# Essential Medical Genetics

# Essential Medical Genetics

J. M. CONNOR
BSc. MB ChB. MD. FRCP
*Professor of Medical Genetics*
*and Director of the West of Scotland*
*Regional Genetics Service.*
*University of Glasgow and*
*Duncan Guthrie Institute*
*of Medical Genetics. Yorkhill*
*Glasgow*

M. A. FERGUSON-SMITH
FRS. FRSE. FRCP. FRCPath
*Professor of Pathology, University of Cambridge*
*and Director. Cambridge University Centre*
*for Medical Genetics. and East Anglian*
*Regional Genetics Service*
*Addenbrooke's Hospital*
*Cambridge*

THIRD EDITION

OXFORD
BLACKWELL SCIENTIFIC PUBLICATIO
LONDON EDINBURGH BOSTON
MELBOURNE PARIS BERLIN VIENNA

Blackwell Scientific Publications
Editorial Offices:
Osney Mead, Oxford OX2 0EL
25 John Street, London WC1N 2BL
23 Ainslie Place, Edinburgh EH3 6AJ
3 Cambridge Center, Suite 208
Cambridge, Massachusetts 02142, USA
54 University Street, Carlton
Victoria 3053, Australia



First published 1984
Reprinted 1986
Second edition 1987
Third edition 1991

Set by Setrite Typesetters, Hong Kong
Printed and bound in Spain by Printek, s.a., Bilbao

DISTRIBUTORS

Marston Book Services Ltd
PO Box 87
Oxford OX2 0DT
(*Orders*: Tel: 0865 791155
Fax: 0865 791927
Telex: 837515)

USA
Mosby-Year Book, Inc.
200 North LaSalle Street
Chicago, Illinois 60601
(*Orders*: Tel: (312) 726–9733)

Canada
Mosby-Year Book, Inc.
5240 Finch Avenue East
Scarborough, Ontario
(*Orders*: Tel: (416) 298–1588)

Australia
Blackwell Scientific Publications
(Australia) Pty Ltd
54 University Street
Carlton, Victoria 3053
(*Orders*: Tel: (03) 347–0300)

British Library
Cataloguing in Publication Data

Connor, J.M. (James Michael)
Essential medical genetics. – 3rd ed.
1. Man. Genetics. Medical aspects
I. Title II. Ferguson-Smith, M.A.
616.042

ISBN 0–632–02872–6

# Preface

This book has been written for those to whom an understanding of modern medical genetics is important in their practice as clinicians, scientists, counsellors and teachers. It is based on the authors' personal experience in both clinical and laboratory aspects of busy regional genetics services over a period of thirty years. This period has seen the emergence of modern cytogenetics and molecular genetics alongside the development of medical genetics from a purely academic discipline into a clinical specialty of relevance to every branch of medicine. As in our teaching to undergraduates, we emphasize the central role of the chromosome and the human genome in understanding the molecular mechanisms involved in the pathogenesis of genetic disease. By genetic disease we include not only the classic Mendelian and chromosomal disorders but also the commoner disorders of adulthood with a genetic predisposition and somatic cell genetic disorders, such as cancer. The modern techniques of molecular biology are not only of use in the determination of genetic anatomy and pathology but are also increasingly being applied to unravelling physiological processes and for attempts at direct gene therapy.

We have followed the format of the previous two editions by retaining the division into basic principles and clinical applications as this approach fits best the preclinical and clinical courses of many medical schools. However, the impact of molecular genetics has been of such importance that it has found its way into every section of the book and is no longer treated separately. This has led to an extensive revision of the text, the incorporation of many new figures and new chapters on DNA diagnosis, the genetics of common disorders, and the genetics of cancer. Also, in response to a number of requests, our descriptions of the disorders most commonly encountered in medical genetics practice have been augmented by further information useful in counselling patients and their families. None the less, we have maintained our principle of passing on what is essential within modern medical genetics and look to the generosity of our readers to help correct our misconceptions and omissions.

The role of genetic counselling, prenatal diagnosis, carrier detection and other forms of genetic screening in the prevention of genetic disease is now well established and this is reflected in the increasing provision of genetic services throughout the world. It is hoped that our book will be useful to those in training for this important task.

J.M.C.
M.A.F.-S.

# Acknowledgements

We wish to thank many people who have influenced the production of this book. These include particularly:

Victor McKusick with whom we both spent the formative periods of our early training in medical genetics.

Our colleagues and students at the Duncan Guthrie Institute and at the Cambridge University Centre for Medical Genetics.

We are grateful for permission to reproduce the following figures:

Plate 1, Figs 8.3, and 13.2: Douglas Wilcox
Fig. 3.2: Avril Morris
Fig. 3.3: Rosemary Akhurst and David Fitzpatrick
Figs 3.4 and 9.10: Elizabeth Boyd and Norma Morrison
Figs 3.9, 3.12 and 3.13: Su Loughlin
Figs 3.11 and 3.15: George Lanyon
Figs 3.17, 19.6 and 20.8: Carolyn Williams, Edward Mayall and Bob Williamson
Fig. 3.18: Sandy Cooke
Figs 3.19 and 3.20: Agapi Kataki
Figs 4.2–4.5, 6.6, 6.8, 13.10, 13.13, 13.14 and 13.15(b): Elizabeth Boyd
Figs 4.8 and 9.12: Nigel Carter
Fig. 4.13: The Editor, *Birth Defects Original Article Series*
Fig. 4.14: The Editor, *Annales de Genetique*
Fig. 4.15: Peter Pearson
Figs 5.2, 5.3, 5.10 and 6.9: The Editor, *Excerpta Medica*
Figs 5.9 and 6.4(d): Anne Chandley
Fig. 6.2: James Galt
Figs 6.4(b) and 6.4(c): The Editor, *Journal of Medical Genetics*
Fig. 6.15 Maj Hulten and N. Saadallah
Figs 6.16 and 6.17: The Editor, *Cytogenetics and Cell Genetics*
Fig. 7.6: Brenda Gibson
Figs 7.10–7.12: Steve Humphries
Fig. 8.9: Susan Holder
Fig. 9.9: The Editor, *Karger*
Fig. 12.11: Paul Debenham (Cellmark Diagnostics)
Fig. 13.4: Blackwell Scientific Publications
Fig. 13.11: Ian Glass
Figs 13.16 and 14.5: John Tolmie
Fig. 13.17: Sue Malcolm
Fig. 14.6: Reed Pyeritz
Fig. 14.13: John Stephenson
Fig. 15.7: Linda Glasgow
Fig. 17.2: Michael Steel
Figs 17.3–17.6: Janet Stewart
Figs 18.5–18.7 and 18.9: Angus Gibson
Figs 18.10 and 19.9(a): Rachel Connor
Fig. 18.12: Robin Winter
Fig. 18.14: Heather May
Figs 19.1–19.2, 19.7 and 19.8: Marie Ferguson-Smith
Fig. 19.9(b): Margaret McNay
Figs 20.1 and 20.2: Nick Wald
Figs 20.3–20.7: David Aitken and Jenny Crossley

# Contents

# Basic Principles

# Chapter 1
# Human Genetics in Perspective

Human genetics is the scientific study of variation in humans, whereas medical genetics is concerned with the application of these principles to the practice of medicine. Although people have long been aware that individuals differ, that children tend to resemble their parents and that certain diseases tend to run in families, the scientific basis for these observations was only discovered during the past 150 years. The clinical application of this knowledge is even more recent, with most progress confined to the past 25 years.

### Mendel's contribution

Prior to Mendel, parental characteristics were believed to blend in the offspring. Whilst this was acceptable for continuous traits such as height or skin pigmentation, it was clearly difficult to account for the family patterns of discontinuous traits such as haemophilia or albinism. Mendel studied clearly defined pairs of contrasting characters in the offspring of the garden pea (*Pisum sativum*). These peas were, for example, either round or wrinkled and were either yellow or green. Pure bred strains for each of these characteristics were available but when cross bred (the first filial or $F_1$ progeny) were all round or yellow. If $F_1$ progeny were bred then each characteristic was re-observed in a ratio of 3 round to 1 wrinkled or 3 yellow to one green (in the second filial or $F_2$ progeny). Mendel concluded that inheritance of these characteristics must be particulate with pairs of hereditary elements (now called genes). In these two examples one characteristic (or trait) was dominant to the other (i.e. all $F_1$ showed it). The fact that both characteristics were observed in the $F_2$ progeny entailed segregation of each pair of genes with one member to one gamete and one to another gamete (Mendel's first law).

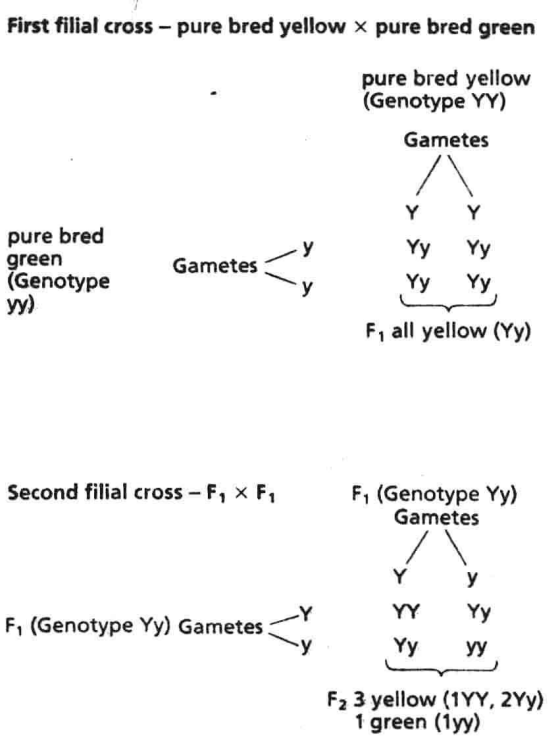


**Fig. 1.1** Example of Mendel's breeding experiments for a single trait (yellow or green seeds)

Figures 1.1 and 1.2 illustrate these experiments with upper case letters used for the dominant characteristic and lower case used for the masked (or recessive) characteristic. If the pair of genes are identical then this is termed homozygous (for the dominant or recessive trait) whereas a heterozygote has one of each type.

In his next series of experiments Mendel crossed pure bred strains with two characteristics e.g. pure bred round/yellow with pure bred wrinkled/green.

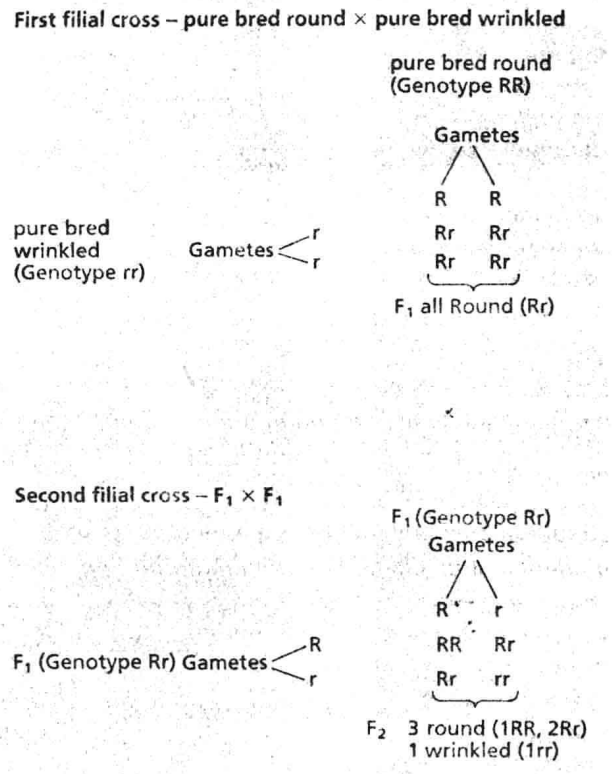


**Fig. 1.2** Example of Mendel's breeding experiments for a single trait (round or wrinkled seeds)

The $F_1$ showed only the two dominant characteristics — in this case round/yellow. The $F_2$ showed four combinations: the original two round/yellow and wrinkled/green in a ratio of 9:1 and two new combinations — wrinkled/yellow and round/green in a ratio of 3:3 (Fig 1.3).

In these experiments there was thus no tendency for the genes arising from one parent to stay together. In other words members of different gene pairs assort to gametes independently of one another (Mendel's second law).

Although Mendel presented and published his work in 1865 the significance of his discoveries was not realized until the early 1900s when three plant breeders, De Vries, Correns, and Tschermak confirmed his findings.

## Chromosomal basis of inheritance

In 1839 Schleiden and Schwann established the concept of cells as the fundamental living units. Hereditary transmission through the sperm and egg became known by 1860, and in 1868 Haeckel, noting that the sperm was largely nuclear material, postulated that the nucleus was responsible for heredity. Flemming identified chromosomes within the nucleus in 1877, and in 1903 Sutton and Boveri independently realized that the behaviour of chromosomes during the production of gametes paralleled the behaviour of Mendel's hereditary elements. Thus the chromosomes were discovered to carry the genes. However, at that time, although the chromosomes were known to consist of protein and nucleic acid, it was not clear which component was the hereditary material.

## Chemical basis of inheritance

Pneumococci are of two genetically distinct strains: rough or non-encapsulated (non-virulent) and smooth or encapsulated (virulent). Griffith in 1928 added heat-killed smooth bacteria to live rough and found that some of the rough pneumococci were transformed to the smooth, virulent type. Avery, MacLeod and McCarty repeated this experiment in 1944 and showed that nucleic acid was the transforming agent. Thus nucleic acid was shown to carry the hereditary information. This stimulated intense interest in the composition of nucleic acids which culminated in Watson and Crick's discovery of the double helical structure for deoxyribonucleic acid (DNA) in 1953.

## Chromosomal disorders

By 1890 it was known that one human chromosome (the X chromosome) did not always have a partner, and in 1905 Wilson and Stevens extended this observation by establishing the pattern of human sex chromosomes. At this time there were believed to be 48 chromosomes in each somatic cell. Tjio and Levan refuted this in 1956 when they showed the normal human chromosome number to be 46. In 1959 the first chromosomal disease in humans, trisomy 21, was discovered by Lejeune and colleagues and by 1970 over 20 different human chromosomal disorders were known. The development of chromosomal banding in 1970 markedly increased the ability to resolve small chromosomal aberrations, and so by 1989 more than 600 different chromosome abnormalities were known, in addition to many normal variants.

## Mitochondrial disorders

Mitochondria have their own chromosomes and these are inherited from a mother to all of her

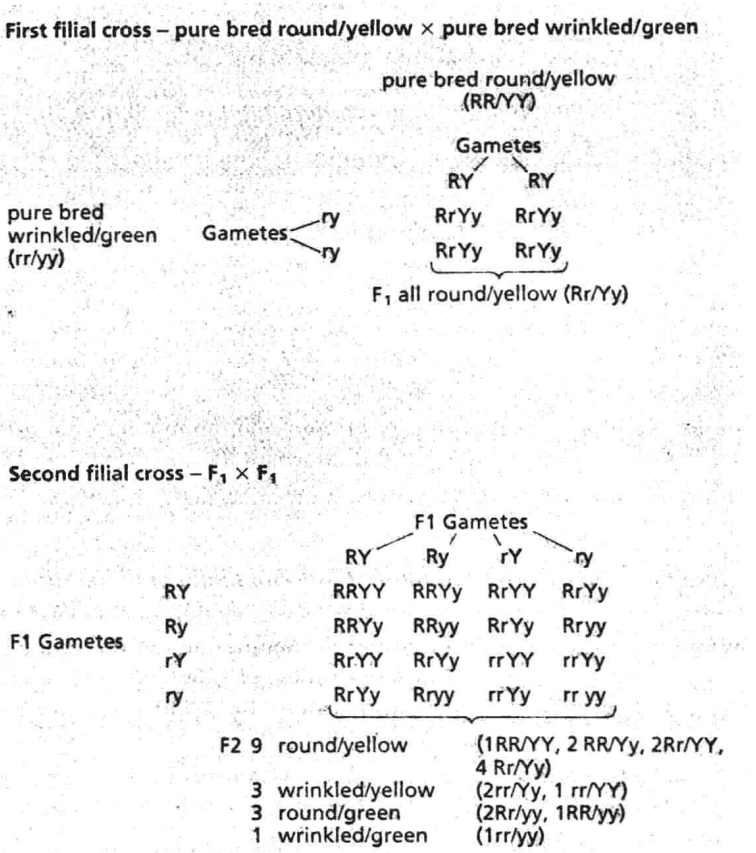


**Fig. 1.3** Example of Mendel's breeding experiments for two traits (yellow or green and round or wrinkled seeds)

children. Faults in these mitochondrial chromosomes can cause disease and this was first shown in 1988 (for a maternally inherited type of blindness — Leber optic neuropathy).

## Single-gene disorders

In 1902 Garrod presented his studies on alkaptonuria, a rare condition in which patients have arthritis and urine which darkens on standing. He found 3 of 11 sets of parents of affected patients to be blood relatives and, in collaboration with Bateson, proposed that this was a Mendelian recessive trait with affected persons homozygous for the underactive gene. This was the first disease to be interpreted as a single-gene trait. Garrod also conceived the idea that patients with alkaptonuria and other inborn errors of metabolism really represented one extreme of human biochemical variations and that other less clinically significant variations were to be expected.

There followed numerous descriptions of distinct human single-gene traits and at the present time more than 4300 human single-gene traits are known (Table 1.1). Pauling in 1949 suspected an abnormal haemoglobin to be the cause of sickle cell anaemia, and this was confirmed by Ingram in 1956 who found an altered haemoglobin polypeptide sequence. This was the first demonstration in any organism that a mutation in a structural gene could produce an altered amino acid sequence. In 1959 only two abnormal haemoglobins were known; now the number exceeds 450. In 1948, Gibson demonstrated the first enzyme defect in an autosomal recessive condition (NADH-dependent methaemoglobin reductase in methaemoglobinaemia). By 1959 five enzyme defects were known, but the number now exceeds 200. The polypeptide product is, however, still unknown in about 85% of human single-gene disorders. Study of these rare, and not so rare, single-gene disorders has provided valuable insight into normal physiological mechanisms, for example, our knowledge of the normal metabolic pathways has been largely derived from the study of inborn errors of metabolism.

Progress has also been made in the assignment of genes to individual chromosomes. Wilson identified the X-linked trait for colour blindness in 1911

**Table 1.1** Human single-gene traits. Growth of knowledge in 22 years (adapted from McKusick, 1988)

| | 1966 | 1971 | 1975 | 1978 | 1982 | 1986 | 1988 |
|---|---|---|---|---|---|---|---|
| Autosomal dominant | 269 (+568) | 415 (+528) | 583 (+635) | 736 (+753) | 934 (+893) | 1172 (+1029) | 1443 (+1114) |
| Autosomal recessive | 237 (+294) | 365 (+418) | 466 (+481) | 521 (+596) | 588 (+710) | 610 (+810) | 626 (+851) |
| X-linked | 68 (+51) | 86 (+64) | 93 (+78) | 107 (+98) | 115 (+128) | 124 (+162) | 139 (+171) |
| Total | 574 (+913) | 866 (+1010) | 1142 (+1194) | 1364 (+1447) | 1637 (+1731) | 1906 (+2001) | 2208 (+2136) |
| Grand total | 1487 | 1876 | 2336 | 2811 | 3368 | 3907 | 4344 |

Numbers in parentheses refer to loci not yet fully identified or confirmed.

and assigned the gene to the X chromosome so making the first human gene assignment. Other X-linked traits rapidly followed and the first autosomal gene to be assigned was thymidine kinase to chromosome 17 in 1967 followed by the Duffy blood group to chromosome 1 in 1968. More than 1446 autosomal genes and more than 171 X chromosomal genes have now been firmly assigned to individual chromosomes. This represents approximately 37% of the total number of Mendelian traits so far identified.

The ability to manipulate DNA *in vitro*, or genetic engineering, has a relatively brief but spectacular history. In 1970 the first sequence-specific restriction enzyme was discovered and the first gene (yeast alanine transfer ribonucleic acid) was synthesized *in vitro*. In 1972 the first recombinant DNA molecules were generated, and in 1977 the first human gene (human placental lactogen) was cloned. By 1989, 945 human structural genes had been cloned, in addition to over 3400 arbitrary DNA segments (one-half of which recognize common DNA sequence variations called restriction fragment length polymorphisms, RFLPs). In 1977 the first protein product (somatostatin) was made by genetic engineering, followed by insulin in 1979, and in 1982 the first product (a vaccine to prevent a diarrhoeal disease in pigs) was marketed.

## Multifactorial disorders

Galton studied continuous human characteristics such as intelligence and physique, which did not seem to conform to Mendel's laws of inheritance and an intense debate ensued, with the supporters of Mendel on the one hand and those of Galton on the other. Finally, a statistician, R.A. Fisher, reconciled the two sides by showing that such inheritance could be explained by multiple pairs of genes, each with a small but additive effect. Discontinuous traits with multifactorial inheritance, such as congenital malformations, were explained by introducing the concept of a threshold effect for the disorder: expression only occurred when the genetic contribution passed the threshold. Many human characteristics are determined in this fashion and usually factors in the environment interact with the genetic background.

Although the genetic contribution to multifactorial disorders is now well accepted the number and nature of the genes involved and their mechanisms of interaction among each other and environmental factors is largely unknown. This is now being investigated by recombinant DNA techniques and progress has been made in identifying the genetic contribution for several of these conditions.

## Somatic cell genetic disorders

Genetic mutations may only arise after birth and be confined to certain somatic cells and their descendents. These genetic faults are now known to play a key role in the causation of many common cancers and may also be involved in autoimmune disorders and the ageing process.

## Clinical applications

Genetically determined disease has become an increasingly important part of ill health in the community now that most infections can be controlled,

**Table 1.2** Important advances in human genetics

| Year | Landmark | Key figure(s) |
|---|---|---|
| 1839 | Cell theory | Schleiden and Schwann |
| 1859 | Theory of evolution | Darwin |
| 1865 | Particulate inheritance | Mendel |
| 1877 | Chromosomes observed | Flemming |
| 1900 | ABO blood groups discovered | Landsteiner |
| 1902 | Biochemical variation | Garrod |
| 1903 | Chromosomes carry genes | Sutton, Boveri |
| 1908 | Inheritance of ABO blood groups | Ottenburg and Epstein |
| 1910 | First US genetic clinic | Davenport |
| 1911 | Linkage in drosophila | Morgan |
| 1911 | First human gene assignment | Wilson |
| 1927 | Mutagenicity of X-rays | Muller |
| 1928 | Transfection | Griffith |
| 1940 | Concept of polymorphism | Ford |
| 1944 | Role of DNA | Avery |
| 1946 | First UK genetic clinic | Roberts |
| 1947 | Transposable elements | McClintock |
| 1949 | Sex chromatin | Barr |
| 1953 | DNA structure | Watson and Crick |
| 1956 | Amino acid sequence of HbS | Ingram |
| 1956 | 46 chromosomes in humans | Tjio and Levan |
| 1959 | First human chromosomal abnormality | Lejeune |
| 1960 | Prenatal sexing | Riis and Fuchs |
| 1960 | Chromosome analysis on blood | Moorhead |
| 1961 | Biochemical screening | Guthrie |
| 1961 | X inactivation | Lyon |
| 1961 | Genetic code | Nirenberg |
| 1964 | Antenatal ultrasound | Donald |
| 1966 | First prenatal chromosomal analysis | Breg and Steel |
| 1967 | First autosomal assignment | Weiss and Green |
| 1970 | Prevention of Rhesus isoimmunization | Clarke |
| 1970 | Chromosome banding | Caspersson |
| 1970 | Sequence specific restriction enzymes discovered | Nathan, Smith |
| 1970 | First gene synthesized *in vitro* | Khorana |
| 1972 | AFP screening | Brock |
| 1973 | HLA disease associations | Terasaki |
| 1977 | First human gene cloned | Shine |
| 1977 | Somatostatin made by genetic engineering | Itakura |
| 1978 | First restriction fragment length polymorphism | Kan |
| 1978 | First DNA diagnosis | Kan |
| 1979 | *In vitro* fertilization | Edwards and Steptoe |
| 1979 | Insulin produced by genetic engineering | Goeddel |
| 1982 | First product of genetic engineering marketed | |
| 1985 | DNA fingerprinting | Jeffreys |
| 1987 | Linkage map of human chromosomes developed | Many contributors |

and now that modern medical and nursing care can save many affected infants who previously would have succumbed shortly after birth. This has led to an increased demand for informed genetic counselling and for screening tests both for carrier detection and to try to identify pregnancies at risk.

Davenport began to give genetic advice as early as 1910 in the USA and the first British genetic counselling clinic was established in 1946 at Great Ormond Street, London. Public demand has since caused a proliferation of genetic counselling centres so that there are now more than 40 in the UK and more than 450 in the USA.

In addition to an accurate assessment of the risks in a family the clinical geneticist also needs to discuss reproductive options. Important advances in this respect have been made with regard to prenatal diagnosis with the option of selective

termination, and this has been a major factor in increasing the demand for genetic counselling. Prenatal diagnosis offers reassurance for couples at high risk of serious genetic disorders and allows many mothers who were previously deterred by the risk the possibility of having healthy children. Genetic amniocentesis was first attempted in 1966 and the first prenatally detected chromosome abnormality was Down syndrome in 1969. Chromosome analysis following amniocentesis is now a routine component of obstetric care and over 30 different abnormalities have been detected. Amniocentesis can also be used to detect biochemical alterations in inborn errors of metabolism. This was first used in 1968 for a pregnancy at risk of Lesch–Nyhan syndrome and has since been used for successful prenatal diagnosis in over 90 IEM. Prenatal diagnosis can also be performed by DNA analysis of fetal samples. This approach was first used in 1978 for a pregnancy at risk of sickle cell disease and has now been used for over 30 single-gene disorders; in many of these, such as cystic fibrosis and Duchenne muscular dystrophy, it has become the main method of prenatal diagnosis.

These prenatal tests which detect chromosomal, biochemical or DNA alterations cannot detect many of the major congenital malformations. Alternative approaches of fetal visualization have been necessary for these. Fetoscopy was first used but this has been largely replaced with the development of high resolution ultrasound scanning. Ultrasound was first used to make a diagnosis of fetal abnormality (anencephaly) in 1972 and since then over 280 different abnormalities have been detected.

The majority of couples are not aware that they are at risk until they have an affected child. This has led to an increased emphasis on prenatal screening, for example by measurement of maternal serum alphafetoprotein and other analytes to detect pregnancies at increased risk of neural tube defects and chromosomal abnormalities. Neonatal screening was introduced in 1961 for phenylketonuria and several other conditions where early diagnosis and therapy will permit normal development, and it is likely that in the future there will be continued development of population screening, prenatal, neonatal, and preconceptional. This should lead to a reduced frequency of many genetic diseases with consequent benefits for individual families and society in general.

## Further reading


Agarwal K.L., Büchs H., Caruthers M.H. *et al.* (1970) Total synthesis of the gene for an alanine transfer ribonucleic acid from yeast. *Nature* **227**, 27–34.

Avery O.T., MacLeod C.M. & McCarty M. (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *J. Exp. Med.* **79**, 137–58.

Boyer S.H. (ed.) (1963) *Papers on Human Genetics.* Prentice Hall, New Jersey.

Donahue R.P., Bias W.B., Renwick J.H. & McKusick V.A. (1968) Probable assignment of the Duffy blood group locus to chromosome 1 in man. *Proc. Natl Acad. Sci. USA* **6**, 949–55.

Garrod A.E. (1902) The incidence of alkaptonuria: a study in chemical individuality. *Lancet* **ii** 1616–20.

Ingram V.M. (1956) A specific chemical difference between the globins of normal human and sickle cell anaemia haemoglobin. *Nature* **178**, 792–4.

Kan Y.W. & Dozy A.M. (1978) Antenatal diagnosis of sickle cell anaemia by DNA analysis of amniotic fluid cells. *Lancet* **ii**, 910–12.

Lejeune J., Gautier M. & Turpin R. (1959) Etude des chromosomes somatiques de neuf enfants mongoliens. *C.R. Acad. Sci. Paris* **248**, 1721–2.

McKusick V.A. (1988) *Mendelian Inheritance in Man. Catalogs of Autosomal Dominant, Autosomal Recessive and X-linked phenotypes*, 8th Edition. The John Hopkins University Press, Baltimore.

Mendel G. (1865) Experiments in plant hybridization. Translated from the German by W. Bateson and reprinted in *Classic Papers in Genetics* (1959) (ed. J.A. Peters), pp. 1–20. Prentice Hall, New Jersey.

Steele M.W. & Breg W.R. (1966) Chromosome analysis of human amniotic fluid cells. *Lancet* **1**, 383–5.

Tjio H.J. & Levan A. (1956) The chromosome number of man. *Hereditas* **42**, 1–6.

Watson D. & Crick F.H.C. (1953) Molecular structure of nucleic acids — a structure for deoxyribose nucleic acid. *Nature* **171**, 737–8.

Wilson E.B. (1911) The sex chromosomes. *Arch. Hikrosk. Anat. Enwicklungsmech* **77**, 249–71.

# Chapter 2
# Nucleic Acid Structure and Function

## Nucleic acid structure

In humans, as in other organisms, nucleic acid is the carrier of genetic information and has a structure which is ideally suited to this function. There are two main types of nucleic acid, DNA (deoxyribonucleic acid) and RNA (ribonucleic aid), which each consist of a sugar–phosphate backbone with projecting nitrogenous bases (Fig. 2.1). The nitrogenous bases are of two types, purines and pyrimidines. In DNA there are two purine bases, adenine (A) and guanine (G), and two pyrimidine bases, thymine (T) and cytosine (C). RNA also contains adenine, guanine and cytosine, but has uracil (U) in place of thymine. In DNA the sugar is deoxyribose, whereas in RNA it is ribose (Fig. 2.2). The nitrogenous bases are attached to the 1′ (one prime) position of each sugar, and the phosphate links 3′ and 5′ hydroxyl groups. Each unit of purine or pyrimidine base together with the attached sugar and phosphate group is called a nucleotide.

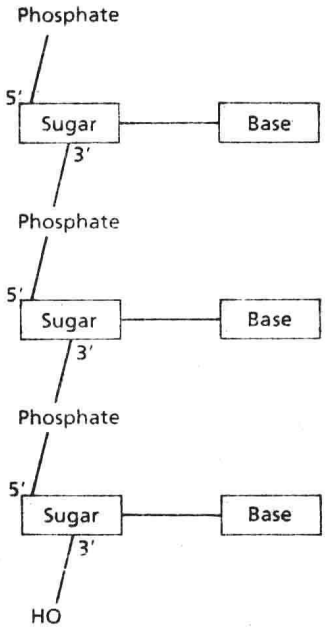


**Fig. 2.1** Diagram of nucleic acid structure. The 5′ phosphate end is at the top and the 3′ hydroxyl group is at the bottom of the molecule

A molecule of DNA is composed of two nucleotide chains which are coiled clockwise around one another to form a double helix with 10 nucleotides per complete turn of DNA (Fig. 2.3). The two chains run in opposite directions and are held together by hydrogen bonds between A in one chain and T in the other or between G and C. This base pairing is very specific, although rarely erroneous combinations may occur. Since A:T and G:C pairing is obligatory the parallel strands must be complementary to one another. Thus if one strand reads ATGA the complementary strand must read TACT. Hence the ratio of A to T is 1 to 1 and of G to C is likewise 1 to 1 (Chargaff's rule). Wide variation exists in the (A + T)/(G + C) ratio. Higher plants and animals tend to have an excess of (A + T) and in humans the ratio is 1.4 to 1.

The unit of length of DNA is the base pair (bp) with 1000 bp in a kilobase (kb) and 1 000 000 bp in a megabase (mb). The total length of DNA in a half set of human chromosomes is 3000 mb ($3 \times 10^9$ bp).

There are an estimated 50 000 human structural genes encoded in the DNA. Each structural gene usually has only one copy in the haploid genome, and if the average gene is 20 kilobases in size then this only accounts for 1000 megabases or about one-third of the total DNA. Much of the remainder consists of repetitive DNA, which may be moderately repetitive with several hundred copies or highly repetitive with many thousands of copies,

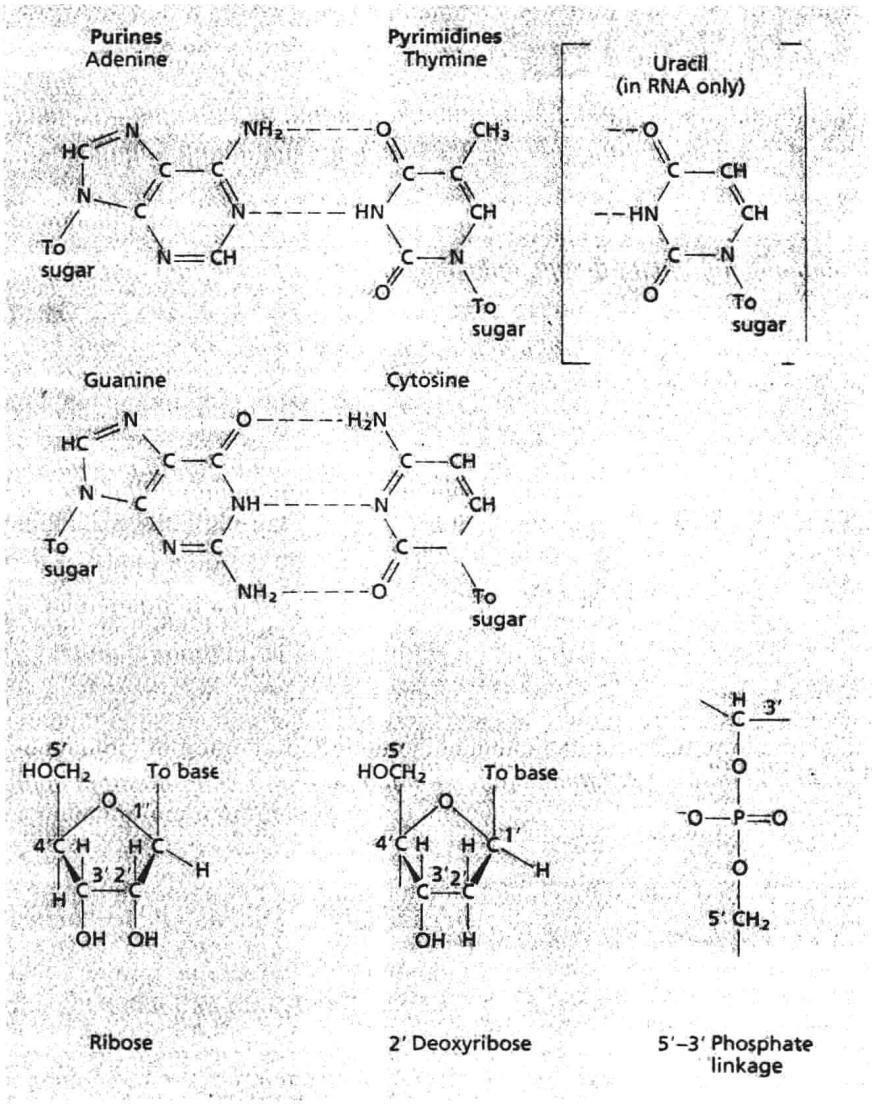


**Fig. 2.2** Chemical structure of purines, pyrimidines, ribose, deoxyribose and the 5′ to 3′ phosphate linkage. The hydrogen bonds between adenine and thymine (or uracil) and guanine and cytosine are indicated

and may be dispersed or occur in clusters. The moderately repetitive DNA includes some functional genes which occur as multiple copies, including ribosomal RNAs (about 2000 copies) and the histone genes. In contrast, the highly repetitive DNA (perhaps 10% of the total DNA) is not transcribed and may have a role in chromosomal pairing alignment and recombination. One family of highly repetitive DNA, called the Alu family, consists of about 500 000 copies of a 300 base pair sequence interspersed throughout the genome and accounts for about 4% of the total DNA. The alphoid family of highly repetitive DNA also accounts for 4% of total human DNA and occurs as short tandem repeats near the centromeres of all chromosomes and is especially abundant in chromosomes 1, 9, 16, and the Y chromosome.

RNA differs in structure from DNA in several respects:

1 The sugar is ribose rather than deoxyribose.
2 Uracil (U) replaces thymine.
3 RNA is single-stranded.
4 Only a single type of DNA is known in humans but five main types of RNA are found (Table 2.1).

Ribosomal RNAs are synthesized directly on a DNA template which occur as multiple clustered copies (about 200 copies for 18S and 28S ribosomal RNAs on the short arms of chromosomes 13, 14, 15, 21 and 22 and about 2000 copies of 5S ribosomal RNA on chromosome 1). The ribosomal RNAs are synthesized as large precursors in the nucleolus and then enzymically cleaved.

Transfer RNAs are also synthesized directly on a DNA template and although 61 different types may

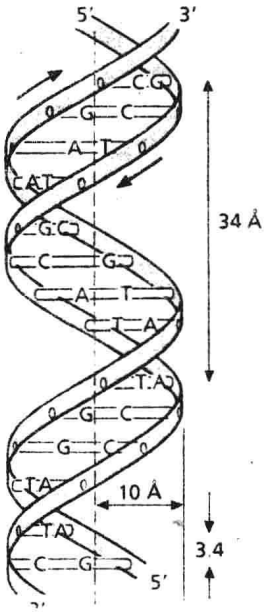


**Fig. 2.3** Diagram of DNA double helix

be expected (Table 2.2) only 40 are found as some tRNAs can bind to more than one codon. The DNA templates for tRNAs tend to occur as multiple copies which may be clustered or dispersed.

## Nucleic acid function

Nucleic acids have two major functions: the direction of all protein synthesis and the accurate transmission of this information from one generation to the next.

Proteins, whether structural components, enzymes, carrier molecules, hormones or receptors, are all composed of a series of amino acids. Twenty amino acids are known, and the sequence of these determines the form and function of the resulting protein. All proteins are encoded in DNA, and the unit of DNA which codes for a protein is, by definition, its gene. Genes vary greatly in size from small genes like the globins to medium sized genes of 15–45 kb to enormous genes such as dystrophin (Table 2.3).

Each set of three DNA base pairs or triplet codes for an amino acid. As each base in the triplet may be any of the four types of nucleotide (A, G, C, T) this results in $4^3$ or 64 possible combinations or codons. The codons for each amino acid are given in Table 2.2, and it is important to note that, by convention, each codon is shown in terms of the messenger RNA, and so the corresponding DNA codon will be complementary.

All amino acids except methionine and tryptophan are coded by more than one codon: hence the code is said to be degenerate. Three of the 64 codons designate the termination of a message and these are called chain terminators (UAA, UGA, UAG), and one codon, AUG (methionine), acts as a start signal for protein synthesis. With a few possible exceptions this code is identical in all species.

The first stage in protein synthesis is transcription. The two strands of DNA separate in the area of the gene to be transcribed. One strand (the *sense* strand — this strand is consistent for a given gene but varies from one gene to another) functions as a template, and messenger RNA (mRNA) is formed with a complementary sequence under the influence of the enzyme RNA polymerase (Fig. 2.4). Transcription proceeds in a 5′ to 3′ direction until the transcription terminator is reached. After some processing and modification the mRNA molecule

**Table 2.1** Types of RNA

| Type | Location | Comments |
|---|---|---|
| Messenger RNA (mRNA) | Nucleus and cytoplasm | Variable size, base sequence complementary to transcribed DNA, about 4% of total cellular RNA, half-life 7–24 hours |
| Transfer RNA (tRNA) | Cytoplasm | Hairpin-loop shape, about 40 types, amino acid-specific, 73–93 nucleotides in each, about 10% of total cellular RNA with tens to hundreds of copies of the genes for each tRNA species |
| Ribosomal RNA (rRNA) | Ribosomes and nucleoli | About 80% of total cellular RNA, synthesized and stored in nucleolus |
| Heterogeneous RNA (hnRNA) | Nucleus | High molecular weight mRNA precursors |
| Small nuclear RNA (snRNA) | Nucleus | Six types (U1–U6) 100–215 nucleotides in each, involved in RNA splicing |