The Development of Chemical and Biological Profiling for the Forensic Provenancing of Norfolk Soils

A thesis submitted to the

School of Chemistry

of the

University of East Anglia

In fulfilment of the requirements for the degree of

Master of Philosophy

Hilary Diane Bathgate

3800563

November 2014

Abstract

Soils are frequently analysed by forensic laboratories by comparing a suspect sample to an especially collected control sample. As yet, they cannot be compared to a central database, unless the area in question has already been identified; with the use databases being highly contested within the field of forensic geosciences. There is a need for a method of soil profiling that allows an unknown sample to be tested and assigned a quantitative likelihood that it originated from a given region. Spatial models can then be created using geographical information systems to house multiple datasets and be used to map soils across geographical areas.

Generally, the more variables available with which to compare items, the greater the certainty a forensic analyst can have when asserting their similarity; this applies to geological materials. Equally, soil profiling methods can be used to exclude soil samples from each other or an area. This research involves a number of chemical and biological profiling methods that have been used to build up a unique signature for soils from different locations across Norfolk. All analyses have been carried out on a single source sample. $^{87}Sr/^{86}Sr$ ratios have been measured using MC-ICP-MS, and trace element concentrations measured using ICP-MS. The $^{87}Sr/^{86}Sr$ ratios are significantly different at each of the sample locations; although there is some variation in the replicates collected at each location this variation is smaller than the regional variation. The correlation between the isotope chemistry of the topsoil and the underlying geology is poor, indicating that other sources such as land-use, vegetation cover and additions to the soil contribute to the $^{87}Sr/^{86}Sr$. Therefore, trace element concentrations have been used to spatially discriminate samples and to investigate the effect of fertilisers on the elemental composition of the topsoil.

The biological techniques used to aid discrimination are soil DNA analysis using the chloroplast-located *mat*K gene and MALDI-ToF-MS, palynology and the creation of Norfolk vegetation maps showing all of the plant species recorded in the area; each additional independent dataset allows for an increasing signature of each sample to be built up which can be used for assessing similarity or exclusionary purposes.

Acknowledgments

This thesis would not have been possible without the help of many people. Firstly, my supervisor, Dr Jurian Hoogewerff, for his advice and support throughout. Also, Dr Andrew Mayes, thank you for all your help and encouragement, especially through the last couple of years when things got very tough. Thanks must also go to the rest of the CFP group, Rob, Aaron, Yvette, Ana, and especially Henriette Ueckermann, for all the hours of technical support and keeping the lab together. Things certainly weren't always easy but we usually got there and I've made some great, lifelong friends.

Dr Jonathan Clarke and Darren Heavens from John Innes Centre, your support with all of the DNA work was invaluable especially when I had more questions than answers; the football chat was pretty good too! Thank you to Martin Horlock of Norfolk Biodiversity, Jim Riding of British Geological Survey and Ines Laube of BfR, Berlin for providing data, helping to develop ideas and techniques, analysing samples, and encouraging my knowledge and understanding.

Dr Louise Botham and Keith Parsley, many thanks for taking the time to read and comment on some of the chapters, I owe you cake!

I am incredibly lucky to have such supportive parents. Simply, thanks for always being there. Finally, and most importantly, Brian; You believed in me when I did not believe in myself, you were there whenever things got tough, when the pain was unbearable, for the long days, through the sleepless nights, always with a hug, making me laugh and I am forever grateful. I can't wait for our next adventure together.

Mum, Dad and Brian, this dissertation stands as a testament to your unconditional love and encouragement.

Table of Contents

Table of Figures

Table of Tables

List of Abbreviations

DNA	Deoxyribonucleic Acid
PCR	Polymerase Chain Reaction
RT-PCR	Real-Time Polymerase Chain Reaction
ICP-MS	Inductively Coupled Plasma Mass Spectrometry
MC-ICP-MC	Multi Collector Inductively Coupled Mass Spectrometry
MALDI	Matrix Assisted Laser Desoprtion Ionisation
ToF-MS	Time of Flight Mass Spectrometry
BGS	British Geological Survey
GIS	Geographical Information Systems
3-HPA	3-Hydroxypicolinic acid
THAP	2,4,6-Trihydroxyacetophenone
DHAC	Di-Ammonium Hydrogen Citrate
SEM	Scanning Electron Microscopy
EDX	Energy Dispersive Xray Analysis
IC	Ion Chromatography
HPLC	High Performance Liquid Chromatography
XRD / XRF	Xray Diffraction / Xray Fluroescence
FTIR	Fourier Transform Infra Red
AFLP	Amplified Fragment Length Polymorphism
TRFLP	Terminal Restriction Fragment Length Polymorphism
RAPD	Random Amplified Polymorphic DNA
RISA	Ribosomal RNA (rRNA) intergenic spacer analysis

Chapter 1 – Introduction to Forensic Geology, Soil and the Sample Area

In a forensic context it can be important to be able to associate (or exlude) a soil with a particular location. The primary aim for many forensic scientists is to establish the certainty with which two samples have been derived from the same source, be these glass, bullets, soil samples, DNA profiles, or pollen. Increasingly, geological materials such as rocks, minerals, sediments and soils are being examined in this way to determine whether they share a common origin with a suspect sample. This research looks at the use of chemical and biological techniques to provenance Norfolk soils, with the aim being the ability to distinguish the soil samples locations from each other using an exclusionary approach.

1:1 Forensic Geology

Forensic geology is the use of geological materials and methods in the analysis of samples and places that may be connected with criminal behaviour (1). Forensic geology primarily uses soil samples to analyse problems (2). Forensic geology has been accepted and widely used within the criminal justice system since the end of the 19th century, in both criminal and civil cases. Like many disciplines of forensic science the idea of applying geology to criminalistics began with the popular Sherlock Holmes works of Sir Arthur Conan Doyle. Conan Doyle's work planted the ideas that now form the basis of forensic geology; the number of different kinds of soil is almost limitless; soils change dramatically over short distances; people may collect soil on their shoes or vehicles by simply coming into contact with earth materials; and the examination of soil evidence may place a person at a specific location (3). Subsequently, the study of soil as forensic evidence was used in criminal cases of Hans Gross, Georg Popp and Oscar Heinrich (1). In a forensic context soil is important due to Locard's theory of every contact leaving a trace (4) meaning that if for example a shoe comes into contact with soil there will be traces of soil on the shoe and by using various methods to analyse the soil a link can be made between the shoe and the crime scene (1). Soil samples can be taken both from and to a crime scene.

1:2 Soil

1:2.1 Soil introduction

In simple terms soil can be defined as layers of largely unconsolidated material that formed at the earth's surface through the combined action of sediment accumulation, biological processes and weathering (5). Soil is often described as the loose surface material in which plants grow and is a general term which represents the unconsolidated,

thin, variable layer of mineral and organic material, which is usually biologically active, that covers most of the earth's terrestrial surface (6). Soil is valuable trace evidence and exists as a conglomerate of inorganic and organic components that are widely distributed upon much of the earth's exterior (7). It is these components that can help to identify where a soil sample has originated from.

A soil sample is a compilation of loose inorganic and organic material found above the bedrock of a given geological unit (8). Soils are a product of both acquired and inherited properties; with their characteristics reflecting an integration of original features and the accumulated influences of the environment (9). The five main soil forming factors, which cause the large variation shown in natural soil properties, are (10):

1. Geological parent material
2. Climate
3. Topography
4. Biological influences
5. Time

It also needs to be noted that the influence of man can have an effect on soil. Human action has resulted in the degradation of some soils; this is caused by the destruction of natural vegetation, ploughing and the subsequent erosion of the surface layers. Soils in some areas have been improved for agricultural processes by the addition of sand, lime, artificial and natural fertilizers and pesticides to the soil (5). Each of these actions has the potential to have an influence upon the chemical and biological fingerprint of the soil.

Some soils are merely the accumulation of weathered rock or sediment. However, in more developed soils a key feature is the presence of a number of distinct layers or horizons (figure 1). The development of these horizons reflects the operation of soil forming processes over time. This includes weathering, the vertical movement of soluble constituents and fine particles from the surface to deeper levels (sometimes this process occurs in reverse), and biological activity that may lead to the accumulation of organic matter in one or more horizons (5). Soil properties change with depth; a soil sample taken from a depth of 3 centimetres can have completely different properties from a sample taken from 12 centimetres (11). For this reason the main layer of the soil that will be used for analyses carried out in this project is the top soil which comes into contact with footwear, tyres etc (see section 1:5.2 for chosen sampling method). However, from a

forensic point of view the deeper layers of soil could be of equal importance, for example in a burial.

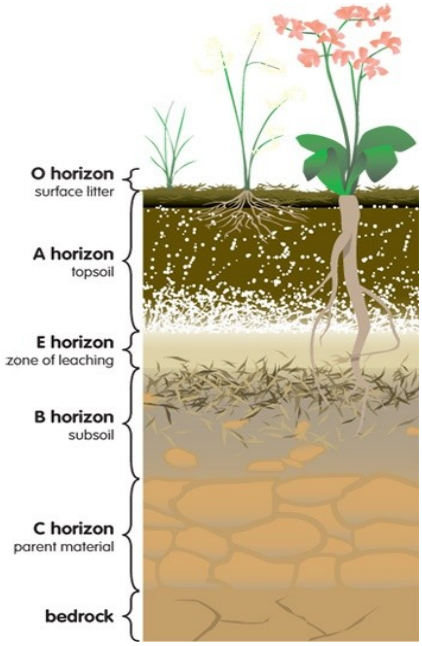


Figure 1: The distinct horizons in a soil profile. It is the topsoil that is used for the analyses carried out in this project (12).

There are four major components of soil: mineral or inorganic, organic, water and the air trapped between the soil particles (see figure 2). The relative proportions of these four soil components vary with soil type and climatic conditions (13).

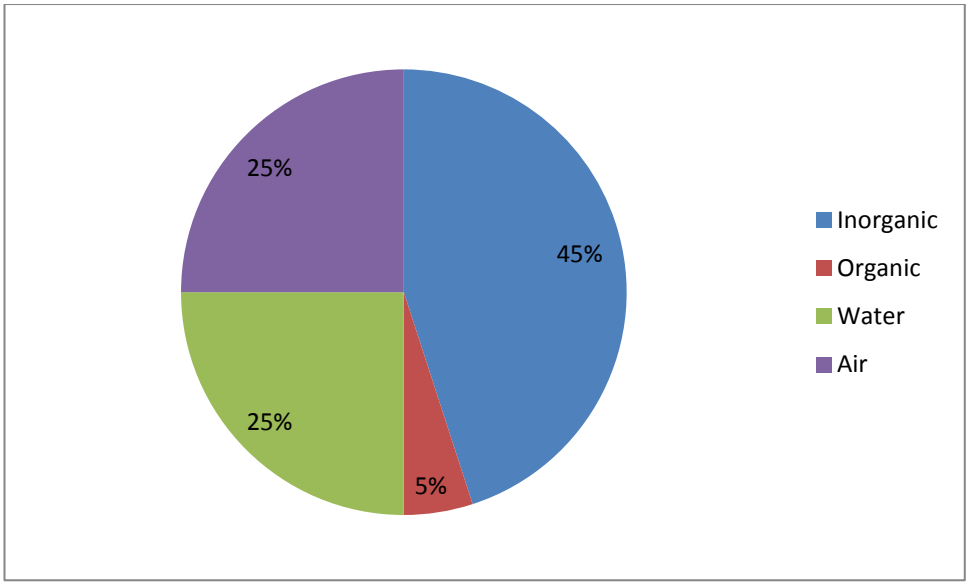


Figure 2: The approximate proportions (by volume) of the four components of a mineral soil (13).

Soils contain both non-living and living matter. Included in the non-living matter are inorganic mineral grains (they are defined depending on their size: coarse and fine materials. The limit between them is arbitrary, depending on the characteristics of every sample, the limit is between 10 and 2 microns) chemical precipitates such as calcium carbonate, dead plant matter in various stages of decomposition, and dead animal matter including the remains of both invertebrates and vertebrates. In some soils, the bones and teeth of small mammals, insect carapaces and shell material are important components. There may also be traces of past microbial communities preserved as chemically altered organic matter present in the soil (5). Tan (14) has discussed how the living matter in soil includes bacteria, soil algae, fungi, roots of higher plants and various species of invertebrates and vertebrate mammals. All of these species, whether they be plant or animal, could add to the DNA present in soil and also have an effect upon the chemical fingerprint of the soil.

1:2.2 Soil in Forensic Science

Soil has been analysed in many forensic cases. Thanasoulias *et al.* (15) mentions that soil can adhere to clothes, shoes, objects and vehicles and can easily be transferred from one particular location to another; this makes soil highly important transfer evidence which could provide information that links a suspect to the crime scene and thus be invaluable in criminal investigations. However Pye *et al.* (5) states that caution is required when analysing forensic soil samples as there is potential for loss and adulteration of soil samples from items such as footwear therefore the sample may not be representative of the parent soil from where it originated.

Cases in forensic science where soil evidence has been used include a suspected badger baiting case described in Morgan and Bull (16), where soil was taken from two spades and analysed using particle size distribution (see section 1:3) and compared to soil taken from a badger set. A link was made between the soil on the spade and the badger set. Soil analysis by XRD has been coupled with pollen analysis in a case described by Brown *et al.* (17) in which the locality of a grave of a missing couple was located. Junger (2) describes how the use of geological evidence has helped to solve several murder and kidnapping cases.

Further details about isotopic, DNA and pollen specific case examples can be found in the relevant chapters.

1:2.3 Outcomes of soil evidence

Brown *et al*. (17) describe how soil evidence can be used in a similar way to pollen evidence in that its main forensic value lies in providing associative evidence which may assist in proving or disproving a link between people or objects with places or other people or objects. There have many been cases where soil evidence has been used and the outcome was that the samples do not compare, that is they are not judged to be associated. There can be several reasons for this outcome including; no soil transfer took place, soil was transferred but was later removed by washing or rubbing, two or more soil samples were transferred resulting in a composite sample, the area has swift soil changes and the sampling was inadequate or that the suspect was not at the crime scene (11).

It has been stated (18) that when soil samples are compared there are three possible outcomes: that the questioned soil definitely did not come from the location of interest therefore the sample can be excluded; that the questioned sample could have come from the location of interest; and that the questioned sample almost certainly came from the location of interest.

Rawlins and Cave (2004) (19) mention that any particles within a soil sample which do not exist naturally are viewed as foreign and those which are specific to a particular location are seen as being unique, and can assist in increasing the discriminatory power of soil evidence. It should also be noted that two samples will rarely (if ever) show a perfect chemical match (20). This is true of samples that have origins only a few centimetres apart and is due to the inherent spatial variation found in soils, and the variance in results which arises in subsampling in the field and laboratory and also due to errors associated with both sample preparation and instrumental measurement. As a general rule, the more chemically similar two samples are the more likely they are to have come from the same source or from a chemically indistinguishable source. In contrast, if a sample is chemically very different to another sample then it is highly unlikely that the two samples came from the same (or even a similar) source. Another scenario to consider is if one sample shows significant bulk compositional similarities to another sample. If this is the case then it needs to be decided if the two samples could have been derived from the same source, if there is just one potential source or if there are several possible sources.

1:2.4 Reliability of soil evidence

The reliability of the outcomes of soil evidence is dependent upon several factors. These include the precision of the methods used to make the comparisons, the degree of variability in the soil properties observed at the location of interest and how common the observed points of similarity or difference are in the wider community (18).

There have been four main problems with the use of forensic geology in real-life forensic investigations. First, the problem of sample preparation which involves homogenisation prior to analysis; second, the use of several analytical techniques and procedures, some of which may be dependent upon one another and therefore do not provide independent corroborative data; third, the use of inappropriate standard samples to justify the discriminatory abilities of a specific technique; and fourth, and arguably the most important, the desire of analysts to match a sample rather than using a forensic protocol of excluding samples from having associations with each other (1, 21). These four points have been addressed when using the soil samples collected in this project and in any analyses used and also in the subsequent inferences made about the samples and their origin by not only looking at the similarity of the soil samples but by using an exclusionary approach in the interpretation of the results.

Within the field of forensic geoscience, it has been suggested that a technique should never be used to match or to suggest a sample almost certainly did come from a certain location. The forensic geoscience rationale should be to exclude a sample from a comparator by means of their biological, chemical and physical components. It is very rare to be able to match two soil samples. By their very nature soils can vary both within and between sites but may also contain similar characteristics to soils from another site. Therefore it can be more important to exclude a soil sample from another; however this does not stop some forensic practitioners from trying to "match" soil samples. It is important when interpreting the results of soil analysis to give care to the exclusion of samples and that those samples which do show similar characteristics are viewed in the context of distinctiveness (16, 21-23).

An exclusionary approach to the interpretation of soil samples was shown in a case about the importation of rare Falcons. Forensic scientists received a piece of climbing rope to analyse, and for comparson purposes, four samples of soils from locations across Mallorca known to bread the Falcons. After analysis, three of the sites were excluded from having

been the souce of the soil on the rope through grain size, SEM and chemical and pollen analysis. It was not possible to exclude the fourth site as having been the source of the soil of the rope. This evidence contributed to a guilty verdict and led to a custodial sentence (16).

1:3 Techniques currently used in forensic soil analysis

Individual soil samples can be characterised by their different properties, processes, profiles, limitations, potentials and management requirements. A soil property is an identifiable trait or quality by which the soil can be classified for example pH or colour. A soil process is a series of actions causing a change in one or several soil properties, for example erosion or weathering (9). To answer questions about the origin of soil samples, a combination of different techniques is often best, although in some cases a single technique can provide sufficient information. The most common cases whereby a single technique can be used are those where one or more exotic particles, one which does not naturally occur in an area, are identified in the sample using either a microscopic or chemical microanalysis. When no unique particles can be identified, the analyst relies on the comparison of bulk sediment properties including multi-element chemical analysis (20).

A wide range of techniques have been used to analyse the chemical, biological and physical properties of soil. One of the commonest methods of soil analysis is through colour. Antoci and Petraco (24) state that colour is a highly significant characteristic of soils and can be used to distinguish them. A Munsell colour chart is used to separate dried and sieved soils based upon hue, value and chroma, however, Janssen *et al* (25), used a filtered clay fraction of soil. A study by Dudley in 1997 (26) described how the use of colour can be applied to soils which have been ashed, increasing the discriminatory power. Croft and Pye (2004) (27) state how colour analysis using a Munsell colour chart is qualitative due to the ubiquitous nature of soil colour and the subjectivity of the user and therefore tried to use an instrumental method for soil colour analysis. The use of a spectrophotometer gives a more quantitative analysis

Another parameter of soil that can be used to provide discrimination is pH. Dudley (28) and Stutter *et al.* (29) determined the pH of soils and found that when combined with colour analysis it can increase the overall discriminatory power of soil analysis. Dudley and

Smalldon (30) found that pH values were consistent within each site sampled but there was little variation between sites therefore the pH values alone proved to be insufficient to provide differentiation between samples from different locations.

The density of a homogenous sample can be a useful property, as it is dependent on the nature and composition of the parent material. It can be used for both identification of soils and comparison of soils of a similar nature. Nute (31) designed apparatus whereby several density gradients could be made simultaneously using gradient tubes. This method also allowed for the removal of individual fractions from the gradient tubes for further analysis. Density gradient techniques were a favoured technique for the comparison of soil samples, but Junger (2) suggested that the technique was in decline due to the availability of more advanced techniques. Although the method developed by Nate is now obsolete, a study by Petraco and Kubic (32) employed density gradient analysis. However, this work suggested the use of Clerici's solution with distilled water rather than the previously used bromoform (BM) and bromobenzene (BB) to prepare the samples. This improved the range of densities within the gradients providing clarity to the weighted mineral fractions.

Another technique utilised to analyse soil is particle size distribution as used by Sugita and Marumo (33). Tthis technique is viewed as a primary assessment of soil in the field and has been utilised to discriminate forensic soil samples. The distribution can be determined using various methods such as traditional dry sieving (7, 34) and laser granulometry (20).

Fitzpatrick and Thornton (35) employed the use of scanning electron microscopy (SEM) to characterise sand samples. This study found that the use of SEM was more appropriate for sand samples than density gradient analysis and stated that sand cannot be treated in the same way as a soil sample. It also highlighted the need for SEM to be used in the context of any other data. Although SEM provides excellent results it is recommended that it is used with other techniques such the analysis of anions by capillary electrophoresis (36). By employing scanning electron microscopy coupled with energy dispersive x-ray analysis (SEM-EDX) elemental composition can be determined quantitatively. SEM–EDX was implemented in the 1989 Ruidoso plane crash, where the analysis determined that the soil found in the plane engine originated from the crash site and therefore could not have contributed to the cause of the crash (37). SEM-EDX was used by Cengiz *et al* (38) to analyse samples that were subjected to 9 tonnes / cm^2 of pressure. This homogenised the samples, which then increased elemental detection and provided a higher level of precision when analysing small samples. This paper also mentioned that SEM analysis (also discussed

in Bull *et al* (21)) has an added advantage that it can prove useful in detecting and determining any foreign particles within the sample such as naturally occurring pollen or spores, which could prove unique to a particular locality and thus provide a higher degree of discrimination. SEM-EDX has also been used by Pye and Croft (39) for the analysis of finely ground powdered samples, therefore eliminating the need to subject the samples to pressure or mounting in a medium and polishing.

Bull *et al.* (40) and Morgan *et al.* (16) suggest the use of ion chromatography (IC) or ion chromatography coupled with mass spectrometry (IC-MS) for the analysis of anions and cations. Siegel and Precord (41) used reverse phase High Performance Liquid Chromatography (HPLC) and found that the chromatograms of soil samples differed from each other quantitatively but not qualitatively and therefore concluded that while this technique is not individualising it could be an excellent presumptive test used in conjunction with other techniques. A more recent study by Bommarito *et al*. (42) used both HPLC and IC in the analysis of forensic soil samples.

Marumo *et al.* (43) used X-ray diffraction (XRD) analysis on the <0.05mm fraction of clays for mineralogical analysis. Prior to this, Graves (44) developed a rapid quantitative method for classifying and comparing the mineral content of a soil sample. XRD was compared with quantitative X-ray diffraction (QXRD) in a study by Ruffell and Wiltshire (45). The two methods were used to determine the mineral composition of fine soils and dusts and it was suggested that XRD is superior, however, both techniques proved to be complimentary to each other. XRF was also utilised by Rawlins *et al.* (46) in a study to establish the provenance of soils. Hiraoka (1994) (47) employed X-ray fluorescence (XRF) coupled with multivariate analysis to detect seven inorganic elements in soil samples for comparison. This proved to be a useful tool and provided high discriminatory power in determining soil provenance but also suggests that these techniques should be coupled with other data.

Fourier Transform Infra Red spectroscopy (FTIR) was used in a study by Liddell in 1997 (48) where spectra from samples were compared to reference spectra to successfully determine the composition of the organic matrix. FTIR was also utilised by Cox *et al.* (49) in the analysis of the organic fraction of the soil samples. The samples were analysed both prior to and after pyrolysis. The spectra produced after pyrolysis was subtracted from the first spectra which resulted in the spectra for the organic fraction. The advantage of this technique was that the inorganic fraction of the sample was not altered in any way and was therefore available for further analysis. Weinger (50) developed a novel approach for the

identification of minerals using an infrared microprobe with a diamond internal reflectron objective. This technique was rapid and reliable and of 96 mineral varieties analysed 77 were differentiated by their attenuated total reflectron (ATR) spectra. ATR was also employed by Madejova (51) to obtain information on wet clay samples.

The use of inductively couple plasma mass spectrometry (ICP) coupled with optical emission spectrometry (OES) or mass spectrometry (MS) is another technique applied for the analysis of soils whereby the concentrations of up to 50 elements can be determined simultaneously (52). ICP-OES was also implemented by Barreto (53) who compared two methods of extraction and concluded and proposed caution when using a single hydrochloric acid (HCl) extraction as it can lead to invalid results. This was confirmed by Navas and Lindhorfer in 2003 (54) who state that extraction can vary in its effectiveness and therefore not all elements present in the soil sample will be extracted successfully, highlighting the need for great caution and care to be taken when choosing an appropriate extraction technique. Additionally, Bull *et al.,* Pye *et al*. and Croft and Pye (18, 40, 55) have analysed carbon and nitrogen isotope ratios, which can be used to compare soil samples but, like IC, can show seasonal variation and therefore need to be used with care.

1:4 Environmental Factors affecting soil

Soil can change over a period of time both in its natural environment and once it has been collected, due to a variety of processes, which are discussed below.

1:4.1 Natural processes

Soil is a medium that is capable of supporting life – the biosphere (see figure 3), and is the product of weathering rock (the geosphere) by air and water (hydrosphere and atmosphere). All four of these spheres are reflected in the way in which soils are generated from rock by air, water and organisms (56).

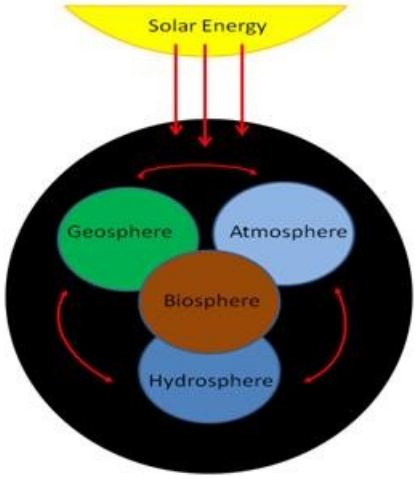


Figure 3: The four spheres which interact and are reflected in the generation of soils (56).

Over time, environmental factors transform geological deposits into soil profiles; this accumulated change is known as differentiation. The natural interactions between soils and their environments are additions, losses, redistributions and transformations, all of which can affect the profile of a soil sample, see table 1 (9).

Table 1: The environmental factors, which, over time change a geological deposit into a soil profile (9).

Interactions	**Examples**
Additions	Deposition – water or wind additions of soil material. Littering – accumulation of organic matter on the soil.
Losses	Leaching – removal of soluble materials. Erosion – removal of the surface layer of the soil.
Redistributions	Eluviation – movement of the material out of a zone. Illuviation – movement of the material into a zone. Pedoturbation – biological or physical mixing of soil materials. Salinization – accumulation of soluble salts. Alkalization – accumulation of alkali elements.
Transformations	Weathering – changes due to exposure to climatic elements. Decomposition – breakdown of mineral and organic materials. Humification – conversion of organic material into humus. Mineralization – release of mineral constituents from organic matter. Synthesis – formation of new mineral or organic species.

1:4.2 Seasonal and longer term changes

The biological (and some chemical) characteristics of soil often exhibit significant seasonal and longer-term change. These changes reflect fluctuations in ambient environmental conditions, especially temperatures, precipitations, humidity and soil moisture content. Therefore in forensic investigations it is important to collect soil samples as soon as possible after a crime has occurred especially as the soil properties in certain soils are

prone to very rapid change after samples are taken from the host environment. For example, some microorganisms may die as a result of a change in pH or oxidation. In order to minimise these effects, soil samples should be analysed as soon as possible after collection or be stored under conditions that minimise changes which occur after collection. If it is not possible to collect soil samples quickly the analysis should focus on the attributes that are least susceptible to change such as the chemical and physical characteristics of the most resistant constituents such as minerals or pollen (5).

It needs to taken into consideration that the drying of a soil sample can cause some of the soil properties to change. For example, any salts present in the sample will be concentrated and may crystallise on the surface, some minerals may oxidise and may change colour, as a sample dries the nitrate content may increase, the populations and activity of microbes may alter dramatically and the colour of the soil sample may lighten upon drying. These factors need to be considered both throughout the analysis and during the interpretation of the data (11).

The effects of temperature upon plant material (often present in soil samples) was investigated in a study by Virtanen *et al*. (57). The study looked at how long DNA stays intact for when it is stored in highly fluctuating environments by placing fresh, moist plant material in paper bags and leaving them in a shed in Finland for 18 months. During the 18 months, temperature varied from -18^{o}C to +25^{o}C. After 18 months DNA was extracted using Qiagen DNeasy Plant mini kit. To test the quality of the DNA, microsatellite PCR was performed for one or three microsatellite loci per sample. If regular PCR amplification and successful genotyping were carried out it was determined that the DNA was of a satisfactory quality. The study showed that all of the samples which were stored for 18 months, even those which had gone mouldy; were successfully able to have DNA extracted from them and be subsequently genotyped. This highlights the importance of plant material as potential forensic evidence even when it is collected after a long time period has elapsed. The study also highlights how plant fragments found on either a suspect or a victim should not be ignored as evidence; the genetic profiling of plants can be a useful aid in forensic investigations especially when other forms of evidence more commonly found such as body fluids are not available for analysis.

The use of plant material for DNA profiling despite an elapsed time period was also discussed by Coyle *et al.* (58) where amplified fragment length polymorphism (AFLP)

patterns of both fresh and dried marijuana samples were compared. It was found that there were no differences in the generated profiles even over an extended time period and when a wide range of template concentrations were used. This, along with the study by Virtanen *et al.*, shows the potential for the use of plant DNA to aid a forensic investigation whether this is fresh or older plant material.

1:5 The sample area – Norfolk

Soil samples were collected from 30 locations across Norfolk. Norfolk was the chosen sample area as the whole of the county could be accessed easily and samples could be transported back to the laboratory for storage promptly.

1:5.1 Norfolk Geography

Norfolk is a one of the largest counties in England and covers over 1.3 million acres. Norfolk lies on the east coast of Britain (see figure 4). To the north and east of Norfolk lie the North Sea, to the north-west lies The Wash. Norfolk is bordered to the south by Suffolk, by Cambridgeshire to the south west and by Lincolnshire to the west. The main towns in Norfolk are Norwich which lies to the east of the county, Kings Lynn in the west, Great Yarmouth on the east coast and Thetford which lies on the southern boundary. Norfolk has many rivers although none of them are very big; the most important one is the Great Ouse which meets The Wash at Kings Lynn, the Yare runs diagonally across the county to the sea at Great Yarmouth, the major tributary of the Yare is the Waveney which forms much of the counties border with Suffolk. Other rivers include the Bure in the north east and the Little Ouse in the south west. Norfolk is a county of diverse habitats and lacks only mountains (59-61). For more detail about Norfolk and its' soil types, vegetation areas and land use see chapter 4 – vegetation data.

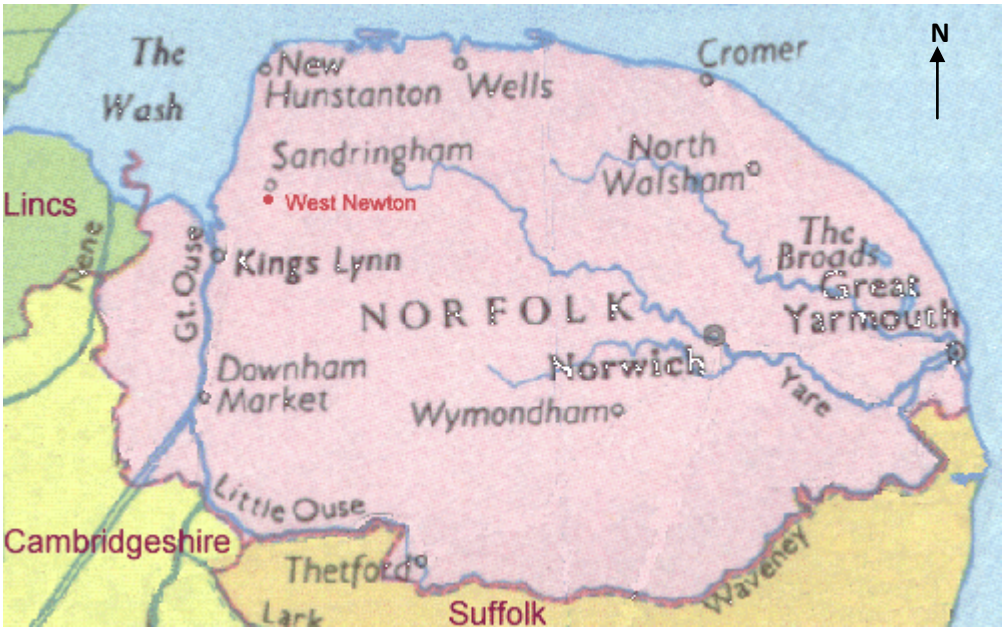


Figure 4: A map of East Anglia, UK showing the geography of Norfolk and its surrounding area. Souce: (60). (Scale not known)

1:5.2 Norfolk Geology

Norfolk is geologically young and is composed of five main geological units, chalk including Red chalk, Amptill Clay, Kimmeridge Clay and Corallian Limestone, Lower Cretaceous, Upper Greensand and Gault Clay, and Norwich Crag, red Crag and Chillesford Clay as can be seen in figure 5. The majority of Norfolk lies upon chalk, which is 98 % pure calcium carbonate. The chalk was formed in the Cretaceous Period 146 to 65 million years ago and is composed of millions of tiny calcareous plates of marine algae called coccoliths. These coccoliths form larger coccospheres which although larger remain microscopic. The Norfolk chalk also contains nodules (a small knobbly rock or mineral cluster) of flint. The flint is chemically very different from the calcium carbonate; it consists of silicon dioxide (also known as silica or quartz), which appears in bands throughout the rock. Red chalk, also found in Norfolk, differs from chalk only in that it contains iron oxide giving rise to its red colouration and hence its name (61-64).

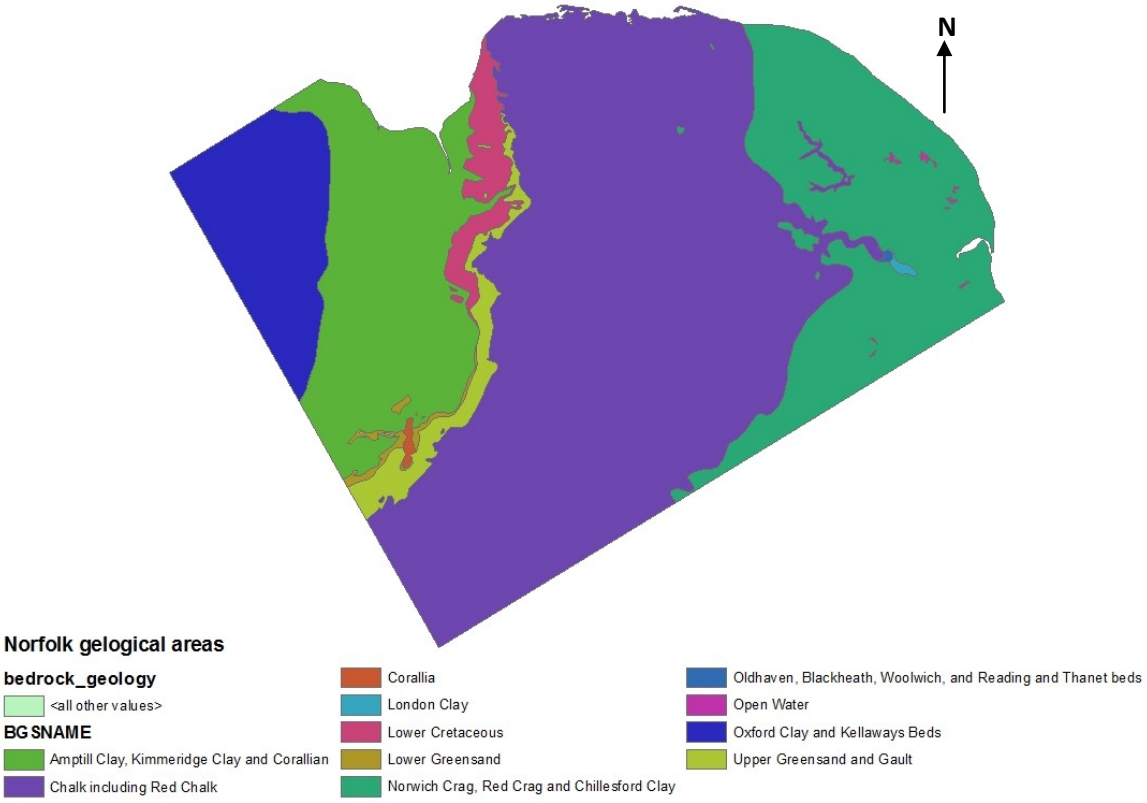


Figure 5: Geological regions of Norfolk: The different geological regions of Norfolk as highlighted by different colours according to the BGS name. (Created on ARC GIS by BGS, projection and scale unkown)

The Cretaceous Period of East Anglia's geological history can be divided into two major phases. The Lower Cretaceous was 146 to 100 million years ago and resulted in the formation of the Lower Greensand, Upper Greensand and Gault. The Upper Cretaceous was 100 to 65 million years ago and resulted in the formation of chalk (63). This was followed by the Quaternary Period; the Pleistocene epoch of which saw the formation of Crag groups and Chillesford Clay. The Waltonian stage of the Pleistocene was 2.5 to 1.8 million years ago and saw shallow water marine sediments forming the Red Crag of East Anglia. After the Baventian stage (1.6 million years ago) there was a break in the geological history of Britain which lasted approximately one million years. This break was followed by a period of marine transgression; the sea level rose relative to the adjacent land causing flooding. This laid down the Icenian Crag of the Bramertonian stage which includes Norwich Crag and Chillesford Clay, these are deposits which consist of estuarine and shallow-water marine sands and silts and gravel with molluscan fauna. The late Pleistocene was during the

last 300 000 years and is composed of three glacial phases where each successive ice sheet destroyed the evidence from the previous one. The Anglian stage occurred 300 000 to 250 000 years ago and it is during this period where the oldest glacial sediments found in East Anglia where great areas of till occur. The till was formed when retreating ice sheets or glaciers leave behind sand and clay, often glacial erratics can also been seen, many of those found in East Anglia have been found to have travelled from Scandinavia (62).

The distinctive geological areas of East Anglia should give rise to soils with unique chemical compositions. Soils from the different geological regions should be able to be discriminated on the basis of their individual elemental profiles. This is discussed in more detail in the strontium and trace element analysis chapter.

1:6 Sampling Techniques

There is little published guidance concerning the number of samples that need to be collected in order to adequately represent an area of interest, how large the collected samples should be and what sampling strategy and pattern should be adopted (18). If a soil sample is to be of forensic evidentiary value it must be representative of some of the properties at the forensic location of interest (19). Pye (65) suggests that at least three samples, ideally five, should be taken from every sampling site, each being analysed in triplicate and that a grid based sampling system aids in the creation of predictive maps from the data. Handheld global positioning systems (GPS) can be used to determine the exact location of each site to ensure accuracy.

In forensic geological work “control samples” should refer to samples of soil that are collected from known geological locations at a known date and time. In case work the locations may include the crime scene and alibi locations from defendants. A “questioned sample” may be obtained from a suspect, such as soil from the bottom of a shoe which was thought to be worn by the suspect at the time and scene of a crime. It can be helpful to know if similar soil is also present on other shoes worn by the suspect, or even by colleagues or members of his family. Such samples taken from other footwear are known as “comparison samples”. Sometimes samples are compared to type examples held in archives, which are well documented in terms of chemical and biological properties. These are known as “reference samples” (11) (65).

The need for careful sampling techniques has been highlighted in work by McKinley and Ruffell (66) who discuss how it is now commonplace for an expert witness to be questioned about how variable the material sampled from a crime scene is, and as statistics are now widely used to question comparative data, the robustness of sample variability and replication will often be questioned within a court of law. The paper looks at the pitfalls of paying inadequate attention to sampling spacing and number, the timing of sampling and also the statistical testing of results and highlights the need to look at the broader issues of sampling such as the location and timing of sampling as well as taking an adequate number of samples rather than just focussing on the type of analysis being carried out. The study states how by taking as many samples as possible in an informed manner, spatial interpolation can then be used to maximise information and make predictions about areas from where no samples are available suggesting for the work carried out as many samples as possible need to be collected to maximise the probability of making correct predictions about other areas. Murray (11) states how the collection of an insufficient sample or an inappropriate sample technique can result in a meaningless time and expense sampling for little contribution of data. Both of the papers highlight factors which were considered when sampling for this project, as enough samples needed to be collected in order to make predictions about the whole of Norfolk.

1:6.1 Different Sampling Techniques

Emphasis is often placed on the need to collect an unbiased set of samples; sufficient samples need to be collected to give an accurate impression of the nature of the background population that is being sampled. Random sampling is often seen as the best way of avoiding systematic bias, however, in order to take into account spatial variations in topography, geology and land use the number and distribution of the samples may be weighted. Purposeful sampling may be useful to test specific hypotheses and also in forensic investigations (65).

When the aim is to determine whether there is a high degree of similarity or a match between a questioned sample and a crime scene, samples are taken from the scene in a controlled manner. For example, samples should be taken from any obvious footwear impressions or from areas of obvious disturbance. For comparison purposes, samples should be taken from a number of other effectively randomly chosen points within the

scene or surrounding area. If it is seen that there is a high degree of similarity between the questioned and control samples, it is then necessary to determine if samples taken at the crime scene can be differentiated from those taken in the surrounding area or any other areas visited by the suspect.

1:6.2 Chosen Sampling Method

At each of the chosen sample locations, the surface litter was carefully removed from the sample location and then three samples of topsoil (see figure 1 (soil horizons)) weighing approximately 1kg each were collected, into Ziploc plastic bags, using a plastic spade (2). The samples were collected from an approximate 5m by 5m square in a pattern shown in figure 6; to allow for inter-site variability to be assessed. At each location photographs of the site were taken and the co-ordinates recorded using a Garmin eTrex Venture HC handheld GPS system. A record of any obvious land use or species present was recorded (see table 2). Between each of the sample locations the plastic spade was washed thoroughly using MilliQ water. Prior to analysis the samples were stored in a cool, dark location and were dried and agitated (67). The soils were then dry sieved (34) using a 2mm disposable mesh over a plastic container, with the mesh being changed between each sample and the container thoroughly cleaned with MilliQ water. The <2mm fraction of the soil was collected and used for subsequent analysis.

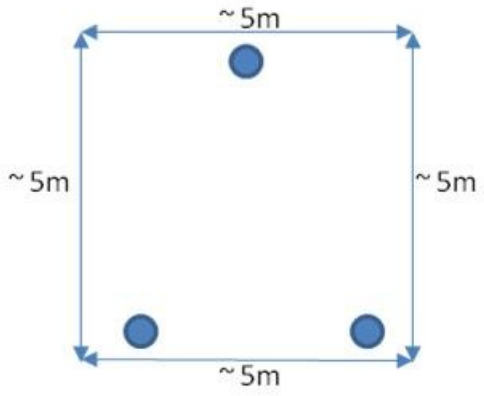


Figure 6: The chosen method of sample collection; three samples collected from within a approximate 5m by 5m square.

1:6.3 Sample Locations

The sample locations were chosen based upon the need for the sample area to cover the whole of Norfolk and the samples to come from different geological and botanical areas so that any potential links between the chemical and biological fingerprint of Norfolk soils could be seen.

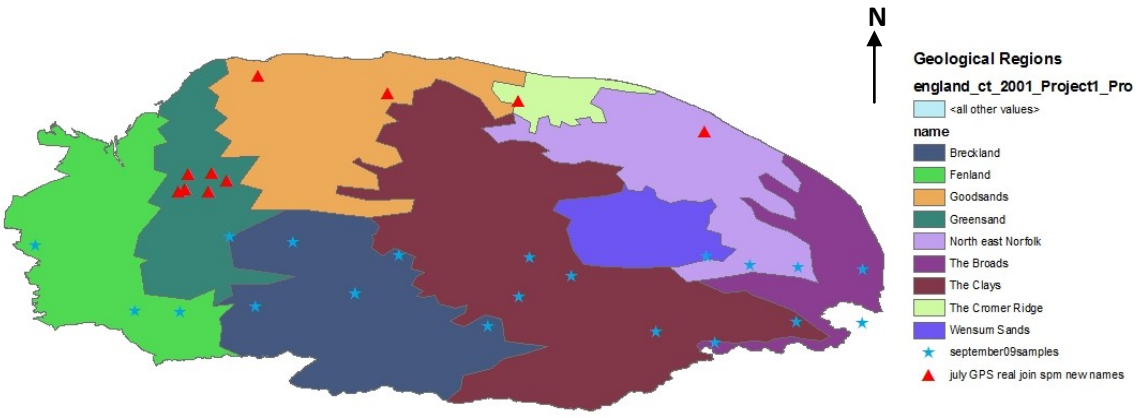


Figure 7: Soil sample locations shown on a geology map of Norfolk. Red triangles highlight sample collected in September 2009 and blue stars July 2010 (locations recorded with GPS). (Map created with ARC GIS 1:50 000 scale)

It should be noted that figure 7 appears different to figure 5 due to a different projection being used; figure 5 was created by the BGS with an unknown projection, and figure 7 created using ARC GIS (see 4:1.7).

The sample locations in respect to the botanical regions of Norfolk are shown in chapter 5. Each of the three samples collected from each location were stored at the University of East Anglia in dark, cool conditions were then used for the chemical and biological techniques discussed in the upcoming chapters.

Table 2: Soil sample locations and details of the soils collected. The sample code reflects the date of the samples collected followed by an underscore and number reflecting the location on each of the sampling days. i.e. 170610_10 was collected on 170610 from location 10.

Sample Code	**Location**	◦**N**	◦**E**	**Weather**	**Landuse / Observations**
160708_5	Knapton	52.51005	001.25157	Dry, some clouds.	Potatoes and weed, some plants starting to flower. Hedgerows and some woodland.
160708_12	Holt	52.54191	001.04591	Dry, sunny	Trees, many ferns, sloped ground.
160708_4	Wighton	52.55115	000.50490	Dry, cloudy	Crops, possibly wheat and rye
160708_9	Titchwell	52.56583	000.36467	Overcast, windy	Crops, possibly wheat
240708_1	Ashwicken	52.74433	000.52485	Very sunny and hot	By public footpath, swedes, ground very hard and sandy.
240708_2	Bawsey Country Park	52.74981	000.48213	Very sunny and hot	Many trees, lots of dried leaf litter. Site to the left of carpark.
240708_3	Bawsey Country Park	52.74916	000.48105	Very sunny and hot	Many trees, lots of dried leaf litter. Site in front of carpark.
240708_4	Bawsey Country Park	52.74563	000.46977	Very sunny and hot	In between a lake and a quarry. Quite sandy.
240708_5	GBase site 440	52.76441	000.55617	Sunny	Crops, near to road.
240708_6	Roydon National Park	52.77705	000.48783	Sunny	Trees, woodland area, horses in nearby fields. Ground quite hard.
240708_7	Roydon	52.77881	000.52918	Very sunny	Crops, ground quite hard.
220909_1	Barford	52.62844	001.10695	Warm and sunny	Harvested wheat fields.
220909_2	Shipdham	52.63206	000.87043	Warm and sunny	Arable fields, ploughed. Dry and rocky.

Sample Code	Location	◦N	◦E	Weather	Landuse / Observations
220909_3	Watton	52.56426	000.79056	Warm and sunny	Wheat fields, harvested. Clay like soil at the edge, more sandy further into field.
220909_3_4	Watton	52.56325	000.79104	Warm and sunny	Sandier soil of same field as 220909_3.
220909_4	Methwold Forest	52.54092	000.61008	Warm and sunny	Deciduous woodland, heavy ground organic cover.
220909_5	Methwold	52.53224	000.47395	Warm and sunny	Onion field
220909_6	Southery	52.53460	000.39293	Warm and sunny	Sugar beet field
220909_7	Emneth	52.64984	000.21145	Warm and sunny	Arable land, ploughed and drilled
220909_8	Marham	52.66529	000.56469	Warm and sunny	Arable, potatoes, harvested.
220909_9	Swaffham	52.65541	000.67855	Warm and sunny	Scrubland, near to EcoCentre.
230909_1	Brundall	52.62982	001.42548	Overcast, cool and breezy	Ploughed wheat field.
230909_2	Lingwood	52.61543	001.50409	Overcast, cool and breezy	Farrow land, gooseberries, quite dry and rocky.
230909_3	Halvergate	52.61152	001.59032	Overcast, cool and breezy	Irrigated agricultural land. Broads, ditches, blackberry bushes, clay like soil with decaying vegetation.
230909_4	Breydon Water	52.60722	001.70926	Overcast, cool and breezy	SSSI, lots of dog walkers, grazing marsh, wet grassland.
230909_5	Corton	52.51330	001.70721	Overcast, cool and breezy	Soil, crops, evidence of fertiliser

Sample Code	Location	◦N	◦E	Weather	Landuse / Observations
230909_6	Haddiscoe	52.51365	001.58944	Overcast, cool and breezy	Soil, no crops, quite hard.
230909_7	Bungay	52.47776	001.44219	Overcast, cool and breezy	Ploughed agricultural land, really fine soil, almost sandy.
230909_8	Woodton	52.49603	001.33462	Overcast, cool and breezy	Very hard, stoney soil
230909_9	Attleborough	52.50575	001.03157	Overcast, cool and breezy	Ploughed field by a road, lots of brambles and nettles as break between road and field. Soil very hard and rocky.
230909_10	Wymondham	52.55899	001.08801	Overcast, cool and breezy	Ploughed, agricultural land. Flint.
230909_11	Hethersett	52.59626	001.18164	Overcast, cool and breezy	By main road, ploughed. Wheat?
170610_1	Horsey	52.73586	001.64726	Warm and sunny	Scrubland in between fields, very dry, nettles. Rape?
170610_2	Ingham	52.78271	001.55101	Warm and sunny	Crops, agricultural land. By a public footpath.
170610_3	Worstead	52.76389	001.40259	Warm and sunny	By a bridle path and road. Crops, corn field.
170610_4	Aylsham	52.78913	001.22740	Warm and sunny	Small crops, irrigation system. Soil getting darker in colour into the distance. Well established hedgerows.
170610_5	Reepham	52.75271	001.10000	Warm and sunny	Wooded, srub, lots of organic litter and leaf cover.
170610_6	West Rudham	52.82618	000.71544	Warm and sunny	Crops, one side of field next to a main road. Well established hedgerows. Soil fairly light in colour.

Sample Code	Location	◦N	◦E	Weather	Landuse / Observations
170610_7	Hillington	52.78069	000.54583	Warm and sunny	Crops, open fields, not much hedgerow.
170610_8	Tottenhill	52.66403	000.44422	Warm and sunny	Forest, ferns, lots of leaf litter. Small dired up ditch between forest and road. Quite peaty.
170610_9	West Dereham	52.59750	000.46285	Warm and sunny	Crops, very large field. Pylons. Not all of the field enclosed by hedgerows. Dry and stoney soil.
170610_10	Thetford	52.39437	000.82798	Warm and sunny	Trees, lots of nettles, dead leaf litter.
170610_11	Diss	52.38443	001.07804	Warm and sunny	Ferns, nettles, trees, lots of leaf cover on ground. Opening by a road, off a footpath.
170610_12	Harleston	52.43122	001.29626	Warm and sunny	Two fields, separated by a private road. Crops. Soil quite dry and rocky.

1:7 Geographical Information Systems

Geological maps are commonly provided in both solid and drift form. Both of these are of value as they assist one in being able to understand the landscape. A solid geological map shows what kind of rocks occur where; they can indicate the locations of old mine workings, subsident grounds, caves as well as the parent geological materials for soils. In comparison, a drift map shows the extent of the overlying materials and hence possible digging locations, and as this overlying material is often loose, they can indicate likely areas whereby this loose material is likely to transfer to the suspect (56).

Spatial data are located on the earth's surface. Geographic Information Systems (GIS) have been developed to allow us to map and analyse this spatial information. GIS may refer to either systems or science. GI systems are essentially tools that have been developed to input, manage and analyse spatial data whereas GI science refers to the problems which arise from handling spatial data in a GI system (56).

GI systems store two kinds of data:

1) Spatial information
2) Non-spatial attribute information.

GIS is different to other information systems and is useful in forensic science due to the spatial component and the ability / need to analyse this spatial data (56).

1:8 Aims of the project

The aim of this project was to develop and validate the use of various chemical and biological profiling techniques for the forensic geographical provenancing of soils as highlighted in figure 8. The soil samples were taken from locations throughout Norfolk. Subsequently attempts were made to develop techniques allowing the analysis of plant DNA present in the samples using the mat*K* gene and traditional sequencing methods and also using matrix assisted laser desorption / ionisation coupled with time-of-flight mass spectrometry (MALDI-ToF-MS). The pollen present in the soil was identified and counted and $^{87}Sr/^{86}Sr$ ratios in the soil were being measured using a multi-collector inductively coupled plasma mass spectrometer (MC-ICP-MS). It was proposed that if possible the data would be collated into a GIS (geographical information system) which could then be used to determine where soil samples originate from, or if they could be excluded from each other based upon their chemical and biological properties.

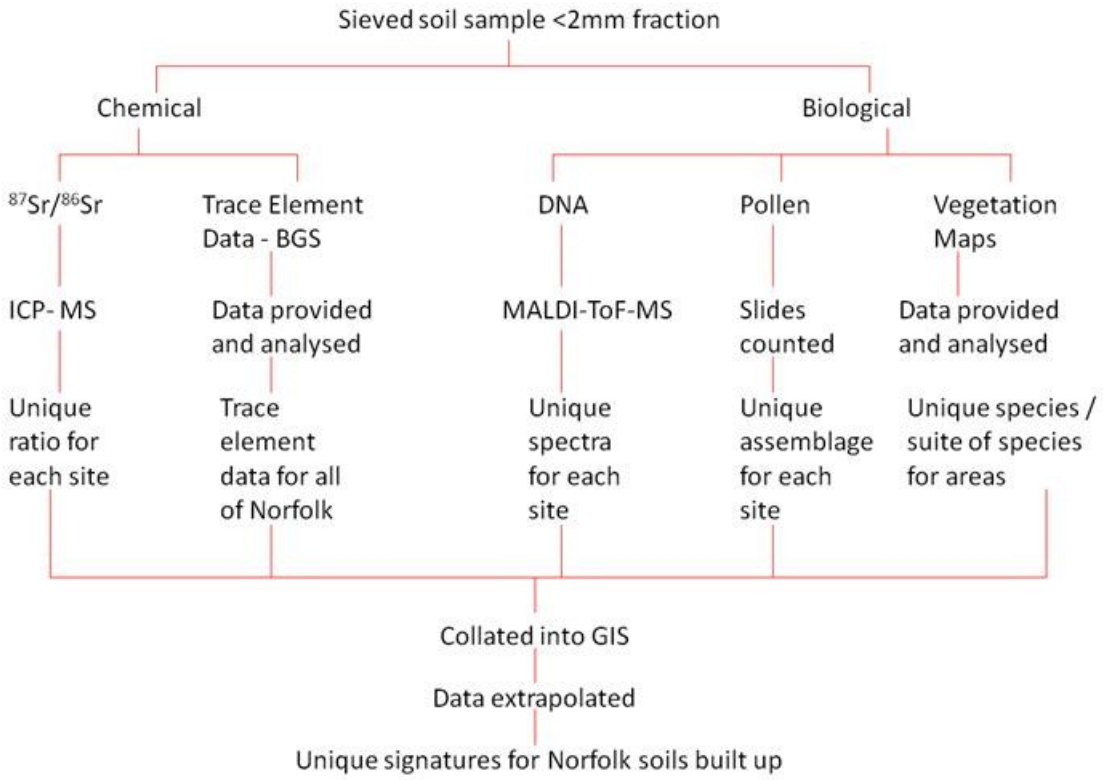


Figure 8: The chemical and biological techniques that were used in order to build up a unique signature for Norfolk soils.

The need for this work has been highlighted for many years. Traditionally, the use of forensic geology, although informative, is rarely applied due to the perception that laboratory and training costs are too high given the limited practical use of the evidence (2). Junger (2) and Sugita and Marumo (68) state how soil is classed as trace evidence, and most current techniques used to compare soil samples are based on geological properties. This highlights the need for new techniques, which can routinely be used to analyse soil, to be developed since, as stated in Fraysier and van Hoven (69), most forensic laboratories cannot afford geologists and therefore the analysis of soil is usually reserved for high-profile, serious crimes. Junger also states how future research needs to be carried out in order for the field of forensic geology to be fully utilised, and also how work needs to be carried out on a small scale as well as a larger scale, i.e. can soil samples be differentiated at a one metre scale? This suggests that when soil samples are collected and subsequently analysed, both inter and intra site variability need to carefully considered. This contradicted a statement in a paper by Petraco *et al.* (8), where it was stated that there is a belief that due to soil being consistent within any geological area it is not a valuable form of evidence. Brown *et al*. (17) state how although a soil type is not unique to a specific location, many of

the components of the soil may be, and more importantly a combination of the soil properties can dramatically reduce the number of possible locations for the origin of a soil sample and can also increase the likelihood of strong associations between a given soil sample and a geographical origin.

Currently, in the field of plant DNA testing, there are primarily two areas being investigated; the linking of marijuana to aid in forensic drug investigations and the linking of plant material to either a suspect or victim to make an association with a crime scene, highlighting the gap for the mapping and predictive modelling of the DNA and pollen of an area which will also aid in making associations between suspect and crime scene (70). Virtanen *et al*. (57) discuss how plant fragments found on either a suspect or a victim should not be ignored as evidence; the genetic profiling of plants can be a useful aid in forensic investigations especially when other forms of evidence more commonly found such as body fluids are not available for analysis. This shows how plant DNA has been used within forensic science and suggests the importance of trying to use this technique more readily. As suggested by Horswell *et al.* (71) and Heath and Saunders (67) most forensic laboratories have the equipment for DNA profiling, and therefore if techniques can be developed to profile the DNA of microbial communities in soil this can routinely be implemented in the analysis of environmental samples.

CHAPTER 2 – The Analysis of Plant DNA in Soil

This research looks at the use of DNA analysis for the analysis of soils, specifically whether or not plant DNA can be extracted from soil samples and analysed using RT-PCR of the *mat*K gene.

2:1 Introduction to DNA

Deoxyribonucleic acid contains all of the information needed to make the molecules from which cells are built. DNA is a polymer which consists of a large number of repeated monomer sequences. The monomer units of DNA are nucleotides, each of which consists of three components:

1. a 5-carbon sugar –deoxyribose
2. a nitrogen containing base attached to the sugar, and
3. a phosphate group.

The phosphate group and the deoxyribose are common to all nucleotides. However, the nitrogen bases vary; they can be one of four types; adenine (A), guanine (G), cytosine (C) and thymine (T). It is the distinct arrangement of these four bases which regulates the production of proteins and enzymes in a cell. As a result of this, DNA is the genotype (genetic identity) of an organism which subsequently leads to particular DNA profiles being associated with particular organisms; an individual of any species has a unique DNA fingerprint (72).

Unique DNA sequences, which are present in all species, are used as biomarkers for the detection of cells from that species. The easiest and most convenient way of detecting these species is through polymerase chain reaction (PCR) based DNA fingerprinting methods (72). Kitts (73) describes how all DNA fingerprinting techniques produce a unique pattern of nucleic acids amplified (by PCR) from a sample; in the case of microbial DNA this pattern reflects the microbial community from an environment. DNA profiling comprises any DNA-based techniques that allow for the identification of DNA from a certain individual or group of individuals within a community of organisms. DNA fingerprints can be used to determine the identity of a specific DNA sample or to assess relatedness between samples. It is hoped that within this project, DNA fingerprinting might be used to identify the DNA from plant species within an area to determine the uniqueness of an area. Most methods

used for DNA fingerprinting are highly specific, highly sensitive and largely independent of the growth state or the organisms in question (72, 73).

2:2 Non-human DNA

Bock and Norris in Butler (74) highlight how, although non-human DNA is not yet routinely used, it has helped to link suspects to crime scenes and aided in criminal investigations. Yoon (75) described the first time that botanical DNA evidence was used in a case. Genetic testing was used on two small samples collected from the back of a pick-up truck and it matched only the tree under which the victims' body was found. This placed the suspect at the scene of the crime and he was found guilty largely based upon this evidence.

DNA technology has long been used in animal wildlife studies for forensic purposes and has two main outcomes; species identification and identification of the individual organism. Tsai, *et al.* state how DNA profiling offers the same prospect for botanical samples; however for many botanical species there is a limited knowledge of the genome, and therefore whole genome screening methods such as DNA fingerprinting, RAPD and AFLP (see section 2:3.3) have been used. These methods are adequate as a means of rapid screening but problems with the reproducibility of these techniques mean they are not yet routinely used in the forensic analysis of botanical evidence (76).

There have been several issues surrounding whether the application of non-human DNA evidence is ready to be presented in court, highlighted in Sensbaugh and Kaye (74). The issues include the novelty of such techniques and the validity of statistical interpretation of the results. Often the techniques used in plant and bacterial DNA fingerprinting have not undergone the same rigorous testing human DNA techniques have been put through. The study also highlights how reference databases used for comparison purposes take many years to build up and may not even be available in some cases; this also affects calculating the probability of a match. A study by Tsai *et al.* (76) has highlighted that forensic botanical evidence can be hampered by the lack of appropriate database for the comparison of samples.

2:3 Analysis of DNA

DNA in a forensic case typically involves samples from both the crime scene, and from suspects and victims (77); the process of DNA analysis is highlighted in figure 9.

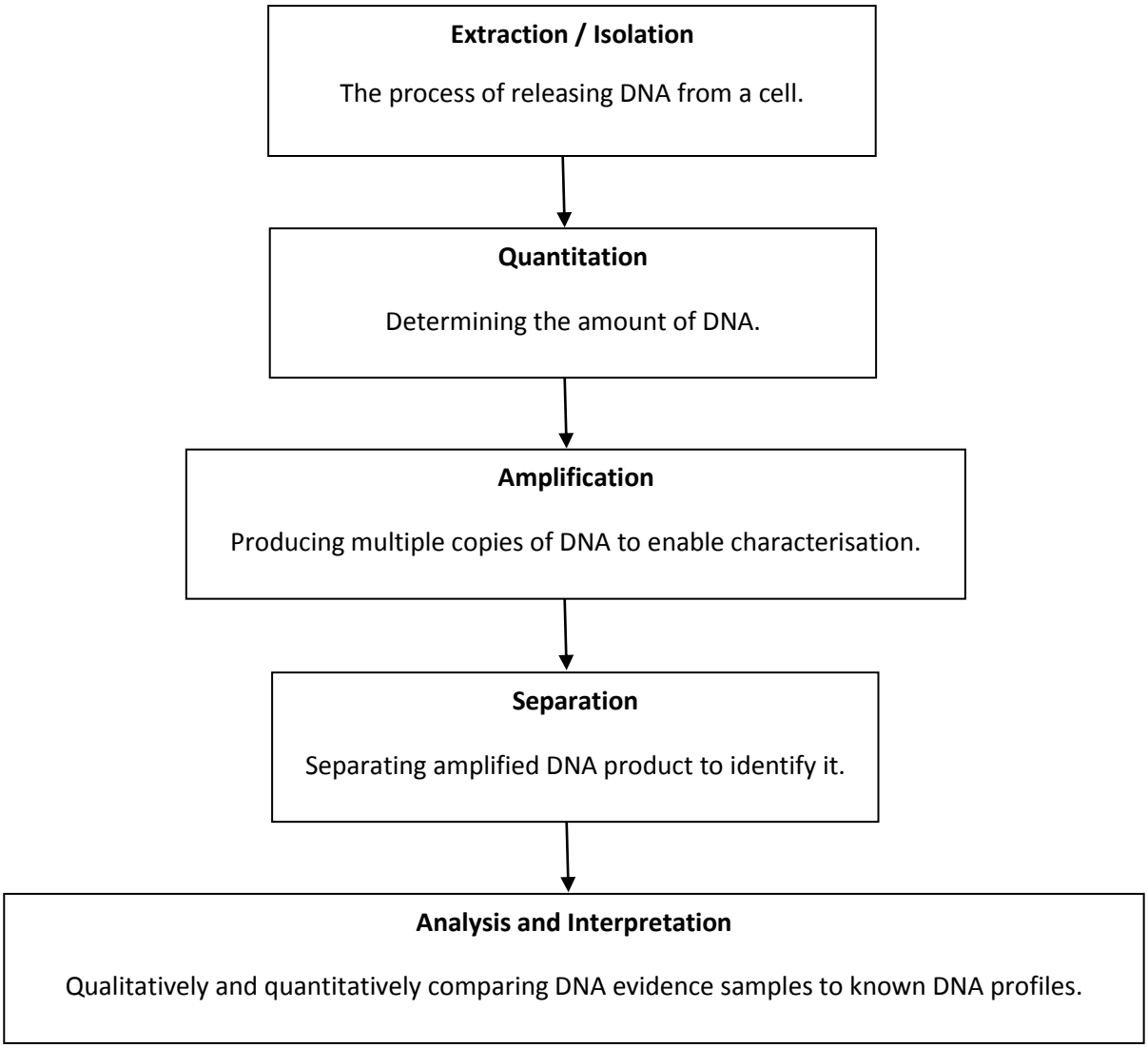


Figure 9: Flowchart showing the main principles of DNA analysis. (78)

2:3.1 DNA extraction

There are two main parts to DNA extraction; cell lysis and the purification of DNA and removal of PCR inhibitors. Cell lysis is needed to break open the cells in order to access the DNA. DNA may be isolated by simply boiling the cells but this may produce DNA which is not of a sufficiently high quality for downstream applications such as PCR. Also, during the boiling process, the DNA is not always completely separated from proteins and other structural elements, which may also inhibit downstream applications. In order to release clean DNA, cell lysis is used to disrupt the phospholipid cell membrane and nuclear membranes. Lysis uses a detergent solution, lysis buffer, which contains sodium dodecyl sulphate (SDS) which disrupts the lipids and thus disrupts the membranes. As DNA is negatively charged (due to the phosphate groups in its structure) and its solubility is charge

dependent, and hence pH dependent, the lysis buffer contains a pH-buffering agent to maintain the pH so that the DNA remains stable. Often the lysis buffer will also contain proteinase, an enzyme to remove proteins bound to the DNA and to destroy cellular enzymes that would otherwise digest the DNA upon cell lysis. Heat and agitation are sometimes used to speed up the enzyme reactions and the lipid solubilisation (79).

The second part of DNA extraction involves the removal of any elements that inhibit either the DNA extraction process or any downstream applications. The inhibitors are mainly problematic in three areas of DNA extraction (79):

1. Interference with cell lysis.
2. Degrading nucleic acids or otherwise preventing their isolation after lysis.
3. Inhibition of polymerase activity during PCR after successful purification.

These two parts of DNA extraction vary slightly depending upon the method being used (see methodology).

2:3.2 DNA Amplification

Polymerase Chain Reaction (PCR) is used to amplify specific DNA sequences. Millions of copies of the target sequence can readily be obtained by PCR if the flanking sequences of the target are known. A PCR cycle consists of three steps:

1. **Strand separation** – the two strands of the parent DNA molecule are separated by heating the solution to 95 °C for 15 seconds.
2. **Hybridisation of primers** – the solution is cooled to 54 °C to allow each primer to hybridize to a DNA strand. One primer hybridizes to the 3’ end of the target on one strand and the other primer hybridizes to the 3’ end of the complementary target strand. Primers are typically from 18 to 30 nucleotides long. Because the primers are present in large excess, parent DNA duplexes do not reform.
3. **DNA synthesis** – the solution is heated to 72 °C, the optimal temperature for *Taq* DNA polymerase (a heat stable polymerase which comes from *Thermus aquaticus*, a thermophilic bacterium that lives in hot springs). The polymerase elongates both primers in the direction of the target sequence because DNA synthesis is in the 5’ to 3’ direction. DNA synthesis takes place on both strands but extends beyond the

target sequence allowing for the exponential acculumaltion of the target product. (80).

PCR allows extremely small quantities of DNA to be amplified, under optimal conditions DNA can be amplified from a single cell. As a result of PCR the sensitivity of DNA profiling has increased and it is now possible to analyse trace evidence and highly degraded samples successfully, even though this does not have a 100% success rate.

PCR takes advantage of the enzymatic processes of DNA replication. During every cell cycle the entire DNA content of a cell is duplicated. In order to amplify specific regions of DNA, this copying of the DNA can be replicated outside of the cell *in vitro*.

The components which are needed for PCR are:

Template DNA:

The amount of DNA which is added to PCR is dependent upon the sensitivity of the reaction. For most forensic purposes the PCR is highly optimised so that the reaction will work with low levels of DNA. Most commercial kits require between 0.5 and 2.5 ng of extracted DNA for optimum results. It has to be noted that as the amount of template DNA is reduced the profiles can become more complex and difficult to interpret.

Taq DNA polymerase:

Taq DNA polymerase is isolated from the thermophilic bacteria *Thermus aquaticus*. The *Taq* polymerase can tolerate the high temperatures that are involved in the PCR and works optimally at 72 – 80 °C. The use of *Taq* DNA polymerase (which is thermostable) greatly amplifies the PCR product and also increases the specificity, sensitivity and yield of the reaction. The *Taq* polymerase exhibits significant activity at room temperature that can lead to the creation of non-specific PCR products. This non-specific binding can be reduced by adding the enzyme to a pre-treated “hot-start” reaction; this also improves the specificity and yield of the PCR.

AmpliTaq Gold polymerase is a modification of the commonly used *Taq* polymerase. The AmpliTaq Gold enzyme is inactive when it is first added to the PCR – it is only activated after 10 minutes of incubation at 95°C. This “hot-start” enzyme allows the PCR to start at an elevated temperature minimising the non-specific binding that can occur at lower temperatures.

Primers:

The primers define the region of the genome that will be amplified. They are short, synthetic pieces of DNA that anneal to the template molecule either side of the target region. The primer sequences are therefore limited to some degree by the DNA sequence that flanks the target region. Primers are normally between 18 and 30 nucleotides long, and have a balanced number of G-C and A-T nucleotides. Primers should not be self-complementary to any of the primers that are in the reaction. Self-complementary regions will result in the primer pairing with itself to form a loop, whereas primers that are complementary will bind to each other to form primer-dimers. The temperature at which primers anneal to the template DNA depends upon their length and sequence – most primers are designed to anneal between 50 and 65 °C. This can be accurately predicted from the sequence.

Magnesium chloride, nucleotide triphosphates and reaction buffer:

Magnesium chloride is critical to PCR as it stabilises the interaction between primers and DNA once the primers bind to the template DNA to form a primer-template duplex. The concentration of the $MgCl_2$ is typically between 1.5 mM and 2.5 mM; the higher the concentration of $MgCl_2$ the more stable the primer-DNA complex becomes. The presence of $MgCl_2$ is also important to PCR as it is required for *Taq* polymerase to work.

The nucleotide triphosphates are incorporated into the nascent DNA strand during replication and the four deoxynucleotide triphosphates are in the PCR mixture at equal concentrations (approximately 200 µM). The reaction buffer maintains the optimal pH and salt conditions required for the PCR reaction (81).

Although PCR is routinely used in forensic science, the technique does have some disadvantages (problems encountered when using the technique) and these must carefully be considered before carrying out any PCR reaction, weighing up the outcomes against the possible disadvantages to ensure that the forensic evidence is maximised. The possible disadvantages include:

1. The target DNA template may not amplify due to the presence of PCR inhibitors in the extracted DNA.
2. Amplification may fail due to sequence changes in the primer-binding region of the genomic DNA template.

3. Contamination from other sources stored beside the forensic samples or from humans or from previously amplified DNA samples is possible without careful laboratory technique and the use of validated protocols (70).

2:3.3 DNA analysis and Interpretation

Following extraction and amplification, the DNA (whether it be plant or microbial) can be analysed using one of several techniques some of which are discussed in more detail in table 3.

Table 3: Summary of the different techniques that can be used for the analysis of soil DNA.

	Applications	Main Steps of Analysis	Advantages	Disadvantages	References
Terminal Restriction Fragment Length Polymorphism (TRFLP)	Ability to characterize the microbial communities from different environments; this allows the determination of spatial patterns. Can be used in criminal forensics to link soil samples. Able to characterize functional diversity in microbial communities.	DNA extraction. PCR amplification of the target gene with fluorescently labelled primers. Digestion of the amplified fragment with restriction enzymes. Separation and visualisation of terminal-labelled fragments using automated capillary based electrophoresis. Data analysis for peak identification.	Rapid- large amounts of information generated rapidly due to the availability of automated systems. A resolution previously unachievable. The potential to use TRF data to search existing databases for matching sequence data. The availability to correlate microbial species data to physical-chemical data of the environment.	PCR can differentially amplify templates preventing quantification of species. The detection of individual TRF is limited by electrophoresis. Patterns which result from a single enzyme digest do not result in accurate species or community identification. Artificial peaks in the TRF pattern can occur due to the incomplete digestion of amplicons. There is the possibility of TRF length overlap. TRFP are destructively sampled leading to complexity in identifying the organisms responsible for a particular element of the profile.	(67, 72, 82-87)

Randomly Amplified Polymorphic DNA (RAPD)	Used for genetic mapping and population studies. Can be used to produce a biochemical fingerprint of a particular species. Has the capability of screening the differences in the DNA sequences of two different plant species.	DNA extraction. Amplification of DNA fragments by PCR using randomly chosen primers at low annealing temperatures – this results in the amplification of multiple loci. The amplified fragments are separated by their polymorphisms by agarose gel electrophoresis. The separated fragments are visualised. It is then possible to clone and sequence the DNA band of interest (longer specific primers which target reproducible amplification and the detection of the differences need to be developed for PCR).	The total analysis can be completed in one day with relative ease. No expensive techniques such as sequencing or cloning are used. Useful for analysing large amounts of sample. If cloning of the RAPD fragments is desired this can be done rapidly and then cloned loci will not contain repetitive sequences.	A lack of specificity due to the easier reaction conditions and low annealing temperatures. On average, only half of the RAPD polymorphisms detected between species would be mapped. Most RAPD markers are dominant which makes them difficult to be used in population studies. RAPD is sometimes not reproducible.	(58, 72, 88-94)
Amplified Fragment Length Polymorphism (AFLP)	Used for "basic" genetic diversity and variation studies. A reliable method for the detection of molecular markers. Used for genetic mapping and	DNA extraction. DNA digestion with two restriction enzymes. PCR amplification of the restricted fragments using primers that are complementary to the adapter sequences. The	PCR technique is fast. No sequence information is required. Large numbers of polymorphisms are generated, and the method is highly sensitive to polymorphism detection at	Some bands overlap and are considered to be homogenous. Problems caused by polyploids. The number of AFLP polymorphisms detected can be affected by the restriction	(72, 95-99)

	evolutionary studies.	restriction fragments are extended in the 3’ end by the addition o0f selected nucleotides; this determines the selectivity and complexity of the amplification. Separation and visualisation of the resulted AFLP fragments.	the total genome level. A high multiplex ratio is possible. Enhanced performance in reproducibility, resolution and time efficiency. Can be automated.	enzyme or primers which are used.	

Of all the different DNA fingerprinting techniques it would appear that TRFLP holds the most potential for the analysis of soil samples, whether the target DNA is plant or bacterial (72). An example of the potential of TRFLP DNA fingerprinting is shown in a study by Horswell *et al.* (71), where by using hypothetical crime scenes bacterial DNA profiling by TRFLP was used to successfully establish differences between soils from different locations. Heath and Saunders (67) looked at the feasibility of comparing soil samples using a bacterial DNA profiling method that could be carried out in all laboratories that perform human DNA profiling and used TRFLP stating that it was the most suitable for forensic applications because as described in Liu *et al.* (100). TFR's can be detected by equipment that has fluorescence detection capabilities. These are routinely found in forensic laboratories.

2:4 DNA Barcoding of plant DNA present in soil:

DNA barcoding is based on the premise that a short standardized sequence can distinguish individuals of a species as genetic variation between species exceeds that within species. (101) DNA barcoding follows the basic principle of taxonomic practice of associating a name with a specific reference collection in conjunction with a functional understanding of species concepts. (102) DNA barcoding datasets are composed of short DNA sequences from several individuals of a large number of species. There is some controversy as to the value of DNA barcoding; there is a perception that this new method of species identification would diminish rather than enhance tradition taxonomy which is morphology based. It is thought that species identification based solely on the amount of genetic divergence could result in incorrect species identification. It is also often perceived that DNA barcoding is a means of reconstructing phylogenies when actually it is a tool largely meant for identification purposes. However, there have been several studies which support the use of DNA barcoding as a useful tool for species identification (103).

It must be remembered that in a DNA sample taken from soil, alongside DNA from any vegetation there will also be microbial and bacterial DNA present and the possibility of the presence of fungal DNA. Bacterial DNA may be present in much larger amounts than the other DNA as bacteria are known to be present in large amounts in soils. Bacteria are simple, tiny organisms with their DNA contained in one long strand. Their DNA has a

tremendous number of base pairs. However, it is the plant DNA which is of interest in this work.

2:5 Possible methods of analysing the DNA present in soil:

2:5.1 Randomly Amplified Polymorphic DNA (RAPD)

RAPD is a PCR-based technique for comparing DNA extracted from soil, however unlike many other approaches for analysing DNA it does not rely on primers that amplify the 16S rRNA gene. Amplification is achieved with short primers of arbitrary sequence that amplify random portions of the template DNA. RAPD can be combined with fluorescent labelling and fragment analysis on automated DNA sequencers achieving high throughput analysis of samples.

An advantage of RAPD is that it avoids 16S primers; this is advantageous as 16S primers may not bind with equal efficiency to all species of a microbial community. A problem with 16S rDNA based techniques is that it can lead to a skewed picture of the diversity within a community. This is not such a problem in forensic applications only concerned with comparing samples when the 16S bias is not important to the interpretation of the results as long as that bias is applied uniformly to all the samples typed by the same method. An important consideration to be made is the requirement of multiple reactions with different primers as each would require a PCR reaction with additional template DNA, which may not always be available in large amounts. There are also concerns surrounding the reproducibility of RAPD results between laboratories (104, 105).

2:5.2 Amplified Fragment Length Polymorphism (AFLP)

AFLP offers many of the advantages of RAPD but with few of the drawbacks. Compared to some of the other techniques used in soil DNA analysis the methodology of AFLP is complex, however the availability of commercial kits makes it simple to practice. For AFLP analysis, DNA extracted from soils is digested with two restriction endonucleases, such as *Mse*I and *Eco*RI. The adapter fragments of DNA are then ligated to the ends of the DNA fragments to generate the template DNA for the subsequent PCR steps. The digestion and ligation can occur in a single reaction considerably simplifying the protocol. The genomic fragments with adapters at each end are amplified using preselective amplification. At this

point, an arbitrary nucleotide may be added to the end of the primers to reduce the number of amplified sequences and simplify the final DNA pattern. Finally, a second PCR reaction known as selective amplification is carried out with fluorescent-dye labelled primers for the detection of DNA fragments.

Advantages of AFLP are that it adapts easily to most automated DNA sequencers and software. However, the kits for performing AFLP profiling on bacteria are intended for analysis of one bacterial type at a time. If run on a community sample, there are often too many bands to easily make comparisons between samples. This may not be such a problem for this work as it can be avoided by some approaches that use primers designed for larger genomes such as plants. These bind less frequently and provide a less complex banding pattern for microbial communities. Larger microbial communities can also be typed by using PCR primers to vary the sensitivity; this can also make the technique applicable to certain strains of bacteria. Like RAPD, AFLP avoids the 16S rRNA gene and presents a less biased view of the microbial community within the soil sample.

2:5.3 Terminal Restriction Fragment Length Polymorphism (TRFLP)

TRFLP is an adaptation of Amplified rDNA Restriction Analysis (ARDRA). It is of high sensitivity and is relatively easy to use and is therefore the commonest technique used in microbial typing. It has been shown by Horswell *et al.* (71) that TRFLP has a great potential for soil analysis in forensic science.

The extracted DNA is amplified with primers (most commonly from the 16S gene), and the PCR products are digested with restriction enzymes. Terminal fragments which are fluorescently labelled are detected. TRFLP is highly sensitive due to the incorporation of fluorescence and the detection of only the terminal fragments, which makes the interpretation of results easier. Typically, the range of fragments that can be analysed ranges from 50 to 600 bases.

Not only is TRFLP rapid and sensitive, other functional genes can be used in addition to the 16S TRFLP to provide an increased power of discrimination which is especially relevant for forensic work. However, a concern with TRFLP analysis is its lack of standardisation. The conditions used for PCR have varied in different studies (71, 100) which could cause complications if results were to be compared between different laboratories. It also needs to be decided how many digestions are required to achieve a match between samples with

enough weight to be useful in court, without unduly consuming the evidence. Also, unlike some of the other techniques, commercial kits are not available and therefore most protocols involve a mixture of PCR kits and outside reagents such as the specially labelled primers.

2:5.4 rRNA Intergenic Spacer Analysis (RISA)

RISA is unlike many protocols that use PCR primers to amplify the gene for the small ribosomal subunit. RISA analysis amplifies the intergenic region located between the genes for the small and large ribosomal subunits. This region is known as the IGS and is highly variable in size between bacterial groups, ranging from 50 base pairs to more than 1.5kb. The amplification products are run on a gel and are silver stained. An adaptation of the technique, automated RISA (A-RISA) uses fluorescently labelled primers for the PCR amplification and then the final analysis is conducted on an automated DNA sequencer.

A-RISA is highly sensitive and reproducible. With the RISA technique, a microbial community may produce hundreds of distinct bands. The protocol is very simple and does not involve enzymatic digestion steps. However, like those techniques which rely on the PCR amplification of the 16S rRNA gene, RISA is subject to similar biases due to it employing PCR primers targeted to a specific region of DNA that it not absolutely conserved across all bacteria.

2:5.5 Most suitable technique for DNA profiling of soil:

Of all the DNA based techniques available, only PCR based methods have the sensitivity required for soil based methods. Due to the fact most laboratories have access to automated DNA sequencers it would seem that AFLP, A-RISA and TRFLP are the methods that can most accurately and sensitively distinguish between microbial communities. AFLP has the advantage of capturing a wider range of soil community members due to the nonspecific range of primers it uses, also it is relatively easy to standardise and there are commercial kits available. A-RISA provides a sensitive analysis and has the shortest turnaround time of the three methods, making it an ideal technique if a high throughput of samples is required. TRFLP has the greatest history of use for the analysis of soil bacterial communities and therefore has gone through extensive scientific, although not forensic,

validation. Also, TRFLP has the ability to target specific genes other than 16S rRNA, which may be important for future applications.

2:5.3 Why plant DNA was chosen over bacterial DNA

Although bacterial DNA is present in large amounts in soil, it was chosen to analyse the plant DNA present in the soil samples. This was because at the starting point of the project there were only single species of plants identifiable through DNA analysis rather than a complex mixture and the aim of this work was to analyse the DNA of suites of species giving a unique DNA fingerprint of that soil. It was also thought that plant DNA was more stable and not as variable over a short space. There was also the potential for a lot of development of primers for different plant species. DNA located on the chlorophyll genome is preferable for forensic analysis compared to mitochondrial DNA due to the size and structure of mitochondrial DNA molecules varying widely even within individual plants (106).

2:6 Methodology

Preliminary work was carried out using soil samples from the TRACE project. TRACE, *tracing the origin of* food, was a five year project funded by the European Commission through the Sixth Framework Programme under the Food Quality and Safety Priority. The project developed generic and sector-specific traceability systems for use within the food industry, using a wide range of samples such as chicken, honey and olive oils as well as soils. The priorities of the project were to enhance consumer confidence in the authenticity of produce, and to assess the perceptions, attitudes and expectations of consumers regarding food production systems, food fraud, food authenticity and their ability to trace the products.

As preliminary DNA work was carried out prior to any soil sampling taking place soil samples were chosen from different TRACE sampling locations. The TRACE samples were readily available in excess in the laboratory. As honey samples contain suites of plant species they were also used in order to see if the methods used for honey DNA extraction at IrF Berlin could be extended to soils and to see if the sites the honeys were collected from could be distinguished from each other based upon their plant DNA.

2:6.1 Preliminary work

There are different methods which can be adapted to allow for the extraction of DNA from soil.

CTAB is a cationic detergent that solubilises membranes and forms a complex with DNA. CTAB is the most widely used method of plant DNA extraction. There are several variations of the CTAB method; in the simplest version ground plant tissue is incubated in hot CTAB buffer, this is then followed by chloroform-Iso amyl alcohol extraction and alcohol precipitation of the CTAB-DNA complex. After centrifugation, the resultant DNA pellet was washed, dried and re-dissolved in TE buffer. More complex methods of CTAB extraction include the addition of RNase to remove RNA and also steps to allow for the removal of any polysaccharides (107).

The Qiagen QIAamp DNA Stool Midi kit is a three step procedure which involves the lysis of bacterial cells and other pathogens present in the soil samples, the adsorption of impurities to an InhibitEX™ matrix (which is supplied in a tablet form in the Stool kit) and the purification of the DNA on QIAamp Midi Spin Columns. Once the impurities and DNA-degrading substances are adsorbed to the InhibitEX™ matrix, the matrix is pellleted by centrifugation and then the supernatant is purified (108).

The purification procedure was carried out on QIAamp Spin Columns and involves the digestion of proteins, the binding of DNA to the QIAamp silica membrane, the washing away of impurities and the elution of pure DNA from the spin column. Like the method used in the extraction of DNA from honey samples, the DNA was washed in two centrifugation steps and optimised buffer solutions are used in the washing process to allow complete removal of any residual impurities without affecting the binding of the DNA to the silica membrane (108, 109).

It should be noted that this was a user-developed protocol and therefore needed to be tested and then optimised to allow for the maximum amount of DNA to be extracted. The DNA which has been extracted was quantified using DyNAqaunt 200 Hoefer technology.

2:6.1 a Protocols

CTAB extraction of plant DNA

200 mg of plant tissue was ground to a fine paste in approximately 500 µl of CTAB buffer. Transfer CTAB/plant extract mixture to a microfuge tube. Incubate the CTAB/plant extract mixture for about 15 min at 55°C in a recirculating water bath. After incubation, spin the CTAB/plant extract mixture at 12000 g for 5 min to spin down cell debris. Transfer the supernatant to clean microfuge tubes. To each tube add 250 µl of Chloroform : Iso Amyl Alcohol (24:1) and mix the solution by inversion. After mixing, spin the tubes at 13000 rpm for 1 min. Transfer the upper aqueous phase only (contains the DNA) to a clean microfuge tube. To each tube add 50 µl of 7.5 M ammonium acetate followed by 500 µl of ice cold absolute ethanol. Invert the tubes slowly several times to precipitate the DNA. Generally the DNA can be seen to precipitate out of solution. Alternatively the tubes can be placed for 1 hr at -20°C after the addition of ethanol to precipitate the DNA. Following precipitation, the DNA can be pipetted off by slowly rotating/spinning a tip in the cold solution. The precipitated DNA sticks to the pipette and is visible as a clear thick precipitate. To wash the DNA, transfer the precipitate into a microfuge tube containing 500 µl of ice cold 70 % ethanol and slowly invert the tube. Repeat (alternatively the precipitate can be isolated by spinning the tube at 13000 rpm for a minute to form a pellet. Remove the supernatant and wash the DNA pellet by adding two changes of ice cold 70 % ethanol). After the wash, spin the DNA into a pellet by centrifuging at 13000 rpm for 1 min. Remove all the supernatant and allow the DNA pellet to dry (approximately 15 min). Do not allow the DNA to over dry or it will be hard to re-dissolve. Resuspend the DNA in sterile DNase free water (approximately 50-400 µl H_2O; the amount of water needed to dissolve the DNA can vary, depending on how much is isolated). RNaseA (10 µg/ml) can be added to the water prior to dissolving the DNA to remove any RNA in the preparation (10 µl RNaseA in 10ml H_2O). After resuspension, the DNA is incubated at 65°C for 20 min to destroy any DNases that may be present and store at 4°C. Agarose gel electrophoresis of the DNA will show the integrity of the DNA, while spectrophotometry will give an indication of the concentration and cleanliness.

Isolation of bacterial DNA from soil using the QIAamp® DNA Stool Mini Kit and QIAamp DNA Blood Midi Kit Protocol:

Weigh up to 5 g soil in a 50 ml BD Falcon™ tube. Add 2–5 ml distilled water to the tube, and mix for 5 min on a shaker. Incubate for 10 min at 95°C. Centrifuge at 3000 rpm for 5 min. Transfer the supernatant to a new tube. Add 7 volumes of Buffer ASL to the supernatant,

and mix well. Add 1 InhibitEX™ tablet to the tube and incubate for 1 min at room temperature (15–25°C) on a shaker. Centrifuge sample at 5000 x g for 5 min. Transfer the supernatant into a new tube. Add 1 volume of Buffer AL to the supernatant, and mix well. Add 1 volume of ethanol (96–100%). Place a QIAamp Midi Spin Column on the QIAvac 24 vacuum manifold. Apply the sample lysate onto the QIAamp Midi Spin Column. Apply maximum vacuum. Wash the column once with 1 ml Buffer AW1. Wash the column once with 1 ml Buffer AW2. Place the QIAamp Midi Spin Column in a 15 ml tube (provided), and centrifuge at 5000 rpm for 15 min to dry the membrane. Place the QIAamp Midi Spin Column in a clean 15 ml tube. To elute the DNA, add 300 µl Buffer AE, and centrifuge at 5000 rpm for 5 min. Reload the eluate onto the membrane of the QIAamp Midi Spin Column, and centrifuge at 5000 rpm for 5 min.

2:6.2 Analysis of Samples using Real-time polymerase chain reaction (real-time PCR)

Following extraction the samples were analysed using real-time PCR. With the use of real-time PCR it is possible to monitor the generation of PCR products as they are synthesized unlike traditional PCR where the products are monitored after 28 – 34 cycles. This enables you to see if a specific species is present in the sample. The TaqMan® real-time PCR system uses two primers and a probe. The probe is within the region defined by the primers and is labelled on the 5' end with a fluorescent molecule, and with a molecule that quenches the fluorescence on the 3' end (see figure 10). As the primers are extended by the *Taq* polymerase, one of them meets the probe, which is degraded by the polymerase (see figure 11), releasing the quencher and the probe into solution (see figure 12). For efficient quenching of the fluorescent molecules to occur they must be in close proximity on the probe molecule. As more PCR products are generated, more fluorescent molecules are released and the amount of fluorescence from the sample increases (81).

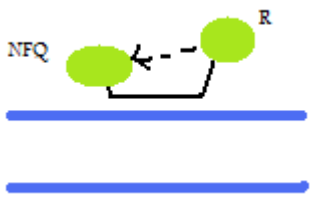


Figure 10: The TaqMan® quantification system consists of two PCR primers and an internal probe that hybridises within the region that is the amplified region (81).

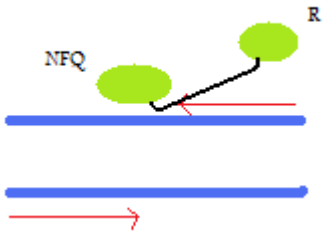


Figure 11: As the primer extends it encounters the probe, the 5' exonuclease activity of the *Taq* polymerase degrades the probe (81).

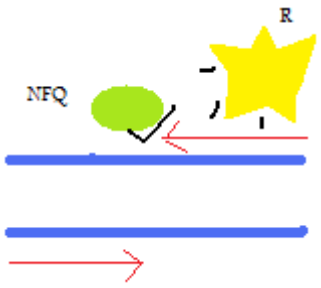


Figure 12: The reporter molecule is no longer in proximity to the quencher and fluoresces. (81)

How RT-PCR can be quantifiable is discussed in section 2.6.8.

2:6.3 Results and Discussion of preliminary work

Real-time PCR produces a CT value (threshold cycle) for the detection of the particular system being run. This CT value indicates the presence of DNA. However, the higher the CT value the more likely there is to be high levels of inhibitors in the PCR mixture which prevent the detection of DNA. Cts < 29 are strong positive reactions indicative of abundant target nucleic acid in the sample. Cts of 30-37 are positive reactions indicative of moderate amounts of target nucleic acid. Cts of 38-40 are weak reactions indicative of minimal amounts of target nucleic acid which could represent an infection state or environmental contamination.

The act system was run for each of the samples first as this is a generic system which detects plant DNA. If the CT value for this system was not too high (preferably below 37) then the plant specific species systems were run as shown in table 4.

Table 4: Plant species present in different honey samples analysed using real-time PCR (Applied Biosystems 7500 Real-Time PCR system)

	Act	Sweet Chestnut	Erica	Rape	Clover	Rockrose
1a_2.5ng_1:2	34.869186	37.434498	37.67544	Undetermined	Undetermined	Undetermined
1a_2.5ng_1:2	34.32173	37.61126	36.26221	Undetermined	Undetermined	Undetermined
1b_6.6ng_1:2	35.507687	37.863	36.56816	Undetermined	Undetermined	Undetermined
1b_6.6ng_1:2	35.32142	36.963165	38.373173	Undetermined	Undetermined	Undetermined
2a_5.5ng_1:2	35.53723	37.43835	37.092297	Undetermined	Undetermined	Undetermined
2a_5.5ng_1:2	35.95094	36.540833	35.530624	Undetermined	Undetermined	Undetermined
2b_6.4ng_1:10	39.09541	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
2b_6.4ng_1:10	37.497173	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
3a_3.4ng_1:5	36.116455	36.723076	35.852703	Undetermined	Undetermined	Undetermined
3a_3.4ng_1:5	37.65835	Undetermined	37.479572	Undetermined	Undetermined	Undetermined
4b_1.5ng_1:5	36.721855	Undetermined	Undetermined	36.85711	Undetermined	Undetermined
4b_1.5ng_1:5	40.003696	Undetermined	Undetermined	37.420387	Undetermined	Undetermined
11a_9.7ng_1:5	35.936787	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
11a_9.7ng_1:5	35.55404	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
11b_9.5ng_1:5	34.809288	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
11b_9.5ng_1:5	35.677338	38.67653	Undetermined	Undetermined	Undetermined	Undetermined
Control	25.773949	26.6517	24.86847	25.309294	30.885826	28.454664
Water	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined

	Sunflower	Oak	Citrus	Linden	Olive
1a_2.5ng_1:2	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
1a_2.5ng_1:2	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
1b_6.6ng_1:2	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
1b_6.6ng_1:2	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
2a_5.5ng_1:2	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
2a_5.5ng_1:2	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
2b_6.4ng_1:10	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
2b_6.4ng_1:10	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
3a_3.4ng_1:5	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
3a_3.4ng_1:5	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
4b_1.5ng_1:5	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
4b_1.5ng_1:5	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
11a_9.7ng_1:5	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
11a_9.7ng_1:5	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
11b_9.5ng_1:5	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
11b_9.5ng_1:5	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined
Control	28.340631	26.427195	23.344416	26.812492	29.35878
Water	Undetermined	Undetermined	Undetermined	Undetermined	Undetermined

Using real-time PCR, it was determined that the same three samples (one, two and three, all from Galicia) were positive for sweet chestnut DNA and heather (Erica) DNA. This conforms to what was written on the labels of the honey which stated the honey was from heather flowers. A different sample (sample four from Cornwall) only tested positive for the presence of rape DNA, as seen in table one. It cannot be determined for certain whether sample 11 from Barcelona is either positive or negative for sweet chestnut DNA. It should be noted that the amounts of DNA were not quantified. No samples tested positive for the presence of clover, rockrose, sunflower, oak, citrus, linden or olive. Probes and primers are also available to test for the presence of acacia, broom, eucalyptus, lavender, rosemary, and maize. These results show the potential for the use of plant specific species being used as a basis of differentiating between honey samples taken from different geographical locations. For example, using the above data, if a honey sample contained rape DNA it may be an indication that that sample was taken from Cornwall; however it may also indicate the sample was taken from somewhere else with rape crops.

Although the DNA extraction and amplification from honey wasn't as successful as had first been hoped this may be due to laboratory procedure. Eppendorf reference (adjustable-volume) pipettes were used to make the dilutions for real-time PCR, the master mix for the real-time PCR reaction and also to pipette the solutions into the 96-well plate ready to be put into the real-time PCR instrument. When the volumes from the pipette were weighed to check for accuracy the weights were slightly different each time; this can account for some of the inconsistencies within the results; especially considering such small volumes of solutions were being pipetted. This could be eliminated by using the Eppendorf research pro pipettes which are electronic and therefore more accurate. The importance of pipetting skills and the need for this to be accurate is highlighted in work by Miller-Coyle (58).

Although large amounts of DNA was successfully extracted from the soil using the CTAB method (see appendix 2), once these samples were analysed using real-time PCR it became apparent that only a small amount of this DNA was from plant species. As the CT value for Act (the generic plant system) was very high (over 40) or undetermined, which meant that any DNA from the specific systems would be undetectable. As a result of this it cannot be determined which plant species the DNA comes from. This suggests that bacterial DNA is present in the soil samples in relatively large amounts. It also has to be considered that no

RNase was used in the CTAB extraction and the DNA was measured using DyNAQuant which measures both DNA and RNA and this could account for the high concentrations.

A Qiagen Stool Kit was also used to extract DNA from the soil samples. This yielded extremely low concentration of DNA, which when analysed using real-time PCR were undetermined for the generic plant system, again indicating that the majority of the DNA present in the soil was bacterial DNA.

2:6.4 DNA Sequences from Species of Interest

Following the preliminary work using RT-PCR, it was decided to try and analyse the plant DNA in soil using the *mat*K gene (see figure 13). The *mat*K gene is located in the chloroplast genome and is therefore more conserved than mitochondrial DNA. The gene is an approximately 1.5 kb protein-coding region located between two highly conserved exons of the *trn*K gene. The *mat*K gene has both conserved and variable segments which aid in DNA amplification and resolution of relationships (110).

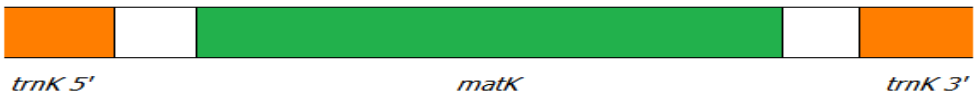


Figure 13: The *mat*K gene: approximately 1.5 kb protein-coding region between two highly conserved exons of the *trn*K gene (110).

The *mat*K gene was chosen because of its location within the chloroplast genome; it is preferable to a mitochondrial gene as these are thought to lack promise for plants due to their low substitution rates (111).

2:6.4a Primer Design

Using Genbank DNA sequences for the *mat*K gene (and if this could not be found the *trn*K gene) were found for the following plant species: wheat, sycamore, pine, maize, horse chestnut, common reed, birch, beech, barley, oak, oilseed rape and sugar beet. The plant species were chosen based upon the likelihood of them occurring throughout Norfolk so that it was possible that the DNA of the species would be present in the soil samples. Once

the *mat*K sequences had been found, unique forward and reverse primers were designed for each of the plant species using Primer 3 technology (table 5).

Table 5: Oligonucleotides for analysis of the *mat*K gene in soil; the sequences are shown and the weight of the oligonucleotides and the amount of water to be added to them.

Oligo Name	Direction	Sequence	Weight / µg	Water added / µl
Wheat	Forward	AACCTGTGGAAATAGTTGTTAGTTG	425	548
	Reverse	TCGAAATAGTCTTTTATTTTCTTTTTG	478	581
Sycamore	Forward	CCAGGGGGCGACTAAATACT	304	494
	Reverse	TGCACACAGCTTTCCCTATG	363	602
Pine	Forward	TGTCGTCCCAAGAGTAAATACG	274	406
	Reverse	TCTGTTAGACTGGTCCAAGACAAT	371	504
Maize	Forward	GTACGAGCCCCATACAGAGC	318	523
	Reverse	TCTTTGTCCCGCATCCTTAC	340	569
Horse Chestnut	Forward	GAGGGGTTTGCAGTCATTGT	342	550
	Reverse	TGGAAGGGTTGTCTCGAACT	356	576
Common Reed	Forward	TCCGAGGAGGAAGTCCACTA	329	535
	Reverse	CTGGAACCTTTCTGGAACGA	326	533
Birch	Forward	TGCGGTTCTTTCTTCATGAGT	406	634
	Reverse	ACCCAAAAAGTCGAGGGAAT	318	515
Beech	Forward	AGCAGCATGTCGTATCAACG	342	558
	Reverse	TGCCCCTTCGAAAGAAGATA	285	466
Barley	Forward	TTGCATTTATTGCGATTCTTTC	352	527
	Reverse	ATCCGACCAAATCGATCAAG	395	650
Oak	Forward	AATACCCTACCCTGCCCATC	367	618
	Reverse	CGCGTGCAGTACTTTTGTGT	389	635
Oilseed Rape	Forward	AAACCGCCCTATTTTCTTGG	401	663
	Reverse	CCTTTTGCGATTGAAACCAT	369	609
Sugar beet (1)	Forward	GGAAAATTGGGTGACAAGA	316	507
	Reverse	CAAGGATCCAACCAGAGGAA	280	455
Sugar beet (2)	Forward	CAAGGATCCAACCAGAGGAA	448	728
	Reverse	GGAAAATTCGGGTGACAAGA	340	546

There are two different sets of sequences for Sugar beet due to a great deal of variance in the *mat*K and *trn*K sequence lists for this species on GenBank. Therefore it was decided to use two different sets of sequences for the sugar beet to optimise the chances of amplifying any sugar beet DNA found in the soil and as a result see which sequence is the correct one for the species.

2:6.5 DNA Extraction from Soil

DNA was extracted from the Norfolk soil samples using the PowerSoil™ DNA Isolation Kit from MoBio Laboratories, Inc. (112) The kit is specially designed specifically for research use and is able to isolate genomic DNA from environmental samples using a novel method. The kit is designed for use with samples that contain a high level of humic acid, including difficult samples such as sediment and manure, but is also effective for more common soil types. The kit is different from other commercially available kits in that it contains a humic substance / brown colour removal step. This method has proven to be effective at removing PCR inhibitors from even the most difficult soil types.

The soil samples are added to a bead beating tube, which allows rapid and thorough homogenization. Cell lysis occurs by both chemical and mechanical methods. Total genomic DNA is captured on a silica membrane within a spin column. The DNA is then washed and eluted from the membrane. The isolated DNA has a high level of purity, which allows for more successful downstream applications such as PCR amplification.

2:6.5a Protocol:

0.25 g of soil was added to the provided PowerBead Tubes. The PowerBead Tube contained a buffer that helped to disperse the soil particles, began to dissolve the humic acid and also protected the nucleic acids from degradation. The sample was then homogenized and lysed by gently vortexing to mix the sample. Vortexing mixed the components in the PowerBead Tube and began to disperse the sample into the PowerBead solution.

60 µl of solution C1 was then added to the sample and inverted several times or vortexed briefly. Solution C1 contained SDS (sodium dodecyl sulphate) and other disruption agents which were required for complete cell lysis. SDS is an anionic detergent that breaks down fatty acids and lipids associated with the cell membrane of several organisms. If it gets cold it will form a white precipitate in the bottle. If this was the case, heating the solution to 60 °C dissolved the SDS and did not harm it or any of the other disruption agents. Solution C1 could be used while it was still warm. After C1 had been added the PowerBead Tubes were secured on a flat-bed vortex mixer with tape and vortexed at maximum speed for 10 minutes. The vortexing step was critical for complete homogenization and cell lysis. Cell

lysis occurred through a combination of chemical agents in the PowerBead tubes and C1, and at this step mechanical shaking was introduced. Through the random shaking of the beads in the presence of disruption agents, the beads collided with the cells, which caused the cells to break open. It was then ensured that the PowerBead tubes rotated freely without rubbing and then centrifuged at 10,000 x g for 30 seconds at room temperature. The speed did not exceed 10,000 x g or the tubes were in danger of breaking.

The supernatant was then transferred to a clean 2 ml collection tube. At this step between 400 and 500 µl of supernatant was expected. The exact amount of supernatant depended upon the absorbency of the starting material, although this was not critical for the procedure to be effective. At this point the supernatant may be dark in colour and still contain some soil particles; this could be expected for many different soil types. The coloration of the supernatant and the carry-over of soil were removed by subsequent steps in the protocol. 250 µl of solution C2 was added and vortexed for 5 seconds, followed by incubation at 4 °C for 5 minutes (tubes were placed into an ice bath and then into a refrigerator) and centrifuged at room temperature at 10,000 x g for one minute. Solution C2 contained a reagent to precipitate non-DNA organic and inorganic material including humic acid, cell debris and proteins. The removal of these contaminants was important because they may have reduced DNA purity and inhibited any downstream applications.

Avoiding the pellet, up to 600 µl of supernatant was transferred to a clean 2 ml collection tube. For the best results the pellet was avoided as at this point it contained non-DNA organic and inorganic material, which included humic acid, cell debris and proteins. 200 µl of solution C3 was added and vortexed briefly and incubated at 4 °C for 5 minutes. The tubes were centrifuged at room temperature at 10,000 x g for one minute. Solution C3 was a second reagent to precipitate any additional non-DNA organic and inorganic material, which included humic acid, cell debris and proteins.

Up to 750 µl of supernatant was transferred to a clean 2ml collection tube. For the best results the pellet was avoided as again, at this point it contained non-DNA organic and inorganic material. 1.2 mL of solution C4 was added to the supernatant and vortexed for 5 seconds. Solution C4 was a high concentration salt solution. DNA binds tightly to silica at high salt concentrations, therefore solution C4 adjusted the DNA salt concentrations to allow the binding of DNA to the Spin Filters, but not any non-DNA organic and inorganic material that may still have been present at very low levels.

Approximately 675 µl was loaded onto a Spin Filter and centrifuged at 10,000 x g for one minute at room temperature. Following centrifugation the flow through was discarded and an additional 675 µl of the supernatant loaded onto the spin filter and centrifuged. The remaining supernatant was then loaded onto the Spin Filter and centrifuged. DNA was selectively bound to the silica membrane in the Spin Filter device in the high salt solution. Any remaining contaminants passed through the filter membrane, leaving only the DNA bound to the membrane.

500 µl of Solution C5 was added to the Spin Filter and centrifuged at 10,000 x g for 30 seconds at room temperature. Solution C5 was an ethanol based wash solution; this was used to further clean the DNA that had bound to the silica membrane in the Spin Filter. The wash solutions removed the residual salts, humic acid and any other contaminants while allowing the DNA to remain bound to the membrane. The flow through was discarded and the spin column centrifuged at 10,000 x g for one minute at room temperature to remove any residual C5 solution. It was essential to remove any residual traces of the C5 solution as it contained ethanol which may have interfered with any downstream applications such as PCR.

The Spin Filter was then carefully placed into a clean 2 mL collection tube, avoiding splashing any of the C5 solution onto the filter. 100 µL of solution C6 was added to the centre of the white filter membrane and centrifuged at 10,000 x g for 30 seconds at room temperature. The C6 solution was placed directly onto the centre of the white filter membrane to ensure that the entire membrane was wetted, resulting in a more efficient and complete release of the DNA from the Spin Filter membrane. As the C6 solution, an elution buffer, passed through the silica membrane, the DNA that was bound in the presence of high salt was selectively released by the C6 solution (10 mM Tris) which lacked salt. The Spin Filter was discarded. The DNA is now ready for any downstream applications. The DNA was stored frozen (-20 °C to -80 °C).

2:6.6 DNA Quantification

The DNA samples were quantified using Hoefer DynaQuant 200. The DynaQuant is a fluorometer designed for the accurate quantification of small volume samples with low DNA concentrations using Hoechst 93528 dye. Dye binds preferentially to double stranded

DNA thus allowing accurate quantification even in the presence of RNA, proteins and nucleotides.

An assay solution was prepared by mixing 10.0 µL stock solution with 10.0 mL TNE buffer and 90 mL distilled water. The DynaQuant was calibrated using 2 µL of a standard DNA solution and 2 mL of assay placed into a cuvette and measured. As the expected concentrations from the DNA soil samples were low the standard DNA solution was diluted in a 1:10 ratio with TNE buffer and distilled water. 2 µL of the DNA extracts were placed into a cuvette with 2 mL of assay solution and measured.

Table 6: DNA concentrations of soil samples extracted using the MoBio PowerSoil DNA extraction kit.

Sample	**DNA Concentration / ng/µL**
160708_04_01	6.8
160708_04_02	4.5
160708_04_03	5.2
160708_04_04	11.2
240708_06_01	3.4
240708_06_02	4.8
240708_06_03	5.8
240708_06_04	5.5

2:6.7 DNA Amplification

Each sample was prepared for PCR reaction for each of the species. For each sample 1 µL 10X PCR buffer, 1 µL dNTP to give a final concentration of 0.2 mM, 1 µL forward primer for the species of interest, 1 µL reverse primer of the species of interest, 1 µL of the DNA extract, 0,1 µL TAQ DNA polymerase and 4.9 µL of MilliQ water were added to the PCR

plate. The PCR reaction cycle was 10 minutes at 96 ◦C to begin with, 30 cycles of 30 seconds at 96◦C, 30 seconds at 58◦C, 45 seconds at 72◦C and 10 minutes at 72◦C to end.

Following PCR, the products were visualised on an agarose gel (see figure 14). From this image it can be seen that there is DNA present at E1 and E8; the presence of wheat and barley in sample 240708-06-01 (Roydon Country Park). The other bands that can be seen are primer dimers and no more action was taken with them.

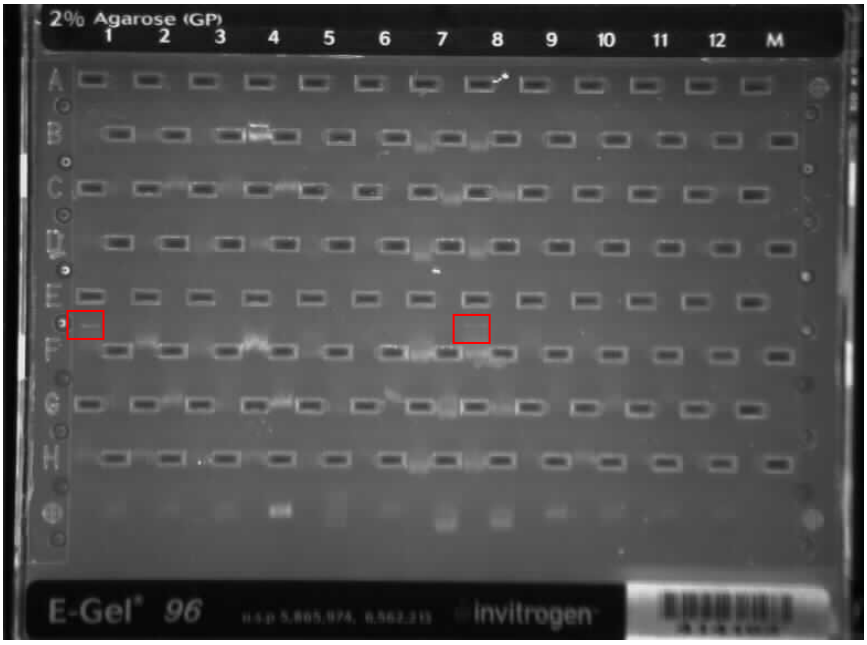


Figure 14: PCR products from samples taken from locations 160708-04 and 240708-06 on an agarose gel.

2:6.8 Quantitative PCR

Quantitative PCR or RT-PCR uses DNA amplification to determine absolute or relative amounts of a known substance of a known sequence in a sample. By using a fluorescent reporter in the reaction it is possible to measure the DNA generation. This is done in the same way as real-time PCR was used to analyse the TRACE samples. No data was able to be obtained for quantitative PCR was obtained due to the decision to send the samples straight for sequencing for method development purposes; however if if was to be carried out it would be done prior to the sequencing taking place.

2:6.9 DNA Sequencing

The bands, which were cut out of the gel (figure 14), were sent for sequencing. Prior to DNA sequencing, the DNA samples were cleaned up by heating to 37 ◦C for 45 minutes followed by heating to 85 ◦C for a further 15 minutes. The sequencing reaction was 2.5 µL of master mix, 1 µL of species specific oligonucleotide, 5 µL DNA extract and 1.5 MilliQ water. The aim of this sequencing was to see if the presence of wheat and barley DNA could be determined as these were the bands present in the original PCR reaction (see figure 14). When the DNA sequences were run through BLAST (Basic Local Alignment Search *Tool)* there was a match with wheat DNA but not barley. It was therefore decided to optimise the PCR reaction to see if this would yield more successful DNA sequencing.

2:6.10 Optimizing the PCR Reaction

Optimizing a PCR reaction is done in order to maximise both specificity and sensitivity of the reaction. This can be done by altering several different parameters involved in the PCR reaction; the primer annealing temperature, the DNA polymerase, the magnesium concentration and the cycle parameters.

By raising the annealing temperature the specificity of the primer annealing is increased by destabilising base pair mismatches. The sensitivity (and yield) of the PCR can be increased by lowering the annealing temperature due to the stabilising of the correct base pairing. By increasing the concentration of the DNA polymerase insufficient product may be made, decreasing the concentration of the polymerase may result in a decrease in reaction specificity. *Taq* polymerase is the most efficient enzyme that can be used but also has the highest error rate; the *Pfu* enzyme has a decreased error rate but on the downside synthesizes the least amount of product. If the magnesium concentration used in the PCR is varied, a higher level of magnesium chloride can stabilize primer annealing and increase sensitivity but can also decrease primer specificity; however a lower level of magnesium chloride can increase the specificity. If the denaturing temperature of the PCR process is elevated it can result in increased sensitivity by allowing complete temperature denaturing especially of G and C rich targets. However, *Taq* polymerase activity decreases rapidly at temperatures above 93°C. Increasing the primer extension times can be beneficial in increasing the sensitivity in long-distance PCR; a method used to generate PCR products of

2-20kb. The number of PCR cycles can also be altered; if the reaction contains less than 10^3 initial target molecules the cycle number can be increased beyond 35. In contrast the number of cycles can be decreased to make the process more quantitative (exponential phase) (58, 113).

It was decided to try to optimise the PCR reaction by altering the temperature gradient. The chosen temperatures were 45°C to 50°C (step two of the cycle, 40 cycles for each step). For each sample 12.5 µL 10X PCR master mix, 1 µL forward primer for the species of interest, 1 µL reverse primer of the species of interest, 1 µL of the DNA extract, 0,1 µL TAQ DNA polymerase and 9.5 µL of MilliQ water were added to the PCR plate. The DNA extracts from table 4 were pooled to try to optimise the DNA content in the samples. Following PCR the DNA was visualised on a gel (figure 15).

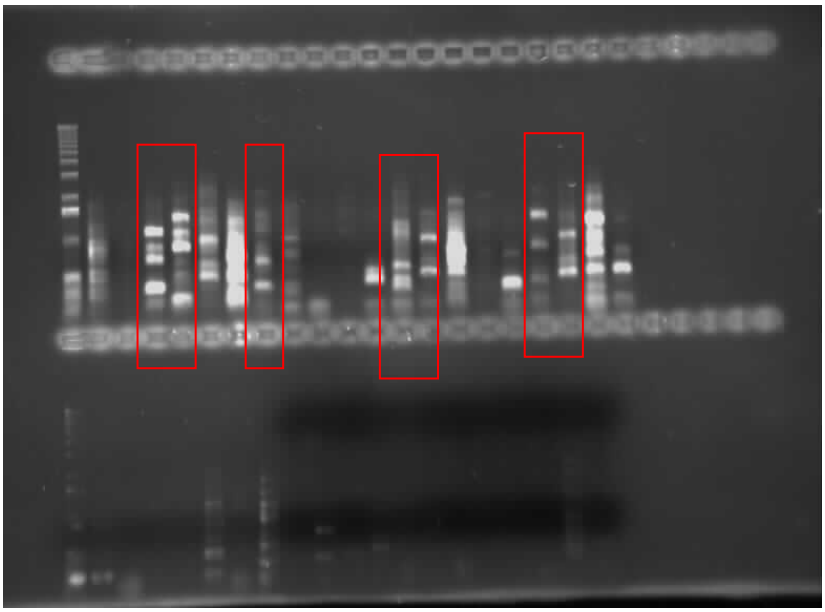


Figure 15: Gel image of the PCR reaction using a temperature gradient of 45°C to 50°C of Norfolk soil samples. The red boxes highlight the bands which are to be extracted and re-run on a gel (figure 16).

Following this reaction using a Qiagen Gel Extraction Kit, the bands highlighted in figure 15 were cut out, extracted and re-run on a gel (figure 16).

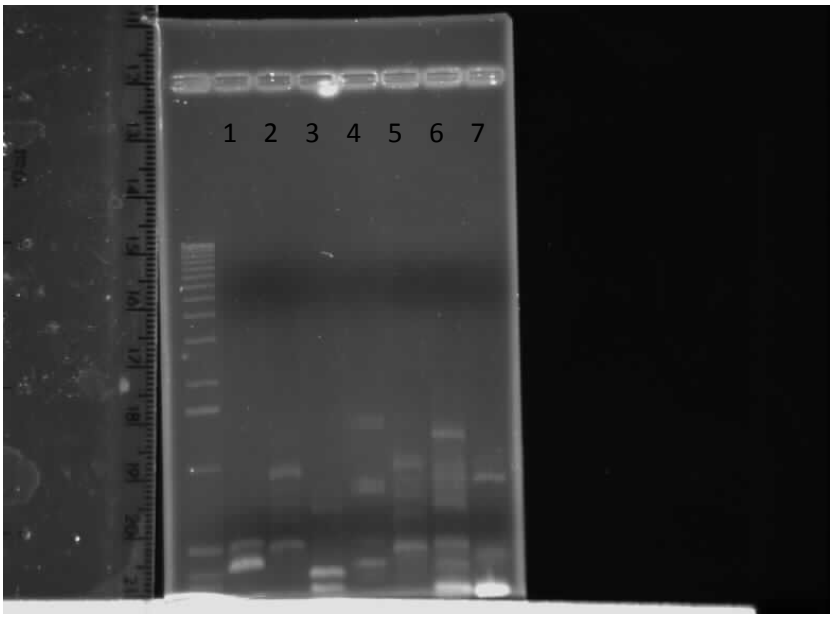


Figure 16: Gel image of the bands cut out from figure 7 and ran on a gel. For each lane see figures 17 to 23)

After viewing the gel, each lane was cut from the gel and viewed next to a ruler in order to determine the presence of the specific species (see figures 17 to 23 with the species found to be present in brackets so they could be sequenced)

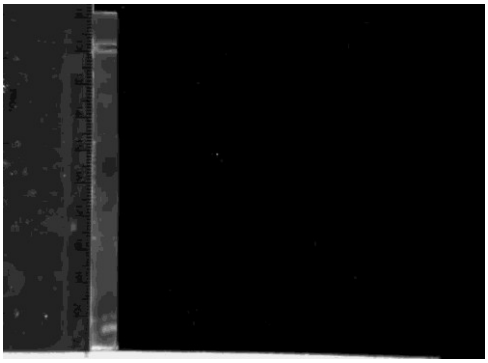


Figure 17: Gel of Lane 1 (sycamore)

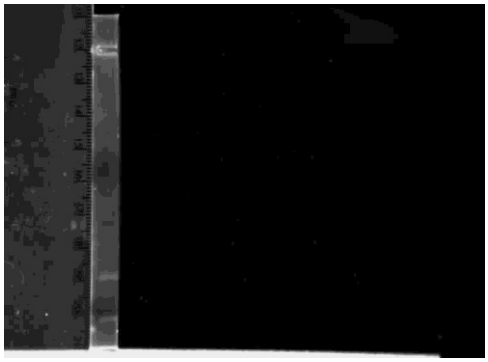


Figure 18: Gel of Lane 2 (Sugar beet)

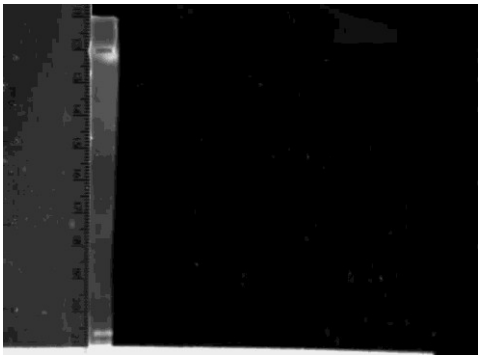


Figure 19: Gel of Lane 3 (Beech)

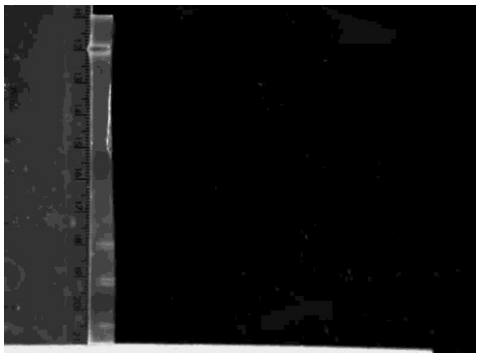


Figure 20: Gel of Lane 4 (Oil Seed Rape)

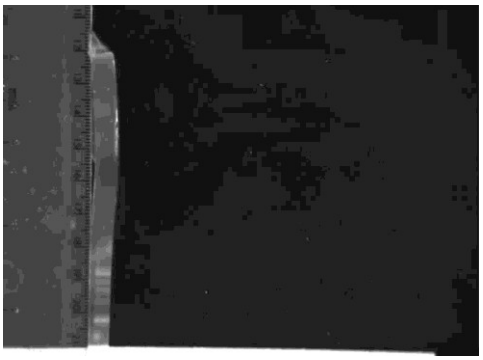


Figure 21: Gel of Lane 5 (Horse Chestnut)

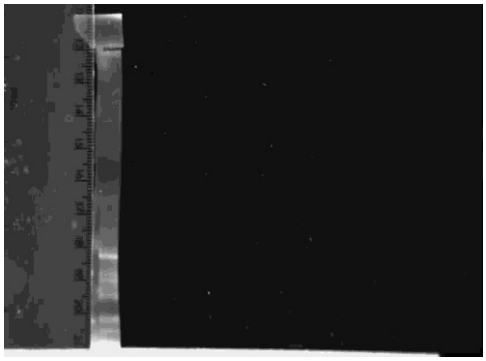


Figure 22: Gel of Lane 6 (Sugar Beet)

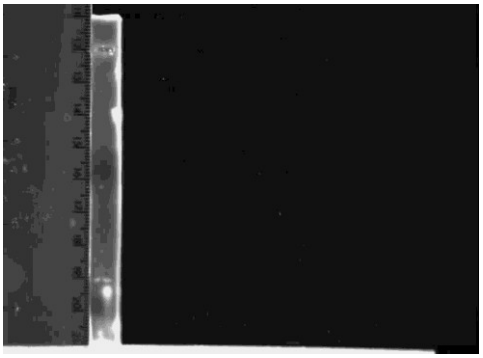


Figure 23: Gel of Lane 7 (Birch)

2:6.11 Further Sequencing

The sequencing reaction was 2.5 µL of master mix, 1 µL of species specific oligonucleotide, 5 µL DNA extract and 1.5 µL MilliQ water.

As the sequencing for the *mak*K gene yielded no viable results for the mixture of plant DNA present in soil, it was decided to try to utilise MALDI-ToF-MS to see if this was a suitable technique (see chapter 3).

2:7 Discussion

2:7.1 Successes of the use of plant DNA in Norfolk soil samples

Positives of RT-PCR

As discussed in section 2:6.3 using the honey samples from the TRACE project for method development, it was possible to identify the origin of some of the samples using RT-PCR. However, it must be remembered that many of the plant species present in the honeys were known as they were stated on the sample labels and were therefore tested for. The Act system of testing for generic plant material potentially could have been useful for soils in giving an idea as to whether there was sufficient plant DNA available for further testing. However it must be remembered that the system is generic for the 10 plant species shown in table 10; and as it is not known if these plant species are present in any of the collected soil samples the potential of the act system is limited. Another positive of RT-PCR is that is quantifiable.

The use of the *mat*K gene could prove to be useful in the analysis of plant DNA in Norfolk soil samples for forensic purposes if the plant species present in the soil sample are known. As can be seen from the sequencing of the wheat it was found that this species was present in the soil sample following running the sequence through BLAST; however as wheat is a common plant throughout Norfolk this does not yield a great deal of useful knowledge in a forensic context. The potential for the use of the *mat*K gene for the analysis of plant DNA was shown in a study by Fazekas *et al* (111) wherein it was stated that using *mat*K it was possible to align sequences of taxa across seed plants but not outside them.

One of the successes with the MoBio Soil Master DNA extraction kit is that it contains a humic acid removal step; this is important as humic acid is an inhibitor of any downstream applications (104).

If a soil sample was thought to come from a specific location within Norfolk primers could be designed for species using the species maps of Norfolk, and then sequencing could be used to see if they are there. However there is lots of bias based on the maps. Also, this would involve the use of many different primers and would not look at suites of species. This highlights the need to use different techniques in tandem, in this case using the DNA sequences alongside the plant species data to give more information.

2:7.2 Problems of the use of plant DNA in Norfolk soil samples

Negatives of RT-PCR

One of the greatest drawbacks to the use of plant DNA in soil for forensic purposes would be the need to design primers for multiple species. This would be very expensive and not a viable method and prior knowledge about the species would need to be available. It would not be possible to analyse a sample for all of the different species especially when a forensic sample is often a small amount. Another issue surrounding primers is about their size; DNA fragments which are too short do not amplify but longer ones will, however this is dependent upon the primer size which is limiting.

Real-time PCR is not good for mixtures of DNA, this links back to extraction and the fact that there is not specific kit or method for the extraction of only plant DNA from soil.

The use of the *mat*K gene, although it has the potential of being useful if the plant species present in the soil sample are known, is very limited in an unknown soil sample such as those used in this project. It is possible that in order for the *mat*K gene to be used successfully, multiple primer sets may be needed for amplification as stated by Fazekas *et al.* (111) and therefore the portion of the region that was sequenced varies among taxonomic groups depending upon the primer sets which were used.

There are four factors that affect the ability to obtain a DNA profile; all of which could have an impact on the inability to sequence the *mat*K gene for the various species from the DNA extracted from the soil samples. The factors are sample quantity, sample degradation, sample purity and the ratio of major to minor contributors in the sample (114). The quality of the DNA in the extracts from the soil samples may have been poor, and also could have been degraded as there is no way of knowing how long the DNA has been present in the

soil, and if it had been there for a long period of time and also been exposed to the elements it could be in various states of decay. The samples may not have been very pure due to the unknown and potential mixed nature of the samples. One major contributing factor appears to be the fact that the plant DNA seems to be a minor contributor to the DNA extract and other sources such as bacterial DNA may be the major contributors. There are various factors within the soil chemistry, which will all impact the DNA profile including pH and alkalinants.

As can be seen from the three different extraction techniques tried (see sections 2:6.1 and 2:6.5) not one of them has been entirely successful. This could be due to several reasons; we know the extraction kit contained a humic acid removal step so therefore humic acid is not likely to be the problem in the PCR reaction and sequencing. Other factors that inhibit the success of the extraction process are inhibitors which lead to low quality DNA extracts, and also the problem of there being no universal method for the extraction of plant DNA from soil, all soils are different (in both properties and the presence of DNA in them) and therefore one protocol will not necessarily work for all soil samples (104).

Problems with DNA extraction have also been encountered when extracting DNA from wood. A study by Rachmayanti *et al.* (115) states numerous factors can affect the success of DNA extraction from wood, the success of which was measured by means of PCR) including the fragment size (shorted fragments amplified more successfully), the state of the wood (unprocessed wood amplified more successfully) and the age of the wood species. Another problem was the number of inhibitors in the wood, which was helped by the addition of polyvinylpyrolidone (PVP) to the lysis buffer during the extraction process. Although Rachmayanti *et al*. work is specific to wood DNA, many of the findings could be applied to this study in that the fragment size of the DNA extracted from the soil may have been longer and therefore not as good at amplifying. Many of the collected soil samples also contained wood fragments (particularly those collected from forest locations) which could have been of vastly differing ages and in various states of decay and although the soil samples were sieved and < 2mm fraction used for analysis some wood fragments may have been in the sample and therefore added to the complications in the DNA extraction process.

One potential way to overcome the problems of the inhibitors present in the DNA extracts from the soil samples is to carry out a purification of the extracts prior to analysis (104).

Although this may prove to be useful; it should be carried out with caution and taking into consideration that the exact components of the DNA extract are unknown and therefore the purification may be unsuccessful.

If a suitable extraction technique could be developed which only extracted plant DNA, or if possible even more specifically chloroplast DNA from the soil, the use of the *mat*K gene could indeed be a viable means of forensic analysis of the plant DNA in soil.

As has been previously discussed one of the main problems appears to be the unknown nature of the DNA extracts; the exact composition of the DNA is unknown and also the age and state of the DNA. This is problematic as it has been shown that young plants and bushes yield the best DNA (116). It may have been advantageous to "spike" some soil samples with a known DNA sample to test the extraction and purification methods.

Due to the problems gaining suitable data for forensic purposes using more traditional methods of DNA analysis it was decided to try to run the DNA extracts on a MALDI-ToF-MS to see if this would gain any usable data (see chapter 3).

2:7.3 Improvements to the method

The *trnL* intron and the *trnL-trnF* intergenic spacer (IGS) have the potential to be used for the analysis of plat DNA present in soil. This gene has been used in a study by Tsai *et al.* (76) and was used to identify samples to the species level. However, in this study it was leaf material that was analysed; this was extracted using a Plant Genomic DNA Miniprep System commercial kit available in Taiwan. As this seems to be a successful way of analysing plant DNA, both the kit and the use of the plant leaf material are things which may be considered in the analysis of plant DNA present in soils.

A more recent study by Wallinger et al, has used the trnT-F gene as a means of rapid plant identification using species and group specific primers targeting chloroplast DNA. Using this gene would mean that a more specific extraction kit could be used for the chloroplast DNA, thus reducing some of the aforementioned problems.

Plant DNA from Soils Summary Box

Unable to successfully utilise plant DNA for the forensic analysis of soils by looking at suites of species.

No suitable extraction method for only plant DNA in soil is currently available.

Specific species were able to be amplified but this depends upon prior knowledge of species likely to be present in the samples and is expensive due to multiple primers needed.

Plant DNA in soil may be present in too low quantities to be forensically viable option of analysis.

CHAPTER 3 – MALDI-ToF-MS of plant DNA from soil

3.1 MALDI introduction

Matrix Assisted Laser Desorption/Ionisation Time of Flight Mass Spectrometry (MALDI-ToF-MS) was developed in the late 1980s by Kras, Hillenkamp *et al.* (117) from similar desorption/ionisation mass spectrometric methods such as Fast Atom Bombardment (FAB) and Laser Desorption Mass Spectrometry (LDMS). At the same time, Tanka *et al.* developed a new related technique whereby an analyte was mixed with a finely ground powder (118, 119) and thus MALDI was developed. In this research MALDI has been used to try to analyse the DNA present in the soil. The decision to implement the use of MALDI was made after more traditional analysis of the plant DNA in the soil using the *mat*K gene was unsuccessful, as described in chapter 3.

3.1.1 Principles of MALDI ToF MS

Matrix-Assisted Laser Desorption/Ionisation (MALDI) is used to get the sample, in this case the DNA extract, into the gas phase so it is suitable for analysis in the MS. Time of Flight Mass Spectrometry (ToF-MS) is a technique that is capable of processing large numbers of samples and has been shown to work with DNA analysis techniques such as STR analysis (70). The analysis of a sample by MALDI can be divided into two steps; firstly the preparation of the sample by mixing the analyte with a molar excess of matrix and secondly, desorption of bulk portions of the solid sample by a short pulse of laser light (120).

In principle, MALDI is a special case of LDI in which a particular sample preparation is used. In the crudest form of MALDI-MS the sample consists of dilute analyte molecules embedded in a matrix of highly light-absorbing, low-mass molecules. The matrix molecules are commonly resonantly excited by a UV, or less commonly by an IR, laser pulse of typically nanosecond duration. The absorbed energy causes an explosive break-up of the sample and ionization of a fraction of the analyte molecules. Subsequently, a volume of the matrix and the trapped analyte molecules are ejected into the gas phase. This ejected material contains both charged and neutral species, which interact with one another during

the expansion of the plume in the ion source as seen in figure 24. A standard MALDI ion source is usually the same as an LDI source except for the matrix preparation, and usually operated under high vacuum conditions, typically a pressure of less than 10^{-6} Torr, and is combined with a ToF mass analyser with axial extraction (119).

Using MALDI-ToF-MS the samples are analysed sequentially by moving the sample plate underneath a fixed laser beam. The sample plate can hold up to 384 samples at any one time (though this is plate dependent) and can be analysed in less than one hour depending upon the number of laser shots collected for each sample and the pulse rate of the laser (121). For a more in-depth discussion of the MALDI process see sections 3:1.1a to f.

(1) $M + e^- \rightarrow M^+ + 2e^-$ (122)

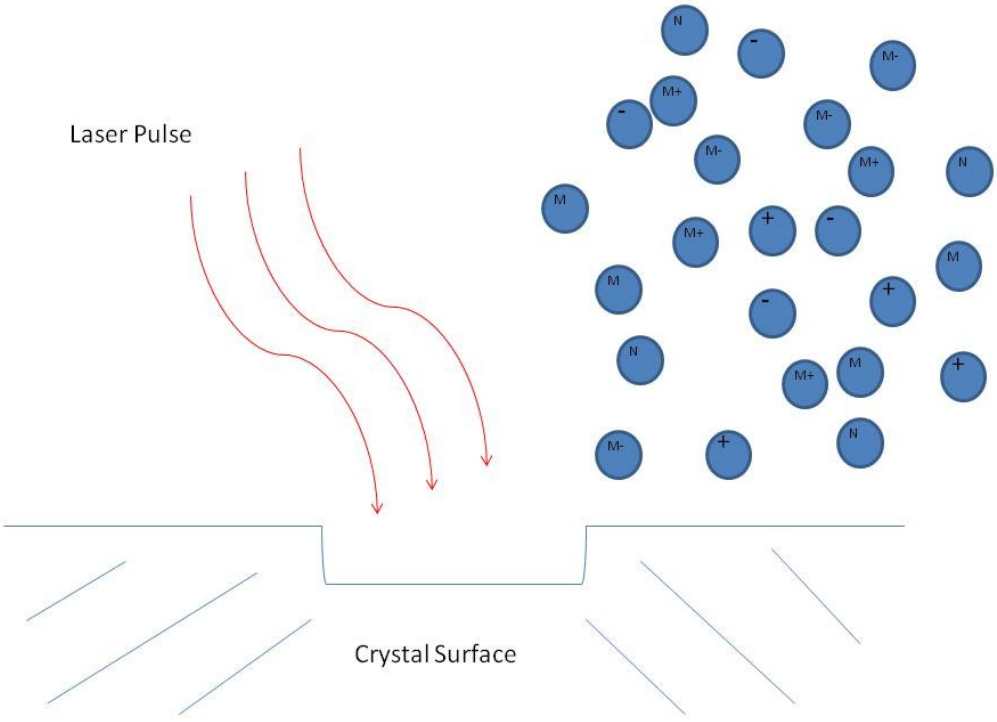


Figure 24: Principles of MALDI: As a result of desorption neutral matrix molecules (M), positive matrix molecules (M+), negative matrix molecules (M-), neutral analyte molecules (N), positive analyte molecules (+) and negative analyte molecules (-) are created and / or are transferred to the gas phase. (Adapted from (123))

MALDI can operate in both positive and negative ion mode, and both generate singly charged ions (this is somewhat dependent upon the matrix), but multiply charged ions are

common (especially for samples with a high mass) and dimers also occur. Due to the soft nature of MALDI it is possible to desorb, ionize and detect large, intact proteins (119). Typicaly, if the sample has functional groups that readily accept a proton (H^+) the positive ion detection is used, for example, for the analysis of amines ($R\text{-}NH_2 + H^+ = R\text{-}NH_3^+$), such as in proteins or peptides. If the sample has functional groups that readily lose a proton then negative ion detection is used. For example, for the analysis of carboxylic acids ($R\text{-}CO_2H = R\text{-}CO_2^-$) or alcohols ($R\text{-}OH = R\text{-}O^-$) such as in saccharides and oligonucleotides (124).

Typically a MALDI analysis will follow the following steps:

3:1.1a Analyte incorporation

The first step of a MALDI analysis is to incorporate the sample within a molar excess of matrix. See section 3.1.4 matrices.

3:1.1b Absorption of the laser radiation

The role of the optical absorption of the matrix in the transfer of energy from the laser beam to the sample is governed by Beer's law:

(2) $H = H_o \times e^{-\alpha z}$

Where H is the laser fluence at a depth z into the sample, H_o is the laser fluence at the sample surface and α is the absorption coefficient. The absorption coefficient is the product of the molar absorption coefficient α_n and the concentration C_n of the absorbing molecules in the sample. α_n is a wavelength dependent molar absorption coefficient which is a property of the matrix compound.

Due to the very shallow ablation depth, a given location of the sample can usually be irradiated many times before the material is exhausted. In the context of this study, this is important so that multiple measurements can be taken of the same sample to see if there is any variation of the sample throughout the spot on the plate and also to see if multiple spectra taken from the same position yield the same results (see experimental section) something important if this technique is to be useful for the analysis of the DNA present in

the soil samples. Thus, these experiments will help to see if this is a technique viable for the forensic provenenancing of soils.

For MALDI, the "density" of energy absorbed per unit volume (E_a/V) of the sample is the process-determining quantity. This is derived from equation (3):

(3) $\frac{E_a}{V} = \alpha \times H$

It is this equation which is at the core of MALDI analysis. If a matrix is chosen with a sufficiently high absorption coefficient α, a relatively low fluence H_0 can be applied. At a fluence of 100 Jm^{-2} and a pulse width of 2 ns the intensity (irradiance) of the laser beam at the surface of the sample is only 10^{11} Wm^{-2} or 10^7 Wcm^{-2}; this is not enough to induce any non-linear absorption such as non-resonant two photon absorption. During linear absorption, using a suitable variable attenuator in the laser beam, the absorbed energy per unit volume can be meticulously controlled. Another essential feature of laser absorption is that the energy is transferred more of uniformly to a macroscopic sample volume (except for the attenuation of the fluence into the sample and the fluence profile), this means the technique is not limited to just smaller sized molecules. This is useful for this project as the size of the molecules in the DNA extract is unknown

The fluence can be converted into a value for the photon-flux i.e. the number of photons impinging on the sample per single laser pulse. A fluence of 100 Jm^{-2} corresponds to a photon flux of $1.7x10^{16}$ photons per cm^2, each carrying an energy of 3.7eV at a wavelength of 337nm of the N_2 laser. A molar absorption coefficient of 10^4 L $mol^{-1}cm^{-1}$ represents a physical absorption cross-section of the chromophore of $1.6x10^{-17}$ cm^2, this equates to an average of 0.7 photons absorbed per matrix molecule for any given laser exposure. This is a very high density of excitation energy close to the solid-state energy stored in all of the intermolecular bonds. Therefore it leads to an explosive ablation of the excited sample volume. However, it renders even resonant two-photon absorption by the matrix unlikely. The high density of excited molecules results in a rather high rate of energy pooling in the sample in which two neighbouring excited molecules pool their energy, with one of them acquiring twice the photon energy and the other falling back to the ground state. In some models, this energy pooling is an important feature for the ionization of the molecules,

which requires at least the energy of two photons for an initial photoionization of the molecules (125).

3:1.1c The Ablation / Desorption Process

Every time the sample is exposed to the laser it leads to the removal of a bulk volume that is many monolayers of matrix molecules of the sample. In some respects the term desorption is somewhat misleading for the process, ablation could be a more correct term, and the two will be used interchangeably throughout this chapter.

The process of material ablation is closely associated with the ionisation of a minor fraction of the matrix and analyte molecules. Both ablation and ionisation take place on a nanosecond time scale. Experimentally, it is very difficult to sort out the complex contributions of the physical process induced by the laser irradiation in great detail.

As discussed previously (3:1.1b) at the threshold fluence for the detection of MALDI ions each laser pulse transfers an amount of energy to the sample, close to the sum of all bond energies in the solid; this is equivalent to the sum of the heat of fusion and evaporation. In all cases this energy will lead to the ablation of the excited volume although different energy dissipation processes need to be taken into account. In both UV and IR MALDI the energy dissipation by heat conduction during the laser pulse is negligible. For a typical UV matrix (as used in this project) for a penetration of laser radiation of 100 nm, the time constant for heat conduction is about 10 ns, this is a factor of three longer than the typical laser pulse width. The rapid heating of the sample by the laser radiation also generates a thermoelastic pressure pulse in the absorbing sample volume which travels out of the excited volume with the speed of sound carrying away part of the depositing energy. This energy is constantly carried away by the pressure wave and only amounts to a very small fraction of the total deposited energy and never substantially influences the ablation process as the pressure in the excited volume never reaches values high enough.

Using theoretical models and molecular modelling it has been suggested that the ablation process generates clusters and material particles besides gaseous components. This has been shown in various plume photographs (117). Although these particles have been seen; they occur only at times late during the ablation process after the generation of ions is complete; thus the particle emission does not seem to be relevant for or affect the MALDI ion generation. Also during modelling (126) of the ablation process it has been shown that clustering occurs during the ablation process but the cluster internal energy does not seem to suffice for a decay by matrix evaporation; this is one of the assumptions for the "lucky survivor" model for ionisation (see section 3:1.1d).

3:1.1d Ionization

The mechanisms of the ionization process are even less understood than those of the material desorption process. To understand the process better it is important to separate the matrix and analyte molecules. No numbers have been determined precisely although it is safe to presume that the ion yield for the matrix (i.e. the ratio of ions to neutrals) is somewhere in the range of 10^{-5} to 10^{-3}. The ion yield of the analytes is much higher, for typical cases in the order of 10^{-3} to 10^{-2} %. The intensity of ion signals, as determined from spectra, are not independent of each other, because charge transfer processes between the two species are taking place in the expanding plume and possibly already in the solid state upon laser irradiation.

Two models for the ionization process have been proposed. The older model (127, 128) assumes the initial step is neutral analyte molecules in the matrix crystals and a photoionization of the matrix molecules followed by a charge transfer to the analyte molecules in the plume. The "lucky survivor" (129) is more recent and assumes that proteins are incorporated into the matrix as charged species; most of which become re-neutralised within desorbed clusters of matrix and analyte.

The ionization process gives rise to molecules, some of which are neutral, and also both positively and negatively charged ones as shown in figure 24.

3:1.1e Fragmentation of the MALDI ions

Fragmentation of MALDI ions, as in all mass spectrometry is both good and bad. Fragmentation can lead to a substantial loss of spectra quality such as loss of mass resolution or even complete loss of signal of the intact parent ion. Alternatively, intrinsic or induced fragmentation can be an indispensible tool for the acquisition of structural information in MS experiments. The nomenclature for the fragmentation is closely related to time-of-flight analysers.

3:1.1f ToF Mass Spectrometry

Various mass analysers can be used in conjunction with MALDI including Time of Flight (ToF), Quadrupole Ion Trap (QIT) and Fourier Transform Ion Cyclotron Resonance. ToF is by far the most suitable analyser and has a high sensitivity which requires no scanning. The use of pulsed ion extraction and / or ion reflectrons (see section 3:1.2) allows ToF to have high resolution and good mass accuracy.

The principle of ToF mass spectrometry is, if ions which are accelerated with the same potential from a fixed point and time are allowed to drift, then the ions will separate according to their mass to charge ratio. The lighter ions will travel faster towards the detector as illustrated in figure 25.

Figure 25: The time of flight tube (Adadpted from (125))

All ions acquire approximately the same terminal kinetic energy as shown in equation 4:

(4) $K = \frac{1}{2}mv^2$

Therefore the ions terminal velocity is dependent upon the mass and thus the lighter ions travel faster than the heavier ones. Ions are readily detected at a certain point in space and flight times are readily assigned to masses as shown in equation 5:

(5) $t = \left(\frac{m}{2qV_0}\right)^{1/2}$L

Where t is the time of flight (s), m is the mass of the ion (kg), q is the charge on the ion (c) V_0 is the accelerating potential (V) and L is the length of the flight tube (m).

The main characteristics of ToF are that it is a pulsed technique, it can acquire a whole spectrum in a single shot, it has high transmission, has an unlimited mass range, it is relatively simple and has a low cost, it can tolerate ion extraction energy from a few KeV upto 25 KeV and has typical timescales of flight times greater than 100 µs and detected peak widths of 10 ns.

The peak width (resolution) of the ToF MS can be affected by ionisation conditions such as laser power and the sample itself, the ion optics, the detector speed including the detector type and channel size, and the data acquisition system speed including whether the acquisition system is analogue or digital and the transient recorder or time-to-digital converter (TOC) bandwidth. Adjustment of these parameters can be key to obtaining a useful and good quality spectrum.

The performance of the ToF MS can be optimised by the use of correctly designed ion optics to direct most of the created ions onto the detector giving a high sensitivity, using a combination of reflectrons and pulsed extraction can shorten single isotopic peak widths further giving a high resolution, power supply accuracy, high sample homogeneity and tight geometric tolerances to ensure the highest possible mass accuracy.

3:1.2 Linear and Reflector modes of detection

A MALDI-ToF-MS can be operated in either linear or reflector mode as shown in figure 26.

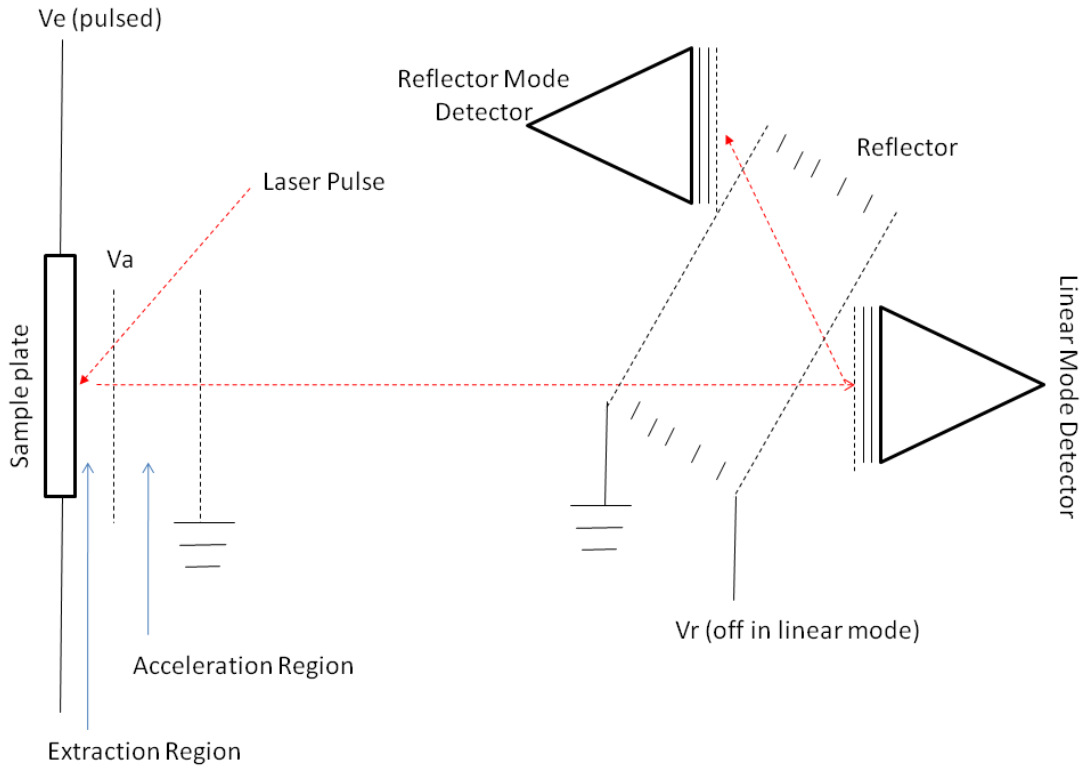


Figure 26: Schematic of the principles of MALDI showing both the reflector and linear modes of detection.

In linear detector mode, ion, under the influence of an electric field will travel in a straight line towards the detector, where the lighter masses arrive before the heavier masses. The mass relates to the time-of-fight and a mass spectrum shows the separation between the masses according to their flight time to reach the detector. For the mass analysis of full length proteins, it is sufficient to use MALDI in linear mode. However, for the analysis of more complex samples such as a peptide mixture (as required for protein identification by peptide mass fingerprinting) the MALDI is used in reflector mode. The use of a reflector allows for a longer flight path resulting in greater separation between the different masses giving a higher resolution in the mass spectrum (123).

The ion reflector was invented by Mamyrin *et al.*, to partly compensate for the initial velocity/energy distribution of MALDI ions. The reflector was also designed to re-focus the ion packets onto the detector. A reflector is a series of ring electrodes which, near their centre, would ideally create a constant electric field through a linear voltage gradient that slows the ions down and turns them around. This only occurs when the potential of the last electrode is larger than the acceleration potential of the ions, if this is the case, the ions are turned around and set back to the second detector as shown in figure 26. The higher-energy ions will arrive at the reflector first and will then go farther up the voltage "hill" to the detector this increasing their time in the time-of-flight tube. If the reflector voltage is adjusted, both the high and low energy ions can be focused to reach the detector at the same time, thus narrowing the bandwidth for the output signal (125).

A reflector can correct the flight time dispersion to the first order and mass resolutions of over 10 000 have been achieved. The dispersions can be corrected to the second order by using a two-stage reflector, a reflector with two different electric fields. A delayed ion extractor compensates mainly for the flight time dispersion due to the initial velocity distribution of the ions. The reflector compensates for the energy dispersion which may result from a sample with a non-flat morphology for example (125). Put simply, the three main things a reflector is used for is; to correct time dispersion due to initial kinetic energy speed which corrects peak broadening, increase the separation of the ions of different mass (resolution) due to the increased flight path and lower flight time and to filter out neutral molecules, which are not reflected in the ion mirror.

Within the scope of this research care must be taken over whether to use the MALDI-ToF-MS in either linear or reflector mode in order to obtain the most useful spectra. This is dependent upon the samples being analysed, but is complicated due to their unknown nature.

3.1.3 Previous applications of MALDI in forensic work

Although MALDI-TOF has been used in forensic science for a wide variety of applications, it is not one of the most well-known or commonly used methods for analysis and data

acquired from the instrument is not routinely presented in a courtroom, unlike data acquired from more commonly used instruments such as GC-MS, HPLC and IR spectroscopy. This is not to say its use within forensic science is limited but greater validation may be required to ensure that it is robust and reliable.

Examples of MALDI being used within forensic science research include the analysis of rosin and modified rosin esters in adhesives (130), where MALDI was successfully used as an analytical method to identify esters and modified esters in 22 adhesives. MALDI was also used in the profiling of trace constituents in condom lubricants in the presence of biological fluids (131), whereby the MALDI data acquired was supported by infra red spectroscopy. MALDI has also been implemented in the analysis of cannabis pollens (132). These examples show the potential for the use of MALDI-TOF within a wide range of forensic disciplines.

3.1.4 MALDI in DNA work

Butler (121) has highlighted the use of MALDI in the analysis of DNA. The method is similar to that discussed in 3:1.1 where a liquid DNA sample is combined with an excess of a matrix compound such as 3-hydroxypicolinic acid. The samples are then spotted onto a metal or silicon plate and the sample allowed to air dry and the DNA and matrix co-crystallize (see 3:14 sample preparation section). The sample plate is then introduced to the vacuum environment of the MS for analysis. A rapid laser pulse initiates the ionisation process and the matrix molecules that surround the DNA protect it from fragmentation during the ionisation process. As previously explained, each of the laser pulses initiates the ionisation of the sample and subsequently the separation of the ions in the ToF tube. The DNA ions travel to the detector in a matter of several hundred microseconds as they separate based on their mass size. Several seconds are taken to analyse each sample as multiple laser pulses are taken and then averaged to form the final mass spectrum.

One of the main advantages of the use of MALDI-ToF-MS to analyse DNA samples is that it enables a new level of high-throughput analysis to be achieved. Another benefit of this

technique is accuracy. For example, STR alleles can be sized with such a degree of accuracy that the use of this technique permits DNA typing without the use of an allelic ladder. (121)

However, there are several challenges surrounding the use of MALDI-ToF-MS for the analysis of DNA; these become apparent when PCR products such as STR's are being analysed. The most significant problem of MALDI-ToF-MS is that the resolution and sensitivity of the mass spectrometer are diminished when either the DNA size or the salt content of the sample is too large. Another major disadvantage of the technique is the cost of the system. Also, the more conventional fluorescence methodologies have a wide-range acceptance in DNA communities and therefore MS technologies may not become commonplace in the analysis of DNA. MS techniques for the analysis of DNA have not been validated for forensic casework and to date have not been presented in a court of law. (121)

MALDI-MS has been used in quality control of oligonucleotides, analysing markers such as SNPS's and analysing the expression of specific alleles. More specifically, UV-MALDI has been used for the analysis of complex mixtures containing large DNA fragments such as DNA sequencing ladders. A lot of the use of MALDI in DNA work has looked at proteins and peptides. However, nucleic acids are harder to detect by MALDI than peptides, this is due to their negatively charged sugar-phosphate backbones.

One of the problems with analysing DNA with MALDI is size dependent fragmentation. As previously discussed, fragmentation is caused by excess energy during laser desorption; this fragmentation results in a loss of signal intensity for intact DNA. RNA is more stable in MALDI and this is due to the additional 2'-hydroxyl group which stabilises the gylcosidic bond leading to significantly reduced depurination and fragmentation of the whole oligomer. Additionally, the use of MALDI for DNA work is limited by the bias of MALDI towards smaller DNA (133).

The use of MALDI for the analysis of DNA has also been reported by Petkovski *et al.* (134) whereby MALDI with its high sensitivity, precision and speed gave a powerful method for

the forensic and anthropological exploitation of biallelic markers. Most recently, Shi *et al.* (135) have utilised MALDI for the identification of genes that affect the metabolism of different drugs in a Chinese population.

3.1.5 Sample preparation

The first step of the MALDI process is the preparation of the sample by mixing the analyte with a molar excess of matrix. Currently, MALDI is based on the laser desorption of solid matrix-analyte deposits, however, this technique has some disadvantages. The technique has a strong dependence on the sample preparation method, a short sample life-span and a low shot-to-shot reproducibility (120).

3.1.5.1 Matrices

An important feature of MALDI is the way the sample and matrix interact. An early surprise in the development of MALDI was that all of the well-functioning matrices incorporate the analyte in the crystals quantitatively (up to a maximum concentration) and in a homogenous distribution. When a UV laser is being used a typical matrix is an aromatic acid with a chromophore that strongly absorbs the laser wavelength. When other lasers are used such as the mid-infrared range where the matrix is energised by vibrational excitation different matrix compounds must be used (120).

A matrix must be able to meet several requirements simultaneously:

- be able to embed and isolate analytes e.g. co-crystallisation

- be soluble in solvents compatible with the analyte

- be vacuum stable

- absorb the laser wavelength

- cause co-desorption of the analyte upon laser irradiation

- promote analyte ionization.(120)

Unfortunately, but not surprisingly, there is no one MALDI matrix or sample preparation method that is suitable for all analytical problems and analytes. There are different matrices of choice for different classes of analytes as seen in table 7.

Table 7: Some of the different matrices available for use in MALDI-ToF MS analysis of biological components. (117, 133)

Matrix	**Quality of Performance**	**What analytes matrix is suitable for and addition comments**
3-Hydroxypicolinic acid (3-HPA)	Excellent	Excellent for oligonucleotides greater than 10 bases, less suitable for smaller oligonucleotides due to matrix adducts.
Picolinic acid	Good	Usually a co-martix of 3-Hydroxypicolinic acid.
6-Aza-2-thiothymine	Good	Little fragmentation, rarely used in practice.
4-Hydroxy-3-methoxycinnamic acid (ferulic acid)	Medium	Rarely used in practice, used for oligonucleotides larger than 3500 Mr
2,5-Dihydroxybenzoic acid (DHB)	Medium	Rarely used for oligonucleotides
2,4,6-Trihydroxyacetophenone (THAP)	Excellent	An alternative to 3-HPA, particularly good for small oligonucleotides. Can be mixed with 2,3,4-THAP.
2,3,4-Trihydroxyacetophenone (THAP)	Excellent	An alternative to 3-HPA, particularly good for small oligonucleotides. Can be mixed with 2,4,6-THAP.
Alpha-cyano-4-hydroxycinnamic Acid	Medium	Particularly for PNA, modified and small oligonucleotides with charge neutral backbones. Limited use of the matrix.

Alpha-cyano-4-hydroxycinnamic acid methyl ester	Excellent	For PNA, modified and small oligonucleotides with charge neutral backbones. Limited use of the matrix
Anthranilic Acid	Medium	Rarely used in practice, rather for oligonucleotides larger than 3500 Mr.
Salicylamide	Medium	Rarely used in practice, rather for oligonucleotides larger than 3500 Mr.

3.1.5.2 Preparation and Drying of sample

A MALDI matrix is applied evenly onto a flat conductive plate which is then introduced into the ionization chamber of the mass spectrometer. The most commonly used plate materials are stainless steel, aluminium / nickel alloys and gold or semi-conductors such silicon. There are several different methods of sample preparation for pipetting the sample and matrix onto the plate for MALDI-ToF-MS as shown in table 8.

Table 8: Sample preparation methods for solid and liquid matrices and special preparations for MALDI-ToF-MS.

Solid Matrix	**Liquid Matrix**	**Special Preparation**
Dried-droplet Vacuum-drying Crushed-crystal Fast-evaporation Overlayer Sandwich Spin-coating Slow-crytallization Electrospray Quick and dirty Matrix-precoated targets	Chemical liquid Particle-doped (two-phase) liquid Chemical-doped liquid	Solid supports MALDI on 2D gels Insoluble samples

As can be seen in table 9 there are many different sample preparation methods for MALDI analysis. The dried-droplet method is the original sample preparation technique developed by Hillenkamp and Karas in 1988 (in (117)). Since this technique has been developed it has only had a few minor modifications. A drop of aqueous matrix solution is mixed with analyte solution and mixed with a solvent or solvent system and dried resulting in a solid-deposit of analyte doped matrix crystal that can be introduced to the MALDI for analysis. In order to remove the non-volatile components of the original solution, the matrix/analyte crystals may need to be washed. This method tolerates the presence of buffers and salts well, making it useful for the analysis of DNA as these may be present in the DNA extract. This method of sample preparation is also a good choice for samples containing more than one component (120). See section 3.3.2 for how this method is used within this study.

The crystals created by the drying of the sample vary in size depending on the matrix, but can vary from submicrometer to several hundred micrometers. Often, surface tension can lead to a non-homogenous distribution of the individual crystals near to the edge of the preparation. The best MALDI spectra are usually achieved from only specific areas of the crystals, which often requires the experimenter to manually search around the crystal looking for the "sweet spot". The sweet spot has been the subject of much debate and it is commonly thought that they are caused by the inhomogeneous distribution of analyte within the crystals. However, this has been disputed by Horneffer *et al.* (136) who found a homogenous distribution of analyte material in the crystals when they were labelled fluorescently by CLSM studies. A cause of the sweet spots may be due to a different ionization state of the analyte at different locations, or heterogeneous orientation of the matrix crystal surface relative to the spectrometer axis and perpendicular to the direction in which the ions are ejected in conjunction with limited angular acceptance of the mass spectrometer (117). This issue surrounding sweet spots is discussed in more detail directly relating to this research in section 3.5.

The addition of the analyte solution should not noticeably change the crystallization behaviour of the neat matrix; this ensures that the solution excess of the matrix with respect to the analyte is maintained in the crystals and helps to keep the contamination levels low. If the solute component of the matrix noticeably changes the behaviour or the crystallisation of the matrix then it will prohibit a successful MALDI analysis (117).

It has been found that the best results (obtained using THAP as a matrix) were gained when the sample was allowed to absorb moisture from the atmosphere following vacuum drying. This method gave a spot with small crystals compared to long needles forming when the matrix was dried under ambient conditions and the sample appearing translucent, almost oil-like, with vacuum drying (137). However, as this research is being carried out for the potential for the use in forensic science, care must be taken to ensure that there is no contamination of the samples, this is something which could occur when the samples are allowed to dry under ambient conditions.

3.2 Preliminary work – different matrices

Within this research, all solvents and chemicals were obtained from Sigma Aldrich. The chosen MALDI plate was a 2 mm thick 96 well stainless steel MALDI target from Shimadzu Biotech UK. The MALDI mass spectra were acquired using an Axima Cfr Plus from Shimadzu Biotech.

The optimisation of MALDI depends upon the identification of both the best matrix and preparation method for the particular analyte and therefore two different matrices were tried. The two matrices, 3-HPA and THAP were chosen based upon their properties and recommendation from Shimadzu Biotech that they may be suitable for the analysis required.

At first the different matrices were spotted out without any sample to give an idea of their properties. A standard DNA was then used as the sample to decide which matrix was most suitable.

3:2.1 3-Hydroxypicolinic acid (3-HPA)

3-HPA is a commonly used matrix in MALDI; it has the chemical formula $C_6H_5NO_3$ and has a molecular weight of 139.11.

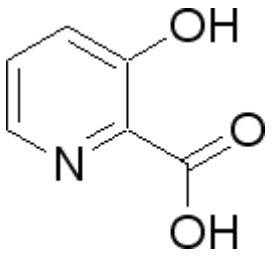


Figure 27: Chemical structure of 3-HPA

As 3-HPA can be used in two different ratios with the solvent, 1:1 and 10:1 (as recommended by Shimadzu Biotech) two different matrices were prepared. To prepare the 1:1 3-HPA 50 mg of 3-HPA was weighed and mixed with 1 mL of 50:50 acetonitrile and MilliQ water. 50 mg of diammonium Hydrogen citrate (DHAC) (see 3:2.3) was weighed and mixed with 1 mL MilliQ water. Then the HPA and DHAC was mixed in a 1:1 volume and finally mixed with the sample in a 1:1 volume and spotted onto the target.

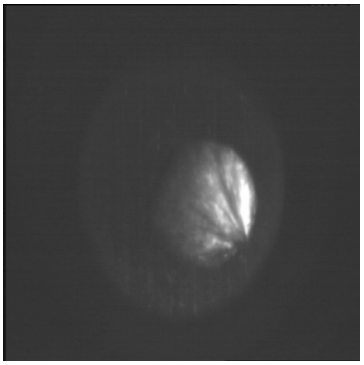


Figure 28: Image taken from the MALDI-ToF-MS of the sample spot for 1:1 3-HPA

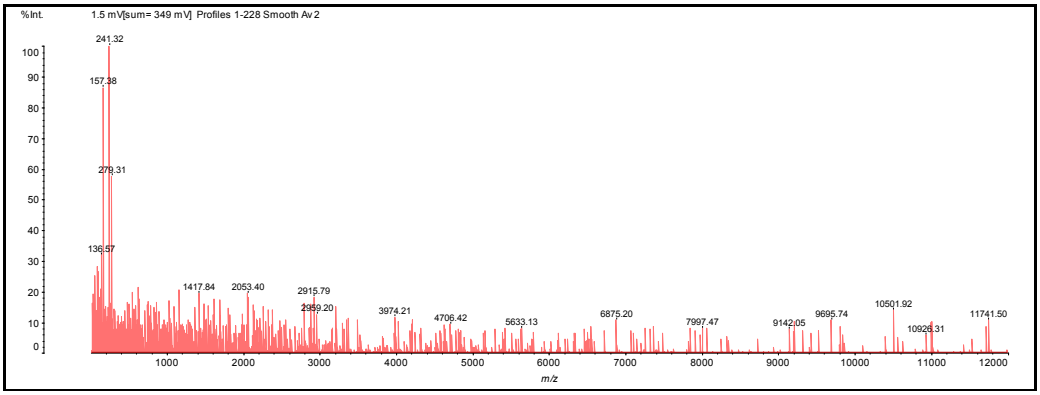


Figure 29: Spectra of 1:1 3-HPA

To prepare the 10:1 3-HPA 50 mg of 3-HPA was weighed and mixed with 1 mL of 50:50 acetonitrile and MilliQ water. Next 50 mg of DHAC was weighed and mixed with 1 mL MilliQ water. Then the HPA and DHAC were mixed in a 10:1 volume and finally mixed with the sample in a 1:1 volume and spotted onto the target.

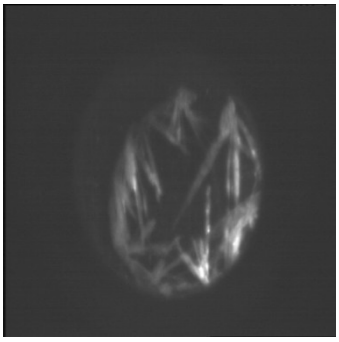


Figure 30: Image taken from the MALDI-ToF-MS of the sample spot for 10:1 3-HPA

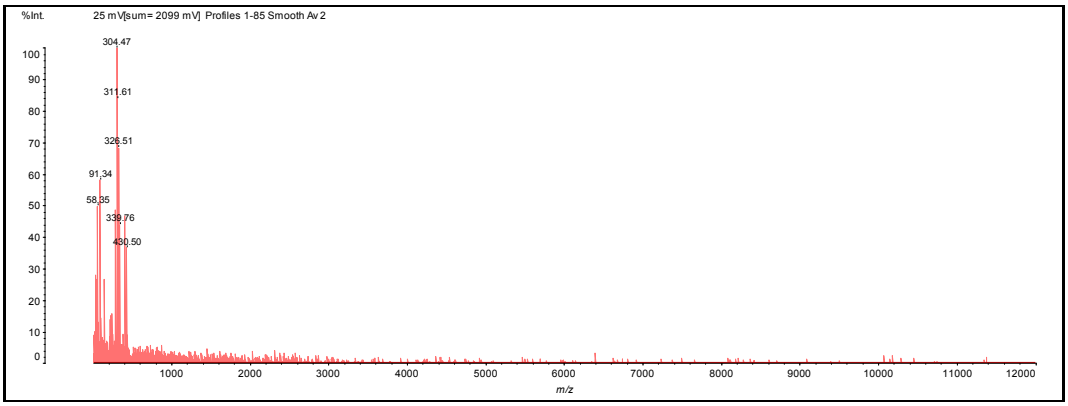


Figure 31: Spectra of 10:1 3-HPA

As can be seen from figures 28 to 31, the 10:1 ratio appeared to give a larger crystal and cleaner spectra. However, the crystal produced by the 1:1 ratio although small appeared more uniform; the 10:1 spot gave slightly differing spectra's across the crystal.

3:2.2 2,4,6-Trihydroxyacetophenone (THAP)

THAP is another fairly common MALDI matrix; it has a chemical formula of $(HO)_3C_6H_2COCH_3$ and a molecular weight of 186.16.

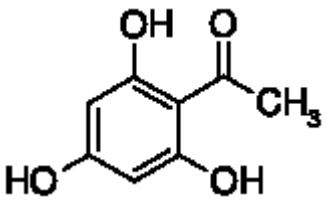


Figure 32: Chemical structure of THAP

To prepare the THAP 12 mg of THAP and 5 mg of DHAC was weighed and then the THAP and DHAC were dissolved in 1 mL of 50:50 acetonitrile and MilliQ water. Finally they were mixed with the sample in a 1:1 volume and spot onto the target.

Although THAP has been the preferred matrix in some studies such as Papac *et al* (1996) (137) due to it providing a lower limit of detection and giving less prompt fragmentation this was compared to matrices such as 2,5-Dihydroxybenzoic acid (DHB) and 6-aza-2-thiothymine and not a direct comparison with 3-HPA and therefore the decision was made to trial both matrices using the same samples so that a direct comparison could be made.

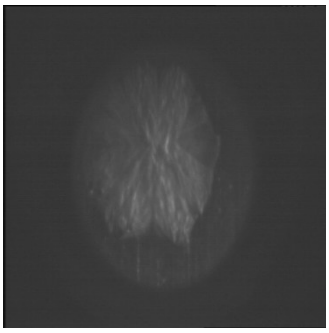


Figure 33: Image taken from the MALDI-ToF-MS of the sample spot for THAP

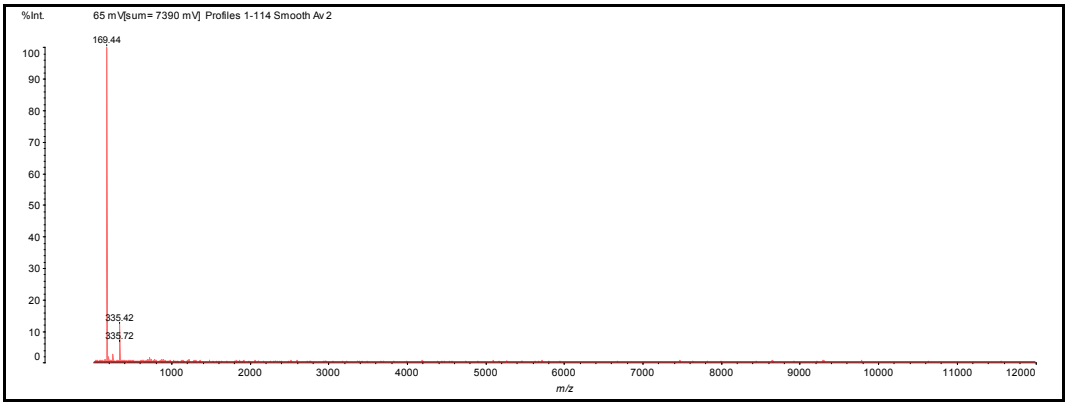


Figure 34: Spectra of THAP

3:2.3 DHAC

DHAC is the solvent used in both the 3-HPA (both 1:1 and 10:1 ratios) and THAP matrices. It has the chemical formula $HOC(CO_2H)(CH_2CO_2NH_4)_2$ and has a molecular weight of 226.18.

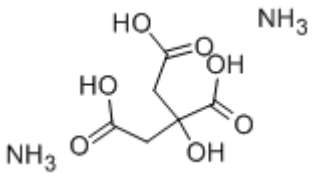


Figure 35: Chemical Structure of DHAC

3.3 Methodology

In order to determine whether or not the use of MALDI-ToF-MS is a viable means of analysing the plant DNA present in the soil DNA extracts were spotted onto a MALDI plate and analysed without any other treatment. The decision was made to not amplify the DNA samples prior to analysis as this can cause further problems in the MALDI analysis as discussed in section 3.1.4, and also due to the unknown nature of the DNA extracts it was unsure as to what primers would be suitable for DNA amplification. Although there was no literature available to support this, the decision was made to just analyse the DNA extracts, with the aim being that this crude methodology would enable one to see if there were any similarities and differences within the DNA MALDI spectra obtained from multiple samples

collected from one location and for the different sample locations. This would be seen through the fragmentation pattern of the samples.

3.3.1 DNA extraction

DNA was extracted from the Norfolk soil samples using the PowerSoil™ DNA Isolation Kit from MoBio Laboratories, Inc. (112) as discussed in section 2:6.5.

3:3.2 Chosen matrix and drying technique

As a result of trying out different matrices and drying techniques the chosen matrix was 3-HPA, and this was used in two different ratios 10:1 and 1:1 with DHAC. Both ratios were used as in the preliminary experiments they yielded a similar quality of crystal and subsequent spectrum and therefore both ratios were used for further samples; however as can be seen from the preliminary results the 10:1 ratio gave a cleaner background spectrum but as has been discussed this changed over the sample spot. 3-HPA was chosen over THAP due to the fact it has been used to detect oligonucleotides of differing sizes both large and small and due to the unknown nature of the samples this was thought to be useful. Also, the THAP matrix did not always produce any crystals when drying and was therefore hard to gain a spectrum from. The DNA extracts from the samples when spotted on their own produced crystals and as it is optimum for the matrix to produce the same crystals as the samples THAP did not appear to be an ideal matrix. A clean background spectrum played a role in the choice of the matrix and the quantities of DNA in the extracts were known to be low and therefore the risk of masking any peaks by matrix ones needed to be eliminated.

In the preliminary experiments different spotting techniques were tried, each time a volume of 0.5 µL was used (of both the sample and matrix); the matrix was spotted on to the plate and allowed to dry and then the sample pipette onto the matrix, this was repeated but followed by another layer of the matrix once the sample had dried and finally the sample and matrix were mixed onto the plate and allowed to dry. Each time the sample plate was left to air dry in a fume hood before any analysis to reduce the risk of

contamination. As well as preparing the samples for analysis a spot was prepared each time of just the matrix and spectrum obtained. Fresh matrix was prepared for each analysis.

3:3.3 Instrument parameters

To try to get the optimal spectra the instrument was run in both linear and reflector modes. The power was approximately 72, chosen after different powers were tried in preliminary experiments. 200 spectra were acquired and overlaid.

When obtaining the spectra the laser beam had to be manually moved around the sample spot to search for a sweet spot.

3:3.4 MALDI-ToF-MS Calibration

Oligonucleotides were used to calibrate the MALDI-ToF-MS. Each of the oligos was produced on a synthesis scale of 0.05 µL, with reverse phase cartridge purification by Sigma Aldrich and were received dried.

Table 9: Oligonucleotides used to calibrate the MALDI-ToF-MS.

	Sequence	**$(M+H)^+$**	**M**
12-mer	ACG TAC GTA CGT	3646.45	3645.40
20-mer	ACG TAC GTA CGT ACG TAC GT	6118.05	6117.00
30-mer	ACG TAC GTA CGT ACG TAC GTA CGT ACG TAC	9192.05	9191.05

The oligionucelotides listed in table 9 were resuspended with MilliQ water to a final concentration of 100 pmol / µL. The stock oligonucletotide was then made into 20 µL aliquots. For the calibration mix one 20 µL aliquot of the 12, 20 and 30-mer vials was taken and to the 30-mer vial 2 µL of 12-mer and 5 µL of 20-mer was added. Then to this 50 µL of MilliQ water was added. This calibration mix was then prepared in the same way as the samples.

3:3.5 Cleaning the MALDI plate

To clean the 96-well target plate, the target was first wiped with methanol to remove all visible samples. The target was then placed in a suitable container such as an Eppendorf box lid and flooded with acetone and placed in an ultrasonic bath for 15 minutes. The solvent was discarded into a waste non-chlorinated solvent bottle. The plate was then flooded with methanol placed in an ultrasonic bath for 15 minutes. The solvent was discarded into a waste non-chlorinated solvent bottle. The plate was then flooded with MilliQ water and placed in an ultrasonic bath for 15 minutes. The water was discarded. The target was finally flooded with methanol and placed on its end to allow the methanol to drain off and then left to dry for at least one hour in a fume hood. When not used, the MALDI plate was stored in an airtight container.

3:4 Results

The spectra obtained by the MALDI-ToF-MS in reflectron mode were exported as mzXML raw data files and were then viewed using OpenMs 1.6.0 software.

First, spectra from two different spots from the same sample spot 2309092_2 using a 10:1 3-HPA matrix were obtained.

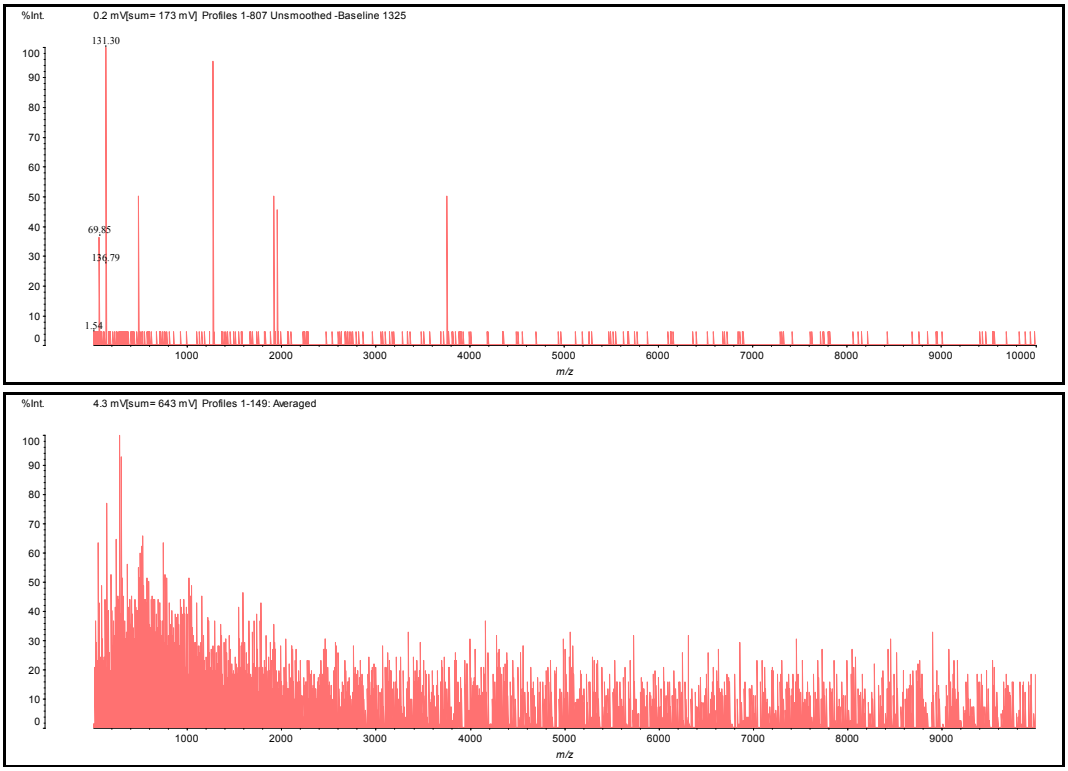


Figure 36: Spectra obtained from the same sample using 10:1 3-HPA with a layered sample preparation.

As can be seen different spectra are obtained from the same sample (figure 36). This also happened when a 1:1 HPA matrix was used; this shows that the spectra differs at different parts of the sample crystal and highlights the need for the sweet spot to be found.

As it was found that different spots on the same sample were producing vastly different spectra, it was decided to spot out the sample multiple times using the same preparation and spotting method to see if reproducible spectra could be produced. This was done using both a layering and a mixing preparation to see if one method produced more reproducible spectra (figures 37 and 38). As can be seen in this was not possible, with very inconsistent spectra generated.

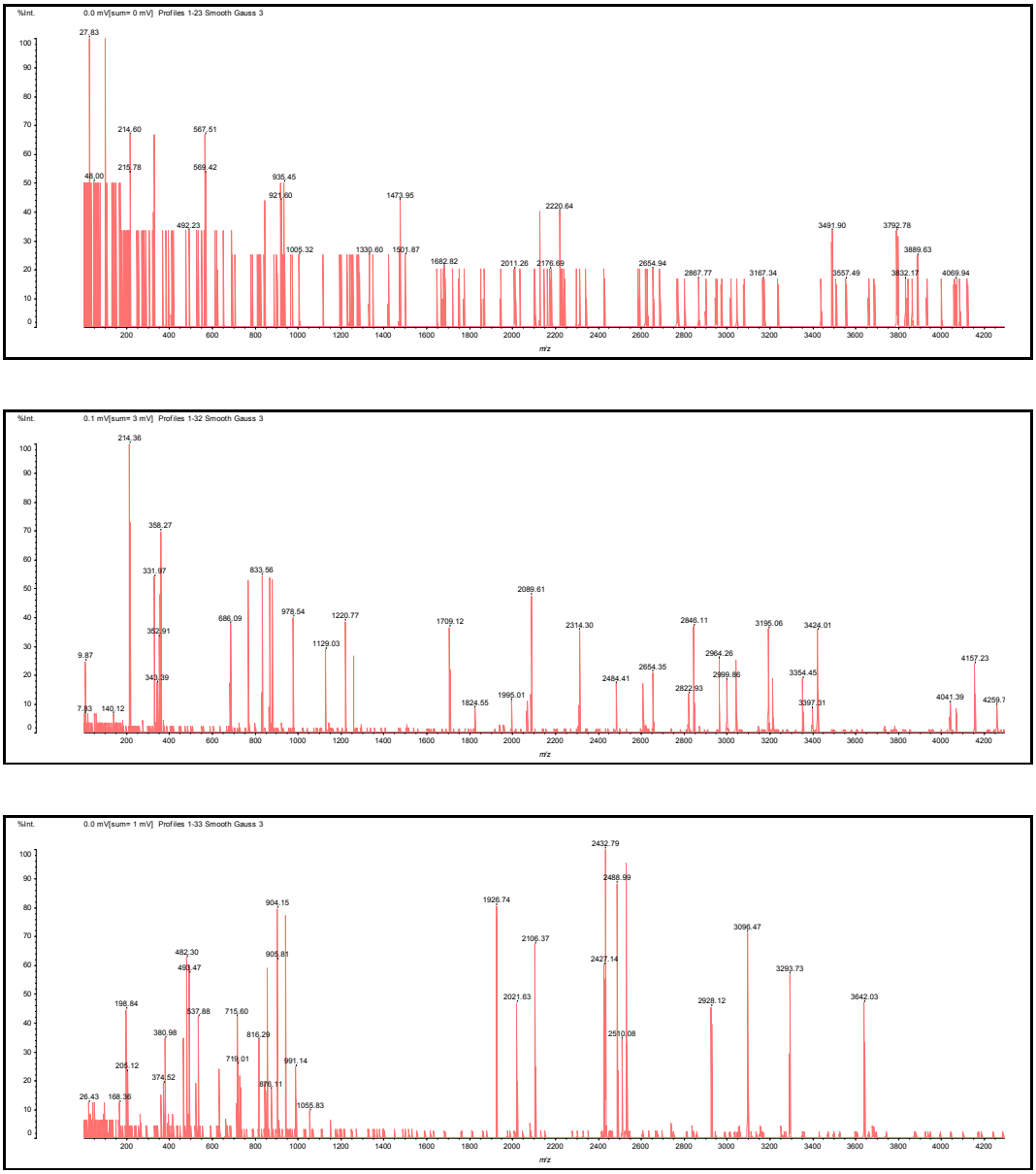


Figure 37: Spectra from the same sample spotted multiple times on plate 2109092_2 prepared by pipetting 0.5 µL of matrix allowing it to dry, then 0.5 µL of sample followed by 0.5 µL of matrix. (10:1 3-HPA matrix)

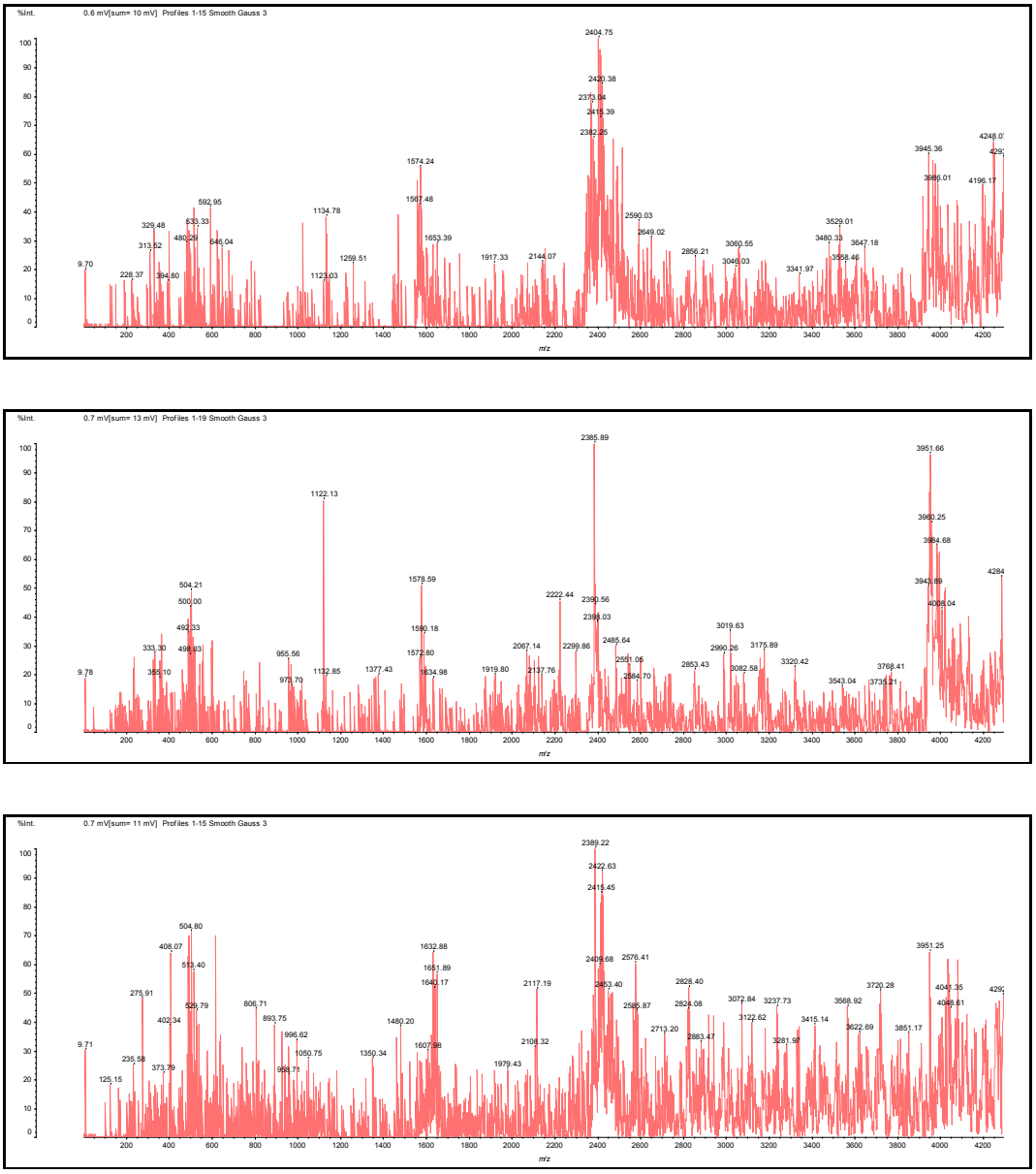


Figure 38: Spectra of the same sample 2109092_2 but different spots on plate prepared by mixing 0.5 µL of 1:1 3-HPA matrix and 0.5 µL of sample on the same plate.

As can be seen when the sample is mixed with the matrix on the plate(figure 38) it appears to give a more similar spectra than when the sample is layered between the matrix (figure 37); however using both methods differences in the spectra can be seen meaning a reproducible spectra from the same sample cannot be obtained. The spectra also have quite differing intensities. As no successful spectra were being obtained; the samples were diluted to varying amounts to see if this yielded any more successful spectra. (Sample number 2309098_1 was used) as seen in figure 39.

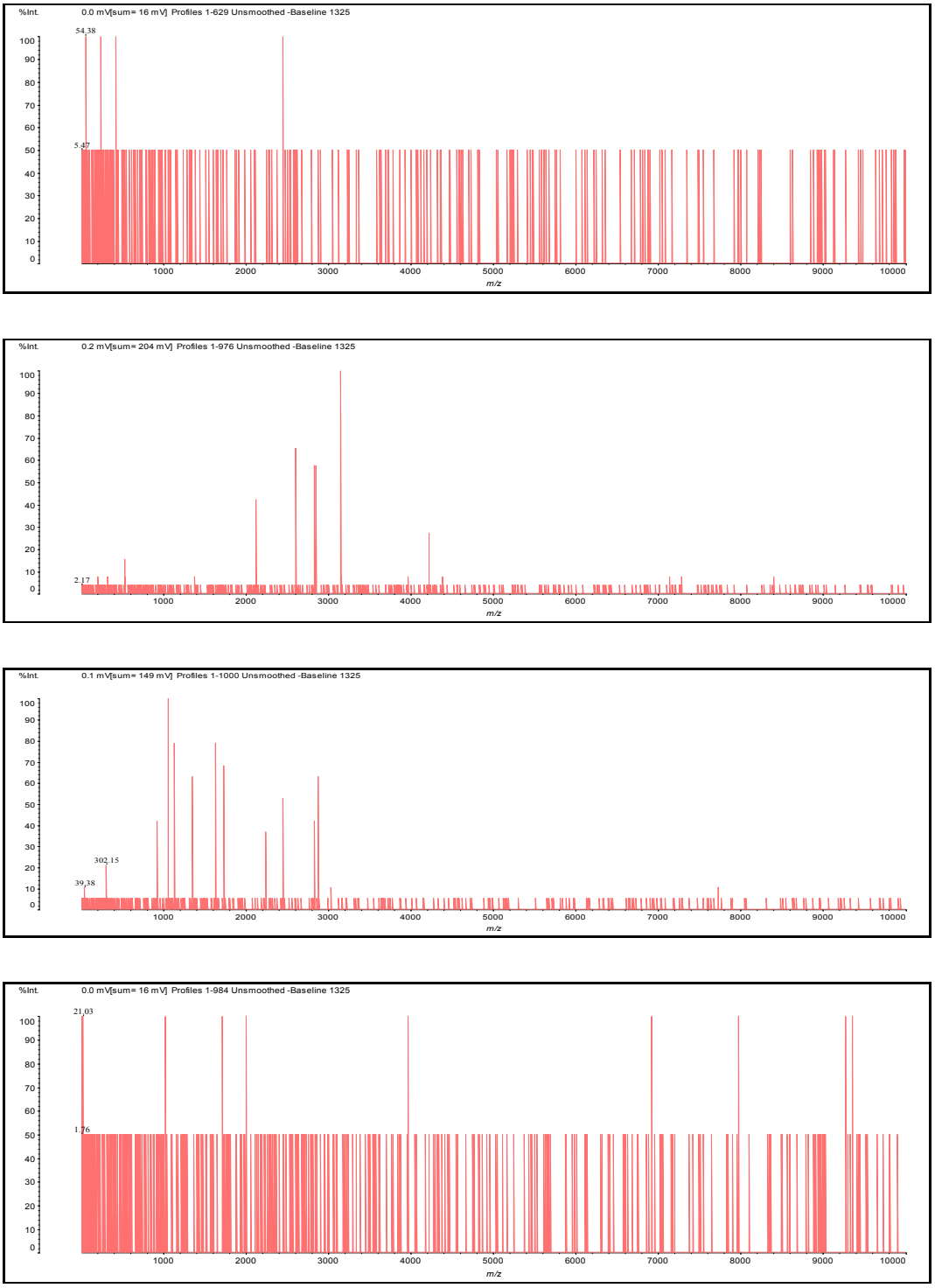


Figure 39: Spectra from different dilutions of the same sample (from top two times dilution, five times dilution, ten times dilution and twenty times dilution).

As can be seen in figure 39 these spectra would appear to be background spectra although they are indeed samples due to the uniformly noisy baseline shown for each of the dilutions. As during the various experiments variations had been seen from spectra from

the same samples; the dilutions of 2309098_1 were run on another occasion to see if the spectra were the same.

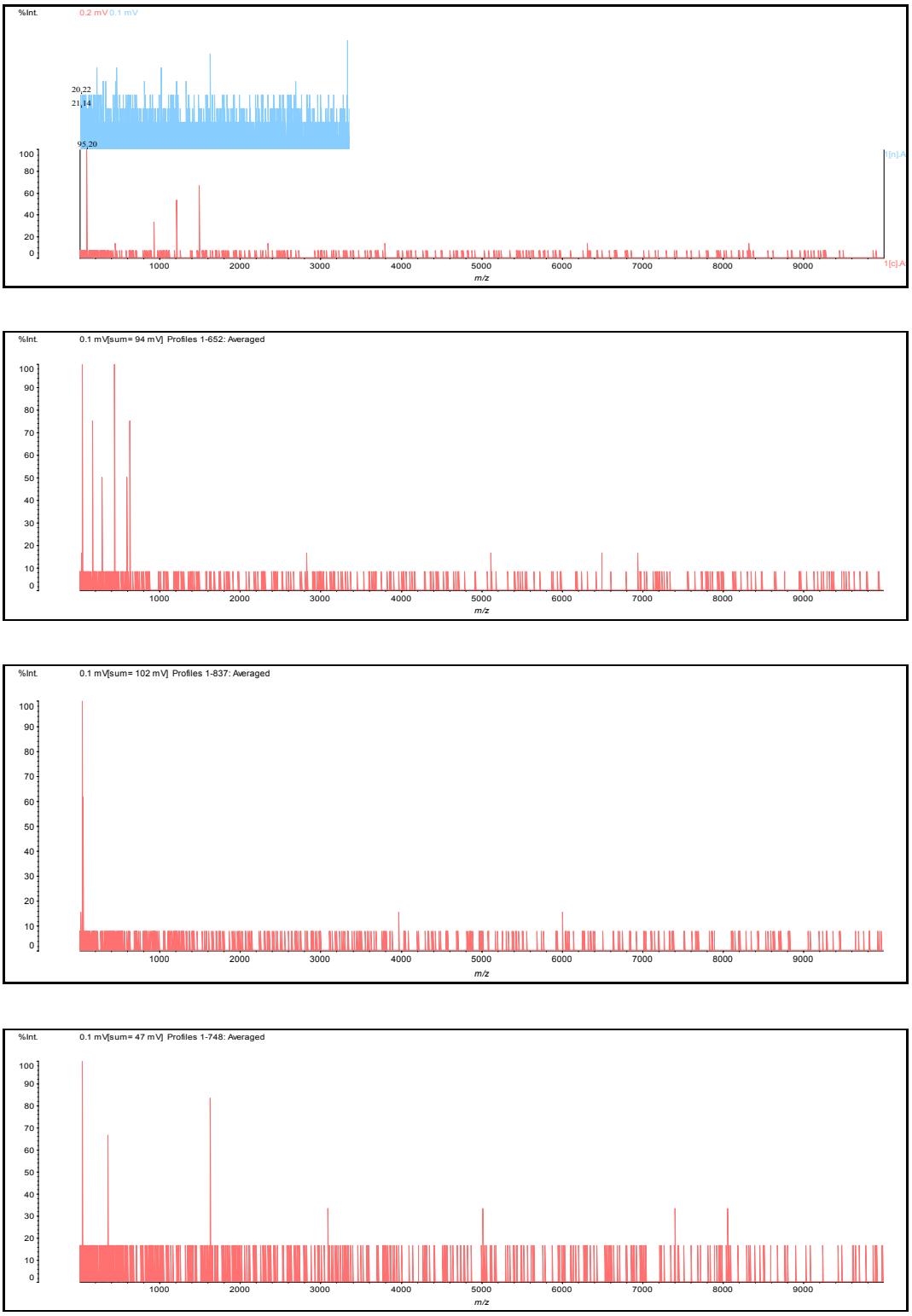
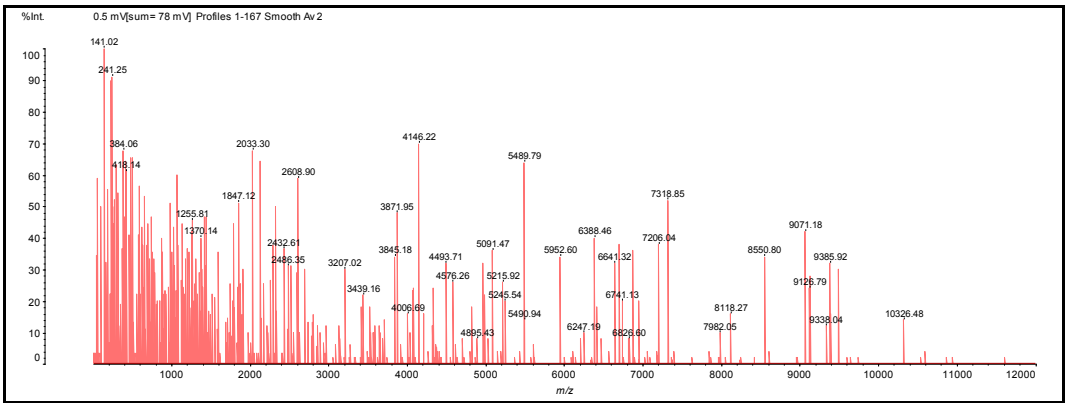


Figure 40: Reproduced spectra from the same samples as the previous figure (39) run on a different occasion.

As can be seen from the two sets of spectra from the same dilutions taken on two separate days using the same matrix (1:1 3-HPA) and the same sample preparation method the spectra are different as seen in figure 39 and 40. However they still appear to have a uniformly noisy baseline; it is thought this is down to a very low amount of DNA being present in the soils (this could be tested further by spiking a soil sample with a known DNA sample if time allowed). This was also the same when two separate sets of dilutions of the same sample were analysed using a 10:1 3-HPA ratio.

In order to determine if different DNA extracts from the samples taken from the same site (intra site) yielded any similarities two samples taken from location 2109096 were analysed.

2109096_1



2109096_2

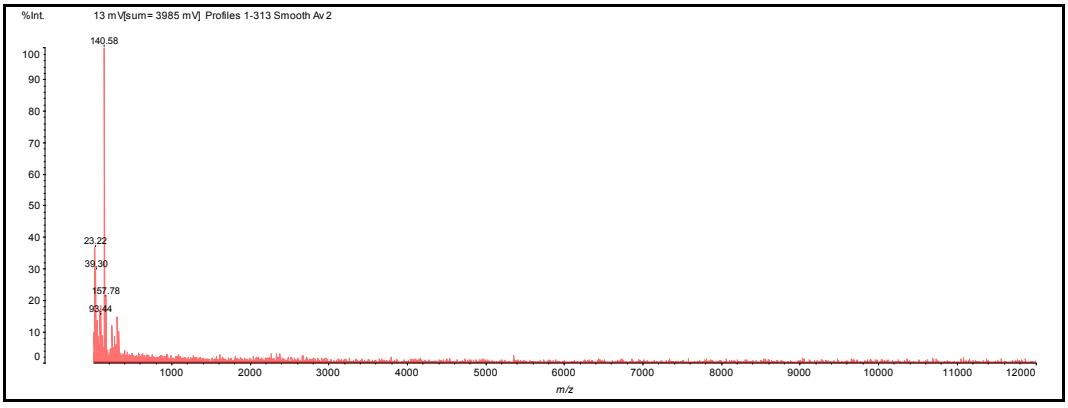


Figure 41: Spectra taken from two samples collected within a 5m by 5m area.

As can be seen in figure 41, both of these spectra are taken from the same site within a 5m by 5m area and have vastly differing spectra, although both samples were extracted and prepared for MALDI in the same way (using a 1:1 3-HPA spectra). This is unexpected; although some intra site variation is expected it was not expected to be to this extent. In order to try to determine if this was a true representation of the sample and that it did vary in its' DNA components or if it was an analytical problem such as some of the problems with variation shown previously, the samples were also run using 10:1 3-HPA and THAP. 10:1 3-HPA also yielded variation (but to a lesser degree) and THAP yielded no spectra.

As there is variability in spectra taken from multiple spots of the same sample at present using this technique it would not be possible to compare the spectra of two different samples from different locations. As reproducible spectra could not be produced it was decided that in the context of this study this method was not feasible with the current knowledge and available methods; but this was not to say that this method is not suitable for more specific analytes and that in the future this method might be possible with additional time, knowledge and experience with both the samples and the technique as discussed in 3:5 and chapter 7.

3:5 Discussion

3:5.1 Matrices

From the experiments carried out in this project it would appear that 3-HPA in a 10:1 ratio is the best matrix as this yielded the best spectra which although still variable was the least variable.

The optimisation of MALDI depends upon the identification of both the best matrix and preparation method for the particular analyte. This is made more difficult due to the unknown samples. More extensive optimisation using known standards would be necessary to develop reproducible methods before any work could be undertaken in unknown samples.

3:5.2 Sample intensity and reproducibility

Some samples did not have enough DNA in them - only 10% intensity was observed on the spectra. As there were only very low levels of DNA in the samples used for trying *mat*K (see chapter 2) DNA LoBIND pipette tips and Eppendorf tubes (from Eppendorf) were used in the extraction process, however this made no difference to the spectra, some samples still had little DNA and others had more. However, exact DNA concentrations of the same samples extracted using normal pipette tips and Eppendorf tubes and the DNA LoBind were not taken and compared using a method such as DynaQuant.

The MALDI spectra were not reproducible, something which has also been seen in work by Williams, 2003 (138, 139). Different spectra can be obtained from different parts of the same preparation – a common problem with all different types of samples.

It was observed that sometimes the same extract gives differing spectra, there appeared to be no apparent reason for this but in order for the technique to have forensic potential this problem would need to be eliminated.

In order to make the spectra more reproducible using a greater volume of sample i.e. more than 0.5 µL was tried. However this did not work and still had the same reproducibility problems. The problem of searching for a suitable spot on the prepared sample for analysis (the sweet spot) may be reduced by using the rastering mode on the instrument, however this was not evaluated.

Although over the years the dried droplet method of MALDI sample preparation or variations of this technique are the most commonly used; it appears that the sample preparation process is a personal choice and can be seen as something of a “black art” (117). This means that for forensic work, this technique may not be suitable as without a verified protocol the methods would not be admissible in a court of law. Krupková *et al*. (140) discuss how MALDI spectra can be confusing due to signals relating to in-source

fragmentation, something to consider when looking at the spectra obtained in this study. Due to the unknown nature of the samples, it is not possible to determine which peaks on the spectra are actually from the sample and which are from fragmentation, if this problem could be overcome the spectra may be much less messy and more meaningful.

3:5.3 Factors affecting the spectra

Arguably, MALDI has been unsuccessful due to the complex and unknown nature of the samples. The DNA extracts were not amplified before being analysed using the MALDI-ToF. Introducing the amplification step may help to make this technique more suitable for the analysis of plant DNA. Specific primers could be used in amplification for different plant species, but again this is problematic in not knowing exactly what plant species are present in the samples although using the vegetation maps (see chapter 5) could aid this. It also needs to be taken into consideration that there have been problems when amplifying the *mat*K gene; Fazekas *et al.* (111) describe how multiple primer sets were needed for amplification and the portion of the region that was subsequently sequenced varied among taxonomic group depending on the primer set used.

Some of the spectra appeared to be very noisy, this may have been the nature of the samples, in some cases noise may also have indicated a degree of contamination but it must be remembered that MALDI can have a substantial amount of noise (141).

3:5.4 Potential for MALDI for the analysis of soil DNA in forensics.

There is the potential for the use of MALDI-ToF for forensic science if one is looking for something more specific and a suitable extraction and amplification method can take place. It is also worth considering exploring the use of other soft ionization methods for the analysis of soil DNA.

From a forensic perspective MALDI is not ideal in many ways. Sigma Aldrich (120) state how the technique has a strong dependence on the sample preparation method, a short sample life-span and a low shot-to-shot reproducibility. Currently, there is no MALDI sample preparation technique which has been validated so the technique would not be admissible

in a court of law. Also, the short sample life-span may be a problem if a case needs to be revisited, and although it only requires a tiny amount of sample 0.5 µL; there may not have been a significant recovery from the original sample to allow for analytical repetition or multiple analyses. Larger volumes of samples may have been used on more traditional DNA sequencing methods reducing the possibility of analytical replicates. Another major disadvantage of MALDI from a forensic perspective is the problem of reproducibility. Due to the possibility of several different spectrum being produced from not only the same sample spotted onto the plate several times, but from several different points on one sample spot, one can never be sure which spectra is "correct" and therefore extreme care needs to be taken when interpreting the spectra and using this technique in a forensic context.

The problem of 'sweet spots' again means this technique may not be suitable for forensic work as certain positions on the preparation give better results than others due to the laser fluence threshold for the desorption/ionization varying from spot to spot. Laser fluence higher than the threshold value decreases mass resolution because the excess energy transferred to the analyte DNA causes metastable decay. Although this can be overcome by searching for the sweet spot this is not ideal and the analyst can never be sure if the spectrum obtained is a true picture of the sample. Another potential way to avoid the problem of 'sweet spots ' is to develop further the sample preparation technique for instance rapidly drying the sample preparation in a vacuum chamber, which could potentially yield smaller analyte-matrix crystals. Also, thin layer preparations may give less variation and thus better mass accuracy and precision. Surface preparation and the use of other types of anchor sample plates can help preparation and therefore help improve the spectrum (117).

UV-MALDI has been used previously for the analysis of complex mixtures containing large DNA fragments, such as allelic ladders that would be routinely measured using Capillary Electrophoresis. The difficulty of measuring numerous DNA fragments is in the amount of ion current available to simultaneously detect different oligonucleotides and the effects this implies on the detection sensitivity and resolution of these analytes. With increasing size of nucleic acids there is a loss of signal intensity and mass resolution and therefore this

technique is limited to nucleic acids smaller than 100 nucleotides. As the size of the DNA fragments in the extracts analysed in this work are unknown it is possible that they are too large and therefore the signal intensity is poor as this has been shown as in some cases the intensity was less than 10%. In contrast to this, masses as small as 9 Mr which is the difference between thymine and adenine are virtually impossible to resolve by convention MALDI ; which again, due to the unknown sizing of the DNA fragments in the samples, may have been a possible cause for the poor quality of the spectra (133).

A study by Williams (138) suggests MALDI should be used in conjunction with other techniques; this affirms what has been discussed in chapter 2 about the need for the analysis of the plant DNA present in soil to be accompanied by another technique. Taranenko (142) shows how MALDI can be used for 16s bacterial DNA, this might mean that for the analysis of DNA in soil, bacterial DNA may be more useful , and is a much more optimisable technique.

A possible way in which to overcome some of the problems of MALDI may be to employ a post-source-decay (PSD). PSD analysis is an extension of MALDI-ToF-MS that allows for the observation and identification of structurally informative fragment ions from decay taking place in the field-free region after leaving the ion source. In work by Brinkmalm and Brinkmalm (143) it was stated that although for more than a decade MALDI-ToF-MS has been favoured due to its speed, sensitivity and reliability it can be difficult to obtain structural information about the sample. However this can be overcome with the use of PSD. This is a technique also employed and discussed in a study by Papac *et al*. (137). Again it might not be useful as in this study we are not trying to gain information about just one molecule, we are trying to gain knowledge of what species are present. However, PSD could prove to be very useful if some information was known as to what species were present in the sample and MALDI was being used to confirm this.

MALDI-ToF-MS of DNA Extracts from Soil Summary Box

Not a useful tool for analysing general plant DNA.

Could be useful for species-specific amplified DNA.

Dependent upon good quality DNA samples; not always possible from soil.

No demonstrable reproducibility; spectra varied over the same sample.

As per more traditional methods of DNA analysis the quality of the extraction method is key to the success of the technique.

4.1 Introduction to chemical analysis of soil

This chapter looks at the chemistry of the soil samples, specifically, the strontium isotope ratios and the trace element profiles and their relationship with the underlying geology and land use.

4.1.1-a A Basic introduction to the chemistry of the earth

The earth is split into different areas. 99 % of the mass of the earth's continental crust is made up from 9 elements (see figure 42). Strontium, the element of interest in this research is the 16th most abundant, 0.0370 % weight of the earth's crust or 370 ppm.

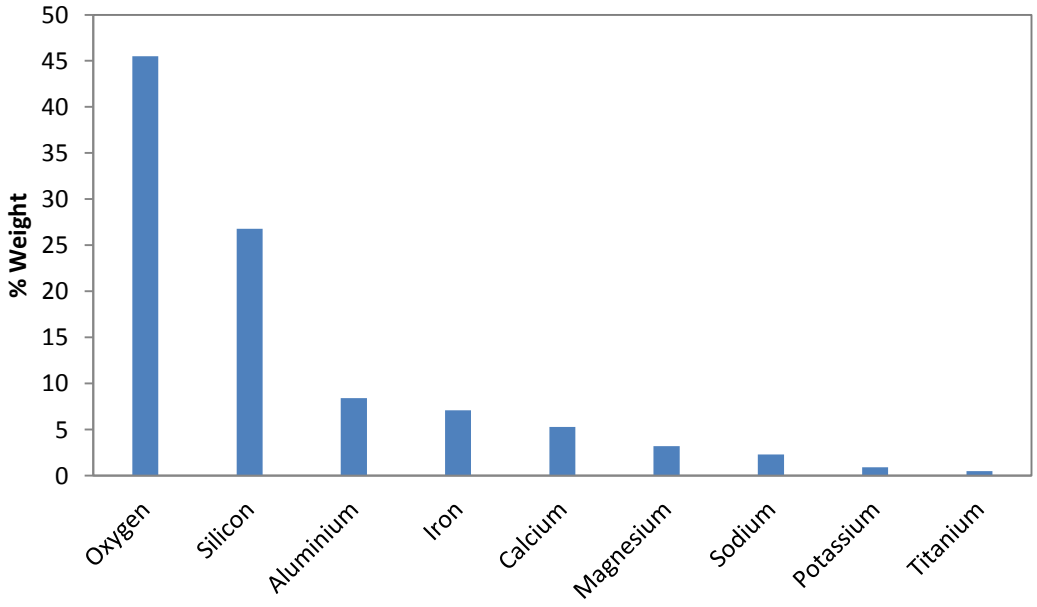


Figure 42: Percentage weight of the 9 most abundant elements which make up 99% of the earth's continental crust (144).

The major elements are those which are the most abundant in the earth's crust, the abundance of these elements is most frequently expressed in weight percent oxide. Silica (SiO_2) comprises 57.3% of the bulk continental crust and alumina (Al_2O_3) constitutes 15.9%. The next most abundant oxides are those of iron, calcium, magnesium, sodium and

potassium. However, it must be remembered that some soils may only contain low levels of these oxides. Trace elements are those that typically have an abundance of less than 0.1 %. The concentration of these elements is commonly expressed as parts per million (ppm) or micrograms per gram (μg/g). However, in some soils the abundance of certain elements classed as trace elements may reach up to seven percent. This high percentage of trace elements can be an indication of hydrothermally altered rocks (145).

4.1.1-b Goldschmidt's Rule of Substitution

Goldschmidt's Rule of Substitution encompasses the theory of one element replacing another if it is of a similar size and charge. The rules of substitution are:

1) The ions of one element can extensively replace those of another in ionic crystals if their radii differ by less than approximately 15 %.
2) Ions whose charge differs by one unit substitute readily for one another provided electrical neutrality of the crystals is maintained. If the charges of the ions differ by more than one unit, substitution is generally slight.
3) When two different ions can occupy a particular position in a crystal lattice, the ion with the higher ionic potential forms a stronger bond with the anions surrounding the site.
4) Substitution may be limited even when the size and charge criteria are satisfied, when the competing ions have different electronegativities and form bonds of different ionic character.

4.1.1-c Mineral transportation into vegetation

Plants use the dissolved minerals found in soil to build the complex molecules they need in order to survive and grow. Poor plant growth due to a lack of minerals in the soil is usually overcome by the addition of fertilisers to the soil. Fertilisers contain nitrogen, potassium, sodium and magnesium as these are the elements that are needed by plants in the greatest quantities. The use of fertilisers must be considered when analysing the elemental profile of soil since they can significantly influence the composition. This is considered in more detail in the discussion section of this chapter.

The concentration of minerals in soil is very low (often trace element level). They dissolve in water and move around the soil in solution. Root hair cells of plants are adapted to absorb the water in soil via osmosis. The cells have a large surface area, thin walls and are close to the xylem cells used for transporting water up the plant. Dissolved forms of minerals cannot be absorbed by osmosis or diffusion. The root hair cells have carrier molecules on their surface, these pick up minerals and move them into the cell against the concentration gradient, in a process known as active transportation.

4.1.2 Radioactivity, radiogenic isotopes and isotope ratios

The nuclei of unstable atoms can undergo spontaneous transformations that involve the emission of particles and radiant energy, this is known as radioactivity (146, 147). Radioactive decay via beta-decay causes a change in the atomic number and the number of neutrons, therefore an atom of one element transforms into that of another element. Radiogenic isotopes are isotopes that originate from the radioactive decay of another isotope e.g. $^{87}Rb \rightarrow {}^{87}Sr$ i.e. a radioactive parent material decays to form a stable daughter material. There are fundamental assumptions used about radiogenic isotopes that are studied in this research. That is that the rate of radioactive decay is independent of external influences such as temperature and pressure and that two isotopes of the same element are chemically identical thus, chemical processes cannot change, or fractionate, the ratio of two isotopes of the same elements (148). Naturally, there are exceptions to these assumptions but these occur under conditions not applicable to this research. Naturally occurring radiogenic isotopes have been moving through the Earth since its creation, and these isotopes can act as natural tracers.

The basic equation of radioactive decay is:

(1) $$\frac{dN}{dt} = -\lambda N$$

where *N* is the number of parent (radioactive) atoms, λ is the decay constant, defined as the probability that a given atom will decay in some time *dt* with units of time^{-1}.

Furthermore, the number of daughter atoms produced (D) is a function of the number of parent atoms present (N) and time (t), In general there are also some daughter atoms present initially, i.e. when t=0 (*D*0). Therefore equation 2 can be written:

(2) $$D = D_0 + N(e^{\lambda t}\text{-}1)$$

The decay of ^{87}Rb to ^{87}Sr is an example that can be explained using equation 2:

(3) $$87_{Sr} = 87_{Sr0} + 87_{Rb}(e^{\lambda t}\text{-}1)$$

For reasons discussed later, ratios of two isotopes rather than absolute abundances are usually measured. The ^{87}Sr is measured to a non-radiogenic isotope, ^{86}Sr, thus equation 3 is re-written as:

(4) $$\frac{87_{Sr}}{86_{Sr}} = (\frac{87_{Sr}}{86_{Sr}})_0 + \frac{87_{Rb}}{86_{Sr}}(e^{\lambda t}\text{-}1)$$

Similar expressions can be written for other decay systems e.g. lead. Equation 4 is a concise statement of strontium isotope geochemistry in that the ^{87}Sr /^{86}Sr ratio of a system depends on:

- The ^{87}Sr /^{86}Sr at time t = 0,
- The assumption that the ^{87}Rb /^{86}Sr ratio of the system is directly proportional to the Rb/Sr ratio
- The time elapsed since *t* = 0.

Forensic isotope geochemistry relies on the subtle differences in the isotopic abundance of elements to characterise a particular material. A unique isotopic composition can be derived from these different isotopic abundances and thus the determination of geographical location of a material can be derived. The two different types of isotopes that define the unique isotopic composition of a material are; isotopes that originate from the radioactive decay of one isotope, e.g., the beta-decay of ^{87}Rb to ^{87}Sr and the isotopes that do not undergo radiogenic decay. When discussing and analysing radiogenic isotopes, the daughter isotopes are the ones that are

measured to characterize the unique signature of the sources, e.g., Sr, and Pb (149). The use of isotopes in forensic science applications is relatively new in comparison to other techniques; this is mainly due to many of the techniques being time consuming and the sample preparation and instrumentation expensive and labour intensive. However, isotope ratios are significantly more sensitive tracers than elemental ratios or concentrations and therefore have a major advantage for a materials characterisation. These tracers can be evident in materials such as soils as used in this research, and also rocks, bones and teeth, feathers, hair, metal ores, pottery, and foods.

4:1.3 Rubidium and Strontium

Most igneous, sedimentary and metamorphic rocks contain rubidium and strontium in detectable amounts; however the concentration of these is usually less than 1 %. Rubidium (Rb) and strontium (Sr) are not major constituents of the common rock-forming silicate materials but strontium does form a carbonate (strontianite, $SrCO_3$) and a sulphate (celestite, $SrSO_4$). These minerals are found in certain sedimentary rocks, especially carbonates. Rubidium and strontium are intrinsically linked in that ^{87}Rb is a naturally occurring radioactive isotope and decays to ^{87}Sr, a stable isotope, via beta-decay. Therefore the level of ^{87}Sr in a rock containing rubidium increases continuously as a function of time. Radiogenic ^{87}Sr can be used as a geological tracer to study certain geological processes. Some of the properties of rubidium and strontium are shown below in table 10 and their position in the periodic table are shown in figure 43.

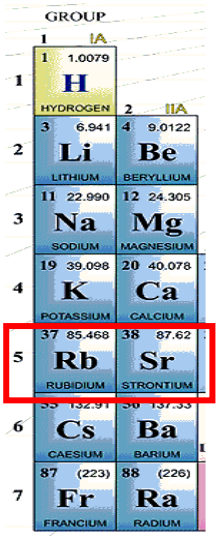


Figure 43: The position of Rubidium and strontium in the periodic table as highlighted.

Table 10: Properties of rubidium and strontium

	Rubidium	Strontium
Atomic Number	37	38
Atomic Weight (based on ^{12}C)	85.467	87.62
Ionic Radius (Å)	1.48	1.13
Radius Ratio (O_2 = 1.40 Å)	1.06	0.81
Co-ordination number in ionic crystals	8.12	8
Electronegativity (Pauling)	0.8	1.0
% ionic character of bond with O_2	87	82

4:1.3a Chemistry of Rubidium

Rubidium is in group 1A of the periodic table, which also includes hydrogen and the alkali metals lithium, sodium, potassium, francium and cesium. Rubidium has a single valence

electron in the *s* orbital outside of the stable electronic configurations. This electron is readily removed to form ions with a +1 charge. Rubidium has a low electronegativity and therefore forms strong ionic bonds with non-metallic elements such as oxygen and halogens.

Rubidium was discovered in 1861 by Bunsen and Kirchoff in the mineral lepidolite. The distribution of rubidium in nature is governed by the fact that Rb^+ ion (ionic radius of 1.48 Å) is small enough to be admitted into K^+ sites (ionic radius =1.33 Å) in all of the important rock-forming minerals that contain potassium. Rubidium is never sufficiently concentrated to form its own minerals but is widely dispersed as a trace element in potassium-bearing minerals.

The principle carriers of rubidium in igneous and metamorphic rocks are the micas (biotite, muscovite and lepidolite) and the potassium feldspars (orthoclase and microcline). Minerals in pegmatites may contain much higher concentrations of rubidium than the same minerals occurring in ordinary igneous or metamorphic rocks. Lepidolite from lithium-bearing pegmatites may contain several percent of rubidium. The concentration of rubidium in plagioclase feldspars is low due to the Rb^+ ion being too large to replace the Na^+ ion (ionic radius (0.95 Å). Rubidium also occurs in other rock forming minerals such as amphiboles and pyroxenes in a concentration commonly less than 10 ppm by weight.

Rubidium has two naturally occurring isotopes; ^{85}Rb and ^{87}Rb. The ratio of ^{85}Rb to ^{87}Rb is 2.593 ± 0.0017, and the corresponding isotopic abundances are, 72.166 ± 0.0132 % for ^{85}Rb and 27.835 ± 0.0132 % for ^{87}Rb. ^{87}Rb has a half-life of 49 billion years (more than three times the estimated age of the universe, therefore on a laboratory scale it can be considered stable. It is a widely held view that all naturally-occurring rubidium has the same isotopic composition, regardless of age, the mineral it is found in or its geochemical history (150).

4:1.3b Chemistry of Strontium

Strontium is in group 2A of the periodic table, which also contains the alkaline earths, beryllium, magnesium, calcium and barium. Strontium has two valence electrons in an *s* orbital outside of stable gas configurations and readily forms ions with a +2 charge.

Strontium has a low electronegativity of -1.0 and therefore forms ionic binds with non-metallic elements including oxygen.

The distribution of strontium in rocks is controlled by the extent to which Sr^{2+} can substitute Ca^{2+} in calcium-bearing minerals and the degree to which potassium feldspar can capture Sr^{2+} in place of K^{+} ions. The ionic radius of Sr^{2+} is 1.13 Å (table 10) which is approximately 15 % greater than that of Ca^{2+} (0.99 Å), and the Sr^{2+} radius relative to O^{-2} is 0.81 whilst Ca^{2+} is 0.71. As a result of this Sr^{2+} favours an eight-fold co-ordination whereas Ca^{2+} can occupy both six-fold and 8-fold co-ordinated lattice position. This means strontium acts as a dispersed trace element in igneous rocks. However, it can be concentrated relative to calcium which enables strontium to form its own minerals in carbonate rocks and hydrothermal deposits. There have been several strontium minerals named, but only two of these; celesite and strontianite are of any importance.

The principle carriers of strontium in igneous rocks are apatite and plagioclase feldspar where Sr^{2+} can replace Ca^{2+} ions. Pyroxenes have a low strontium concentration due to calcium being bonded to six oxygen atoms therefore the lattice being too small for the Sr^{2+} ions. Potassium feldspar can capture Sr^{2+} ions in place of K^{+} ions; this substitution is coupled with the replacement of Si^{4+} by Al^{3+} in the silica tetrahedral space in order to maintain electrical neutrality. In micas, the capture of Sr^{2+} for K^{+} is not favoured, as potassium in micas has a 12-fold co-ordination, making this site too large for strontium.

During magma crystallisation, strontium initially enters calcic plagioclase by substitution for Ca^{2+}. If differentiation progresses and potassium-feldspar begins to form, Sr^{2+}, ions are captured from the residual liquid by K^{+} sites. Consequently, in many series of differentiated igneous rocks the strontium concentration decreases with increasing degree of fractionation of the magma. If pyroxene or olivine forms before plagioclase, the strontium concentration of the residual liquid may rise to a maximum before decreasing.

Strontium has four stable isotopes, ^{88}Sr, ^{87}Sr, ^{86}Sr and ^{84}Sr and 14 short-lived radioactive isotopes have been artificially made. The isotopic composition of strontium in nature is not constant, it is dependent upon the Rb/Sr ratio of the sample from which the strontium is extracted and on the time period it has been associated with rubidium as discussed in section 4:1.2 (150). The relative abundance of ^{87}Sr is commonly expressed as the atomic ratio $^{87}Sr/^{86}Sr$. Nier (151) reported that the value for $^{86}Sr/^{88}Sr$ ratio is 0.11940 and this value is still used as a reference to eliminate the effect of isotopic fractionation from measured

values of the $^{87}Sr/^{86}Sr$ ratio in geological materials. In order to measure the differences in the ^{87}Sr abundances in different rocks or soil samples the ^{87}Sr is normalised to the non-radiogenic isotope, ^{86}Sr. This is used because the number of atoms in ^{86}Sr remains constant as ^{86}Sr is stable and not produced by the decay of a naturally occurring isotope of another element (150). Beard (152) reports that using the $^{87}Sr/^{86}Sr$ ratio instead of absolute ^{87}Sr abundances enables the removal of variations in the ^{87}Sr that are a reflection of the natural variations in total strontium. Therefore, by using the $^{87}Sr/^{86}Sr$ ratio it allows for the variations in the ^{87}Sr abundances which are a function of the beta-decay of ^{87}Rb to ^{87}Sr to be isolated.

4:1.3c Relationship between Rubidium and Strontium

The general chemical coherence of rubidium with potassium and strontium with calcium is highlighted tables 10 and 11. Table 11 also highlights the relationship between rubidium and strontium. Generally, the rocks that have a higher rubidium concentration have a lower strontium concentration and vice versa. This means that the Rb/Sr ratios have a wide range which allows the differentiation of rock types (144, 153). It is this Rb-Sr system that allows for a crude transformation from a geological map of bedrock types and ages into a map of the measured and expected $^{87}Sr/^{86}Sr$ ratios; a parameter hoped to be developed in this research.

Table 11: The properties, and thus, the relationship between rubidium and potassium and strontium and calcium in different rock types.

Rock Type	**Concentration / (ppm)**			
	Rubidium	Potassium	Strontium	Calcium
Ultramafic	0.2	40	1	25000
Basaltic	30	8300	465	76000
High-Ca Granite	110	25200	440	25700
Low-Ca Granite	170	42000	100	5100
Syenite	110	48000	200	18000
Shale	140	26600	300	22100
Sandstone	60	10700	20	39100
Carbonate	3	2700	610	302300
Deep-sea carbonate	10	2900	2000	312400
Deep-sea clay	110	25000	180	29000

The $^{87}Sr/^{86}Sr$ ratio is defined as a function of the initial $^{87}Sr/^{86}Sr$ ratio, the initial Rb/Sr ratio and time as can be seen in equation 4. This therefore shows that strontium isotopic signatures have evolved over geological time, a principle that will be used in this research; as it is expected that the strontium isotope ratios will show some correlation to the geology from which the samples were collected. $^{87}Sr/^{86}Sr$ isotope ratios are currently used as geological tracers and are also used in petrogenesis (the study of the origin of mineral rocks). The type of rock and therefore the age of the rock is proportional to the $^{87}Sr/^{86}Sr$ ratios; rubidium poor and young rocks have low $^{87}Sr/^{86}Sr$ ratios whereas rubidium rich or older rocks have considerably higher $^{87}Sr/^{86}Sr$ ratios (146). The $^{87}Sr/^{86}Sr$ reflects the average of all of the strontium that has contributed to the sample (154).

4:1.3d Typical strontium isotope ratios / Strontium in Soil

Strontium isotope ratios for granites from the various continents are highly variable, ranging from 0.709 to 0.890 and more. The strontium ratios for basalt are much more uniform, ranging from 0.7020 to 0.7070 and for oceanic basalts from 0.7022 to 0.7045. Analytical precision on the $^{87}Sr/^{86}Sr$ ratio is at least $1.5x10^{-4}$ (relative) and therefore on ratios close to 0.7 the uncertainty is better than 10^{-4} (absolute). This means that it is possible to distinguish between a sample with a ratio of 0.7022 from one with a ratio of 0.7021. (155) As has been previously discussed the age of the rocks results in differing strontium isotope ratios, this is shown in the Quaternary Basalts from 1.6 million years ago to the present in areas of Iceland which have $^{87}Sr/^{86}Sr$ ratios from approximately 0.7020 to 0.7060, and in the Palaezoic granites of Scandinavia which are 245 to 270 million years old and have higher $^{87}Sr/^{86}Sr$ ratios from approximately 0.7110 to 0.7800 (146).

Strontium isotopic signatures are conveyed from the eroding geological materials into the soil and subsequently into the food chain. However, there are additional, non-geological sources of strontium in the biosphere which can contribute to the strontium isotopic composition of the soil. The major contributions to strontium in soil are erosion, mineral weathering, ground and stream water, atmospheric deposition, and also the addition of fertilisers to the soil (154).

The expected strontium isotope ratios in Norfolk are not expected to be diverse however based upon the analytical precision stated above, it is thought that they will be able to be distinguished from each other. Data from archaeological samples show a strontium isotope range from 0.7066 – 0.7140 (taken from tooth enamel but expected to also be shown in geology) as mentioned by Evans *et al*. The map of biosphere strontium isotope variation in the paper shows a smaller variation in Norfolk in comparison to other areas of the UK (156).

4:1.3e Bioavailable Strontium

The term bioavailability is often used to describe the fraction of a substance that is actually taken up by organisms and can be quantified by the effect of the substance on the organisms (157). Sillen *et al*. (158) have shown large differences in the values of bioavailable strontium and whole soil strontium and suggest that any future use of

strontium isotope ratio relationships should use biologically available strontium as it shows a good relationship with the samples. It is also highlighted that the strontium isotope ratios of soils are generally higher than those of plants as is shown in figure 44. In this paper it is also suggested that the geological strontium isotopic composition could be explained by the presence of a significant dust component, from differential weathering; i.e. highly radiogenic muscovites and potassium feldspars present in rocks and whole soils may minimally contribute to available soil strontium when compared to less radiogenic components.

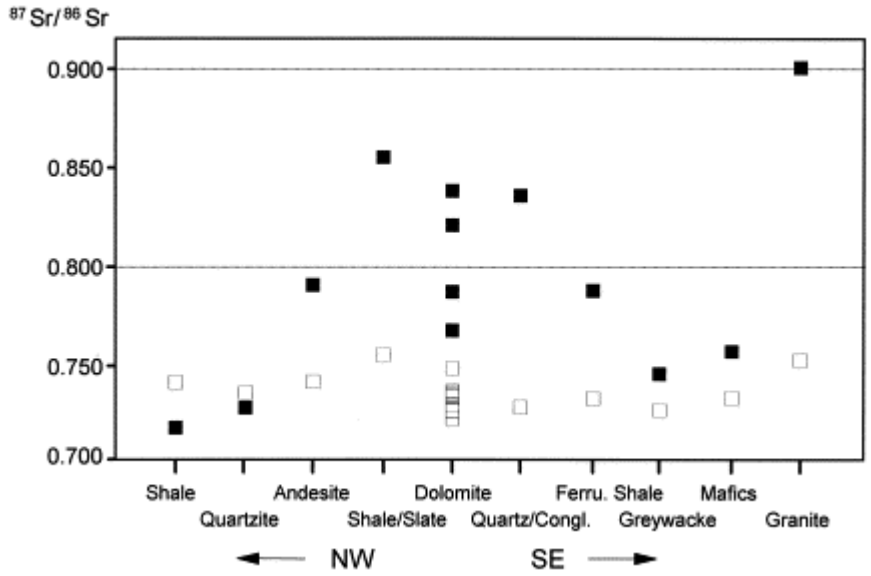


Figure 44: Typical strontium isotope ratios from plants and soils from an area in Swartkrans (soils shown by filled squares and plants by hollow squares). Whole soils are generally more radiogenic than the plants that are growing on them. Taken from SIllen *et al.* (158).

4:13f Fractionation

In order to eliminate the effect of natural and instrumental isotopic fractionation the $^{87}Sr/^{86}Sr$ ratio is normalised against the $^{86}Sr/^{88}Sr$ ratio of 0.1194 (see MC-ICP-MS section) (150). This normalisation allows the strontium isotope ratios found in plants, and the animals eating the plants, to be related to the bioavailable part of the mineralised

strontium from the underlying lithology, which is the sum of the individual minerals and geological areas (159).

4:1.3g Strontium Isotope ratios in this research

It is the aim of the project to use geology maps to show the correlation between the geological units and the bioavailable $^{87}Sr/^{86}Sr$ isotope ratios. Whether or not this correlation occurs is the basis for the research, and if there is a correlation if it can be subsequently used to predict the origins of the soil samples. In order to investigate the influence of any fertilisers added to the soil upon bioavailable $^{87}Sr/^{86}Sr$ ratio, the $^{87}Sr/^{86}Sr$ ratios will also be plotted on land-use maps to see any correlations.

4.1.3h Work where strontium ratios have been or are currently used

Strontium isotope ratios have been extensively used in work carried out by the TRACE project (for more information on TRACE see chapter 2), where they were used to link various food stuffs including chicken, garlic, mineral waters, olive oil and honeys to geologies. Strontium isotopes from soil were also measured. This work showed correlations between the strontium isotope ratios of various foodstuffs and the underlying geologies of the areas where they originated from.

The use of strontium isotope ratios for mapping has been highlighted in work by Evans *et al.* (156). The paper states how a number of different materials can be used to characterise the strontium isotope characteristics, specifically the study uses present day water, dilute acetic acid soil lechates, present day plant material and river water samples. The study suggests that in clay-carbonate terrains direct sampling of plant material is the most reliable way of deriving the isotope composition of the biosphere in a particular area. Plants, and soils, have the advantage of being readily available and can be precisely sourced. However, a disadvantage is that they provide the average value for only a few centimetres of ground from which they derived their nutrients. Strontium concentrations broadly follow a pattern of diminishing concentration with rising trophic level but such comparisons must be restricted to similar geographical and climatic regions.

4:1.4 Major and Trace Elements

The distinctive geological areas of East Anglia (see introduction) should give rise to soils with a particular chemical composition and it should be possible to disyinguish between these soils based on their elemental profile. As previously stated, the major elements are those that are most abundant in the earth's crust (> 1 % abundance in a sample), minor elements are those that generally have an abundance of 0.1 to 1.0 % and trace elements are those that have < 0.1 % abundance. It should be noted, that an element, which is a major component in one sample, could be a minor component in another, however, this can aid in the discrimination of samples.

There have been many different methods developed and used to determine the concentration of both major and trace elements in soils. These methods include Atomic Absorption spectrometry (AAS) (160), X-ray Fluorescence Spectroscopy (XRF), Neutron Activation Analysis (NAA) (160) and Inductively Coupled Plasma Spectrometry (ICP). XRF is a popular technique in forensic geochemistry and uses a beam of primary radiation produced in an x-ray tube to excite a secondary x-ray emission from the sample. The emitted x-rays have a characteristic energy for each element, allowing for quantitative and qualitative elemental abundance analysis.

In a study by Rawlins and Cave (161) elemental abundance data for 19 elements, obtained by XRF analyses, was used to investigate the extent to which samples derived from the same geological parent material could be distinguished from each other. For 13 of the 19 individual elements they found that on average, it was possible to discriminate between more than 80% of the samples within parent material groups, but when using the elements in combination, more than 99.8% of samples could be discriminated from one another. The most discriminatory element was found to differ between parent material groups, sometimes being obviously influenced by the type of bedrock e.g. in a region of Cretaceous Chalk, over 96% of samples could be discriminated using the abundance of CaO. In general, it was found that the most useful elements for discrimination were, Mn (achieved 93% sample discrimination), Mg (90%), Rb (89%), Zr (89%), Sr (87%), Ni (87%) and Zn (86%). It was also demonstrated that it was possible to obtain a multi-element geochemical signature for soils from each parent material by calculating the Mahalanobis distances between samples. The Mahalanobis distance is an extension of the Euclidean in that it also takes into account the correlation between variables when computing distances. The soil

samples in this study were collected by the British Geological Survey (BGS) as part of the Geochemical Baseline Survey of the Environment (G-BASE).Therefore, there are significant differences between these samples and methods of collection, and those which might be expected in forensic work and those implemented in this research, and therefore care must be taken when comparing data. The predominant difference in the sampling methods is that a far larger amount of sample is taken during a soil survey than would generally be available for forensic examination with errors in sampling increasing as the amount of sample taken to represent a specimen decreases. It has been shown previously that approximately 1 gram of rock powder, which is made up of the same minerals that later form soils, was required to accurately determine major element concentrations that were representative of the bulk from which the samples had been taken, and an even larger amount was required for trace elements (162).

Recently, the use of inductively coupled plasma – atomic emission spectrometry (ICP-AES), optical emission spectrometry (ICP-OES) and mass spectrometry (ICP-MS) have become more prevalent in geochemical analysis. These techniques allow analysis of a very small amount of sample and the simultaneous determination of a large number of elements with high precision and good sensitivity. Pye and Blott (163) conducted a study comparing the elemental composition of control soils and suspect samples taken from a wellington boot to investigate the ability of discriminating between them in this way. They determined the abundance of 49 elements by ICP-OES and ICP-MS and demonstrated that the abundance of certain elements e.g. Ce, and ratios of particular pairs of elements e.g. Ce/La, can help discriminate between sets of samples. They discuss the effects of particle size on the chemical composition, highlighting that some elements e.g. silicon are more abundant in the coarser fractions and that the trace metals are generally more concentrated in the finer fractions. They state that in their experience, analysis of the < 150 μm fraction allows adequate discrimination between samples and does provide an indicative measure of sample composition.

In addition to the geological trends within elemental abundance data, there is also a certain amount of variation resulting from anthropogenic factors. Archaeologists are beginning to use multi-element analysis for site prospection, in the determination of land use and the

evaluation of ancient human activities. Studies in Guatemala have shown that elevated levels of barium, phosphorous and manganese were associated with areas of organic waste disposal, whilst mercury and lead were associated with craft production areas (164). In addition, elevated strontium and calcium levels have been found in field areas and high concentrations of potassium, rubidium and thorium have been shown to be indicators of settlements on small farms (165). Care should be taken when examining elemental abundance data from soil surveys as seemingly anomalous values could occur as a result of anthropogenic effects on an isolated area.

Webster (166) has stated that due to the incompleteness of our understanding of the factors which determine the concentration of an element in the soil at a specific location in the natural landscape, it means that the properties of soil appear to be random. It should also be taken into consideration that it is very rare that two soil samples show perfect chemical similarity (as discussed in the introduction) due to inherent spatial variation of soil samples as well as the errors associated with both sample preparation and instrumental measurements (163). Therefore in this research the trace element data will be used alongside the strontium isotope ratios and the biological parameters to try to distinguish between the different soil samples and determine any relationships between the soil samples and underlying bedrock geology.

It should also be considered that the literature has questioned the use of major and trace elements for forensic analyses; this is due to the often mixed provenance of forensic soil samples. However for this research the soils samples analysed have a single provenance, and in the interpretation of the data thought is given to the fact that in a criminal investigation the samples may have a mixed provenance. Careful consideration must also be given to the homogenisation of soil samples. The homogenisation of soil samples prior to analysis can mean that it is impossible to assess whether the results derived exclude the sample from having the same or similar provenance as the crime scene, or whether the exclusion is a false negative result (21, 22, 40, 167).

4.1.5 British Geological Survey data

4:1.5a British Geological Survey (BGS)

The British Geological Survey has provided soil parent material data, superficial deposit data and bedrock material data.

4:1.5b Bedrock Geology Theme

The bedrock geology is formerly known as the solid geology and is the main mass of rocks forming the earth. Bedrock geology is present everywhere; in some areas it is exposed at the surface in outcrops and in other areas it is concealed below water and surface deposits. The bedrock geology is formed over long geological time periods and includes many lithologies, which are often classified on the origin of the rocks i.e. metamorphic, igneous, and sedimentary (168).

4:1.5c Superficial Deposits Theme

The superficial deposits are formerly known as drift. These are the youngest geological deposits, which have been formed during the most recent period of geological time, the Quaternary Period and rest on the older deposits, the bedrock (169).

4:1.5d Soil Parent Material Model

The parent material is a name for a weathered rock or deposit from and within which a soil has formed. The parent material provides the basic foundations and building blocks of soils and influences their chemistry, structure, texture and drainage (170).

The BGS have developed a digital parent material map (PMM) at a 1:50 000 scale. This map details the distribution of the physiochemical properties of the weathered and unweathered parent materials of the United Kingdom. The PMM facilitates spatial mapping of UK soil properties, identifies soils and landscapes sensitive to erosion, provides a national overview of soil resources and develops a better understanding of weathering properties and processes. (170) It is an extraction from this map that has been used as a basis for the mapping in this project.

4:1.6 How the data will be used

Data provided by the BGS has enabled the geology and superficial deposit maps to be created using ESRI® ARCMap 9.2. These maps have been used as a base layer for the maps to show the $^{87}Sr/^{86}Sr$ data and were also used initially when determining the location of the sample sites (see figure 7). Trace element data was collected to use alongside the $^{87}Sr/^{86}Sr$ data to try to provide a greater degree of certainty when predicting from where a soil sample originates. The trace element data has also been used to determine if any individual elements, pair of elements or a suite of elements can be used to determine the origin of a soil sample.

4:2 Methods

4:2.1 Reagents

All preparations and separations were conducted inside an Airclean 4000 workstation (Airclean systems). All solutions were prepared using high purity deionised milli-Q water by passing water through a laboratory-reagent grade water system (Ultrapure (Type 1) Millipore, USA). Analytical grade nitric acid (Merck) was purified by distillation. All samples were weighed on an analytical balance (Mettler Toledo AG245 Mettler, Switzerland). A strontium selective extraction resin (Eichrom, USA) was used for separations. Separations were performed using polyethylene (PE) micro-columns (1 mL volume) loaded with 0.2 mL of the resin. Soil preparation was carried out, with minor amendments as indicated, according to TRACE standard operating procedure (SOP) which employs the German DIN 19730 (1997) extraction methods. All calibration standards and final solutions for analysis were prepared in 15 mL PP test tubes (Cellstar, cat no. 188285). Eppendorf research pro pipettes were used for all solution transfers. Samples were only handled wearing disposable (PE) gloves to minimize secondary contamination.

4:2.2 Extraction of bioavailable fraction

This procedure used the soil fraction of <2 mm as it has been shown that results of analysis on this fraction is not statistically different from the <150 μm fraction (19) and 1mol/L ammonium nitrate solution as set out in the TRACE protocol. The ammonium nitrate solution was made by dissolving 80.04 g of ammonium nitrate per litre of water and adding 10 μL of 1000 ppm Rhenium internal standard. 50 mL (±0.5 mL – this seems to give a large source of potential error but was the stated protocol) was then added to 20 g of the soil sample (± 0.2 g). The samples were then shaken for two hours at room temperature to ensure that all of the solid components are suspended. The closed bottles then were left to stand for one hour to allow the solid constituents to settle. An aliquot of the supernatant was then filtered through a 0.45 μm PTFE membrane filter. The first part of the filtrate (approximately 1mL) was discarded and the remainder collected in a 15 mL falcon tube for storage. At this stage the filtrate was usually a clear solution and was stabilised by adding 10 μL of concentrated nitric acid per mL of filtrate.

4:2.3 Strontium columns

4:2.3.a Resin

In order to reduce molecular and isobaric interferences, the strontium in the samples needed to be separated from the soil matrix. This was done by ion chromatography using a strontium-specific resin. The resin used was Eichrom's Strontium Resin Extractant System; a 1M 4,4'(5')-di-t-butylcyclohexano 18-crown-6 ether in 1-ocatnol sorbed onto an inert chromatographic support (see figure 45) (171).

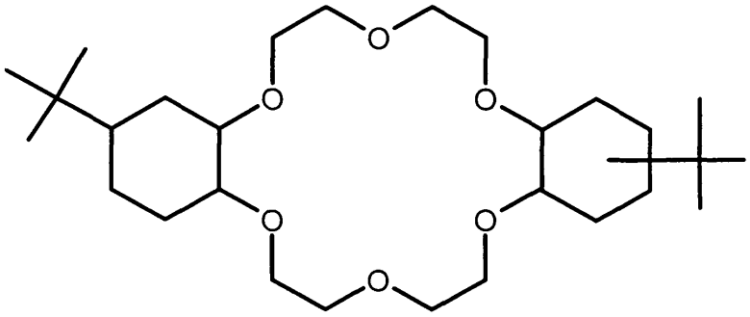


Figure 45: The chemical structure of the Eichrom strontium resin extractant system.

The strontium specific resin is capable of selectively retaining the strontium over a range of nitric acid concentrations; this allows for the separation of strontium from the potentially interfering calcium, radium and barium, and other matrix components. The advantages of this strontium selective method over conventional ion-exchange techniques includes high selectivity and high recoveries, overall it is simple to employ and only uses very small reagent volumes of a single acid, nitric.

4:2.3.b Method

For the purification of the sample for Sr isotope analysis nitric acid cleaned syringes were prepared. The columns were packed with acid cleaned glass wool and 500 µL of resin was pipetted on top. Another layer of glass wool was packed into the column on top of the resin and pushed down firmly as shown in figure 46. The columns were then washed with 1.8 mL of milliQ water. In order to prepare the columns for the extraction 1.8 mL of 3M nitric acid was pipetted through. This was followed by the samples (3M concentration) and finally the columns were rinsed with 3.2 mL 3M nitric acid which washed out everything except the strontium. In order to collect the sample savilex cups were placed under the columns and the columns rinsed with 1.8 mL milliQ water and the fraction for strontium isotope analysis collected. As the samples were to be first analysed using the Agilent ICP-MS, (they) were dried on a hotplate and then made up to 1.8 mL in 2% nitric acid.

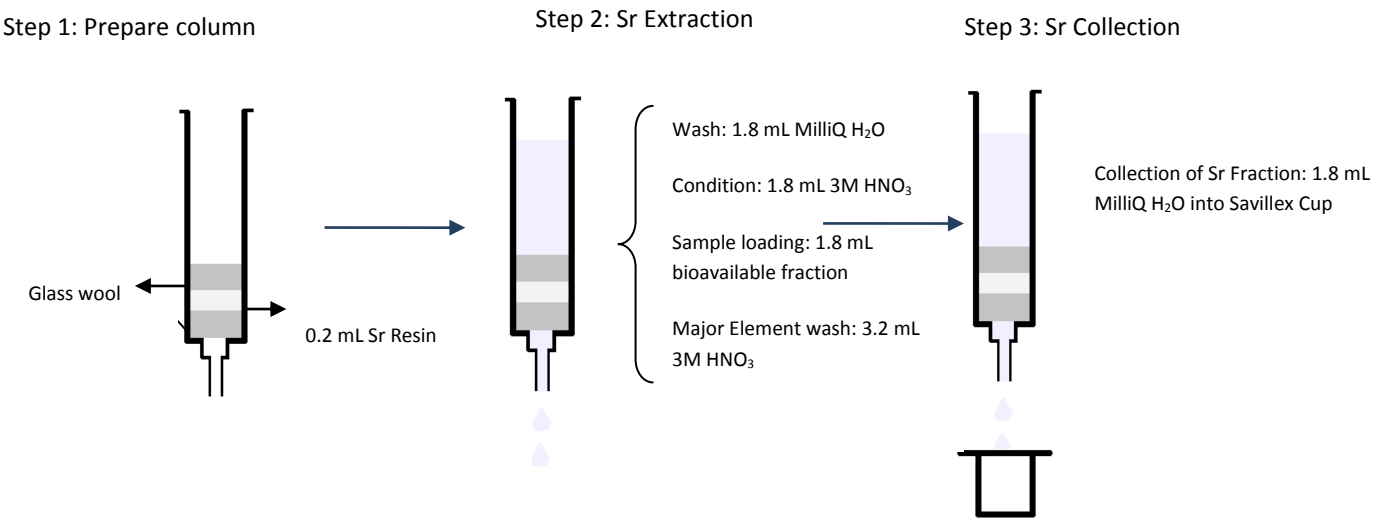


Figure 46: Schematic of the preparation of the columns and subsequent conditioning and wash steps for the extraction of the strontium fraction.

4:2.4 Sample Preparation for ICP-MS analysis

Samples collected from the Sr extraction were dried down and then taken up in 2 % HNO_3 (3 mL) and gently heated on a hotplate in order to get all the sample back into solution. 1 mL of this solution was diluted further with 2% ultrapure HNO_3, containing Re, Ge and Pt internal standard, to make up a 5 mL solution. The masses of all solutions were accurately recorded in order to allow for accurate calculations of the dilution factor for each solution.

The soil samples were extracted as described above and then prepared for trace element analysis. 0.4 mL of the extracted sample solution was diluted to 10 mL using 2 % HNO_3 containing internal standard GeRhPt (a 25 times dilution).

4:2.4a Standards for ICP-MS analysis

The ICP-MS was calibrated using three sets of multi-element standards. A thirty element ICP multi-element standard VI solution (Merck) was diluted to a range of six concentrations from 100 – 20,000 times dilution with ultra-pure 2 % HNO_3 containing Ge, Rh and Pt internal standard. A seventeen rare earth elements solution (Claritas) was diluted to a range of four concentrations from 1000 – 20,000 times dilution with

ultra-pure 2 % HNO_3 containing Ge, Rh and Pt internal standard. A seven high field strength elements solution (laboratory prepared) containing Zr, Sb, W, T, Sn, Hg and Cs was diluted to a range of four concentrations from 1000 – 20,000 times dilution with ultra-pure 2 % HNO_3 containing Ge, Rh and Pt internal standard. Blanks were prepared by diluting ultra-pure 2 % HNO_3 (1 mL) with ultra-pure 2 % HNO_3 containing Ge, Rh and Pt internal standard (9 mL). The elemental composition and masses of each batch of standard solutions were dated and tabulated.

4:2.5 Sample Preparation for MC-ICP-MS Analysis

Samples collected from the Sr elution were dried down and then taken up in 2 % HNO_3 (3 mL). The samples were then analysed by ICP-MS to determine the Sr concentration present in the samples. The concentration data was then used to dilute the samples down to 100 ppb using HNO_3 to analyse the samples by MC-ICP-MS to allow for Sr isotope ratio determination.

4:2.6 Instrumentation - Agilent (ICP-MS) – Strontium concentration and Trace Element Data

The first commercially available ICP-MS was by Perkin Elmer in 1983. ICP-MS has been used for the determination of major, minor and trace elements in almost every analytical field since the first mass spectra acquired from a plasma was published in 1975. A quadrupole ICP-MS can measure most of the elements of the periodic table including trace elements with detection limits at or below ppt range (172). The use of ICP-MS is a widely accepted technique due to its low detection level and its ability to scan more than seventy elements in order to determine the composition of unknown samples. Within this work all experiments have been carried out on an Agilent 7500ce ICP-MS coupled with an octopole reaction system (Agilent, USA) fitted with an automated Cetac ASX-510 autosampler (Cetac, USA) as shown in figure 47

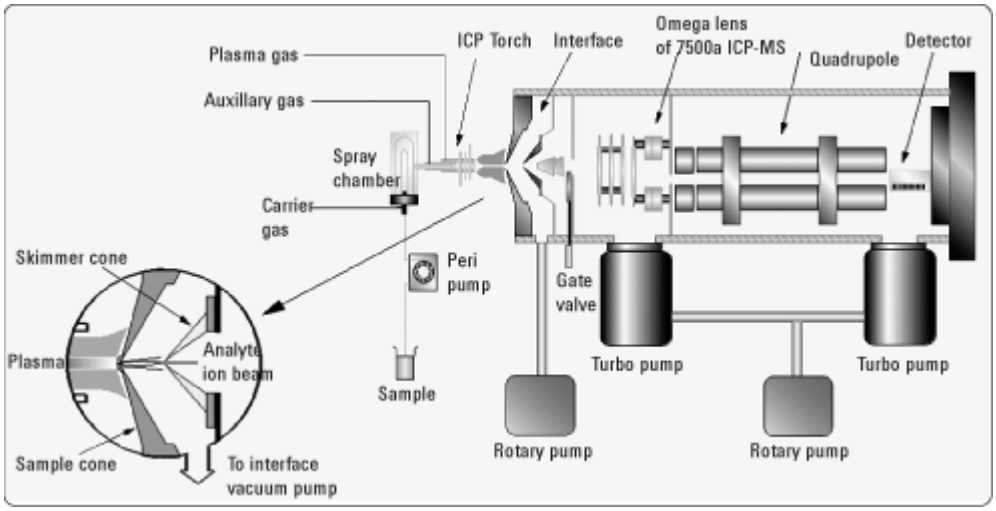


Figure 47: Schematic of the Agilent 7500 series ICP-MS as used in this research (172)

4:2.6a Sample Introduction

The samples were introduced into the ICP-MS via an autosampler; the sample was introduced as an aerosol, which was produced by passing the liquid sample through a pneumatic nebulizer. Excess sample and larger aerosol droplets were removed from the gas stream (they collided with the chamber walls) via a spray chamber and the remaining smaller droplets of sample were transferred into the central channel of the argon plasma as shown in figure 48. The Agilent was fitted with a Scott-type double pass spray chamber which was manufactured from high-purity quartz. The spray chamber temperature was maintained via a thermoelectric (Peltier) device. This allowed the prevention of signal drift due to fluctuations in room temperature and reduction of large solvent loading onto the plasma, which leads to a higher plasma temperature, reduction of oxide interferences and assistance in matrix decomposition.

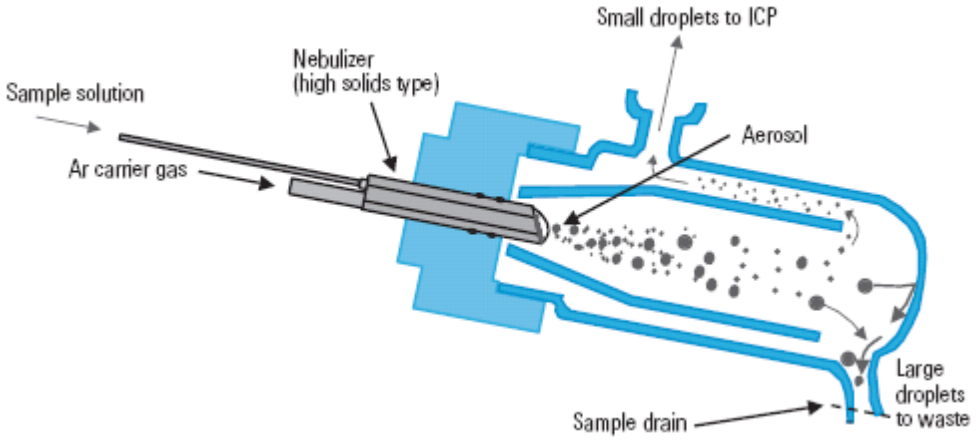


Figure 48: Schematic of a spray chamber (172).

4:2.6 b Ion Generation in the ICP

From the spray chamber the fine aerosol was transported into the plasma torch where positively charged ions were generated. The plasma was generated by passing argon (Ar) through the torch consisting of a series of concentric quartz tubes, wrapped at one end by a radio frequency (RF) coil which induces a magnetic field, as shown in figure 49. The magnetic field caused collisions between Ar atoms and free electrons to produce ions and more electrons until a stable, high temperature plasma was formed. This stable plasma (at a temperature of up to 10,000K) caused aerosol droplets to rapidly dry, decompose, vaporize and atomize before finally being ionised due to the absorption of energy from the plasma, releasing one electron to form a singly charged ion. The resulting ion beam then continued to the ion filters and quadrupole where the sample ions were separated by their mass to charge ratio as shown in figure 50.

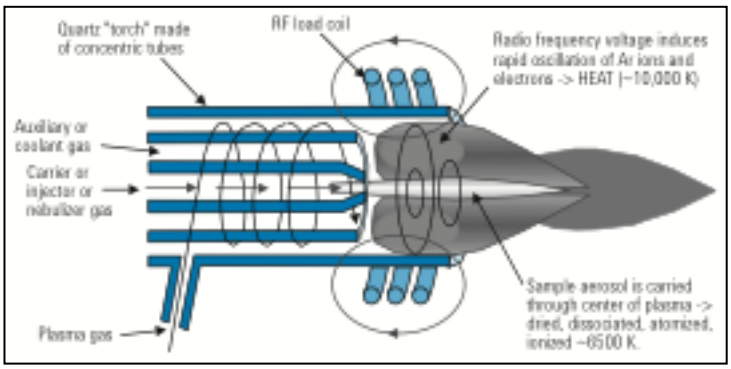


Figure 49: Schematic of the ICP-MS plasma torch based on the Fassel design (172)

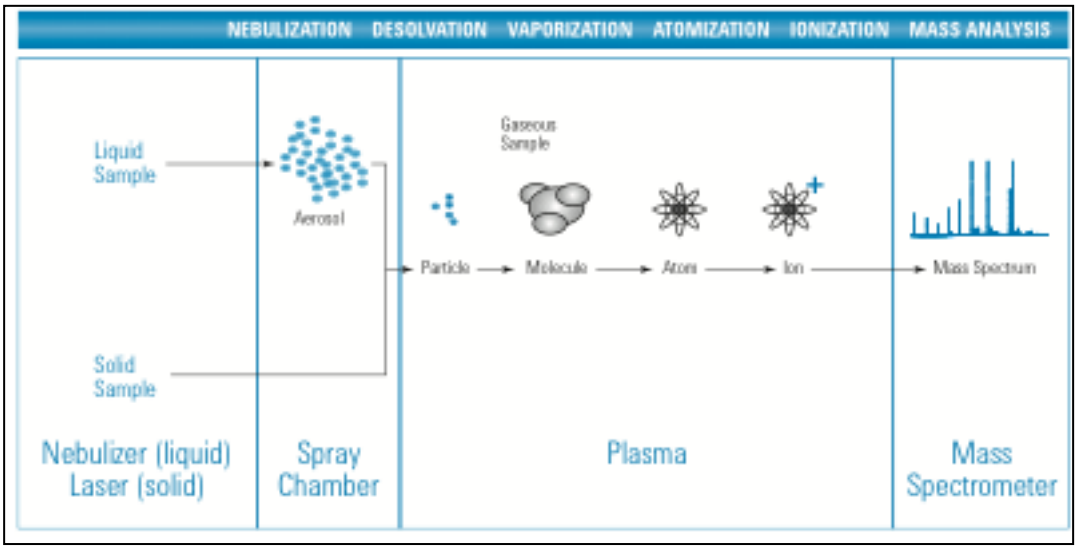


Figure 50: Representation of the processes involved in ICP-MS from sample introduction to mass analysis. (172)

4:2.6 c Interface

Once the positively charged ions had been produced in the plasma; they were directed to the mass spectrometer via the interface region. The plasma/vacuum interface consisted of two cones, a sampling cone and a skimmer cone (see figure 51). Firstly the analyte ions were extracted into the first vacuum stage by passing through the sampling cone into the interface. Secondly, the ions passed through the skimmer cone, which acted as a differential aperture between the interface and the intermediate vacuum stage. The purpose of the interface region was to extract a representative sample of the ion population and transfer this efficiently to the high vacuum regions where the ion focusing, mass separation and detector systems were located. Another purpose is to maintain the high vacuum within the mass analyser; a low vacuum would mean short free paths for ions therefore more collisions and a lower sensitivity. The function of the skimmer cones is to remove most of the solvent; however, as a result of this fractionation occurs.

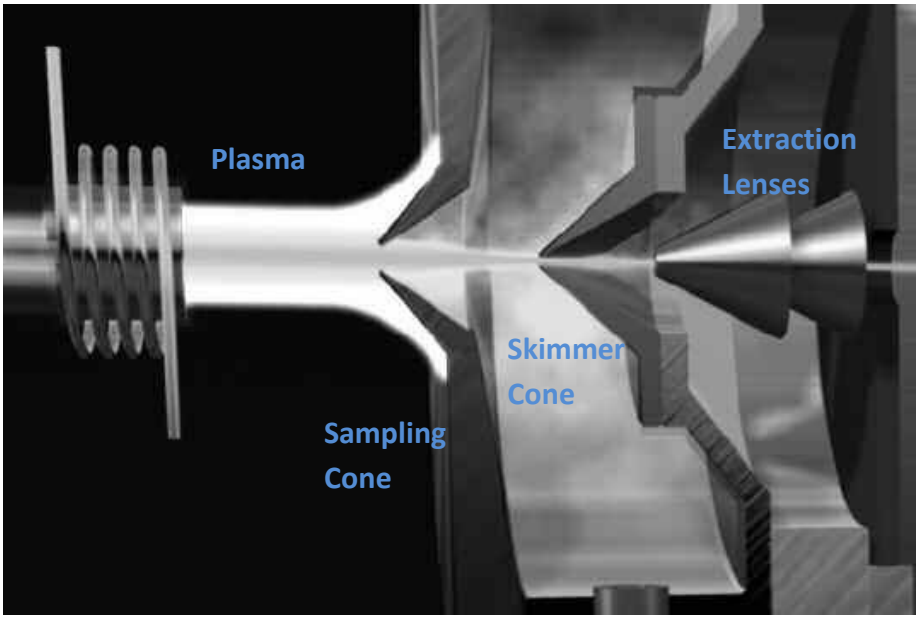


Figure 51: Schematic of the principle components of the Agilent interference showing the sampling and skimmer cones (172).

4:2.6 d Ion Focusing

Following successful ion extraction from the plasma/vacuum interface, the ions were directed into the main vacuum chamber by a series of electrostatic lenses known as ion optics. These optics focused the ion beam achieving a high signal sensitivity before they entered the quadrupole. A focused ion beam was formed using a positively charged metallic cylinder which focussed the ions into a narrow beam. The focussed ion beam was then bent by an off-axis Omega lens (see figure 52) in order to separate the positively charged ions from any unwanted neutral particles, and photons, to prevent them from reaching the detector helping to reduce any background noise.

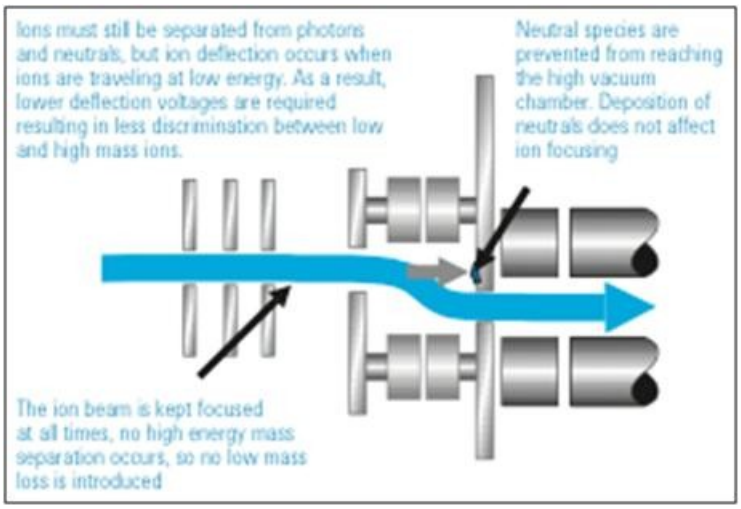


Figure 52: Schematic of the Omega ion lens design used in ion focusing (172).

4:2.6 e Collision Cell / Reaction Cell (CRC): Octopole Reaction System (ORS)

The CRC is a way of removing any spectral interferences and consists of an ion guide, which is enclosed in a cell that can be pressurized with a gas. The ion beam interacts with the gas removing polyatomic interferences by either reaction or collision mode. Reaction mode is where the gas reacts with the interference and thus converts it to a different species. The collision mode is where the gas and the interference collide, causing the interference to lose energy, which can then become separated from the high-energy analyte by energy discrimination. The ORS is located after the ion optics and before the quadrupole (see figure 53) and operated effectively in both reaction and collision modes and is thus able to use both reactive and non-reactive gases. Argon was used in this study. Therefore, the ORS is suitable for interference removal in high-purity semiconductor reagents and for the removal of variable and unidentified interferences in complex matrices. This is an advantage in this study, as soil is a complex matrix and although separation is employed prior to analysis there may still be interferences, which need to be removed in order to obtain usable data.

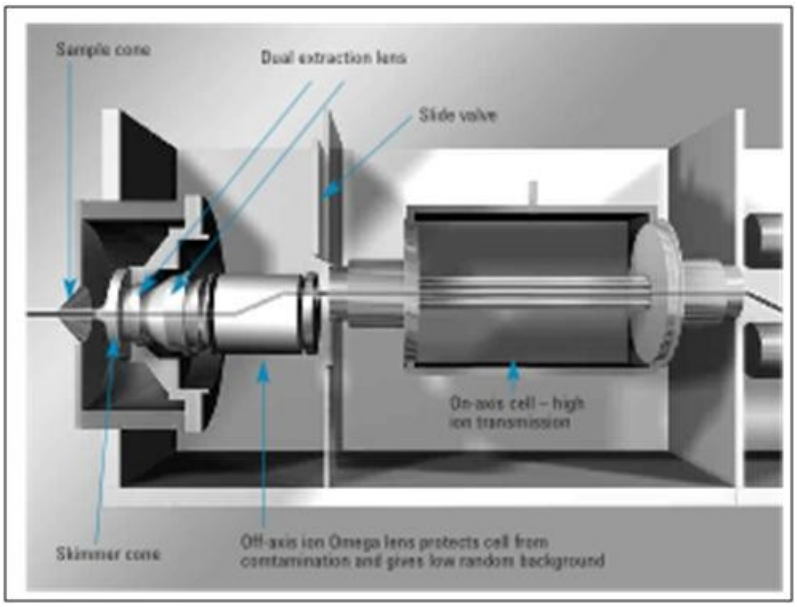


Figure 53: Schematic of the Agilent Octopole Reaction System (taken from the ICP-MS 30 minute guide)

4:2.6 f Quadrupole Mass Spectromter

The ion beam now travels through the mass spectrometer; the main function of which is to allow the analyte ions with a particular mass to charge ratio through to the detector and to filter out all of the non-analyte, interfering and matrix ions. The mass analyser used in this study was the quadrupole. It consists of two pairs of metal cylindrical rods approximately 20 cm in length and 1 cm diameter. The quadrupole separates ions using a combination of AC (alternating current) and DC (direct current) electrical fields. By applying the same, but out of phase, AC voltage between the opposite pairs of rods and positive DC voltages on one pair of rods and negative DC voltages on the other pair a hyperbolic electric field is created. This allows the quadrupole to act as a mass filter; only ions of a certain mass-to-charge ratio pass through the centre of the quadrupole, ions of different masses enter an unstable trajectory and collide with the rods. The voltages applied to the rods can be varied and increased very rapidly to scan the entire mass range of 2 – 260 amu in just 100 ms, passing each mass of interest to the electron multiplier detector sequentially.

4:2.6 g Electron Multiplier Detector

The ion signals, which have passed through the quadrupole, are measured by the electron multipler detector, which is made up of dynodes. An ion hits the first dynode and generates a shower of electrons, this process continues on through sucessive dynodes, initiating the amplification process. By the time the electron cascade reaches the final dynode, a large pulse has been accumulated and thus a measureable and accurate ion count. By counting the pulses generated by the detector the system counts the ions that hit the first dynode creating a mass spectrum. The magnitude of each peak in the mass spectrum is directly proportional to the concentration of an element in the sample. By comparing signal intensities of the samples, to those produced by calibration standards, quantitative results are acheived.

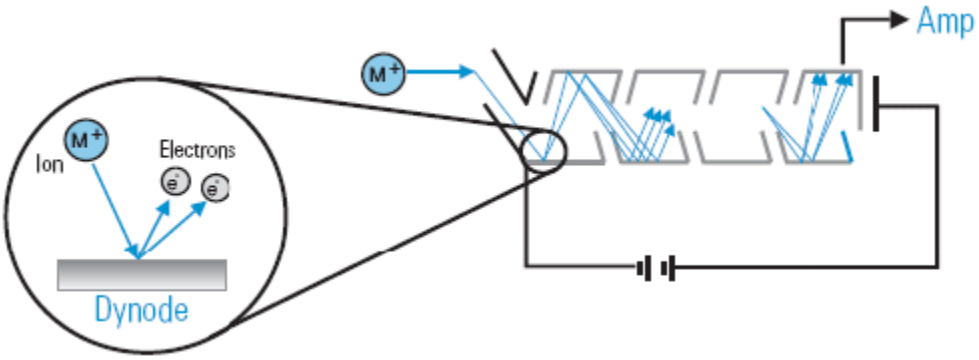


Figure 54: Schematic of electron multiplier detector (172)

Table 12: The optimised Aglient operating parameters.

Plasma Conditions		**Octopole Parameters**	
RF Power	1500 W	Oct P RF	150 V
RF Matching	1.61 V	Oct P Bias	-6 V
Sample Depth	7.6 mm		
Torch-H	0 mm	**Detector Parameters**	
Torch-V	-0.8 mm	Discriminator	8.0 mV
Carrier Gas	0.9 L min^{-1}	Analogue HV	1800 V
Makeup Gas	0.13 L min^{-1}	Pulse HV	1230 V
Optional Gas	0 %		
Nebuliser Pump	0.2 rps	**Q-Pole Parameters**	
Sample Pump	-- Rps	Amu Gain	128
S/C Temperature	2 °C	Amu Offset	125
		Axis Gain	1.0005
Ion Lenses		Axis Offset	-0.12
Extract 1	0 V	QP Bias	-2.5 V
Extract 2	-140V		
Omega Bias – ce	-22 V	**Reaction Cell**	
Omega Lens – ce	0.6 V	Reaction Mode	0 V
Cell Entrance	-26 V	H_2 Gas	0 mL min^{-1}
QP Focus	2 V	He Gas	0 mL min^{-1}
Cell Exit	-40V	Optional Gas	-- %

The instrument parameters are listed above in table 12; these were tuned to maximum sensitivity by adjusting the torch position, gas flows and ultrasonic nebuliser settings and tuning the lens system. The system was calibrated using the standards described in section 4:2.4a.

4:2.7 Instrumentation - Isoprobe (MC-ICP-MS) - $^{87}Sr/^{86}Sr$ ratios

An Isoprobe (ex. Micromass) magnetic sector field MC-ICP-MS equipped with a micro centric nebulizer and Ni and Pt sample and skimmer cones was used for the experiments carried out in this research as shown in figure 55.

Multi-collector ICP-MS (MC-ICP-MS) has proven to be very useful for strontium isotope analysis (173). The Isoprobe (used in this study) is a single-focussing, magnetic sector MS, which has an IC plasma source and a multiple-cup collector. Unlike the Agilent ICP-MS, where the quadrupole uses a single detector, the Isoprobe employs a detector array of multiple Faraday cups. The use of a separate detector for each measured mass therefore allows for the simultaneous analysis of all of the chosen isotopes. This results in high accuracy and precision by combining the high ionization efficiency of the ICP and eliminating the problems encountered with sequential scanning and plasma flicker.

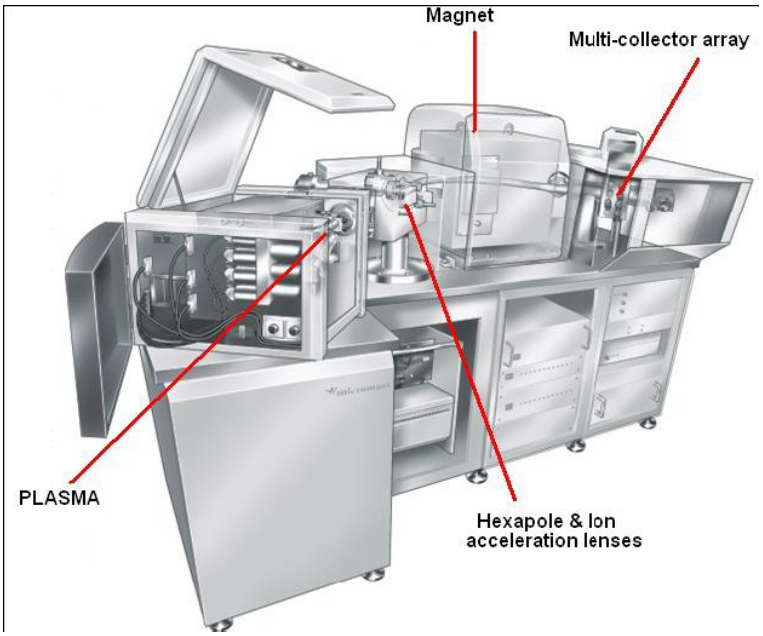


Figure 55: Diagram showing the Isoprobe multiple collector inductively coupled plasma mass spectrometer (MC-ICP-MS) used in this research. (Micromass Isoprobe User's Guide)

As can be seen in figure 56, the Isoprobe consists of an ICP source and a magnetic mass analyser with a unique ion transfer system based on an RF-only hexapole cell used as the collision cell. The ions from the plasma interact with a collision gas introduced into the hexapole reducing their energy spread from approximately 15 V to less than 1 V; this makes the ions suitable (after acceleration) for direct entry into a magnetic-sector mass analyser. It is the low energy spread of the ions, which allows the refocused beam width to be less than the width of the collector slits and thus produce the essential flat top peaks. These peaks mean that the ion beam measured at any point on the flat peak will have the same intensity; resulting in ion intensities that can be measured accurately despite any small fluctuations in magnetic field or ion focusing (172).

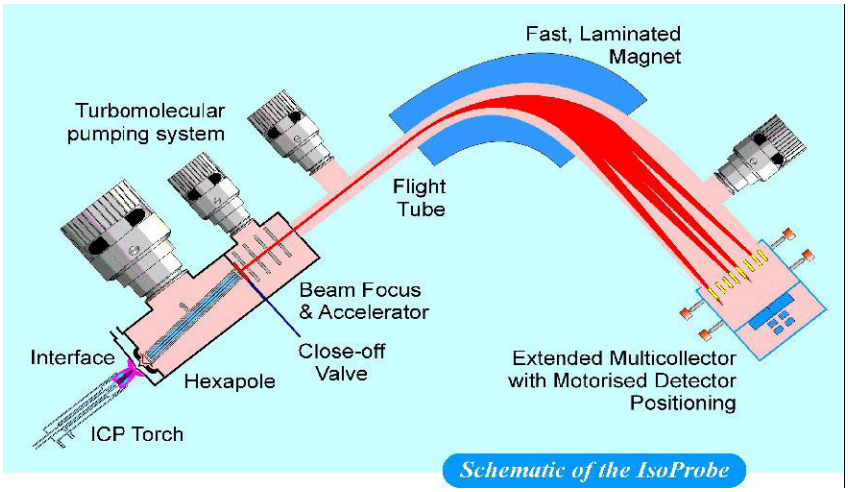


Figure 56: Schematic of the Micromass (now GV Instruments) Isoprobe MC-ICP-MS (Micromass Isoprobe User's Guide (172))

As can be seen in figure 57, the ions are produced in a plasmas torch and then pass through the sampling cone into the first chamber; the expansion chamber. The ions are then further selected through a skimmer cone into a second chamber; the intermediate chamber. The ions are then accelerated into the hexapole where they are thermalised by collisions with a small amount of collision gas, which is introduced into the hexapole chamber. The gas used is argon (Ar), and as mentioned previously, reduces the energy spread of the ions. Once the ions exit the hexapole, they pass into a high vacuum section where they are collimated and accelerated to 6 kV before they enter the magnetic sector. Due to the narrow energy

spread of the ions leaving the hexapole all of the ions are transmitted to the magnet at the same potential.

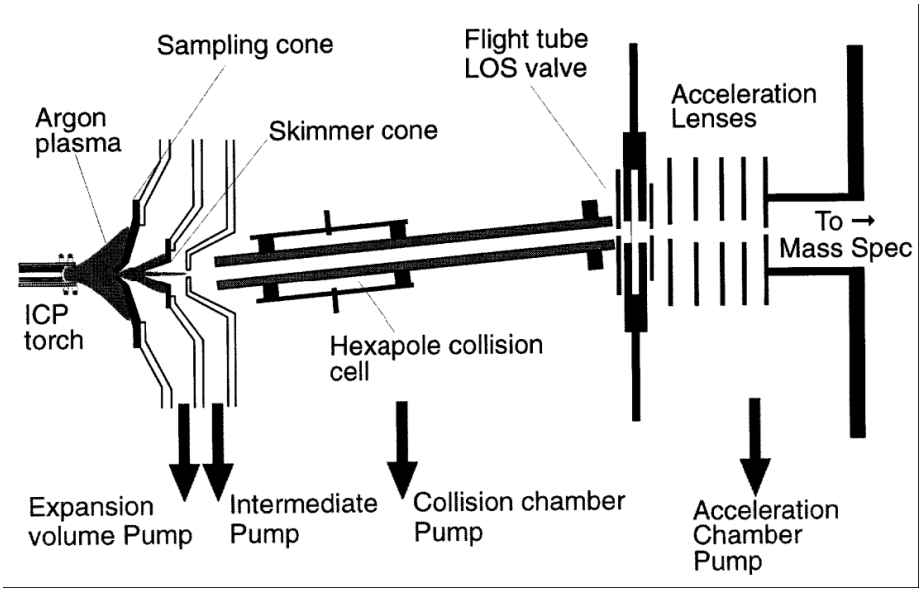


Figure 57: Schematic of the plasma interface and the hexapole collision cell of the Isoprobe MC-ICP-MS. (Source: Rehkämper *et al*, 2006)

The Isoprobe is fitted with a laminated magnet capable of the high-speed mass-scanning required for elemental analysis. The magnet separates the ions in the beam according to their mass-to-charge ratio. A wide flight tube ensures the measurement of a mass window of 17 %; up to 17 amu at mass 100.

4:2.7 a Faraday Cups

The Isoprobe was fitted with nine Faraday cups which enable the simultaneous collection of the separated isotopes, which in turn cancels out the “noisy” signal effect of the isotope ratio measurement. The use of the Faraday cups and the single static simultaneous collection also eliminates the errors associated with plasma instability, which can occur with sequential beam measurements. Each of the Faraday cups can be independently altered to allow for the analysis of a wide range of elements whose isotopes display different mass dispersions. The configuration of the Faraday cups for the analysis of strontium isotope ratios is shown in table 13.

Table 13: The configuration of the Faraday cups in the Isoprobe for the analysis of $^{87}Sr/^{86}Sr$.

Faraday Cup	Isotope Collected
H5	^{88}Sr
H4	^{87}Sr
H3	^{86}Sr
H2	^{85}Rb
H1	^{84}Sr
Axial	^{82}Kr

4:2.7 b Mass Bias of Isotope Ratio Measurements

Mass bias is the inability of a system to transmit all isotopes of the same element at the same rate. Generally, the measured mass of an ion is slightly lighter than its actual mass. This is due to the preferential transport of the heavier isotopes into the mass analyser, and therefore the effect is more pronounced for the low-mass elements. This is highlighted in the plot of the mass bias (figure 58) by Rehkämper *et al*. (174)

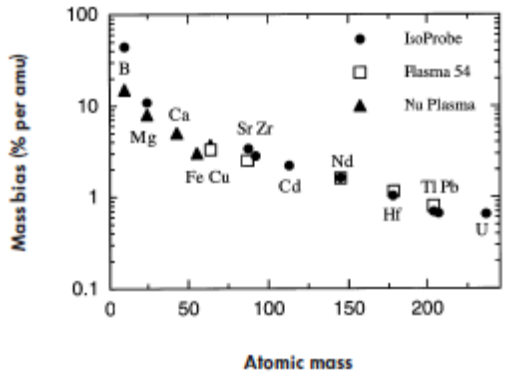


Figure 58: Plot of the mass bias per amu (%) verses Atomic mass for different elements measured on different MC-ICP-MS instruments. (Source: Rehkämper *et al.* (174))

Processes, such as the space charge effects in the region of the skimmer cone, are thought to contribute to the deviation of the mass from the actual value.

Generally, mass bias is expressed as the deviation of a measured isotope ratio (R_{meas})from the reference value (R_{true}) normalised to a mass difference (ΔM) of 1 amu as shown in equation 5 (175).

5 $$Mass\ bias = \frac{{}^{R_{meas}}/_{R_{true}} - 1}{\Delta M}$$

4:2.7 c Mass Bias Correction by Internal Normalisation

Measurements of radiogenic isotope compositions such as $^{87}Sr/^{86}Sr$ can be corrected for by normalizing the measured ratio to an invariant isotope ratio of the same element (174). When an element has two stable isotopes occurring in a known constant ratio, the measurements can be corrected internally for mass fractionation. The $^{86}Sr/^{88}Sr$ ratio is constant in nature and can therefore be used to correct the mass bias. The difference between the measured and the theoretical $^{86}Sr/^{88}Sr$ ratio (from NIST SRM 987, value of 0.1194) can be calculated via equation 1. Subsequently the measured $^{87}Sr/^{86}Sr$ ratio can then be accordingly corrected.

4:3 Results

4:3.1 Strontium isotope ratios

A summary of the strontium isotope ratios is shown in the table below as measured by MC-ICP-MS.

Table 14: Summary of the 87Sr/86Sr ratios for all of the sample sites, with three samples analysed from each site (taken within a 5m by 5m area).

	87 / 86 Sr ratio			
Site	1	2	3	**St Dev**
160708-09	0.70795	0.70785	0.70796	0.00006
240708-1	0.70783	0.70784	0.70775	0.00005
240708-2	0.70986	0.71016	0.71007	0.00015
240708-4	0.70800	0.71031	0.70896	0.00116

240708-7	0.70888	0.70824	0.70825	0.00037
160708-12	0.71091	0.71115	0.71069	0.00023
16070-84	0.70788	0.70790	0.70789	0.00001
160708-5	0.70764	0.70780	0.70783	0.00010
240708-6	0.70851	0.70794	0.70861	0.00036
240708_3	0.71051	0.71049	0.71050	0.00001
240708_5	0.70777	0.70775	0.70778	0.00001
220909_5	0.70808	0.70809	0.70808	0.00000
220909_7	0.70941	0.70938	0.70943	0.00003
220909_9	0.70808	0.70809	0.70808	0.00000
230909_1	0.70867	0.70864	0.70859	0.00004
230909_2	0.70870	0.70870	0.70886	0.00009
230909_3	0.70911	0.70904	0.70908	0.00003
230909_4	0.70897	0.70905	0.70892	0.00007
230909_5	0.70891	0.70891	0.70868	0.00013
230909_6	0.70823	0.70819	0.70821	0.00002
230909_7	0.70821	0.70806	0.70811	0.00008
230909_8	0.70878	0.70878	0.70902	0.00014
230909_9	0.70853	0.70851	0.70850	0.00002
230909_10	0.70828	0.70829	0.70829	0.00000
230909_11	0.70872	0.70871	0.70875	0.00002
170610-10	0.70838	0.70852	0.70837	0.00009
170610-11	0.70910	0.70894	0.70906	0.00008
170610-12	0.70865	0.708654	0.70858	0.00004
170610-6	0.70841	0.708329	0.708329	0.00005
170610-7	0.70809	0.70814	0.70819	0.00005
170610-8	0.70780	0.71010	0.71013	0.00134
160610-9	0.70847	0.708639	0.70856	0.00008
170610-1	0.70927	0.70927	0.70938	0.00007
170610-2	0.70857	0.70859	0.70848	0.00006
170610-3	0.70847	0.70850	0.70845	0.00002
170610-4	0.70830	0.70825	0.70826	0.00002
170610-5	0.70964	0.70976	0.70959	0.00009
210909-2	0.70855	0.70859	0.70852	0.00004
210909-4	0.70901	0.70875	0.70862	0.00020
210909-8	0.70772	0.70772	0.70769	0.00002
210909-1	0.70859	0.7086	0.70856	0.00002
210909-3	0.70800	0.708045	0.70778	0.00014
210909-6	0.70917	0.70915	0.70912	0.00002

Firstly the data was sorted into a histogram to show the spread of the strontium isotope ratios (figure 59).

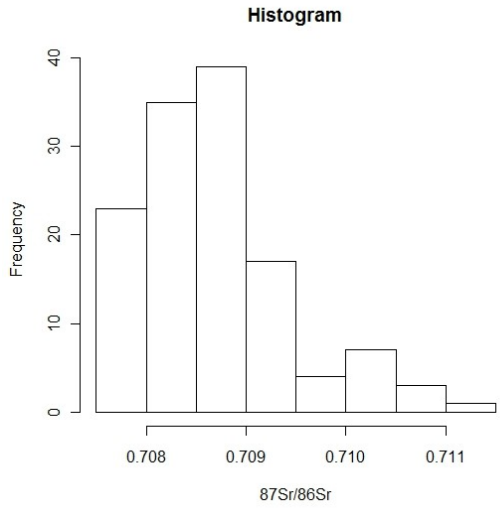


Figure 59: Histogram to show the range of the $^{87}Sr/^{86}Sr$ ratios measured with MC-ICP-MS.

In order to see if there was any correlation between the strontium isotope ratios and their underlying geologies a boxplot was created with the data arranged according to geology (figure b). Sites A to Q are from a chalk geology, R-AE Upper Greensand and Gault , AF-AI Norwich Crag, AJ-AO Lower Cretaceous and AP-AQ Amtptil Clay.

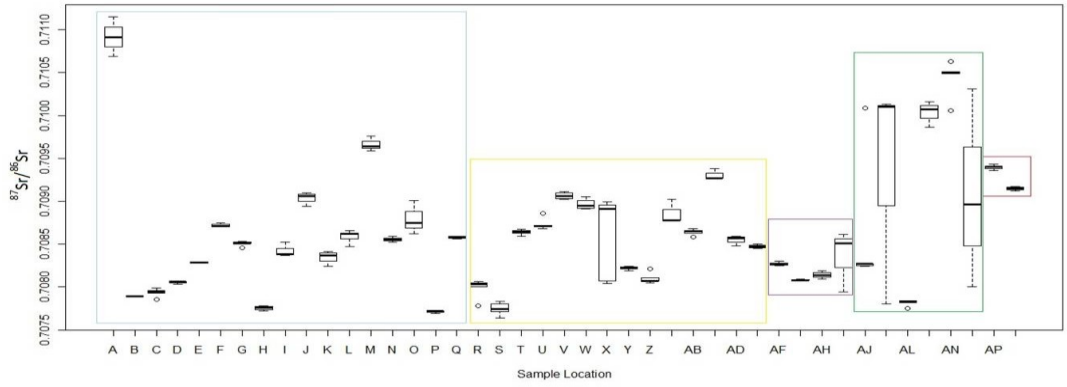


Figure 60: Boxplot showing the $^{87}Sr/^{86}Sr$ variations between different sample locations grouped according to geology, shown by Chalk in blue, Upper Greensand and Gault in yellow, Norwich Crag in purple, Lower Cretaceous in green and Amptill Clay in red. As can be seen there is an outlier in the chalk geology (site A). This has been reanalysed and is discussed in more detail later.

The data was then shown spatially across a Norfolk geology map with graduated markers for the strontium isotope ratios.

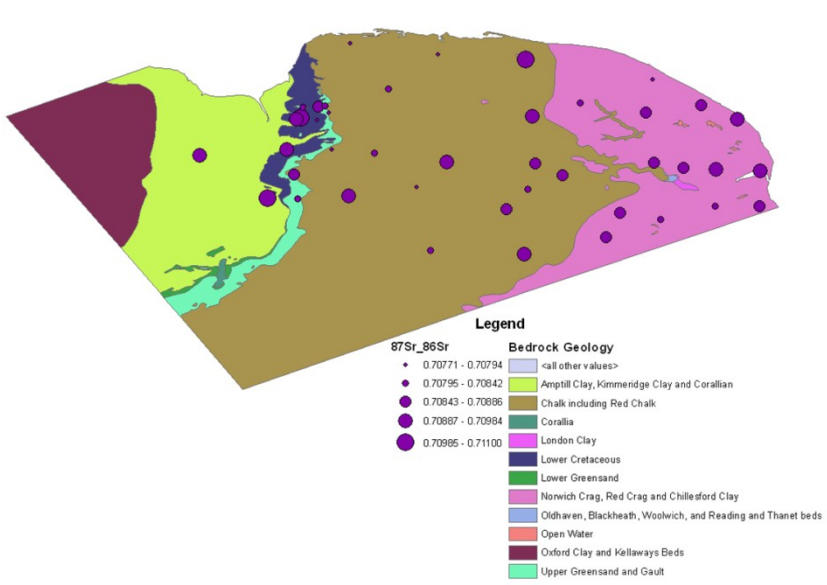


Figure 61: The geological spatial variation of the $^{87}Sr/^{86}Sr$ (shown by the graduated purple dots) across Norfolk.

In order to show the variations between the geological regions the $^{87}Sr/^{86}Sr$ ratios were grouped according to their underlying geologies and a boxplot made.

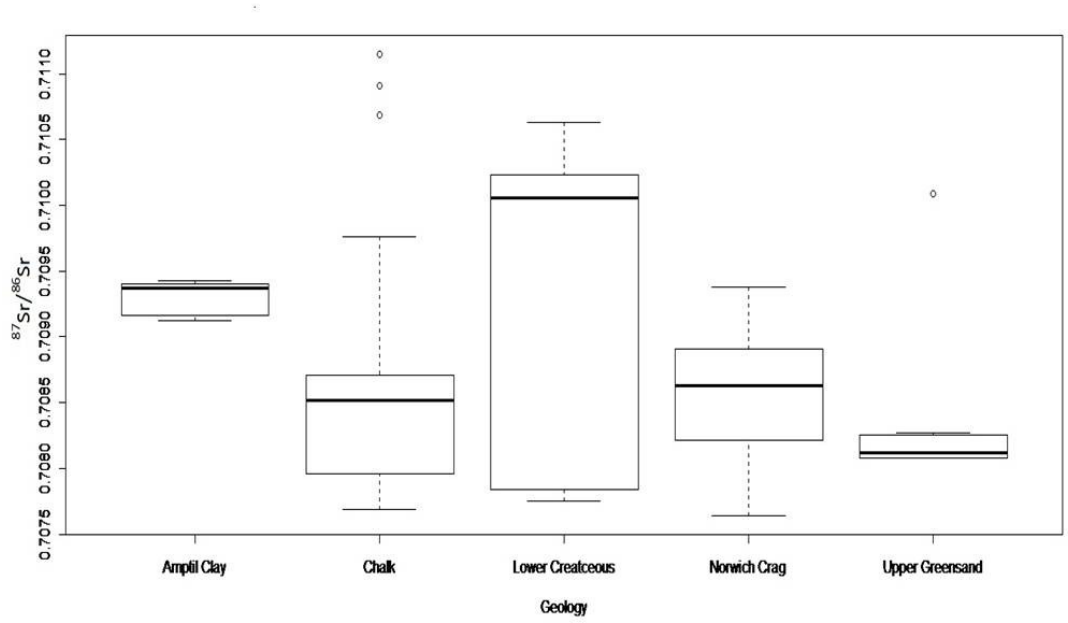


Figure 62: Boxplot showing the $^{87}Sr/^{86}Sr$ variations between the different geological regions.

As can be seen in figure 62 there are no clear differences between geologies, although Amptil Clay and Upper Greensand and Gault can be distinguished from each other. The differences are confirmed in a Tukey HSD plot (figure 63). It should also be noted that mean $^{87}Sr/^{86}Sr$ for the Lower Creatacous is higher than that of all of the other geologies; something that could perhaps be exploited for forensic purposes.

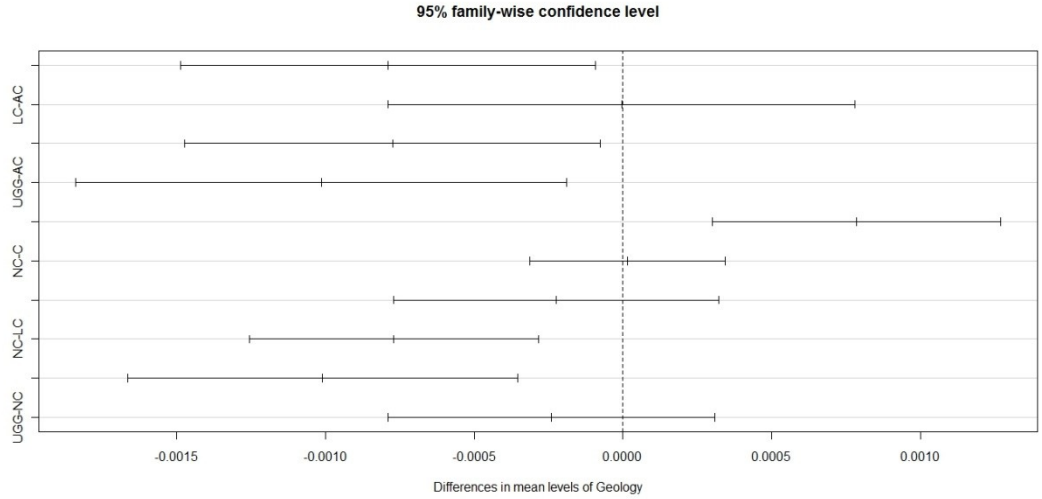


Figure 63: Tukey HSD plot showing where the significant differences occur between the $^{87}Sr/^{86}Sr$ of the different geologies.

As no clear differences between the strontium isotope ratios according to their underlying geologies were seen across Norfolk it was decided to investigate whether they had any correlation with the botanical regions. The $^{87}Sr/^{86}Sr$ data were plotted spatially across the Norfolk botanical regions with graduated markers showing the isotope ratios (figure 64)

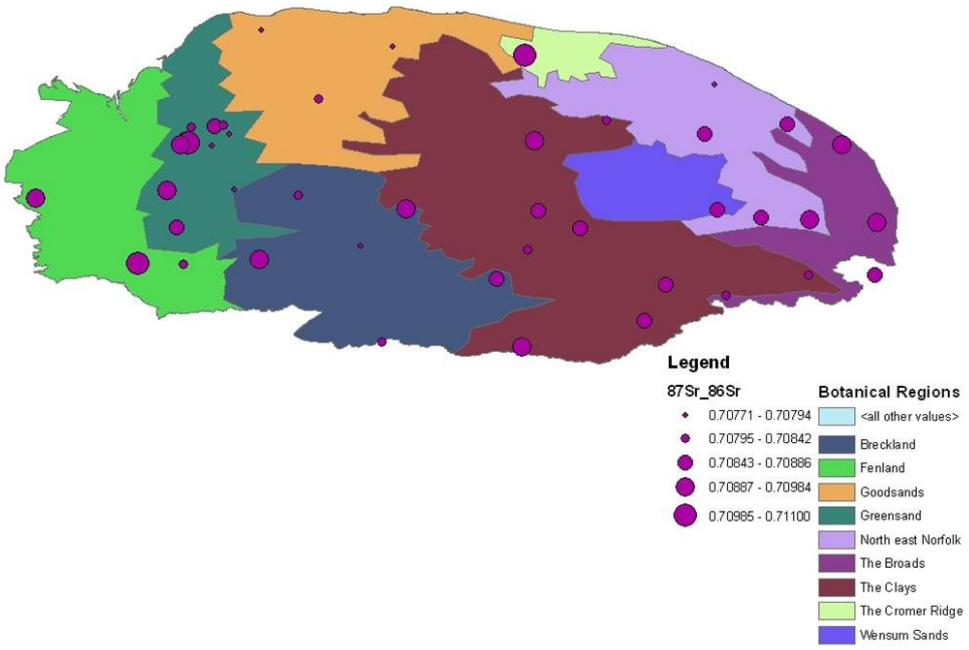


Figure 64: The botanical spatial variation of the $^{87}Sr/^{86}Sr$ (shown by the graduated pruple dots) across Norfolk.

As before, in order to show the variations between botanical regions the $^{87}Sr/^{86}Sr$ ratios were grouped according to their botanical regions and a box plot made (figure 65)

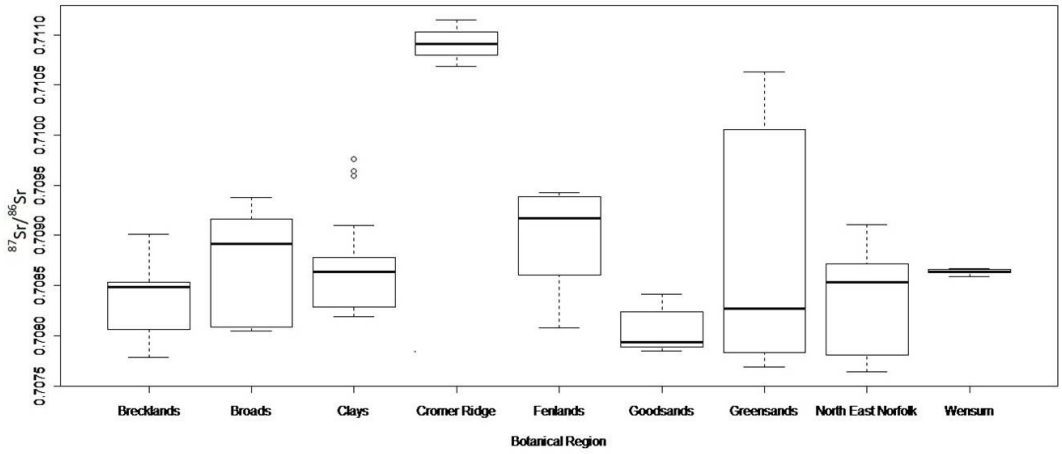


Figure 65: Boxplot showing the$^{87}Sr/^{86}Sr$ variations between the different botanical regions.

Again, as per geology, all of the botanical regions cannot be separated from each other. However, Cromer Ridge is different from all of the other botanical regions, Wensum Sands is different from Good Sands, and depending upon errors there is the potential for Goodsands to be differentiated from Fenlands. This is confirmed in the Tukey HSD plot of the botanical regions (figure 66).

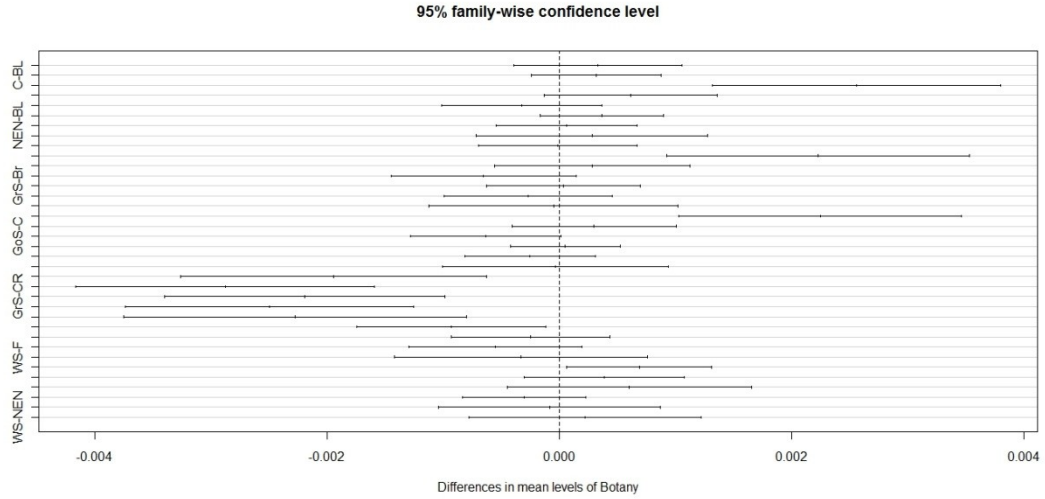


Figure 66: Tukey HSD plot showing where the significant differences occur between the $^{87}Sr/^{86}Sr$ of the different botanical regions.

Intra and inter site variation

In order to show inter and intra site variations of the $^{87}Sr/^{86}Sr$, a simple 2 way ANOVA with replication was performed. The intra site variation gave a *p*-value of 3.06×10^{-9} and the inter site a *p*-value of 0.674. This shows that the variation within the sites is much smaller than the variations between the different sampling locations.

Triplicate Samples

Triplicate preparations were made of some soil samples in order to measure the analytical precision, table 15.

Table 15: Summary statistics of the errors associated with the triplicate samples analysed.

	Triplicates	Mean	Standard Deviation	% RSD	Standard Error	Standard Error / ppm
16070844a	0.70789					
16070844b	0.70788					
16070844c	0.70790	0.70789	1.03261E-05	0.0015%	0.00001	0.842
16070853a	0.70780					
16070853b	0.70783					
16070853c	0.70781	0.70781	1.56598E-05	0.0022%	0.00001	1.277
16070854a	0.70772					
16070854b	0.70774					
16070854c	0.70771	0.70772	1.4593E-05	0.0021%	0.00001	1.190
16070893a	0.70799					
16070893b	0.70793					
16070893c	0.70794	0.70795	3.04598E-05	0.0043%	0.00002	2.484
2209097_2a	0.70940					
2209097_2b	0.70938					
2209097_2c	0.70936	0.70938	2.35139E-05	0.0033%	0.00001	1.914
2309091_2a	0.70866					
2309091_2b	0.70864					
2309091_2c	0.70863	0.70864	1.29134E-05	0.0018%	0.00001	1.052
2309092_1a	0.70871					
2309092_1	0.70872	0.70870	1.79652E-05	0.0025%	0.00001	1.464

b						
2309092_1c	0.70868					
2309095_3a	0.70899					
2309095_3b	0.70809					
2309095_3c	0.70898	0.70868	0.000516438	0.0729%	0.00030	42.073
2309096_1a	0.70823					
2309096_1b	0.70822					
2309096_1c	0.70824	0.70823	9.5096E-06	0.0013%	0.00001	0.775
2309097_2a	0.70805					
2309097_2b	0.70807					
2309097_2c	0.70807	0.70806	1.13409E-05	0.0016%	0.00001	0.925
2309099_3a	0.70851					
2309099_3b	0.70846					
2309099_3c	0.70852	0.70850	3.10714E-05	0.0044%	0.00002	2.532
24070812a	0.70783					
24070812b	0.70782					
24070812c	0.70784	0.70783	7.64905E-06	0.0011%	0.00000	0.624
24070832a	0.71049					
24070832b	0.71050					
24070832c	0.71051	0.71050	9.38276E-06	0.0013%	0.00001	0.762
24070851a	0.70777					
24070851b	0.70775					
24070851c	0.70778	0.70777	1.33137E-05	0.0019%	0.00001	1.086
24070871a	0.71009					
24070871b	0.70827	0.70888	0.001050638	0.1482%	0.00061	85.570

2407087 1c	0.70827					
2109091-1a	0.70859					
2109091-1b	0.70858					
2109091-1c	0.70864	0.70860	3.26064E-05	0.0046%	0.00002	2.657
170610-12-2a	0.70866					
170610-12-2b	0.70863					
170610-12-2c	0.70868	0.70865	2.7189E-05	0.0038%	0.00002	2.215
170610-6-2a	0.70824					
170610-6-2c	0.70836					
170610-6-2b	0.70838	0.70833	7.69036E-05	0.0109%	0.00004	6.268
170610-9-2a	0.70866					
170610-9-2b	0.70862					
170610-9-2c	0.70863	0.70864	1.92933E-05	0.0027%	0.00001	1.572
170610-3-3a	0.70839					
170610-3-3b	0.70878					
170610-3-3c	0.70882	0.70866	0.000233095	0.0329%	0.00013	18.990
2109093-1a	0.70805					
2109093-1b	0.70806					
2109093-1c	0.70803	0.70804	1.5681E-05	0.0022%	0.00001	1.279

Samples with high standard error as highlighted in red were re-analysed and gave a very similar result. Generally the error is very similar to that of the ICP-MS and is also similar to the error seen on other samples analysed on the MC-ICP-MS.

ISMA Standards

Analysis of the precision of the MC-ICP-MS was performed by analysing a NIST SRM of known isotopic composition before each sample. NIST SRM 987 was analysed giving the data shown in table 16.

Table 16: Summary statistics of the errors associated with the ISMA standards analysed (the average of eleven ISMA standards were analysed in total across all analytical runs).

Mean	Standard Deviation	% RSD	Standard Error	Standard Error / ppm
0.70917	0.00003	0.00461%	0.00001	16.31526

As can be seen from tables 15 and 16 the standard error of most of the triplicates is less than that of the eleven ISMA standards run.

4:3.2 Trace Element Data

Table 16: Summary of the trace element data acquired using ICP-MS for the bioavailable fraction of all soil samples.

	Mean concentration /ppb	Standard Deviation	Confidence Level(95.0%)
Li	9.93187	12.94753	1.88325
Be	0.08631	0.35099	0.05105
B	105.87982	176.77203	25.71191
Na	6860.01495	7904.58766	1149.74097
Mg	23237.49511	23608.44808	3433.90461
Al	303.83480	1569.69131	228.31531
K	85418.33668	101497.26305	14763.01696
Ca	607721.59891	591510.13862	86036.54861
Sc	0.76126	0.59889	0.08711
Ti	2.61608	5.38386	0.78310
V	1.56751	2.04696	0.29774
Cr	0.05145	0.69795	0.10152
Mn	1200.48791	5318.87813	773.64340
Fe	87.71673	337.96016	49.15710
Co	3.57732	8.88998	1.29307

Ni	9.69493	16.15751	2.35015
Cu	22.25868	34.12220	4.96315
Zn	152.20562	481.65725	70.05819
Ga	0.02847	0.13068	0.01901
As	5.12807	5.34490	0.77743
Se	0.84453	0.93934	0.13663
Rb	63.74675	55.59907	8.08702
Sr	1789.08404	1900.96149	276.49934
Y	0.36731	1.22706	0.17848
Zr	0.49512	1.86239	0.27089
Mo	2.07092	3.46851	0.50450
Ag	1.03712	3.29667	0.47951
Cd	1.01144	2.02339	0.29431
Sb	0.41124	0.88735	0.12907
Te	0.01570	0.07401	0.01076
Cs	1.43518	1.34178	0.19517
Ba	2703.13079	2351.66404	342.05510
La	0.52842	1.89081	0.27502
Ce	0.67947	2.67387	0.38892
Pr	0.08024	0.30622	0.04454
Nd	0.33701	1.19249	0.17345
Sm	0.07017	0.22069	0.03210
Eu	0.67222	0.56927	0.08280
Gd	0.09376	0.26784	0.03896
Tb	0.00828	0.03400	0.00495
Dy	0.04789	0.18258	0.02656
Ho	0.01155	0.03472	0.00505
Er	0.02616	0.08847	0.01287
Tm	0.00359	0.01140	0.00166
Yb	0.01496	0.06147	0.00894
Lu	0.00340	0.00897	0.00131
Hg	0.69701	3.37472	0.49086
Tl	0.76043	0.73301	0.10662
Pb	7.82893	39.27470	5.71260
Bi	0.00871	0.02296	0.00334
Th	0.06326	0.32583	0.04739
U	0.20673	0.28327	0.04120

For tables of trace element data grouped according to geology and botanical areas see appendix 2.

As can be seen in the above table some of the trace elements have a very varied mean concentration. Some of the elements seem to have a particularly low concentration (Tm, Lu and Bi), but it must be remembered that these are the results from the bioavailble fraction of the soil, and a higher concentration would be found if the bulk soil had been analysed. This data is in accordance to trace element data acquired in the Trace project, where a low concentration was found for many elements in the bioavailable fraction of the soil.

In order to see whether the trace elements show any correlation with the underlying geology and land use the Bray-Curtis dissimilarity has been used as shown in the equation below (see pollen chapter for full explanation of this statistical method); and then ordination plots made using the Bray-Curtis distances.

$$d^{BCD}(i,j) = \frac{\sum_{k=0}^{n-1} \left| y_{i,k} - y_{j,k} \right|}{\sum_{k=0}^{n-1} \left(y_{i,k} + y_{j,k} \right)}$$

As can be seen when grouped according to both geology and land use, the rare earth elements give a better separation than the whole suite of elements. The ordiplots also show that both geology and land use can be influencing factors upon the trace element concentrations. Rare earth elements, from fertilisers and pesticides may explain the observations.

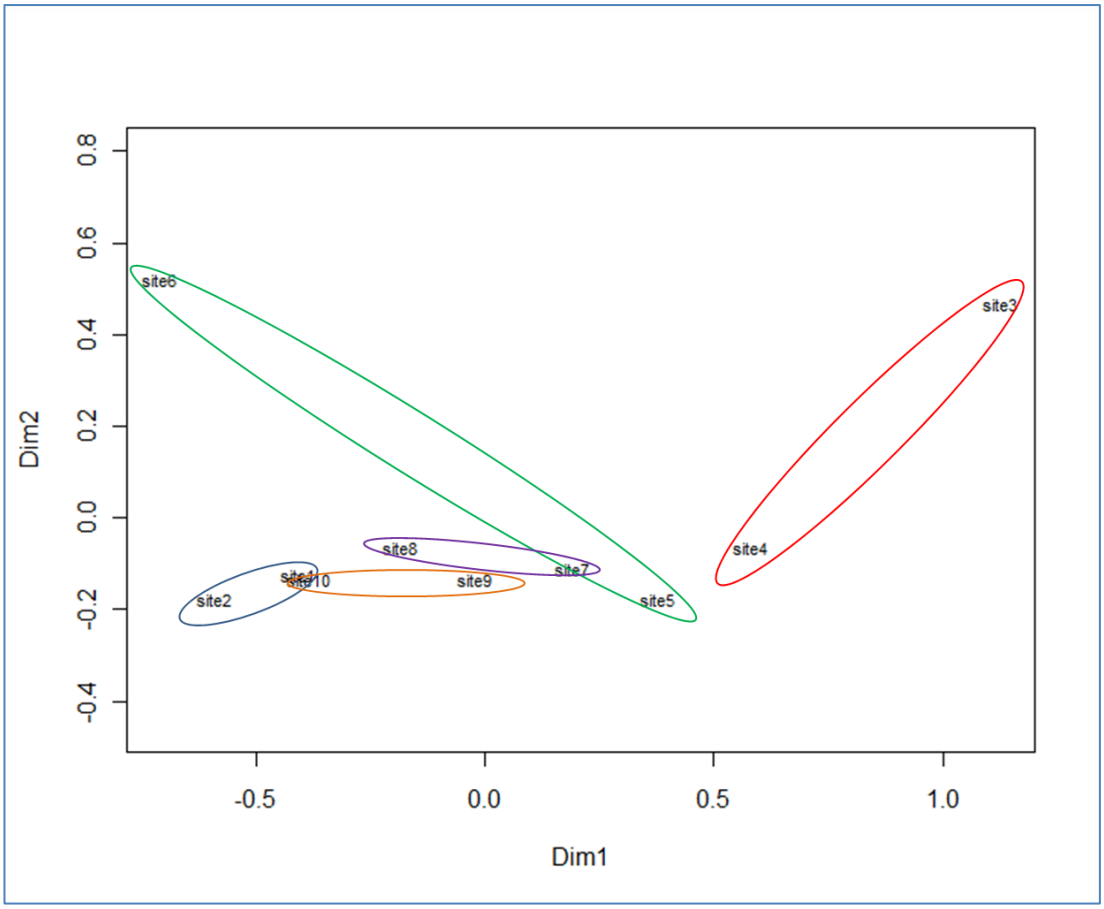


Figure 67: Ordination plot for trace elements for geology: Ordination plot showing the 95% confidence intervals values for the Bray-Curtis equation and then plotted using NMDS (Non-metric multi-dimensional scaling) to show the range of the trace elements for each geology, grouped by blue for Amptill Clay, red for Lower Cretaceous, green for Upper Gault and Greensand, purple for Chalk and orange for Norwich Crag.

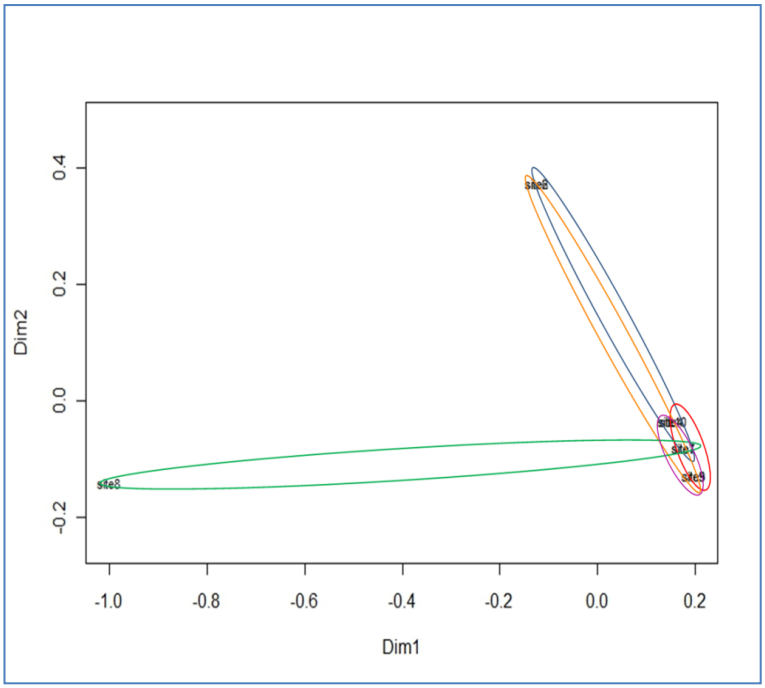


Figure 68: Ordination plot of trace elements for land-use: Ordination plot showing the 95% confidence intervals value for the Bray-Curtis equation and then plotted using NMDS (Non-metric multi-dimensional scaling) to show the range of the trace elements for each land use, grouped by blue for Forrest, red for Agriculture, green for Quarry, purple for Scrubland and orange for Park land.

As can be seen by the above plots (figures 67 and 68) the trace element data cannot be used to differentiate all of the samples, and it was therefore decided to investigate the use of fertilisers and see if this could be used to explain some of the similarities and variation within the data.

4:3.3 The effect of fertilisers.

As it was unclear which factors were having an effect upon the $^{87}Sr/^{86}Sr$ ratio and the trace element concentrations, it was decided to look at land use and the use of fertilisers to see if this gave some separation of the different sample sites. Fertilisers are substances that are added to soil, which supply plant nutrients or amend the soil fertility. Fertilisers typically provide three primary macronutrients, nitrogen, phosphorus and potassium; three secondary macronutrients, calcium, sulphur and magnesium; and micronutrients (trace elements), boron, chlorine, manganese, iron, zinc, copper, molybdenum and selenium.

Inorganic fertilisers may not replace trace mineral elements in the soil, which gradually become depleted by crops. Organic and inorganic fertilisers, which contain nitrogen, can cause soil acidification; this can lead to a decrease in nutrient availability, which can be offset by liming.

The rare earth elements can be representative of fertilisers, however some minerals are more transient in soil than others so will move down the soil faster and in greater quantities, something which must be considered when looking at the concentrations in the bioavailable topsoil. The concentration of the trace elements in the bioavailable fraction will be lower than the bulk extracts.

The primary macronutrients are essential for plant growth. Nitrogen is used by plants to produce leafy growth and the formation of stems and branches. Plants which mostly need nitrogen are grasses and leafy vegetables. Plants cannot utilise the nitrogen in the atmosphere. However, legumes have nodules on their roots where bacteria live that obtain nitrogen from the airand convert it to ammonia. Nitrogen fertilisers are quickly washed out of the soil by rain and need renewing annually. Phosphorus is essential for seed germination and root development, and is needed particularly by young plants forming their root systems. Phosphates remain in the soil for two to three years after application. Potassium promotes flower and fruit production and is vital for maintaining growth and helping plants to resist disease. It is used in the process of building starches and sugars, and is needed in fruit and vegetables. A shortage of potassium leads to a low resistance to disease, scorching of leaves and a poor fruit yield. Potassium usually lasts for two to three years in the soil after application but for vegetable production more will be required.

Secondary macronutrients are those which plants require in small quantities. Calcium is required for plants to utilise and transport other nutrients internally, particularly phosphorus. A shortage of calcium results in stunted growth. In soils, where the pH is correct for growing (usually between 5.5 and 7.5), calcium is usually available. A shortage of calcium is easily rectified by the addition of lime (primarily calcium carbonate) or gypsum. Sulphur is vital for protein production and management within the plant. A

sulphur deficiency yields symptoms similar to those of nitrogen deficiency; low growth rates, yellowing of leaves etc. Sulphur deficiency is not usually a problem, it is a component of artificial fertilisers and previously it used to be present in the atmosphere and fall to the ground from coal fires and power stations, although with modern legislation this has dramatically reduced. Sulphur is lost from the soil due to leaching and in the removal of crops. Green manures will prevent leaching and return sulphur simultaneously. Magnesium is essential for the formation of chlorophyll. Magnesium is the central atom in the chrlophyll molecule ($C_{55}H_{72}O_5N_4Mg$) and deficiency is common. A lack of magnesium results in a lack of chlorophyll and causes a reduced yield and stunted growth, increased susceptibility to disease and eventually the death of the plant. Plants quickly absorb magnesium so treating a deficiency is straight forward.

Micronutrients are those which are required in very small quantities but are still necessary for plant growth. Boron is necessary for calcium to perform its functions within the plant, too much boron can be harmful to the plant however. Signs of boron deficiency are poor development of the growing tip of the plant; this is more likely in soils with a pH over 6.5. Excess use of magnesium sulphate will also cause a boron deficiency. Copper deficiency is rare but can occur in sandy, peaty and chalky soils which have high pH levels. Copper is required for root formations. Excess copper is very toxic to both people and plants. Iron deficiency can cause yellowing of the leaves and a general lack of vigour in the plant. It is very rare but can be overcome with an iron sulphate fertiliser. Manganese deficiency is often caused be over-liming of the soils, and is most often found in peaty and sandy soils with a high pH. The addition of sulphur to the soil will increase the acidity of the soil (decrease the pH) and solve the problem. Molybdenum is rarely deficient in plants and soils and as it is only required in minute amounts, an excess can be equally as harmful as a deficiency. A zinc deficiency is more likely in soils with a high pH than low pH, and those crops which are most sensitive are tomoatoes, onions and beans (176, 177).

As the following figures show, ordination plot showing the 95% confidence interval values for the Bray-Curtis equation and then plotted using NMDS (Non-metric multi-dimensional scaling) to show the range of the rare earth elements for both land-use and geology have been created. This is because the rare earth elements can be indicative of fertiliser use and

both geology and land-use plots have been created so that they can be compared to those plots containing all of the trace element concentration data.

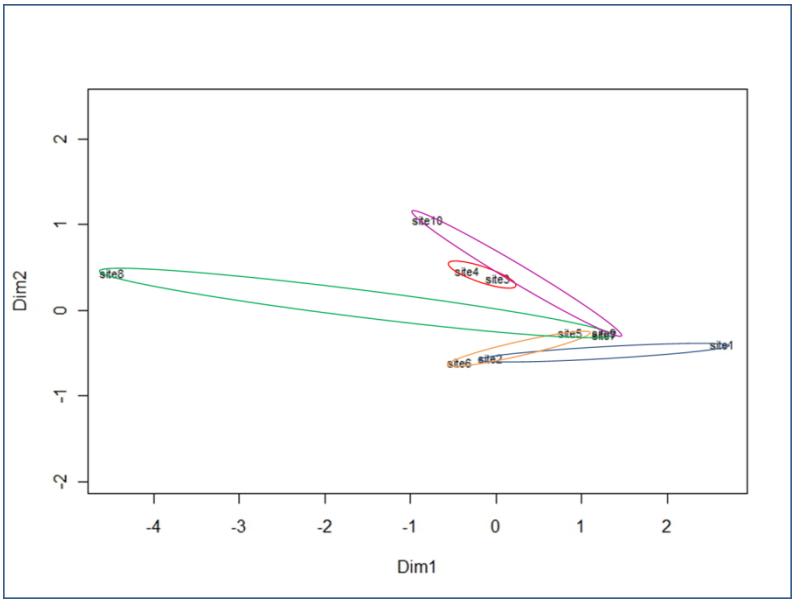


Figure 69: Ordination plot of rare earth elements for land-use: Ordination plot showing the 95% confidence interval values for the Bray-Curtis equation and then plotted using NMDS (Non-metric multi-dimensional scaling) to show the range of the rare earth elements for each land-use, grouped by blue for Forrest, red for Agriculture, green for Quarry, purple for Scrubland and orange for Park land.

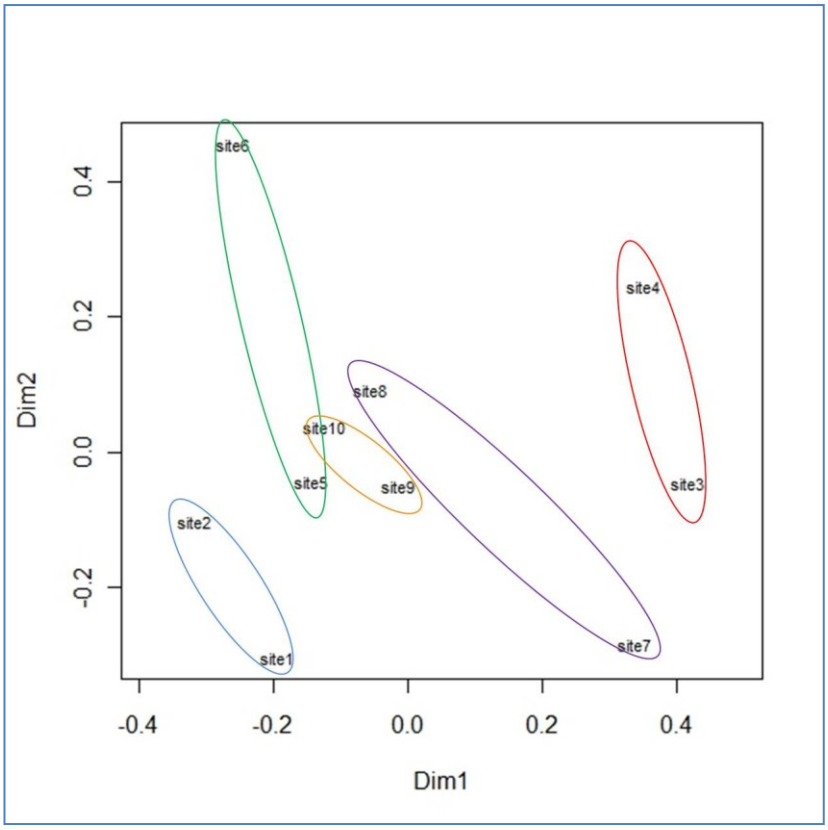


Figure 70: Ordination plot of rare earth elements for geology: Ordination plot showing the 95% confidence interval values for the Bray-Curtis equation and then plotted using NMDS (Non-metric multi-dimensional scaling) to show the range of the rare earth elements for each geology, grouped by blue for Amptill Clay, red for Lower Cretaceous, green for Upper Gault and Greensand, purple for Chalk and orange for Norwich Crag.

5:0 Discussion

Pye and Blott (178) discuss the effects of particle size on the chemical composition, highlighting that some elements e.g. silicon are more abundant in the coarser fractions and that the trace metals are generally more concentrated in the finer fractions. They state that in their experience, analysis of the < 150 µm fraction allows adequate discrimination between samples and does provide an indicative measure of sample composition.

When considering the trace elements in soil, the concentration of the rare earth elements in the bioavailable extract of the soil was very low, often below 2ppb. This has also been shown in the analysis of the bioavailable soil for samples in the TRACE project, and in analysis of GEMAS samples. The analysis of the rare earth elements in the bulk soil may

yield a higher concentration. However, although the concentrations are so low, there do appear to be significant differences between the different geologies based upon the analysis of the trace elements.

Using the $^{87}Sr/^{86}Sr$ data when grouped according the botanical region (see figures e to g); great care must be taken when making any conclusions about any correlations as within some of the botanical regions there is only one sampling site and therefore a maximum of three samples taken from that particular site so it is not necessarily representative of the whole area. This is because when the initial sampling locations were chosen they were done so on the basis of them having different underlying geologies; it was only later when very poor correlations were seen with the $^{87}Sr/^{86}Sr$ and the underlying geology was it decided to investigate vegetation and land-use and therefore the sampling locations are not always very representative of the areas.

As yet the $^{87}Sr/^{86}Sr$ and trace element spatial data cannot systematically reproduce the spatial structure of the observations from first principles. There is a poor correlation between the isotope chemistry of the topsoil and the underlying geology indicating that other sources such as land use, vegetation cover and additions to the soil contribute to the $^{87}Sr/^{86}Sr$. Trace element concentrations show some correlation with both geology and land-use. Rare earth element concentrations suggest additions to the soil influence the topsoil chemistry. Glacial drift, which occurred during the ice ages, is a mass-transport process whereby soil is transported to another location; hence, soil composition is not necessarily proportional to the underlying bedrock. Thus one can use isotopic and elemental data for provenancing by using empirical data but this requires high density sampling at present.

The rare earth elements give better separation using the Bray-Curtis distance than all of the trace elements collectively, as these are often found in fertilisers. This suggests a link to the topsoil chemistry being affected by the additions to the soil as well as the underlying geology.

Based upon the strontium isotope ratio and trace element data in a forensic investigation these analyses could be useful; it would be feasible to use these techniques in any investigation looking to determine the provenance of soil samples and also the comparison of suspect and control samples. The data could be used to exclude samples from each other; so if a forensic analyst has been given a suspect soil sample and five control samples

the chemical analysis could be used to exclude three or four of the control samples from the suspect sample based upon the strontium isotpe ratio and suite of trace elements. This chemical analysis would be implemented after any physical investigations of the soil had been carried out. However, if there was only a very small amount of soil present for analysis thought must be given to whether any other analysis such as palynology or grain size may be more beneifical to the investigation.

It must be remembered that from a forensic point of view, while looking at the macronutrients and trace element concentrations can be useful in helping to associate or disassociate sample locations this needs to be done with care, as the concentrations of the elements may alter over time through agricultural practice and further additions to the soil. So, any databases or maps created may be useful in a current context but would need to be continually updated if it was feasible to do so. The support of the farming community to aid in the understanding of what fertilisers have been added to what locations and when could also prove to be useful for forensic purposes.

In this research the topsoil has been used, this is because in a forensic context this is the most important soil horizon as this is the part of the soil which most frequently comes into contact with a suspect, or forensic exhibits can be found amongst the topsoil. However, future research needs to consider how the elements leach through the soil and if the trace element concentrations are indeed an accurate means of determining soil origin. This would be done by taking a soil aliquot from each of the different soil horizons for analysis; this would be of forensic interest in a case where the soil from a grave was being examined. In a forensic investigation it must also be remembered that a soil sample may not have a single provenance (unlike control samples), or may be a mixture of soil horizons and therefore care must be taken when making any similarity or exclusion conclusions.

Future work to expand the potential for the use of strontium isotope ratios and trace elements for forensic provenancing would include analysing a soil core and looking at the profile of the elements down through the soil, and see how this differs over time. Also it would be useful to return to the chosen sample locations and see how the strontium isotope ratios and trace element concentrations have altered over time; this would also help to look into the effects of seasonality. Another expansion of the work would be to compare the strontium isotope data with the trace element data from each of the different individual samples and then the different sites and see if this could be used to further

exclude any of the samples from each other. The combined datasets could be statistically analysed using the Bray-Curtis method which has been extensively discussed and the ordination plots used to see if any of the different sites could be excluded from each other, or if landuses / botanical region / geology could be excluded bases upon the combination of strontium isotope and trace element data.

Strontium Isotope Ratios and Trace Element Summary Box.

As yet the $^{87}Sr/^{86}Sr$ and trace element spatial data cannot systematically reproduce the spatial structure of the observations from first principles

REE elements give the best separation of the trace elements both according to geology and botanical region.

Although not all geologies or botanical regions can be distinguished from each other, certain ones can be.

Bray-Curtis statistics can be useful in showing site separation.

Chapter 5 – Norfolk Vegetation data

Norfolk vegetation records provided by Norfolk Biodiversity have been used to try to aid in the ability to predict where a soil sample has come from; the data has been used to see if any plants are unique or absent or if any suites of species are found to be from a specific area and subsequently can help provenance soil. It was also hoped that this information would be helpful in both the DNA and pollen analyses also carried out in the project.

The hypothesis is that there will be plant species unique to the character areas of Norfolk, species unique to all but one character area of Norfolk, and suites of species representing different habitats. It is also hypothesised that the vegetation data of each character areas will be reflected in the pollen data from chapter 6.

5.1 Norfolk Biodiversity

Norfolk Biodiversity was established in 1996 and is based at County Hall, Norwich. The Norfolk Biodiversity Partnership brings together local authorities, statutory agencies and voluntary groups in pursuit of a shared vision; "the conservation, enhancement and restoration of the counties biological diversity". The partnership is involved in a wide range of work including preparing and implementing action plans for some of Norfolk's most threatened habitats and species and managing a Biodiversity Project Fund which is used to support high priority recommendations contained in the species and habitat action plans. The partnership also works closely with the Norfolk Biodiversity Information Service to improve the quality and availability of any biodiversity information and promotes the integration of biodiversity into strategies, plans and programs; these include Local Development Frameworks and Sustainable Community Strategies.

A database of every species that has been recorded in Norfolk since 1985 has been provided by Norfolk Biodiversity; it is this data that will be used to see if it can be of use to aid in the forensic provenancing of soils for forensic investigations. This database is based upon vascular plants, and samples are recorded in 2 km by 2 km tetrads. On advice and discussion with Martin Horlock of Norfolk Biodiversity It was decided to gain all of the

records of plants recorded since 1985 as this should include all of the plants that are currently grown in the county and also information about those that were perhaps previously grown. It was important to gain information for this period of time as those plants that are currently grown could aid in a forensic investigation if for example they are only found in a small area of Norfolk. It was also though that the information about the plants that have previously been grown in the area may prove to be useful for the plant DNA and pollen aspects of the research, as even if the plants are no longer growing in the area their DNA and pollen has the potential to still be present.

Amongst other information, the plant database included details about who collected the data and when, and also both the Latin and common names of the plants and the longitude and latitude of where the plant was recorded.

It should be noted that within this research the Latin and common names of plant species are used interchangeably; in appendix 3, a list showing the translations between the two names can be found.

5.2 Different character areas of Norfolk

For the purposes of this research the character areas of Norfolk have been defined as Fenland, Greensand, Breckland, The Clays, Wensum Sands, Goodsands, Cromer Ridge, North East Norfolk and The Broads as shown in figure 71 and described in section 5:2.1. In the research the areas are interchangeably described as being the vegetation areas and botanical regions of Norfolk.

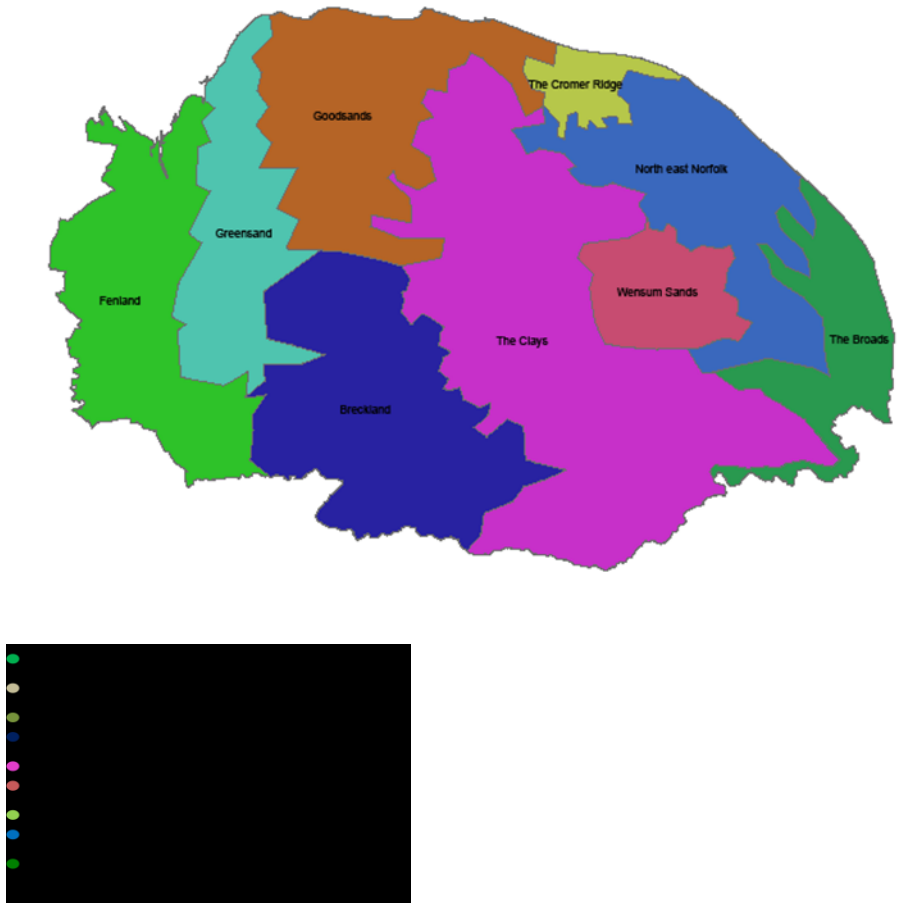


Figure 71: The different regions of Norfolk, chosen based upon their latitude, distance from the sea, soil type and drainage; which may give rise to the occurrence of different plant species.

5.2.1 How the botanical character areas were decided upon

Using the "A Flora of Norfolk" book by Beckett *et al*. (61) Norfolk plant species book, information provided by Norfolk Biodiversity and the Natural England website; it was decided that Norfolk could be divided into the nine character regions previously listed and shown in figure 10; this is based upon their latitude, distance from sea, soil type and drainage. It was thought that these areas may give rise to differing plant types. A more detailed description of each area, giving information about each of these parameters is given in section 5:2.2. As a result of these classifications, all of the botanical data provided by Norfolk Biodiversity could be divided into these vegetation areas and thus see if the data could aid in forensic investigations. This is a technique not previously employed in forensic

science research or casework to date, with limited local botanical knowledge usually available.

It should also be noted that these character areas of Norfolk have also been used in other aspects of the research (see chapter 4).

5.2.2 The individual character areas

There are ten different botanical character areas of Norfolk.

5.2.2 a Fenland

The Fenlands is divided into two areas; the Silt Fens and the Peat Fens which can each give rise to different vegetation types. The more northerly parts of this botanical area are made up largely of silt deposited by the rivers which flow through the wash (see figure in chapter 1 Norfolk Geography section). Behind the silt barrier lie several villages (between Kings Lynn and Wisbech), and south of these is Marshland. This Marshland is mainly silt but sometimes interspersed with saltmarsh representing the history of the area sometimes being inundated by the sea. Further south are the Peat Fens. In recent times these have been reclaimed and are intensively farmed with areas separated by dykes. In this area the only settlements are on former river banks or scattered remote farms positioned on firm ground. Botanically, the Peat Fens are now very poor in species, however, the flood lands of the River Ouse Wash (in the central to lower Fenlands between the Silt and Peat) are much more botanically rich.

5.2.2 b Greensand

In this area, sands and clays of the Lower Creataceous period overlap with the upper layers of the Jurassic period. Above the Sandringham sands lies iron-rich strata which continues to the Lower Greensands where it is hard enough to produce a local building stone. These deposits are free draining and associated with a heathland flora although now much of the area has been overtaken by trees through both self-seeding and commercial plantations.

5.2.2 c Breckland

The Brecklands adjoin the Fenlands in south-west Norfolk. In this area, chalk is covered by varying depths of sand meaning sharp drainage. In this area, the climate is closest to being continental, there are high summer temperatures, low rainfall and a tendency for frequent night frosts. There are large areas of afforestation, but until recently agricultural practice has been extensive rather than intensive. In the past twenty years, supermarket led agriculture has brought vegetable growing and subsequent irrigation to the area. Typical plants of the grass-heath, traditional of the Brecklands, now only survive on reserves, the wider range verges in the forest and in the large military training area.

5.2.2 d The Clays

The Clays cover a large area of central Norfolk where the dominant soil type is clay, being more acidic in the north but becoming more limey as they merge with the chalky boulder clays of south-east Norfolk. Much of this clay lies on Norfolk's watershed and often level ground can still become flooded if drainage is neglected. In the past, it was avoided using this land for agricultural purposes, especially in the least fertile parts. However, it is in this area that some of the greatest concentration of ancient woodland can be found, sometimes adjacent to the heathlands of The Cromer Ridge or Wensum Sands. The chalky boulder clay region in the south of the area is now mainly wheat, and has suffered more than any other area of the county in terms of hedgerow removal and the ploughing of its former wide verges. This had led to the decrease in what was once a distinctive flora which now only survives, rather perilously, on the few remaining areas of common land.

5.2.2 e Wensum Sands and The Cromer Ridge

These two different areas are a second region of acid sands which run from the coast southwards to just below Norwich. This two areas have a heathland flora and their history of landuse is very similar to the Greensands, but with soils of glacial origin.

5:2.2 f Goodsands

The Goodsands lie to the north of the Brecklands in north-west Norfolk. This area has many of the characteristics of the Brecklands, but has a marly layer between sand and chalk and digging into this layer and spreading it over the land at the time of enclosure gave rise to both body and nutrients in the soil. This method has now been replaced by fertilisers but the clay layer still prevents the soils from drying as rapidly when there are droughts. As this area is at a greater elevation and close to the coast it has a notably higher rainfall than the nearby Breckland or Fenlands. This area is now predominantly barley and sugar beet, but has been saved from becoming all arable land by game shooting which has meant hedgerows and copses have been maintained for pheasants and areas of grassland for partridges. Throughout the Goodsands there are several large parks with woods, grasslands and lakes.

5.2.2 g North East Norfolk

This is another area of Norfolk which has fertile soil and is highly agricultural although some of the many streams throughout the area still have grazing land beside them. Due to the areas proximity to the north-east coast it has a strong maritime air. This area is another which has lost a great deal of hedgerows and verges and it can be said it has no distinctive botanical character away from the coast.

5.2.2 h The Broads

The Broadland is arguably the best known part of Norfolk. The botanical interest in this area lies mainly in the areas of fen and marsh which lie on the level ground close to the rivers. Several of the areas are still traditionally cut for reed and used for grazing. Some areas have been drained and are now arable land bringing increasing problems with nutrient rich water reaching the dykes, broads and rivers. Many of the largely manmade broads, have now become polluted, particularly with nitrogen from arable land and phosphorus from the sewage works. (It should be noted that this may be reflected in trace element data from samples taken from area.) Found in the Broads are the Halvergate Levels, a large area of grazing land which lies between the main rivers and Breydon Water.

Efforts have been made to keep this area as grassland and some of the area may be bought by conservation bodies.

5.2.3 Other vegetation information – habitats

Habitat information for Norfolk was obtained as it was felt that this may yield some information as to what plant types may be present in certain areas and give a starting point for looking for suites of species which may be unique to different habitats. It was also thought the habitat information could be used alongside the pollen data as the land type was known at each of the sample points and also the pollen species present and thus is could be seen if any parallels between the two could be drawn.

5.2.3a Woodlands and Scrub

Compared to many other parts of England, Norfolk has a low proportion of woodland cover and a large percentage of this can be found in the recently planted conifer plantations in Breckland. However, Norfolk still has surprisingly diverse woodland.

Oak woodlands are particularly developed on the sands and gravels of the North Coast, and are associated with species *Quercus petraea, Sorbus aucuparia* and *Ilex aquifolium.* The flora on the ground near to oak woodland is often associated with bluebells, bracken and *Convallaria*.

Beech woodland occur at the Cromer-Holt ridge at Fellbrigg and Northrepps. At these areas the ground flora is mainly absent; with brambles occurring in some places.

Mixed broadleaved woodland encompasses the majority of the woodland found in the Clays. The flora of these woodlands is often dominated by *Mercurialis perennis, Primula vulgaris, Hyacinthoides non-scripta, Anemone nemorosa,* and *Allium ursinum.* Rarer species found are *Gagea lutea, Melampyrum cristatum, Paris quadrifolia, Platanthera chlorantha* and *Campanula trachelium.*

Alder woodland is found in Swanton Novers and in the floodplains of the Broadlands. The woodlands are dominated by *Alnus glutinosa* with *Urtica dioica* abundant in the herb layer.

The number of scrubland communities is widespread and includes species such as *Crataegus* and *Myrica.*

5.2.3b Grassland

Chalk Grassland

There is very little typical chalk grassland in Norfolk despite chalk underlying 60% of the county. The main sites where it is found is north-west Norfolk and the Brecklands. The main chalk communities are *Festuca ovina-Avenula pratensis* associated with the species *Scabiosa columbaria, Briza media, Carex flacca, Anacamptis pyramidalis, Asperula cynanchica, Hippocrepis comosa* and *Linum catharticum.* In the Brecklands the *Festuca ovina-Hieracium pilosella-Thymus pulegioides* community is characteristic.

Mesotrophic Grassland

Areas of mesotrophic grassland occur in the heavier clays of south east Norfolk. Here the *Cynosurus cristatus-Centaurea nigra* community is a major component with *Centaurea nigra, Ophioglossum vulgatum, Silaum silaus, Trisetum flavescens, Leucanthemum vulgare* and *Primula veris* species present. The habitat can also be found in the sandy, slightly calcareous soils of the Brecklands with *Bromus hordeacus* abundant and *Artemisia campestris* present

Acid Grassland

Acid grassland is mostly found in the Brecklands with variants of the *Festuca ovina-Agrostis capillaris-Rumex acetosella* (a grassland community) being widespread.

5.2.3c Heathland

Heathland is defined as vegetation dominated by *Calluna, Erica* or *Ulex* species. In Norfolk, it is developed on the sandier parts of the Greensands, and on the glacial gravels and sands in the Cromer Ridge area. Areas of *Calluna* heath can also be found amongst the grass heaths of the Brecklands. In the north-east of Norfolk the heaths are largely dominated by *Calluna vulgaris* and *Ulex gallii*; this is more commonly found in western Britain. In parts of

the Brecklands and on the Greensand ridge the dry heathis dominated by *Calluna vulgaris-Festuca ovina* which is likely to be a reflection of the dry climate in the Brecklands. However, it is not easy to explain the difference between the Greensand heaths and those found in north-east Norfolk. One possible explanation is that some of the north-eastern heaths have a clay deposit below the sand and gravel which may result in a less well drained, more humid heath. In many areas, a recent lack of management has led to a succession of birch and oak woodland. Often, these drier heaths can down grade into wet heaths, these are characterised by *Erica tetralix-Sphagnum compactum*. In areas such as Roydon, Sheringham and Beeston Common, the wet heath can grade into valley mires.

5.2.3d Wetlands

Although the total rainfall in Norfolk is low, the county has a wealth and diversity of wetlands which are probably unmatched by any other area of lowland Britain. This wealth of wetland is due to topography and geology, with chalk overlaid by a complex series of acid and calcareous drift deposits. This results in there being many National Vegetation Communities (NVC) (179) present. These can be broadly grouped into flood plain wetlands, which are primarily irrigated by a river, and valley mires, which occupy slopes as well as the valley floor and are maintained primarily by seepages and springs. The main floodplain wetlands are concentrated in the Broadlands. The valley mires can be found scattered throughout Norfolk, but especially near the headwaters of the main rivers and in the Greensands area.

Broadland Floodplain Fens and Swamps

The herbaceous fens of the Broadlands form the largest expanse of species-rich fen in lowland Britain. They have been described by Wheeler (1978) (180); it is particularly interesting to note the *Phragmites australis-Peucedanum palustre* community and its' sub communities, which are almost wholly confined to the broads. This area covers over 550 hectares, which is approximately 20 % of the whole fen area, and a number of rare species are associated with this area; *Peucedanum palustre, Cicuta virosa, Sium latifolium and Ranunculus lingua*. In addition, fen meadow communities occur; these are also found on the peaty soils of the drained marshes. Common species are *Juncus subnodulosus-Cirsium*

palustre covering 110 hectares, *Molinia caerulea-Cirsium dissectum* covering 34 hectares and *Schoenus nigricans-Juncus subnodulosus*.

Valley Mires

The Valley Mires of Norfolk are matched only by those in Anglesey and the Lleyn peninsula in North Wales. They can be sub-divided into three groups: calcareous valley mires, mixed valley mires and acid valley mires. However, it is best to consider these sub groups as a continuum from acid to calcareous depending upon the nature and source of the ground water feeding the site. The valley mires generally have a wide range of vegetation types with some zonation from the driest to wettest areas. The zonation is reflected in the vegetation with areas of tall herb or swamp vegetation, mire and fen meadow.

The mixed mires arise due to the complex nature of the drift geology over the chalk which means that water derived from several sources can feed a site. As a result, the mires exhibit a perplexing range of calciphilous and acidphilous communities alongside one another. As mentioned previously, the communities include mire, fen meadow and tall herbaceous fens. The most extensive mire community is *Schoenus nigricans-Juncus subnodulosus* which is of major importance for its floristic richness with species such as *Carex dioica, Pinguicula vulagris* and *Epipactis palustris*. Other communities include *Carex rostrata-Calliergon cuspidatum* and *Carex diandra-Calliergon giganteum* which is confined to wet ground. The *Juncus subnodulosus-Cirsium palustre* is found in the fen meadows and is characteristic of the drier parts of the calcareous flushes. Tall herb fen communities are often associated with a lack of grazing and include the *Filipendula ulmaria-Angelica sylvestris*.

Acid Mires

The valley mires which are developed on the Greensands are more acidic than those elsewhere; although the vegetation can still be influenced by patches of chalky boulder clay such as at Roydon Common. The most acidic mires such as Roydon Common and Dersingham have a series of communities which are well developed with *Sphagnum auriculatum* and *Sphagnum cuspidatum / recurvum* bog pool communities present in the wettest areas. This gives rise to species such as *Eriophorum angustifolium* and *Drosera rotundiflora.* Where there are areas of flushed calcareous groundwater the rare *Schoenus nigricans-Narthecium ossifragrum* community occurs. Species associated with this

community include a rich bryophyte layer with *Aneura pinguis, Sphagnum subnitens* and *Campylium stellatum.* At both Roydon Common and Dersingham, the *Narthecium ossifragrum-Sphagnum papillosum* community is well developed with species such as *Carex lepidocarpa, Potamogeton polygonifolius, Drosera angelica* and *Menyanthes trifoliata.* Two rare species which are associated with this community are *Rhynchospora alba* and *Hammarbya paludosa*.

5.2.3e Open waters

Open water communities occur throughout the Broads (although these are affected by pollution) and in the dyke systems of the fens and grazing marshes. The river systems are diverse, ranging from calcareous fast-flowing chalk streams to slow-flowing ones in the fens. The dyke systems found in the Broadland fens and some of the grazing marshes support the *Potamogeton natans, Hydrocharis morsus-ranae-Stratiotes aloides* and *Spirodela polyrhiza-Hydrocharis morus-ranae* communities. Rare species associated with these communities include *Stratiotes aloides* and *Potamogeton acutifolius*.

The Breckland meres provide a unique habitat of fluctuating water bodies which can be dry for several years. These areas generally exhibit a clear zonation and are often dominated by sands of *Myosoton acquaticum* in dry periods. When the beds are dry large areas of the grasses *A. Geniculatus* and *Alopecurus aequalis* can be found. Along the margins of some of these meres, the rare moss *Physcomitrium eurystomum* can be found. Different areas of the rivers Nar, Tad and Wensum can also support communities; the fast flowing middle sections of the Nar and Tad and the upper section of the Wensum support communities with the species *Ranunculus pencillatus ssp. pseudofluitans* and found only on the Nar *Hippuris vulgaris*. On the river Wensum, the fast slowing sections below mills are dominated by communities which *Potamogeton perfoliatus* and *P. pectinatus* are dominant, and in the slower sections above mills there are *Nuphar* and *Sagittaria sagittifolia* dominated communities.

5.2.3f Coastal

Norfolk has approximately 1100 hectares of sand dune systems; these are located in two main areas between Holme and Blakeney Point on the north coast and between Eccles and Great Yarmouth on the east coast. One of the features of the north coast dunes is that there are low dunes. The dune systems range from being slightly calcareous at Holme to being very acidic at Winterton. Although in the remit of this research no soil samples have been collected from the sand dunes situated on the Norfolk coast, the vegetation data from these areas is still being considered as to whether the coastal areas have a unique flora which may be utilised for forensic purposes.

Drift line and foredunes

Along the north coast there are well established communities such as *Cakile maritima* and *Honkeyna peploides* and also other annually occurring communities such as *Atriplex laciniata* and *Salsola kali*.

Semi fixed dunes

The more fixed dunes are dominated by communities such as *Ammophila arenaria-Festuca rubra*.

Dune grasslands

In small areas of a number of sand dunes the *Festuca rubra-Galium verum* community can be found. This community has characteristic species such as *Calium verum, Polypodium vulgare, Crepis capillaris* and *Pilosella officinarum.* At Blakeney Point the rare grass *Corynrphorus canescens* can be found. There are areas of *Ligustum vulgare* scrub which occur generally throughout the dune grasslands of Norfolk, and at Hunstanton and Holme there are areas of *Hippophae* scrub. At Holkham, a large proportion of the dunes were planted in the 18th and 19th centuries, and now support naturally regenerating *Pinus nigra ssp. maritima* and *Quercus ilex* woodlands. In these woodlands, *Monotropa hypopitys* and *Goodyera repens* occur.

Dune heath

At Winterton the largest area of this habitat can be found; it is characterised by *Calluna vulgaris*, also commonly present are *Festuca ovina* and *Carex arenaria* both of which are

frequently occurring and are co-dominant. Lichens are bryophytes can also be frequently occurring.

5.2.3g Shingle

There are extensive areas of shingle habitat between Weybourne and Blakeney. The habitat also often underlies dunes elsewhere of the North Norfolk coast. Between Cley and Weybourne the shingle supports an open community with little other than *Glaucium flavium*. Where there is more sand *Honkenya peploides* and *Silene uniflora* is developed with the species *Sedum acre, Phleum arenarium, Glaucium flavum* and *Ammophilia arenaria.* Where there is a higher sand content in addition to the above species *Carex arenaria* can be found. At Snettisham this species is present with *Dicranum scoparium, Festuca rubra* and *Plantago lanceolata* can be found.

5.2.3h Saltmarsh

Norfolk has approximately 2800 hectares of salt marsh. This is concentrated in three main areas, The Wash, North Norfolk Coast and Breydon Water. Communities on The Wash are *Spartina, Suaeda-salicornia, Aster* and *Puccinellia.* Within this habitat in North Norfolk, nationally scarce communities can be found; *Limonium-Armeria, Suaeda vera* and *Frankenia laevis Limonium bellidifolium.*In summer *Limonium* can be found.

5.2.3i Cliffs

The coastal cliffs between Happisburgh and Weybourne are mainly composed of Pleistocene sands and clays. The cliffs are relatively vertical where little water drains from them although they are still prone to erosion. Between Cromer and Mundesley where water issues from the cliffs they are unstable with frequent landslides which creates a mosaic of habitats from scrub with *Hippophae* near Cromer, through relatively stable areas of grassland which are often dominated by a community in which *Festuca rubra* and *Anthyllis vulneraria* are dominant to bare mud which is colonised by species such as

Tussilago farfara. In places *Phragmites* beds occur in area on impeded drainage on the cliffs. The rare *Orobanche purpurea* can be found on the cliffs but is largely confined to the mesotrophic grassland on the cliff tops and never occurs on the cliff slopes.

5.2.3j Common Lands

DEFRA carried out a survey of the common lands of Norfolk; this forms part of a wider study of the inventory of commons in England and Wales (181). The main objective of the study is to draw together an array of biological data from a variety of sources including mapping of vegetation and habitat data. The main habitats found in the common lands of Norfolk are summarised in table 18. It is important to give consideration to this habitat data as the individual habitats may give rise to specific combinations of plant species which are indicative of and unique to the habitat. Also, some of the samples; particularly those taken from Roydon Common, are taken from common lands.

Table 18: Summary of habitats recorded during a DEFRA survey. 170 sites were included in the survey, and 53 different habitats were recorded on 4460 hectares of land.

Habitat	**Area (hectares)**	**Number of Commons**	**% Common Land**
Woodland and Scrub	814.0	321	18.3
Grassland	852.2	220	19.2
Tall Herb and Fern	184.7	108	4.2
Heathland	104.0	19	2.4
Bog, Flush and Fen	144.2	59	3.3
Open Water	94.0	48	2.2
Coastland	2039.7	41	45.6
Anthropogenic Habitats	224.4	150	5.0
	4457.2	966	100.2

Of the woodland habitat of the common land of Norfolk, established and emerging semi-natural broadleaved woodlands (scrub) constitute the largest part of this habitat. Other

commons include planted woodland, namely coniferous woodland found in eight areas and broad-leaved woodland found in ten areas. Flordon Common is an example of a semi-natural broad-leaved woodland and has *Alnus glutinosa* dominated woodland around fenland. Other species found in this habitat are *Betula* dominated heathland in Snettisham and *Quercus robur* dominated woodland over acidic grassland in the Roydon Common.

A wide range of grasslands can be found on the Norfolk Commons. Calcareous grassland is arguable the least evident grassland but the most important. This is found at Flordon Common on higher ground above calcareous spring-fed fenland. Other grassland habitats are semi-improved calcareous grassland, marshy grassland, and improved grassland.

In the tall herb and fen habitats dense bracken is found in 40 of the commons. In these habitats bracken is invading open areas especially former heaths.

5.3 Methods

5.3.1 How the Norfolk vegetation data was used / manipulated

The database file obtained from Norfolk Biodiversity was opened and any manipulations made in Microsoft Excel 2007. Any maps were made in Esri Arc Map 9.0. An overview of how this data was subsequently used is given below.

The data was first sorted into the character areas of Norfolk; this allowed one to see if the data points had been recorded accurately, and if not allowed for these to be removed.

The data has also been split into those plant species recorded before 1990 and those recorded after 1990, to see if there is any difference over time in the prevalent species in Norfolk. It was decided to do this as land use inevitably changes over time and this will have an impact on the species that may be found in an area. Historically, man has farmed within the limits imposed by Norfolk's soils; there were no attempts to grow water-demanding crops in the dry Breckland soils for example. It has only been in the last sixty years that a combination of economical and technical advances has allowed for the landscape to change due to farming. The other important historical impact upon the Norfolk landscape is ice advances. The last ice advance left a landscape which was far from uniform, and depended largely upon the moisture content and acidity and alkalinity of the

soils. This is the state we now call natural vegetation. Since the ice advance, the history of the vegetation is a result of man's efforts to use or change the landscape to his own advantage with he has done with increasing success (61).

The habitat information has been used in this study to see if the expected species have been recorded in the specific areas.

In order to see if the data was useful for a forensic investigation the database was also used to see which species were unique to certain character areas and which species were found in all but one of the character areas. It was also seen if there were suits of species that were unique to the different areas.

For more details on how the data has been used and the results of the analyses see section 5:4.

5.3.2 GIS mapping

ESRI Arc Gis 9.0 has been used to create the maps showing the Norfolk vegetation data. It was decided to use the maps to show the data alongside tables, lists and graphs as the visual images can help to portray the information in a clearer, more concise, manner than numbers. It was also thought that if this information was to ever be used in a forensic investigation and eventually a jury the maps created by Arc Map can instantly show information that may take much longer to explain using a complementary table; that is not to say for explaining some of the data tables etc are not more appropriate.

5.4 Results

The data provided by the Norfolk Biodiversity has been analysed and is presented below in a variety of different graphs, maps and tables.

5.4.1 Graphs and Maps

Firstly all of the data acquired from Norfolk Biodiversity was overlaid onto a character map of Norfolk showing the nine different character regions using Esri Arc Gis 9.0 as seen in

figure 11, the data was then joined to the map and exported as an spreadsheet file so that it could be determined in which character area the plant species occurred. The base layer character map and the plant data were both projected using (WGS 1984 projection). As can be seen in figure 72 there are obviously some discrepancies in the recording of the plants as some are recorded in the North Sea. The outliers were removed before any further data manipulation took place.

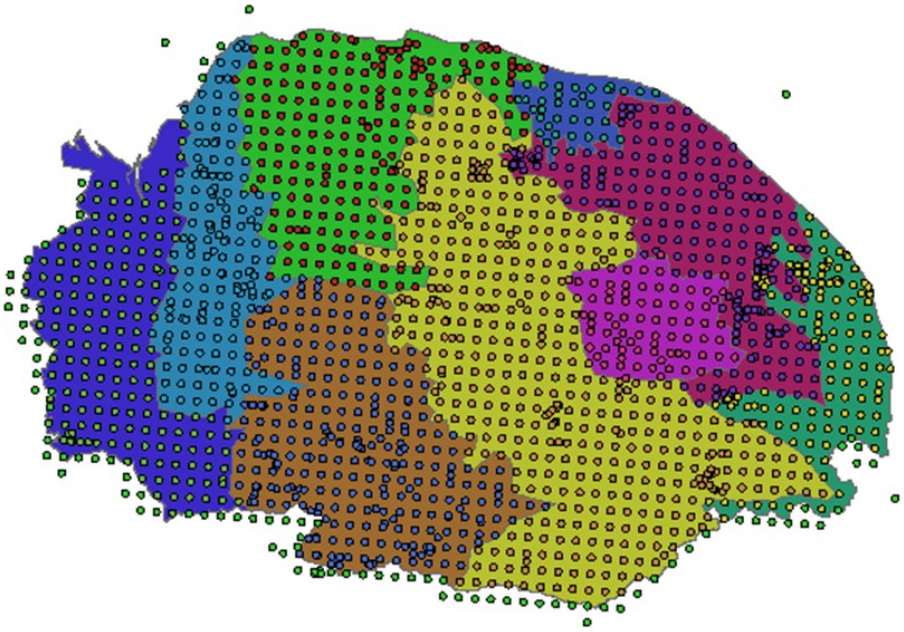


Figure 72: All of the tetrad data provided by the Norfolk Biodiversity overlaid on the different botanical regions of Norfolk; from left to right the character areas are shown by dark blue Fenlands, light blue Greensand, brown Brecklands, light green Goodsands, yellow The Clays, lilac The Cromer Ridge, dark purple N.E. Norfolk, light purple Wensum Sands and dark green The Broads.

As a simple look at how many different plant species occurred in the botanical regions, any duplicated records were removed as shown in figure 73 below. The data was then summarised simply (table 19) to show the percentage of unique plant species in each area.

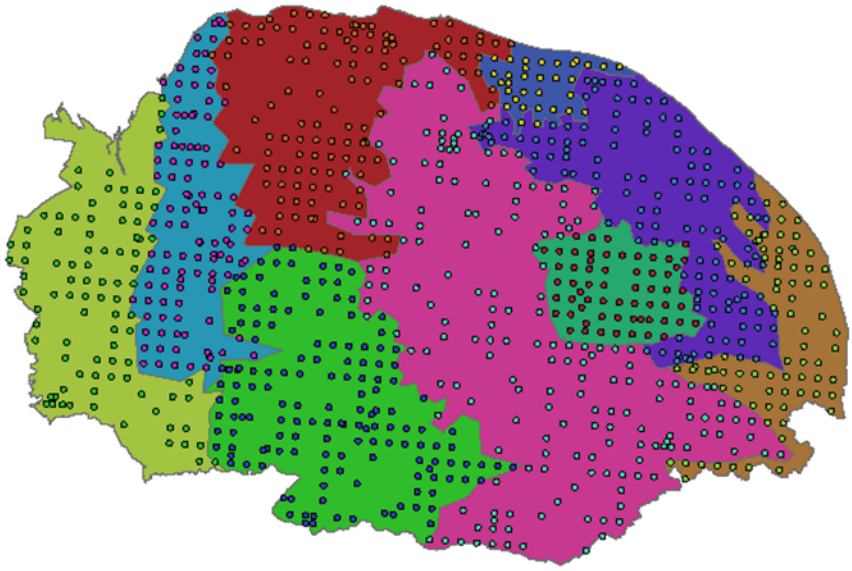


Figure 73: Map to show the number of different plant species in each area; the coloured dots represent the areas in which the different plant species were recorded overlaid onto each of the botanical character areas of Norfolk from left to right the character areas are shown by olive Fenlands, light blue Greensand, green Brecklands, red Goodsands, pink The Clays, dark blue The Cromer Ridge, purple N.E. Norfolk, jade green Wensum Sands and dark brown The Broads.

Table 19: Summary of the number of plants, different plant species and the unique plant species recorded in each of Norfolk's botanical regions.

	Total Plants	Total Species	Unique Species	% unique species
Fenlands	5016	228	5	2.19
Goodsands	6123	266	7	2.63
Greensands	5849	282	15	5.32
North East Norfolk	7769	288	8	2.78
The Brecklands	10239	272	15	5.51
Broads	3880	259	12	4.63
Clays	19066	322	14	4.35
Cromer Ridge	1674	240	2	0.83
Wensum Sands	3178	232	3	1.29

As can be seen in table 19 the areas with the greatest numbers of recorded plant species are the Clays and The Brecklands. The areas with the least number are Wensum Sands and Cromer Ridge, however, these areas are the smallest geographically so this is perhaps not surprising. Also Wensum Sands covers Norwich and it's surrounding area and therefore there will be less countryside.

5:4.2 Unique Species

The numbers of unique species are shown in table 19, are are listed in the following table (table 20). These unique species are of forensic importance as if they are present in a soil sample (whether in a botanical, pollen or DNA context) they can help to narrow down an area from a regional level to a more localised botanical region.

It must be remembered that although a species might be unique, it may not appear throughout the botanical region and may only appear once, or, the species may only have been recorded once and not ever been recorded again.

Table 20: The species which are unique to each of the botanical regions; these could be used to determine where a sample has come from.

Wensum Sands	**Brecklands**	**Fenlands**	**Goodsands**	**Greensands**	**Clays**	**North East Norfolk**	**Cromer Ridge**	**Broads**
Deodar	Adders Tongue	Autumn Millet	Bristle Oat	Bearded Crouch	Coastal Redwood	Blue-Eyed Grass	Wood Horsetail	Alpine Rush
Italian Garlic	Bog Arum	Grass	Corsican Pine	Bog Asphodel	Dactylorhiza	Cape Pondweed	Yellow-Eyed-Grass	Baltic Stonewort
Swamp Meadow Grass	Common Whittowgrass	Large-Flowered Waterweed	Drooping Brome	Bog Orchid	Dawn Redwoos	Eastern Gladiolus		Bulbous Foxtail
	Eastern Hemlock Spruce	Marsh Clubmoss	Elymus	Wintoni	English Iris	Greater Quaking Grass		Convergent Stonewort
	Garen Daffodil	Ribbon Fern	Frog Rush	Greater Red-Hot-Poker	Golden Crocus	Horsetail		Field Garlic
	Hard Fescue		Maritime Pine	Hairy Finger-Grass	Highland Bent	Hosta		Great Tassel Stonewort
	Maidenhair Fern		Sharp Rush	Hermerocallis	Japanese Red-Cedar	Irish Yew		Intermediate Stonewort
	Narrow Small-			Hybrid	Neapolitan	X Cupresso		Lesser Bearded

	Reed			Sweet-Grass	Garlic	Cyparis		Stonewort
	Nonesuch Daffodil			Juncus	Ostrich Fern			Opposite Stonewort
	Nootka Cypress			Man Orchid	Ptendophyta			Pillwort
	Oak Fern			Molinia	Rye Brome			Rough Stonewort
	Onion			Pale Yellow-Eyed-Grass	Shama Millet			Starry Stonewort
	Rugged Stonewort			Pinaceae	Small Flowered Hair-Grass			
	Tassel Hyacinth			Pinus	Thin-Spiked Wood Sedge			
	Yellow Star-Of-Bethlehem			Scentless Mayweed ag.				

As well as species that are present in only one of the character areas of Norfolk, it was investigated if any species were only found in two of the areas as it was thought that this could also prove to be useful in helping to identify where a soil sample has come from. It was found that there are 35 plants, which are found only in two character areas of Norfolk, as shown in table 21. This information may not give a definitive conclusion as to where a soil sample has come from based upon the plant species present in the same way as a unique species may give intelligence. For example, if a specific species was known to only be found in two areas, and a crime had been committed at a location known to be in a different character area, and traces of the plant was found in a soil sample taken from a suspect it could be determined that they may have been in one of the two locations where the plant species is known to have occurred / be occurring and not necessarily at the crime scene. Conversely, if a plant species was present in a sample it can aid in placing a suspect at a crime scene.

Table 21: Plant species that are found in only two of the character areas of Norfolk.

Species	**Area species is found in**
Ammophila	Goodsand, Greensand
Autumn Ladys-Tresses	Goodsands, Cromer Ridge
Awned Canary Grass	Greensands, Clays
Birds Nest Orchid	Clays, Cromer Ridge
Cedrus	Goodsands, Wensum Sands
Chewings Fescue	Goodsand, Greensand
Couch	Fenlands, Goodsands
Delicate Stonewart	Breckalnds, Greensands
Dog's-Tooth Violet	North-East Norfolk, Clays
Fen Orchid	North-East Norfolk, Broads
Floating Water-Plantain	North-East Norfolk, Broads
Fragile Stonewort	Greensands, Broads
Galingale	Goodsands, Clays
Green Algae	North-East Norfolk, Broads
Green Flowered Helleborine	Clays, Wensum Sands
Hares Tail Cottongrass	Goodsands, Greensands

Martagon Lily	Clays, Wensum Sands
May Lily	Clays, Cromer Sands
Monterey Cypress	Broads, Clays
Moonwort	Broads, Goodsands
Pale Galingale	Broads, Wensum Sands
Pampas Grass	North-East Norfolk, Clays
Pheasants Eye	Clays, Wensum Sands
Picea	Brecklands, North-East Norfolk
Purple-Stem Cat's Tail	Brecklands, Fenlands
Rare-Spring-Sedge	Brecklands, Fenlands
Rescue Brome	Brecklands, Clays
Small-Fruited Yellow Sedge	Greensands, Cromer Ridge
Smooth Stonewort	Greensands, North-East Norfolk
Stags Horn Clubmoss	Greensands, Cromer Ridge
Swamp Cypress	Goodsands, Clays
Tassel Stonewort	Goodsands, Broads
Weymouth Pine	Broads, Clays
Wild Tulip	Clays, North-East Norfolk
Yellow Crocus	Goodsands, Clays

5:4.3 Species present in all but one area

If a sample is not present in one area and is present in all other areas, in a forensic context this is equally as important as a unique species as it can indicate a sample may be found in several different areas and therefore not forensically useful, or, may not be from a specific area. For example as shown in table 22, smooth meadow grass is found everywhere across Norfolk except in the Broads. If a soil sample was found and had meadow grass in it, it could be determined that it was unlikely that the sample originated from the Broads. Although when looking at figure 71 it may appear that this information only eliminates a small area of Norfolk it can help police intelligence in an investigation.

Table 22: The species that are present in all but one character area; the presence of the species is indicated by a *.

	Clays	Wensum Sands	Brecklands	Fens	Goodsands	Greensand	N. E. Norfolk	Broads	Cromer Ridge
Bottle Sedge	*	*	*		*	*	*	*	*
Broad-Leaved Helleborine	*	*	*	*	*		*	*	*
Bulbous Rush	*	*	*	*	*	*		*	*
Butchers Broom	*	*	*	*	*	*	*	*	
Common Cottongrass	*	*	*		*	*	*	*	*
Creeping Bent	*	*	*		*	*	*	*	*
Deergrass	*	*	*	*		*	*	*	*
Distant Sedge	*		*	*	*	*	*	*	*
Flattened Meadow Grass	*	*	*	*	*	*	*	*	
Flowering Rush	*	*	*	*		*	*	*	*
Giant Fir	*	*	*		*	*	*	*	*

	Clays	Wensum Sands	Brecklands	Fens	Goodsands	Greensand	N.E. Norfolk	Broads	Cromer Ridge
Grape Hyacinth	*	*	*	*	*	*	*	*	
Mat-Grass	*	*	*		*	*	*	*	*
Meadow Barley	*		*	*	*	*	*	*	*
Narcissus	*	*	*		*	*	*	*	*
Ramsons	*	*	*	*	*	*	*		*
Royal Fern	*	*	*	*	*	*		*	*
Sand Sedge	*		*	*	*	*	*	*	*
Sae Crouch	*		*	*	*	*	*	*	*
Smooth Meadow Grass	*	*	*	*	*	*	*		*
Spring Sedge	*	*	*	*	*	*		*	*
Star Sedge	*	*	*		*	*	*	*	*
Tufted Sedge	*	*	*	*	*	*	*	*	

Western Hemlock	*	*	*		*	*	*	*	*
Wild Onion	*		*	*	*	*	*	*	*
Total Absent Species	0	5	0	8	2	1	3	2	4

5:4.4 Suites of Species

As the soil samples were primarily collected on the basis they would have different geologies; not too much detail was taken about the habitats. However; it was decided to look at the samples based upon the land use types used for the trace element data analysis as seen in chapter 4. Seven samples were found to be from forested areas, three from parkland or a SSSI, two from scrubland, one from a quarry and the remaining samples from agricultural land.

However based upon the description in section 5:3.2 it was decided to see where in Norfolk the common habitats occurred and what species were expected and whether this was reflected in the data.

Table 23 shows the different suites of species which are indicative of certain areas. These suites of species can be compared to the pollen data to see if the recorded species data is reflected in the pollen samples. These suites of species may also be important to forensic science in more of a "blind" investigation. For example, if an unknown soils sample contained *Calluna, Erica* and *Ulex* it could suggest that the sample is from an area of heathland.

Table 23: Suites of species indicative of different areas / habitats

Area	**Plant species present**
Forested area and scrub - North Coast	*Quercus petraea* *Sorbus aucuparia* *Ilex aquifolium* *Convallaria*
Beech Woodland - Cromer-Holt Ridge	Brambles
Clays	*Mercurialis perennis* *Primula vulgaris* *Hyaccinthoides non-scripta* *Anemorle nemorosa* *Allium ursinum*
Broadlands	*Alnus glutinosa* *Uritca dioica*
Scrubland	*Crataegus* *Myrica*
Grassland - North-West Norfolk & Brecklands	*Festuca-ovina*
Grassland - South-East Norfolk	*Cynosurus cristatus*
Heathland	*Calluna* *Erica* *Ulex*

5:4.5 Date separated plant species

As discussed in 5:3.1 the plant species data was divided into those species which were recorded before 1990 and those which were recorded after in order to see if there were any differences in the recordings. As can be seen in figures 74 and 75 there is a clear separation; pre 1990 most of the plant species are recorded in the west of the county, and post 1990 the majority are recorded in the east. Pre 1990 there is very minimal sampling in the east of the county with most occurring in the Fenlands and Greensands, partially in the Goodsands and Brecklands, a small number of recordings in the Clays and a few sporadic recordings further east. As can be seen in figure 74 there is more sampling across the whole of the county compared to pre 1990 (figure 75). However the majority of the sampling is clustered over the west of the county with less extensive sampling in the east; there recordings are in the Cromer Ridge, North East Norfolk, the Broads and Wensum Sands, the majority of the Clays and sporadic recordings in the Fendlands, Greensands, Goodsands and the Brecklands. The main difference in the two different time segments is that pre 1990 in some of the character areas (Cromer Ridge and Wensum Sands) there is just one plant recording, whilst post 1990, there is a minimum of twelve recordings in each of the character areas.

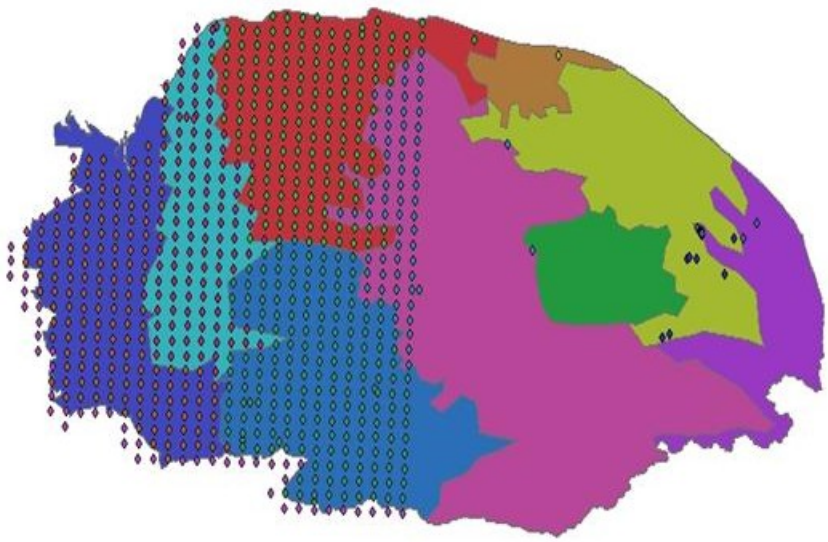


Figure 74: Plant species data recored before 1990 shown on the botanical character areas.

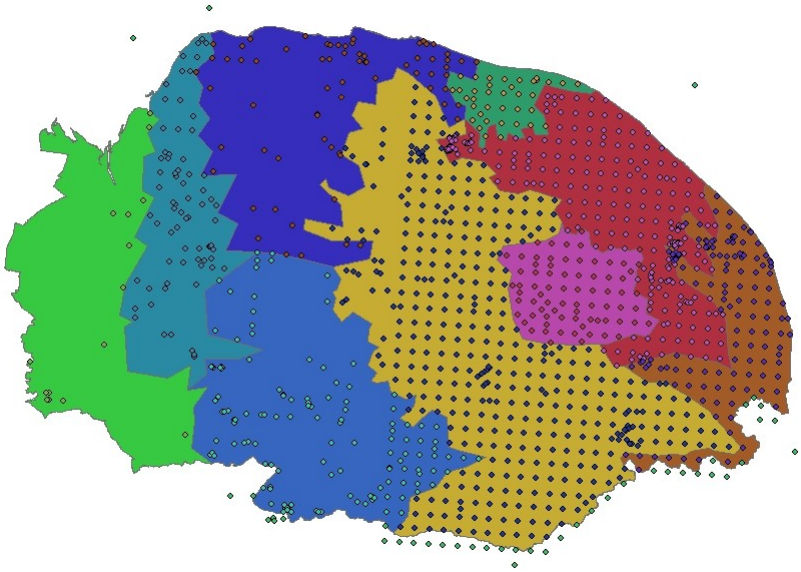


Figure 75: Plant species data recored after 1990 shown on the botanical character areas.

Once the plant species data had been separated into pre and post 1990 recordings it was decided to look at the number of plants and the total number of species recorded. As can be seen in figures 76 and 77; the most plants have been recorded after 1990; this is promising for this research as it means there is a greater chance of the plants still being present in the areas and thus being able to aid in the identification of the origin of a soil sample.

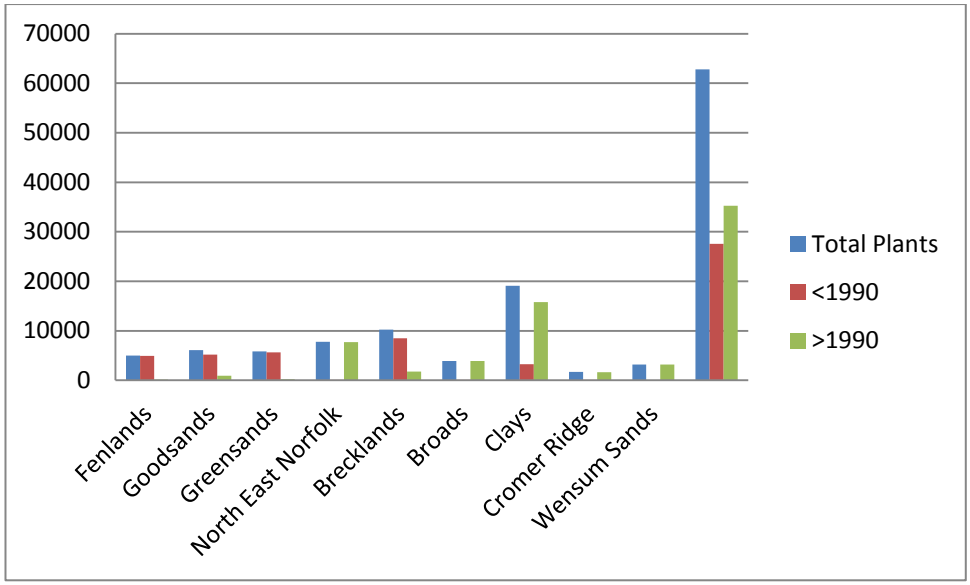


Figure 76: Barchart showing the total number of plants recorded in Norfolk, the plants recorded prior to 1990 and those recorded after 1990.

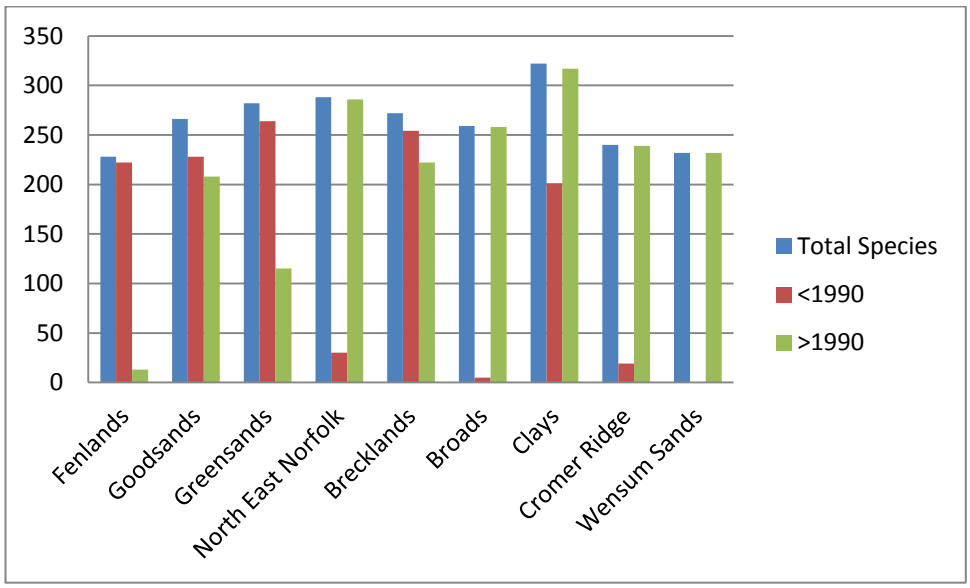


Figure 77: Barchart showing the total number of plant species recorded in Norfolk, the plants recorded prior to 1990 and those recorded after 1990.

5:4.6 Most Common Plant Species

Whilst the focus of this research has been on the plant species which may be able to help identify where a soil sample has originated from, some note must be taken of those plant species which occur prevalently across the whole of the county. There are 135 plant species which can be found throughout Norfolk; some of these are common throughout the whole of the county such as Perennial Rye-Grass and some, are more predominant in one area but do occur throughout the whole county. For example, Timothy is most predominant in Breckland but is found throughout Norfolk. The most common plant species are shown in appendix 4.

5.5 Discusssion

The database of plants recorded in Norfolk since 1985 and the subsequent analysis has shown that they can be useful in a forensic context; however there are limitations. The presence of absence of a plant species or suite of species can help to show if a sample can possibly have originated from a specific area. This is discussed in more detail below.

The Norfolk vegetation data can be compared to the pollen data set to see if the species recorded in the pollen data match with those recorded by Norfolk Biodiversity (see Chapter 6). The plant species data could provide an insight into what species to look for in DNA profiles if you were to amplify a sample for specific species.

Using the data; to some extent all of the character areas can be distinguished based upon the plant species which can be found there. However, it must be considered that the plant species may not grow throughout the area so care must be taken so that the plant data is not misleading as mentioned in section 5:4.

In a forensic context vegetation data can be used to give intelligence to a forensic investigation. If any fragments of plant species were found in a soil sample and they were identifie, they could then be compared to the database and any areas of Norfolk where the

plants were not found could be excluded from the investigation. A similar principle can be applied if the vegetation databases are used alongside the pollen data; if the pollen data cannot provide a conclusive link between samples then the pollen species found in a sample can be compared to a database and areas which do not contain those plant speces excluded from the investigation. However, when taking this approach thought must be given to whether the sample is from a single provenance or if it has been mixed and formed from several different locations. It must also be remembered that whilst this data is available for Norfolk this might not be the case for all of the UK, and the databases may not always be as up to date and as well maintained as others.

5:5.1 Unique Species

As is highlighted in table 20, there is a low percentage of unique species in each of the character areas of Norfolk. This is not too unusual as Norfolk is not geologically too diverse (see introduction chapter and Sr and TE chapter for more detail on Norfolk geology) and subsequently does not give rise to greatly differing soil types and thus similar plants may be found across the county. The differences throughout the county are subtle; but this is not to say the differences are not of forensic importance.

5:5.2 Absent Species

There are some character areas of Norfolk which cannot be discriminated upon the premise of an absent species; namely the Clays and the Brecklands as shown in table 22. However the rest of the character areas have at least one species which can be found everywhere else in Norfolk apart from that specific area.

5:5.3 Suites of Species.

The suites of species arguably could provide more information that unique species. Those species which occur throughout the area can also provide information when in conjunction with others. Suites of species can be indicative of certain habitats and can include those species which may be commonly found throughout Norfolk but when found alongside

another species the origin of the sample can be narrowed down. Using the species in suites also eliminates some of the uncertainties which occur when relying upon plant species which are unique to an area. Although suites of species can be useful, care must be taken as some of the different botanical regions are similar such as Goodsands and Brecklands and thus there will be a degree of uncertainty about pinpointing the location of a sample based upon the plant species alone.

5:5.4 Limitations of the use of plant data

As can be seen from figure 72; there are some inaccuracies in the recording of the data, as can be seen some plant species have been recorded some distance into the sea around the north and east of the county. This raises the question whether other species although they have been recorded in the different areas of Norfolk have actually been growing there. This highlights the need for the information to be used with some caution as it is recorded by people working voluntarily and with no formal training.

Some plant species can look very similar to others, therefore it must be considered that what is recorded may not actually be what was growing at the area. Again this comes down to human error and the nature in which the records for the database have been collected by untrained volunteers.

Some plants such as the most common ones or least common ones can be over recorded. Equally really common plants could have been under-recorded as volunteers focused on looking for / recording the more unusual species. Some recorders could have a bias towards certain species and may look out for them more. There also may have been bias towards some of the areas where plants were recorded; certain areas such as SSSIs may have had their plants recorded much more frequently. Areas which are remote or difficult to access may have been recorded less often. The expertise of the volunteer recorders must also be taken into consideration, some records may not be accurate with one species being mistaken for another or also may not recorded in the correct location (see 5.4.1).

The age of the records also has to be taken into consideration, although the plant species may have been recorded at one point, it may no longer be present in the area, and so to some extent the data could be somewhat misleading. As can be seen in figures 74 and 75 showing the division across the east and west of Norfolk of when the data was collected there is a definite bias of the majority of the data in the east of the county being collected after the majority of the data recorded in the west of Norfolk. Some plants may only be present at certain times of year and thus due to when volunteers chose to collect data may be over recorded. Also, there is potential for there to be plants which have not been recorded as they were not in season when the locations were recorded.

Consideration needs to be taken into the likelihood of fragments of these plants being found in a soil sample, and whether these would be able to be identified. For example if soil is retrieved from a suspect shoe, the amount of soil may be very small, and there may be only a minute amount of plant material within this soil sample. In order to establish this control samples from a specific site may be taken along with samples retrieved from the soles of shoes with different tread which have been walked over the area for a certain time period to gain a "suspect" sample, and the control and suspect samples analysed and compared to see the different plant species present.

Norfolk Vegetation Summary

Norfolk only has a small number of plant species which are unique to their botanical regions.

The presence of a plant species can be just as important as the absence of a plant.

Suites of different species can be indicative of a botanical area.

Vegetation data is limited by when the samples were recorded, who the data was recorded by, and the accuracy of where the plants were recorded.

CHAPTER 6 – Pollen Analysis of Soil Samples

Pollen analysis is utilised in this study to primarily complement the other data sets; it is hoped that the pollen spectra for each site will help to increase the certainty with which it is possible to predict the origin of a soil sample. The other aim of the study is to determine on what scale the pollen data in Norfolk is useful, for example does it vary within a sample location.

6:1 Palynology introduction

Palynology began in the early 20th century when the Swedish geologist Lennart von Post represented the results of pollen analysis as stratigraphical diagrams which demonstrated similarities in the pollen distribution from small areas and differences between larger areas.

6:1.1 Pollen grains and spores

Palynology is the study of pollen grains and spores. It is concerned with both the structure and the formation of pollen grains and spores, their dispersal and their preservation under differing environmental conditions. Pollen grains are produced by seed plants, angiosperms and gymnosperms. Spores are produced by pteridophytes, bryophytes, algae and fungi. Pollen grains and spores differ greatly in their function; however both of these groups (with the exception of some algae and fungi) result from cell division involving a reduction by half of the number of chromosomes (meiosis). Both groups are also similar in that they need to be transported (to a plant) in order to adequately perform their functions.

The pollen grain hosts the male gametophyte generation of the angiosperm or gymnosperm. The spore is a resting and dispersal phase of the cryptogam. Both the pollen grain and the spore require dispersal in space, but, the pollen grain can only be seen as successful once it has arrived at the stigma (or micropyle) of a plant of the same species and germinates there with the subsequent fertilisation of an egg. In comparison spores are

only required to arrive at any site where they can germinate, but this site must be suitable for the resulting gametophyte plant to establish, survive and produce gamete-bearing structures (182, 183).

The widely accepted basic principles of pollen analysis are:

- Pollen and spores are produced in great abundance by plants.
- A very small fraction of the pollen fulfils its' natural function of fertilisation, the majority of the pollen falls to the floor.
- Unless the processes of biological decay are inhibited pollen will decay in places such as those which are poor in oxygen.
- Pollen found in the atmosphere is well mixed (pollen rain).
- Pollen is related numerically to vegetation.
- A sample of pollen rain is an index of the vegetation at that time and space.
- Pollen is identifiable to various taxonomic levels.

It is these basic principles of pollen analysis which lends the technique to aid the identification of soil samples in this study and the use of pollen analysis in forensic science (see section 6:2). The principles of pollen analysis mean that pollen grains can also be utilised in other scientific studies including Taxonomy, genetic and evolutionary studies, honey studies, allergy studies, tracing vegetation history in both Individual species and communities, correlating deposits and assigning tentative dates, climatic change studies and the study of human impact on vegetation.

6:1.2 Pollen structure

Pollen and spores are of similar size (often around 20 – 40 μm); and are both surrounded by tough resistant walls that are often sculptured in unique ways. Pollen and spores are thought of together under the discipline of palynology as aerodynamically, wind-transported pollen and spores and they behave in similar ways.

The structure of a pollen grain can be divided into three parts; the inner living cell, middle intine and the exine. The inner living cell germinates on stigma and forms the pollen tube and is not preserved. The middle intine envelopes the whole pollen grain and consists

mainly of cellulose and also pectins, callose, proteins, enzymes, antigens and polysaccharides. The exine is made of sporopollenin and is formed by oxidative polymerisation of carotenoids and carotenoid esters. The basic formula of the sporopollenin is $[C_{90}H_{42}O_{36}]^n$, it is a polymer with saturated and unsaturated hydrocarbons and phenols. Sporopollenin is also present in some algae and fungi and is related to lignin.

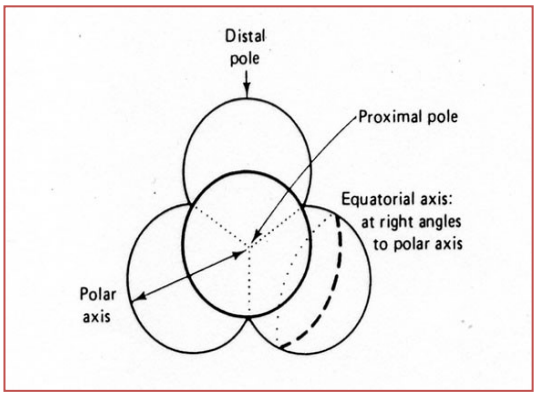
Pollen grains and spores are useful in such a wide range of studies due to their tough outer coat (the exine) being made of sporopollenin. Sporopollenin is one of the most resistant substances known to man and means the pollen grains and spores survive better and longer than many other biological materials. Another feature that makes pollen suited to forensic analysis is that the pollen may be preserved after other evidence types have disintegrated. The chemistry of the outer coat renders the pollen and spores resistant to decay, and where microbial activity is depressed whether it be due to wetness, salinity, low oxygen availability or drought, there is a chance of the pollen or spores surviving. This means that pollen and spores are of great value in the study of vegetation history and its applications in dating and climatology. Another feature which makes pollen and spores so useful is the variation in the form and sculpture of the resistant coat making them uniquely identifiable to the trained eye (183).

The small size of pollen grains and spores is necessary for their ease of transport; for pollen in the movement of genetic material and for spores in the invasion of new territories. Both pollen grains and spores are unreliable in finding their targets and are therefore produced in very large numbers in order to be effective. As pollen grains and spores are produced in much greater quantities than needed the excess material is what palynology depends upon. Pollen sedimenting from the atmosphere is washed out by the rain, and finds its way into soils, streams, lakes and mires where it may lie for long periods of time. Due to the large abundance of pollen grains and spores, statistical work can be carried out which can lead to precise environmental reconstruction (183). Again, this is a factor which makes pollen appealing to forensic science and why it was chosen for the analysis of the soils in this research.

6:1.3 Pollen morphology

Pollen grains of various species can vary quite a lot in size (often around 20 to 40 micrometer but can vary from about 10 to nearly 100 micrometer; exceptions are the thread-shaped pollen grains of some eelgrass) and in aspect: round, oval, disc or bean-shaped and sometimes filamentous. The natural colour is mostly white, cream, yellow or orange. The texture of the cell wall shows also great variations, from smooth to spiky

Pollen grains can be arranged in a tetrad which shows the symmetry of the grains as shown in figures 78 and 79.



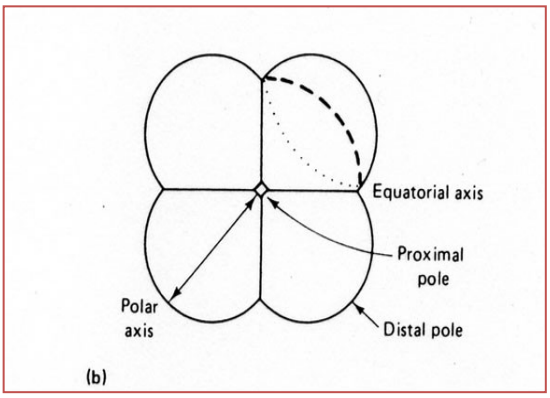


Figure 78: Dicotyledon arrangement

Figure 79: Monocotyledon arrangement

Commonly a three-apertured arrangement can be seen in pollen grains as shown in figure 80. These can be a Triporate structure, e.g. *Corylus avellana* which has three pores arranged on equatorial axis, a Tricolpate structure, e.g. *Quercus petraea* which has Three furrows (colpi) arranged longitudinally on equatorial axis or a Tricolporate structure, e.g. *Fagus sylvatica* which has three furrows each with a pore on the equatorial axis.

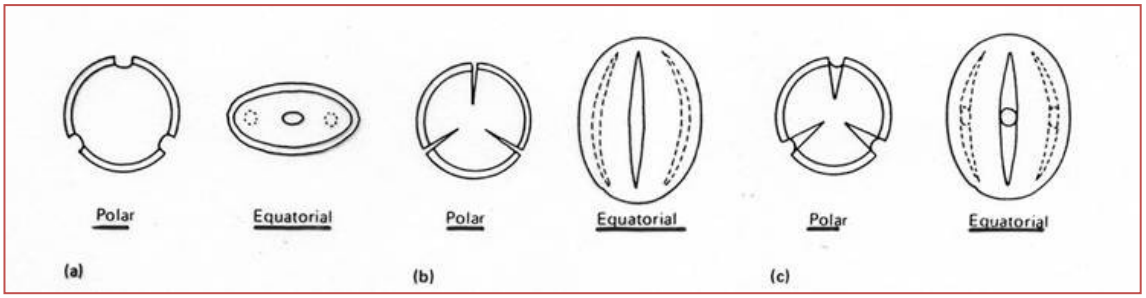


Figure 80: Common three aperture arrangement of pollen grains showing (a) triporate, (b) tricolpate and (c) tricolporate.

6:1.4 Pollen Identification

In the first instance of a palynological analysis the preservation of the pollen will be noted. The pollen grains can be described as either being well preserved, corroded pollen (distinctive etching on the exine, generally on the exine only), degraded pollen (the exine has undergone structural rearrangement with elements partially fused and the whole exine is affected), broken pollen (where the exine is ruptured) and crumpled pollen (where the pollen grains are badly folded, wrinkled or collapsed).

There are two main categories when identifying a pollen grain; identifiable grains and unidentifiable grains. Unidentifiable pollen grains can then be sub categorized into unknown grains and indeterminable grains which can be either concealed and cannot be turned to aid identification or deteriorated grains (as a result of corrosion, degradation, breakage or crumpling).

The success of the comparison of a pollen grain with modern reference material depends upon:

- Choice of reference material
- Variability of reference material
- Size and extent of reference material
- Discriminating powers of available morphological characters

- Preservation and quantity of pollen grains available for study
- Bias involved in selection and scoring of characters
- Experience and knowledge of the analyst.

Identification of pollen grains can be described as a mental “discrimination analysis” or assignment of unknowns, and can be done numerically.

Problems with identifying a pollen grain can also arise around the measurement of the grains. There can be a considerable variation in pollen size within the same species which can be due to clonal variation, physiological conditions and soil nutrients including the availability of water and racial differentiation.

6:1.5 Conventions for indicating the degree of confidence when identifying a pollen grain

In palynological studies, the taxonomic identification of an individual pollen grain or of a group of grains may be in doubt in the mind of the investigator. The confidence an investigator feels in his/her identification will depend on a variety of factors. For example, (1) the investigator may lack sufficient competence or experience with the taxonomic group concerned, (2) they may lack the time or the equipment required to study the pollen in sufficient detail, (3) the pollen observed may lack some essential diagnostic characters because of poor preservation, (4) the pollen may belong to a taxonomic group that is inadequately understood or described within the region of his/her study, and (5) the extent and the observed variability of the reference material with which the pollen are being compared may limit the taxonomic precision to which the pollen can be identified.

The degree to which the investigator attempts to overcome these problems will depend on her/his judgement of the degree of taxonomic precision required in the study. As she/he cannot know to what uses his/her data may be put by others, it is essential that the degree of confidence in the identification be clearly stated and defined in his/her notes and in subsequent publications. If this is done, equal weight in interpretation will not be given to both certain (i.e. beyond reasonable doubt) and doubtful identification.

6:2 Palynology in Forensic Science

Along with the basic principles of palynology, the relative abundance of pollen in the environment, its recognisable character and persistence has lead to its use within forensic science. Pollen grains can be found within soil, leaf litter and even dust and may provide clues as to the geographical region from which a sample originates. Soil taken from clothing, scrapings from fingernails or dust retrieved from clothes may yield enough pollen for analysis and the reconstruction of recent movements. Palynology has been successfully used in a wide range of criminal cases; murder investigations (184), allegations of chemical warfare (185) the origin of illicit drugs (186), and more recently in fire investigation (187) and document analysis (188). As previously mentioned pollen analysis is suited to forensic analysis due to its resistant nature of pollen and the fact it can be treated statistically.

Once the sample has reached the laboratory, the processes which will be undertaken include (182):

- Storage of the sample
- Subsample extraction
- Preparation of the sample for counting
- Staining and mounting of the sample
- Microscopy

Forensic palynology is based upon the presumption that the pollen assemblage for a sample will be distinctive for a particular location. This is because every locality is though to have a distinctive palynological profile that is due to the huge variability in vegetation and taphonomic factors which both affect the distribution of pollen and spores (189, 190). Using this principle the comparison of a control sample from the crime scene with the forensic (suspect) samples ought to be able to establish the probability that the samples did or did not originate from the same source.

There have been many high profile cases, involving forensic palynology and environmental analyses that have received publicity, meaning that this branch of forensic science is

assured. Several studies discuss cases whereby palynology has ensured a conviction (17, 191).

Despite the use of palynology in forensic casework, few studies have been undertaken to test some of the basic assumptions upon which the use of forensic palynology are based. In a study by Horrocks *et al*. (192) the variability of multiple samples collected from within a small area such as a crime scene was investigated and also the samples were compare palynologically to other samples collected from small areas with a similar vegetation type. The study found that the soil surface samples from the principle control site were dominated by grass pollen and bracken spores and overall the multiple samples had a similar pollen and spore spectrum which was demonstrated statistically. However, it was found that when these samples were compared with those from a different area which had a similar vegetation cover the pollen and spore content was somewhat different. In a further study by Horrocks *et al*. (193) soil samples were collected from consecutive footprints made with clean shoes within a localised area and the pollen analysed. Pollen samples from and between the prints and from the shoes were analysed and it was found that there was a homogenous pollen assemblage but that a perfect match did not occur due to minor differences within the spectra.

The use of palynology for forensic purposes is potentially limited by the mixing or diluting of the pollen assemblage from a crime scene with that from sites visited both before and after the crime occurred. Taking this into consideration, the pollen assemblage taken from footwear worn at a crime scene will never perfectly match the pollen assemblage of the crime scene; this something which isn't a problem when an exclusionary approach (i.e. can a sample be excluded from being the same as another based upon its pollen profile) is being used as is the case in this research. Unless a suspect is apprehended immediately and footwear and clothing seized the adherent pollen assemblage will be mixed with new pollen assemblages from subsequent areas visited. However, it has been found that pollen is retained very effectively by footwear over considerable periods of time even after items have been cleaned. The mixing of pollen depends upon several factors, including the characteristics of the footwear and clothing, the ground surfaces at each of the locations visited such as the soil texture and the water content, and the nature of the contact that

occurs between the surface and clothing. The ideal situation for optimum pollen collection is one whereby a new pair of footwear is worn which has a deep tread and the site is damp with clay-rich soil. Although to date it has not been scientifically tested it is assumed that the pollen assemblage on a shoe will predominantly reflect the site that was last visited (189, 190).

6:2.1 Contamination Issues

It is of upmost importance to avoid contamination of the sample, particularly when the work is for forensic purposes. In order to do this, if possible, the laboratory should be separate from the room in which the microscopy work is carried out. The cleanliness of the laboratory should be the same as that of a microbiology or specialist forensic laboratory. The air should be filtered pollen free and guarded against secondary contamination. If the air is not filtered the contamination can be serious; in spring and summer the air contains substantial amounts of pollen which may be caught in the preparations and affect the results. It is especially important to avoid contamination when working with samples that contain low levels of pollen or if the analyst is looking for the presence of uncommon pollen types. However, the risk of contamination is greatly reduced if the sample in question is treated with great care from the moment it is received into the laboratory to the time analysis is completed (182, 194).

Another factor which needs to be accounted for when preparing a soil sample for pollen analysis is problems with pollen size. The size of pollen grains plays an important part with aiding with their identification. However, it has been found that many of the procedures in preparing a sample for identification can cause alterations in the size of the pollen grains and some changes in their form. Many of these problems are unavoidable but can be circumvented by some extent by ensuring that all of the samples are treated in exactly the same way (183).

6:3 Methodology – Pollen Analysis

The pollen was extracted from the chosen soil samples and subsequently mounted, identified and counted at the British Geological Survey in Keyworth, UK. The BGS commonly analyse pollen samples and therefore are established to do this in a way to reduce contamination between samples and also are trained to analyse and identify and count the pollen. Only a small subset of the soils samples could be sent for analysis. Minimum information was given to the BGS in order to enable the unbiased analysis of the samples.

6:3.1 Samples for pollen analysis

The following samples (see table 24) were sent for pollen analysis; the rationale for chooseing the samples was based upon their different locations and also having replicates from within the same location in order to determine if pollen analysis was usable for forensic work on both an inter and intra site level. The chosen samples are from locations right across Norfolk (for thie location on a map see the red triangles on figure 7) meaning that they are from different geological and botanical regions. This meant that the soil samples should have a different palynological profile. For more detail about the location and description of samples and the sampling method utilised refer to Chapter 1, Table 2. Multiple samples from each of the locations were used for pollen analysis in order to determine if the pollen assemblage changed within the small (approximate 5 m by 5 m) sample area in order to help determine to what level could pollen analysis be useful in a forensic investigation and meant inter and intra variability could be determined.

Table 24: Samples codes, replicate samples and sample locations of the samples analysed for the presence of pollen.

Sample Code	Triplicates	Location
160708_5	5a, 5b, 5c	Knapton
160708_12	12a, 12b, 12c	Holt
160708_4	4a, 4b, 4c	Wighton
160708_9	9a, 9b, 9c	Titchwell
240708_1	1a, 1b, 1c	Ashwicken
240708_2	2a, 2b, 2c	Bawsey Country Park
240708_3	3a, 3b, 3c	Bawsey Country Park
240708_4	4a, 4b, 4c	Bawsey Country Park
240708_5	5a, 5b, 5c	GBase site 440
240708_6	6a, 6b, 6c	Roydon National Park
240708_7	7a, 7b, 7c	Roydon

6:3.2 Extraction of pollen from soil samples

The soil samples were sent to the British Geological Survey at Keyworth; pollen was extracted from the samples in their purpose fit laboratories by the following protocol (permission for the protocol to used in this thesis obtained from the BGS).

Obtaining pollen in a mountable form from a sample such as soil does not require extraction as such, more concentration. The aim of this concentration process (as outlined below in figure 81) is to produce pollen rich samples which once mounted onto microscope slides will allow the accurate identification of as many as possible of the visible grains and the counting of an adequate number of pollen grains and spores to provide a representation of the total population in the original sample. A variety of chemical processes have been developed for the treatment of pollen samples in relation to the different matrices in which the pollen may be embedded. A combination of these processes can then be used determined by the precise nature of the sample under investigation (183).

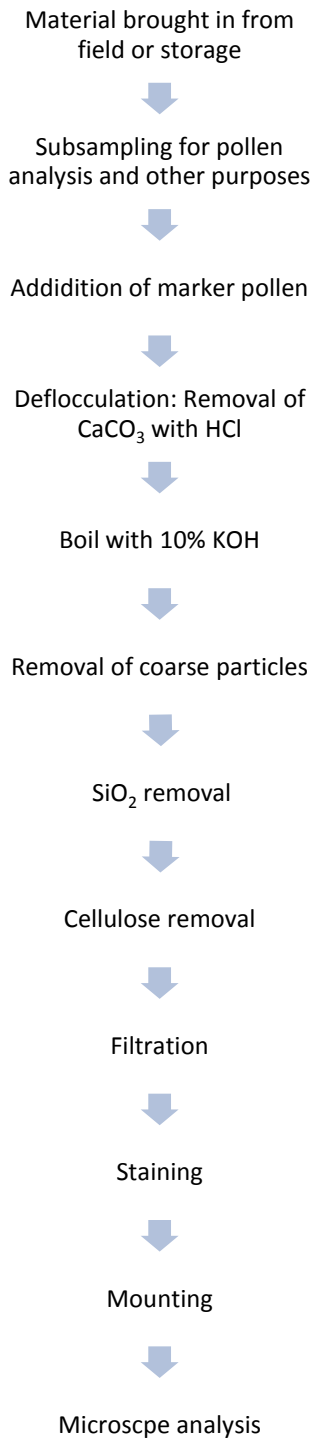


Figure 81: Flowchart of preparation of material for microscopic pollen analysis (182).

As can be seen in figure 81, once the marker pollen has been added to the sample the next step was the removal of any calcium carbonates with hydrochloric acid (HCl). This treatment should always be used first and then followed by any other necessary

treatments such as hydrofluoric acid (HF) or potassium hydroxide (KOH) digestion. The HCl treatment is simple; cold 10 % HCl was added in excess to the sample. The HCl can be added warm but this carries a danger that the pollen wall might be corroded. The froth which was formed during the reaction was subdued with a drop of 96 % alcohol (183).

The next step in the preparation of the pollen sample was KOH digestion to remove humic acids. It should be noted that from certain peats this treatment alone can produce a reasonable concentration of pollen, however, the quality of the extracted pollen and spores is of a very poor quality and therefore this procedure should be accompanied by acetolysis to remove any cellulose wherever possible. In this step approximately 1 cm of the sample was placed into a boiling tube and 10 mL 10 % KOH added (the precise quantities of sample and reagent are not critical unless pollen densities are to be determined). The boiling tube was then placed in a boiling water bath for 10 – 15 minutes, and stirred occasionally with a glass stirring rod to break up the material. In order to ensure that the concentration of the KOH did not rise above 10 % distilled water was added occasionally. During this step if possible prolonged boiling was avoided as it could cause the pollen grains to swell. This digestion process broke up the matrix and dissolved the humic material and after 15 minutes in the boiling water bath the sample was be a dark brown solution (182, 183).

The sample was then passed through a fine sieve (100 to 120 μm) to allow all the pollen and spores to pass through but retain the large particulate matter. The suspension which passed through the sieve was collected in a polypropylene centrifuge tube. The sample(s) was then centrifuged at 3000 rpm for three minutes. (If possible a swing-out head centrifuge should be used rather than a solid-head one as the latter produces a pellet in the form of a smear up the side of the tube which can very easily be lost in decanting.) Following centrifugation the liquid was decanted and discarded. The pellet was resuspended in distilled water. If any clumping of the fine detritus fraction occurred a few drops of detergent such as 5 % sodium lauryl sulphate was added. The sample was then centrifuged and resuspended in distilled water (182, 183).

In order to remove silicates in the sample hydrofluoric acid (HF) treatment was used; this is important as an abundance of silica in the sample can obscure the pollen once the sample has been mounted. Before this treatment is carried out it was of upmost importance to remove any carbonates and any remaining KOH. Also, as HF is being used suitable safety procedures must be adhered to at all times. Approximately 6 mL of concentrated (30 – 40 %) HF was added to the washed pellet (in a polypropylene centrifuge tube). The pellet was then resuspended carefully using a polypropylene stirrer. The sample was then placed in a boiling water bath or an electrically heated thermostatic block at 100 °C for 15 minutes. During this time any grittiness in the tube was checked for with a stirring rod. If necessary the sample could be boiled for an hour or more to ensure that all the silica had been removed; this caused no damage to the pollen. Once all of the silica has been removed the sample was centrifuged whilst still hot, this was done inside of the fume cupboard and with caps on the centrifuge tubes. The supernatant was decanted and once cooled was neutralised with a suspension of sodium carbonate. Meanwhile the pellet was resuspended in 10 % HCl and warmed in order to remove any silicofluorides produced during the HF treatment. This was centrifuged in the fume cupboard and the supernatant discarded as before. The pellet was washed twice with distilled water to remove any traces of HF (182, 183).

The most effective way of removing any cellulose in the samples was by acetolysis. As the reagents used in acetolysis react vigorously with water the sample needed to be dehydrated before this step. In order to dehydrate the sample, the washed pellet was resuspended in glacial acetic acid. This was then centrifuged and decanted and the supernatant discarded. Approximately 6 mL of the acetolysis mixture (acetic anhydride mixed with a concentrated sulphuric acid in a 9:1 ratio) was added to the pellet and resuspended by stirring. The suspension was placed in a boiling water bath for three minutes; any longer may have been harmful to the pollen and spores. The sample was then centrifuged and the supernatant decanted into running water. The pellet was resuspended in acetic acid and once again centrifuged and decanted. The pellet was now ready for rehydration without the danger of exothermic reactions. Once the pellet was resuspended in distilled water it was centrifuged and the supernatant decanted, then this step repeated. If any of the residual acetic acid was not thoroughly washed out crystals would form during the mounting process if glycerine jelly is used (182, 183).

It should be noted that acetolysis usually leaves grains with a slight yellow colour which may prove adequate for some palynologists, however for some a stronger contrast may be required. The most widely used staining material is safranine but fuchsin is also used. During the final wash step of acetolysis it is helpful to add a few drops of 10 % KOH solution. This will neutralise any residual acid and ensure the stain will take in a more effective manner. Two mL of distilled water was added to the clean pellet plus two drops of 5 % aqueous safranine solution and care was taken to ensure that the pellet was fully resuspended. The sample was then topped up with distilled water, stirred, centrifuged and decanted. At this point the pellet was red in colour and comprised only of fine organic material. If there was a mineral layer at the base HF treatment needed to be repeated. The sample was now ready to be mounted. There are various mounting materials available including glycerine jelly and silicon oil. Of those available, glycerine jelly has one major advantage in that it is easy to use and has excellent optical properties. However, it also has one main disadvantage in that it absorbs water from the atmosphere which causes the pollen grains to swell when mounted, often increasing in size by 1.25 to 1.5 times. Therefore over long time periods the grains deteriorate and their wall structures become unclear. For this reason silicon oil was the mounting medium of choice. Once mounted the pollen sample was then ready to be viewed under a microscope and the pollen grains and spores be identified and counted (182, 183). Silicon oils are also known to have good optical properties and are often used in refractive index matching for example.

6:4 Results

All the pollen analyses were carried out by the British Geological Survey at Keyworth, and the final report containing the pollen and spore counts prepared by James B. Riding. Due to the samples being prepared and analysed externally no images of the slides are available.

6:4.1 Descriptions of pollen assemblages at each sample site

The palynofloras from the eight locations are described in this section. The samples generally yielded relatively sparse palynomorph assemblages of relatively low diversity; the preservation proved to be fair. It was found that tree pollen grains are consistently more abundant than herb pollen and fern/moss spores.

6:4.1a Ashwicken

The three samples from Ashwicken were collected from an arable farmland locality characterised by hard, sandy soil. At the time of sampling, a crop of swedes were growing. All three samples produced relatively abundant organic residues which are dominated by various types of light, poorly-sorted plant tissue. Wood fragments were rare, except in sample 2 where they proved relatively common. Fungal materials proved extremely sparse.

The palynomorph spectrum mirrors the palynofacies with the pollen/spore associations in samples 1 and 3 being remarkably similar. Arboreal (tree) pollen is dominant in samples 1 and 3; *Alnus, Betula, Corylus* and *Pinus* are all relatively prominent. The herbaceous pollen is largely represented by Gramineae and *Taraxacum officinale*. The two dominant palynomorph taxa are Gramineae and *Pinus*. This dominance of Gramineae and *Pinus* was also observed at Roydon (crops), Tichwell, and Wighton, although taxa such as *Alnus, Betula* and *Corylus* are rare at these localities.

6:4.1b Bawsey Country Park

Nine samples from Bawsey Country Park were collected from two localities in this area. These samples are all overwhelmingly dominated by *Pinus* pollen. In terms of this study, this abundance is unique and it probably directly reflects the presence of pine trees in the forest.

Samples 4 to 9 - forest

Samples 4 to 9 were taken from a forested area; the leaf litter was cleared prior to sampling. The organic residues were all dominated by light plant tissue with lesser proportions of fungal material (hyphae etc.), insect parts and wood fragments. The relative prominence of fungal material, insects and wood is typical of a litter-strewn forest floor. *Pinus* pollen is extremely prominent and other relatively prominent arboreal pollen include *Alnus, Betula, Corylus* and *Tilia*. *Erica* (heather) pollen is also present in small numbers, reflecting some input of heathland elements. Herbaceous pollen and spores are sparse. This assemblage would be expected in a forest setting.

Samples 10 to 12 – near lake

Samples 10 to 12 were collected from close to a lake near a quarry. The residues are dominated by light plant tissue with subordinate fungal material and wood fragments. Insect parts proved rare. These three samples are also rich in *Pinus* and other arboreal taxa such as *Alnus, Betula* and *Corylus*. *Corylus* and the shrub pollen *Erica* are very prominent, thereby indicating significant stands of hazel and heather in this area. These three samples proved relatively rich in spores. *Sphagnum* is especially common with lesser levels of *Dryopteris*, *Lycopodium* and indeterminate forms. This relative prominence of fern and moss spores is typical of low-lying, damp areas for example those close to lakes. A possible dinoflagellate cyst (possibly derived from the lake) and a questionable specimen of the Carboniferous spore (*Densosporites*) were observed from samples 10 and 12 respectively. The latter specimen was reworked from underlying bedrock (Quaternary till).

6:4.1c BGS G-BASE Site 440

The three samples from this locality, a field growing crops, produced residues dominated by varied plant tissues and significantly lesser proportions of wood fragments. Pollen and spores proved relatively sparse. *Pinus* pollen was common in samples 13 and 15. A small, indeterminate reticulate pollen grain proved relatively common. This grain may be derived from an agricultural crop.

6:4.1d Roydon National Park

The three samples from Roydon National Park were collected from a wooded area close to a grassed field. The organic residues are dominated by light, poorly-sorted plant tissue, with lesser proportions of fungal material and wood fragments. Common fungal material is typical of a forest floor. Pollen and spores proved relatively sparse, especially in sample 17. *Pinus* pollen was common in samples 16 and 18. Other arboreal pollen include *Betula, Corylus* and *Tilia*, and the shrub *Erica* was also relatively prominent. Herbaceous pollen and fern/moss spores are rare. This pollen assemblage is typical of a forest setting.

6:4.1e Roydon

The three samples collected from a field of crops at Roydon produced relatively abundant organic residues which are dominated by light plant tissues. Wood fragments and fungal materials were relatively rare except in sample 20, where fungal remains proved common. The palynomorph assemblages are sparse. Arboreal (tree) pollen, with prominent *Pinus*, dominates all three samples. However the tree pollen associations are extremely low in diversity. Herbaceous pollen is present in low proportions and includes Caryophyllaceae, Gramineae and *Taraxacum officinale*. The two dominant palynomorph taxa are Gramineae and *Pinus*. This dominance of Gramineae and *Pinus* was also observed at Ashwicken Ground, Tichwell, and Wighton. A single, questionable specimen of the Mesozoic dinoflagellate cyst *Cribroperidinium* was observed; this was reworked from underlying Quaternary till.

6:4.1f Titchwell

Three samples were taken from Titchwell which proved to have somewhat variable residues. Sample 22 yielded large levels of amorphous organic material and sample 24 is dominated by plant tissues. Fungal material and woody fragments are relatively rare. The pollen/spore assemblages are sparse and diversity is low. *Pinus* and Gramineae are dominant. This prominence of Gramineae and *Pinus* was also observed at Ashwicken Ground, Roydon (crops), and Wighton.

6:4.1g Wighton

Three samples were collected from a field of cereals at Wighton. All the samples produced relatively abundant residues dominated by light, poorly-preserved plant tissue with subordinate wood fragments and palynomorphs. The samples have a pollen/spore diversity is extremely low; *Pinus* and Gramineae are the dominant elements. This dominance of Gramineae and *Pinus* was also observed at Ashwicken Ground, Roydon (crops), and Tichwell. Single specimens of the Carboniferous spore *Densosporites* were recorded from samples 26 and 27; these were reworked from underlying Quaternary till.

6:4.1h Holt

Three samples were collected from an area near Holt in North Norfolk where ferns are a significant part of the vegetation. Each of the three samples produced abundant organic residues dominated by light, poorly-sorted plant tissue with subordinate wood fragments, palynomorphs, fungal materials and insect parts. Tree pollen grains are relatively common with prominent *Alnus, Betula, Corylus, Pinus* and *Tilia*. Herbaceous input was also noted with relatively prominent Caryophyllaceae and Gramineae. It was stated that ferns are prominent locally. However the only record of fern spores were two specimens of *Polypodium vulgare* in sample 30. This means that the stands of ferns are assumed to be dominantly composed of *Polypodium* ferns.

6:4.1i Knapton Hedges

Three samples were collected from Knapton Hedges. The vegetation from this locality was a mixture of pine trees, grass, harvested wheat, potatoes and an abundance of weeds. All three samples yielded organic residues dominated by light plant tissue with subordinate wood fragments, palynomorphs and fungal material. Pollen and spores were sparse; *Pinus* is the most prominent taxon as was to be expected from the sampling. Other tree pollen comprise *Corylus* and the shrub *Erica* was also recorded. Herb pollen was also recorded; this includes Caryophyllaceae, Gramineae and relatively prominent *Taraxacum officinale*. Spores are rare, however *Lycopodium* was recorded in low numbers throughout. A single specimen of the Carboniferous spore *Densosporites* was recorded from sample 31; this was reworked from underlying Quaternary till.

This locality yielded a characteristic pollen/spore spectrum. It is unique in this set of samples due to the fact that the dominance of *Pinus* is not as marked. Moreover, herb pollen, particularly *Taraxacum officinale* was relatively prominent and *Lycopodium* spores were recorded consistently and in significant numbers.

6:4.2 Pollen Diagrams comparing all sites and multiple samples taken at one site

The following pollen diagrams have been created by converting the counts for each of the different species into a percentage of the total number of pollen grains counted at the sites. The data is displayed as bar graphs allowing comparison of both the similarities and differences both between all of the different and between the multiple samples collected at each of the individual sites (see figures 82 to 90).

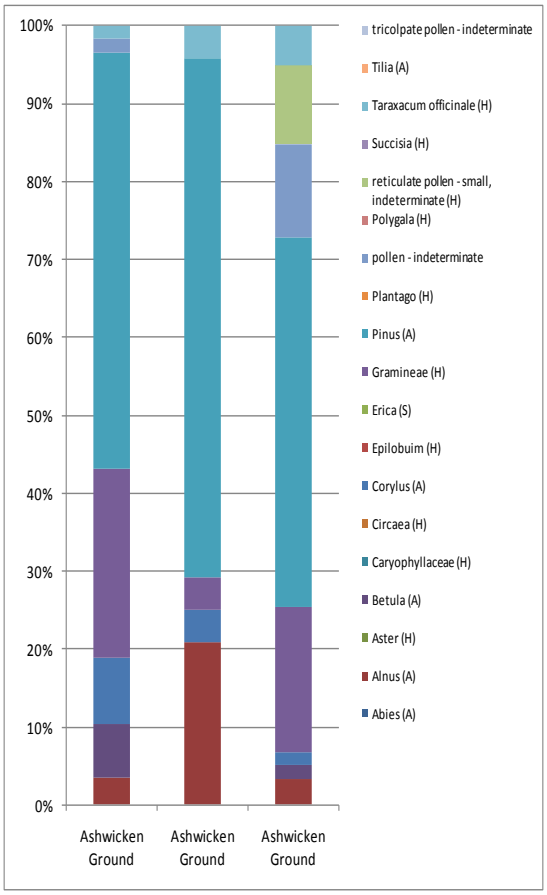


Figure 82: Pollen diagram of Ashwicken

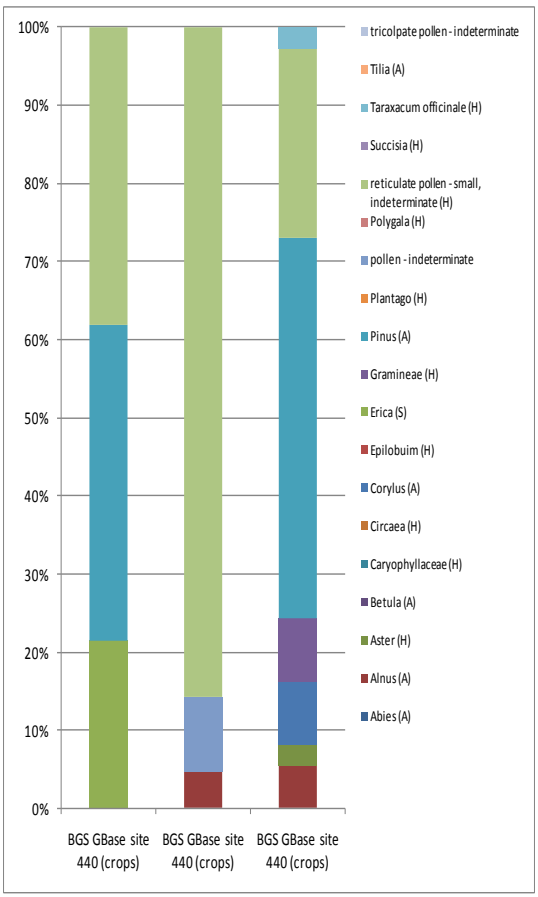


Figure 83: Pollen diagram of BGS Gbase site 440

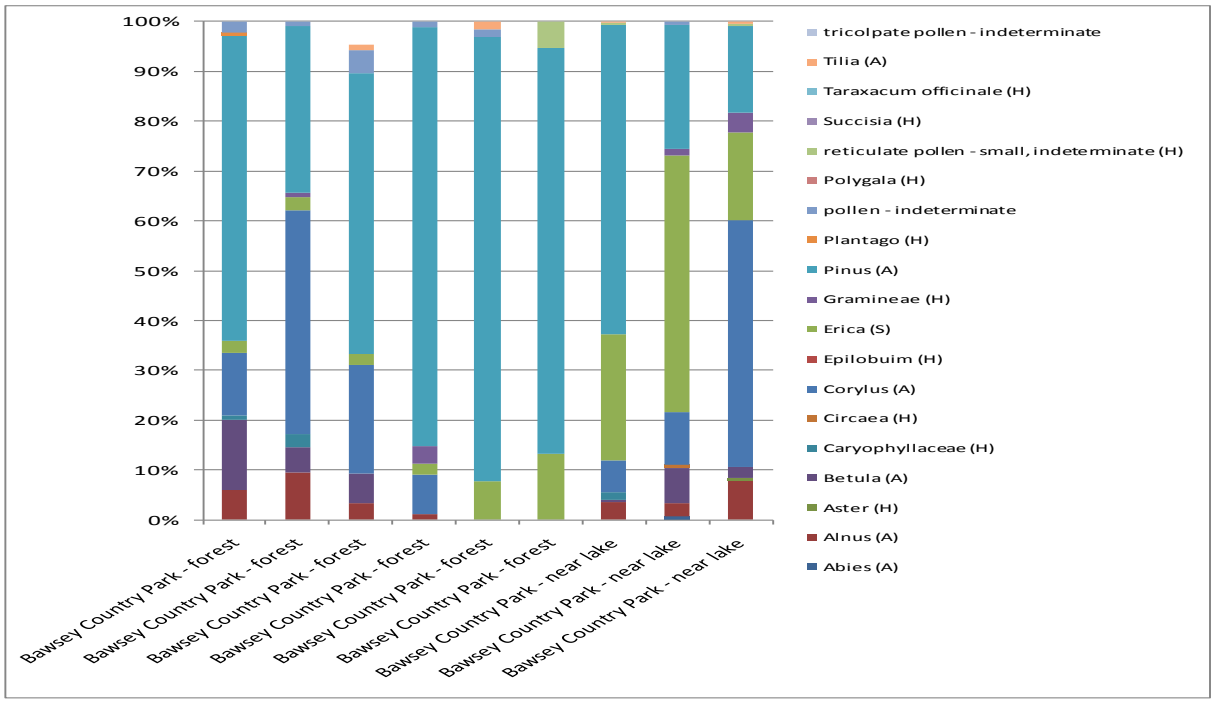
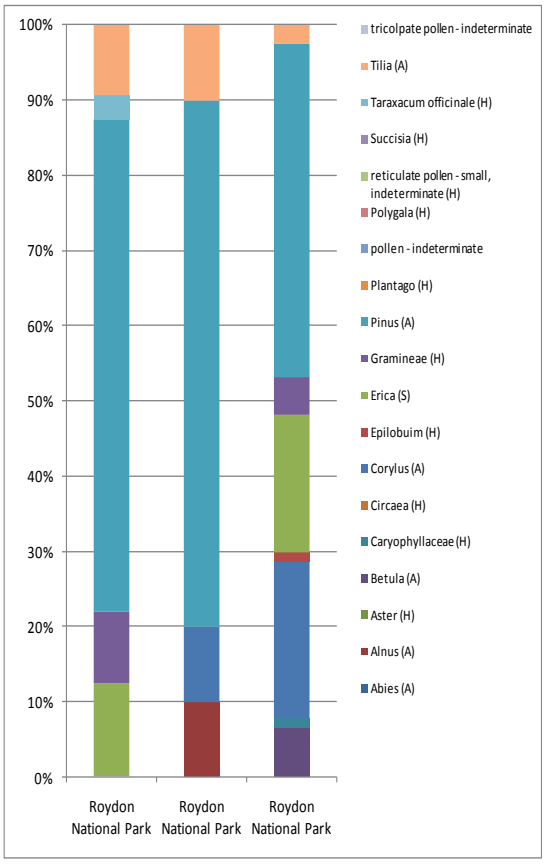


Figure 84: Pollen diagram of two sets of samples taken from two different areas of Bawsey Country Park



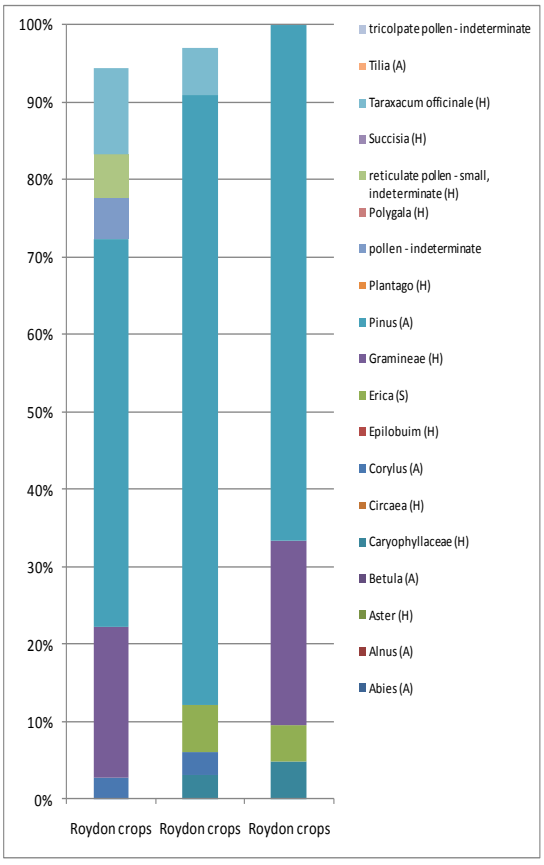


Figure 85: Pollen diagram of Roydon National Park **Figure86:** Pollen diagram of Roydon

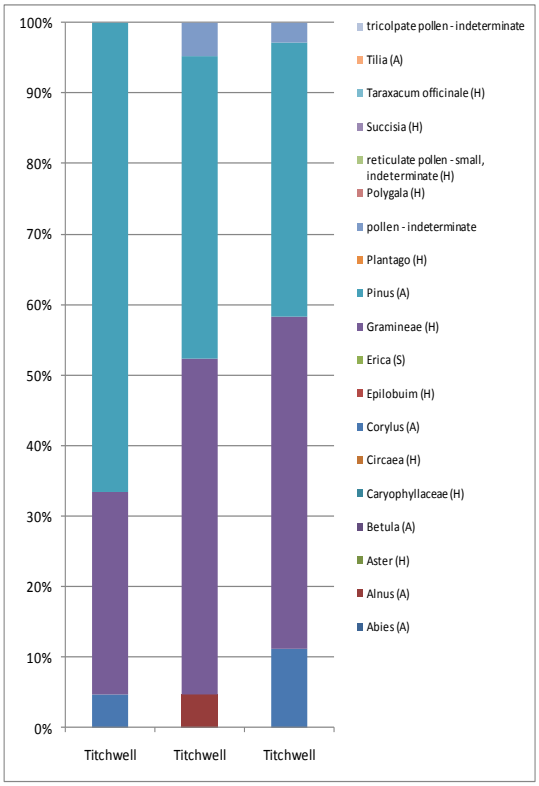


Figure87: Pollen diagram of Titchwell

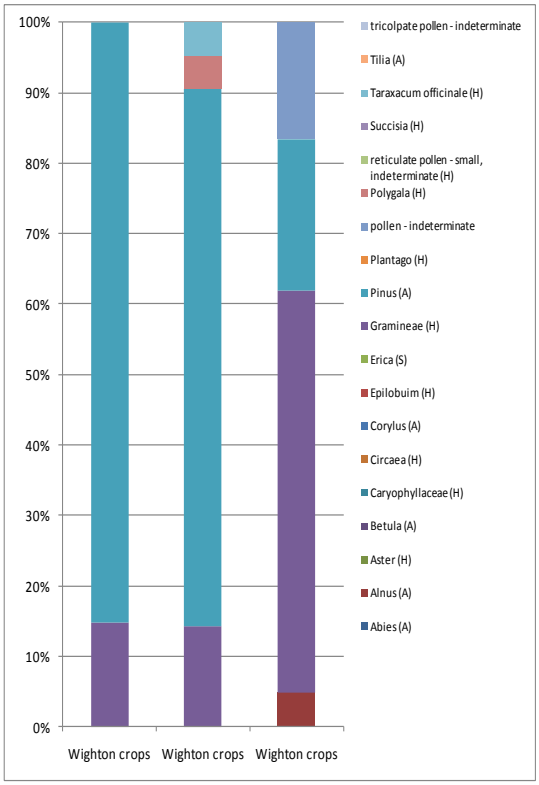


Figure 88: Pollen diagram of Wighton

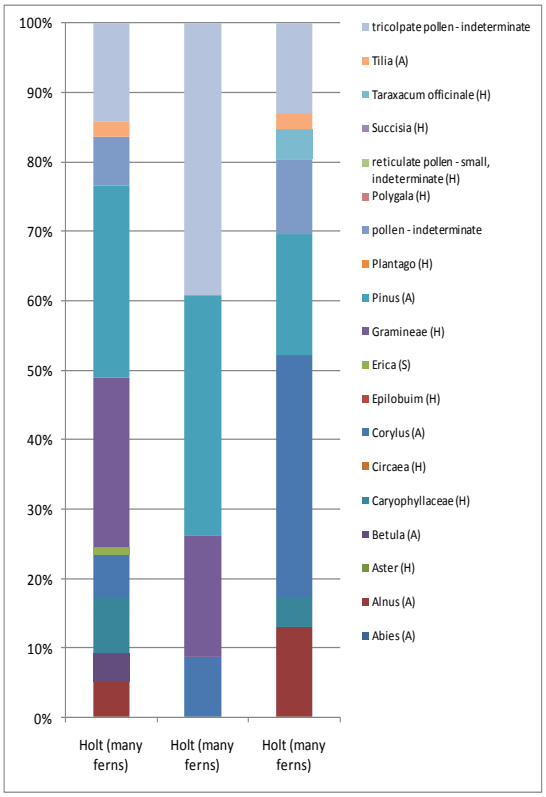


Figure 89: Pollen diagram of Holt

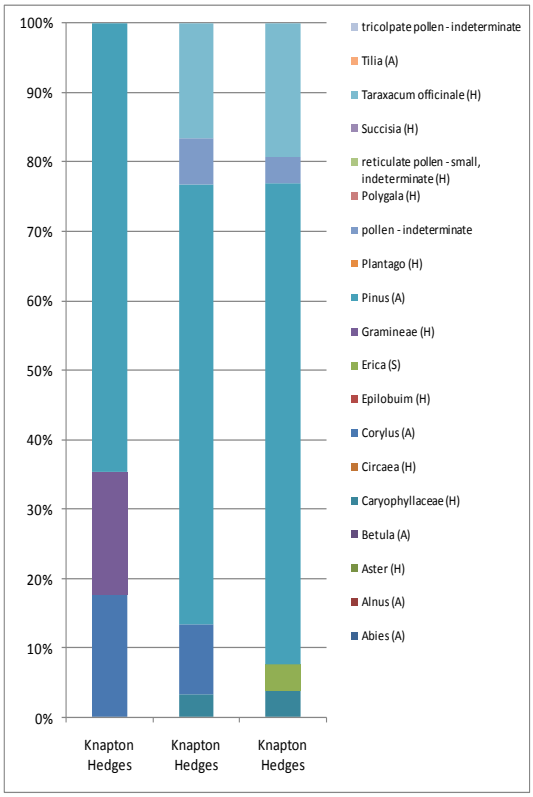


Figure 90: Pollen diagram of Knapton Hedges

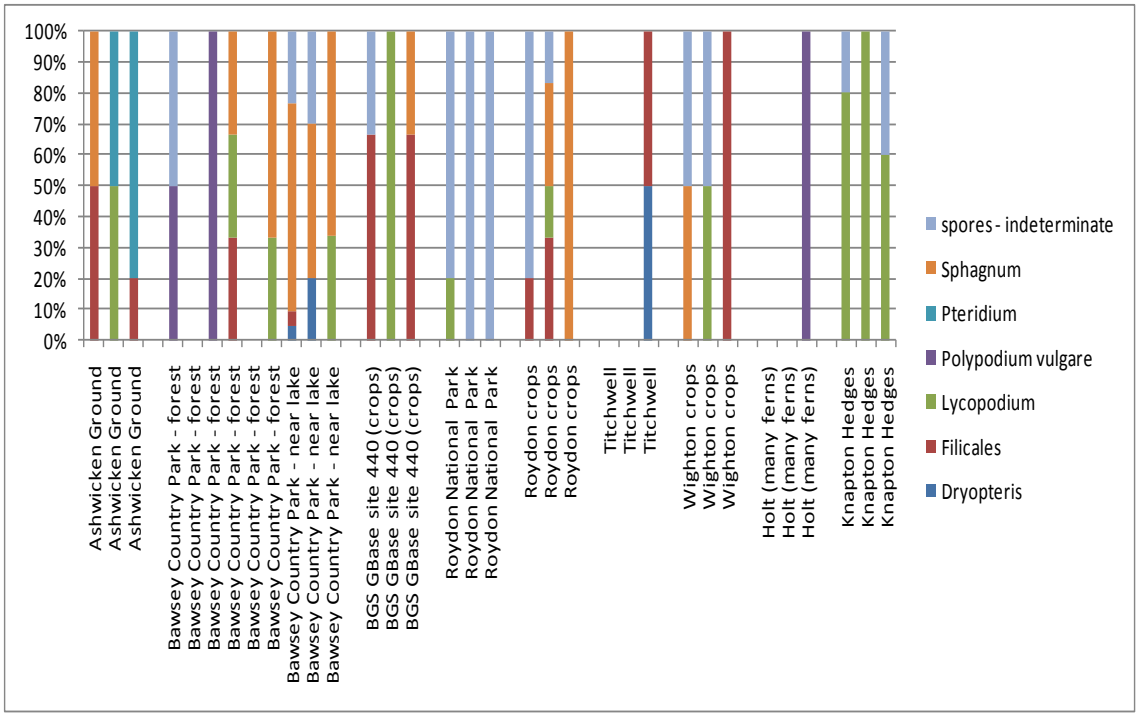


Figure 91: Spore diagram of all the different sample sites

As can be seen from the pollen diagrams, the results show some variation in the pollen and spore assemblages found in the triplicate samples from one location. However, it is also clear that there are differences between the different sample locations. As described in 6.4.1 Wighton, Askwicken and Titchwell all have a dominance of Gramineae and Pinus.

6:4.3 Spatial statistics using the pollen data

Following the protocol set out by Bhattacharya *et al.* (195), the pollen data was expressed as a percentage of total pollen in the sample, as was used in the previous pollen diagrams. The paper excludes all of the taxa that are less than approximately 1% in at least two samples; however it needs to be remembered that for forensic work the absence of a taxa can be equally as important as the presence of a taxa.

To determine whether there are significant differences between sampling site pollen spectra, the Bray-Curtis dissimilarity was calculated between each pair of samples. This allows for the exclusionary approach which is needed for forensic geosciences to be

employed for the pollen data. The Bray-Curtis index has values which range from zero (indicating an identical comparison) to one (indicating complete dissimilarity), and calculates the ecological dissimilarity between the taxa represented in two samples by summing the difference in representation of a taxon between samples divided by the total amount of that taxon present in both samples. This index tends to minimise the contribution of rare taxa to dissimilarity between samples. The main advantage of this technique is that it is better able to handle high levels of diversity; turnover of pollen types between samples, that other dissimilarity indices do not.

Following the calculation of the Bray-Curtis index the data can then be ordinate with non-metric multi-dimensional scaling (NMDS). NMDS moves sample points in ordination space to produce a final configuration where ranked inter-sample distance in ordination space closely approximates rank-order Bray-Curtis dissimilarity in species composition between samples.

This statistical method has also be applied to the vegetation data, strontium isoptpe ratios and trace element data (see chapters 4 and 5)

As can be seen in figure 92, the ordination plot of the Bray-Curtis distances of all of the different species counts before they were converted into percentages would appear to have given a considerable separation of all of the different sites (the triplicates are included in this analysis). However, this plot also shows that there are differences between the triplicates i.e. sites 1, 2 and 3 all being from Ashwicken, and therefore no clear groupings as to each of the locations; there is overlap between the multiple samples from different locations. In places the plots are hard to read due to the overlapping of different sites; this is unavoidable due to several of the samples not being able to be excluded from one another. It should be noted that the text on plots 94 to 96 is very small, however the text just depicts the different samples collected within one site.

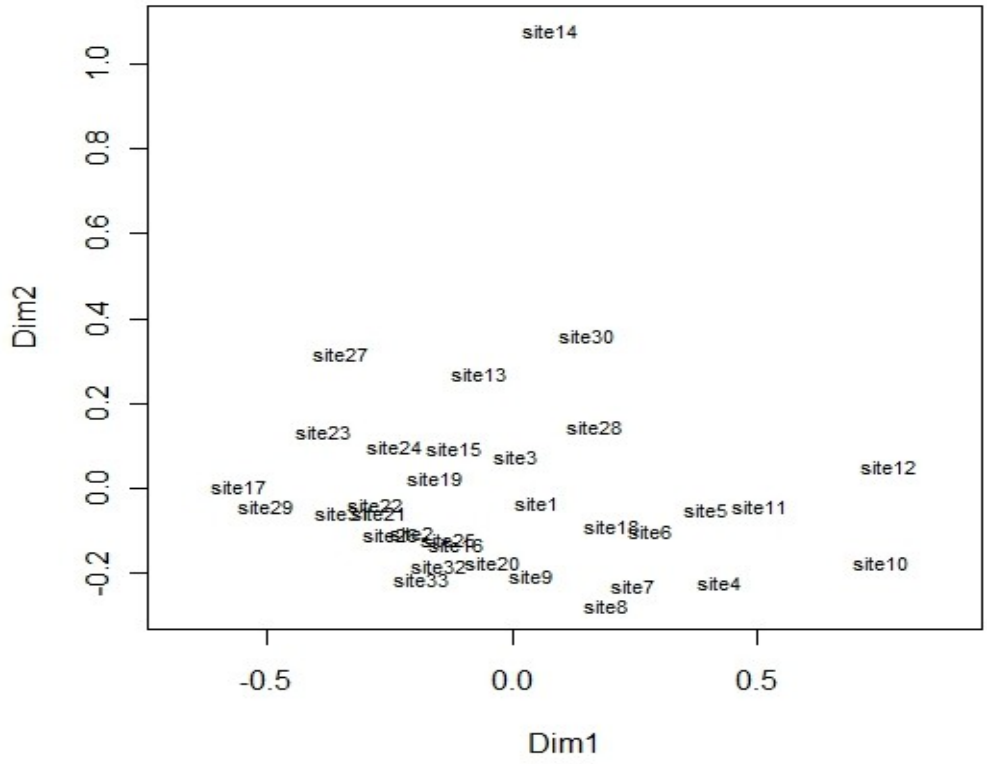


Figure 92: Ordination plot using the Bray-Curtis distances between all of the different pollen counts for all species for all different locations. (Sites 1 to 3 from Ashwicken, sites 4 to 12 Bawsey Country Park, sites 13 to 15 GBase site 440, sites 16 o 18 Roydon National Park, sites 19 to 21 Roydon, sites 22 to 24 Titchwell, sites 25 to 27 Wighton, sites 28 to 30 Holt, sites 31 to 33 Knapton Hedges)

In order to be able to discuss this data in respect to other work, the pollen counts were converted into percentages of the total count for each site. This was then plotted in the same manner as figure 92 in order to see any differences between the two figures. As can be seen figure 93, it appears to have a similar distribution to figure 92 (as would be expected), although the distributions appear to have shifted approximately ninety degrees to the right.

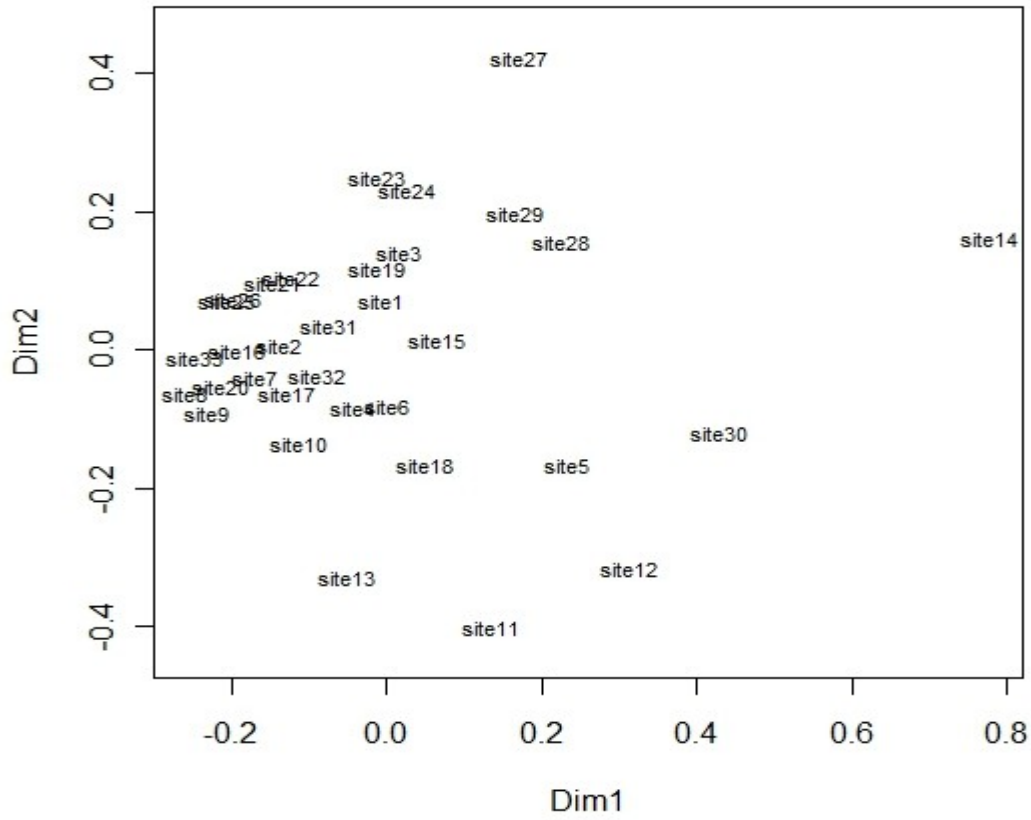


Figure 93: Ordination plot using the Bray-Curtis distances between all percentage pollen for all species for all different locations. (Sites 1 to 3 from Ashwicken, sites 4 to 12 Bawsey Country Park, sites 13 to 15 GBase site 440, sites 16 o 18 Roydon National Park, sites 19 to 21 Roydon, sites 22 to 24 Titchwell, sites 25 to 27 Wighton, sites 28 to 30 Holt, sites 31 to 33 Knapton Hedges)

As it appears not to have been studied previously, it was next decided to plot the ordination plots for each of the sample locations (showing the Bray-Curtis distances between the multiple samples) so they could be compared to the pollen diagrams. As can be seen in figures 93 - 102, no multiple samples from the same location show complete similarity, this is shown in both the pollen diagrams and the ordination plots. However, it should be noted that the scales on the ordination plots are low and show a small level of differences between the multiple samples, and that these differences are smaller than the ones shown in figures 91 and 92 (by comparison of the ordination plots scales) as would be

expected. Thus, the samples show a smaller intra-site variability than an inter-site variability.

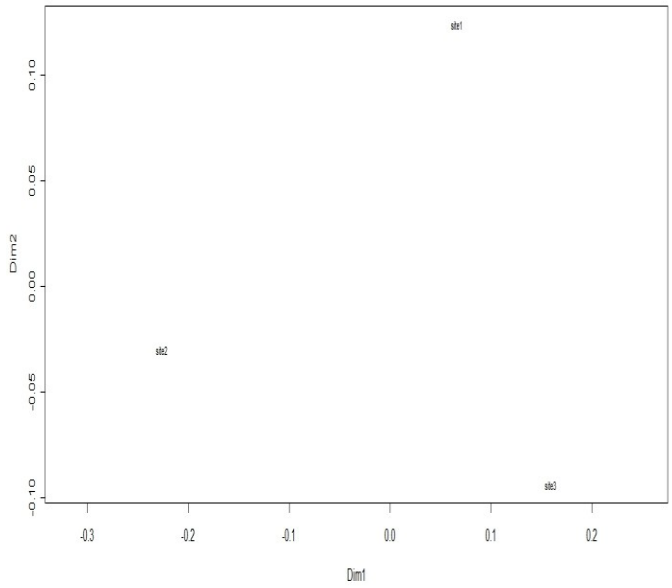


Figure 94: Ordination plot of the Bray-Curtis distances between the samples taken from Ashwicken.

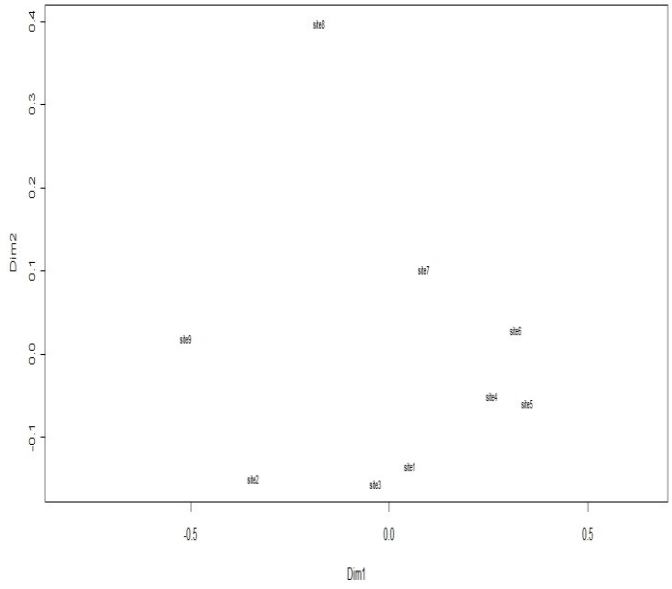


Figure 95: Ordination plot of the Bray-Curtis distances between the samples taken from Bawsey Country Park.

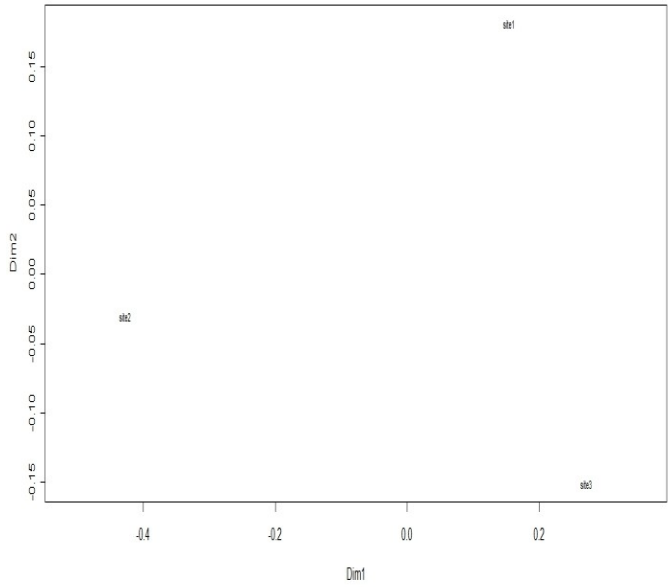


Figure 96: Ordination plot of the Bray-Curtis distances between the samples taken from GBase site 440.

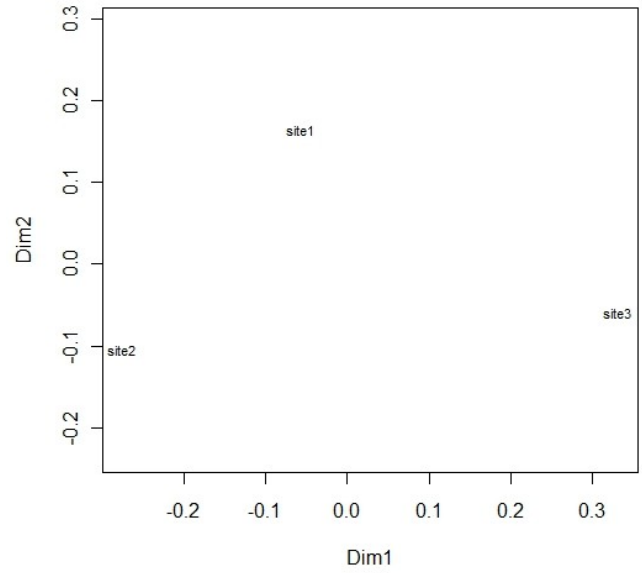


Figure 97: Ordination plot of the Bray-Curtis distances between the samples taken from Roydon National Park.

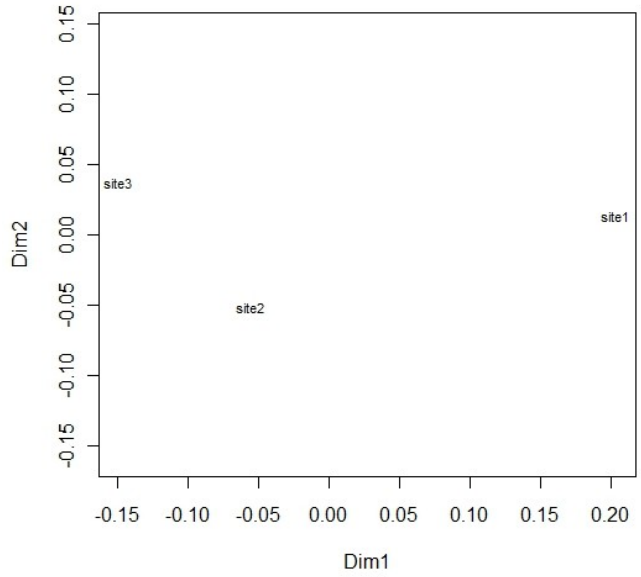


Figure 98: Ordination plot of the Bray-Curtis distances between the samples taken from Roydon.

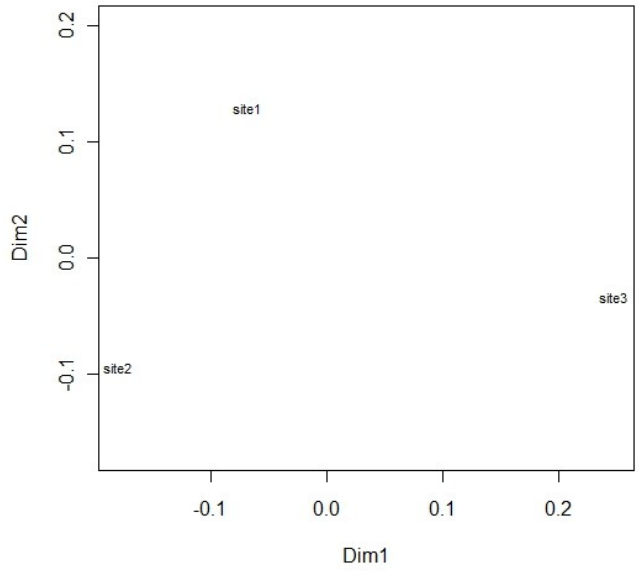


Figure 99: Ordination plot of the Bray-Curtis distances between the samples taken from Titchwell.

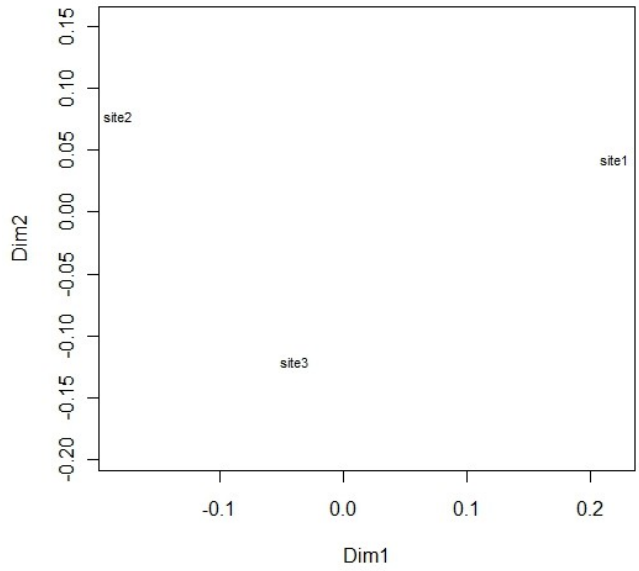


Figure 100: Ordination plot of the Bray-Curtis distances between the samples taken from Wighton.

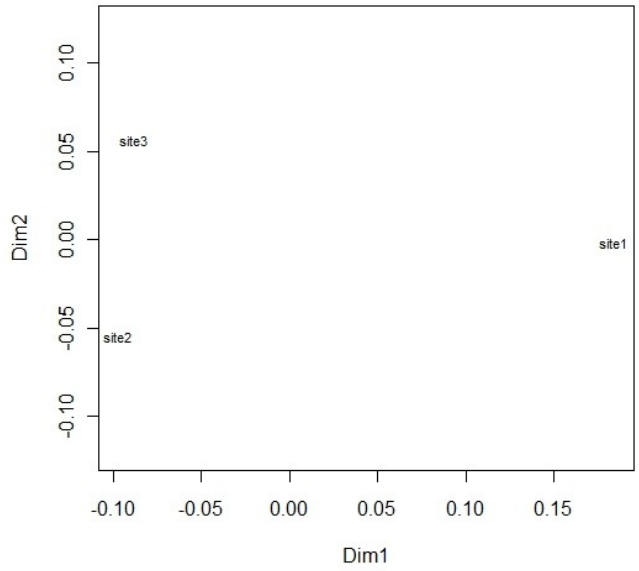


Figure 101: Ordination plot of the Bray-Curtis distances between the samples taken from Holt.

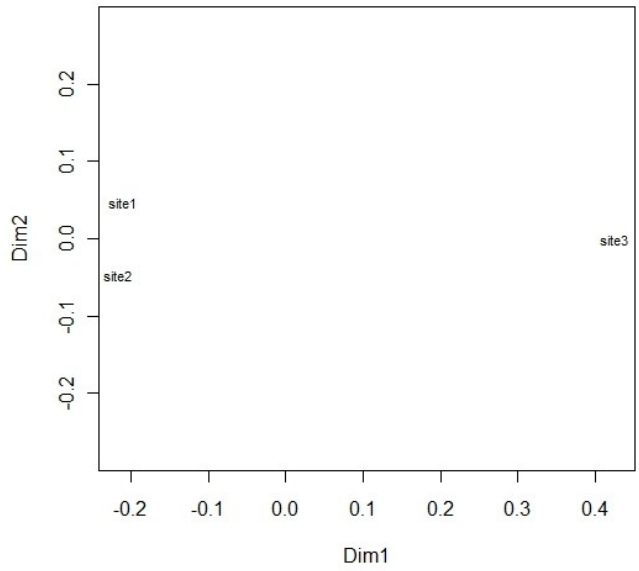


Figure 102: Ordination plot of the Bray-Curtis distances between the samples taken from Knapton Hedges.

When comparing the ordination plots it must be noted that they are not all of the same scale; as can be seen in the previous figures the Dim 1 and Dim 2 scales differ.

Finally, in order to determine whether the Bray-Curtis equation and the subsequent ordination plots could be used to differentiate between any of the different sample locations, an ordination plot of the average pollen percentages (for all of the different species) was produced (see figure 103). As can be seen on the plot, sites 3 and 8 (GBase site 440 and Holt respectively) are the only two sites which can conclusively be distinguished from the other seven sample locations.

When looking at the pollen diagrams, the pollen which may be responsible for GBase site and Holt being significantly different from each other is more reticular pollen at GBase than any other site and more tricolpate pollen at Holt than any other site.

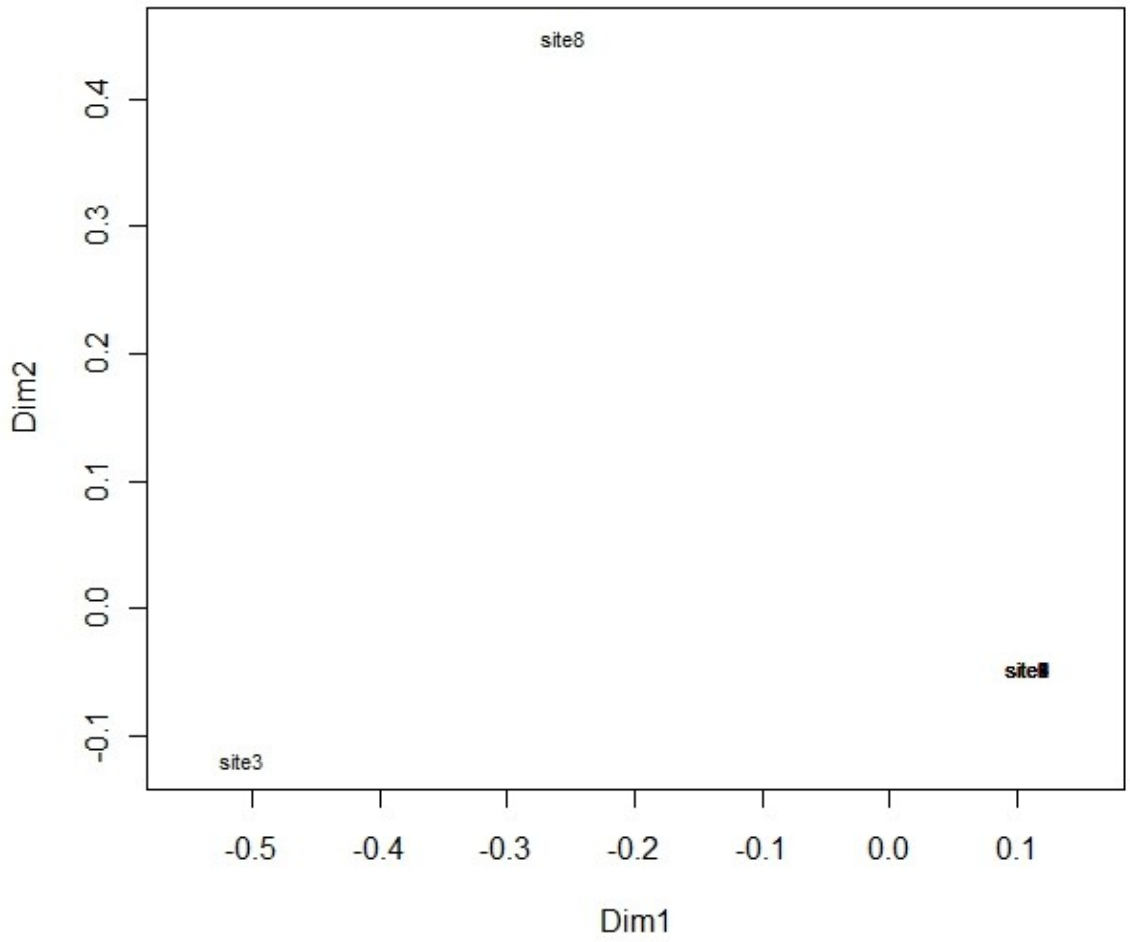


Figure 103: Ordination plot produced using the Bray-Curtis distances between the average pollen percentages of the multiple samples (for all pollen species) between the different sample locations (site 1 Ashwicken, site 2 Bawsey Country Park, site 3 GBase site 440, site 4 Roydon National Park, site 5 Roydon, site 6 Titchwell, site 7 Wighton, site 8 Holt, site 9 Knapton Hedges).

In figure 103, sites 3 and 8 are separated and all of the other sites overlap in the bottom right of the plot.

6:4.4 Conclusions drawn from pollen data

The pollen/spore assemblages and palynofacies of 33 samples collected from eight localities in Norfolk are unsurprisingly typical of temperate woodland/grassland habitats. They exhibit significant variability in terms of diversity, productivity and taxonomic spectra. This indicates the applicability of palynology for characterising both locality types within a biome in a general sense, and specific localities within a given locality type (189).

The samples taken from the arable agricultural settings at Ashwicken Ground, Roydon, Tichwell and Wighton produced similar palynomorph spectra. These are all dominated by arboreal pollen such as *Alnus, Betula, Corylus* and *Pinus*, with subordinate herbaceous pollen, largely comprising Gramineae and *Taraxacum officinale*. The two dominant palynomorph taxa are consistently Gramineae and *Pinus* which are both abundant, wind-pollinated types. BGS Gbase site 440 also represents an arable setting. Pollen and spores here proved sparse; however a small, indeterminate reticulate pollen grain proved common. Clearly, there are variations in arable localities, even within a specific area such as North Norfolk.

Six samples were taken from a forested area in Bawsey Country Park. Here, the relative prominence of fungal material, insect fragments and wood is typical of a litter-strewn forest floor. Unsurprisingly, arboreal pollen was prominent and *Erica* pollen was noted; herbaceous pollen and spores were sparse. Similar palynomorphs and palynofacies were recorded from a wooded area in Roydon National Park. Three samples were also collected from a locality close to a lake at Bawsey Country Park. Arboreal taxa again proved common, and *Corylus* and *Erica* were prominent, indicating stands of hazel and heather in the vicinity. Significantly, these samples proved rich in fern/clubmoss spores; this is typical of low-lying, damp areas.

A locality near Holt where ferns are prominent was sampled. Tree pollen grains were relatively common, and herbaceous pollen was subordinate. The only fern spores recorded were two specimens of *Polypodium vulgare* in one of the three samples. Hence the

abundant ferns were not reflected in the palynomorph associations, however they appear to be dominantly *Polypodium* ferns. The three samples from Knapton Hedges produced sparse pollen and spores. The normal dominance of *Pinus* was not as significant here. The herb pollen *Taraxacum officinale* was relatively prominent, and *Lycopodium* spores were consistently recorded in low numbers. Rare reworked Carboniferous and Mesozoic palynomorphs from the underlying Quaternary till were sporadically observed throughout.

In summary, the differences in both palynomorph content and palynofacies noted between settings such as arable fields, woodland and lakesides in North Norfolk soils indicates the applicability of palynology for differentiating the type of environment at a minimum of two levels. This is in a generic sense (i.e. the environment type), and potentially more accurate (i.e. to differentiate individual localities).

6:5 Discussion of pollen data

Pollen is a type of trace evidence and as such is transferred between different materials. Research has shown that in the first instance, pollen is shed from clothing rapidly (minutes) and then subsequently shed more slowly, with varying percentages of pollen remaining after several hours. The overall effect of the decay curve appears to be one of exponential decrease. However, this is dependent upon the size and texture of the material being transferred. This may affect the persistence of the pollen and any subsequent links this to the type of surface on which the material is being retained. In a forensic context care must also be given as to whether the any attempt has been made to remove the pollen.Whilst it is well known that pollen can provide valuable spatial and temporal information when interpreting pollen data, thought must also be given to the temporal changes during both the annual and decadal timescales and if this can account for any differences in a palynologcal profile when making inferences about the data. Thought must also be given to the conditions is which the pollen has been subjected to as these will have an impact upon the quality and quantity of the preserved pollen, however pollen is extremely resistant to decomposition due to its multi-layered cell wall (187, 188, 192, 193).

General strengths of forensic palynology is that pollen can be found on an assortment of unusual surfaces including condoms, counterfeit banknotes, paintings, grease on guns, in the lungs and stomach as well as on more common surfaces such as clothing and footprints. The amount of sample required for pollen analysis is very small yielding an ideal forensic process.

Weaknesses of forensic palynology are that in the past many police forces have not been aware of forensic pollen studies and thus are often sceptical and do not collect the appropriate samples from a crime scene. Proper evidential sample collection must be carefully taken from the scene to avoid contamination. This requires trained forensic palynologists. Control samples need to be selected to encompass the potential area represented by the samples and also to eliminate other areas that could not be represented by the samples. However, it needs to be considered that control samples may be skewed by an overabundance of one pollen type. A further weakness is that pollen evidence is purely circumstantial in that it may indicate where an event has taken place and who was present but not that the accused had actually committed the crime.

The pollen diagrams show variation within the triplicates from each sampling location; as was expected as previous research has shown that localised areas of similar vegetation type, even within the same geographic region, have significantly different pollen assemblages (194). As has previously been discussed in relation to other techniques no multiple samples from the same location show complete similarity, this is shown in both the pollen diagrams and the ordination plots. However, it should be noted that the scales on the ordination plots are low and show a small level of differences between the multiple samples, and that these differences are smaller than the ones shown in figures a and b (by comparison of the ordination plots scales) as would be expected. Thus, the samples show a smaller intra-site variability than an inter-site variability. Whilst the samples show a small intra-site varaiability in a forensic context it would be very rare to be looking at samples on such a small scale as 5m by 5m from which the intra-site samples were taken. It is more likely to be comparing pollen samples from different sites to determine if they can be distinguished from each other.

A factor which needs to be considered when discussing the variation in pollen assemblages is that of the nutrients found in the soil. For example, boron has an effect on pollen formation and germination, and molybdenum also has a significant effect of pollen formation. A shortage of either of these (or other) soil nutrients will have an impact on the pollen assemblage. However, this could also be exploited for forensic purposes. Differences in the pollen assemblages due to nutrient deficiencies are likely to differ in different soil types and therefore potentially can be used to distinguish between soils.

There is a high yield of pollen from trees in the analysed samples this is due to tree pollen grains being more abundant than herbs and ferns this may have had an impact upon why not all of the different sites can be excluded from one another. Although not all samples could be excluded from each other some inferences could be made. Samples which have been taken from arable agricultural settings at Ashwicken Ground, Roydon, Tichwell and Wighton all produced similar palynomorph spectra which were are all dominated by arboreal pollen such as *Alnus, Betula, Corylus* and *Pinus*. These could be differentiated from those samples from Bawsey Country Park which were dominated by fungal spores and aboreal pollen. In a forensic context this means that a sample from Baswey Country Park could be excluded from one from Ashwicken for example based upon the fungal content and the presence of aboreal pollen.

The samples were analysed at the BGS and therefore there is no knowledge on the counting of the pollen grains. This raises the point (similar to that raised in the vegetation chapter) that some counters may be more familiar with some pollen grains than others and subsequently there may be some degree of under and over recording and also some subjectivity. Also, the morphologies of some pollen grains are very similar leading to them easily being incorrectly identified. However, in the hands of a skilled forensic palynologist the subtle differences in pollen can become a powerful tool. Incorrect identification may also be linked to the preservation of the pollen and other factors as discussed in 6:1.4. Some soil types are most suited to the retaining of pollen grains. Clay soils are the best for optimum pollen collection.

Forensic palynology can play a role as a supporting tool in a multi-disciplinary approach with other techniques. Although the sample amount required for palynology is very small there is often no control over the size of the evidential samples which may vary from a few 100ths of a gram to larger amounts. Most of the sample will then be prioritised to be used for other better suited analytical techniques.

Forensic Palynology Summary Box

Not all of the sites could be distinguished from each other based upon their palynological profile.

Intra-site variability was shown to be less than inter-site variability.

Bray-Curtis statistical analysis is a suitable technique for seeing whether pollen assemblages match; this can be exploited for forensic purposes.

When coupled with other techniques forensic palynology is a versatile and convenient tool.

CHAPTER 7 – Forensic Interpretation of the Data, Conclusions and Future Work

The chemical and biological analyses carried out in this research have shown the potential importance of soil analysis within forensic science. Soils may constitute evidence that connects a person or an object to a particular location. Soil has a forensic value due to it ubiquity and ease of transfer to an object (196). Due to the complex nature of soil, the analysis of its inorganic and organic components can provide complementary and independent types of information about geological origin, dominant vegetation, management and environment as has been shown by the various techniques used.

The soil samples analysed in this research were collected at two different times; the intial set of soil samples focused mainly on the different geological regions of Norfolk. As the research progressed, it expanded into also looking in botanical and land use regions so more samples were collected to not only cover the rest of the county but other region than just geology too. Although collected at different times, care was given to collect the samples in the same seasonal period (Summer) to eliminate the influence of seasonality. All samples were chemically analysed and a subset for pollen and DNA analysis due to analyses being carried out at external businesses. A subset of samples have been analysed by each of the techniques used in this research allowing for the discrimination by each technique to be assessed.

7:1 Chemical and biological analysis of Norfolk Soils

As has been discussed in the previous chapters each of the individual chemical or biological techniques have strengths and weaknesses when it comes to their use within forensic science. For example the use of plant DNA found in the soil is limited by the unknown nature of the DNA and the present limits of the technology. This also has an impact upon using MALDI-ToF-MS for the analysis of the DNA. The vegetation data provided by the Norfolk Biodiversity can help with intelligence as to which areas samples may or may have come from based upon the presence or absence of specific species; but this is limited by the nature and accuracy of the records and how the data has been collated. The vegetation

data can be used alongside the pollen data to give intelligence as to what habitat or land-use a sample may have originated from. As yet the $^{87}Sr/^{86}Sr$ and trace element spatial data cannot systematically reproduce the spatial structure of the observations from first principles. However, strontium isotope ratios, trace element data and pollen data have all provided some means of discrimination of either, the different sample sites, geology, land-use or botanical region with different degrees of success. Each of techniques showed a smaller intra than inter site variation and some spatial differences between the different geologies and botanical areas. However, no technique was able to successfully differentiate between all of the soil samples on any spatial level, whether that is site specific, geology or botanical region. The use of Bray-Curtis statistical analysis of the different datasets has shown to have the potential to be a powerful tool in forensic science. The use of GIS maps can aid in showing the variations in the data spatially.

The strengths and weaknesses of different methods of soil analysis is something which has been discusses by Dawson and Hillier (196). The study states how individual analytical techniques each have different scales of resolution and relevance on the nature of the criminal case and context. Each method has its strengths and weaknesses, something which has been previously discussed about each of the individual techniques used in this research.

7:2 Forensic Interpretation of the data

The main focus of this research has been about the exlusion of different samples from each other as is important in forensic geosciences. The exclusion of some, but not all, samples can be shown by each of the differentbioloigcal and chemical techniques used in this research.

Strontium isotope data alone has potential to be a great aid in forensic investigations as it is clear that some sample locations have vastly different strontium isotope ratios and therefore are not from the same source. This is shown using box plots, Tukey plots and confirmed using Bray-Curtis statistics. It is also possible to differentiate between some of

the different character areas, be those; geology or vegetation, using $^{87}Sr/^{86}Sr$ which again could prove useful in a forensic investigation. It is clear from this research that $^{87}Sr/^{86}Sr$ shows some correlation with landuse; this has the potential to be exploited for forensic use.

Like the $^{87}Sr/^{86}Sr$ data set, the trace elemental profile of Norfolk soils does not distinguish between all of the different sample locations but as some sites have clearly different compositions from others it shows that they are unlikely to be from the same source. Again, like with the $^{87}Sr/^{86}Sr$, the trace element data can be used to differentiate between some but not all of the character areas of Norfolk. When the Rare Earth Elemental profile of the soil samples is used the levels of separation between the different character areas increases, improving the intelligence for a forensic investigation.

As was discussed in the introduction, two samples will rarely (if ever) show a perfect chemical fingerprint (20). This was shown not only with the three samples taken from each location, but with the triplicate analyses done on one sample. However, using the chemical profiles of the soils the intra-site profile shows a greater homogeneity than the inter-site profiles. This can aid in forensic science as it can mean that whilst it may not be able to differentiate all of the sample locations, it is possible to say if two samples are likely to be from the same 5m by 5m area.

Although statistically the Bray-Curtis equation and subsequent ordination plots have shown to be a useful tool, court rooms often now use a probabilistic approach in order for a jury to understand the forensic context of some evidence sufficiently. Therefore this may limit the potential for any of these techniques to be used in a forensic context until a probability approach of explaining if soils coming from different areas is possible.

At present it is not possible to make any forensic interpretations of any DNA profiles of plant DNA found in soil samples due to limitations of the techniques. However, the *mat*K gene is a viable option for DNA profiling if it is known what plant species are in the soil

(knowledge of which can be gained from the Norfolk vegetation data) and a suitable extraction technique is developed. As a DNA profile can be sought from a very low amount of sample, if this technique proved to be successful in the future, it could be a great advantage to a forensic investigation. It could also be useful for the analysis of any plant material found at a crime scene, or on a suspect, not just soil samples.

Whilst no MALDI spectra were gained which were suitable for forensic interpretation, if following successful development of the technique, MALDI-ToF-MS could be a useful tool for forensic analysis due to the ease of which a sample can be prepared for analysis once a DNA sample has been extracted and the speed of gaining a spectra for interpretation. However, this assumes that sample variation and sweet spot location variances can be overcome.

Not all of the sites could be distinguished from each other based upon their palynological profile. Intra-site variability was shown to be less than inter-site variability and Bray-Curtis statistical analysis is a suitable technique for seeing whether pollen assemblages match; this can be exploited for forensic purposes. Forensic palynology is known to be a highly valuable, accurate and effective mean of forensic reconstruction and has been used by a number of experts to provide intelligence and evidence. However, consideration should be given to the disproportionally low body of forensic palynology literature looking at the value and reliability of the techniques and evidence in practice. The majority of the literature is case examples (197).

It has to remembered that any one dataset whether it be chemical or biological or the combination of data may seem to be unique to that area of Norfolk but it must be remembered that it is not to say that that combination may not be found elsewhere nationally and indeed globally. But still the data can be a useful tool for intelligent lead forensics.

Whilst it was not possible to unequivocally separate all of the soil samples with any technique, there were instances the data could be used to exclude an area from an investigation. The importance of exclusion in forensic geosciences has been highlighted by Morgan and Bull (22). Forensic geosciences require techniques of exclusion rather than inclusion and an acknowledgment that analytical techniques may be diagnostic only in very specific situations. Realising the limits of the techniques used in soil analysis leads to conjunctive analysis by other techniques, something which has been highlighted in this research. Pringle *et al.* (23) also touch on the subject of the importance of exclusion. As well as potentially being able to exclude areas, in some cases geological trace evidence can help to identify search areas. Therefore available geological trace evidence can aid an intelligence led search.

7:3 Combining the different datasets

The use of a multidisciplinary approach for forensic soil examination has been highlighted in many literature papers (23). As was discussed in the introduction the aim of the project was to use chemical and biological techniques which could be collated and combined to give a unique signature of Norfolk soils as shown in figure 104 below.

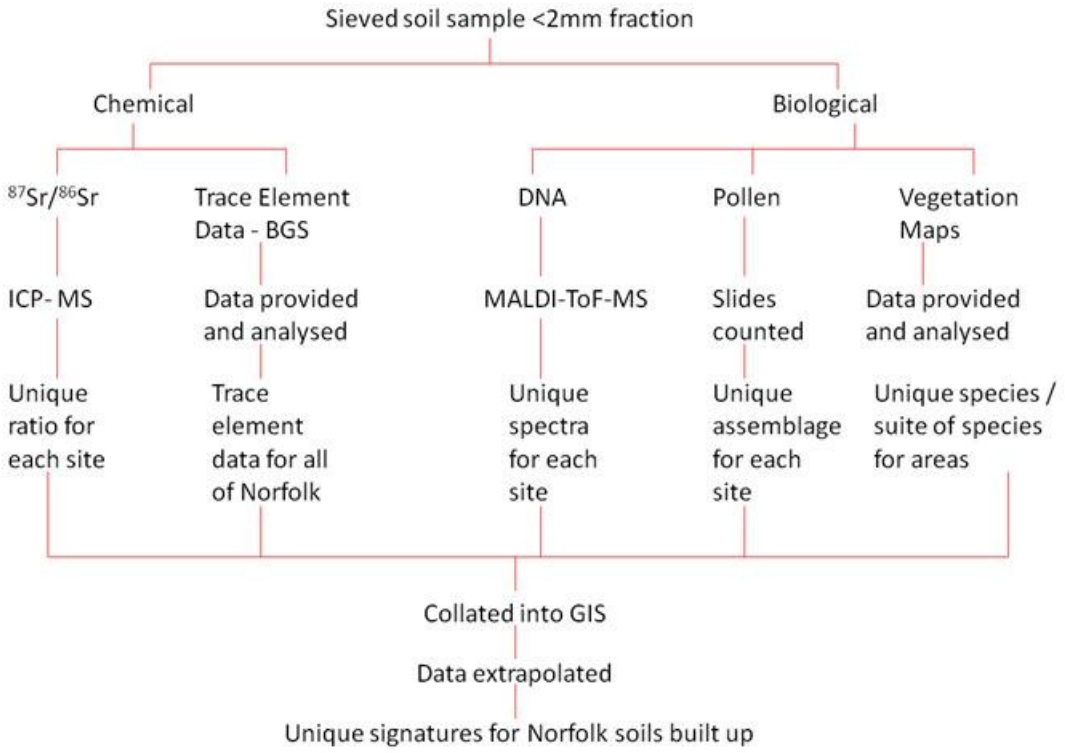


Figure 104: The proposed use of chemical and biological techniques as set out at the start of the research.

Due to limitation of all of the techniques and the inability to gain full datasets in the case of DNA and pollen, unique signatures for all of the Norfolk soils were not built up. However, as shown in figure 105, the techniques used can be used alongside each other in other ways, often providing intelligence, to determine whether or not a soil sample is from a certain origin. The diagram also shows how currently only four of the five techniques tried yield suitable data for any further statistical analysis.

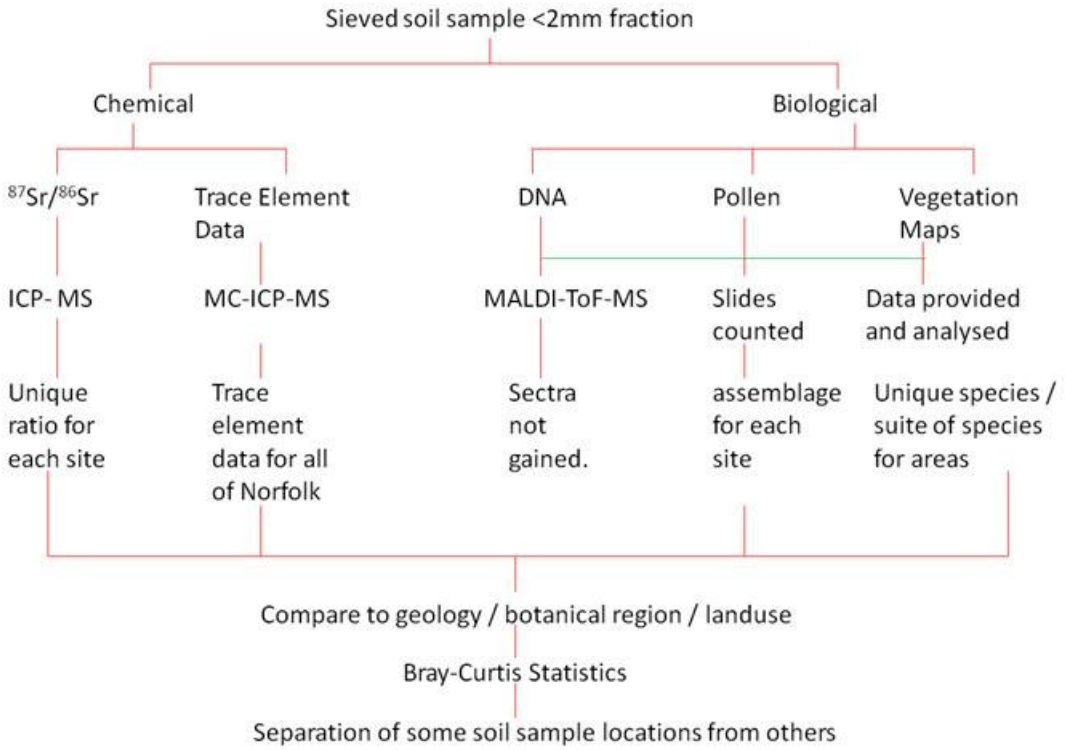


Figure 105: Flow chart showing how the biological and chemical techniques have been used for the analysis of Norfolk soils. The green line highlights the intelligence which gained from the vegetation maps can aid in the other biological techniques.

As can be seen in figure 105 the different techniques can be linked for example pollen and vegetation data can be used in conjunction with each other, with the vegetation data yielding intelligence as to what pollen may be present, and an indication as to the presence of which species would indicate a sample is from a specific area. Sites can be differentiated solely based upon one parameter but with knowledge gained from other parameters the confidence with which sites can be differentiated may be increased.

In the future it may be possible to combine all of the data and to see if this increases the degree of certainty with which it is possible to discriminate samples by using the Bray-Curtis equation and subsequent non-dimensional ordination plots have been used. As the Bray-Curtis equation is not unit dependant the multiple datasets for each of the sample points can all be combined.

A study by Fitzpatrick (198) discussed how a systematic approach for forensic soil examination is applied in three stages. Stage 1 is the rapid characterisation of composite soil particles in whole soil samples for the quick screening of samples. Stage 2 is the detailed characterisation and quantification and individual soil particles following sample selection, size fractionation and detailed mineralogical and organic matter analysis using advanced analytical methods. The final stage, stage 3, is the integration and extrapolation of soil information from one scale to next to build a coherent model of soil information from microscopic observations to the landscape scale. Stages two and three are highlighted in figure 105 above. Each of the individual chemical and biological techniques fit into stage 2, and the use of statistics to look at the soil profiles in relation to the underlying geologies and botanical regions relates to stage 3. If in the future, the datasets were extrapolated within a GIS, this would also fit into stage 3 of a forensic soil analysis. The use of these stages for the analysis and interpretation of the different datasets allows for geographic sourcing to identify the origin of a crime scene soil sample by placing constraints on the environment from which the sample originated.

7:4 Future Research

The main area for future research is the application of exclusionary statistics within the forensic analysis of soils. This has been approached with the use of Bray-Curtis statistics, which have previously been limited to ecological work. It is important to be able to exclude samples from one another in a forensic context and this can be applied to case work. Whilst the field of forensic science is always (and rapidly) advancing and new techniques being developed, though must also be given to the interpretation of any data and how this will help in a criminal investigation and a court of law. Since 2007, the exclusionary approach to forensic geoscience has been discussed indepth (21, 22, 197); this research compliments and adds to this work and highlights the different techniques this approach can be applied to and introduces a new statistical method to give weight to such exclusionary inferences.

In order to exclude samples from one another consideration needs to be given to both the chemical and biological techniques in order to further advance the use of them for the analysis of soils.

7:4.1 Biological Techniques

The analysis of plant DNA found in soils has proven to be very difficult. The use of Twist DX could be advantageous for the DNA analysis of Norfolk soils. TwistDX is a new technology for the amplification of DNA and is based upon Recombinase Polymerase Amplification (RPA). TwistDX has shown to be effective using material generated with sample preparation techniques from a variety of biological sources. This new technique requires only an inexpensive, battery-operated device. It operates at constant body temperature and the process is robust to temperature fluctuations. TwistDX can be combined with simple assays in portable devices. One of the main advantages of this new technology is its capability of amplifying from a single DNA molecule to detectable levels within 15 minutes. The method does not require thermal energy melting steps or any involved setup procedures eliminating the need for extraction methods prior to the amplification. The technology can also operate at reduced ambient temperature.

Benefits of RPA technology:

Portability / point of care

Operates at low / constant temperatures and does not require the initial melting of sample DNA (37 °C). The procedure can be supported by body heat. The process is also fully robust to low-temperature set-up. At a typical ambient temperature (25 °C) the process works, albeit more slowly, and results can typically be obtained in 45 to 90 minutes.

Speed

At temperatures of around 37 °C, RPA is capable of amplifying from a few molecules to detectable levels within 15 minutes. When this performance is combined with the very simple and rapid sample preparation and the decentralized point-of-use device capability, there is the potential for results to be obtained within 30 minutes.

Sensitivity

RPA can detect single copies of DNA and tens of copies of RNA.

Specificity

RPA is so specific it can operate to single molecule levels in the presence of hundreds of nanograms of unrelated complex genomic DNA. This allows for the detection of trace levels of targets even within extremely complex nucleic acid samples.

Broad applicability

RPA can readily be applied to any DNA or RNA target. If reverse transcriptase is included ultra-high sensitivity RNA detection is also possible within the one-pot system.

Simultaneous multiple target detection

Several different potential targets may be detected with a single RPA test as multiple detection primers can be combined in one tube.

Ease of use

RPA has been shown not to require a clean sample (such as those generated by single-step lysis) and thus will be available for portable systems used by untrained personnel.

Low upfront cost burden

Minimal or no device requirements broaden users compared to the more restricted high cost equipment available.

No refrigeration

The reaction system can be stabilized in a dried format permitting transportation and storage without refrigeration (199).

RPA uses five main ingredients; a sample of the DNA to be amplified, a primer-recombinase complex which initiates the copying process when it attaches to the template, nucleotides from which to form the new strands, a polymerase which brings together the nucleotides in

the right order and single-stranded DNA-binding proteins (SSB's) which help keep the original DNA from zipping back together while the new DNA is being made (200).

RPA works by the primer-recombinase complex attaching to the double-stranded DNA which eliminates the need to heat the mixture. Once this complex is in place, it disassembles which allows the DNA polymerase to start to synthesize a new strand of DNA complementary to the template. The SSB's then attach to the displaced strand to stabilize it. Under the right conditions (a precise mixture of process-regulating chemicals) this process will repeat automatically which results in an exponential increase in the DNA sample (200).

Although when the sensitivity and the speed of RPA was tested and it was found to be rapid and accurate some problems were discovered. It was noticed that when using RPA to detect the presence of a specific type of DNA such as a specific pathogen, at low or zero concentrations the primer produced an artefact effect similar to some PCR reactions. This was counteracted with the use of a probe-based detection method which causes the sample to glow in the presence of the DNA being tested for but not in the presence of the primer alone (200).

One of the main areas for further work for the DNA profiling of soil samples would be through the development of suitable extraction methods. Thought should be given to the presence of fungal and grass DNA in a soil sample. If their DNA could be successfully extracted from a sample, then their profiles may be able to be used to differentiate different sample locations.

A recent study by Wallinger *et al.* (201) has looked at the rapid identification of plant species using chloroplast DNA and group specific primers. The paper concedes, like this research, that currently there are limitations to the identification of plant DNA. Currently, there is a lack of broadly applicable methods for the identification of plants when no morphological characteristics are available. This has impeded in the progress made in

ecological disciplines including soil ecology. The study developed a PCR-based approach which identified a variety of plant taxa based on the *trn*T-F cpDNA region using PCR assays for different species. These assays allowed for the identification of plants based on size specific amplicons. The high sensitivity and specificity of the technique enables identification from small amount of plant DNA. Specific primers were combined in multiplex PCR to increase the efficacy of identification of plant species from a large number of samples. The method is suitable for botanical trace evidence and using a set of primers could be applied to forensic botany.

7:4.2 Chemical Techniques

In order for the trace elements and strontium isotopes signature of soils to be used in forensic science routinely much more intense sampling needs to be considered. With a much increased level of samples, a greater degree of separation between different character areas may be found increasing the potential for the use of chemical profiling for forensic provenancing.

It would also be beneficial to look at the chemical profiles at different depths of the soil – agricultural practice will automatically involve a certain turnover of the soil via ploughing so look at a core of the soil and see if there are any more significant correlations as you go down the core of the soil towards the bedrock geology. Look at the different soil horizons.

The increased use of soil databases in the future would be as advantage to forensic science. As more methods have become digital and quantitative, their use in combination as digital profiles will help to characterise soils more accurately and at different scales. Database development will help to refine and narrow probable origin of a questioned sample as well as giving increasingly robust sample comparisons for evidence (196). Databases are best used alongside expertise in assessing whether the landscapes can give rise to an occurrence such as land being too soft for a burial to take place.

Further work with respect to the production of spatial predication maps with ArcGIS could be carried out. Extrapolation of the results in a GIS for example by using kriging (202) can be used to help the predict the provenance of a soil sample.

7.5 Conclusions

In conclusion, at present whilst each of the techniques both biological and chemical have their advantages they are not wholly suitable for forensic provenancing of Norfolk soils. Currently, the profiling of plant DNA in soils is not suitable for purpose until a suitable extraction technique can be developed, this, alongside the development of the *trn* gene and multiplex PCR means that in future it could aid in forensic provenanicng. On a regional scale, all of the soil samples cannot be distinguished from each other using either their biological or chemical profile. However, some of the different character areas of Norfolk, whether geological and botanical, can be distinguished from each other using either strontium isotope ratios, trace elements or pollen analysis. The study has also found that looking at the chemical profiles of soils in respect to land use rather than the more traditional underlying geology can also help to aid in the provenancing.

The use of Bray-Curtis statistical models has shown to be a useful way in which to clearly show whether two soil samples can be excluded from being from the same source using either their chemical or biological profile. The combination of different biological and chemical techniques for the analysis of soil samples has the potential to be a useful tool in forensic science.

References

1. Ruffell A. Forensic pedology, forensic geology, forensic geoscience, geoforensics and soil forensics. Forensic Science International.202(1-3):9-12.
2. Junger EP. Assessing the unique characteristics of close-proximity soil samples, just how useful is soil evidence? Journal of Forensic Science. 1996;41(1):27-34.
3. Murray RC, Tedrow JCF. Forensic Geology. 2nd ed. New Jersey: Prentice-Hall, 1992.
4. Jackson ARW, Jackson JM. Forensic Science: Pearson Education Limited, 2004.
5. Pye K. Geological and Soil Evidence Forensic Applications: CRC Press, 2007.
6. Singer MJ, Munns D, N. Soils An Introduction. 4th ed. New Jersey: Prentice-Hall, Inc., 1999.
7. Wanogho S, Gettingby G, Caddy B, Robertson J. Determination of particle size distibution of soils in forensic science using classical and modern insturmental methods. Journal of Forensic Sciences. 1989;34(4):823 - 35.
8. Petraco N, Kubic TA, Petraco NDK. Case studies in forensic soil examinations. Forensic Science International. 2008;178:23 - 7.
9. Thien SJ, Graveel JG. Laboratory Manual for Soil Science 7th ed: Wm. C. Brown Publishers, 1997.
10. Jenny H. Factors of Soil Formation. A system of Quantitative Pedology: McGraw Hill Book Company, New York, 1941.
11. Murray RC. Evidence from the Earth Forensic Geology and Criminal Investigation. Missoula, Montana: Mountain Press Publishing Company, 2004.
12. Spark-Notes. Environmental Studies. 2009 [updated 2009; cited]; Available from: http://sparkcharts.sparknotes.com/gensci/envsci/section3.php.
13. Krzic M, Wiseman K, Gaumont-Guay D, Dampier L, Grand S. SoilWeb. [cited]; Available from: http://www.landfood.ubc.ca/soil200/components/index.htm.
14. Tan KH. Environmental Soil Science. 2nd ed. New York: Marcel Dekker, Inc., 2000.
15. Thanasoulias NC, Piliouris ET, Kotti MSE, Evmiridis NP. Application of multivariate chemometrics in forensic soil discrimination based on the UV-Vis spectrum of the acid fractions of humus. Forensic Science International. 2002;130:73 - 82.
16. Morgan RM, Wiltshire P, Parker A, Bull PA. The role of forensic geoscience in wildlife crime detection. Forensic Science International. 2006;162:152 - 62.
17. Brown AG, Smith A, Elmhurst O. The Combined Use of Pollen and Soil Analyses in a Search and Subsequent Murder Investigation. Journal of Forensic Sciences. 2002;47(3):614 - 8.
18. Pye K, Blott SJ, Croft DJ, Carter JF. Forensic comparison of soil samples: Assessment of small-scale spatial variability in elemental composition, carbon and nitrogen isotope ratios, colour, and particle size distribution. Forensic Science International. 2006;163(1-2):59-80.
19. Rawlins BG, Cave M. Investigating multi-element soil geochemical signatures and their potential for use in forensic studies. In: Pye K, Croft DJ, editors. Forensic Geoscience: Principles, Techniques and Applications: Geological Society, London, Special Publications; 2004;197 - 206.
20. Pye K, Blott SJ. Comparison of soils and sediments using major and trace element data. In: Pye K, Croft DJ, editors. Forensic Geoscience Principles, Techniques and Applications: Geological Society London Special Publications; 2004.
21. Bull PA, Morgan RM, Freudiger-Bonzon J. A critique of the present use of some geochemical techniques in geoforensic analysis. Forensic Science International. 2008;178(2-3):e35-e40.

22. Morgan R, Bull PA. The philosophy, nature and practice of forensic sediment analysis. Progress in Physical Geography. 2007;31(3):43-58.
23. Pringle JK, Ruffell A, Jervis JR, Donnelly L, McKinley J, Hansen J, et al. The use of geoscience methods for terrestrial forensic searches. Earth-Science Reviews.114(1â€“2):108-23.
24. Antoci PR, Petraco N. A technique for comparing soil colors in the forensic laboratory. Journal of Forensic Sciences. 1993;38:437 - 41.
25. Janssen DW, Ruhf W, Prichard WW. the use of clay for soil color comparisons. Journal of Forensic Sciences. 1983;28(3):773 - 6.
26. Dudley RJ. The use of colour in the discrimination between soils. Journal of Forensic Sciences. 1977;15:209.
27. Croft DJ, Pye K. Colour theory and the evaluation of an instrumental method of measurement using geological samples for forensic applications studies In: Pye K, Croft DJ, editors. Forensic Geoscience: Principles, Techniques and Applications: Geological Society, London; 2004;49 - 62.
28. Dudley RJ. A simple method for determining the pH of small soil samples and its use in Forensic Science. Journal of Forensic Sciences. 1976;16:21 - 7.
29. Stutter MI, Deeks LK, Billet MF. Spatial variability in soil ion exchange chemistry in a granitic upland catchment Journal of the American Soil Science Society. 2004;68:1304 - 14.
30. Dudley RJ, Smalldon KW. The evaulation of methods for soil analysis under simulated scenes of crime conditions. Forensic Science International. 1978;15.
31. Nute HD. An improved density gradient system for forensic soil studies. Journal of Forensic Sciences. 1975;20(4).
32. Petraco N, Kubic T. A density gradient technique used in forensic soil analysis. Journal of Forensic Sciences. 2000;45(4):872 - 3.
33. Sugita R, Marumo Y. Screening of soil evidence by a combination of simple techniques: validity of particle size distribution. Forensic Science International. 2001;122:155 - 8.
34. Wanogho S, Gettinby G, Caddy B, Robertson J. Some factors affecting soil sieve analysis in forensic science 1. Dry Sieving. Forensic Science International. 1987;33:129-37.
35. Fitzpatrick F, Thornton JI. Forensic science characterization of sand. Journal of Forensic Sciences. 1975;20(3).
36. Cengiz S, Sakul O, editors. Capillary Electrophoresis in Forensic Soil Analysis. Proceedings of American Academy of Forensic Sciences 53rd; 1996.
37. Daugherty LA. Soil science contribution to an airplane crash investigation, Ruidoso, New Mexico. Journal of Forensic Sciences. 1997;42(3):401 - 5.
38. Cengiz S, Karaca AC, Cakir I, Uner HB, Sevindik A. SEM-EDS analysis and discrimination of forensic soil. Forensic Science International. 2004;141:33 - 27.
39. Pye K, Croft DJ. Forensic analysis of soil and sediment traces by scanning electron micrscopy and energy dispersive X-ray analysis: an experiment investigation. Forensic Science International. 2007;165:52 - 63.
40. Bull PA, Parker A, Morgan RM. The forensic analysis of soils and sediment taken from the cast of a footprint. Forensic Science International. 2006;162:6 - 12.
41. Siegel JA, Precord C. The analysis of soil samples by reverse phase high performance liquid chromatography using wavelength ratioing. Journal of Forensic Sciences. 1985;30(2):511 - 25.
42. Bommarito CR, Sturdevant AB, Szymanski DW. Analysis of forensic soil samples via high-performance liquid chromatography and ion chromatography. Journal of Forensic Sciences. 2007;52(1).

43. Marumo Y, Nagatsuka S, Oba Y. Clay mineralogical analysis using the <0.05mm fraction for forensic science investigation, its application to volcanic ash soils and yellow-brown forest soils. Journal of Forensic Sciences. 1986;31(1):92 - 105.
44. Graves WJ. A mineralogical soil classification technique for the forensic scientist. Journal of Forensic Sciences. 1979;24(2).
45. Ruffell A, Wiltshire P. Conjunctive use of quantitative and qualitative X-ray diffraction analysis of soils and rocks for forensic analysis Forensic Science International. 2004;145:13 - 23.
46. Rawlins BG, Kemp SJ, Hodgkinson EH, Riding JB, Vane CH, Poulton C, et al. Potentials and Pitfalls in Establishing the Provenance of Earth-Related Samples in Forensic Investigations. Journal of Forensic Science. 2006;51(4):832-45.
47. Hiraoka Y. A possible approach to soil discrimination using X-ray fluorescence analysis Journal of Forensic Sciences. 1994;39(6):1381 - 92.
48. Liddell CM. Field Sampling and Chemical Analysis. Journal of Forensic Science. 1997;42(3):398 - 400.
49. Cox RJ, Peterson HL, Young J, Cusik C, Espinoza EO. The forensic analysis of soil organic by FTIR. Forensic Science International. 2000;108(2):107-16.
50. Weinger BA, Reffner JA, De Forest PR. A Novel Approach to the Examination of Soil Evidence: Mineral Identification Using Infrared Microprobe Analysis. Journal of Forensic Science. 2009;54(4).
51. Madejová J. FTIR techniques in clay mineral studies. Vibrational Spectroscopy. 2003;31(1):1-10.
52. Pye K, Blott SJ, Wray DS. Elemental analysis of soil samples for forensic purposes by inductively coupled plasma spectrometry -- precision considerations. Forensic Science International. 2006;160(2-3):178-92.
53. Barreto SRG, MNozaki J, Oliveira E, Filho VF, Aragao PHA, Scarminio IS, et al. Comparison of metal analysis in sediments using EDXRF and ICP-OES with HCl and Tessie extraction methods Talanta. 2004;64:345 - 54.
54. Navas A, Lindhorfer H. Geochemical speciation of heavy metals in semiarid soils of the central Ebro Valley (Spain) Environmental International. 2003;29:61 - 8.
55. Croft DJ, Pye K. Multi-technique comparison of source and primary transfer soil samples: an experimental investigation. Science & Justice. 2004;44(1):21-8.
56. Ruffell A, McKinley J. Geoforensics. Chichester: John Wiley and Sons, Ltd, 2008.
57. Virtanen V, Korpelainen H, Kostamo K. Forensic Botany: Usability of bryophyte material in forensic studies. Forensic Science International. 2007;172:161-3.
58. Coyle HM, Shutler G, Tully L, Pagliaro E, Harper A, Palmbach T, et al. Validation of a DNA Method for the Individualization of Plant Evidence. In press 2005.
59. Council NC. Where I Live. 2007-8 [updated 2007-8; cited]; Available from: www.whereilive.norfolk.gov.uk/LocalView/.
60. Fisher J. Genealogy and Norfolk, England. [cited]; Available from: http://homepages.nildram.co.uk/~jimella/norfolk.htm.
61. Beckett G, Bull A, Stevenson R. A Flora of Norfolk. Thetford: Jarrold Book Printing, 1999.
62. Toghill P. Chalk Seas and the end of the Dinosaurs - The Creataceous Period. The Geology of Britain - An Introduction: Airlife Publishing; 2006;145 - 53.
63. Toghill P. Ice Ages and Mammoths - the Quaternary Period. The Geology of Britain: An Introduction: Airlife Publishing; 2006;169 - 84.
64. British Geological Survey. Cutting the Ice in East Anglis Unravelling the Glacial History. Earthwise. 2004(20):20 - 1.

65. Pye K. Sampling and Sample Handling. Geological and Soil Evidence Forensic Application: CRC Press; 2007.
66. McKinley J, Ruffell A. Contemporaneous spatial sampling at scenes of crime: Advantages and disadvantages. Forensic Science International. 2007;172:196-202.
67. Heath L, Saunders V. Assessing the Potential of Bacterial DNA Profiling for Forensic Soil Comparisons. Journal of Forensic Science. 2006;51(5).
68. Sugita R, Marumo Y. Validity of colour examination for forensic soil identification. Forensic Science International. 1996;83(3):201-10.
69. Fraysier H, Van Hoven H. A simpler mineralogical soil analysis method using dispersion staining. Microscope. 1993;40(3):107-10.
70. Butler JM. Forensic DNA Typing. 2nd ed: Elsevier Academic Press, 2005.
71. Horswell J, Cordinger S, Maas E, Martin T, Sutherland B, Speir T. Forensic Comparisons of Soils by Bacetrial Communities. Journal of Forensic Science. 2001;47(2):350-3.
72. Balouet JC, Oudijk G, Petrisor I, Morrison R. Emerging Forensic Techniques. In: Murphy BL, Morrison RD, editors. Introduction to Environmental Forensics
Elsevier Inc.; 2007.
73. Kitts CL. Terminal restriction fragment patterns: A tool for comparing microbila communities and assessing community dynamics. Current Issues in Intestinal Microbiology. 2001;2(1):17-25.
74. Sensabaugh G, Kaye DH. Non-Human DNA Evidence. Jurimetrics Journal. 1998;38:1-16.
75. Yoon CK. Forensic Science. Botanical witness for the prosecution. Science. 1993;260:894-5.
76. Tsai L-C, Yu Y-C, Hsieh H-M, Wang J-C, Linacre A, Lee JC-I. Species identification using sequences of the trnL intron and the trnL-trnF IGS of chloroplast genome among popular plants in Taiwan. Forensic Science International. 2006;164(2-3):193-200.
77. Korpelainen H, Virtanen V. DNA Fingerprinting in Mosses. Journal of Forensic Science. 2003;48(4).
78. Government US. Presidents DNA Initiative. 2007 [updated 2007; cited]; www.dna.gov
]. Available from.
79. Nagy M. Automated DNA extraction techniques for forensic analysis. In: Rapley R, Whitehouse D, editors. Molecular Forensics: Wiley; 2007.
80. Berg J, Tymoczko J, Stryer L. Biochemistry. 5th edition ed: Freeman, 2006.
81. Goodwin W, Linacre A, Hadi S. An Introduction to Forensic Genetics: Wiley, 2007.
82. Cano R, Borucki M. Revival and identification of bacterial spores in 25 to 40-million-year-old Dominican amber. Science. 1995;268:1060-4.
83. Franklin R, Garland J, Bolster C, Mills A. Impact of dilution on microbial community structure and functional potential: Comparison of numerical simulations and batch culture experiments. Applied Environmental Microbiology. 2001;67:702-12.
84. Kaplan C, Kitts CL. Bacterial succession in a petroleum land treatment unit. Applied and Environmental Microbiology. 2004;March 1777-86.
85. Kaplan CW, Astaire JC, Sanders ME, Reddy BS, Kitts CL. 15s Ribosomal DNA terminal restriction fragment pattern analysis of bacterial communities in feces of rats fed lactobacillus acidophilus NCFM. Applied and Environmental Microbiology. 2001;April:1935-9.
86. Brodie E, Edwards S, Clipson N. Soil fungal community structure in a temperate upland grassland soil. FEMS Microbiology Ecology. 2003;45(2):104-14.

87. Cappellato R, Peters N, Meyers T. Above-ground sulfur cycling in adjacent coniferous and deciduous forests and watershed sulfur retention in the Georgia Piedmont, USA. Water, Air and Soil Pollution. 1998;103(1-4):151-71.
88. Heun M, Schafer-Pregl R, D, Castagna R, Accerbi M, Borghi B, Salamini F. Site of einkorn wheat domestication identified by DNA fingerprinting. Science. 1997;278:1312-4.
89. Law J, Donini P, Koebner R, Jones C, Cooke R. DNA profiling and plant variety registration III: The statistical assessment of distinctness in wheat using amplified fragment length polymorphisms. Euphytica. 1998;102:335-42.
90. Lee CL, Coyle HM, Lee HC. Genetic Analysis of Individual Seeds by Amplified Fragment Length Polymorphism Method. Croatian Medical Journal. 2007;48:563-5.
91. Mace E, Gebhardt C, Lester R. AFLP analysis of genetic relationships in the trible Datureae (Solanaceae). Theoretical and Applied Genetics. 1999;99:634-41.
92. Melcher U. Molecular Genetics. 2001 [updated 2001; cited]; Available from: http://opbs.okstate.edu.
93. Russell JR, Webber JC, Booth A, Powell W, Sotelo-Montes C, Dawson IK. Genetic Variation of Calycophyllum spruceanum in the Peruvian Amazon Basin, revealed by amplified fragemnt length polymorphism (AFLP). Molecular Ecology. 1999;8:199-204.
94. Xu M, Melchinger A, Xia X, Lubberstedt T. High-resolution mapping of loci conferring resistance to sugarcase mosaic virus in maize using RFLP, SSR and AFLP markers. Molecular and General Genetics. 1999;261:574-81.
95. Dergam J, Paiva S, Schaeffer C, Godinho A, Vieira F. Phylogeography and RAPD-PCR variation in Hoplias malabaricus in southeastern Brazil. Genetics and Molecular Biology. 2002;25(4).
96. Krane D, Sternberg D, Burton G. Randomly amplified polymorphic DNA profile-based measures of genetic diversity in crayfish correlated with environmental impacts. Environmental Toxicology and Chemistry. 1999;18:504-8.
97. Manners MM, Morvillo N, RFrederick C, Wagner A. RAPD-PCR answers some long-standing questions about rose identification. Acta Horticulturae (ISHS). 2004;634:85-9.
98. Shianna K, Rytter R, Spanier J. Randonly amplified polymorphic DNA PCR analysis of bovine Cryptosporidium parvum strains isolated from the watershed of the Red River of the North. Applied and Environmental Microbiology. 1998;64(6):2262-5.
99. Srivastava AK, Ara A, Bhargava P, Mishra Y, Rai SP, Rai C. A rapid and cost effective method of genomic DNA isolation from cyanpbacterial culture, mat and soil suitable for genomic fingerprinting and community analysis. Journal of Applied Phycology. 2007;19:373-82.
100. Liu W, Marsh T, Cheng H, Forney L. Characterisation of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16s rRNA. Applied Environmental Microbiology. 1997;63(11):4516-22.
101. Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. TRENDS in Genetics. 2007;23(4).
102. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(23):8369-74.
103. Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, et al. Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well. Plos One. 2008;3(7).
104. Hill J, Strausbaugh L, Graf J. Soil DNA Typing in Forensic Sience. In: Miller Coyle H, editor. Nonhuman DNA Typing Theory and Casework Applications: CRC Press; 2008.

105. Franklin RB, Taylor DR, Mills AL. Characterization of microbial communities using randomly amplified polymorphic DNA (RAPD). Journal of Microbial Methods. 1999;35:225-35.
106. Nybom H, Weising K, Rotter Br. DNA fingerprinting in botany: past, present, future. Investigative Genetics.5:1-.
107. Weising K, Nybom H, Wolff K, Kohl G. DNA Fingerprinting in Plants Principles, Methods and Applications: CRC Press, 2005.
108. Qiagen. QIAamp DNA Stool Handbook. 2nd ed, 2007.
109. Qiagen. DNeasy Blood and Tissue Kit. 2006.
110. Wilson CA. Phylogeny of Iris based on chloroplast matK gene and trnK intron sequence data. Molecular Phylogenetics and Evolution. 2004;33(2):402-12.
111. Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, et al. Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well. PLoS ONE. 2008;3(7):e2802.
112. MOBio Laboratories Inc. PowerSoil™ DNA Isolation Kit Instruction Manual, 2007.
113. Miesfeld RL. Lecture 12 - The Polymerase Chain Reaction (PCR). 2001 [updated 2001; cited]; Available from: www.biochem.arizona.edu.
114. Miller-Coyle H, editor. Forensic Botany Principles and Applications to Criminal Casework: CRC Press, 2005.
115. Rachmayanti Y, Leinemann L, Gailing O, Finkeldey R. DNA from processed and unprocessed wood: Factors influencing the isolation success. Forensic Science International: Genetics. 2009;3(3):185-92.
116. Craft KJ, Owens JD, Ashley MV. Application of plant DNA markers in forensic botany: Genetic comparison of Quercus evidence leaves to crime scene trees using microsatellites. Forensic Science International. 2007;165(1):64-70.
117. Hillenkamp F, Karas M. The MALDI Process and Method. In: Hillenkamp F, Peter-Katalinic J, editors. MALDI MS A Practical Guide to Instrumentation, Methods and Applications: Wiley-VCH; 2007.
118. Tanka C, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T. Protein and Polymer Analysis up to m/z 100,000 by Laser Ionization Time of FLight Mass Spectrometry. Rapid Communications in Mass Spectrometry. 1988;2:151 - 3.
119. Westman-Brinkmalm A, Brinkmalm G. A Mass Spectrometer's Building Blocks. In: Ekman R, Silberring J, Westman-Brinkmalm AM, Kraj A, editors. Mass Spectrometry Instrumentation, Interpretation and Applications. New Jersey: John Wiley & Sons, Inc.; 2009.
120. Aldrich S. MALDI Mass Spectrometry. Analytix Advances in Analytical Chemistry. 2001.
121. Butler JM. Forensic DNA Typing: Biology, Technology and Genetics of STR Markers. 2nd ed: Elsevier Acadmenic Press; 2005.
122. Hammes GC. Spectroscopy for the Biological Sciences. New Jersey: John Wiley & Sons, Inc, 2005.
123. Biosciences A. Ettan MALDI-TOF Method Handbook. Edition AA ed: Amersham Biosciences, 2000.
124. Ashcroft AE. An Intorduction to Mass Spectrometry. Unknown [updated Unknown; cited]; Available from: www.ashbury.leeds.ac.uk/facil/MStut/MStutorial.htm.
125. O'Connor PB, Hillenkamp F. MALDI Mass Spectromery Insturmentation. In: Hillenkamp F, Peter-Katalinic J, editors. MALDI MS: A Practical Guide to Instrumentation, Methods and Applications Wiley-VCH; 2007.
126. Knochenmuss R. Photoionization pathways and free electrons in UV-MALDI. Analytical Chemistry. 2004;16:1287 - 94.

127. Karas M, Bachmann D, Bahr U, Hillenkamp F. Matrix-assisted ultraviolet laser desportion of non-volatile compounds. International Journal of Mass Spectrometry and Ion Processes. 1987;78:53 - 68.
128. Ehring H, Karas M, Hillenkamp F. Role of photoionization processes or organic molecules and relevance for matrix-assisted laser desorption ionization mass spectrometry. Organic Mass Spectrometry. 1992;27:472 - 80.
129. Karas M, Gluckmann M, Schaefer J. Ionization in matrix-assisted laser desoprtion/ionization: singly charged molecule ions are lucky survivors. Journal of Mass Spectrometry. 2000;35:1 - 12.
130. Kumooka Y. Analysis of rosin and modified rosin esters in adhesives by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Forensic Science International. 2008;176(2):111-20.
131. Spencer SE, Kim SY, Kim SB, Schug KA. Matrix-assisted laser desorption/ionizationâ€“time of flight-mass spectrometry profiling of trace constituents of condom lubricants in the presence of biological fluids. Forensic Science International.207(1):19-26.
132. Tanaka H, Degawa M, Kawata E, Hayashi J, Shoyama Y. Identification of Cannabis pollens using an allergic patient's immunoglobulin E and purification and characterization of allergens in Cannabis pollens. Forensic Science International. 1998;97(2):139-53.
133. Sauer S. The essence of DNA sample preparation for MALDI mass spectrometry. Journal of Biochemical and Biophysical Methods. 2007;70:311 - 8.
134. Petkovski E, Keyser-Tracqui C, Hienne R, Ludes B. SNPs and MALDI-TOF MS: Tools for DNA Typing in Forensic Paternity Testing and Anthrpology. Journal of Forensic Science. 2005;50(3).
135. Shi Y, Xiang P, Li L, Shen M. Analysis of 50 SNPs in CYP2D6, CYP2C19, CYP2C9, CYP3A4 and CYP1A2 by MALDI-TOF mass spectrometry in Chinese Han population. Forensic Science International. 2011;207(1):183-7.
136. Horneffer V, Strupat K, Hillenkamp F. Localization of Noncovalent Complexes in MALDI-Preparations by CLSM. Journal of the American Society for Mass Spectrometry. 2006;17(11):1599-604.
137. Papac DI, Wong A, Jones AJS. Analysis of Acidic Oligosaccharides and Glycopeptides by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. Analytical Chemistry. 1996;68:3215 - 23.
138. Williams JB, Chapman TM, Hercules DM. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Discrete Mass Poly(butylene glutarate) Oligomers. Analytical Chemistry. 2003;75(13):3092-100.
139. Williams TL, Andrzejewski D, Lay JO, Musser SM. Experimental factors affecting the quality and reproducibility of MALDI TOF mass spectra obtained from whole bacteria cells. Journal of the American Society for Mass Spectrometry. 2003;14(4):342-51.
140. Krupkova A, Cermak J, Walterova Z, Horsky J. Quantitative Interpretation of MALDI-TOF Mass Spectra of Imperfect Carbosilane Dendrimers. Analytical Chemistry. 2007;79:1639 - 45.
141. Ressom HW, Varghese RS, Goldman L, Loffredo CA, Abdel-Hamid M, Kyselova Z, et al. Analysis of MALDI-TOF Mass Spectrometry Data for Detection of Glycan Biomarkers. Pacific Symposium on Biocomputing. 2008;13:216 - 27.
142. Taranenko NI, Hurt R, Zhou JZ, Isola NR, Huang H, Lee SH, et al. Laser desorption mass spectrometry for microbial DNA analysis. Journal of Microbiological Methods. 2002;48(2-3):101-6.

143. Westman-Brinkmalm A, Brinkmalm G. Tandem Mass Spectrometry. In: Ekman R, Silberring J, Westman-Brinkmalm AM, Kraj A, editors. Mass Spectrometry Instrumentation, Interpretation and Applications. New Jersey: John Wiley & Sons, Inc.; 2009.
144. Faure G. Principles and Applications of Geochemistry. 2nd ed: Prentice Hall, 1998.
145. Pye K. Bulk Properties of Geological and Soil Materials. Geological and Soil Evidence Forensic Applications: CRC Press, Taylor and Francis Group; 2007;59 - 118.
146. Faure G. Principles of Isotope Geology. New York: John Wiley and Sons, 1986.
147. Faure G, Mensing TM. Isotope Principles and Applications: John Wiley and Sons, 2005.
148. White WM. Radiogenic Isotope Geochemistry. Geochemistry: Online Textbook: http://www.geo.cornell.edu/geology/classes/geo455/Chapters.HTML; 2007.
149. Aggarwal J, Habicht-Mauche J, Juarez C. Application of heavy isotopes in forensic isotope geochemistry. Applied Geochemistry. 2008;23.
150. Faure G, Powell LJ. Strontium Isotope Geology: Springer-Verlag, 1972.
151. Nier AD. Isotopic Constitution of Sr, Ba, Bi, Tl and Hg. Physical Review. 1938;54:275 - 8.
152. Beard B. Strontium isotope composition of skeletal material can determine the birth place and geographical mobility of animals and humans. Journal of Forensic Science. 2000;45(5):1049 - 61.
153. Turekian KK. Distribution of the elements in some major units of the Earth's crust. Geological Society of America Bulletin. 1961;72:175 - 82.
154. Alexander Bentley R. Strontium Isotopes from the Earth to the Archaeological Skeleton: A Review. Journal of Archaeological Method and Theory. 2006;13(3):135-87.
155. Allegre CJ. Isotope Geology: Cambridge University Press, 2008.
156. Evans JA, Montgomery J, Wildman G, Boulton N. Spatial variations in biosphere 87Sr/86Sr in Britain. Journal of the Geological Society. 2010;167:1-4.
157. Bondarento O, Rolova T, Kahru A, Ivask A. Bioavailibility of Cd, Zn and Hg in soil to nine recombinant luminescent metal sensor bacteria. Sensors. 2008;8:6899 - 923.
158. Sillen A, Hall G, Richardson S, Armstrong R. 87Sr/86Sr ratios in modern and fossil food-webs of the Sterkfontein Valley: implications for early hominid habitat preference. Geochimica et Cosmochimica Acta. 1998;62(14):2463-73.
159. Kelly S, Heaton K, Hoogewerff J. Tracing the geographical origin of food: The application of multi-element and multi-isotope analysis. Trends in Food Science & Technology. 2005;16(12):555-67.
160. Hoffmann CM, Brunelle RI, Snow KB. Forensic comparisons of soils by neutron activation and atomic absorption analysis. Journal of Criminal Law, Criminology and Police Science. 1969;60:395 - 401.
161. Rawlins BG, Cave M. Investigating multi-element soil geochemical signatures and their potential for use in forensic studies. In: Pye K, Croft DJ, editors. Forensic Geoscience: Principles, Techniques and Applications: The Geological Society of London; 2004;197-206.
162. Wilson AD. The Sampling of Silicate Rock Powders for Chemical Analysis. Analyst. 1964;89:18-30.
163. Pye K, Blott SJ. Comparison of soils and sediments using major and trace element data. In: Pye K, Croft DJ, editors. Forensic Geoscience: Principles, Techniques and Applications: Geological Society, London, Special Publications; 2004;183-96.
164. Parnell JJ, Terry RE, Nelson Z. Soil Chemical Analysis Applied as an Interpretive Tool for Ancient Human Activities in Piedras Negras, Guatemala. Journal of Archaeological Science. 2002;29:379-404.

165. Entwistle JA, Abrahams PW, Dodgshon RA. Multi-Element Analysis of Soils from Scottish Historical Sites. Interpreting Land-Use History Through the Physical and Geochemical Analysis of Soil. Journal of Archaeological Science. 1998;25:53-68.
166. Webster R. Is soil variation random? Geoderma. 2000;97:149-63.
167. Morgan R, Bull A. Forensic Geoscience and Crime Detection. Minerva Medica Legal. 2007;127:73-9.
168. British Geological Survey. Digital Geology - Bedrock Geology Theme. 2009 [updated 2009; cited]; Available from: www.bgs.ac.uk/science.
169. British Geological Survey. Digital Geology - Superficial Deposits Theme. 2009 [updated 2009; cited]; Available from: www.bgs.ac.uk/science.
170. British Geological Survey. Soil Parent Material Model (PMM). 2009 [updated 2009; cited]; Available from: www.bgs.ac.uk/science.
171. Eichrom. Radiochemistry:Resin. [cited]; Available from: http://www.eichrom.com/products/info/sr_resin.cfm.
172. Micromass. Isoprobe Users Guide: Micromass Ltd.
173. Makishima A, Nath BN, Nakamura E. New sequential separation procedure for Sr, Nd and Pb isotope ratio measurement in geological material using MC-ICP-MS and TIMS. Geochem J. [Article]. 2008;42(3):237-46.
174. Rehkamper M, Mezger K. Investigation of matrix effects for Pb isotope ratio measurements by multiple collector ICP-MS: verification and application of optimized analytical protocols. Journal of Analytical Atomic Spectrometry. 2000;15:1451-60.
175. Gillson GR, Douglas DJ, Fulford JE, Halligan KW, Tanner SD. Nonspectroscopic interelement interferences in inductively coupled plasma mass spectrometry. Analytical Chemistry. 1988 2014/11/17;60(14):1472-4.
176. Harrison J. Allotment Vegetable Growing. 2004-2011 [updated 2004-2011; cited]; Available from: www.allotment.org.uk.
177. Kabala-Pendias A. Trace Elements in Soils and Plants. 3rd Edition ed: CRC Press, 2000.
178. Pye K, Blott SJ. Development of a searchable major and trace element database for use in forensic soil comparisons. Science and Justice. 2009.
179. Committe JNC. National Vegetation Classification (NVC). Unknown [updated Unknown; cited]; Available from: http://jncc.defra.gov.uk/page-4259.
180. Wheeler B. The Wetland Plant Communities of the River Ant Valley, Norfolk. Transactions of the Norfolk and Norwich Naturalist Society. 1978;24:153 - 87.
181. DEFRA. The Common Land of Norfolk - A Biological Survey. 1998.
182. Faegri K, Iversen J. Textbook of Pollen Analysis. 4th ed: The Blackburn Press, 1989.
183. Moore PD, Webb JA, Collinson ME. Pollen Analysis. 2nd Edition ed: Blackwell Scientific Publications, 1991.
184. Erdtman G. Handbook of Palynology. Copenhagen: Munksgaard, 1969.
185. Nowicke JW, Meselson M. Yellow rain - a palynological analysis. Nature. 1984;309:205-6.
186. Mildenhall DC. Forensic Palynology in New Zealand. Review of Palaeobotany and Palynology. 1990;64:227 - 34.
187. Morgan RM, Flynn J, Sena V, Bull PA. Experimental forensic studies of the preservation of pollen in vehicle fires. Science and Justice. 2014/11/20;54(2):141-5.
188. Morgan RM, Davies G, Balestri F, Bull PA. The recovery of pollen evidence from documents and its forensic implications. Sci Justice. Dec;53(4):375-84.
189. Riding JB, Rawlins BG, Coley KH. Chnages In Soil Pollen Assemblages on Footwear Worn at Different Sites. Palynology. 2007;31:135 - 51.

190. Wiltshire PEJ. Consideration of some taphonomic variables of relevance to forensic palynological investigation in the United Kingdom. Forensic Science International. 2006;163(3):173-82.
191. Mildenhall DC, Wiltshire PED, Bryant VM. Forensic Palynology: Why do it and how it works. Forensic Science International. 2006;163:163-72.
192. Horrocks M, Coulson SA, Walsh KAJ. Forensic Palynology: Variation in the Pollen Content of Soil Surface Samples. Journal of Forensic Science. 1998;43(2):320-3.
193. Horrocks M, Coulson SA, Walsh KAJ. Forensic palynology: variation in the pollen content of soil on shoes and in shoeprints in soil. Journal of Forensic Science. 1999;44(1):119 - 22.
194. Horrocks M. Sub-sampling and Preparing Forensic Samples for Pollen Analysis. Journal of Forensic Science. 2004;49(5):4.
195. Bhattacharya T, Beach T, Wahl D. An analysis of modern pollen rain from the Maya lowlands of northern Belize. Review of Palaeobotany and Palynology. 2011;164:109 - 20.
196. Dawson LA, Hillier S. Measurement of soil characteristics for forensic applications. Surface and Interface Analysis. 2010;42(5):363-77.
197. Morgan RM, Flynn J, Sena V, Bull PA. Experimental forensic studies of the preservation of pollen in vehicle fires. Sci Justice. 2012 Mar;54(2):141-5.
198. Fitzpatrick RW. Soil: Forensic Analysis. Wiley Encyclopedia of Forensic Science: John Wiley & Sons, Ltd; 2009.
199. Armes NA, Piepenburg D. TwistDX. 2008 [updated 2008; cited]; Available from: www.twistdx.com.
200. Hoff M. DNA Amplification and Detection Made Simple (Relatively). PLoS Biology. 2006;4(7).
201. Wallinger C, Juen A, Staudacher K, Schallhart N, Mitterrutzner E, Steiner E-M, et al. Rapid Plant Identification Using Species- and Group-Specific Primers Targeting Chloroplast DNA. PLoS ONE. 2012;7(1):e29473.
202. Lark RM, Rawlins BG. Can we predict the provenance of a soil sample for forensic purposes by refernce to a spatial database? European Journal of Soil Science. 2008;59:1000-6.

Appendix 1

Strontium Isotope Ratios

This table shows all of the strontium isotope ratios taken from each site, the standard deviations of the measurements and their subsequent underlying geology and botanical regions.

Site	Replicates								St dev	Geology	Vegetation
160708_12	0.710913	0.711147	0.710692						0.000228	Chalk	Cromer Ridge
160708_4	0.707884	0.7079	0.707888	0.707878	0.707898				9.38E-06	Chalk	Goodsands
160708_5	0.707645	0.707797	0.707827	0.707805	0.70772	0.70774	0.707711		6.43E-05	Norwich Crag	NE Norfolk
160708_9	0.707854	0.707962	0.707989	0.707933	0.707939				5.05E-05	Chalk	Goodsands
220909_5	0.708082	0.708086	0.708077						4.64E-06	Upper Greensand	Fenlands
220909_7	0.709406	0.709404	0.709383	0.709357	0.709432				2.83E-05	Amptil Clay	Fenlands
220909_9	0.708068	0.708064	0.70803						2.05E-05	Chalk	Breckalnds
230909_1	0.708674	0.708656	0.708643	0.70863	0.708588				3.23E-05	Norwich Crag	Wensum Sands
230909_10	0.708283	0.708287							2.83E-06	Chalk	Clays
230909_11	0.708752	0.708714	0.708699	0.708713					2.3E-05	Chalk	Clays
230909_2	0.708707	0.708719	0.708683	0.708697	0.708864				7.37E-05	Norwich Crag	NE Norfolk
230909_3	0.709108	0.709044	0.709077	0.70902					3.86E-05	Norwich Crag	NE Norfolk
230909_4	0.708973	0.709049	0.708915	0.708922					6.19E-05	Norwich Crag	Broads
230909_5	0.708911	0.708044	0.708914	0.708046	0.70893	0.708988	0.708087	0.708975	0.000459	Norwich Crag	

230909_6	0.708228	0.708222	0.708241	0.708193	0.708206				1.88E-05	Norwich Crag	Clays
230909_7	0.70821	0.708051	0.708074	0.708065	0.708107				6.41E-05	Norwich Crag	Broads
230909_8	0.708784	0.708782	0.709017						0.000136	Norwich Crag	Clays
230909_9	0.708527	0.708511	0.708506	0.708462	0.708521				2.57E-05	Chalk	Breckalnds
240708_1	0.707829	0.707824	0.707839	0.707841	0.707754				3.61E-05	Lower Creaceous	Greensand
240708_2	0.709861	0.710157	0.710073						0.000153	Lower Creaceous	Greensand
240708_3	0.710489	0.710499	0.710508	0.710628	0.71006				0.000218	Lower Creaceous	Greensand
240708_4	0.708001	0.710314	0.708961						0.001162	Lower Creaceous	Greensand
240708_5	0.707773	0.707754	0.707779	0.707716					2.86E-05	Chalk	Greensand
240708_6	0.70851	0.707938	0.708614						0.000364	Upper Greensand	Greensand
240708_7	0.710089	0.708272	0.708267	0.708238	0.708248				0.00082	Upper Greensand	Greensand
170610_10	0.708381	0.708522	0.708368						8.53E-05	Chalk	Breckland
170610_11	0.709098	0.708944	0.70906						8.03E-05	Chalk	Clays
170610_12	0.708648	0.708656	0.708626	0.70868	0.708579				3.84E-05	Norwich Crag	Clays
170610_6	0.708411	0.70824	0.708364	0.708381					7.49E-05	Chalk	Goodsands

170610_7	0.708089	0.708143	0.70819						5.05E-05	Upper Greensand	Greensand
170610_8	0.707795	0.710102	0.710129						0.00134	Lower Creataceous	Greensand
170610_9	0.708469	0.70866	0.708623	0.708632	0.708557				7.67E-05	Chalk	Greensand
170610_1	0.709269	0.709266	0.709384						6.7E-05	Norwich Crag	Broads
170610_2	0.708566	0.708591	0.708476						6.05E-05	Norwich Crag	NE Norfolk
170610_3	0.708466	0.708501	0.708453						2.47E-05	Norwich Crag	NE Norfolk
170610_4	0.708295	0.708248	0.708257						2.5E-05	Norwich Crag	Clays
170610_5	0.709645	0.709762	0.70959						8.81E-05	Chalk	Clays
210909_2	0.708549	0.708587	0.708516						3.54E-05	Chalk	Breckland
210909_4	0.709015	0.708745	0.708619						0.000202	Chalk	Breckland
210909_8	0.707718	0.707719	0.707691						1.6E-05	Chalk	Greensand
210909_1	0.708589	0.708579	0.708586	0.708559					1.33E-05	Chalk	Clays
210909_3	0.708047	0.708059	0.708028	0.707999	0.707785				0.000113	Chalk	Breckland
210909_6	0.709151	0.709123	0.709168						1.95E-05	Amptil Clay	Fenlands

Appendix 2

Trace Element Data

The following tables show all of the trace element data, firstly grouped according to geology and secondly according to land use.

Geology	Li	Be	B	Na	Mg	Al	K	Ca	Sc	Ti	V	Mn	Fe
Amtill Clay	25.94958	0	190.0292	11148.75	81008.33	0	231420.8	919458.3	1.986625	4.796792	4.349875	31.425	12.95208
Lower Cretaceous	2.622683	0.00284	45.41067	2557.432	11787.61	452.6426	17026.56	190946.9	0.260988	1.675498	0.893285	2012.208	112.6368
UGG	17.46591	0	58.96692	4080.558	17296.66	0	60469.8	1187597	0.636396	1.965179	0.766397	26.38429	69.68145
Chalk	7.643026	0.0048	91.28206	7559.336	20243.16	632.1077	69037.53	566024.7	0.716144	2.740497	1.681997	1927.802	111.1163
Norwich Crag	12.53669	0	108.5108	9091.664	27076.54	0	130010.9	698335.6	0.876029	1.581256	1.390727	368.3715	22.39769

Co	Ni	Cu	Zn	Ga	As	Se	Rb	Sr	Y	Zr	Mo	Ag
2.114625	10.19958	54.91917	0	0	11.02708	0.614513	157.3083	2538.875	0.075565	0.152985	5.739583	0.623458
4.690059	14.09647	5.572224	464.1703	0.032508	4.580242	0.216662	28.05389	523.9946	0.416811	0.594887	0	0.781139
2.23542	6.125431	13.61323	1.244326	0	2.765596	0.957176	43.7717	4494.269	0.078512	0.223237	1.763123	0.699138
4.727317	10.71397	28.36149	154.75	0.012116	4.459472	0.86574	68.15552	1404.937	0.656861	0.657754	1.287354	1.344724
2.51828	7.775146	21.47419	29.27606	0	4.033371	0.885229	70.44148	2268.695	0.172815	0.28701	2.880661	0.930973

Cd	Sb	Cs	Ba	La	Ce	Pr	Nd	Sm	Eu	Gd	Ho
0.192397	1.004125	3.42125	5011.25	0.052233	0.01643	0	0.01982	0	1.28375	0.058762	0
1.951104	0	0.762487	903.4843	0.658058	1.12201	0.122226	0.480773	0.091444	0.239614	0.105737	0.014178
0.127267	0.023967	0.836609	2295.013	0.045816	0.010466	0	0.013499	0	0.557807	0.009208	0.000787
1.325045	0.333172	1.376655	2975.451	1.016325	1.206192	0.142845	0.565266	0.098955	0.754565	0.143707	0.018353
0.534727	0.229965	1.838618	3536.377	0.17235	0.153922	0.019774	0.098733	0	0.849034	0.043211	0.005878

Lu	Tl	Pb	Th	U
0.001467	1.312125	0.017297	0	0.091562
0.001752	0.503693	15.97596	0.118266	0
0	0.641827	0	0.000935	0.270154
0.003664	0.737621	14.64499	0.088649	0.161952
0	0.732532	0.256363	0.004633	0.233044

Land Use	Li	Be	B	Na	Mg	Al	K	Ca	Sc	Ti	V	Cr	Mn
Forrested	3.391136	0.561307	107.3318	6138.267	36029.45	2214.037	42990.76	480287.2	0.749484	6.579818	2.170767	0.430341	5259.641
Agricultural	11.03945	0.011363	110.0894	6721.852	20331.12	8.491738	93563.65	655050.2	0.77766	1.915602	1.395706	0	482.3251
Park Land / SSSI	12.80256	0.078371	93.55025	7221.361	26166.4	213.68	86800.25	505558.8	0.685275	2.646863	1.487601	0	1215.965
Quarry	1.193583	0.069542	51.57792	4837.767	15532.92	261.5708	23891.67	338225	0.42825	1.421333	1.501255	0	2005.205
Scrubland	10.51875	0.015558	116.0125	11118.75	37480	44.46917	122400	650125	1.04725	4.419167	3.364	0	1168.479

Fe	Co	Ni	Cu	Zn	Ga	As	Se	Rb	Sr	Y	Zr
364.9102	11.37939	25.69263	24.55001	568.9263	0.136555	4.07983	0.794182	79.68256	1138.748	1.76529	1.490351
19.61488	1.946716	6.870226	21.38828	16.39978	0.009855	4.505555	0.874061	62.68692	1952.995	0.131579	0.206939
140.3588	4.814477	9.353025	28.25326	513.7811	0.036695	7.613775	0.723399	49.3375	1785.809	0.361388	1.018919
373.715	2.917027	7.014167	6.248547	382.4583	0.029047	12.30379	1.044792	37.95	692.9208	0.235047	1.509493
91.94708	7.040917	17.415	25.53583	134.55	0.00745	6.462917	0.62825	107.1708	1817.458	0.5163	0.372292

Mo	Ag	Cd	Sn	Sb	Te	Cs	Ba	La	Ce	Pr	Nd
0.722064	0.828458	3.620648	0	0.161684	0	1.350683	2012.072	2.764526	3.630449	0.408124	1.588145
2.028159	1.109185	0.584763	0	0.478618	0.014457	1.431715	3098.609	0.179373	0.153694	0.021614	0.108282
4.173938	1.077839	0.9583	0	0.288075	0	1.556726	1750.572	0.408209	0.784946	0.099	0.427152
0.637042	0.979113	0.75411	0	0.10535	0.173125	0.591238	719.6867	0.383465	0.766572	0.085583	0.352275
2.567708	0.472083	1.216625	0	0.654805	0	2.437833	1783.042	0.482497	0.863542	0.087648	0.416867

Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	Hg
0.287252	0.579915	0.360983	0.043056	0.236675	0.045961	0.119486	0.013786	0.075295	0.011489	1.134659
0.027714	0.755991	0.045554	0.00158	0.011822	0.005018	0.008398	0.001392	0.002747	0.001601	0.697012
0.09997	0.429452	0.113665	0.012399	0.069899	0.015025	0.035833	0.005849	0.025655	0.004865	0.7035
0.081867	0.188227	0.072667	0.009655	0.044352	0.011657	0.02075	0.003257	0.013558	0.002798	0
0.092255	0.481147	0.114135	0.011305	0.071198	0.015665	0.044837	0.005885	0.023373	0.00882	0

Tl	Pb	Bi	Th	U
0.957641	57.0382	0.030541	0.349572	0.176945
0.716605	0.627025	0.00698	0.005615	0.207961
0.6691	4.198146	0	0.100016	0.14105
0.721333	3.122717	0	0.177627	0.176102
1.4155	0.453958	0.006325	0.027483	0.466935

Appendix 3

Plant latin and common names

Please not these are listed alphabetically according to the latin name

Latin Name	**Common Name**
Ammophila arenaria	Marram Grass
Anacamptis pyramidalis	Pryamidal Orchid
Aneura pinguis	Dumort
Artemsia campestris	Mugwort
Briza media	Quacking Grass
Bromus hordeacus	Soft-Brome
Calluna	Heather
Calluna Vulgaris	Common Heather or Ling
Campylium stellatum	Star Campylium Moss
Carex arenaria	Sand Sedge
Carex dioica	Dioecious Sedge
Carex flacia	Glaucous Sedge
Carex diandra-Calliergon giganteum	Lesser Panicled Sedge-Giant Moss
Carex lepidocarpa	Slender Sedge
Carex rostrata-Calliergon cuspidatum	Bottle Sedge-Moss
Cicuta virosa	Cowbane or Northern Water Hemlock
Dicranum scoparium	Broom Moss
Drosera angelica	English Sundew or Great Sundew
Drosera rotundiflora	Common Sundew or Round-Leaved Sundew
Epipactis Palustre	Marsh Helleborine

Erica	Heath / Heather (Winter Heather to distinguish from *Calluna*)
Erica tetralix	Cross-Leaved Heath
Eriophorum angustifolium	Common Cottongrass
Filipendula ulmaria-Angelica sylvestris	Meadow Sweet-Angelica Mire
Festuca ovina	Sheep's Fescue
Festuca rubra	Red-Fescue
Frankenia leavis limonium bellidifolium	Sea Heath-Matted Sea Lavender
Glaucium flavum	Yellow Horned Poppy
Hippocrepis comusa	Horseshoe Vetch
Honkenya peploides	Sea Sandwort
Juncus subnodulosus-Cirsium palustre	Fen Meadow-Blunt Flowered Rush-Marsh Thistle
Limonium-Armeria	Sea Lavender
Linum catharticum	Farry Flax
Menyanthes trifoliata	Bogbean
Molina caerulea-Cirsium dissectum	Fen Meadow Purple Moorgrass-Meadow Thistle
Narthecium ossifragrum-Sphagnum papillosum	Bog Asphodel-Papillose Bog Moss
Ophioglossum vulgatum	Adders Tongue
Peucedanum palustre	Milk Parsley
Phragmites australis-Peucedanum palustre	Tall herb fen community
Pinguicula vulgaris	Common Butterwort
Plantago lanceolata	Ribwart Plantain
Potamogeton polygonifolius	Bog Pond Weed
Primula veris	Cowslip
Puccinellia	Salt-Marsh Grass
Ranunculus lingua	Greater Spearwort

Rhynchospora alta	White Beak Sedge
Rumex acetosella	Grassland community
Sedum acre	Biting Stonecrop
Scabiosa columbaria	Small Scabious / Butterfly Blue
Schoenus nigricans-Juncus subnodulosus	Black Bog-Rush Blunt Flowered Rush (Fen Meadow)
Schoenus nigricans-Narthecius ossifragum	Black Bog-Rush – Bog Asphodel
Silium silaus	Pepper-Saxifrage
Sium latifolium	Greater Water Parsnip
Spartina	Common Gord-Grass
Sphagnum auriculatum	Bog Pool Community
Sphagnum compactum	Bogmoss
Sphagnum cuspidatum	Bog Pool Community
Sphagnum recurvum	Toothed Sphagnum
Sphagnum subnitens	Peat Moss
Suaeda-Salicornia	Glasswort-Seablite
Suaeda-Vera	Shrubby-Seablite
Trisetum flavescens	Yellow Oat-Grass
Ulex	Gorse
Ulex Gallii	Western Gorse

Appendix 4

Most Common Plant Species Recorded in Norfolk

The most common plant species which are found throughout Norfolk and how often they are recorded are shown here.

GOODSANDS	Annual Meadow-Grass Count	159
	False Oat-Grass Count	159
	Perennial Rye-Grass Count	159
	Yorkshire-Fog Count	159
	Cock's-Foot Count	158
	Barren Brome Count	157
	Common Couch Count	153
	Creeping Bent Count	153
	Soft-Brome Count	147
	Timothy Count	139
FENLANDS	Cock's-Foot Count	156
	Perennial Rye-Grass Count	156
	Annual Meadow-Grass Count	155
	Barren Brome Count	152
	False Oat-Grass Count	152
	Common Couch Count	149
	Wall Barley Count	146
	Common Reed Count	143
	Creeping Bent Count	141
	Yorkshire-Fog Count	141
CROMER RIDGE	Bluebell Count	38
	Annual Meadow-Grass Count	28
	Barren Brome Count	28
	Bracken Count	28
	Cock's-Foot Count	28
	False Oat-Grass Count	28
	Perennial Rye-Grass Count	28
	Yorkshire-Fog Count	28
	Scots Pine Count	27
	Sweet Vernal Grass Count	27
THE CLAYS	Cock's-Foot Count	396
	Annual Meadow-Grass Count	395
	False Oat-Grass Count	390

	Yorkshire-Fog Count	390
	Perennial Rye-Grass Count	387
	Barren Brome Count	386
	Lords-And-Ladies Count	382
	Field Horsetail Count	380
	Common Couch Count	379
	Wild Oat Count	356
BROADS	Annual Meadow-Grass Count	84
	Cock's-Foot Count	84
	Perennial Rye-Grass Count	84
	False Oat-Grass Count	83
	Yorkshire-Fog Count	82
	Common Reed Count	78
	Creeping Bent Count	78
	Wall Barley Count	78
	Barren Brome Count	76
	Common Couch Count	74
BRECKLANDS	False Oat-Grass Count	224
	Cock's-Foot Count	219
	Annual Meadow-Grass Count	216
	Perennial Rye-Grass Count	214
	Yorkshire-Fog Count	214
	Barren Brome Count	209
	Common Couch Count	202
	Timothy Count	197
	Creeping Bent Count	195
	Soft-Brome Count	191
NORTH EAST NORFOLK	Soft Rush Count	153
	Cock's-Foot Count	148
	False Oat-Grass Count	148
	Annual Meadow-Grass Count	147
	Perennial Rye-Grass Count	147
	Yorkshire-Fog Count	147
	Bluebell Count	142
	Common Couch Count	142
	Barren Brome Count	140
	Field Horsetail Count	140
GREENSANDS	Cock's-Foot Count	118
	Annual Meadow-Grass Count	117
	Perennial Rye-Grass Count	117

	Yorkshire-Fog Count	117
	False Oat-Grass Count	115
	Barren Brome Count	114
	Common Couch Count	112
	Creeping Bent Count	111
	Soft-Brome Count	110
	Timothy Count	105
WENSUM SANDS	Bluebell Count	69
	Annual Meadow-Grass Count	59
	Barren Brome Count	59
	Bracken Count	59
	Cock's-Foot Count	59
	False Oat-Grass Count	59
	Perennial Rye-Grass Count	59
	Yorkshire-Fog Count	59
	Scots Pine Count	58
	Wall Barley Count	58

Appendix 5

Pollen Data

The following data is the raw data files received from the British Geological Survey showing the pollen and spore counts for each of the different samples provided.

				POLLEN																			
NO.	**MPA NO.**	**UEA NO.**	**Locality**	***Abies* (A)**	***Alnus* (A)**	***Aster* (H)**	***Betula* (A)**	**Caryophyllaceae (H)**	***Circaea* (H)**	***Corylus* (A)**	***Epilobuim* (H)**	***Erica* (S)**	**Gramineae (H)**	***Pinus* (A)**	***Plantago* (H)**	**pollen - indeterminate**	***Polygala* (H)**	**reticulate pollen - small, indeterminate (H)**	***Succisia* (H)**	***Taraxacum officinale* (H)**	***Tilia* (A)**	**tricolpate pollen - indeterminate**	**TOTAL POLLEN GRAINS COUNTED**
1	59515	240708-01-03	Ashwicken Ground	…	2	…	4	…	…	5	…	…	14	31	…	1	…	X	…	1	…	…	**58**
2	59516	240708-01-02	Ashwicken Ground	…	5	…	…	…	…	1	…	…	1	16	…	…	…	…	…	1	…	…	**24**
3	59517	240708-01-04	Ashwicken Ground	…	2	…	1	…	…	1	…	…	11	28	…	7	…	6	…	3	…	…	**59**
4	59518	240708-02-01	Bawsey Country Park – forest	…	8	…	19	?1	…	17	…	3	…	82	?1	3	…	…	…	…	…	…	**134**
5	59519	240708-02-02	Bawsey Country Park – forest	…	11	…	6	?3	…	52	…	3	1	39	…	1	…	…	…	…	…	…	**116**

6	59520	240708-02-03	Bawsey Country Park – forest	…	3	…	5	…	…	19	…	2	…	49	…	4	…	…	…	…	1	…	**87**
7	59521	240708-03-02	Bawsey Country Park – forest	…	1	…	…	…	…	7	…	2	3	74	…	1	…	…	…	…	…	…	**88**
8	59522	240708-03-03	Bawsey Country Park – forest	…	…	…	…	…	…	…	…	5	…	57	…	1	…	…	…	…	1	…	**64**
9	59523	240708-03-04	Bawsey Country Park – forest	…	…	…	…	…	…	…	…	5	…	31	…	…	…	2	…	…	…	…	**38**
10	59524	240708-04-01	Bawsey Country Park - near lake	X	11	…	1	4 + ?1	…	19	…	76	?X	187	…	X	…	1	…	…	1	…	**301**
11	59525	240708-04-02	Bawsey Country Park - near lake	1	4	…	11	…	1	16	…	79	2	38	…	1 + X	…	…	…	…	…	…	**153**
12	59526	240708-04-03	Bawsey Country Park - near lake	…	29	1	8	X	…	179	…	64	14	63	X	X	…	1	X	…	2	…	**361**
13	59527	240708-05-01	BGS GBase site 440 (crops)	…	…	…	…	…	…	…	…	9	…	17	…	…	…	16	…	…	…	…	**42**
14	59528	240708-05-02	BGS GBase site 440 (crops)	…	1	…	…	…	…	…	…	…	…	…	…	2	…	18	…	…	…	…	**21**
15	59529	240708-05-04	BGS GBase site 440 (crops)	…	2	1	…	…	…	3	…	…	3	18	…	…	…	9	…	1	…	…	**37**

16	59530	240708-06-01	Roydon National Park	…	…	…	…	…	…	…	…	4	3	21	…	…	…	…	…	1	3	…	**32**
17	59531	240708-06-02	Roydon National Park	…	1	…	…	…	…	1	…	…	…	7	…	…	…	…	…	…	1	…	**10**
18	59532	240708-06-04	Roydon National Park	…	…	…	5	?1	…	16	1	14	4	34	…	…	…	…	…	…	2	…	**77**
19	59533	240708-07-01	Roydon crops	…	?2	…	…	…	…	1	…	…	7	18	…	2	…	2	…	4	…	…	**36**
20	59534	240708-07-02	Roydon crops	…	…	…	…	1 + ?1	…	1	..	2	…	26	…	…	…	…	…	2	…	…	**33**
21	59535	240708-07-03	Roydon crops	…	…	…	…	?1	…	…	…	1	5	14	…	…	…	…	…	…	…	…	**21**
22	59536	160708-09-02	Titchwell	…	…	…	…	…	…	1	…	…	6	14	…	…	…	…	…	…	…	…	**21**
23	59537	160708-09-03	Titchwell	…	1	…	…	…	…	…	…	…	10	9	…	1	…	…	…	…	…	…	**21**
24	59538	160708-09-04	Titchwell	…	…	…	…	…	…	4	…	…	17	14	…	1	…	…	…	…	…	…	**36**
25	59539	160708-04-01	Wighton crops	…	…	…	…	…	…	…	…	…	4	23	…	…	…	…	…	…	….	…	**27**
26	59540	160708-04-03	Wighton crops	…	…	…	…	…	…	…	…	…	3	16	…	…	?1	…	…	1	…	…	**21**
27	59541	160708-04-04	Wighton crops	…	2	…	…	…	…	…	…	…	24	9	…	7	…	…	…	…	…	…	**42**
28	59542	160708-12-01	Holt (many ferns)	…	5	…	4	8	…	6	…	1	24	27	…	7	…	…	…	…	2	14	**98**
29	59543	160708-12-02	Holt (many ferns)	…	…	…	…	…	…	2	…	…	4	8	…	…	…	…	…	…	…	9	**23**
30	59544	160708-12-03	Holt (many ferns)	…	6	…	…	2	…	16	…	…	…	8	…	5	…	…	…	?2	1	6	**46**

31	59545	160708-05-02	Knapton Hedges	…	…	…	…	…	…	3	…	…	3	11	…	…	…	…	…	…	…	**17**
32	59546	160708-05-03	Knapton Hedges	…	…	…	…	1	…	3	…	…	…	19	…	2	…	…	…	5	…	**30**
33	59547	160708-05-04	Knapton Hedges	…	…	…	…	1	…	…	…	1	…	18	…	1	…	…	…	5	…	**26**
NO.	**MPA NO.**	**UEA NO.**	**Locality**	***Abies* (A)**	***Alnus* (A)**	***Aster* (H)**	***Betula* (A)**	**Caryophyllaceae (H)**	***Circaea* (H)**	***Corylus* (A)**	***Epilobuim* (H)**	***Erica* (S)**	**Gramineae (H)**	***Pinus* (A)**	***Plantago* (H)**	**pollen - indeterminate**	***Polygala* (H)**	**reticulate pollen - small, indeterminate (H)**	***Succisia* (H)**	***Taraxacum officinale* (H)**	***Tilia* (A)**	**tricolpate pollen - indeterminate**	**TOTAL POLLEN GRAINS COUNTED**
				POLLEN																		

				SPORES															
NO.	**MPA NO.**	**UEA NO.**	**Locality**	***Dryopteris***	**Filicales**	***Lycopodium***	***Polypodium vulgare***	***Pteridium***	***Sphagnum***	**spores - indeterminate**	**TOTAL SPORES COUNTED**		**?Quaternary dino. cysts**		**Carb. spores - indeterminate**	***Densosporites* spp.**		**Mesozoic dinoflagellate cysts**
1	59515	240708-01-03	Ashwicken Ground	…	1 + ?1	…	…	…	2	…	**4**		…		…	…		…
2	59516	240708-01-02	Ashwicken Ground	…	…	1	…	?1	…	…	**2**		…		…	…		…
3	59517	240708-01-04	Ashwicken Ground	…	?1	…	…	?4	…	…	**5**		…		..	…		…
4	59518	240708-02-01	Bawsey Country Park – forest	…	…	…	1	…	…	1	**2**		…		…	…		…
5	59519	240708-02-02	Bawsey Country Park – forest	…	…	…	…	…	…	…	**0**		…		…	…		…
6	59520	240708-02-03	Bawsey Country Park – forest	…	…	…	?2	…	…	..	**2**		…		…	…		…

7	59521	240708-03-02	Bawsey Country Park – forest	…	1	1	…	…	1	…	**3**		…		…	…		…
8	59522	240708-03-03	Bawsey Country Park – forest	…	…	…	…	…	…	…	**0**		…		…	…		…
9	59523	240708-03-04	Bawsey Country Park – forest	…	…	1	X	…	2	…	**3**		…		…	…		…
10	59524	240708-04-01	Bawsey Country Park - near lake	2	2	X	…	…	29	10	**43**		?X		…	…		…
11	59525	240708-04-02	Bawsey Country Park - near lake	2	…	…	…	…	5	3	**10**		…		…	…		…
12	59526	240708-04-03	Bawsey Country Park - near lake	…	…	16	…	…	31	…	**47**		…		?X	…		…
13	59527	240708-05-01	BGS GBase site 440 (crops)	…	6	…	…	…	…	3	**9**		…		…	…		…
14	59528	240708-05-02	BGS GBase site 440 (crops)	…	…	4 + ?2	…	…	…	…	**6**		…		…	…		…
15	59529	240708-05-04	BGS GBase site 440 (crops)	…	4	…	…	…	2	…	**6**		…		…	…		…
16	59530	240708-06-01	Roydon National Park	…	…	1	…	…	…	4	**5**		…		…	…		…
17	59531	240708-06-02	Roydon National Park	…	…	…	…	…	…	1	**1**		…		…	…		…

18	59532	240708-06-04	Roydon National Park	…	…	…	…	…	…	4	**4**		…		…	…		…
19	59533	240708-07-01	Roydon crops	…	1	…	…	…	…	4	**5**		2		…	…		…
20	59534	240708-07-02	Roydon crops	…	2	1	…	…	2	1	**6**		…		…	…		1
21	59535	240708-07-03	Roydon crops	…	…	…	…	…	1	…	**1**		…		…	…		…
22	59536	160708-09-02	Titchwell	…	…	…	…	…	…	…	**0**		…		…	…		…
23	59537	160708-09-03	Titchwell	…	…	…	…	…	…	…	**0**		…		…	…		…
24	59538	160708-09-04	Titchwell	1	1	…	…	…	…	…	**2**		…		…	…		…
25	59539	160708-04-01	Wighton crops	…	…	…	…	…	1	1	**2**		?1		…	…		…
26	59540	160708-04-03	Wighton crops	…	…	1	…	…	…	1	**2**		…		…	1		…
27	59541	160708-04-04	Wighton crops	…	1	…	…	…	…	…	**1**		…		…	1		…
28	59542	160708-12-01	Holt (many ferns)	…	…	…	…	…	…	…	**0**		…		…	…		…
29	59543	160708-12-02	Holt (many ferns)	…	…	…	…	…	…	…	**0**		…		…	…		…
30	59544	160708-12-03	Holt (many ferns)	…	…	…	2	…	…	…	**2**		…		…	…		..
31	59545	160708-05-02	Knapton Hedges	…	…	4	…	…	…	1	**5**		…		…	1		…
32	59546	160708-05-03	Knapton Hedges	…	…	4	…	…	…	…	**4**		…		…	…		…

NO.	MPA NO.	UEA NO.	Locality	*Dryopteris*	Filicales	*Lycopodium*	*Polypodium vulgare*	*Pteridium*	*Sphagnum*	spores - indeterminate	TOTAL SPORES COUNTED		?Quaternary dino. cysts		Carb. spores - indeterminate	*Densosporites* spp.		Mesozoic dinoflagellate cysts
				SPORES														
33	59547	160708-05-04	Knapton Hedges	…	…	3	…	…	…	2	**5**		…		…	…		…