

PREFACE

In the auricular structure introduced by this University for students of Post-Graduate degree programme, the opportunity to pursue Post-Graduate course in Subject introduced by this University is equally available to all learners. Instead of being guided by any presumption about ability level, it would perhaps stand to reason if receptivity of a learner is judged in the course of the learning process. That would be entirely in keeping with the objectives of open education which does not believe in artificial differentiation.

Keeping this in view, study materials of the Post-Graduate level in different subjects are being prepared on the basis of a well laid-out syllabus. The course structure combines the best elements in the approved syllabi of Central and State Universities in respective subjects. It has been so designed as to be upgradable with the addition of new information as well as results of fresh thinking and analysis.

The accepted methodology of distance education has been followed in the preparation of these study materials. Co-operation in every form of experienced scholars is indispensable for a work of this kind. We, therefore, owe an enormous debt of gratitude to everyone whose tireless efforts went into the writing, editing and devising of a proper lay-out of the materials. Practically speaking, their role amounts to an involvement in invisible teaching. For, whoever makes use of these study materials would virtually derive the benefit of learning under their collective care without each being seen by the other.

The more a learner would seriously pursue these study materials the easier it will be for him or her to reach out to larger horizons of a subject. Care has also been taken to make the language lucid and presentation attractive so that they may be rated as quality self-learning materials. If anything remains still obscure or difficult to follow, arrangements are there to come to terms with them through the counselling sessions regularly available at the network of study centres set up by the University.

Needless to add, a great deal of these efforts is still experimental—in fact, pioneering in certain areas. Naturally, there is every possibility of some lapse or deficiency here and there. However, these do admit of rectification and further improvement in due course. On the whole, therefore, these study materials are expected to evoke wider appreciation the more they receive serious attention of all concerned.

Professor (Dr.) Manimala Das
Vice-Chancellor

Third Reprint : June, 2015

Printed in accordance with the regulations and financial assistance of the
Distance Education Bureau of the University Grants Commission.

POST GRADUATE : ZOOLOGY
[M.Sc]

PAPER : GROUP
PGZO - 3 : B

	Writer	Editor
Part-I		
Units 1-4	☐ Dr. Kamalesh Mishra	
Units 5-8	☐ Prof. Chandrasekhar Chakraborty	Prof. Chandrasekhar Chakraborty
Units 9-12	☐ Prof. Anil Kumar Saha	
Part-II		
Units 1-8	☐ Dr. Sanjib Kumar Das	Dr. Subir Chandra Dasgupta

Notification

All rights reserved. No part of this book may be reproduced in any form without permission in writing from Netaji Subhas Open University.

Professor (Dr.) Debesh Roy
Registrar



GROUP B(I)

Part-I : Cytogenetics

Unit 1	□ Biology of Chromosomes	9-33
Unit 2	□ Sex Chromosomes, Sex Determination and Dosage Compensation	34-48
Unit 3	□ Imprinting of Genes, Chromosomes and Genomes	49-53
Unit 4	□ Somatic Cell Genetics	54-61
Unit 5	□ Human Cytogenetics	62-111
Unit 6	□ Cytogenetic Implications and Consequences of Structural Changes and Numerical Alterations of Chromosomes	112-117
Unit 7	□ Microbial Genetics	118-141
Unit 8	□ Cytogenetic Effects of Ionizing and Non-ionizing Radiations	142-143
Unit 9	□ Molecular Cytogenetic Techniques	144-152
Unit 10	□ Genome Analysis	153-161
Unit 11	□ Linkage Map, Cytogenetic Mapping	162-169
Unit 12	□ Genetics of Cell Cycle	170-181

**GROUP
B(II)
Molecular Biology**

Unit 1	□ History and Scope of Molecular Biology	185 - 188
Unit 2	□ DNA Replication	189-218
Unit 3	□ Prokaryotic Transcription	219 - 244
Unit 4	□ Post Transcriptional Modification of RNA	245 - 263
Unit 5	□ Translation	264 - 286
Unit 6	□ Antisense and Riobozyme Technology	287 - 303
Unit 7	□ Recombination and Repair	304 - 320
Unit 8	□ Molecular Mapping of Genome	321 - 339

**GROUP
B (I)**

Part-I : Cytogenetics

Unit 1 □ Biology of Chromosomes

Structure

- 1.1 Molecular anatomy of eukaryotic chromosomes
- 1.2 Metaphase chromosome : Centromere, Kinetochore, Telomere and its maintenance
- 1.3 Heterochromatin and Euchromatin
- 1.4 Giant chromosomes : Polytene and Lampbrush chromosomes
- 1.5 Suggested questions

1.1 Molecular anatomy of eukaryotic chromosomes

The proteins that bind to the DNA to form eukaryotic chromosomes are traditionally divided into two general classes: *histones and the nonhistone chromosomal proteins*. The complex of both classes of protein with the nuclear DNA of eukaryotic cells is known as chromatin. The total mass of histones in chromatin is about equal to that of the DNA. Histones are responsible for the first and most basic level of *chromosome* organization, the nucleosome, which was discovered in 1974. At interphase nuclei most of the chromatin is in the form of a fibre with a diameter of about 30 nm. If this chromatin is subjected to treatments that cause it to unfold partially, it can be seen under the electron microscope as a series of “beads on a string”. The string is DNA, and each bead is a “nucleosome core particle” that consists of DNA wound around a protein core formed from histones. The beads on a string represent the first level of chromosomal DNA packing.

The structural organization of nucleosomes was determined after first isolating them from unfolded chromatin by digestion with particular enzymes (called nucleases) that break down DNA by cutting between the nucleosomes. After digestion for a short period, the exposed DNA between the nucleosome core particles, the linker DNA, is degraded. Each individual nucleosome core particle consists of a complex of eight histone proteins—two molecules each of histones H2A, H2B, H3, and H4—and double-stranded DNA that are 146 nucleotide pairs long. The histone octamer forms a protein core around which the double-stranded DNA is wound (Fig. 1.1).

Each nucleosome core particle is separated from the next by a region of linker DNA, which can vary in length from a few nucleotide pairs up to about 80. (The term nucleosome technically refers to a nucleosome core particle plus

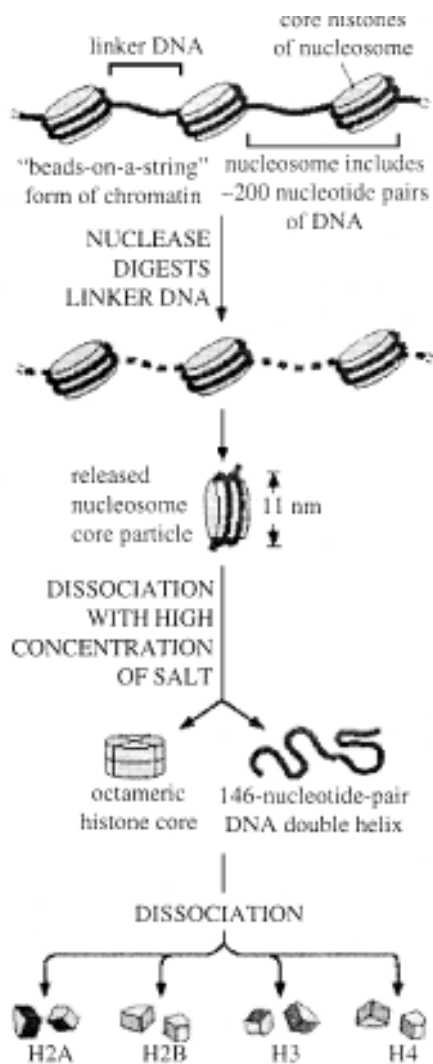


Fig. 1.1 Structural organization of the nucleosome A nucleosome contains a protein core made of eight histone molecules. As indicated, the nucleosome core particle is released from chromatin by digestion of the linker DNA with a nuclease, an enzyme that breaks down DNA. (The nuclease can degrade the exposed linker DNA but cannot attack the DNA wound tightly around the nucleosome core.) After dissociation of the isolated nucleosome into its protein core and DNA, the length of the DNA that was wound around the core can be determined. This length of 146 nucleotide pairs is sufficient to wrap 1.65 times around the histone core

one of its adjacent DNA linkers, but it is often used synonymously with nucleosome core particle.) On average, therefore, nucleosomes repeat at intervals of about 200 nucleotide pairs.

The structure of the nucleosome core particle reveals how DNA is packaged

The high-resolution structure of a nucleosome core particle, solved in 1997, revealed a disc-shaped histone core around which the DNA was tightly wrapped 1.65 turns in a left-handed coil. All four of the histones that make up the core of the nucleosome are relatively small proteins (102-135 amino acids), and they share a structural motif, known as the *histone fold*, formed from three α helices connected by two loops (Figure 1.2).

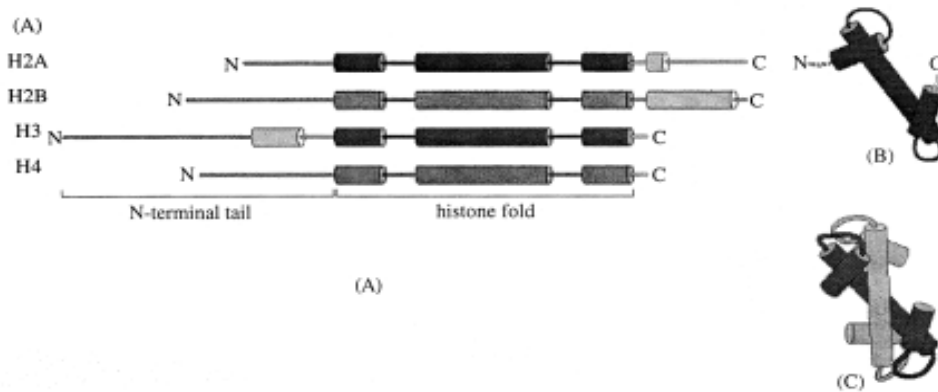


Fig. 1.2 The overall structural organization of the core histones, (A) Each of the core histones contains an N-terminal **tail**, which is subject to several forms of covalent modification, and a histone fold region, as indicated. (B) The structure of the histone fold, which is formed by all four of the core histones. (C) Histones 2A and 2B form a dimer through an interaction known as the “handshake.” Histones H3 and H4 form a dimer through the same type of interaction, as illustrated in Figure 1.3

In assembling a nucleosome, the histone folds first bind to each other to form H3-H4 and H2A-H2B dimers, and the H3-H4 dimers combine to form tetramers. An H3-H4 tetramer then further combines with two H2A-H2B dimers to form the compact octamer core, around which the DNA is wound (Fig. 1.3).

The interface between DNA and histone is extensive ; 142 hydrogen bonds are formed between DNA and the histone core in each nucleosome. Nearly half of these bonds form between the amino acid backbone of the histones and the phosphodiester backbone of the DNA. Numerous hydrophobic interactions and salt linkages also hold DNA and protein together in the nucleosome. These numerous interactions explain in part why DNA of virtually any sequence can be bound on a histone octamer core. The path of the DNA around the histone core is not smooth; rather, several kinks are seen in the DNA, as expected from the nonuniform surface of the core. In addition to its histone fold, each of the core histones has a long N-terminal amino acid “tail”, which extends out from the DNA-histone core (see Figure 1.3). These histone tails are subject to several different types of covalent modifications, which control many aspects of chromatin structure. The histones are among the most highly conserved eukaryotic proteins. This strong evolutionary conservation suggests that the functions of histones involve nearly all of their amino acids, so that a change in any position is deleterious to the cell. Despite the high conservation of the core histones, many

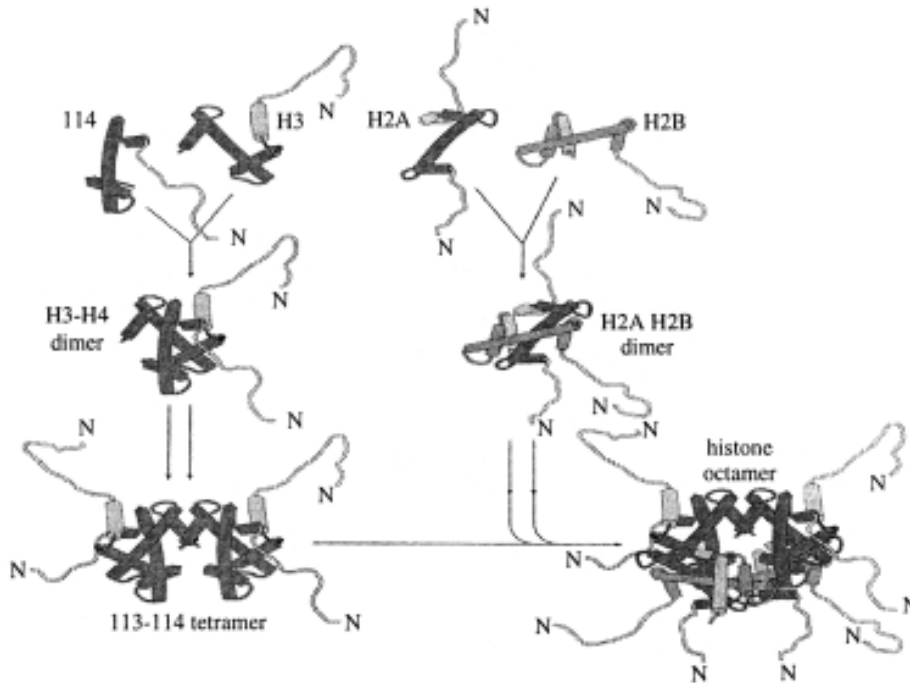


Fig. 1.3 The assembly of a histone octamer. The histone H3-H4 dimer and the H2A-H2B dimer are formed from the handshake interaction. An H3-H4 tetramer forms the scaffold of the octamer onto which two H2A-H2B dimers are added, to complete the assembly. Note that all eight N-terminal tails of the histones protrude from the disc-shaped core structure. In the x-ray crystal, most of the histone tails were unstructured (and therefore not visible in the structure), suggesting that their conformations are highly flexible. (Adapted from figures by J. Waterborg.)

eukaryotic organisms also produce specialized variant core histones that differ in amino acid sequence from the main ones. It is thought that nucleosomes that have incorporated these variant histones differ in stability from regular nucleosomes, and they may be particularly well suited for the high rates of DNA transcription and DNA replication that occur during these early stages of development.

The positioning of nucleosome on DNA is determined by both DNA flexibility and other DNA-bound proteins

Two main influences determine where nucleosomes form in the DNA. One is the difficulty of bending the DNA double helix into two tight turns around the outside of the histone octamer, a process that requires substantial compression of the minor groove of the DNA helix. Because A-T-rich sequences in the minor groove are easier to compress than G-C-rich sequences, each histone octamer tends to position itself on the DNA so as to maximize A-T-rich minor grooves on the inside of the DNA coil. Thus, a segment of DNA that contains short

A-T-rich sequences spaced by an integral number of DNA turns is easier to bend around the nucleosome than a segment of DNA lacking this feature. In addition, because the DNA in a nucleosome is kinked in several places, the ability of a given nucleotide sequence to accommodate this deformation can also influence the position of DNA on the nucleosome.

The second, and probably most important, influence on nucleosome positioning is the presence of other tightly bound proteins on the DNA. Some bound proteins favour the formation of a nucleosome adjacent to them. Others create obstacles that force the nucleosomes to assemble at positions between them. Finally, some proteins can bind tightly to DNA even when their DNA-binding site is part of a nucleosome. The exact positions of nucleosomes along a stretch of DNA therefore depend on factors that include the DNA sequence and the presence and nature of other proteins bound to the DNA.

Nucleosomes are usually packed together into a compact chromatin fibre

The nucleosomes are packed on top of one another, generating regular arrays in which the DNA is even more highly condensed. Thus, when nuclei are very gently lysed onto an electron microscope grid, most of the chromatin is seen to be in the form of a fiber with a diameter of about 30 nm, which is considerably wider than chromatin in the “beads on a string” form. Several models have been proposed to explain how nucleosomes are packed in the 30-nm chromatin fiber; the one most consistent with the available data is a series of structural variations known collectively as the zigzag model. In reality, the 30-nm structure found in chromosomes is probably a fluid mosaic of the different zigzag variations.

Several mechanisms probably act together to form the 30-nm fiber from a linear string of nucleosomes. First, an additional histone, called histone H1, is involved in this process. H1 is larger than the core histones and is considerably less well conserved. A single histone H1 molecule binds to each nucleosome, contacting both DNA and protein, and changing the path of the DNA as it exits from the nucleosome. Although it is not understood in detail how H1 pulls nucleosomes together into the 30-nm fiber, a change in the exit path in DNA seems crucial for compacting nucleosomal DNA so that it interlocks to form the 30-nm fibre (Fig. 1.4).

A second mechanism for forming the 30-nm fiber probably involves the tails of the core histones, which, as we saw above, extend from the nucleosome. It is thought that these tails may help attach one nucleosome to another—thereby allowing a string of them, with the aid of histone H1, to condense into the 30-nm fibre (Fig. 1.5).

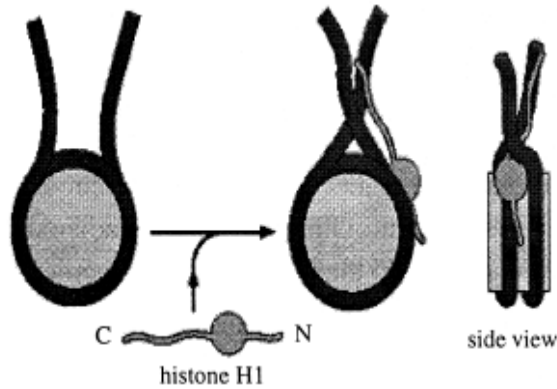


Fig. 1.4 A speculative model for how histone H1 could change the path of DNA as it exits from the nucleosome. Histone H1 (green) consists of a globular core and two extended tails. Part of the effect of H1 on the compaction of nucleosome organization may result from charge neutralization: like the core histones, H1 is positively charged (especially its C-terminal tail), and this helps to compact the negatively charged DNA. Unlike the core histones, H1 does not seem to be essential for cell viability; in one ciliated protozoan the nucleus expands nearly two fold in the absence of H1, but the cells otherwise appear normal

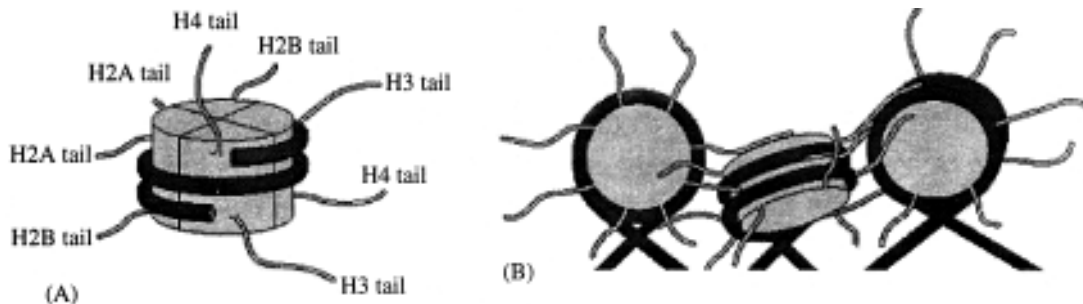


Fig. 1.5 A speculative model for histone tails in the formation of the 30-nm fiber. (A) The approximate exit points of the eight histone tails, four from each histone subunit, that extend from each nucleosome. In the high-resolution structure of the nucleosome, the tails are largely unstructured, suggesting that they are highly flexible. (B) A speculative model showing how the histone tails may help to pack nucleosomes together into the 30-nm fiber. This model is based on (1) experimental evidence that histone tails aid in the formation of the 30-nm fiber, (2) the x-ray crystal structure of the nucleosome, which showed that the tails of one nucleosome contact the histone core of an adjacent nucleosome in the crystal lattice, and (3) evidence that the histone tails interact with DNA

ATP-driven chromatin remodeling machines change nucleosome structure

Eukaryotic cells contain chromatin remodeling complexes, protein machines that use the energy of ATP hydrolysis to change the structure of nucleosomes temporarily so that DNA becomes less tightly bound to the histone core. The remodeled state may result from movement of the H2A-H2B dimers in the nucleosome core; the H3-H4 tetramer is particularly stable and would be difficult to rearrange (Fig. 1.3).

The remodeling of nucleosome structure has two important consequences. First, it permits ready access to nucleosomal DNA by other proteins in the cell, particularly those involved in gene expression, DNA replication, and repair. Even after the remodeling complex has dissociated, the nucleosome can remain

in a “remodeled state” that contains DNA and the full complement of histones—but one in which the DNA-histone contacts have been loosened; only gradually does this remodeled state revert to that of a standard nucleosome. Second, remodeling complexes can catalyze changes in the positions of nucleosomes along DNA (Fig. 1.6); some can even transfer a histone core from one DNA molecule to another.

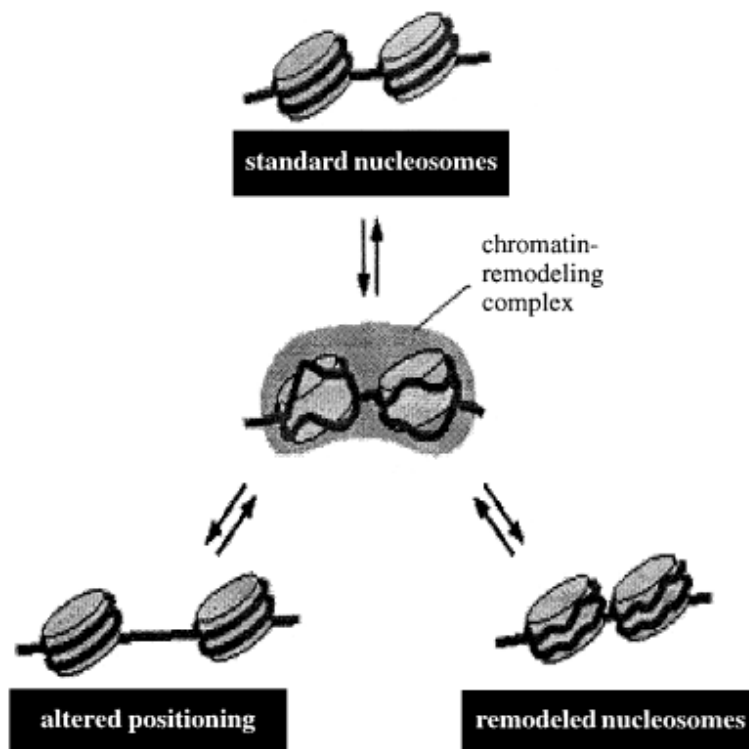


Fig. 1.6 Model for the mechanism of some chromatin remodeling complexes. In the absence of remodeling complexes, the interconversion between the three nucleosomal states shown is very slow because of a high activation energy barrier. Using ATP hydrolysis, chromatin-remodeling complexes (green) create an activated intermediate (shown in the center of the figure) in which the histone-DNA contacts have been partly disrupted. This activated state can then decay to any one of the three nucleosomal configurations shown. In this way, the remodeling complexes greatly increase the rate of interconversion between different nucleosomal states. The remodeled state, in which the histone-DNA contacts have been loosened, has a higher free energy level than that of standard nucleosomes and will slowly revert to the standard nucleosome conformation, even in the absence of a remodeling complex. Cells have many different chromatin remodeling complexes, and they differ in their detailed biochemical properties; for example, not all can change the position of a nucleosome, but all use the energy of ATP hydrolysis to alter nucleosome structure. (Adapted from R.E. Kingston and G.J. Narlikar, *Genes Dev.* 13:2339-2352, 1999.)

Cells have several different chromatin remodeling complexes that differ subtly in their properties. Most are large protein complexes that can contain more than ten subunits. It is likely that they are used whenever a eucaryotic cell needs direct access to nucleosome DNA for gene expression, DNA replication, or DNA repair. Different remodeling complexes may have features specialized for each of these roles. It is thought that the primary role of some remodeling complexes is to allow access to nucleosomal DNA, whereas that of others is to re-form nucleosomes when access to DNA is no longer required (Fig. 1.7).

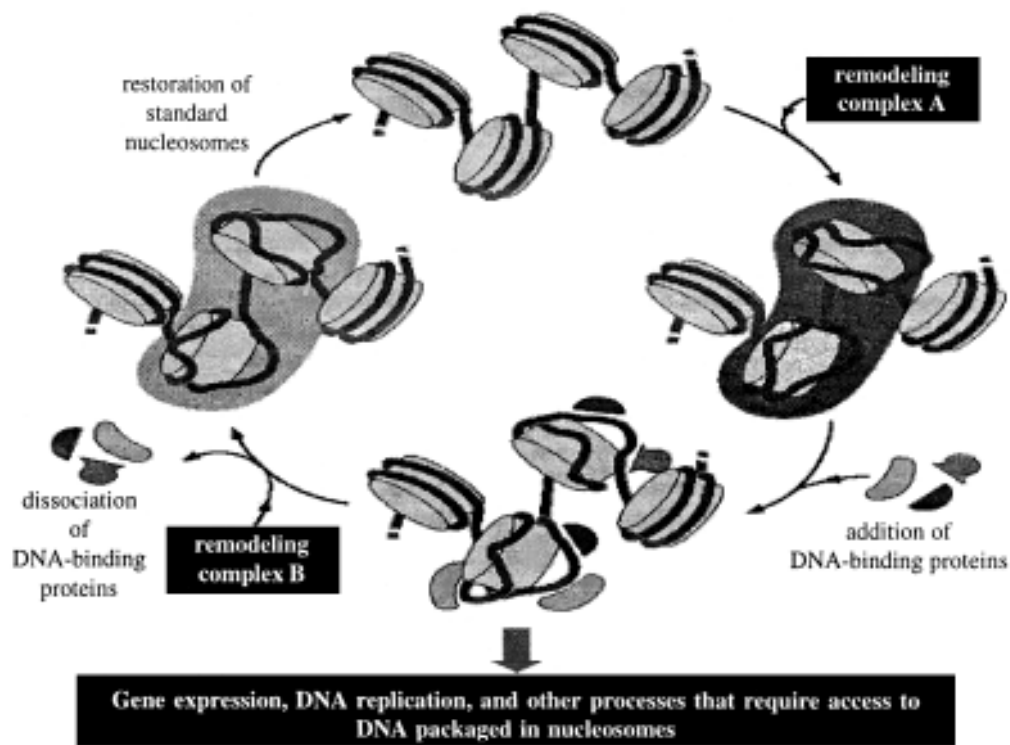


Fig. 1.7 A cyclic mechanism for nucleosome disruption and re-formation. According to this model, different chromatin remodeling complexes disrupt and re-form nucleosomes, although, in principle, the same complex might catalyze both reactions. The DNA-binding proteins could function in gene expression, DNA replication, or DNA repair, and in some cases their binding could lead to the dissociation of the histone core to form nucleosome-free regions of DNA, (Adapted from A. Travers, *Cell* 96:311-314, 1999.)

Covalent modification of the histone tails can profoundly affect chromatin

The N-terminal tails of each of the four core histones are highly conserved in their sequence, and perform crucial functions in regulating chromatin structure.

Each tail is subject to several types of covalent modifications, including acetylation of lysine: methylation of lysines, and phosphorylation of serines (Fig. 1.8).

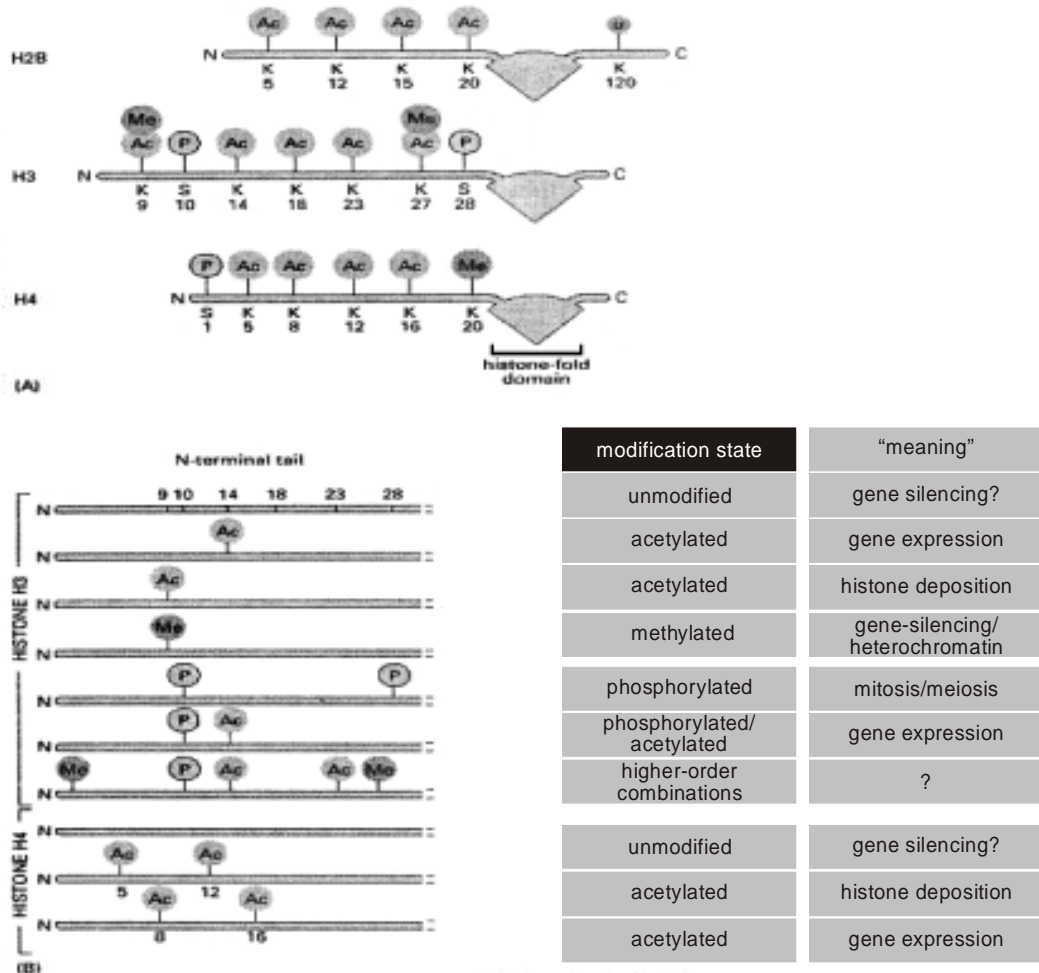


Fig. 1.8 Covalent modification of core histone tails

Histones are synthesized in the cytosol and then assembled into nucleosomes. Some of the modifications of histone tails occur just after their synthesis, but before their assembly. The modifications that concern us, however, take place once the nucleosome has been assembled. These nucleosome modifications are added and removed by enzymes that reside in the nucleus; for example, acetyl groups are added to the histone tails by histone acetyl transferases (HATs) and taken off by histone deacetylases (HDACs). The various modifications of the histone tails have several important consequences, histone acetylation tends to destabilize chromatin structure, perhaps in part because adding an

acetyl group removes the positive charge from the lysine, thereby making it more difficult for histones to neutralize the charges on DNA as chromatin is compacted. However, the most profound effect of modified histone tails is their ability to attract specific proteins to a stretch of chromatin that has been appropriately modified.

The enzymes that modify (and remove modifications from) histone tails are usually multisubunit proteins, and they are tightly regulated. They are brought to a particular region of chromatin by other cues, particularly by sequence-specific DNA-binding proteins. It is likely that histone-modifying enzymes and chromatin remodeling complexes work in concert to condense and recondense stretches of chromatin; for example, evidence suggests that a particular modification of the histone tail attracts a particular type of remodeling complex. Moreover, some chromatin remodeling complexes contain histone modification enzymes as subunits, directly connecting the two processes.

1.2 Metaphase chromosome : Centromere, Kinetochore, Telomere and its maintenance

1.2.1 Centromere

The region of the chromosome that is responsible for its segregation at mitosis and meiosis is called the **centromere**. It is associated with two important features :

- It contains the site at which the sister chromatids are held together prior to the separation of the individual chromosomes.
- The term “centromere” historically; has been used in both the functional and structural sense to describe the feature of the chromosome responsible for its movement.

The centromere is essential for segregation, as shown by the behavior of chromosomes that have been broken. **Acentric fragment** does not become attached to the mitotic spindle. There can be only one centromere per chromosome. In some species the centromeres are “diffuse”, which creates a different situation. Only discrete centromeres have been analyzed at the molecular level,

The regions flanking the centromere often are rich in satellite DNA sequences and contain a considerable amount of constitutive heterochromatin.

1.2.2 Kinetochore

Within the centromeric region, a darkly staining fibrous object of diameter or length ~400 nm can be seen. This object is called as Kinetochore. This

Kinetochores appear to be directly attached to the microtubules. The Kinetochores provide the MTOC on a chromosome.

Genetic engineering has produced plasmids of yeast that are replicated like chromosomal sequences. However, they are unstable at mitosis and meiosis, segregate erratically. Fragments of chromosomal DNA have been isolated by virtue of their ability to confer mitotic stability on these plasmids.

A *CEN* fragment is defined by its ability to confer stability upon such a plasmid. A *CEN* fragment derived from one chromosome can replace the centromere of another chromosome with no apparent consequence. This suggests that centromeres are interchangeable. *They are used simply to attach the chromosome to the spindle, and play no role in distinguishing one chromosome from another.*

The sequences required for centromeric function fall within a stretch of ~120 bp. The centromeric region is packaged into a nuclease-resistant structure, and it binds a single microtubule. The *S. cerevisiae* centromeric region has three types of sequence element that may be distinguished in the *CEN* region.

- CDE-I is a sequence of 9 bp that is conserved with minor variations at the left boundary of all centromeres.
- CDE-II is a >90% A-T-rich sequence of 80-90 bp found in all centromeres; its function could depend on its length rather than exact sequence. Its base composition may cause some characteristic distortions of the DNA double helical structure.
- CDE-III is an 11 bp sequence highly conserved at the right boundary of all centromeres. Sequences on either side of the element are less well conserved, and may also be needed for centromeric function.

A 240 kD complex of three proteins, called Cbf-III, binds to CDE-III. Mutations in the components of the genes coding for Cbf-III block chromosome movement at mitosis. A protein complex with motor activity connects the centromeric region of a chromosome to microtubules and contributes to movement on the mitotic spindle. The yeast *S. pombe* have the centromeres within regions of 40-100 kb that consist largely or entirely of repetitive DNA. The significance of the difference between the short centromeric regions in *S. cerevisiae* and the long regions in *S. pombe* is not clear. The common feature is that the DNA consists of noncoding sequences that are repetitive. Attempts to localize centromeric functions in *Drosophila* chromosomes suggest that they are dispersed in a large region, consisting of 200-600 kb. The large size of this type of centromere suggests that it is likely to contain several separate specialized functions, including sequences required for Kinetochores assembly, sister chromatid pairing, etc.

The primary modification comprising the constitutive heterochromatin of primate centromeres is a satellite DNA, which consists of tandem arrays of a 170 bp repeating unit. There is significant variation between individual repeats, although those at any centromere tend to be better related to one another than to members of the family in other locations. It is not clear whether the satellite sequences themselves provide this function, or whether other sequences are embedded within the satellite arrays.

1.2.3 Telomere

Essential feature in all chromosomes is the **telomere**. This “seals” the end. Telomere must be a special structure, because chromosome ends generated by breakage are “sticky” and tend to react with other chromosomes, whereas natural ends are stable. Two criteria in identifying a telomeric sequence :

- It must lie at the end of a chromosome (or, at least, at the end of an authentic linear DNA molecule).
- It must confer stability on a linear molecule.

Several telomeric sequences have been obtained from genomes of lower eukaryotes. In plant and man the construction of the telomere seems to follow a universal principle. Each telomere consists of a long series of short, tandemly repeated sequences. Table 1.1 lists the repeating units that have been identified at the ends of the linear DNA molecules. All can be written in the general form $C_n (A/T)_m$, where $n > 1$ and m is 1-4.

Within the telomeric region is a specific array of discontinuities, taking the form of single-strand breaks whose structure prevents them from being sealed by the ligase enzyme. They may be organized in a hairpin so that they are not recognized by nucleases.

Table 1.1 Telomeres have a common type of short tandem repeat. The repeating unit gives the sequence of one strand, in the 5'-3' direction from the telomere toward the centromere

Telomere	Repeating unit
Ciliate (<i>Tetrahymena</i>) macronucleus	CCCCAA
Ciliate (<i>Oxytricha</i>) macronucleus	CCCCAAAA
Trypanosoma minichromosome	CCCTA
Slime molds (<i>Dictyostelium</i>) rDNA	CCCTA
Yeast (<i>Saccharomyces</i>) chromosome	$C_{2-3} A(CA)_{1-3}$
Plant (<i>Arabidopsis</i>) chromosome	$C_3 TA_3$
Human chromosome	$C_3 TA_2$

Addition of telomeric repeats to the end of the chromosome in every replication cycle could solve the problem of end replicating. The addition of repeats by de novo synthesis would counteract the loss of repeats resulting from failure to replicate up to the end of the chromosome. Extension and shortening would be in dynamic equilibrium.

The overall length of the telomere is under genetic control. If telomeres are continually being lengthened (and shortened), their exact sequence may be irrelevant. All that is required is for the end to be recognized as a suitable substrate for addition. This explains how the ciliate *telomere* functions in yeast. Extracts of *Tetrahymena* contain an enzyme, called telomerase, that uses the 3'-OH of the G+T telomeric strand as a primer for synthesis of tandem TTGGGG repeats. Only dGTP and dTTP are needed for the activity. The telomerase is a large ribonucleoprotein. It contains a short RNA component, 159 bases long in *Tetrahymena*, 192 bases long in *Euplotes*. Each RNA includes a sequence of 15-22 bases that is identical to two repeats of the C-rich repeating sequence given in Table 1.1. This RNA provides the template for synthesizing the G-rich repeating sequence, to which it is complementary. Bases are added individually, in the correct sequence. The enzyme progresses discontinuously. The telomerase is a specialized example of a reverse transcriptase. The protein component provides the catalytic activity of reverse transcriptase, and is (presumably) confined to acting upon the RNA template provided by the nucleic acid component.

The structure of the telomere is organized as single-stranded extension of the G-T-rich strand, usually for 14-16 bases.

A model for the structure of the end proposes the existence of a "quartet" of G residues, formed by an association of one G from each repeating unit. The association between the G residues requires that two of them change the orientation of the base with regard to the sugar (from the usual anti to be usual syn configuration). Since each repeating unit has more than one G, more than one quartet could be formed if other G residues associate, in which case quartets might be stacked upon one another in a helical manner.

It is not known how the complementary (C-A-rich) strand of the telomere is assembled, but we may speculate that it could be synthesized by using the 3'-OH of a terminal G-T hairpin as a primer for DNA synthesis.

1.3 Heterochromatin and Euchromatin

1.3.1 Heterochromatin

Light-microscope studies in the 1930s distinguished between two types of chromatin in the interphase nuclei of many higher eukaryotic cells: a highly condensed form, called ***heterochromatin***, and all the rest, which is less condensed, called ***euchromatin*** which is composed of the types of chromosomal structures—30-nm fibers and looped domains. Heterochromatin, in contrast, includes additional proteins and probably represents more compact levels of organization that are just beginning to be understood. In a typical mammalian cell, approximately 10% of the genome is packaged into heterochromatin. Although present in many locations along chromosomes, it is concentrated in specific regions, including the centromeres and telomeres.

Heterochromatin is classified as :—

- (i) Constitutive
- (ii) Facultative

(i) Constitutive heterochromatin :

It consists of specific regions that are not expressed. They include satellite DNAs, and could play a structural role in the chromosome. Often these sequences are concentrated in specific regions, typically around the centromere,

(ii) Facultative heterochromatin :

It takes the form of entire chromosomes that are inactive in one cell lineage, although they can be expressed in other lineages. The example *parexcellence* is the mammalian X-chromosome, one copy of which (selected at random) is entirely inactive in a given female. (This compensates for the presence of two X chromosomes, compared with the one present in males.) The inactive X chromosome is perpetuated in a heterochromatic state, while the active X chromosome is part of the euchromatin. Here it is possible to see a correlation between transcriptional activity and structural organization when the *identical DNA sequences are involved in both states*.

Most DNA that is folded into heterochromatin does not contain genes. However, genes that do become packaged into heterochromatin are usually resistant to being expressed, because heterochromatin is unusually compact. Regions of heterochromatin are responsible for the proper functioning of telomeres and centromeres (which lack genes), and its formation may even help protect the genome from being overtaken by “parasitic” mobile elements of DNA. Moreover, a few genes require location in heterochromatin regions if they are to be expressed. In fact, the term *heterochromatin* (which was first defined cytologically) is likely to encompass several distinct types of chromatin structures

whose common feature is an especially high degree of organization. When a gene that is normally expressed in euchromatin is experimentally relocated into a region of heterochromatin, it ceases to be expressed, and the gene is said to be *silenced*. These differences in gene expression are examples of position effects, in which the activity of a gene depends on its position along a chromosome. Many **position effects** exhibit an additional feature called position effect *variegation*, which result from patches of cells in which a silenced gene has become reactivated; once reactivated, the gene is inherited stably in this form in daughter cells. The study of position effect variegation has revealed two important characteristics of heterochromatin. First, heterochromatin is dynamic; it can “spread” into a region and later “retract” from it at low but observable frequencies. Second, the state of chromatin—whether heterochromatin or euchromatin—tends to be inherited from a cell to its progeny. These two features are responsible for position effect variegation, as explained in Figure 1.9.

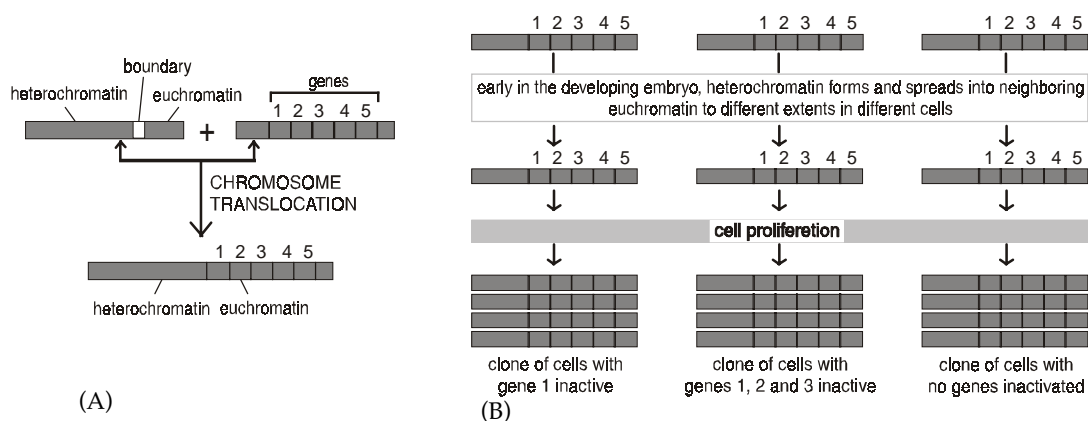


Fig. 1.9 The cause of position effect variegation in *Drosophila*. (A) Heterochromatin is normally prevented from spreading into adjacent regions of euchromatin by special boundary DNA sequences. In flies that inherit certain chromosomal rearrangements, however, this barrier is no longer present. (B) During the early development of such flies, heterochromatin can spread into neighboring chromosomal DNA, proceeding for different distances in different cells. This spreading soon stops, but the established pattern of heterochromatin is inherited, so that large clones of progeny cells are produced that have the same neighboring genes condensed into heterochromatin and thereby inactivated. Although “spreading” is used to describe the formation of new heterochromatin near previously existing heterochromatin, the term may not be wholly accurate. There is evidence that during expansion, heterochromatin can “skip over” some regions of chromatin, sparing the genes that lie within them from repressive effects. One possibility is that heterochromatin can expand across the base of some DNA loops, thus bypassing the chromatin contained in the loop

The ends of chromosomes have a special form of heterochromatin

The molecular nature of heterochromatin is probably best understood in the simple yeast *S. cerevisiae*. Mutations in any one of a set of yeast Silent information regulator (Sir) proteins prevent the silencing of genes located near

telomeres, thereby allowing these genes to be expressed. Analysis of these proteins has led to the discovery of a telomere-bound Sir protein complex that recognizes underacetylated N-terminal tails of selected histones (Fig. 1.10A). One of the proteins in this complex is a highly conserved histone deacetylase known as Sir2, which has homologs in diverse organisms, including humans, and presumably has a major role in creating a pattern of histone under acetylation unique to heterochromatin.

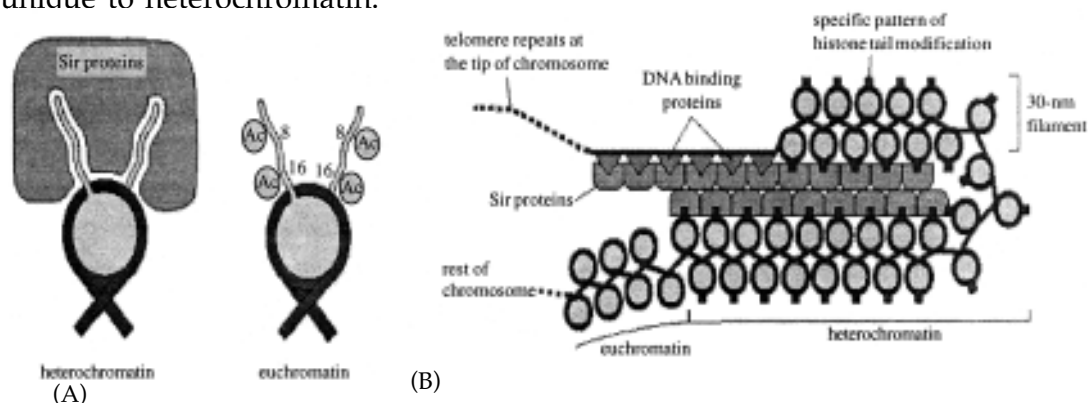


Fig. 1.10 Speculative model for the heterochromatin at the ends of yeast chromosomes. (A) Heterochromatin is generally underacetylated, and underacetylated tails of histone H4 are proposed to interact with a complex of Sir proteins, thus stabilizing the association of these proteins with nucleosomes. Although shown as fully unacetylated, the exact pattern of histone H4 tail modification required to bind to the Sir complex is not known with certainty. In some organisms, the methylation of lysine 9 of histone H3 is also a critical signal for heterochromatin formation. In euchromatin, histone tails are typically highly acetylated. Those of H4 are shown as partially acetylated but, in reality, the acetylation state varies across euchromatin. (B) Specialized DNA-binding proteins (*blue triangles*) recognize DNA sequences near the ends of chromosomes and attract the Sir proteins, one of which (Sir2) is a NAD^+ dependent histone deacetylase. This then leads to the cooperative spreading of the Sir protein complex down the chromosome. As this complex spreads, the deacetylation catalyzed by Sir2 helps create new binding sites on nucleosomes for more Sir protein complexes. A “fold back” structure of the type shown may also form

But how is the Sir2 protein delivered to the ends of chromosomes in the first place? A DNA-binding protein that recognizes specific DNA sequences in yeast telomeres also binds to one of the Sir proteins, causing the entire Sir protein complex to assemble on the telomeric DNA. The Sir complex then spreads along the chromosome from this site, modifying the N-terminal tails of adjacent histones to create the nucleosome-binding sites that the complex prefers. This “spreading effect” is thought to be driven by the cooperative binding of adjacent Sir protein complexes, as well as by the folding back of the chromosome on itself to promote Sir binding in nearby regions (see Fig. 1.10B). In addition, the formation of heterochromatin probably requires the action of chromatin remodeling complexes to readjust the positions of nucleosomes as they are packed together.

Whatever the precise mechanism of heterochromatin formation, it has become clear that covalent modifications of the nucleosome core histones have a critical role in this process. Of special importance in many organisms are the *histone methyl transferases*, enzymes that methylate specific lysines on histones including lysine 9 of histone H3. This modification is “read” by heterochromatin components (including HP1 in *Drosophila*) that specifically bind this modified form of histone H3 to induce the assembly of heterochromatin. It is likely that a spectrum of different histone modifications is used by the cell to distinguish heterochromatin from euchromatin.

Centromeres are also packaged into heterochromatin

In many complex organisms, including humans, each centromere seems to be embedded in a very large stretch of heterochromatin that persists throughout interphase. The structure and biochemical properties of this so-called *centric heterochromatin* are not well understood, but, like other forms of heterochromatin, it silences the expression of genes that are experimentally placed into it. It contains, in addition to histones (which are typically underacetylated and methylated in heterochromatin), several additional structural proteins that compact the nucleosomes into particularly dense arrangements.

Heterochromatin may provide a defense mechanism against mobile DNA elements

DNA packaged in heterochromatin often consists of large tandem arrays of short, repeated sequences that do not code for protein, as we saw above for the heterochromatin of mammalian centromeres. This suggests that some types of repeated DNA may be a signal for heterochromatin formation. This feature, called *repeat-induced gene silencing*, may be a mechanism that cells have for protecting their genomes from being overtaken by mobile genetic elements. These elements can multiply and insert themselves throughout the genome.

Once a cluster of such mobile elements has formed, the DNA that contains them would be packaged into heterochromatin to prevent their further proliferation. The same mechanism could be responsible for forming the large regions of heterochromatin that contain large numbers of tandem repeats of a simple sequence, as occurs around centromeres.

1.3.2 Euchromatin

Euchromatin is a lightly packed form of **chromatin** that is rich in gene concentration, and is often (but not always) under active **transcription**. Unlike **heterochromatin**, it is found in both **eukaryotes** and **prokaryotes**.

Structure

The structure of euchromatin is reminiscent of an unfolded set of beads along a string, where those beads represent nucleosomes. Nucleosomes consist of eight proteins known as histones, with approximately 145bp of DNA wound around them; in euchromatin this wrapping is loose so that the raw DNA may be accessed. Each core histone possesses a 'tail' structure which can vary in several ways; it is thought that these variations act as "master control switches" which determine the overall arrangement of the chromatin. In particular, it is believed that the presence of methylated lysine 4 on the histone tails acts as a general marker for euchromatin. The exact organization of the DNA within euchromatin is not known, but with the electron microscope it is possible to see loops of DNA within the euchromatin regions, each loop between 40 and 100 kb in length and predominantly in the form of the 30 nm chromatin fiber. The loops are attached to the nuclear matrix via AT-rich DNA segments called **matrix-associated regions (MARs)** or **scaffold attachment regions (SARs)**.

Appearance

Euchromatin generally appears as light-colored bands when stained in **GTG banding** and observed under an **optical microscope**; in contrast to **heterochromatin**, which stains darkly. This lighter staining is due to the less compact structure of euchromatin. It should be noted that in **prokaryotes**, euchromatin is the only form of chromatin present; this indicates that the heterochromatin structure evolved later along with the **nucleus**, possibly as a mechanism to handle increasing genome size and therefore a decrease in safety/managability.

Function

Euchromatin participates in the active transcription of DNA to mRNA products. The unfolded structure allows gene regulatory proteins and RNA polymerase complexes to bind to the DNA sequence, which can then initiate the transcription process. Not all euchromatin is necessarily transcribed, but in general that which is not is transformed into **heterochromatin** to protect the genes while they are not in use. There is therefore a direct link to how actively productive a cell is and the amount of euchromatin that can be found in its nucleus. It is thought that the cell uses transformation from euchromatin into heterochromatin as a method of controlling gene expression and replication, since such processes behave differently on densely compacted chromatin- this is known as the 'accessibility hypothesis' .

1.4 Giant chromosomes : Polytene and Lampbrush chromosomes

1.4.1 Polytene chromosomes

In dividing diploid cells the DNA synthetic phase (S phase) is regularly followed by mitosis (M phase). The alternation of G₁, S, G₂, M and G₁ phases is called the **cell cycle**. The process of recurrent duplication cycle without consequent mitosis is called endoreduplication. Many of the cells of certain fly larvae grow to an enormous size through multiple cycles of DNA synthesis without cell division. The resulting giant cells contain as much as several thousand times the normal DNA complement. Cells with more than the normal DNA complement are said to be **polyploid** when they contain increased numbers of standard chromosomes. In several types of secretory cells of fly larvae, however, all the homologous chromosome copies are held side by side, creating a single polytene chromosome. The fact that, in some large insect cells, polytene chromosomes can disperse to form a conventional polyploid cell demonstrates that these two chromosomal states are closely related, and that the basic structure of a polytene chromosome must be similar to that of a normal chromosome. Instances of polyploid chromosomes in *Drosophila* include ovary nurse cells, follicle cells surrounding oocytes, abdominal histoblasts (see **Escargot**), fat body cells, gut cells, and cells of the late prepupal **salivary gland**. During the process of polyploidization, chromosomes become multistranded.

Polytene chromosomes are large and precisely aligned side-by-side adherence of individual chromatin strands greatly elongates the chromosome axis and prevents tangling. Polyteny has been most studied in the salivary gland cells of *Drosophila* larvae, in which the DNA in each of the four *Drosophila* chromosomes has been replicated through 10 cycles without separation of the daughter chromosomes, so that **1024** (2^{10}) identical strands of chromatin are lined up side by side.

When viewed in the light microscope, distinct alternating dark **bands** and light **interbands** are visible. Each band and interband represents a set of **1024** identical DNA sequences arranged in register. About **95 %** of the DNA in polytene chromosomes is in bands, and **5%** is in interbands. The chromatin in each band appears dark, either because it is much more condensed than the chromatin in the interbands, or because it contains a higher proportion of proteins, or both. Depending on their size, individual bands are estimated to contain 3000-300,000 nucleotide pairs in a chromatin strand. The bands of *Drosophila* polytene chromosomes can be recognized by their different thickness and spacings, and each one has been given a number to generate a chromosome "map." There are

approximately 5000 bands and 5000 interbands in the complete set of *Drosophila* polytene chromosomes.

Polyploid chromosomes exhibit a banded structure that is reproducible from individual to individual. In *Drosophila* there are thousands of recognizable bands. In situ hybridization of cloned complementary DNA of identified genes to banded polyploid chromosomes allows the localization of genes to individual chromosome bands. Chromosomal rearrangements are easily documented by comparing the order of bands between individuals, lines or even species. The degree of rearrangement observed between species is indicative of their evolutionary distance. *Drosophila melanogaster* has four pairs of chromosomes, three pairs of autosomes and a pair of sex chromosome.

The reference system proposed by Bridges divides the limbs of salivary gland chromosomes into 102 sections called "divisions" designated by number from 1 to 102. Each of the five main limbs (X, 2L, 2R, 3L, and 3R) contains 20 divisions; the short chromosome 4 contains only two divisions. The divisions are started with a prominent band and divided further into 6 subdivisions, each designated with capital letters from A to F. Each subdivision starts with a sharp band. Thus each individual band of salivary gland chromosomes can be identified by giving the division number, subdivision, and the number of the band starting from the beginning of the subdivision. Bridges presents the following minimum numbers of bands for the salivary gland chromosomes of *Drosophila melanogaster*: 537 bands for the X chromosome, 1032 bands for the second chromosome, 1047 bands for the third chromosome, and 34 bands for the fourth chromosome, totalling a minimum of 2650 bands for the whole genome. In this initial count doublets were listed as single bands; more recent interpretations give the total number of bands as 3286 (Sorsa, 1988).

In late prepupal salivary gland chromosomes, not all DNA in each of the chromosomes is polyploid. Approximately a third of the *Drosophila* genome is represented by heterochromatin, and heterochromatic regions are underrepresented in polytene chromosomes as these regions do not undergo endoreduplication. For example, the *rolled* locus is found in a heterochromatic region of chromosome 2 that is considered to remain condensed (and for the most part transcriptionally inactive) throughout all or most of the cell cycle, *rolled* lies in what is considered to be alpha heterochromatin, a chromosome region that makes up the chromocenter of polytene salivary gland chromosomes. The chromocenter is thought to be made up of DNA and protein in a dense, tightly knit structure that is transcriptionally inactive. Such heterochromatic regions, which make up 30% of the *Drosophila* genome, have a much lower density of genes as compared to euchromatin. *Rolled* gene activity is unusual in that it requires the surrounding heterochromatin for gene function. Rolled gene

activity is severely impaired by bringing *rolled* close to any euchromatic position. However, these position effects can be reversed by chromosomal rearrangements that bring the *rolled* gene closer to any block of autosomal or X chromosome heterochromatin (Eberl, 1993).

Both bands and interbands in Polytene chromosomes contain genes

Since the number of bands in *Drosophila* chromosomes was once thought to be roughly equal to the number of genes in the genome, it was initially thought that each band might correspond to a single gene; however, we now know this simple idea is incorrect. There are nearly three times more genes in *Drosophila* than chromosome bands, and genes are found in both band and interband regions. Moreover, some bands contain multiple genes, and some bands seem to lack genes altogether.

It seems likely that the band-interband pattern reflects different levels of gene expression and chromatin structure along the chromosome, with genes in the less compact interbands being expressed more highly than those in the more compact bands. The remarkable appearance of fly polytene chromosomes is thought to reflect the heterogeneous nature of the chromatin compaction found along all interphase chromosomes. The remarkable appearance of fly polytene chromosomes is thought to reflect the heterogeneous nature of the chromatin compaction found along all interphase chromosomes.

Individual Polytene chromosome bands can unfold and refold as a unit

A major factor controlling gene expression in the polytene chromosomes of *Drosophila* is the insect steroid hormone ***ecdysone***, the levels of which rises and falls periodically during larval development. When ecdysone concentrations rise, they induce the expression of genes coding for proteins that the larva requires for each molt and for pupation. As the organism progresses from one developmental stage to another, distinctive ***chromosome puffs*** arise and old puffs recede as new genes become expressed and old ones are turned off. Most puffs arise from the decondensation of a single chromosome band.

Puffing is the term that describes structural changes in polytene chromosomes. If one observes polytene chromosomes during the late prepupal stage, different bands appear to be puffed up. Puffs, then, afford a view of the temporal sequence of gene activation. A temporal pattern to puffing in the salivary glands of late prepupal flies is inducible by ecdysone injection and is therefore under control of the **ecdysone receptor**. A small number of genes react by puffing within minutes of exposure to ecdysone, and a much larger number (>100) react within hours. It is hypothesized that the time sequence of puffing represents a genetic hierarchy of gene activation. Early puffs are independent of protein

synthesis while late puffs require prior protein synthesis (Ashburner, 1990).

In recent years, transcription factors and chromosomal proteins have been localized to various bands. Binding of these proteins is thought to have functional significance and to reflect the activity of these proteins in gene regulation. For more information on the binding of various proteins and RNA species to bands, see **HP1/ Su(var)205, polycomb, male sex lethal 2, and suppressor of hairy wing.**

An example of binding of specific proteins to polytene chromosomes is found in a study of the protein CHD1 (chromo-ATPase/helicase-DNA-binding domain). Proteins related to CHD1 via the helicase domain have been shown to exist in large multiprotein complexes. For example SNF2/SWI2/Brm proteins are thought to participate in ATP-dependent remodeling of chromatin. Antibodies to CHD1 localize this protein to extended chromatin (interbands) and regions associated with high transcriptional activity (puffs) on polytene chromosomes from salivary glands. These observations support the idea that CHD1 functions to alter chromatin structure in a way that facilitates gene expression (Stokes, 1996).

Polyploidization by endoreduplication requires regulation of the **cell cycle**. What makes one region of the chromosome become polyploid while another remains underreplicated. Information about the roles of cell cycle genes in the regulation of polyploidization can be found in **cyclin E, Escargot, and origin recognition complex 2.**

Electron micrographs of certain puffs, called Balbiani rings, of *Chiwnomus* salivary gland polytene chromosomes show the chromatin arranged in loops, much like those observed in the amphibian lampbrush chromosomes. Each loop contains a single gene. When not expressed, the loop of DNA assumes a thickened structure, possibly a folded 30-nm fiber, but when gene expression is occurring, the loop becomes more extended. Both types of loops contain the four core histones and histone H1.

It seems likely that the default loop structure is a folded 30-nm fiber and that the histone modifying enzymes, chromatin remodeling complexes, and other proteins required for gene expression all help to convert it to a more extended form whenever a gene is expressed.

1.4.2 Lampbrush chromosomes

In 1882 Fleming first observed these chromosomes in urodele amphibian ovary. Riickert (1892) first described in great detail in shark oocytes. He coined the name "Lampbrush Chromosome" because of their brush-like appearance. These chromosomes' occur at diplotene stage of meiotic prophase in oocytes of all animal species, in spermatocytes of several species and even in giant nucleus

of unicellular algae *Acetabularia*. These chromosomes are characterized by several lateral projections called "Lateral Loops". They are very large and best seen in salamander oocytes because of their high DNA content.

Morphology : Lampbrush chromosomes are extensible and elastic. These chromosomes with well-developed lateral loops can be stretched to about 2½ times of original length. Since, these chromosomes are found in meiotic prophase they are present in the form of bivalents in which maternal and paternal chromosomes are held together by chiasmata. The axis of each chromosome consists of a row of granules or chromomeres and from which lateral loops extend.

(i) Centromeres : These are round, smooth, and Fuelgen positive and bear no lateral loop. In many species of urodele, centromeres are identifiable chromosome landmarks as "axial bars", formed by the amalgamation of neighbouring chromomeres, whereas in certain species of urodeles centromeres are not flanked by axial bars and are difficult to be identified, flank them. In such urodele species partner centromere do not fuse whereas urodele having axial bars centric fusion occurs.

(ii) Telomeres : Ends of Lampbrush chromosomes are occupied by distinctive telomeres consisting of a small fuelgen positive part closely applied to the surface of a smooth round Fuelgen negative part. Fuelgen negative material can be digested by proteolytic enzyme. Like centromere, Telomere do not possess lateral loop. They are of different sizes, large in *T. c. cristatus* and small in *T. c. karelini*. In some urodele fusion b/w telomeres are common in *T. c. aristatus* while in *T. c. karelini* fusions are rare.

(iii) Lateral loops : Loops are always symmetrical. Each chromosome having two of them, one for each chromatid. Loops can be distinguished by size, thickness and by several other morphological characteristics. Each loop appears at a constant position in the chromosome and there are about 10,000 loops / chromosome set. Each loop has axis formed by a single DNA molecule. About 5-10% DNA is present in the lateral loops the rest is condensed in chromomeres of chromosome axis, which is transcriptionally inactive.

Types of loops :

1. **Normal loop :** Most loops can form one pattern termed as normal loop while other loops are distinguished by their matrix deposition.

2. **Granulous loop :** They are so called because they accumulate granule at the distal end of the fine fibers projection from other classes of lateral loop differ from granulous loop in material accumulation.

(a) In some cases matrix plastering the axis but leaving the tips of the projecting fibers visible.

(b) In other cases, matrix fusion is irregular over the loop length. Such loops have uneven outlines.

3. Lumpy Loop : Situated on either side of centromere much degree of matrix fusion, usually so great that the loop pattern is wholly observed. Sometimes sister lumpy loop may fuse together so that instead of a pair of loops a single amorphous body is present.

4. Giant Loops : They are much larger than lumpy are matrix that they accumulate is exceedingly heterogeneous in texture. Each loop has its own developmental sequence of extension and regression. For example, giant granular loops are already of full size in very young oocytes and remain the same throughout oocyte development. The giant fusing loops are small in small oocytes and regress only just before ovulation. The granulous loops are largest in young oocytes is not regress early.

Unineme theory and C. value paradox :

The results of most of the earlier studied have revealed that each loop has just 1 DNA molecule as a major finding because it showed a single thread of DNA runs through each chromatid and lead to the elaboration of unineme model concept of chromosome structure.

A matrix covers each loop that consists of RNA transcripts with RNA binding proteins attached to them. In general ribonucleoprotein matrix is asymmetrical being thickness at one end of the loop than at the others. RNA synthesis starts at the thinner end and progresses to the thicker end. Many of the loops correspond to a single transcriptional unit while the other loop contains several units of transcription. Some t unite on lampbrush chromosome are extremely long i.e. over 100um in length i.e. (1u-m of DNA = 3000 bases) why they are so enormous? Even more puzzling is that the length of loop increases with C. value as a result Salamander has 10 times longer transcription units than those of a frog although both code for a similar gene product. We have no answer to this paradox but it may be possibly connected with inefficient termination of transcription in oocytes.

The results of *in situ* hybridization studies have revealed that long Lampbrush transcripts are due to failure of termination i.e. transcription eventually stops where the next t. unit is reached. However the function of these long transcripts remain unexplainable, though the majority of them are degraded in the nucleus but presumably some RNA has some role to play in preparing an oocyte for the journey that an egg undertakes after fertilization i.e. development of a new organism.

Anyway, at the cellular level of analysis both Lampbrush chromosome and

p.c. provide remarkably favorable opportunities to study the mechanisms responsible for gene ordered synthesis. Most of the recent exciting advances in our understanding of the nature and mode of action of the genetic material have come from the genetic studies on microorganisms. Cytologists have contributed rather little to this advances.

I.5 Suggested questions

1. Describe the structure of nucleosome along with diagrams.
2. State and explain the ways of chromatin remodeling.
3. State the different types of covalent modifications in histone tails and its significance.
4. Explain the significance of centromeric sequence in chromosomal segregation.
5. How is the length of telomere maintained in eukaryotic systems?
6. Explain "position effect variegation" with example.
7. How is heterochromatinization brought about?
8. Elucidate polytene chromosome structural organization.
9. Explain chromosomal puffs.
10. State the morphology of lampbrush chromosomes.
11. Validate uncinome theory by lampbrush chromosome structure.

Unit 2 □ Sex Chromosomes, Sex Determination and Dosage Compensation

Structure

- 2.1 Introduction
- 2.2 Sex determination and dosage compensation in *Caenorhabditis elegans*
- 2.3 Sex determination and dosage compensation in *Drosophila*
- 2.4 Genetic regulation of sex determination and gonadal differentiation in humans
- 2.5 Suggested questions

2.1 Introduction

In multicellular organisms sex is determined by many different mechanisms, which vary greatly. Of the various mechanisms of sex determination known till date, sex-chromosomal method of determination is perhaps the best understood and intriguing. Here, sex of an individual is determined by the presence or absence of its species-specific sex chromosomes. In this system of sex determination there are defined set of autosomes and well-defined pair of *allosomes* (sex chromosomes). The allosomes may be of one kind (e.g. in *C. elegans*, Grasshopper etc. has only X chromosome; thus two sexes are determined by either XX or XO) or of two different kinds (e.g. *Drosophila* has both X and Y chromosomes and sexes are determined by XX or XY).

The paradox for such mechanism of sex determination is in the fact that either of the two sexes have different sex chromosomal constitution, leading to differential allosomal gene dosages. In many organisms there are two X chromosomes in female and one X in male. Therefore, it is essential to make a balance between the products of the genes of two X chromosomes and the products of one X chromosome. The mechanism by which the balance between two dosages and one dose is maintained is known as dosage compensation. This is done either by suppressing the activities of the genes of one of the two X chromosomes of the female (inactivation of one of the female X chromosome) or by hyperactivation of the male X chromosome. This would thus require dosage compensation to negate the genie imbalance for the sex chromosomes. Although sex determination pathway and dosage compensation are different pathways they may have few steps in common but must not be considered to be same under any circumstances.

2.2 Sex determination and dosage compensation in *Caenorhabditis elegans*

Caenorhabditis elegans has two sexes: hermaphrodites and males. Hermaphrodites are essentially female animals that produce sperm during larval development and oocytes during adulthood. Hence, hermaphrodites are capable of self-fertilization, as well as cross-fertilization by males. Although some adult structures such as the pharynx are similar in males and hermaphrodites, most tissues and many aspects of behavior are different.

The pathway is not as linear and that several loops and branches in the pathway play important roles in specifying sexual development.

2.2.1 Control of *xol-1* by the X:A Ratio

The primary signal for sex determination is the ratio of X chromosomes to sets of autosomes, which causes XX animals to become hermaphrodites and XO animals to become males. Early in development this ratio regulates the activity of *xol-1* (Fig.2.1), a key developmental switch gene that controls both sex determination and dosage compensation, *xol-1* encodes a novel protein, and during early embryogenesis, high levels of XOL-1 protein activity promote male development and low levels promote hermaphrodite development. The male specifying *xol-1* transcript is not needed after the end of gastrulation.

The early time at which *xol-1* acts strongly suggests that it is a direct target of the X : A signal. This signal must involve elements on the X and elements autosomes that are compared.

The X chromosome signal is polygenic, and that the combined action of these X signal element is required to inhibit *xol-1* activity in hermaphrodites. At least four different regions, regions 1-4, of the X contain signal elements, and two of these elements have been identified molecularly: *sex-1* and *fox-1*. Increasing

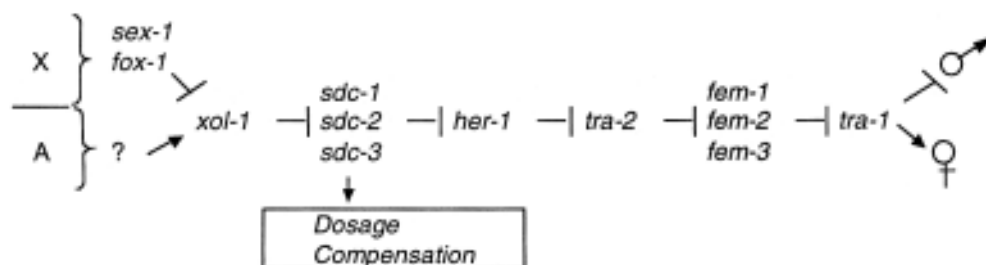


Fig. 2.1 The basic sex determination and dosage compensation pathways in *C. elegans*

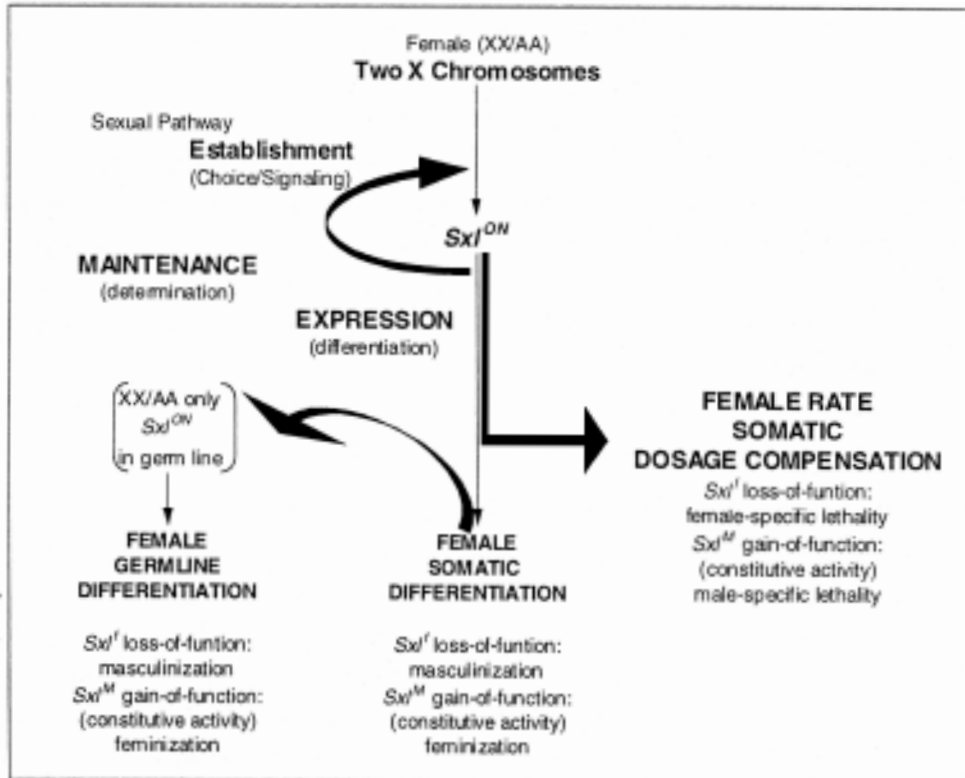


Fig. 2.2 The female-specific developmental switch gene, *Sex-lethal*, counts X chromosomes early in development to establish the choice between male and female alternative pathways of development at the cellular level. Through an autoregulatory feedback loop, *Sxl* subsequently maintains this choice throughout development, and ultimately, directs sexually dimorphic aspects of differentiation through its effects on different sets of subordinate genes downstream. Because one of these sets controls the vital process of X chromosome dosage compensation, misregulation of *Sxl* caused by upsets in the X chromosome counting process is lethal to one sex or the other. This lethality obscures potential effects on sexual phenotype. Loss-of-function (f-type) mutations in this gene are deleterious to chromosomal females (XX), while gain-of-function mutations (M-type) lead to constitutive expression and are deleterious to chromosomal males (XY). Control of *Sxl* in the germ line requires sex-specific input from the soma

the dose of these elements in XO animals represses *xol-1*, promotes hermaphrodite development and causes death because dosage compensation is activated. Decreasing their dose in XX animals activates *xol-1*, promotes male development, and causes death due to failure to initiate dosage compensation. To date, no autosomal signal elements have been identified. Evidence indicates that *sex-1* regulates the transcription of *xol-1*.

In contrast to *sex-1*, *fox-1* and region 2 act posttranscriptionally to regulate *xol-1* expression. The *fox-1* gene encodes a protein with ribonuclear protein (RNP)

motifs, suggesting that it might bind the *xol-1* RNA. It is possible that the OF-1 protein regulates *xol-1* alternative splicing, or it might govern another aspect of *xol-1* mRNA metabolism.

Combinatorial effect of these regulatory mechanisms allows the worm to discriminate accurately between small differences in the X : A.

2.2.2 Control of the *sdc* genes by *xol-1*

Three genes are required in XX animals to promote both hermaphroditic development and dosage compensation—*sdc-1*, *sdc-2* and *sdc-3*. The primary means by which XOL-1 transmits the X:A signal appears to be by negative regulation of *sdc-2*, as *sdc-2* is not expressed in wild-type XO embryos, but is expressed in *xol-1* XO embryos.

Null mutations in *sdc-2* and *sdc-3* have no effect on XO animals but cause complete reversal of sexual fate in XX animals; null mutations in *sdc-1* cause only a partial reversal of sexual fate. The *sdc* genes control XX hermaphrodite development by regulating the expression of the downstream sex-determining gene, *her-1*, a gene required for male development.

SDC-2 and SDC-3 might act in a complex to directly repress *her-1* transcription. This is supported by the finding that SDC-2, is highly charged and contains coiled-coil motifs, is targeted to transgenic copies of the *her-1* promoter. Moreover, this localization is blocked by specific *sdc-3* mutations, called *sdc-3* (*Tra* alleles). SDC-3 is a novel protein that contains two functional domains. The first is a zinc finger motif that is required for dosage compensation but not for sex determination. The second resembles a **myosin** ATPase domain, and is necessary for sex determination but not for dosage compensation. The *sdc-3* (*Tra*) mutations affect this latter domain. The role of SDC-1 in regulating sexual development is less clear. SDC-1 has seven zinc fingers and resembles TFIIIA.

2.2.1 Somatic sex determination

Regulation of TRA-2A by HER-1

her-1 is required for male development, as mutations in *her-1* cause XO animals to develop as hermaphrodites but do not affect dosage compensation. The *her-1* gene is predicted to encode a novel protein with an amino-terminal signal sequence and protein cleavage and glycosylation sites, suggesting that HER-1 is a secreted protein.

HER-1 promotes male development by repressing the activity of *tra-2*. Major transcript of this gene, *tra-2*, encodes a transmembrane protein, a direct interaction between secreted HER-1 and TRA-2A.

Regulation of sexual fate by *tra-2*

TRA-2A might be processed to release its intracellular domain, TRA-2ic. Production of TRA-2ic might occur by the action of TRA-3, a member of the calpain protease family. The *tra-3* gene is necessary for hermaphroditic development. TRA-3 can proteolytically cleave TRA-2A to release TRA-2ic in insect cells. How HER-1 inhibits TRA-2A activity is unclear. One possibility is that HER-1 inhibits production of TRA-2ic by TRA-3. *tra-2* activity is controlled not only at the level of protein processing or protein interaction, but also at the translational level by two elements called *tra* *Gli* elements (TGEs) which are located in the 3' untranslated region (3'UTR) of the *tra-2* message. Epistasis test show that *tra-2* promotes hermaphrodite development by inhibiting the activity of three genes, *fem-1*, *fem-2* and *fem-3*. TRA-2A does not transcriptionally regulate these genes, as they are all expressed at high levels in both sexes. Instead it appears that TRA-2A or TRA-2ic inhibits FEM activities by protein-protein interaction. FEM-2 can bind FEM-3, so they might interact to promote male development. In addition, TRA-2A and TRA-2ic can bind FEM-3, which they might inactivate to allow hermaphrodite development.

Regulation of TRA-1A activity by the FEM proteins

The final gene in the sex-determination pathway is *tra-1*, which acts cell autonomously to promote hermaphrodite development. Although genetic experiments indicate that the FEM proteins promote male development by inhibiting TRA-1 A activity, how they do so is a mystery. They are unlikely to act transcriptionally, as *ra-1* mRNA levels do not differ between males and hermaphrodites. The phosphatase activity of Fem-2 is necessary for its activity, so it is possible that FEM-2 control the activity of TRA-1 A by altering its phosphorylation state. Alternatively the FEM proteins might control sexual

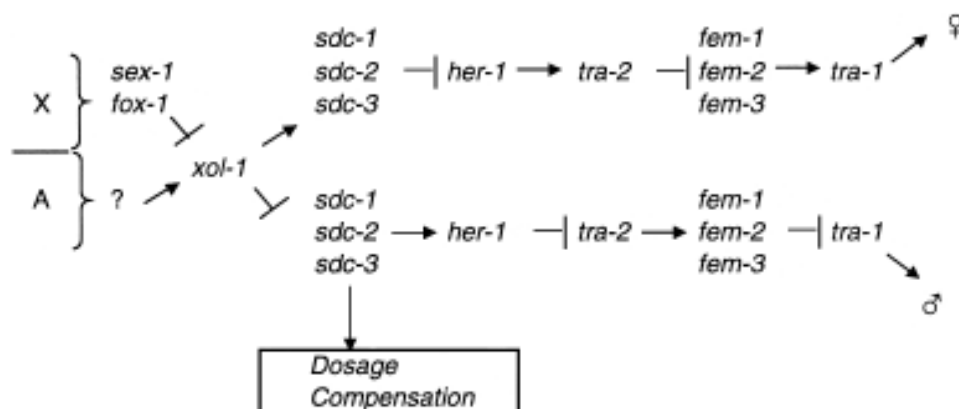


Fig. 2.3 Schematic diagram representing both male and female specific sex determination cascade in *C. elegans*

development by regulating the nuclear levels of TRA-1A. Recent analyses have revealed that hermaphrodite tissues have higher TRA-1A nuclear levels than male tissue. Thus TRA-1A transcriptional regulatory activity might be specified by nuclear versus cytoplasmic distribution of the protein. Furthermore, the FEM proteins might regulate nuclear import or export of TRA-1A, as TRA-1A is almost completely nuclear in loss-of-function *fem-1* animals. The entire sex-determination cascade can be summarized as Figure 2.2.

2.3 Sex determination and dosage compensation in *Drosophila*

In 1921 Bridges proposed that ‘..Sex in *D. melanogaster* is determined by a balance between the genes contained in the X chromosomes and those contained in the autosomes. It is not the simple possession of two X chromosomes that makes a female, and of one that makes a male’.

The discovery of numerator and denominator elements appears to validate the concept of sex determination of; as a ratio-measuring process. Proposed regulation cascade for *Drosophila* sex determination is schematically given below (Fig. 2.4)

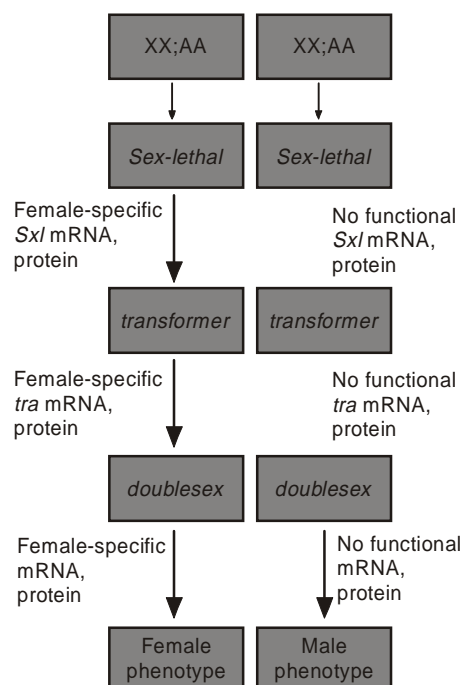


Fig. 2.4 Sex determination in *Drosophila*. This simplified scheme shows that the X-to-autosome ratio is monitored by the *Sex-lethal* gene. If this gene is active, it processes the *transformer* mRNA into a functional female-specific message. In the presence of the female-specific Transformer protein, the *doublesex* gene transcript is processed in a female-specific fashion. The female-specific Doublesex protein is a transcription factor that leads to the production of the female phenotype. If the *transformer* gene does not make a female-specific product (i.e., if the *Sex-lethal* gene is not activated), the *doublesex* transcript is spliced in the male-specific manner, leading to the formation of a male-specific Doublesex protein. This is a transcription factor that generates the male phenotype

***Sxl*, the target of the somatic sex determination signal**

Sexual differentiation and dosage compensation are a consequence of *Sxl* being turned on in diplo-X individuals and remaining off in haplo-X individuals. In somatic cells, the active (female) state is maintained by a positive feedback loop (Fig.2.2), rather than continued input from the X-linked genes that trigger the initial activation.

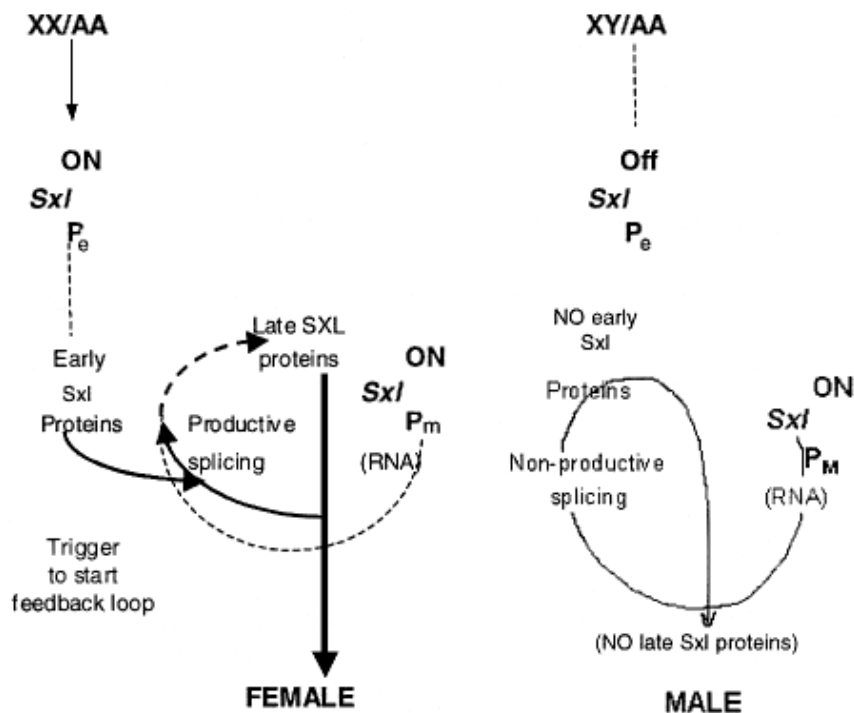


Fig. 2.5 Molecular steps in the operation of *Sxl* to establish and maintain a sexual pathway choice in somatic cells that is appropriate for their X chromosome dose

Figure 2.5 summarizes the current model for how, in the soma, a brief, very early effect of X chromosome dosage has a long-lasting effect on the activity of *Sxl*. The double dose of X chromosomes in females activates an *Sxl* 'establishment' promoter, P_e before the blastoderm stage, resulting in production of the 'early' *Sxl* proteins. In contrast, the single X chromosome dose in males leaves P_e inactive; hence males fail to produce early *Sxl* proteins.

Transition from the sexual pathway establishment (signaling) level of *Sxl* regulation to the pathway maintenance (determination) level reflects a switch in *Sxl* promoters and an attendant shift from transcriptional regulation to regulation at the level of RNA splicing. At the blastoderm stage, P_e shuts down and a 'maintenance' promoter, P_m , located 5 kb upstream, becomes active in both sexes and remains active throughout the rest of development. In contrast to

transcripts from P_e , transcripts from P_m can be spliced into mRNA encoding full-length *Sxl* proteins only if *Sxl* proteins (early or late) are already present.

The early burst of P_e expression in diplo-X individuals generates a pulse of *Sxl* early protein that directs the splicing of P_m -derived transcripts to eliminate a male-specific, translation-terminating exon that would otherwise block synthesis of active *Sxl* proteins. The *Sxl* proteins generated from P_m -derived transcripts then maintain productive RNA splicing through a positive auto regulatory feed back loop, and act on downstream genes to elicit female differentiation and suppress X chromosomes hyper activation.

In the absence of *Sxl* protein, downstream gene targets involved in both sexual differentiation and dosage compensation are expressed in a male-specific manner.

Numerator and denominator elements

The first candidate was *sisterless-a*. A zygotically acting positive regulator of *Sxl*. The gene was subsequently defined as 'numerator in experiments at that also revealed a second numerator elements, *sis-b*. the behavior of these two genes is characterized by reciprocal, zygotic dose-dependent *Sxl*-based, sex specific lethality.

One additional numerator element, the apir-rule segmentation gene *runt*, the first known denominator element, the semi lethal behavioral mutant *deadpan*. As a denominator element, *dpn* shows reciprocal, zygotic dose-dependent, *Sxl*-based, sex specific lethality that is the inverse of that for numerator elements; duplications, rather than deletions of this autosomal gene kill females.

Maternal involvement in progeny sex determination

Mother plays an important role in, sex determination by building into the egg the biochemical machinery the embryo needs to count its X chromosomes and decide whether to activates *Sxl*. Maternal daughterless *da* is a positive regulator that is necessary but not sufficient for *Sxl* activation. Without *da* activity cannot activate *Sxl* after fertilization, regardless of their X chromosome dosage. All progeny develop as males.

Analysis of the cDNA from *Sxl* mRNA shows that the *Sxl* mRNA of males differs from the *Sxl* mRNA of females. This is the result of differential RNA processing. Moreover, the *Sxl* protein appears to bind to its own mRNA precursor to splice it in the female manner. Since males do not have any available *Sxl* protein, their new 5x7 transcripts are processed in the male manner. The male *Sxl* mRNA is non functional. While the female-specific *Sxl* transcript contains a translation termination codon (UGA) after amino acid 48. The differential RNA processing that puts this termination codon into the male-specific mRNA. In

males, the nuclear transcript is spliced in a manner that yields three exons, and the termination codon is within the central exon. In females, RNA processing yields only two exons, and the male-specific central exon is now spliced out as a large intron. Thus, the female-specific mRNA lacks the termination codon. The protein made by the female-specific *Sxl* transcript can be predicted from its nucleotide sequence. This protein would contain two regions that are important for binding to RNA. Bell and colleagues (1988) have proposed that there are two targets for the RNA-binding protein encoded by *Sxl*. One of these targets is the pre-mRNA of *Sxl* itself. This would be the mechanism that would maintain the female state of the pathway after the initial activating event had passed. The second target of the female-specific *Sxl* protein would be the pre-mRNA of the next gene on the pathway, *transformer*.

The transformer genes

The *Sxl* gene regulates somatic sex determination by controlling the processing of the *transformer* gene transcript. The *transformer* gene (*tra*) is alternatively spliced in males and females. There is a female-specific mRNA and also a nonspecific mRNA that is found in both females and males. The non-specific *tra* mRNA contains a termination codon early in the message, making the protein non-functional. The second exon of the non-specific mRNA has the termination codon. This exon is not utilized in the female-specific message. The female-specific protein from the *Sxl* gene activates a female-specific 3' splice site in the transformer pre-mRNA, causing it to be processed in a way that splices out the second exon. To do this, the *Sxl* protein blocks the binding of splicing factor U2AF to the nonspecific splice site by specifically binding to the polypyrimidine tract adjacent to it. This causes U2AF to bind to the lower-affinity (female-specific) 3' splice site and generate a female-specific mRNA. The protein encoded by this message is critical in female sex determination.

The female-specific *tra* product acts in concert with the *transformer-2* (*tra2*) gene to help generate the female phenotype. The *tra2* gene is constitutively active and makes the same protein product in both males and females. This Tra2 protein, like that of the female-specific *Sxl* protein, contains an RNA-binding domain. It is proposed that the *tra2* gene can bind to the transcript of the *doublesex* gene, but only in the presence of the female-specific Tra protein.

Doublesex : the switch gene of sex determination

The *doublesex* gene is active in both males and females, but its primary transcript is processed in a sex-specific manner. Male and female transcripts are identical through the first three exons. The 3' exons differ markedly. What is an exon for the female-specific transcripts is part of the untranslated 3' end of the male-specific message.

The alternative RNA processing appears to be the result of the *transformer* genes. The Tra2 and female-specific Tral proteins bind specifically to a DNA sequence adjacent to the female-specific 3' splice site of the *dsx* pre-mRNA, and they recruit nonspecific splicing factors to this site. If *tra* is not produced, the *doublesex* transcript is spliced in the male-specific manner. The downstream 3' splice site is used, and a male-specific transcript is made. This encodes an active protein that inhabits female traits and promotes male traits. The Transformer proteins bind to sequences within the female-specific exon and activate the female-specific 3' splice site. This activation of an otherwise unused female-specific 3' splice site produces an mRNA encoding a female-specific protein that activates female-specific genes (such as those of the yolk proteins) and inhibits male development.

The functions of the Doublesex proteins can be seen in the formation of the *Drosophila* genitalia.

Target genes for the sex determination cascade

Numerous proteins in *Drosophila* are present in one sex and not the other. In females, these include yolk proteins and eggshell (chorion) proteins. In males,

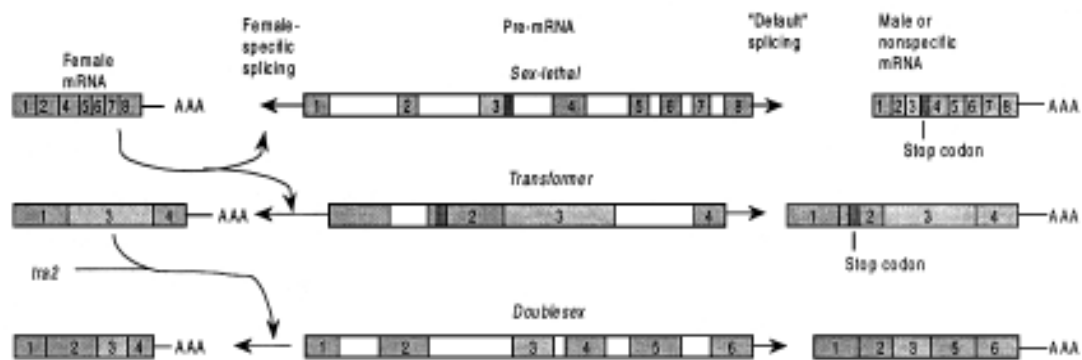


Fig. 2.6 The pattern of sex-specific RNA splicing in three major *Drosophila* sex-determining genes. The pre-mRNAs are located in the center of the diagram and are identical in both male and female nuclei. In each case, the female-specific transcript is shown at the left, while the default transcript (whether male or non-specific) is shown to the right. Exons are numbered, and the positions of the termination codons and poly(A) sites are marked. (After Baker, 1989)

the sex combs of the legs are sex-specific structure. Both the male and female *doublesex* transcripts bind to three sites within the 127-base-pair enhancer of the *yolk* protein genes. Their binding and mutagenesis studies demonstrate that the male-specific Doublesex product inhibits transcription by its binding to these sites, whereas the female-specific Doublesex protein activates gene transcription from the same sites.

2.4 Genetic regulation of sex determination and gonadal differentiation in humans

It is clear that only a small region of the Y-chromosome is endowed with the gene(s) for sex determination. The mystery of the sex reversed (Sxr) male mice having XX-chromosome constitution was resolved unambiguously by Singh and Jones (1982) who showed that one of the 2 X-chromosomes in the Sxr males carried a minute segment of the short arm of the Y-chromosome.

Testis determining gene on the Y-chromosome (TDY)

Subsequent studies led to the discovery of SRY gene (sex determining region of the Y) that coded for an HMG domain DNA binding protein. SRY is present on the short arm of human Y, and is conserved. The expression of Sry occurred from the onset of differentiation of medulla (primordial testis) in the genital ridge. Clinching evidence in favour of Sry as the male determining gene came from the sex reversal of the XX embryos to male through insertion of only a 14kb fragment of Y-chromosomal DNA having Sry. Transgenic mouse males are sterile the role of Sry as the male determining factor is confirmed. This single-exon gene codes for a protein that has a 79 amino acid long HMG (High mobility group proteins)-domain. The proteins harbouring this domain constitute a SOX (Sry-box) family of transcription factors that bind to a heptamer of nucleotides A/TAACAAT. Mutations in the HMG domain or in the upstream promoter region of Sry have been shown to result in XY pseudohermaphrodites, gonadal dysgenesis and other gonadal pathologies. The mouse Sry carries a polyglutamine stretch, which is absent in the human Sry.

Autosomal genes

XY individuals have however been reported with anomalies of gonad and urogenital system, in spite of having completely normal SRY (Sry) gene. More than 75% of the XY individuals suffering from acute dwarfism due to Campomelic dysplasia (a rare skeletal disorder) are hermaphrodite with ambiguous genitalia. This disorder is caused by mutation in a gene SOX9. Importance of SOX9 (Sox9 in mouse) in testis differentiation as well as subsequent organogenesis of the male genital system has since been established. An X-linked transcription factor, DAX1 (Xp21.2-22.2), has repressive effect on male development. The XY, SRY-positive pseudohermaphrodite individuals were found to have a duplication of DAX1 on the single X-chromosome, showing dosage-dependent effect of DAX1. Individuals suffering from WAGR (Wilms aniridia, genitourinary malformation, mental retardation) and Denys Drash syndromes both fail to have genital as

well as the renal development and are caused due to mutation in the WT-1 gene. Loss of function mutation in the autosomal WT-1, the Wilms1 tumour gene, also leads to gonadal dysgenesis. Similarly, mutation in SF-1 (orphan steroid factor-1 gene), has a broader effect on the mesonephric gonadal complex.

Interaction of genes in sexual differentiation

SF-1 and Sox9 bring about the activation of MIS gene, whose product leads to the regression of MD. These two transcription factors bind to the MIS promoter to induce its activity. WT-1 and GATA-4 act as cofactors of SF-1 and facilitate its binding to the MIS promoter. The X-linked, Dax-1, acts as a repressor of SF-1 in the female. Thus in the absence (or low level) of Sox9 and Sf-1, MIS (Mullerian Inhibiting Substances) is not produced and there is no regression of the Mullerian duct. Dax-1 is expressed also in males coincidentally with Sry, Sox9 and Sf-1 but its level is much lower than that in the female, the dosage of Dax-1 vis-a-vis Sry, Sox9, perhaps makes the difference between the differentiation of the male and female sex.

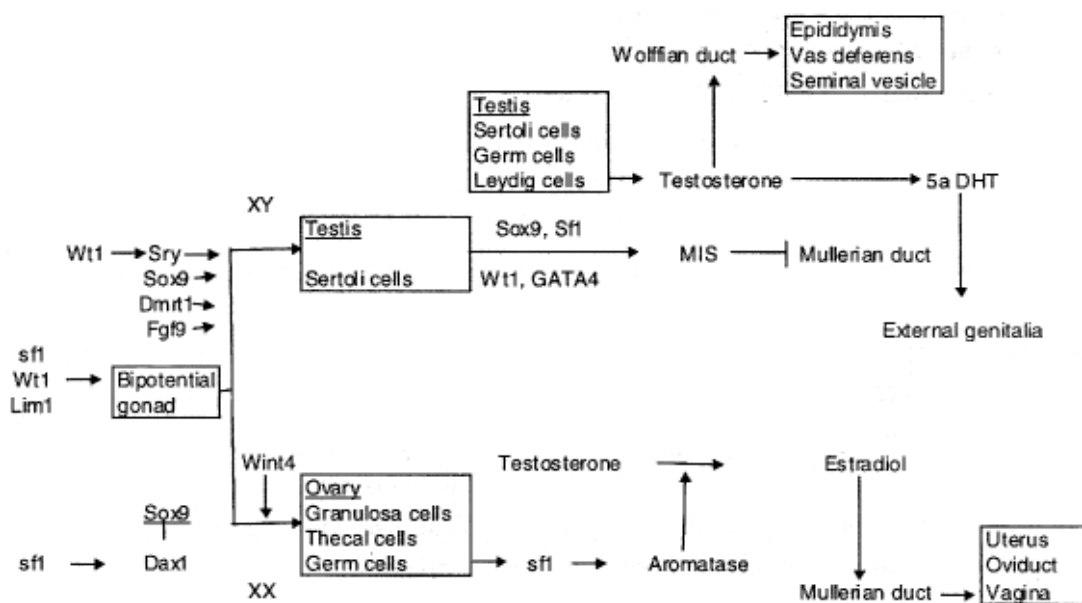


Fig. 2.7 Cartoon giving a simplified scheme of the genes participating in sex determination, and differentiation of the urogenital system in mammals

Obviously, *Dax1* is an “anti-male” rather than “female” gene. Its elevated level in female must be blocking not only SF-1 induced activation of *MIH* but also testicular development.

More genes in the gonadal pathway

In addition a number of cases of sex reversals and ambiguous genitalia have been observed through mutations in other parts of the genome. *Wnt4*, coding for a cell-signaling molecule, is likely to have a positive role in female determination, unlike the “anti-testis” *DAX1*.

Like SF-1 and WT-1, genes, *LIM1* and *LIM9*, also act in the mesonephros to trigger the differentiation of GR (GR = Gonadal Ridge). *Fgf9* mediates the migration of mesonephric cells in the testicular sex cords in GR of XY fetus. Therefore its absence is not detrimental to ovary in females. The XY-male specific expression of *Vanin-1*, a cell membrane asso-

ciated G-protein, during testis differentiation may also be playing an important role in directing the mesonephric cells towards the testicular sex cords in the genital ridge.

It is obvious from the foregoing that though autosomal and X-linked genes play a role in the differentiation of the urogenital system, *SRY* is the key gene that switches the bipotential genital ridge towards testis determination. DNA as well as RNA binding functions are assigned to *Sry* (and other *Sox* proteins). It is intriguing that it occurs only in mammals. Functional *SRY* initiates a chemotactic action that leads to the migration of mesonephric cells into GR. It has also been demonstrated that the *Sry* expression in GR is not only temporally but also spatially restricted. The most important role of *Sry* in gonadogenesis is to decisively direct the bipotential precursor cells in GR to form the Sertoli cells instead of the follicular cells that form ovary. Centrality of *SRY* in testis determination is established. Nevertheless, which genes are the *Sry*-specific target gene(s) still remains a mystery. In a recent experiment, introduction of WT1 - promoter-driven *Sox9* fusion gene succeeded in imparting male phenotype to the XX mouse, suggesting that if the *Sox9* level could be raised then *Sry* was redundant for testis formation. The immediate conclusion from this evidence is that *SOX9* may be the immediate target of *SRY*. It also suggests that WT-1 which is expressed upstream of *Sry* is involved in its activation.

Although many sex-chromosomal as well as autosomal genes play a role in mammalian (human) sex determination pathway, only few of them have been tabulated below (Table 2.1).

Table 2.1 Genes Commonly Involved in Sex Determination in Man, Other Mammals and Lower Vertebrates

Genes	Gene product	Mutant Phenotype in man/mouse	Mutational mechanisms	Orthologue in other vertebrates
SRY	protein with an HMG domain	XY gonadal dysgenesis	loss of function	None (C. versicolor?)
WT1	Zn finger	WAGR syndrome Denys-Drash syndrome, Agenesis of gonads, kidney	Haplo-insufficiency, Dominant negative, loss of function	TSD-reptiles, birds
SF-1	orphan steroid receptor	Agenesis of urogenital system	loss of function	TSD-reptiles, birds
SOX9	SRY-like HMG domain	XY-sex reversal, Campomelic dysplasia,	loss of function	fish (Sox9a, b), reptiles, birds
DAX1/DSS	nuclear receptor	AHC, HGG XY SEX reversal	loss of function duplication	reptiles, birds
MIS	TGF- β gene family	persistance of Mullerian duct	loss of function	birds, TSD-reptiles
DMRT1	protein with dm-vdomain	XY-sex reversal	loss of function	fish, amophibia, reptiles, birds
Aromafase	enzyme in steroid biosynthesis	female to male sex reveral in birds & reptiles	loss of function	reptiles, birds
Wint4	signal transduction	XX, Leydig like cells	loss of function	—
Lim1, Lirn9	homebox	Lim-1-loss of gonad, kidney Lim9-gonadal agenesis	loss of function	—
Fgf9	growth factor	XY abnormal gen italia	loss of function	—
Vanin2	G-protein	XY-abnormal genitalia	loss of function	—

2.5 Suggested questions

1. Describe the molecular mechanisms that regulate *Sxl*- promoter activity and its maintenance.
2. With the help of diagrams, describe the events of alternative splicing that lead to sex determination in *Drosophila*.
3. Briefly describe the cascade of events that lead to sex determination in *C. elegans*.
4. Explain the role of *xol-J* in *C. elegans* sex determination.
5. State and explain the effect on sex determination and dosage compensation due to mutation in the following genes:
 - (a) Loss of function of *fox-1* in XX
 - (b) Gain of function of *Xol* in XO
 - (c) Gain of function of *Sxl* in XY
 - (d) Loss of function of *da* in XX and XY
 - (e) Gain of function of *Sdc* in XX
6. With the help of diagram briefly describe the initial pathway that regulates sex determination in Human.

Unit 3 □ Imprinting of Genes, Chromosomes and Genomes

Structure

3.1 Introduction

3.2 Genomic imprinting

3.3 Uniparental disomy and genomic imprinting

3.4 Suggested questions

3.1 Introduction

In humans and other mammals, several biallelic genes are known where the expression of one parental allele, either the paternal or the maternal allele but not both, is *normally* repressed in some cells (**allelic exclusion**). In such cells the relevant gene is said to exhibit functional hemizyosity; *even although the sequences of both parental alleles are perfectly consistent with normal gene expression and may even be identical*. In some cases the allelic exclusion may be a property of select cells or tissues while in other cells of the same individual both alleles may be expressed normally. A variety of different expression mechanisms can be involved and two broad classes of mechanism are involved :

- *Allelic exclusion according to parent of origin (imprinting)*. The choice of which of the two inherited copies is expressed is not random. This means that for some genes the allele whose expression is repressed is always the paternally inherited allele; in others it is always the maternally inherited allele.
 - *Allelic exclusion independent of parent of origin*. Here the decision as to which of the two alleles is repressed is initially made randomly, but afterwards that pattern of allelic exclusion is transmitted stably to daughter cells following cell division. A variety of different mechanisms may be involved. A unique form of control is the programmed DNA rearrangements.
-

3.2 Genomic imprinting

Genomic imprinting is an epigenetic phenomenon, which, in most cases, is believed to occur during gametogenesis. Genomic imprinting occurs when both maternal and paternal alleles are present, but one allele will be expressed while the other remains inactive. The most prominent assumption is that this process is necessary for development and may somehow regulate growth in the embryo and neonate. Some

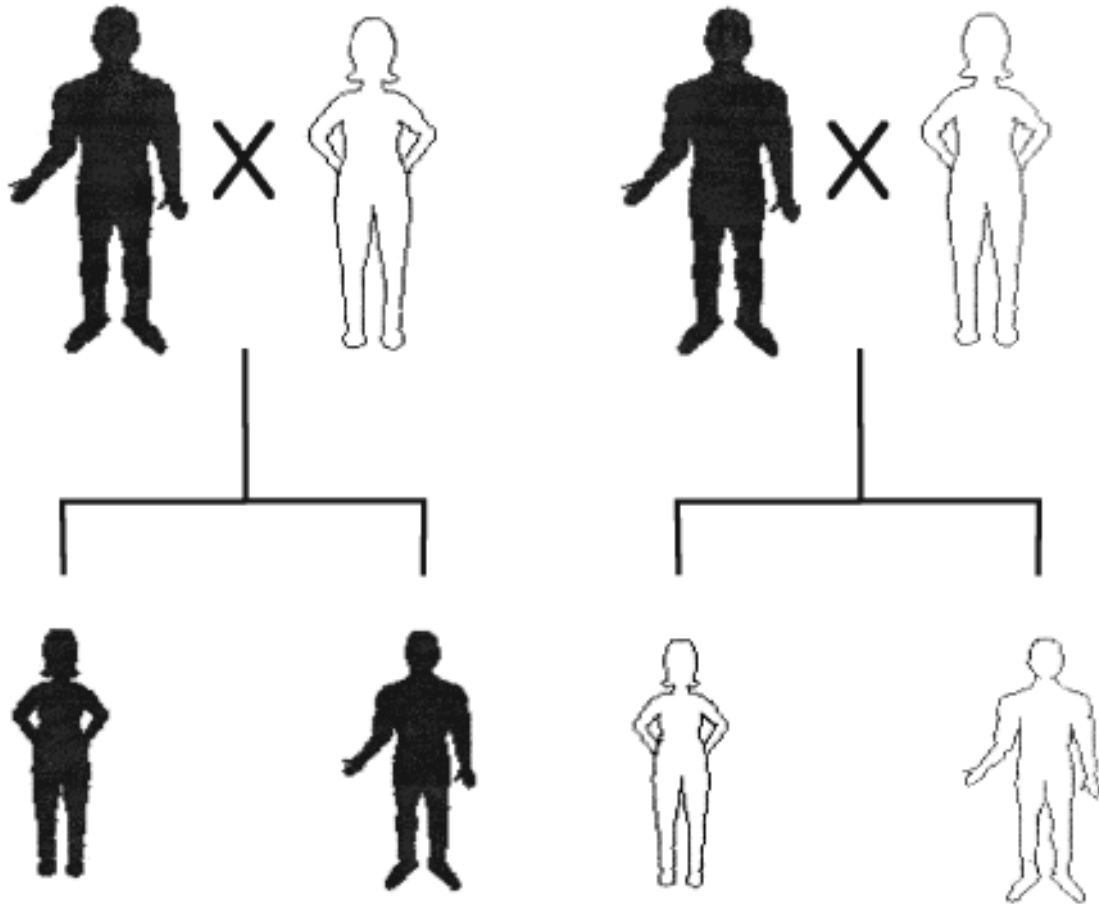


Fig. 3.1 Two examples of a hypothetical imprinted gene responsible for body color. (LEFT) In this example the pigment gene is maternally imprinted (maternal allele is inactivated). Matings between a male who possesses the allele for pigment and a female who possesses the allele for no pigment produces offspring that show only the pigmented phenotype. In this example, the mother's allele is imprinted and inactivated in the offspring. Therefore, the only actively-expressing allele is the father's pigment allele, which is not imprinted in the offspring. (RIGHT) In this example the pigment gene is paternally imprinted (paternal allele is inactivated). Matings between a male who possesses the allele for pigment and a female who possesses the allele for no pigment produces offspring that show only the pigmented phenotype. In this example, the father's allele is imprinted and inactivated in the offspring. Therefore, the only actively expressing allele is the mother's no pigment allele, which is not imprinted in the offspring. (Figure courtesy of Ross McGowan, Dept. Zoology, University of Manitoba)

mechanism must be able to distinguish between maternally and paternally inherited alleles: as chromosomes pass through the male and female germlines they must acquire some imprint to signal a difference between paternal and maternal alleles in the developing organism (Fig. 3.1).

An optimal method for gene imprinting, at least in maintaining the imprinted status, is allele-specific DNA methylation. The imprinting of several

imprinted genes has been shown to be disrupted in mutant mice that are deficient in the ***Dnmt 1*** cytosine methyltransferase gene and all imprinted genes are characterized by CG-rich regions of differential methylation. This process is carried out with the enzyme DNA methyltransferase (DNA MTase) in mammals. DNA MTase acts on the DNA sequence 5'-CpG-3'. Some organisms (primarily higher eukaryotes) have aggregates of CpG (known as CpG islands) in their genomes. These islands are rarely methylated in animal cells. This may be due to the bound transcription factors that block DNA MTase. De novo methylation and maintenance of methylation are two distinct processes that are required for establishment and mitotic inheritance of tissue specific methylation patterns. *Dnmt1* is the major maintenance methyltransferase. ***Dnmt3a*** and ***Dnmt3b*** are essential for de novo methylation. And Sequences that are methylated are usually not active (Gold and Pedersen, 1994). Recent investigations, however, have shown that this is not always the case (Li *et al.*, 1993).

It has been postulated that if a mutation was introduced to the DNA MTase gene in the embryonic stem cells of mice, the methylation of CpG would be abnormal, and gene expression would be affected. The mutation of the DNA MTase gene was caused by homologous recombination. The three genes used in this experiment were H19, Igf2 (insulin-like growth factor) and Igf2r (Igf2 receptor).

For the H19 gene, it is the maternal allele that is expressed, while the paternal allele is silent. It should be noted that the inactive paternal allele is methylated while the maternal allele is not. It was shown that typical DNA methylation is a requirement to keep the paternal allele inactive for the H19 gene, a result that is consistent with the hypothesis.

In contrast to the H19 gene, the Igf2 gene is expressed only from a methylated paternal allele. It has now been concluded that a normal level of DNA methylation is needed for expression of the paternal Igf2 allele.

The gene Igf2r (insulin like growth factor receptor) is expressed from a methylated maternal allele. DNA methylation by DNA MTase is a requirement for the expression of the Igf2r gene.

Disruption of primary imprinting during oocyte growth leads to the modified expression of imprinted genes during embryogenesis. Thus far, experiments have not demonstrated how the imprinting process is regulated. It has been proposed that further research should be attempted to discover and isolate imprinting genes. Already, some progress is being made in these areas. For instance, genomic imprinting has been implicated in cancer and shown to be involved in chromosomal deletion syndromes, such as Prader-Willi and Angelman Syndromes (Peterson and Sapienza, 1993).

3.3 Uniparental disomy and genomic imprinting

Uniparental disomy refers to the presence of two copies of a chromosome (or part of a chromosome) from one parent and none from the other. Several additional disorders resulting from uniparental disomy of single genes or multiple genes (including whole chromosomes) have been reported. A readily detectable adverse outcome of uniparental disomy is the consequence of a newly recognized phenomenon called *genetic or genomic imprinting*.

The first recognized example of such human abnormality resulting from the presence of uniparental disomy of an imprinted part of the genome was in Prader-Willi syndrome (PWS). Uniparental maternal disomy for chromosome 15

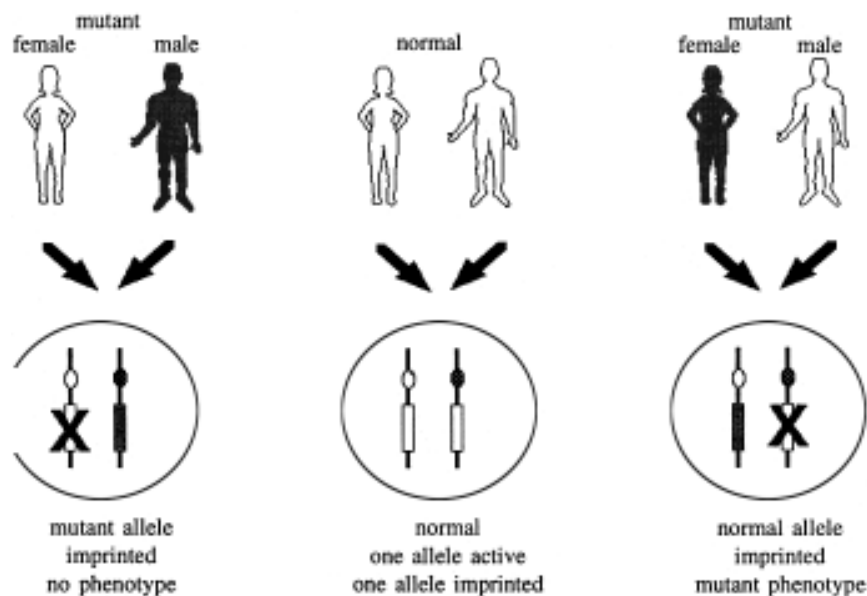


Figure 3.2 Schematic representation of the phenotypic effects of maternal imprinting of a mutant allele. Darkened body indicates individual that is mutant for the hypothetical imprinted locus. A cross is used to indicate the imprinted/inactive allele. (CENTER) Both parents are homozygous for the normal allele at the imprinted locus. Although only one allele is active (the paternal copy) in the offspring produced from these parents, it must be a normal allele and therefore all offspring will have a normal phenotype. (LEFT) The mother is homozygous mutant at the imprinted locus, and the father is normal. Since this hypothetical locus is maternally imprinted, the maternal mutant copy will be inactivated in their offspring and the paternal normal copy will be the only active allele. The offspring will be phenotypically normal, and the mutant allele will appear to be a recessive mutation. (RIGHT) The mother is homozygous normal at the imprinted locus, and the father is homozygous mutant. The maternal normal allele is imprinted and inactivated in the offspring of these parents. The only allele that is active is the mutant paternal copy. Therefore, all offspring produced from these parents will display the mutant phenotype, and the mutant allele will appear to be a dominant mutation. (Figure courtesy of Ross McGowan, Dept. Zoology, University of Manitoba)

is thought to cause Prader-Willi syndrome because there is absence of needed paternally contributed genes in the critical PWS region (del 15q 11 -q 13). The paternal contribution is hypothesized to be necessary because the homologous maternally derived genes are inactivated or imprinted (perhaps by methylation).

Interestingly, a very different disorder called Angelman syndrome also involves imprinting of the same chromosome region - only in Angelman syndrome the maternal contribution of the critical region is missing. The terminology used to describe the role of imprinting in these two disorders is somewhat confusing but goes as follows.

It is hypothesized that the critical genetic region which determines Prader-Willi syndrome is *maternally* imprinted (i.e. inactivated when inherited from the mother), whereas the critical region which determines Angelman syndrome is *paternally* imprinted (i.e. inactivated when inherited from the father). Both disorders result when the expected active genetic contribution from one parent is missing, either by deletion or uniparental disomy (Fig.3.2).

Interestingly, a number of human congenital tumors show evidence of genomic imprinting. For example, in cells from Wilms' tumor, loss of the maternal chromosome 11 is common. This suggests that the maternal chromosome 11 has a tumor suppressor role not present on the paternal 11. This phenomenon in relation to cancer is referred to as "loss of heterozygosity".

3.4 Suggested questions

1. How is genomic imprinting established and maintained?
2. Explain the phenomena and importance of uniparental disomy.
3. Differentiate paternal and maternal imprinting.

Unit 4 □ Somatic Cell Genetics

Structure

- 4.1 Cell fusion and hybrids-agents and mechanisms of fusion
- 4.2 Heterokaryon-selecting hybrids and chromosome segregation
- 4.3 Radiation hybrids, hybrid panels and gene mapping
- 4.4 Suggested questions
- 4.5 Suggested books

4.1 Cell fusion and hybrids-agents and mechanisms of fusion

Cultured animal cells infrequently undergo cell fusion spontaneously. The fusion rate, increases greatly in the presence of certain viruses that have a lipoprotein envelope similar to the plasma membrane of animal cells. Cell fusion is also promoted by polyethylene glycol, which causes the plasma membranes of adjacent cells to adhere to each other and to fuse. As most fused animal cells undergo cell division, the nuclei eventually fuse, producing viable cells with a single nucleus that contains chromosomes from both “parents.” The fusion of two cells that are genetically different yields a hybrid cell called a **heterokaryon** (Fig. 4.1).

Because some **somatic cell** from animals can be cultured from single cells in a well-defined medium, it is possible to select for genetically distinct cultured animal cells. Genetic studies of cultured animal cells are called *somatic-cell genetics* to distinguish them from *classical genetics*, which deals with whole organisms derived from **germ cells** (sperm and eggs).

Assigning genes to chromosomes

The technique of somatic cell hybridization is extensively used in human genome mapping, but it can in principle be used in many different animal systems. The procedure uses cells growing in culture. A virus called the Sendai virus has a useful property that makes the mapping technique possible.

If suspensions of human and mouse cells are mixed together in the presence of Sendai virus that has been inactivated by ultraviolet light, the virus can mediate fusion of the cells from the different species. The initial fusion products are described as heterokaryons because the cells contain both a human and a rodent nucleus. Eventually, heterokaryons proceed to mitosis, and the two nuclear

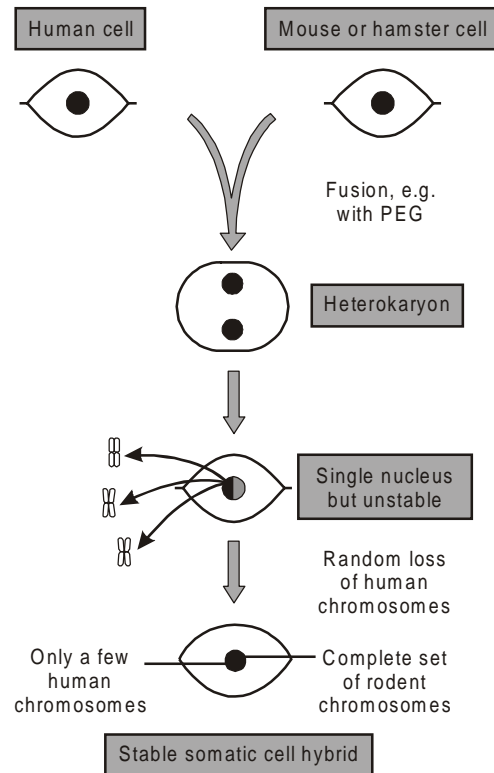


Fig. 4.1 Fusion of cells from different species can result in stable somatic cell hybrids. The example shows how stable human-rodent somatic cell hybrids can be generated following initial fusion using polyethylene glycol (PEG). For reasons that are not understood, human chromosomes are selectively lost from the initial fusion products. The loss occurs essentially at random so that eventually the stable products of a single fusion experiment will include a variety of cells with different complements of human chromosomes. They can be cloned to establish individual cell lines with a specific complement of human chromosomes. The identity of the human chromosomes can be established by PCR-based typing for chromosome-specific markers

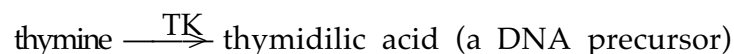
envelopes dissolve. Because the mouse and human chromosomes are recognizably different in number and shape, the two sets in the hybrid cells can be readily distinguished. However, in the course of subsequent cell divisions, for unknown reasons the human chromosomes are gradually eliminated from the hybrid at random.

The loss of human chromosomes can be arrested in the following way to encourage the formation of a stable partial hybrid. The cells used are mutant for some biochemical function; so, if the cells are to grow, the missing function must be supplied by the other genome. This selective technique results in the maintenance of hybrid cells that have a complete set of mouse chromosomes

and a small number of human chromosomes, which vary in number and type from hybrid to hybrid but which always include the human chromosome carrying the wild-type allele defective in the mouse genome.

4.2 Heterokaryon-selecting hybrids and chromosome segregation

In cells, DNA can be made either de novo (“from scratch”) or through a salvage pathway that uses molecular skeletons already available. The selective technique involves the application of a chemical, aminopterin that blocks the de novo synthetic pathway, confining DNA synthesis to the salvage pathway. Two essential salvage enzymes, thymidine kinase (TK) and hypoxanthine-guanine phosphoribosyl transferase (HGPRT), are relevant to the system, as shown in the following two reactions :



The mouse cell line to be fused is genetically unable to make TK because it is homozygous for the allele *tk⁻*, whereas the human cell line is genetically unable to make HGPRT because it is homozygous at another locus for the allele *hgprt⁻*. So the genotypes of the two fusing cell lines are :

Mouse : *tk⁻ / tk⁻; hgprt⁺ / hgprt⁺*

Human : *tk⁺ / tk⁺; hgprt⁻ / hgprt⁻*

Because each is deficient for one enzyme, neither the mouse nor the human cells are able to make DNA individually. In the hybrid cells, however, the *tk⁺* allele complements the *hgprt⁺* allele, so the cells can make both enzymes. Therefore, DNA is synthesized and the cells can proliferate. Most human chromosomes are eliminated from the hybrid cell cultures because their loss has no effect on the cultures’ ability to grow. But, to continue to grow in medium containing hypoxanthine, aminopterin, and thymidine (*HAT medium*), a hybrid culture must retain at least one of the human chromosomes that carry the *tk⁺* allele.

4.2.1 Selecting for the chromosome contents of hybrids

Hybrids can be selected for retention of a given human chromosome or chromosome fragments if it corrects an otherwise lethal abnormality in the rodent cell. Frequently used systems include :

- **Hybrid cells often are selected on HAT medium :** The medium most often used to select hybrid cells is called HAT *medium*, because it contains hypoxanthine (a purine), aminopterin, and thymidine. Normal cells can grow in HAT medium because even though aminopterin blocks de novo synthesis of purines and TMP, the thymidine in the media is transported into the cell and converted to TMP by TK and the hypoxanthine is transported and converted into usable purines by HGPRT. On the other hand, neither TK nor HGPRT cells can grow in HAT medium because each lacks an enzyme of the salvage pathway. However, hybrids formed by fusion of these two mutants will carry a normal TK gene from the HGPRT parent and a normal HGPRT gene from the TK parent. The hybrids thus will produce both functional salvage-pathway enzymes and grow on HAT medium. Likewise, hybrids formed by fusion of mutant cells and normal cells can grow in HAT medium. Somatic cell hybrids can be forced to retain human chromosome 17 by using thymidine kinase deficient (TK) rodent cells and growing the hybrids in HAT (hypoxanthine-aminopterin-thymidine) medium.
- **G418 selection :** Hybrids can be selected for the presence of a particular human chromosome segment if it has been tagged by incorporation of a neomycin resistance (*neo^R*) gene. The neomycin analog G418 kills nonresistant cells. Neo^R is a typical example of a dominant selectable marker.

Selecting for the chromosome contents of hybrids

Hybrids can be selected for retention of a given human chromosome or chromosome fragments if it corrects an otherwise lethal abnormality in the rodent cell. Frequently used systems include :

- **HAT selection :** Somatic cell hybrids can be forced to retain human chromosome 17 by using thymidine kinase deficient (TK) rodent cells and growing the hybrids in HAT (hypoxanthine-aminopterin-thymidine) medium. TK cells are killed in HAT medium, but are rescued by the human TK gene on chromosome 17.
- **G418 selection :** Hybrids can be selected for the presence of a particular human chromosome segment if it has been tagged by incorporation of a neomycin resistance (*neo^R*) gene. The neomycin analog G418 kills nonresistant cells. Neo^R is a typical example of a dominant selectable marker.

4.2.2 Somatic cell hybrid panels can permit chromosomal localization of any human DNA sequence

The human chromosomes in somatic cell hybrids can conveniently be identified by PCR screening with sets of chromosome-specific primers. By

collecting hybrid cell lines with different human chromosome contents it is possible to generate a hybrid cell panel that can be used to map any human DNA sequence to a specific chromosome. To do this, each of the hybrid cell lines is tested for the presence of the human sequence of interest. A PCR assay can be used with primers specific for that sequence or the relevant DNA sequence can be labeled and used as a hybridization probe.

4.3 Radiation hybrids, hybrid panels and gene mapping

Subchromosomal mapping is possible using hybrid cells containing defined portions of a human chromosome

Conventional somatic cell hybrids are a relatively crude tool for physical mapping. More refined mapping is possible using hybrids that contain only part of a particular human chromosome. Translocation hybrids and deletion hybrids are made using donor human cells that have a chromosomal translocation or deletion. To be useful, the hybrids must lack the normal homolog of the chromosome of interest. Such hybrids can be used for subchromosomal mapping of a human sequence-tagged site or biochemical marker. (Fig. 4.2). They are especially useful for defining the sequences removed by microdeletions, by segregating the deletion-carrying chromosome away from its normal homolog.

Chromosome-mediated gene transfer

One of the first techniques to use this approach was **chromosome-mediated gene transfer (CMGT)**. Fragments of purified mitotic chromosomes from a donor, such as a human fibroblast, are coprecipitated with calcium phosphate on to the surface of a recipient rodent cell line in monolayer culture. Human chromosome fragments enter the recipient cells, such as mouse fibroblasts, and integrate into the chromosomes, resulting in stable transformation. As a result, hybrids can be established that retain segments of human DNA (**transgenomes**) of a size that is useful for mapping (usually in the range of 1-50 Mb). However, the transgenomes are prone to frequent rearrangements, so CMGT is more suited to functional assays of complex loci than as a mapping tool.

Irradiation fusion gene transfer

The most valuable hybrids for gene mapping are radiation hybrids. Donor cells are subjected to a lethal dose of radiation, which fragments their chromosomes. The average size of a fragment is a function of the dose of radiation. After irradiation the donor cells are fused with recipient cells of a different species. A selection system is used to pick out recipient cells that have taken

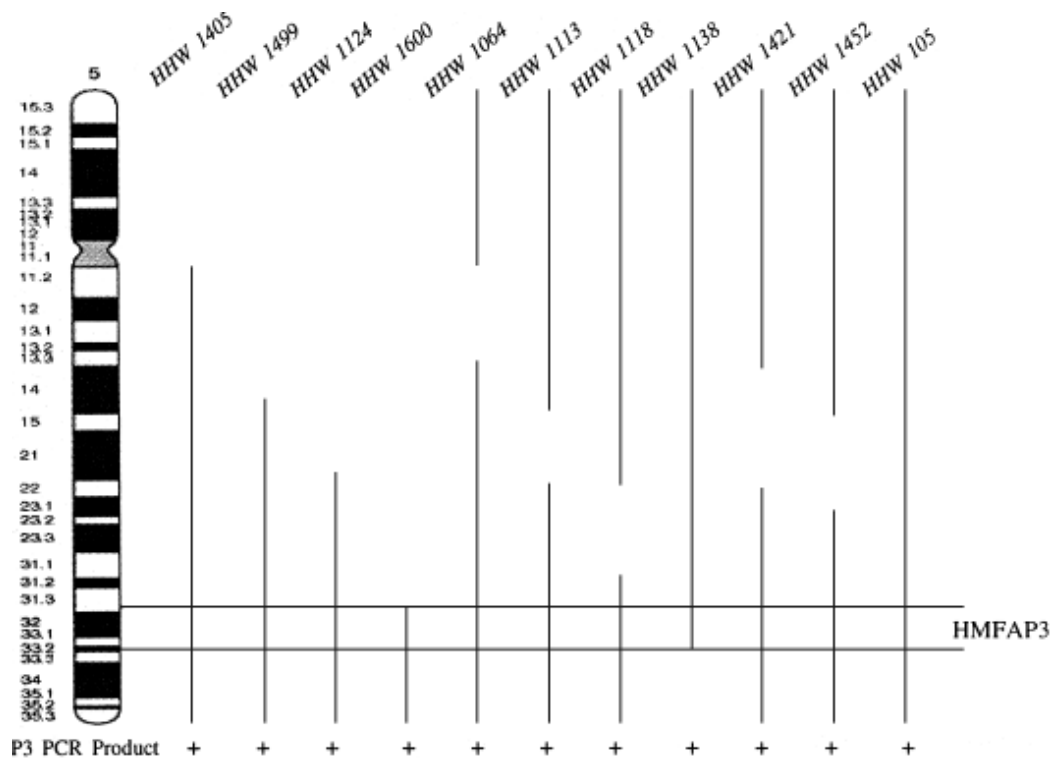


Fig. 4.2 Subchromosomal localization can be achieved by mapping against a panel of hybrid cells containing translocation or deletion chromosomes. The figure illustrates PCR-based mapping of the human microfibrillar protein MFAP3 using a panel of 5q translocation and deletion hybrids. Vertical black bars to the right indicate the extent of human chromosome 5 sequences, which are retained in the hybrids. Hybrids HHW1405, 1499, 1124 and 1600 contain translocation chromosomes with 5q breakpoints and retention of the segment distal to the breakpoint. By contrast, translocation hybrid HHW1138 retains material proximal to the 5q breakpoint. Hybrids HHW1064, 1113, 1118, 1421 and 1452 have different interstitial deletions of 5q. The solid blue vertical bar to the left indicates the inferred subchromosomal location as defined by breakpoints in hybrids HHW1600 and HHW1138 (blue horizontal lines near bottom). Reproduced from **Abrams et al. (1995)** *Genomics*, **26**, pp. 47-54, with permission from Academic Press, Inc

up some of the donor chromosome fragments. These cells are useful for mapping insofar as they have taken up a random set of other chromosome fragments from the donor, as well as the selected fragment. Stably incorporated donor fragments are either integrated into rodent chromosomes or are assembled into novel human minichromosomes formed around fragments containing a functional centromere. When a set of DNA markers from the human chromosome is assayed in a panel of such radiation hybrids, the patterns of cross-reactivity can be used to construct a map. (Fig. 4.3).

The principle is very similar to meiotic linkage analysis : the nearer together two DNA sequences are on a chromosome, the lower the probability that they

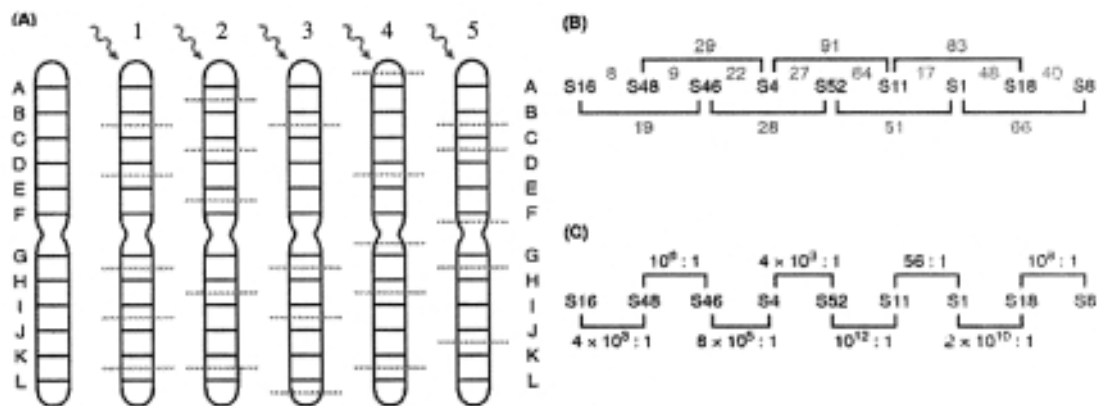


Fig. 4.3 Constructing radiation hybrid maps. (A) Breakpoints occur randomly. Five possible examples of breakpoints (dashed blue lines) on the same type of chromosome are shown. Markers close together will tend to occur on the same fragment, e.g. A and B in all cases other than example 2. Thus, if a radiation hybrid contains marker A it will frequently also contain marker B, but rarely a distant marker such as L. (B) Ordering of markers on human 21q. The order of markers *D21S16-D21S8* as inferred by *Cox et al* (1990) from radiation hybrid mapping is shown. Figures on the top panel refer to distances between markers in centiRays8000. For example, the SO 6-S48 interval is 8 cR8000: at a radiation dose of 8000 rad, there is 8% frequency of breakage between them, and so a 92% chance they will occur together on one fragment. (C) Odds ratios refer to the likelihood of the indicated order for pairs of markers compared with that with the markers inverted. For example, the calculated likelihood for the order S16-S48-S46-S4 is 106 times greater than for the order S16-S46-S48-S4

will be separated by the chance occurrence of a breakpoint between them. The frequency of breakage between two markers can be defined by a value, analogous to the recombination frequency in meiotic mapping. The value, varies from 0 (the two markers are never separated) to 1.0 (the two markers are always broken apart). As in meiotic mapping, the value underestimates the distance between markers that are far apart on the same chromosome, in this case because a cell can take up two markers on separate fragments. A more accurate estimate is provided by a mapping function, $D = -\ln(1 - \text{value})$, which is analogous to the Haldane mapping function used in meiotic linkage analysis. D is measured in centiRays (**cR**). D is dependent on the dosage of radiation, so it is referenced against the number of rads. For example, a distance of 1 cR₈₀₀₀ between two markers represents a 1% frequency of breakage between them after exposure to 8000 rad of X-rays.

Radiation hybrids derived from monochromosomal hybrid donor cells have been superseded by whole-genome radiation hybrids where the donor is an irradiated normal human diploid cell. The first such panel consisted of 199 hybrids made by fusing an irradiated 46,XY human fibroblast cell line to TK

hamster cells (Walter *et al.*, >1994), Gyapay *et al.* (1996) used 404 microsatellite markers of known location to show that this hybrid panel could generate accurate maps, and then used it to map 374 unmapped ESTs. A subset of 93 of the hybrids has been made widely available as the Genebridge 4 panel. The 93 hybrids average 32% retention of any particular human sequence, with an average fragment size of 25 Mb. Laboratories can map any unknown STS by scoring the 93 Genebridge hybrids and comparing the pattern with patterns of previously mapped markers held on a central server.

This has turned into an extremely powerful and convenient tool for physically mapping any STS or EST. A second human-hamster panel, Stanford G3, was made using a higher dose of radiation, so that the average human fragment size is smaller. The 83 hybrids in G3 average 16% retention of the human genome, with an average fragment size of 2.4 Mb. Thus G3 can be used for finer mapping. The impressive results of large-scale use of these panels can be accessed at <http://www.ncbi.nlm.nih.gov/genemap98/>.

4.4 Suggested questions

1. What is heterokaryon? What are the methods of heterokaryon selection?
2. Explain the principle of mapping by radiation hybrids.
3. How can genes be assigned to specific regions on chromosomes by low-resolution mapping (radiation hybrids)?

4.5 Suggested books

1. Alberts, B. *et al.* (2003) *Molecular Biology of the Cell*, Fourth edition; Garland Sciences, New York.
2. Lewin, B. (2004) *Genes VIII*, John Wiley, New York.
3. Klug, W, and Cummings, M. (2003) *Concepts of Genetics*; Seventh Edition; Pearson Education, Singapore.
4. Karp, G. (2005) *Cell and Molecular Biology*, Fourth Edition; John Wiley, New York.
5. Gilbert, S. (1997) *Developmental Biology*, Fifth Edition, Sinauer Associates Publishers, Massachusetts.
6. Russell, P. (1998) *Genetics*, Fifth Edition; Addison Wesley Longman, New York.
7. Strachan, T. and Read, A. P. (2004) *Human Molecular Genetics*, Third Edition; Oxford University Press, Oxford.

Unit 5 □ Human Cytogenetics

Structure

- 5.1 Techniques in human chromosome analysis—molecular cytogenetic approach
- 5.2 Human karyotype—banding—nomenclature
- 5.3 Numerical and structural abnormalities of human chromosomes—syndromes.
- 5.4 Human genome

5.1 Techniques in human chromosome analysis

5.1.1 Introduction

The correct chromosome number for man was established only after the *application of tissue culture methods* to cytogenetics. Before 1956, the chromosome number of man was considered to be 48. Tjio and Levan (1956), analyzing cultures of human embryonic lung fibroblasts, found a consistent chromosomal number of 46. At the same time and independently Ford and Hamerton (1956), using meiotic cells obtained from testicular biopsy material, found only 23 pairs of chromosomes. Mitotic cells from the same specimens contained 46 chromosomes. These two papers mark the beginning of modern human cytogenetics. The 15-year period from 1956 to 1971 saw the development of a standardized system of nomenclature which became more refined as the identification of human chromosomes became more precise. The delineation of most of the syndromes associated with chromosomal abnormalities occurred during this period.

5.1.2 Terminologies used in the identification of human chromosomes karyotype & idogram

Human chromosomes can be arranged in an orderly fashion to produce a *karyotype* (Rowley, 1969). A karyotype is composed of individual chromosomes from a particular cell; the chromosomes are aligned in pairs and identified according to the standard nomenclature accepted by cytogeneticists. Karyotypes of different cells will reflect the variations in chromosomal morphology present in these cells. An *idiogram* is the schematized drawing of a composite of many karyotypes and is not directly related to any particular cell. Idiograms generally are not prepared for clinical purposes. But are used for comparing chromosomal patterns of different species.

Metaphase chromosome

Metaphase chromosomes differ from one another in size and shape (Fig. 5.1) Each metaphase chromosome is identified by its size, shape, and specific banding pattern. The absolute size of any chromosome varies with the stage of mitosis. Chromosomes are longer and less coiled in prophase and shorter and more compact at the end of metaphase. The duration of treatment with mitotic blocking agents and the type of hypotonic solution also influence the absolute size of the chromosomes. In general, the longest human metaphase chromosome is about 7-8 μm in length, whereas the shortest is about 2 μm long. Each metaphase chromosome is composed of two *chromatids* joined at the *centromere* (the site of attachment of the spindle fibre). The position of the centromere is specific for

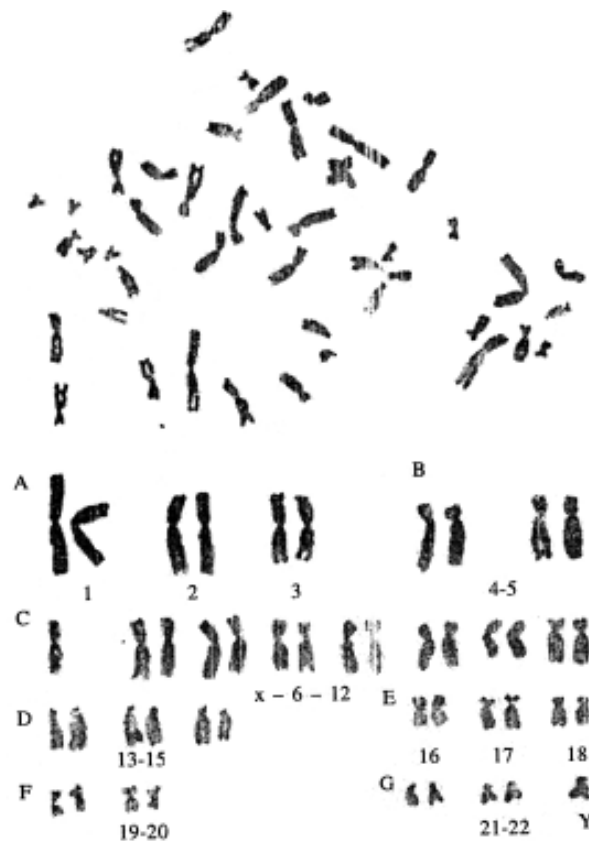


Fig. 5.1 Intact metaphase plate from a normal male with 46 chromosomes from bone marrow specimen (top) and karyotype of cell (bottom). Chromosomes are arranged in seven groups of morphologically similar chromosomes. The X chromosome is included with group C from which it cannot be distinguished. The Y chromosome in this cell is about the size of G-group chromosomes, but it can be differentiated from them by the size of the short arm

each chromosome and divides it into a long and short arm. The relative length of the two arms (arm ratio) is important for the identification of chromosomes.

A chromosome with a centromere in the middle that divides it into two equal arms is called *metacentric*. When the centromere is somewhat nearer to one end of the chromosome, so that there is a distinct long and short arm, the chromosome is said to be submetacentric. If the centromere is very near to one end of the chromosome, which thus only a very short small arm and a relatively longer long arm, the chromosome is called *acrocentric* (Fig. 5.2). A *telocentric* chromosome, not normally found in human cells, has the centromere at the end.

5.1.3 Standardization of nomenclature of chromosome

Four conferences were held between 1960 and 1971 to revise the nomenclature used to identify human chromosomes along with the improvement of technologies. The last two conferences were held in Chicago and Paris. When the first reports on human chromosomal abnormalities appeared in 1956, each group of investigators used its own system of arranging and numbering the chromosomes. The necessity for adopting a uniform system was generally recognized, and a standard nomenclature of human mitotic chromosomes was adopted at the Denver Conference in 1960. Numbers were assigned to each pair of autosomes as nearly as possible in descending order of length. The sex chromosomes, XX pairs of autosomes (44) plus two sex chromosomes, making a total of 46 chromosomes. Seven groups of morphologically similar chromosomes could be distinguished (Table 5.1). It soon became evident, however, that not all pairs of chromosomes could be identified with certainty, even in preparations of the highest technical quality.

Patau (1960) proposed that the seven groups of morphologically similar chromosomes be identified by capital letters A through G, an arabic number being added only when the individual chromosome could be identified with certainty. This system had the great advantage of flexibility, since it permitted general recognition of the group to which a chromosome belonged without implying identification of the specific chromosome involved in an abnormality. Patau's recommendations were accepted at the London Conference on the normal human karyotype (1963).

Whereas individual investigators were using the same systems (Denver and London) for identifying individual chromosomes, they were using different systems for describing chromosomal abnormalities. This led to confusion when data from different laboratories were collated. A major achievement of the Chicago Conference in 1966 was the development of a uniform system of notation designed to facilitate coding for data retrieval. It was agreed that the analysis of the karyotype would be recorded with the total chromosomal number first, followed

by the sex chromosomes, and finally by any additional abnormalities. Thus the karyotype of a normal male was written 46, XY; a normal female was 46, XX. The recommended nomenclature symbols used in describing normal or abnormal chromosomes are summarized in Table 5.2.

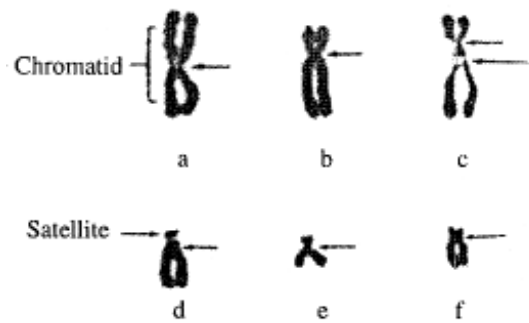


Fig. 5.2 Centromere position in typical chromosomes, a, Two chromatids joined at centromere or primary constriction. Centromere (arrow) is median and divides chromosomes in two equal arms. Chromosome is metacentric. b, Submedian centromere (arrow) divides the chromosome into short arm and long arm. Chromosome is submetacentric. c, Submetacentric chromosome with secondary constriction (long arrow) in long arm. d, Large acrocentric chromosome with subterminal centromere (short arrow). Satellites (long arrow) are separated from short arm by secondary constriction, e and f, G-group chromosome and Y from same cell. Long arm chromatids of Y are close together and short arm is larger than G-group chromosome

Table 5.1 Systems of identification of human metaphase chromosomes*

Denver			Chromosome description	
Patau		Chromosome	Centromere	
group	Group	number	position	Morphology
A	1-3	1,3	Median	Metacentric
		2	Submedian	Less metacentric than above
B	4-5	4,5	Submedian	Submetacentric
C	X 6-12	X,6,7,9,11	Submedian	Submetacentric but more metacentric than remainder
		8,10,12	Submedian	Less metacentric than above
D	13-15	13,14,15	Subterminal	Acrocentric—all may have satellites
E	16-18	16	Median	Metacentric
		17,18	Submedian	Submetacentric
F	19-20	19,20	Median	Metacentric
G	21-22	21,22	Subterminal	Acrocentric—all may have satellites
		Y	Subterminal	Acrocentric—similar in size to group G but usually morphologically distinct; long arms tend to be close together: satellites are not present

*Rowley (1969)

Table 5.1 Nomenclature symbols

		<i>Chicago Conference</i>
A—G		the chromosome groups
1—22		the autosome numbers
X,Y		the sex chromosomes
diagonal(/)		separates cell lines in describing mosaicism
?		questionable identification of chromosome or chromosome structure
*		chromosome explained in text or footnote
ace		acentric
cen		centromere
dic		dicentric
end		endoreduplication
h		secondary constriction or negatively staining region
i		isochromosome
inv		inversion
mar		marker chromosome
mat		maternal origin
p		short arm chromosome
pat		paternal origin
q		long arm of chromosome
r		ring chromosome
s		satellite
t		translocadon
s t repeated symbols		duplication of chromosome structure

*Paris Conference***A. Recommended changes in Chicago Conference nomenclature**

- +
1. The + and — signs should be placed *before* the appropriate symbol where they mean additional or missing whole chromosomes. They should be placed *after* a symbol where an increase or decrease in length is meant. Increases or decreases in the length of secondary constriction, or negatively staining regions, should be distinguished from increases or decreases in length owing to other structural alterations by placing the symbol h between the symbol for the arm and the + or — sign (e.g., 16qth +).
 2. All symbols for rearrangements are to be placed before the designation of the chromosome (s) involved in the rearrangement, and the rearranged chromosome (s) always should be placed in parentheses, e.g., r(18), i (Xq), die (Y).

B. Recommended additional nomenclature symbols

del	deletion
der	derivative chromosome
dup	duplication
ins	insertion
inv ins	inverted insertion
rep	reciprocal translocation*
rec	recombinant chromosome
rob	Robertsonian translocation* (“centric fusion”)

tan	tandem translocation*
ter	terminal or end (“p ter” for end of short arm; “q ter” for end of long arm)
:	break (no reunion, as in a terminal deletion)
::	break and join
→	from-to

* Optional, where greater precision is desired than that provided by the use of t as recommended by the Chicago Conference.

5.1.4 Recommendations of the Paris Conference

The fluorescent karyotype, published by Caspersson et al (1971), was accepted as the basis for the assigning of numbers to each chromosome.

A. Definitions

The bands seen with the fluorescent dyes (quinacrine) were called Q-bands and were accepted as the reference bands. Those bands of chromatin stained by methods that demonstrate “constitutive heterochromatin” were called C-bands and they are mainly confined to the centromeric region. The bands stained with basic dyes such as Giemsa were called G-bands and, except in one technique, they correspond quite well with Q-bands. One of the techniques using Giemsa, the exception just noted, gives patterns that are opposite in intensity to the G-bands; these were called R-bands.

A *band* was defined as a part of a chromosome which is clearly distinguishable from its adjacent segments by appearing darker or lighter with the Q, G, R, or C staining methods. By definition there were no “interbands.” In the construction of the chromosome map, each band was referred to by its midline and not by its margins. A chromosome *landmark* was defined as a consistent and distinct morphological feature that is an important diagnostic aid in identifying a chromosome. A region was defined as any area of a chromosome lying between two adjacent landmarks. A chromosome arm lacking a prominent landmark consists of only one region.

B. Band numbering

Regions and bands are numbered consecutively from the centromere outward along each chromosome arm (Fig. 5.3). A band used as a landmark is considered as belonging entirely to the region distal to the landmark and is accorded the band number of “1” in the region. A band bisected by the centromere is considered as two bands, each being labeled as band 1, in region 1, of the appropriate chromosome arm.

For the designation of a particular band, four items are required: the chromosome number, the arm symbol, the region number and the band number within that region. These items are given in order without spacing or punctuation.

For example, Ip33 indicates chromosome No. 1, short arm, region 3, band 3. If a band defined in the present chromosome map has to be subdivided, the original band designation will be followed by a decimal point and the sub-bands will then be numbered sequentially from the centromere outward, e.g., Ip33.1; Ip33.2; Ip33.3, indicating that the original band 33 in the short arm of chromosome No. 1 has been divided into three sub-bands, 33.1 being proximal and 33.3 distal to the centromere. This system is thus relatively simple and yet sufficiently flexible to accommodate further refinements as the banding techniques are improved.

C. Characterization of chromosomes by the various banding techniques

The technical quality of the chromosomes is of the utmost importance for characterizing chromosomes by their banding patterns; in fact, it is much more important than previously adopted methods when morphology was the sole criterion for preparing the karyotype. The most distinct banding patterns are obtained when the chromosomes are relatively long and free of overlaps. It is important to note that the morphological characteristics of size and centromere position remain critical parameters used in the identification of chromosomes. Thus, the cytogeneticist uses banding patterns as well as overall morphology to distinguish individual chromosomes. In general, the number of distinct bands increases with increasing length of the chromosome. It is thus, meaningless to mention the absolute number of bands in a chromosome arm, since the number varies with the state of contraction, the quality of the preparation and the type of treatment and stain. Once identification of the individual chromosomes by means of the major landmarks is mastered, careful observation of very good preparations can reveal a number of fine bands which can be used, among other things, for defining the site of chromosomal breakage and rejoining.

The intensity of fluorescence is influenced by the position of the chromosome in the metaphase. Chromosomes in the center of the cell frequently fluoresce more brightly than homologous chromosomes that are on the periphery. It is therefore necessary to consider the *pattern* of the bands in a particular chromosome and to appreciate the fact that the overall intensity of fluorescence of homologs may be different. There may also be "spreading" of fluorescence from brighter to duller chromosomes, such that a 19 or a 22 adjacent to an X or the long arm of the Y may appear to be much brighter than normal.

The terms "distal" and "proximal" refer to the position of a band with respect to the centromere. The following terms were used in the Paris report to indicate the approximate intensity of fluorescence:

negative	no or almost no fluorescence
pale	as on distal lp
medium	as the two broad bands on 9q
intense	as the distal half of 13q
brilliant	as on distal Yq

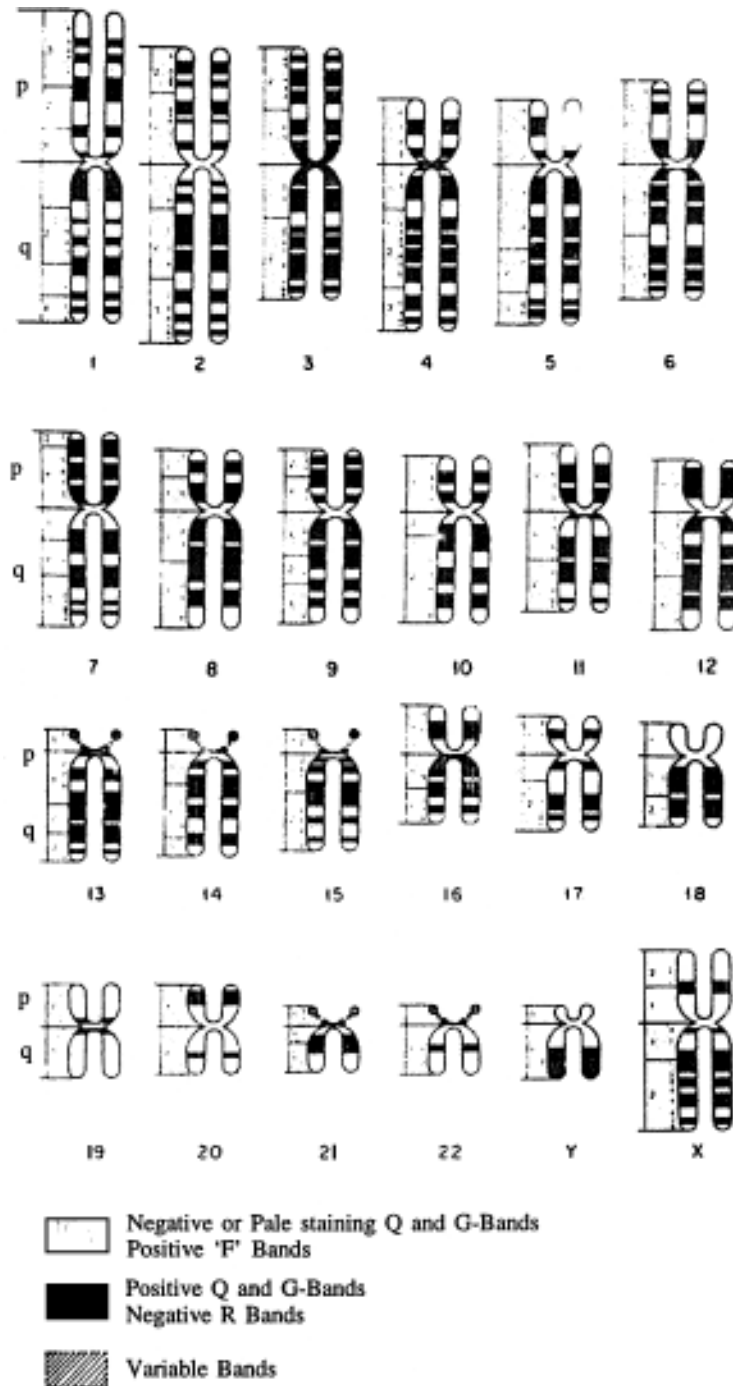


Fig. 5.3 Diagrammatic representation of chromosome bands as observed with the Q-, G-, and R-staining methods; centromere representative of Q-staining method only. Reproduced from the report of The Paris Conference (1972)

D. Characterization of chromosomes on the basis of different banding techniques :

Chromosome 1 is the largest chromosome and is usually metacentric. The distal 40% of the short arm shows pale fluorescent bands (32 to 36), and the proximal segment shows two bands of medium fluorescence; the more proximal band divides region 1 from 2, and the more distal band divides region 2 from 3 (Fig. 5.4). The area of the secondary constriction, a poorly staining gap, which by definition is in the long arm, is adjacent to the centromere and shows negative fluorescence; it constitutes region 1. The long arm also contains a central intense band which divides region 2 from 3, with a less intense band distal to it which divides region 3 from 4.



Fig. 5.4 Idiogram of fluorescent bands in human chromosomes (see text)

The large block of densely stained material in the long arm adjacent to the centromere is the most prominent feature of No. 1 in cells treated to produce G-bands (Fig. 5.5). This area corresponds to the negatively quinacrine-stained secondary constriction. The two proximal bands in the short arm and the very

darkly stained central band and less darkly stained distal band in the long arm are also present. The end of the short arm is faintly stained.

The technique for staining C-bands reveals the same densely staining region of the secondary constriction, generally called "constitutive heterchromatin." as do the G-band techniques. The R-banding technique (Dutrillaux and Lejeune, 1971) demonstrates moderately staining material in this region; elsewhere along the chromosome, however, dark bands appear pale and vice versa. This reversal of staining intensity as compared with G-bands is particularly noticeable at the end of the short arm.



Fig. 5.5 Idiogram of Giemsa bands in human chromosomes (see text)

Chromosome 2 is a large chromosome, less metacentric than No. 1, and lacking conspicuous landmarks. It shows a number of bands of medium fluorescence along the entire length; the central band (14 to 16) in the short arm and the two central bands (22 to 24) of equal intensity in the long arm are most prominent. In the short arm, the dull band distal to the central band divides

region 1 from 2. In the long arm, the dull band proximal to the proximal central band divides region 1 from 2, whereas the dull band distal to distal central band divides region 2 from 3.

Chromosome 3 is a large metacentric chromosome which is smaller than No. 1. It has a nearly symmetrical banding pattern. There is a distinct band of pale fluorescence in the center of each arm which separates two broad bands of medium intensity. This pale band separates region 1 from 2 in each arm. The distal medium band in the short arm is narrower, but frequently appears more intensely fluorescent than that in the long arm, whereas the terminal pale band is longer in the short than in the long arm.

Similar morphological features are observed in cells treated to produce G-bands. The centromeric regions stain darkly, but the variation in intensity of stain, seen with fluorescence, is not observed in G-bands.

Chromosome 4 is a long submetacentric chromosome that, similar to No. 2, lacks prominent landmarks. It has one band (15) of medium fluorescence in the short arm and four or five relatively evenly spaced bands of medium intensity in the long arm. The long arm is divided into three regions by a proximal dull band separating region 1 from 2 and a distal dull band separating 2 from 3.

Chromosome 5 has more distinctive characteristics than No. 4 and is frequently the first pair of B-group chromosomes identified in the cell. The central band (14) of medium fluorescence in the short arm is brighter and frequently wider than that in No. 4. There is a broad central band of medium fluorescence which separates region 1 from 2 in the long arm, with a prominent distal pale band which separates region 2 from 3. Frequently, the terminal portion of the long arm of No. 5 is paler than No. 4.

Chromosome 6 is the largest and one of the last submetacentric C-group chromosomes. The most prominent feature is a distinct band of pale fluorescence in the middle of the short arm, which separates two bands of medium intensity and which also separates region 1 from 2. The long arm contains a number of bands of medium intensity; those bands near the centromere are frequently brighter and more distinct. A dull band in the middle of the long arm separates region 1 from 2.

Chromosome 7 is slightly smaller than No. 6, but their centromere positions are similar. The two bands of intense fluorescence in the center of the long arm are the most prominent feature of this chromosome. The proximal band divides regions 1 and 2, whereas the distal band divides regions 2 and 3. There is a distinct band of medium fluorescence at the end of the short arm which divides region 1 from 2.

Chromosome 8 is one of the most submetacentric C group chromosomes, and it lacks distinctive landmarks. The short arm shows less intense fluorescence

than the long arm, and a central pale band which divides region 1 from 2 may be seen in good preparations. A distal pale band (22) separates two medium bands in the long arm, The more proximal of these medium bands separates region 1 from 2.

Chromosome 9 is the middle-sized C chromosome and is more metacentric than Nos. 8 or 10. The long arm shows a negatively staining centromeric region (12) with two distal evenly spaced bands of medium intensity. The proximal band of the pair may appear wider and separates region 1 from 2; the distal band separates 2 from 3. The short arm has a characteristic heart-shaped appearance with a central band of medium intensity which divides region 1 from 2, Both G and Q techniques show a prominent, faintly staining region near the centromere. The R-bands are the reverse, except for the centromeric region which is pale; the R-bands are similar in staining intensity to Q- and G-bands. The negatively fluorescing and pale-staining region in the long arm near the centromere shows a large block of material that stains intensely with Giemsa after treatment to produce C-bands. This C-band material presumably represents one of the types of constitutive heterochromatin, the size of which may vary with individuals.

Chromosome 10 is one of the smaller, less metacentric C-group chromosomes which can be identified by the three evenly spaced bands in the long arm. The proximal band is most intense (and divides region 1 from 2) and the distal is the least intense. The short arm shows medium fluorescence.

Chromosome 11 is one of the least submetacentric C group chromosomes; and it may be very slightly larger than No. 10. It is somewhat similar to No. 9, but can be distinguished on the basis of the following features: There is a narrow medium band (12) in the middle of the negatively staining centromeric portion of the long arm. In poor preparations, this narrow band may be very faint or not apparent. A broad band of medium fluorescence is present in the middle of the long arm. A narrow dull band in the middle of this broad band separates region 1 from 2. This broad band usually appears as a *single* band that distinguishes it from the definite double band in the long arm of No. 9. The short arm shows medium fluorescence (14) and has a rather squarish appearance. In contrast, the short arm of No. 9 tends to taper near the centromere.

The C-banding technique reveals that the amount of centromeric staining material in No. 11, and the X is second only to that in No. 9.

Chromosome 12 is the most submetacentric of the C-group chromosomes; it is similar in size to No. 10. The short arm shows a band of medium fluorescence (12) which is smaller than that in No. 11, because the short arm is smaller. The band (12) of medium intensity in the long arm near the centromere is wider than that found in No. 11. A short band (13) of negative fluorescence separates band

ql2 from a distal long segment of medium intensity which divides region 1 from 2. This distal segment and the terminal dull band are both longer than the corresponding bands of No. 11.

The X *chromosome* is the third largest chromosome in the C group; it is, with 6 and 11, among the least submetacentric chromosomes in this group. The X chromosome is frequently the most fluorescent of the C-group chromosomes. Both Xs in the female show identical fluorescence patterns. In both arms, the region proximal to these medium bands shows pale fluorescence. Two other evenly spaced, less distinct and less intense bands of fluorescence are found in the distal long arm.

Chromosome 13 is the largest of the D-group chromosomes and shows intense fluorescence of the distal half of the long arm. In good preparations this segment is seen as two bands; the more proximal band divides regions 1 and 2 and the more distal divides 2 and 3. Fluorescence of the satellites and short-arm region shows variable intensity and is an inherited variant; a proximal intense band in the long arm has also been observed as an inherited variant.

Chromosome 14 shows a broad medium to intense band in the proximal half of the long arm which divides regions 1 and 2 and a narrow medium and, dividing regions 2 and 3, close to the distal end of the long arm. Variable fluorescence of the satellites and short-arm region has been noted.

Chromosome 15 is the smallest D chromosome and shows the last intense fluorescence of its group. The proximal half of the long arm shows medium fluorescence which divides region 1 from 2, whereas the distal half is pale. This chromosome is distinguished from No. 14 by the absence of a distal *medium* band, although there may be a distal band of faint fluorescence in No. 15. In poor preparations, it is frequently difficult or impossible to differentiate No. 14 from 15. The satellites and short-arm region show variable fluorescence.

Chromosome 16 is the largest and most metacentric of the E-group chromosomes and is one of the few chromosomes that could be identified solely by morphology. The short arm shows a band (12) of medium fluorescence which is less intense than the central band which separates region 1 from 2 in the long arm. The long arm contains a proximal segment of negative fluorescence (11) which corresponds to the secondary constriction.

The G-band technique reveals the medium bands in the long and short arm and the densely staining secondary constriction. The R-bands are the reverse of Q-bands, except that the region of the secondary constriction is pale. Obanding reveals a large block of constitutive heterochromatin in the region of the secondary constriction which varies in length in different individuals.

Chromosome 17 is the palest staining of the E-group chromosomes. It has a

single distal band (22) of medium fluorescence in the long arm. The dull band proximal to this band divides region 1 from 2.

Chromosome 18 is the smallest E-group chromosome. The long arm contains two bands of medium intensity (12 and 22), the proximal band being the brighter and wider of the two. The dull band between these two bands separates regions 1 and 2.

Chromosome 19 is the most weakly fluorescent chromosome in the karyotype and it is difficult to distinguish the long and short arms except in very good preparations. There is a fluorescent spot (12) in each arm adjacent to the centromere; the spot in the short arm is longer and brighter than that in the long arm.

G-bands are similar to Q-bands, and the chromosome shows the same very pale staining except for the centromeric region which is well stained.

Chromosome 20 also shows weak fluorescence, but more than No. 19. The short arm is brighter than the long arm. The C-band is medium sized and smaller than in No. 19.

Chromosome 21 is the smaller of the G-group chromosomes and is much brighter than No. 22. The long arm shows a proximal intense segment which divides regions 1 and 2, and distal pale segment. There is variable fluorescence of the short-arm region and the satellites; this is an inherited variant. This variability has been useful in determining the source of the meiotic error that results in Down's syndrome (Robinson, 1973).

Chromosome 22 is the larger of the G-group chromosomes and shows very dull fluorescence, similar to No. 19. In fact, in some preparations it may be difficult to distinguish Nos. 19 and 22. A narrow pale band may be observed in the middle of the long arm. A bright fluorescent band in the short arm and variable fluorescence of the satellites are inherited variants.

G-bands are similar to Q-bands, although the centromere region stains darkly with Giemsa. The C-band is medium sized and larger than in No. 21.

The *Y chromosome* was the first human chromosome identified with quinacrine fluorescence (Zech, 1969) because of the brilliance of the distal long arm in cells from all but a few males. The variability in the length of the Y chromosome is well established; it is now evident that this polymorphism is correlated with variation in the length of the brilliant segment.

The Y has a variable appearance with G-banding techniques, but may show two distal bands; it is relatively pale when stained with the R-banding technique. The size of the distal C-band in the long arm is directly related to the length of the brilliant segment on fluorescence.

5.2 Human karyotype—banding—nomenclature

5.2.1 Introduction

Metaphase chromosomes show little morphological differentiation in conventional preparations. The size, the position of the centromere, and occasional secondary constriction(s) are the only criteria that can be employed for recognizing chromosomes. In species with high diploid numbers, chromosome pairs with similar morphology become increasingly common, thus making the identification of individual pairs extremely difficult. The human karyotype is comparatively favorable because at least the chromosomes can be classified, according to morphology, into seven groups, and a few pairs can be identified unequivocally. In the karyotype of the laboratory mouse, all chromosomes are acrocentric and do not even allow grouping.

Cytologists have attempted a variety of ways, such as distribution of chemical-i induced breaks, unstained chromosome regions induced by low temperature, and differential DNA replication time revealed by autoradiography, to further differentiate the chromosomes longitudinally, but all these methods are laborious and the results are ambiguous. The first break through was recorded not long ago when Caspersson I and his collaborators (1969a, b) found that certain fluorochromes, e.g., quinacrine mustard, when applied to cytological preparations and observed with ultraviolet optics, produced characteristically bright and dark bands. Later Caspersson *et al.* (1970a,b) applied the technique to human chromosome preparations and found that the fluorescent banding pattern is likewise specific for each chromosome pair.

Credit to the second major advance must go to Joseph G. Gall and Mary Lou Pardue who perfected the *in situ* DNA/RNA hybridization technique. In their studies on the cytological locations of the satellite DNA of the laboratory mouse (Pardue and Gall, 1970), they treated the cytological preparations with a series of chemicals in order to achieve molecular hybridization. In these preparations, the centromeric areas of the mouse chromosomes stained more deeply with Giemsa than the Chromosome arms. They regarded the densely stained centromeric areas as heterochromatin. The discovery of a simple staining procedure led to an explosive activity in inventing new procedures, particularly regarding the chromosomes of man and other mammals.

5.2.2 Classification of banding patterns :

The Q-bands. Florescent banding with quinacrine mustard or quinacrine dihydrochloride.

The C-bands. Constitutive heterochromatin revealed by the PardueGall *in situ* hybridization procedure or its modifications.

The G-bands. Crossbands of chromosomes revealed by a variety of procedures. These bands coincide well with Q-bands, i.e., deeply stained G-bands are brightly fluorescent in Q-band preparations.

The R-bands. The “reverse” banding pattern following the procedure of Dutrillaux and Lejeune (1971).

5.2.3 Cell harvest and slide preparations

Arresting

Harvesting cells is strictly conventional. Bone marrow cell culture or any cell population containing a high incidence of mitosis are suitable. Agents such as colchicine, colcemid, and vinca alkaloid can be used to accumulate mitosis. However, over condensed chromosomes yield very poor banding, so that prolonged mitotic arrest should not be done.

Fixation

The cell populations should be treated with a hypotonic solution prior to fixation. It matters little which kind of hypotonic solution is used. The cells, after hypotonic solution treatment, are fixed according to the type of slide preparations to be made, viz., “Carnoy” fixative (1 glacial acetic acid: 2 methanol) for air-dried slides. For cell populations which require squash technique (e.g., many solid tissues), the fixative to use is 45—50% acetic acid. However, the preparations are not suitable for G-bands, though they are excellent for C-bands.

It is suggested that the air-dried slides be incubated at 37°C for 1 hour (without covering) and thereafter be kept in air-tight slide boxes containing a drying agent such as silica gel. The slides may be used immediately after this incubation period or may be stored for a few weeks.

5.2.4 C-Banding

In good C-band preparations, the constitutive heterochromatin should stain deeply and the euchromatin should show only a faint outline of the chromosome. However, flame-dried preparations, when improperly treated, will show G-bands as well as C-bands, which is confusing.

The original procedure (Pardue and Gall, 1970) and the modifications thereof (Arrighi and Hsu, 1971; Yunis et al., 1971) are useful for distinguishing constitutive heterochromatin and euchromatin in mammalian chromosomes. The procedure to be described is a simplified version designed to reveal C-bands using air-dried and flame-dried preparations, although some comments will also be made for squash preparations.

A. Reagents

1. HCL : prepare 0.2 N solution.

2. NaOH : Prepare a 0.07 N solution.
3. SSC : Prepare a 10x concentrate (a solution of 0.15 M sodium citrate and 1.5 M NaCl) and dilute with distilled water to the desired concentration.
4. Giemsa staining solution.
5. Phosphate buffer solution (0.01 M Sorensen's phosphate buffer, pH 7.0).

B. Procedure

1. Treat the slides with HCL at room temperature for 15 minutes. Rinse with distilled water three times.
2. Treat the slides with NaOH for 2 minutes. Rinse with 70%, then 95% ethanol three times for a period of 5 minutes each. Air dry the slides.
3. Place slide horizontally, with cell-side up, in a moist chamber, and add either 2x or 6x SSC to the cell area of the slide. Place a coverglass over the SSC solution.
4. Incubate the moist chambers containing the slides at 60°—65°C for 16-20 hours.
5. Rinse in either 2x or 6x SSC (three times, 5 minutes each), 70% ethanol (three times, 5 minutes each), 95% ethanol (three times, 5 minutes each), and air dry.
6. Stain in Giemsa solution.

C. Comments

1. After a considerable amount of experimentation, it is generally opined that the HCL treatment is an important step in eliminating the G-bands in C-band preparation, particularly when air-dried and flame-dried slides are used. In squash preparations, HCL treatment is not a vital step and may be omitted for C-band preparations.

2. The concentration of NaOH and the duration of the NaOH treatment are important. As a standard, one may start with 0.07 N for 2 minutes. This combination may be too strong and the resulting euchromatic chromosomes may appear bloated and show an empty appearance. If such a result is obtained, one must experiment with reduced concentration of NaOH solution (0.02 N, 0.01 N) and time of treatment (1 minute, 30 seconds, or even 15 seconds). If the chromosomes still appear distorted, one should then try the solution suggested by Stefos and Arrighi (1970): a 2 x SSC solution with pH adjusted to 12 by NaOH. This solution is particularly useful for small chromosomes such as the microchromosomes of the birds and the chromosomes of *Drosophila* (Hsu, 1971). Conversely, some C-bands require a prolonged NaOH treatment.

3. Many laboratories use Coplin jars filled with 2 x or 6 x SSC for the overnight incubation. This is undesirable because the glass slides will stain heavily with Giemsa, thus interfering with the observations on the chromosome banding.

Incubating slides in moist chambers eliminates this defect. If, however, such defect is not observed, Coplin Jars are of course convenient.

A simple moist chamber can be constructed as follows. Use a Petridish of suitable size. If square (120 mm each side) Petri dishes are used, 10 ml of either 2 x or 6 x SSC is placed in the bottom (15 ml for overnight treatment at 65°C). Next a stand for the slides is placed in the bottom half. The stand should be as small as possible and of sufficient height so that the slide is above the salt solution in the bottom half of the Petri dish. The slide(s) is placed on the stand. A few drops of the solution are placed on the slide to cover the cells. A coverglass is then placed over the solution, the Petri dish is covered, and the entire chamber is placed in an oven set at the desired temperature.

4. The stock Giemsa-staining solution is diluted with phosphate buffer, and the concentration varies with each new lot of stain. Usually concentrations varying from 2 to 10% have been used and stained the slides from 5 to 30 minutes.

5. Slides should be of good quality and should be cleaned in some manner. Slides can be cleaned in 95% ethanol, soap, and dilute HCL. AH seemed to be acceptable. Coverglasses should also be cleaned.

D. Squash preparation

This section is for squash preparations only. If squash preparations are used for C-bands, the slides should be dipped into a solution of 0.1 % gelatin and 0.01 % chrome alum and dried prior to squashing. This thin coat prevents cellular loss during the treatments. However, if the HCL treatment is omitted, the slides should be treated with RNase (100 pg/ml. diluted in 2 x SSC) at 37°C for 1 hour using the moist chamber method. Rinse the slides for 5 minutes each in three changes of 2 x SSC, 70% ethanol, and 95% ethanol and air dry. Treat the slides with NaOH solution (0.07 N NaOH or 2 x SSC, pH 12). Try several treatment times, e.g., 1 minute, 2 minutes, 4 minutes, etc. if 0.07 //NaOH is used, rinse in 70% ethanol, 95% ethanol, dry, and incubate in 6 x SSC as usual. If 2 x SSC at pH 12 is used, rinse slides in three changes of 2 x SSC for 10 minutes each. Do not dry but drain and immediately place the slides at 65°C in a moist chamber, as suggested earlier.

5.2.5 G-Banding

G-bands are the crossbands of various width and shades stained with Giemsa, Leishman's, Wriht's, or similar stains. They usually correspond to the Q-bands but do not always correespond to the C-bands. In some cases, the C-bands and the G-bands may be opposite in staining behavior. For example, the C-band of human chromosome 9 is relatively unstained with G-band techniques.

The Y chromosome of man shows a distinct C-band in the distal portion of the long arm, but the same chromosome is somewhat variable in G-band staining, usually deeply stained throughout. Thus, G-banding does not replace C-banding in assessing information.

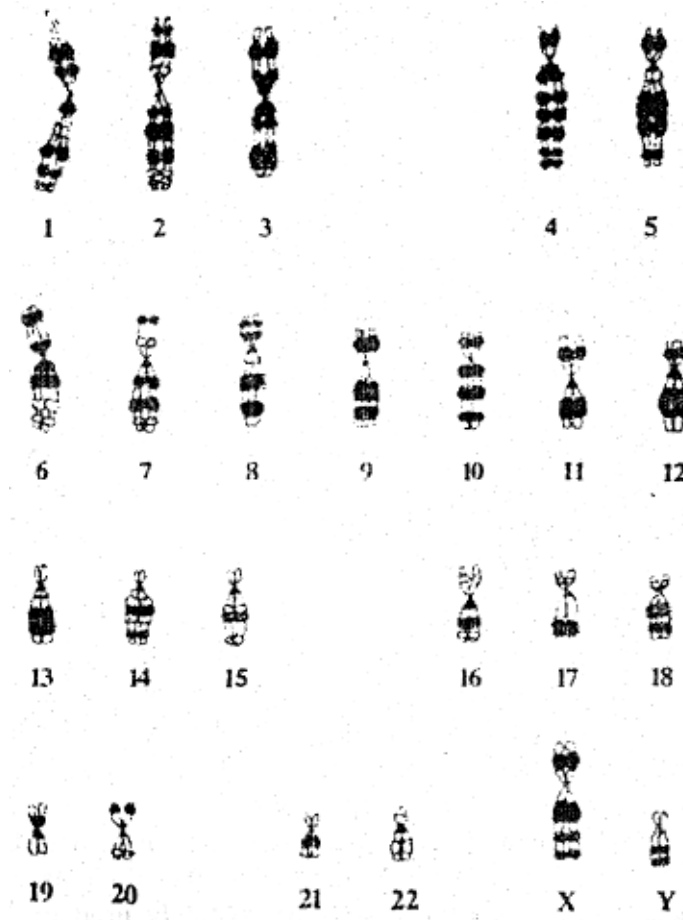


Fig. 5.6 Karyotype of human male showing C-banding pattern from a flame-dried preparation. Note the large amount of heterochromatin at the centromere areas of A1, C9, E16, and the distal portion of the Y. Variations in amounts exist in the two homologs of A1, Courtesy of Mrs. Ann Craig-Holmes and Dr. M. W. Shaw

There were many papers published in 1971 and 1972, each proposing a certain procedure to reveal crossbands in mammalian chromosomes (Summer et al., 1971; Drets and Shaw, 1971; Patil et al, 1971; Schned, 1971a,b; Seabright, 1971; Wang and Fedoroff, 1972; Kato and Yosida, 1972; Utakoji, 1972). Since the results of the various procedures are similar to one another, it is unnecessary to present the procedures for all of them. The trypsin procedure described here more or

less follows the one devised by Seabright (1971) with recommendations for individual laboratory modifications.

A. Reagents

1. Trypsin solution : Seabright uses Bacto trypsin (Difco Catalogue No. 0153) prepared by adding 10 ml of sterile distilled water or isotonic saline to each vial as the stock solution. This stock solution is diluted 1 : 10 with saline before use.

It is really not necessary to use the particular brand of trypsin recommended by Seabright. Most laboratories carry monolayer cell cultures which require trypsin to dislodge the cells for harvest or for subculturing. Usually it is a crude trypsin solution (0.20—0.25% dissolved in a balanced salt solution without Ca^{2+} and Mg^{2+}). In some laboratories, purified trypsin solution (0.01—0.02%) is used. Whatever the kind and the concentration, the trypsin solution routinely used in the cell culture laboratory can be considered as the "stock solution". The trypsin solution used for G-bands is prepared by diluting the stock trypsin solution with saline, balanced salt solution, or "rinsing solution" (balanced salt solution without Ca^{2+} and Mg^{2+}). In our laboratory we use the rinsing solution.

For laboratories using trypsin solution for G-banding only, it is advisable to dispense the stock trypsin solution in small containers and store them in a freezer. Keep only a small amount in the refrigerator for immediate use.

2. Rinsing solution : Physiological saline or balanced salt solution without Ca^{2+} and Mg^{2+} . This solution is used diluent of the stock trypsin solution as well as for rinsing the slides after the trypsin treatment.

3. 95% Ethanol

4. Giemsa staining solution : See C-bands above.

5. Phosphate buffer : Like that of C-bands.

B. General principles

Many factors may influence the success of the G-band staining by trypsin treatment. It is, therefore, pointless to follow a set recipe without knowing these factors because the preparations may give excellent results if one knows how to modify the procedure. The success of G-banding depends primarily on the combination of the concentration of the trypsin solution and the duration of treatment, but the following factors dictate the correct combination :

- 1. The method for preparing the slides** : The flame-dried preparations are more resistant to the trypsin treatment than air-dried preparations.
- 2. The age of the slides** : The longer the slides are stored, the more resistant the cells are to the treatment. Cells of very old slides often give spotty, instead

of banded, chromosomes. Refixing of the slides in the Carnoy fixative sometimes helps.

3. **Heating the slides :** Air-dried slides be heated at 37°C for 1 hour (without covering). This procedure seems to give more consistent results.
4. **The salt composition of the trypsin solution (including the diluent) :** The presence of divalent cations in the solution slow the reaction but do not prevent it.
5. **The temperature of the trypsin solution (the higher the temperature, the faster the reaction) :** In laboratories with air conditioning, room temperature is suitable. The trypsin solution should be stabized at room temperature for approximately 30 minutes prior to use. For laboratories without room temperature control, it is probably a good practice to stabilize the trypsin solution at 4°C (refrigerator) or even in an ice bucket (Deaven and Petersen, 1973).
6. **Trypsin concentration :** We suggest a dilution of 1:5 or 1:10. If these are too strong, dilute further. If too weak (as in the case of flamelried preparations), the concentration may be raised to 1:2 or the undiluted trypsin used.
7. **Time of treatment :** The time of treatment of course depends on all the factors mentioned above. As a general principle, it should be adjusted to give good results in not more than 2 minutes but not less than 30 seconds.

C. Procedure

1. Prepare trypsin solution in a Coplin jar. Using a Coplin jar is somewhat more convenient than flooding the slides with the solution but either way is acceptable. Use two or three slides and vary the duration of trypsin treatments as the initial monitor. Since for best results monitoring is necessary, it is advisable to prepare at least 10 to 12 slides of good quality.
2. Rinse with physiological saline or rinsing solution. Seabright suggested, at this stage, inspection of the wet slides by phase-contrast microscopy to determine the effect of trypsin treatment. The chromosomes should appear slightly swollen. The preparations can be treated again with trypsin if necessary.
3. Rinse with 95% ethanol and let dry.
4. Stain with diluted Giemsa (2% Giemsa solution in phosphate buffer) for 4— 10 minutes. Overstaining may obliterate some of the lighter bands.
5. The slides can be pulled out of Giemsa, rinsed quickly with deionized water, and air dried. It is not necessary to mount the slides.

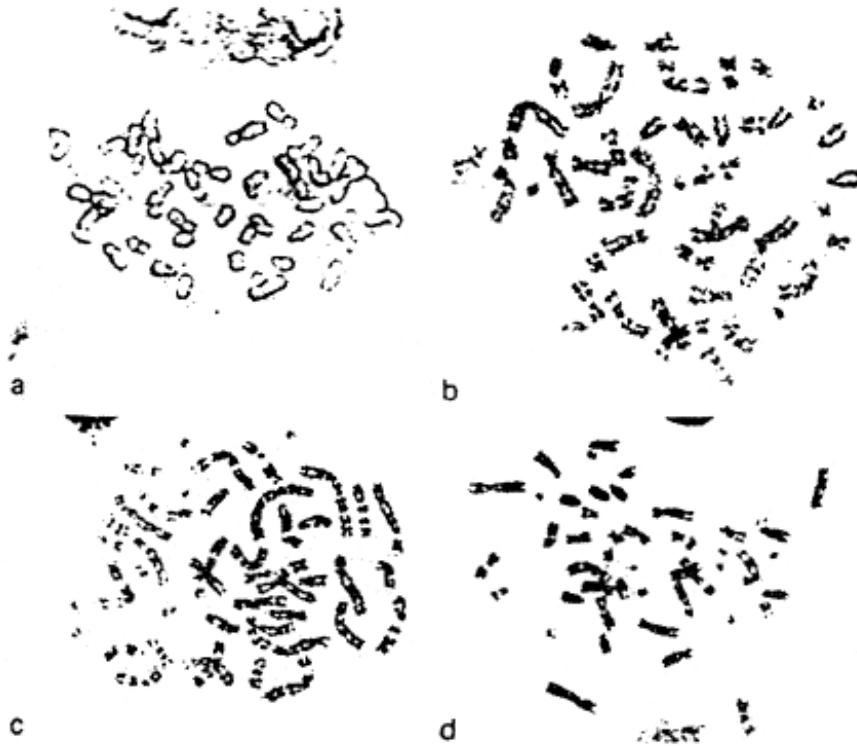


Fig. 5.7 Metaphase plates of human cells showing results of various durations of trypsin treatment from extreme over treatment to proper timing, (a) Extreme overtreatment; (b) overtreatment, highly unsatisfactory; (c) slight overtreatment; chromosomes are fuzzy but discrete bands can be seen; (d) proper treatment

Examine the stained preparations to determine the proper duration of the trypsin treatment. The chromosomes in undertreated preparations, in overtreated preparations, the chromosomes will show a series of appearances ranging from completely “ghost” chromosomes (Fig 5.7a) to those with poorly differentiated crossbands and fuzzy outlines (Fig. 5.7b). The appearance of these cells indicates that the treatment time or the concentration of trypsin solution should be reduced. Figure 5.7c shows a metaphase with reasonably good but slightly overtreated chromosomes, and Fig 5.7d shows proper G-bands.

6. Once the proper combination of trypsin concentration and the duration of treatment is determined by the preliminary monitor, treat the rest of the slides according to the best combination in the same day using the same solutions. The solutions in the Coplin jars should be discarded each day.

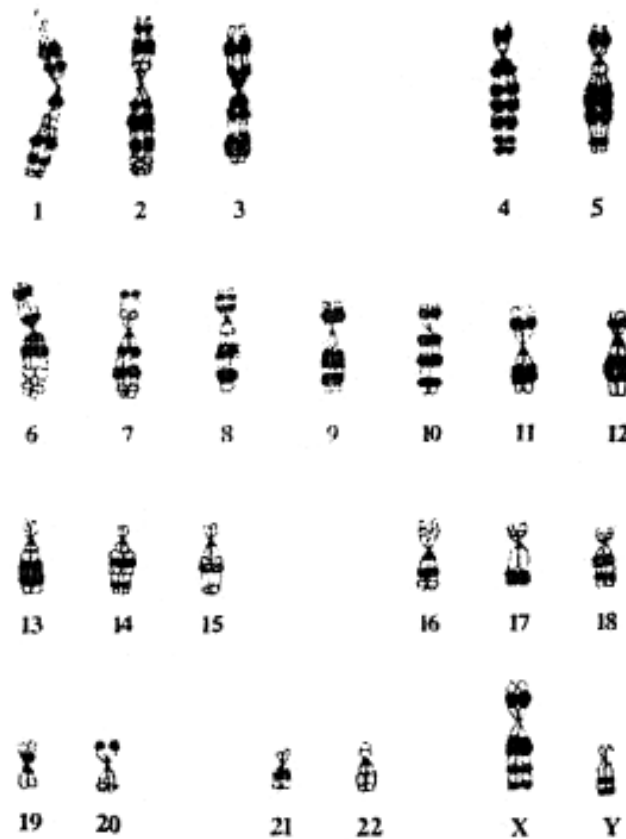


Fig. 5.8 Karyotype of human male showing G-banding pattern produced by the trypsin technique. Courtesy of Mrs. Marina Seabright

7. Figure 5.8 presents a male human karyotype showing the distribution of G-bands.

D. Comments

1. If the trypsin procedure does not give satisfactory G-banding, it may be worthwhile to try the urea procedure of Kato and Yosida (1972). Prepare a stock urea solution (8 M aqueous) and mix 3 parts of this stock solution with 1 part of Sorensen's phosphate buffer (0.15 M, pH 6.8), heat to 37UC. Treat the slides at 37°C for 10 minutes. Rinse in tap water and stain directly without drying. For old slides, elevate the temperature to 60°C.

2. For combination staining, e.g., Q-banding and C-banding of the same cell, it is advisable to perform Q-banding first. The same slide can be then used for C-banding. It is recommend, however, separate runs for C-banding and G-banding. From each sample, one can always prepare a sufficient number of slides for slides for al treatments as well as some for storage.

5.3 Numerical and structural abnormalities of human chromosomes—syndromes.

5.3.1 Introduction

Fifty years has elapsed since Tjio and Levan in 1956 established that human somatic cells have 46 chromosomes. Their technique employing treatment of cells grown in culture with colchicine and hypotonic solution, was applied by numerous workers over the next half-dozen years, demonstrating the chromosomal findings in such disorders as trisomy 21, trisomy 13, trisomy 18, Turner's syndrome, and Klinefelter's syndrome. During this same period, the significance of the Barr body was realized, and lyonization of the X chromosome was hypothesized.

The next few years saw the application of autoradiography for better identification of chromosomes, the delineation of more subtle chromosome disorders such as the cri-du-chat (5p—) syndrome, Wolf-Hirschhorn (4p—) syndrome, 18p—, 18q—, 13q—and a host of other anomalies. Amniocentesis was employed for prenatal diagnosis.

Following 6 years have been marked by advances in our understanding of heterochromatin, chromosome banding, the identification of the Y body, and documentation of previously unrecognized alterations (small translocations, inapparent inversions etc.).

From the first conference held in Denver in 1958 to the one, convened in Paris in 1971, chromosome nomenclature has developed *pari passu* with advances in our understanding of these disorders.

Gross human chromosome abnormalities are not rare. Over 25% of human abortuses lost before the eighth week of pregnancy have abnormal karyotypes. Large surveys on newborns have consistently shown that about 0.5% exhibited aneuploidy (Sergovich *et al.*, 1969; Lubs and Ruddle, 1970).

5.3.2 Autosomal abnormalities of human chromosome—syndromes

A. 4p— Syndrome (Wolf-Hirschhorn Syndrome)

The syndrome described independently by Wolf *et al.* (1965) and Hirschhorn *et al.* (1965), results from partial deletion of the short arm of one of the late replicating no. 4 chromosomes. It is much less common than the 5p— syndrome. About 25 cases have been described to date, all sporadic. Translocation has not been demonstrated with the possible exception of the case described by Wilson *et al.* (1970). Parental age has been somewhat increased (Fryns *et al.*, 1973). A 4r syndrome has been described (Carter *et al.* 1969).

The disorder is characterized by severe psychomotor and growth

retardation. Birth weight is usually about 2000 gm in spite of normal gestation time. Fetal activity is diminished. Most infants are hypotonic.

The skull is microcephalic and often there is cranial asymmetry. In a few cases, midline scalp defects have been noted (Hirschhorn *et al.*, 1965; Wolf *et al.*, 1965; Miller *et al.*, 1970). Hemangioma on the brow is frequent. A prominent glabella and ocular hypertelorism are almost constant features. Divergent strabismus, antimongoloid obliquity of the palpebral fissures have been noted in about half the cases. Iris coloboma has been found occasionally.

The ears have narrow external canals and are low set and simplified in form. The nose is misshapen or beaked with a broad base. The philtrum is short with a down-turned cleft lip or, especially, cleft palate and micrognathia have been noted in most cases.

Males commonly exhibit cryptorchidism and, especially, hypospadias. Absent uterus and streak gonad have been described. Congenital heart malformations, most often atrial or ventricular septal defects, have been noted in about 50% (Wolf *et al.*, 1965, Arias *et al.*, 1970, Guthrie *et al.*, 1971) and may result in death during the first year.

In several patients there has been dimpling of the skin over the sacrum and elsewhere, such as shoulders, elbows, or knuckles. The pelvic and carpal bones are late in ossification. Pseudoepiphyses are seen in the phalanges and at the base of each metacarpal.

B. 5p—Syndrome (Cri-du-Chat Syndrome)

Described initially by Lejeune *et al.* (1963), over 150 examples have been documented to date. The syndrome is present in about 1% of institutionalized individuals with intelligence quotients less than 35. The syndrome results from deletion of 35—55% of the short arm of one of the early replicating B group chromosomes (German *et al.*, 1964; Miller *et al.*, 1969). Most deletions are thought to occur as a result of two breaks. If these occur in the short arm, an interstitial deletion results. If deletion occurs in both arms, a ring chromosome is produced (Rohde and Tompkins, 1965). Maternal age is not elevated. About 70% of those identified at birth are females; however, most older patients have been male (Breg *et al.*, 1970). The reason for this discrepancy is not evident. Mosaicism has also been described; patients having all the stigmata of the full-blown syndrome (Zellweger, 1966; Mennicken *et al.*, 1968). About 10—15% result from translocation (Warburton and Miller, 1967). Pericentric inversion has also been described (Faed *et al.*, 1972).

As the name implies, the syndrome is characterized by a catlike, weak, shrill cry in infancy caused by hypoplasia of the larynx (Ward *et al.*, 1968). However, the cry usually disappears with time, even within a few weeks of age.

(Gordon and Cooke, 1968; Breg *et al*, 1970). The cry, almost one octave higher than normal, is quite monotone in quality (Schroeder *et al*, 1967).

The infant face is characterized by microcephaly, round form, hypertelorism, antimongoloid obliquity of palpebral fissures, epicanthus, bilateral alternating strabismus, broad nasal bones, and low-set ears. Most patients have mild micrognathia. However, the roundness of the face and the ocular hypertelorism disappear with age. The face becomes thin and the philtrum short. Premature graying of the hair has been noted in about 30%. Dental malocclusion is common (Breg *et al*, 1970, Gordon and Cooke, Niebuhr, 1971).

There is usually severe mental retardation (I. Less than 25), failure to thrive, and hypotonia in infancy. Birth weight is usually less than 2500 gm in spite of normal gestation time. Adult height usually ranges from 124 to 168 cm. (49 to 66 inches) Various musculoskeletal anomalies have included hypotonia, flat feet, mild scoliosis, large frontal sinuses, small ilia, syndactyly, and short metacarpals and metatarsals (Mennicken *et al*, 1968).

Dermatoglyphic alterations include simian creases in about 35%. Eight or more whorls have been noted in about 30%.

C. Group C deletion, trisomy, trisomy mosaicism, and partial trisomy

There are seven pairs of C-group chromosomes and hence many possible types of trisomy, partial trisomy, or deletion involving this group. These states, with few exceptions, have not been clinically recognized and, thus, are probably lethal. The use of newer banding techniques on the chromosomes of abortuses may shed light on this question (Hirschhorn *et al*, 1973).

There is a paucity of information concerning anomalies of chromosome no. 6 documented by banding, de Grouchy *et al*. (1968) described a child with bulbous nose, preauricular tubercle, hernias, hypospadias, undescended testes, deep acromial dimples, and psychomotor retardation.

Deletion of the long arms of chromosome no. 7 was reported by Shokeir *et al*. (1973). The child exhibited psychic and somatic retardation, urinary malformations, flexion contractures at the elbows, and low finger ridge count. The most striking aspects of trisomy 8 or trisomy 8 mosaicism syndrome are mental retardation, abnormally shaped skull, reduced joint mobility, various vertebral anomalies, supernumerary ribs, strabismus, absent patellae, short neck, long slender trunk, cleft palate, and marked palmar and plantar creases (Oikawa *et al.*, 1969; Lejeune *et al*, 1969; Emberger *et al*, 1970; Riccardi *et al.*)

Deletion of the short arms of chromosome no. 9 was noted by Alfi *et al*. (1973) and probably by Kistenmacher and Punnett (1970). The patients had trigonocephaly, mental retardation, ocular hypertelorism, anteverted nostrils, malformed pinnae, long philtrum, short neck, hypertonia, congenital heart disease,

and an increased number of digital whorls.

Trisomy for the short arm of no. 9 was defined by Rethore *et al.*, (1973) who reviewed earlier cases. Clinical features included mental retardation, microcephaly, enophthalmos, hypertelorism, mongoloid palpebral fissures, bulbous nose, abnormal pinna, hypoplasia of the phalanges, and abnormal finger creases.

Partial trisomy for the long arm of chromosome no. 10 was reported by de Grouchy *et al.*, (1972). Facial dysmorphia was evident with microcephaly, large forehead, flat round face, arched and wideset eyebrows. Antimongoloid palpebral fissures, microphthalmia, cleft palate, small nose with depressed bridge, malformed pinnas, short neck, micrognathia, various skeletal anomalies (osteoporosis, various rib abnormalities, scoliosis), congenital heart disease, and genitourinary defects. All patients had severe mental retardation.

Partial trisomy for the short arm of chromosome no. 11 was described by Falk *et al.* (1973) and Sanchez *et al.* (1974). Findings common to both cases were mental retardation, marked frontal bossing, nystagmus, antimongoloid palpebral fissures, strabismus, broad fingers or toes, and cleft lip and/or palate.

D. Trisomy 13 syndrome (Patau's syndrome, trisomy D₁)

Trisomy 13 was first recognized by Patau *et al.* (1960), although Bartholin in 1657 may have given the first description of the clinical features (Warburg, 1960). The phenotype is so striking that diagnosis is usually made on clinical ground before the karyotype has been made. The incidence has been estimated to be about 1 per 6000 births (Conen and Erkrnan, 1966).

Arhinencephaly, apneic spells, seizures, feeding difficulties, severe mental retardation, and deafness are common. Any of the holoprosencephalic states (cyclopia, ethmocephaly, cebocephaly, and premaxillary agenesis) may be associated with trisomy 13 (Conen *et al.*, 1966; Fujimoto *et al.*, 1973). Moderate microcephaly with sloping forehead and wide sagittal suture and fontanels have been noted in over 60%.

Microphthalmia or iris coloboma with retinal dysplasia, ocular hypertelorism, and malformed pinnas occur in about 80% (Cogan and Kuwabara, 1964). Capillary hemangiomas in the glabellar region and localized scalp defects in the parieto-occipital area have been described in about 75%. Cleft lip and/or cleft palate and micrognathia have been noted in 60-70% (Conen *et al.*, 1966; Taylor, 1968).

Musculoskeletal abnormalities include postaxial polydactyly of the hands or feet with overlapping flexed fingers (about 75%) with hyper convex narrow fingernails. The calcaneus is often prominent and frequently there are rockerbottom feet.

At least 80% have congenital heart defects, genital anomalies include cryptorchidism (over 90%) in males and bicornuate uterus (about 50%) and hypoplastic ovaries in females.

Polymorphonuclear neutrophils frequently (25—80%) have nuclear projections in cases of trisomy 13 owing to primary nondisjunction. Excellent ultrastructural study of the projections has been carried out (Waltzer *et al.*, 1966, Lutzner and Hecht, 1966). Fetal hemoglobin, Hb-Gower and other hemoglobins have been elevated but there is good evidence that those changes disappear with age and merely represent general delayed maturity (Marden and Yunis, 1967).

DNA replication studies have demonstrated that the D-group chromosome involved is number 13, which is the longest and the latest of the pairs to replicate (Yunis and Hook, 1966).

1. Trisomy 13 Caused by Primary Nondisjunction

About 75% of cases of 13 trisomy are caused by primary nondisjunction. There is no sex predilection. The mean age for mothers of infants with 13 trisomy caused by this type is elevated (32.4 years), far higher than for cases caused by translocation or mosaicism (Magenis *et al.*, 1968; Taylor *et al.*, 1970).

There have been several examples of 13 trisomy occurring with other chromosomal abnormalities in the same sibship (Klinefelter's syndrome, Turner's syndrome, Down's syndrome and triploidy), but this may be chance association (Visfeldt, 1969).

2. Translocation D_1

About 20% of the cases of trisomy 13 are caused by translocation, far more common than occurs in Down's syndrome (Magenis *et al.*, 1968; Taylor *et al.*, 1970). In at least 85%, the translocation has occurred between two D chromosomes. Maternal age is not elevated (25.6 years). There appears to be definite male predilection. Fertility and intelligence in balanced carriers are quite variable (Wilson, 1971).

3. Mosaicism

About 5% of the cases of D_1 trisomy are caused by mosaicism. About half of these examples are caused by an extra chromosome 13 in proportion of the cells. The remainder result from a complex assortment of chromosomal abnormalities (Magenis *et al.*, 1968; Taylor *et al.*, 1970).

As in translocation D_1 trisomy, the age of the mother of a D_1 trisomy mosaic is not elevated (25.4 years) in contrast to mothers of G_1 trisomy mosaics. The clinical stigmata, as expected, are less severe than in those of children with classic trisomy 13 (Bain *et al.*, 1965).

4. Partial Trisomy

Partial trisomy for the distal segment of the long arm of chromosome 13 was documented by banding techniques by Taysi *et al.* (1973) and by Escobar *et al.* (1974). The latter authors reviewed several case reports which had been documented prior to the advent of banding. Common clinical characteristics included psychomotor retardation, seizures, microcephaly, *frontal bossing*, open anterior fontanel, short neck, inguinal and umbilical hernias, polydactyly, rocker bottom feet, distal axial triradius, and elevated fetal hemoglobin. Life expectancy over a year was frequent. Absent were cleft lip and palate, sloping forehead, microphthalmia, and neutrophil drumsticks, common findings in trisomy 13.

E. Dq— and Dr syndromes

Over 60 case reports have been published in which the patient had deficiency of part of the long arm of a D-group chromosome (Dq—) or in which a D-group chromosome was replaced by a ring (Dr) (Lejeune *et al.*, 1968; Gilgenkrantz *et al.*, 1971; Niebuhr and Ottosen, 1973). Although these cases may represent a heterogeneity, there is good evidence to suggest that most involve No. 13 (Wilson *et al.*, 1973).

Only a few examples of Dr have been described in which the chromosome has been identified as no. 15 (Jacobsen, 1966; Emberger *et al.*, 1971). The phenotype in these cases was not striking: short stature, mental retardation, and microcephaly. Mean survival has been 39 months for Dq— cases and 89 months for Dr examples (Taylor, 1970).

All patients have exhibited mental and somatic retardation and many have been hypotonic.

Musculoskeletal abnormalities have included bilateral hip dislocation, focal lumbar vertebral agenesis, inguinal hernia, coxa valga, and synostosis of the fourth and fifth metacarpals.

F. Trisomy 18 syndrome (Edwards syndrome)

In 1960, Edwards *et al.* and, almost simultaneously, Patau *et al.* (1961) described a new syndrome associated with the presence of an extra chromosome in the E group which was subsequently shown to be a no. 18 chromosome (Yunis *et al.*, 1964).

The most constant features of this syndrome, noted in over 75% of the cases, include: developmental retardation, failure to thrive, feeding difficulties, hypertonia, limited hip abduction, flexion deformities (usually ulnar deviation) of fingers, short sternum, congenital heart disease (ventricular septal defect—90%, patent ductus arteriosus—70%, and atrial septal defect—20%), short

dorsiflexed halluces, rockerbottom feet, calcaneovalgus deformity of feet, and cryptorchidism (Weber and Sparkes, 1970).

Craniofacial anomalies almost always present include prominent occiput, low-set malformed pinnae, and micrognathia. Severe anomalies found at autopsy, apart from the cardiac anomalies noted above include Meckel's diverticulum, heterotopic pancreatic tissue, thin diaphragm with eventration, and various renal anomalies.

Dermatoglyphic alterations are frequent. Over 85% of finger prints are simple arches. Over 30% have a simian palmar crease and over 40% have a single flexion crease in the fifth finger.

Trisomy 18 has an uncommon but yet definite association with aplasia of the radius and thrombocytopenia.

1. *Trisomy 18*

The incidence of trisomy 18 in the more recent surveys has varied from 1 per 3500 to 1 per 7000 births (Taylor, 1968, Benady and Harris, 1969; Garfinkel and Porter, 1971). Mean maternal age is elevated, 32 years (Taylor, 1968).

There is a 3 : 1 female predilection caused, in large part, by a greater male fatality rate during the first few weeks of life (Weber, 1967).

The mother often exhibits small weight gain during pregnancy and indicates that fetal movements were feeble. Most examples are postmature. Mean birth weight is less than 2300 gm. The placenta is often small with umbilical artery, and hydramnios has been noted in over 50%.

Thirty percent fail to survive more than 1 month, 50% succumb by 2 months, and less than 10% live more than 1 year. Mean survival time is about 70 days (females—134 days, males—15 days).

2. *Double Trisomies*

Double primary nondisjunction has been observed in 5—10% of cases (Hamerton, 1971). Mean survival time for double trisomies has been 3 weeks. Maternal age is markedly increased in this group.

3. *Trisomy 18 Caused by Translocation*

Translocation is usually sporadic but examples of familial translocation have been recorded (Hamerton, 1971). Mean maternal age is lower than for those with trisomy 18 caused by nondisjunction.

G. 18p—syndrome

Deletion of the short arms of chromosome 18 is associated with a variable phenotype. Maternal age is elevated. There is a 2 : 1 female sex predilection (Parker *et al*, 1973).

Mental retardation is a constant feature but of variable degree. Birth weight is low and somatic growth retarded. There is no characteristic facial dysmorphism. Frequently, however, hypertelorism, epicanthal folds, strabismus, and ptosis of lids are noted. The ears are low-set, large, floppy, and poorly formed.

H. Trisomy 21 (Down's syndrome)

Langdon Down (1866) first extensively described the syndrome which has received his name, calling it "Mongoloid idiocy" or "Mongolism." In 1959, Lejeune demonstrated that the disorder was associated with an extra chromosome in the G group. In 1960, Polani et al. described translocation Down's syndrome, and in 1961, Clarke et al. discovered mosaicism for an extra G-group chromosome. Yunis et al. (1965), by means of autoradiography, identified the chromosome as one of the no. 22 chromosomes, although by this time the term trisomy 21 had been so extensively employed that it has remained.

The incidence of trisomy 21 is between 1 and 2 per 1000 live births among various populations (Mikkelsen, 1971). Over 95% of the cases are caused by nondisjunction, the remainder resulting from translocation.

The skull is brachycephalic with shortening of the anteroposterior diameter and flattening of the occiput in about 75%. The cephalic index (normally 0.75—0.80) is usually greater than 0.80 and may exceed 1.00. (Roche et al., 1961). In infants with trisomy 21, the fontanelles are larger than normal and closure is late. In those over 10 years of age, a patent metopic suture is found in 65% of males (normal—9%) and in 40% of females (normal—12%). An extremely common feature (over—90%) is absence of frontal and sphenoid sinuses and hypoplasia of the maxillary sinuses (Spitzer *et al.*, 1961, Betlejewski *et al.* 1964). There is poor development of the bones of the middle face, producing a relative prognathism and ocular hypertelorism (Gerald and Silverman, 1965).

The profile is flattened owing to hypoplasia of the nasal bones. The palpebral fissures are oblique, the outer canthus being slightly higher than the inner. Epicanthal folds are extremely common. Speckled iris (Brushfield's spots) and lens opacity are present in about 85% and 60% of patients, respectively.

Various other anomalies include missing or malformed teeth (especially maxillary lateral incisors and mandibular second premolars), delayed eruption, increased periodontal destruction, and malocclusion (Cohen and Cohen, 1971).

The hands are characteristically short and broad, the fifth finger usually being abbreviated and clinodactylous, and having a single flexion crease in about 20% of the cases. There is usually greater space than normal between the hallux and the rest of the toes.

Hypotonia, especially marked in infancy, improves with age. Joints are usually hyperextensible. The penis and scrotum are usually small and about

25% have cryptorchidism Pubic hair is straight. Congenital cardiac anomalies is present in about 40% Down Syndromic persons.

Diastasis recti, duodenal atresia, or umbilical hernia occur in about 10% (Butterworth *et al.*, 1964).

Radiographic changes include reduced iliac and acetabular angles in the young infant (Nicolis and Sacchetti, 1963) and hypoplastic middle phalanx of the fifth finger.

Intelligence quotients range from 25 to 70, most Down's syndrome patients 3 years of age or less having I.Q. s of 50—59 but slipping with increasing age to 25-49 (Penrose and Smith, 1966).

Dermatoglyphic anomalies include distal axial triradius in the palm (over 80%), bilateral simian creases (30%), single flexion crease in fifth finger (20%), 10 ulnar loops (30%), and hallucal arch tibial (70%) or small loop distal (30%) patterns (Preus and Fraser, 1972).

Because of susceptibility to respiratory infection, early mortality used to be great. With the introduction of antibiotics, the means survival age is almost 20 years. There is a twentyfold increased association with acute leukemia (Conand Erkman. 1966).

Numerous attempts have been made to established specific biochemical alterations. However et al, (1965) found decreased blood serotonin and increased galactose phosphate uridylyltransferase leukocyte alkaline with Down's syndrome.

1. *Trisomy G₁ due to Primary Nondisjunction*

As discussed above, about 95% of cases of Down's syndrome are sporadic primary trisomics, resulting from nondisjunction which is age dependent. This occurs at the first meiotic division in the mother (Robinson, 1973). If the mother is less than 20 years of age at time of conception, the risk of producing a child with trisomy 21 is about 1 per 2500 live births. This risk gradually increases until 35 years, after which there is a more marked increase in frequency such that a mother over 45 years has about 1 chance in 50 or less of having a child with Down's syndrome.

2. *Association of Down's Syndrome with Other Primaiy Nondisjunctions*

Individuals with trisomy 21 have been occasionally (about 1 per 200) found to have another extra chromosome (double primary nondisjunction), the most frequent type being 48,XXY, G+ (Hamerton *et al*, 1965; Taylor and Moores, 1967). Other forms such as 48,XXX,G+ and 48,XYY,G+ have also been described (Yunis *et al*, 1964, Uchida *et al.*, 1966). This association is much higher than might be expected by chance.

3. *Translocation Down's Syndrome*

Down's syndrome patients born to young mothers as well as those with affected relatives often have the extra G] chromosome attached to another chromosome. This has been designated translocation and comprises about 3.5% of cases of Down's syndrome. It may be sporadic or familial. Translocation Down's syndrome is not age dependent. About 8% of Down's syndrome patients born to mothers less than 30 years of age have exhibited translocations as opposed to 1.5% born to mothers over 30 years old. It is widely accepted that the short arms of acrocentric chromosomes have nucleolar organizers and that these points are likely to break, producing a high frequency of structural chromosome aberrations.

In familial translocation Down's syndrome, one of the parents has 45 chromosomes instead of the normal 46. One of the small G-group chromosomes is "missing" since it has been translocated to another chromosome. The parent carrying the translocation chromosome is phenotypically normal, since no significant amount of genetic material has been lost in the translocation process. In most cases, the extra G(chromosome is phenotypically normal, since no significant amount of genetic material has been lost in the translocation process.

4. *Down's Syndrome Mosaicism*

Patients having two different cell populations, one trisomic for chromosome G; and another normal, constitute about 2—3% of patients with Down's syndrome. This condition is usually suspected when the phenotypic expression of trisomy 21 is not fully expressed or when the intelligence of the patient is higher than expected. In addition, they may have children with Down's syndrome (Weinstein and Warkany, 1963). Individuals having trisomy Gt mosaician may vary in phenotype from typical trisomy 21 to normal. There is no age dependency (Richards, 1969). One cannot correlate the percentage of trisomic blood cells with intelligence. Richards (1969) found about 20% more trisomic cells in fibroblasts than in lymphocytes.

If mosaicism is found in one of the parents of a child with Down's syndrome, meiotic study of ovary or testis should be carried out. There is evidence that if half the cells are abnormal, about 25% of the children will have Down's syndrome (Mikkelsen, 1971a).

I. Nonmongoloid "Trisomy G"

Several cases of nonmongoloid "trisomy G" have been published (Uchida et al, 1968; Al-Aish, 1969; Lozzio, 1969; Mikkelsen, 1969). Some have been designated as having trisomy 22 to contrast with trisomy 21 (Down's syndrome). At this point in time, within this group, with two possible exceptions cited

below, there seems to be no characteristic phenotype and it would appear likely that some of these represent centric fragments that may come from several different chromosomes.

J. G deletion syndromes

There are at least two relatively distinct phenotypes presumably representing monosomy or deletion of a portion of the long arm of two different G-group chromosomes (Warren *et al.*, 1973).

1. The G₁ Deletion Syndrome (Antimongolism)

This syndrome consists of mental and growth retardation, hypertonia, nail anomalies, skeletal malformations, cryptorchidism, hypospadias, inguinal hernia, pyloric stenosis, thrombocytopenia, eosinophilia, and hypogammaglobulinemia. Facial and oral manifestations include microcephaly, large low-set ears, antimongoloid obliquity of palpebral fissures, highly arched or cleft, and micrognathia. Dermatoglyphic analysis has shown a marked increase in radial loops (Schindeler and Warren, 1973).

2. The G₁ Deletion Syndrome

This syndrome has less distinctive features: severe mental retardation, hypotonia, soft tissue syndactyly of the second and third toes, and clinodactyly of the fifth finger. Facial and oral manifestation include large, low-set ears, epicanthal folds, ptosis of eyelids, highly arched palate, and bifid uvula (1971; Stoll *et al.*, 1973). Dermatoglyphic analysis has shown a marked increase in whorls, a decrease in ulnar and radial loops, a distal axial triradius, and hypothenar patterns (Schindeler and Warren, 1973).

K. Triploidy

Triploidy, as noted later in this chapter, is a frequent cause of fetal wastage prior to the eighth intrauterine week. Diploid/triploid mosaicism is occasionally compatible with survival and there have even been several examples of pure triploidy.

5.3.3 Sex chromosomal abnormalities

A. Klinefelter syndrome

In 1942, Klinefelter *et al.*, described a syndrome in post pubertal males consisting of small firm testes with tubular by a lincization but with a normal number of Leydig cells, azospermia, gynecomastia, elevated urinary gonadotropies and low concentrations of urinary 17-ketosteroids. Several years later, Bradleury *et al.* (1956) and Plunkett and Barr (1956) noted Chromatin-positive nuclei in the tissues of such patients, and Jacob and Strong (1959) described an

XXY sex chromosome complement in chromatin-positive klinefelters syndrome.

Chromatin-positive males have been found to comprise about 2 per 1000 live male births. These also contains XXY, XXYY, XY/XXY, and other rarer forms of Klinefelter syndrome. It has been estimated that about 80% are XXY, 10% are mosaics, and the rest are XXYY and the more unusual types.

1. XXY Klinefelter syndrome

The clinical features of XXY klinefelter syndrome do not become apparent until after puberty. Body proportions usually do not appear remarkably abnormal, however, the lower extremities tend to be and about 60% have a span that exceeds their height by 3 cm or more.

The Prepubertal testes are normal size and microscopic appearance but during adolescence they fail to enlarge and remain small and firm, averaging less than 2cm in length. The seminiferous tubules are usually shrunken, hyalinized, and irregularly arranged. Elastic fibres are absent around the tunica propria of the tubules. Ledig cells are clumped, Rarely spermatogenesis can be demonstrated. In nearly all cases the testes descend. The penis is usually of normal size but may be somewhat shorter than normal. The prostate is smather than normal. Gynecomastia develops after puberty in about 50% and facial hair is sparse in about 60-75%. Axillary hair may also be less.

The prepubertal testes are of normal size and microscopic appearance but during adolescence they fail to enlarge and remain small and firm, averaging less than 2 cm in length. The seminiferous tubules are usually shrunken, hyalinized, and irregularly arranged. Those tubules which are not sclerotic are immature and lined exclusively with Sertoli cells. About 50% have a female pubic pattern. Libido and potency are usually decreased. There is some evidence of increased tendency to pulmonary disorders, varicose veins, and, possibly, breast cancer. There is the same frequency of color blindness among XXY patients as in normal females. Although intelligence may be reduced, at least 75% of XXY males have normal intelligence. Personality is usually passive.

The incidence of XXY Klinefelter's syndrome is about 1.3 per 1000 live male births. Maternal age is significantly increased for XXY but not for XXYY, XXXY, or XXXXY patients. About 60% of XXY males are $X^M X^M Y$, while 40% are $X^M X^{PY}$. The $X^M X^M Y$ state arises from nondisjunction either during oogenesis or at an early postzygotic division. The $X^M X^{PY}$ condition probably has its origin in nondisjunction during the first meiotic division.

2. XXYY Klinefelter syndrome

Patients with the XXYY variant tend to be about 4 cm taller than average height, more aggressive, and more mentally retarded than those with XXY Klinefelter's syndrome (Borgaonkar *et al*, 1970). Otherwise the phenotype is quite

similar: small firm testes, eunuchoid body build, sparse, body hair, gynecomastia, and elevated gonadotropins. Almost all XYY males described to date have been mentally retarded and many have been aggressive (Schlegel *et al.*, 1965).

As mentioned above, there is no increase in parental age in contrast to XXY Klinefelter's syndrome. The disorder is most likely due to nondisjunction in both the first and second meiotic divisions during spermatogenesis with production of an XYY sperm. One cannot, however, rule out the less likely possibility of nondisjunction at the second meiotic division in both parents.

Dermatoglyphic studies have shown that digital arch patterns are more common in the XYY patient than in the XXY individual who, in turn, has more than the normal male.

A child with an XYYY sex chromosome complement was noted to have mental retardation, lordosis, flexed index and fifth fingers, pes planus, and aggressive personality (Gracey and Fitzgerald, 1967).

3. XXXY Klinefelter's syndrome

Over 25 cases of XXXY Klinefelter's syndrome have been published (Vormittag and Weninger, 1972). All have been mentally retarded. The phenotype is similar to that of the XXY male but the size of the penis is small (McGann *et al.*, 1970).

Two late-labeling X chromosomes have been demonstrated. However, two Barr bodies are seen in only a proportion of cells (Vormittag and Weninger, 1972). The condition may arise from successive nondisjunction in either the maternal or paternal meiotic divisions (Pfeiffer and Sanger, 1973). Dermatoglyphic findings have not been consistent. An XXXYY male has been described by Bray and Josephine (1963).

4. XXXXY Klinefelter's syndrome

There have been over 70 cases of 49,XXXXY males published since Fraccaro and Lindsten documented the first example in 1960. Nearly all have been severely mentally retarded intelligence quotients ranging from 20 to 60. A marked difference between the XXY and XXXXY male is the poor development of the external genitalia in the latter. The penis is always minute and the testes very small and underserved with hypoplastic Leydig cells and absence of germ cells. The scrotum is usually hypoplastic.

Mild microcephaly, ocular hypertelorism (90%), myopia (25%), strabismus (50%), mild nongonoid obliquity of palpebral fissures (35%), epicanthus (80%), and short neck with redundant skin on the nape have been noted. Skeletal anomalies present in over half the cases. Congenital heart disorders have been noted in about 20% of cases. Gonadotropins have not been elevated.

Autoradiographic evidence has shown three heavily labeled X chromosomes

(Hsu and Lockhart, 1965). Three Barr bodies may be found in a proportion of interphase nuclei (Miller and Warburton, 1968).

Parental age is not elevated. Postzygotic nondisjunction in an XXY zygote appears to be the cause for the XXXXY state, all the X chromosomes coming from the mother (Murken and Scholz, 1967; Race and Sanger, 1969).

5. *XX Klinefelter syndrome*

Less than 30 cases have been published of males having 46,XX karyotypes. They exhibit many of the stigmata of Klinefelter's syndrome and, hence, will be considered here (Anderspn *et al.*, 1972).

All have small testes, are infertile, and rarely shave. About 70% have gynecomastia and elevated gonadotropin levels. Plasma testosterone levels are very low (Neuwirth *et al.*, 1972). The penis and scrotum have been small in about half the cases. All are of normal intelligence and have normal skeletal proportions.

6. *Klinefelter's syndrome Mosaicism*

About 15% of patients with Klinefelter's syndrome have been found to have two or more chromosomally distinct cell populations. In each of these individuals, one of the cell populations generally has an XXY, XXXY, or XXXXY sex chromosome constitution while the other is XX or XY. The clinical expression of mosaicism for Klinefelter's syndrome depends on the type of sex constitution present at a critical time of development. Thus, one can find, for example, and XY/XXY mosaic who is phenotypically normal, provided the XY cells exerted the predominant genetic effect.

In a study of XY/XXY mosaics, Gordon *et al.* (1972) found that only one-half exhibited azospermia and about one-third had gynecomastia and elevated gonadotropins. About one-quarter had germinal epithelium. Among 6 patients with XX/XXY mosaicism, Ferguson-Smith (1969) noted comparable findings.

B. XYY syndrome

Although the presence of an extra Y chromosome had been described as early as 1961 (Sandberg *et al.*, 1961), interest was markedly aroused by a finding of a disproportionately high percent (usually 2—4%) of such individuals in prisons and mental hospitals (Casey *et al.*, 1968; Jacobs *et al.*, 1968; Marinello *et al.*, 1969; Hook 1973). It was soon noted that most XYY patients are excessively tall and not uncommonly mildly mentally retarded (mean intelligence quotient —90) (Valentine *et al.*, 1971). However, the frequency of the condition among newborn male infants is about 1 per 700 births (Ratcliffe *et al.*, 1970) and few of these individuals lead other than quite routine lives. The adult height of an XYY individual is usually over 180 cm while XYY children are usually above the 90th

percentile in height by 6 years of age. Leg length and trunk length are increased but the leg/trunk ratio is normal (Keutel and Dauner, 1969). Muscle weakness (especially of the pectoralis major) and poor coordination are commonly noted. Phenotypical alterations are subtle: mild facial asymmetry, mild pectus excavatum, and mild scapular winging. The ears tend to be long and often there is a bony chin point. Most have exhibited normal sexual development (Court Brown, 1969). There are no characteristic dermatoglyphic alterations (Hubbell *et al.*, 1973).

The disorder probably arises from paternal nondisjunction during the second meiosis. Retarded intelligence (I.Q.—70), impulsive aggressive behaviour, bilateral simian creases, clinodactyly of fifth fingers, retarded bone age with pseudoepiphyses at the bases of the metacarpals and metatarsals, and lack of patellar epiphyseal calcification were described by Schoepflin and Centerwall (1972).

Ridler *et al.* (1973) noted low normal intelligence, behavior problems with aggressive outbursts, repeated pulmonary infections, hypotrophic testes, sparse body hair, and acne in a 48,XYYY patient. Conversely, Hunter and Quaife (1973) described no stigmata other than sterility.

C. Turner's syndrome

In 1938, Turner described a syndrome in postpubertal female consisting of sexual infantilism, short stature, webbed neck, and cubitus valgus. Albright *et al.* (1942) showed that these patients had an elevated urinary excretion of gonadotropins, and Wilkins and Fleischmann (1944) described "streak" gonads devoid of ovarian follicles in such cases. Polani *et al.*, (1954) and Wilkins *et al.*, (1954) demonstrated that most cases are chromatin-negative, and Ford *et al.* (1959) first described the XO karyotype.

Turner's syndrome has been estimated to occur 1 per 2500 female births (Maclean *et al.*, 1964; Mikamo, 1968) and has been frequently noted in abortuses. Parental age is not increased.

Variation in phenotype has led to some confusion concerning nomenclature. Since the most common features are short stature, streak gonads, and X monosomy or short arm loss of X chromosomal material, all patients with these features are classified here as examples of Turner's syndrome (Yunis, 1965). Deletion of the long arm of the Y chromosome has occasionally been associated with the Turner syndrome. Cases with streak gonads and sexual infantilism but of normal or increased stature and normal female or male sex chromosome complement will be referred to as having "pure gonadal dysgenesis" or, more accurately, XY or XX gonadal dysgenesis.

Primary amenorrhea and sterility are almost constant features of the XO Turner syndrome although exceptions have been noted. Breast development is

poor, the chest is broad with seemingly widely spaced, hypoplastic, at times, inverted nipples. The external genitalia are infantile and pubic hair is sparse.

The histological pattern of the dysgenetic gonad found in Turner's syndrome consists of long streaks of white wavy connective tissue stroma without follicles. Follicles are present, however, in fetal and infantile ovaries of patients with Turner's syndrome (Gordon and O'Neill, 1969).

Adult height is usually less than 57 inches (144 cm). Various skeletal anomalies include cubitus valgus (about 75%), short fourth metacarpals (about 65%), deformity of medial fibial condyle (about 40%) osteoporosis (about 50%), hypoplasia of first cervical vertebra (about 40%), and small carpal angle.

Birth weight is below the 3rd percentile in about half the cases. In infants, excess skin on the nape and peripheral lymphedema have been noted in 15—50% of the cases. During embryonic life, neck blebs or cystic hygroma are common (Singh and Carr, 1966; Rushton *et al.*, 1969). Toenails are frequently hypoblastic. With age, the excess skin on the nape metamorphoses into pterygium colli and, with improvement in deep lymphatic circulation, the peripheral lymphedema gradually disappears. Increased numbers of cutaneous nevi are found in about 60%.

Epicanthal folds, ptosis of upper eyelids, prominent ears, and micrognathia are common facial features. The hairline is low at the nape.

Thyroid antibodies are elevated in XO Turner's syndrome but less frequently than in the X-iso X mosaic and glucose intolerance occurs with greater frequency in patients with Turner's syndrome and in their parents than in the normal population (Rimoin, 1973).

1. *X Monosomy*

Patients with an XO sex complement appear to comprise about 60% of the cases of Turner's syndrome. Furthermore, they appear to be more severely affected clinically than other forms of the disorder.

2. *XO/XX and XO/XXX Mosaicism*

Patients with Turner's syndrome may have two different cell populations, one having an XO sex constitution, the other a normal XX sex complement. Such individuals are called XO/XX mosaics and constitute about 7% of the cases of Turner's syndrome. The two cell population types may appear in every tissue of the body or only in certain ones.

Presumptive evidence for mosaicism lies in a discrepancy between sex chromatin pattern and karyotype, or through observing a low percentage of chromatin-positive nuclei (5—15%) in phenotypic females. The clinical spectrum of XO/XX mosaicism is wide and may vary from cases quite typical of Turner'

syndrome with many associated anomalies to cases with normal gonads and normal stature. About 20% menstruate (Ferguson-Smith, 1969). In contrast to patients with XO Turner's syndrome who are prone to aortic coarctation, those with XO/XX karyotypes are likely to have pulmonic stenosis with or without atrial septal defect (Nora *et al.*, 1970), being similar to patients with Noonan's syndrome.

The usually accepted explanation for XO/XX mosaicism is loss of an X chromosome during cleavage in the early embryo.

About 5% of the cases of Turner's syndrome are XO/XXX mosaics. Clinically they resemble the XO/XX mosaic. Patients having three stem lines XO/XX/XXX have been reported but are quite similar phenotypically to XO/XX mosaics.

3. *Isochromosome X (XXqi)*

About 20% of patients having Turner's syndrome have an X isochromosome, i.e., replication of the long arm of the late replicating X chromosome. They exhibit short stature, sexual infantilism, primary amenorrhea, and skeletal anomalies.

The Barr body and polymorphonuclear neutrophils drumsticks are larger than normal. Drumsticks are also more numerous (Taft *et al.*, 1965; Sparkes and Motulsky, 1967).

4. *Short and Long Arm Deletion of an X chromosome*

Deletion of the short arm of an X chromosome (XXp—) results in the Turner phenotype. They are as short as individuals with the XO Turner syndrome but are less likely to have associated malformations (Atkins *et al.*, 1965).

Deletion of the long arm of an X chromosome (XXq—) is far less likely to be associated with short stature. The girl was 159 cm tall and exhibited no stigmata of Turner's syndrome. She never menstruated and her ovaries were not palpable. As expected, her Barr bodies were smaller than normal. Xg1 studies showed that the Xpi was of maternal origin. However, Turner's syndrome has been reported in association with XXq—.

5. *Y deletion*

At least a dozen cases of Turner's syndrome associated with a dicentric Y Chromosome have been published (Armendares *et al.*, 1972; Cohen *et al.*, 1973). All have short stature, female phenotype, and most have associated anomalies. A patient with long arm isochromosome Y had drumsticks have been noted in polymorphonuclear leukocytes.

5.4 Human genome

5.4.1 The Human Genome Project

As the recombinant DNA, gene cloning and DNA sequencing technologies improved in the 1970s and early 1980s, scientists began discussing the possibility of sequencing all 3×10^9 nucleotide pairs in the human genome. These discussions led to the launching of the **Human Genome Project** in 1990. The initial goals of the Human genome project was to construct a detailed physical map of the entire human genome, and to determine the nucleotide sequences of all 24 human chromosomes by the year 2005. Scientists soon realized that this huge undertaking should be a worldwide effort. Therefore, an international **Human Genome Organization** (HUGO) was organized to coordinate the efforts of human geneticists around the world.

James Watson, who with Francis Crick, discovered the double-helix structure of DNA, was the first director of this ambitious project, which was expected to take nearly two decades to complete and to cost in excess of \$3 billion. In 1993, Francis Collins, led the research teams that identified the cystic fibrosis gene, replaced Watson as director of the Human Genome Project. In addition to work on the human genome, the Human Genome Project has served as an umbrella for similar mapping and sequencing projects on the genomes of several other organisms, including the bacterium *E. coli*, the yeast *S. cerevisiae*, the fruit fly *D. melanogaster*, the plant *A. thaliana*, and the worm *C. elegans*.

5.4.2 Bacterial Genomes

In 1995, *Haemophilus influenzae* was the first bacterium to have its genome sequenced in its entirety. By mid-2001, the complete sequences of 32 bacterial genomes were available in the public databases (collections of the sequences of genes, chromosomes and genomes). The genomes range in size from 580,070 bp for *Mycoplasma genitalium*, which is thought to have the smallest genome of any self-replicating organism, to 4,411,529 bp for *Mycobacterium tuberculosis*, which causes more human deaths than any other infectious bacterium, to 4,539,221 bp for *Escherichia coli*, the best-known cellular microorganism. The genome of *M. genitalium* is of special interest because it may approximate the “minimal gene set” for a self-replicating organism—the smallest set of genes that will allow an organism to reproduce itself. Of course, the genome of *M. tuberculosis* is of great interest because of the pathogenicity of this organism and the hope that a complete understanding of its metabolism will suggest ways to prevent tuberculosis in humans: The need for new ways to combat this pathogen has been enhanced by the recent evolution of antibiotic-resistant strains of *M. tuberculosis*.

Of the bacterial genomes sequenced to date, the genome of *E. coli* (Fig. 5.9)

has undoubtedly caused the most excitement among geneticists. *E. coli* is the most planet. Geneticists, biochemists, and molecular biologists have utilized *E. coli* as the preferred model organism for decades. Most of what is known about bacterial genetics was learned from research on *E. coli*.



Fig. 5.9 Sequence-based map of the chromosome of *Escherichia coli*. The blue arrows mark the halves of the chromosome traversed by the two replication forks. The outer concentric circle gives the position of genes encoding proteins similar to bacteriophage proteins. The second concentric circle shows the location of genes that are transcribed clock-wise (gold) from one strand or counterclockwise (yellow) from the complementary strand. The sunburst in the center is a histogram in which the length of each ray is proportional to the randomness of codon usage within each coding sequence

The *E. coli* genome contains 4288 putative protein-coding sequences or genes. About one-third of these are well-studied genes encoding known products, whereas 38 percent are of unknown function. Putative protein-coding sequences that are not known to encode proteins are called **open reading frames** or **ORFs**. An **ORF** is a nucleotide sequence that begins with a translation-initiation signal (usually ATG), continues with a sequence of base triplets specifying amino acids, and ends with one of the three translation-termination signals.

The average distance between genes (size of intergenic regions) in the *E. coli* genome is 118 bp. Known and putative genes specifying proteins and stable RNAs make up 87.8 percent and 0.8 percent of the genome, respectively, and noncoding repetitive elements account for 0.7 percent of the genome. Thus, 10.7 percent of the genome must involve regulatory sequences and sequences with unknown functions.

Once the complete sequence of a bacterial genome is available, it can be searched using computers for similarities with other sequenced genomes. Such sequence comparisons can often be used to gain inferences about gene function. Because so much is known about gene function in *E. coli*, comparisons with other sequenced bacterial genomes are often very informative. For example, a comparison of the genomes of *Treponema pallidum*, the parasitic spirochete that causes syphilis, and *E. coli* shows that *T. pallidum* contains the genes that encode proteins involved in DNA replication and repair, transcription, and translation, but carries few genes encoding biosynthetic enzymes and transport proteins.

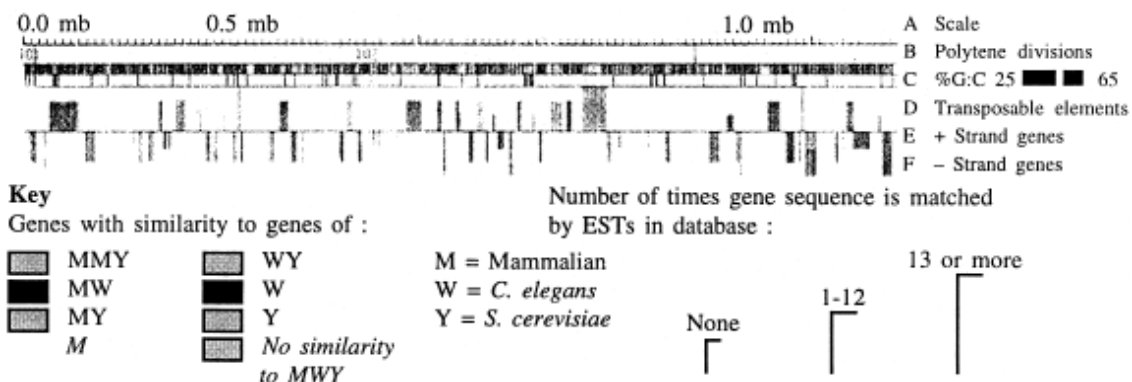


Fig. 5.10 Sequence-based map of chromosome 4 of *Drosophila*. The top line (A) gives map position in megabase pairs (1 mb = 1 million base pairs). The second panel (B) gives the polytene chromosome band number. The third panel (C) shows the percentage of G:C base pairs, and the fourth (D) shows the positions of transposable genetic elements. The bottom two panels show the positions of genes where transcription occurs with the plus strand as template (E) or with the minus strand as template (F). The color of each gene box in panels E and F indicates its similarity to genes of mammals, *C. elegans*, or *S. cerevisiae*, as shown by the key below the map, and the height of each gene box in panels E and F indicates the frequency of the sequence in the expressed-sequenced tag (EST) database for *Drosophila*.

Transposable genetic elements make up about 10 percent of the *Arabidopsis* genome, far less than the 50 to 80 percent of the corn genome estimated to be derived from transposable elements.

The next challenge is to determine the functions of the *Arabidopsis* gene products. There is where this little weed really shines as a model system. It is ideally suited for genetic dissections of biological processes. The goal of the *Arabidopsis* research community is to determine the function of all 25,498 genes in the next 10 years.

5.4.3 The Human Genome

Recall that the initial goals of the Human Genome project were (10 to map all the human genes, to construct detailed physical maps of all 24 chromosomes, and to sequence the entire genome by 2005. All of these goals are likely to be achieved ahead of schedule. With two first-draft sequences of the human genome already published in February 2001 a complete sequence of the euchromatic portion of the human genome will certainly be available long before 2005.

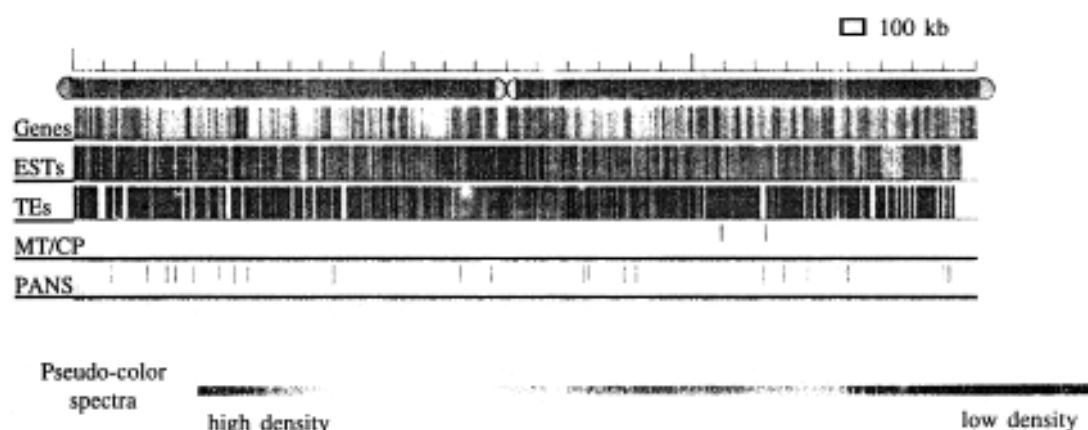


Fig. 5.11 Sequence-based map of chromosome 1 of *Arabidopsis*. Distance in 100-kb units is shown at the top. The chromosome is represented by the top bar, with sequenced regions in red and the unsequenced centromere and telomeres in blue. The densities of genes (top panel below chromosome), matches to expressed sequence tags (ESTs; second panel), transposable small nuclear RNAs (bottom panel) are color coded, with red representing the highest density and dark blue the lowest density

Progress in mapping the human genome has been excellent. Complete physical maps of chromosomes Y and 21 and detailed RFLP maps of the X chromosome and all 22 autosomes were published in 1992. By 1995, the genetic map contained markers separated by, on average, 200 kb. A detailed microsatellite map of the human genome was published in 1996, and a comprehensive map of 16,354 distinct loci was released in 1997. All of these maps have proven invaluable to researchers cloning genes based on their locations in the genome.

Unfortunately, the resolution of genetic mapping in humans is quite low—in the range of 1-10 mb. The resolution of fluorescent *in situ* hybridization (FISH) is also approximately 1 mb. Higher resolution mapping (down to 50 kb) can be achieved by **radiation hybrid mapping**, which is a modification of the somatic-cell hybridization mapping procedure. Standard somatic-cell hybridization involves the fusion of human cells and rodent cells growing in culture and the correction of human gene products with human chromosomes retained in the hybrid cells.

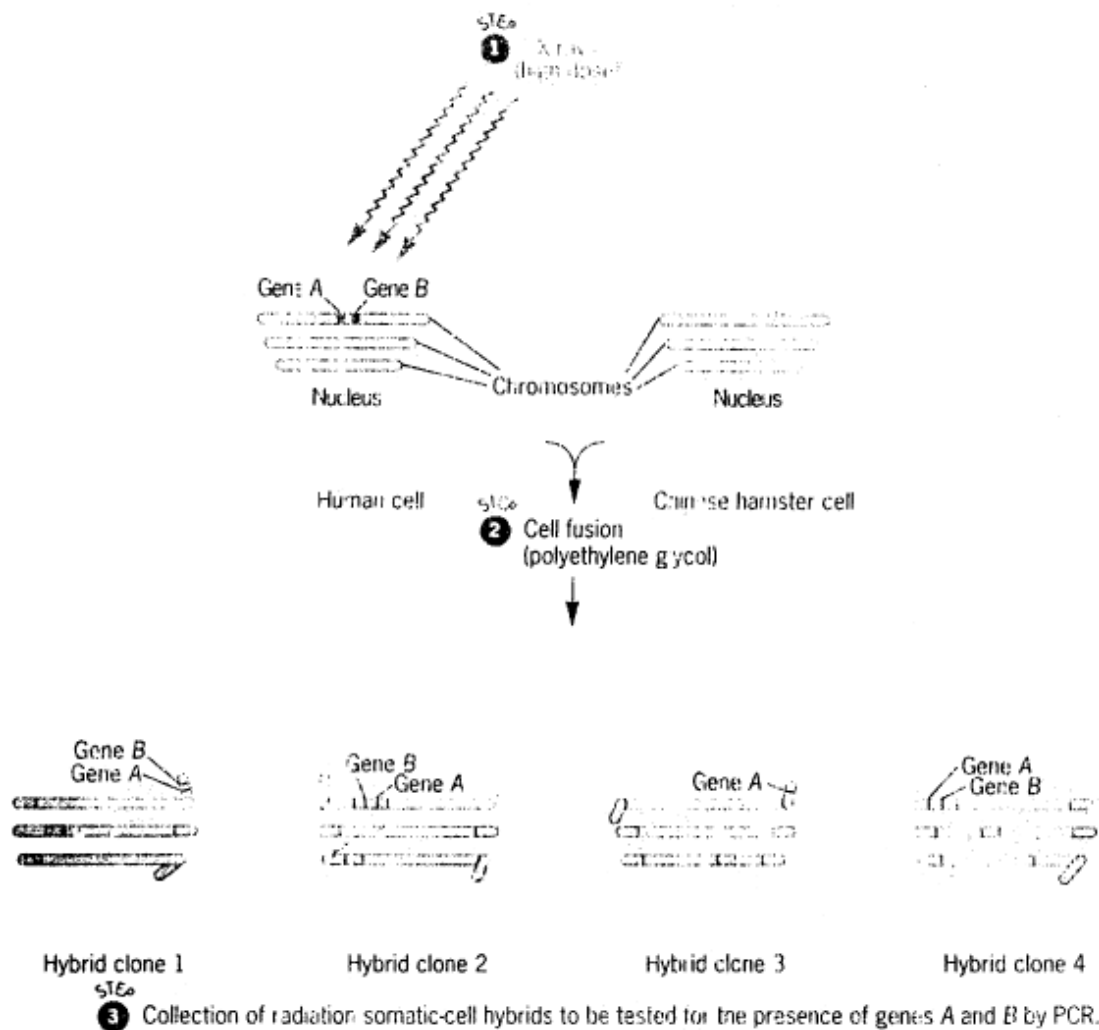


Fig. 5.12 The use of radiation somatic-cell hybrids for high-resolution mapping of the human genome. The rationale behind radiation hybrid maps is that the probability of an X ray-induced break between genes *A* and *B* is directly proportional to the distance between them on the DNA molecule. Note that genes *A* and *B* have remained together in hybrid clones 1, 2, and but were separated by an X ray-induced break during the formation of hybrid clone 3

Radiation hybrid mapping is done by fragmenting the prior to cell fusion (Fig. 5.12). The irradiated human cells are then fused with Chinese hamster (or other rodent) cells growing in culture, usually in the presence of a chemical such as polyethylene glycol to increase the efficiency of cell fusion. The human—Chinese hamster somatic-cell hybrids are then identified by growth in an appropriate selection medium.

Many of the human chromosome Chromosome 1 fragments become integrated into the Chinese hamster chromosomes during this process and are transmitted to progeny cells just like the normal genes in the Chinese hamster chromosomes. The polymerase chain reaction is used to screen a large panel of the selected hybrid cells for the presence of human genetic markers. Chromosome maps are constructed based on the assumption that the probability of an X-ray-induced break between two markers is directly proportional to the distances separating them in chromosomal DNA.

Several groups have used the radiation hybrid mapping procedure to construct high-density maps of the human genome. In 1997, Elizabeth Stewart and coworkers published a map of 10,478 STSs based on radiation hybrid mapping; their map of human chromosome 1 is shown in Fig. 5.13.

Whereas the gene mapping work advanced quickly, progress towards sequencing the human genome initially lagged behind schedule. However, that all changed rapidly beginning in 1998. During May of 1998, J. Craig Venter announced that he had formed a private company, Celera Genomics, with the goal of sequencing the human genome in just three

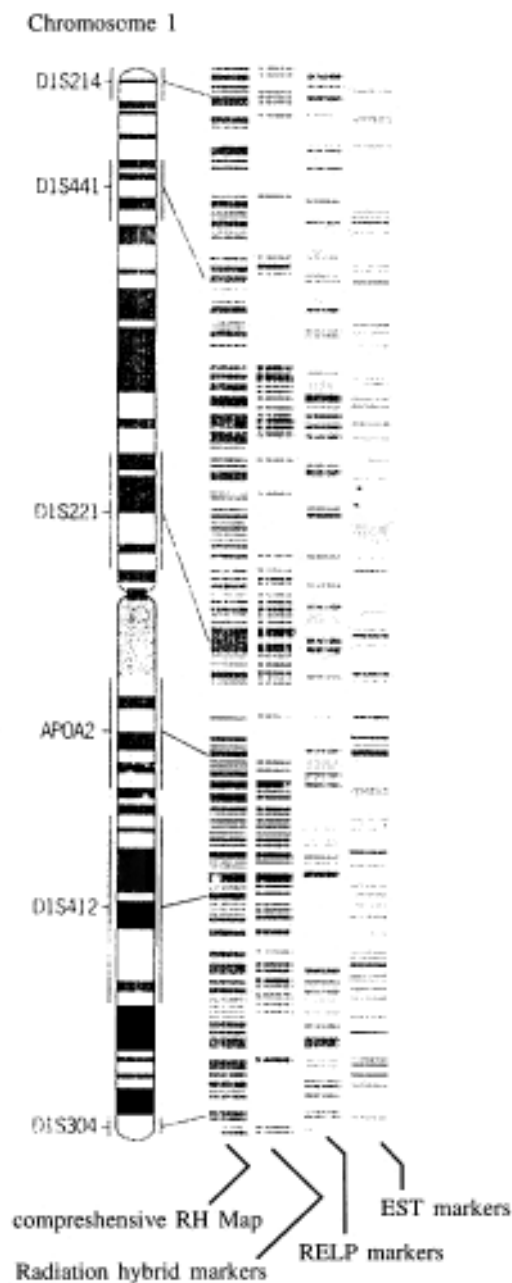


Fig. 5.13 A high-resolution radiation hybrid map of human chromosome 1. The cytogenetic map of chromosome is shown on the left with the locations of the comprehensive radiation hybrid map showing all the markers (red lines), the high confidence radiation hybrid markers (blue lines), the high confidence radiation hybrid markers (blue lines), the high confidence radiation hybrid markers (blue lines), the RFLP markers (green lines), and the ESTs (purple lines)

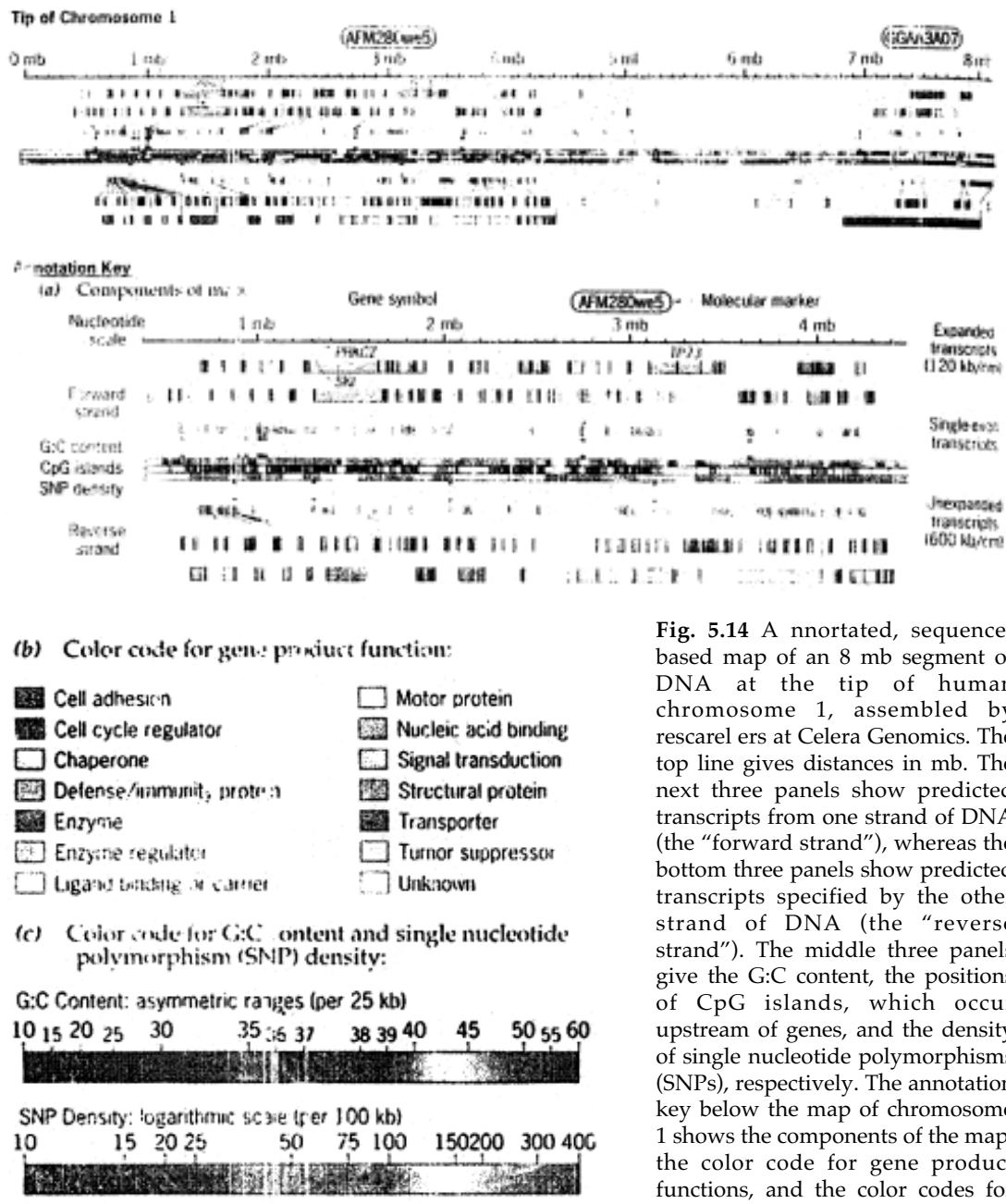


Fig. 5.14 An annotated, sequence-based map of an 8 mb segment of DNA at the tip of human chromosome 1, assembled by rescarleters at Celera Genomics. The top line gives distances in mb. The next three panels show predicted transcripts from one strand of DNA (the “forward strand”), whereas the bottom three panels show predicted transcripts specified by the other strand of DNA (the “reverse strand”). The middle three panels give the G:C content, the positions of CpG islands, which occur upstream of genes, and the density of single nucleotide polymorphisms (SNPs), respectively. The annotation key below the map of chromosome 1 shows the components of the map, the color code for gene product functions, and the color codes for G:C content and SNP density

years. Shortly thereafter, the leaders of the public Human Genome Project’s sequencing laboratories announced that they had revised their schedule and planned to complete the sequence of the human genome by 2003—two years earlier than originally proposed. From this point in time, everything accelerated.

The complete sequence of the first human chromosome—small chromosome 22—was published in December 1999. The complete sequence of human chromosome 21 followed in May of 2000. Then, with the intervention of the White house, Venter, of Celera Genomics, and Francis Collins, Director of the public Human Genome Project, agreed to publish first drafts of the sequence of the human genome at the same time. The Celera and public sequences were both published in February 2001. Figure 5.14 shows an annotated, sequence-based map of an 8mb segment at the tip of the short arm of human chromosome 1. This map illustrates the positions and orientations of known and predicted genes in one small portion of the human genome. For summary maps of the entire human genome, see the February 15, 2001, issue of *Nature* and the February 16, 2001, issue of *Science*.

The amount of information in these first drafts of the human genome was quite overwhelming including the sequence of over 2650 megabase pairs of DNA (over 2,650,000,000 bp). The human genome is more than 25 times the size of the previously sequenced *Drosophila* and *Arabidopsis* genomes, and it is eight times the sum of all previously sequenced genomes.

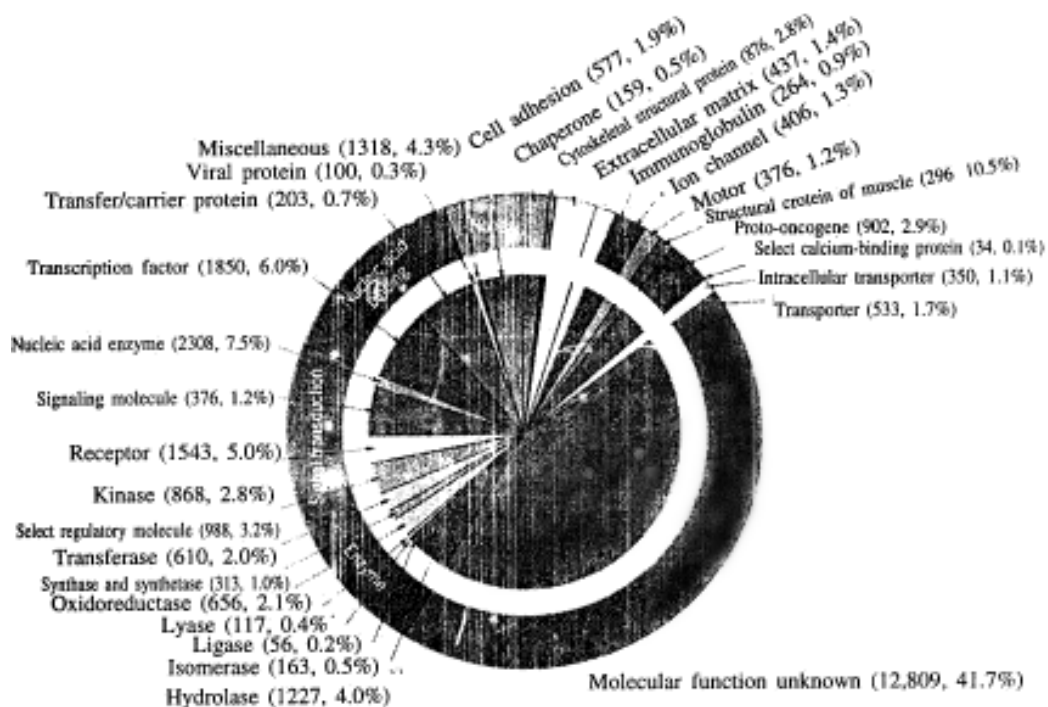


Fig. 5.15 Functional classification of the 26,283 genes predicted by Celera Genomics' first draft of the sequence of the human genome. Each sector gives the number and percentage of gene products in each functional class in parentheses. Note that some classes overlap: a proto-oncogene, for example, may encode a signaling molecule

The sequence of the human genome provided one surprise: there appear to be only about 30,000 to 35,000 genes rather than the estimated 50,000 to 120,000 genes suggested by earlier studies. The distribution of functions for the 26,383 genes predicted by the Celera sequence is shown in Figure 5.15. About 60 percent of the predicted proteins have similarities with proteins of other species whose genomes have been sequenced (Figure 5.16). Over 40 percent of the predicted human proteins share similarities with *Drosophila* and *C. elegans* Proteins. The picture is quite different for families of closely related proteins, which tend to perform important basic cellular functions. Only 94 of 1278 protein families predicted by the sequence of the human genome are specific to vertebrates. The rest have evolved from domains of proteins in distant ancestors, including prokaryotes and unicellular eukaryotes.

On average, there is one gene per 60 to 85 kb in the human genome, although there is some clustering of highly expressed genes in euchromatic regions of specific chromosomes. Exons make up only 1.1 percent of the genome, whereas introns make up 24 percent, with 75 percent of the genome being intergenic DNA. Of the intergenic DNA, at least 44 percent is derived from transposable genetic elements. The initial drafts of the human genome are far from complete, and the immediate goal will be to fill in the gaps in the genome and produce a finished sequence in the next year or so. The other major goal is to determine the structure and function of the human proteome (all of the proteins encoded by the human genome).

Knowledge of the nucleotide sequences of entire genomes has spawned the development of a new scientific discipline, *bioinformatics*, a fusion of computer science and biology, with the goal of developing new tools with which to analyze

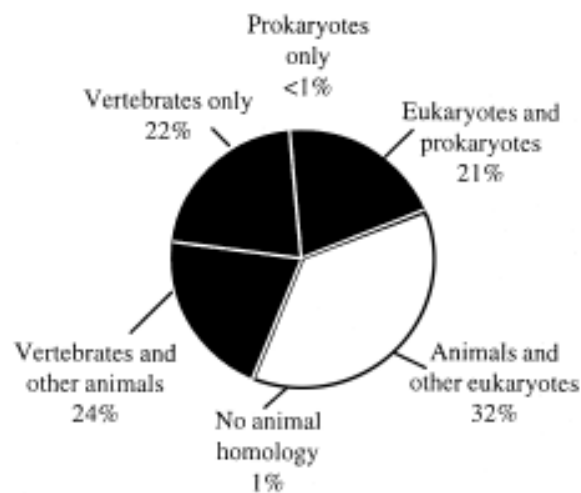


Fig. 5.16 Pie chart showing homology of predicted human proteins to proteins of other species for those where homologues were detected by computer searches of the public databases

the welth of data that genomics in providing. These new tools allow scientists to search genome databases for specific sequences or structural features, to compare various features of different genomes, and to make inferences about the evolution of genomes. Indeed, with the sequence of the human genome approaching completion, the question being asked is which genomes to sequence next—the mouse genome, the chimpanzee genome, and so on. One point is very clear: comparative genomics is providing unprecedented information about the evolution of species.

Unit 6 □ Cytogenetic Implications and Consequences of Structural Changes and Numerical Alterations of Chromosomes

Structure

6.1 Introduction

6.2 Chromosomes and cancer

6.3 Diseases associated with spontaneous chromosome aberrations

6.1 Introduction

Between 20—50% of human abortuses have been shown to have a chromosome abnormality (Rashad and Kerr, 1965; Thiede and Metcalfe, 1966; Carr, 1967; Larson and Titus, 1970; Kajii *et al.*, 1973). An even wider range has been reported (8—50%), but the studies have not been comparable (Carr, 1971a).

Analysis of over 350 cases compiled from several published series has demonstrated that 45,XO is the most frequent single anomaly, constituting about 20% of the cases. Triploidy is only slightly less frequent, i.e., about 15%, while tetraploidy has been demonstrated in 5%. Trisomies, as a group, have been found in about 50% of abortuses (E, 15%; G, 15%; D, 10%; C, 5%; A, B, F, 5%). The remainder are mosaics, or translocations (Carr, 1965, 1967; Inhorn 1967; Larson and Titus, 1970). only rarely is autosomal monosomy found (Kajii *et al.*, 1973).

It should be pointed out that within the E group, trisomy 16 and not trisomy 18 largely comprises this number (Carr, 1967; Waxman *et al.*, 1967). Use of recently developed banding techniques have shown trisomies for chromosomes 2, 3, 4, 6, 7, 8, 9, 10, 14, 15, 16, 18, 21 and 22 (Lauritsen *et al.*, 1972; Kajii *et al.*, 1973).

It is likely that findings in spontaneous abortuses do not necessarily reflect the frequency of clinical anomalies at conception since the more lethal ones probably never survived past a few cell divisions. This may explain the absence of viable trisomic states for A, B, most C, 14, 15, 16, 17, and F chromosomes being a more likely explanation than low frequency of meiotic or mitotic error for these chromosomes. The apparent high lethality of the 45,XO embryo cannot be explained. There is suggestions that embryos with chromosomal abnormalities are more likely to be aborted earlier than those with normal karyotypes (Carr, 1965, 1967; Szulman, 1965; Dhadi *et al.*, 1970).

The gross appearance of abortuses have some correlation with their chromosome status. A recognizable fetus is most frequent in the 45, XO group. Not uncommonly they can be recognized by cystic hygromas (neck blebs) in older fetuses. Triploid abortuses characteristically exhibit hydatidiform degeneration of villi and only rarely contain an embryo (Szulman, 1965; Carr, 1965, 1971b; Singh and Carr, 1967; Boue *et al.*, 1967). Trisomic abortuses do not have any specific phenotype with the possible exception of those having D-group trisomy, which not uncommonly have facial clefts. Presumably most of these are 13 trisomics (Roux, 1970). Kajii *et al.* (1973) found no example, however, of trisomy 13 in their large series.

When analyzed for mean maternal age, polyploid and XO abortuses have been found to be from younger mothers while trisomics have been from older mothers (Carr, 1965, 1971a; Szulman, 1965; Kerr *et al.*, 1966). However, the mean maternal age for 45,XO abortuses has been higher than that of survivors (Dhadi *et al.*, 1970). Arakaki and Waxman (1970) found an increase in mean maternal age in cases of 16 trisomy abortuses.

In those couples having a history of two or more spontaneous abortions, 1 in 26 couples was found to have a translocation. This contrasts with the 0.4% found in the general population (Lucas *et al.*, 1972).

6.2 Chromosomes and cancer

Cancer cells may have bizarre karyotypes which may be hypodiploid, hyperdiploid, triploid, hypertriploid, hypotriploid, etc. Many unusual structural abnormalities have been described. Cells having over 1000 chromosomes have been documented. On the other hand, cancer cells have been described with normal karyotype and no evidence of structural abnormalities. While one may conclude that not all neoplasia is associated with gross chromosomal anomalies, one cannot exclude point mutations, gene deletions or duplications, or hidden rearrangements.

Review of chromosomal alterations concerning specific tumors is beyond the scope of this review and the reader is referred to Cervenka and Koulischer (1973).

A. Acute leukemias

In no form of acute leukemias have any specific chromosome abnormalities been described and, in at least half the cases, normal karyotypes have been found (Sandberg *et al.*, 1968; Krogh-Hensen, 1969; Whang Peng *et al.*, 1969). Furthermore, ostensibly identical clinical types of acute leukemia may manifest different chromosomal patterns. Karyotypic changes, when present, are confined to the leukemic cells of the marrow or other organs. Long-term culture of

leukocytes from the blood of patients with acute leukemias is rarely successful. When an abnormal karyotype is discerned, it seems to exhibit more hypoploid cell lines while acute lymphoblastic leukemias have more hyperploid lines. Karyotype analysis alone cannot be employed either for diagnosis or for prognosis concerning survival (Cervenka and Koulischer, 1973). During remission, aneuploid cells may disappear from the marrow only to reappear on relapse (Sandberg and Hossfeld, 1970).

B. Chronic myelogenous leukemia

In 1960, Nowell and Hungerford found deletion of part of the long arm of a G chromosome associated with chronic myelogenous leukemia (CML). This unique structural abnormality, termed the Philadelphia chromosome (Ph^1), has been noted in over 90% of patients with CML (De Nava, 1969). Chronic myelogenous leukemia without the ph^1 chromosome and the ph^1 chromosome without chronic myelogenous leukemia have been thoroughly reviewed by Cervenka and Kouliche (1973).

The Ph^1 chromosome represents deletion with translocation to the long arm of a chromosome no. 9 (see Chapter 3). With the use of quinacrine mustard fluorescent technique, it has been shown to be a G_{22} chromosome (Caspersson *et al.*, 1971; O'Riordan *et al.*, 1971). Its occurrence is limited to hematopoietic cells of all the granulocytic, erythrocytic, and megakaryocytic types (Tough *et al.*, 1963; Clein and Flemans, 1966). Other tissues, such as skin fibroblasts, do not contain the ph^1 chromosome.

The best technique for demonstration of the ph^1 chromosome is by direct study of bone marrow (Sandberg and Hossfeld, 1970).

It is an acquired, not an inherited, characteristic as demonstrated by its presence in only one of monozygotic twins with chronic myelogenous leukemia and not in the healthy co-twin (Jacobs *et al.*, 1966, Kosenow and Pfeiffer, 1969). An unusual subgroup of ph^1 -positive CML patients are males who are missing the Y chromosome in all or in a portion of their marrow cells. However, fibroblasts and blood lymphocytes contain the Y chromosome (Lawler and Galton, 1966; Pedersen, 1968). Two or more ph^1 chromosomes appearing in marrow cells either heralds or accompanies the transformation of CML to a blastic phase (Smalley, 1966).

A ph^1 -like chromosome has been found in a small proportion of marrow cells of patients with acute myeloblastic leukemia, polycythemia, thrombocytopenia, myeloid metaplasia with myelofibrosis, and erythroleukemia (Sandberg and Hossfeld, 1970). Khan (1973) reported two ph^1 -like chromosomes in acute myeloid leukemia.

C. Solid tumors

Most malignant tumors have aneuploid karyotypes ranging from hypodiploidy to extreme hyperdiploidy. Human tumor cell populations are clonal in nature, some tumors having but a single clone, others of two or more. No consistent cytogenetic findings have been described, but various markers have been noted, for example, a missing no. 22 chromosome in meningiomas (Zaftg and Singer, 1967; Mark *et al.*, 1972) and microchromosome in various neurogenic tumors (Cox *et al.*, 1965; Levan *et al.*, 1968, Kucheria, 1968). Metastatic cells tend to have a higher ploidy and more variability in chromosome number (Sandberg *et al.*, 1967). In about 50% of the cases, abnormal ("marker") chromosome have been found in metastatic cancer cells. In general, there is a tendency toward relatively few chromosomes with distally placed centromeres, i.e., fewer B-, D-, and G-group chromosome and more A-3, C-group, and E 16 chromosomes (Atkin, 1970). Manolov and Manolova (1972) described a marker band in a chromosome 14 in Burkitt's lymphoma.

Precancerous lesions, largely of the uterine cervix, have shown that dysplastic lesions exhibit chiefly pseudo- or near diploid karyotypes while carcinoma in situ shows an increase in ploidy and aberrations. Invasive carcinomas exhibit near diploid patterns, showing that progression does not depend on high chromosome counts (Atkin *et al.*, 1967). Benign tumors have normal diploid karyotypes.

D. Waldenstrom's macroglobulinemia

Waldenstrom (1944) described a disorder characterized by intractable anemia and increased amounts of macroglobulin in serum, accompanied by fatigue, epistaxis, gingival hemorrhage, disturbances in vision, moderate lymphadenopathy, high sedimentation rate, and bone marrow lymphocytosis (Kok *et al.*, 1963). It is presently classified in the group of gammopathies. The disease usually appears after age 40 and is more frequent in males. Its relationship to lymphosarcoma and leukemia is not clear but patients diagnosed as having the disease sometimes develop chronic lymphatic leukemia or lymphoma.

Bottura *et al.* (1961) first described the presence of 47 chromosomes in about 50% of the cells, the supernumerary being about the size of an A-group chromosome. This finding was soon confirmed by German *et al.* (1961) and Benirschke *et al.* (1962), who employed the term "W" chromosome.

The morphology of the marker chromosome is not constant. It usually has been large with the centromere varying from metacentric to subterminal, but in some cases it has been as small as an F-group chromosome (Spengler *et al.*, 1966). The marker has been noted in both marrow and in peripheral cells in form 0—50% of cells (De Nava, 1969).

The abnormality is apparently acquired. Spengler *et al.* (1966) demonstrated the marker in one monozygotic twin who had Waldenstrom's macroglobulinemia but not in his normal co-twin. Interesting also are the findings of Lustman *et al.* (1968), who described an affected female with the marker whose otherwise healthy son had a normal karyotype but had an elevated γ -globulin peak. Elves and Brown (1968) described the marker in 4 of 6 relatives of a patient with the disorder. Only one of the individuals had an elevated γ_1 fraction.

6.3 Diseases associated with spontaneous chromosome aberrations

At least seven inherited diseases have been found to be associated with spontaneous chromosome aberrations and increased frequency of leukemia or other neoplasias. The chromosome aberrations consist of gaps (achromatic regions), chromatid and chromosome breaks, fragments, reunion or translocation figures, ring chromosomes, and dicentric chromosomes. It should be emphasized that chromosome breakage may be very rarely seen in cells of ostensibly normal people. The enzyme deficiencies in the inherited disorders may result either in increased frequency in which openings appear in the DNA strands or in decreased speed with which such breaks are healed. Higurashi and Conen (1973) demonstrated greater *in vitro* chromosomal sensitivity in several of these disorders.

Fanconi's anemia, inherited as an autosomal recessive trait, is characterized by generalized skin pigmentation, pancytopenia with marrow hypoplasia, thumb and radius anomalies, hypogenitalism, and microcephaly (Fanconi, 1967). In 1964, Schroeder *et al.* noted that more than 40% of analyzed metaphases from peripheral blood cultures of patients with Fanconi's anemia exhibited chromatid gaps and breaks and chromosomal rearrangements. Direct bone marrow preparations have shown about 10% aberrant metaphases, usually involving B and C group chromosomes (Hirschman *et al.*, 1969; Shahid *et al.*, 1972). Of 41 cases subsequently studied, 36 were found to have similar findings. Among 170 known cases, four have terminated in leukemia and one had skin cancer (Swit and Hirschhorn, 1966; Swift, 1971). Heterozygotes have an increased frequency of leukemia (Gmyrek *et al.*, 1967; Swift, 1971). Occasionally quadriradials and dicentric forms are noted but far less frequently than in Bloom's syndrome (vide infra).

Bloom's syndrome, consisting of growth retardation, sensitivity to sunlight, and telangiectatic erythema, was reported by German (1969) to have chromosome breaks. Of 35 cases, four were found to have subsequently developed leukemia

or cancer, especially gastrointestinal. Cell lines with an abnormal karyotype have been described in cultured fibroblasts from a patient with Bloom's syndrome (Rauh and Soukup, 1968).

Quadriradial figures, i.e., a four-armed figure derived from two chromosomes, each arm consisting of sister chromatids of one of two homologous chromosomes. The autosomes most often involved are No. 1 and either No. 19 or 20. Asymmetric dicentric chromosomes, triradials, and abnormal new monocentric chromosomes can also be found. Heterozygotes may have the same types of figures, but less frequently than the homozygote.

Ataxia-telangiectasia inherited as an autosomal recessive trait is characterized by retarded growth, progressive cerebellar ataxia, telangiectasia especially about the face and bulbar conjunctiva, increased sonopulmonary infections, and decreased immunoglobulins (especially IgA and IgE). Approximately 10% develop lymphomas (Pfeiffer, 1970). Hecht *et al.* (1966) reported a high frequency of in vitro chromosome breakage, a finding supported by Groop and Flatz (1967), Pfeiffer (1970), and German (1972).

Lesser well-documented associations are with glutathione reductase deficiency anemia, pernicious anemia, Kostmann's agranulocytosis (Schroeder and Kurth, 1971), and possibly xeroderma pigmentosum (German *et al.*, 1970).

Matsaniotis *et al.* (1966) found approximately 20% aberrant cells from direct bone marrow preparations of a baby with Kostmann's agranulocytosis. Krogh-Jensen and Friis-Møller (1967) and Bottura and Continho (1988) described in vivo demonstration of chromosomal aberration in untreated pernicious anemia. The evidence for dominantly inherited glutathione reductase deficiency is less solid and seems to depend on the stage of the disease (Hampel *et al.*, 1969).

German *et al.* (1970) detected a tendency toward the formation of pseudodiploid clones in cultured fibroblasts from a patient with xeroderma pigmentosum, an autosomal recessively inherited disorder, in which there is a proclivity toward development of skin cancer. Failure of DNA repair following ultraviolet light exposure has been demonstrated. Repair failure results from deficiency of ultraviolet-specific endonuclease.

Unit 7 □ Microbial Genetics

Structure

- 7.1 Introduction
- 7.2 Bacterial mutation
- 7.3 Conjugation—method of genetic recombination in bacteria
- 7.4 Transformation—Process leading to genetic recombination in bacteria
- 7.5 Transduction is virus—mediated bacterial DNA transfer

7.1 Introduction

Main constituent in microbial genetics are the bacteria. Bacteria reproduce asexually. However **parasexual** reproduction is also found among bacteria. In fact, genetic information is transferred from one bacterium to another by three totally distinct processes—**transformation, conjugation and transduction**.

Gene transfer in Bacteria :

Transformation (Definition) : This is the process by which a donor DNA molecule is taken up from the external environment and incorporated into the genome of a recipient cell.

Conjugation (Definition): This is the process by which bacterial cells make direct contact with each other, and DNA is transferred from one cell (the donor) to the other (the recipient cell).

Transduction (Definition) : This is the process by which DNA is transferred from one bacterial cell to another by a bacterial virus, or bacteriophage.

7.2 Bacterial mutation

Infection by the bacteriophage leads to the reproduction of the virus at the expense of the bacterial cell, which is lysed or destroyed. If a plate of *E. coli* is homogeneously sprayed with T1, almost all cells are lysed. Rare *E. coli* cells, however, survive infection and are not lysed. If these cells are isolated and established in pure culture, all their descendants are *resistant* to T1 infection. This might be argued that the mutations responsible for T1 resistance were “induced” by the presence of the T1 viruses, and that, in the absence of the T1 viruses, the mutations would not have occurred. In 1943 Salvador Luria and Max Delbruck elegantly proved that such T1-resistant cells result from **spontaneous mutation**.

Bacterial cells that bear spontaneous mutations, such as T1 resistance, can be isolated and established independently from the parent strain by means of various selection techniques. As a result, one can now induce and isolate mutations for almost any desired characteristic. Because bacteria and the viruses that infect them are haploid, all mutations are expressed directly in the descendants of mutant cells, adding to the ease with which these microorganisms can be studied.

Bacteria are grown in either a liquid culture medium or in a petri dish on a semisolid agar surface. If the nutrient components of the growth medium are very simple and consist only of an organic carbon source (such as a glucose or lactose) and a variety of inorganic ions, including Na^+ , K^+ , Mg^{++} , and NH_4^+ , present as inorganic salts, it is called **minimal medium**. To grow on such a medium, a bacterium must be able to synthesize all essential organic compounds (e.g., amino acids, purines, pyrimidines, sugars, vitamins, and fatty acids). A bacterium that can accomplish this remarkable biosynthetic feat—one that we ourselves cannot duplicate—is termed a **prototroph**. It is said to be wild type for all growth requirements and can grow on minimal medium. On the other hand, if a bacterium loses, through mutation, the ability to synthesize one or more organic components, it is said to be an **auxotroph**. For example, a bacterium that loses the ability to make histidine is designated as a *bis⁻* auxotroph, as opposed to its prototrophic *bis⁺* counterpart. For the *bis⁻* bacterium to grow, this amino acid must be added as a supplement to the minimal medium. The medium that has been extensively supplemented is referred to as complete medium.

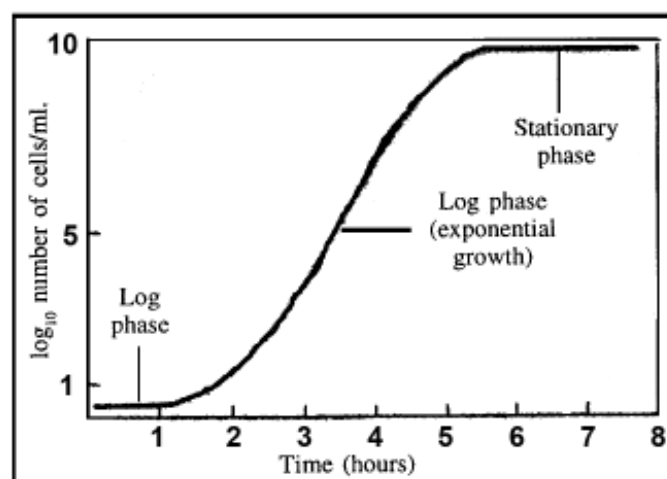


Fig. 7.1 A typical bacterial population growth curve illustrating the initial lag phase, the subsequent log phase where exponential growth occurs, and the stationary phase that occurs when nutrients are exhausted

To study mutant bacteria in a quantitative fashion, an inoculum (e.g., 0.1 ml, 1.0 ml) of bacteria is placed in liquid culture medium. The bacteria exhibit a characteristic growth pattern, as illustrated in Figure 7.1. Initially, during the lag phase, growth is slow. Then, a period of rapid growth follows called the **log phase**, during which cells divide many times with a fixed time interval between cell divisions, resulting in logarithmic growth. When the bacteria reach a cell density of about 10^9 cells per milliliter, nutrients and oxygen become limiting and cells enter the **stationary phase**. As the doubling time during the log phase may be as short as 20 minutes, an initial inoculum of a few thousand cells added to the culture can easily achieve a maximum cell density overnight.

Once cells are grown in liquid medium they can be quantitated. First, the bacteria are plated on (transferred to) semi-solid medium in a petri dish where, following incubation and many divisions, each cell gives rise to a colony visible on the surface of the medium. From the number of colonies that subsequently grow, it is possible to estimate the number of bacteria present in the original culture. If the number of colonies is too great to count, then serial dilutions of the original liquid culture can be made and plated, until the colony number is reduced to the point where it can be counted (Figure 7.2).

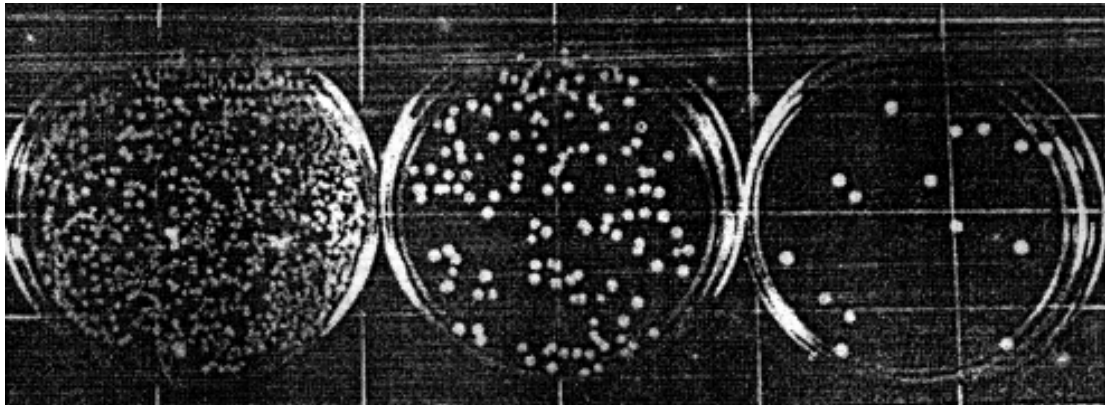


Fig. 7.2 Results of the serial dilution technique and subsequent culture of bacteria. Each of the dilutions varies by a factor of 10. Each colony was derived from a single bacterial cell

For example, assume that the three petri dishes in Figure 7.2 represent dilutions of 10^{-3} , 10^{-4} , and 10^{-5} , respectively (left to right). We select the dish where there are few enough colonies to be accurately counted. Because each colony presumably arose from a single bacterium, the number of colonies times the dilution factor represents the number of bacteria in the initial milliliter. In this case, the dish farthest to the right contains 15 colonies. Since it represents a dilution of 10^{-5} , we can estimate the initial number of bacteria to be 15×10^5 per milliliter. Calculations such as these are useful in a number of studies.

7.3 Conjugation is one of the methods of genetic recombination in bacteria

7.3.1 Introduction

Development of techniques that allowed the identification and study of bacterial mutations led to detailed investigations of the arrangement of genes on the bacterial chromosome. Such studies began in 1946 when Joshua Lederberg and Edward Tatum showed that bacteria undergo conjugation, a parasexual process in which the genetic information from one bacterium is transferred to and recombined with that of another bacterium. Like meiotic crossing over in eukaryotes, genetic recombination in bacteria provided the basis for the development of methodology for chromosome mapping. Note that the term **genetic recombination**, as applied to bacteria and bacteriophages, leads to the *replacement* of one or more genes present in one strain with those from a genetically distinct strain. While this is somewhat different from our use of genetic recombination in eukaryotes, where the term describes crossing over that results in *reciprocal exchange events*, the overall effect is the same : Genetic information is transferred from one chromosome to another, resulting in an altered genotype. Two other phenomena that result in the transfer of genetic information from one bacterium to another, **transformation** and **transduction**, have also served as a basis for determining the arrangement of genes on the bacterial chromosome.

Lederberg and Tatum's initial experiments were performed with two multiple auxotroph strains (nutritional mutants) of *E. coli* K12. As shown in Figure 7.3, Strain A required methionine (met) and biotin (bio) in order to grow, whereas strain B required threonine (thr), leucine (leu), and thiamine (thi). Neither strain would grow on minimal medium. The two strains were first grown separately in supplemented media, and then cells from both were mixed and grown together for several more generations. They were then plated on minimal medium. Any bacterial cells that grew on minimal medium are **prototrophs** (wild-type bacteria that did not need nutritional supplements). It was highly improbable that any of the cells that contained two or three mutant genes would undergo spontaneous mutation simultaneously at two or three gene sites. Therefore, any prototrophs recovered must have arisen as a result of some form of genetic exchange and recombination.

In this experiment, prototrophs were recovered at a rate of $1/10^7$ (10^{-7}) cells plated. The controls for this experiment involved separate plating of cells from strains A and B on minimal medium. No prototrophs were recovered. On the basis of these observations, Lederberg and Tatum proposed that, while the events were indeed quite rare, genetic recombination had occurred.

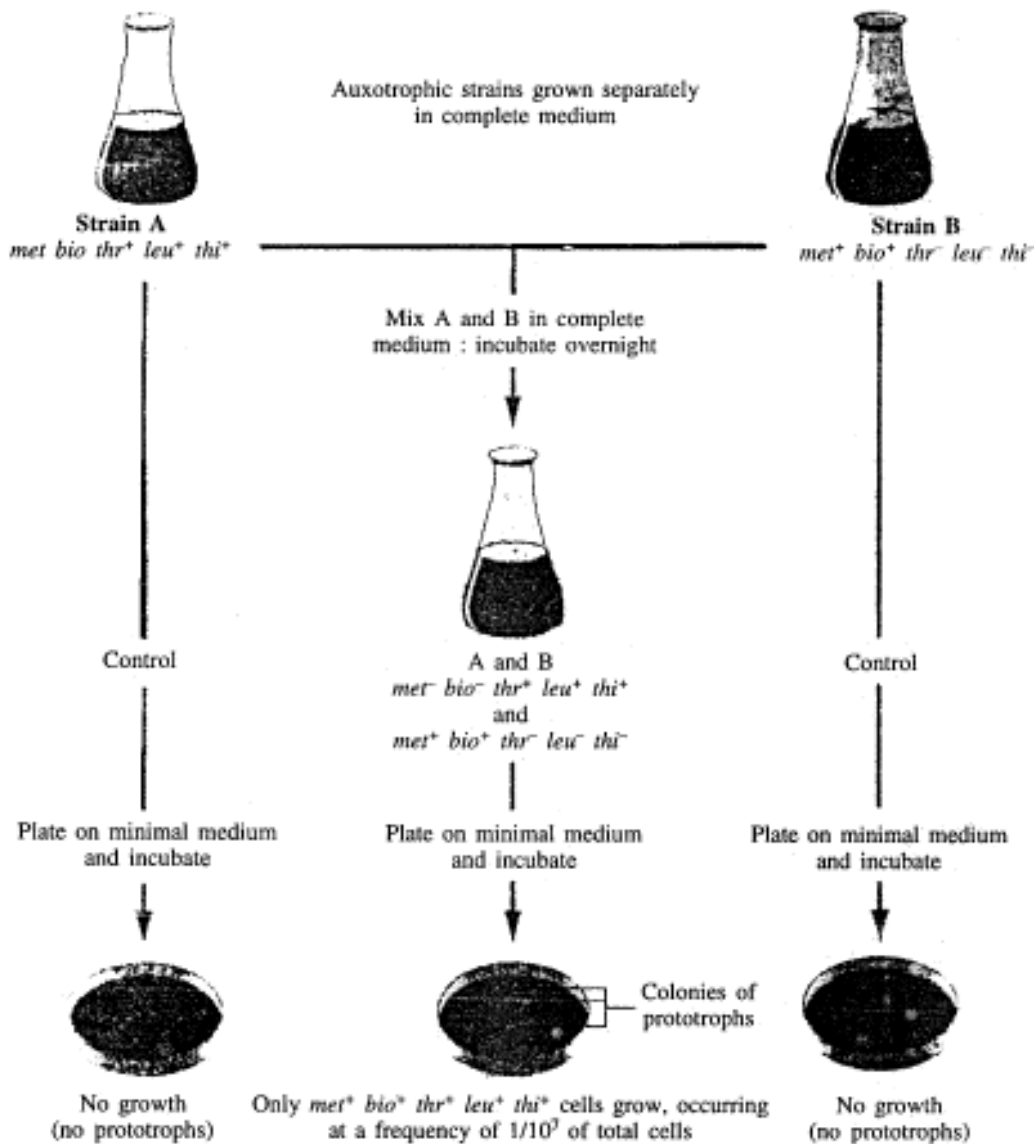


Fig. 7.3 Genetic recombination involving two auxotrophic strains producing prototrophs. Neither auxotroph will grow on minimal medium, but prototrophs will, suggesting that genetic recombination has occurred

7.3.2 F⁺ and F⁻ Bacterial strains

Lederberg and Tatum's findings were soon followed by numerous experiments designed to elucidate the genetic basis of conjugation. It quickly became evident that different strains of bacteria were involved in a unidirectional

transfer of genetic material. When cells serve as donors of parts of their chromosomes, they are designated as **F⁺ cells** (F for “fertility”). Recipient bacteria receive the donor chromosome material (now known to be DNA), and recombine it with part of their own chromosome. They are designated as **F⁻ cells**.

7.3.3 Conjugation in bacteria

It was established subsequently that cell contact is essential to chromosome transfer. Support for this concept was provided by Bernard Davis, who designed a U-tube in which to grow F⁺ and F⁻ cells (Figure 7.4). At the base of the tube was a glass filter with pores large enough to allow passage of the liquid medium, but too small to allow the passage of bacteria. The F⁺ cells were placed on one side of the filter and F⁻ cells on the other side. The medium was moved back and forth across the **filter** so that the bacterial cells essentially shared a common medium during incubation. Samples from both sides of the tube were then plated independently on minimal medium, but no prototrophs were found. Davis concluded that *physical contact is essential to genetic recombination*. Such physical interaction is the initial stage of the process of conjugation and is mediated through a conjugation tube called the F, or sex, pilus. Bacteria often have many pili, which are microscopic tube-like extensions of the cell. Different types of pili perform different cellular functions, but all pili are involved in some way with adhesion. After contact has been initiated between mating pairs via the F pili (Figure 7.5), transfer of DNA begins.

Later evidence established that F⁺ cells contained a fertility factor (called the F factor) that confers the ability to donate part of their chromosome during conjugation. Experiments by Joshua and Esther Lederberg and by William Hayes and Luca Cavalli-Sforza showed that certain conditions could eliminate the F factor in otherwise fertile cells. However, if these “infertile” cells were then grown with fertile donor cells, the F factor was regained.

The conclusion that the F factor is a mobile element was further supported by the observation that, following conjugation and genetic recombination, recipient cells always become F⁺. Thus, in addition to the *rare* cases of transfer of genes from the bacterial chromosome (genetic recombination), the F factor itself is passed to *all* recipient cells. On this basis, the initial crosses of Lederberg and Tatum (Figure 7.3) may be designated.

STRAIN A	X	STRAIN B
F ⁺		F ⁻
DONOR		RECIPIENT

Isolation of the F factor confirmed these conclusions. Like the bacterial chromosome, though distinct from it the F factor has been shown to consist of

a circular, double-stranded DNA molecule, equivalent to about 2 percent of the bacteria] chromosome (about 100,000 nucleotide pairs). Contained in the F factor, among others, are 19 genes, the products of which are involved in the transfer of genetic information (*tra* genes). These include those essential to the formation of the sex pilus.

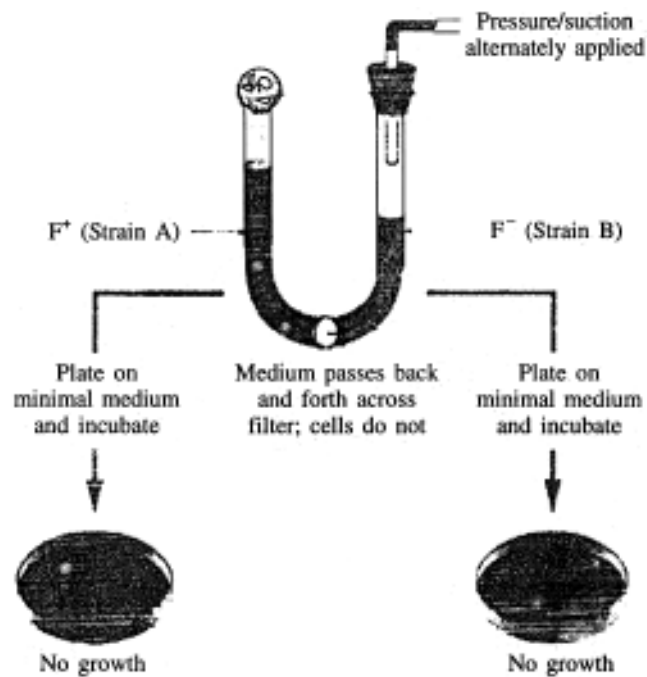


Fig. 7.4 When strain A and B auxotrophs are grown in a common medium but separated by a no genetic recombination occurs and no prototrophs are produced. This apparatus is called a Davis U-tube

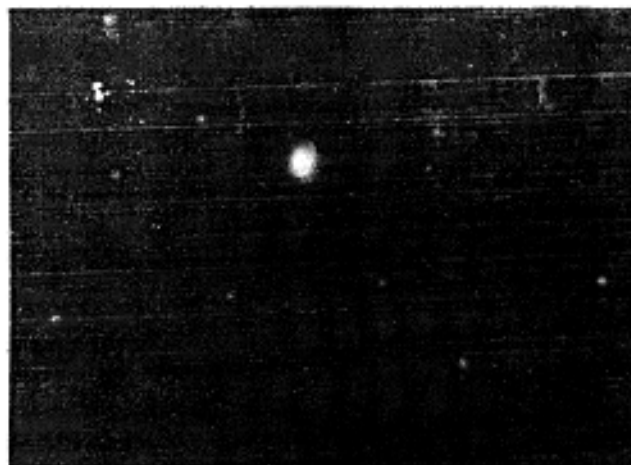


Fig. 7.5 An electron micrograph of conjugation between an F^+ *E. coli* cell. The sex pilus linking them is clearly visible

As we soon shall see, the F⁻ factor is in reality an autonomous genetic unit referred to as a plasmid. However, in our historical coverage of its discovery, we will continue in this chapter to refer to it as a "factor."

It is believed that the transfer of the F factor during conjugation involves separation of two strands of the F factor double helical DNA and the movement of one of the two strands into the recipient. The other strand remains in the donor cell. Both of these parental strands serve as templates for DNA replication, resulting in two intact F factors, one in each of the two cells. Both cells are now F⁺ (See in Figure 7.6).

To summarize, *E. coli* cells may or may not contain the F factor. When it is present, the cell is able to form a sex pilus and potentially serve as a donor of genetic information. During conjugation, a copy of the F factor is almost always transferred from the F⁺ cell to the F⁻ recipient, converting it to the F⁺ state. The question remains as to exactly how a very low percentage of F⁻ cells undergo genetic recombination. The answer awaited further experimentation. Subsequent

discoveries not only clarified how genetic recombination occurs but also defined a mechanism by which the *E. coli* chromosome could be mapped.

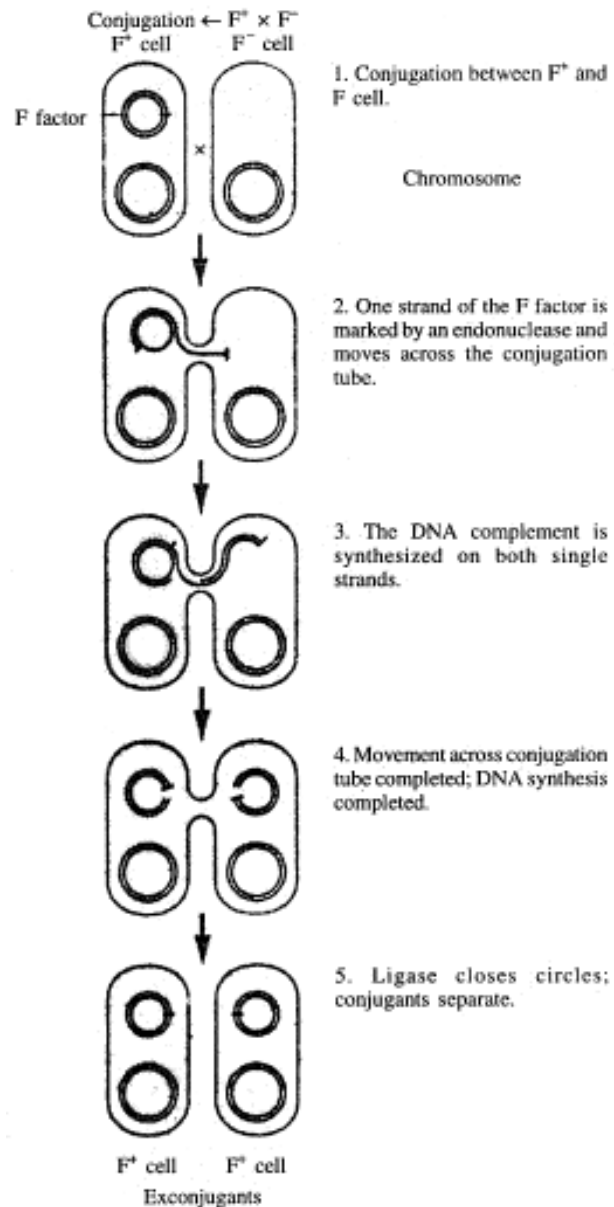
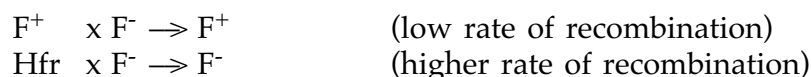


Fig. 7.6 An F⁺ × F⁻ mating demonstrating how the recipient F⁻ cell is converted to F⁺. During conjugation, the DNA of the F factor is replicated with one new copy entering the recipient cell converting it to F⁺. The black bar has been added to the F factors to follow their clockwise rotation during replication

7.3.4 Hfr bacteria and chromosome mapping

In 1950, Cavalli-Sforza treated an F⁺ strain of *E. coli* K12 with nitrogen mustard, a potent chemical known to induce mutations. From these treated cells, he recovered a strain of donor bacteria that underwent recombination at a rate of 1/10⁴ (10⁻⁴), 1000 times more frequently than the original F⁺ strains. In 1953, William Hayes isolated another strain demonstrating a similar elevated frequency. Both strains were designated **Hfr**, or **high-frequency recombination**. Because Hfr cells behave as chromosome donors, they are a special class of F⁺ cells.

Another important difference was noted between Hfr strains and the original F⁺ strains. If the donor is an Hfr strain, recipient cells, while sometimes displaying genetic recombination, never become Hfr; that is, they remain F⁻. In comparison, then,



Perhaps the most significant characteristic of Hfr strains is the nature of recombination. In any given strain, certain genes are more frequently recombined

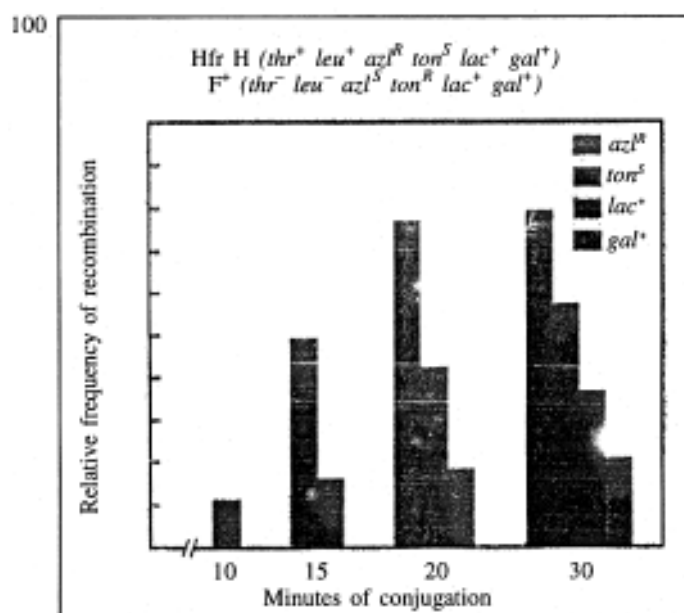


Fig. 7.7 The progressive transfer during conjugation of various genes from a specific Hfr strain of *E. coli* to an F strain. Certain genes (*azi* and *ton*) are transferred sooner than others and recombine more frequently. Others (*lac* and *gal*) take longer to be transferred and recombine with a lower frequency. Others (*thr* and *leu*) are always transferred and are used in the initial screen for recombinants

than others, and some do not recombine at all. This *nonrandom pattern of gene transfer* was shown to vary from Hfr strain to Hfr strain. While these results were puzzling, Hayes interpreted them to mean that some physiological alteration of the F factor had occurred, resulting in the production of Hfr strains of *E. coli*.

In the mid-1950s, experimentation by Ellie Wollman and Francois Jacob explained the differences between cells that are Hfr and those that are F⁺ and showed how Hfr strains allow genetic mapping of the *E. coli* chromosome. Wollman and Jacob first incubated a culture containing a mixture of an Hfr strain and an F strain. To facilitate the recovery of only recombinants, the Hfr strain was sensitive to an antibiotic while the recipient strain was resistant. At various intervals, the researchers removed samples and placed them in a blender. The shear forces created in the blender separated conjugating bacteria so that the transfer of the chromosome was effectively terminated. To assay the cells for genetic recombination following the blender treatment, they were grown on medium containing the antibiotic in order to ensure the recovery of only recipient cells.

7.3.5 Interrupted mating technique

This process, called the interrupted mating technique, demonstrated that specific genes of a given Hfr strain were transferred and recombined sooner than others. Figure 7.7 illustrates this point. During the first 8 minutes after the two strains were initially mixed, no genetic recombination could be detected. In cells assayed at about 10 minutes, recombination of the *azi*^R gene could be detected, but no transfer of the *ton*^S, *lac*⁺, or *gat*⁺ genes was noted. By 15 minutes, 70 percent of the recombinants were *azi*^R; 30 percent were now also *ton*^S; but none was *lac*⁺ or *gat*. Within 20 minutes, the *lac*⁺ gene was found among the recombinants; and within 30 minutes, *gat*⁺ was also being transferred. Wollman and Jacob had demonstrated an *oriented transfer of genes* that was correlated with the length of time conjugation was allowed to proceed.

It appeared that the chromosome of the Hfr bacterium was transferred linearly and that the gene order and distance between genes, as measured in minutes, could be predicted from such experiments (Figure 7.8). This information served as the basis for the first genetic map of the *E. coli* chromosome. "Minutes" in bacterial mapping are equivalent to "map units" in eukaryotes.

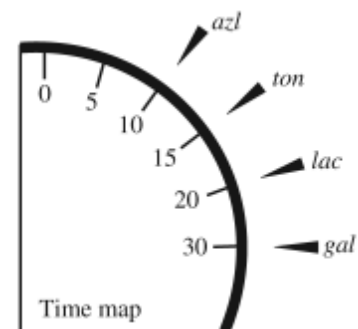


Fig. 7.8 A time map of the genes studied in the experiment depicted in Figure 7.7

Wollman and Jacob repeated the same type of experimentation with other Hfr strains, obtaining similar results with one important difference. Although genes were always transferred linearly with time, as in their original experiment, which genes entered first and which followed later seemed to vary from Hfr strain to Hfr strain [Figure 7.9(a)].

When they reexamined the rate of entry of genes, and thus the different genetic maps for each strain, a definite pattern emerged. The major difference between each strain was simply the point of the origin and the direction in which entry proceeded from that point [Figure 7.9(b)].

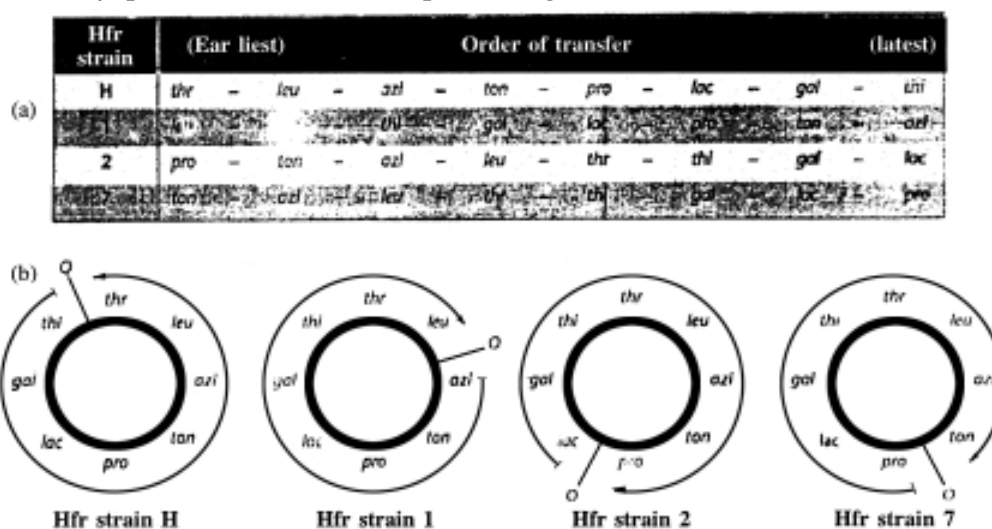


Fig. 7.9 (a) The order of gene transfer in four Hfr strains, suggesting that the *E. coli* chromosome is circular, (b) The point where transfer originates (*O*) is identified in each strain. Note that transfer can proceed in either direction, depending on the strain. The origin is determined by the point of integration into the chromosome of the F factor, and the direction of transfer is determined by the orientation of the F factor as it integrates

To explain these results, Wollman and Jacob postulated that the *E. coli* chromosome is circular. If the point of origin (*O*) varied from strain to strain, a different sequence of genes would be transferred in each case. But what determines *O*? They proposed that in various Hfr strains, the F factor integrates into the chromosome at different points and its position determines the *O* site. One such case of integration is shown in Figure 7.10 (Step 1). During conjugation between this Hfr and an F cell, the position of the F factor determines the initial point of transfer (Step 2 and 3). Those genes adjacent to *O* are transferred first. *The F factor becomes the last part to be transferred* (Step 4). Apparently, conjugation rarely, if ever, lasts long enough to allow the entire chromosome to pass across the conjugation tube (Step 5). This proposal explains why recipient cells, when mated with Hfr cells, remain F⁻.

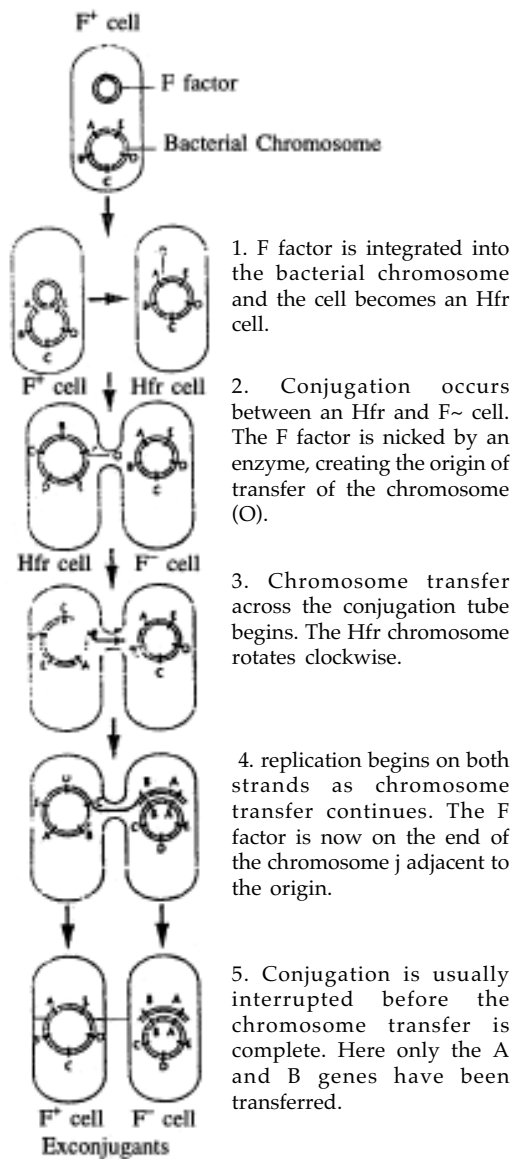


Fig. 7.10 Conversion of F⁺ to an Hfr state occurs by the integration of the F factor into the bacterial chromosome. The point of integration determines the origin (*O*) of transfer, during conjugation, the F factor, now integrated into the host chromosome, is nicked by an enzyme, initiating transfer of the chromosome at that point. Conjugation is usually interrupted prior to complete transfer. Above, only the A and B genes are transferred to the F⁻ cell, which may recombine with the host chromosome

Figure 7.10 also depicts the way in which the two strands making up a DNA molecule unwind during transfer, allowing for the entry of one of the strands of DNA into the recipient (Step 3). Following replication, the entering DNA now has the potential to recombine with its homologous region of the host chromosome. The DNA strand that remains in the donor also undergoes replication.

The use of the interrupted mating technique with different Hfr strains has provided the basis for mapping the entire *E. coli* chromosome. Mapped in time units, strain K12 (or *E. coli* K12) is 100 minutes long. Over 900 genes have now been placed on the map. In most instances, only a single copy of each gene exists.

7.3.6 Recombination in F⁺ × F⁻ matings : A review

The above model has helped geneticists to better understand how genetic recombination occurs during the F⁺ × F⁻ matings. Recall that recombination occurs much less frequently in them than in Hfr × F⁻ matings, and that random gene transfer is involved. The current belief is that when F⁺ and F⁻ cells are mixed, conjugation occurs readily and that each F⁻ cell involved in conjugation with an F⁺ cell receives a copy of the F factor, *but that no genetic recombination occurs*. However, at an extremely low frequency in a population of F⁺ cells, the F factor integrates spontaneously from the cytoplasm to a random point in the bacterial chromosome, converting the F⁺

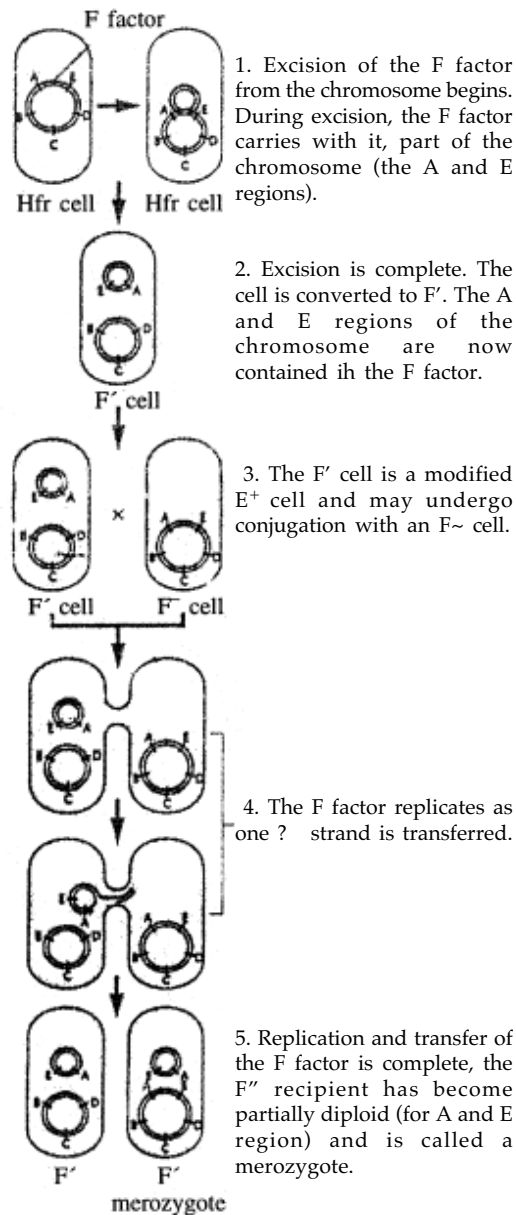


Fig. 7.11 Conversion of an Hfr bacterium to F' and is subsequent mating with an F⁻ cell. The conversion occurs when the F factor loses its integrated status. During excision from the chromosome, it carries with it one or more chromosomal genes (A and E) following conjugation with an F⁻ cell, the recipient cell becomes partially diploid and is called a merozygote. It also behaves as an F⁻ donor ones

cell to the state as we saw in Figure 7.10. Therefore, in F⁺ x F⁻ crosses, the extremely low frequency of genetic recombination (10⁻⁷) is attributed to the rare, newly formed Hfr cells, which then undergo conjugation with F⁻ cells. Because the point of integration of the F factor is random, a nonspecific gene transfer ensues, leading to the low-frequency, random genetic recombination observed in the F⁺ x F⁻ experiment. Unless the recipient cell simultaneously or subsequently undergoes conjugation with a separate F⁺ cell, it will remain F⁻. Most often, the recombinants become F⁺.

7.3.7 The F' State and merozygotes

In 1959, during experiments with Hfr strains of *E. coli*. Edward Adelberg discovered that the F⁺ factor could lose its integrated status, causing the cell to revert to the F⁺ state (Figure 7.11 Step 1). When this occurs, the F factor frequently carries several adjacent bacterial genes along with it (Step 2). Adelberg labeled this condition F' to distinguish it from F⁺ and Hfr. F⁺ like Hfr, is thus another special case of F⁺. This conversion is described as one from Hfr to F⁺.

The presence of bacterial genes within a cytoplasmic F factor creates an interesting situation. An F' bacterium behaves like an F⁺ cell, initiating conjugation with F⁻ cells (Figure 7.11-Step 3). When this occurs, the F factor, containing chromosomal genes, is transferred to the F⁺ cell (Step 4). As a result, whatever chromosomal genes are part of the F factor are now present in

duplicate in the recipient cell (Step 5), because the recipient still has a complete chromosome. This creates a partially diploid cell called a merozygote. Pure cultures of F' merozygotes can be established. They have been extremely useful in the study of bacterial genetics, particularly in genetic regulation.

7.3.8 Bacterial recombination is dependent on rec proteins

Once researchers established that a unidirectional transfer of DNA occurs between bacteria, they were interested in determining how the actual recombination event occurs in the recipient cell. Just how does the donor DNA replace the comparable region in the recipient chromosome? As with many systems, the biochemical mechanism by which recombination occurs was deciphered through genetic studies. Major insights were gained as a result of the isolation of a group of mutations representing genes called *rec*.

The first relevant observation in this case involved a series of mutant genes labeled *recA*, *recB*, *recC*, and *recD*. The first mutant gene, *recA*, was found to diminish genetic recombination in bacteria 1000-fold, nearly eliminating it altogether. The other *rec* mutations reduced recombination by about 100 times. Clearly, the normal wild-type products of these genes play some essential role in the process of recombination.

By looking for a functional gene product present in normal cells but missing in mutant cells, researchers subsequently isolated several gene products and showed that they played a role in genetic recombination. The first is called the RecA protein.* The second is a more complex protein called the RecBCD product, an enzyme consisting of polypeptide subunits encoded by three other *rec* genes. The roles of these proteins have now been elucidated *in vitro*. As a result of this genetic research, our knowledge of the process of recombination has been extended considerably. These discoveries underscore the value of isolating mutations, establishing their phenotypes, and determining the biological role of the normal, wildtype gene as a result of subsequent investigation.

7.3.9 F factors are plasmids

In the preceding sections we have examined the extra-chromosomal heredity unit called the F factor. When it exists autonomously in the bacterial cytoplasm, the F factor is composed of a double-stranded closed circle of DNA [Figure 7.12(a)]. These characteristics place the F factor in the more general category of genetic structures called plasmids. These structures contain one or more genes—often, quite a few. Their replication depends on the same enzymes that replicate the chromosome of the host cell, and they are distributed to daughter cells along with the host chromosome during cell division.

Plasmids are generally classified according to the genetic information specified by their DNA. The F factor confers fertility and contains genes essential

*Note that the names of bacterial genes begin with lowercase letters and are italicized. The names of the corresponding gene products (proteins) begin with an uppercase letter and are not italicized. For example, the *m-4* gene encodes the RecA protein.

for sex pilus formation, upon which genetic recombination depends. Other examples of plasmids include the R and the Col plasmids.

Most R plasmids consist of two components : the RTF (resistance transfer factor) and one or those r-de-terminants [Figure 7.12(b)]. The RTF encodes genetic information essential to transfer of the plasmid between bacteria, and the r-determinants are genes conferring resistance to antibiotics.

The **Col plasmid**, ColEI, derived from *E. coli*, is clearly distinct from R plasmids. It encodes one or more proteins that are highly toxic to bacterial strains that do not harbor the same plasmid. These proteins, called **colicins**, may kill neighboring bacteria. Bacteria that carry the plasmid are said to be colicinogenic. Present in 10 to 20 copies per cell, the plasmid also contains a gene encoding an immunity protein that protects the host cell from the toxin. Unlike an R plasmid, the Col plasmid is not usually transmissible to other cells.

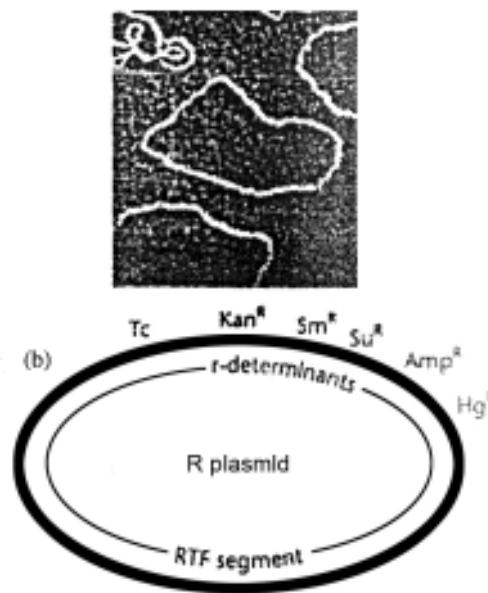


Fig. 7.12 (a) Electron micrograph of a plasmid isolated from *E. coli*; (b) diagrammatic representation of an R plasmid containing resistance transfer factors (RTFs) and multiple r-determinance (Tc, tetracycline; Kan, kanamycin; Sm, streptomycin; Su, sulfonamide; Amp, ampicillian; and Hg, mercury)

7.4 Transformation is another process leading to genetic recombination in bacteria

Transformation is another process that provides a mechanism for the recombination of genetic information in some bacteria. In transformation, small pieces of extracellular DNA are taken up by a living bacterium, ultimately leading to a stable genetic change in the recipient cell.

7.4.1 The Transformation Process

Transformation (Figure 7.13) consists of numerous steps that can be divided into two main categories: (1) entry of DNA into a recipient cell, and (2) recombination of the donor DNA with its homologous region in the recipient chromosome. In a population of cells, only those in a particular physiological state, referred to as competence, take up DNA. Entry is thought to occur at a limited number of receptor sites on the surface of the bacterial cell. Passage

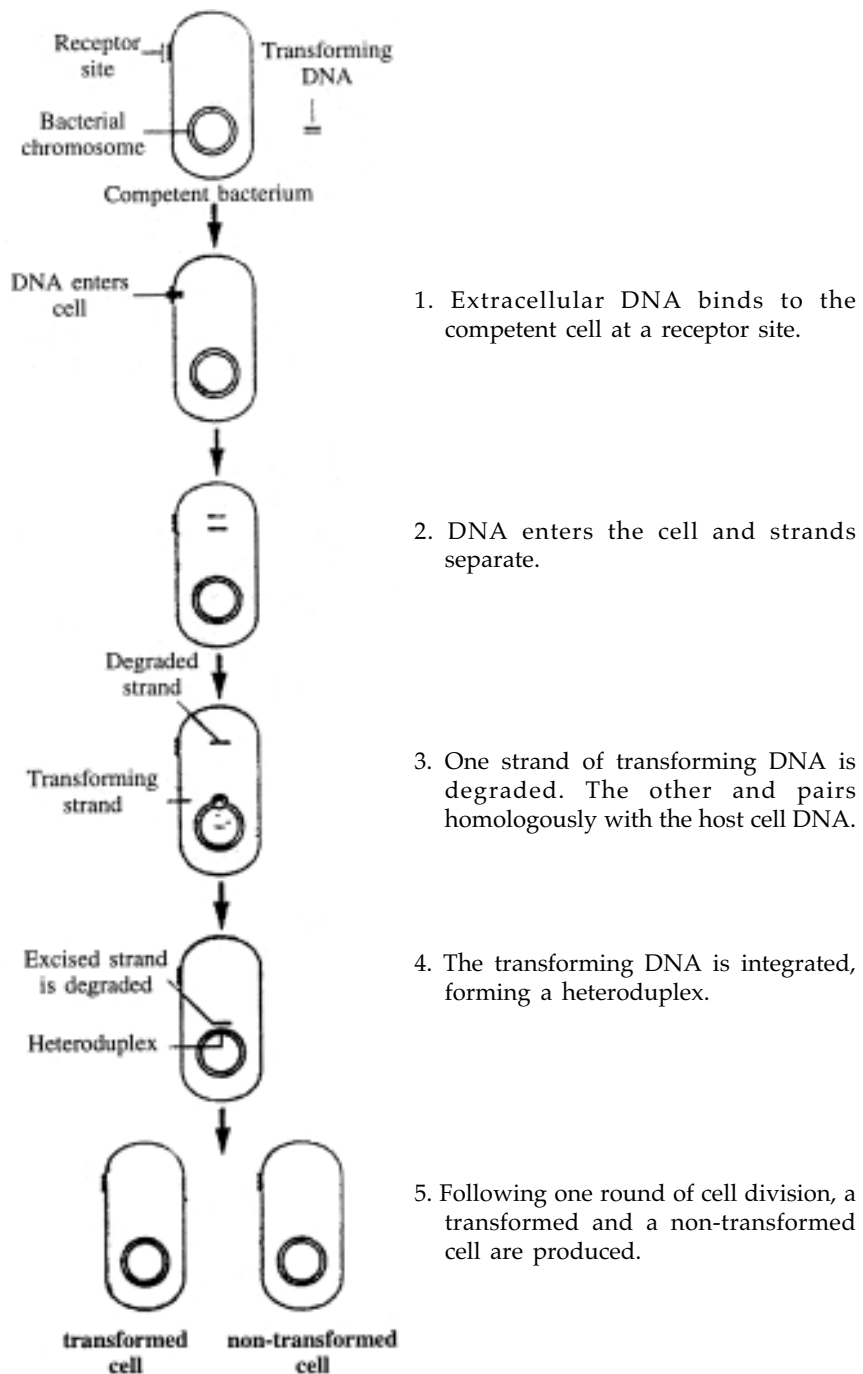


Fig. 7.13 Proposed steps leading to transformation of a bacterial cell by exogenous DNA. Only one of the two strands of the entering DNA is involved in the transformation event, which is completed following cell division

across the cell wall and membrane is an active process requiring energy and specific transport molecules. This concept is supported by the fact that substances that inhibit energy production or protein synthesis in the recipient cell also inhibit the transformation process.

During the process of entry, one of the two strands of the invading DNA molecule is digested by nucleases, leaving only a single strand to participate in transformation (Step 2 and 3). The surviving DNA strand aligns with its complementary region of the bacterial chromosome. In a process involving several enzymes, this segment of DNA replaces its counterpart in the chromosome, which is excised and degraded (Step 4).

For recombination to be detected, the transforming DNA must be derived from a different strain of bacteria, bearing some genetic variation. Once integrated into the chromosome, the recombinant region contains one DNA strand from the bacterial chromosome and one from the transforming DNA. Because these strands are not genetically identical, this helical region is referred to as a **heteroduplex**. Following one round of replication, one chromosome is restored to its original configuration, identical to that of the recipient cell, and the other contains the transformed gene. Cell division produces one host cell and one transformed cell (Step 5).

7.4.2 Lysogeny

The relationship between virus and bacterium does not always result in viral reproduction and lysis. As early as the 1920s, it was known that some bacteriophages could enter a bacterial cell and establish a symbiotic relationship with it. The precise molecular basis of this symbiosis is now well understood.

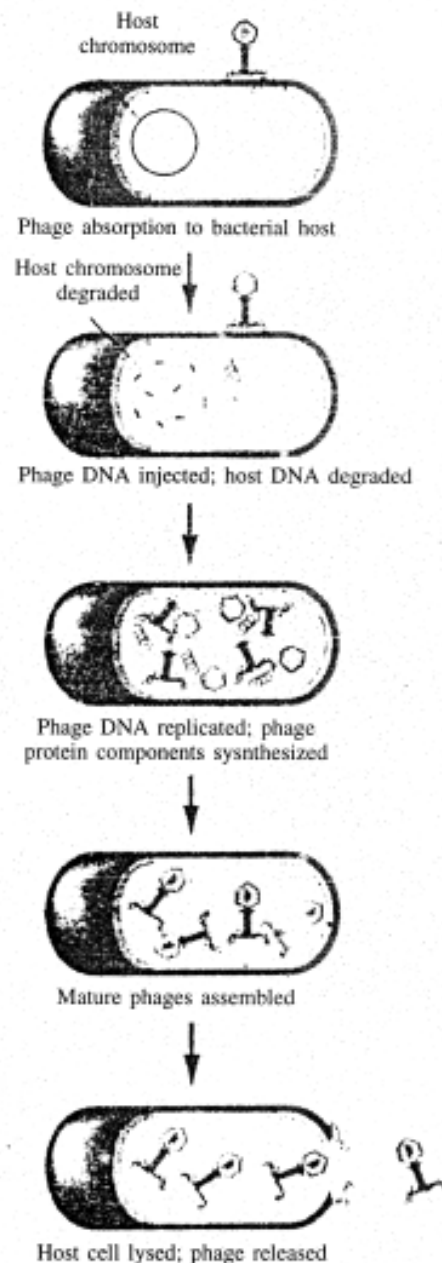


Fig. 7.15 Life cycle of bacteriophage T4

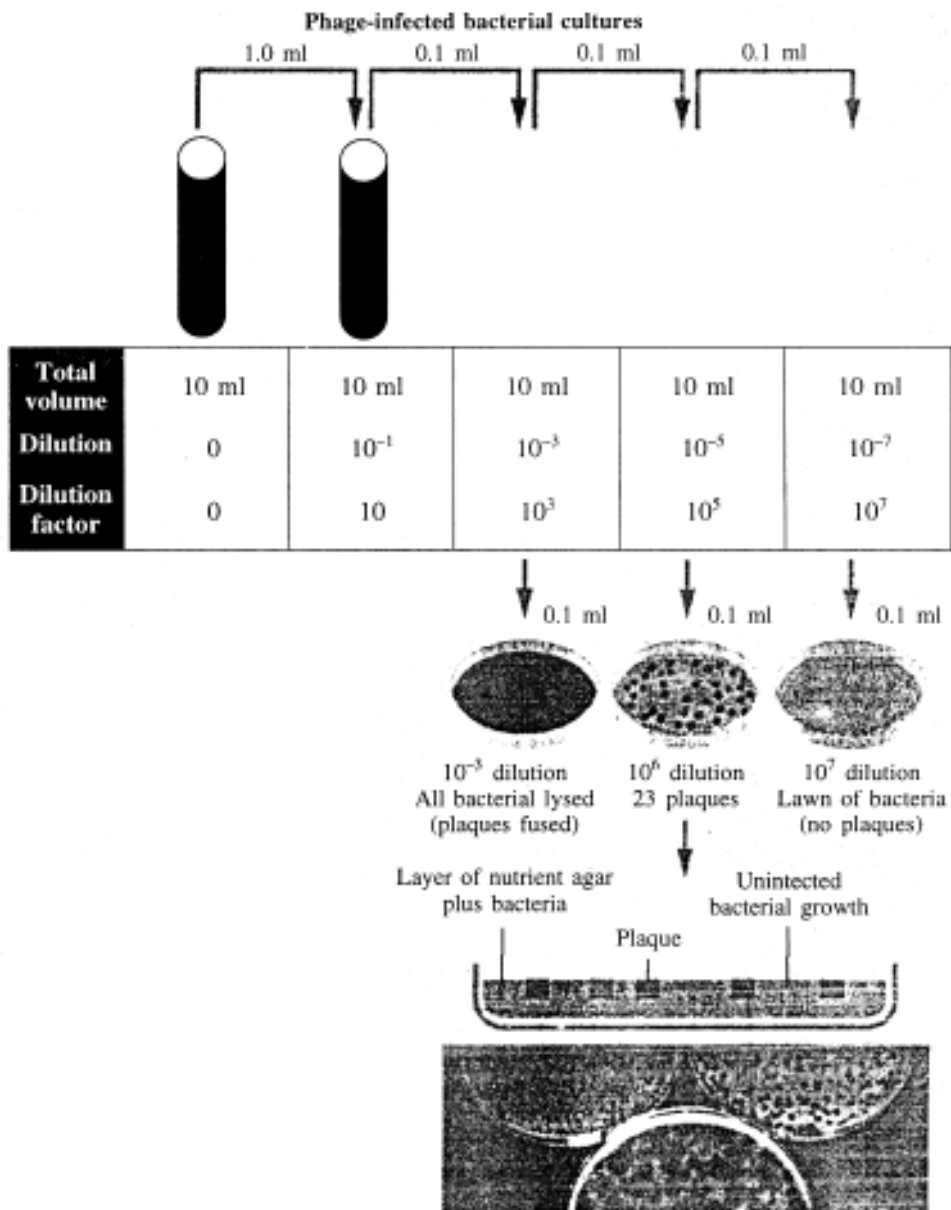


Fig. 7.16 Diagrammatic illustration of the plaque assay for bacteriophage analysis. Serial dilutions of a bacterial culture infected with bacteriophages are first made. Then three of the dilutions (10^3 , 10^6 and 10^7) are analyzed using the plaque assay technique in each case 0.1 ml of the diluted culture is used. Each plaque represents the initial infection of one bacterial cell by one bacteriophage. In the 10^3 dilution, so many phages are present that all bacterial are lysed. In the 10^{-5} dilution 23 plaques are produced. In the 10^{-7} dilution, the dilution factor is so great that no phages are present in the 0.1 ml sample, and thus no plaques form. From the 0.1 ml sample of the 10^{-3} dilution, the original bacteriophage density can be calculated as 23×10^4 or 23×10^5 phages/ml. The photograph illustrates phage T2 plaques on lawns of *E. coli*

Upon entry, the viral DNA, instead of replicating in the bacterial cytoplasm, is integrated into the bacterial chromosome, step that characterizes the developmental stage referred to as lysogeny. Subsequently, each time the bacterial chromosome is replicated, the viral DNA is also replicated and passed to daughter bacterial cells following division. No new viruses are produced and no lysis of the bacterial cell occurs. However, in response to certain stimuli, such as chemical or ultraviolet-light treatment, the viral DNA may lose its integrated status and initiate replication, phage reproduction, and lysis of the bacterium. (Fig. 7.15)

Several terms are used to describe this relationship. The viral DNA integrated into the bacterial chromosome is called a **prophage**. Viruses that can either lyse the cell or behave as a prophage are called **temperate**. Those that can only lyse the cell are referred to as virulent. A bacterium harboring a prophage has been lysogenized and said to be lysogenic; that is, it is capable of being lysed as a result of induced viral reproduction. The viral DNA, which can replicate either in the bacterial cytoplasm or as part of the bacterial chromosome, is sometimes classified as an **episome**.

7.5 Transduction is virus-mediated bacterial DNA transfer

In 1952, Norton Zinder and Joshua Lederberg were investigating possible recombination in the bacterium *Salmonella typhimurium*. Although they recovered prototrophs from mixed cultures of two different auxotrophic strains, subsequent investigations revealed that recombination was occurring in a manner different from that attributable to the presence of an F factor, as in *E. coli*. What they discovered was a process of bacterial recombination mediated by bacteriophages and now called transduction.

7.5.1 The Lederberg-Zinder experiment

Lederberg and Zinder mixed the *Salmonella* auxotrophic strains LA-22 and LA-2 together and, when the mixture was plated on minimal medium, they recovered prototroph cells. LA-22 was unable to synthesize the amino acids phenylalanine and tryptophan (*phe⁻ trp⁻*), and LA-2 could not synthesize the amino acids methionine and histidine (*met⁻ his⁻*). Prototrophs (*phe⁺ trp⁺ met⁺ his⁺*) were recovered at a rate of about $1/10^5$ (10^{-5}) cells.

Although these observations at first suggested that the recombination involved was the type observed earlier in conjugative strains of *E. coli*, experiments using the Davis U-tube soon showed otherwise (Figure 7.17). The two auxotrophic strains were separated by a glass-sintered filter, thus preventing cell contact but allowing growth to occur in a common medium. Surprisingly, when samples were removed-foam both sides of the filter and plated independently on minimal medium, prototrophs were recovered only from the side of the tube containing LA-22 bacteria.

Since LA-2 cells appeared to be the source of the new genetic information (phe^+ and trp^+), how that information crossed the filter from the LA-2 cells to the LA-22 cells allowing recombination to occur, was a mystery. The unknown source was designated simply as a **filterable agent (FA)**.

Three subsequent observations were useful in identifying the FA :

1. The FA was produced by the LA-2 cells only when they were grown in association with LA-22 cells. If LA-2 cells were grown independently and that culture medium was then added to LA-22 cells, recombination did not occur. Therefore, LA-22 cells play some role in the production of FA by LA-2 cells and do so only when the two share common growth medium.
2. The presence of DNase, which enzymatically digests DNA, did not render the FA ineffective. Therefore, the FA is not naked DNA, ruling out transformation.
3. The FA could not pass across the filter of the Davis U-tube when the pore size was reduced below the size of bacteriophages.

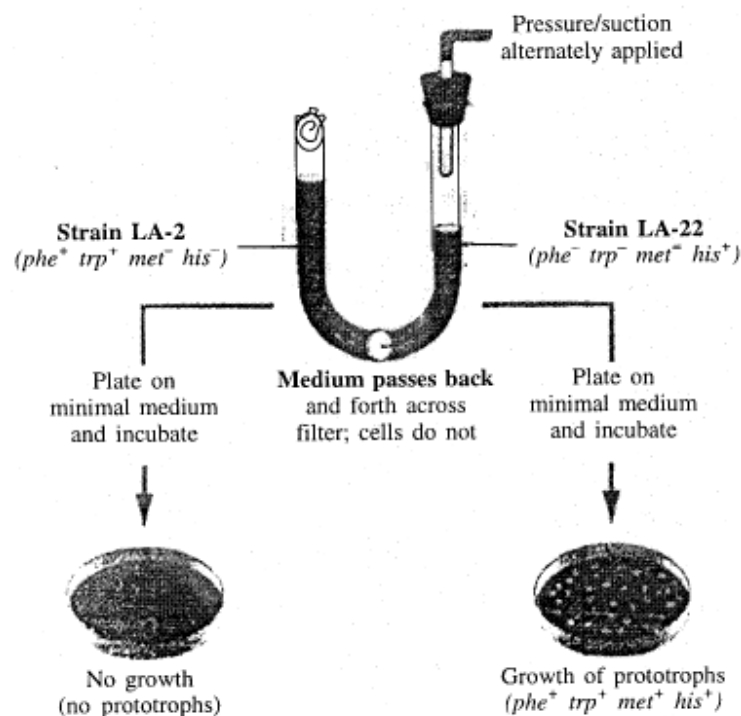


Fig. 7.17 The Lederberg-Zinder experiment using *Solmonelid*. After placing two auxotrophic strains on opposite sides of a Davis U-tube. Lederberg and Zinder recovered prototrophs from the side containing the LA-22 strain but not from the side containing the LA-2 strain. These initial observations led to the discovery of the phenomenon called transduction

Added by these observations and aware of temperate phages that could lysogenize *Salmonella*, researchers proposed that the genetic recombination event was mediated by bacteriophage P22, present initially as a prophage in the chromosome of the LA-22 *Salmonella* cells. It was hypothesized that rarely P22 prophages might enter the vegetative or lytic phase, reproduce, and be released by the LA-22 cells. Such phages, being much smaller than a bacterium, were then able to cross the filter of the U-tube and subsequently infect and lyse some of the LA-2 cells. In the process of lysis of LA-2, the P22 phages occasionally packaged in their heads a region of the LA-2 chromosome. If this region contained the *phe*⁺ and *trp*⁺ genes, and if the phages subsequently passed back across the filter and infected LA-22 cells, these newly lysogenized cells would behave as prototrophs. This process of transduction, whereby bacterial recombination is mediated by bacteriophage P22, is diagrammed in Figure 7.18.

7.5.2 The nature of transduction

Further studies revealed the existence of transducing phages in other species of bacteria. For example, *E. coli* can be transduced by phages PI and X. *Bacillus subtilis* and *Pseudomonas aeruginosa* can be transduced by the phages SPO1 and F1 16, respectively. The details of several different modes of transduction have also been established. Even though the initial discovery of transduction involved a temperate phage and a lysogenized bacterium, the same process can occur during the normal lytic cycle. Sometimes a small piece of bacterial DNA is packaged

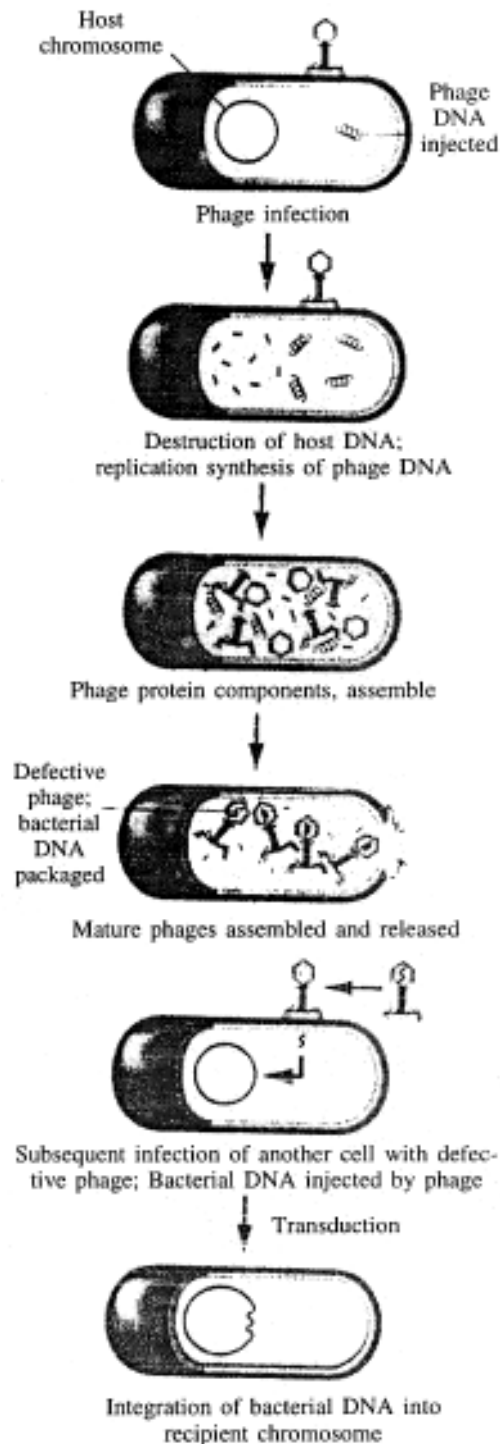


Fig. 7.18 Generalized transduction

along with the viral chromosome so that the transducing phage contains both viral and bacterial DNA. In such cases, only a few bacterial genes are present in the transducing phage. However, when *only* bacterial DNA is packaged, regions as large as 1 percent of the bacterial chromosome may become enclosed in the viral head. In either case, the ability to infect is unrelated to the type of DNA in the phage head, making transduction possible.

When bacterial rather than viral DNA is injected into the bacterium, it can either remain in the cytoplasm or recombine with the homologous region of the bacterial chromosome. If the bacterial DNA remains in the cytoplasm, it does not replicate but is transmitted to one of the progeny cells following each division. When this happens, only a single cell, partially diploid for the transduced genes, is produced—a phenomenon called **abortive transduction**. If the bacterial DNA recombines with its homologous region of

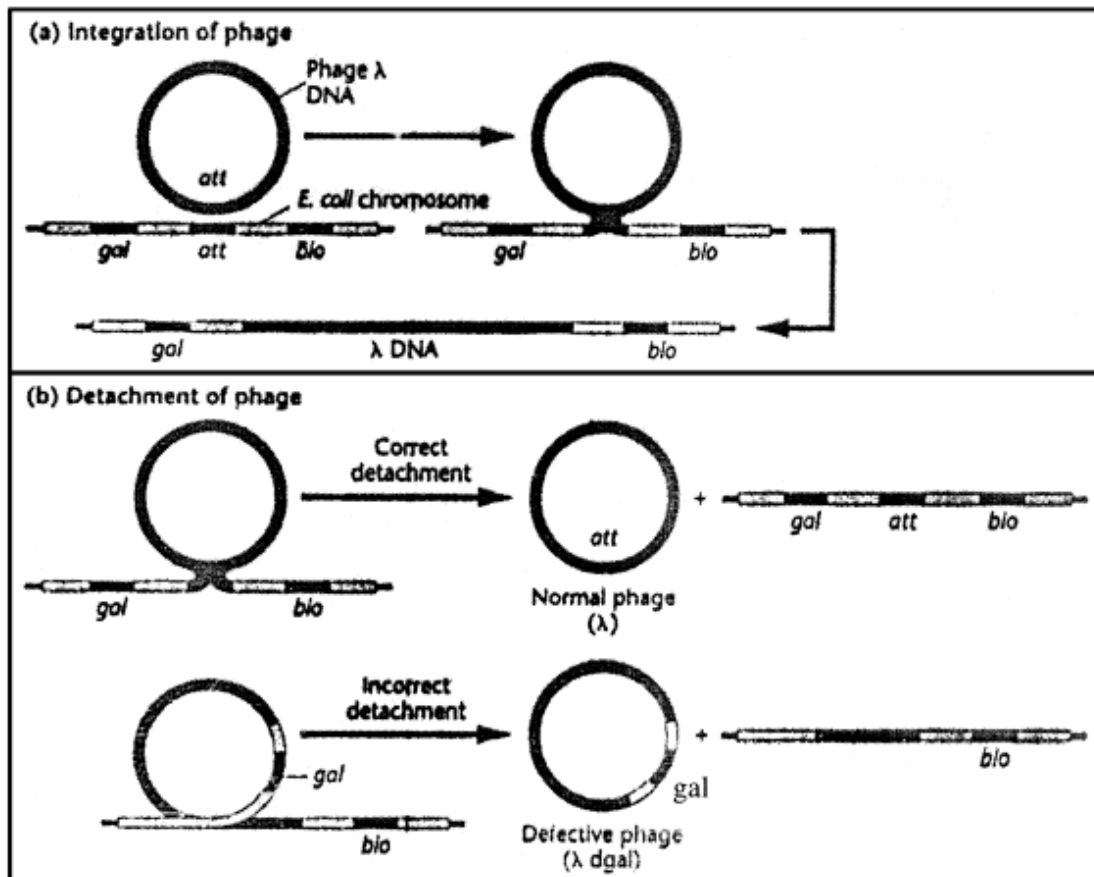


Fig. 7.19 The production of defective phage λ_{dgal} , which can result in specialized transduction following another round of infection of *E. coli*. If detachment occurs correctly, no transduction will result

the bacterial chromosome, the transduced genes are replicated as part of the chromosome and passed to all daughter cells. This process is called **complete transduction**.

7.5.3 Transduction and mapping

Like transformation, generalized transduction has been used in linkage and mapping studies of the bacterial chromosome. The fragment of bacterial DNA involved in a transduction event is large enough to include numerous genes. As a result, two genes that are closely aligned (linked) on the bacterial chromosome may be simultaneously transduced, a process called cotransduction. Two genes that are not close enough to one another along the chromosome to be included on a single DNA fragment require two independent events in order to be transduced into a single cell. Since this occurs with a much lower probability than cotransduction, linkage can be determined.

By concentrating on two or three linked genes, transduction studies can also determine the precise order of these genes. The closer to each other linked genes are, the greater the frequency of cotransduction. Mapping studies involving three closely aligned genes can be executed. The analysis of such an experiment is predicated on the same rationale underlying other mapping techniques.

7.5.9 Specialized transduction

In some instances, only certain genes are recombined, a situation called **specialized transduction**. This is in contrast to generalized transduction described above, where all genes have an equal probability of being recombined. One of the best examples involves transduction of *E. coli* by the temperate phage λ .

In this case, transduction is restricted to the *gal* (galactose) or *bio* (**biotin**) genes. The reason why transduction involves only these genes became clear when it was learned that the λ DNA always integrates into the region of the *E. coli* chromosome between these two genes at a site called *att* [Figure 7.19(a)]. Phage λ DNA that is integrated in the bacterial chromosome can subsequently detach from it, reproduce, and lyse the host cell [Figure 7.19(b)]. Sometimes the excision process occurs incorrectly and carries either the *gal* or *bio* *E. coli* genes in place of part of the viral DNA [Figure 7.19(b)]. The resulting phage chromosome is defective because it has lost some of its own genetic information, but it is nevertheless replicated and packaged during the formation of mature phage particles. The virus can subsequently inject the defective chromosome into another bacterial cell.

In this case of specialized transduction, the defective phage chromosome

is integrated into the bacterial, chromosome and is replicated along with it. Such bacterial cells contain the defective phage DNA, making them diploid for the *gal* or *bio* genes. The presence of this transducing *gal'* or *bio'* DNA causes these auxotrophs to revert to a *gal'* or *bio'* phenotype.

7.5.5 Bacteriophages undergo intergenic recombination

Around 1947, several research teams demonstrated that genetic recombination also occurs in bacteriophages. These studies relied on the discovery of numerous phage mutations that could be visualized or assayed. Before considering recombination in these bacterial viruses, we will briefly introduce several of the mutations that were studied.

Phage mutations often affect the morphology of the plaques formed following lysis of bacterial cells. For example, in 1946 Alfred Hershey observed unusual T2 plaques on plates of *E. coli* strain B. Where the normal T2 plaques are small and have a clear center surrounded by a diffuse (nearly invisible) halo, the unusual plaques were larger and possessed a more distinctive outer perimeter (Figure 7.20). When the viruses were isolated from these plaques and replated on *E. coli* B cells, the resulting plaque appearance was identical. Thus, the plaque phenotype was an inherited trait resulting from the reproduction of mutant phages. Hershey named the mutant *rapid lysis* (*r*) because the plaques were larger, apparently resulting from a more rapid or more efficient life cycle of the phage. It is now known that in wild-type phages, reproduction is inhibited once a particular-sized plaque has been formed. The mutant T2 phages are able to overcome this inhibition, producing larger plaques.

Unit 8 □ Cytogenetic effects of Ionizing and Non-ionizing Radiations

Structure

- 8.1 Introduction
- 8.2 Radiation
- 8.3 Viruses
- 8.4 Chemical clastogens

8.1 Introduction

Various agents shown to cause chromosome breaks have been termed “clastogens” by Shaw (1970). These include physical agents (X-rays, ultraviolet light, cold shock, magnetic fields, and sound waves), biological agents (certain genes, viruses and protozoa), and a host of chemical agents. It should be emphasized that most of these clastogens produced these effects *in vitro* by the addition of the agent to cultured lymphocytes and/or fibroblasts for varying times and concentration. In but few cases is there evidence for *in vivo* chromosome breakage.

8.2 Radiation

Survivors of the atomic bomb blasts in Japan have developed leukemia in proportion to the amount of radiation received. Furthermore, increased numbers of chromosome breaks and rearrangements have been found in lymphocytes of nonleukemic survivors (Bloom *et al.*, 1967). Similar anomalies (translocations and inversions) have been demonstrated in lymphocytes of individuals who have received X-ray therapy to the spine or injections of thorotrast (Buckton *et al.*, 1982; Court Brown *et al.*, 1967).

Although ultrasound can effect chromosomal breaks *in vitro*, there is no evidence that it does *so in vivo* (Macintosh and Davey, 1972).

Fibroblasts cultured from skin in the path of X-radiation have manifested chromosome abnormalities (Engel *et al.*, 1964; Visfeldt, 1966). Marrow cells may exhibit abnormalities even after many years following primary exposure (Goh, 1971). Leukemia is also more likely to develop in individuals who have received chronic exposure to radiation (Lewis, 1970). Maternal irradiation before and during the reproductive period increases the incidence of chromosomally abnormal conceptuses. However, most are nonviable and lost early in pregnancy (Alberman *et al.*, 1972).

8.3 Viruses

There is insufficient evidence currently available to directly implicate viruses in effecting human chromosome abnormalities. However, a number of investigators have studied the effect of SV40 virus on cultured human fibroblasts and have observed altered growth patterns to the haphazard growth and alterations, within several months an emergence of a few stable heteroploid cells becomes evident (Moorhead, 1970).

Numerous other viruses (Rous sarcoma, vaccinia, rubella, herpes zoster, poliomyelitis, influenza, polyoma, etc.) have been shown to produce chromosome breaks in infected cells *in vitro*. Furthermore, several viruses can produce abnormalities in metaphase chromosome in circulating lymphocytes during natural human infections (measles, chicken pox, mumps, and hepatitis) (Moorhead, 1970). The effects have been of at least three types: single breaks, pulverization, and fusion and spindle abnormalities. The mechanism is unknown but may be related to addition of the virus like particles (Epstein-Barr virus) have been demonstrated, often exhibit a long submetacentric marker (Gripenberg *et al.*, 1969).

8.4 Chemical clastogens

Over 200 drugs or chemical shown to cause chromosome breaks *in vitro* (shaw, 1970) can be grouped into several categories with a few illustrations in each group: (a) nucleic acid related compounds (6-mercaptapurine and 5-fluorodeoxyuridine), (b) antibiotics (mitomycin C, streptomycin, actinomycin D and daunomycin), (c) central nervous system drugs (meprobamate, chlorpromazine, mescaline, lysergic acid diethylamide, and scopolamine), (d) food derivatives and additives (caffeine, cyclamate, theobromine, and theophylline), (e) air and water pollutants (chloramine T and ozone), (f) pesticides (captan and thioTEPA), (g) alkylating agents (nitrogen mustards and cytoxan), (h) mitotic poisons (colchicine), (i) photodynamic dyes (acridine orange and neutral red), (j) antifolate compounds (methotrexate and aminopterin), (k) organic solvents (benzene and mercaptoethanol), (l) inorganic substances (lead and arsenic), and (m) miscellaneous compounds (Imuran and piperazine).

However, it should be emphasized that few chemical clastogens have been implicated in chromosomal breakage *in vivo*. To cite but a few examples: Ambient exposure to benzene has been noted to be associated with both chromosome breakage and subsequent development of leukemia (Tough and Court Brown, 1965; Hartwich *et al.*, 1969). On the other hand, LSD, while producing chromosome breaks *in vitro*, has not been shown to be effective *in vivo* (Stenchever and Jarvis, 1970).

Unit 9 □ Molecular Cytogenetic Techniques

Structure

- 9.1 FISH, GISH
- 9.2 DNA finger printing
- 9.3 Flow cytometry
- 9.4 Chromosome painting

9.1 FISH, GISH

9.1.1 *In situ* hybridization technique

Under normal temperature and ionic conditions DNA remains in a duplex state by the base pairing through the hydrogen bonds. By heating in buffer solution or by increasing pH the two strands can be separated. But if again the temperature is lowered or pH is reduced then the separated strands will join again and reassociate.

This fact was shown by Julius Marmur and Paul Doty in 1960. This type of reassociation of DNA strands is called molecular hybridization or nuclear hybridization. It may also take place between the complementary strands of DNA or RNA or between DNA and RNA.

Hybridization technique which can be used to localize specific nucleic acid fragments that reside in their original site (*in situ*) within cells, is known as *in situ* hybridization.

In situ hybridization is a version of hybridization analysis, in which an intact chromosome is examined by probing it with a labeled DNA molecule.

For this technique to work, DNA in the chromosome must be made single stranded and denatured by breaking the hydrogen bonds between the base pairs. Only then the DNA or RNA probe can be hybridized with the chromosomal DNA. For this purpose, without destroying the chromosome morphology a dry preparation is made into a glass microscope-slide and then treated with formaldehyde.

The position where the hybridization occurs provides the information about the map location of this gene, thus physical mapping of chromosome could be done by this method.

To locate the region of hybridization two types of markers viz., radioactive marker and fluorescent marker are used. Besides these other types of markers are also used for this purpose.

To prepare a DNA profile, the nucleotides are synthesized in which one of the phosphorus atoms replaced by ^{32}P or ^{33}P . One of the oxygen atom in the phosphate group is replaced with ^{35}S or one or more of the hydrogen atom is replaced with ^3H . Radioactive nucleotides act as substrates for DNA polymerases and so are incorporated into a DNA molecule by any strand synthesis reaction catalyzed by a DNA polymerase.

The labeled nucleotides or individual phosphate groups can also be attached to one or both ends of a DNA molecule by the reaction catalyzed by T4 polynucleotide kinase or terminal deoxy nucleotidyl transferase.

The radioactive signal can be detected by scintillation counting, but for most molecular biology, the position of the hybridization is detected by exposure of an X-ray sensitive film (autoradiography) or radiation sensitive phosphorescent screen (phosphorus imaging). ^{32}P has a high emission energy and the resolution is lower. But low emission isotopes such as ^{35}S or ^3H give less sensitivity but greater resolution.

9.1.2 Fluorescent *in situ* hybridization (FISH)

Concept

To solve these problems in 1980's non radioactive fluorescent DNA labels were developed. These labels combine high sensitivity with high resolution and are ideal for *in situ* hybridization. The different fluorolabels of different colours have been designed for the probes and it is possible to hybridize the chromosome and the individual hybridization signals enable the location of relative position of the probe sequence to be mapped.

Fluorescent dyes (Fluorochromes) like quinacrine (Q) and quinacrine mustard (QM) are used to obtain specific patterns of cross striations or bands appear with alternate fluorescent and non fluorescent bands. The bands obtained with quinacrine are called Q bands while those obtained with quinacrine mustard are called Q M-bands. Bands are obtained when fluorescent dyes attach to the specific regions of the chromosomes (See Fig. 9.2).

Technique

During the *in situ* hybridization a sample of dividing cells is dried on a microscope slide and treated with formaldehyde so that the chromosomes become denatured but do not lose their characteristic metaphase morphologies. The probe is added to the denatured chromosomes, which will be hybridized with the complementary DNA region of the chromosome. The position at which the probe hybridizes to the chromosomal DNA is visualized by detecting the fluorescent signal emitted by the labeled DNA.

Problem

If a probe be a long fragment of DNA, then a potential problem is that it is likely to contain repetitive DNA sequences and so may hybridize many portions of the chromosome. If these sequences are not blocked then the probe will hybridize non specifically to any copy of these genome and will repeat in the target DNA. To block the repeat sequences, the probe is pro-hybridized with a DNA fraction enriched for repetitive DNA.

Advantage of FISH over other methods

Health and environmental issues have meant that radioactive markers have become less popular in recent years and they are now largely superseded by non-radioactive alternatives.

Other drawbacks of radioactive markers in *in situ* hybridization and that the radioactive label has high emission energy (eg., P^{32}) and then it scatters its signal and so gives poor resolution. On the other hand , if H^3 is used, the emission energy is less but its sensitively is so low that lengthy exposures are needed.

Disadvantage

The metaphase chromosomes are highly condensed and a fluorescent signal obtained by FISH is marked by measuring its position relative to the end of the short arm of the chromosome (the fletcher value). The two markers having at least 1Mb apart to be resolved as separate hybridization signal (Trask *et al*, 1991). Therefore, using FISH, the highly condensed nature of metaphase chromosomes means that only low resolution mapping is possible. Therefore, FISH provides a fough idea of its map position.

More advanced FISH

In 1996 Heiskanen *et al* solved the problem of FISH and a range of higher resolution FISH technique has been developed. Higher resolution is achieved by changing the condensing pattern of the metaphase chromosome. There are two ways of doing this.

(1) Mechanically stretching of metaphase chromosomes : Centrifugation generates shear forces, which can result in the chromosomes becoming stretched upto 20 times of the normal length. Thus the resolution is significantly improved and markers that are only 200-300 Kb apart can be distinguished.

(2) Taking non-metaphase chromosomes : Attempts have been made to use prophase nuclei because in this stage the chromosomes are not still sufficiently condensed for individual ones to be identified and so provides no advantage.

Interphase chromosomes become more useful because, then the

chromosomes again become less condensed. Using anaphase stage the resolution down to 25 kb is possible but their chromosome morphology is lost, so there are no external reference point against which the position of the probe to be mapped. This technique is used to construct map in small region of the chromosome after obtaining preliminary information.

Fibre-FISH

Interphase chromosomes are most unpacked of all cellular DNA. To improve the resolution of FISH better than 25 kb, it is therefore, necessary to abandon intact chromosomes. This approach is called fiber FISH. In this approach the DNA is prepared by gel stretching or molecular combing. This can distinguish markers that are less than 10 Kb.

To carry out gel stretching (Fig. 9.1a), molten agarose containing chromosomal DNA molecules is pipetted into a microscope slide, coated "with a restriction enzyme (Schwartz *et al*, 1993). As the gel solidifies, the DNA molecules become stretched. It is not understood why this happens but it is thought that fluid movement on the glass surface during gelation might be responsible. Once the gel is solidified it is washed with $MgCl_2$ solution, which activates the restriction enzyme. A fluorescent dye such as DAPI (4,6 diamine 2-phenylindole dihydrochloride) is added which stains the DNA so that the fibres can be seen when the slide is examined with a high power fluorescence microscope. The restriction enzyme cuts the DNA molecule. As the molecules gradually coil up, the gaps representing the cut sites become visible. The relative positions of the cuts are to be recorded.

In molecular combing (Michalet *et al*, 1997) (Fig. 9.1b), the DNA fibres are prepared by dipping a silicon-coated coverslip into a solution of DNA. It takes 5 minutes to attach DNA to the coverslip by their ends. After that coverslip is removed at a constant speed of 0.3 mm/sec. The force required to pull the DNA molecule through the meniscus causes to line up. Once in the air the surface of the coverslip dries, DNA molecules are arranged as a parallel fibre thus producing a comb of parallel molecules.

9.1.3 Biotin-labelling *in situ* hybridization

Recent advances in nucleic acid technology offer alternative to radioactive-labelling probes. One procedure that is becoming increasingly popular is biotin-labelling of nucleic acid. This is nontoxic, whose half life is longer and can be prepared in advance in bulk and stored at $-20^{\circ}C$ for (repeated) use.

Drosophila salivary gland chromosomes can be hybridized with a biotin labelled nucleic acid probe. After washing, detection can be done by adding a biotin-binding protein called ovidin which is covalently bound to alkaline

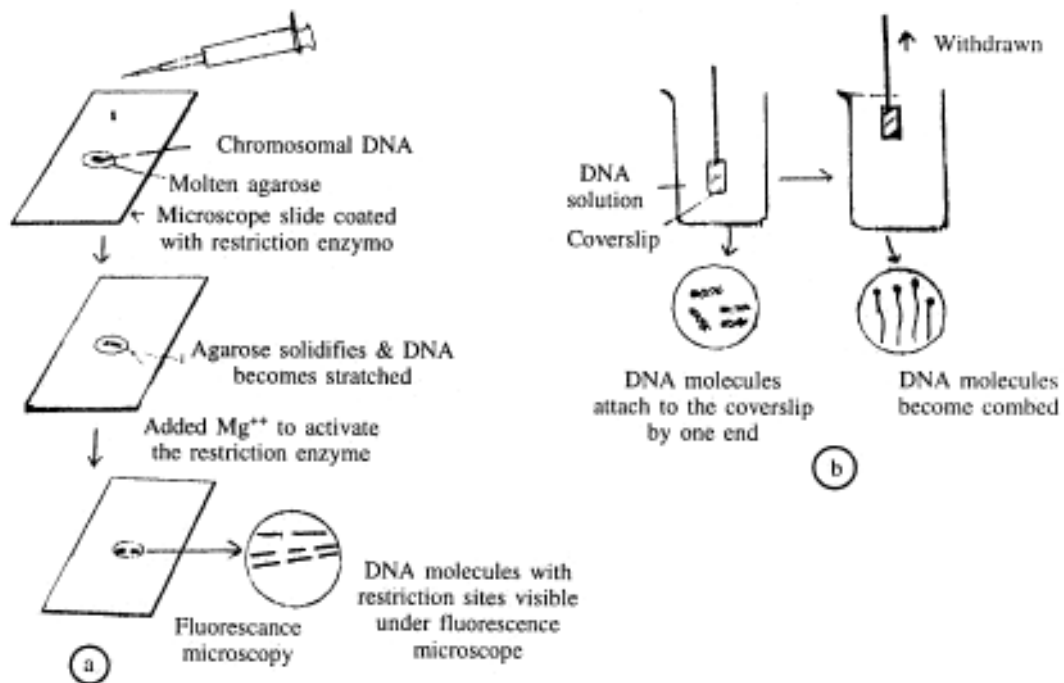


Fig. 9.1 (a) Gel stretching and (b) molecular combing in fibre FISH

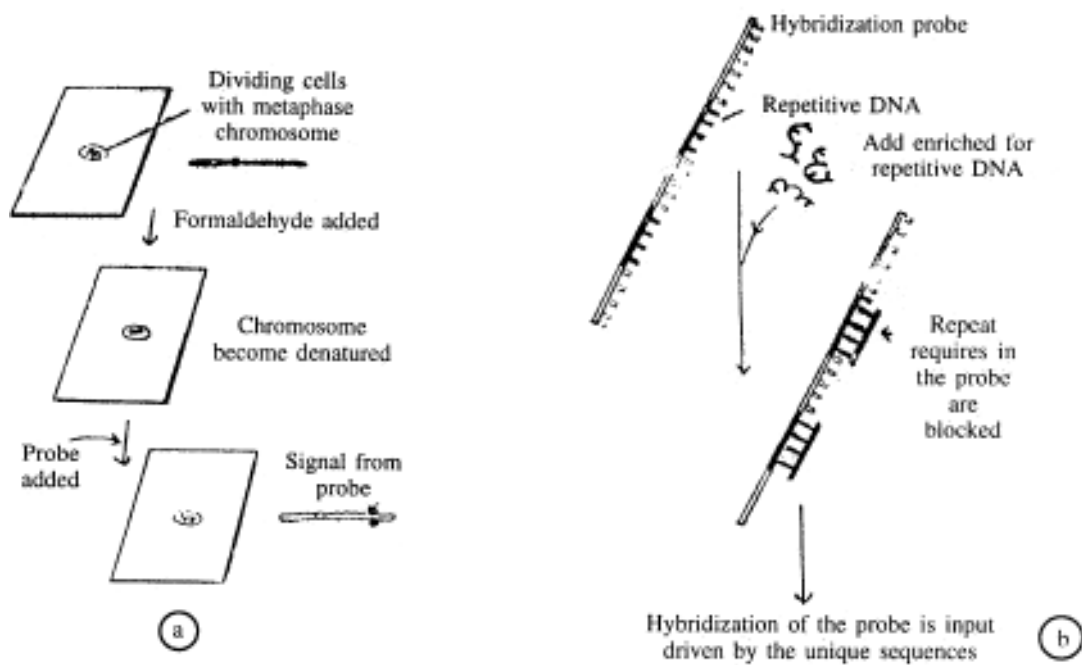


Fig. 9.2 (a) FISH ; (b) Blocking of repetitive DNA sequences in a hybridization probe

phosphatase. After addition of a soluble substrate, the enzyme catalyses a reaction that results in formation of an insoluble blue coloured precipitate at that site of hybridization. The intensity is proportional to the amount of biotin in the hybrid.

In situ hybridization with a biotin labelled probe has been particularly useful in chromosome mapping of DNA clones in *Drosophila* because the logical map of the poly-tene chromosomes of this organism is known at high resolution.

9.2 DNA finger printing

The techniques of DNA diagnosis have found application in a quite different area, the identification of medicine. This is important in areas as diverse as identifying cell cultures, determining family relationships in studies of animal behaviour, immigration problems to identify criminals or murderer, disputed parentage and in forensic medicine.

The most accurate method of identification technique based on recombinant technology is called DNA finger printing or DNA typing or DNA profiling.

Principle

DNA finger printing is based on **sequence polymorphism** that occurs in human genome and the genome of other organisms. The sequence polymorphisms are slight sequence differences, usually single base pair changes, that occur from individuals to individuals once in every few hundred base pairs on average. Each difference from the consensus human genome sequence is generally present in only a fraction of the human population but every individual has some of them. These polymorphic locus is called minisatellite or VNTR (variable numbers tandem repeat) thus forming a haplotype which shows mendelian inheritance among their offsprings. This locus is made up of a variable number of identical sequences joint together in tandem.

One family of minisatellites in the human genome share a common "Core" sequence, The core is G-C rich sequence of 10-15 bp showing on asymmetry of purine/ pyrimidine distribution on the two strands. These repeats are written as (C-A)_n (G-T)_n, occur in 100000 blocks in every genome and appear to be uniformly distributed throughout the genome (value of n varies from 1 to 40). The successful application of these (C-A)_n (G-T)_n repeats has led to the use of a variety of other di-, tri- **and** tetranucleotide sequences for mapping.

Technique

(1) DNA of the sample is first isolated whose DNA finger printing has to be made. Usually in forensic case the DNA is prepared from dried stains, sperms

in vaginal swabs that had been stored for as long as two years. Sufficient DNA can be isolated from freshly pulled hair roots, polymorphism in mitochondrial DNA and class II HLA gene DQ have been analysed from the shed hairs of several months old containing less than 1 ng of DNA.

(2) Sufficient quantities of intact DNA from forensic samples will always be a problem. PCR may have a great impact in this area. From a small quantities of DNA PCR can produce a large number DNA. These are used as probes. This probe is more redioactive.

(3) The DNA from the individual whose DNA is to be compared with the forensic sample is isolated and are cut into fragments by restriction enzymes.

(4) DNA fragments after digestion of DNA from the genome are first separated according to their size by agarose gel electrophoresis. These ds DNA are denatured by soaking the gel in alkali to make it ss DNA.

(5) The DNA fragments are transferred to nito-cellulose paper by the southern blot technique. The paper is then immersed in a solution containing a radioactively labeled DNA probe. Fragments to which the probe hybridizes are revealed by autoradiography.

Detection of forensic problem :

The power of DNA fingerprinting was demonstrated by Alec Jeffreys in 1985, when a man had been aceured of two rape murders committed three years apart and had made a confussion. Lastly the real murderer was caught and DNA finger printing confirmed the identification.

DNA from a semen sample obtained from a rape and murder victim was analysed along the DNA samples from the victim and two suspects.

Each of the DNA samples was cleaved into fragments and separated by gel electrophoresis.

Radioactive DNA probes were used to identify a small subset of these fragments that contained sequences complimentary to the probe. The sizes of the fragments identified varied from one individual to the next. The different patterns for the three individuals (victim and two suspects) tested. One rape suspects DNA exhibits a banding pattern identical to that of the semen sample taken from the victim.

More than one probe may be used to make a positive identification.

9.3 Separation of chromosomes by flow cytometry

The dividing cells with condensed chromosomes are carefully broken open so that a mixture of intact chromosomes is obtained. The chromosomes are then stained with fluorescent dye. The amount of dye that a chromosome binds

depends on its size. Thus larger chromosomes bind more dye and fluorescence more brightly than smaller ones. The mixture of chromosomes is diluted and passed through a fine aperture, producing a stream of droplets, each one containing a single chromosome. The droplets are passed through a detector that measures the amount of fluorescence and thus identifies which droplets contain the particular chromosome (being sought). An electric charge is applied to these droplets by a charger and then the droplets reach the electric plates, the charged ones are deflected into a separate beaker.

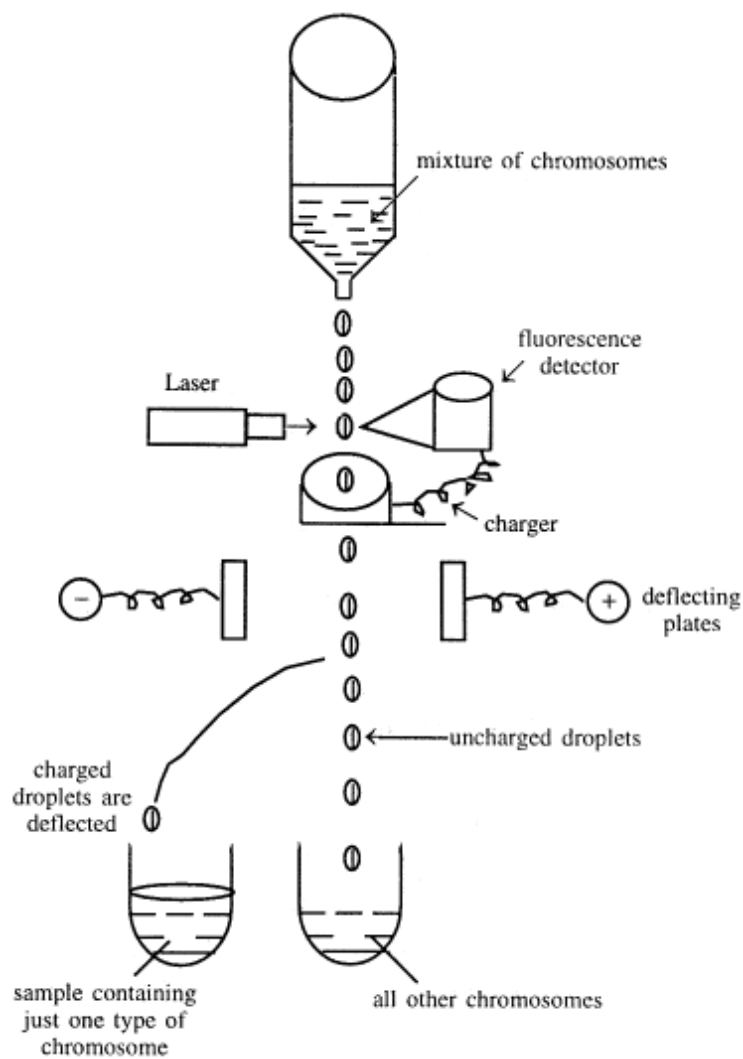


Fig. 9.3 Flow cytometry

If two chromosomes are of equal size as in human chromosome number 21 and 22, then the dye, thiochast 33258 and chlomomycin-A preferably bind to the A-T and G-C rich DNA respectively and by this differential staining activity, these chromosomes can be distinguished properly (Fig. 9.3).

9.4 Chromosome painting

This is a method for visualizing each of the chromosomes in distinct bright colours and thus it simplifies greatly the distinction between chromosomes of similar size and shape and the karyotyping of the chromosome. Usually such a painting is done at the stage in the cell cycle (mitosis) when chromosomes are specially compact and easy to visualize, usually at mitosis. But sometimes, selective chromosome painting may be done in the interphase stage to see its orientation in the nucleus.

Method

(1) The probes which are used for chromosome painting are specific for sites scattered along the length of each chromosome.

(2) The probes are labeled with one of two dyes that fluoresce at different wave lengths. For example, DNA molecules derived from chromosome-1 are labeled with one specific dye combination, chromosome-2 with another and so on.

(3) The labeled probe can hybridize only the chromosome from which it was derived, each chromosome is differently labeled. After the probes are hybridized to chromosomes the excess is removed, the sample is placed in a fluorescent microscope in which a detector determines the fraction of each dye present at each fluorescing position in the microscopic field. The information may be conveyed, to a computer and a special program assigns a false colour image to each type of chromosome.

Use

Chromosome painting is very useful in karyotyping the chromosomes. It can be done in interphase stage to locate the specific chromosomes and their arrangement in the nucleus. A combination of chromosome banding with FISH, called multicolor FISH can detect chromosomal translocations which are associated with certain genetic disorder and specific types of cancers. For example, in chronic myeloid leukemia (CML), the lymphocytes contain the Philadelphia chromosome (small), which is produced by the translocation of chromosome no-22 and chromosome no-9. The translocations can be detected by classical banding analysis technique.

Unit 10 □ Genome Analysis

Structure

10.1 C-value paradox

10.2 Satellite DNA

10.3 Complexity

10.1 C-value paradox

The haploid DNA content in an individual is described as its C-value.

The anomalies of the gene contents by two different methods one on the basis of knowledge about the rate of mutation per locus and other on the basis of general method used for DNA content, is called C-value paradox.

DNA content in eukaryotic cells are much higher than that in the prokaryotic cells and a wide range of variations are observed among different species even among same species. The content of DNA also depends on the number of chromosomes in the cells (i.e., ploidy of the chromosome). Example : See Table 10.1.

Table 10.1 DNA content of some organisms

Class of organisms	Species	Haploid DNA (Picogram)	Dalton	Base pairs
Phages	Ø × 174 T4	2.6×10^{-6}	1.7×10^6	5400
		20.7×10^{-5}	126×10^6	200000
Animal virus	Adenovirus	21.7×10^{-6}	13×10^6	21000
Prokaryotes	<i>E. coli</i>	4.4×10^{-3}	2.7×10^9	4.2×10^6 .
Unicellular-eukaryotes	<i>S. cerevisial</i>	14×10^{-3}	8.5×10^9	1.4×10^7
Multicellular enkaryotes	<i>D. melanogaster</i>	0.18	0.11×10^{12}	0.17×10^9
	<i>Homo sapiens</i>	2.8	18×10^{12}	2.8×10^9

C-value paradox takes its name from the inability to explain the content of a genome in terms of an anticipated function. There are two aspects of the

paradox. First there are huge variations in C-values between certain species whose apparent complexity does not vary correspondingly. There can be rather substantial variations even between certain closely related species.

The range of C-values is found in different evolutionary phyla. There is some increase in the minimum genome size that is found in each phylum as the complexity increases.

For example, in prokaryotes, the genome size are very small. In eukaryotes, a vast increase occurs in genome size. In yeast, *Saccharomyces cerevisiae* has a genome size of 2.3×10^7 bp, only 5 times greater than that of *E. coli*.

The modest increase in genome size just over two folds is adequate to support the slime mold, *D. discoideum*, able to live in either unicellular or multicellular modes.

Another increase in complexity is necessary to produce the first fully metazoan organisms. For example, *C. elegans* has a DNA content of 8×10^7 bp. Then any close relationship between complexity of the organisms and content of DNA is obscure, although it is necessary to have a genome of more than 10^8 bp, to make an insect of more than 4×10^8 bp, to assemble an echinoderm of more than 8×10^8 bp, to produce a bird or amphibians and more than 2×10^8 bp to develop into a mammals.

In some cases the spread of genome size is quite small. For example, birds, reptiles and mammals, all show a little variation within the phylum, with a range of genome size in each case about two fold. But in other cases, there is quite a wide range of values, often more than 10 folds. This reflects some surprising discrepancies between genome size and complexity of the organism.

An extraordinary C-value is found in amphibian where the smallest is below 10^9 bp while the largest are almost 10^{11} bp. It is hard to believe that this could reflect a 100 fold variation in number of genes in different amphibians.

There are some cases where rather closely related species show surprising variations in total genome size. For example two amphibian species may have 10 fold increase where morphologies are very similar. Yet if the gene number is roughly similar most of the DNA in the species with the larger genome cannot be concerned with coding for protein. So the question, what could be its function?

The second aspect of C-value paradox is the apparent absolute excess of DNA compared with the amount that could be expected to code for proteins.

Actually, eukaryotic DNA has an excess length and the excess is encountered because genes are much larger than the sequences needed to code for proteins. For example, human haploid cell has DNA amount equal to 1.8×10^{12} dalton = 87 cm of DNA which is equal to 2.8×10^9 base pairs, then this genome could contain approximately as many as 3×10^6 genes assuming 1000 bp per gene coding for nearly 300 amino acids.

However, the number of genes estimated in humans on the basis of the rate of mutation per locus, as estimated by Muller (1967), the frequency of deleterious mutations per locus in human is 10^{-5} to 10^{-6} in each generation. If the number of gene is 3×10^6 (as calculated in general method), then it will yield 30 deleterious mutations in each generations at the rate of 10^{-5} mutations per locus. This will be an unbearable genetic load. The actual frequency of deleterious mutations per generation per individual has been estimated to be 0.5 against an expected frequency of 30. This mutation frequency at the rate of 10^{-5} per locus will be an estimate of 5×10^4 genes.

Thus it is supposed that the actual number of genes in human should be 50000 and not 3 million as estimated from DNA content.

This anomalous situation has been described by some workers as C-value paradox.

10.2 Satellite DNA

Large proportions of DNA in eukaryotes has been shown to be present in the form of multiple copies of identical DNA sequences, thus is called repetitive DNA or Satellite DNA. The remaining DNA in the cell is found in the form of single copy of DNA sequences which is known as unique DNA.

When the denatured DNA (single stranded) is led to reassociate, then it was observed that from the heterogenous populations, the smaller molecular weight DNA associate easily. Britten and his Coworkers (1966, 68) have demonstrated that many vertebrate DNAs reassociates easily when it is broken into small pieces. This observation gave rise to the hypothesis that certain short sequences of bases are repeated hundred times in DNA, this is the **satellite DNA**.

This repetitive DNA generally contributes at least 20% DNA and can reach upto 90% in some cases. It is believed that these repetitive sequences do not carry any genetic informations and therefore do not form genes, but play some other structural or regulatory role.

Repetitive DNA consists of short identical genes which are repeated in tandem, several hundred or thousand times. Such DNA is found in the region of the chromosome adjacent to the centromere. In many case the base compositions of the repeating sequences are unlike that of the rest of the DNA. It is, therefore easy to separate repetitive DNA by ultracentrifugation.

The satellite DNA can be isolated by density gradient centrifugation in neutral caesium chloride as they have distinctive bouyant densities. The fractions can be separated as a band from the main band of DNA, this band is called satellite band.

The satellite band is found on the left of the main band if lighter and on the right side if heavier than the DNA of the main band.

In *Drosophila virilis*, there are three highly repetitive DNA each consisting of a repeating sequence of seven nucleotide pairs and about 25% of the DNA is satellite DNA (See Table 10.2).

Table 10.2 Satellite DNA of *Drosophila virilis* (Gall *et al.*, 1974)

Repetitive DNA	Buoyant density	Repeat sequence
Sat DNA-I	1.692	5' ACAAACT 3' 3' TGT'TTGÅ5'
Sat DNA-II	1.688	5' ATAAACT 3' 3' TÅTTTGA 5'
Sat DNA-III	1.671	5' ATAAATT 3' 3' TÅTTTÅÅ 5'

In human, 30% of DNA is repetitive and is designated as sat I, II and III.

In mouse, 10% of the DNA is highly repetitive and renatures within a few seconds, 20% is moderately repetitive and reassociates at an intermediate rate, 70% is single copy DNA which renatures very slowly. There are about a million copies of repeating sequences of about 300 bp.

In prokaryotes, the repeated base sequence is not found.

Two remarkable features of satellite DNA are—

- (I) Remarkable (relative) uniformity within the same species.
- (II) Great variability between closely related species.

The satellite DNA often lies in heterochromatic region of chromosomes and its location can be demonstrated by cytological hybridization by incubating the cells in the radioactive solution and is determined by autoradiography.

The function of highly repetitive DNA is unknown. This can replicate but cannot transcribe RNA for protein synthesis. This is probably because the short sequences lack promoter sites on which RNA chains can be initiated by RNA polymerase. Repetitive DNA is, therefore, inert and is partly dispensable.

In the African clawed toad, *Xenopus laevis*, the genes for 40S precursor RNA which give rise to 28S and 18S RNA are repeated about 450 times. The genes are tandemly arranged and are separated by a spacer region of about 5000 bp. Genes for 5S rRNA are also separated by spacer regions and are arranged

in clusters of 100 to 1000 repetitive units at the ends of the most of the 18 chromosomes.

10.3 Complexity

Complexity is the total length of different sequences of DNA present in a given preparation. The double stranded DNA is denatured and converted into single stranded DNA by heating the DNA solution. This is accompanied with increase in optical density, which is called hyperchromicity.

Again when it is allowed to cool, the single stranded DNA is transformed into a double stranded DNA, again the optical density is decreased, this is called hypochromicity. The 50% resaturation is achieved usually at a specific temperature which is called melting temperature (T_m). The formation of double stranded DNA is actually measured over different values of a parameter which is described as $C_0 \cdot t$ (concentration x time). It is the product of DNA concentration and time of incubation in a reassociation reaction.

Complexity of the genome can be described under two heads—

A. Kinetic complexity

The reassociation of DNA in the solution depends on the random collisions between the complementary strands, which follow the second order of **kinetics**, since concentration of both the complementary strand will influence the rate of reaction. The rate of reaction, when expressed through differential calculus is as follows :—

$$\frac{dc}{dt} = -Kc^2 \quad \text{where, } C = \text{Concentration of single stranded DNA at time 't'}$$

$K =$ reassociation rate constant

$$\text{or} \quad \int \frac{dc}{c^2} = -K \int dt$$

$$\text{or,} \quad -\frac{1}{C} = -Kt + A \quad A \text{ is constant}$$

$$\text{or,} \quad \frac{1}{C} = Kt - A$$

When $t = 0$, then $C = C_0$

$$\text{Now,} \quad \frac{1}{C_0} = K \cdot 0 - A$$

$$\text{or, } A = -\frac{1}{C_0}$$

The equation is

$$\frac{1}{C} = Kt + \frac{1}{C_0}$$

When the reaction is half complete then time is ($t_{1/2}$) and in $t_{1/2}$ time the concentration is C. Then—

$$\frac{1}{C} = Kt_{1/2} + \frac{1}{C_0}$$

$$\text{or, } \frac{C_0}{C} = k \cdot C_0 \cdot t_{1/2} + 1$$

$$\text{We know } C = \frac{1}{2}C_0$$

$$\text{then, } 1 + K \cdot C_0 \cdot t_{1/2} = 2$$

$$\text{or, } K = \frac{1}{C_0 \cdot t_{1/2}}$$

$$\text{or, } C_0 \cdot t_{1/2} = \frac{1}{k}$$

So, during reassociation of DNA occurs at the rate constant K (nt. moles / lit/see) in equal to the reciprocals of $C_0 \cdot t_{1/2}$

$C_0 \cdot t_{1/2}$ is the product of DNA concentration and time (t) required to proceed to half completion of the reaction; it is directly proportional to the unique length of reassociating DNA.

The $C_0 \cdot t_{1/2}$ of a reaction indicates the total length of different sequence that are present. This is described as the complexity. It is usually given in base pairs, but can be expressed in daltons or any other mass unit.

A higher $C_0 \cdot t_{1/2}$ means, slower reaction and lower $C_0 \cdot t_{1/2}$ means faster reaction.

If there is no repetitive DNA (because the repetitive DNA reassociates faster), the $C_0 \cdot t_{1/2}$ of a reaction will be directly proportional to the DNA content. In view of this $C_0 \cdot t_{1/2}$ will indicate the length of all the different sequences in a genome, which will be less than the length of the total DNA in a genome when

there is repetition. This is called kinetic complexity.

Kinetic complexity is the complexity of a DNA component measured by the kinetics of the DNA association.

This can be calculated by knowing the $C_0 \cdot t_{1/2}$.

For example, *E. coli* has a genome = 0.004 pg DNA = 4.2×10^6 bp with $C_0 \cdot t_{1/2} = 4$.

The Kinetic complexity of the genome

$$= \frac{C_0 \cdot t_{1/2} \text{ of DNA of any genome} \times \text{base pair}}{C_0 \cdot t_{1/2} \text{ of } E. coli \text{ genome}}$$

$$= \frac{4.2 \times 10^6 \times C_0 \cdot t_{1/2} \text{ of the genome}}{4}$$

In eukaryotes, the genome contain more than one pure components. For example. calf DNA has two component, each with characteristics $C_0 \cdot t_{1/2}$ value. In wheat, more than two such components are found. Proportion of each

component is determined by using the formula $1 - \frac{C}{C_0}$, where C = concentration

at $t_{1/2}$ for corresponding component.

Form this proportion, chemical complexity on the component can be determined.

B. Chemical Complexity : Chemical complexity is the amount of a DNA component measured by chemical assay.

For example, if the genome size is 12×10^8 bp and the component represents 25% of the genome then the chemical complexity of this component is 3×10^8 bp.

Chemical complexity = size of the genome \times % of the component in this group.

If the kinetic complexity is known from earlier equation, then repetition frequency (f) of repetitive DNA component can be determined using the following formula—

$$f = \frac{\text{chemical complexity}}{\text{kinetic complexity}} = \frac{C_0 \cdot t_{1/2} \text{ of non repetitive DNA}}{C_0 \cdot t_{1/2} \text{ of repetitive DNA}}$$

Following table shows the reassociation of a eukaryotic genome starting at a $C_0 \cdot t$ of 10^4 and terminating at $C_0 \cdot t$ of 10^4 . Reaction falls into three types of components and their results are as follows—

	Fast Component	Intermediate Component	Slow Component
% of genome	25	30	15
$C_0 \cdot t_{1/2}$	0.0013	1.9	630
ECinetic complexity (bp)	340	6×10^5	3×10^8
Repetition frequency	500000	350	1

There is a good relationship between the kinetic complexity and chemical complexity of eukaryotic genome.

Usually *E. coli* is used as a standard. Its components are taken to identical with the length of the genome. Thus, complexity of any DNA can be determined by comparing its $C_0 \cdot t_{1/2}$ with that of standard DNA of known DNA complexity.

$$\frac{C_0 \cdot t_{1/2} \text{ of (DNA of any genome)}}{C_0 \cdot t_{1/2} (E.coli \text{ DNA})} = \frac{\text{Complexity of any genome}}{4.2 \times 10^6 \text{ bp}}$$

According to the table, the slow component represents 45% of the total DNA, so the concentration in the reassociation reaction is 0.45 of the measured concentration (total amount of DNA).

If DNA were isolated as a pure component, free of other fractions, it would renature with $C_0 \cdot t_{1/2}$ of $0.45 \times 630 = 283$.

Suppose that under these conditions *E. coli* DNA reassociates with a $C_0 \cdot t_{1/2}$ of 4.0. Comparing these two clues, we see :

$$\begin{aligned} \text{The kinetic complexity} &= \frac{C_0 \cdot t_{1/2} \text{ of DNA of any genome} \times 4.2 \times 10^6}{4} \\ &= \frac{0.45 \times 630 \times 4.2 \times 10^6}{4} \\ &= 3 \times 10^8 \text{ bp} \end{aligned}$$

$$\text{Then the whole genome is} = 3 \times 10^8 \times \frac{1}{0.45} = 6.6 \times 10^8 \text{ bp}$$

$$\text{First component complexity} = \frac{0.0013 \times 0.25 \times 4.2 \times 10^6}{4} = 340$$

$$\text{Second component complexity} = \frac{1.9 \times 0.30 \times 4.2 \times 10^6}{4} = 6 \times 10^5$$

Reversing the argument if we took three DNA preparations, each containing a unique sequence of the appropriate length 340 bp, 6×10^5 bp and 3×10^8 bp respectively and mix them in the proportions = 25 : 30 : 45, each would renature as though it was a single component, together the mixture would display the same kinetics as those determined for the whole genome.

Non repetitive DNA complexity can estimate the genome size. The complexity of the slow components comprise sequences that are unique in the genome upon denaturation each single stranded sequence is able to renature only with the corresponding complementary sequences. It is usually the major component in eukaryotes. It is called non-repetitive DNA. According to the table the complexity of non-repetitive DNA is 3×10^8 bp. If this fraction is unique and represent 45% of the genome, then the whole genome would have a size of $3 \times 10^8 \div 0.45 = 6.6 \times 10^8$ bp. This provides an independent assessment of genome size. The value is approximately 7×10^8 , obtained from the result of chemical complexity.

Eukaryotic genomes certainly contain repetitive sequences. Intermediate component occupies 30% of the genome. According to chemical complexity the total amount is $0.30 \times 7 \times 10^8 = 2.1 \times 10^8$ bp.

But kinetic complexity of this component is only 6×10^5 pp.

$$\text{Thus repetition frequency} = \frac{2.1 \times 10^8}{6 \times 10^5} = 350$$

Thus intermediate components behaves as though consisting of a sequence of 6×10^5 bp that present in 350 copies in every genome. Repetition frequency (f) is the number of copies present per genome.

Highly repetitive DNA takes the name from the very large number of copies of the basic reassociating sequence present. The fast component consists 310 bp long in 500000 copies per genome. Because of the short length of the reassociating unit sometimes this is also referred to as simple sequence DNA.

$$\begin{aligned} \text{The repetition frequency (f)} &= \frac{C_0 \cdot t_{1/2} \text{ of non repetitive DNA}}{C_0 \cdot t_{1/2} \text{ of repetitive DNA}} \\ &= \frac{630}{0.0013} = 500000 (\text{Approx.}) \end{aligned}$$

Unit 11 □ Linkage Map, Cytogenetic Mapping

Structure

11.1 Physical and molecular maps

11.2 Restriction mapping of genes

11.3 DNA foot printing

11.4 Micro satellite mapping

11.1 Physical and molecular maps

STS, i.e. Sequence Tagged Site is a short DNA sequence generally between 100-500 bp in length that is easily recognisable and occurs in the chromosome only once (i.e. unique).

STS mapping is a physical mapping procedure that locates the positions of sequence tagged sites (STSs) in a genome.

11.1.1 Qualities of STS

(1) Its sequence must be known, so that a PCR (polymerase chain reaction) assay can be set up to test the presence or absence of the STS on different DNA fragments.

(2) It must have a unique location in the chromosome being studied once in the genome. If the STS sequence occurs in more than one position, then the mapping data will be ambiguous. So STSs do not include sequences found in repetitive DNA.

11.1.2 Sources of STS

(1) **Expressed sequence tags (ESTs)** : These are short sequences obtained by the analysis of cDNA clones. cDNA is prepared by converting mRNA into dsDNA. Thus ESTs represent the genes that are expressed in the cell.

(2) **SS4Ps** : These are also used in genetic mapping as well as physical mapping of genes.

(3) **Random genomic sequences (RGS)** : These are obtained by sequencing random pieces of cloned genomic DNA.

11.1.3 Principles of STS mapping

(1) To map a set of STSs, a collection of overlapping DNA fragments from single chromosome or from the entire genome is needed.

(2) The data from which the map will be derived are obtained by determining fragments which contain STSs. This can be done by hybridization analysis but PCR is generally used which is quieter and automated process.

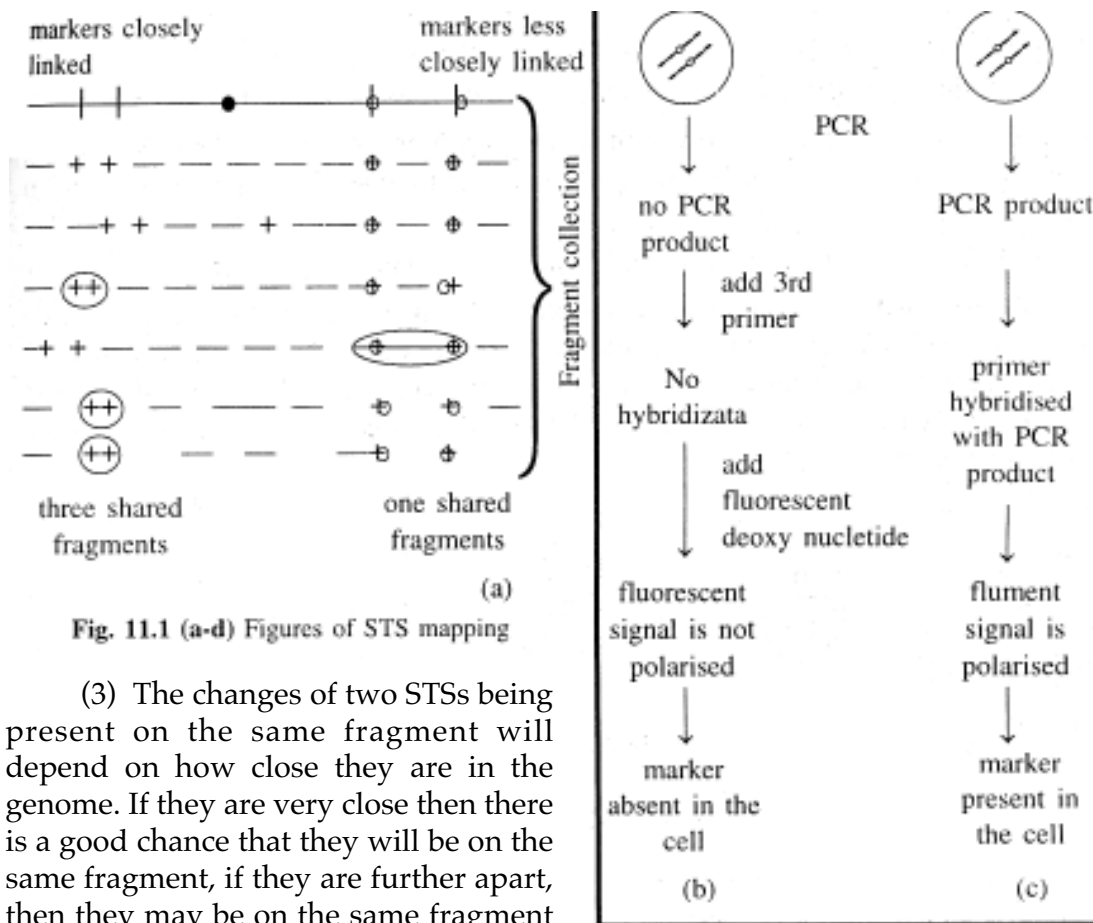


Fig. 11.1 (a-d) Figures of STS mapping

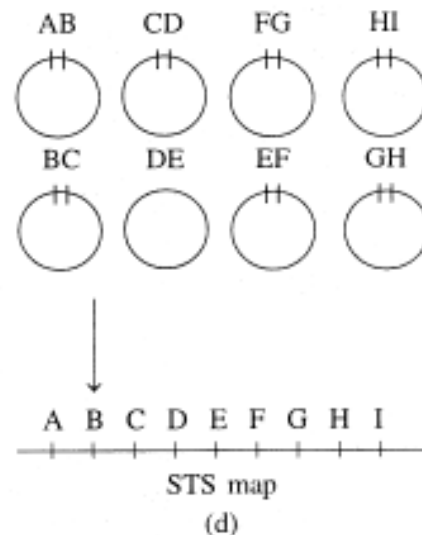
(3) The changes of two STSs being present on the same fragment will depend on how close they are in the genome. If they are very close then there is a good chance that they will be on the same fragment, if they are further apart, then they may be on the same fragment or not.

(4) The data can be used to calculate the distance between two markers, where map distance is based on the frequency at which breaks occur between two markers.

11.1.4 Physical mapping of chromosomes by screening YAC clones of STSs

Segments of human DNA upto 1000 kb long can be cloned in yeast artificial chromosomes (YACs).

(1) Aliquots of DNA prepared from each YAC clone viz., A, B and C are



subjected to PCR amplification using primer pairs (1-6) corresponding to the ends of various STSs. Only those clones containing STSs with ends complementary to particular primers will be amplified.

(2) Electrophoretic analysis then shows that YAC clones contain STSs.

(3) The illustration is very simple showing 6 primer pairs. Clone-A contains STSs no. 1,3 and 5; clone-B contains 2 and 1 and clone-C contains 3, 4, 5 and 6 primer pairs.

(4) The three YAC clones can be ordered showing their relative positions. In the mapping of human chromosome-21, about 1,20,000 clones from two separate YAC libraries were screened. In addition, 14,000 YACs isolated from a library prepared specifically from chromosome-21 were screened individually. By the use of 198 STSs, researchers identified 810 positive clones and ordered them into a contiguous map.

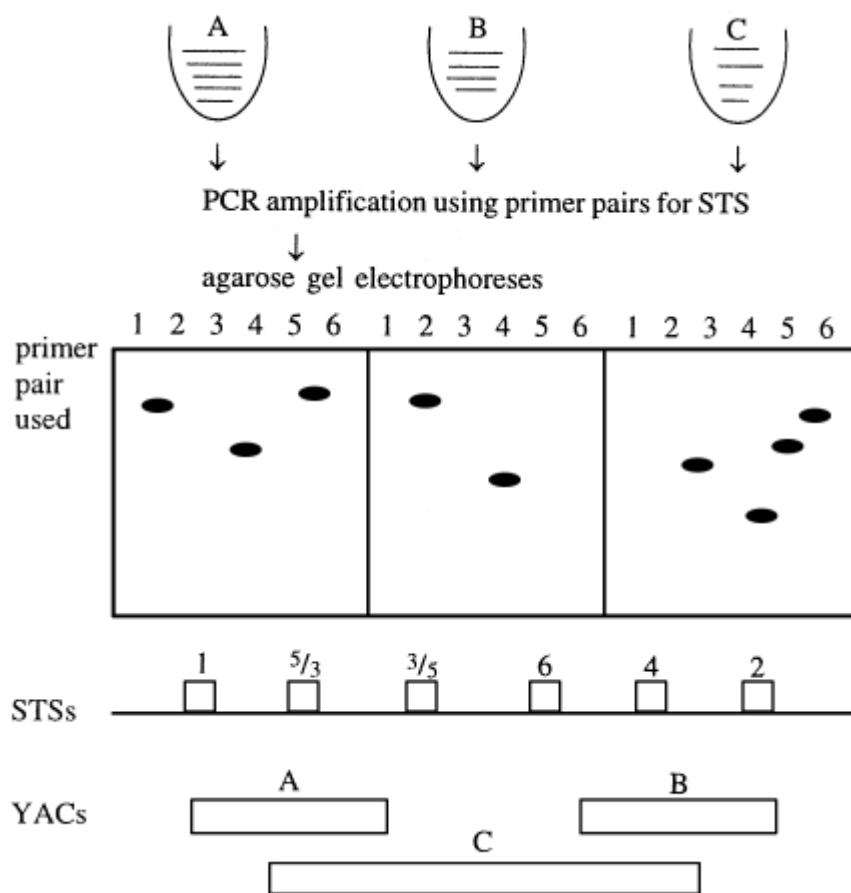


Fig. 11.2 Ordering of contiguous overlapping YAC

11.1.5 Fragments of DNA for STS mapping

At first, collection of DNA fragments sparing the chromosome or genome is required. The collection is called mapping reagent. There are two methods as follows—

A. Radiation hybrid method :

(1) Human cells are exposed to X-ray doses of 3000-8000 rads, which causes the chromosomes to break randomly into fragments. Higher doses produce smaller fragments.

(2) These fragments can be propagated if the irradiated cell is fused with non-irradiated hamster (or other rodent) cells. Fusion is achieved chemically with polyethylene glycol or by exposure to Sendai virus.

(3) The hamster cell line that is unable to make either thymidine kinase (TK) or hypoxanthine phosphoribosyl transferase (HPRT) is used for the purpose. Those cells incapable of taking up human chromosome fragments are unable to survive in HAT (hypoxanthine, aminopterin and thymidine) medium. Those cells taking up chromosome fragments can synthesize TK and HPRT and are able to grow in HAT. Thus the hybrid cells are collected. These hybrid cells are then used as a mapping reagent in STS mapping.

(4) To map, as many markers as possible are used. A pair of primers is designed for every DNA marker that was to be tested. Each primer pair is specific for one marker and will not give a PCR product with any other part of the genome.

(5) The success or failure of PCR is determined by agarose gel electrophoresis. The presence of a band of the expected size in the gel indicates that the PCR has worked.

(6) Another procedure has been designed where a third specific primer for each marker is added to the reaction mixture along with a fluorescently labelled dideoxy-nucleotide. If PCR has been successful, then the third primer anneals to the product and is extended by the fluorescent dideoxy nucleotide and emits signals from this.

B. Use of clone library as the mapping reagent for STS analysis

A clone library can also be used as a mapping reagent in STS analysis. The clone library is prepared by using the genome or chromosome which is broken into fragments and are integrated into a high capacity vector.

The single specific chromosome can be isolated by flow cytometry technique and a clone library of a chromosome is made possible.

The data obtained from STS analysis is used for preparation of the physical map.

11.2 Restriction mapping of genes

Restriction mapping of genes is a physical mapping which locates the relative positions on a DNA molecule of the recognition sequences for restriction endonucleases.

Genetic mapping using RFLPs as DNA markers can locate the positions of polymorphic restriction sites within a genome, but very few of the restriction sites in a genome are polymorphic, so many sites are not mapped by this technique. Restriction mapping is very useful to solve a problem but the limitation of the technique is that it is applicable only to relatively small DNA molecules.

Two methods are employed for restriction mapping *viz.*, partial digestion method and double digestion method.

(1) In **partial** digestion method, the circular or linear DNA is treated with a particular restriction enzyme. But the enzyme is prevented from the complete digestion of DNA molecule. So the DNA is incubated for a short time or using a suboptimal incubation temperature. This leads to the partial digestion and will produce many cut products and uncut products. If the DNA molecule is linear, the terminal end of the DNA is labeled with P^{32} by the polynucleotide kinase reaction prior to cleavage and only radioactively labelled fragments are considered in agarose gel electrophoresis, the other fragments are ignored/discarded of the molecular weight of each enhanced band is invariably the sum of the molecular weights of two fragments which are considered to be adjacent. Thus the relative positions of the fragments can be ordered.

(2) Second method of the restriction mapping is the double digestion method, although partial digestion method is usually followed. In double digestion method, three samples of a particular DNA species are taken. Each two of these is treated with two separate restriction enzymes and the third sample is treated with both the restriction enzymes. Thus the three sets of fragments are compared following the agarose gel electrophoresis. The terminal end of the DNA molecule is also labeled with radioactive P^{32} prior to the restriction enzyme cleavage.

11.3 DNA foot printing

When transcription factor binds to a DNA sequence, it protects that sequence from digestion by nucleases. Researchers take advantage of this property by isolating chromatin from cells and treating it with DN-digesting enzymes, such as DNAase-I, that destroy sections of the DNA that are not protected by bound transcription factors. Once the chromatin has been digested, the bound protein

is removed and the DNA sequences that had been protected are identified. This method is called DNA foot printing. This is also used to locate the binding sites of proteins on RNA.

Method

- (1) A pure DNA fragment that is labeled at one end with ^{32}P is isolated.
- (2) This molecule is then cleaved with a nuclease or a chemical that makes random single-stranded cuts in the DNA.
- (3) The DNA molecule is then denatured to separate into two strands.
- (4) The resultant fragments from the labeled strand are separated on a gel and detected by autoradiography.

The pattern of bands from DNA cut in the presence of a DNA-binding protein is then compared with that from DNA cut in its absence. The protein covers the nucleotides at its binding site and protects from DNA ase. The labeled fragments that shows no cleavage will show an area which is missing in the electrophoretic gel, is leaving a gap is called "foot print".

11.4 Gene mapping by human pedigree analysis (microsatellite mapping)

Recombination (CO) mapping is very difficult in human because

- (1) It is impossible to preselect the genotypes of parents and set up crosses.
- (2) The data for the calculation of recombination frequencies have to be obtained by examining the genotypes of the members of successive generations of existing families.
- (3) The data obtained are very limited and their interpretation is often - difficult because in human test cross rarely occurs and the number of family members and offsprings are not many.

Therefore, gene mapping in human may be done by pedigree analysis.

Let us, suppose a family of two parents and six children were studied with a genetic disease. The diseased state is due to one allele and the healthy state is due to second allele. Diseased allele is dominant over healthy allele.

The pedigree showed that mother is affected because four of her children are affected by the disease (3♂ + 1♀).

The grand mother is affected. The grand father is dead but? We can assume that he was also affected. We can include them in the pedigree analysis.

The aim is to map the position of the gene for the genetic disease. For **that** purpose one is studying its linkage to a microsatellite marker M, four alleles of which *viz.*, M₁, M₂, M₃ and M₄ are present in the living family members. One has now to calculate the number of children who are recombinants.

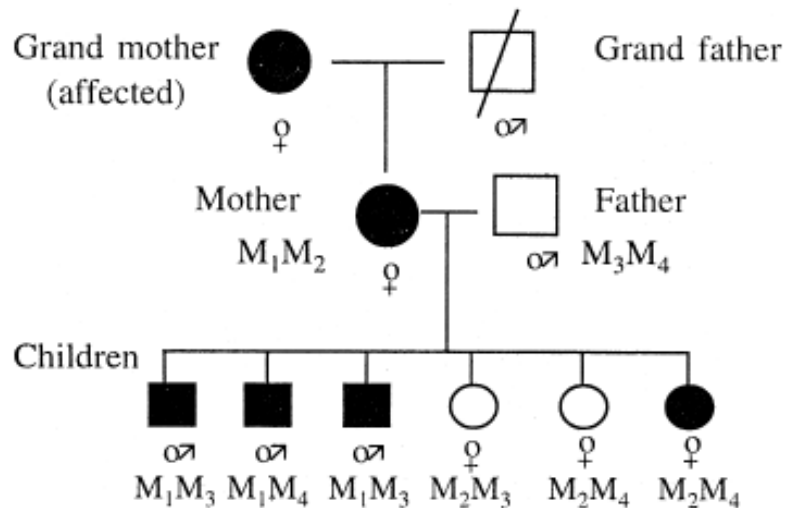


Fig. 11.3 Pedigree of a family with inheritance of genetic disease (solid means affected person)

The pedigree can be interpreted by two different hypotheses because child-1, 2 and 3 have the disease allele and microsatellite allele M₁, the child-4 and 5 have the healthy allele and microsatellite allele M₂.

In hypothesis-1, the mother would have the genotype = $\frac{\text{Disease allele} - M_1}{\text{Healthy allele} - M_2}$

The child no-1, 2, 3, 4 and 5 all would have parental genotype. Only the child no-6 would be a recombinant. The recombination percentage is 16.66 i.e., the disease gene and **the** microsatellite allele are relatively closely linked.

In hypothesis- II, the genotype of the mother would be = $\frac{\text{Healthy} - M_1}{\text{Disease} - M_2}$

Here the child **no-1**, 2, 3, 4, 5 are recombinants while child no-6 is with parental genotype. The recombination percentage is 83.33, which means that the disease gene and microsatellite gene are far apart on the chromosome.

		Possible genotypes of the mother	
		Hypothesis-I	Hypothesis-II
		Disease-M₁	Healthy-M₁
		=====	=====
Child-1	Disease-M₁	Healthy-M₂	Disease-M₂
Child-2	Disease-M₁	Parental	Recombinant
Child-3	Disease-M₁	Parental	Recombinant
Child-4	Healthy-M₂	Parental	Recombinant
Child-5	Healthy-M₂	Parental	Recombinant
Child-6	Disease-M₂	Recombinant	Parental
	Recombination frequency	16.66%	83.33%

Fig. 11.4 Probable interpretation of the pedigree

Imperfect pedigrees are analysed statistically by using a measure called “lod score” (Morton, 1955). This stands for logarithms of the odds that genes are linked. This is used to determine whether the two markers lie on the same chromosome or not. If lod analysis establishes the linkage then the data will give confidence about their recombination frequencies.

If the number of the family members are larger the result would be more satisfactory. Atleast three generations are to be tested. Atleast four grand parents and atleast eight second generation children could be sampled.

Unit 12 □ Genetics of Cell Cycle

Structure

12.1 Genetic regulation of cell division in yeast and eukaryotes

12.2 Molecular basis of cellular Checkpoints

12.1 Genetic regulation of cell division in yeast and eukaryotes

12.1.1 Introduction

A cell reproduces by performing an orderly sequence of events in which it duplicates its contents and then divides into two. This cycle of duplication and division is known as the cell cycle. The basic organization of the cell cycle and its control system are essentially the same in all eukaryotic cells. Three **eukaryotic** systems in which cell-cycle is commonly studied are yeasts, frog embryo and cultured mammalian cells.

Yeasts are tiny, single-celled fungi. Two species are generally used in studies of cell cycle. The fission yeast viz. *Schizosaccharomyces Pombe*, is a rod shaped cell that grows by elongation at its ends. Division occurs by the formation of a septum or cell plate in the centre of the rod. It has a typical eukaryotic cell cycle with G_1 , S, G_2 and M phases. In contrast with that happening in higher eukaryotic cells, the nuclear envelope of the yeast cell does not break down during M-phase. The microtubules of the mitotic spindle are formed inside the nucleus and are attached to spindle pole bodies (SPB) at its periphery. The cell divides by forming a partition (cell plate) and splitting into two. The mitotic chromosomes are readily visible.

The budding yeast *Saccharomyces cerevisiae*, also called baker's yeast is a oval cell and divides by forming buds which first appears during G_1 and grows until it separates from the mother cell after mitosis. It has normal G_1 and S-phase but does not have a normal G_2 phase. The microtubule based spindle begins to form inside the nucleus early in the cycle during S-phase. Nuclear envelope remains intact during mitosis and the spindle forms within the nucleus.

12.1.2 Genetic regulation of cell cycle in *S.pombe* (Fig. 12.1)

(1) *ede 2* is identified as a crucial regulator by its involvement at both stages of cell cycle block, i.e. between G_2 and M-phase and in G_1 at start.

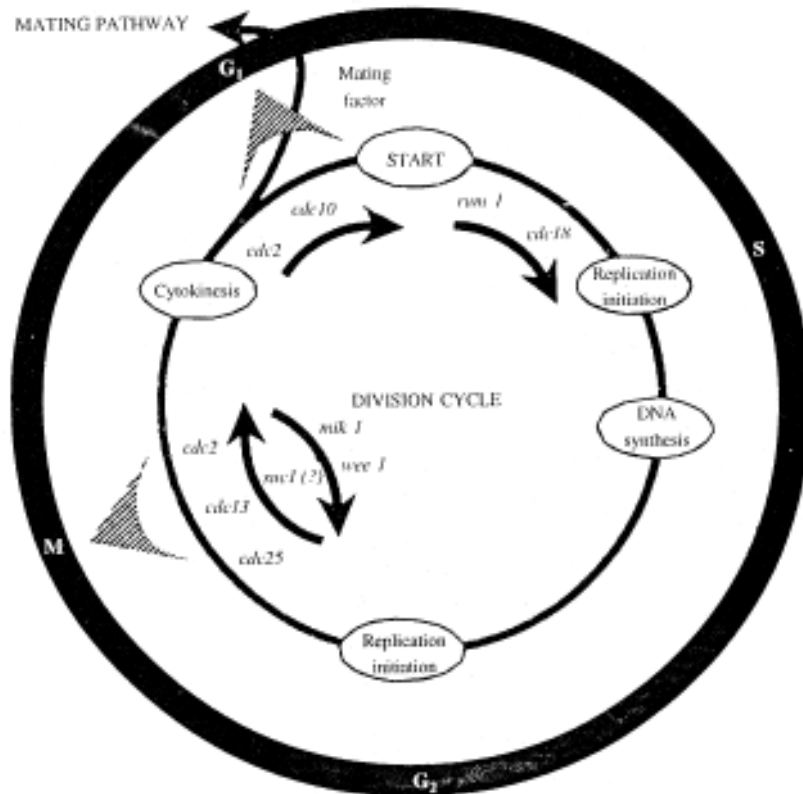


Fig. 12.1 The cell cycle in *S. pombe* requires *cdc* genes to pass specific stages, but may be retarded by genes that respond to cell size (*wee 1*). Cells may be diverted into the mating pathway early in G1

(2) During G_2 and M-phase *cdc 2* has a partner which is the product of *Cdc 13* generating an M-phase kinase (resembling the p^{34} -B cyclin dimer of animal cells). The activity of the *cdc 2* catalytic subunit in these dimers are controlled by phosphorylation (in the same way as p^{34} in animal cell).

But the difference that in yeast there is no Thr-14 so there are only two relevant sites Tyr-15 where phosphorylation is inhibitory and Thr-161 where phosphorylation is required.

(3) Under normal conditions, the cell division cycle is related to the size of the cell. In the poor growth condition the cells increase in size more slowly. The genes involved in control cell size are identified, *wee 1* usually inhibits cells from initiating mitosis until their size is adequate. It has been suggested that *wee 1* is part of a check point that prevents the activation of *cdc 2* until an adequate mass has been attained. *wee 1* codes for a kinase and can phosphorylate serine or

threonine and tyrosine. It inactivates cdc 2 by phosphorylating Tyr-15. Another gene viz., *mik 1* has similar effects.

- (4) The product of *cdc 25* is required for the dephosphorylation of cdc 2 in the cdc 2/cdc 13 dimer. It is probably responsible for the key dephosphorylating event in activating the M-phase kinase. The level of cdc 25 increases at mitosis and its accumulation over a threshold level could be important, cdc 25 executes the checkpoint that ensures s-phase to be completed before M-phase can be activated.
- (5) The products of *wee 1* and *cdc 25* plays antagonistic roles. The kinase activity of *wee 1* acts on Tyr-15 to inhibit cdc 2 function. The phosphatase activity of *cdc 25* acts on the same site to activate cdc 2.
- (6) During mitosis the cdc 2 of cdc 2/cdc 13 dimer is in the active state that locks the phosphate at Tyr 15 and has the phosphate at Thr 161. At the end of mitosis kinase activity is lost and cdc 13 is degraded cdc 2 does not change at this point.
- (7) During G, the active form of cdc 2 has a different partner, the B-like cyclin, cig 2 encoded by cig 2 gene. The dimer is converted from inactive state to the active state by dephosphorylation of Tyr-15 residue of cdc 2.
- (8) Progression of G₁ into S is controlled by activation of cdc 2-G₁ cyclin. In *S. pombe* it is cdc 2/ cig 2.
- (9) Transcription of *cdc 18* is activated as a consequence of passing START and cdc 18 is required to enter S-phase. Over expression of cdc 18 allows multiple cycle of DNA replication without mitosis.
- (10) For cdc 18 to be active, cdc 2/cdc 13 must be inactive. Again when M-phase kinase is active, it causes cdc 18 to be inactive possibly by phosphorylating it and prevents initiation of another S-phase.
- (11) Activity of cdc 2/cdc 13 M phase kinase is influenced by a factor rum-1, which controls entry into S-phase. When rum-1 is depleted, premature entrance into mitosis occurs and over expression causes cells to fail to enter mitosis. This suggests that M_{m-1} is an inhibitor of the M-phase kinase. It is expressed between G₁ and G₂ and keep M-phase kinase in an inactive state.

12.1.3 Genetic regulation of cell cycle in *S.cerevisiae* (Fig. 12.2)

Cell cycle in *S.cerevisiae* consists of three cycles that separate after START and join before cytokinesis. The cells may be diverted into the mating pathway early in G₁.

(1) **Chromosome cycle:** In this cycle, duplication and separation of chromosomes, completion of S-phase and nuclear division occurs. Mutation of *cdc 8* stops this cycle in S-phase. Mutation in the chromosome cycle do not stops the cytoplasmic cycle.

(2) **Cytoplasmic cycle :** It consists of bud emergence and nuclear migration into the buds. This cycle can be halted before bud emergence by *cdc 24* mutation but the mutation does not prevent chromosome replication.

(3) **Centrosome cycle :** This cycle consists of duplication and separation of spindle polar body (SPB) and organizes microtubules to allow chromosome segregation within the nucleus. Blocking of the cycle by *cdc 31* does not prevent S-phase or bud emergence.

Completion of entire cell cycle requires all three constituent cycles because nucleokinesis needs both chromosome and centrosome cycles but cytokinesis requires all the 3 cycles.

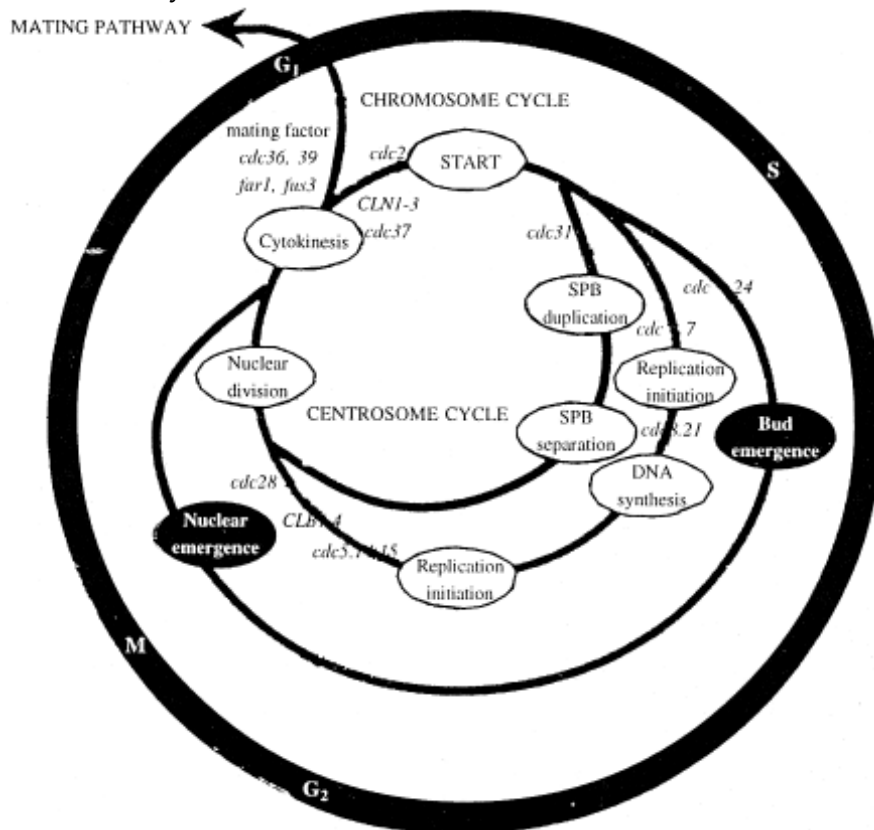
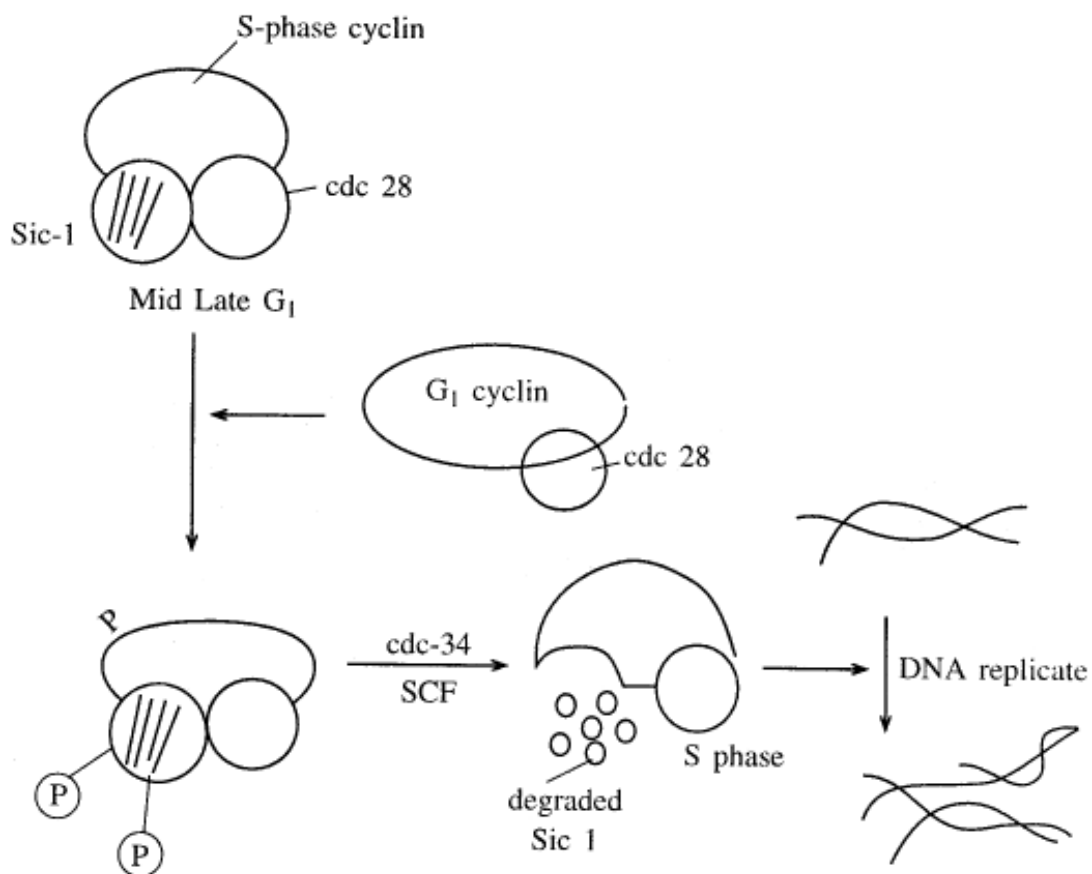


Fig. 12.2 The cell cycle in *S. cerevisiae* consists of three cycles that separate after START and join before cytokinesis. Cells may be diverted into the mating pathway early in G₁

(1) *S. cerevisiae* expresses a single cyclin dependent protein kinase (cdk) encoded by *cdc 28* gene which interacts with several cyclins during different phases of the cell cycle.

(2) Just after cytokinesis, the decision on whether to initiate a division cycle is made before the START. The cells can be diverted into mating type pathway by mating factors and *cdc 36* and *cdc 39* which appear to block the cell cycle before START and really function in the mating type pathway. Mutants block cell cycle by diverting cells into mating even in absence of the mating gene.

(3) After cytokinesis the mother cell and the bud both the cell remain in the G_1 phase of the cell cycle. The ability to pass START is determined by environmental conditions. Prior to bud formation, spindle body duplication and DNA replication, the yeast cell exhausts its nutrients. When *S.cerevisiae* cells in G_1 have grown sufficiently in the growth medium, they begin a programme of



Fig, 12.3 Mechanism of Genetic regulation

gene expression that leads to entry into mitosis. Once G_1 cells reach the critical size, they become committed to completing cell cycle even if they are shifted to low nutrient medium.

(4) The crucial gene in passing START is *cdc-28* which is homologous to *cdc-2* in *S.pombe*. Three cyclins are active in G_1 . Cln 1, Cln 2 and Cln 3 are encoded by *CLN-1*, *CLN-2*, *CLN-3* respectively. Mutations in any one or two of these genes fail to block the cell cycle, thus the *CLN* genes are functionally redundant.

(5) The complexes formed between *cdc 28* and the three G_1 cyclins (Cln 1, Cln 2 and Cln 3) have protein kinase activity and constitute the hypothesized S phase promoting factor (SPF_S). In wild type yeast cells, Cln 3 is expressed at a nearly constant level throughout the cell cycle. Cln 1 and Cln 2 are expressed during the second half of G_1 and they increase rapidly and when their accumulation exceeded a critical threshold level, triggers the passage of *start* (START) into S-phase. After that its concentration declines gradually and are eliminated by the time of mitosis.

(6) *cdc 28*-Cln 3 phosphorylates and activates SBF and MBF. These induce transcription of *CLN 1* and *CLN 2* genes as well as several other genes required for DNA replication, including genes encoding DNA polymerase, RPA (ssDNA binding proteins), DNA ligase and certain enzymes acquired for deoxyribonucleoside triphosphate synthesis.

(7) *cdc 28*-Cln 1 and *ede 28*-Cln 2 phosphorylate APC in late G_1 and inactivate it.

(8) Two B-type cyclin genes *CLB 5* and *CLB 6* are also regulated by MBF and transcribed beginning in late G_1 . The corresponding proteins Clb 5 and Clb 6 accumulate because of the inactivation of APC.

(9) At late G_1 *ede 28*-Clb 5 and *ede 28*-Clb 6 heterodimers accumulate and are inactivated by Sic-1 (an S-phase inhibitor), but it has no effect on *ede 28*-Cln complexes. Sic 1 is degraded following polyubiquitination by E_2 associated with E_3 . Once Sic is degraded, *ede 28*-Clb 5 and *ede 28*-Clb 6 kinases induce DNA replication.

(10) Initiation of DNA replication needs both assembly of pre replication complex and an active *cdc 28*-Cln complex. A second heterodimeric protein kinase *ede 7*-Dbf 4, which is expressed in G_1 is also required to trigger initiation. Once replication has initiated, Mem proteins and *ede 45* move away from the origin along with DNA polymerases. Mem proteins are homologous to helicase, associated with replication fork movements.

(11) Later in S-phase, transcription of the genes *CLB 3* and *CLB 4* begins, encoding two additional B-type cyclins, Clb 3 and Clb 4, which also form

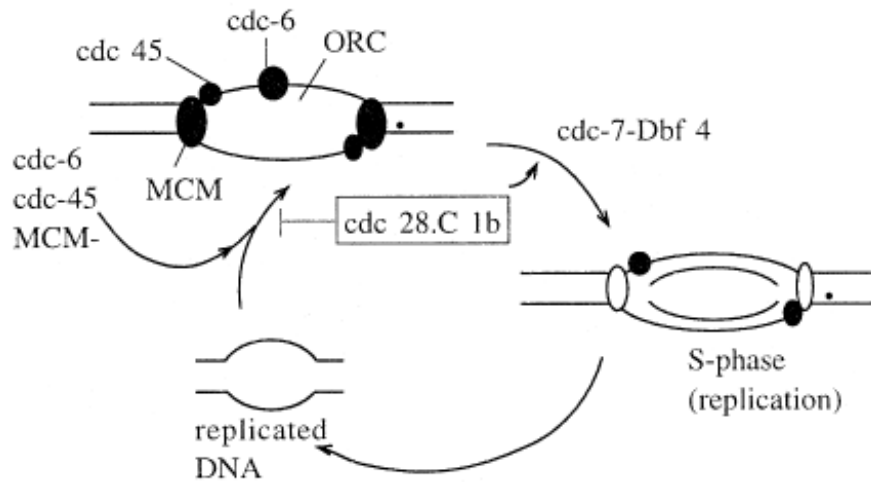


Fig. 12.4 Role of cdc in replication

heterodimeric protein kinases with cdc 28. These two cdc 28-Clb 3 and cdc 28-Clb 4 complexes also initiate the formation of mitotic spindle at the beginning of mitosis.

(12) As cells complete chromosome replication and enter G_2 , two more B-cyclins are expressed. These are Clb 1 and Clb 2 encoded by *CLB 1* and *CLB 2* genes. These function as mitotic cyclins, associating with cdc 28 to form complexes that are required for chromosome segregation and nuclear division.

12.2 Molecular basis of cellular checkpoints

12.2.1 Introduction

In most cells there are several points in the cell cycle, called *checkpoints* in which the cycle can be arrested if previous events have not been completed. Four checkpoint control can arrest the passage through cell cycle. These are (a) G_1 -arrest due to DNA damage, (b) S-arrest due to unreplicated DNA, (c) C_2 -arrest due to DNA damage, (d) M-arrest due to improper spindle formation.

In the checkpoints, the control system can be regulated by extra cellular signals from other cells. These signals either promote or inhibit cell proliferation. Checkpoints generally operate through negative intracellular signals.

12.2.2 The DNA replication checkpoint

Most of the cells by DNA replication checkpoint mechanism, avoided entry into cell division until the last nucleotide in the genome has been copied. The

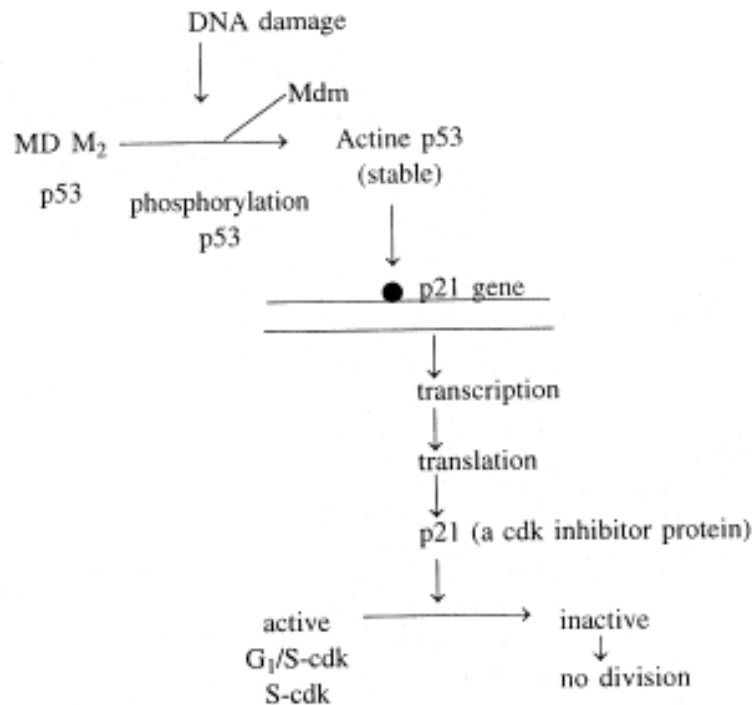


Fig. 12.5(a) Mechanism of G₁ arrest checkpoint

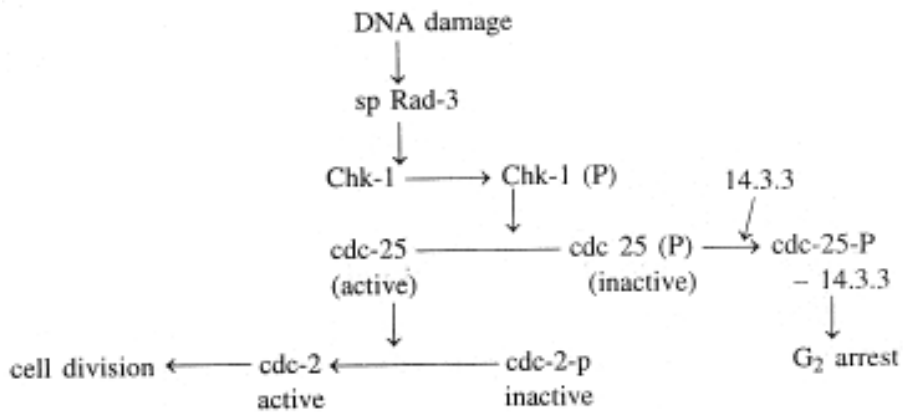


Fig. 12.5(b) Mechanism of G₂ arrest checkpoint

molecular mechanism has not been discovered but any signal from unreplicated DNA or unfinished replication forks send a negative signal to the cell cycle control system that blocks the activation of M-Cdk. Thus normal cells treated with chemical inhibitors of DNA synthesis viz., hydroxyurea, do not progress

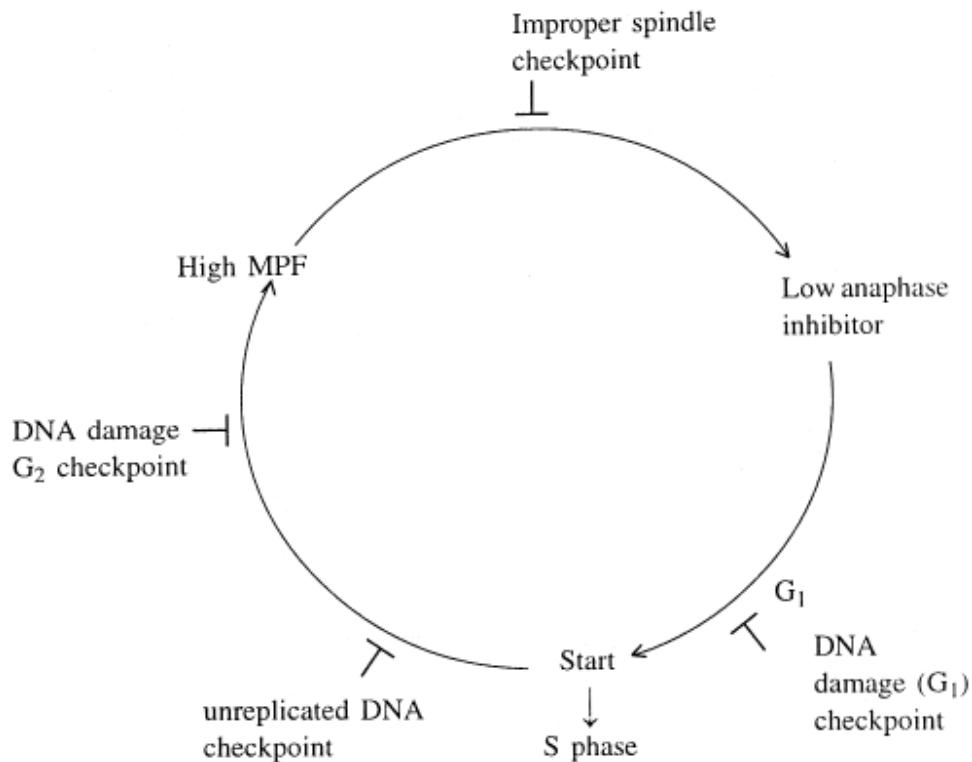


Fig. 12.5(c) DNA replication checkpoint

into mitosis. The block activates a checkpoint mechanism that arrests the cells in S-phase, delaying mitosis. But if caffeine is added along with hydroxyurea checkpoint mechanism fails and the cells proceed into mitosis according to their normal schedule with incompletely replicated DNA. As a result, the cells die.

12.2.3 The spindle attachment checkpoint

The effect of colchicin which inhibits spindle assembly, shows the presence of this checkpoint. In most cell types, a spindle attachment checkpoint mechanism operates to ensure that all chromosomes are properly attached to the spindle before sister chromatid separation occurs. During metaphase kinetochore regions of the chromosomes are attached to microtubules and any kinetochore that is not properly attached to the spindle, sends out a negative wait signal to the cell cycle control system that blocks APC activation which is needed for sister chromatid separation. The nature of signal is not clear. But it has been seen that several proteins, including Mad 2 are recruited to unattached kinetochores which

are required for spindle attachment checkpoint to function. Even a single unattached kinetochore results inhibition of cdc 20-APC activation by binding with Mad 2.

In mice **MAD 2** and **BUB 1** and in humans **MAD 2** genes have been recently identified. MAD 2 protein remains concentrated at the kinetochore until completion of microtubules attachment. This protein is continuously migrating into the cytoplasm and broadcasting signal throughout the cytoplasm. In mammals MAD 2 is associated with p⁵⁵ CDC but in budding yeast MAD 2 is with cdc 20 and in fission yeast Mad 2-slp 1 (sip = sleepy) form a large complex with APC. Hct 1, another protein is associated with APC. APC is kept in check by MAD 2 and when APC is active the cell initiate anaphase by catalyzing degradation of the two proteins with the help of Hct and cdc 20.

cdc 20/slp 1 promotes degradation of CUT2 but the activity of cdc 20/slp 1 is inhibited by MAD 2. Hct 1 promotes degradation of CLB-1, APC catalyzes this degradation by ubiquitination. This ubiquitination is inhibited by spindle attachment checkpoints.

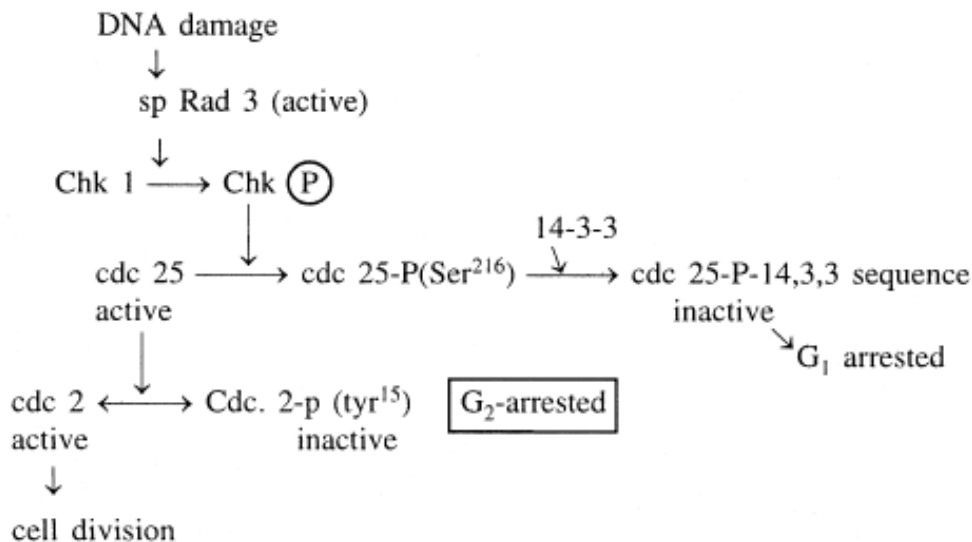


Fig. 12.6 DNA damage checkpoint

In mammals BUB 1 and BUB 3 proteins along with CNEP-E protein induce changes that lastly shut down the transmission of wait signal. MAD 2 dissociates from p⁵⁵ cdc and loses its control on APC. APC is activated and catalyzes breakdown of cyclins and facilitates anaphase separation of chromatid.

12.2.4 DNA damage check points

The cell cycle control system can readily detect DNA *damage* and *arrest* the cell cycle at DNA damage checkpoints. These two checkpoints are one is in late G₁ which prevents entry into S-phase and the other is in late G₂ which prevents entry into, mitosis.

G₂ checkpoint depends on a (*similar*) mechanism that delays entry into mitosis in response to incomplete DNA replication. The damaged DNA sends a signal to a series of protein kinases that phosphorylate and inactivate the phosphatase-cdc 25. This blocks the dephosphorylation and activation of M-cdk, thereby blocking entry into mitosis when the damaged DNA is repaired, the inhibitory signal is turned off and the cell division continues.

In the yeast *S.pombe*, the checkpoints sense the DNA damage and transduce inhibitory signal. Four genes including *RAD 9*, *RAD 17*, *RAD 24* and *MEC 3* will sense this damage. The model proposed that DNA damage activates a protein sp Rad 3 (sp prefix means *S. pombe*). It brings about phosphorylation of Chk 1 and Chk 1-Ⓟ functions as a kinase, which brings about phosphorylation at Ser²¹⁶ of Cdc 25. It then promotes binding of cdc 25 to a protein 14-3-3, coded by *rad 24* and *rad 25*, leading to sequestration of cdc 25. Then *cdc 25* is not available for activation of *cdc 2 P (tyr¹⁵)* and the cell division is arrested at G₂-M transition.

G₁ checkpoint blocks progression into S-phase by inhibition of G₁S-Cdk and S-Cdk complex.

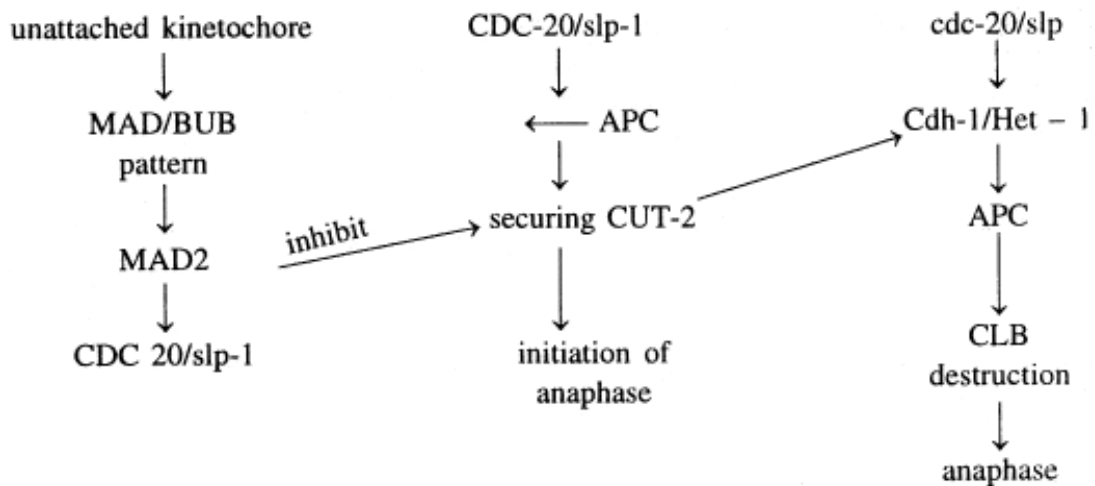


Fig. 12.7 Spindle attachment checkpoint

(1) In mammalian cells a gene regulatory protein p^{53} is being activated by DNA damage and it stimulates transcription of many genes.

(2) A CKI protein, called p^{21} is encoded by such activated genes which binds to G_1/S -Cdk and S-Cdk and inhibits their activity and thus blocks entry into mitosis.

(3) Actually in an undamaged cell, p^{53} is very unstable and is present at a low concentration. It interacts with another protein Mdm2, that causes destruction of p^{53} by ubiquitination mechanism. Damaged DNA activates protein kinase, phosphorylates p^{53} and reduce the binding of Mdm2. As a result p^{53} concentration rises and stimulates gene transcription of p^{21} which inactivates G_1/S -Cdk and S-Cdk activity.

(4) A rare genetic disease, Ataxia telangiectasia is caused by a defect in one of the protein kinases that phosphorylates and activates p^{53} in response to radiation and due to the loss of the DNA damage checkpoints, they suffer from increased rate of cancer.

