WILLIAM GOODWIN
ADRIAN LINACRE
SIBTE HADI

AN INTRODUCTION TO Forensic Genetics

WILEY

An Introduction to Forensic Genetics

William Goodwin
University of Central Lancashire, UK

Adrian Linacre
University of Strathclyde, UK

Sibte Hadi
University of Central Lancashire, UK

John Wiley & Sons, Ltd

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wileyeurope.com or www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 33 Park Road, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 6045 Freemont Blvd, Mississauga, Ontario, L5R 4J3, Canada

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Anniversary Logo Design: Richard J. Pacifico

Library of Congress Cataloging-in-Publication Data

Goodwin, William, Dr.
An introduction to forensic genetics / William Goodwin, Adrian Linacre, Sibte Hadi.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-0-470-01025-9 (alk. paper) – ISBN 978-0-470-01026-6 (alk. paper)
1. Forensic genetics. I. Linacre, Adrian. II. Hadi, Sibte. III. Title.
[DNLM: 1. Forensic Genetics–methods. 2. DNA Fingerprinting.
3. Microsatellite Repeats. W 700 G657i 2007]

RA1057.5.G67 2007
614′.1–dc22 2007019041

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN (HB) 9780470010259
ISBN (PB) 9780470010266

Typeset in 10.5/12.5pt Times by Aptara, New Delhi, India
Printed and bound in Great Britain by Antony Rowe Ltd. Chippenham, Wiltshire
This book is printed on acid-free paper responsibly manufactured from sustainable forestry
in which at least two trees are planted for each one used for paper production.

An Introduction to Forensic Genetics

Preface

It is strange to consider that the use of DNA in forensic science has been with us for just over 20 years and, while a relatively new discipline, it has impacted greatly on the criminal justice system and society as a whole. It is routinely the case that DNA figures in the media, in both real cases and fictional scenarios.

The increased interest in forensic science has led to a burgeoning of university courses with modules in forensic science. This book is aimed at undergraduate students studying courses or modules in Forensic Genetics.

We have attempted to take the reader through the process of DNA profiling from the collection of biological evidence to the evaluation and presentation of genetic evidence. While each chapter can stand alone, the order of chapters is designed to take the reader through the sequential steps in the generation of a DNA profile. The emphasis is on the use of short tandem repeat (STR) loci in human identification as this is currently the preferred technique. Following on from the process of generating a DNA profile, we have attempted to describe in accessible terms how a DNA profile is interpreted and evaluated. Databases of DNA profiles have been developed in many countries and hence there is need to examine their use. While the focus of the book is on STR analysis, chapters on lineage markers and single nucleotide polymorphisms (SNPs) are also provided.

As the field of forensic science and in particular DNA profiling moves onward at a rapid pace, there are few introductory texts that cover the current state of this science. We are aware that there is a range of texts available that cover specific aspects of DNA profiling and where there this is the case, we direct readers to these books, papers or web sites.

We hope that the readers of this book will gain an appreciation of both the underlying principles and application of forensic genetics.

About the Authors

William Goodwin is a Senior Lecture in the Department of Forensic and Investigative Science at the University of Central Lancashire where his main teaching areas are molecular biology and its application to forensic analysis. Prior to this he worked for eight years at the Department of Forensic Medicine and Science in the Human Identification Centre where he was involved in a number of international cases involving the identifications of individuals from air crashes and from clandestine graves. His research has focused on the analysis of DNA from archaeological samples and highly compromised human remains. He has acted as an expert witness and also as a consultant for international humanitarian organisations and forensic service providers.

Adrian Linacre is a Senior Lecturer at the Centre for Forensic Science at the University of Strathclyde where his main areas of teaching are aspects of forensic biology, population genetics and human identification. His research areas include the use of non-human DNA in forensic science and the mechanisms behind the transfer and persistence of DNA at crime scenes. He has published over 50 papers in international journals, has presented at a number of international conferences and is on the editorial board of Forensic Science International: Genetics. Dr Linacre works as an assessor for the Council for the Registration of Forensic Practitioners (CRFP) in the area of human contact traces and is a Registered Practitioner.

Sibte Hadi is a Senior Lecture in the Department of Forensic and Investigative Science at the University of Central Lancashire. His main teaching areas are Forensic Medicine and DNA profiling. He is a physician by training and practised forensic pathology for a number of years in Pakistan before undertaking a PhD in Forensic Genetics. Following this he worked at the Department of Molecular Biology Louisiana State University as a member of the Louisiana Healthy Aging Study group. He has acted as a consultant to forensic service providers in the USA and Pakistan. His current research is focused on population genetics, DNA databases and gene expression studies for different forensic applications.

Contents

Preface ix
About the Authors x

1 Introduction to forensic genetics 1
Forensic genetics 1
A brief history of forensic genetics 2
References 5

2 DNA structure and the genome 7
DNA structure 7
Organization of DNA into chromosomes 7
The structure of the human genome 9
Genetic diversity of modern humans 11
The genome and forensic genetics 11
Tandem repeats 12
Single nucleotide polymorphisms (SNPs) 13
Further reading 14
References 14

3 Biological material – collection, characterization and storage 17
Sources of biological evidence 17
Collection and handling of material at the crime scene 19
Identification and characterization of biological evidence 19
Evidence collection 20
Sexual and physical assault 21
Presumptive testing 21
Storage of biological material 23
References 24

4 DNA extraction and quantification 27
DNA extraction 27
General principles of DNA extraction 27
DNA extraction from challenging samples 30
Quantification of DNA 32
DNA IQ™ system 36
References 36

5 The polymerase chain reaction **39**
The evolution of PCR-based profiling in forensic genetics 39
DNA replication – the basis of the PCR 40
The components of PCR 40
The PCR process 42
PCR inhibition 44
Sensitivity and contamination 45
The PCR laboratory 46
Further reading 48
References 48

6 The analysis of short tandem repeats **51**
Structure of STR loci 51
The development of STR multiplexes 51
Detection of STR polymorphisms 54
Interpretation of STR profiles 56
Further reading 61
References 61

7 Assessment of STR profiles **65**
Stutter peaks 65
Split peaks (+/−A) 65
Pull-up 67
Template DNA 68
Overloaded profiles 68
Low copy number DNA 68
Peak balance 70
Mixtures 70
Degraded DNA 71
References 73

8 Statistical interpretation of STR profiles **75**
Population genetics 75
Deviation from the Hardy–Weinberg equilibrium 76
Statistical tests to determine deviation from the Hardy–Weinberg equilibrium 77
Estimating the frequencies of STR profiles 78
Corrections to allele frequency databases 78
Which population frequency database should be used? 83
Conclusions 83
Further reading 84
References 84

9 The evaluation and presentation of DNA evidence **87**
Hierarchies of propositions 87
Likelihood ratios 89
Two fallacies 93
Comparison of three approaches 94
Further reading 95
References 95

10 Databases of DNA profiles **97**
The UK National DNA database (NDNAD) 97
International situation 102
References 104

11 Kinship testing **105**
Paternity testing 105
Identification of human remains 111
Further reading 112
References 112

12 Single nucleotide polymorphisms **115**
SNPs – occurrence and structure 115
Detection of SNPs 115
SNP detection for forensic applications 117
Forensic applications of SNPs 119
SNPs compared to STR loci 120
Further reading 121
References 121

13 Lineage markers **125**
Mitochondria 125
Applications of mtDNA profiling 127
The Y chromosome 130
Forensic applications of Y chromosome polymorphisms 131
Further reading 132
References 133

Appendix 1 Forensic parameters **137**

Appendix 2 Useful web links **139**

Glossary **141**

Abbreviations **145**

Index **147**

1 Introduction to forensic genetics

Over the last 20 years the development and application of genetics has revolutionized forensic science. In 1984, the analysis of polymorphic regions of DNA produced what was termed 'a DNA fingerprint' [1]. The following year, at the request of the United Kingdom Home Office, DNA profiling was successfully applied to a real case, when it was used to resolve an immigration dispute [2]. Following on from this, in 1986, DNA evidence was used for the first time in a criminal case and identified Colin Pitchfork as the killer of two school girls in Leicestershire, UK. He was convicted in January 1988. The use of genetics was rapidly adopted by the forensic community and plays an important role worldwide in the investigation of crime. Both the scope and scale of DNA analysis in forensic science is set to continue expanding for the foreseeable future.

Forensic genetics

The work of the forensic geneticist will vary widely depending on the laboratory and country that they work in, and can involve the analysis of material recovered from a scene of crime, paternity testing and the identification of human remains. In some cases, it can even be used for the analysis of DNA from plants [3, 4], animals [5, 6] and microorganisms [7]. The focus of this book is the analysis of biological material that is recovered from the scene of crime – this is central to the work of most forensic laboratories. Kinship testing will be dealt with separately in Chapter 11.

Forensic laboratories will receive material that has been recovered from scenes of crime, and reference samples from both suspects and victims. The role of forensic genetics within the investigative process is to compare samples recovered from crime scenes with suspects, resulting in a report that can be presented in court or intelligence that may inform an enquiry (Figure 1.1).

Several stages are involved with the analysis of genetic evidence (Figure 1.2) and each of these is covered in detail in the following chapters.

In some organizations one person will be responsible for collecting the evidence, the biological and genetic analysis of samples, and ultimately presenting the results to a court of law. However, the trend in many larger organizations is for individuals to be

An Introduction to Forensic Genetics W. Goodwin, A. Linacre and S. Hadi

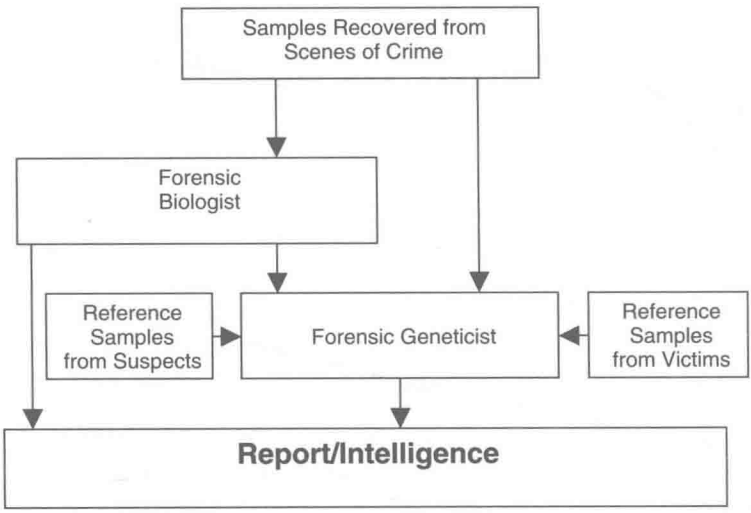

Figure 1.1 The role of the forensic geneticist is to assess whether samples recovered from a crime scene match to a suspect. Reference samples are provided from suspects and also victims of crime

responsible for only a very specific task within the process, such as the extraction of DNA from the evidential material or the analysis and interpretation of DNA profiles that have been generated by other scientists.

A brief history of forensic genetics

In 1900 Karl Landsteiner described the ABO blood grouping system and observed that individuals could be placed into different groups based on their blood type. This was the first step in the development of forensic haemogenetics. In 1915 Leone Lattes published a book describing the use of ABO typing to resolve a paternity case and by 1931 the absorption–inhibition ABO typing technique that became standard in forensic laboratories had been developed. Following on from this, numerous blood group markers and soluble blood serum protein markers were characterized and could be analysed in combination to produce highly discriminatory profiles. The serological techniques were a powerful tool but were limited in many forensic cases by the amount of biological material that was required to provide highly discriminating results. Proteins are also prone to degradation on exposure to the environment.

In the 1960s and 1970s, developments in molecular biology, including restriction enzymes, Sanger sequencing [8], and Southern blotting [9], enabled scientists to examine DNA sequences. By 1978, DNA polymorphisms could be detected using Southern blotting [10] and in 1980 the analysis of the first highly polymorphic locus was reported [11]. It was not until September 1984 that Alec Jeffreys realized the potential forensic application of the variable number tandem repeat (VNTR) loci he had been studying [1, 12]. The technique developed by Jeffreys entailed

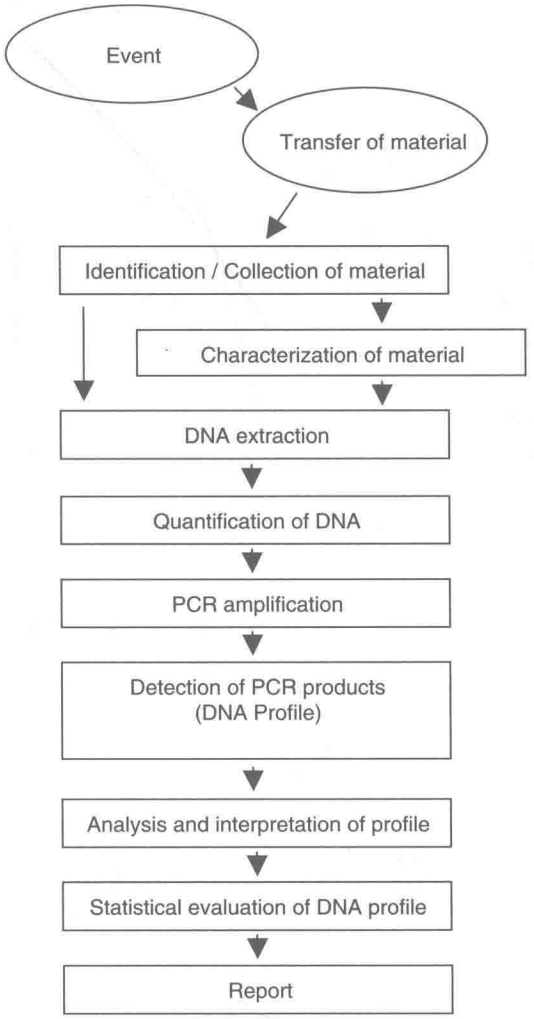

Figure 1.2 Processes involved in generating a DNA profile following a crime. Some types of material, in particular blood and semen, are often characterized before DNA is extracted

extracting DNA and cutting it with a restriction enzyme, before carrying out agarose gel electrophoresis, Southern blotting and probe hybridization to detect the polymorphic loci. The end result was a series of black bands on X-ray film (Figure 1.3). VNTR analysis was a powerful tool but suffered from several limitations: a relatively large amount of DNA was required; it would not work with degraded DNA; comparison between laboratories was difficult; and the analysis was time consuming.

A critical development in the history of forensic genetics came with the advent of a process that can amplify specific regions of DNA – the polymerase chain reaction (PCR) (see Chapter 5). The PCR process was conceptualised in 1983 by Kary Mullis, a chemist

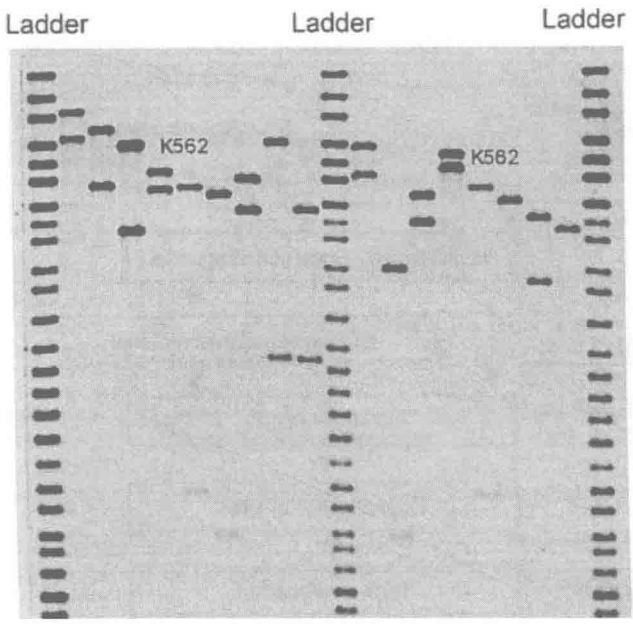

Figure 1.3 VNTR analysis using a single locus probe: ladders were run alongside the tested samples that allowed the size of the DNA fragments to be estimated. A control sample labelled K562 is analysed along with the tested samples

working for the Cetus Corporation in the USA [13]. The development of PCR has had a profound effect on all aspects of molecular biology including forensic genetics, and in recognition of the significance of the development of the PCR, Kary Mullis was awarded the Nobel Prize for Chemistry in 1993. The PCR increased the sensitivity of DNA analysis to the point where DNA profiles could be generated from just a few cells, reduced the time required to produce a profile, could be used with degraded DNA and allowed just about any polymorphism in the genome to be analysed. The first application of PCR in a forensic case involved the analysis of single nucleotide polymorphisms in the DQα locus [14] (see Chapter 12). This was soon followed by the analysis of short tandem repeats (STRs) which are currently the most commonly used genetic markers in forensic science (see Chapters 6 to 8). The rapid development of technology for analysing DNA includes advances in DNA extraction and quantification methodology, the development of commercial PCR based typing kits and equipment for detecting DNA polymorphisms.

In addition to technical advances, another important part of the development of DNA profiling that has had an impact on the whole field of forensic science is quality control. The admissibility of DNA evidence was seriously challenged in the USA in 1987 – '*People* v. *Castro*' [15]; this case and subsequent cases have resulted in increased levels of standardization and quality control in forensic genetics and other areas of

forensic science. As a result, the accreditation of both laboratories and individuals is an increasingly important issue in forensic science. The combination of technical advances, high levels of standardization and quality control have led to forensic DNA analysis being recognized as a robust and reliable forensic tool worldwide.

References

1. Jeffreys, A.J. *et al.* (1985) Individual-specific fingerprints of human DNA. *Nature* **316**, 76–79.
2. Jeffreys, A.J. *et al.* (1985) Positive identification of an immigration test-case using human DNA fingerprints. *Nature* **317**, 818–819.
3. Kress, W.J. *et al.* (2005) Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 8369–8374.
4. Linacre, A. and Thorpe, J. (1998) Detection and identification of cannabis by DNA. *Forensic Science International* **91**, 71–76.
5. Parson, W. *et al.* (2000) Species identification by means of the cytochrome b gene. *International Journal of Legal Medicine* **114** (1–2), 23–28.
6. Hebert, P.D.N. *et al.* (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, S96–S99.
7. Hoffmaster, A.R. *et al.* (2002) Molecular subtyping of *Bacillus anthracis* and the 2001 bioterrorism-associated anthrax outbreak, United States. *Emerging Infectious Diseases* **8**, 1111–1116.
8. Sanger, F. *et al.* (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–5467.
9. Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503–517.
10. Kan, Y.W. and Dozy, A.M. (1978) Polymorphism of DNA sequence adjacent to human B-globin structural gene: relationship to sickle mutation. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 5631–5635.
11. Wyman, A.R. and White, R. (1980) A highly polymorphic locus in human DNA. *Proceedings of the National Academy of Sciences of the United States of America* **77**, 6754–6758.
12. Jeffreys, A.J. and Wilson, V. (1985) Hypervariable regions in human DNA. *Genetical Research* **45**, 213–213.
13. Saiki, R.K. *et al.* (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle-cell anemia. *Science* **230**, 1350–1354.
14. Stoneking, M. *et al.* (1991) Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. *American Journal of Human Genetics* **48**, 370–382.
15. Patton, S.M. (1990) DNA fingerprinting: the Castro case. *Harvard Journal of Law and Technology* **3**, 223–240.

2 DNA structure and the genome

Each person's genome contains a large amount of DNA that is a potential target for DNA profiling. The selection of the particular region of polymorphic DNA to analyse can change with the individual case and also the technology that is available. In this chapter a brief description of the primary structure of the DNA molecule is provided along with an overview of the different categories of DNA that make up the human genome. The criteria that the forensic geneticist uses to select which loci to analyse are also discussed.

DNA structure

DNA has often been described as the 'blueprint of life', containing all the information that an organism requires in order to function and reproduce. The DNA molecule that carries out such a fundamental biological role is relatively simple. The basic building block of the DNA molecule is the nucleotide triphosphate (Figure 2.1a). This comprises a triphosphate group, a deoxyribose sugar (Figure 2.1b) and one of four bases (Figure 2.1c).

The information within the DNA 'blueprint' is coded by the sequence of the four different nitrogenous bases, adenine, guanine, thymine and cytosine, on the sugar-phosphate backbone (Figure 2.2a).

DNA normally exists as a double stranded molecule which adopts a helical arrangement – first described by Watson and Crick in 1953 [1]. Each base is attracted to its complementary base: adenine always pairs with thymine and cytosine always pairs with guanine (Figure 2.2b).

Organization of DNA into chromosomes

Within each nucleated human cell there are two complete copies of the genome. The genome is 'the haploid genetic complement of a living organism' and in humans contains approximately 3 200 000 000 base pairs (bp) of information, which is organized into 23 chromosomes. Humans contain two sets of chromosomes – one version of each chromosome inherited from each parent giving a total of 46 chromosomes (Figure 2.3). Each chromosome contains one continuous strand of DNA, the largest – chromosome

An Introduction to Forensic Genetics W. Goodwin, A. Linacre and S. Hadi

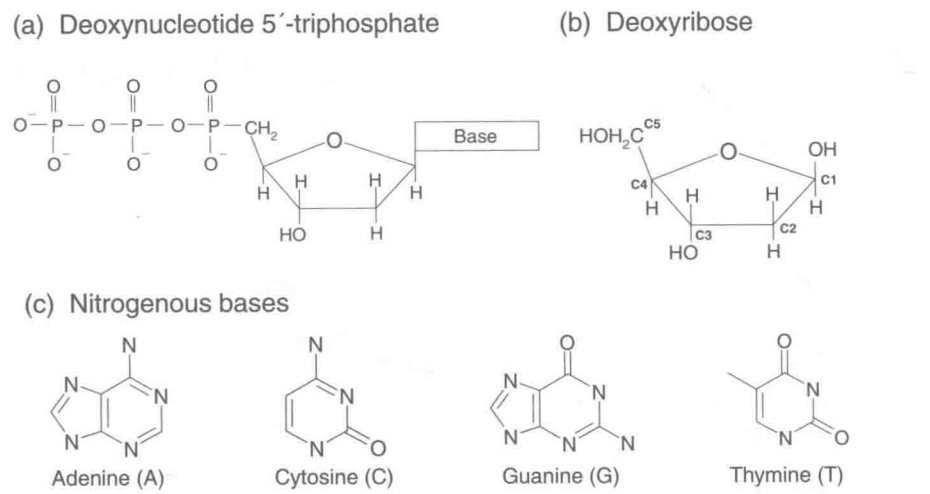

Figure 2.1 The DNA molecule is built up of deoxynucleotide 5′-triphosphates (2.1a). The sugar (2.1b) contains five carbon atoms (labelled C1 to C5); one of four different types of nitrogenous base (2.1c) is attached to the 1 prime (1′) carbon, a hydroxyl group to the 3′ carbon and the phosphate group to the 5′ carbon

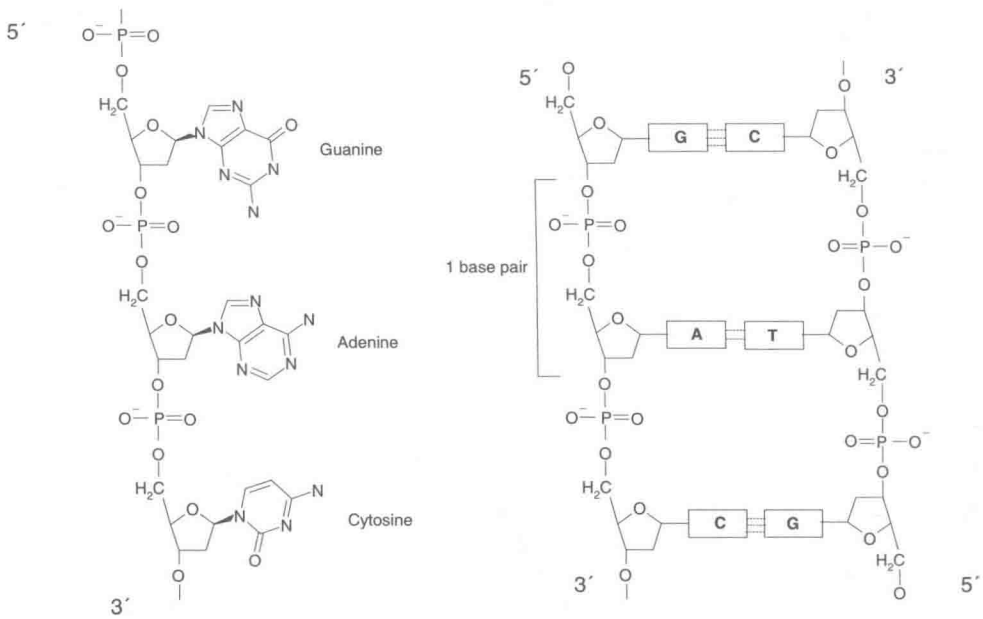

Figure 2.2 In the DNA molecule the nucleotides are joined together by phosphodiester bonds to form a single stranded molecule (2.2a). The DNA molecule in the cell is double stranded (2.2b) with two complementary single stranded molecules held together by hydrogen bonds. Adenine and thymine form two hydrogen bonds while guanine and cytosine form three bonds