result in an individual organism.

5. *Describe one way you could identify organisms that does not involve either culturing of microscopic methods.*

A technique called fluorescent in situ hybridization (FISH) can be used to identify organisms. One method of identifying the presence of certain microorganisms is the use of nucleic acid probes. This method involves a DNA or RNA oligonucleotide that is complementary to a sequence in the target gene. These probes can be dyed fluorescent colors and used to identify organisms containing a complementary sequence.

6. Give an example, not already listed in ch. 18, of how environmental genomics has discovered a known metabolism in a new organism, although that organism has never been cultured.

A good example of this is found in an article describing the identification of a novel group of bacteria. The sample that this bacterium was found in was collected from a deteriorated biological phosphorus removal reactor. It was tested using methods other than direct cultivation. Comparative analysis of the DGGE results of PCR-amplifyed 16s ribosomal DNA was used to identify the present organisms. The partial 16s rRNA sequences indicated that one of the four sequences formed a novel group of the gamma

subclass of Proteobacteria. It had no close relationship to any previously described or sequenced species. This group made up about 75% of the amplified DNA. The bacteria were coccoid shape, 3-4 μm in diameter, and represented approximately 35 % of the total poplation. (Nielsen et al., 1251)

References

Krawetz, S. A., 1989, Sequence errors described in GenBank: a means to determine the accuracy of DNA sequence interpretation, Nucleic Acids Research, Vol. 17 No. 10, p 3951-3957

Madigan, M.T., Martinko, J.M., 2006, *Brock Biology of Microorganisms*, Pearson Prentice Hall, Upper Saddle River, 1992

Muyzer, G., De Waal, E.C., Uitterlinden, A.G., 1993, Profiling of Complex Microbial Populations by Denaturing Gradient Gel Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes Coding for 16S rRNA, Applied and Environmental Microbiology, Vol. 59 No. 3, p. 695-700

Nielsen, A.T., Liu, W., Filipe, C., Grady Jr., L., Molin, S., Stahl, D.A., 1999, Identification of a Novel Group of Bacteria in Sludge from a Deteriorated Biological Phosphorus Removal Reactor, Applied and Environmental Microbiology, Vol. 65 No. 3, p. 1251-1258

Nubel, U., Garcia-Pichel, F., Muyzer, G., 1997, PCR Primers To Amplify 16S rRNA Genes from Cyanobacteria, Applied and Environmental Microbiology, Vol. 63, No. 8, p 3327-3332

Wintzingerode, F.V., Gobel, U.B., Stackebrandt, E., 1997, Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis, FEMS Microbiology Reviews, Vol. 21, p. 213-229