DNA Fingerprinting for identification of British tree species

Amy Finlayson
Dalkeith High School

Nuffield Science Bursary Project Report
Centre for Ecology & Hydrology, Biosystems Management Section

Quercus palustris leaves (Pin Oak)

Contents

1. Abstract

The aim of the project was to produce a database of the DNA fingerprints of a large variety of native and naturalised British tree species. Various techniques and procedures were used to set up the database. Leaves from a large range of different species of tree were collected from the local area and preserved by drying on silica gel. DNA was then extracted from each sample. Using PCR (polymerase chain reaction) a short fragment of DNA from the chloroplast genome was amplified. Next, the DNA fragments were cut up into smaller lengths by incubation with restriction enzymes. These short fragments were then visualised by polyacrylamide gel electrophoresis (PAGE), which separates fragments out according to their size (length of DNA sequence). By staining the PAGE gel with a DNA-binding dye and illumination under UV light, the fragment patterns can be seen and photographed as a gel image. Interspecific differences, due to differences in DNA fragment length and sequence between species, can then be taken from the gel image. Once each sample had been fingerprinted, each of the digested fragments from the gel image was sized, to create a database for species identification.

2. Introduction

Background Information

In Britain, there is a huge variety of tree species growing naturally and being cultivated. Not only are there 63 native species and over 22 naturalized species (Johnson & More, 2005), there are many hybrids and varieties growing across Britain. In Europe, there exists over 6000 species of tree.
Every living thing has DNA in its cells. In plants, DNA exists primarily in the nucleus of cells but also in the mitochondria and chloroplasts (the organelles) of plant cells. In this project, the target was the chloroplast DNA molecule ($_{cp}$DNA).
DNA is a unique code made of 4 possible bases: Guanine, Thymine, Adenine and Cytosine (GTAC). The DNA molecule is a double stranded helix with each strand containing thousands of bases: each base on a strand matches a complementary base on the other strand (Guanine always pairs with Cytosine, and Adenine always pairs with Thymine: known as base pairs). In any species, bases occur in a unique order and this is the sequence referred to in this report. Most trees within a species will have very similar codes with long stretches of identical base pairing because the sequence ultimately leads to protein synthesis and the specific proteins which are produced depend on the order of the bases. So individual trees within a species will have similar proteins synthesised and therefore have a similar DNA sequence. However, due to random mutations, every individual tree will have slight differences in its DNA sequence, usually just 4 or 5 different bases out of a hundred. Between species, there is a lot more genetic variation. The analysis of this genetic diversity and determining how similarities and differences in DNA sequences reveal relatedness is the central focus of many biological scientists. There are numerous techniques now that allow us to sequence DNA and find variations and similarities. A great deal of effort has been made in recent years to sequence DNA and discover more and more about genetics interspecifically (between species) and intraspecifically (within species), and the complete genomes of many organisms have now been sequenced. Therefore the range of information available to scientists is vast and growing.

Aims

The project aims to:

- Use chloroplast DNA to produce a fingerprint for important UK tree species
- Produce a database of the fingerprints of these species
- Be able to use the database to identify unknown species
- Ensure each fingerprint is unique and distinguishable from others

Applications

The database which was produced can be used for various things. Trees in urban environments can sometimes be pests. Roots can damage the foundation of housing and trees may need to be cut down to avoid collapsing of foundations, but how can you know which tree is the problem? A sample would be taken of the root in question and after following the process of DNA fingerprinting and coming up with a gel image of the fragments, the database produced in this project will tell you what species of tree the root in question comes from. It can then be cut down safely. The other side to that is if someone owns a tree privately and wants to prove it is not a problem to avoid losing it. Also, in the case where a branch falls off a tree and causes damage, the tree responsible may have to be identified to assign liability for the insurance claim. The database could provide a match to which tree the branch fell from.

Other applications include police forensics. For example, if a criminal is being traced and there are pollen grains on his clothes or skin, the DNA from these can be extracted and fingerprinted. Then, using the database, the particular species the pollen came from can be identified. This is sound proof that the criminal has been in a particular flowerbed or under a specific tree. This could help in identifying for example, who broke entry into the house with the pollinating Lawson cypress hedge along its perimeter.

There will never be 100% proof with DNA fingerprinting as, even within a species, there can be variation in DNA sequence and it could be hard to find a match. Also, the tree in question could be some kind of hybrid variant and not match up to anything on the database. However, the closest match can be identified and then you would know what you’re looking for, for example, a variant of white willow will still have many characteristics of the white willow itself. Normally, this level of proof is sufficient to resolve most problems in a professional way. As the process is a quick and easy way to sequence a tree’s DNA and the database has been focussed on the species which most often cause problems, it can be useful in a wide range of cases.

3. Methodology

3.1 DNA Sampling

A large range of species and a number of samples per species were required to be sampled so young leaves were collected from a variety of species.

Samples were taken from various points around the CEH building and Bush Estate and from local botanical gardens in Edinburgh. Trees which were sampled were ideally spread around randomly and when repeats were done, the samples were not taken from the same tree.

It is quite important that the leaves taken are young because they contain more cells per weight and contain less polysaccharides and polyphenolics than older leaves which build up these chemicals over time. These chemicals can interfere with the extraction process and make things difficult so it is a lot easier to try to collect as young leaves as possible. Once each leaf had been taken, it was placed immediately into a sealed plastic bag containing silica gel to dry it out (see **Figure 1)**. This preserves the DNA and stops the leaf from degrading too much. Once a set of samples had been collected, they were taken out of the silica gel and cut using a scalpel into small squares. They were then weighed using a precise balance. The weighed sample was then placed into a mortar along with a small amount of sand. A pestle was used to grind it up until it consisted of a very fine powder (see **Figure 2)**. This powder was then placed carefully by use of folded filter paper into a plastic eppendorf tube (see **Figure 3**) and stored immediately in the freezer. This is once again to avoid degradation and preserve the DNA.

Figure 1. Sample bag containing silica gel

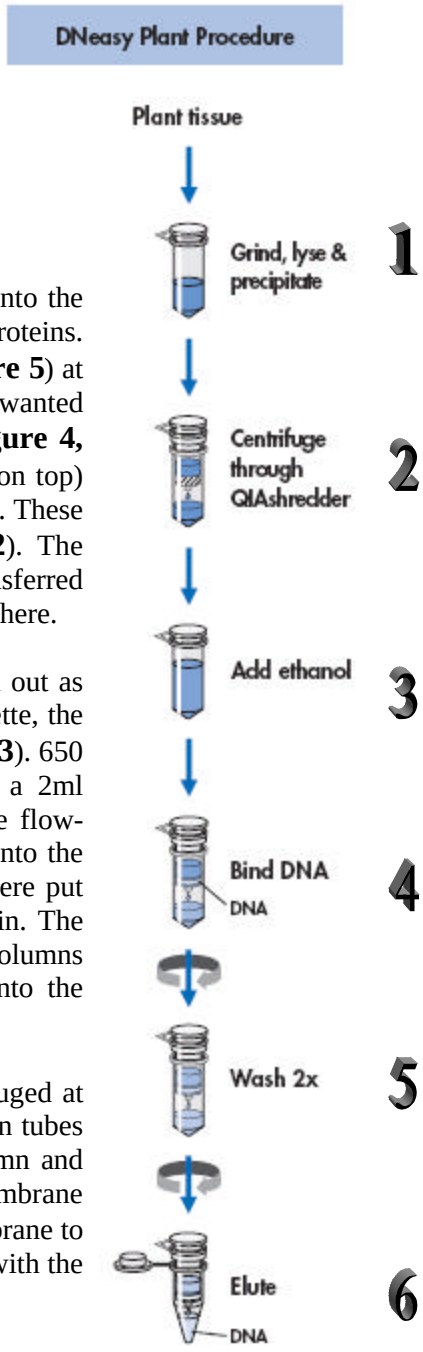
Figure 2. Mortar and pestle

Figure 3. Plastic eppendorf tube

3.2 DNA Extraction

A kit called a DNeasy Plant Mini Kit (QIAGEN) was used for the extraction process. This process gets rid of all of the plant material except the target DNA molecules. A detailed protocol for the extraction is given below - the kit includes all of the buffers referred to. (Buffer AP1 had formed a precipitate whilst in storage. After warming in the oven on a gentle shaker at 65ºC the precipitate was dissolved.) The powdered samples were taken out of the freezer and the extraction process started as follows:

Figure 4. DNA Extraction



(QIAGEN, 2004)
400µl of Buffer AP1 was added to each sample along with 4µl of RNase. The mixtures were vortexed to avoid clumps of plant tissue accumulating. Clumped tissue would not lyse (break up) properly so would therefore have reduced the amount of DNA yielded. The mixtures were placed in the oven at 65ºC and left for 30 minutes. They were in a large cylinder which rotated gently so continually inverted the mixtures. The Buffer AP1 lyses the cells and then RNase works to digest unwanted RNA. The mixtures were at this point called lysates.

130µl of Buffer AP2 was added to the lysates, mixed and then put into the freezer for 5 minutes. This precipitated out unwanted materials like proteins. As an extra precipitation step, the lysates were centrifuged (see **Figure 5**) at 14,000rpm for 5 minutes. Once finished, there was a layer of unwanted material sitting at the bottom of the eppendorf - the precipitate (**Figure 4, stage 1**). Using a pipette the supernatants of the lysates (the liquid on top) were transferred to QIAshredder spin columns in 2ml collection tubes. These were centrifuged at 14,000rpm for 2 minutes (**Figure 4, stage 2**). The liquid which passed through into the collection tubes was then transferred into eppendorf tubes. Approximately 450µl was recovered per sample here.

675µl of Buffer AP3/E was added to each sample. This was worked out as 1.5 volumes of the lysate which was recovered (450µl). Using a pipette, the Buffer was mixed into each sample as it was added (**Figure 4, stage 3**). 650 µl of each mixture was put into a DNeasy mini spin column in a 2ml collection tube. This was centrifuged at 8000rpm for 1 minute. The flow-through in the collection tube was discarded and the tube put back onto the DNeasy column. The remaining sample from the eppendorf tubes were put into the spin columns and centrifuged at 8000rpm for 1 minute again. The collection tubes were discarded with the flow-through. The spin columns were put onto new collection tubes. The DNA had now bound onto the DNeasy column membrane (**Figure 4, stage 4**).

500µl of Buffer AW was put onto each DNeasy column and centrifuged at 8000rpm for 1 minute. The flow-through was discarded and collection tubes re-used. Another 500µl of Buffer AW was put onto each spin column and this time centrifuged at 14,000rpm for 2 minutes. This washed the membrane (**Figure 4, stage 5**) and the longer centrifugation dried out the membrane to avoid residual ethanol being present since this would have interfered with the next step.

The DNeasy columns were placed into eppendorf tubes and 100µl of Buffer AE was pipetted onto each one. They were left to incubate at room temperature for 5 minutes and then centrifuged at 8000rpm for 1 minute. This eluted the DNA so it was present in the eluates in the eppendorf tubes (**Figure 4, stage 6**). Another 100µl of Buffer AE was added to the DNeasy column and left to incubate for 5 minutes then centrifuged at 8000rpm for 1 minute. The elution was then complete. The DNeasy columns were disposed of. The DNA was then extracted and ready for the next step.

To check that DNA was present at this point, an agarose gel was prepared. 5µl of each sample was mixed with 2µl of loading buffer and put into the wells of the gel. An electric current was run through the gel for 30 minutes (see **Figure 6**). This is called electrophoresis: the process works because DNA molecules are charged and move toward one electrode when the current is applied. As molecules of different sizes have different mobilities in the gel, molecules will separate out (travel different distances across the gel) according to size (see **Figure 7**). The gel contained ethidium bromide which binds to DNA and, once bound, fluoresces under UV light, making fragments visible. Therefore, once the electrophoresis was complete, the gel was scanned in a UV scanner and a picture taken. This showed that DNA was present in the samples.

Figure 5. Centrifuge – tubes go into the circular middle section and are spun at high speeds.

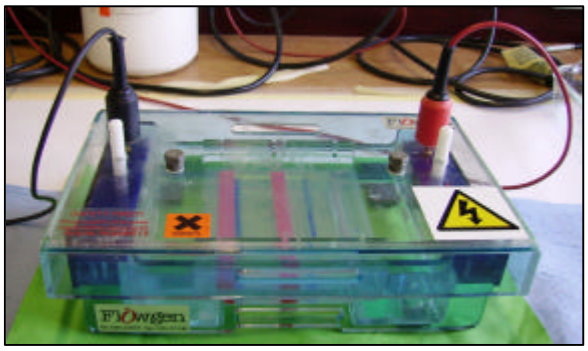


Figure 6. Electrophoresis on an agarose gel.

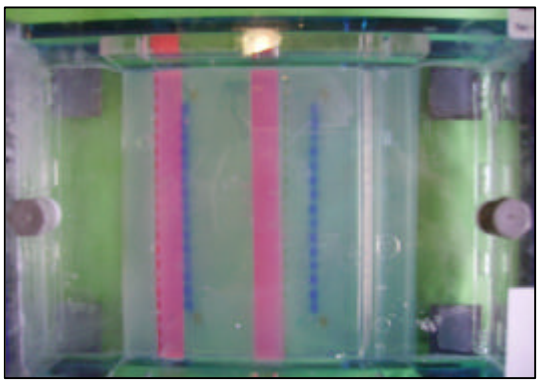


Figure 7. Samples running through the gel

3.3 PCR

PCR (polymerase chain reaction) was used to amplify (make copies of) specific regions of the DNA which was extracted from the leaves. It is a cyclical process which involves repeated denaturing (breaking apart of DNA strands at high temperature) and copying of short regions of DNA code. The region to be copied is identified by using two short 'primer' sequences which bind (anneal) to specific sites on the DNA molecule corresponding to their own sequence: the two primers bind at either end of the target DNA region. The new copies of DNA are then made by a DNA synthesis enzyme known as Taq polymerase. As it is repetitive, the PCR process is programmable and run by a computer. The machine which does the PCR is called a thermal cycler. Before PCR was started, a mix of chemicals needed to be made. The chemicals involved are:

- Water
- 10x Buffer
- Two oligonucleotide primers
- dNTPs (deoxynucleotides: N = G, A, T or C)
- Taq (thermostable DNA polymerase)
- The DNA extracts

The buffer includes salts and is present to keep the pH of the reaction stable and allow the enzyme to function. The primers anneal to the DNA at opposite ends, working in opposite directions. The primers used were TF_CFor and TG_FRev (see **Table 1**).

Table 1: Primer names and sequences for amplification of *trn*L region of cpDNA.

Primer name	Location	Sequence	Annealing temp. / ºC	Reference
TF_CFor	*trn*L (UAA) – 5' exon	CGAAATCGGTAGACGCTACG	57	Taberlet *et al.* (1991)
TG_FRev	*trn*L (UAA) – 3' exon	GGGGATAGAGGGACTTGAAC		

2µl of each DNA sample was put into a well in a sample plate. A mastermix of all of the chemicals required was then made up. This was calculated to allow 25µl to be put into each well, including the DNA. All measurements per sample were the same, 16.3µl water, 2.5µl 10x Buffer, 1.5µl each primer, 1µl dNTPs and 0.2µl Taq. According to how many samples needed PCR, these figures were multiplied by the number of samples plus 2 to allow for measurement errors. This avoided pipetting very small volumes and minimised the possible measurement errors. Once the mastermix was made up and the DNA was in the wells ready, the mastermix was divided out putting 23µl into each well with the 2µl DNA sample. The plate was then sealed using sticky foil (see **Figure 9**) and placed onto the PCR machine and the programme started.

In PCR, the DNA was first subjected to a very high temperature: 94ºC. This denatured the DNA so that it divided into 2 strands. This lasted for 3 minutes then it entered the cyclical phase. This involved the denaturing at 94ºC for 1 minute followed by 1 minute at 57ºC which was when the primers annealed to the DNA (stuck to it). 2 minutes of 72ºC followed this and was called the extension phase, when dNTPs were used by the Taq enzyme to build up the DNA fragment. After this, the DNA had been copied and was back to a double stranded molecule. This was when the cycle began again with denaturing. For the samples in this project, PCR was optimal when it cycled 40 times. After these 40 cycles there was a final extension phase of 72ºC for 10 minutes, to complete any remaining partial DNA fragments, followed by incubation at 4ºC to preserve the PCR products, which was maintained until the PCR machine was switched off and the samples recovered.

The whole process took around 4 hours to complete.

Once DNA amplification was complete, an agarose gel was made up and as before, 5µl of each sample was mixed with 2µl of loading buffer and put into the wells of the gel. Once the electrophoresis was complete, the gel was scanned in a UV scanner and a picture taken. This checked how well the PCR had worked and what size the fragments at this stage were.

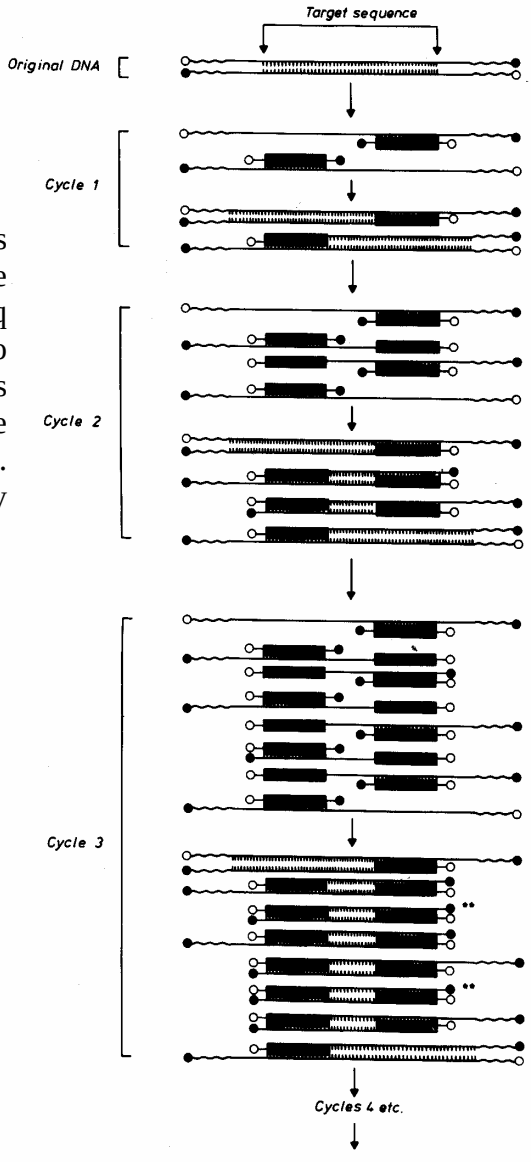


Figure 8. (Weising *et al.*, 1995)

This shows first the original DNA strands being denatured and entering cycle 1. The black rectangles are the primers. The Taq builds up the DNA using the DNTPs and two new strands are made. The DNA then enters cycle 2 where the same process copies the DNA. 4 DNA molecules are now present. PCR continues like this exponentially increasing the number of DNA molecules.

Figure 9. Sticky foil covering used wells to avoid evaporation on heating

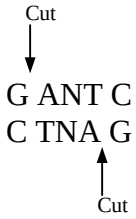
3.4 Digestion

After the DNA had been amplified and a gel image taken, they needed to be digested. This cut the fragment of DNA in specific places. On average, 4 fragments per sample were found after the digestion, so 3 cuts had been made.

For the digestion process, a mastermix of chemicals needed to be prepared. This consisted of:

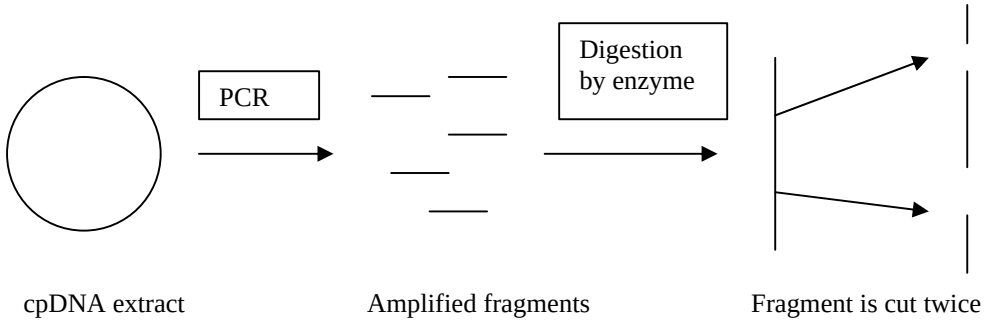
- 12.6µl Water
- 2µl 10x Buffer
- 0.4µl Enzyme

The enzyme used was one called *Hinf*I, which cuts DNA when it finds the 4 base sequence:



Different enzymes recognise different sequences which determine the points at which the cuts are made, so by using a different enzyme, a very different result would have been achieved. The approach of using different enzymes also allows different parts of the fragment to be investigated.

Once the mastermix was prepared, 5µl of each DNA sample from the PCR was put into a well and then 15µl of the mastermix was added. The wells were sealed with sticky foil and put onto the PCR machine to be incubated. The PCR machine was programmed to heat the samples to 37ºC for 5 hours. After this time, the samples were all fully digested and ready for the next step.



When visualised on polyacrylamide gel (PAGE), the fingerprint may have looked like:

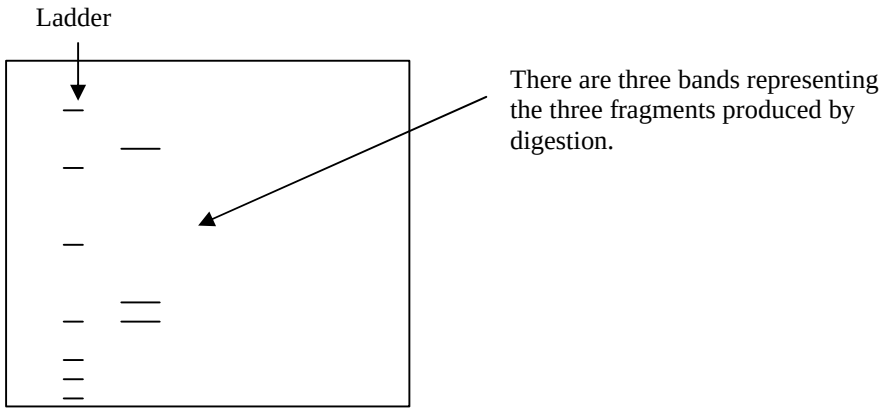


Figure 10. Graphical representation of fingerprinting processes. The top half shows a fragment of cpDNA being amplified by PCR then cut up by digestion. The bottom half shows the digested fragments being visualised by PAGE and compared to a DNA ladder which consists of known sizes of fragments.

3.5 PAGE

PAGE (polyacrylamide gel electrophoresis) was the final stage in the DNA fingerprinting process. It allowed the DNA fragments to be separated and visualized then accurately sized and compared to be stored in a database for future use.

First, the gel rig needed to be prepared. This involved sandwiching two glass plates together and clamping them together. A tiny gap of a few millimetres was left for the gel to go into. Masking tape was used to seal the bottom end of the rig and Vaseline covered the corners of the masking tape to reduce the chances of leakage occurring. A comb was then put into the top of the gel rig to form sample wells. This needed to be repeated with a second rig. Once the rigs were set up cams were used to tighten them into the holder and the gel solution could be prepared. This had to be done in a fume cupboard because acrylamide and TEMED (two of the constituents of polyacrylamide gel) are toxic substances. Once they are in a gel and the acrylamide has polymerised, they are a lot safer to handle.

52.5ml of deionised water was measured out in a measuring cylinder and put into a beaker. 15ml of 40% acrylamide solution was measured then added to this. 7.5ml of 10x TBE Buffer was measured using a pipette and added to the beaker. 60µl of TEMED was pipetted into the beaker then 400µl of 10% APS. The TEMED and APS are catalysts which polymerise the acrylamide. The beaker was mixed briefly by pouring into the measuring cylinder then back into the beaker twice. A small pipette tip was then placed into the gap in the gel rig between the glass plates. Carefully, the mixture was poured into the pipette tip and filled the gel rig (see **Figure 11)**. This was repeated for the second gel. Once filled, they needed to be checked regularly and if leakage occurred, APS was put onto the leak to polymerise it and the gel was topped up. As the gel polymerised and set, a pattern (known as a “Schlieren” pattern) could be seen around the teeth of the combs. After just over an hour, the gel was set.
The combs were then removed and the wells washed twice with deionised water, then twice with 1x TBE Buffer which was left in the wells the second time ready for loading. The samples were then mixed with a small amount of loading buffer and loaded into the wells. Three 100bp DNA sizing standards (ladders) were also put into the wells, one at each end and one in the middle. The upper part of the rig was put onto the two gels and clamped into place with the cams. This was then filled with 1x TBE Buffer until it covered the electrode. The whole assembly was then put into the electrophoresis unit (see **Figure 12)** which was also filled with 1x TBE Buffer up to a marked level. The lid was put in place and the electrophoresis ran for 2 hours at 300V and 90mA.

Once the electrophoresis was finished it was stopped and the gels were removed from the rig. Two water baths were set up, one with deionised water to rinse the gels and the other with deionised water and 200µl of ethidium bromide to bind to the DNA fragments so that they fluoresced on the UV scanner. The gels were each carefully removed from the glass plates and put into the ethidium bromide bath for 5 minutes then in the pure water for 2 minutes to rinse off any excess ethidium bromide. They were then placed individually onto the UV scanner screen and an image taken of each. Gels were then disposed of into the toxic waste bin and the scanner screen wiped down with 70% ethanol.

The fragments could then be sized from the 100bp ladder and a database produced with the sizing of each fragment.

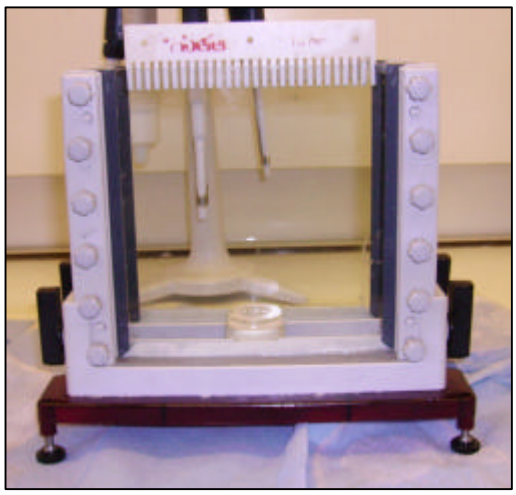


Figure 11. Gel rig with well comb placed in top.

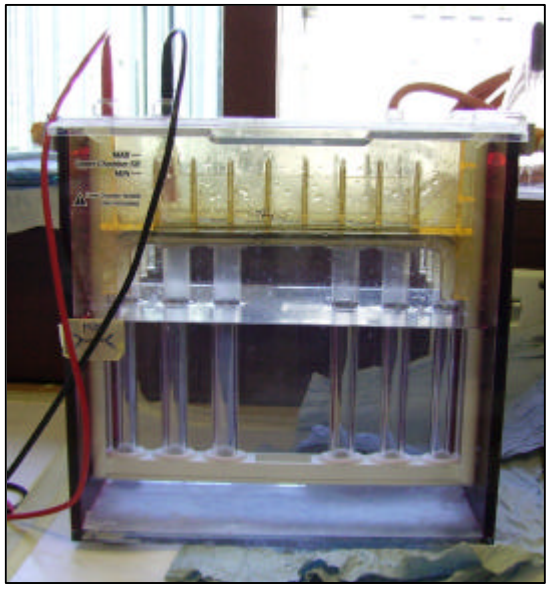


Figure 12. Unit into which the gel rigs are placed

4. Results

4.1 DNA Sampling

The samples were taken from various families:

Common Name	Scientific Name
Oak	Quercus
Maple	Acer
Ash	Fraxinus
Willow	Salix
Prunus	Prunus
Poplar	Populus

Common Name	Scientific Name
Lime	Tilia
Cypress	Chamaecyparis
Horse Chestnut	Aesculus
Chestnut	Castanea
Southern Beech	Nothofagus
Beech	Fagus

Common Name	Scientific Name
Service Trees	Sorbus
Whitebeam	Sorbus
Rowan	Sorbus
Walnut	Juglans
Hawthorn	Crataegus
Elm	Ulmus
Hazel	Corylus

A variety of different species from each family was collected in 4 sets as follows:

Date/set/location	No.	Common Name	Scientific Name
24-Jul Sample Set 1 Bush Estate	1	English Oak	*Quercus robur*
	2	Norway Maple	*Acer platanoides*
	3	Common Ash	*Fraxinus excelsior*
	4	Goat Willow	*Salix caprea*
	5	Wild Cherry	*Prunus avium*
	6	White Poplar	*Populus alba*
	7	Common Lime	*Tilia x europaea*

Date/set/location	No.	Common Name	Scientific Name
27-Jul Sample Set 2 Bush Estate	1	Chestnut variant	
	2	Common Rowan	*Sorbus aucuparia*
	3	Sycamore	*Acer pseudoplatanus*
	4	Common Hawthorn	*Crataegus monogyna*
	5	Common Walnut	*Juglans regia*
	6	Common Beech	*Fagus sylvatica*
	7	White Willow variant	*Salix alba*
	8	Turkey Oak	*Quercus cerris*
	9	Red Oak	*Quercus rubra*
	10	Horse Chestnut	*Aesculus hippocastanum*
	11	Sweet Chestnut	*Castanea sativa*
	12	Common Whitebeam	*Sorbus aria*
	13	Wild Service	*Sorbus torminalis*
	14	Lawson Cypress	*Chamaecyparis lawsoniana*
	15	Lawson Cypress	*Chamaecyparis lawsoniana*
	16	Lawson Cypress	*Chamaecyparis lawsoniana*
	17	Cypress Hybrids	
	18	Cypress Hybrids	
	19	Cypress Hybrids	

Date/set/location	No.	Common Name	Scientific Name
03-Aug Sample Set 3 RBGE	1	Red Oak	*Quercus rubra*
	2	Sessile Oak	*Quercus petraea* 'Mespilifolia'
	3	Pyrenean Oak	*Quercus pyrenaica*
	4	English Oak	*Quercus robur*
	5	Turkey Oak	*Quercus cerris*

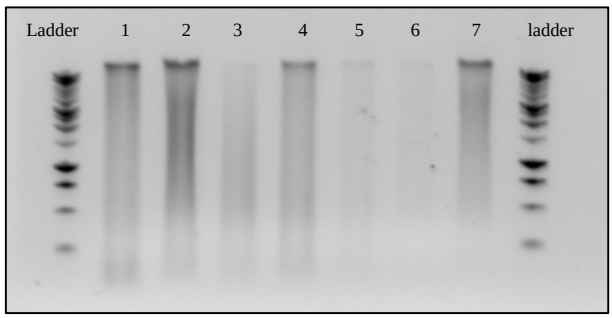
Date/set/location	No.	Common Name	Scientific Name
03-Aug Sample Set 3 RBGE	6	Hungarian Oak	*Quercus frainetto*
	7	Black Oak	*Quercus velutina*
	8	Downy Oak	*Quercus pubescens*
	9	Caucasian Oak	*Quercus macranthera*
	10	Hybrid Oak	*Quercus x rosacea* (*petraea x robur*)
	11	Lebanon Oak	*Quercus libani*
	12	Heldreich's Maple	*Acer heldreichii*
	13	Grey Snake-bark Maple	*Acer rufinerve*
	14	Sycamore	*Acer pseudoplatanus*
	15	Norway Maple	*Acer platanoides*
	16	Wild Black Poplar	*Populus nigra ssp. Betulifolia*
	17	Scarlet Oak	*Quercus coccinea*
	18	Silver Maple	*Acer saccharinum*
	19	Sugar Maple	*Acer saccharum*
	20	Oregon Maple	*Acer macrophyllum*
	21	Field Maple	*Acer campestre*
	22	Manna Ash	*Fraxinus ornus*
	23	White Poplar	*Populus Alba*
	24	Holm Oak	*Quercus ilex*
	25	Pin Oak	*Quercus palustris*
	26	Ubame oak	*Quercus phillyreoides*
	27	Pere David's Maple	*Acer davidii*
	28	Red Ash	*Fraxinus pennsylvanica*
	29	Black Cherry	*Prunus serotina*
	30	Montpelier Maple	*Acer monspessulanum*
	31	Columnar English oak	*Quercus robur 'Fastigiata'*
	32	Korean oak	*Quercus glandulifera*
	33	Vallonea Oak	*Quercus macrolepis*
	34	Sessile Oak	*Quercus petraea*
	35	Chinese Cork Oak	*Quercus variabilis*
	36	Sawthorn oak	*Quercus serrata*
	37	California Black Oak	*Quercus kelloggi*
	38	Japenese Chestnut Oak	*Quercus acutissima*
	39	Hybrid Oak	*Quercus x hickelii (pontica x robur)*
	40	Lime-leaved Maple	*Acer distylum*

Date/set/location	No.	Common Name	Scientific Name
04-Aug Sample Set 4 Bush Estate	1	Goat Willow	*Salix caprea*
	2	Wych Elm	*Ulmus glabra*
	3	Common Ash	*Fraxinus excelsior*
	4	Common Hazel	*Corylus avellana*
	5	Common Beech	*Fagus sylvatica*
	6	Common Hawthorn	*Crataegus monogyna*
	7	Horse Chestnut	*Aesculus hippocastanum*
	8	Copper Beech	*Fagus sylvatica var. purpurea*
	9	Swedish Whitebeam	*Sorbus intermedia*
	10	Common Lime	*Tilia x europaea*
	11	Wild Cherry	*Prunus avium*
	12	Plum	*Prunus domestica*
	13	Rauli	*Nothofagus procera*
	14	Roble Beech	*Nothofagus obliqua*
	15	Antarctic Beech	*Nothofagus antarctica*

4.2 DNA Extraction

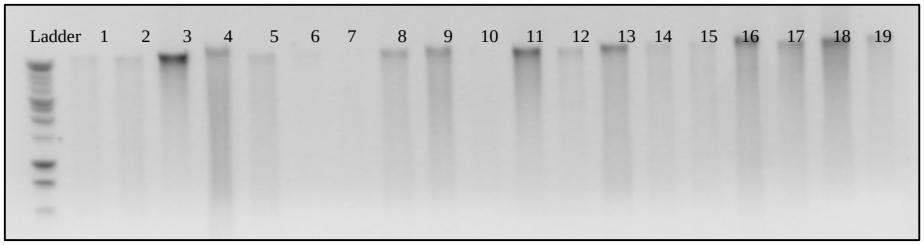
SS1

This is the extractions from sample set 1. All samples had DNA present. Streaking indicated that the samples had begun to degrade.



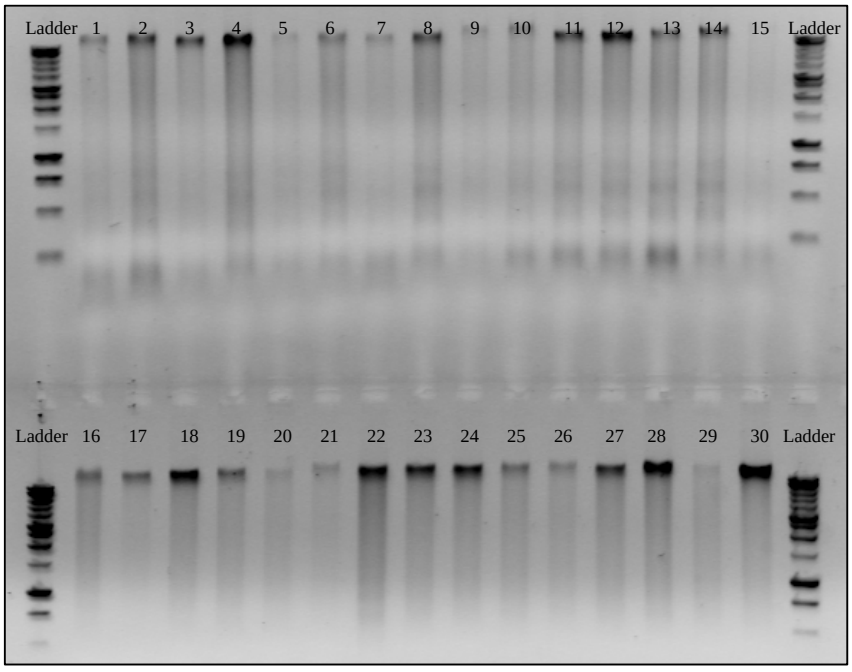
SS2

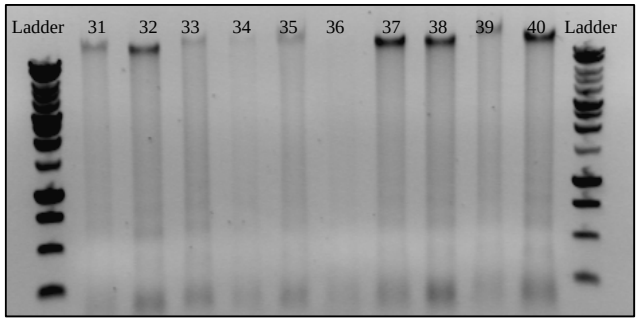
These are the sample set 2 extractions. Some of the samples had very weak levels of DNA present. Number 7 had no DNA present. 6 and 10 were very weak. Again, streaking shows degradation.



SS3

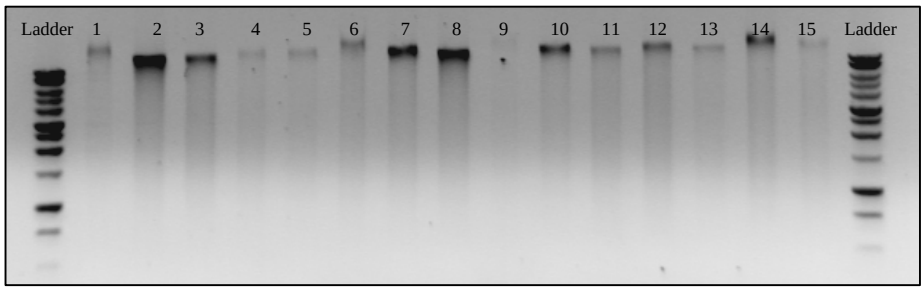
This is sample set 3 extractions. All samples show DNA presence. Samples 15, 34 and 36 are a bit weak but still present. This will improve on amplification.





SS4

In sample set 4 all samples have DNA present. Sample 9 is particularly weak but still visible.

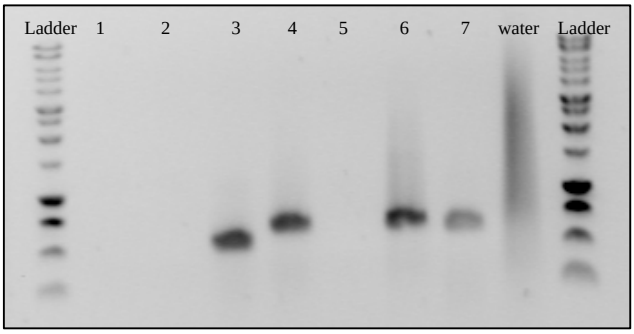


Out of these 81 samples, only 1 did not show any sign of DNA being present. This could have been due to an error in the extraction process.

4.3 PCR

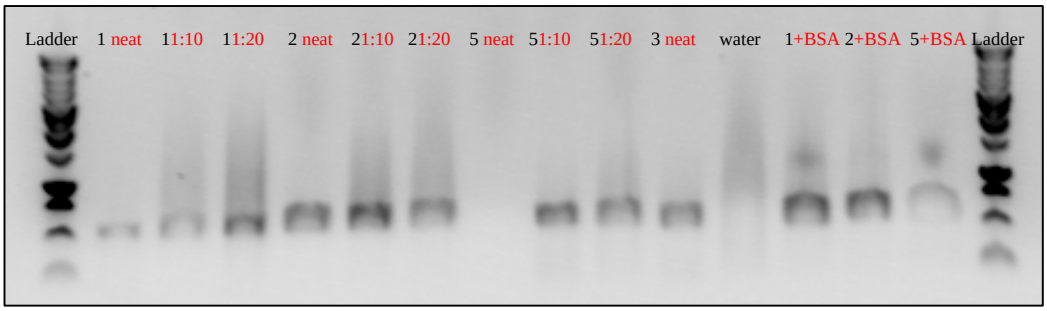
SS1

After PCR on sample set 1 the DNA only showed up on samples 3, 4, 6 and 7. Made up diluted DNA from the samples which did not work then repeated a PCR on them. Already, variation can be clearly seen between samples 3 and 4.



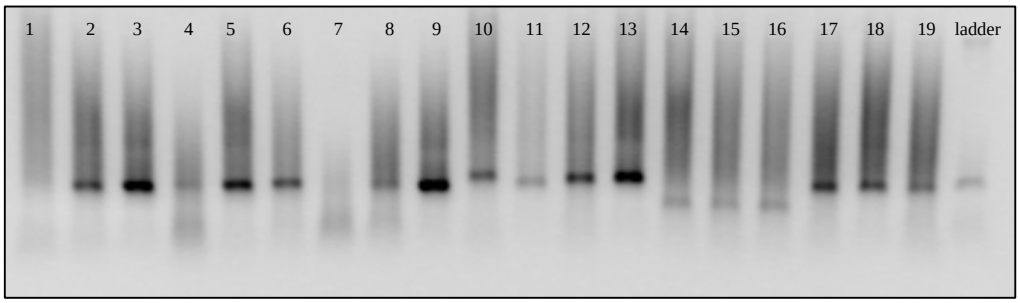
SS1 dilutions

PCR worked this time for each sample on a 1:10 dilution so used these samples for the digestion.



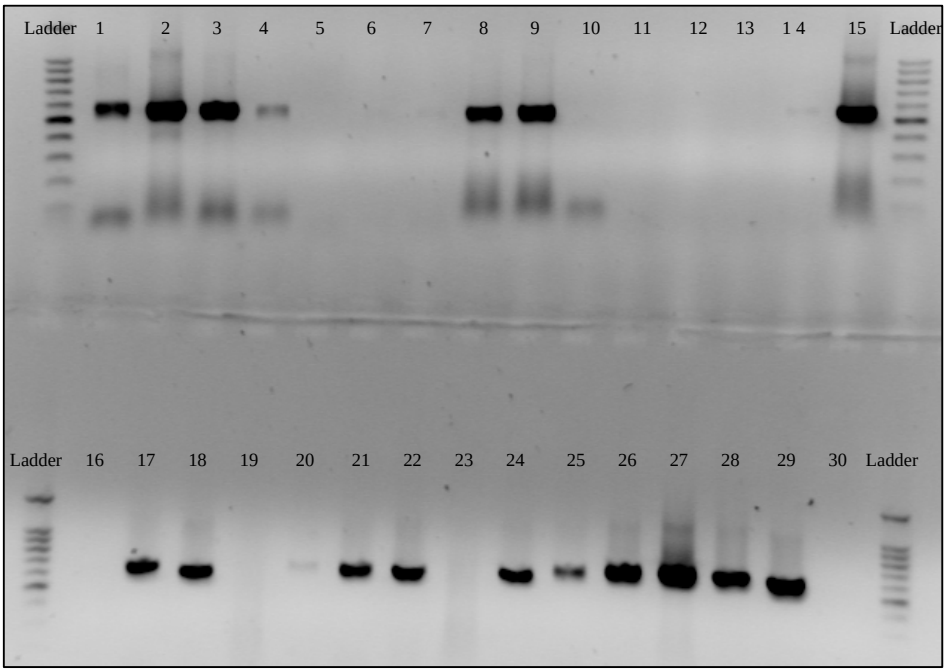
SS2

Sample 7 did not work at all. The rest despite some weak ones did work. On close inspection sample 1 does have a faint band present.



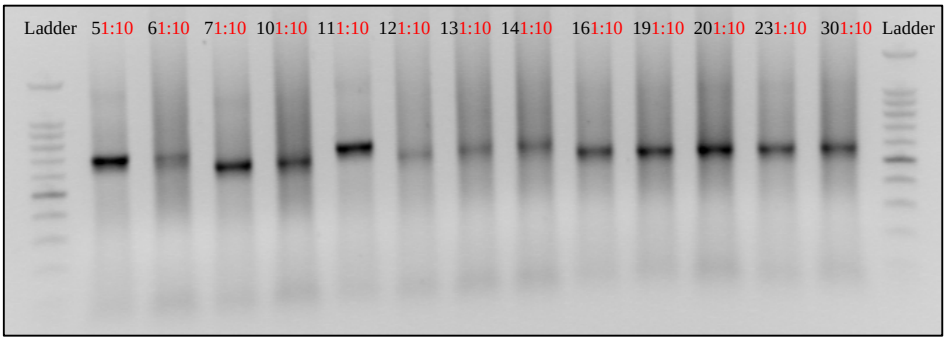
SS3 1-30

Some samples did not work at all. These ones were repeated with diluted extracts. The samples which did work here, showed up strongly.



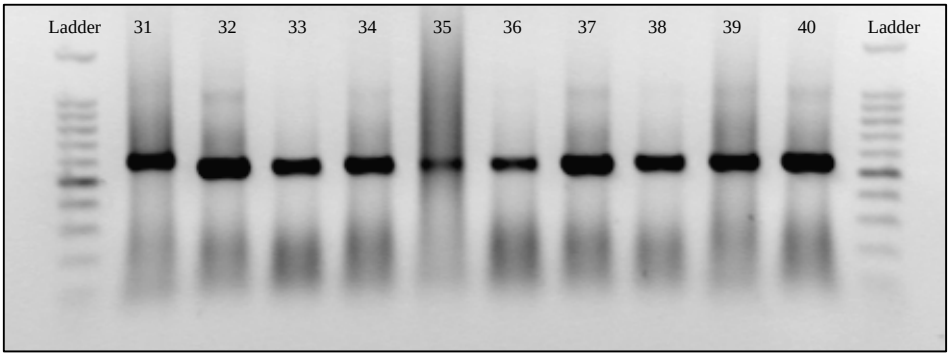
SS3 Dilutions

These all worked very well and were used for the digestion.



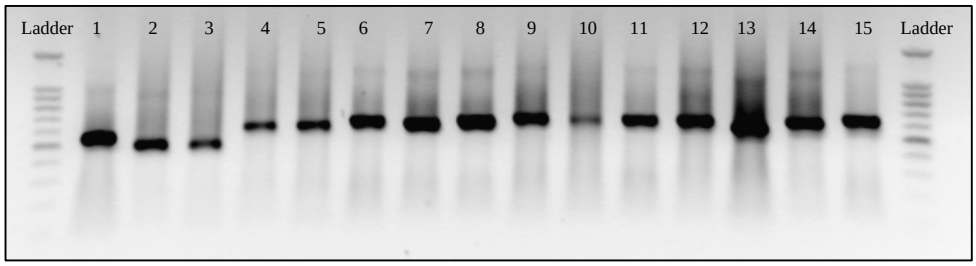
SS3 30-40

All samples showed clearly and strongly on this gel. These were all diluted on a 1:10 ratio to try to optimise the PCR. The 100% success rate proved that this worked well.



SS4

All samples again showed up clearly and strongly. These were all diluted on a 1:10 ratio.



4.4 PAGE

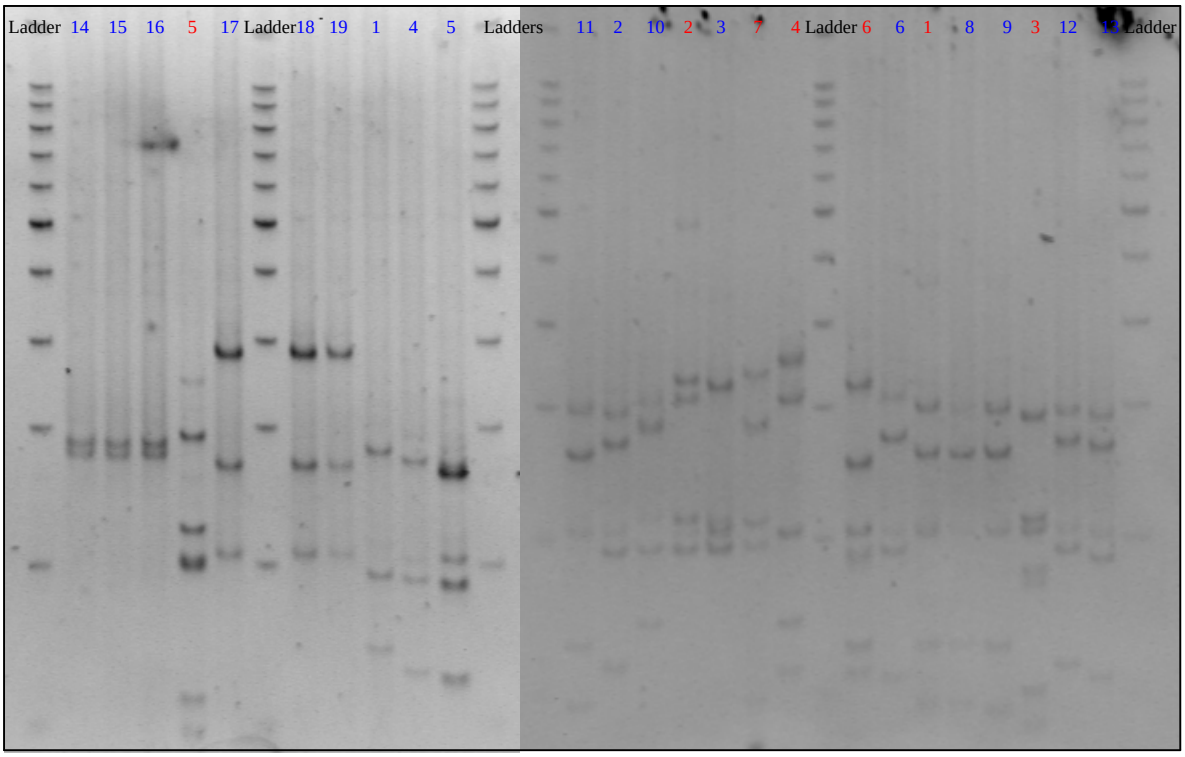


Figure 13. Sample sets 1 & 2
This shows that each sample has successfully been digested with usually 4 or 5 bands present. Samples 1, 8 and 9 have very similar sizes - they are all Oaks so this indicated that Oak diversity needed to be investigated further.

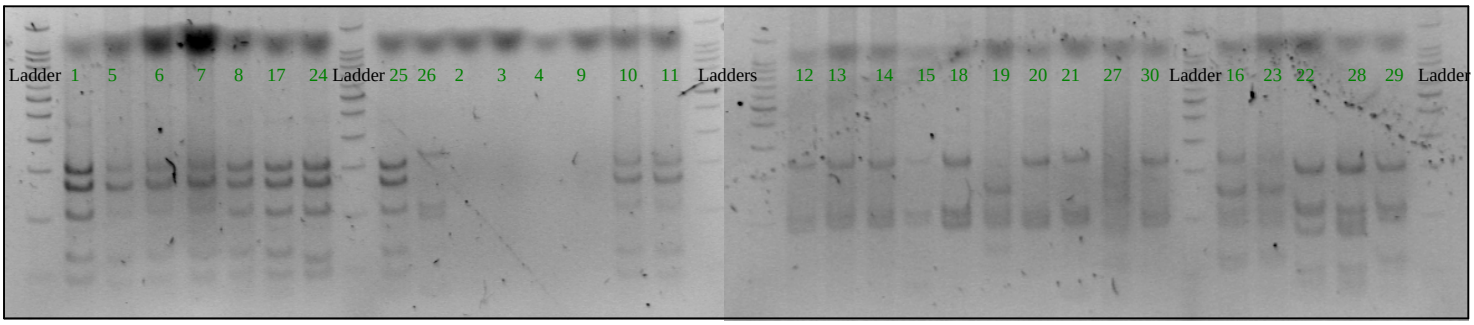
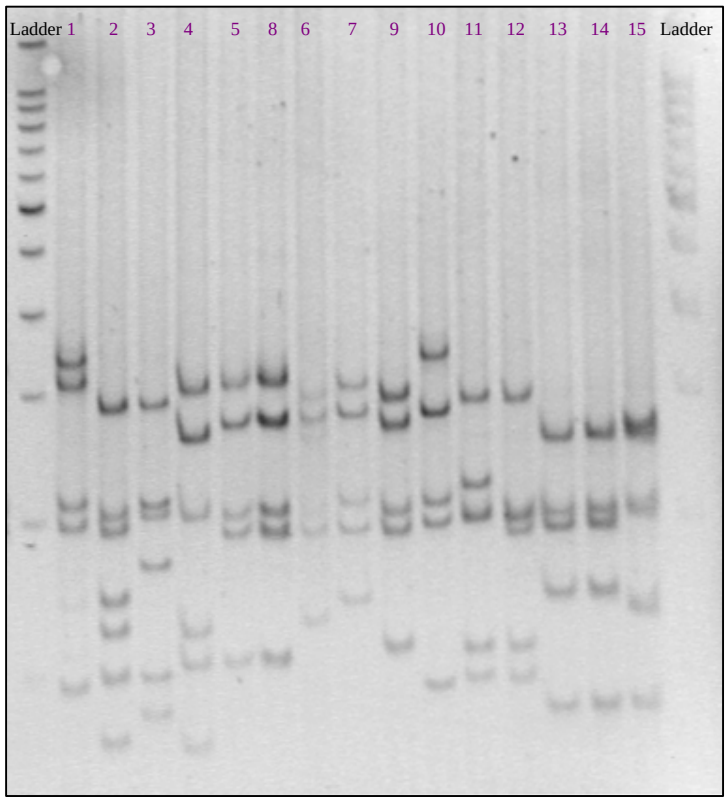


Figure 14. Sample set 3.
This shows that most samples have successfully digested. The problems evident in SS3 PCR carried on and showed in the PAGE. The first Gel samples are all Oak species and most have extremely similar sized fragments. Sample 26 has a different fingerprint compared to the other Oaks. However, there is a clear Oak fingerprint present. In the second gel all samples worked. The fingerprints between the first and second set of DNA ladders are all Maple species. Excluding sample 19, a clear Maple fingerprint is evident. Samples 16 and 23 are both Poplars and show the fingerprint for Poplar trees. Samples 22 and 28 are both Ashes and show an Ash fingerprint.

Figure 15. Sample set 4



All samples here showed their fragment sizes clearly. The digest was successful for every sample. Samples 13, 14 and 15 have similar sizes; they are all Southern Beech trees. Samples 5 and 8 are both Beech trees so are also similar in fragment sizes to each other.

These gel images were put onto a program on a computer called Gene ImagIR. This program is used to size each sample's fragments. Each one was recorded and a database built up.

4.5 Database

The database was produced in Microsoft Excel. It lists each species fingerprinted with its fragment sizes in order so that the most similar ones are consecutive.

Figure 16. Database of fingerprints

B	C	D	E	F	G	H	I	J	K	L	M	N
Genus	Species	Scientific Name	PCR Band Length	Digest no. frags	Sizes							Total
Cypress	Lawson Cypress	*Chamaecyparis lawsoniana*	390	2	184	175						359
Chestnut	Chestnut variant	-	0	3	178	94	65					337
Maple	Montpelier Maple	*Acer monspessulanum*	582	3	226	103	90					419
Maple	Silver Maple	*Acer saccharinum*	626	3	226	109	93					428
Oak	Ubame oak	*Quercus phillyreoides*	670	3	227	107	95					429
Maple	Pere David's Maple	*Acer davidii*	679	3	228	188	179					595
Cypress	Cypress Hybrids	-	610	3	283	165	105					553
Walnut	Common Walnut	*Juglans regia*	675	4	159	103	90	56				408
Maple	Sycamore	*Acer pseudoplatanus*	668-685	4	219	111	104	95				529
Maple	Oregon Maple	*Acer macrophyllum*	603	4	225	110	103	95				533
Maple	Lime-leaved Maple	*Acer distylum*	560	4	225	111	106	94				536
Maple	Field Maple	*Acer campestre*	664	4	229	110	105	96				540
Maple	Sugar Maple	*Acer saccharum*	610	5	152	110	106	96	64			528
Beech	Rauli	*Nothofagus procera*	495	5	157	103	96	66	35			457
Beech	Roble Beech	*Nothofagus obliqua*	528	5	160	105	97	66	35			463
Beech	Antarctic Beech	*Nothofagus antarctica*	538	5	166	104	100	60	35			465
Sorbus	Wild Service	*Sorbus torminalis*	685	5	187	159	101	89	48			584
Oak	Korean oak	*Quercus glandulifera*	554	5	189	149	99	54	38			529
Oak	Chinese Cork Oak	*Quercus variabilis*	569	5	190	150	99	54	39			532
Oak	Sessile Oak	*Quercus petraea*	561	5	191	152	100	54	38			535
Oak	Vallonea Oak	*Quercus macrolepis*	562	5	192	148	100	53	38			531
Sorbus	Common Whitebeam	*Sorbus aria*	690	5	192	164	103	93	51			603
Prunus	Plum	*Prunus domestica*	557	5	193	101	93	49	41			477
Oak	Columnar English oak	*Quercus robur 'Fastigiata'*	598	5	193	154	102	54	39			542
Oak	Japenese Chestnut Oak	*Quercus acutissima*	553	5	194	155	103	56	40			548
Sorbus	Common Rowan	*Sorbus aucuparia*	695	5	194	164	104	94	52			608
Oak	Downy Oak	*Quercus pubescens*	540	5	195	154	103	57	41			550
Oak	Hybrid Oak	*Quercus x hickelii (pontica x robur)*	562	5	195	154	103	56	39			547
Oak	Red Oak	*Quercus rubra*	569-650	5	195	155	104	58	42			554
Chestnut	Sweet Chestnut	*Castanea sativa*	675	5	196	154	104	57	42			553
Oak	Holm Oak	*Quercus ilex*	642	5	196	154	102	50	39			541
Oak	Sawthorn oak	*Quercus serrata*	567	5	196	154	101	54	39			544
Oak	Pin Oak	*Quercus palustris*	686	5	196	155	103	56	40			550
Oak	California Black Oak	*Quercus kelloggi*	555	5	197	154	102	55	39			547
Oak	Hybrid Oak	*Quercus x rosacea (petraea x robur)*	625	5	197	155	109	55	40			556
Oak	Scarlet Oak	*Quercus coccinea*	641	5	197	155	103	56	40			551
Sorbus	Swedish Whitebeam	*Sorbus intermedia*	577	5	198	169	105	94	49			615
Hawthorn	Common Hawthorn	*Crataegus monogyna*	559-750	5	198	170	105	94	57			624
Oak	English Oak	*Quercus robur*	565	5	199	155	104	57	42			557
Oak	Turkey Oak	*Quercus cerris*	660-673	5	200	155	104	57	40			556
Oak	Lebanon Oak	*Quercus libani*	701	5	202	154	106	56	41			559
Oak	Hungarian Oak	*Quercus frainetto*	666	5	203	161	110	58	41			573
Oak	Black Oak	*Quercus velutina*	614	5	205	160	103	56	40			564
Chestnut	Horse Chestnut	*Aesculus hippocastanum*	532-725	5	209	179	111	95	64			658
Beech	Common Beech	*Fagus sylvatica*	524-720	5	210	170	104	93	47			624
Beech	Copper Beech	*Fagus sylvatica var. purpurea*	542	5	216	174	107	94	47			638
Maple	Grey Snake-bark Maple	*Acer rufinerve*	666	5	222	111	105	96	30			564
Maple	Heldreich's Maple	*Acer heldreichii*	646	5	227	111	105	98	30			571
Willow	Goat Willow	*Salix caprea*	440-711	5	244	209	107	64-98	40-50			664-708
Lime	Common Lime	*Tilia x europaea*	575-667	5	243-231	180	109	97	42			659-671
Ash	Manna Ash	*Fraxinus ornus*	637	6	191	111	104	80	44	37		567
Ash	Red Ash	*Fraxinus pennsylvanica*	663	6	192	106	86	78	45	37		544
Prunus	Black Cherry	*Prunus serotina*	593	6	192	109	103	93	50	42		589
Hazel	Common Hazel	*Corylus avellana*	538	6	206	158	104	54	46	29		597
Poplar	Wild Black Poplar	*Populus nigra ssp. Betulifolia*	616	6	221	143	101	92	55	48		660
Poplar	White Poplar	*Populus Alba*	588-737	6	223	146	104	92	56	50		671
Maple	Norway Maple	*Acer platanoides*	531-648	6	227	208	110	103	96	36		780
Prunus	Wild Cherry	*Prunus avium*	560-618	6	249	193	120	100	50	42		754
Elm	Wych Elm	*Ulmus glabra*	381	7	186	103	96	66	55	43	30	579
Ash	Common Ash	*Fraxinus excelsior*	404-566	7	189	110	104	83	79	43	36	644

5. Discussion & Conclusions

The database that was produced can be used by anyone who has fingerprinted an unknown species and wishes to identify it. It will never be a 100% proof of identity but will easily locate a ‘closest match’ to a species.

In the genus ‘Quercus’, which is the Oak family, we collected a huge variety of species and were able to average out a general Oak fingerprint but also pick up very small differences between the species, making it a little more possible to identify which Oak an unknown could be. Further work could be done to determine the differences between these species by using a different enzyme in the digestion stage. This would allow for bigger differences in band sizes to be produced and therefore a more diverse range of fingerprints. This would also allow for a basic fingerprinting to be done, then when the unknown is identified by the general Oak fingerprint, the different enzyme used for digestion and then the species more reliably identified.

The same was seen for the Maple family ‘Acer’. Once again, more research could go into producing a more accurate database to identify differences in the family.

Between the different families there proved to be noticeable differences making it relatively easy to distinguish between the families. The database can therefore be used successfully to identify the genus of an unknown and even the species depending on how close a match it is.

6. References & Acknowledgements

References

Johnson O. & More D. (2004) Collins Tree Guide. HarperCollinsPublishers Ltd, London.

Taberlet P., Gielly L., Pautou G. and Bouvet J. (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology Reporter*. 17: 1105-1109.

Wesing K., Hilde Nymbom, Kirsten Wolff & Wieland Meyer (1995) DNA Fingerprinting in Plants and Fungi. CRC Press, Inc, Florida

QIAGEN (2004) DNeasy® Plant Mini and DNeasy Plant Maxi Handbook.

Acknowledgements

Thanks to Stephen Cavers for supervising and guiding me, everyone at CEH for helping me become one of the team, especially Bob Munro and Katy Walker. Also thanks to Frances Chapman and the Nuffield Foundation for her support and their funding, and Mr Gray at Dalkeith High School for introducing me to this bursary project scheme.