A Study on Nucleic Acid Degradation in Drying and Dried Bloodstains as a Means to Determine the Time since Deposition

Wesley Morta

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (Forensic Science) in the School of Chemical Sciences

University of Auckland

2012

Abstract

Forensic evidence often plays a crucial role in linking a suspect or victim to a crime. DNA evidence can provide the gender of the victim or offender and can also indicate the strength of the DNA evidence in favour of a probable source. Forensic investigators will often identify and secure body fluids such as blood, semen and saliva stains as a source of evidence. However, it is not always clear if the evidence relates to the particular incident under investigation or another unrelated event. Therefore, determining the time since deposition of the sample can clarify the circumstances of the investigation.

Bloodstains are the most plentiful forms of evidence at a crime scene, and presently there is no method available that can accurately determine when they were deposited. Previous studies have shown that it is possible to correlate ribonucleic acid degradation to bloodstain age, but this correlation was only apparent in stains that were more than six months old; recent bloodstains could not be distinguished from six month old bloodstains.

The aim of this study was to determine whether it is possible to use the degradation profile of nucleic acids to provide a better estimate to when a bloodstain was deposited. The main hypothesis was that as a bloodstain ages, nucleic acids are degraded and become fragmented. By examining the degradation of beta hemoglobin (HBB) DNA, Glucose-6-phosphate dehydrogenase (G6PD) mRNA and 18S rRNA in aged bloodstains by capillary electrophoresis, these fragments could be used to determine bloodstain age.

In order to test this hypothesis, a real-time quantitative PCR method to standardise the input amount of cDNA analysed was developed as previous studies have not taken sample variation into account. The pattern of degradation of these nucleic acids was then investigated as bloodstains transitioned from wet to dry when exposed to different storage conditions to determine the directionality of degradation, whether it was from the 5' end, 3' end or random. This was achieved by developing a multiplex PCR assay with primers and probes specific to different regions of the HBB DNA, G6PD mRNA and 18S rRNA sequences.

Regression analysis was performed on the peak heights of these marker sequences from aged bloodstains deposited on cotton cloth to determine which fragments of which markers were statistically relevant to age prediction. The result was an age prediction equation which used the peak heights of the G6PD 5' medium, 3' long and 3' medium fragments and the central long fragment of 18S rRNA.

The equation was evaluated using blind samples and on bloodstains found on common forensic type substrates such as carpet and articles of clothing stored under different conditions.

The study found that the prediction equation was able to discern age differences between recent and old stains on the substrates tested. However, the equation was more accurate at predicting the age of bloodstains found on cotton cloth stored at room temperature with exposure to sun. The method can be used to age bloodstains that are days, weeks and even months old.

Acknowledgments

I would like to thank my supervisors, SallyAnn Harbison, Douglas Elliot and Rachel Fleming for their guidance throughout the duration of this project. Also, special thanks to our volunteers for various body fluids, without whom this work would not have been possible. My thanks also go to the Institute of Environmental Science and Research capability fund that funded this study. I would also like to thank my graduate colleagues in the Forensic program for their help and support. Finally, a special thanks to my family and partner Kerrin, for their never wavering support and encouragement.

Glossary

AAR	aspartic acid racemisation
AFM	atomic force microscopy
AVP	arginine-vasopressin
B2M	beta-2 microglobulin
BLAST	basic local alignment search tool
bp	base pair
CD3G	cluster of differentiation gamma-3 molecule
cDNA	complementary deoxyribonucleic acid
Ct	threshold cycle
CyPA	cyclophilin A
DNA	deoxyribonucleic acid
ELISA	enzyme-linked immunosorbant assay
EPO	erythropoietin
EPR	electron paramagnetic resonance
FAM	Fluorescein amidite
G6PD	glucose-6-phosphate dehydrogenase
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GC	guanine-cytosine content
GlycoA	glycophorin A
GOT	glutamic oxaloacetic transaminase
GPT	glutamic pyruvic transaminase
Hb	de-oxyhemoglobin
Hb02	oxyhemoglobin
HBB	beta hemoglobin
HIF1A	hypoxia-inducible factor 1 alpha
HK2	transcription elongation factor
HPLC	high performance liquid chromatography
HTN3	histatin 3
IPC	internal positive control
kPA	kilopascal
LDH	lactate dehydrogenase
Met-HB	methemoglobin
MMP11	matrix metalloproteinase 11
mRNA	messenger ribonucleic acid
mtDNA	mitochondrial DNA
PCR	polymerase chain reaction
PMI	postmortem interval
PRM2	protamine 2
qPCR	quantitative real-time polymerase chain reaction
RFU	relative fluorescent units
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
RT-PCR	real time polymerase chain reaction
SD	standard deviation
STATH	statherin
STR	short tandem repeat
TEF	transcription elongation factor
TBP	TATA-binding protein
TGM4	transglutaminase 4
Tm	melting temperature
tRNA	transfer ribonucleic acid
UBC	ubiquitin C
UCE	ubiquitin conjugating enzyme
UV	ultraviolet
UV-VIS	ultraviolet-visible
VEGF	vascular endothelial growth factor

Table of Contents

Chapter 3 – Development of a Method to Standardise the Amount of cDNA Added to a PCR Reaction

Chapter 4 – Design and Development of Novel Multiplex PCR Assays to Assess the Degradation of mRNA, rRNA and DNA in Bloodstains

Chapter 5 – A Method to Determine the Time since Deposition of a Bloodstain

Chapter 6 – A Method to Determine the Age of a Bloodstain – a Study on Forensic Type Samples

Chapter 7 – Final Discussion

Appendices

List of Tables

List of Figures

Chapter 1

1.1 Introduction

With current technologies in forensic science, a significant amount of genetic information can be obtained from an individual biological stain found at a crime scene. Not only is it possible to obtain a DNA profile of the stain's donor, but new methods are also being developed to determine body fluid origin using RNA based methods, and also to predict physical characteristics including eye colour, hair colour, skin pigmentation, height, weight and relative age (Bauer & Patzelt, 2008; Fleming & Harbison, 2010; Juusola & Ballantyne, 2003; Ruiz, Phillips, Gomez-Tato *et al.*, 2012; Saeed, Berlin, & Cruz, 2012). The probative information obtained from these samples can aid investigators in cases where there are no known suspects. The ability to determine when a biological stain was deposited would aid in the reconstruction of a timeline of events associated to a crime that may be limited to eyewitnesses.

During a crime, evidence is deposited through interactions between the perpetrator and the victim. It is often difficult to determine if the DNA evidence found at the scene was left during the crime or deposited during a previous interaction. Through short tandem repeat (STR) analysis, the person from whom the sample originated can be determined; however, to link or exclude the source from the crime requires a witness or other evidence. When proof of the time of deposition is lacking, the reliability of the evidence itself may be called into question in a court case.

If the victim and suspect had a history of previous personal contacts, it is not uncommon for the suspect’s DNA to end up on or around the victim or vice versa. It may be argued that the DNA sample was left prior to the commission of the crime due to social relationships between suspect and victim. Being able to distinguish when a biological sample was deposited would help to resolve this dilemma.

1.2 The problem under investigation

Bloodstains are some of the most valuable forms of evidence encountered at crime scenes due to their frequency and variety of uses. Blood splatters can be used in pattern analysis to help reconstruct a series of events leading up to the deposition, and can aid in verifying a suspect's identity through DNA-profiling if a reference sample is available (Aronson, 2007). However, to date it is not possible to determine at what point in time the evidence was deposited. The events leading up to deposition could have taken place days, months or even years before or after the crime was committed.

Temporal correlation between the evidence and the commission of a crime is especially important in situations involving suspects and victims with close personal ties. It is not uncommon to find biological material from a suspect in the home, vehicle, or other relevant location associated with the victim. For example, if the victim's blood was found in a vehicle owned by the suspect, the accused could argue that the sample had previously been deposited in an unrelated event. At present there is no scientific way to prove otherwise. A method that can accurately determine when the blood was deposited is essential to help investigators verify witness statements, assess alibis and rule out particular suspects.

If investigators are able to identify the age of a bloodstain, they could also use this information to estimate the likely whereabouts of an individual or suspect who may have been injured, for example during a police pursuit. This would give investigators an idea of how far the suspect may have travelled given the time since deposition.

1.3 Composition of bloodstains

Various components of a bloodstain have been assessed for age estimation. Bloodstains originate from droplets of whole blood that have dried out. Whole blood contains blood cells, proteins and amino acids suspended in plasma.

There are three major types of blood cells and cell fragments: red blood cells, white blood cells and platelets (Whittemore, 2004).

Red blood cells are the most numerous; they contain no DNA and largely consist of the oxygen transport protein hemoglobin. Hemoglobin molecules are mainly present in two forms *in vivo*: de-oxyhemoglobin (Hb) and oxy-hemoglobin (HbO2). When blood is expelled from the body, the ambient oxygen in the atmosphere binds to and saturates the hemoglobin. As the availability of cytochrome b5 is diminished *ex vivo*, it means that conversion from Hb02 to methemoglobin (met-Hb) to Hb can no longer occur. Consequently, chemical and physical changes in red blood cells have been used as an attempt to determine bloodstain age.

White blood cells are less numerous than red blood cells but they constitute the main source of nuclear material (DNA). Bauer *et al.* and Anderson *et al.* have attempted to use RNA for age estimation (Anderson, Hobbs, & Bishop, 2005; Bauer, Gramlich, Polzin *et al.*, 2003) but have had limited success due to the poor resolution times between samples.

Platelets are involved in the clotting and coagulation process that occurs within seconds of injury. There are no reports in the literature where platelets have been used for age estimation.

1.4 Techniques developed to determine the age of a bloodstain

The earliest account in the literature is by Louis Tomellini (Tomellini, 1907) where he developed a chart that illustrated the change in colour of a fresh bloodstain up to a period of one year. The limitation of this method was that it was based on visual observations and the colour was defined by the observer. Leers later defined the temporal change in bloodstains as a transformation in the hemoglobin spectra after comparing the spectra of dried blood to fresh blood (Leers, 1910). This observation was from a physiological point of view and only as the bloodstain dried. Schwarzacher performed water solubility tests on bloodstains in an attempt to link solubility to age (Schwarzacher,

1930). He determined that the solubility of bloodstains in distilled water decreases as a function of age. Schwarzacher was also the first person to conclude that light has an effect on the character and rate of aging of bloodstains. Schwarzacher noted that the decrease of solubility in distilled water and the time of exposure to light could be represented through a curve that resembled an exponential curve. Through the solubility method, Schwarzacher was able to say that if a bloodstain is placed in direct sunlight, the maximum age estimate is 20 hours and the maximum age estimate is several weeks if kept in complete darkness (for bloodstains deposited at the same time).

In 1937, Schwarz used a guaiacum-based assay to determine the catalase and peroxidase activity of hemoglobin in bloodstains (Schwarz, 1937). Guaiacum, in the presence of hemoglobin, would catalyse a guaiacum bluing reaction with hydrogen peroxide. This assay showed that the intensity of the reaction's colour varied with the age of the bloodstain but could only differentiate between very recent (less than 24 hours) and old bloodstains (weeks). In the 1950s, Rauschke discovered that large amounts of met-Hb could only be found in bloodstains 19-24 hours after deposition, a result that was not observed in frozen samples even after a period of months (Rauschke, 1951).

Patterson later used photo-spectrometry and recorded the reflectance spectra of bloodstains. He observed that temperature, light and humidity have an impact on the rate of colour change (Patterson, 1960). He later derived a time dependent "α ratio" that is independent of the amount of blood present which improved age estimation (Kind, Patterson, & Owen, 1972). In 1962, Fiori proposed a method that correlated the age of a bloodstain with the progressive diffusion of chloride ions around the stain by fixing with AgCl (Fiori, 1962). Upon reduction, a black border was observed around stains that were more than two months old. There were incremental increases to the size of the border up to a period of nine months. This method does not take into account the size of the stain and does not differentiate between human and other blood.

In 1974, Nuoreteva presented a case report where the age of a bloodstain on a discarded shirt was determined by entomological analysis whereby the number of days the fly larvae found on the shirt took to hatch was counted (Nuoreteva, 1974).

1.5 Techniques developed for red blood cells

1.5.1 High performance liquid chromatography (HPLC)

HPLC is a separation technique that uses retention times to identify and quantify the individual components of a mixture. The retention time is the time taken for a particular compound to travel through the separation column to the detector. This time is measured from the time at which the sample is injected to the point at which the display shows a maximum peak height for that compound (Clark, 2007). Different compounds have different retention times. In dried bloodstains, the products of heme degradation (Kapitulnik, 2004) can be quantified by HPLC and used as markers for age estimation (Andrasko, 1997; Inoue, Fukutaro, Iwasa *et al.*, 1991; Inoue, Fukutaro, Iwasa *et al.*, 1992). The various components of a mixture, such as the proteins in blood, have different retention times and it is the integrated peak area of each component that gives an indication of the amount present. Inoue *et al.* observed a linear decrease in the ratio of the peak areas of the hemoglobin α-globin chain to that of the heme protein as the age of the bloodstain increased. They also found that neonatal bloodstains could be differentiated from adult bloodstains by the presence of γ-globin chains in neonates up to 32 weeks old (Inoue *et al.*, 1991). An improved age indicator was found following the discovery of the presence of an unknown protein ‘X’, which is only detected in aged bloodstains (Inoue *et al.*, 1992). The peak area of protein ‘X’ increases with the age of the bloodstain (Andrasko, 1997; Inoue *et al.*, 1992) and it is not affected by temperature (between 0 – 37°C). The ratio of protein 'X' to heme is 0 for fresh blood and increases to 0.3 for bloodstains stored for 52 weeks in the dark at 37°C.

HPLC requires that the bloodstain is first removed from the substrate using a wet swab which is then dissolved in distilled water. As a result, the method suffers from variations in the concentration of hemoglobin derivatives between samples that equate to large standard deviations that make age prediction estimates difficult.

1.5.2 Reflectance spectroscopy

The colour of bloodstains visibly change with age. Optical spectroscopy methods that are able to quantify these changes could be used to age bloodstains. Numerous attempts have been made based on the visible reflectance spectrum (450-700 nm) of the bloodstain to its age (Bremmer, Nadort, Van Leeuwen *et al.*, 2011; Hanson & Ballantyne, 2010; Kind *et al.*, 1972; Patterson, 1960). The standard approach appears to be measuring the reflectance at two different wavelengths and calculating the ratio, thus estimating age.

This is the least invasive technique, as the method requires no sample preparation and there is no physical contact with the sample. However, for the method to work, the bloodstain must be on a white background.

1.5.3 Oxygen electrodes

Matsuoka *et al.* demonstrated that bloodstain age could be determined from the ratio of Hb02 to Hb using an oxygen electrode immersed in water in which the oxygen had been depleted. The total hemoglobin was then determined by conventional colorimetry (cyanomethemoglobin method) where blood is mixed with solution containing potassium ferricyanide and potassium cyanide. Potassium ferricyanide oxidizes iron to form methemoglobin and potassium cyanide then combines with methemoglobin to form cyanmethemoglobin which is a stable colour pigment read photometrically at a wave length of 540nm (Matsuoka, Taguchi, & Okuda, 1995). By monitoring bloodstains for up to ten days at different temperatures, they showed that the decay of HbO2 at room temperature is rapid at first but decreases after a few hours;

bloodstains stored at higher temperatures resulted in a faster decay of HbO2. No further decay was observed after 24 hours.

The method dictates that the sample is first dissolved in a saline solution, which not only means that the sample is consumed in the process but that the method is also affected by different volumes of blood as the total amount of hemoglobin would differ.

1.5.4 Electron paramagnetic resonance (EPR)

Electron paramagnetic resonance is used to measure the spin state change of the iron ion in hemoglobin as it denatures (Fujita, Tsuchiya, Abe *et al.*, 2005; Miki, Kai, & Ikeya, 1987; Sakurai & Tsuchiya, 1989). Bloodstains yield four significant EPR signals due to ferric high-spin, ferric non-heme, ferric low-spin and free radical species (Miki *et al.*, 1987). *Fujita et al.* found that by plotting double logarithms of the EPR intensity ratio of ferric low-spin to ferric non-heme versus days since deposition, a linear relationship of up to 432 days (+/- 25%) could be established. This relationship however, is affected by environmental factors such as temperature and light exposure. Bloodstains or samples subjected to this method must first be cooled to -196°C, a practice not commonly suited to routine forensic analysis and not all substrates can withstand such temperatures.

1.5.5 Atomic force microscopy (AFM)

Atomic force microscopy is a microscopic technique that provides a high resolution, three-dimensional profile of a surface using a sensitive probe attached to a cantilever. AFM, in the case of red blood cells, can be used to determine the elasticity of a surface by determining the cantilever's resonance frequency. Strasser *et al.* used AFM to investigate the elasticity changes of a red blood cell as a function of time. They found that the elasticity of a red blood cell increases from 40 kPA to 300 kPA 1.5 h after bleeding, 600 kPA within 30 h and 2.5 GPa after 30 days (Strasser, Zink, Kada *et al.*, 2007).

As with reflectance spectroscopy, AFM does not require any sample preparation but the substrate on which the bloodstain has been deposited must be very flat. The method is therefore only applicable to bloodstains deposited on glass and tiles.

1.6 Techniques developed for blood plasma

1.6.1 Immunoelectrophoresis

Rajamannar found that serum blood proteins present in bloodstains gradually deteriorate over time. The serum protein profile of a bloodstain was examined by studying the decomposition of various α, β, and γ-globulins, as well as albumin ranging from recently deposited up to a year old (Rajamannar, 1977). In this study, the presence of proteins was determined by immunoelectrophoresis. It showed that a characteristic pattern of the disappearance of the various globulin proteins and albumin was evident. All globulins and albumin were present in fresh bloodstains but after fifteen days, the α-globulins and albumin had decomposed. The β and γ-globulins gradually decomposed up to a year where no globulins were found in the bloodstain. Other work by Sensabaugh *et al*. that also utilised immunoelectrophoresis, tested for the presence of eleven specific globular proteins in a sample of dried blood that had been stored for eleven years. They found, in contrast, that eight of them (including albumin) were recoverable and exhibited some degree of antigenic/enzymatic activity (Sensabaugh, Wilson, & Kirk, 1971).

1.6.2 UV absorbance photometry

Tsutsumi *et al.* used UV absorbance photometry to quantify the activity of lactate dehydrogenase (LDH), glutamic oxaloacetic transaminase (GOT) and glutamic pyruvic transaminase (GPT) (Tsutsumi, Yamamoto, & Ishizu, 1983) in bloodstains. They were able to estimate the age of the bloodstain up to a period of 12 weeks using a combination of the LDH/GOT and GOT/GPT ratios. The GOT/GPT ratio remained stable over the first few weeks, while the

LDH/GOT ratio decreases over eight weeks (as referenced by (Andrasko, 1997; Inoue *et al.*, 1992)).

1.6.3 Aspartic acid racemisation

Decaying amino acids from blood plasma can also be used for age estimation through aspartic acid racemisation (AAR) (Arany & Ohtani, 2011). This method is only suited for older stains (> 10 years) because amino acids decay very slowly (Alkass, Buchholz, Ohtani *et al.*, 2010). The AAR method is based on the estimation of the D-L-aspartic acid ratio in tissues with a slow turnover, such as teeth, to predict the age of an individual. However, Arany *et al.* have shown that the soluble protein fraction of a bloodstain produces a strong correlation between time and D-aspartic content (Arany & Ohtani, 2011).

1.6.4 Enzyme-linked immunosorbant assay (ELISA)

Ackerman *et al.* have shown that it is possible to determine at what point in the 24 hour day a bloodstain was deposited (Ackerman, Ballantyne, & Kayser, 2010). Using an ELISA, they were able to detect melatonin and cortisol (two circadian biomarkers) in bloodstains that had been stored for a period of four weeks.

The majority of these approaches display limitations in their methodologies and applicability to current forensic practices. Most methods are incapable of differentiating between bloodstains from different species, are limited to blood stains, and may construe results due to the inability to control for different amounts of blood. The restrictions of these approaches highlight the need for continued research in order to accurately determine the age of bloodstains.

1.7 Effective means of estimating age

If the crime involves murder, it is possible to determine an approximate time of death. The most commonly used technique to relate the post mortem interval

to the commission of crime, in the case of the presence of the body, is measuring body temperature (Al-Alousi, 2002; Henssge, Brinkmann, & Puschel, 1984; Kaliszan, Hauser, & Kernbach-Wighton, 2009; Ritz-Timme, Cattaneo, Collins *et al.*, 2000). Archaeologists use carbon-14 dating as an effective means of dating historic organic samples (Bowman, 1990). Due to the stability of carbon-14, with a half-life of thousands of years, this method cannot be used to determine the time since deposition.

1.8 Stability of nucleic acids

Nucleic acids are the building blocks of life. Deoxyribonucleic acid (DNA) contains the genetic instructions used in the development and functioning of all living organisms (Venter, Adams, Myers *et al.*, 2001). DNA is typically used in crime scene-analysis. Ribonucleic acids (RNA) fulfil a variety of roles, messenger RNA (mRNA) direct the synthesis and assembly of proteins on ribosomes where ribosomal RNA (rRNA) join the amino acids together to form proteins (Blackburn, 2006). Due to the variety of roles, RNA is more abundant than DNA as there are thousands of copies per cell compared to the two copies of the majority of DNA sequences (Butler, 2005). The presence of hydroxyl groups in RNA makes it less chemically stable than DNA and this has been exploited in the past for developing methods to age stains (Al-Alousi, 2002; Hampson, Louhelainen, & McColl, 2011).

1.8.1 Mitochondrial DNA

In mammalian cells, genetic information is stored in two locations, in the nucleus and in the mitochondria. Nuclear DNA is organized into chromosomes of which two sets are present per cell, one paternal and one maternal. In contrast, mitochondrial DNA (mtDNA) inheritance is exclusively maternal with a few thousand copies per cell (Schaffer & Suleiman, 2007).

Mitochondrial DNA has been successfully isolated and amplified from aged and degraded forensic samples including skeletal remains (O'Rourke, Hayes, & Carlyle, 2000), fingernails (Cline, Laurent, & Foran, 2003) and hair (Melton

& Nelson, 2001). Numerous hypotheses accounting for the differences in the recovery of DNA from nuclear and mtDNA have been proposed. Foran suggested that differences are a result of being located in a different cellular organelle with different enzymes. Additionally, nuclear DNA is linear which may render it prone to exonucleases that would not digest mtDNA, a circular molecule (Foran, 2006). He examined the relative degradation of mitochondrial DNA and nuclear DNA in an attempt to understand why mitochondrial DNA can often be isolated from old or degraded samples when nuclear DNA could not (Foran, 2006). Two hypotheses were tested. The first was that mitochondrial DNA has a higher copy number compared nuclear DNA. The second was that mitochondrial DNA analysis is successful because of its cellular location within the mitochondrion; resulting in a higher level of protection then the nucleus affords chromosomal DNA. He found that copy number does influence successful analysis but it was the cellular environment that played a key role in mitochondrial DNA survival. In whole tissues, the Cytochrome b (Cyt b) gene sequence was found to degrade less than nuclear DNA. When the tissue was homogenized, Cyt b degraded faster than nuclear DNA.

1.8.2 Nuclear DNA

Although DNA is the carrier of genetic information, it has limited chemical stability. Hydrolysis, oxidation and nonenzymatic methylation of DNA occur at significant rates *in vivo*, and are counteracted by specific DNA repair processes. The spontaneous decay of DNA is likely to be a major factor in aging, carcinogenesis and mutagenesis (Lindahl, 1993). There are two main features of DNA that make it susceptible to damage. The first is the linkage between the deoxyribose carbon atom and the base, a bond that is highly susceptible to hydrolysis. The other feature of DNA that contributes to its chemical instability is the presence of a large number of oxygen and nitrogen atoms in the bases.

Hydrolysis of the deoxyribose carbon base linkage results in the complete loss of a purine or pyrimidine base, resulting in an abasic site. Hydrolysis is much

more likely to occur at purine bases, resulting in depurination of the DNA (Lindahl & Nyberg, 1972). Loss of DNA bases occurs slowly at normal physiological pH and temperatures. Hydrolysis is the main reason why DNA within dead tissues gradually loses its integrity.

Another common type of DNA damage that occurs under physiological conditions is the hydrolytic deamination of cytosine to form uracil (Lindahl, 1993). Cytosine deamination, like the formation of an abasic site, is caused by hydrolysis and is probably present in the DNA extracted from many sources. Unlike depurination, the rate of cytosine deamination is slowed in double-stranded DNA compared to single stranded DNA.

A third and common type of damage to DNA is oxidation. As in the case of hydrolytic damage, most DNA samples are susceptible to oxidation, as they are exposed to oxygen throughout storage. Many types of base modifications are created by oxidation, but the conversion of guanine to 8-oxo-guanine is one of the most common (Lindahl, 1993).

Kaiser *et al.* examined DNA degradation in human bones with post-mortem intervals (PMI) ranging from 1 to more than 200 years that had been kept under comparable conditions (Kaiser, Bachmeier, Conrad *et al.*, 2008). Following bone separation into the three different zones of interest (inner, middle and outer segments) they determined the quantity of total DNA in each region. The degree of DNA fragmentation was then determined by amplifying 150, 507 and 763 bp fragments with primers specific to the human multicopy β-actin-gene.

By examining the three different zones, they found a significant correlation between the zone of interest and the amount of DNA recovered. They concluded that the protected middle part of the bone is most useful in obtaining DNA due to the fact that extrinsic DNA (bacterial DNA) was present in high amounts on the extrinsic surface. They found no correlation between the amount of DNA and PMI but did find an inverse correlation between fragment length and PMI (Kaiser *et al.*, 2008). In this instance, post-mortem DNA degradation into increasingly smaller fragments reveals a time-

dependent process. This method has the potential to be used as a predictor of age in other sources of evidence, provided that environmental conditions are known.

1.8.3 RNA

As previously described, studies have used PCR amplification to detect DNA targets of different lengths as an assessment of degradation in samples of forensic interest (Alaeddini, Ahmadi, Walsh *et al.*, 2010; Swango, Hudlow, Timken *et al.*, 2007; Swango, Timken, Chong *et al.*, 2006). Investigations into the use of RNA and have also been attempted (Bauer, Gramlich*, et al.*, 2003; Bauer, Polzin, & Patzelt, 2003). Other non-nucleic acid methods such as electron paramagnetic resonance spectroscopy (Fujita *et al.*, 2005) and aspartic acid racemisation (Arany & Ohtani, 2011) have also been applied but these have been affected by temperature and light, or have failed to mimic real world samples by artificially accelerating the degradation process.

Ribonucleic acids are relatively new to the field of forensic science and their potential role into routine investigative work is vast (Bauer, 2007). It is therefore important to understand more of their behavioural properties because the type of environmental exposure will vary between samples taken from a crime scene. Ribonucleic acids are generally believed to be more unstable and prone to degradation, compared to DNA, due to the action of ribonucleases and environmental factors such as pH, UV light and humidity.

Recent advances in molecular forensics have revealed that a number of RNA markers can be successfully amplified under certain post-mortem and/or *in vitro* conditions (Inoue, Kimura, & Tuji, 2002; Ohshima & Sato, 1998; Zhao, Zhu, Ishikawa *et al.*, 2006). DNA has been successfully analysed from old forensic specimens or those that have been in their dried state for many years (Adler, Haak, Donlon *et al.*, 2011; King, Gilbert, Willerslev *et al.*, 2009; Marrone & Ballantyne, 2010). The level of nucleic acid degradation may therefore infer information relating to age.

RNA has become a widely studied biological molecule for the use in evidence analysis and forensic science investigation. RNA has recently been studied to estimate the post mortem interval (PMI) or time since death of a victim (Bauer, Polzin*, et al.*, 2003; Zhao *et al.*, 2006), to establish the type of biological stain deposited at a crime scene (Fleming & Harbison, 2010; Juusola & Ballantyne, 2003), and more recently in an attempt to establish the age of a biological sample (Anderson *et al.*, 2005; Bauer, Polzin*, et al.*, 2003). The methods aimed at determining the age of a bloodstain have all confirmed changing chemical and physical properties of biochemical components but none have demonstrated the precision required for forensic applications. Although RNA does not have distinguishing properties to differentiate between two people, when mixed with STR genotyping it could be used to help answer the questions: who left the biological stain and when? Not only would this strengthen evidence in court, but it may also help to exclude non-relevant suspect and or samples, allowing the investigators to focus their efforts on other leads.

Most RNA species have a well-characterized half-life and highly regulated mechanisms of decay *in vivo* (Elliott & Ladomery, 2011). These mechanisms, as they relate to mRNA, are diverse and dependent on factors such as secondary structure, specific base sequences, the 5' cap, the 5' untranslated region and the poly(A) tail (Beelman & Parker, 1995; Caponigro & Parker, 1995; Lemm & Ross, 2002; Parker & Song, 2004; Ross, 1995; Sachs, 1993; van Hoof & Parker, 2002). Although these mechanisms are numerous, the *ex vivo* mechanisms of RNA decay have not been investigated.

The majority of information regarding RNA stability has been obtained through post-mortem studies in various animal and human studies (Bauer, Gramlich*, et al.*, 2003; Fontanesi, Colombo, Beretti *et al.*, 2008; Gopee & Howard, 2007; Heinrich, Lutz-Bonengel, Matt *et al.*, 2007; Inoue *et al.*, 2002; Kuliwaba, Fazzalari, & Findlay, 2005; Marchuk, Sciore, Reno *et al.*, 1998; Noguchi, Arai, & Iizuka, 1991; Zhao *et al.*, 2006; Zubakov, Hanekamp, Kokshoorn *et al.*, 2008).

Inoue *et al.* profiled the post-mortem degradation of mRNA by extracting and analysing total RNA extracted from the brain, lung, heart and liver of rats. They found that total RNA was most stable in the brain and most unstable in the liver. Real-time RT-PCR analysis showed that the degradation rates of three housekeeping genes in the brain tissue were similar (Inoue *et al.*, 2002). Marchuk *et al.* investigated the integrity of RNA in dense connective tissue over time by extracting total RNA extracted from rabbit ligament, tendon and cartilage. They found no degradation of rRNA or loss in the integrity of mRNA for genes of low and high copy number up to 96 hours post-mortem at 4°C and up to 48 hours post-mortem at room temperature (Marchuk *et al.*, 1998). Noguchi *et al.* examined the post-mortem stability of arginine-vasopressin (AVP) mRNA in the rat brain by comparing changes with those in the recovered amounts of total RNA and rRNA. They found that the amount of AVP mRNA and rRNA decreased with post-mortem time, whereas the amount of total RNA did not. Analysis of the ratio of AVP mRNA to 18S-rRNA suggested that AVP mRNA was also degraded more rapidly (Noguchi *et al.*, 1991). Fontanesi *et al.* evaluated the effect of post-mortem times on the quality of porcine skeletal muscle total RNA in order to consider the possibility to use post-mortem material for gene expression analysis. By analysing the 28S:18S rRNA peak ratios at various post-mortem intervals, they found that RNA degradation was present at 48 hours post-mortem (Fontanesi *et al.*, 2008).

Gopee *et al.* examined the effect of PMI on the integrity of total RNA and the levels of representative mRNA species in murine cutaneous tissue. They found that there was no apparent change to the 28S/18S rRNA ratio or RNA integrity number at time points up to 60 minutes post-mortem nor were there any statistical differences in the relative expression of the five genes studied (Gopee & Howard, 2007).

Heinrich *et al.* investigated whether commonly used endogenous control genes were stable over various post-mortem intervals (Heinrich *et al.*, 2007).

They extracted RNA from three different human tissues of five post-mortem intervals ranging from 15 to 118 hours. They then measured the C_t values obtained by real-time RT-PCR from five commonly used endogenous control genes: beta actin (β-actin), Beta-2 microglobulin (B2M), cyclophilin A (CyPA), TATA-binding protein (TBP), and ubiquitin C (UBC). They found that RNA in skeletal muscle was relatively stable, whereas all endogenous controls were highly variable in heart and brain tissue. A key finding of this study was that post-mortem mRNA degradation is a complex process, and that the use of one single endogenous control in gene expression studies of post-mortem tissue could lead to erroneous data interpretation and further studies should be performed.

Bauer *et al.* developed a method for quantitating mRNA degradation by multiplex-RT–PCR coupled to laser-induced fluorescence capillary electrophoresis. Their study included whole blood samples from living individuals and post-mortem blood and brain samples which were stored at 4°C for up to 5 days. They found that RNA degradation is significantly correlated with the storage interval of blood samples taken from living individuals and with the PMI in autopsy cases (Bauer, Gramlich*, et al.*, 2003).

Zhao *et al.* aimed to establish quantitative assays of oxygen-regulated factors including erythropoietin (EPO), vascular endothelial growth factor (VEGF) and hypoxia-inducible factor 1 alpha (HIF1A) mRNAs, and to investigate the post-mortem stability of these mRNA transcripts in forensic autopsy materials. They performed relative quantification of EPO, VEGF and HIF1A mRNAs on autopsy tissue specimens from the heart, brain, kidney and lung stored at room temperature for various times. They found that VEGF and HIF1A mRNA progressively degraded in patterns similar to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA used as the endogenous reference. The relative quantification of VEGF/GAPDH and HIF1A/GAPDH changed little up to 48 hours post-mortem in tissue samples from the brain, kidney and lung (Zhao *et al.*, 2006).

Kuliwaba *et al.* examined the stability of total RNA and bone-specific mRNA

from bone samples obtained from routine autopsies and at surgery. They found that with appropriate storage and handling, RNA can reliably be extracted from human bone obtained at post-mortem and surgery that would aid in *ex vivo* gene expression studies (Kuliwaba *et al.*, 2005).

Zubakov *et al.* performed whole-genome gene expression analyses on a series of time-wise degraded blood and saliva stain samples to generate sets of stable RNA markers for reliable identification of these body fluids. They identified nine stable mRNA markers for blood and five stable mRNA markers for saliva detection showing tissue-specific expression signals in stains aged up to 180 days of age (Zubakov *et al.*, 2008).

1.8.4 Factors that affect nucleic acid stability and degradation

RNA

A large number of factors can affect RNA stability but they can generally be divided into 2 groups: 1) all cellular factors, 2) the primary, secondary and tertiary structure of RNA. There are several factors required for oligoribonucleotide cleavage that can also be divided into 2 groups (Bibillo, Figlerowicz, Ziomek *et al.*, 2000). The first comprise the necessary environment of the RNA molecule and include polyamines and non-specific cofactors. The second group consists of requirements of the RNA structure: 1) single stranded RNA, 2) the presence of the labile phosphodiester bond, 3) the position of the scissile internucleotide bond within the oligoribonucleotide and 4) the presence of certain functional groups on the nucleotides flanking the cleaved phosphodiester bonds. The nucleotide sequence composition or folding energy can also play a role in determining the half-life of a transcript (Kaukinen & Mikkola, 2002). Transcripts with a high guanine-cytosine (G+C) content tend to be more stable on average (the G+C content for 18S rRNA is lower than G6PD: 56.2 % and 62.6 % respectively) because they have an additional hydrogen bond (Galtier & Lobry, 1997; Katz & Penman, 1966).

Although all RNA molecules present in a given cell are exposed to the same RNA degradation machinery, the half-life of RNA can range from hours to seconds (Lorentzen & Conti, 2006). Our knowledge about different factors affecting RNA degradation is very limited and it is not currently possible to predict the stability of individual RNAs. Some short-living mRNAs have been shown to carry adenosine and uracil rich cis-acting elements (ARE) usually located in their 3' untranslated region (UTR), these molecules undergo rapid degradation by the ARE-mediated mRNA decay (Houseley & Tollervey, 2009). The primary, secondary and tertiary structures of RNA molecules could also significantly affect their degradation but current experimental methods do not allow for studying all aspects of this process.

In contrast to mRNA, the degradation pathways of stable RNAs (rRNA, tRNA), accounting for more than 90% of total cellular RNA, have been poorly characterised. Recent reports have shown that rRNA and tRNA can undergo cleavage in response to oxidative stress or under developmental regulation (Thompson, Lu, Green *et al.*, 2008).

Oligoribonucleotides must be single stranded for specific, non-enzymatic hydrolysis (Bibillo *et al.*, 2000). However, the length of the single stranded fragment required to undergo specific cleavage is not established.

The chemical stability of individual phosphodiester bonds in RNA is important for many biological processes, including RNA degradation, site specific cleavage, and splicing. Kierzek (Kierzek, 2001) reported that some single stranded RNAs have phosphodiester bonds that are very susceptible to hydrolysis when in the presence of certain cofactors. The results described suggested that an internal single stranded UA sequence is unusually susceptible to hydrolysis.

Kierzek (2001) designed an experiment to determine how long a single RNA must be in order to undergo specific cleavage. He developed synthetic RNA molecules that were designed to be complementary to each other (a perfect duplex) plus 1, 2 and 3 nucleotides shorter at the 5' end. As a result, different

bases of the 3' end were exposed as single strands. Specific cleavage (of UA) only took place in the presence of the shortest complementary oligomer (the one with the greater number of exposed 3' bases). Thus, for effective cleavage, both nucleotides participating in the formation of a labile phosphodiester bond must be unpaired. The lack of complementary base pairing increases the flexibility of the labile phosphodiester bond. This allows an in-line orientation of the 2'-hydroxyl of uridine, the phosphorous atom, and the 5'-oxygen atom of adenosine, as required for cleavage of oligoribonucleotides (Bibillo *et al.*, 2000).

Generally, the phosphodiester bond instabilities are in the following order, with the least stable linkages first: UA > CA > YC > YG > YU (Y = pyrimidine). A polyamine is also necessary for nonenzymatic cleavage (unbound amount in humans (Desser, Hocker, Weiser *et al.*, 1975): 1.8±1.4 nmol putrescine/10^8 cells, 3±0.9 nmol spermidine/10^8 cells and 12.9±3.8 /10^8 cells spermine/10^8 cells). The amount of polyamines required for the maximum rate of specific cleavage are: 0.1 mM for spermine, 1.0 mM for spermidine, and 10 mM for putrescine (Bibillo *et al.*, 2000; Kierzek, 2001).

Kierzek (2001) tested the biological relevance on the stability of transfer RNA (tRNA) under nonenzymatic cleavage conditions. Transfer RNA was chosen because it is relatively small, well characterised and the secondary structure is well defined (Neidle, Schneider, & Berman, 2005). The dominating cleavage position was the phosphodiester bond between C75 and A76 (3' end degradation) and didn't allow observation of the cleavage of other phosphodiester bonds, most of which are 5' to this position. Non-enzymatic cleavage could account for the observed reduction of the peak height for the various sized fragments as they fail to amplify.

Blazewicz *et al.* (2011) focused on the structural factors affecting RNA stability and their work involved the synthesis of two artificial RNA molecules in an attempt to describe (predict) the RNA degradation process using bioinformatics. Both of the artificial RNA molecules were designed in such a way that they contained well-defined unstable regions as described by Kierzek.

As a result, they proposed a new algorithm (called the RNA Partial Degradation Problem) that is capable of reconstructing an RNA molecule based on the length of degraded fragments which in turn determines the location of cleavage sites and thus identifying the unstable regions which are most susceptible to degradation.

1.8.5 Using the stability of nucleic acids to age bloodstains

As described earlier, there are too few reliable and accurate methods to approximate the time of deposition of these dried biological stains. RNA is not as stable as DNA; therefore RNA degradation measurements seem a suitable candidate for age determination of bloodstains. RNA degradation in bloodstains has been attempted for age determination by Bauer *et al.* (Bauer, Polzin*, et al.*, 2003), and Anderson *et al.* (Anderson *et al.*, 2005).

Bauer *et al.* used a semi-quantitative duplex and competitive real-time PCR method to estimate the approximate age of bloodstains, with the assumption that degradation of mRNA occurs from the 5' end. However this study did not attempt to investigate direction of mRNA degradation in dried biological stains in order to verify that this assumption is correct, validating their proposed method. They reported that a difference of 4-5 years between samples is required in order to detect a significant difference in the RNA obtain from bloodstains. As a consequence, estimations produced using this method are too large to be forensically useful.

Anderson *et al.* (2005) examined the ratios of β-actin mRNA and 18S rRNA as a function of time using real-time PCR. A relationship between the age of the bloodstain and RNA degradation was found when the ratio of 18S rRNA to β-actin mRNA was investigated. The Ct value of 18S rRNA did not change over the course of 150 days, but the Ct values for β-actin mRNA became significantly higher, thus the relative ratio of 18S rRNA to β-actin mRNA increased over time. Whilst this method is more compatible with the current capabilities of most forensic laboratories, the analysis produces crude time estimates as samples were only examined every four weeks. This study was

simple in its approach as it only measured the ratios of β-actin mRNA and 18S rRNA using the Ct value from RT-PCR. It provides no analysis detailing the extent or type of degradation occurring to the RNA species that may prove useful in obtaining a more accurate estimation of time of deposition.

Anderson *et al.* (2011) improved upon their earlier methodology by using a multiplex qPCR assay to determine the relative stability of different sized segments of the same RNA species as well as incorporating the differences they observed between 18S rRNA and β-actin mRNA (Anderson, Hobbs, & Bishop, 2011). They found that multivariate analysis of the changing ratio of the different 18S rRNA and β-actin mRNA segments improved the accuracy of bloodstain age estimation and could be used to differentiate between samples of different ages in a defined population.

The age estimation by Anderson *et al.* (2005) was limited to half a year of aging, while Bauer *et al.* showed that quantification of degraded RNA is possible for bloodstains stored up to 15 years. It was later determined that for highly degraded biological samples, the amount of detection using PCR was directly correlated to the size of the target. It was found that for highly degraded samples, a smaller sequence was detected more than a larger sequence (Swango *et al.*, 2006). It was hypothesized that if the factors responsible for loss of RNA-derived signal acted in a random, stochastic fashion, larger segments will be more susceptible to degradation. This will cause the larger segments to degrade more rapidly and therefore amplification through PCR will be significantly reduced over time. Anderson *et al.* found that the larger segments did indeed degrade at a faster rate than smaller segments (Anderson *et al.*, 2011).

Qi *et al.* (2013) investigated the 18S rRNA: β-actin mRNA ratio in bloodstains for a period of 28 days under uncontrolled room conditions to observe the progressive degradation of the two RNA species (Qi, Kong, & Lu, 2013). Their methods involved aliquoting 10 µl of blood onto cotton cloth and "similar-sized samples of dried bloodstains were cut from the fabric" for analysis. The authors mentioned that control cDNA was used to adjust for experimental variation caused by the real-time machines but no attempts were made to

address experimental variation between different samples and time points as a result of analysing “similar-sized” stains. Nevertheless, they found that the ratio changed in a linear fashion and the gender of the originating stain affected the relationship between time and the RNA ratio.

Simard *et al.* (2012) tested the stability of four different RNA targets (18S rRNA, β-actin mRNA, GAPDH mRNA and PPIA mRNA) in blood, semen, and saliva stains in order to enable the determination of the time since deposition of these bodyfluids (Simard, DesGroseillers, & Sarafian, 2012) when stored at room temperature in the dark for up to six months. Unlike Anderson *et al.* they found that the decay rates of rRNA compared to the mRNA did not exhibit significant differences over the time points studied. They were not able to establish a good correlation between relative mRNA/18S rRNA ratios and time of storage. Simard *et al.* concluded that the degradation profiles of individual RNA markers may be useful in estimating the age of sample as an exclusionary test based on their detection.

Bremmer *et al.* (2012) gave an extensive review of the literature and techniques that address the issue of determining the age of a bloodstain (Bremmer, de Bruin, van Gemert *et al.*, 2012). Figure 1 shows that most techniques (mentioned above) are complementary to each other, for both recent and long term age determination. They also addressed the need for standardised experimental guidelines for future experiments.

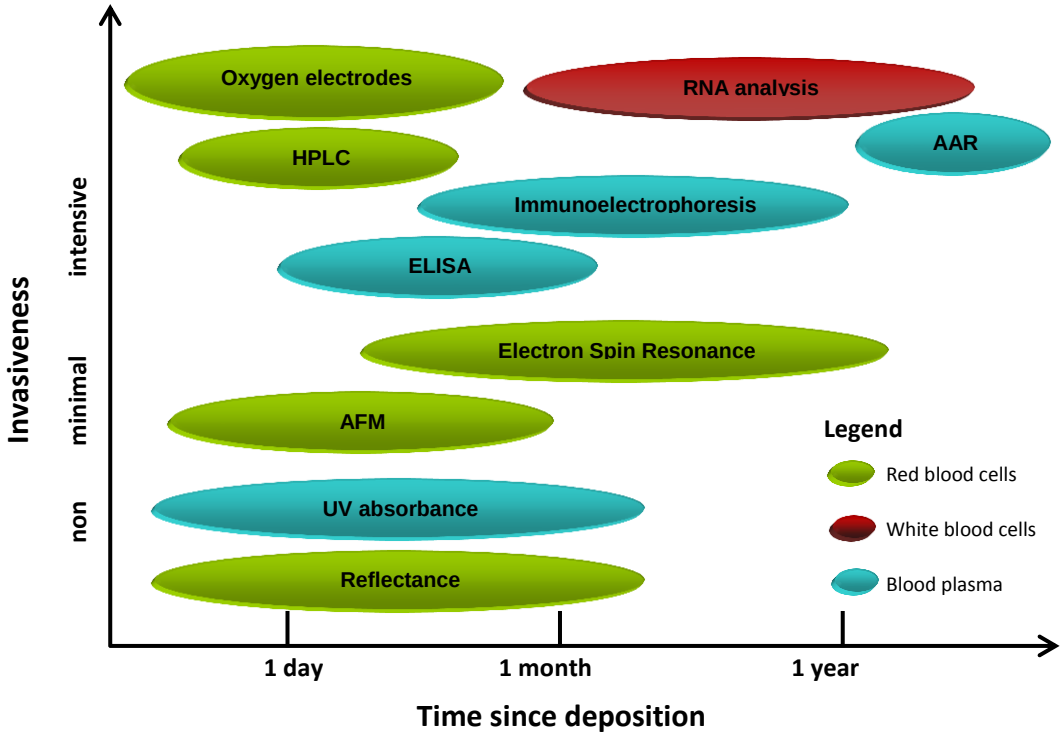


Figure 1.1. Overview of techniques currently used for bloodstain age estimation. The horizontal axis shows the time scale for which each technique is most suited, ranging from less than one day to more than one year. The vertical axis depicts the level of invasiveness of the technique (adapted with permission from Bremmer *et al.* (2012)). The relevant methods for blood plasma, red blood cells and white blood cells are discussed in sections 1.5, 1.6 and 1.8 respectively.

1.9 Research aims

The aim of this thesis was to develop a method that can be used to determine the time since deposition of a bloodstain and to test this method on forensic type samples. To achieve this aim, nucleic acid degradation of DNA, rRNA and mRNA was evaluated with the hypothesis that the degradation profile of nucleic acids in dried bloodstains can be used to determine the time since deposition. This is based on the assumption that biochemical reactions occur during the drying phase that damage and degrade the nucleic acids in the bloodstain. As the stain dries, the damage and degradation to the nucleic acids can be studied to give a characteristic "profile" that is indicative of its age. The study also included the development of a quantification method to standardise the amount of cDNA added to each assay to reduce the amount of variability between different time points. The aims of this thesis are summarised below:

- Develop a method to standardise the amount of cDNA template used for analysis.
- Develop multiplex assays to assess degradation in DNA, rRNA and mRNA.
    - Determine if nucleic acid degradation occurs in a predictable manner such that TSD can be estimated.
- Statistical analysis of nucleic fragments following deposition.
    - Develop a prediction equation.
- Test the equation on forensic type samples

Chapter 2
General Methods

The materials and methods described in this chapter relate to the extraction and processing of nucleic acids from bloodstains that are used for downstream applications. Specific materials and methods or any deviations from this process can be found or are discussed in their relevant chapter.

2.1 Ethics and sample collection

Ethics approval was obtained from the University of Auckland Human Participants Ethics Committee for the collection of body fluid samples from volunteers.

2.2 Co-isolation and purification of DNA and RNA

The basic steps of nucleic acid isolation are disruption of the cellular structure to create a lysate, separation of the soluble nucleic acid from cell debris and other insoluble material and purification of the nucleic acid of interest from other soluble proteins or other nucleic acids. Historically, this has been done using organic extraction followed by ethanol precipitation. Alvarez *et al.* developed a method to co-extract mRNA and DNA from the same physiological stain based on a modified phenol:chloroform extraction method (Alvarez, Juusola, & Ballantyne, 2004). However, these methods are time consuming and use a variety of hazardous reagents and have consequently been replaced by newer techniques such as magnetic bead separation.

The DNA IQ™ System is a DNA isolation system for the forensic and paternity community. This kit employs a novel technology with paramagnetic beads to prepare clean samples for short tandem repeat (STR) analysis. The DNA IQ™ System can be used to extract DNA from stains or liquid samples such as blood. One of the main benefits of using the DNA IQ™ System is that the DNA IQ™ resin reduces PCR inhibitors and contaminants frequently encountered in forensic casework. Each sample was extracted using the DNA IQ™ System (Promega, Madison, WI) following the manufacturer's instructions and the extraction process is described in Figure 2.1.

The separation of DNA and RNA using the DNA IQ™ System has been previously described by Bowden *et al.* (Bowden, Fleming, & Harbison, 2011), with modifications to the RNA purification process for the purpose of this thesis which are described below. DNA was eluted in 50 µl of elution buffer that was contained in the kit. Twenty microliters of freshly drawn blood was used as a positive control and an extraction blank (cotton cloth) was included as a negative control for each experiment.

2.3 RNA purification

RNA recovered from the DNA IQ™ lysis buffer retentate was prepared using the Clean and Concentrator kit™ (Zymo Research, Irvine, CA) according to the manufacturer's protocol. The RNA Clean and Concentrator kit™ provides a simple and reliable method for the rapid preparation of high-quality RNA. This procedure is based on the use of a RNA binding buffer and Fast-Spin column technology. All of the reagents were provided in the kit.

Two volumes of RNA Binding Buffer were added to each volume retentate followed by adding one volume of 100% ethanol to the mixture.

The solution was then transferred to a Zymo-Spin™ Column and centrifuged at ≥12,000 x *g* for 30 seconds. The samples were DNase treated using the in-column DNase I digestion protocol for 45 minutes.

Four hundred microliters of RNA Prep Buffer was added to the column and centrifuged at ≥12,000 x *g* for 1 minute and the flow through was discarded.

Four hundred microliters of RNA Wash Buffer was added to the column and centrifuged at ≥12,000 x *g* for 30 seconds, the flow through was discarded and the process was repeated.

The Zymo-Spin™ Column was then centrifuged in an emptied collection tube at ≥12,000 x *g* for 2 minutes to ensure the complete removal of the wash buffer.

The column was then transferred into an RNase-free tube. The purified RNA was eluted in 20 µL of nuclease free water that was part of the kit following centrifugation at 13,000 x *g* for 1 minute.

One microliter of the eluted RNA was used to quantify genomic DNA and the remainder was stored at -20°C prior to cDNA synthesis.

2.4 Quantification of genomic DNA

Although the Clean and Concentrator kit™ had specified DNA free RNA, the eluted RNA samples were tested with the Quantifiler™ human DNA assay (Applied Biosystems by Life Technologies, Carlsbad, CA) to make sure that there was no residual DNA in the purified RNA. The manufacturer's instructions were followed using 1 µL RNA in a 12.5 µL reaction ("Quantifiler Human DNA Quantification Kit,"). If samples contained DNA, they were re-treated with DNase I and quantified again.

2.5 Complementary DNA (cDNA) synthesis

Complementary DNA was synthesized in 20 µL reactions using the High Capacity RNA-cDNA Master Mix kit (Applied Biosystems by Life Technologies, Carlsbad, CA), using 16 µL of the purified RNA and 4 µL of Master Mix according to the manufacturer's instructions. The cDNA was stored at -20°C until required.

2.6 CellTyper

CellTyper is a multiplex PCR system designed by Fleming and Harbison (Fleming & Harbison, 2010) that uses mRNA to identify blood, saliva, semen and menstrual blood in individual stains or in mixtures of body fluids. They have identified mRNA transcripts specific to each type of body fluid and developed a multiplex reverse transcriptase-polymerase chain reaction assay to identify these body fluids along with three housekeeping genes. The mRNA multiplex is able to definitively identify circulatory blood, menstrual blood, semen and saliva in forensic casework samples, whilst co-extracting

the nuclear DNA for profiling. This method was used to validate the qPCR assay described in chapter 3.

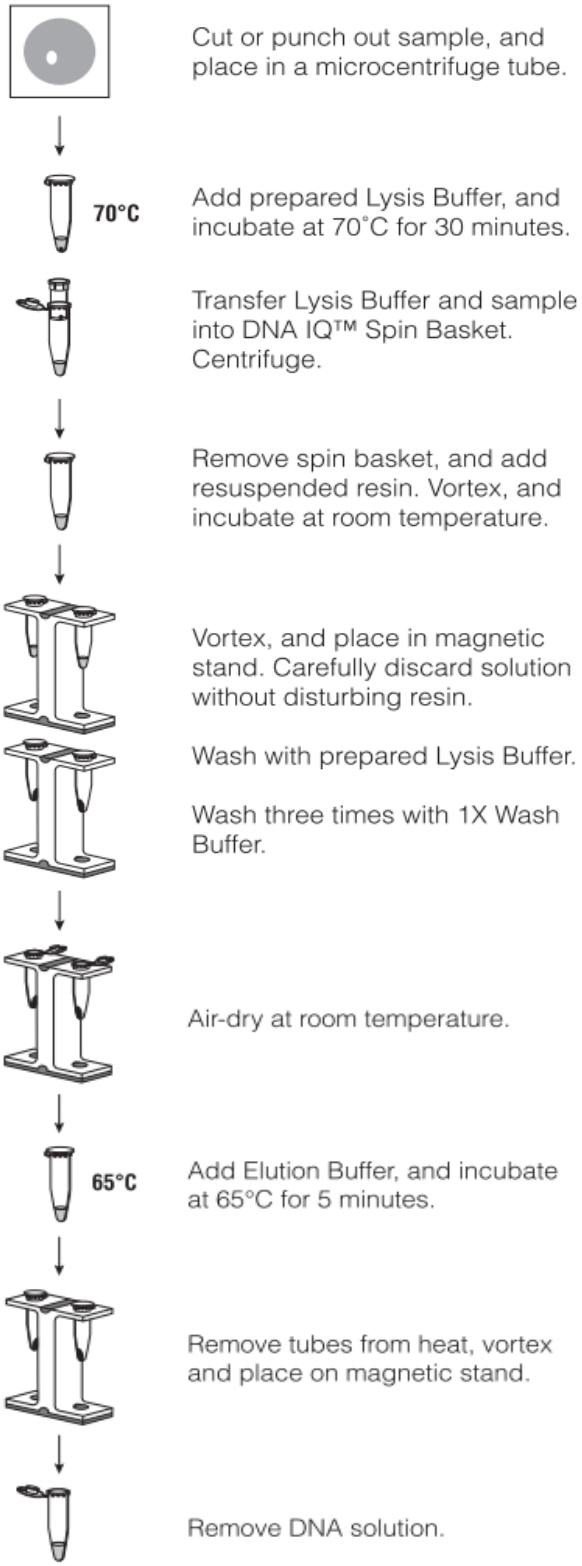


Figure 2.1. Schematic of DNA isolation from stains on solid material using the DNA IQ™ System (Promega).

2.6.1 Body fluid genes

The multiplex assay is designed to amplify glycophorin A (GlycoA) for blood, matrix metalloproteinase 11 for menstrual blood (MMP11), histatin 3 (HTN3) and statherin (STATH) for saliva, protamine 2 (PRM2) for spermatozoa and transglutaminase 4 (TGM4) for seminal fluid. Three housekeeping genes were also included in the multiplex; transcription elongation factor 1α (TEF), glucose 6-phosphate dehydrogenase (G6PD) and ubiquitin conjugating enzyme (UCE). Table 2.1 lists the primer sequences used in the multiplex assay and Figure 2.2 shows the position of the amplicons.

Table 2.1. Body fluid genes and primer sequences used in the multiplex PCR reaction to identify the origin of the body fluid.

Gene	Genbank accession number	Primer sequence	Amplicon size (bp)
GlycoA	NM_002099	Forward: Hex - CAGACAATGATACGCACAAACG	188
		Reverse: CCAATAACACCAGCCATCACC	
MMP11	NM_005940	Forward: FAM - CAAGACTCACCGAGAAGGGG	173
		Reverse: TAGCGAAAGGTGTAGAAGGCG	
HTN3	NM_000200	Forward: FAM - TGGGGCATGATTATGGAGGTT	233
		Reverse: CAGAAACAGCAGTGAAAACAGCTT	
STATH	NM_003154	Forward: Hex - CTTGAGTAAAAGAGAACCCAGCCA	162
		Reverse: TTCTGGAACTGGCTGATAAGGG	
PRM2	NM_002762	Forward: FAM - CGTGAGGAGCCTGAGCGA	201
		Reverse: CGATGCTGCCGCCTGT	
TGM4	NM_003241	Forward: Hex - TGAGAAAGGCCAGGGCG	215
		Reverse: AATCGAAGCCTGTCACACTGC	
TEF	M81601	Forward: FAM - TGGGCCATCAACTGAGAAAGA	206
		Reverse: TCTCCCTACACTTCAACTGCACA	
G6PDH	X03674	Forward: FAM - ATCATCGTGGAGAAGCCCTTC	181
		Reverse: GTTCCAGATGGGGCCGA	
UCE	U39317	Forward: Hex - AATGATCTGGCACGGGACC	241
		Reverse: ATCGTAGAATATCAAGACAAATGCTGC	

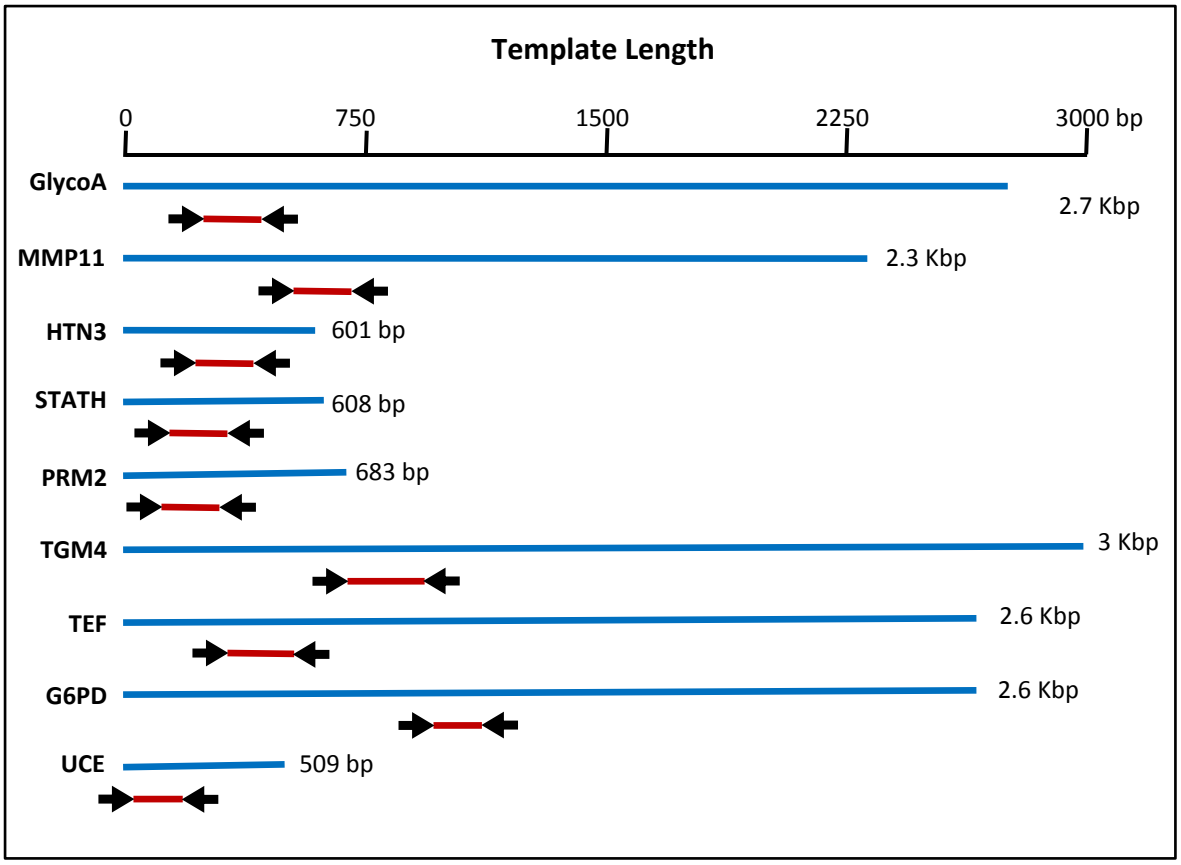


Figure 2.2. Bodyfluid multiplex assay design. Position of each amplicon used in the CellTyper assay for the amplification of glycophorin A (GlycoA) for blood, matrix metalloproteinase 11 for menstrual blood (MMP11), histatin 3 (HTN3) and statherin (STATH) for saliva, protamine 2 (PRM2) for spermatozoa and transglutaminase 4 (TGM4) for seminal fluid. Three housekeeping genes were also included in the multiplex; transcription elongation factor 1α (TEF), glucose 6-phosphate dehydrogenase (G6PD) and ubiquitin conjugating enzyme (UCE).

2.6.2 cDNA multiplex polymerase chain reaction

The Qiagen multiplex PCR Kit (Qiagen, Germantown, MD) was used for the CellTyper assay. The reaction mix contained 12.5 µL of reaction buffer, 2.5 µL of 10 × primer stocks and 8 µL sterile water. The final optimised concentrations of primers (forward and reverse) in the multiplex reaction were: GlycoA 0.2 µM, MMP11 0.1 µM, HIS 0.25 µM, STATH 0.25 µM, PRM2 0.05 µM, TGM4 0.1 µM, TEF 0.025 µM, G6PDH 0.2 µM and UCE 0.125 µM. The PCR conditions were 95 °C for 15 min, 94 °C for 30 s, 58 °C for 30 s, and 72 °C for 90 s for a total of 30 cycles and a final step of 72 °C for 45 min.

2.7 Bloodstain age estimation

The development of the age estimation method is described in chapters 3 to 5. A summary of the method is presented here.

2.7.1 cDNA quantification

2.7.1.2 Standard curve

To create the standard curve for cDNA quantification, 1 µL of DNA-free leukocyte RNA (human blood peripheral leukocytes total RNA, 1 µg/µL (ClonTech, Mountain View, CA) was reverse transcribed into cDNA as described in section 2.5. The leukocyte cDNA was serially diluted by one third such that the starting concentration was 50 ng/µL and the lowest was 0.023 ng/µL creating 8 different concentrations (50, 16.7, 5.56, 1.85, 0.62, 0.21, 0.066, 0.023 ng/µL). One microliter of each cDNA standard was added to the qPCR assay for the detection of the cluster of differentiation gamma-3 molecule (CD3G) mRNA and 18S rRNA markers.

2.7.1.3 Real-time quantitative PCR analysis (qPCR) of samples

Primers and probes for a 63 bp segment corresponding to the mRNA for the CD3G complex (GenBank accession number NM_000073.2) were: forward primer GCATTTTCGTCCTTGCTGTTG, reverse primer AGCTCTCGACTG GCGAACTC, FAM probe TCTACTTCATTGCTGGACAGG Primers and probes for the 18S rRNA molecule were obtained from the 18S ribosomal RNA control kit from Applied Biosystems (Applied Biosystems by Life Technologies, Carlsbad, CA).

The final optimised primer and probe concentrations used in the reactions were as follows: 200 nM 18S rRNA probe (VIC dye) and 50 nM for both the forward and the reverse 18S rRNA primers; 250 nM CD3G mRNA probe (FAM dye), 300 nM forward and 100 nM reverse CD3G mRNA primers. Realtime qPCR analysis was performed using an ABI-7500 SDS instrument and 5x qARTA EvaGreen qPCR mix, with ROX as the reference dye (qARTA Bio, Carson, CA) following the manufacturer's instructions.

The thermal cycling conditions consisted of an initial denaturation step at 95°C for 15 minutes, denaturation at 95°C for 15 seconds, annealing at 60°C for 20 seconds and an elongation step at 72°C for 1 minute for 40 cycles. Duplicate reactions were run for each standard and unknown cDNA sample, which included a negative control. Real-time qPCR reactions for the standards and unknown samples were all performed in 20 µL reaction volumes containing 1 µL of cDNA.

2.7.1.4 Real-time PCR data analysis

The qPCR data was analysed using SDS software (v1.2; Applied Biosystems by Life Technologies, Carlsbad, CA). The fluorescence threshold was set in the middle of the exponential (linear) phase (phase 2) of the PCR with a manual baseline determined for each probe.

2.8. Multiplex assays for age determination

2.8.1 G6PD mRNA and 18S rRNA multiplex assay

Six hundred picograms of cDNA template was added to the G6PD assay and 200 pg of cDNA was added to the 18S rRNA assay and amplified in a total reaction volume of 25 µl. The Qiagen multiplex PCR kit was used for each assay. The reaction mix contained 12.5 µl of master mix, 2.5 µl of 10x primer stock (Table 2.2) and made up to 25 µl with sterile water.

The following PCR thermal cycling conditions were used: 95°C for 15 minutes, 94°C for 30 seconds, 58°C for 30 seconds and 72°C for 90 seconds for a total of 30 cycles and a final elongation step of 72°C for 10 minutes. Prior to capillary electrophoresis analysis, the amplified 18S products were diluted to assist with analysis. One microliter of amplified product was added to 99 µl of sterile DNase free water in a sterile PCR tube (1:100 dilution) and vortexed for 30 seconds.

Amplification products were separated on a 3130 genetic analyser (Life Technologies, Carlsbad, CA). One microliter of template was loaded with 9 µl

of HiDi™ Formamide, POP-4 polymer was used as the separation medium and Multi-capillary DS-30 (Dye set D) was used as the matrix standard. GeneScan™ 500 LIZ was used as the internal size standard and results were analysed using Peak Scanner (Life Technologies, Carlsbad, CA) version 1.0. No template controls were included and all samples were amplified in duplicate.

2.9. Age prediction

After analysing the experimental data, it was determined that the following fragments/amplicons should be included in the final model: The G6PD 5' medium, 3'medium, 3' long fragment and the 18S central long fragments.

Once the peak heights of the G6PD 5' medium, 3'medium, 3' long fragment and the 18S central long fragments were obtained from test samples, they were substituted into the prediction equation and the peak height of each fragment was multiplied by the corresponding coefficient. The prediction equation and model coefficients are shown in Table 2.3 and are derived from the experimental data given in Chapter 5.

Table 2.2. Multiplex assay design. Primer sequences used in each multiplex assay for 18S rRNA and G6PD mRNA. (5') – 5' region, (C) – central region, (3') – 3' region, (5M) – 5' medium fragment, (CL) – central large fragment, (3M) – 3' medium fragment and (3L) – 3' large fragment.

	Primer sequences	Final conc. (µM)	Expected size (bp)
18S rRNA *(NR_003286.2)*	Forward: FAM – TTA GAG TGT TCA AAG CAG GCC (C)	0.05	
	Reverse: CCC TTC CGT CAA TTC CTT T (CL)	0.02	399
G6PD mRNA *(NM_000402.3)*	Forward: NED – GCC TTC CAT CAG TCG GAT AC (5M)	0.015	
	Reverse: GAA GGG CTC ACT CTG TTT GC (5')	0.1	189
	Forward: VIC – GGC AGA CGA GCT GAT GAA GA (3')	0.2	
	Reverse: GAA TGT GCA GCT GAG GTC AA (3M) GGA GTG AGT GGA GG AGG TGA (3L)	0.1 0.8	177 331

Table 2.3. Regression model equation and associated coefficients used for age prediction.

Ln_time = c + B1(G6PD_5M) + B2(G6PD_3L) + B3(G6PD_3M) + B4(18S_CL)

	Coefficient	**Value**
Intercept	*c*	4.994519031
G6PD_5M	*B1*	-0.001872919
G6PD_3L	*B2*	-0.003621702
G6PD_3M	*B3*	0.000332205
18S_CL	*B4*	-0.001228488

Chapter 3

Development of a Method to Standardise the Amount of cDNA Added to a PCR Reaction

3.1 Introduction

RNA analysis has been under the spotlight over the last few years because of its potential to supplement current DNA analysis. Its use has been trialled in a variety of applications including identifying the source of body fluids in forensic samples, determining the age of wounds and stains and thus complement or even substitute current techniques (Anderson *et al.*, 2005; Bauer, 2007; Bauer, Kraus, & Patzelt, 1999; Juusola & Ballantyne, 2003; Nussbaumer, Gharehbaghi-Schnell, & Korschineck, 2006; Ohshima & Sato, 1998). As with DNA, quantification of the recovered genetic material is important for the quality assurance and sufficiency for the RNA analysis process.

The generalised process for RNA analysis is to co-extract RNA with DNA from the samples, remove the DNA from the RNA fraction using DNase I followed by converting the RNA into complementary DNA (cDNA) which is used in downstream applications. A simplified schematic of the overall process is shown in Figure 3.1.

Assessing nucleic acid recovery (DNA and RNA) after extraction can prove useful in determining the approach for downstream applications. Due to the limited nature of biological evidence that may be recovered from some crime scene samples, decisions often have to be made whether or not to proceed with testing low amounts of template. Nucleic acids from almost any source, for example blood or fingerprints, can be used as a template for analysis. However, the success of the analyses in terms of yield and reliability are highly dependent on the way the sample was isolated and stored. Successful analysis would also depend on factors such as the degradation of the DNA or RNA.

The template source can also have an adverse effect on PCR - DNases and polysaccharides can be present in intact tissue and heme can be carried over from the extraction of DNA from blood, and inhibit PCR. The Quantifiler™

human DNA assay kit (Applied Biosystems by Life Technologies, Carlsbad, CA) for example, has an internal positive control (IPC) that may be used to detect for the presence of inhibitors that may inhibit downstream analysis. Inhibited samples would exhibit an IPC with a higher than expected Ct value. Samples exhibiting potential inhibition may also result in lower than expected quantification results requiring further sample clean-up or dilution in order to obtain accurate quantification. Neglecting to re-quantitate an inhibited sample after clean-up or dilution may result in off-scale data for smaller fragments when run on a capillary electrophoresis instrument, impairing the final results ("Minifiler FAS Questions,").

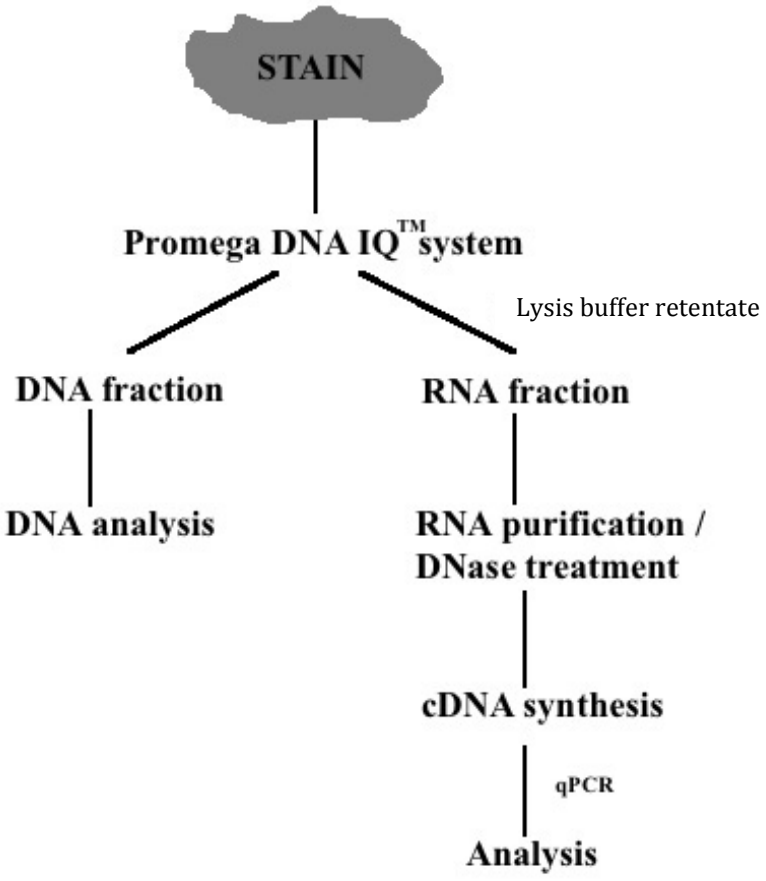


Figure 3.1. The co-extraction of DNA and RNA. Overview of the method used for the simultaneous analysis of DNA and RNA.

Methods for the quantification of DNA from forensic type case samples (Nielsen, Mogensen, Hedman *et al.*, 2008) and RNA do exist (Aranda, Dineen, Craig *et al.*, 2009; Jones, Yue, Cheung *et al.*, 1998). Due to the limited amount of sample encountered in forensic casework and the integrity of the RNA, there are inherent pitfalls for RNA quantification (Fleige & Pfaffl, 2006). Common methods used for the quantification of RNA include: spectrophotometric ultraviolet-visible (UV–VIS) assays, which determine the amount of total RNA present in any given sample based on absorbance readings; and fluorometric assays which determine the concentration of the molecule based on the amount of intercalated dye but suffer from interference caused by impurities and dye stability (Dell'Anno, Fabiano, Duineveld *et al.*, 1998; Lee, 1997; Lightfoot, 2002). These systems provide good estimates of total RNA (provided that the samples are free from genomic DNA contamination), but will often overestimate the amount of material suitable for analysis because mRNA typically only constitutes about 3% of total cellular RNA (Miura, Ichikawa, Ishikawa *et al.*, 1996). The RNA quantification methods are often not sensitive enough, prone to interference from contaminants present in the sample such as organic residues, metal ions and lipids that can interfere with the reverse transcription reaction (Imbeaud, Graudens, Boulanger *et al.*, 2005) and are therefore not suitable for downstream applications.

Alternatively, the presence and quantity of RNA in a sample may be determined based on the transcript level of constitutively expressed housekeeping genes (Barber, Harmer, Coleman *et al.*, 2005; Bustin, 2000). Tissue or cell-specific RNA transcripts may be quantified relative to the housekeeping genes using quantitative real-time polymerase chain reaction (qPCR), which is based on the amount of fluorescence emitted during amplification (Bustin, 2000).

This chapter describes the development of a qPCR method that determines the amount of blood specific cDNA added to a PCR reaction. The method allows for the amount of cDNA, reverse transcribed from RNA, to be normalized for addition to subsequent PCR reactions and standardise assays

by removing a variable. Most forensic research involving the use of RNA does not take sample-to-sample variation between assays into account and use fixed volumes of cDNA for analysis; this has been addressed in the methods section of this chapter.

Two markers were selected to create standard curves – the cluster of differentiation gamma-3 molecule (CD3G) and 18S ribosomal rRNA. The cluster of differentiation gamma-3 is one of four peptides that form the CD3 complex that associates with a T – cell receptor to activate T lymphocytes and was chosen because of its stable and characteristic expression in blood (Flanagan, Wotton, Tuck-Wah *et al.*, 1990). Ribosomal 18S rRNA has a ubiquitous expression pattern and is present in all eukaryotic cells, making it suitable for testing with non-human samples. These genes for mRNA and rRNA were targeted because they are utilised in other assays described later in this thesis.

3.2 RNA quantification

As previously described, a number of methods are available for the measurement of mRNA levels including RNase protection assays, northern blots, *in situ* hybridisation, and real-time PCR. Real-time PCR is the only method that is specific and sensitive enough to measure low template amounts of RNA (<10 copies/reaction) through exponential amplification (Dallman, Montgomery, Larsen *et al.*, 1991). This allows for the detection of RNA from small amounts of originating material, mixed cell populations and rare RNAs (Schmittgen, Zakrajsek, Mills *et al.*, 2000). Other methods such as RiboGreen® RNA Quantitation (Lightfoot, 2002) require large amounts of RNA (up to 10 ng) which is difficult to obtain from forensic type samples.

During real-time PCR, forward and reverse primers hybridise to specific sequences of target cDNA with the probe hybridizing to a target sequence internal to these primers ("Real-Time PCR Vs. Traditional PCR,"). During the extension phase of the PCR cycle, the DNA polymerase enzyme cleaves the probe. A TaqMan probe, for example, is an oligonucleotide with a fluorescent reporter dye located at the 5’ end and a quencher located at the 3’ end. The

proximity of the reporter dye to the quencher dye in an intact probe, results in the inhibition of fluorescence. Since DNA polymerase cleaves from the 5' to the 3' end, the fluorescent reporter dye and quencher dye are separated, resulting in an increase of fluorescence. As a result, the amount of fluorescence is proportional to the amount of template present in the originating sample.

The threshold cycle (Ct) is the cycle number at which exponential amplification is detected; differences in the Ct value correspond to differences in the amount of template. All current qPCR detection systems use fluorescent technologies and there are a variety of different template detection methods (Table 3.1) (La Paz, Esteve, & Pla, 2007). SYBR® Green is the most commonly used dye for non-specific detection. It is a double-stranded DNA intercalating dye that fluoresces once bound to DNA. The main disadvantage of this technique is that the SYBR® Green will bind to any amplified double stranded DNA and is not suitable for qualitative qPCR.

One of the main advantages of qPCR is that it benefits the user because it has the ability to simultaneously amplify more than one target in a process called multiplexing. Detection of multiple targets in a single tube is possible because each probe is labelled with a different fluorescent reporter dye that absorbs different wavelengths.

Table 3.1. qPCR detection methods. All current qPCR detection systems use fluorescent technologies and there are a variety of different template detection methods with associated advantages and disadvantages.

Detection method	Advantages	Disadvantages
SYBR® Green	Low cost	Non specific
TaqMan® probes	Multiplex capabilities	Expensive
LNA® Double-Dye probes	High specificity and reproducibility	-
Molecular Beacon probes	Melting curve possible	-
Scorpions® primers	Compatible with any dye	-
Hybridisation probes	Machine specific	Machine specific
TaqMan® MGB® probes	High specificity	-
MGB Eclipse® probes	High specificity	-

3.3 Materials and Methods

3.3.1 Sample Collection

Healthy, consenting volunteers donated biological fluids. Samples were collected and processed during the same time of the day to minimize any circadian rhythm–based expression differences. All samples were left to dry at room temperature for 24 hours after which RNA was extracted from them. All purified RNA samples were tested to make certain that they were free of any genomic contamination, converted into cDNA and 1 µL of each was assayed unless specified.

3.3.2 Blood samples

Blood samples were freshly taken from five volunteers. Each volunteer gave 20, 10 and 5 µL of fresh liquid blood in duplicate that were deposited on sterile cotton cloth. All blood samples were processed as described in sections 2.2 to

2.5.

3.3.2 Specificity studies

Saliva and semen samples were obtained from two volunteers and these samples were processed as described in sections 2.2 to 2.5. Blood from various domestic animals (dog, chicken, pig, rat) were also obtained and these were processed as described in sections 2.2 to 2.5. Although the quantification method referenced in section 2.4 is a human specific assay, the animal samples were still quantified to determine if there was any human genomic contamination present that could otherwise skew the final qPCR result. Twenty microliters of each sample was deposited on sterile cotton cloth in duplicate.

3.3.3 Creation of the standard curve

The stock solution of leukocyte RNA (human blood peripheral leukocytes total RNA, 1 µg/µL (ClonTech, Mountain View, CA) used for creating the standard curve for cDNA quantification was tested for genomic DNA as referenced in section 2.4.

To create the standard curve for cDNA quantification, 1 µL of DNA-free leukocyte RNA was reverse transcribed into cDNA in a 20 µL reaction as described in section 2.5. The cDNA was either used immediately or stored at -20°C until required[1].

The standard curves were generated by amplifying the leukocyte cDNA with primers and probes for the CD3G mRNA and 18S rRNA markers. The leukocyte cDNA concentration was estimated nominally as 50 ng/µL (in a 20 µL cDNA reaction) – assuming that all of the RNA was reverse transcribed into cDNA.

[1] The cDNA was resynthesised after 3 freeze/thaw cycles as exploratory work showed an increase in the initial C_T value of each standard that also corresponded to a reduction in amplification efficiency.

The leukocyte cDNA was serially diluted by one third such that the starting concentration was 50 ng/µL and the lowest was 0.023 ng/µL creating 8 different concentrations (50, 16.7, 5.56, 1.85, 0.62, 0.21, 0.066, 0.023 ng/µL). One microlitre of each cDNA standard was added to the qPCR assay for the detection of the CD3G mRNA and 18S rRNA markers.

The performance of the standard curve was determined by running each standard in duplicate nine times to give a total of 18 values per standard. The C_{TFAM} (CD3G) and C_{TVIC} (18S rRNA) values for each standard were collated, and the mean and standard deviation C_T values were calculated.

3.3.4 Real-time quantitative PCR analysis (qPCR) of samples

Real-time qPCR reactions for the standards and unknown samples were all performed in 20 µL reaction volumes containing 1 µL of cDNA as described in Section 2.7.1.3.

3.3.5 Real-time PCR data analysis

The qPCR data was analysed using SDS software (v1.2; Applied Biosystems by Life Technologies, Carlsbad, CA). The fluorescence threshold was set in the middle of the exponential (linear) phase (phase 2) of the PCR with a manual baseline determined for each probe.

3.3.6 Sensitivity studies

Sensitivity studies were performed by quantifying 1 µL of cDNA obtained from extracting RNA from 20, 10 and 5 µL bloodstains in duplicate. A 1:10 and 1:100 dilution was also performed on the cDNA and 1 µL of each was quantified.

A dilution was performed on control cDNA (obtained from 20 µL of fresh blood that had been quantified by the assay) in the following series: 5.56, 1.85, 0.62, 0.21, 0.066 and 0.023 ng. One microlitre of each was tested with CellTyper in triplicate as described below.

3.3.7 Assessing the benefit of standardised reactions

Fleming and Harbison developed a multiplex PCR system (CellTyper) using mRNA that can identify blood, saliva, semen and menstrual blood in individual stains or in mixtures of body fluids (Fleming & Harbison, 2010). This multiplex assay was used to assess the effects of using a standardised cDNA concentration versus using a fixed volume of human blood cDNA.

Samples were run in duplicate according to the protocol established by Fleming and Harbison (Fleming & Harbison, 2010). The amount of blood cDNA added to CellTyper was adjusted to approximately 200 pg (as determined by the quantification assay). One microliter of cDNA (including a 1:10) dilution was also tested with CellTyper to compare the effect of adding set volumes to PCR reactions, instead of adjusting by cDNA amount.

3.4 Results and Discussion

3.4.1 Creation of the standard curve

The ability of the cDNA quantification assay to accurately determine the concentration of an unknown sample depends on how precise (or how variable) the C_{TFAM} and C_{TVIC} measurements are, as the unknowns are extrapolated from these values (Green, Roinestad, Boland *et al.*, 2005). The results of the C_T performance experiment are shown in Table 3.2.

The mean C_{TFAM} and C_{TVIC} results showed the expected inverse relationship with sample cDNA concentration ranging from 10.26 to 11.03 for the highest standard (50 ng) and 23.37 to 25.10 for the lowest standard (0.023 ng) respectively. In order to quantify a two-fold dilution in more than 99.7% of cases, the average standard deviation (SD) should be ≤ 0.167. In more than 95% of cases the SD should be ≤ 0.250 (Applied Biosystems, application note: Real-Time PCR) ("Real-time PCR: Understanding Ct,"). The average

SD for the 18 values of the 8 quantification standard dilutions was 0.173 for C_{TFAM} (CD3G) and 0.166 for C_{TVIC} (18S).

Table 3.2. Assessment of the utility of the quantification standards in the CD3G (FAM) mRNA and 18S rRNA (VIC) quantitation assay. Quantification standards were prepared from cDNA made from peripheral leukocyte RNA. The performance of the standard curve was determined by running each standard in duplicate nine times to give a total of 18 values per standard. The C_{TFAM} (CD3G) and C_{TVIC} (18S rRNA) values for each standard were collated, and the mean and standard deviation C_T values were calculated.

Estimate cDNA conc. of standards (ng/µL)	CD3G		18S		N
	MEAN C_T	SD† C_T	MEAN C_T	SD† C_T	
50.00	10.255	0.218	11.026	0.194	18
16.70	11.661	0.119	12.988	0.158	18
5.56	13.386	0.090	15.001	0.211	18
1.85	15.240	0.178	17.061	0.161	18
0.62	17.648	0.232	19.121	0.121	18
0.21	19.812	0.178	20.982	0.126	18
0.066	21.367	0.167	22.894	0.210	18
0.023	23.374	0.206	25.097	0.151	18
0.00	34.553	0.320	38.555	0.308	18

† SD = standard deviation

The co-efficient of variation (R^2) value between the standards ranged from 0.993 to 0.996 for CD3G and 0.991 to 0.999 for 18S. An R^2 value of > 0.99 provides good confidence in correlating two values (Applied Biosytems, application note: Real-Time PCR). The average efficiency (Efficiency = -1+10(-1/slope)) of the qPCR was between 93% and 96% based on an average slope of -3.5 and -3.4 for 18S and CD3G respectively. The recommended efficiency is from 90 to 110%. Threshold cycle values were observed for the standards with no cDNA template (34.553 for CD3G and 38.555 for 18S) due to background levels of fluorescence (approximately a 2^{11} and 2^{14} fold difference between 0.00 ng/µl and 0.023 ng/µl for CD3G and 18S respectively.

Table 3.2 shows that variability between replicates does exist (SD C_T). This could be attributed to pipetting small volumes of cDNA template (1 µl) used for analysis, the efficiency of the reverse transcriptase during the amplification steps and experimental variation caused by 7500 real-time machines.

Using 18S rRNA for the quantification of cDNA would be expected to have no effect on the downstream application of the assay (as used in the multiplex assays described in Chapter 4) as 1) the later assays are standardised on the input amount of CD3G cDNA and 2) different primers and amplicons are used in each multiplex assay.

In summary, the results demonstrate that the standard curves generated from the leukocyte cDNA using the CD3G mRNA and 18S rRNA primers and probes met the required criteria for efficiency (slope = ~3.3, $R^2 > 0.99$ that results in a 2-fold increase of the target for each cycle) and precision (≤ 0.167 SD – ≤ 0.250 SD). Reactions with no input cDNA and/or no reverse transcriptase all gave high C_T values (> 34 cycles) that were equivalent to background levels of detection and classified as “not detected” by the analysis software.

3.4.2 Quantification assay specificity

Although this quantification assay was not designed to be specific for human blood, the assay was tested with blood samples from other animals and with other human body fluids to determine the extent of specificity.

Of the two markers used, 18S rRNA is a ubiquitous housekeeping gene expressed in all eukaryotic cells and is not expected to be either human or tissue specific. The cluster of differentiation gamma-3 molecule has a mammalian expression pattern that is higher in whole blood and regions of immune cell development such as the thymus and lymph nodes. There is limited expression in other tissues (Flanagan *et al.*, 1990).

In assessing the specificity of the quantification assay, cDNA samples each of 20 µL derived from human and animal blood, human saliva and semen were tested. The results are summarised in Table 3.3.

All samples gave a C_T value, with much lower detection in the domestic animal species (dog, pig, chicken and rat) for both CD3G and 18S rRNA. Lower detection was interpreted as a high C_T result (the average C_T of duplicate reactions was 24 for dog, 21 for pig, 24 for chicken and 23 for rat compared to 14 for human blood). To put this C_T difference into perspective, 7 C_T unit would correspond to a difference in cDNA quantity of 2^7 (or 128-fold) when amplification is 100% efficient. This shows that the CD3G and 18S rRNA markers chosen for the assay are more sensitive for human blood than the other animal blood tested. This is important because in the forensic context, blood is often the most visible and abundant form of evidence. When the source is unknown it can be assayed against a control sample of human blood, the evidence could then be included or excluded based on a relatively high C_T value (the average C_T of 6 replicates of a 1:100 dilution of cDNA derived from 5 µL bloodstains was 24.4 with an average of 0.04 ng/µL compared to 22.9 and 0.04 for cDNA derived from 20 µL of animal blood).

When different human body fluids were assayed, there was lower detection of the markers in semen and saliva (the average C_T of duplicate reactions was 19 for semen and 22 for saliva), which is consistent with the low expression level of CD3G in other tissues.

Table 3.3. **Specificity of the quantification assay using different sources of human and non-human cDNA.** The specificity of the quantification assay was tested on human blood, saliva and semen samples as well as blood from various domestic animals (dog, chicken, pig, rat).

cDNA source	Marker	Average Ct value of replicates	Average amount (ng/µL)
20 µL human blood	**CD3G mRNA**	**14.39**	**4.58**
	18S rRNA	15.51	4.38
10 µL human blood	**CD3G mRNA**	**15.08**	**1.63**
	18S rRNA	17.04	1.54
5 µL human blood	**CD3G mRNA**	**16.38**	**0.79**
	18S rRNA	18.10	0.88
20 µL dog blood	**CD3G mRNA**	**23.81**	**0.02**
	18S rRNA	24.62	0.03
20 µL pig blood	**CD3G mRNA**	**20.98**	**0.10**
	18S rRNA	21.70	0.13
20 µL chicken blood	**CD3G mRNA**	**23.88**	**0.02**
	18S rRNA	24.95	0.03
20 µL rat blood	**CD3G mRNA**	**22.93**	**0.03**
	18S rRNA	23.93	0.05
20 µL human saliva	**CD3G mRNA**	**21.67**	**0.02**
	18S rRNA	23.38	0.02
20 µL human semen	**CD3G mRNA**	**18.57**	**0.12**
	18S rRNA	19.83	0.17

3.4.3 Quantification assay sensitivity

Real-time quantitative PCR analysis (qPCR) of samples

Assessment of the amplification performance with a range of cDNA concentrations would be helpful in understanding the potential limitations of the qPCR system; particularly where the recovery of RNA/cDNA is expected to be low.

In order to achieve this, a range of cDNA templates derived from 20, 10 and 5 µL bloodstains were quantified. One microliter of template was added to the quantification assay in duplicate along with 1 µL of a 1:10 and 1:100 dilution of the template cDNA. The results are shown in Figure 3.2.

Figure 3.2 indicates that when a range of template cDNA concentrations is quantified (17 to 0 ng), all of the data points are located along a straight line (R^2> 0.99). This shows that not only is the assay sensitive to low amounts of cDNA (starting material) it is also very accurate at the lower levels of detection where linear amplification still occurs.

Applications of the quantification assay

The utility of the quantitative assay was assessed by comparing the results obtained from different bloodstain volumes (10 and 5 µL) when either a standard volume of cDNA or a standard concentration of cDNA (based on the qPCR result) was added to CellTyper for body fluid identification (Fleming & Harbison, 2010). One microliter of each and the equivalent of 200 pg was added to CellTyper and analysed in duplicate. The blood specific marker in the CellTyper multiplex is glycophorin A (Glyco A). Figure 3.3 shows a typical electropherogram when 500 pg of CD3G cDNA derived from human blood is assayed with CellTyper. A summary of Glyco A peak heights for 1 µL (including a 1:10 dilution) and 200 pg are shown in Table 3.4.

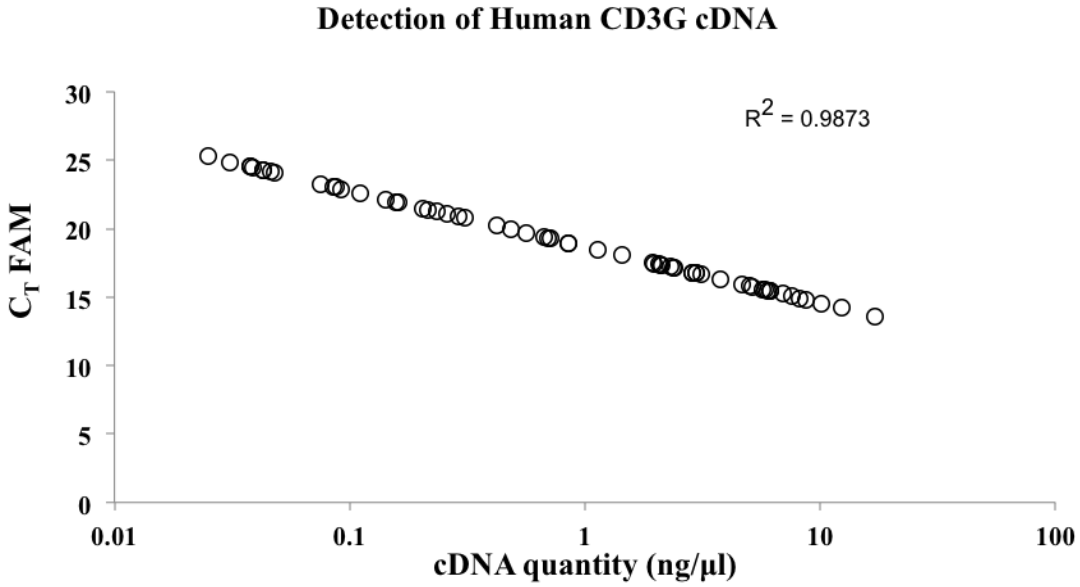


Figure 3.2. Sensitivity of the cDNA quantification assay. Different volume bloodstains were extracted (5, 10 and 20 µL) and converted into cDNA. The cDNA (and 1:10 and 1:100 dilutions of each) were quantified and the resulting C_T FAM values were plotted against cDNA concentration.

By using a fixed volume of cDNA (1 µL), the Glyco A peaks were over-amplified (above 5000 RFU) for both the 5 and 10 µL bloodstains. The results could not be statistically evaluated because the actual peak heights are difficult to determine when they are above 7000 RFU (software limitation). However, by diluting the cDNA samples 1:10, the average peak heights were able to be determined. The SD was 1256.85 and 253.71 RFU for the 10 and 5 µL stains respectively.

No over-amplification was observed when the cDNA concentration was determined and a fixed concentration of CD3G cDNA was added to the assay; this was irrespective of the bloodstain volume. The SD for the standardised samples was much lower at 122.52 RFU, with all samples falling within a much tighter RFU range (1015 to 1443). The noted difference in the cDNA yield between the 5 and 10 µL stains could be attributed to the presence of cellular debris carried over from the lysis solution that interfered with the capacity of the RNA to bind to the purification columns.

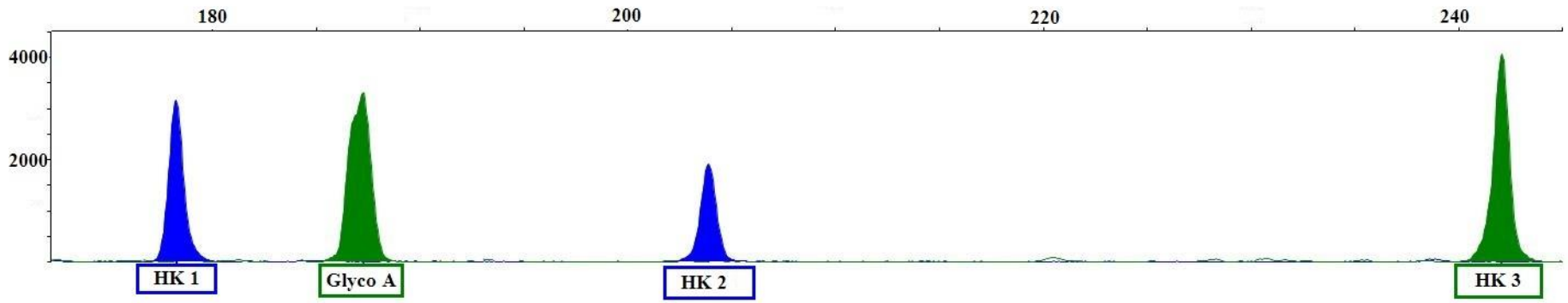


Figure 3.3. CellTyper Electropherogram. Electropherogram of the body fluid multiplex for the detection of a blood specific marker – Glycophorin A (Glyco A) when 500 pg of CD3G cDNA is added to the reaction. HK1 – glucose 6-phosphate dehydrogenase, HK2 – transcription elongation factor and HK3 – ubiquitin conjugating enzyme, these are housekeeping genes in the CellTyper multiplex.

Table 3.4. **Application of the CD3G quantification assay.** Comparison studies between fixed cDNA volume (A) and standardised input cDNA concentration (B) in a CellTyper reaction.

A)

	Blood stain volume (µL)	Final CD3G cDNA concentration	1:10 CD3G cDNA concentration	Volume of cDNA added to CellTyper (µL)	Height of Glyco A peak (final / 1:10) (RFU)
Donor 1	10	1.95	0.22	1	6698 / 1704
	10	3.11	0.14	1	8009 / 574
	5	1.43	0.42	1	7907 / 896
	5	2.12	0.31	1	7622 / 1493
Donor 2	10	6.14	0.49	1	5714 / 3318
	10	5.02	0.70	1	5350/ 2982
	5	2.84	0.26	1	6404 / 1304
	5	2.84	0.29	1	5885 / 1327
Negative control	-	0.00	0.00	1	N/A

B)

	Blood stain volume (µL)	1:10 final CD3G cDNA concentration	Amount of cDNA added to Celltyper (pg)	Height of Glyco A peak (RFU)
Donor 1	10	0.27	200	1443
	10	0.14	200	1245
	5	0.42	200	1221
	5	0.31	200	1259
Donor 2	10	0.49	200	1183
	10	0.69	200	1191
	5	0.26	200	1015
	5	0.29	200	1121
Negative control	-	0.00	1 µL	N/A

Figure 3.4 shows that the assay can also be used to determine the optimum concentration of cDNA (CD3G, or “blood”) based on the amplification of a certain target (Glyco A in this instance). When increasing amounts of CD3G cDNA template are added to CellTyper, the assay is saturated (the Glyco A peak is over-amplified) when the amount of CD3G cDNA exceeds 600 pg. Where samples were not over-amplified (i.e. < 600 pg), the average standard deviation was 141.85 RFU, which is similar to those samples standardised to 200 pg (122.52 RFU). Lower amounts of CD3G cDNA (≤0.023 ng) were below the detection threshold required to conclude a positive result for Glyco A (100 RFU). This is not unexpected as qPCR is more sensitive and efficient than end-point PCR ("Real-Time PCR Vs. Traditional PCR,") due to the small amplicon size (64 bp for CD3G versus 188 bp for Glyco A) and the increased number of cycles (35 cycles for CellTyper versus 40 for qPCR).

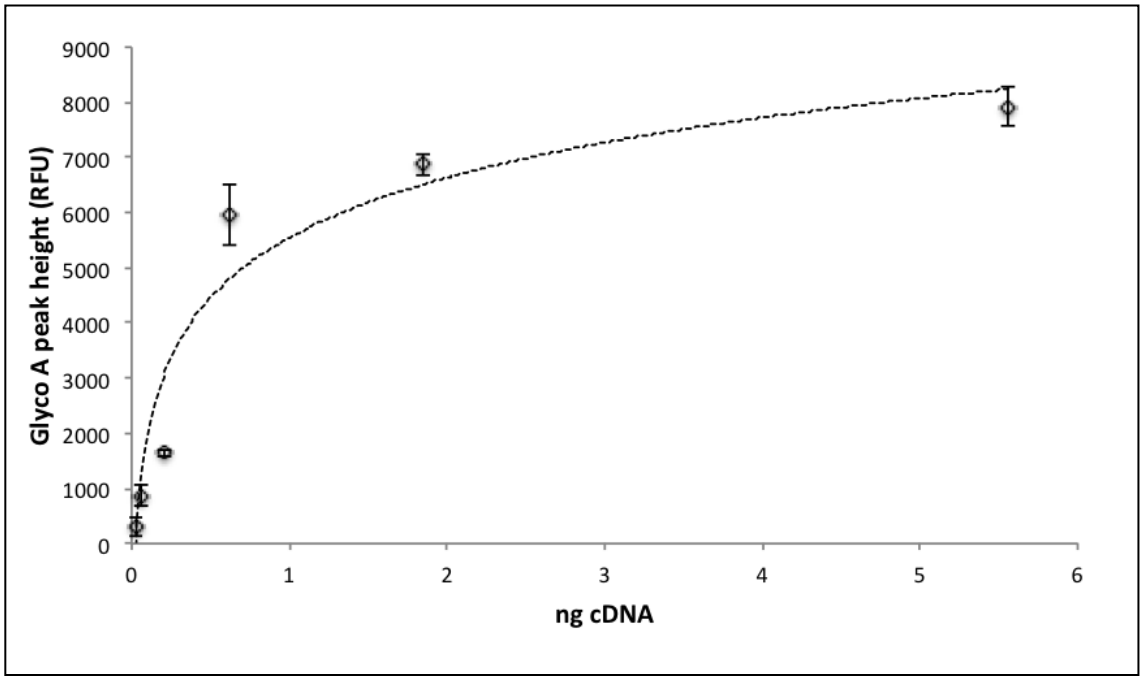


Figure 3.4. **Sensitivity of the cDNA quantification assay in relation to the detection of Glyco A in the CellTyper assay.** Control cDNA from a 20 µL bloodstain was diluted to 5.56, 1.85, 0.62, 0.21, 0.066 and 0.023 ng. One microliter of each was added to CellTyper (in triplicate) and the resulting Glyco A peak height was plotted against the amount of cDNA. The error bars are indicative of the standard deviation between triplicates.

3.5 Conlusion

Ribonucleic acid has emerged as an important forensic tool in recent years. The development of sensitive, reliable and accurate method to quantify RNA would aid in the development of novel probative techniques. Real-time qPCR can be used to establish the presence and quantity of human-specific mRNA in a given sample. Previous forensic studies (Bauer & Patzelt, 2008; Juusola & Ballantyne, 2003) have based the quantity of RNA in a sample on the results obtained using a GAPDH amplicon. This could have a detrimental effect, as the amplicon could be too abundant for assessing samples that are essentially low in concentration or prone to degradation. This would likely result in increasing the occurrence of the false-negative results and would likely prevent a sample from being included for subsequent analysis.

Due to the uncertain nature of biological stains deposited at crime scenes, it is important that enough biological material is recovered to allow analysis to proceed with analysis. The advantage of using a cDNA quantification system is that a PCR assay can be optimised based on the input concentration of cDNA and allows relative comparisons between different samples to be made. There have been instances where comparisons have been made between two different markers based on the relative expression of each (Anderson *et al.*, 2005; Hampson *et al.*, 2011). These methods have failed to standardise the concentration of cDNA added to each subsequent PCR. Each PCR assay was volume based and in some instances, the amount of cDNA has been estimated visually (Bauer, Polzin*, et al.*, 2003); making it difficult to make true comparisons between different samples.

The quantification assay developed here is more sensitive to the quantification of human blood cDNA than cDNA obtained from the blood of other animals or human body fluids. However, if the aim were to ascertain if the source is human, a human specific assay would need to be conducted (Hudlow, Chong, Swango *et al.*, 2008; Kanthaswamy, Premasuthan, Ng *et al.*, 2012). This is not applicable for the purpose of this thesis as all samples are known to be either of human or animal origin; but it may be an issue with

evidence collected from crime scenes. Based on the detection threshold of the qPCR assay, it can be used on samples where the amount of blood recovered is low (blood splatter for example) or where the storage conditions are unknown (i.e. samples are likely to be degraded).

The development of this quantification assay was necessary as a method to standardise samples required for fragment degradation and statistical analysis where known amounts of cDNA are required. Work involving the presence or absence of a bio-marker is more qualitative and would not be affected by over amplification, even when comparing different samples. The work involved in developing such a method is qualitative and it is important to know whether degradation is occurring as a result of storage conditions for example, and not a result of inadvertently limiting the amount of cDNA in the reaction.

Chapter 4

Design and Development of Novel Multiplex PCR Assays to Assess the Degradation of mRNA, rRNA and DNA in Bloodstains

4.1 Introduction

As discussed in Chapter 1, previous attempts have been made to determine the age of a bloodstain. However, there are no viable or practical biological methods for forensic use that are capable of determining the time since deposition of a bloodstain. Methods that measure DNA degradation (Alaeddini *et al.*, 2010) or the amount of damage to DNA (Swango *et al.*, 2006) or RNA in dried bloodstains have so far been able to establish a limited correlation with age (Bauer, Polzin*, et al.*, 2003). The limited correlation equates to inaccurate deposition estimates, and is most likely the result of not taking into account the way in which degradation occurs in bloodstains as they dry and age.

This chapter has attempted to characterise the degradation pattern of different nucleic acids as a bloodstain dries, in order to assess whether it is possible to use the degradation of nucleic acids to determine the time since blood deposition.

This study investigated which regions of the nucleic acid sequence were more susceptible to degradation and which regions were more stable by using capillary electrophoresis to detect amplified hemoglobin beta gene (HBB), 18S rRNA and glucose-6-phosphate dehydrogenase (G6PD) mRNA fragments. Fragment analysis of each region was performed over a period of 24 hours and the effects of ultra-violet (UV) light, drying and temperature were observed.

By observing changes in peak height of various amplified fragments, the pattern of degradation for DNA, rRNA and mRNA was investigated to determine if a consistent pattern could be established. Multiplex assays were developed around the HBB DNA, 18S rRNA and G6PD mRNA sequences to investigate the directionality of degradation as the bloodstain dried. Each

assay was designed to amplify, through PCR, various sized fragments from the 5', 3' and central portions of the sequence that were detected and analysed through capillary electrophoresis.

4.2 Materials and Methods

4.2.1 Blood samples

Blood samples were obtained from three female and two male volunteers by means of a finger prick. Twenty microliter samples were deposited onto sterile cotton cloth in duplicate and left under the following conditions: room temperature (20 minutes to 24 hours), 37°C and under a UV light source (20 to 60 minutes). The bloodstains were extracted immediately after each time point. Control bloodstains were prepared from 20 µl of fresh blood that was extracted and processed straight away. Blank pieces of sterile cotton cloth were extracted and processed at the same time as negative controls. All bloodstains were extracted and processed as described in sections 2.2 to 2.5. The amount of DNA recovered from each bloodstain was determined by the method described in section 2.4. The amount of RNA that had been converted into cDNA was determined by the method described in section 3.3.4.

4.2.2 UV exposure

Bloodstains were exposed to UV irradiation using a UV lamp (Osram HNS 15 OFR). According to the manufacturer, the UV lamp's spectrum is composed of two main wavelengths at 254 and 185 nm with a radiated power of 0.3 Jcm^{-2}. The bloodstains were placed 20 cm below the lamp and irradiated for 20 to 60 minutes.

4.2.3 Drying at room temperature and exposure to heat

Bloodstains were incubated at room temperature for 20 to 60 minutes and 24 hours and at 37°C for 20 to 60 minutes. The bloodstains were completely dry to the touch after 30 minutes at room temperature and 20 minutes at 37°C.

4.2.4 Exposure of extracted nucleic acids

Extracted nucleic acids were analysed to determine if degradation differed in the absence of the local cellular environment.

Ten volumes of 20 µl liquid blood samples were obtained from volunteers and were immediately extracted as described in sections 2.2 to 2.3. Ten microliters of the extracted DNA and purified RNA were deposited into sterile 200 µl PCR amplification tubes in duplicate that were then placed on their side with the lids open to allow the extracts to dry under the conditions studied.

The extracts were incubated at room temperature for 20 to 60 minutes and 24 hours (n=8) and at 37°C for 20 to 60 minutes (n=6). The extracts were also exposed to UV for 20 to 60 minutes (n=6). If the extracts were dry at the conclusion of each time point, they were resuspended in 20 µl of sterile water.

The RNA extracts were then converted into cDNA as described in section 2.5. Both the cDNA and DNA templates were quantified as described in sections 2.4 and 2.7.1. Duplicate samples from each condition/time point were analysed. One nanogram of DNA was added to the HBB DNA, 600 pg of cDNA to the G6PD mRNA assay and 200 pg of cDNA to the 18S rRNA assay.

4.2.5 Volume study

In order to assess the difference in each assay when a fixed volume of template was used, 2 µl of DNA or cDNA was added to each assay instead of a standardised concentration. Each multiplex assay was performed in a 25 µl reaction volume that included 12.5 µl of Qiagen multiplex PCR master mix,

2.5 µl of the 10x primer stock and 8 µl of sterile water. The thermal cycling conditions for each assay were unchanged and are described in Section 4.2.6.

4.2.6 Multiplex assays

4.2.6.1 Multiplex PCR assay design

Sequences for design work (HBB: NG_000007.3 from bases 70545 to 72150, 18S rRNA: NR_003286.2, G6PD: NM_000402.3) were downloaded from the GenBank resource at the National Centre for Biotechnology Information (NCBI) website. Primers for the multiplex were designed with Primer3 (Rozen & Skaletsky) with the following considerations in mind: to be 18-30 oligo nucleotides in length as the longer end of this range allows higher specificity (Wu, Ugozzoli, Pal *et al.*, 1991). The melting temperature (Tm) of the primers used were not more than 5°C different from each other. A Tm between 62 and 67°C for each primer over the region of hybridisation. The guanine-cytosine (GC) content of each primer was in the 40-60% range for optimum PCR efficiency. Potential primers were also checked to make sure that they were not self-complementary or complementary to the other primers in the reaction mixture, as this will encourage the formation of hairpins and primer dimers and will compete with the template for the use of primer and reagent. The nucleotide sequences and primer alignments are shown in Appendices 1 to 3.

The primer sequences and optimised primer concentrations are given in Table 4.1. These primers were checked for specificity using the basic local alignment search tool (BLAST) on the NCBI website ("Basic Local Alignment search tool,").

Primers were designed to amplify short (60-78 bp), medium (149-189 bp) and large fragments (254-401 bp) from each of the 5', central and 3' portions of the sequence (Figure 4.1). Common primers were designed to be away from the 3' ends under the hypothesis that degradation is from the 3' end meaning that the 3' products would disappear first followed by the central products and

then the 5' product. This approach was taken to investigate the directionality of degradation and to determine whether it was from the 5' end, the 3' end or random.

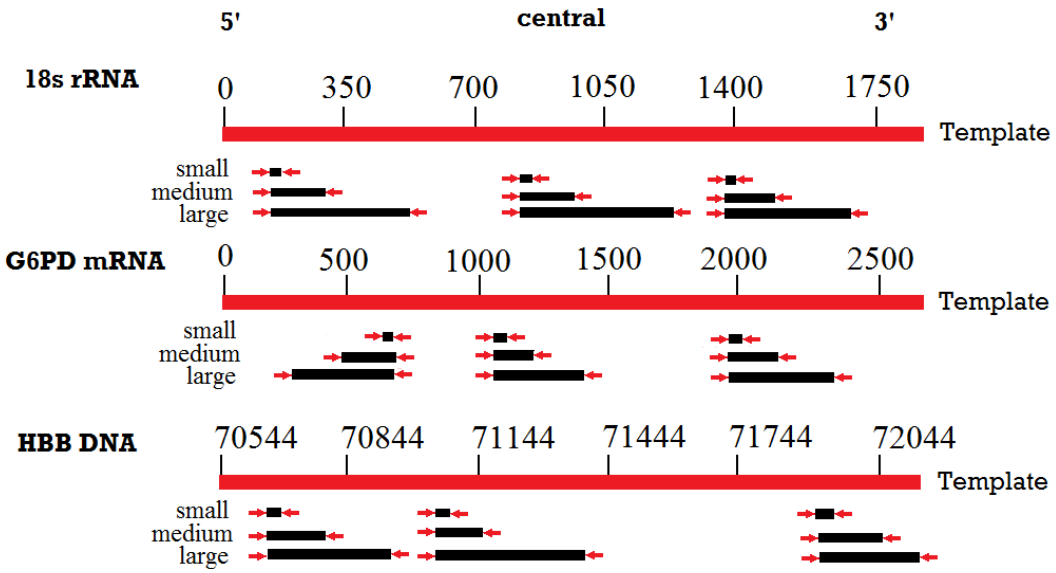


Figure 4.1 Multiplex assay design to investigate the directionality of degradation. Primers were designed to amplify fragments of different sizes from the 5', central and 3' regions of the nucleic acid sequence. The schematic representation shows the position of the fragments. Small fragments were 60 to 80 bp in length, medium fragments were 149 to 183 bp in length and the large fragments were 300 to 400 bp in length.

4.2.6.2 G6PD mRNA and 18S rRNA multiplex assay

Six hundred picograms of CD3G cDNA was added to the G6PD assay and 200 pg of CD3G cDNA was added to the 18S rRNA assay and amplified in a total reaction volume of 25 µl. The difference in input amount was because of the higher amount of rRNA compared to mRNA in a typical cell; rRNA constitutes 80 – 90 % of the total RNA (Lindberg & Lundeberg, 2010). The Qiagen multiplex PCR kit (Qiagen, Germantown, MD) was used for each assay. The reaction mix contained 12.5 µl of master mix, 2.5 µl of 10x primer stock (Table 4.1) and made up to 25 µl with sterile water.

	Forward primer	Final conc. (μM)	Reverse primer	Final concentration (μM)	Expected size (bp)
18S rRNA	(5') NED- GTCGCTCGCTCCTCTCCTACT	0.05	CCCGTCGGCATGTATTAGCT (5S) AGTCACCAAAGCCGCCG (5M) TTTTTCGTCACTACCTCCCCG (5L)	0.006 0.0125 0.0125	63 172 401
	(C) FAM-TTAGAGTGTTCAAAGCAGGCC	0.05	CGCGGTCCTATTCCATTATT (CS) GCCGGTCCAAGAATTTCAC (CM) CCCTTCCGTCAATTCCTTT (CL)	0.4 0.02 0.02	67 164 399
	(3') VIC-CGTTCTTAGTTGGTGGAGCGAT	0.05	TTAGCATGCCAGAGTCTCGTT (3S) AGACCTGTTATTGCTCAATCTCG (3M) GCTTATGACCCGCACTTACTGG (3L)	0.1 0.05 0.05	66 158 330
HBB DNA	(5') VIC-GCCCTGGGCAGGTTGGTATC	0.05	ACATGCCCAGTTTCTATTGGTCTCCT (5S) CCAAGGGTAGACCACCAGCAGC (5M) GGTGCCCTTGAGGTTGTCCAGGTG (5L)	0.15 0.25 0.015	66 164 304
	(C) FAM-AGCTGCACGTGGATCCTGAGA	0.05	AAGAAGGGGAAAGAAAACATCAAGCGT (CS) TGCAATCATTCGTCTGTTTCCCA (CM) TGCAAATAAGCACACATATATTCCAA (CL)	0.15 0.025 0.15	69 149 343
	(3')VIC- ACCTCTTATCTTCCTCCCACAGCTCC	0.05	TGATGGGCCAGCACACAGACC (3S) TAGAAATTGGACAGCAAGAAAGCGAGC (3M) AATCCAGATGCTCAAGGCCCTTCATA (3L)	0.025 0.25 0.15	60 178 252
G6PD mRNA	AACACCTTCATCGTGGGCTA (5S) GCCTTCCATCAGTCGGATAC (5M) CCGGAAACGGTCGTACACT (5L)	0.025 0.015 0.8	GAAGGGCTCACTCTGTTTGC-NED (5')	0.1	72 189 329
	(C) FAM-TCTACCGCATCGACCACTACC	0.24	CGAAGATCCTGTTGGCAAAT (CS) GGATGATCCCAAATTCATCG (CM) TCGTCCAGGTACCCTTTGGT (CL)	0.018 0.1 1.6	78 183 388
	(3') VIC-GGCAGACGAGCTGATGAAGA	0.2	AGAGCTTGTGGGGGTTCAC (3S) GAATGTGCAGCTGAGGTCAA (3M) GGAGTGAGTGGAGGAGGTGA (3L)	0.4 0.1 0.8	74 177 331

Table 4.1. Multiplex assay design. Primer sequences used in each multiplex assay for 18S rRNA, HBB DNA and G6PD mRNA. (5') – 5' region, (C) – central region, (3') – 3' region, (5S) – 5' small fragment (5M) – 5' medium fragment, (5L) – 5' large fragment, (CS) – central small fragment (CM) – central medium fragment, (CL) – central large fragment, (3S) – 3' small fragment, (3M) – 3' medium fragment and (3L) – 3' large fragment.

The following PCR thermal cycling conditions were used: 95°C for 15 minutes, 94°C for 30 seconds, 58°C for 30 seconds and 72°C for 90 seconds for a total of 30 cycles and a final elongation step of 72°C for 10 minutes. Prior to capillary electrophoresis analysis, the amplified 18S products were diluted 1:100 in sterile water to assist with analysis. This was a necessary step due to over amplification.

Amplification products were separated on a 3130 genetic analyser (Life Technologies, Carlsbad, CA). One microliter of template was loaded with Hi-Di™ Formamide, POP-4 polymer was used as the separation medium and Multi-capillary DS-30 (Dye set D) was used as the matrix standard. GeneScan™ 500 LIZ was used as the internal size standard and results were analysed using Peak Scanner (Life Technologies, Carlsbad, CA) version 1.0. No template controls were included and all samples were amplified in duplicate.

4.2.6.3 HBB DNA assay

The multiplex assay was performed in a 25 µl reaction volume that included 12.5 µl of Qiagen multiplex PCR master mix, 2.5 µl of 10x primer stock (Table 4.1) and 1 ng of DNA template. The volume of sterile water was adjusted accordingly.

The following thermal cycling conditions were used: 95°C for 15 minutes, 94°C for 30 seconds, 61°C for 30 seconds and 72°C for 90 seconds for a total of 30 cycles and a final elongation step of 72°C for 10 minutes. No template controls were included and all samples were amplified in duplicate.

4.2.7 Determining the direction of degradation of nucleic acids in dried bloodstains

In order to determine if degradation was occurring from the 5’ end, 3’ end or if it was random, analysis was performed on each replicate (n=20) and time point (n=4). The height of each amplified fragment was taken from the

capillary electrophoresis data and the average peak height was calculated for each fragment at each time point.

Comparisons between the averages of each fragment were made between the following time points: 0 – 20 minutes, 20 – 40 minutes, 40 – 60 minutes and 60 minutes – 24 hours (in the room temperature study). When the different time points were compared, the averaged peak heights were used to determine the direction of degradation. The most stable fragments would exhibit the lowest reduction in peak height, whilst the least stable fragments would exhibit the highest reduction in peak height. As the fragments were amplified from specific regions of the nucleic acid sequence, this would inform the direction of degradation.

4.3 Results and Discussion

In order to investigate the immediate pattern of degradation following bloodstain deposition, the usefulness of the short fragments were evaluated on 150 day old bloodstains as Bauer *et al.* reported that smaller sized fragments could be amplified after many years (Bauer, Polzin*, et al.*, 2003). Figure 4.2 shows that when small fragments from all three markers were analysed from fresh blood and 150 day old stains (as used in the extended depositions trials of Chapter 5), they were found to be stable. As there were negligible changes to these fragments over a period of 150 days (Table 4.2) they were removed from the multiplex. This observation was based on the average of 3 replicates of each marker at each time point (t= 0 and t=150 days). The multiplex PCR assays with the medium fragments (149 to 183 bp in length) and large fragments (300 to 400 bp in length) amplified are shown in Figure 4.3.

The differences in peak height (as measured by the change in relative fluorescent units; RFU) between each fragment under the different conditions are shown in Figure 4.4 (HBB DNA gene fragments), Figure 4.5 (18S rRNA fragments) and Figure 4.6 (G6PD mRNA fragments). The error bars show the average variation in RFU between replicates.

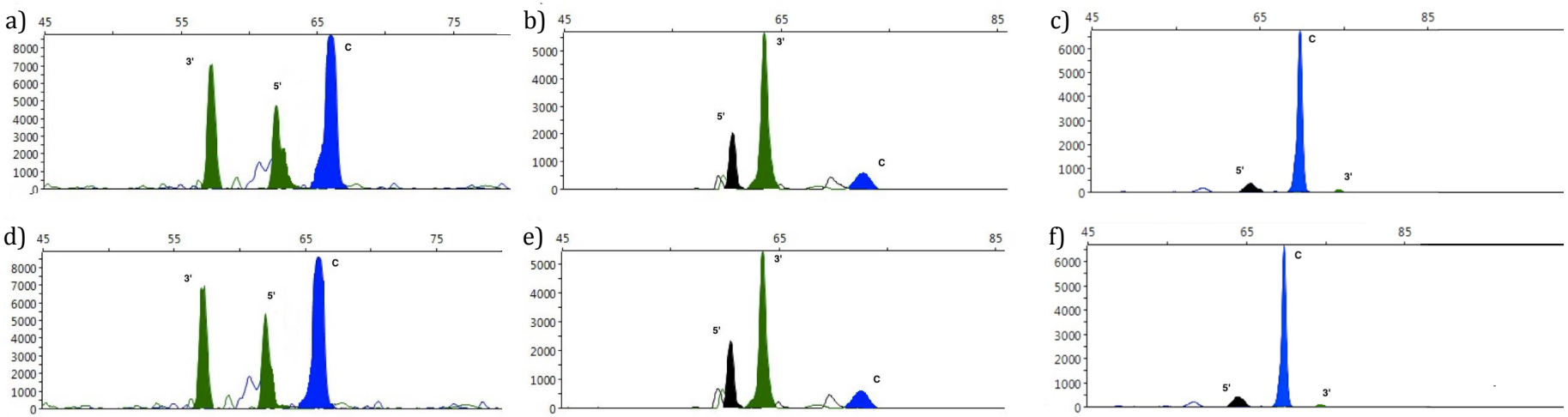


Figure 4.2. Stability of small nucleic acid fragments. The usefulness of the short fragments for degradation studies were evaluated in 150 day old bloodstains. a) HBB DNA small fragments at T=0, b) 18S rRNA small fragments at T=0, c) G6PD mRNA small fragments at T=0, d) HBB DNA small fragments at T=150 days, e) 18S rRNA small fragments at T=150 days, f) G6PD mRNA small fragments at T=150 days.

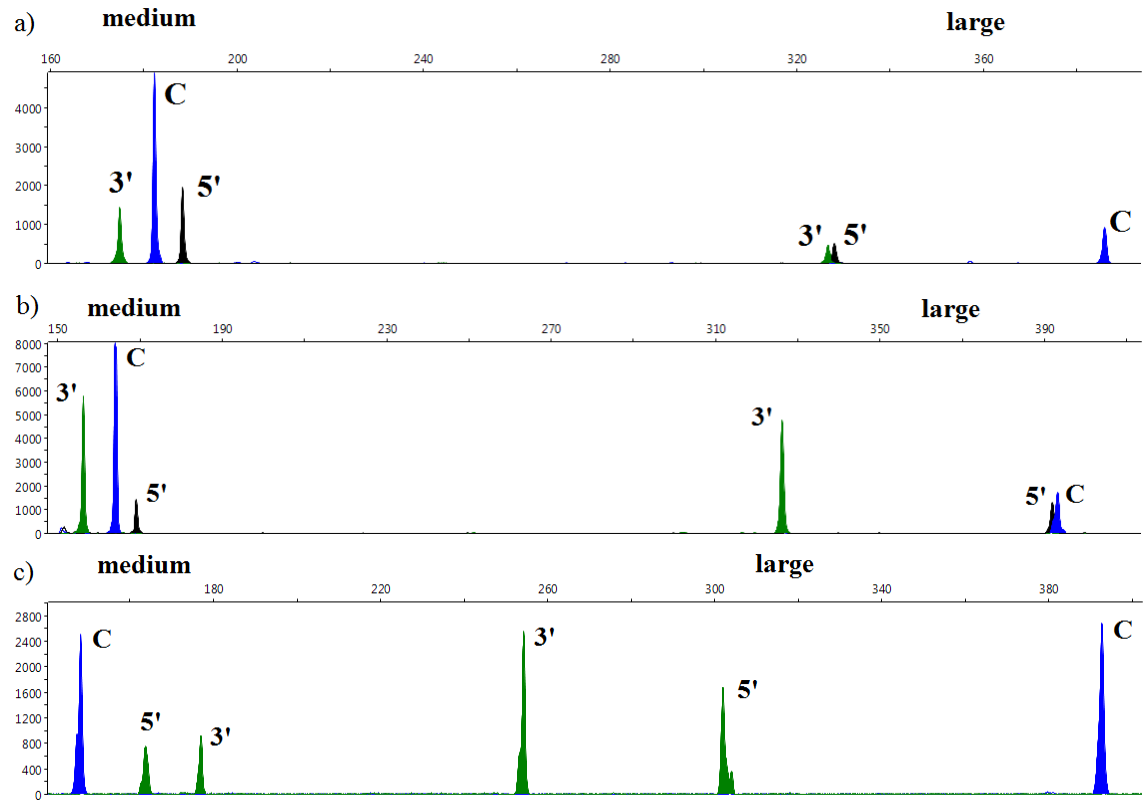


Figure 4.3. Multiplex assays for the detection of HBB, 18S and G6PD fragments. Electropherogram showing the amplified fragments from fresh blood (Time = 0) control for the G6PD mRNA (a), 18S rRNA (b) and HBB DNA (c) multiplex assays. 3' – from the 3' region of the sequence, C – from the central region of the sequence and 5' – from the 5' region of the sequence.

Previous studies have shown that longer exposure to the environment during drying affects the integrity of DNA. McNally et al. showed that bloodstains that take longer to dry undergo autolysis or, have bacteria and other contaminants present to a greater extent than stains that dry faster. Conversely, faster drying stains were shown to have a higher level of non-degraded DNA present (McNally, Shaler, Baird *et al.*, 1989). The latter scenario would be typical of a scene where, for example, a shooting has taken place and heavy bleeding has occurred.

Table 4.2. Stability of short nucleic acid fragments over a deposition time course. Short nucleic acid fragments (60 – 80 bp in length) designed to amplify regions of the hemoglobin beta gene (HBB), 18S ribosomal RNA (18S rRNA) and glucose-6-phosphate dehydrogenase mRNA (G6PD mRNA) were amplified in triplicate in 150 day old bloodstains to evaluate their potential use in determining the directionality of nucleic acid degradation in drying and dried bloodstains. The results show the average RFU and standard deviation of 3 replicates.

Fragment	Peak height (RFU) at t = 0 and (SD)	Peak height (RFU) at t = 150 days and (SD)
HBB DNA central small	8710 (78)	8581 (45)
HBB DNA 3' small	7061 (486)	6956 (336)
HBB DNA 5' small	3736 (382)	3932 (337)
18S rRNA 3' small	5661 (536)	5432 (572)
18S rRNA central small	583 (107)	588 (55)
18S rRNA 5' small	2042 (491)	2324 (299)
G6PD mRNA 3' small	366 (154)	386 (195)
G6PD mRNA central small	6748 (30)	6657 (40)
G6PD mRNA 5' small	103 (57)	99 (40)

Figure 4.4 shows the degradation pattern of DNA fragments of the HBB gene sequence in bloodstains under the environmental conditions studied: A) room temperature, B) 37°C and C) UV treatment, over a period of 20 minutes to 24 hours when the multiplex PCR assay was standardised on 1 ng of DNA template. The results are presented as a decrease in average peak height for each fragment, observed in each replicate at the indicated time point. Where no decrease was observed, no result appears on the histogram (ΔRFU = 0).

For example, Figure 4.4A shows that there was no observed change for the central medium fragment (CM) between 0-20 minutes and between 60 minutes-24 hours, the largest decrease was between 20-40 minutes.

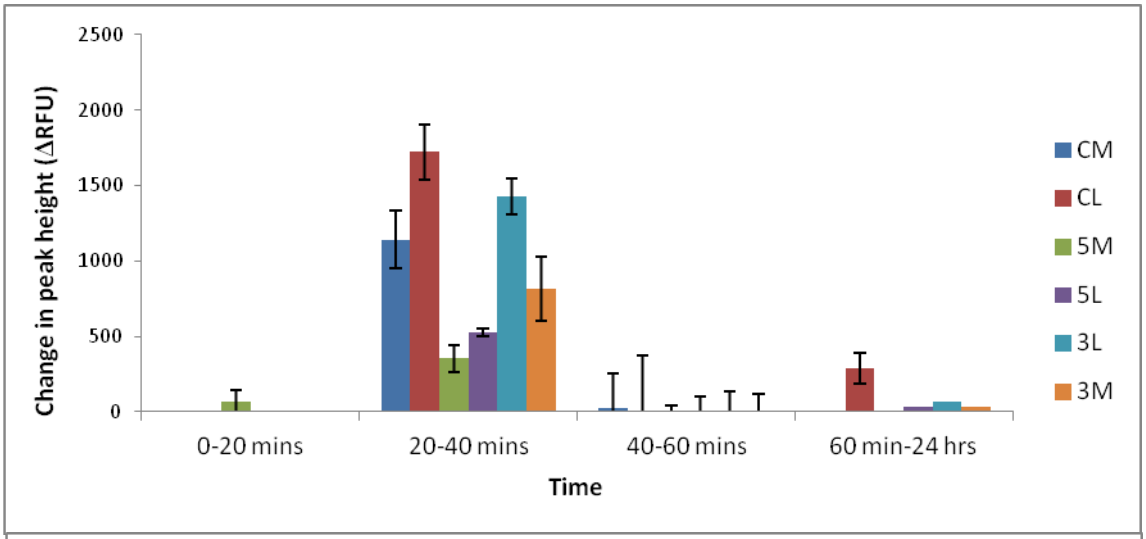
The HBB DNA fragments were less susceptible to degradation (lowest reduction in RFU) when the bloodstains were dried and left at 37°C (Figure 4.4B) compared to leaving them to dry at room temperature (Figure 4.4A) or drying whilst exposed to UV (Figure 4.4C). These results are consistent with the findings of Madisen *et al.* (Madisen, Hoar, & Holroyd, 1987) who reported that that high molecular weight DNA could be extracted from samples incubated for up to 26 weeks at 37°C.

When bloodstains were left to dry at room temperature (Figure 4.4A), more degradation (an increase in the reduction of RFU) occurred earlier on in the drying process (the 20-40 minute time frame), with little change occurring after the longer time points.

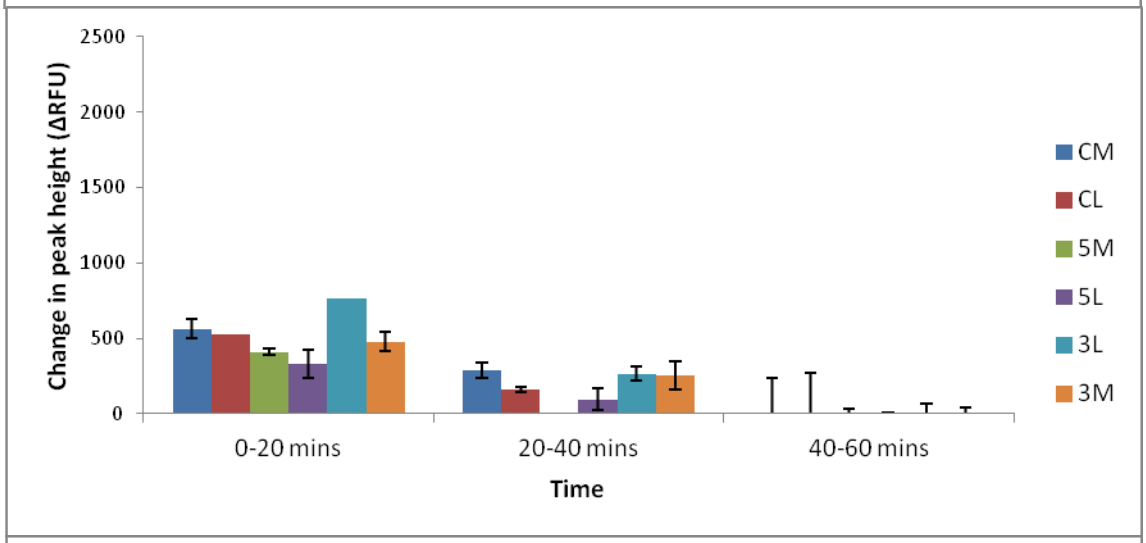
When exposed to UV light (Figure 4.4C), the greatest decrease in peak height for all fragments occurred in the first 20 minutes, with negligible change thereafter.

By means of the lowest average reduction in fragment peak height, the results show that the 5' region appears to be the most stable under the conditions studied.

a)



b)



c)

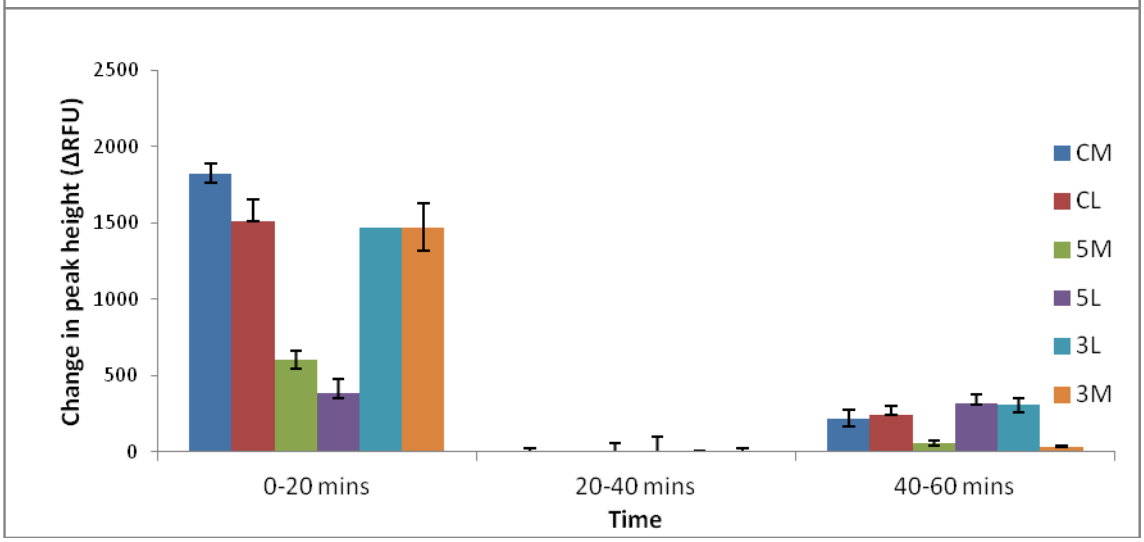


Figure 4.4. Effect of drying on the HBB DNA fragments. Decrease in fragment height a) bloodstain dried at room temperature, between T=0 and 24 hours; b) bloodstains dried at 37°C between T=0 and 60 Minutes; c) Bloodstain dried under UV light between T=0 and 60 minutes. CM-central medium fragment, CL- central long fragment, 5M-5' medium fragment, 5L 5' long fragment, 3L 3' long fragment, 3M 3' medium fragment.

For the 18S rRNA fragments assayed from bloodstains (Figure 4.5), the greatest decrease and thus the greatest amount of degradation was most apparent when the bloodstains were left to dry at 37°C (Figure 4.5B) or under UV light (Figure 4.5C) compared to drying at room temperature. Like the DNA fragments, more degradation occurred earlier in the degradation process than later (0-20 minutes at 37°C and 20-40 minutes under UV). The results show that the central region of the 18S rRNA sequence appears to be the most stable under the conditions studied as indicated by the negligible change in peak height across all samples in this region (Figure 4.5: CM and CL fragments).

The results for the G6PD mRNA fragments (Figure 4.6) show that mRNA was more stable when dried at 37°C (Figure 4.6B) or under UV light (Figure 4.6C) compared to those analysed from bloodstains dried at room temperature (Figure 4.6A). Unlike DNA and rRNA, mRNA did not appear to be affected by UV exposure when drying. On average, each fragment was 450 RFU higher than the control samples of fresh blood that were extracted and processed immediately. This is in line with the findings of Gowrishankar *et al.* (Gowrishankar, Winzen, Bollig *et al.*, 2005) as they have shown that degradation of mRNA in cells that were exposed to UV-B light was inhibited and the removal of the Poly-A tail was reduced. The increase in the average peak height could be attributed to the up-regulation of proteins derived from normally unstable mRNAs, such as inflammatory cytokines, before the bloodstain has dried. Like DNA, the 5' region appears to be the most stable.

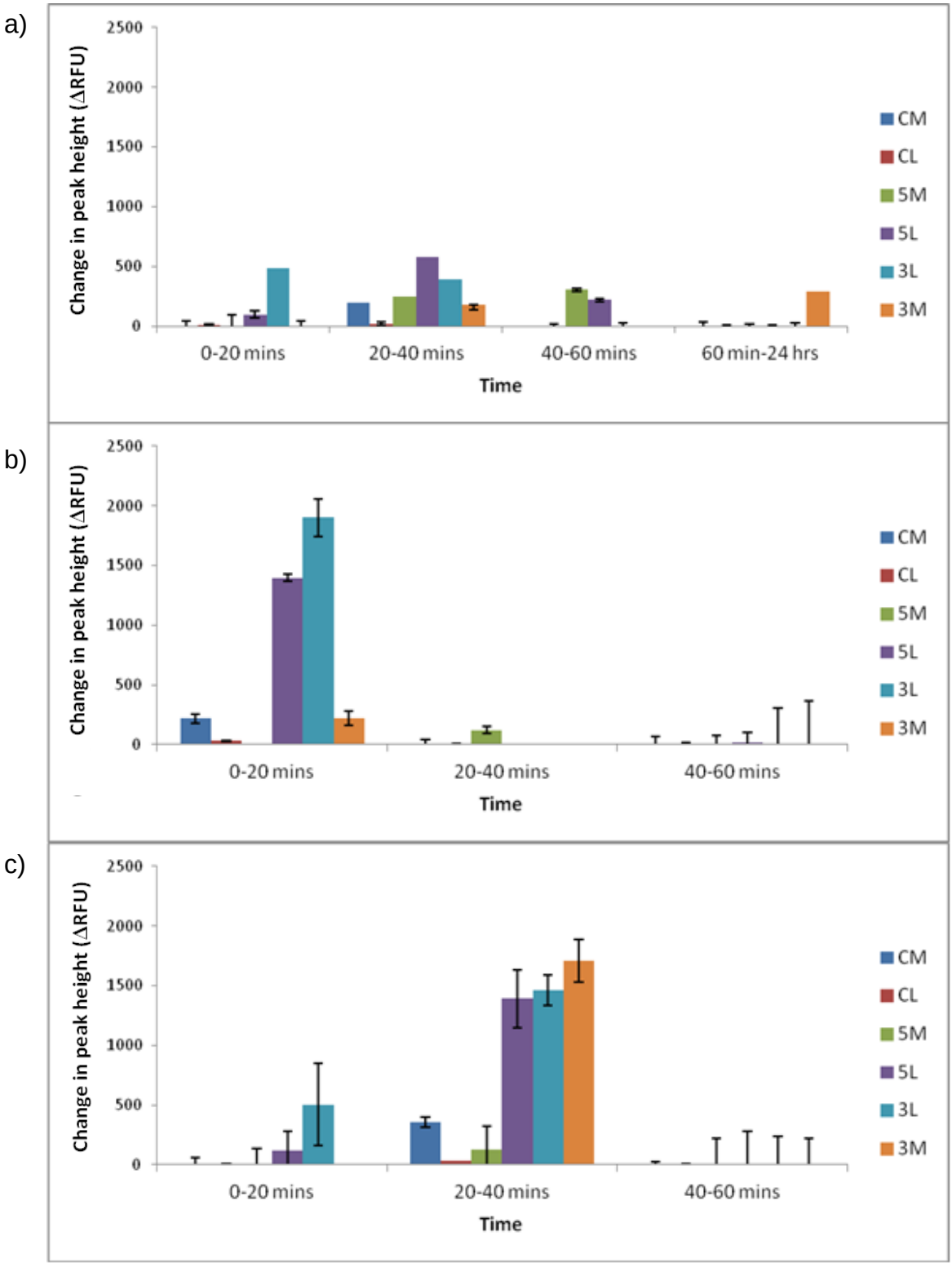
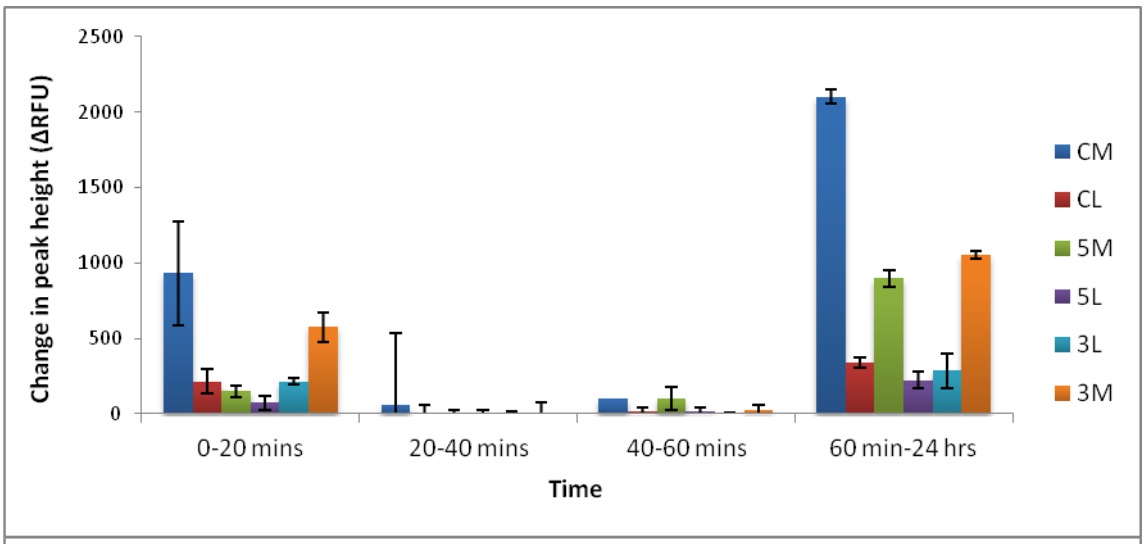
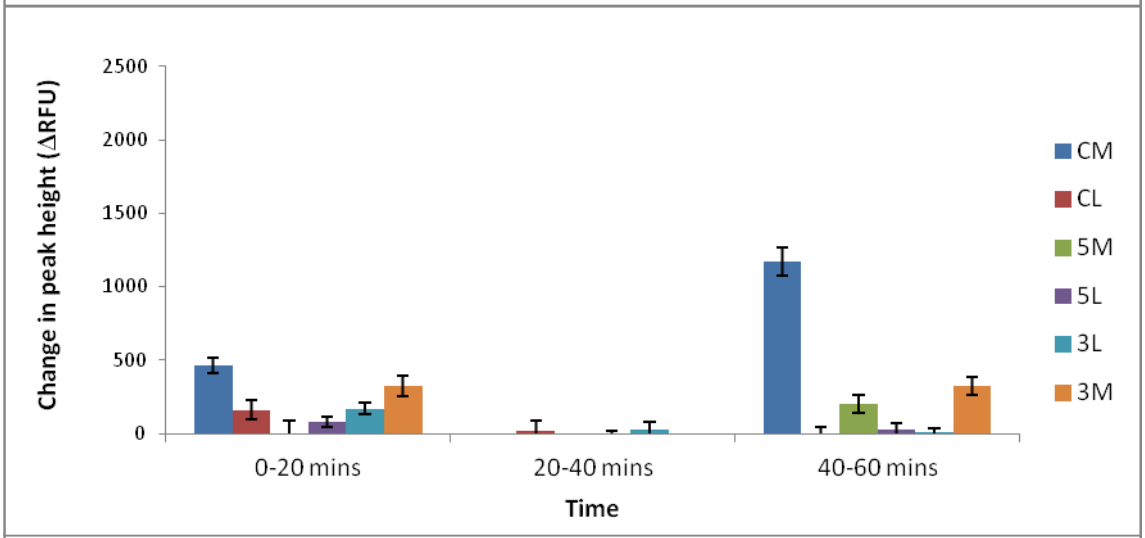


Figure 4.5. Effect of drying on the 18S rRNA fragments. Decrease in fragment height a) bloodstain dried at room temperature, between T=0 and 24 hours; b) bloodstains dried at 37°C between T=0 and 60 Minutes; c) Bloodstain dried under UV light between T=0 and 60 minutes. CM-central medium fragment, CL- central long fragment, 5M-5' medium fragment, 5L 5' long fragment, 3L 3' long fragment, 3M 3' medium fragment.

a)



b)



c)

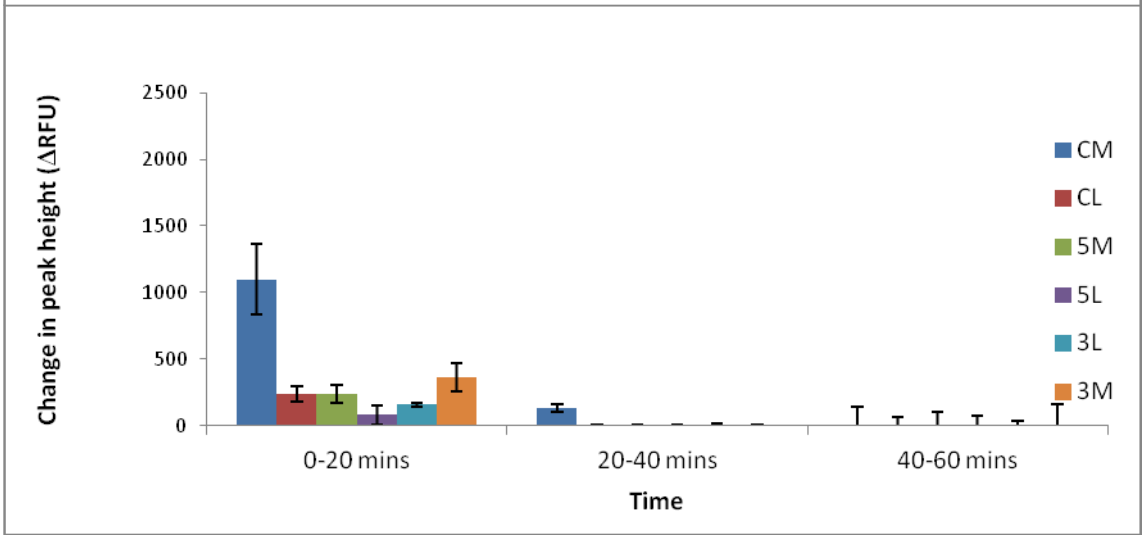


Figure 4.6. Effect of drying on the G6PD mRNA fragments. Decrease in fragment height a) bloodstain dried at room temperature, between T=0 and 24 hours; b) bloodstains dried at 37°C between T=0 and 60 Minutes; c) Bloodstain dried under UV light between T=0 and 60 minutes. CM-central medium fragment, CL- central long fragment, 5M-5' medium fragment, 5L 5' long fragment, 3L 3' long fragment, 3M 3' medium fragment.

4.3.1 Pattern of degradation

Table 4.3A shows the order of degradation of the 5', central and 3' fragments for each type of nucleic acid analysed from 10 bloodstains analysed in duplicate under the different drying conditions. The results in Table 4.2A were based on standardised input concentrations of CD3G cDNA (600 pg for G6PD mRNA assay, 200 pg for 18S rRNA assay) and DNA (1 ng) in each assay and are derived from the overall (average) reduction in peak height (Figures 4.7 to 4.9).

When each bloodstain was analysed under the different conditions, the peak height data were recorded in Microsoft Excel and the average reduction in RFU of each fragment was calculated. For example, Figure 4.7A shows that the average reduction (n=20) of the HBB "CL" (central large) fragment was 2000 RFU over 24 hours when left to dry at room temperature.

Based on the average decrease in peak height of each fragment (n=20 for each storage condition) over time, the pattern of degradation was determined to be random. If degradation was proceeding from the 3' end, the process would result in the loss of the 3' large fragment, 3' medium fragment, the central large fragment, central medium fragment, the 5' large and 5' medium fragments. That is to say that the 3' large fragment would exhibit the greatest reduction in peak height and the 5' medium fragment would exhibit the least. If degradation was proceeding from the 5' end then the opposite would be expected. As this was neither the case for HBB DNA, 18S rRNA or G6PD mRNA, the direction of degradation was deemed to be random under these conditions.

4.3.2 Extracted nucleic acids (naked nucleic acids)

Naked nucleic acids were also examined to determine what kind of affect the cellular environment plays in protecting nucleic acids from degradation.

Nucleic acids from 20 µl of fresh blood were extracted and left to dry under the same conditions as the bloodstains over the same time periods as described in section 4.2.4.

Figure 4.10 shows an example electropherogram of (1ng) HBB DNA stored under UV for 20 to 60 minutes when the template was from whole blood (Figure 4.10A to C) and extracted blood (Figure 10D to F). Figure 4.10C shows that after 60 minutes of UV exposure, less degradation has occurred to the HBB template in whole blood compared to Figure 4.10F where the peak heights are noticeably lower.

Figure 4.11 shows a similar result with the 18S rRNA template with a noticeable difference in Figures 4.11E and F where a loss of fragments is observed.

This trend is also observed in Figures 4.12D to F, where after 60 minutes of exposure to UV, the G6PD mRNA template is severely degraded (Figure 4.12F)

A key finding was observed in the UV study in that naked nucleic acids showed greater amounts of degradation compared to bloodstains as shown in Figures 4.10 to 4.12. This implies that the cellular environment does offer some degree of protection against nucleic acid degradation.

Table 4.3. Pattern of nucleic acid degradation. Pattern of nucleic acid degradation as a bloodstain dries under different conditions when a) each assay is standardised using fixed DNA or cDNA concentration and b) using a fixed volume of DNA or cDNA in the multiplex reactions. CL – central large fragment, CM – central medium fragment, 3L – 3' large fragment, 3M – 3' medium fragment, 5L – 5' large fragment, 5M – 5' medium fragment.

a)

	Room temperature	37°C	UV exposure
HBB DNA	(least stable) CL > CM > 3L > 3M > 5L > 5M (most stable)	(least stable) 3L > CM > CL > 3M > 5M > 5L (most stable)	(least stable) CM > CL > 3M > 3L > 5M > 5L (most stable)
18S rRNA	5L > 3L > 5M > CM > 3M > CL	3L > 5L > 3M > CM > CL > 5M	3M > 3L > 5L > CM > 5M > CL
G6PD mRNA	CM > 3M > 5M > CL > 3L > 5L	CM > 3M > 3L > CL > 5L > 5M	CM > 3M > 5M > CL > 3L > 5L

b)

	Room temperature	37°C	UV exposure
HBB DNA	Below detection threshold	CM > CL > 3L	CM > CL > 3L > 3M > 5L > 5M
18S rRNA	3L > 3M > 5M > 5L > CM > CL	3L > 3M > 5M > 5L > CM > CL	3L > 3M > 5L > CM > CL > 5M
G6PD mRNA	5L > 3M > CL > 3L > 5M > CM	3M > 5L > CL > 3L > 5M > CM	5M > 3M > CL > 5L > CM > 3L

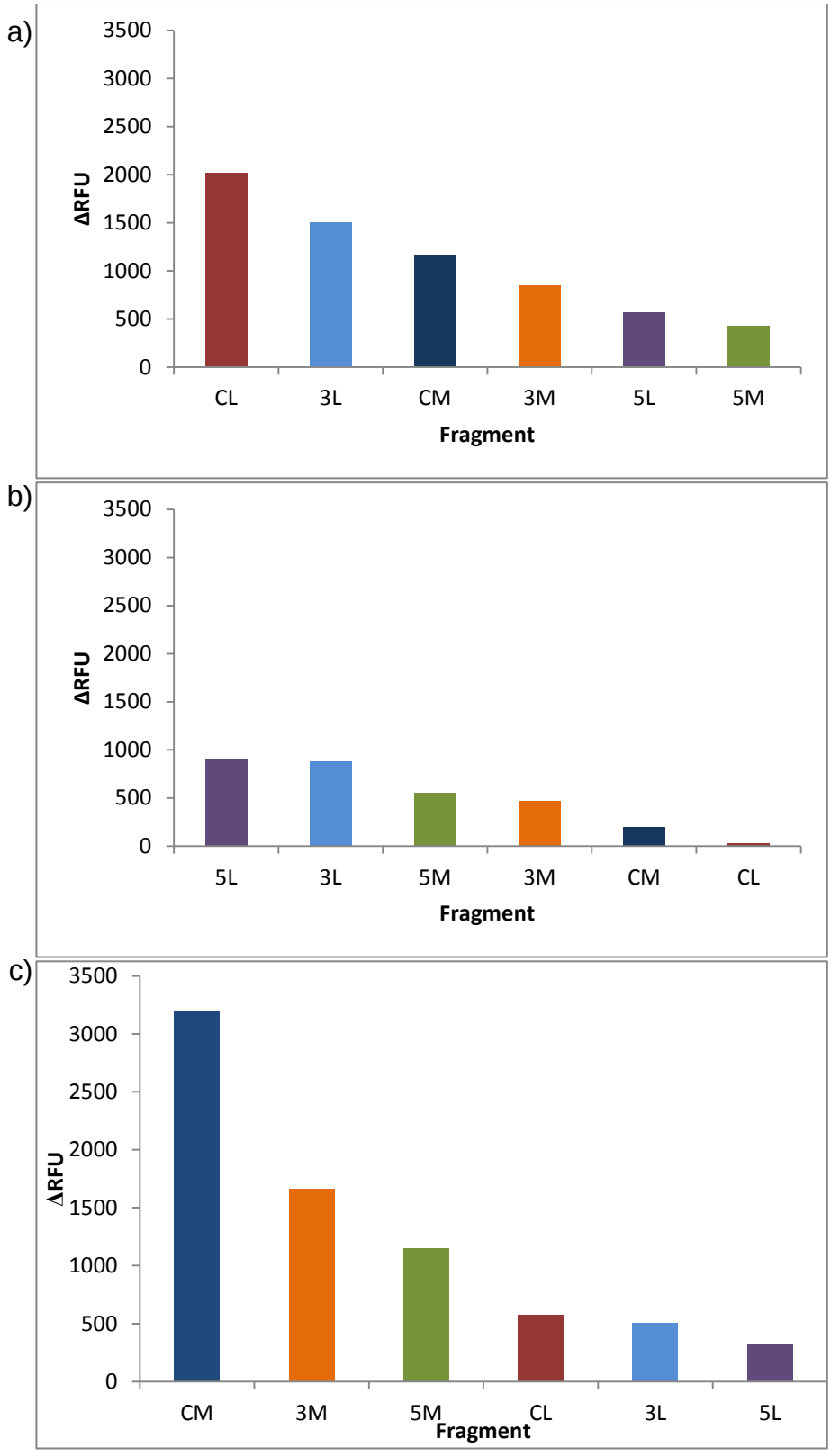


Figure 4.7. Degradation of nucleic acid fragments as a bloodstain dries at room temperature. a) overall reduction of the HBB DNA fragments over a 24 hour period, b) overall reduction of the 18S rRNA fragments over a 24 hour period, c) overall reduction of the G6PD mRNA fragments over a 24 hour period.

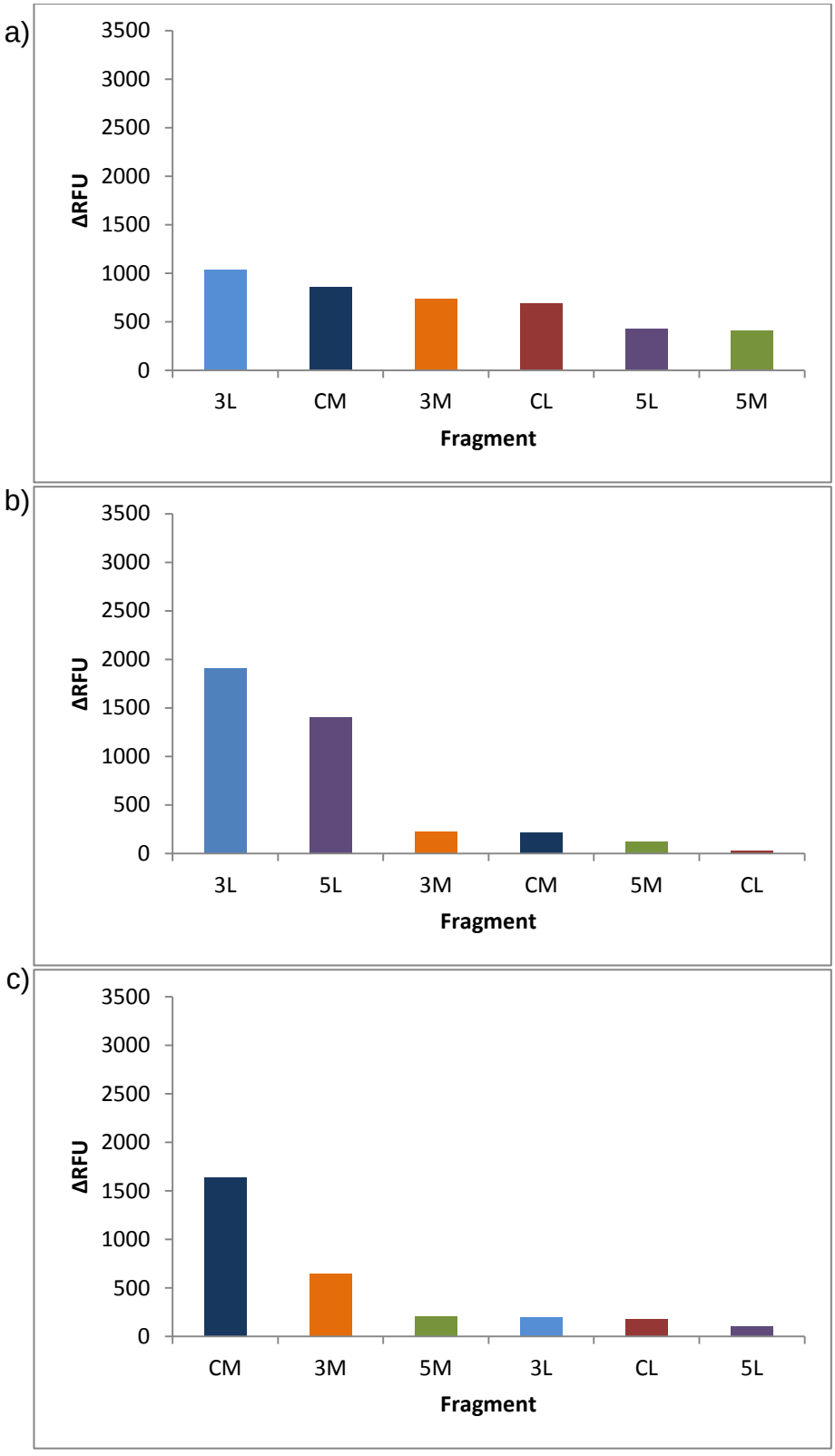


Figure 4.8. Degradation of nucleic acid fragments as a bloodstain dries at 37 °C. a) overall reduction of the HBB DNA fragments over a 60 minute period, b) overall reduction of the 18S rRNA fragments over a 60 minute period, c) overall reduction of the G6PD mRNA fragments over a 60 minute period.

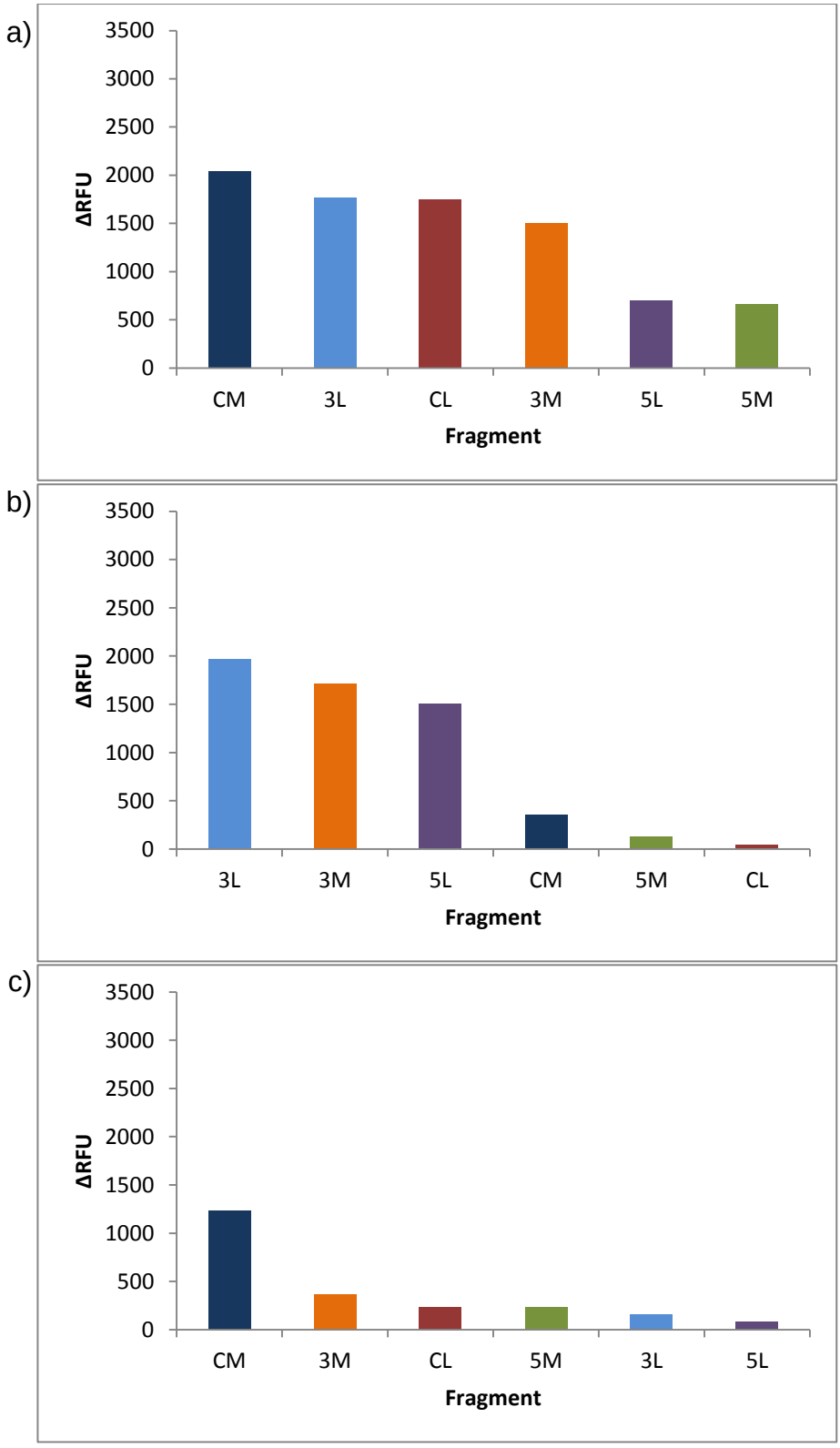


Figure 4.9. Degradation of nucleic acid fragments as a bloodstain dries under UV light. a) overall reduction of the HBB DNA fragments over a 60 minute period, b) overall reduction of the 18S rRNA fragments over a 60 minute period, c) overall reduction of the G6PD mRNA fragments over a 60 minute period.

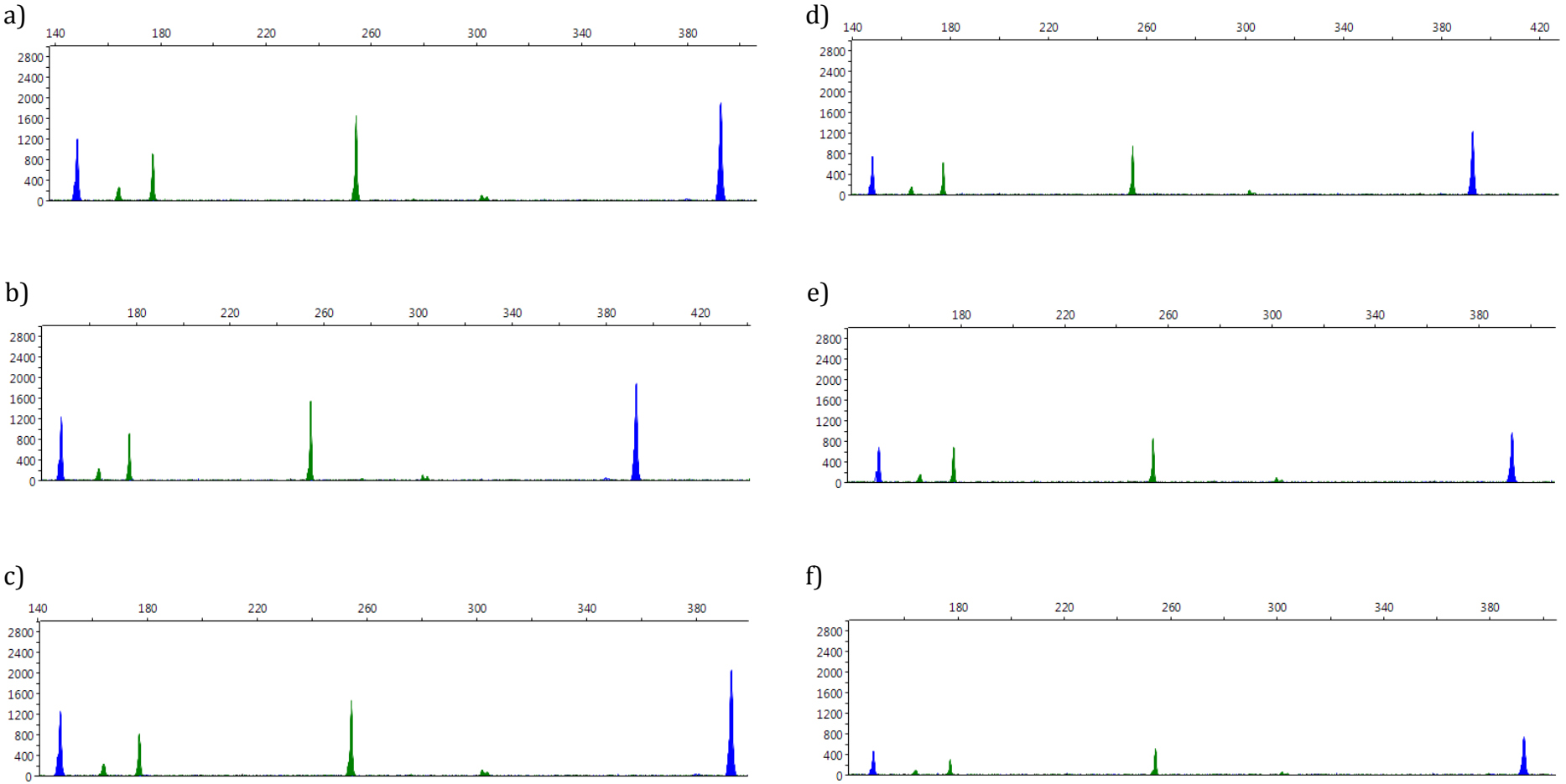


Figure 4.10 Stability of DNA in bloodstains compared to naked DNA. 20 µl bloodstains were deposited on sterile cotton cloth and left to dry under UV for a) 20 minutes, b) 40 minutes, c) 60 minutes; and 20 µl of extracted DNA left to dry under UV d) 20 minutes, e) 40 minutes, f) 60 minutes.

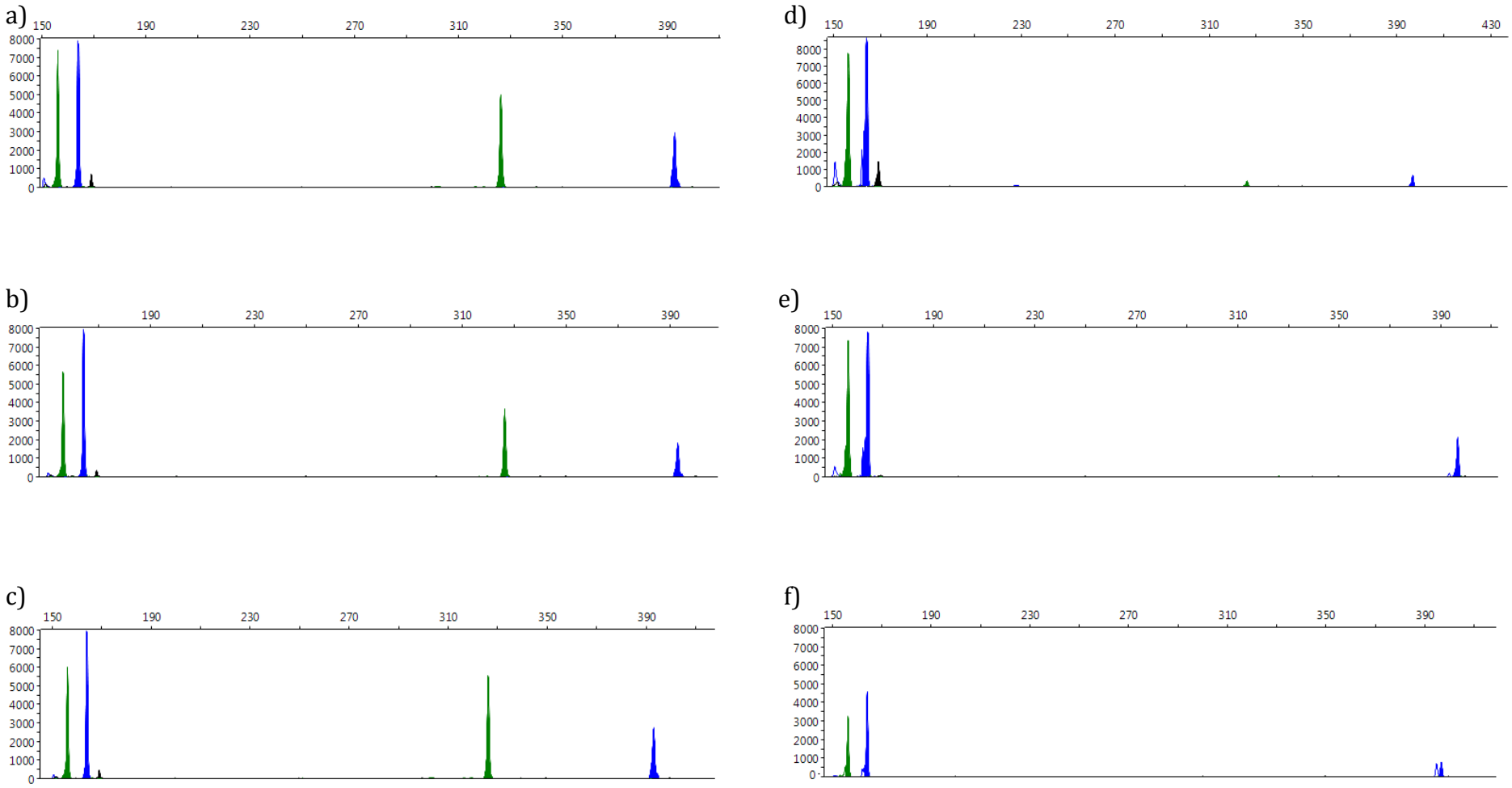


Figure 4.11 Stability of rRNA in bloodstains compared to naked rRNA. 20 µl bloodstains were deposited on sterile cotton cloth and left to dry under UV for a) 20 minutes, b) 40 minutes, c) 60 minutes; and 20 µl of extracted RNA left to dry under UV d) 20 minutes, e) 40 minutes, f) 60 minutes.

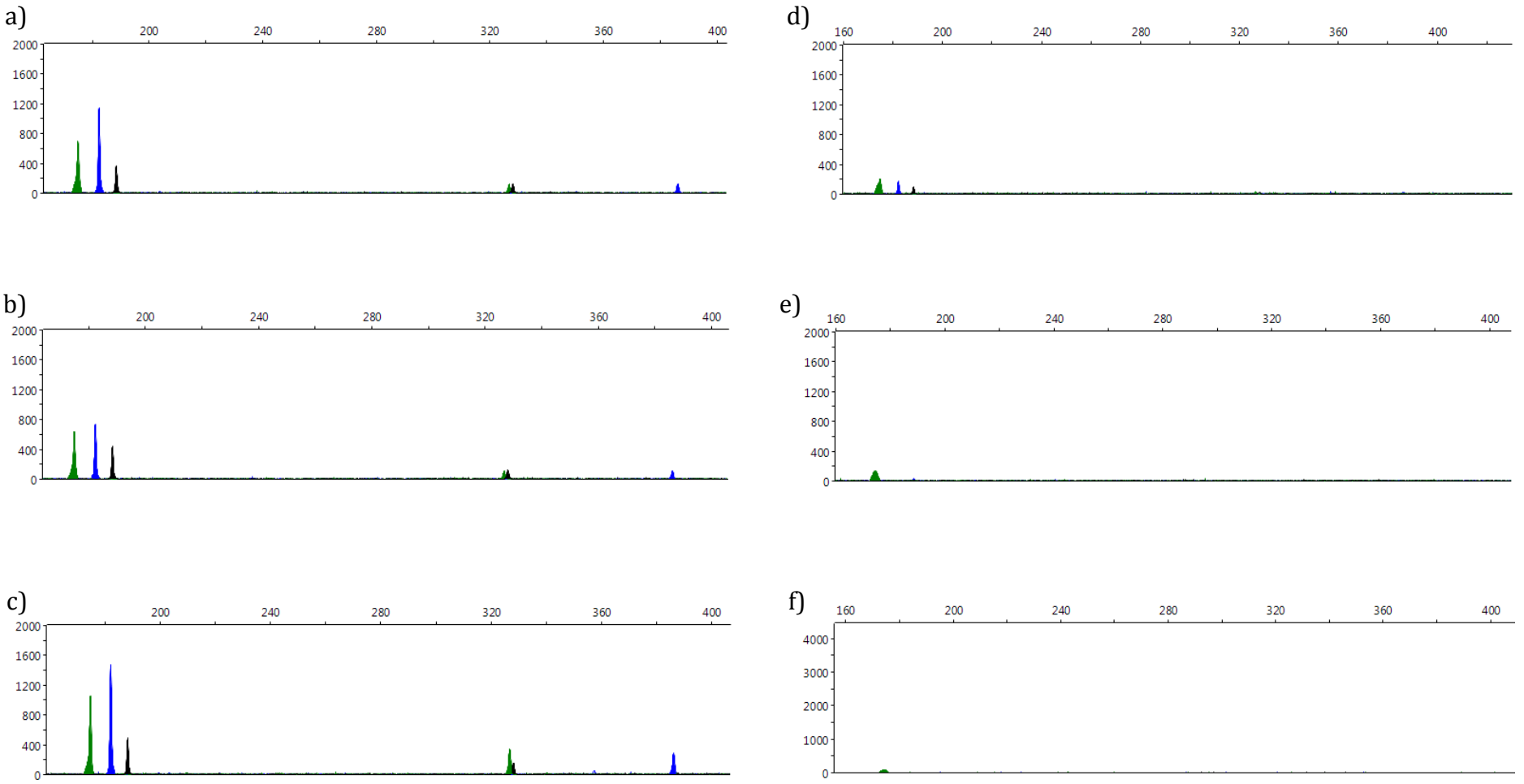


Figure 4.12 Stability of mRNA in bloodstains compared to naked mRNA. 20 µl bloodstains were deposited on sterile cotton cloth and left to dry under UV for a) 20 minutes, b) 40 minutes, c) 60 minutes; and 20 µl of extracted RNA left to dry under UV d) 20 minutes, e) 40 minutes, f) 60 minutes.

4.3.3 Standardisation of template input for PCR

When the same samples were analysed after standardising the volume of template to 1 µl (1 µl of cDNA or DNA was added to each assay) (Table 4.3B), it was evident that volume based analysis is not suitable for degradation studies, especially where comparisons are to be made between samples. This is because the template concentrations were different for each sample (as determined previously by the quantification step) and, as a result, some fragments were below the detection threshold (the detection threshold for each assay was set to 20 RFU) and were not able to be detected.

Consider the following scenario: A method to determine the age of bloodstain (which is dependent on relative input amounts of template) is being validated. A technician has been sent blind samples for analysis; they load 2 µL of template into the assay as described by the method. Analysis shows (either by way of an increased C_T value or drop in peak height) that the sample is below the set threshold. The sample could therefore be interpreted as a very old stain as it is highly degraded and skew the result. In effect, this outcome could be a result of simply not adding enough template for analysis. There would be no conclusive way of telling, especially if the positive control is unaffected and there is no indication of inhibition.

If nucleic acid analysis is solely based on input volume such of those of Anderson *et al.* (Anderson *et al.*, 2005, 2011), inherent differences should be expected between samples and the results should be used with caution. These studies somewhat negate the need to standardise samples by stating that analysis is based on the ratio of the different amplified products and this ratio is unaffected by the starting volume of the originating material. This rationale is debatable as the assays are generally designed to amplify more than one amplicon. For example, they simultaneously amplified 18S rRNA with β-actin mRNA. Ribosomal RNA is more abundant than β-actin mRNA and the competition of reagents in the assay would favour the amplification of 18S rRNA. The subsequent ratio of 18S rRNA to β-actin mRNA would always tend to favour 18S rRNA. In reality, when older stains are analysed and the

site of the target amplicon is either degraded, cleaved, or there is generally less competing starting material, it would constitute a reduction in the ratio of the two nucleic acids and it would therefore infer that the originating stain is generally older.

The order of fragment degradation was once again based on the highest reduction in average peak height (preferential degradation of that particular fragment) to the lowest (most stable). Where the assays were standardised using 1 µl of template and where results were obtained, the directionality of degradation was different (Table 4.2B) when compared to those samples standardised on template amount. These results demonstrate the importance of standardising the assay for the reasons mentioned above.

In order to relate the secondary structure of 18S rRNA and G6PD mRNA to the results obtained in this chapter; that the directionality of degradation is random as a result of the high occurrence of single stranded bases in certain areas of the sequence and the presence of labile phosphodiester bonds, the Mfold web server for nucleic acid folding and hybridization prediction (Zuker, 2003) was used. Mfold uses an algorithm to predict structure plots of the nucleic acid sequence based on minimum free energies. Mfold was also used to compare the "pattern" of degradation to the "theoretical" models based on the work of Bibillo (Bibillo *et al.*, 2000) and Kierzek (Kierzek, 2001).

Once the secondary structures of the RNAs were predicted, the primer binding sites were superimposed on the structure to see if they spanned any single stranded bases or labile phosphodiester bonds and therefore likely to undergo non-enzymatic cleavage. One or more unstable bonds could control the rate of RNA decay and influence the manner in which it is degraded. General analyses showed that the 5' region of 18S rRNA and central region of G6PD mRNA contained the most single stranded regions and are therefore likely to be the least stable. A summary of the results are shown in Table 4.4.

Table 4.4. RNA secondary structure modelling. Number of potential cleavage sites for 18S rRNA and G6PD mRNA based on their secondary structure.

	Region	Number of single stranded bases	Number of UA cleavage sites
18S rRNA	5' large	41	5
	5' medium	20	2
	Central large	44	2
	Central medium	30	1
	3' large	31	-
	3' medium	17	1
G6PD mRNA	5' large	34	-
	5' medium	20	-
	Central large	102	-
	Central medium	17	-
	3' large	41	1
	3' medium	16	-

DNA

The degree and spectrum of DNA damage depends on the sample source and the type of environment to which it was exposed. Some types of damage are ubiquitous and can potentially be present in all extracted DNA, while other types of damage are the result of exposure to a specific source (see Table 4.5). Under physiological conditions the most labile bond in DNA is the N-glycosyl bond that attaches the base to the deoxyribose backbone. This is in contrast to RNA in which the phosphodiester bond in the backbone is the least stable under the same conditions.

Hydrolysis of the N-glycosyl bond results in the loss of a base leaving an apurinic/apyrimidinic (AP) site that decomposes into a nick. AP sites are expected in all stored DNA samples because the reactive species is water. DNA is unstable in the presence of water because of its sensitivity to the addition of water across its phosphodiester bonds, which produce strand breakage (Komiyama, Takeda, & Shigekawa, 1999). If water is absent, the rate of strand breakage becomes very slow for DNA (and RNA).

Absence of water also reduces the rate of UV-light damage in DNA (as seen in the DNA of air dried seeds (Setlow, 1992)). In the air-dried state at less than 70% relative humidity, DNA remains a double helix but shifts to the A form of the helix. The A form is a right-handed double helix made up of antiparallel strands held together by Watson-Crick base-pairing. The A helix is wider and shorter than the B form, and its base pairs are tilted rather than perpendicular to the helix axis. The A-form of DNA is incompatible with ordinary UV light induced thymidine dimer formation (Lamola & Mittal, 1966). In the absence of that primary lesion, the only remaining source of UV-light photo damage in DNA is pyrimidine bases (see Table 4.5 for types of DNA damage). This could account for the observed increase in DNA degradation in during the transition from wet to dry when exposed to UV light.

Marrone *et al.* have stated that changes in dry state hemoglobin over time do not increase the potential for oxidative DNA damage in dried blood (Marrone

& Ballantyne, 2009). Oxygenated Hb converts over time to oxidised Hb, but this happens more quickly in the dried state than in the hydrated state. The Hb molecule in the dried state undergoes oxidative changes and releases reactive Fe(II) cations. There is no evidence that Hb becomes more prone to generating OH· as it ages in either the hydrated or dried states.

Bonnet *et al.* (Bonnet, Colotte, Coudy *et al.*, 2009) described *ex vivo* studies where DNA is degraded mainly by water and reactive oxygen species. Hydrolysis mainly leads to single-strand breaks following depurination. It also causes base deamination that is not followed by chain breakage (e.g. cytosine converted into uradine (Frederico, Kunkel, & Shaw, 1990)). Other pathways including the removal of pyrimidines (Lindahl & Nyberg, 1974) or phosphodiester hydrolysis (Schroeder, Lad, Wyman *et al.*, 2006) are much slower. Oxidation concerns both sugar and bases: attacks on the bases generate a wide variety of modifications, some lead to the rupture of the N-glycosidic or phosphodiester bond, others do not. If enough AP sites are present, amplification reactions will fail as typical PCR polymerases stall at the AP site preventing further replication (Loeb & Preston, 1986) . The breakdown of AP sites into nicks further compounds the problem as it eventually leads to the fragmentation of the DNA that can be seen as a decrease in peak height.

Establishment of the “dried state” leads to a strong decrease in molecular mobility and consequently in chemical reactivity. Upon dehydration (90-70% relative humidity), DNA undergoes conformational changes from form B to form A (Ayala-Torres, Chen, Svoboda *et al.*, 2000). Below 70-50% relative humidity and at room temperature, DNA undergoes natural denaturation such as base destacking and the rupture of hydrogen bonds (Bonnet *et al.*, 2009).

Table 4.5. Types of DNA damage. The different types of potential and sources of DNA damage.

Source of DNA	Potential Damage	Comments
Ancient DNA	Abasic sites, deaminated cytosine, oxidized bases, fragmentation, nicks(Gilbert, Binladen, Miller *et al.*, 2007; Hofreiter, Jaenicke, Serre *et al.*, 2001)	Cytosine deamination has been reported to be the most prevalent cause of sequencing artefacts in ancient DNA.
Environmental DNA	Fragmentation, nicks(Qiu, Wu, Huang *et al.*, 2001)	Nicks and fragmentation can increase the formation of artifactual chimeric genes during amplification.
Source of Damage		
Exposure to Heat	Fragmentation, nicks, abasic sites, oxidized bases, deaminated cytosine, cyclopurine lesions (Bruskov, Malakhova, Masalimov *et al.*, 2002)	Heating DNA accelerates the hydrolytic and oxidative reactions in aqueous solutions.
Phenol/Chloroform Extraction	Oxidized bases (Finnegan, Herbert, Evans *et al.*, 1995)	Guanine is more sensitive to oxidation than the other bases and forms 8-oxo-G. 8-oxo-G can base pair with A making this damage mutagenic.
Exposure to Light (UV)	Thymine dimers, (cyclobutane pyrimidine dimers) pyrimidine (6–4) photo products (Cadet, Sage, & Douki, 2005; Pfeifer, Young-Hyun, & Besaratinia, 2005)	UV trans-illumination commonly used to visualize DNA causes thymine dimer formation.
Mechanical shearing	Fragmentation, nicks	Normal DNA manipulations such as pipetting or mixing can shear or nick DNA.
Desiccation	Fragmentation, nicks, oxidized bases (Mandrioli, Borsatti, & Mola, 2006)	
Storage in Aqueous Solution	Abasic sites, oxidized bases, deaminated cytosine, nicks, fragmentation (Lindahl & Andersson, 1972; Lindahl & Nyberg, 1972)	Long term storage in aqueous solution causes the accumulation of DNA damage.

4.4 Conclusion

The purpose of this chapter was to investigate the factors affecting the degradation of nucleic acids in a bloodstain as it dries and apply the principle of standardising each assay to determine a pattern of degradation.

DNA found in a dehydrated bloodstain is present in a nucleoprotein complex within the cellular infrastructure. Therefore, dehydration affords DNA some protection against the harmful effects of environmental damage. Ribosomal RNA is typically part of a ribosomal subunit, which cluster into polysomes, so like DNA, this also affords some degree of protection. This could be seen in the "naked" nucleic acid study where nucleic acids that were exposed to UV light were rapidly degraded (Figures 4.10 to 4.12). Messenger RNA on the other hand is single stranded and despite a higher copy number, this may make it more prone to degradation but not the effects of UV (Section 4.3 and Figure 4.6C).

Collectively, the results indicate that the nucleic acids in bloodstains are somewhat protected against the damaging effects of the environment (such as sun exposure and elevated temperatures) when compared to "naked" nucleic acids (section 4.3.2 and Figures 4.10 to 4.12) and this makes further analysis of nucleic acids feasible. This protection could be due to the dehydrated state of the nucleic acids in the stain, the local cellular environment of the DNA and RNA, or a combination of both.

When bloodstain evidence is recovered from a scene, the environment in which the evidence has been exposed to may be unknown. For example, a bloodstain might have been recovered from a dashboard of a car that has been parked in the sun over a number of days. As the results of this study indicate, the "pattern" of degradation is random and as a result, a multiplex assay design would be preferable over a one-targeted amplicon approach as that single amplicon may be situated on part of the sequence that has been cleaved during the degradation process.

By targeting different sized amplicons through the use of a multiplex assay, this study shows that in the *ex vivo* environment, degradation of nucleic acids appears to occur at random sites along the nucleic acid strand as measured by a decrease in peak height of specific fragments. By investigating the effects of drying and temperature, this study shows that the rate of degradation is fastest as the stain dries (within the first 60 minutes). Degradation then proceeds at a slower rate after a 24-hour period. Hampson *et al.* have also reported a sharp initial decrease in the relative expression of β-actin and 18S rRNA between time 0 and day 1 in hair (Hampson *et al.*, 2011). This could be a result of increased intra-cellular enzymatic activity, which is dependent on the natural aqueous environment common to all living cells.

It was interesting to find that the work of McNally *et al.* (McNally *et al.*, 1989) that investigated the recovery of intact DNA from bloodstains recovered from a variety of environments and substrates, found that suitable DNA for restriction fragment length polymorphism analysis could be recovered from small bloodstains as they most likely dried faster. His finding is now also transferrable to smaller stains and not just to DNA but also RNA. Those stains that were exposed to elevated temperatures dried faster and showed less degradation (based on the average change in fragment peak height as shown in Figure 4.8).

Chapter 5

A Method to Determine the Time since Deposition of a Bloodstain

5.1 Introduction

Having found out more about the behaviour of nucleic acids in drying bloodstains through the use of three novel multiplex assays, the same approach was taken to develop a quantitative method to determine the time since deposition of aged bloodstains.

This chapter is a pilot study that describes the development of a statistical method. The study investigated how nucleic acids degrade as a function of time, and how certain fragments within each multiplex assay can be used to estimate when a bloodstain was deposited.

5.2 Methods

5.2.1 Blood collection and sampling

Thirty-four bloodstains (20 µL) were obtained from eight volunteers over a period of months and deposited onto sterile cotton cloth. They were left to age at room temperature with exposure to sunlight. Stains were extracted and processed in duplicate at the following time points: 10 at t = 4 days, 8 at t = 20 days and 16 at t = 150 days (n = 68). These time points were selected to determine if there was any significant difference in peak height (between days, weeks and months) when they were used in the development of the statistical model. All bloodstains were extracted and processed as described in sections 2.2 to 2.5.

5.2.1 Data collection

The amount of DNA or cDNA used for fragment analysis was standardised using the method described in sections 3.3.4 to 3.3.5. Two hundred picograms of CD3G cDNA was added to each 18S rRNA multiplex assay and 600 pg to each G6PD multiplex assay, 1 ng of DNA was added to each HBB DNA assay as described in sections 4.2.6.2 and 4.2.6.3. The small fragments

of each assay were omitted as section 4.3 describes negligible changes to fragment peak height up to 150 days (Figure 4.2 and Table 4.2).

Sixty eight measurements of bloodstain age (from the three time points) were therefore used for statistical analysis to determine if bloodstain age could account for changes in peak height of the HBB DNA, 18S rRNA and G6PD mRNA fragments.

5.2.2 Statistical analysis of dried blood

The statistical model described here was developed using the three time points mentioned above and then subsequently tested (Chapter 6) using varied time points.

5.2.2.1 Data analysis technique

The data were analysed to identify the least number of fragments that best predict the age of the bloodstain. The data was analysed in 5 steps: 1) outlier detection and elimination, 2) exploratory data analysis, 3) correlational analysis, 4) regression modelling and 5) multicollinearity diagnosis. All data analysis was performed with a statistical package for the social sciences (SPSS) (IBM Corp. Released 2012. IBM SPSS Statistics, Version 21.0. Armonk, NY: IBM)

5.2.2.2 Outlier detection and elimination

Outliers can be described as data values that fall significantly outside of acceptable ranges. Outliers can increase the errors associated with the regression model and can have a dramatic effect on the model fit (MacCallum, 1995). Box plot graphs were used to identify any potential outliers. A box plot graph divides the data into quartiles (25, 50 and 75%) and identifies minimum and maximum acceptable values (Martinez & Martinez, 2004). Data points outside the minimum and maximum were considered outliers.

5.2.2.3 Exploratory data analysis

The purpose of exploratory data analysis is to obtain an understanding of the data associated with each variable in the study. This step involved the generation of a number of descriptive statistics such as standard deviation, skewness, kurtosis and mean. The independent x variables (peak heights of the nucleic acid fragments) were also tested for normality since this is an assumption required for linear regression modelling. A one-sample Kolmogorov-Smirnov (K-S) test was used to determine if the independent variables approximated a normal distribution. If the *p*-value of the K-S test was less than 0.5, the null hypothesis stating the distribution was not normal was rejected (Corder & Foreman, 2009).

5.2.2.4 Correlational analysis

An important assumption involved in regression analysis is that there is at least some degree of linear relationship between each independent x variable (peak height) and the dependent *y* variable (time since deposition). This assumption was tested by calculating the Pearson Product-Moment Correlation Coefficient (*R*) for each independent and dependent variable combination. The *R* value can range from -1 to +1. A value of +1 means there is a perfect, positive relationship between the two. A value of -1 means there is a perfect negative relationship (Salkind & Green, 2003).

Besides calculating the *R* value for each pairwise combination, SPSS was used to determine the significance of the linear relationship using the bivariate correlations procedure. If the significance of the relationship was found to be less than 0.05, then there was a linear relationship between the peak heights of the nucleic acid fragments and the time since the bloodstain was deposited. If the pairwise combination of $r \neq 0$ or the significance was more than 0.05, the null hypothesis would not be rejected, meaning that there was no linear relationship of the peak heights of the nucleic acids fragments. A weak correlation implies that the independent variable should potentially be excluded from the regression model.

5.2.2.5 Regression modelling

On completion of the correlation analysis, a regression model was developed using SPSS. The model was constructed by adding all relevant independent variables (nucleic acid fragments and their corresponding peak heights) simultaneously. From this initial regression model, an R, R^2, and adjusted R^2 value was calculated. The R value is a measure of multiple correlations between the independent and dependent variables. The R^2 value is a measure of variance in the outcome that is influenced by the predictor variables (fragment peak height), whilst the adjusted R^2 value represents the estimated value of variance.

Once the R, R^2, and adjusted R^2 values were calculated for the initial regression model, a backward stepwise method was used to eliminate fragments that did not significantly contribute to the model fit. This method involved reviewing the p-value for each independent variable (nucleic acid fragments and their corresponding peak heights). If the p-value was greater than 0.1, the variable was removed from the model and the model was re-calculated with the remaining variables. If the removal of the variable significantly weakened the model then the independent variable was re-entered otherwise it was deleted. This procedure continued until each remaining fragment explains a significant partial amount of the variability in the age of the bloodstain. Forward stepwise regression was then performed on the remaining fragments, starting with one fragment and then adding the next, to assess the contribution of each fragment to the overall model.

Analysis of Variance (ANOVA) was used to investigate the relationship between the age of the bloodstain and the remaining nucleic acid fragments. Analysis of Variance uses an F test (comparing the ratio of two variances) to compare the means of the different fragments and determine if a statistical significance exists (at the $p<0.05$ level). The F-test values from the ANOVA results were used to compare different regression models to determine the contribution of each fragment in the regression model. The F value is the ratio of the mean regression sum of squares divided by the mean error sum of

squares. The F-value will range from zero to an arbitrarily large number, so the larger the F-value, the better the model (as it will have greater predictive power).

5.2.2.6 Hypotheses

The null hypotheses for regression analyses were as follows:

H_{01}: $R^2 = 0$	There is no linear relationship between fragment degradation and time.
H_{02}: bG6PD_5L = 0	The R value of the G6PD mRNA 5' medium fragment is 0.
H_{03}: bG6PD_5M = 0	The R value of the G6PD mRNA central large fragment is 0.
H_{04}: bG6PD_CL = 0	The R value of the G6PD mRNA central medium fragment is 0.
H_{05}: bG6PD_CM = 0	The R value of the G6PD mRNA 3' large fragment is 0.
H_{06}: bG6PD_3L = 0	The R value of the G6PD mRNA 3' medium fragment is 0.
H_{07}: bG6PD_3M = 0	The R value of the G6PD mRNA 5' large fragment is 0.
H_{08}: b18S_5L = 0	The R value of the18S rRNA 5' large fragment is 0.
H_{09}: b18S_5M = 0	The R value of the 18S rRNA 5' medium fragment is 0.
H_{010}: b18S_CL = 0	The R value of the 18S rRNA central large fragment is 0.
H_{011}: b18S_CM = 0	The R value of the 18S rRNA central medium fragment is 0.
H_{012}: b18S_3L = 0	The R value of the 18S rRNA 3' large fragment is 0.
H_{013}: b18S_3M = 0	The R value of the 18S rRNA 3' medium fragment is 0.
H_{014}: bHBB_5L = 0	The R value of the HBB DNA 5' large fragment is 0.
H_{015}: bHBB_5M = 0	The R value of the HBB DNA 5' medium fragment is 0.
H_{016}: bHBB_CL = 0	The R value of the HBB DNA central large fragment is 0.
H_{017}: bHBB_CM = 0	The R value of the HBB DNA central medium fragment is 0.
H_{018}: bHBB_3L = 0	The R value of the HBB DNA 3' large fragment is 0.
H_{019}: bHBB_3M = 0	The R value of the HBB DNA 3' medium fragment is 0.

That is for each hypothesis tested, there is no correlation between the height in RFU of the named fragment (x variable) and the age of the bloodstain (y variable).

5.2.2.7 Multicollinearity analysis

Multicollinearity is a situation where two or more independent variables in a regression model are highly correlated such that neither of them are useful when the other variables are also in the model, making it difficult to determine the importance of a specific independent variable. Multicollinearity limits the size of *R* of the regression model (section 5.2.2.5), meaning that the true linear relationship between "x" and "y" is not known, and increases the standard error of the regression coefficients (Field, 2005) meaning that the ability to predict "y" from "x" results in large error estimates. The collinearity statistic provided by SPSS, namely the Variance Inflation Factor (VIF), was used to identify potential instances of multicollinearity. Rule of thumb values of VIF range from 4 to 10; which indicate how much the variance has been inflated by lack of independence (O'brien, 2007). When multicollinearity is found, it often results in removing one of the independent variables as it is important for the independent variables to be correlated with the age of the bloodstain and not highly correlated among themselves. Multicollinearity is not a problem if the goal of the regression model is simply to predict the age of the bloodstain from a set of nucleic acid fragments as the overall R^2 value shows how well the model predicts age. However, as the goal is to understand how the various fragments contribute to age estimation, then multicollinearity is a problem for the reasons mentioned above.

There is no formal VIF value for determining presence of multicollinearity (O'brien, 2007). Values of VIF that exceed 10 are often regarded as indicating multicollinearity. A VIF value of 10 was therefore used to identify instances of collinearity between independent variables.

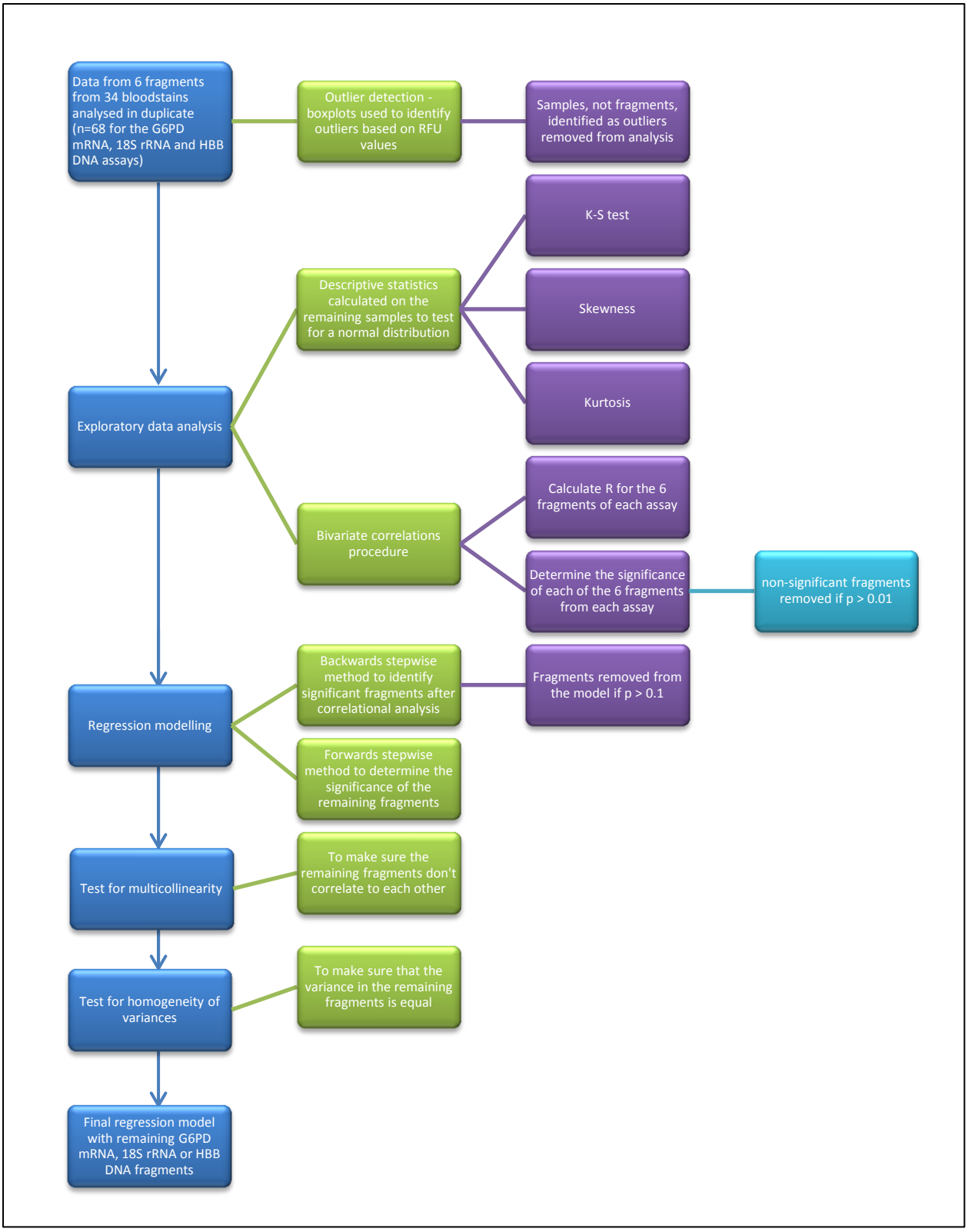


Figure 5.1. Summary of the statistical analysis used in the development of a method to determine the age of a bloodstain.

A summary of the statistical methods used in the development of a method to determine the age of a bloodstain are shown in Figure 5.1.

5.3 Results and Discussion

Figures 5.2 to 5.4 illustrate the changes in fragment peak heights and the complete degradation of others (as seen in mRNA in Figure 5.3) over the 150 day time course. These visual changes gave credence to the theory that fragment degradation is linked to time. Figure 5.2 shows a drop in peak height of the central medium, 3' large and central large fragments (for peak identification see Figure 4.3C) up to 150 days.

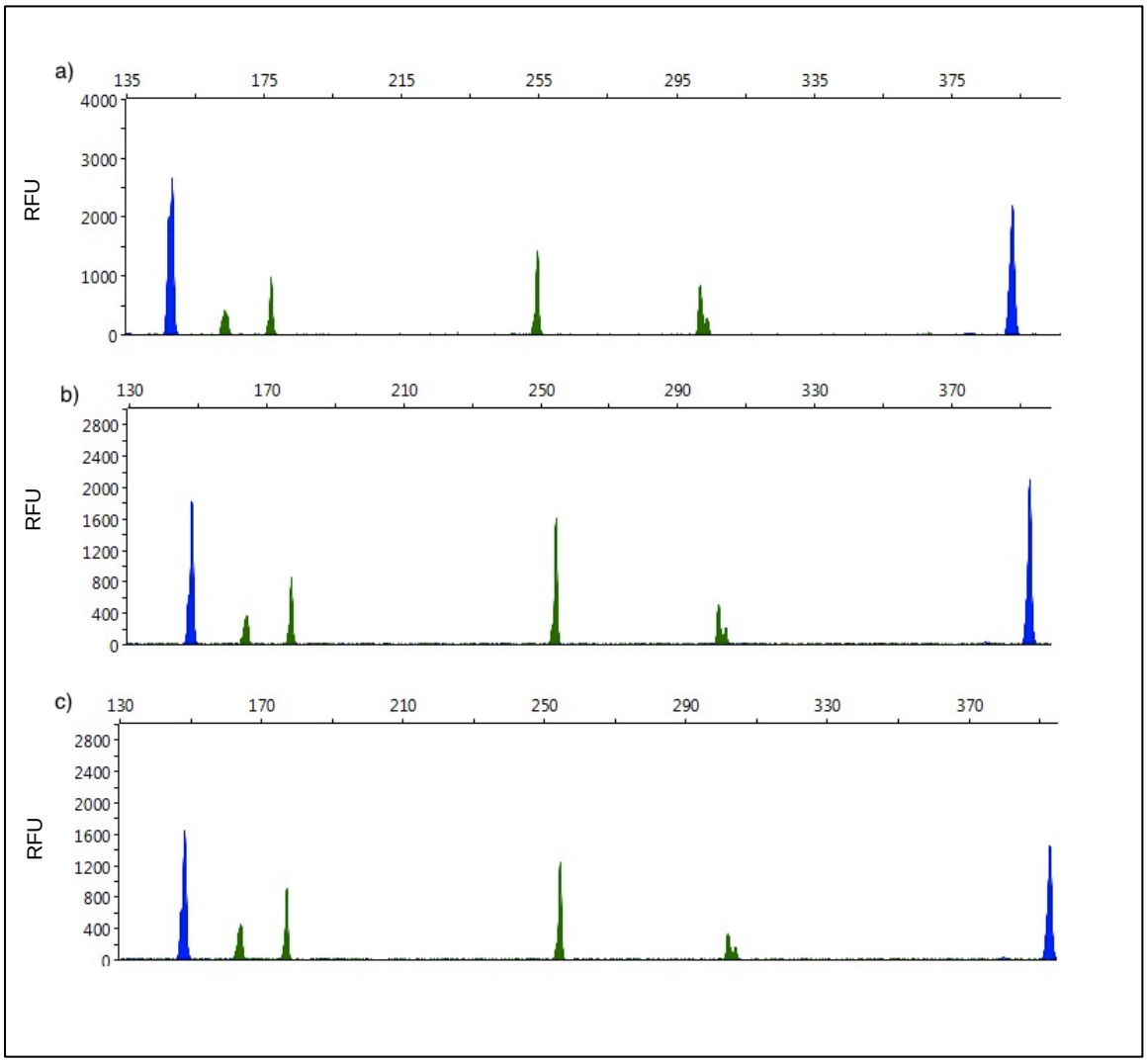


Figure 5.2. HBB DNA multiplex assay electropherograms of different bloodstain ages. HBB DNA time course showing changes in fragment peak height between a) t= 4, b) t= 20 and c) t= 150 days.

The changes in 18S rRNA shown in Figure 5.3 indicate that the central medium, central large, 3' large and 5' large fragments decreased in peak height over the time course (for peak identification see Figure 4.3B).

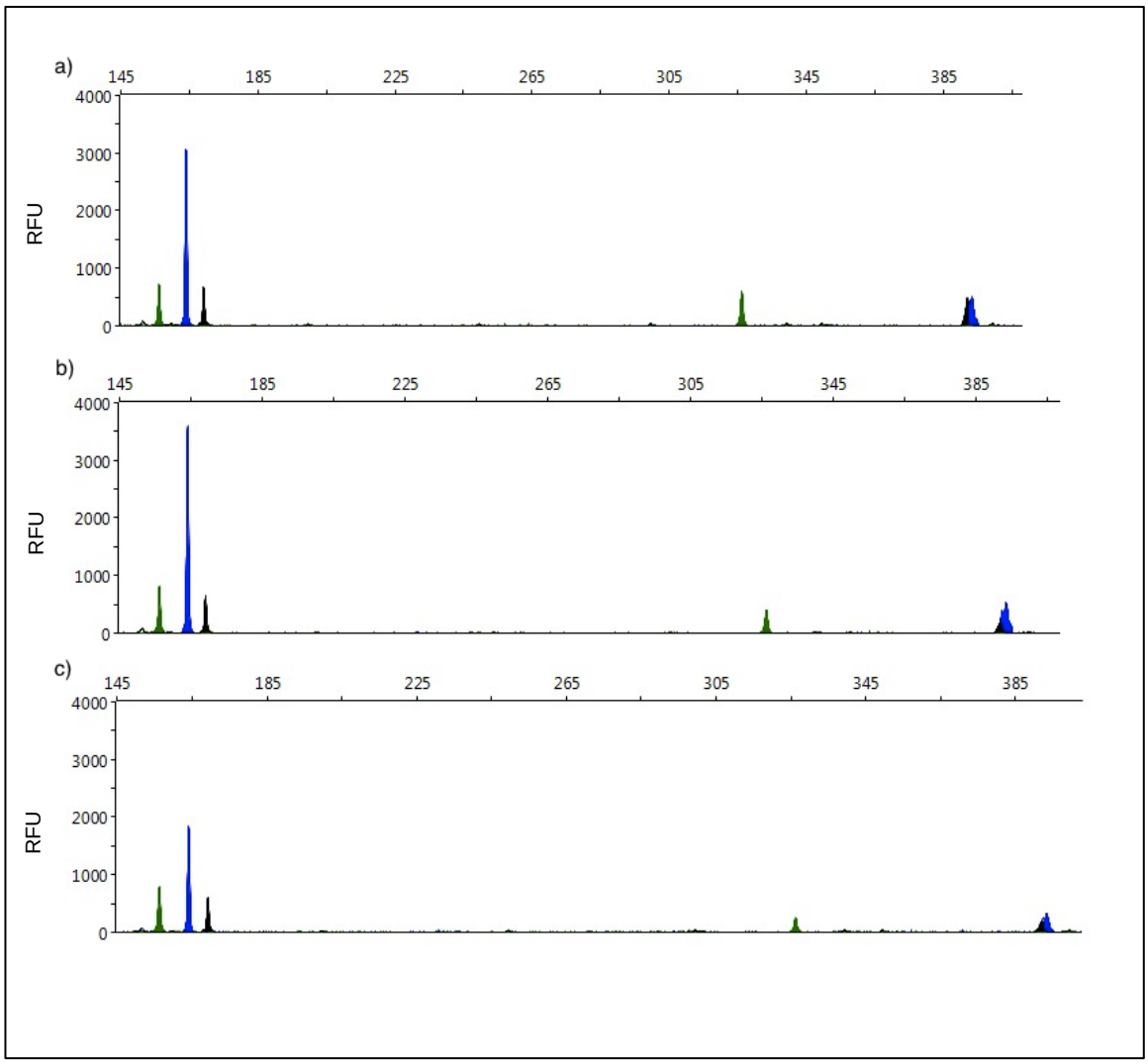


Figure 5.3. 18S rRNA multiplex assay electropherograms of different bloodstain ages. 18S rRNA time course showing changes in fragment peak height between a) t= 4, b) t= 20 and c) t= 150 days.

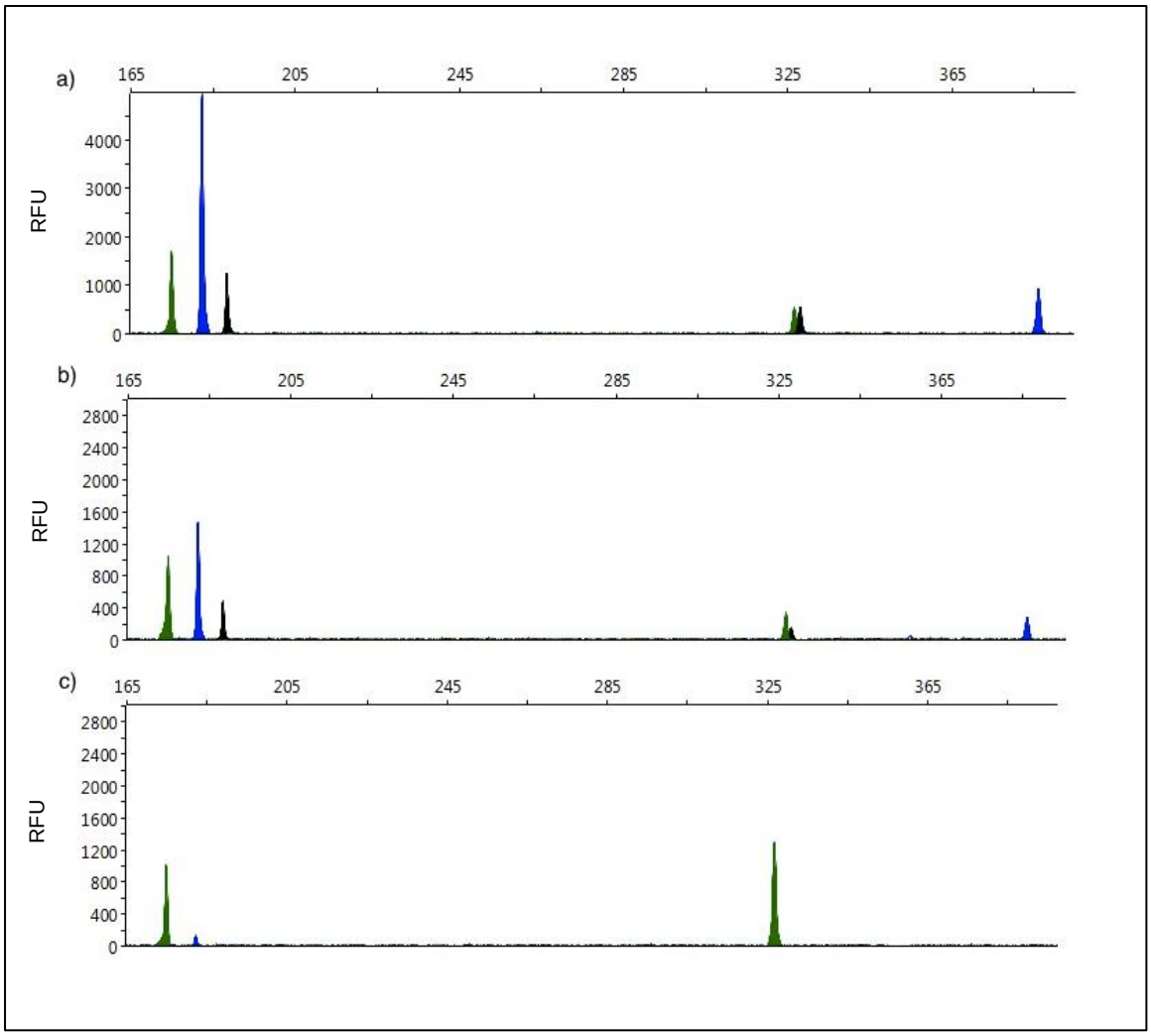


Figure 5.4. G6PD mRNA multiplex assay electropherograms of different bloodstain ages. G6PD mRNA time course showing changes in fragment peak height between a) t= 4, b) t= 20 and c) t= 150 days.

The changes in G6PD mRNA shown in Figure 5.4 indicate that degradation (reduction in peak height and fragment loss) was much more apparent over the time course compared to the DNA and rRNA markers. Figure 5.4C shows that at t = 150 days, there was the complete loss of the central long, 5' long and 5' medium fragments (for peak identification see Figure 4.3A).

This research focused on two variables. The independent variables were the nucleic acid fragments (5' medium fragment, 5' large fragment, central medium fragment, central large fragment, 3' medium fragment and the 3' large fragment) whose peak height in RFU is hypothesised to correspond to the dependent variable, time.

Through the use of linear regression, this study aimed to determine which nucleic acid fragments could be used in the development of an equation that could be used to estimate or predict the age of a bloodstain. Regression analysis is dependent on four principle assumptions: 1) Assumption of linearity: there is a linear relationship between independent variables and the dependent variable, 2) Assumption of independence: there is no serial correlation between independent variables, 3) Assumption of constant variance: there are no equal statistical variances, and 4) Assumption of normality: errors are normally distributed ("Regression with SPSS,"). A violation of any one of these assumptions could lead to the development of a misleading model. When these assumptions are satisfied, it means all of the information available from the patterns in the data has been used. When a violation occurs, it means that there is a pattern to the data that has not been included the model, and there could be a model that fits the data better.

Linear regression analysis is an analytical technique that determines if one or more independent variables can predict an outcome in a statistically significant way. In its' simplest form, linear regression analysis fits a straight line through a set number of points so that the sum of the squared residuals of the model are as small as possible. This is called the "least squares regression equation" ("Regression with SPSS,"). The basic assumption rests in the ability to model the variables using a linear function:

$$Y = a + bx \qquad (1)$$

In equation (1), Y is the dependent variable, in our case the age of the bloodstain, the constant, *b* is the regression weight that is multiplied against each observation of the fragment peak height in RFU (*x*) to plot a line. The constant and regression weight are derived from and are dependent on the data used for analysis.

Linear regression can be expanded by the inclusion of multiple independent variables. The linear multiple regression model is:

$$Y = a + b1x1 + b2x2 + b3x3 + \ldots byxy \qquad (2)$$

In equation (2), the coefficients of each independent variable (x, RFU) needed to predict the age of the bloodstain (Y) are represented by *b* ("Regression with SPSS,").

To evaluate the goodness of fit of a multiple linear regression model, the *R*-squared value (R^2) is commonly used. If there is no relationship between the independent variables and the dependent variables, the R^2 value would be 0. The value of R^2 will always fall between 0.0 and 1.0. However, the R^2 value is not a true measure of significance of the model. The overall *p*-value of the model is used to determine if the model is a good fit. The rule used is "if *p*-value > α, the null hypothesis stating that $r^2 = 0$ is rejected", thus indicating that the model is a poor fit. To achieve a good fit, the model required that the *p*-value be less than 0.05 (α). A *p*-value was generated for each coefficient (*b*y) and compared against the α value for each variable (*x*y) to indicate whether the variable was significant or not in the overall model.

5.3.1 Presentation of findings

SPSS statistics was used to perform statistical analyses and create graphical representations of the raw data (Appendix 4, Tables 1 to 3). The results of each step are explained below.

5.3.2 Analysis of outliers

Outlier analysis was performed on the data obtained (RFU) on each fragment (6 per sample) for each mRNA, rRNA and DNA assay (10 at t = 4 days, 8 at t = 20 days and 16 at t = 150 days). Thirty-four bloodstains were analysed in duplicate (n = 68) for each assay. The results are shown in Figures 5.5 to 5.7. The numbers in each plot represent the corresponding RFU values that fell

outside the minimum and maximum acceptable values (outliers). The boxplot is a graphical display of the data that shows: (1) the median as a black line, (2) the middle 50% of RFU values in red, (3) the top and bottom 25% of the RFU values for each fragment extending out of the red region, (4) the smallest and largest (non-outlier) RFU values for each fragment are the horizontal lines at the top/bottom of the boxplot, and (5) outliers. The boxplots show both "mild" outliers and "extreme" outliers. Mild outliers were classified as any fragment heights (RFU) that were more than 1.5x the interquartile range (IQR) from the rest of the RFU values, and are indicated by open dots. The interquartile range is the middle 50% of the RFU values. Extreme outliers were classified as any RFU values more than 3x the IQR from the rest of the RFU values, and are indicated by stars. These benchmarks were arbitrarily chosen, similar to how $p<0.05$ is arbitrarily chosen (Stapel, 2004).

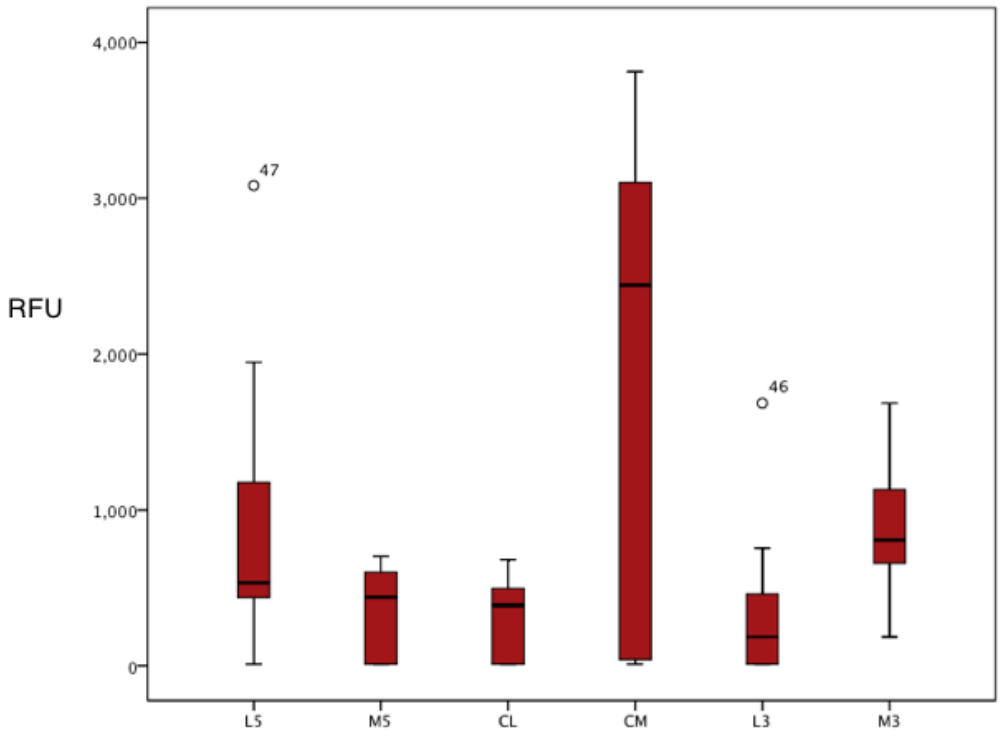


Figure 5.5. Box plot of RFU distribution for the six G6PD mRNA fragments. L5 = 5' large fragment, M5 = 5' medium fragment, CL = central large fragment, CM = central medium fragment, L3 = 3' large fragment and M3 = 3' medium fragment.

As shown in Figures 5.5 to 5.7 mild and extreme outliers were detected. Figure 5.4 shows that in the G6PD mRNA assay. Two samples had RFU values that were within 1.5x the IQR from the rest of the RFU values, and were therefore mild outliers. These samples were from bloodstains sampled at 150 days and were the 5' large fragment (sample 47) and 3' large fragment (sample 46).

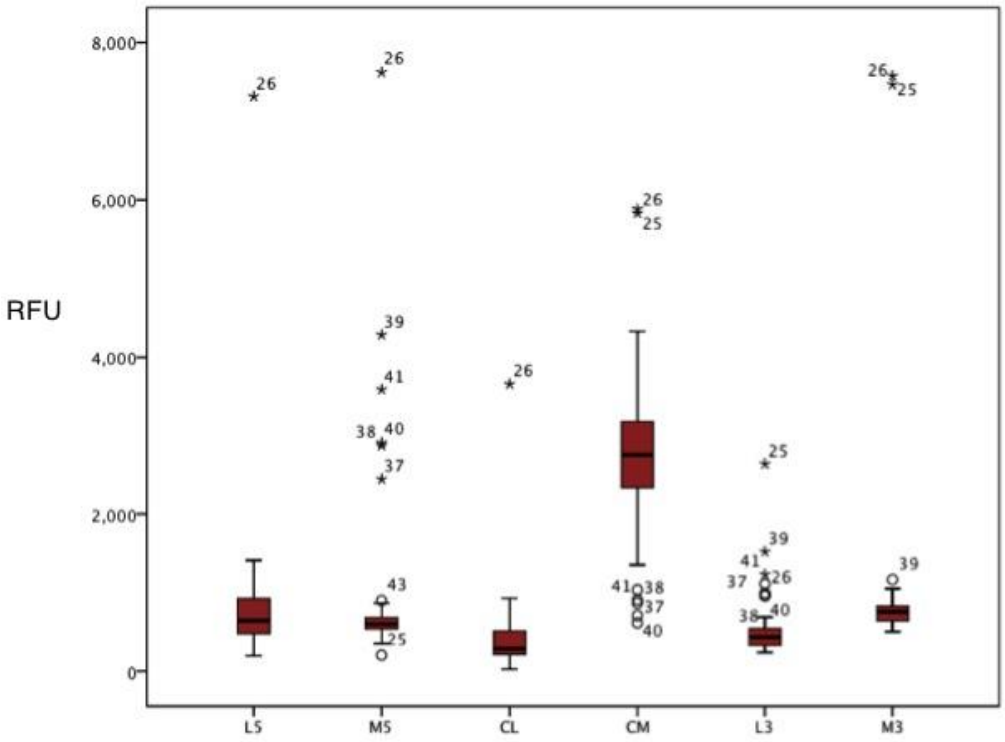


Figure 5.6. Box plot of RFU distribution for the six 18S rRNA fragments. L5 = 5' large fragment, M5 = 5' medium fragment, CL = central large fragment, CM = central medium fragment, L3 = 3' large fragment and M3 = 3' medium fragment.

Figure 5.6 shows that in the 18S rRNA assay, 10 RFU values were mild outliers. One sample at 20 days (sample 25) and one from 150 days (sample 43) were within 1.5x the IQR of the rest of the RFU values for the 5' medium fragment. Four samples at 150 days (samples 37, 38, 40, 41) were within 1.5x the IQR of the rest of the RFU values for the 3' long fragments. Sample 39 was 1.5x the IQR from the rest of the RFU values for the 3' medium fragment.

Fourteen RFU values were extreme outliers. Sample 26 at 20 days was 3x the IQR from the rest of the RFU values for the 5' long fragment. Samples 26, 37, 38, 39, 40 and 41 at 150 days were 3x the IQR from the rest of the RFU values for the 5' medium fragment. Sample 26 was 3x the IQR from the rest of the RFU values for central large fragment. Sample 25 and 26 at 20 days, were 3x the IQR from the rest of the RFU values for the central medium fragment. Samples 25, 26 and 39 were 3x the IQR from the rest of the RFU values for the 3' large fragment and samples 25 and 26 were 3x the IQR from the rest of the RFU values for the 3' medium fragment. In summary, all samples at 4 days were kept, 2 at 20 days were removed, and 6 at 150 days were removed.

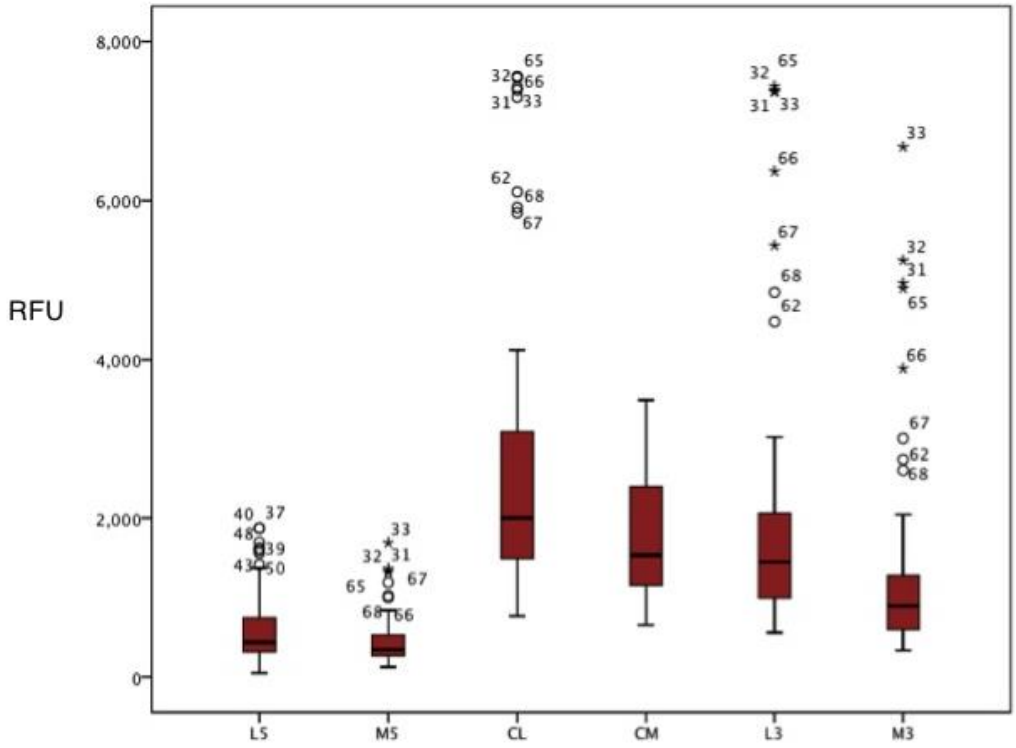


Figure 5.7. Box plot of RFU distribution for the six HBB DNA fragments. L5 = 5' large fragment, M5 = 5' medium fragment, CL = central large fragment, CM = central medium fragment, L3 = 3' large fragment and M3 = 3' medium fragment.

Figure 5.7 shows that in the HBB DNA assay, 23 RFU values were mild outliers. Samples 37, 39, 40, 43, 48 and 50 at 150 days were 1.5x the IQR

from the rest of the RFU values for the 5' large fragment. Samples 65, 66, 67, and 68 at 150 days were 1.5x the IQR from the rest of the RFU values for the 5' medium fragment. Samples 31, 32 and 33 at 20 days, and 62, 65, 66, 67 and 68 at 150 days were 1.5x the IQR from the rest of the RFU values for the central large fragment. Samples 62 and 69 were 1.5x the IQR from the rest of the RFU values for the 3' large fragment. Samples 62, 67 and 68 were 1.5x the IQR from the rest of the RFU values for the 3' medium fragment.

Fourteen RFU values were extreme outliers. Samples 31, 32 and 33 were 3x the IQR from the rest of the RFU values for the 5' medium fragment. Samples 31, 32, 33, 65, 66 and 67 were 3x the IQR from the rest of the RFU values for the 3' large fragment. Samples 31, 32, 33, 65 and 66 were 3x the IQR from the rest of the RFU values for the 3' medium fragment. In summary, all samples at 4 days were kept, 3 at 20 days were removed, and 10 at 150 days were removed.

This figure also shows that between samples and compared to the G6PD mRNA and 18S rRNA assays, the HBB DNA fragments were more variable which indicated that DNA fragment degradation might not be suitable for age prediction.

When outliers were detected, the raw RFU data and electropherograms were reviewed. The electropherograms of the samples in question were checked for fragment "pull up", odd peaks or artefacts produced by the dye matrix that could account for any increase in peak heights. Where samples exhibited higher RFU values and no "pull up" or obvious peak artefacts then the cDNA quantification data was reviewed. Standardising the amount of DNA or cDNA added to the assay limited the variation in fragment peak height as a result of different concentrations (Chapter 3, Table 3.4). The calculations for template concentration were checked. If there were no miscalculations at the input level (of which there were none), then the source of the outliers was deemed to have arisen as a result of different PCR efficiencies at the fragment amplification step and the corresponding samples were removed before additional analysis was performed.

As shown in Table 5.1, two samples were removed from the G6PD mRNA bloodstain RFU data (samples 46 and 47 at 150 days). A total of 66 samples of G6PD mRNA RFU data therefore remain for analysis; 20 at 4 days, 16 at 20 days and 30 at 150 days. Eight samples were removed from the 18S rRNA RFU data; samples 25 and 26 at 20 days, and samples 37, 38, 39, 40, 41 and 43 at 150 days. A total of 60 samples of 18S rRNA RFU data therefore remain for analysis; 20 at 4 days, 14 at 20 days 26 at 150 days. Thirteen samples were removed from the HBB DNA RFU data; samples 31 to 33 at 20 days, and 39, 40, 43, 48, 50, 62, 65, 66, 67 and 68 at 150 days. A total of 55 samples of HBB DNA RFU data therefore remain for analysis; 20 at 4 days, 13 at 20 days and 22 at 150 days.

Table 5.1. Summary of outlier analysis.

	Time point (days)	**Number of samples removed**	**Number of samples remaining (all 6 fragments)**
G6PD mRNA assay	4 20 150	0 0 2	20 16 30
Total		**2**	**66**
18S RNA assay	4 20 150	0 2 6	20 14 26
Total		**8**	**60**
HBB DNA assay	4 20 150	0 3 10	20 13 22
Total		**13**	**55**

5.3.3 Assumption of normality

The purpose of exploratory data analysis was to determine how the data was distributed as normality is a key assumption in regression modelling. A data set is normally distributed when the data follows a bell-shaped curve that is symmetric about its mean. Each RFU value remaining after the removal of outliers was analysed and tested for normality.

5.3.3.1 G6PD mRNA fragment assessment of normality

A one sample K-S test was performed on the G6PD fragment height distribution and the results of the test were significant with one exception, the 3' medium fragment. As a result, the implied null hypotheses that the distributions of fragments are normally distributed were rejected (Table 5.2).

Table 5.2. One-sample K-S test performed on G6PD fragment height distribution.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of L5 is normal with mean 774.64 and standard deviation 492.25.	One-Sample Kolmogorov-Smirnov Test	.001	Reject the null hypothesis.
2	The distribution of M5 is normal with mean 325.29 and standard deviation 287.27.	One-Sample Kolmogorov-Smirnov Test	.000	Reject the null hypothesis.
3	The distribution of CL is normal with mean 282.38 and standard deviation 258.19.	One-Sample Kolmogorov-Smirnov Test	.000	Reject the null hypothesis.
4	The distribution of CM is normal with mean 1,717.74 and standard deviation 1,528.34.	One-Sample Kolmogorov-Smirnov Test	.000	Reject the null hypothesis.
5	The distribution of L3 is normal with mean 220.77 and standard deviation 230.68.	One-Sample Kolmogorov-Smirnov Test	.000	Reject the null hypothesis.
6	The distribution of M3 is normal with mean 863.17 and standard deviation 322.98.	One-Sample Kolmogorov-Smirnov Test	.199	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

If the assumption of normality fails, it implies that any statistical methods later used in the development of a linear regression model (such as the F-test in analysis of variance) would be speculative if the data is used because the errors in sampling are not normally distributed. The assumption of normality holds especially true for data obtained from specific time points, such as peak heights, as errors in measurement might change with time.

Additional statistical analysis was then performed with a focus towards skewness and Kurtosis. Skewness describes the asymmetry of a distribution. A positive skewness shows an asymmetric distribution with the largest tail to the right, while a negative skewness shows an asymmetric distribution with the largest tail to the left (Brys, Hubert, & Struyf, 2008). The skewness value of a normal distribution is 0.

The skewness-kurtosis test for normality is designed to only reject specific departures from normality that fall outside of the confidence intervals associated with the variables, meaning that the test approximates a normal distribution rather than fit it exactly. The skewness-kurtosis test is designed to determine whether it is possible to use the RFU fragment data. Passing this test is sufficient justification to use the RFU fragment data even if the K-S test has failed. Kurtosis measures the concentration of observations in the middle of distribution that results in a peaked shape. The kurtosis value of a normal, standard distribution is 3. The general consensus is to treat a skewness value of between -2 and 2 and a kurtosis value of -3 to 3 as an approximate normal distribution. Table 5.3 shows the key descriptive statistics of the fragment distribution of G6PD mRNA.

Table 5.3. G6PD mRNA descriptive statistics. Standard deviation, Skewness and Kurtosis of G6PD mRNA fragments. L5 – 5' large, M5 – 5' medium, CL – central large, CM – central medium, L3 – 3' large, M3 – 3' medium.

Descriptive Statistics

	N	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
L5	66	492.24846	0.698	0.295	-0.485	0.582
M5	66	287.27204	-0.077	0.295	-1.899	0.582
CL	66	258.19431	-0.008	0.295	-1.827	0.582
CM	66	1528.33701	-0.082	0.295	-1.905	0.582
L3	66	230.67852	0.619	0.295	-1.066	0.582
M3	66	322.97581	0.233	0.295	-0.450	0.582

Sixty six observations (N) were included in the analyses for each fragment. The SD shows that the central medium fragment was the most variable, meaning that over the time course this fragment showed the greatest change in peak height. Skewness statistics ranged from -0.008 to 0.698 which is between -2 and 2. Kurtosis values ranged from -1.905 to -0.485, which even if the standard errors are included, demonstrate that they fall between -3 and 3 and approximate a normal distribution.

The results indicate that the distribution determined for the fragment peak heights approximate a normal distribution and were subsequently included in the tentative regression model.

5.3.3.2 18S rRNA fragment assessment of normality

Table 5.4 shows the results of the one sample K-S test performed on the remaining 60 RFU values for the 18S rRNA fragments. The results of all six fragments were not significant thus the null hypothesis that the distributions of fragments represent an approximate normal distribution was retained and the 18S rRNA fragments were included in the tentative regression model.

Table 5.4. 18S rRNA fragment normality test. One-sample K-S test performed on 18S rRNA fragment height distribution.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of L5 is normal with mean 688.40 and standard deviation 273.07.	One-Sample Kolmogorov-Smirnov Test	.147	Retain the null hypothesis.
2	The distribution of M5 is normal with mean 591.28 and standard deviation 93.56.	One-Sample Kolmogorov-Smirnov Test	.855	Retain the null hypothesis.
3	The distribution of CL is normal with mean 365.14 and standard deviation 207.42.	One-Sample Kolmogorov-Smirnov Test	.074	Retain the null hypothesis.
4	The distribution of CM is normal with mean 2,764.24 and standard deviation 559.55.	One-Sample Kolmogorov-Smirnov Test	.982	Retain the null hypothesis.
5	The distribution of L3 is normal with mean 422.40 and standard deviation 121.82.	One-Sample Kolmogorov-Smirnov Test	.690	Retain the null hypothesis.
6	The distribution of M3 is normal with mean 726.79 and standard deviation 125.18.	One-Sample Kolmogorov-Smirnov Test	.762	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

5.3.3.3 HBB DNA fragment assessment of normality

Table 5.5 show the results of the one sample K-S test performed on the remaining 55 RFU values for the HBB DNA fragments. The results of the six

fragments were not significant therefore the null hypothesis that the distributions of fragments represent an approximate normal distribution was retained and HBB DNA fragments were included in the tentative regression model.

Table 5.5 HBB DNA fragment normality test. One-sample K-S test performed on HBB DNA fragment height distribution.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of L5 is normal with mean 409.17 and standard deviation 131.63.	One-Sample Kolmogorov-Smirnov Test	.992	Retain the null hypothesis.
2	The distribution of M5 is normal with mean 276.90 and standard deviation 42.92.	One-Sample Kolmogorov-Smirnov Test	.834	Retain the null hypothesis.
3	The distribution of CL is normal with mean 1,489.40 and standard deviation 267.25.	One-Sample Kolmogorov-Smirnov Test	.502	Retain the null hypothesis.
4	The distribution of CM is normal with mean 1,227.03 and standard deviation 253.87.	One-Sample Kolmogorov-Smirnov Test	.970	Retain the null hypothesis.
5	The distribution of L3 is normal with mean 1,030.90 and standard deviation 228.63.	One-Sample Kolmogorov-Smirnov Test	.563	Retain the null hypothesis.
6	The distribution of M3 is normal with mean 628.50 and standard deviation 94.27.	One-Sample Kolmogorov-Smirnov Test	.395	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

5.3.4 Assumption of linearity

The pearson product-moment correlation coefficient (*R*) assesses the degree to which different quantitative variables are linearly related in a sample. One of the assumptions underlying the correlation test is that the fragment heights are bivariately normally distributed (Salkind & Green, 2003). Since all of the

independent variables (RFU values of the different nucleic acid fragments) of each assay approximate a normal distribution (section 5.3.3), this assumption was met. Sixty six G6PD mRNA, 60 18S rRNA and 55 HBB DNA RFU values and their corresponding fragments were tested with the bivariate correlations procedure which is used to generate *R* along with the significance levels of each fragment to determine the probability that the correlation is real and not a chance occurrence.

5.3.4.1 Assumption of linearity – G6PD mRNA fragments

A logarithmic time scale was used for all assays due to the large spread of possible predicted deposition times. This scale was chosen because it is able to display the correlation of both small and large changes of deposition times and the corresponding changes to the nucleic acid fragments (∆RFU) in each assay.

Table 5.6. G6PD mRNA fragment correlation analysis. Bivariate Correlations between ln_time and the G6PD mRNA independent variables.

Correlations

Statistical attribute	5' large	5' medium	Central large	Central medium	3' large	3' medium
Pearson Correlation	0.606**	-0.890**	-0.868**	-0.862**	-0.965**	0.589**
Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000
N	66	66	66	66	66	66

**. Correlation is significant at the 0.01 level (2-tailed).

Sixty six observations were included for analysis (N) and the results show that all six fragments were significant and correlated with time (Table 5.6). The *R* values (Pearson Correlation) ranged from 0.589 for the 3' medium fragment to -0.965 for the 3' large fragment and there was less than a 1/100 chance that the correlation was due to chance. The p-value (significance) was less than 0.01 for all six fragments. The R values also indicated that the 5' medium,

central large, central medium and 3' large fragments had a negative correlation with time. As the bloodstain aged, the height of these fragments reduced in RFU whereas the 5' large and 3' medium fragments increased.

As a result, null hypotheses H_{02-7} were rejected pending further development of the regression model. This meant that all six G6PD mRNA fragments had a linear correlation with time.

5.3.4.2 Assumption of linearity – 18S rRNA fragments

Sixty observations were included for analysis (N) and Table 5.7 shows that the correlation between time and the change in peak heights of the 5' medium, central large, 3' large and 3' medium fragments of 18S rRNA was significant. The R value for the 5' medium fragment was 0.310 and there was less than a 1 in 20 chance that the correlation was due to a random event as the p-value of this correlation was 0.018 (which is more than 0.01 but less than 0.05) which implies that it has a weak correlation to time. The R value also indicated that the RFU of this fragment increased with the age of the bloodstain. The central large, 3' large and 3' medium fragments had a stronger correlation with time as shown by the higher R values (0.643 to -0.860) and there was less than a 1 in 100 chance that the correlation was due to chance as the p-values of the correlation were less than 0.01 for all three fragments. The R values of the central large and 3' large fragments indicated that the RFU of these fragments reduced as the bloodstain aged whereas the 3' medium and 5' medium increased.

As a result, null hypotheses H09, 10, 12 and 13 were rejected pending further development of the regression model. The differences in peak height of the 5' large and central medium fragments with regard to time were not significant and hypotheses H08 and 11 were retained. This meant that the 5' medium, central large, 3' large and 3' medium fragments had a linear correlation with time whereas the 5' large and central medium fragments did not.

Table 5.7. 18S rRNA fragment correlation analysis. Bivariate Correlations between ln_time and the 18S rRNA independent variables.

Correlations

Statistical attribute	5' large	5' medium	Central large	Central medium	3' large	3' medium
Pearson Correlation	-0.063	0.310*	-0.821**	-0.161	-0.860**	0.643**
Sig. (2-tailed)	0.640	0.018	0.000	0.227	0.000	0.000
N	58	58	58	58	58	58

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

5.3.4.3 Assumption of linearity – HBB DNA fragments

Fifty five observations were included for analysis (N) and Table 5.8 shows that the correlation between time and the change in peak heights of the 5' medium and 3' large fragments of HBB DNA were significant. The R value for the 5' medium was 0.505 and there was less than a 1/100 chance that the correlation was due to chance. The p-value (Sig.) of this correlation was 0.004. The 3' large fragment had R value of -0.442 and there was less than a 1/20 chance that the correlation was due to chance. The p-value (Sig.) of this correlation was 0.014. The R values indicated that the RFU of the 5' medium fragment increased with age whereas the 3' large reduced.

Table 5.8. HBB DNA fragment correlation analysis. Bivariate Correlations between ln_time and the HBB DNA independent variables.

Correlations

Statistical attribute	5' large	5' medium	Central large	Central medium	3' large	3' medium
Pearson Correlation	-0.144	0.505**	-0.172	0.047	-0.442*	-0.242
Sig. (2-tailed)	0.448	0.004	0.364	0.806	0.014	0.197
N	55	55	55	55	55	55

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

As a result, null hypotheses H015 and 18 were rejected pending further development of the regression model. Null hypotheses H014, 16, 17 and 19 were retained as the 5' large, central large, central medium and 3' medium fragments did not correlate with time. This meant that the 5' medium and 3' large fragments had a linear correlation with time whereas the 5' large, central large, central medium and 3' medium did not.

To summarise, the 18S rRNA 5' large and central medium fragments were removed from the preliminary regression model due to their lack of linear relationship with time (Table 5.7) as were the 5' large, central large, central medium and 3' medium fragments of HBB DNA (Table 5.8). All of the G6PD mRNA fragments were included in the regression analysis (Table 5.6), as were the 5' medium, central large, 3' large and 3' medium fragments of 18S rRNA, and the 5' medium and 3' large fragments of HBB DNA.

5.3.5 Regression modelling

Two regression analyses were conducted to predict the time since deposition based on the change in fragment height of DNA, rRNA and mRNA.

One analysis included all twelve independent variables determined from the previous analyses (all of the G6PD mRNA fragments, the 5' medium, central large, 3' large and 3' medium fragments of 18S rRNA, and the 5' medium and 3' large fragments of HBB DNA) inserted into the model simultaneously.

The other analysis used a backward-stepwise approach, where variables previously entered simultaneously were removed one at a time if the variable's significance was found to be greater than 0.1 (as specified by the regression criteria in SPSS). The purpose of this latter method was to remove those fragments (of the twelve) that were not statistically significant to predict age. In both cases SPSS was used to solve the linear multiple regression equation (3)

Ln_time = a + *b*1(G6PD 5'L) + *b2*(G6PD 5'M) + *b3*(G6PD CL) + *b4*(G6PD CM) + *b5*(G6PD 3'L) + *b6*(G6PD 3'M) + *b7*(18S 5'M) + *b8*(18S CL) +*b9*(18S 3'L) + *b10*(18S 3'M) + *b11*(HBB 5'M) + *b12*(HBB 3'L)

(3)

Where:

a = constant

b = regression coefficient of each fragment

G6PD 5'L = peak height of G6PD 5' long fragment at time y

G6PD 5'M = peak height of G6PD 5' medium fragment at time y

G6PD CL = peak height of G6PD central large fragment at time y

G6PD CM = peak height of G6PD central medium fragment at time y

G6PD 3'L = peak height of G6PD 3' long fragment at time y

G6PD 3'M = peak height of G6PD 3' medium fragment at time y

18S 5'M = peak height of 18S 5' medium fragment at time y

18S CL = peak height of 18S central large fragment at time y

18S 3'L = peak height of 18S 3' large fragment at time y

18S 3'M = peak height of 18S 3' medium fragment at time y

HBB 5'M = peak height of HBB 5' medium fragment at time y

HBB 3'L = peak height of HBB 3' large fragment at time y

First, the combination of the twelve nucleic acid fragments entered simultaneously was significant. Table 5.9 shows that the multiple correlation coefficient (*R*) was 0.988, indicating that approximately 97.7 % of the variance (R square) in the age of the bloodstain could be accounted for by fragment peak height (a high degree of correlation). The adjusted R square is used in multiple regression to account for the use of more than predictor and indicates that 97.2 % of the variance can be accounted for by the combination of the twelve nucleic acid fragments.

Table 5.9. Regression model summary with 12 independent variables. Regression model statistics using twelve nucleic acid fragments.

Model Summary[b]

R	R Square	Adjusted R Square
0.988[a]	0.977	0.972

a. Predictors: (Constant), HBB_3L, G6PD_3M, HBB_5M, rRNA_5M_18S, G6PD_CL, rRNA_CL_18S, rRNA_3M_18S, rRNA_3L_18S, G6PD_5L, G6PD_3L, G6PD_CM, G6PD_5M

b. Dependent Variable: ln_time

5.3.5.1 Backward stepwise regression modelling

The backward-stepwise approach was conducted by removing one independent variable at a time based on its lack of significance ($p > 0.1$). The associated "change statistics" or the impact of removing one or more fragments are discussed below.

5.3.5.1.1 HBB DNA 3' large fragment removal

The variable HBB_3L (HBB DNA 3' large fragment) was removed with little change in the model (Table 5.10 Model 2). The change in the R^2 (R square) value was zero (no change from 0.977). The standard error of the estimate (SE) value was 0.26223 which meant that age prediction was more accurate when this fragment was removed as the standard deviation dropped from 0.26460. The *p* value (Sig. F change) of 0.851 meant that this fragment did not significantly predict age as the p value was more than 0.1.

Table 5.11 shows that with the removal of HBB_3L, the F statistic (F) was calculated as 208.605 (Table 5.11 Model 2) which was higher than 187.809 (Table 5.11 Model 1) and this change was significant as it was less than 0.05 (Sig. 0.000). Overall, this meant a higher predictive power was achieved by removing this fragment from the model.

Table 5.10. Regression model statistics using the Backward Stepwise Method.

Model Summary[a]

Model (fragment removed)	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1 None	0.988[a]	0.977	0.972	0.26460	0.977	187.809	12	53	0.000
2 (HBB_3L)	0.988[b]	0.977	0.972	0.26223	0.000	0.036	1	53	0.851
3 (18S_5M)	0.988[c]	0.977	0.973	0.26039	0.000	0.230	1	54	0.633
4 (HBB_5M)	0.988[d]	0.977	0.973	0.25868	0.000	0.267	1	55	0.608
5 (18S_3L)	0.988[e]	0.977	0.973	0.25769	0.000	0.565	1	56	0.455
6 (G6PD_5L)	0.988[f]	0.976	0.974	0.25633	0.000	0.390	1	57	0.535
7 (G6PD_CM)	0.988[g]	0.976	0.973	0.25848	-0.001	1.995	1	58	0.163
8 (G6PD_3M)	0.987[h]	0.974	0.972	0.26226	-0.001	2.768	1	59	0.101

a. Dependent Variable: ln_time

1. Predictors: (Constant), HBB_3L, G6PD_3M, HBB_5M, 18S_5M, G6PD_CL, 18S_CL, 18S_3M, 18S_3L, G6PD_5L, G6PD_3L, G6PD_CM, G6PD_5M
2. Predictors: (Constant), G6PD_3M, HBB_5M, 18S_5M, G6PD_CL, 18S_CL, 18S_3M, 18S_3L, G6PD_5L, G6PD_3L, G6PD_CM, G6PD_5M
3. Predictors: (Constant), G6PD_3M, HBB_5M, G6PD_CL, 18S_CL, 18S_3M, 18S_3L, G6PD_5L, G6PD_3L, G6PD_CM, G6PD_5M
4. Predictors: (Constant), G6PD_3M, G6PD_CL, 18S_CL, 18S_3M, 18S_3L, G6PD_5L, G6PD_3L, G6PD_CM, G6PD_5M
5. Predictors: (Constant), G6PD_3M, G6PD_CL, 18S_CL, 18S_3M, G6PD_5L, G6PD_3L, G6PD_CM, G6PD_5M
6. Predictors: (Constant), G6PD_3M, G6PD_CL, 18S_CL, 18S_3M, G6PD_3L, G6PD_CM, G6PD_5M
7. Predictors: (Constant), G6PD_3M, G6PD_CL, 18S_CL, 18S_3M, G6PD_3L, G6PD_5M
8. Predictors: (Constant), G6PD_CL, 18S_CL, 18S_3M, G6PD_3L, G6PD_5M

5.3.5.1.2 18S rRNA 5’ medium fragment removal

The 18S_5M variable (18S rRNA 5’ medium fragment) was then removed with similar effect. The change in the R^2 value was zero (Table 5.10 Model 3). The SE value was 0.26039 which meant that age prediction was more accurate when this fragment was removed and the *p* value (Sig. F change) of the 18S_5M fragment (0.633) meant that this fragment did not significantly predict age as it was more than 0.1.

Table 5.11 shows that with the removal of 18S_5M fragment the F statistic was calculated as 232.699 (compared to 208.605) (Table 5.11 Model 3) which means an even higher predictive power was achieved by removing this fragment.

A total of seven variables were removed (Table 5.10 and 5.11, models 2 to 8): HBB_3L, 18S_5M, HBB_5M, 18S_3L, G6PD_5L, G6PD_CM, and G6PD_3M.

Table 5.11. Backward stepwise regression analysis of variance.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	157.791	12	13.149	187.809	0.000[b]
	Residual	3.711	53	0.070		
	Total	161.502	65			
2	Regression	157.789	11	14.344	208.605	0.000[c]
	Residual	3.713	54	0.069		
	Total	161.502	65			
3	Regression	157.773	10	15.777	232.699	0.000[d]
	Residual	3.729	55	0.068		
	Total	161.502	65			
4	Regression	157.755	9	17.528	261.956	0.000[e]
	Residual	3.747	56	0.067		
	Total	161.502	65			
5	Regression	157.717	8	19.715	296.896	0.000[f]
	Residual	3.785	57	0.066		
	Total	161.502	65			
6	Regression	157.691	7	22.527	342.859	0.000[g]
	Residual	3.811	58	0.066		
	Total	161.502	65			
7	Regression	157.560	6	26.260	393.044	0.000[h]
	Residual	3.942	59	0.067		
	Total	161.502	65			
8	Regression	157.375	5	31.475	457.616	0.000[i]
	Residual	4.127	60	0.069		
	Total	161.502	65			

a. Dependent Variable: ln_time
1. Predictors: (Constant), HBB_3L, G6PD_3M, HBB_5M, 18S_5M, G6PD_CL, 18S_CL, 18S_CM, 18S_3L, G6PD_5L, G6PD_3L, G6PD_CM, G6PD_5M
2. Predictors: (Constant), G6PD_3M, HBB_5M, 18S_5M, G6PD_CL, 18S_CL, 18S_3M, 18S_3L, G6PD_5L, G6PD_3L, G6PD_CM, G6PD_5M
3. Predictors: (Constant), G6PD_3M, HBB_5M, G6PD_CL, 18S_CL, 18S_3M, 18S_3L, G6PD_5L, G6PD_3L, G6PD_CM, G6PD_5M
4. Predictors: (Constant), G6PD_3M, G6PD_CL, 18S_CL, 18S_3M, 18S_3L, G6PD_5L, G6PD_3L, G6PD_CM, G6PD_5M
5. Predictors: (Constant), G6PD_3M, G6PD_CL, 18S_CL, 18S_3M, G6PD_5L, G6PD_3L, G6PD_CM, G6PD_5M
6. Predictors: (Constant), G6PD_3M, G6PD_CL, 18S_CL, 18S_3M, G6PD_3L, G6PD_CM, G6PD_5M
7. Predictors: (Constant), G6PD_3M, G6PD_CL, 18S_CL, 18S_3M, G6PD_3L, G6PD_5M
8. Predictors: (Constant), G6PD_CL, 18S_CL, 18S_3M, G6PD_3L, G6PD_5M

No other variables could be removed without significantly reducing the predictive power of the model, as measured by the R^2 value or the F statistic. As a result, the tentative model contains five fragments (The 5' medium, central large and 3' large fragments of G6PD mRNA, and the 3' medium and central large fragments of 18S rRNA) which were found to show statistically significant changes over time as measured by the R^2, the F and p values . Table 5.10 (Model 8) shows that the R^2 value of the final model was 0.974 and the adjusted R^2 value was 0.972. This means that 97.2 % of the variance can be accounted for by the combination of the G6PD central large, 3' large, 5' medium fragments and the 18S central large and 3' medium fragments. Table 5.11 (Model 8) shows the regression model's significance statistic for the F-test ($F= 457.616$, $p < 0.001$) indicates that there is less than a 1 in 1,000 chance that the observed correlation between the remaining five fragments and time is due solely to random sampling error.

5.3.5.2 Forward stepwise regression modelling

The forward stepwise regression procedure was used to assess each of the remaining fragments (The 5' medium, central large and 3' large fragments of G6PD mRNA, and the 3' medium and central large fragments of 18S rRNA) in the following way: each time a fragment was added to the model, the significance of the other fragments in the model were re-examined. That is, at each step in the forward selection procedure, the significance of each of the fragments currently in the model was tested.

Table 5.12 shows that the 3' large fragment of G6PD mRNA (Model 1) accounted for 93.1 % of the variance (as measured by the R square change). Model 2 shows that the 5' medium fragment accounted for 3.3 % and the 18S rRNA central large fragment for 0.6% (Model 3). The central large fragment of G6PD mRNA accounted for 0.3 % (Model 4) whilst the 18S rRNA 3' medium fragment accounted for an additional 0.2 % of the variance (Model 5).

Table 5.12. Regression model summary of five predictor variables. Regression model summary using the Forward Stepwise method showing the contribution of each fragment to the overall model.

Model Summary[a]

Model	Change Statistics				
	R Square Change	F Change	df1	df2	Sig. F Change
1	0.931	865.280	1	64	0.000
2	0.033	56.761	1	63	0.000
3	0.006	12.461	1	62	0.001
4	0.003	5.722	1	61	0.020
5	0.002	4.759	1	60	0.033

a. Dependent Variable: ln_time
1. Predictors: (Constant), G6PD_3L
2. Predictors: (Constant), G6PD_3L, G6PD_5M
3. Predictors: (Constant), G6PD_3L, G6PD_5M, rRNA_CL_18S
4. Predictors: (Constant), G6PD_3L, G6PD_5M, rRNA_CL_18S, G6PD_CL
5. Predictors: (Constant), G6PD_3L, G6PD_5M, rRNA_CL_18S, G6PD_CL, rRNA_3M_18S

5.3.5.3 Proposed model: conclusions

Since the tentative regression models' R^2 is not equal to zero (Table 5.10 model 8) with a *p* value less than 0.05 (Table 5.11 model 8), the primary hypothesis H_{01} was rejected. This meant that there is a linear relationship between fragment degradation and time. Formula (4) shows the proposed regression model before multicollinearity analysis.

$$Ln\_time = c + B1(G6PD\_5M) + B2(G6PD\_CL) + B3(G6PD\_3L) + B4(18S\_CL) + B5(18S\_3M) \quad (4)$$

The parameters that correspond to each variable can be found in Table 5.13. The table shows the coefficients used for each fragment to predict age that were obtained from the backward stepwise regression method (Table 5.10 and 5.11

model 8). The constant represents the Y intercept (the predicated age value when all fragments are 0). The "B" coefficients are the values for the regression equation (4) for predicting the age of a bloodstain from the fragments. The p values show the significance of each fragment as they relate to age predication (all significant at $p < 0.05$)

Table 5.13. Regression model parameter estimates.

Coefficients[a]

	Coefficients		t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error			Lower Bound	Upper Bound
(Constant)	**(c)** 4.65347	0.30444	15.285	0.000	4.045	5.262
G6PD_5M	**(1)** -0.00392	0.00077	-5.119	0.000	-0.005	-0.002
G6PD_CL	**(2)** 0.00220	0.00084	2.612	0.011	0.001	0.004
G6PD_3L	**(3)** -0.00378	0.00037	-10.102	0.000	-0.005	-0.003
rRNA_CL_18S	**(4)** -0.00085	0.00030	-2.863	0.006	-0.001	0.000
rRNA_3M_18S	**(5)** 0.00079	0.00036	2.181	0.033	0.000	0.002

a. Dependent Variable: ln_time

5.3.6 Assumption of independence

5.3.6.1 Test for multicollinearity

Multicollinearity relates to a situation where two or more independent variables in the model are correlated and provide redundant information about the response. A VIF value obtained in the test of over 10 and a tolerance below 0.1 indicates such a result. All of the fragments from the final model were tested at once and Table 5.14 shows the multicollinearity coefficients for the main nucleic acid fragments. The Table shows that a correlation exists between the G6PD 5' medium and central large fragments (VIF of 45.701 and 44.502 respectively and a tolerance of 0.022 for both). The model was consequently re-analysed with the omission of each fragment in turn.

Table 5.14. Multicollinearity coefficients for the main effect independent variables.

Coefficients[a]

Variables	Collinearity Statistics	
	Tolerance	VIF
G6PD_5M	0.022	**45.701**
G6PD_CL	0.022	**44.502**
G6PD_3L	0.142	**7.041**
rRNA_CL_18S	0.315	**3.170**
rRNA_3M_18S	0.593	**1.687**

a. Dependent Variable: ln_time

Removal of the G6PD 5' medium fragment

When the G6PD 5' medium fragment was removed from the data set for analysis, the regression statistics were re-calculated once again with the backward stepwise method. The result was a poorer predictive model (R^2 of 0.971 compared to 0.974 and an F statistic of 332.235 compared to 457.616). The removal of the G6PD 5' medium fragment also significantly weakened the model which meant that fragments that were previously determined to be non-significant (Table 5.10) were re-entered altogether. Table 5.15 shows that G6PD central long fragment was removed to counteract this effect. The G6PD central medium and 3' medium fragments, and the 18S rRNA 5' medium fragment were also re-entered. Although the VIF values had been reduced, the 18S 5'medium and 3' medium fragments were not significant ($p > 0.05$) so the proposed model with the G6PD 5' medium fragment removed was rejected.

Table 5.15. Coefficient summary for the main effect independent variables with the removal of the G6PD 5' medium fragment.

	Sig.	Collinearity Statistics	
		Tolerance	VIF
G6PD_CM	0.000	0.284	**3.523**
G6PD_3L	0.000	0.122	**8.219**
G6PD_3M	0.005	0.616	**1.623**
rRNA_5M_18S	0.071	0.639	**1.564**
rRNA_CL_18S	0.000	0.293	**3.412**
rRNA_3M_18S	0.090	0.441	**2.270**

Removal of the G6PD central large fragment

The model was re-analysed with the removal of the G6PD central large fragment (Table 5.16). The result was a better predictor model. The F statistic was 547.006 compared to 457.616, with a slight decrease in R^2 to: 0.973 compared to 0.974. The VIF values of the new model were also all below 10, the tolerance values were all higher than 0.1 and the fragments were all significant ($P < 0.05$) (Table 5.16).

As a result, by removing the G6PD central long fragment, the correlation with the G6PD 5' medium fragment was removed and it was concluded that there was no interrelationship between the remaining independent variables and the model was accepted.

Table 5.16. Coefficient summary for the main effect independent variables with the removal of the G6PD central large fragment.

	Sig.	Collinearity Statistics	
		Tolerance	VIF
G6PD_5M	0.000	0.302	**3.311**
G6PD_3L	0.000	0.147	**6.810**
G6PD_3M	0.011	0.657	**1.522**
rRNA_CL_18S	0.000	0.325	**3.074**

5.3.7. Assumption of constant variance (homoscedasticity)

Homoscedasticity refers to the assumption that that the dependent variables (nucleic acid fragments) exhibit similar amounts of variance across a range of values for the independent variable (time) (Tuffery, 2011). Table 5.17 shows the result of Levene's Test of Homogeneity of Variance, which tests for similar variances (Tuffery, 2011) for the model. If the significance value is greater than

0.05 (found in the Sig. column) then the assumption of constant variance is met. Table 5.17 shows that Levene's F Statistic has a significance value of 0.128 for the G6PD 5' medium fragment, 0.093 for the G6PD 3' Large fragment, 0.232 for the G6PD 3' medium fragment and 0.076 for the 18S rRNA central large fragment. The assumption of constant variance was therefore met. This is important because the constant variance assumption means that no matter what time point is chosen (or no matter how old the stain is), the corresponding RFU values of the G6PD 5' medium, 3' large, 3' medium and 18S central long fragments will always have the same variation.

Table 5.17. Test for Homogeneity of Variances.

	Levene Statistic	**df1**	**df2**	**Sig.**
G6PD_5M	**2.138**	2	55	.128
G6PD_3L	**2.478**	2	55	.093
G6PD_3M	**1.503**	2	55	.232
rRNA_CL_18S	**2.837**	2	27	.076

5.3.8 Regression summary

Table 5.18 shows the final regression model summary for bloodstains stored on cotton cloth at ambient temperatures with exposure to sunlight. Since the regression models' R^2 value was not equal (0.973) to zero with a *p* value of less than 0.05, the primary hypothesis H_{01} was once again rejected. Formula (5) shows the final regression model after the collinearity and homoscedasticity testing.

A summary of the final model and the parameters that correspond to each variable can be found in Table 5.18.

Table 5.18. Final regression model summary for bloodstains stored on cotton cloth at ambient temperatures with exposure to sunlight.

$$Ln\_time = c + B1(G6PD\_5M) + B2(G6PD\_3L) + B3(G6PD\_3M) + B4(18S\_CL) \quad (5)$$

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
0.986[a]	0.973	0.971	0.26797

a. Predictors: (Constant), rRNA_CL_18S, G6PD_3M, G6PD_5M, G6PD_3L

ANOVA[a]

	Sum of Squares	df	Mean Square	F	Sig.
Regression	157.121	4	39.280	547.006	0.000[b]
Residual	4.380	61	0.072		
Total	161.502	65			

a. Dependent Variable: ln_time

b. Predictors: Predictors: (Constant), G6PD_3M, rRNA_CL_18S, G6PD_3L, G6PD_5M

Coefficients[a]

	Coefficients		t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error			Lower Bound	Upper Bound
(Constant)	4.9945190	0.14987	33.326	0.000	4.695	5.294
G6PD_5M	-0.0018729	0.00021	-8.896	0.000	-0.002	-0.001
G6PD_3L	-0.0036217	0.00038	-9.632	0.000	-0.004	-0.003
G6PD_3M	0.0003322	0.00013	2.616	0.011	0.000	0.001
rRNA_CL_18S	-0.0012285	0.00030	-4.095	0.000	-0.002	-0.001

a. Dependent Variable: ln_time

5.4 Conclusion

Through regression modelling, it was found that the height of the G6PD 5' medium fragment, 3' large fragment, 3' medium and the 18S central large fragment in combination, were statistically significant as they relate to bloodstain age. This study shows that both rRNA and mRNA markers are suitable indicators of age. Although DNA did exhibit changes in peak height over time (Figure 5.1), the HBB DNA fragments had a limited correlation with time (Table 5.8) and the fragments heights were variable (Figure 5.6), regression analysis showed that these DNA fragments were not suitable for age prediction.

This study took into account the entire length of each transcript and investigated how parts of it degrade as a function of time. By looking at specific amplicons throughout the sequence, linear regression analysis revealed that certain fragments could be used to predict age. Regression analysis is dependent on four principle assumptions: assumption of linearity, assumption of independence, assumption of constant variance, and assumption of normality. As these four assumptions were met, a combination of the 5' medium fragment, the 3' large and 3' medium fragments of G6PD mRNA and the central large fragment of 18S rRNA appear to be suitable markers for estimating age of a bloodstain.

The results presented here should be considered a pilot study as only three time points were used for analysis (t = 4, 20 and 150 days). Additional time points and bloodstains (>34) would help improve the predictive power and resolution of the regression model. The model is also based on bloodstains stored at room temperature with exposure to sunlight, further analysis would be required for different storage conditions.

Chapter 6

A Method to Determine the Age of a Bloodstain – a Study on Forensic Type Samples

6.1 Introduction

The focus of the studies done in this chapter are aimed at analysing RNA isolated from dried bloodstains using the regression model (prediction equation) developed in the Chapter 5. The main focus is to test the prediction equation on forensic type samples and determine its applicability for practical forensic use.

6.2 Methods

6.2.1 Blood samples

Twenty microliters of blood were deposited on various substrates (examples of the different substrates are shown in Figure 6.1) and stored under different environmental conditions (described in Table 6.1) for varying lengths of time up to a period of 180 days. Blind samples were also included for analysis. The age of these blind samples were only known to a third party and were used to assess the accuracy of the prediction equation. In order to replicate real-world scenarios, the various substrates used in this study were not sterilised in any manner. Bloodstains on cotton cloth and cotton swabs were cut away and placed directly into the DNA-IQ lysis buffer contained in the DNA IQ™ kit lysis. All other bloodstains were swabbed from the substrate using sterile cotton swabs dipped in DNA-IQ lysis buffer. The swab heads were then placed in lysis buffer. All samples were processed as described in Sections 2.2 to 2.5. Samples were quantified as described in Sections 3.3.3 to 3.3.5. Six hundred picograms of CD3G cDNA from each extract was added to the G6PD mRNA and 200 pg of CD3G cDNA to the 18S rRNA assays for analysis as described in section 4.2.6.

Table 6.1. Bloodstain storage conditions. Bloodstains were stored on a variety of substrates and under different environmental conditions.

Substrate	Temperature	Exposure to direct sunlight
Sterile cotton cloth (n=14)	Ambient room temperature	yes
Sterile cotton cloth (n=4)	37°C	no
Sterile cotton cloth (n=2)	4°C	no
Sterile cotton cloth (n=2)	Ambient room temperature	no
Wood (n=4)	Ambient room temperature	yes
Leather (n=3)	Ambient room temperature	yes
Plasterboard (n=4)	Ambient room temperature	yes
Wool carpet (n=4)	Ambient room temperature	yes
Cement paving slab (n=2)	Ambient room temperature	yes
Cotton swab*[1] (n=3)	Ambient room temperature	no
Glass*[2] (n=1)	Ambient room temperature	no
Sports shoe (n=2)	Ambient room temperature	yes

*[1]blind samples “K”, “A”, “B”
*[2] blind sample “S”

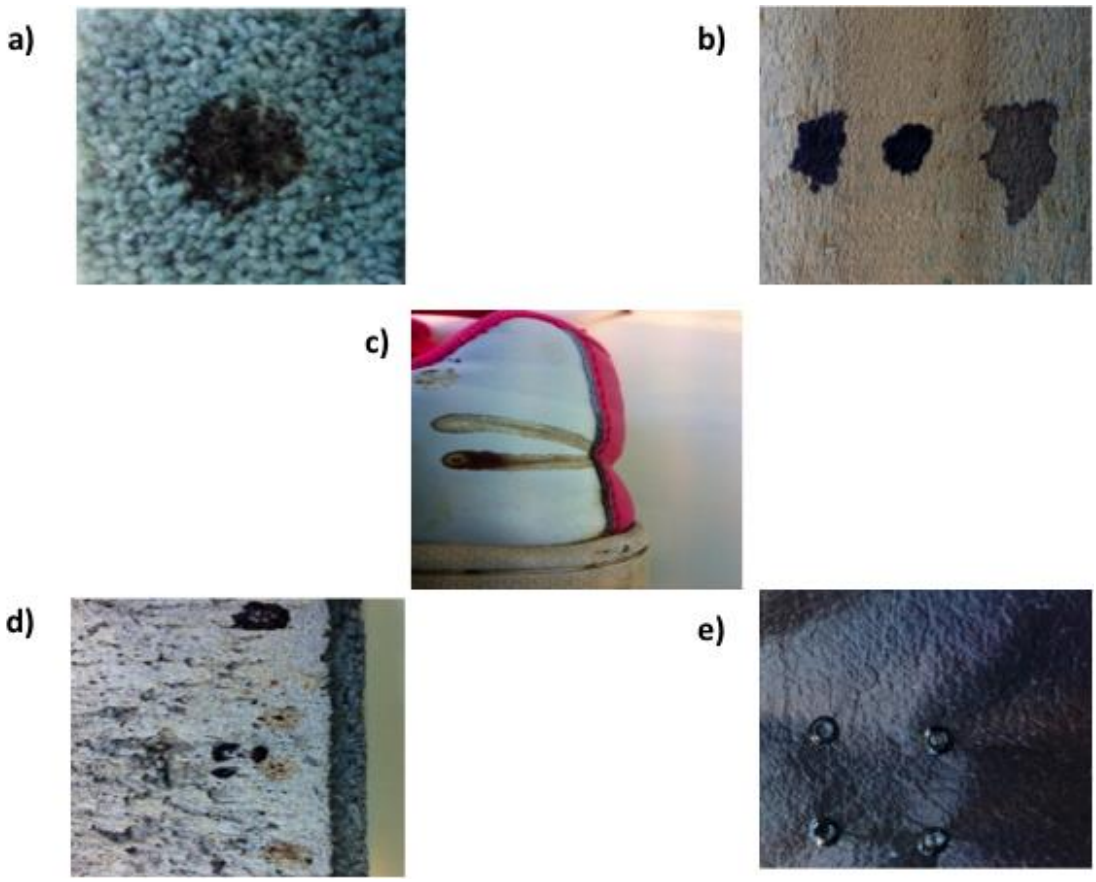


Figure 6.1. **Examples of substrates used for blood deposition.** 20 µL of blood on a) wool carpet, b) wood, c) sports shoe, d) cement paving slab and e) leather.

6.2.2 Age prediction

The peak heights of the G6PD 5' medium, 3'medium, 3' large fragment and the 18S central large fragments were obtained from each extract analysed (Tables 6.2 and 6.3), the values were substituted into the prediction equation derived from the previous chapter (Equation 5, Chapter 5). The peak heights of each fragment were multiplied by their corresponding coefficient and these were obtained from Table 5.18. The fragment peak heights were not averaged for analysis as individual values were substituted into the equation. Where fragments were completely degraded, they were prescribed an RFU value of 10 (baseline). The prediction equation and model coefficients are shown in Table 6.4 for convenience.

Table 6.2. Raw fragment RFU data for age prediction when deposited on cotton cloth. Where a RFU of 10 is specified the fragment has been degraded (baseline).

Condition	G6PD 5' medium RFU	G6PD 3' large RFU	G6PD 3' medium RFU	18S central large RFU
Ambient temperature, light	614.5	641	651.5	522
	614.5	641	651.5	505.5
	570.5	477.5	648.5	693
	570.5	477.5	648.5	545
	460	271.5	727	271.5
	460	271.5	727	369.5
	625	164	792	329.5
	638	245	813.5	231.5
	15	18.5	269	191.5
	100	10	185.67	187.33
	10	10	1164	177.5
	10	10	1242.5	270.5
	10	10	1272	176
	43.63	10	1630.75	171
4°C	144	676	604	310
	145	221	453	527
37°C	195	794	820	209
	961	330	758	493
	184	10	553	429
	750	10	663	532
Dark	289	10	301	398
	278	10	735	394

Table 6.3. Raw fragment RFU data for age prediction when deposited on different substrates. Where a RFU of 10 is specified the fragment has been degraded (baseline).

Substrate	G6PD 5' medium RFU	G6PD 3' large RFU	G6PD 3' medium RFU	18S central large RFU
Wool carpet	238	10	2218	222
	275	10	4081	193
	218	10	1106	167
	148	10	955	80
Plasterboard	108	10	2603	225
	295	10	827	256
	320	10	1392	155
	114	10	3333	250
Wood	258	10	1479	183
	340	10	2228	184
	207	10	1185	102
	130	10	1992	129
Leather	171	10	4520	278
	54	10	1817	60
	200	10	1143	277
Cement paving slab	574	748	1022	305
	274	390	532.5	407
Sports shoe	114	920	1131	379
	259	308	453	246

Table 6.4 Regression model equation and associated coefficients used for age prediction.

Ln_time = c + B1(G6PD_5M) + B2(G6PD_3L) + B3(G6PD_3M) + B4(18S_CL)

	Coefficient	Value
Intercept	*c*	4.994519031
G6PD_5M	*B1*	-0.001872919
G6PD_3L	*B2*	-0.003621702
G6PD_3M	*B3*	0.000332205
18S_CL	*B4*	-0.001228488

6.3 Results

6.3.1 Bloodstain storage on different substrates and conditions

Table 6.5 shows that RNA, which can be converted into cDNA, can be recovered from bloodstains deposited on different substrates for varying lengths of time. As up to 600 pg of CD3G cDNA was required for analysis (as described in Section 4.2.6), this was easily obtained from each substrate.

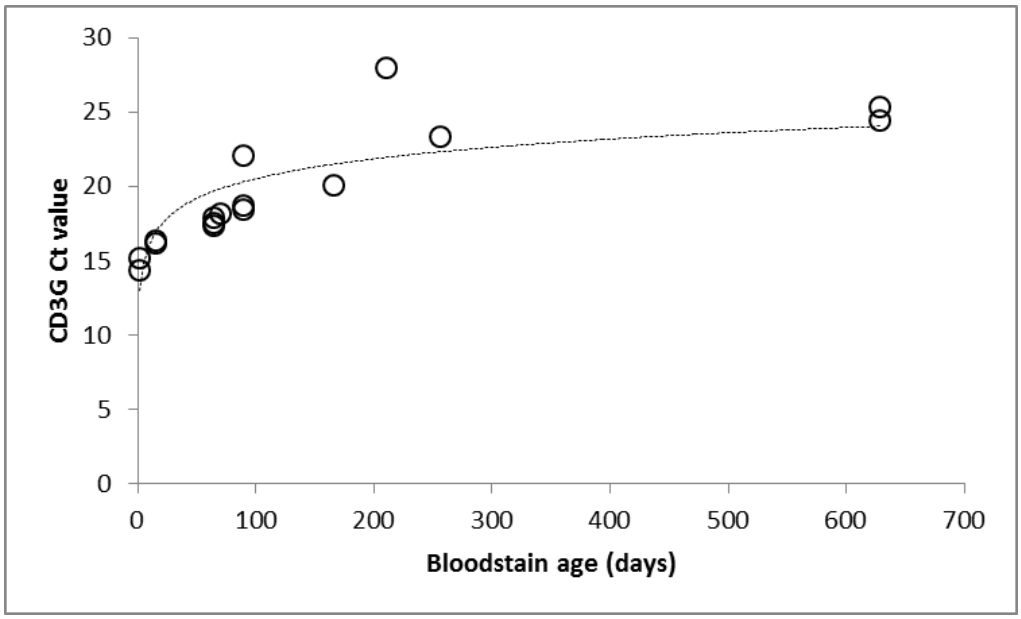
Figure 6.2 shows the correlation between bloodstain age and the Ct value of CD3G, and the concentration of CD3G. Figure 6.2A shows that as the age of the bloodstain increased, there was an associated increase in the Ct value of CD3G. This meant that in older bloodstains, the amount of CD3G had been reduced (degraded) and this was shown in Figure 6.2B. Figure 6.2B shows that the concentration of CD3G reduced as a function of time.

Example electropherograms of different bloodstain ages on various substrates are shown in Figures 6.3 to 6.9. The results of analysing the peak heights in the prediction equation are shown in Tables 6.6 and 6.7.

Table 6.5. Results of CD3G cDNA quantification assay from RNA obtained from different substrates.

Sample	Ct	Qty (ng/µl)
Wool carpet 64 days	18.19	0.568
	16.81	1.440
Wool carpet 90 days	19.76	0.199
	17.63	0.826
Plasterboard 64 days	18.08	0.613
	17.66	0.811
Plasterboard 90 days	18.58	0.439
	18.33	0.519
Wood 166 days	19.85	0.187
	20.31	0.138
Wood 70 days	17.24	1.080
	19.13	0.304
Leather 64 days	16.02	2.440
	18.68	0.409
Leather 90 days	21.23	0.074
	22.85	0.025
Paving slab 15 days	15.78	2.850
	16.63	1.620
Paving slab 2 days	14.53	6.590
	14.21	8.170
Sports shoe 15 days	16.11	2.280
	16.65	1.600
Sports shoe 2 days	16.44	1.840
	13.94	9.790
37°C 150 days	18.39	0.498
	18.18	0.574
37°C 7 days	15.21	4.190
	14.88	5.210
4°C 8 days	14.82	5.420
	14.24	8.000
180 days in the dark	17.79	0.745
	16.55	1.710
Blind sample “K”	23.32	0.11
Blind sample “A”	25.33	0.04
Blind sample “B”	24.45	0.06
Blind sample “S”	28.00	0.10

a)



b)

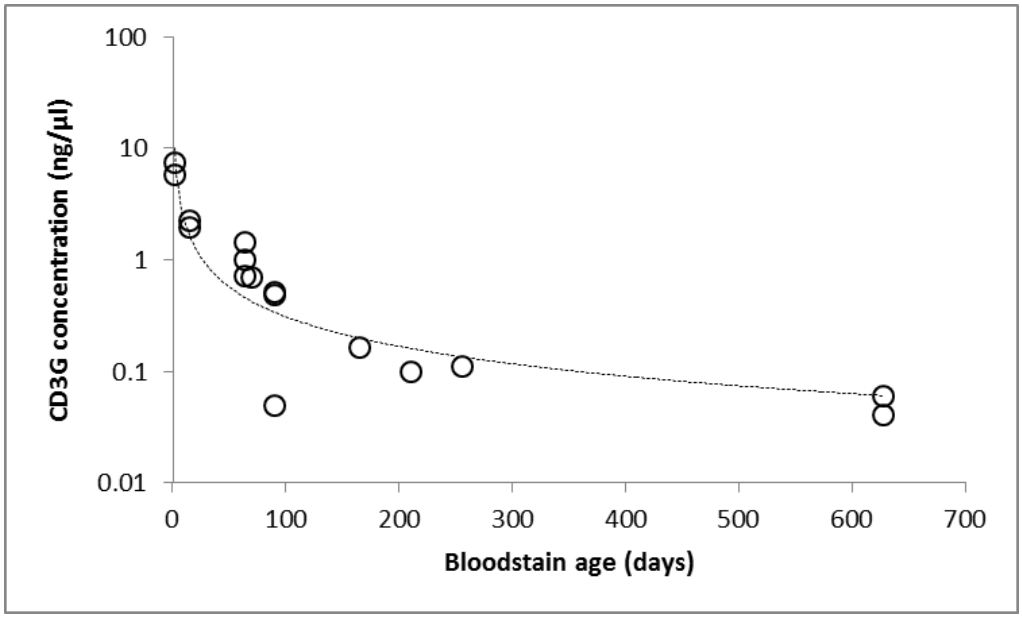


Figure 6.2. Correlation between bloodstain age and CD3G qPCR results. A) bloodstain age and CD3G cDNA Ct value, b) bloodstain age and CD3G cDNA concentration.

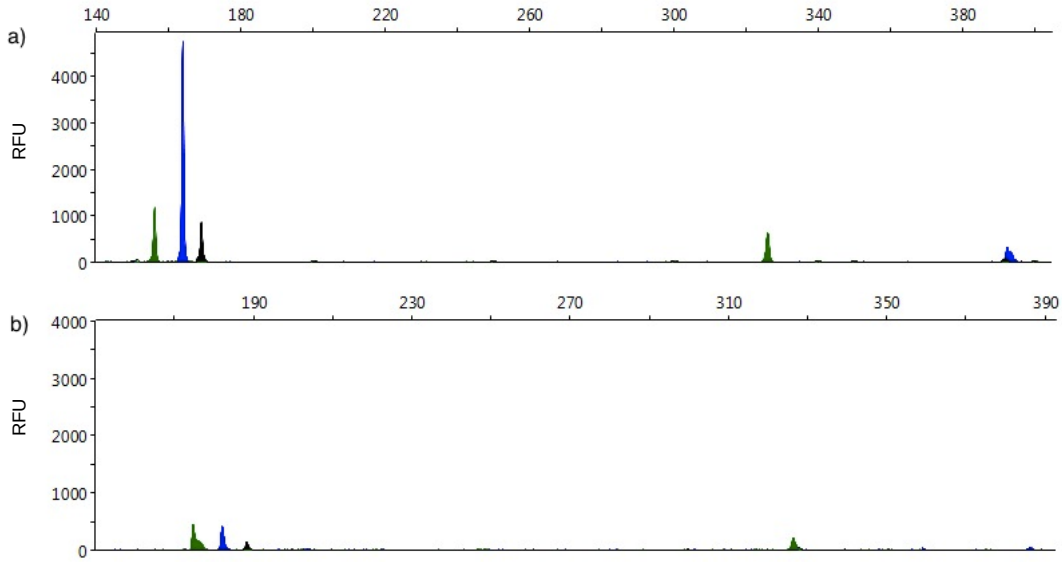


Figure 6.3. Electropherogram of an 8-day-old bloodstain. Electropherogram of an 8-day-old bloodstain stored at 4°C on cotton cloth and then analysed with the a) 18S rRNA multiplex assay, b) G6PD mRNA multiplex assay.

Figure 6.3 shows that all six 18S rRNA fragments could be amplified in eight day old bloodstains stored at 4°C, whereas only five G6PD mRNA fragments could be amplified under the same conditions (no 5' large fragment). Considering the age of the bloodstain, the RFU values of the fragments were lower than expected (compared to Figure 6.6 for example) which meant that the lower temperature and an increased drying time attributed to more degradation of the nucleic acid fragments.

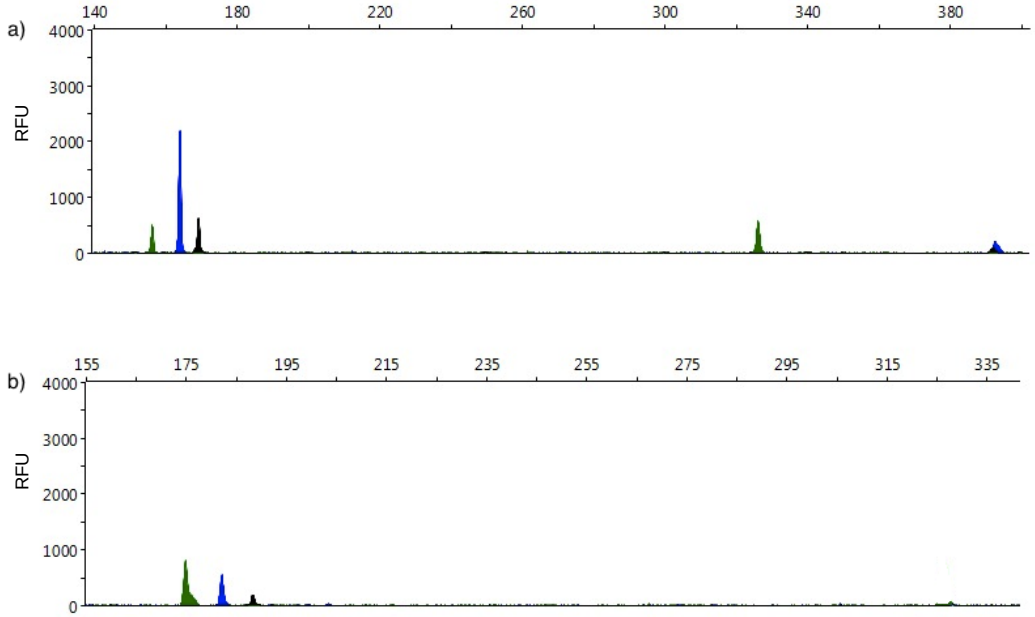


Figure 6.4. Electropherogram of a 150-day-old bloodstain. Electropherogram of a 150-day-old bloodstain stored at 37°C on cotton cloth and then analysed with the a) 18S rRNA multiplex assay, b) G6PD mRNA multiplex assay.

Figure 6.4 shows that in a 150 day old bloodstain stored at 37°C, all six 18S rRNA were amplified whereas only three G6PD mRNA fragments could be amplified under the same conditions (loss of the 3', 5' and central large fragments).

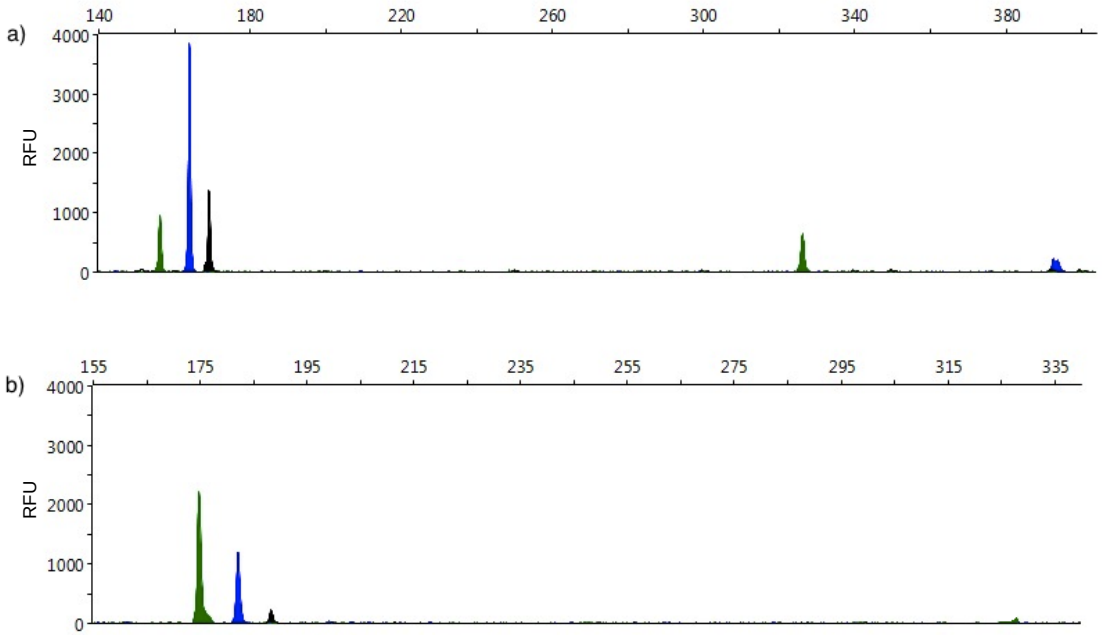


Figure 6.5. Electropherogram of a 64-day-old bloodstain. Electropherogram of a 64-day-old bloodstain stored at ambient room conditions on wool carpet exposed to sunlight and then analysed with the a) 18S rRNA multiplex assay, b) G6PD mRNA multiplex assay.

Figure 6.5 shows that in a sixty-four day old bloodstain stored on wool carpet with exposure to sunlight at ambient room temperature, all six 18S rRNA fragments could be amplified whereas only three G6PD mRNA fragments could be amplified under the same conditions (loss of the 3', 5' and central large fragments).

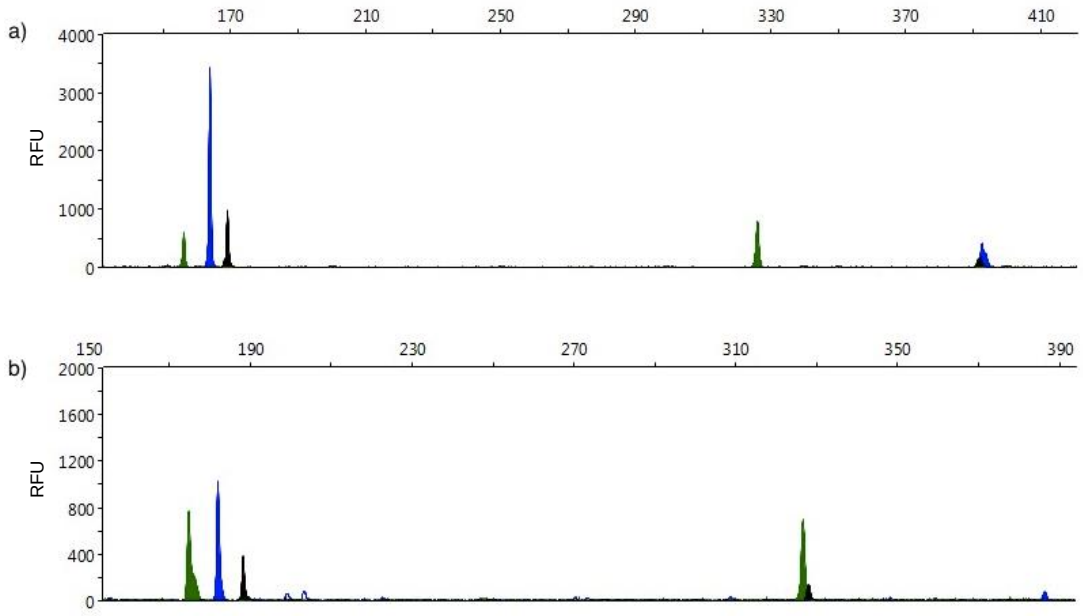


Figure 6.6. Electropherogram of a 2-day-old bloodstain. Electropherogram of a 2-day-old bloodstain stored at ambient room conditions on a cement-paving slab exposed to sunlight and then analysed with the a) 18S rRNA multiplex assay, b) G6PD mRNA multiplex assay.

Figure 6.6 shows that in a two day old bloodstain stored on a cement paving slab, exposed to sunlight and ambient room temperature, all six 18S rRNA and G6PD mRNA fragments could be amplified.

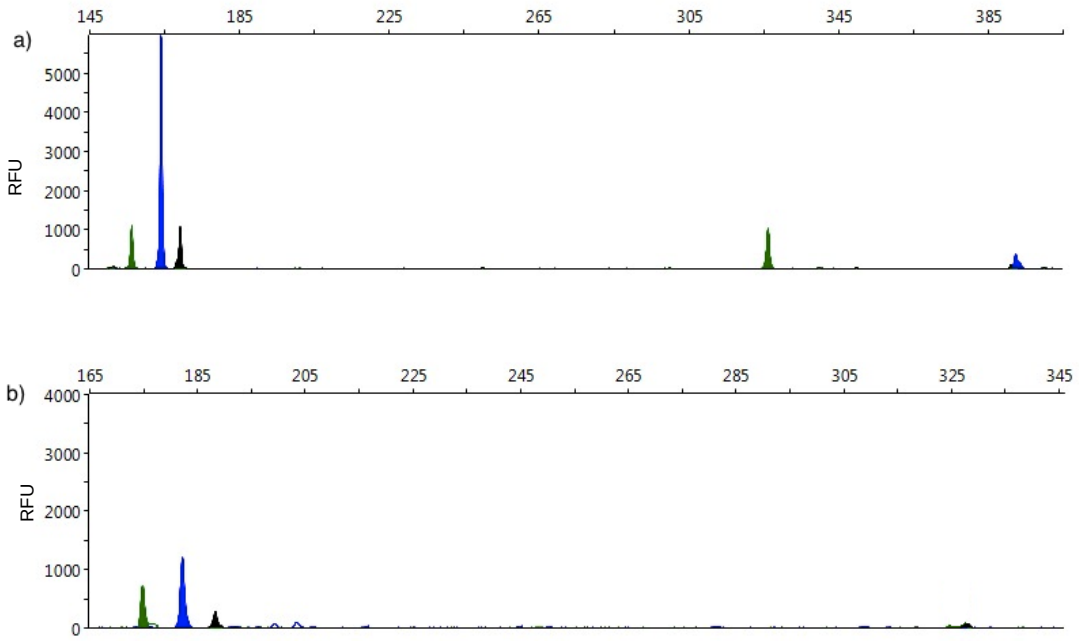


Figure 6.7. Electropherogram of a 180-day-old bloodstain. Electropherogram of a 180-day-old bloodstain stored at ambient room conditions on cotton cloth in the dark and then analysed with the a) 18S rRNA multiplex assay, b) G6PD mRNA multiplex assay.

Figure 6.7 shows that in a six month old bloodstain, stored in the dark and ambient room temperature, all six 18S rRNA could be amplified whereas there was a loss of the 3', 5' and central large G6PD mRNA fragments under the same conditions.

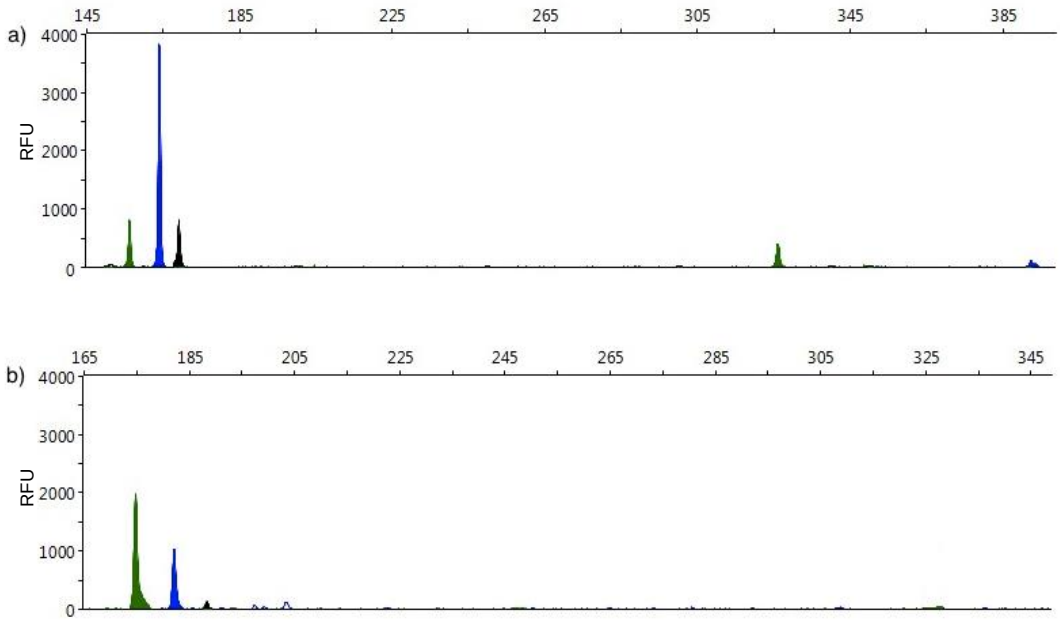


Figure 6.8. Electropherogram of a 64-day-old bloodstain deposited on leather. Electropherogram of a 64-day-old bloodstain stored at ambient room conditions on leather exposed to sunlight and then analysed with the a) 18S rRNA multiplex assay, b) G6PD mRNA multiplex assay.

Figure 6.8 shows that in a two month old bloodstain stored on leather that was exposed to sunlight and ambient temperatures, five 18S rRNA fragments were amplified (loss of the 5' large fragment) and three G6PD mRNA fragments could be amplified under the same conditions. Compared to Figure 6.5, Figure 6.8 shows that different substrates (at equivalent time points) have different effects on nucleic acid fragment degradation.

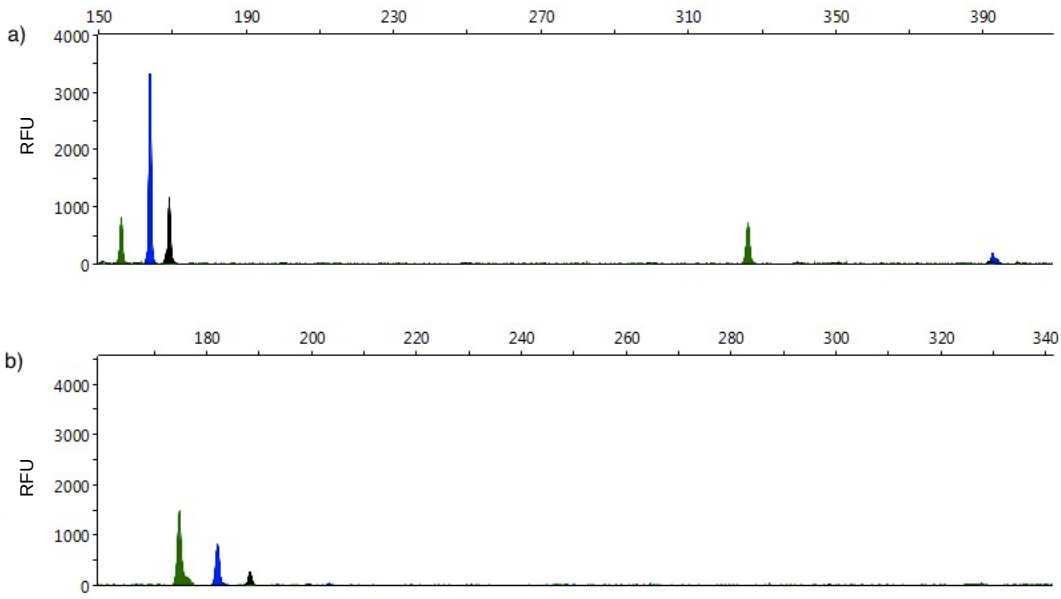


Figure 6.9. Electropherogram of a 166-day-old bloodstain. Electropherogram of a 166-day-old bloodstain stored at ambient room conditions on wood exposed to sunlight and then analysed with the a) 18S rRNA multiplex assay, b) G6PD mRNA multiplex assay.

Figure 6.9 shows that in a five and half month old bloodstain, the degradation pattern (RFU values and loss of certain fragments) of bloodstain stored on wood is similar to that of a two month old stain (Figure 6.8) stored under the same conditions. These figures show that different substrates do affect nucleic acid degradation.

Table 6.6. Bloodstain age prediction when deposited on cotton cloth. Stains of known age (actual age) were extracted and analysed using the methods described in sections 6.2.1 to 6.2.2, the prediction equation was used to estimate the age of the bloodstain (predicted age).

Condition	Predicted age (days)	Actual age (days)
Ambient temperature, sunlight	3	1
	3	1
	5	7
	6	7
	21	14
	18	14
	22	25
	20	25
	116	90
	100	90
	165	160
	151	160
	172	180
	183	180
4°C	8	8
	31	8
37°C	6	7
	5	7
	72	150
	23	150
Ambient temperature, dark	56	180
	67	180

Table 6.7. Bloodstain age prediction when deposited on different substrates. Stains of known age (actual age) were extracted and analysed using the methods described in sections 6.2.1 to 6.2.2, the prediction equation was used to determine the age of the bloodstain (predicted age).

Substrate	Predicted age (days)	Actual age (days)
Wool carpet	145	64
	260	64
	111	90
	134	90
Plasterboard	209	64
	79	64
	103	90
	256	90
Wood	155	166
	126	166
	126	70
	185	70
Leather	330	64
	219	64
	102	90
Cement paving slab	3	2
	16	15
Sports shoe	4	2
	26	15

Table 6.8. Summary table for bloodstain age estimation.

Substrate	Temperature	Sunlight	Time	
			Predicted	Actual
cotton cloth	RT*	y	3	1
cotton cloth	RT	y	3	1
cotton cloth	RT	y	5	7
cotton cloth	RT	y	6	7
cotton cloth	RT	y	21	14
cotton cloth	RT	y	18	14
cotton cloth	RT	y	22	25
cotton cloth	RT	y	20	25
cotton cloth	RT	y	116	90
cotton cloth	RT	y	100	90
cotton cloth	RT	y	165	160
cotton cloth	RT	y	151	160
cotton cloth	RT	y	172	180
cotton cloth	RT	y	183	180
cotton cloth	RT	n	56	180
cotton cloth	RT	n	67	180
cotton cloth	37°C	n	6	7
cotton cloth	37°C	n	5	7
cotton cloth	37°C	n	72	150
cotton cloth	37°C	n	23	150
cotton cloth	4°C	n	8	8
cotton cloth	4°C	n	31	8
wood	RT	y	155	166
wood	RT	y	126	166
wood	RT	y	126	70
wood	RT	y	185	70
leather	RT	y	330	64
leather	RT	y	219	64
leather	RT	y	102	90
Plasterboard	RT	y	209	64
Plasterboard	RT	y	79	64
Plasterboard	RT	y	103	90
Plasterboard	RT	y	256	90
wool carpet	RT	y	145	64
wool carpet	RT	y	260	64
wool carpet	RT	y	111	90
wool carpet	RT	y	134	90
cement	RT	y	3	2
cement	RT	y	16	15
shoe	RT	y	4	2
shoe	RT	y	26	15

*RT – room temperature

6.3.2 Blind samples

All of the bloodstains examined thus far were of a known age. In order to test the accuracy of the prediction equation, bloodstains whose age was only known to a third party were tested. These bloodstains were deposited on the head of a cotton swab and a glass microscope slide. Figures 6.10 to 6.13 show the electropherograms of the blind samples following analysis with the G6PD mRNA and 18S rRNA multiplex assays. Table 6.9 shows the fragment data that was used for analysis (age prediction). After the age was predicted, the third party was consulted, the actual ages were obtained and the results are shown in Table 6.10.

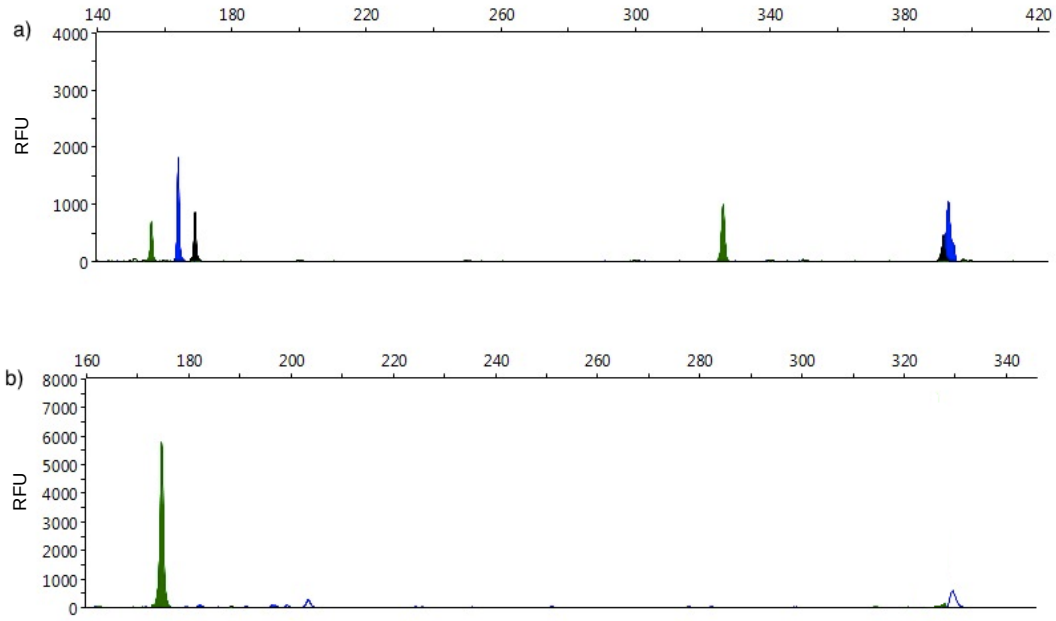


Figure 6.10. Electropherogram of blind sample "A". Electropherogram of blind blood sample "A" deposited on a cotton swab and analysed with a) 18S rRNA multiplex assay, b) G6PD mRNA multiplex assay.

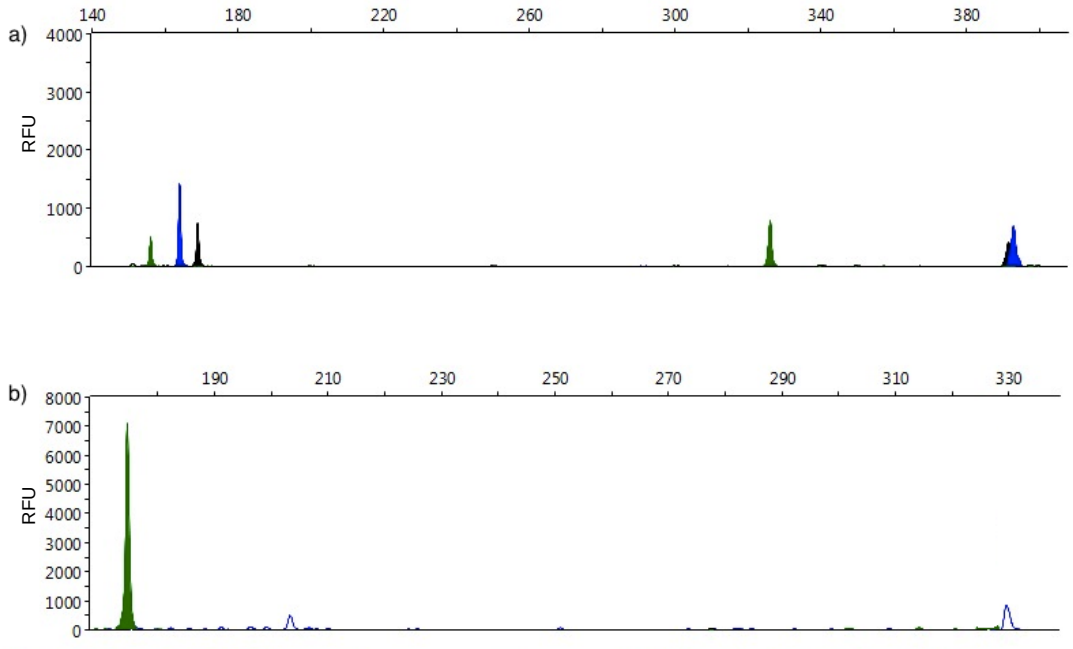


Figure 6.11. Electropherogram of blind sample "B". Electropherogram of blind blood sample "B" deposited on a cotton swab and analysed with a) 18S rRNA multiplex assay, b) G6PD mRNA multiplex assay.

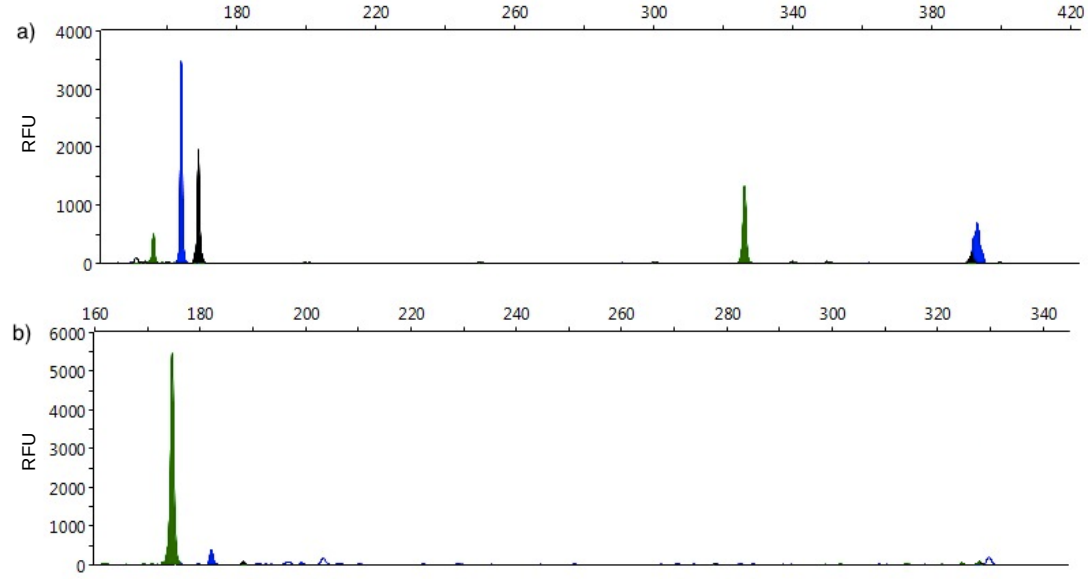


Figure 6.12. Electropherogram of blind sample "K". Electropherogram of blind blood sample "K" deposited on a cotton swab and analysed with a) 18S rRNA multiplex assay, b) G6PD mRNA multiplex assay.

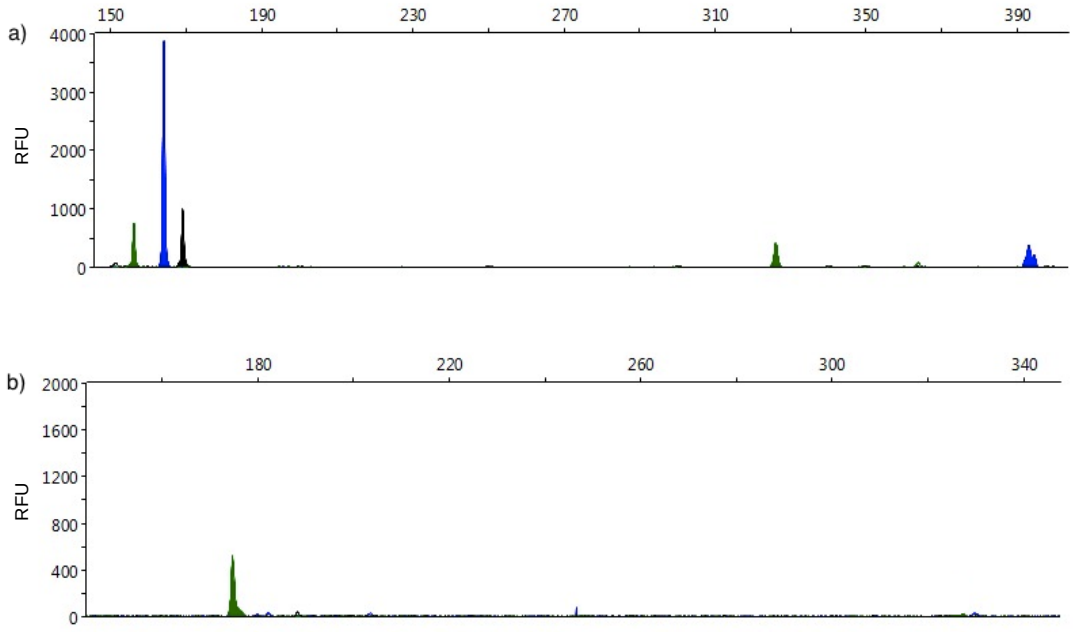


Figure 6.13. Electropherogram of blind sample "S". Electropherogram of blind blood sample "S" deposited on a glass microscope slide and analysed with a) 18S rRNA multiplex assay, b) G6PD mRNA multiplex assay.

Table 6.9 Blind sample fragment data. Raw fragment RFU data of blind samples used for age prediction. Where a RFU of 10 is specified the fragment has been degraded below the detection threshold.

	G6PD 5' medium	G6PD 3' large	G6PD 3' medium	18S central large
K	10	10	5473	693
A	10	10	5805	810
B	10	10	7116	692
Glass	10	10	532	382

Table 6.10 Results of blind study. The ages of the blind samples were predicted using the prediction equation and then compared to the actual age.

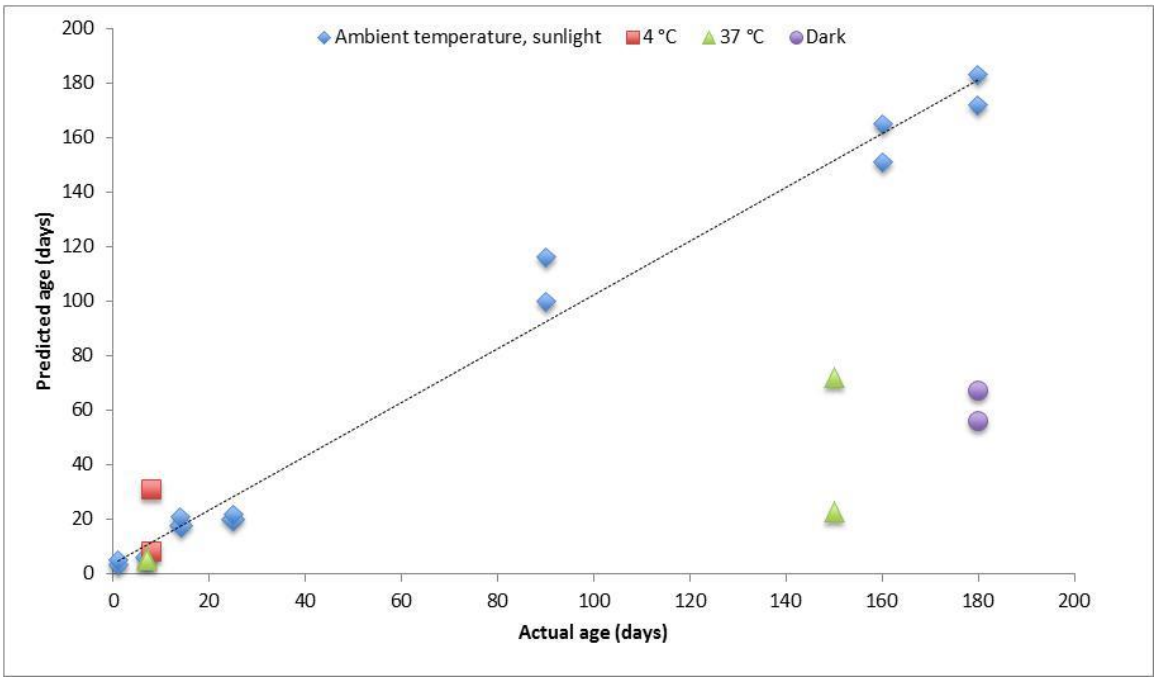
Sample	Predicted age (days)	Actual age (days)
A	355	628
B	635	628
K	367	256
S	104	211

6.4 Discussion

Table 6.6 and Figure 6.14 show the results when bloodstains are recovered and analysed from cotton cloth. The dotted line in Figures 6.14 and 6.15 represent the regression model. The data points in blue represent a new test set under the same conditions from which the regression model was generated.. Table 6.6 and Figure 6.14 show that temperature and sunlight appear to have a negative effect on bloodstain age because when the bloodstains were stored at 4°C, the average predicated age of an 8 day old stain was 20 days. In essence, the stain appeared older than it actually was, possibly due to the fact that at 4°C, the bloodstain took longer to dry (more than 60 minutes) and it is in the transition from wet to dry that most of the degradation occurs (Figure 4.5A and 4.6A). When the bloodstains were stored at 37°C, the inverse relationship was observed and the average predicated age of 150 day old stains was 48. The stains would have dried faster and after a period of 7 or more days, degradation proceeded at a much slower rate (as inferred by Figures 4.5B and 4.6B). Bloodstains that were not exposed to sunlight and stored in ambient room conditions exhibited a lower rate of degradation. The average predicated age of 180 day old stains was 62, inferring that sunlight also contributes to RNA degradation. Overall, these results could relate to the manner in which the prediction equation was developed. The samples used to derive the prediction were exposed to sunlight and ambient room temperature so any deviations in age could be a result of the storage conditions. Multivariate analysis would need

to be performed on a larger number of samples to ascertain the effect of each storage condition.

Figure 6.14. Dot plot of actual bloodstain age plotted against predicted age when bloodstains were deposited on cotton cloth. Bloodstains were deposited on cotton cloth and stored in different conditions for varying lengths of time.

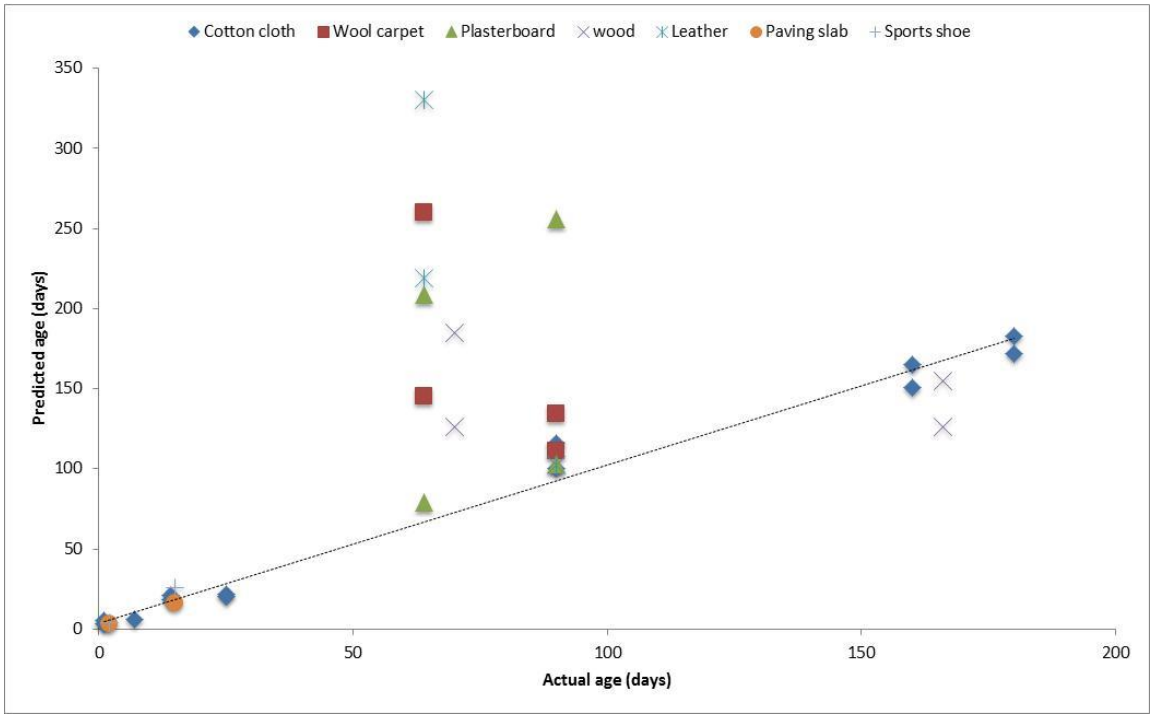


Of the different substrates tested (Table 6.7 and Figure 6.15), the wool carpet, plasterboard and leather cut offs appear to have the biggest effect on predicting bloodstain age. All predicted bloodstain ages were older and this implies that certain substrates plus a combination of factors (such as exposure to sunlight) increase RNA degradation. As these substrates were not sterilised in any manner, surface bacteria, dust, plus a host of other contaminants could have attributed to accelerated RNA degradation.

Bloodstains of an unknown age were analysed (Table 6.9). The results showed that the stains exceeded the age of the 150 day old stains that the prediction equation was developed with. Despite this, the equation was still able to predict and show differences between the various bloodstains. Most importantly, the

results show that prediction equation was able to determine that the stains were months and even years old.

Figure 6.15. Dot plot of actual stain age plotted against predicted age when bloodstains were deposited on different substrates. Bloodstains were deposited on various substrates varying lengths of time at ambient room temperature with exposure to sunlight.



6.4.1 Uncertainty of age prediction

Analysis showed that there is a degree of uncertainty associated with the prediction equation and the age of the bloodstain (as shown by Table 6.8 and 6.10). The goal of developing the regression equation was to differentiate between recent and old bloodstains (those that are days old compared to those that are months old). Table 6.10 shows when samples “A” and “B” were analysed, the predicted ages were 355 and 635 days respectively. Both samples were 628 days old. In terms of being able to distinguish between recent and old stains, the predication equation is shown to be capable of accomplishing this task. Without knowing the age of the samples and an investigator was trying to link these to a recent crime (a few days), analysis shows that samples “A” and “B”

would not be related to the investigation. A general statement could be issued whereby the lower 95% confidence intervals for the different coefficients (Table 5.18) are used in the prediction equation stating that “the bloodstain is at least (“x” amount) of days old”. When used in conjunction with the prediction equation, a degree of certainty could be established. Sample “B” for example would be at least 140 days old but is more likely to be 635 days old.

6.5 Conclusion

When the prediction equation was used to determine the age of deposited bloodstains on various substrates, stored under different conditions it was found that there is scope for answering the question “how old is this bloodstain?” often posed by forensic investigators, which has so far gone unanswered. When bloodstains were analysed from pieces of cotton cloth, such as those associated with bloodied t-shirts or garments recovered from crime scenes, the prediction equation was clearly able to show a distinguishable difference between stains deposited recently (days) to those deposited some time ago (weeks and months). This highlights the fact that relative age determination is possible.

Bloodstains were recovered and analysed from substrates common to household items such as a carpet in a living room, a leather couch, a paved driveway or garage; and those more common to a building site such as plasterboard and wood. The results show that the substrate upon which the blood is deposited on and the conditions that it is exposed to affect the rate of nucleic acid degradation. However, with the use of the prediction equation it is still possible to distinguish between recent and old bloodstains. The substrates analysed here were left “as is” to best mimic these real world scenarios of where one can expect to find blood. As demonstrated with the blind sample tests, RNA can be obtained from bloodstains which are years old, and still give somewhat accurate age predictions (e.g. a predicted age of 635 days compared to the actual age of 628 days or the

age of the predicated stains were all months old) and still distinguish between recent and old (or older as was the case).

The scenarios presented here are typical of a household environment where room temperatures of 18 – 22°C can be expected and there is adequate sunlight. The results show that prediction equation is applicable to these types of scenarios.

The results from this chapter show that by analysing amplified fragments of the 18S rRNA and G6PD mRNA sequences with the prediction equation, the method can be used to determine the age of the bloodstain within limits, making the technique described in this thesis of practical forensic use. By using the prediction equation it is possible to determine between fresh and old stains, or distinguish between day, week, month and year old bloodstains.

Chapter 7

Final Discussion

7.1 Prelude

This research was conducted in an active forensic laboratory. The methods used to address the aims of this research were designed to be incorporated into the daily running of the laboratory. If the method is to be implemented for routine use, little or no change would be required in terms of equipment or consumables, as such, the approach used is quite specific, using a narrow range of techniques which are already instigated.

Capillary electrophoresis was chosen as the preferred method over RT-PCR because RT-PCR multiplexing is limited to four targets/reporter dyes, a limitation not shared by capillary analysis. For the most part, commercial primers and probes were not used in this study due to their proprietary nature. As a main focus of the study was to investigate the degradation pattern in drying bloodstains and relate it to age prediction, specific primers were developed and used to amplify known regions of the various sized amplicons of the target sequences required for analysis. This has the added benefit because the exact location of the amplified sequences are known. This would not have been possible if commercial primers and probes were used. As DNase is routinely used in the laboratory on extracted RNA for the removal of DNA contamination prior to cDNA synthesis, this step was incorporated into the methodology of the study so that primers did not have to be designed to span introns. This allowed for amplification of the desired sized fragments required for analysis.

7.2 Chapter Summaries

Forensic evidence is the “invisible witness” when crimes are committed. Trace evidence often plays a crucial role in linking a suspect or victim to a crime.

Through DNA evidence, it is possible to determine the gender of the victim or offender and can also indicate the strength of the DNA evidence in favour of a probable individual. Forensic scene investigators will often secure pieces of evidence for body fluid identification to be used in cases of sexual assault, assault and homicide. However, it is not always clear if the evidence relates to the particular incident under investigation or another unrelated event. By determining the time since deposition of the sample, it would be possible to clarify the circumstances of the investigation.

Bloodstains are the most visible forms of evidence at a crime scene and presently there is no method available that can accurately determine when they were deposited. Previous studies have shown that it is possible to confirm changing chemical and physical properties in bloodstains (Chapter 1) but none have shown the precision required for forensic applications.

The aim of this study was to determine whether it is possible to use the degradation profile of nucleic acids to provide a better estimate in bloodstain age estimation. The main hypothesis was that as a bloodstain ages, whole nucleic acid transcripts are degraded and become fragmented as they are no longer regulated or repaired by normal repair pathways. By examining the degradation of HBB DNA, G6PD mRNA and 18S rRNA in aged bloodstains through PCR, not all fragments could be used to determine bloodstain age.

Chapter three discussed the development of a real-time qPCR assay that was used to standardise the concentration of cDNA added to the 18S rRNA and G6PD mRNA multiplex assays. This was significant because no other study to date that has investigated bloodstain age has done this.

When different bloodstains or samples are compared, it is important that the method used to analyse them accounts for variation in template amounts as a result of the efficiency of the extraction procedure and cDNA synthesis. By

standardising the template amount, the procedure limits the amount of variation between samples and assays, helping to ensure that any differences arise from age and not the experimental technique. No method is 100% accurate but standardising the template limits some of the variability. This method can be used in applications where the exact determination of RNA/cDNA amount may be necessary and the blood origin of the stain is already established (e.g. bloodstain age determination).

In its development, there were several factors that needed to be taken into consideration. Using DNA (cDNA) as a standard for quantification is possible but it does have a downside as there is no control for the efficiency of the reverse transcription (RT) step and must therefore be assumed to be 100%. The efficiency of an RT reaction might vary from sample to sample as well as from kit to kit. One way to negate this problem was to aliquot the cDNA used in the standard curve creation, which meant that repeated freeze/thawing cycles were avoided and each batch of aliquots allowed for the creation of ten standard curves. Table 3.2 also demonstrated that the cDNA standards met the required criteria for efficiency and precision over a number of qPCR runs (Section 3.4.1). Results demonstrated that the developed assay is more sensitive towards human blood samples (Table 3.3) which means that the method cannot be applied when the origin of the sample is not known, which for the purpose of this thesis, was not an issue as all samples were known to be of human origin. Further developments to this method would involve the synthesis of human specific primers and probes. This would provide greater applicability to the field of forensics. Another possible improvement would involve developing a G6PD mRNA quantification assay to determine the concentration of G6PD cDNA to use in the aging bloodstain model.

Chapter four discussed the development of three multiplex assays for the detection and amplification of HBB DNA, 18S rRNA and G6PD mRNA fragments. By observing changes in peak height of various amplified fragments, the pattern

of degradation for DNA, rRNA and mRNA was investigated. This was also important because the methods involving RNA and bloodstain age have not taken the directionality of degradation into account when designing primers and probes. The assays showed that degradation occurs in all three nucleic acids (Figures 4.4 to 4.6) and the process is random. The results also indicated that the nucleic acids in bloodstains are somewhat protected against the damaging effects of the environment (such as sun exposure and elevated temperatures) when compared to “naked” nucleic acids (Section 4.3.2 and Figures 4.10 to 4.12) and this made further analysis of nucleic acids feasible with the development of a regression model.

Chapter 5 discussed the development of a statistical model that could be used to explain the relationship between fragment peak height and time. Regression analysis was performed on aged bloodstains to determine which fragments were statistically relevant to age prediction. Linear regression analysis revealed that a combination of the 5’ medium fragment, the 3’ large and 3’ medium fragments of G6PD mRNA and the central large fragment of 18S rRNA were suitable markers for estimating the age of a bloodstain (Table 5.18). The results of the HBB DNA assay showed that DNA degradation did not correlate well with age and was therefore excluded from the final model. The result was an age prediction equation that was evaluated using aged blind samples and on bloodstains found on forensic type substrates stored under different conditions.

Chapter six found that the prediction equation, which used a combination of mRNA and rRNA fragments, was good at predicting the age of bloodstains. The model was good at predicting differences between days, weeks, months and even years old. Environmental conditions and substrate play were also found to play a role in nucleic acid degradation (Figures 6.14 and 6.15), providing opportunities for further study.

In summary, the results presented in this thesis indicate that the different types of nucleic acids and different sized regions of nucleic acids decay at different rates

in *ex vivo* dried bloodstains. Different primer sets were used to examine different sized regions between HBB DNA, 18S rRNA and G6PD mRNA isolated from bloodstains that had been aged up to 150 days. The results indicate that degradation to nucleic acids occurs most rapidly as the bloodstain transitions from wet to dry and that certain regions of a transcript are more susceptible to degradation than others. Glucose-6-phosphate dehydrogenase mRNA degraded more rapidly than 18S rRNA and HBB DNA and the degradation of certain fragments of this transcript were shown to have a strong correlation to time (Table 5.18).

7.3 Key findings

7.3.1 Effect of the environment on bloodstain age

Preliminary results indicated that the nucleic acids isolated from dried bloodstains exposed to high temperatures and UV are less stable then the RNA isolated from dried bloodstains exposed to lower temperatures and sunlight (Figures 4.4 to 4.6). Full spectrum light may have an effect on nucleic acid degradation as bloodstains stored in darkness were “younger” then their sunlight counterparts (Figure 6.13).

When blood remains hydrated, nucleic acid degradation occurs more rapidly presumably due to the continued presence of enzymatic activity. Once the bloodstain dries, the cells become desiccated and the nucleic acids present within them appear to be more stable. This observation was evident when bloodstains were stored at different temperatures (Figures 4.4B, 4.5B, 4.6B and 6.13).

7.3.2 Effect of substrate on bloodstain age

The different substrates on which blood was deposited did appear to influence stain age (Figures 6.14 and 6.15). Bloodstains were perceived to be generally older because degradation on some substrates was greater than others. Leather for example, is treated with a variety of chemicals and the darker colour (as used in this study) would have absorbed more light and would have therefore been a

better radiator of heat, enhancing degradation. As the surfaces were not sterile, other factors such as bacteria may play a role.

7.4 Contribution of this study to the field of forensic investigation

By applying the prediction formula, the age estimate of a bloodstain deposited under different environmental conditions and substrates may provide a temporal link between the deposition of the blood and the time a crime was committed. Conversely, these results may be helpful in excluding material that does not correspond to the time when a crime was committed. For example, if a bloodstain is determined to be over three months old and the crime was committed two weeks ago, that particular bloodstain could be excluded from the crime scene evidence. In the absence of other knowledge, this technique may also be of use to approximate when the crime was committed based on the criteria mentioned above and demonstrated by the blind study (Table 6.10).

Prior to this research, one day old bloodstains could not be distinguished from six month old blood. This research provides a foundation to elucidate a more accurate means of dating a bloodstain.

7.4 Future direction of the research and improvements

The results that have been presented in this thesis support the hypothesis that nucleic acid degradation can be used to determine the age of a bloodstain. However, as there are currently no standardised guidelines in place for bloodstain age determination experiments these had to be established during this study and hence there are certain limitations of the study design. The model was developed with bloodstains deposited on cotton cloth stored at ambient temperatures with exposure to sunlight. Such conditions were chosen as they best mimic the interior of a household where most crimes are committed and articles of clothing are often sources of evidence. Cotton cloth was chosen because the majority of studies described in Chapter 1 had chosen cotton before. As the substrate studies indicated, results do vary and different substrates with

longer time points should be examined. The time points of 4, 20 and 150 days were chosen to determine if there was any detectable difference in nucleic acid fragment degradation for the method to be of any use in the forensic context. As there were differences, the model could be re-analysed with additional time points and increased sample sizes to improve the resolution of the prediction equation and reduce the standard deviation between time points or samples (therefore providing a better model) at shorter and longer intervals (for example 1 day, 3 days, 7 days, 12 days, 15 days, 20 days, 30 days, 40 days, 60 days, 90 days, 130 days etc.) The effects of sunlight, temperature and humidity on bloodstain age would also need to be investigated separately. An internal standard could also be incorporated into each assay to reduce the variability during analysis.

There is also potential for this method to be applied to tissues other than blood, such as hair, semen and salvia, with tissue-specific mRNA profiles to determine the tissue from which the sample is derived. This would involve designing different multiplex assays, applying the aspect of standardising each sample and then analysing the results with the regression analysis principle described here to determine which fragments are relevant to age prediction in that particular body fluid or hair sample.

Appendices

Appendix 1. HBB DNA nucleotide sequence showing primer alignment.

>gi|224589802:c5248301-5246696 Homo sapiens chromosome 11, GRCh37.p9 Primary Assembly

ACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCATCTGACTCCTGA
GGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGC
AGGTTGGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGAAG
ACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATTGGTCTATTTTCCCACCCTTAGGCTGCTGG
TGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGG
CAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGAC
AACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACT
TCAGGGTGAGTCTATGGGACGCTTGATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCATGTCATAG
GAAGGGGATAAGTAACAGGGTACAGTTTAGAATGGGAAACAGACGAATGATTGCATCAGTGTGGAAGTCT
CAGGATCGTTTTAGTTTCTTTTATTTGCTGTTCATAACAATTGTTTTCTTTTGTTTAATTCTTGCTTTCT
TTTTTTTTCTTCTCCGCAATTTTTACTATTATACTTAATGCCTTAACATTGTGTATAACAAAAGGAAATA
TCTCTGAGATACATTAAGTAACTTAAAAAAAAACTTTACACAGTCTGCCTAGTACATTACTATTTGGAAT
ATATGTGTGCTTATTTGCATATTCATAATCTCCCTACTTTATTTTCTTTTATTTTTAATTGATACATAAT
CATTATACATATTTATGGGTTAAAGTGTAATGTTTTAATATGTGTACACATATTGACCAAATCAGGGTAA
TTTTGCATTTGTAATTTTAAAAAATGCTTTCTTCTTTTAATATACTTTTTTGTTTATCTTATTTCTAATA
CTTTCCCTAATCTCTTTCTTTCAGGGCAATAATGATACAATGTATCATGCCTCTTTGCACCATTCTAAAG
AATAACAGTGATAATTTCTGGGTTAAGGCAATAGCAATATCTCTGCATATAAATATTTCTGCATATAAAT
TGTAACTGATGTAAGAGGTTTCATATTGCTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTT
ATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTCATACCTCTT
ATCTTCCTCCCACAGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCA
CCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCACAAGTATCA
CTAAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACT
GGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGC

Appendix 2. 18S rRNA nucleotide sequence showing primer alignment.

>gi|225637497|ref|NR_003286.2| Homo sapiens RNA, 18S ribosomal 5 (RNA18S5), ribosomal RNA

TACCTGGTTGATCCTGCCAGTAGCATATGCTTGTCTCAAAGATTAAGCCATGCATGTCTGAGTACGCACG
GCCGGTACAGTGAAACTGCGAATGGCTCATTAAATCAGTTATGGTTCCTTTGGTCGCTCGCTCCTCTCCT
ACTTGGATAACTGTGGTAATTCTAGAGCTAATACATGCCGACGGGCGCTGACCCCCTTCGCGGGGGGGAT
GCGTGCATTTATCAGATCAAAACCAACCCGGTCAGCCCCTCTCCGGCCCCGGCCGGGGGGCGGGCGCCGG
CGGCTTTGGTGACTCTAGATAACCTCGGGCCGATCGCACGCCCCCCGTGGCGGCGACGACCCATTCGAAC
GTCTGCCCTATCAACTTTCGATGGTAGTCGCCGTGCCTACCATGGTGACCACGGGTGACGGGGAATCAGG
GTTCGATTCCGGAGAGGGAGCCTGAGAAACGGCTACCACATCCAAGGAAGGCAGCAGGCGCGCAAATTAC
CCACTCCCGACCCGGGGAGGTAGTGACGAAAAATAACAATACAGGACTCTTTCGAGGCCCTGTAATTGGA
ATGAGTCCACTTTAAATCCTTTAACGAGGATCCATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAAT
TCCAGCTCCAATAGCGTATATTAAAGTTGCTGCAGTTAAAAAGCTCGTAGTTGGATCTTGGGAGCGGGCG
GGCGGTCCGCCGCGAGGCGAGCCACCGCCCGTCCCCGCCCCTTGCCTCTCGGCGCCCCCTCGATGCTCTT
AGCTGAGTGTCCCGCGGGGCCCGAAGCGTTTACTTTGAAAAAATTAGAGTGTTCAAAGCAGGCCCGAGCC
GCCTGGATACCGCAGCTAGGAATAATGGAATAGGACCGCGGTTCTATTTTGTTGGTTTTCGGAACTGAGG
CCATGATTAAGAGGGACGGCCGGGGGCATTCGTATTGCGCCGCTAGAGGTGAAATTCTTGGACCGGCGCA
AGACGGACCAGAGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCGGAGGTTCGAA
GACGATCAGATACCGTCGTAGTTCCGACCATAAACGATGCCGACCGGCGATGCGGCGGCGTTATTCCCAT
GACCCGCCGGGCAGCTTCCGGGAAACCAAAGTCTTTGGGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAA
CTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAA
CCTCACCCGGCCCGGACACGGACAGGATTGACAGATTGATAGCTCTTTCTCGATTCCGTGGGTGGTGGTG
CATGGCCGTTCTTAGTTGGTGGAGCGATTTGTCTGGTTAATTCCGATAACGAACGAGACTCTGGCATGCT
AACTAGTTACGCGACCCCCGAGCGGTCGGCGTCCCCCAACTTCTTAGAGGGACAAGTGGCGTTCAGCCAC
CCGAGATTGAGCAATAACAGGTCTGTGATGCCCTTAGATGTCCGGGGCTGCACGCGCGCTACACTGACTG
GCTCAGCGTGTGCCTACCCTACGCCGGCAGGCGCGGGTAACCCGTTGAACCCCATTCGTGATGGGGATCG
GGGATTGCAATTATTCCCCATGAACGAGGAATTCCCAGTAAGTGCGGGTCATAAGCTTGCGTTGATTAAG
TCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTGGATGGTTTAGTGAGGCCCTCGGATCGGC
CCCGCCGGGGTCGGCCCACGGCCCTGGCGGAGCGCTGAGAAGACGGTCGAACTTGACTATCTAGAGGAAG
TAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTA

Appendix 3. G6PD mRNA nucleotide sequence showing primer alignment.

>gi|108773794|ref|NM_000402.3| Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), transcript variant 1, mRNA

AGAGGCAGGGGCTGGCCTGGGATGCGCGCGCACCTGCCCTCGCCCCGCCCCGCCCGCACGAGGGGTGGTG
GCCGAGGCCCCGCCCCGCACGCCTCGCCTGAGGCGGGTCCGCTCAGCCCAGGCGCCCGCCCCCGCCCCCG
CCGATTAAATGGGCCGGCGGGGCTCAGCCCCCGGAAACGGTCGTACACTTCGGGGCTGCGAGCGCGGAGG
GCGACGACGACGAAGCGCAGACAGCGTCATGGCAGAGCAGGTGGCCCTGAGCCGGACCCAGGTGTGCGGG
ATCCTGCGGGAAGAGCTTTTCCAGGGCGATGCCTTCCATCAGTCGGATACACACATATTCATCATCATGG
GTGCATCGGGTGACCTGGCCAAGAAGAAGATCTACCCCACCATCTGGTGGCTGTTCCGGGATGGCCTTCT
GCCCGAAAACACCTTCATCGTGGGCTATGCCCGTTCCCGCCTCACAGTGGCTGACATCCGCAAACAGAGT
GAGCCCTTCTTCAAGGCCACCCCAGAGGAGAAGCTCAAGCTGGAGGACTTCTTTGCCCGCAACTCCTATG
TGGCTGGCCAGTACGATGATGCAGCCTCCTACCAGCGCCTCAACAGCCACATGAATGCCCTCCACCTGGG
GTCACAGGCCAACCGCCTCTTCTACCTGGCCTTGCCCCCGACCGTCTACGAGGCCGTCACCAAGAACATT
CACGAGTCCTGCATGAGCCAGATAGGCTGGAACCGCATCATCGTGGAGAAGCCCTTCGGGAGGGACCTGC
AGAGCTCTGACCGGCTGTCCAACCACATCTCCTCCCTGTTCCGTGAGGACCAGATCTACCGCATCGACCA
CTACCTGGGCAAGGAGATGGTGCAGAACCTCATGGTGCTGAGATTTGCCAACAGGATCTTCGGCCCCATC
TGGAACCGGGACAACATCGCCTGCGTTATCCTCACCTTCAAGGAGCCCTTTGGCACTGAGGGTCGCGGGG
GCTATTTCGATGAATTTGGGATCATCCGGGACGTGATGCAGAACCACCTACTGCAGATGCTGTGTCTGGT
GGCCATGGAGAAGCCCGCCTCCACCAACTCAGATGACGTCCGTGATGAGAAGGTCAAGGTGTTGAAATGC
ATCTCAGAGGTGCAGGCCAACAATGTGGTCCTGGGCCAGTACGTGGGGAACCCCGATGGAGAGGGCGAGG
CCACCAAAGGGTACCTGGACGACCCCACGGTGCCCCGCGGGTCCACCACCGCCACTTTTGCAGCCGTCGT
CCTCTATGTGGAGAATGAGAGGTGGGATGGGGTGCCCTTCATCCTGCGCTGCGGCAAGGCCCTGAACGAG
CGCAAGGCCGAGGTGAGGCTGCAGTTCCATGATGTGGCCGGCGACATCTTCCACCAGCAGTGCAAGCGCA
ACGAGCTGGTGATCCGCGTGCAGCCCAACGAGGCCGTGTACACCAAGATGATGACCAAGAAGCCGGGCAT
GTTCTTCAACCCCGAGGAGTCGGAGCTGGACCTGACCTACGGCAACAGATACAAGAACGTGAAGCTCCCT
GACGCCTACGAGCGCCTCATCCTGGACGTCTTCTGCGGGAGCCAGATGCACTTCGTGCGCAGCGACGAGC
TCCGTGAGGCCTGGCGTATTTTCACCCCACTGCTGCACCAGATTGAGCTGGAGAAGCCCAAGCCCATCCC
CTATATTTATGGCAGCCGAGGCCCCACGGAGGCAGACGAGCTGATGAAGAGAGTGGGTTTCCAGTATGAG
GGCACCTACAAGTGGGTGAACCCCCACAAGCTCTGAGCCCTGGGCACCCACCTCCACCCCCGCCACGGCC
ACCCTCCTTCCCGCCGCCCGACCCCGAGTCGGGAGGACTCCGGGACCATTGACCTCAGCTGCACATTCCT
GGCCCCGGGCTCTGGCCACCCTGGCCCGCCCCTCGCTGCTGCTACTACCCGAGCCCAGCTACATTCCTCA
GCTGCCAAGCACTCGAGACCATCCTGGCCCCTCCAGACCCTGCCTGAGCCCAGGAGCTGAGTCACCTCCT
CCACTCACTCCAGCCCAACAGAAGGAAGGAGGAGGGCGCCCATTCGTCTGTCCCAGAGCTTATTGGCCAC
TGGGTCTCACTCCTGAGTGGGGCCAGGGTGGGAGGGAGGGACAAGGGGGAGGAAAGGGGCGAGCACCCAC
GTGAGAGAATCTGCCTGTGGCCTTGCCCGCCAGCCTCAGTGCCACTTGACATTCCTTGTCACCAGCAACA
TCTCGAGCCCCCTGGATGTCCCCTGTCCCACCAACTCTGCACTCCATGGCCACCCCGTGCCACCCGTAGG
CAGCCTCTCTGCTATAAGAAAAGCAGACGCAGCAGCTGGGACCCCTCCCAACCTCAATGCCCTGCCATTA
AATCCGCAAACAGCC

Appendix 4. Raw nucleic acid data

The data presented in Tables 1-3 is the raw peak height data (RFU) used for statistical analysis and regression modelling.

Table 1. HBB DNA fragment RFU data.

Time (days)	ln(time)	Central Medium	Central Large	5' medium	3' medium	3' large	5' large
4	1.386	926	1076	198	493	867	502
4	1.386	660	771	133	377	605	358
4	1.386	1162	1310	228	660	1225	385
4	1.386	799	1253	238	691	1008	133
4	1.386	926	1076	198	493	867	278
4	1.386	1061	1160	221	769	1119	348
4	1.386	708	823	126	335	618	217
4	1.386	707	813	141	368	660	228
4	1.386	1027	1249	219	608	999	483
4	1.386	1155	2030	345	1012	1471	459
4	1.386	1043	1715	264	831	1191	832
4	1.386	1381	1674	268	639	1285	681
4	1.386	1340	1809	309	749	1106	604
4	1.386	1278	1593	260	582	1180	617
4	1.386	1155	2030	345	1012	1471	1087
4	1.386	2054	2476	425	1074	2069	699
4	1.386	1043	1715	264	831	1191	408
4	1.386	1381	1674	268	639	1285	436
4	1.386	1278	1593	260	582	1180	384
4	1.386	1363	1551	233	580	1232	414
20	2.996	1173	1638	281	574	786	339

20	2.996	1572	1798	321	634	1196	529
20	2.996	1240	1600	281	569	881	390
20	2.996	1820	2097	364	852	1617	500
20	2.996	1173	1638	281	574	786	310
20	2.996	1718	1959	352	750	1486	456
20	2.996	1240	1600	281	569	881	330
20	2.996	2203	2113	462	1044	1606	541
20	2.996	1322	1613	328	690	875	289
20	2.996	2090	2003	531	1176	1669	427
20	2.996	2938	7544	1367	4961	7373	163
20	2.996	2412	7390	1334	5247	7438	325
20	2.996	3360	7427	1691	6674	7391	347
20	3.401	2203	2113	462	1044	1606	541
20	3.401	1322	1613	328	690	875	289
20	3.401	2090	2003	531	1176	1669	427
150	5.011	2601	3411	844	1788	2703	1875
150	5.011	2254	2150	523	1014	1763	1367
150	5.011	2395	2634	458	998	1909	1416
150	5.011	3000	3622	696	1503	2532	1874
150	5.011	2406	2490	440	1152	1891	1376
150	5.011	2886	3161	728	1361	2255	1695
150	5.011	2838	3728	674	1401	2149	1598
150	5.011	2616	3085	551	995	1598	1200
150	5.011	2077	2200	412	971	1433	846
150	5.011	2331	2414	321	946	1358	962
150	5.011	2674	3018	456	1031	1699	1125
150	5.011	2593	3363	458	1334	2064	1624

150	5.011	2712	2382	420	1079	1463	797
150	5.011	3492	3301	766	1576	2061	1565
150	5.011	2916	3237	584	1224	1762	1344
150	5.011	1406	1250	296	588	982	320
150	5.011	1501	1430	328	608	1021	443
150	5.011	826	1413	263	569	787	495
150	5.011	1077	1134	287	591	847	486
150	5.011	1069	1045	290	665	995	654
150	5.011	1060	1278	271	439	560	214
150	5.011	1501	1430	328	608	1021	328
150	5.011	826	1413	263	569	787	230
150	5.011	2146	2440	461	1132	2056	494
150	5.024	1921	4120	595	2046	3025	80
150	5.024	2022	6110	772	2607	4477	131
150	5.024	1495	3104	398	1347	2133	48
150	5.024	1869	2766	523	1697	2586	64
150	5.081	2956	7562	1322	4896	7363	337
150	5.081	2548	7299	1192	3886	6368	260
150	5.081	2820	5913	1025	3009	5434	183
150	5.081	2243	5848	1003	2742	4847	180

Table 2. 18S rRNA fragment RFU data.

Time (days)	ln(time)	Central Medium	5' large	3' medium	3' large	5' medium	central large
4	1.386	2841	478	557	550	557	527
4	1.386	2835	500	612	568	592	517
4	1.386	3805	681	746	674	672	755
4	1.386	2861	391	509	440	467	498
4	1.386	2890	442	597	504	522	509
4	1.386	2777	495	584	497	532	522
4	1.386	3002	498	601	539	557	555
4	1.386	2606	399	521	459	487	456
4	1.386	3397	596	681	619	627	653
4	1.386	3002	498	601	539	557	555
4	1.386	2957	616	736	604	711	672
4	1.386	2509	430	683	524	613	523
4	1.386	2193	1271	616	639	608	928
4	1.386	1901	900	588	570	530	759
4	1.386	2460	1117	712	690	576	692
4	1.386	2306	1361	568	609	553	694
4	1.386	2243	1175	587	528	524	801
4	1.386	1826	879	505	440	413	595
4	1.386	3108	686	657	575	668	512
4	1.386	3035	565	690	522	620	578
20	2.996	3568	492	791	478	496	284
20	2.996	3167	445	663	402	424	259
20	2.996	2620	958	663	485	400	394
20	2.996	2316	889	571	412	350	345
20	2.996	5884	384	7575	2643	204	23

20	2.996	5832	7313	7464	996	7621	3657
20	2.996	3197	608	815	430	662	186
20	2.996	3230	647	823	440	656	196
20	2.996	3483	471	770	367	589	214
20	2.996	3613	530	830	404	653	249
20	2.996	2552	1019	660	392	485	314
20	2.996	2614	1090	709	441	555	345
20	2.996	3483	471	770	367	589	214
20	2.996	3613	530	830	404	653	249
20	2.996	3042	781	759	415	596	248
20	2.996	3577	1100	825	448	655	298
150	5.011	617	894	758	964	2449	156
150	5.011	870	1031	869	1114	2876	227
150	5.011	1038	1413	1168	1523	4284	232
150	5.011	712	827	702	967	2912	148
150	5.011	911	1007	787	1232	3587	182
150	5.011	3336	226	714	276	603	39
150	5.011	4330	234	995	365	902	50
150	5.011	2418	462	629	295	494	48
150	5.011	3147	487	838	397	735	69
150	5.011	1396	633	698	278	511	280
150	5.011	2689	650	1051	395	507	210
150	5.011	2890	450	918	242	650	170
150	5.011	2655	468	936	245	654	185
150	5.011	3309	486	931	291	677	209
150	5.011	2141	965	798	267	553	235
150	5.011	1939	916	815	269	524	255

150	5.011	2794	908	915	323	589	286
150	5.011	2368	1037	805	309	553	315
150	5.011	2686	818	898	289	569	286
150	5.011	2184	1123	830	319	663	322
150	5.011	2609	301	710	237	660	175
150	5.011	2738	515	725	279	694	177
150	5.011	1866	763	595	263	574	283
150	5.011	2365	979	841	381	736	360
150	5.011	2975	425	833	281	604	134
150	5.011	2615	807	797	432	856	294
150	5.011	1358	703	546	341	568	247
150	5.011	2413	844	848	497	710	227
150	5.011	2884	922	782	440	756	241
150	5.011	3830	193	867	264	815	101
150	5.011	2707	644	801	346	746	228
150	5.011	3342	927	934	425	871	298

Table 3. G6PD mRNA fragment RFU data.

Time (days)	ln(time)	Central Medium	Central Large	3' medium	5' large	5' medium	3' large
4	1.386	2841	478	557	550	557	527
4	1.386	3805	681	746	674	672	755
4	1.386	2890	442	597	504	522	509
4	1.386	2861	391	509	440	467	498
4	1.386	2777	495	584	497	532	522
4	1.386	3035	497	599	533	550	562
4	1.386	3467	607	688	621	626	671
4	1.386	2602	387	510	445	474	454
4	1.386	3035	497	599	533	550	562
4	1.386	2660	402	522	457	484	468
4	1.386	2509	430	683	524	613	523
4	1.386	2507	420	614	437	528	432
4	1.386	3035	565	690	522	620	578
4	1.386	2684	472	662	494	587	511
4	1.386	2988	553	704	544	638	585
4	1.386	2379	391	620	445	536	437
4	1.386	2048	357	518	385	456	387
4	1.386	3248	578	787	591	702	623
4	1.386	2742	486	655	484	578	507
4	1.386	3011	559	694	528	625	580
20	2.996	3568	492	791	478	496	284
20	2.996	3167	445	663	402	424	259
20	2.996	3560	490	790	478	490	280
20	2.996	3169	449	669	402	420	258
20	2.996	3197	608	815	430	662	186

20	2.996	3230	647	823	440	656	196
20	2.996	3483	471	770	367	589	214
20	2.996	3613	530	830	404	653	249
20	2.996	3030	569	796	408	622	169
20	2.996	2997	530	788	398	628	159
20	2.996	3813	608	857	436	687	276
20	2.996	3483	471	770	367	589	214
20	2.996	3413	452	803	372	619	222
20	2.996	3397	635	842	462	696	213
20	2.996	3232	649	825	442	658	198
20	2.996	3192	603	811	433	661	186
150	5.011	40	10	352	20	10	27
150	5.011	23	10	186	10	20	10
150	5.011	40	10	200	10	96	10
150	5.011	65	10	221	10	99	10
150	5.011	10	10	1167	1390	10	10
150	5.011	308	10	1057	1415	10	10
150	5.011	132	10	1010	1297	10	10
150	5.011	333	10	1244	1742	90	10
150	5.011	348	10	928	1087	10	10
150	5.011	280	10	1060	1385	30	1685
150	5.011	10	10	1685	3082	10	10
150	5.011	10	10	1487	1759	10	10
150	5.011	10	10	1148	1948	10	10
150	5.011	278	10	993	1482	10	10
150	5.011	113	10	906	1435	10	10
150	5.011	300	10	644	1202	10	10

150	5.011	119	10	1328	1075	10	10
150	5.011	10	10	1000	793	10	10
150	5.011	10	10	1380	1092	10	10
150	5.011	10	10	1477	1155	10	10
150	5.011	10	10	1207	1537	10	10
150	5.011	10	10	1278	1663	10	10
150	5.011	37	10	1296	1029	22	10
150	5.011	71	10	1455	1897	28	10
150	5.011	149	10	1362	1433	17	10
150	5.011	119	10	1328	1075	10	10
150	5.011	98	10	1151	1191	10	10
150	5.011	10	10	1000	793	10	10
150	5.011	10	10	1380	1092	10	10
150	5.011	10	10	1012	1162	10	10
150	5.011	10	10	1477	1155	10	10
150	5.011	10	10	1114	1250	10	10

References

Ackerman, K., Ballantyne, K., & Kayser, M. (2010). Estimating trace deposition time with circadian biomarkers: a prospective and versatile tool for crime scene reconstruction. *International Journal of Legal Medicine, 124*, 387-395.

Adler, C., Haak, W., Donlon, D., & Cooper, A. (2011). Survival and recovery of DNA from ancient teeth and bones. *Journal of Archaeological Science, 38*(5), 956-964.

Al-Alousi, L. (2002). A study of the shape of the post-mortem cooling curve in 117 forensic cases. *Forensic Science International, 125*, 237-244.

Alaeddini, R., Ahmadi, M., Walsh, S., & Abbas, A. (2010). Semi-quantitative PCR Analysis of DNA Degradation. *Australian Journal of Forensic Sciences, 43*(1), 53-64.

Alkass, K., Buchholz, B., Ohtani, S., Yamamoto, T., Druid, H., & Spalding, K. (2010). Age estimation in forensic sciences: application of combined aspartic acid racemization and radiocarbon analysis. *Molecular and Cellular Proteomics, 9*, 1022-1030.

Alvarez, M., Juusola, J., & Ballantyne, J. (2004). An mRNA and DNA co-isolation method for forensic casework samples. *Analytical Biochemistry, 335*(2), 289-298.

Anderson, S., Hobbs, G., & Bishop, C. (2005). A method for determining the age of a bloodstain. *Forensic Science International, 148*(1), 37-45.

Anderson, S., Hobbs, G., & Bishop, C. (2011). Multivariate Analysis for Estimating the Age of a Bloodstain. *Journal of Forensic Sciences, 56*(1), 186-193.

Andrasko, J. (1997). The Estimation of the Age of Bloodstains by HPLC Analysis. *Journal of Forensic Science 42*, 601-607.

Aranda, R., Dineen, S., Craig, R., Guerrieri, R., & Robertson, J. (2009). Comparison and evaluation of RNA quantification methods using viral, prokaryotic, and eukaryotic RNA over a 104 concentration range. *Analytical Biochemistry, 387*(1), 122-127.

Arany, S., & Ohtani, S. (2011). Age estimation of bloodstains: A preliminary report based on aspartic acid racemization rate. *Forensic Science International, 212*(1), 36-39.

Aronson, J. (2007). *Genetic witness*. New Brunswick: Rutgers University Press.

Ayala-Torres, S., Chen, Y., Svoboda, T., Rosenblatt, J., & Van Houten, B. (2000). Analysis of Gene-Specific DNA Damage and Repair Using Quantitative Polymerase Chain Reaction. *Methods, 22*(2), 135-147.

Barber, R., Harmer, D., Coleman, R., & Clark, B. (2005). GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiological Genomics, 21*(3), 389-395.

Basic Local Alignment search tool. Retrieved April, 2012, from http://blast.ncbi.nlm.nih.gov/

Bauer, M. (2007). RNA in forensic science. *Forensic Science International: Genetics, 1*(1), 69-74.

Bauer, M., Gramlich, I., Polzin, S., & Patzelt, D. (2003). Quantification of mRNA degradation as possible indicator of postmortem interval, a pilot study. *Legal Medicine, 5*(4), 220-227.

Bauer, M., Kraus, A., & Patzelt, D. (1999). Detection of epithelial cells in dried blood stains by reverse transcriptase-polymerase chain reaction. *Journal of Forensic Sciences, 44*(6), 1232-1236.

Bauer, M., & Patzelt, D. (2008). Identification of menstrual blood by real time RT-PCR: Technical improvements and the practical value of negative test results. *Forensic Science International, 174*(1), 54-58.

Bauer, M., Polzin, S., & Patzelt, D. (2003). Quantification of RNA degradation by semi-quantitative duplex and competitive RT-PCR: a possible indicator of the age of bloodstains? *Forensic Science International, 138*(1-3), 94-103.

Beelman, C., & Parker, R. (1995). Degradation of mRNA in eukaryotes. *Cell, 81*, 179-183.

Bibillo, A., Figlerowicz, M., Ziomek, K., & Kierzek, R. (2000). The Nonenzymatic Hydrolysis of Oligoribonucleotides VII. Structural Elements Affecting Hydrolysis. *Nucleosides, Nucleotides and Nucleic Acids, 19*(5-6), 977-994.

Blackburn, M. (2006). *Nucleic acids in chemistry and biology* (3rd ed.). Cambridge: Royal Society of Chemistry.

Bonnet, J., Colotte, M., Coudy, D., Couallier, V., Portier, J., Morin, B., & Tuffet, S. (2009). Chain and conformation stability of solid-state DNA: implications for room temperature storage. *Nucleic Acids Research*.

Bowden, A., Fleming, R., & Harbison, S. (2011). A method for DNA and RNA co-extraction for use on forensic samples using the Promega DNA IQ™ system. *Forensic Science International: Genetics, 5*(1), 64-68.

Bowman, S. (1990). *Radiocarbon Dating*. Berkeley University of California Press.

Bremmer, R., de Bruin, K., van Gemert, M., van Leeuwen, T., & Aalders, M. (2012). Forensic quest for age determination of bloodstains. *Forensic Science International, 216*(1–3), 1-11.

Bremmer, R., Nadort, A., Van Leeuwen, T., & Van Gemert, M. (2011). Age estimation of bloodstains by hemoglobin derivative determination using reflectance spectroscopy. *Forensic Science international, 206*, 166-171.

Bruskov, V., Malakhova, L., Masalimov, Z., & Chernikov, A. (2002). Heat-induced formation of reactive oxygen species and 8-oxoguanine, a biomarker of damage to DNA. *Nucleic Acids Research, 30*(6), 1354-1363.

Brys, G., Hubert, M., & Struyf, A. (2008). Goodness-of-fit tests based on a robust measure of skewness. *Computational Statistics, 23*(3), 429-442.

Bustin, S. (2000). Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology, 25*(2), 169-193.

Butler, J. (2005). *Forensic DNA typing : biology, technology, and genetics of STR markers*. London: Elsevier Academic Press.

Cadet, J., Sage, E., & Douki, T. (2005). Ultraviolet radiation-mediated damage to cellular DNA. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 571*(1), 3-17.

Caponigro, G., & Parker, R. (1995). Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast. *Genes and Development, 9*, 2421-2432.

Clark, J. High performance liquid chromatography Retrieved March, 2012, from http://www.chemguide.co.uk/analysis/chromatography/hplc.html

Cline, R., Laurent, N., & Foran, D. (2003). The fingernails of Mary Sullivan: Developing reliable methods for selectively isolating endogenous and exogenous DNA from evidence. *Journal of forensic Sciences, 48*(2), 328-333.

Corder, G., & Foreman, D. (2009). *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*: Wiley.

Dallman, M., Montgomery, R., Larsen, C., Wanders, A., & Wells, A. (1991). Cytokine Gene Expression: Analysis using Northern Blotting, Polymerase Chain Reaction and in situ Hybridization. *Immunological Reviews, 119*(1), 163-179.

Dell'Anno, A., Fabiano, M., Duineveld, G., Kok, A., & Danovaro, R. (1998). Nucleic acid (DNA, RNA) quantification and RNA/DNA ratio determination in marine sediments: Comparison of spectrophotometric, fluorometric, and high-performance liquid chromatography methods and estimation of detrital DNA. *Applied and Environmental Microbiology, 64*(9), 3238-3245.

Desser, H., Hocker, P., Weiser, M., & Bohnel, J. (1975). The content of unbound polyamines in blood plasma and leukocytes of patients with polycythemia vera. *Clinica Chimica Acta, 63*, 243-247.

Elliott, D., & Ladomery, M. (2011). *Molecular biology of RNA*. Oxford: Oxford University Press.

Field, A. (2005). *Discovering statistics using SPSS : (and sex, drugs and rock'n'roll)* London: Sage.

Finnegan, M., Herbert, K., Evans, M., & Lunec, J. (1995). Phenol isolation of DNA yields higher levels of 8-oxodeoxyguanosine compared to pronase E isolation. *Biochemical Society Transactions, 23*(3), 430s.

Fiori, A. (1962). Detection and indentification of bloodstains: methods of forensic sciences. *Interscience, 1*, 243-290.

Flanagan, B., Wotton, D., Tuck-Wah, S., & Owen, M. (1990). DNase hypersensitivity and methylation of the human CD3G and D genes during T-cell development. *Immunogenetics, 31*(1), 13–20.

Fleige, S., & Pfaffl, M. (2006). RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular Aspects of Medicine, 27*(2-3), 126-139.

Fleming, R., & Harbison, S. (2010). The development of a mRNA multiplex RT-PCR assay for the definitive identification of body fluids. *Forensic Science International: Genetics, 4*(4), 244-256.

Fontanesi, L., Colombo, M., Beretti, F., & Russo, V. (2008). Evaluation of post mortem stability of porcine skeletal muscle RNA. *Meat Science, 80*(4), 1345-1351.

Foran, D. (2006). Relative Degradation of Nuclear and Mitochondrial DNA: An Experimental Approach. [10.1111/j.1556-4029.2006.00176.x]. *Journal of Forensic Sciences, 51*(4), 766-770.

Frederico, L., Kunkel, T., & Shaw, B. (1990). A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry, 29*(10), 2532-2537.

Fujita, Y., Tsuchiya, K., Abe, S., Takiguchi, Y., Kubo, S., & Sakurai, H. (2005). Estimation of the age of human bloodstains by electron paramagnetic resonance spectroscopy: Long-term controlled experiment on the effects of environmental factors. *Forensic Science International, 152*(1), 39-43.

Galtier, N., & Lobry, J. (1997). Relationships Between Genomic G+C Content, RNA Secondary Structures, and Optimal Growth Temperature in Prokaryotes. *Journal of Molecular Evolution, 44*(6), 632-636.

Gilbert, M., Binladen, J., Miller, W., Wiuf, C., Willerslev, E., Poinar, H., . . . Schuster, S. (2007). Recharacterization of ancient DNA miscoding lesions: insights in the era of sequencing-by-synthesis. *Nucleic Acids Research, 35*(1), 1-10.

Gopee, N., & Howard, P. (2007). A time course study demonstrating RNA stability in postmortem skin. *Experimental and Molecular Pathology, 83*(1), 4-10.

Gowrishankar, G., Winzen, R., Bollig, F., Ghebremedhin, B., Redich, N., Ritter, B., . . . Holtmann, H. (2005). Inhibition of mRNA deadenylation and degradation by ultraviolet light. *Biological Chemistry, 386*(12), 1287-1293.

Green, R., Roinestad, I., Boland, C., & Hennessy, L. (2005). Developmental validation of the Quantifiler™ real-time PCR kits for the quantification of human nuclear DNA samples. *Journal of Forensic Sciences, 50*(4), 809-825.

Hampson, C., Louhelainen, J., & McColl, S. (2011). An RNA Expression Method for Aging Forensic Hair Samples. *Journal of Forensic Sciences, 56*(2), 359-365.

Hanson, E., & Ballantyne, J. (2010). Blue Spectral Shift of the Hemoglobin Soret Band Correlates with the Age (Time Since Deposition) of Dried Bloodstains. *PLoS One, 5*, e12830.

Heinrich, M., Lutz-Bonengel, S., Matt, K., & Schmidt, U. (2007). Real-time PCR detection of five different ‚Äúendogenous control gene‚Äù transcripts in forensic autopsy material. *Forensic Science International: Genetics, 1*(2), 163-169.

Henssge, C., Brinkmann, B., & Puschel, K. (1984). Determination of the time of death by measurement of rectal temperature of corpses suspended in water. *Journal of legal medicine, 92*, 255-276.

Hofreiter, M., Jaenicke, V., Serre, D., & Haeseler, A. (2001). DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Research, 29*(23), 4793-4799.

Houseley, J., & Tollervey, D. (2009). The Many Pathways of RNA Degradation. *Cell, 136*(4), 763-776.

Hudlow, W., Chong, M., Swango, K., Timken, M., & Buoncristiani, M. (2008). A quadruplex real-time qPCR assay for the simultaneous assessment of total human DNA, human male DNA, DNA degradation and the presence of PCR inhibitors in forensic samples: A diagnostic tool for STR typing. *Forensic Science International: Genetics, 2*(2), 108-125.

Imbeaud, S., Graudens, E., Boulanger, V., Barlet, X., Zaborski, P., Eveno, E., . . . Auffray, C. (2005). Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Research, 33*(6), 56-56.

Inoue, H., Fukutaro, T., Iwasa, M., & Maeno, Y. (1991). Identification of fetal hemoglobin and simultaneous estimation of bloodstain age by high-performance liquid chromatography. *International Journal of Legal Medicine, 104*, 127-131.

Inoue, H., Fukutaro, T., Iwasa, M., Maeno, Y., & Seko, Y. (1992). A new marker for estimation of bloodstain age by high performance liquid chromatography. *Forensic Science International, 57*, 17-27.

Inoue, H., Kimura, A., & Tuji, T. (2002). Degradation profile of mRNA in a dead rat body: basic semi-quantification study. *Forensic Science International, 130*(2), 127-132.

Jones, L., Yue, S., Cheung, C., & Singer, V. (1998). RNA quantitation by fluorescence-based solution assay: RiboGreen reagent characterization. *Analytical Biochemistry, 265*(2), 368-374.

Juusola, J., & Ballantyne, J. (2003). Messenger RNA profiling: a prototype method to supplant conventional methods for body fluid identification. *Forensic Science International, 135*(2), 85-96.

Kaiser, C., Bachmeier, B., Conrad, C., Nerlich, A., Bratzke, H., Eisenmenger, W., & Peschel, O. (2008). Molecular study of time dependent changes in DNA stability in soil buried skeletal residues. *Forensic Science International, 177*(1), 32-36.

Kaliszan, M., Hauser, R., & Kernbach-Wighton, G. (2009). Estimation of the time of death based on the assessment of post mortem processes with emphasis on body cooling. *Legal Medicine, 11*, 111-117.

Kanthaswamy, S., Premasuthan, A., Ng, J., Satkoski, J., & Goyal, V. (2012). Quantitative real-time PCR (qPCR) assay for human, dog and cat species identification and nuclear DNA quantification. *Forensic Science International: Genetics, 6*(2), 290-295.

Kapitulnik, J. (2004). Bilirubin: An Endogenous Product of Heme Degradation with Both Cytotoxic and Cytoprotective Properties. *Molecular Pharmacology, 66*(4), 773-779.

Katz, L., & Penman, S. (1966). Association by hydrogen bonding of free nucleosides in non-aqueous solution. *Journal of Molecular Biology, 15*, 220-231.

Kaukinen, U., & Mikkola, S. (2002). The reactivity of phosphodiester bonds within linear single-stranded oligoribonucleotides is strongly dependent on the base sequence. *Nucleic Acids Research, 30*(2), 468-474.

Kierzek, R. (2001). Nonenzymatic cleavage of oligoribonucleotides *Methods in Enzymology* (Vol. Volume 341, pp. 657-675): Academic Press.

Kind, S., Patterson, D., & Owen, G. (1972). Estimation of the age of dried blood stains by spectrophotometric method. *Forensic Science, 1*, 27-54.

King, G., Gilbert, M., Willerslev, E., Collins, M., & Kenward, H. (2009). Recovery of DNA from archaeological insect remains: first results, problems and potential. *Journal of Archaeological Science, 36*(5), 1179-1183.

Komiyama, M., Takeda, N., & Shigekawa, H. (1999). Hydrolysis of DNA and RNA by lanthanide ions: mechanistic studies leading to new applications. *Chemical Communications*(16), 1443-1451.

Kuliwaba, J., Fazzalari, N., & Findlay, D. (2005). Stability of RNA isolated from human trabecular bone at post-mortem and surgery. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 1740*(1), 1-11.

La Paz, J., Esteve, T., & Pla, M. (2007). Comparison of Real-Time PCR Detection Chemistries and Cycling Modes Using Mon810 Event-Specific Assays as Model. *Journal of Agricultural and Food Chemistry, 55*(11), 4312-4318.

Lamola, A., & Mittal, J. (1966). Solution Photochemistry of Thymine and Uracil. *Science, 154*(3756), 1560-1561.

Lee, K. (1997). Quantification of total RNA by ethidium bromide fluorescence may not accurately reflect the RNA mass. *Journal of Biochemical and Biophysical Methods, 34*(2), 147-154.

Leers, O. (1910). Modern Medicine. *The Journal of the American Medical Association, 55*, 151.

Lemm, I., & Ross, J. (2002). Regulationof c-myc mRNA Decay by Translational Pausing in a Coding Region Instability Determinant. *Molecular and Cellular Biology, 22*, 3959-3969.

Lightfoot, S. (2002). *Quantitation comparison of total RNA using the Agilent 2100 Bioanalyzer, ribogreen analysis and UV spectrometry.* Palo Alto, CA: Agilent Technologies.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature, 362*, 709-715.

Lindahl, T., & Andersson, A. (1972). Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid. *Biochemistry, 11*(19), 3618-3623.

Lindahl, T., & Nyberg, B. (1972). Rate of depurination of native deoxyribonucleic acid. *Biochemistry, 11*(19), 3610-3618.

Lindahl, T., & Nyberg, B. (1974). Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry, 13*(16), 3405-3410.

Lindberg, J., & Lundeberg, J. (2010). The plasticity of the mammalian transcriptome. *Genomics, 95*(1), 1-6.

Loeb, L., & Preston, B. (1986). Mutagenesis by Apurinic/Apyrimidinic Sites. *Annual Review of Genetics, 20*(1), 201-230.

Lorentzen, E., & Conti, E. (2006). The Exosome and the Proteasome: Nano-Compartments for Degradation. *Cell, 125*(4), 651-654.

MacCallum, C. (1995). Model specification: procedures, strategies, and related issues. *Structural equation modelling: concepts, issues, and applications.* (pp. 37-55). Thousands Oaks: Sage Publications,.

Madisen, L., Hoar, D., & Holroyd, C. (1987). DNA banking: The effects of storage of blood and isolated DNA on the integrity of DNA. *American Journal of Medical Genetics, 27*(2), 379-390.

Mandrioli, M., Borsatti, F., & Mola, L. (2006). Factors affecting DNA preservation from museum-collected lepidopteran specimens. *Entomologia Experimentalis et Applicata, 120*(3), 239-244.

Marchuk, L., Sciore, P., Reno, C., Frank, C., & Hart, D. (1998). Postmortem stability of total RNA isolated from rabbit ligament, tendon and cartilage. *Biochimica et Biophysica Acta (BBA) - General Subjects, 1379*(2), 171-177.

Marrone, A., & Ballantyne, J. (2009). Changes in Dry State Hemoglobin over Time Do Not Increase the Potential for Oxidative DNA Damage in Dried Blood. *PLoS ONE, 4*(4), e5110.

Marrone, A., & Ballantyne, J. (2010). Hydrolysis of DNA and its molecular components in the dry state. *Forensic Science International: Genetics, 4*(3), 168-177.

Martinez, W., & Martinez, A. (2004). *Exploratory Data Analysis with MATLAB*: Chapman and Hall.

Matsuoka, T., Taguchi, T., & Okuda, J. (1995). Estimation of bloodstain age by rapid determinations of oxyhemoglobin by use of oxygen-electrode and total hemoglobin. *Biological and Pharmaceutical Bulletin, 18*, 1031-1035.

McNally, L., Shaler, C., Baird, M., Balazs, I., Kobilinsky, L., & De Forest, P. (1989). The effects of environment and substrata on deoxyribonucleic acid (DNA): The use of casework samples from New York City. *Journal of Forensic Sciences, 34*(5), 1070-1077.

Melton, T., & Nelson, K. (2001). Forensic mitochondrial DNA analysis: Two years of commercial casework experience in the United States. *Croatian medical journal, 42*(3), 298-303.

Miki, T., Kai, A., & Ikeya, M. (1987). Electron spin resonance of bloodstains and its application to the estimation of time after bleeding. *Forensic Science International, 35*, 149-158.

Minifiler FAS Questions. Retrieved August, 2012, from http://marketing.appliedbiosystems.com/mk/get/minifiler_fas_questions

Miura, Y., Ichikawa, Y., Ishikawa, T., Ogura, M., DeFries, R., Shimada, H., & Mitsuhashi, M. (1996). Fluorometric determination of total mRNA with oligo(dT) immobilized on microtiter plates. *Clinical Chemistry, 42*(11), 1758-1764.

Neidle, S., Schneider, B., & Berman, H. (2005). Fundamentals of DNA and RNA Structure *Structural Bioinformatics* (pp. 41-73): John Wiley & Sons, Inc.

Nielsen, K., Mogensen, H., Hedman, J., Niederstatter, H., Parson, W., & Morling, N. (2008). Comparison of five DNA quantification methods. *Forensic Science International: Genetics, 2*(3), 226-230.

Noguchi, I., Arai, H., & Iizuka, R. (1991). A study on postmortem stability of vasopressin messenger RNA in rat brain compared with those in total RNA and ribosomal RNA. *Journal of Neural Transmission, 83*(3), 171-178.

Nuoreteva, P. (1974). Age determination of a bloodstain in a decaying shirt by entomological means. *Forensic Science, 3*, 89-94.

Nussbaumer, C., Gharehbaghi-Schnell, E., & Korschineck, I. (2006). Messenger RNA profiling: A novel method for body fluid identification by Real-Time PCR. *Forensic Science International, 157*(2), 181-186.

O'Rourke, D., Hayes, M., & Carlyle, S. (2000). Ancient DNA studies in physical anthropology. *Annual Review of Anthropology, 29*(1), 217-242.

O'brien, R. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity: International Journal of Methodology, 41*(5), 673-690.

Ohshima, T., & Sato, Y. (1998). Time-dependent expression of interleukin-10 (IL-10) mRNA during the early phase of skin wound healing as a possible indicator of wound vitality. *Int J Legal Med, 111*(5), 251-255.

Parker, R., & Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. *Nature Structural and Molecular Biology, 11*, 121-127.

Patterson, D. (1960). Use of reflectance measurements in assessing the colour changes of ageing bloodstains. *Nature, 187*, 688-689.

Pfeifer, G., Young-Hyun, Y., & Besaratinia, A. (2005). Mutations induced by ultraviolet light. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 571*(1), 19-31.

Promega. DNA IQ™ System - Small Sample Casework Technical Bulletin Retrieved September, 2012, from https://www.promega.com/resources/protocols/technical-bulletins/101/dna-iq-systemsmall-sample-casework-protocol/

Qi, B., Kong, L., & Lu, Y. (2013). Gender-related difference in bloodstain RNA ratio stored under uncontrolled room conditions for 28 days. *Journal of Forensic and Legal Medicine, 20*(4), 321-325.

Qiu, X., Wu, L., Huang, H., McDonel, P., Palumbo, A., Tiedje, J., & Zhou, J. (2001). Evaluation of PCR-Generated Chimeras, Mutations, and Heteroduplexes with 16S rRNA Gene-Based Cloning. *Applied and Environmental Microbiology, 67*(2), 880-887.

Quantifiler Human DNA Quantification Kit. Retrieved April, 2012, from http://tools.invitrogen.com/content/sfs/manuals/cms_041395.pdf

Rajamannar, K. (1977). Determination of the Age of Bloodstains using Immunoelectrophoresis. *Journal of Forensic Science, 22*(1).

Rauschke, J. (1951). Beitrage zur frage de alterbestimmung von blutspuren. *International Journal of Legal Medicine, 40*, 578-584.

Real-Time PCR Vs. Traditional PCR. Retrieved May, 2011, from http://www6.appliedbiosystems.com/support/tutorials/pdf/rtpcr_vs_tradpcr.pdf

Real-time PCR: Understanding Ct. Retrieved May, 2011, from http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_053906.pdf

Regression with SPSS. Retrieved July 2012, from https://www.ats.ucla.edu/stat/spss/webbooks/reg/chapter2/spssreg2.htm

Ritz-Timme, S., Cattaneo, C., Collins, M., Waite, E., Schutz, H., & Kaatsch, H. (2000). Age estimation: the state of the art in relation to the specific demands of forensic practise. *International Journal of Legal Medicine, 113*, 129-136.

Ross, J. (1995). mRNA Stability in Mammalian Cells. *Microbiological Reviews, 97*, 423-450.

Rozen, S., & Skaletsky, H. Primer3 Retrieved April, 2012, from http://primer3.sourceforge.net

Ruiz, Y., Phillips, C., Gomez-Tato, A., Alvarez-Dios, J., Casares de Cal, M., Cruz, R., . . . Lareu, M. (2012). Further development of forensic eye color predictive tests. *Forensic Science International: Genetics*.

Sachs, A. (1993). Messenger RNA degradation in eukaryotes. *Cell, 74*, 413-421.

Saeed, M., Berlin, R., & Cruz, T. (2012). Exploring the utility of genetic markers for predicting biological age. *Legal Medicine*.

Sakurai, H., & Tsuchiya, K. (1989). Dating of Human Blood by Electron Spin Resonance Spectroscopy. *Naturwissenschaften, 76*, 24-25.

Salkind, N., & Green, S. (2003). *Using SPSS for Windows and Macintosh : analyzing and understanding data.* N.J.: Prentice Hall.

Schaffer, S., & Suleiman, M. (2007). *Mitochondria: the dynamic organelle*. New York: Springer.

Schmittgen, T., Zakrajsek, B., Mills, A., Gorn, V., Singer, M., & Reed, M. (2000). Quantitative Reverse Transcription, Polymerase Chain Reaction to Study mRNA Decay: Comparison of Endpoint and Real-Time Methods. *Analytical Biochemistry, 285*(2), 194-204.

Schroeder, G., Lad, C., Wyman, P., Williams, N., & Wolfenden, R. (2006). The time required for water attack at the phosphorus atom of simple phosphodiesters and of DNA. *Proceedings of the National Academy of Sciences of the United States of America, 103*(11), 4052-4055.

Schwarz, F. (1937). Quantitative Untersuchungen der Katalase und Peroxydase im Blutfleck,. *International Journal of Legal Medicine, 27*, 1-34.

Schwarzacher, W. (1930). Determination of the age of bloodstains. *American Journal of Police Science, 1*(4), 374-380.

Sensabaugh, G., Wilson, A., & Kirk, P. (1971). Protein stabiliy in preserved biological remains. *International Journal of Biochemistry, 2*, 545-557.

Setlow, P. (1992). I will survive: protecting and repairing spore DNA. *Journal of Bacteriology, 174*(9), 2737-2741.

Simard, A., DesGroseillers, L., & Sarafian, V. (2012). Assessment of RNA Stability for Age Determination of Body Fluid Stains. *Canadian Society of Forensic Science Journal, 45*(4).

Stapel, E. Box-and-Whisker Plots: Interquartile Ranges and Outliers: Purplemath Retrieved February, 2012, from http://www.purplemath.com/modules/boxwhisk3.htm

Strasser, S., Zink, A., Kada, G., Hinterdorfer, P., Peschel, O., Heckl, W., . . . Thalhammer, S. (2007). Age determination of blood spots in forensic medicine by force spectroscopy. *Forensic Science International, 170*, 8-14.

Swango, K., Hudlow, W., Timken, M., & Buoncristiani, M. (2007). Developmental validation of a multiplex qPCR assay for assessing the quantity and quality of nuclear DNA in forensic samples. *Forensic Science International, 170*(1), 35-45.

Swango, K., Timken, M., Chong, M., & Buoncristiani, M. (2006). A quantitative PCR assay for the assessment of DNA degradation in forensic samples. *Forensic Science International, 158*(1), 14-26.

Thompson, D., Lu, C., Green, P., & Parker, R. (2008). tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *RNA, 14*(10), 2095-2103.

Tomellini, L. (1907). De l'emploi d'une table chromatique pour les taches du sang. *Archives d'antropologie criminelle de Criminologie, 14*, 2.

Tsutsumi, A., Yamamoto, Y., & Ishizu, H. (1983). Determination of the age of bloodstains by enzyme activities in blood cells. *Nihon Hoigaku Zassh, 37*, 770-776.

Tuffery, S. (2011). Data Exploration and Preparation *Data Mining and Statistics for Decision Making* (pp. 43-91): John Wiley & Sons, Ltd.

van Hoof, A., & Parker, R. (2002). Messenger RNA Degradation: Beginning at the End. *Current Biology, 12*, 285-287.

Venter, J., Adams, M., Myers, E., Li, P., Mural, R., Sutton, G., . . . Milshina, N. (2001). The Sequence of the Human Genome. *Science, 291*(5507), 1304-1351.

Whittemore, S. (2004). *The circulatory system*. New York: Chelsea House.

Wu, D., Ugozzoli, L., Pal, B., Qian, J., & Wallace, R. (1991). The Effect of Temperature and Oligonucleotide Primer Length on the Specificity and Efficiency of Amplification by the Polymerase Chain Reaction. *DNA and Cell Biology, 10*(3), 233-238.

Zhao, D., Zhu, B., Ishikawa, T., Quan, L., Li, D., & Maeda, H. (2006). Real-time RT-PCR quantitative assays and postmortem degradation profiles of erythropoietin,

vascular endothelial growth factor and hypoxia-inducible factor 1 alpha mRNA transcripts in forensic autopsy materials. *Legal Medicine, 8*(2), 132-136.

Zubakov, D., Hanekamp, E., Kokshoorn, M., van IJcken, W., & Kayser, M. (2008). Stable RNA markers for identification of blood and saliva stains revealed from whole genome expression analysis of time-wise degraded samples. *International Journal of Legal Medicine, 122*, 135–142.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research, 31*(13), 3406-3415.