Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Faculty of Natural Resources and Agricultural Sciences

Development of PCR primers for community studies of bacterial hydrogenases in environmental samples

Ndobe Ernest Ndobe

Department of Plant Biology and Forest Genetics, Uppsala BioCenter
Independent project EX0542 • 15 HEC • Advanced E-level
One-years´s Master Programme in Biotechnology • Project report No. 140 • ISSN 1651-5196
Uppsala, Sweden 2013

Development of PCR primers for community studies of bacterial hydrogenases in environmental samples

Ndobe Ernest Ndobe

Supervisor: Associate Professor Per-Olof Lundquist,
Swedish University of Agricultural Sciences (SLU),
Uppsala BioCenter, Dept of Plant Biology and Forest Genetics,
P.O. Box 7080, SE-750 07 Uppsala, Sweden.

Assistant Supervisor: PhD student Enid Ming Zhao,
Swedish University of Agricultural Sciences (SLU),
Uppsala BioCenter, Dept of Plant Biology and Forest Genetics,
P.O. Box 7080, SE-750 07 Uppsala, Sweden

Examiner: Professor Christina Dixelius
Swedish University of Agricultural Sciences (SLU),
Uppsala BioCenter, Dept of Plant Biology and Forest Genetics,
P.O. Box 7080, SE-750 07 Uppsala, Sweden.

Credits: 15 HEC
Level: Advanced E
Course title: Independent project in Biology
Course code: EX0542
Program/education: One-years´s Master Programme in Biotechnology

Place of publication: Uppsala, Sweden
Year of publication: 2013
Title of series: Project reports **No:** 140
ISSN: 1651-5196
Online publication: http://stud.epsilon.slu.se

Key Words: *bacterial community, hupL, nitrogen fixation, PCR, symbiosis, uptake hydrogenase,*

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

The Faculty of Natural Resources and Agricultural Sciences
Department of Plant Biology and Forest Genetics, Uppsala BioCenter

Abstract

Soil microorganisms interact with plants in ways that affect nutrient availability, hormonal regulation and often lead to improved plant growth. In the N_2-fixing root nodule symbiosis between legume plants and bacteria generally called rhizobia the plant gains access to atmospheric nitrogen. In that process H_2 is released from the N_2-fixing rhizobium to the surrounding plant and soil and is used by the microbial community as energy source. The goal of this master's thesis was to design PCR primers targeting the gene for the large subunit of the uptake hydrogenase (*hupL*) that could be used to identify and quantify the part of that community that is able to use H_2. Protein sequences of uptake hydrogenases from three phyla (Proteobacteria, Firmicutes and Actinobacteria) known to interact with plant roots were used for phylogenetic analysis and for designing degenerate primers. Sets of primers for obtaining short and long PCR products were designed for each of the three phyla using the software iCODEHOP. The designed primers were verified against their gene sequences, manually edited and controlled for their quality. Using PCR analysis it was possible to generate PCR products of the expected size of 310 to 330 bp using the primers specific for Proteobacteria and Actinobacteria and an environmental DNA sample as the template. Further studies will show identities and diversity of the *hupL* sequences and the wider usefulness of the designed primers.

Populärvetenskaplig sammanfattning

Vätgas bildas i flera olika miljöer t ex i kvävefixerande symbiotiska rotknölar. Dessa rotknölar bildas av baljväxter efter infektion av bakterier som tillhör en grupp som allmänt kallas rhizobium. När dessa bakterier omvandlar kvävgas (N_2) från luften till ammoniumkväve som värdväxten kan använda som kväve-källa bildas samtidigt vätgas och den passerar ofta ut till omgivande miljön. Jord-mikroorganismer kan dra nytta av denna vätgas om de har enzymet upptags-hydrogenas. För att identifiera vilka bakterier som har denna gen och hur vanliga dessa bakterier är kan man använda polymeras-kedje-reaktionen (PCR) för att amplifiera DNA från prover. Dessa DNA-molekyler kan sedan sekvenseras och genom att jämföras med databaser visa vilka organismer som är de vanligaste att ha denna gen. Målet med detta examensarbete var att designa PCR primers riktade mot genen för den stora subenheten av upptags-hydrogenase (*hupL*). Dessa primers blir då alltså mycket användbara för att identifiera och kvantifiera den del av det bakteria-samhället som har kapacitet att bilda upptags-hydrogenas. Man kan designa PCR primers genom att jämföra DNA sekvenser, eller som i detta fall protein-sekvenser av upptags-hydrogenaser från bakterier som välkända. Regioner i sekvenserna som finns hos alla upptagshydrogenaser i databasen och som är specifika för detta protein används för att avgöra vilka sekvenser PCR primers ska ha. I detta arbete designades tre par av primers som är specifika för tre olika grupper av bakterier. De primers som designats i detta arbete kan sedan användas för olika slags prover för att studera förekomsten av genen för upptags-hydrogenas. För studier av kvävefixerande symbiotiska rotknölar kommer de att kunna bidra till en ökad förståelse för hur symbiosen fungerar och vilka organismer som är inblandade.

Contents

1. Introduction 6

1.1 H_2 production in N_2-fixing root nodules 6

1.2 The structural composition of the hydrogen uptake gene 6

1.3 Classification of hydrogenases based on their metal content 7

1.4 Plant-bacteria associations 7

1.5 Identification of bacteria 8

1.6 PCR and primers 9

1.7 Aims of this study 10

1.8 Significance of this work 10

2. Materials and Methods 10

2.1 Database searches 10

2.2 Alignments and phylogenetic analysis 10

2.3 Primer design 11

2.4. DNA extraction 11

2.5 PCR analysis 11

3. Results 13

3.1 Phylogeny of uptake hydrogenases 13

3.2 Design of degenerate PCR primers 14

3.3 PCR using the designed primers 15

4. Discussion 16

5. Conclusions 18

6. Acknowledgement 18

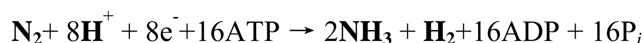
7. References 18

1. Introduction

1.1 H_2 production in N_2-fixing root nodules

In several ecological situations of Bacteria, Archaea and Eucaryae, in e.g. marine and freshwater environments (Barz et al. 2010) hydrogen gas (H_2) is part of energetic metabolism either as a product during anaerobiosis or as energy source. Hydrogenases play an important role in hydrogen metabolism and some enzymes can function in a reversible mechanism and others in the heterolytic breakdown of the H_2 according to: $H_2 \rightarrow 2H^+ + 2e^-$ as is the case of uptake hydrogenases.

Biological N_2 fixation carried out by rhizobia in the root nodules of leguminous plants is a key process that leads to the production of NH_3 to the plant. During the N_2 fixation reaction catalyzed by the nitrogenase enzyme, H_2 is produced as an inevitable end product. A simple reaction stoichiometry for N_2 fixation and the production of H_2 by the nitrogenase enzyme is:



$$\mathbf{N_2} + 8\mathbf{H}^+ + 8e^- + 16ATP \rightarrow 2\mathbf{NH_3} + \mathbf{H_2} + 16ADP + 16P_i$$

The production of H_2 requires the use of 4 ATP molecules (Dixon 1972) and represents a significant cost. If H_2 is lost to the surrounding soils this thus causes a reduction in the efficiency of the N_2 fixation system (Maier and Triplett 1996). However, it has been observed that some rhizobia, for example *Rhizobium leguminosarum* bv. *viciae* UPM791 (Albrecht et al. 1979), do have the uptake hydrogenase enzyme that is capable of capturing the H_2 produced during nitrogen fixation. This enzyme is encoded by the uptake hydrogenase gene (*hup*) that is present in the genomes of some rhizobia (Ruiz-Argueso et al. 1989). However, although they may have the gene for uptake hydrogenase the level of expression varies among rhizobia in the symbiotic stage. The H_2 produced in many legume nodule symbioses is therefore released from the rhizobium which makes the N_2 fixation process less efficient.

1.2 The structural composition of the hydrogen uptake gene

The hydrogen uptake genes that are involved in the synthesis of an active hydrogenase are composed of 18-24 gene clusters (Ruiz-Argueso et al. 2000). For example in *Rhizobium leguminosarum* the gene clusters are ***hup***SLCDEFGHIJK***hyp***ABFCDE***hox***X (Baginsky et al. 2002). The functions of each cluster vary with the bacterial type. For example in *Rhizobium leguminosarium* the *hup* gene cluster is important in giving the enzyme structure, whereas the *hyp* and the *hox* gene clusters function in regulating the enzyme activity by recruiting and

incorporating nickel and some other metal groups into the active site of the hydrogenase enzyme (Baginsky et al. 2002). The *hupL* and *hupS* genes encode the large and the small subunits of uptake hydrogenase, respectively.

1.3 Classification of hydrogenases based on their metal content

There exist three major categories of hydrogenases and whose classification is based on the metal content in the active site of the enzyme (Vignais and Billoud 2007). The three major categories of hydrogenases are NiFe-, FeFe- and the Fe-hydrogenases. The NiFe-hydrogenase is common in Archaea and Bacteria. A sub-type of NiFe-hydrogenase is the NiFeSe hydrogenase. The Fe-hydrogenase is common among the Archaea, whereas the FeFe-hydrogenase is common in anaerobic eukaryotes and bacteria. The chemical structure of the FeFe and the NiFe is such that they do have a metalloorganic cluster that is essential for the transfer of electrons from the H_2 to the enzyme active site. These Fe-S clusters occur at long distant to each other (Shima et al. 2010). However in the Fe hydrogenases the metalloorganic clusters are absent, therefore conduction of electrons from hydrogen is done directly to the enzyme active site through short distances. The NiFe hydrogenases are mostly uptake hydrogenases, the Fe hydrogenases are mostly methanogens, while Fe-Fe are mostly H_2 evolving and they are sensitive to O_2.

1.4 Plant-bacteria associations

During N_2-fixation, H_2 produced in some root nodules leak out to the soils which may lead to increased soil metabolism and CO_2 sequestration in soils (Dong and Layzell 2001). Dong et al. (2009) performed a series of experiments in the greenhouse, lab and in the field on different soil types treated with H_2. The aim of their investigation was to observe the manner in which H_2 introduction in the different soils would have on microbial life. They found that the soils treated with H_2 contained large communities of H_2-oxidizing bacteria and the control soils treated with air had no H_2–oxidizing bacteria. Later, they used the 16S rDNA terminal restriction fragment technique to assess the microbial community structure of the different soils and found that growth of some bacterial species was stimulated by H_2 exposure while others were inhibited (Zhang et al. 2009). To assess the bacterial community that colonized the H_2-treated soils three different bacterial species were identified; *Burkholderia*, *Flavobacterium* and *Variovorax* (Dong et al. 2007). To investigate the nature of the

interaction between the different bacterial strains and plants using laboratory techniques they inoculated *Burkholderia* and *Variovorax* strains to *Arabidopsis thaliana* plants and found that the plant biomass increased by 11-27% compared to the control. A similar experiment on the spring winter wheat crop caused an increase in root growth by 57 to 254%. The results of their findings therefore made them to conclude that H_2-oxidizing bacteria enhance soil fertility (Dong et al. 2007). Baginsky et al. (2005) assessed the H_2-oxidation rates of different Bradyrhizobium strains with a specific legume host plant. They observed that inoculants of different strains to a specific legume host yielded different H_2 uptake abilities, thus that different plant-bacterial associations vary in their H_2 uptake abilities. In a related experiment, they mutated the *hupL* gene in the bacteria, inoculated them into their legume host plant and observed that the H_2-uptake rates stopped and the nodules evolved H_2.
Soil bacteria in general interact with plants in many ways independent of H_2 metabolism. Some bacteria improve the ability of the plants to assimilate nutrients from the soil by stimulating root development and as well as root colonization by the beneficial root symbiosis with arbuscular mycorrhizal fungi (Bharadwaj et al. 2008) that improves nutrient uptake, stress resistance and pathogen resistance. Soil bacteria associated with plant roots may also increase the availability of essential nutrients like phosphorous from the soil by aiding in their solubilization (Kuklinsky-Sobral et al. 2004, Bharadwaj et al. 2008). Some plant-associated bacteria may also function as biocontrol agents against fungal and bacterial plant diseases (Bharadwaj et al. 2008). Because plant-bacteria associations have shown to influence the plant's growth, it is therefore important to identify the bacterial community involved. This may help us to exploit some of the existing traits in the bacteria to improve the plants health.

1.5 Identification of bacteria

Studies of microbial communities and identification of bacteria has long depended on techniques that are based on isolation and *in vitro* growth of free-living bacteria. However, these techniques have been shown to be limited in covering the full diversity of bacteria. The use of DNA-sequence based methods for identification of bacteria has proven to be very reliable and efficient when compared to the cultivation-dependent technique (Vilgalys 2011). In studies of bacterial communities recent applications of deep-sequencing techniques such as pyrosequencing have revealed highly diverse communities (e.g. Handl et al. 2011). In identification studies and community studies it is mostly the 16S rRNA gene that has been exploited. The use of the 16S rDNA gene in microbial identification is reasonable because it

is present in all organisms and a large amount of information has become available over time for Bacteria and Archaea. The gene has highly conserved regions while other regions are less conserved and therefore random changes in its sequences over time may be good to measure evolution. This gene is 1500 bp long making it large enough to provide information (Patel 2001). The identification of bacteria using sequences like the 16S rDNA is straightforward as there are several PCR primers developed and well tested.

To study diversity of functional traits based on presence of certain enzymes requires designing primers using protein-coding sequences. This is challenging as the sequences of such genes of different organisms vary because of the degenerate genetic code in addition to larger evolutionary changes. However, by studying functional genes at the community level it is possible to obtain information regarding which organisms contribute certain functions in the communities. Special tools are needed to design primers to these coding genes to make them valid and hopefully include amplification of homologous genes of as yet unidentified organisms.

1.6 PCR and primers

The polymerase chain reaction (PCR) is a technique that targets a specific region in a DNA for amplification (Mullis et al. 1986). To conduct a PCR it is important that the reaction mixture contains the DNA template to be amplified, DNA polymerase, primers (oligonucleotides), deoxyribonucleotides, and that the reaction is set at an annealing temperature suitable for the forward and reverse primers. Optimal annealing temperatures of the primers should be investigated by doing gradient PCR. Also, because the PCR yields are very sensitive to the DNA template concentration it would be necessary to investigate effects of DNA template concentrations. With regards to the PCR primers that should target a sequence from a range of organisms it is needed to design degenerate primers. Such primers have certain positions in their sequences with more than one nucleotide and are in practice a mix of primers (Linhart and Shamir 2005).

Protein-coding gene sequences are more variable among organisms compared to other sequences used in taxonomy such as the 16S rDNA due to that their biological function is at the protein level and the genetic code is degenerate. It is therefore less likely to obtain conserved DNA sequences to design primers when based on protein-coding genes. Also, if the goal is to possibly identify organisms not yet in any sequence databases it should stand a higher chance if conserved regions in protein sequences are used compared to DNA

sequences. Therefore to develop primers for such coding sequences it should be best to use protein sequences as starting point for construction of degenerate PCR primers and use special softwares such as iCODEHOP (Boyce et al. 2009).

1.7 Aims of this study

The aim of this work was to design degenerate primers for the *hupL* gene encoding the large subunit of uptake hydrogenase to be used for studies of bacterial communities.

1.8 Significance of this work

The work will provide a non-cultivation-dependent tool for further studies of bacterial communities that can give knowledge about identity and diversity of the bacteria that carry the *hup* gene in environmental samples such as symbiotic root nodules. By knowing the identity of such bacteria it will provide a starting-point for studies of detailed functions and significance of those particular organisms.

2. Materials and Methods

2.1 Database searches

The search term ”large subunit uptake hydrogenase” was used to search the UniProt protein sequence database (www.uniprot.org) in the phyla Proteobacteria, Firmicutes, Actinobacteria and Cyanobacteria, respectively. Sequences were selected so that the sequences represented diverse sub-taxa distributed over each phylum. The obtained sequences were retrieved in FASTA format.

2.2 Alignments and phylogenetic analysis

Sequences for each phylum were aligned using the multiple sequence alignment tool ClustalW and phylogenetic analysis were conducted using MEGA version 5 (Tamura et al. 2011). The evolutionary history was inferred using the Neighbor-Joining method and bootstrapping. The analysis involved 44 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 472 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al. 2011).

2.3 Primer design

Degenerate primers were designed using the software iCODEHOP (Boyce et al. 2009) http://dbmi-icode-01.dbmi.pitt.edu/i-codehop-context/iCODEHOP/view/PrimerAnalysis at default settings except for the codon usage table which was selected based on the phylum. The codon usage tables of *Sinorhizobium meliloti*, *Bacillus subtilis*, *Streptomyces coelicolor* and *Synechococcus* sp were used for Proteobacteria, Firmicutes, Actinobacteria and Cyanobacteria, respectively. From the series of forward and reverse primers suggested by iCODEHOP, primer pairs were selected to have to have good scores to improve the chances of specific primer–template binding, a degeneracy below 64, similar annealing temperatures, a GC content of 40 to 60% and lengths in the range of 18-30 bp. One set of primer pairs should when used in PCR yield a product size of 250 to 400 bp and another an almost full-length gene product of 1700 bp.

2.4. DNA extraction

Root nodules of groundnut plants inoculated with field soil were used as an environmental sample. DNA from 157 mg root nodules was extracted using the FastDNA spin kit for soil (MP Biomedical, USA). The nodules were homogenized in the FastPrep machine (Fp 120, BIO 101 Savant) for 40 sec, run at a speed of 6.0 and the homogenization was repeated 3 times. The homogenized sample was then extracted according to the protocol of the manufacturer. The DNA was finally eluted with 50 µl nuclease-free water, its concentration measured using a Nanodrop spectrophotometer and stored at +4 ^{o}C.

2.5 PCR analysis

All PCR reactions were run in a MyCycler PCR machine (BioRAD). The PCR reactions had a total volume of 20 µl and contained 1X Ex Taq buffer, 0.2 mM dNTP, 0.5 µM of each of the primers and 0.025 units ExTaqHS DNA polymerase (hotstart Taq polymerase; Takara, Japan). The amount of template DNA was in the range 0.1, 1, 10, 100 and 1000 ng per reaction. The wide range of template DNA was used because the DNA extract is a mixture of plant DNA, rhizobial DNA and other unidentified bacterial DNA. Negative controls in which no DNA was added were used for all primer pairs.

To test suitable annealing temperatures gradient PCR was carried out. To choose the range of annealing temperatures for the gradient PCR, the melting temperatures (Tm) for the

primers were calculated using the software OligoAnalyzer (http://eu.idtdna.com/analyzer/applications/oligoanalyzer/) set to consider the presence of 2 mM Mg^{2+}. The annealing temperatures that were used were approximately 3˚C below the Tm for the different species of the degenerate primers. The gradient PCR was run for each primer pair in the range of 55 to 62 °C. Three identical PCR reactions for each primer pair were prepared by setting up 60 µl reactions that were divided into three 20 µl reactions and each of them were placed in the PCR machine so that they experienced either high, medium or low annealing temperatures.

The temperature cycling conditions for the short product PCR were; 95 °C initial denaturation and inhibitor inactivation for 4 mins, then followed by 30 cycles of denaturation at 95 °C for 30 sec, annealing in the range 54 to 66 °C for 30 sec and extension at 72 °C for 30 sec followed by a final extension at 72 °C for 10 min. The reaction cycle for the long product PCR was the same except for that the extension time was for 100 sec.

To investigate the optimal DNA concentration, PCR was run in total volumes of 50 µl and with 0.1, 1, 10, 100 and 1000 ng DNA per reaction. The concentrations of reagents was as above and only the proteobacterial short primers were investigated with temperature cycling conditions as above for the short PCR with an annealing temperature of 58°C.

To investigate the presence of the *hupL* gene in the environmental sample used in this study the three sets of primers designed to target *hupL* of bacteria in the phyla Proteobacteria, Firmicutes and Actinobacteria, respectively, was carried out. The sets of primers for obtaining short products were used. Based on the calculations of the different melting temperatures and the results of the gradient PCR the annealing temperatures were set to 62 °C for the proteobacterial primers, 54 °C for the firmicute primers and 58 °C for actinobacterial primers. The amount of DNA per 50 µl reaction was 10 ng DNA. The PCR reactions were fractionated by gel electrophoresis in a 1.5% agarose gel run at 100 V and with 4 µl of PCR products loaded into each well.

3. Results

3.1 Phylogeny of uptake hydrogenases

A search for uptake hydrogenase in the UniProt database was carried out for Proteobacteria, Firmicutes, Actinobacteria and Cyanobacteria, respectively. Within each of these phyla a subset of sequences from diverse sub-taxa was retrieved. In total 44 sequences, of Proteobacteria (15), Firmicutes (6), Actinobacteria (16) and Cyanobacteria (5) were aligned by ClustalW. A phylogenetic tree was inferred from the uptake hydrogenase protein sequences (Fig. 1). The phylogeny of the uptake hydrogenase protein sequences shows a high diversity among the taxa and that it followed the taxonomy of the organisms.

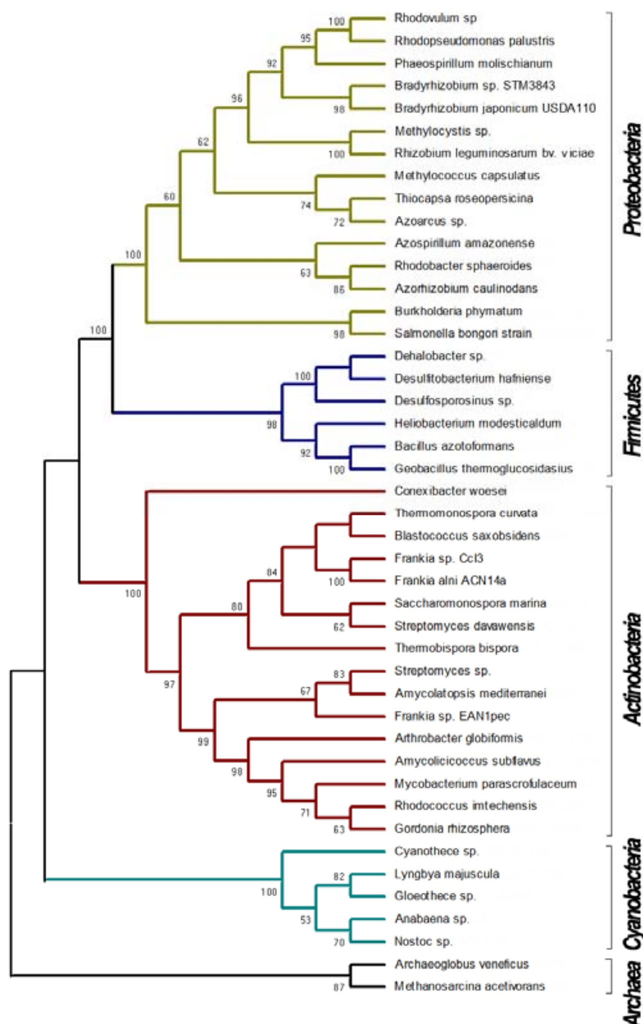


Fig. 1. Phylogeny of uptake hydrogenase protein sequences in Proteobacteria, Firmicutes, Actinobacteria and Cyanobacteria.Protein sequences were retrieved from the UNIPROT database and analyzed in MEGA5. All the above uptake hydrogenases are classified as large subunit NiFe/NiFeSe-hydrogenases. The phylogeny was inferred using the Neighbor-Joining method and bootstrapping and shown as topology tree. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches if >50.

3.2 Design of degenerate PCR primers

The alignments of uptake hydrogenase large subunit protein sequences showed several conserved domains as exemplified by the aligned sequences of Proteobacteria (Fig. 2). Several conserved domains were thus available to be used for designing primers. A higher number of conserved domains were observed among the proteobacterial sequences than among the cyanobacterial sequences.

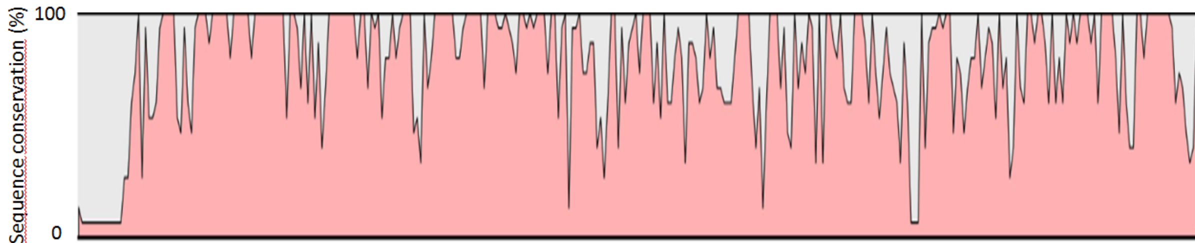


Fig. 2. *Sequence conservation plot of the alignment of the 15 proteobacterial uptake hydrogenase protein sequences used in this study. The plot shows the percent of the sequences having identical amino acids in the same position.*

The iCODEHOP software successfully designed several primers for the Proteobacteria, Firmicutes and Actinobacteria. It failed to suggest any primer for the Cyanobacteria because the sequence alignment sequences from ClustalW had low sequence conservation. From the suggested primers, pairs were chosen that would give the desired amplification product lengths. i.e. primers giving short and long products of the *hupL* gene sequence in the bacteria. The primer sequences were verified against DNA sequences of the same organisms. The primer sequences were adjusted to reduce degeneracies and make the 3' ends start with a C or G and this resulted in a set of forward and reverse primers targeting each the three taxa (Appendix). The primers were further analyzed using the OligoAnalyzer software for features like primer dimerization, hairpin structures, %GC content and melting temperatures. The calculated melting temperatures of the primer pairs designed for these three phyla were found to be different between the phyla. Therefore, in the PCR reactions that were carried out the annealing temperatures for each primer pair were chosen to accommodate to those differences in annealing temperatures.

3.3 PCR using the designed primers

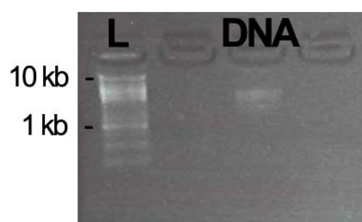


Fig. 3. *Agarose gel showing the presence of DNA in the root nodule extract. L= lane for the ladder and DNA= the lane for the root nodule DNA extract.*

Total DNA from root nodules (Fig. 3) was used as template in PCR reactions. Gradient PCR using annealing temperatures in the range of 58 to 62 ^{o}C showed amplification of a short PCR product of the expected size ca 330 bp when the primer pairs designed to obtain a short PCR product was used (data not shown).

The amount of DNA template had a clear effect on the amplification result (Fig. 4) and a clear amplification product of about 330 bp was only obtained when the amount of template DNA was 10 ng per 50 µl PCR reaction (Fig. 4). Less abundant larger size products were also seen. At higher amounts of DNA template there were smears of amplified DNA.

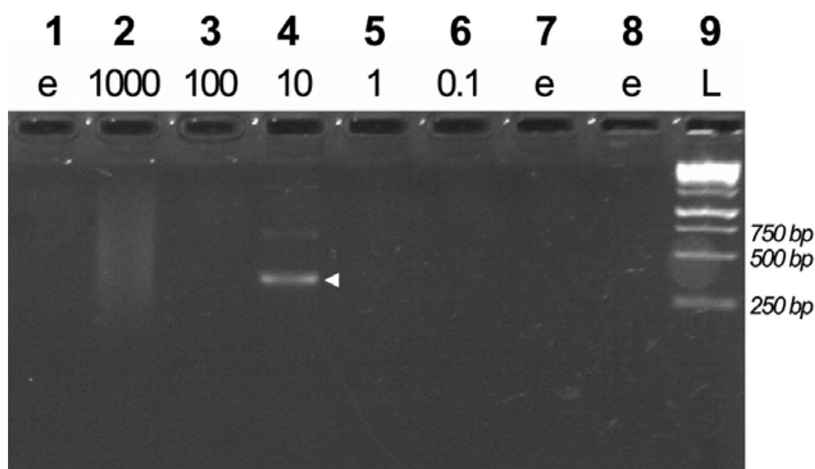


Fig. 4. *Effect of amount of template DNA on PCR amplification of the hupL gene using the primers designed for proteobacteria. Wells 2 to 6 in the agarose gel contained PCR reactions in which 1000, 100, 10, 1 and 0.1 ng DNA, respectively, was used as template in the PCR. Wells 1, 7 and 8 had no sample added to them. The L= lane for the DNA size marker GeneRuler 1 kb DNA Ladder. Lane 4 with a DNA template of 10 ng per reaction showed a most abundant product of about 330 bp indicated by arrow.*

To evaluate the three separate sets of primers for Proteobacteria, Firmicutes and Actinobacteria, respectively, for the short PCR amplified product, a PCR was set up using 10 ng root nodule DNA as template. The results show that conducting PCR with annealing temperature fixed at 58°C and amount of DNA template fixed at 10 ng per 20 µl reaction gave the desired product size of approximately 310-330 bp for Proteobacteria and for Actinobacteria (Fig. 5), but not for Firmicutes. In the Actinobacteria the PCR reaction generated two additional larger sized products.

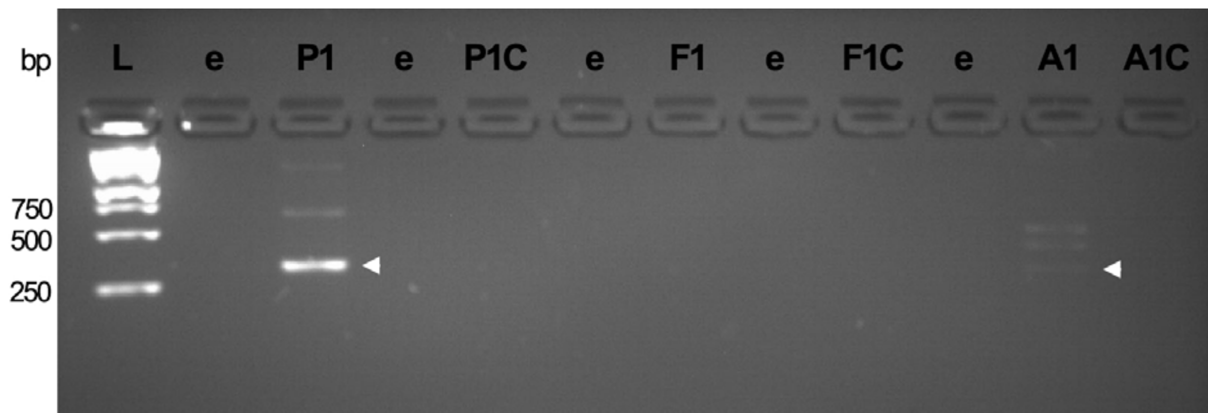


***Fig. 5**. Agarose gel showing the results from the PCR run with primers designed to target a short product (310-330bp) of the hupL gene. The three primer pairs were designed specifically for proteobacteria, firmicutes and actinobacteria, respectively. The results are from PCR using the annealing temperature of 57.2 °C and DNA template concentration of 10 ng per 50 µl. The L= lane for the DNA size marker GeneRuler 1 kb DNA Ladder, P1= is lane for proteobacterial primers, P1C= is lane with proteobacterial primers control with no DNA template, F1= is lane for firmicute primers, F1C= is lane with firmicute primers with no DNA template, A1= is lane for actinobacterial primers and A1C= Actinobacterial primers control with no DNA template, e= empty wells.*

4. Discussion

Since a large fraction of microorganisms are not possible to cultivate *in vitro* it is necessary to use other methods to obtain a better view of the communities. Bacterial communities associated with plants can be identified by using ubiquitous and conserved genes such as the 16SrDNA gene. However, to begin to understand the roles played by these bacteria it is necessary to target genes encoding certain functions such as enzyme activities and to try to identify the bacteria that carry those genes. Gene and protein sequence databases offer tremendous information that can be used to design PCR primers. However, an organism not yet characterized may have additional sequence diversity. Due to the higher conservation of protein sequences compared to DNA sequences it is an advantage to use protein sequences as a starting point rather than DNA sequences when designing primers that should be used to

investigate the full diversity of bacteria carrying the *hupL* gene, including trying to discover the presence of unknown bacteria carrying the *hupL* gene.

The uptake hydrogenase enzyme makes it possible for organisms to use H_2 and is widespread among various bacterial phyla. The phylogeny of uptake hydrogenase protein sequences (Fig. 1) showed that this gene is diverse and that the phylogeny is consistent with the taxonomy of the organisms. For that reason it was decided to design primers that were specific for each of the three phyla.

Based on the alignments, iCODEHOP results and manual verification in relation to DNA sequences for several available bacterial genomes of each phylum, it was for each of the three taxa possible to design one forward and two reverse primers. The domains in the protein that were used to design the primers were located in three domains which have high conservation among all uptake hydrogenases. One region was in the N terminus, one in a domain about 100 amino acids downstream from the first domain and finally one domain in the C terminus. The forward primer in pair with either of the reverse primers would thus theoretically yield a short PCR product and a long PCR product.

The primer pairs giving the short PCR product were initially chosen to be used in PCRs to investigate suitable DNA concentrations and annealing temperatures. Short products were obtained, which could be due to primer matching in the sample, but also very likely be due to the higher amplification efficiency in PCR for short products than for long products (Taqman Life technologies 2012). Further laboratory work on optimal primer annealing, product extension and DNA template concentration should help to improve the PCR amplification reaction.

Using the primers designed in this study and the DNA extracted from root nodules as template it was possible to obtain amplification products for the proteobacterial and actinobacterial primers. Since the root nodules are formed in the interaction with the symbiotic N_2-fixing bacterium *Bradyrhizobium* sp*,* and that *Bradyrhizobium* is very likely the most abundant bacterium in these root nodules, it is not surprising that an amplification product was obtained for the proteobacterial primers. The amplification product obtained for the actinobacterial primers suggest that actinobacteria carrying *hupL* could be present in the root nodules. Further work may show their identity and make possible to evaluate the roles of *hupL* and these bacteria in the root nodules.

5. Conclusions

Three sets of primers based on the large subunit of the uptake hydrogenase protein sequences from Proteobactera, Actinobacteria and Firmicutes were designed to amplify a short PCR product of about 310 to 330 nt. Using an environmental sample (a N_2-fixing root nodule of a plant grown with field soil) as template it was possible to obtain PCR products of the expected size when using the primers designed based on proteobacterial and actinobacterial uptake hydrogenase. Further studies are needed to identify the bacteria that carry *hupL* in this sample and to evaluate the roles of *hupL* and these bacteria in the root nodules.

6. Acknowledgement

Many thanks go to my supervisor Per-Olof Lundquist for his helpful explanations, discussion and in the laboratory work. I will also like to thank the co-supervisor Enid Ming Zhao for her helpful explanations and also guiding me in the laboratory work. Many thanks also go to my mum, Ndobe Hannah Mesode and to my father, Ndobe Manfred Mpako for assisting me in all aspects in my endeavors. The last but not the least is to thank my beloved wife for the wonderful job she did in caring for our two children Ndobe Favour Mesode and Ndobe Joel Mpako during these years I have spent abroad for studies.

7. References

Albrecht, S. L., Maier, R. J., Hanus F. J., Russell, S. A., Emerich, D. W. and Evans, H. J. 1979. Hydrogenase in *Rhizobium japonicum* increase nitrogen fixation by nodulated soybeans. Science 203: 1255-1257

Baginsky, C., Brito, B., Imperial, J., Palacios, J. M. and Ruiz-Arguso, T. 2002. Diversity and Evolution of hydrogenase systems in rhizobia. Applied and Environmental Microbiology 68: 4915-4924

Baginsky, C., Brito, B., Imperial, J., Ruiz-Arguso, T. and Palacios, J. M. 2005. Symbiotic hydrogenase activity in *Bradyrhizobium sp* (Vigna) increases nitrogen content in *Vigna unguiculata* plants. Applied and Environmental Microbiology 71: 7536-7538

Barz, M., Beimgraben, C., Staller, T., Germer, F., Opitz, F., Marquardt, C., Schwarz, C., Gutekunst, K., Vanselow, K. H., Schmitz, R., LaRoche, J., Schulz, R., Appel, J. 2010.

Distribution analysis of hydrogenases in surface waters of marine and freshwater environments. PLoS ONE 5(11): e13846

Bharadwaj, D.P., Lundquist, P.-O. and Alström S. 2008. Arbuscular mycorrhizal fungal spore-associated bacteria affect mycorrhizal colonization, plant growth and potato pathogens. Soil Biology & Biochemistry 40: 2494-2501

Boyce, R., Chilana, P. and Rose, M. 2009. iCODEHOP: a new interactive program for designing Consensus-Degenerate Hybrid Oligonucleotides Primers from multiply aligned protein sequences. Nucleic Acids Research 37 (web server issue): W222-W228

Dixon R.O.D. 1972. Hydrogenase in legume root nodule bacteroids: occurrence and properties. Archives of Microbiology 85: 193-201

Dong, Z. He, X, and Zhang, Y. 2009. Effect of hydrogen on soil bacteria community structure in two soils as determined by terminal restriction fragment length polymorphyism. Plant and Soil 320: 295-305

Dong, Z. and Layzell, D. B. 2001. H_2 oxidation, O_2 uptake and CO_2 fixation in hydrogen treated soils. Plant and Soil 229: 1-12

Dong, Z. and McLearn, N. 2002. Microbial nature of the hydrogen-oxidizing agent in hydrogen-treated soil. Biology and Fertility of Soils 35: 465-469

Dong, Z., Peoples, M., Layzell, D. B., Cen, Y-P, Yang, J., Zhang, Y. and Maimaiti, J. 2007. Isolation and characterization of hydrogen–oxidizing bacteria induced following exposure of soil to hydrogen gas and their impact on plant growth. Environmental Microbiology 9: 435-444

Evans, J. H. and Schubert, R. K.1976. Hydrogen Evolution: A major factor affecting the efficiency of nitrogen fixation in nodulated symbionts. Proceedings of the National Academy of Science of the United States of America. 73:1207-1211

Handl, S., Dowd, S.E., Garcia-Mazcorro, J.F., Steiner, J.M. and Suchodolski, J.S. 2011. Massive parallel 16S rRNA gene pyrosequencing reveals highly diverse fecal bacterial and fungal communities in healthy dogs and cats. FEMS Microbiology Ecology 76: 301-310

Kuklinsky-Sobral, J. Araujo, W. L., Mendes, R., Geraldi, I. O., Pizzirani-Kleiner, A. A., Azevedo, J. L. 2004. Isolation and characterisation of soyabean-associated bacteria and their potential for plant growth promotion. Environmental Microbiology 6: 1244-1251

Linhart, C. and Shamir, R. 2005. The degenerate primer design problem: theory and applications. Journal of Computational Biology 12: 431-456

Maier, R. J. and Triplett, E. W. 1996. Towards more productive, efficient and competitive nitrogen fixing symbiotic bacteria. Critical Reviews in Plant Science 15: 191-234

Mullis, F.K., Scharf. S., Saiki, G., Horn and Erlich, H. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor Sympiosa on Quantitative Biology 51: 263-273

Patel, J. B. 2001. 16S rRNA gene sequencing for bacteria pathogen identification in the clinical laboratory. Molecular Diagnostic 6: 313-321

Ruiz-Argueso, T., Imperial, J. and Palacios, J. M. 2000. Uptake hydrogenases in root nodule bacteria. p. 489–507. In E. W. Triplett (ed.), Prokaryotic nitrogen fixation: a model system for analysis of a biological process. Horizon Scientific Press, UK.

Ruiz-Argueso, T., Murillo, J., Villa, A. and Chamber, M. 1989. Occurrence of H_2-uptake hydrogenases in *Bradyrhizobium* sp. (Lupinus) and their expression in nodules of lupines spp. and *Ornithopus compressus*. Plant Physiology l89:1781-85

Shima, S., Vogt, S., Göbels, A. and Bill, E. 2010. Iron-Chromophore Circular Dichroism of (Fe)-hydrogenases. The conformational Change Required For H_2 Activation. Angewandte Chemie International edition. 49: 9917-9921.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28: 2731-2739

Taqman life technologies. 2012. Amplification efficiency of the TaqMan Gene Expression Assays. Application Note.

Vignais, P. M. and Billoud, B. 2007. Occurrence, Classification and Biological function of hydrogenases: An overview. Chemical Reviews 107: 4206-4272

Vilgalys, R. http://www.biology.duke.edu/fungi/mycolab/primers.htm. Accessed online 2013-07-07.

Zhang Y., He X. and Dong Z. 2009. Effect of hydrogen on soil bacterial community structure in two soils as determined by terminal restriction fragment length polymorphism. Plant and Soil 320: 295-305

ISSN: 1651-5196 Nr 140
Uppsala 2013

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

The Faculty of Natural Resources and Agricultural Sciences
Department of Plant Biology and Forest Genetics
Uppsala BioCenter
Box 7080
SE-750 07 UPPSALA
Sweden