Chromosomal changes associated with changes in development

By CHRISTOPHER J. BOSTOCK

MRC Mammalian Genome Unit, King's Buildings, West Mains Road, Edinburgh EH9 3JT, U.K.

TABLE OF CONTENTS

Introduction
What is meant by genome change and developmental change?
Rearrangements not involving gain or loss of DNA
  Mating-type locus of yeast
  Variant surface glycoprotein genes of trypanosomes
  Immunoglobulin genes of mammals
  Chromosomal rearrangements associated with the tumour phenotype
Rearrangements involving quantitative changes
  Amplification of ribosomal RNA genes during oogenesis
  Genomic happenings in ciliates
    Ribosomal DNA and genome rearrangements in Tetrahymena
    Gene amplification and diminution in hypotrichous ciliates
    Differential gene amplification in polytene tissues of insects
      Chorion protein genes in Drosophila
      DNA 'puffs'
  Gene amplification in mammalian cells.
    In vitro selection
    Non-selected tumour cells
  DNA sequence diminution or elimination
Concluding remarks

SUMMARY

In the past there has been a tendency to dwell on aspects of chromosomes which stress constancy of structure, number and content; even to the extent of dismissing exceptions as 'aberrations' or 'oddities'. It is now becoming clear that genomes can be quite plastic, and that structural changes to chromosomes are an important and often necessary part of normal differentiation and development. Elimination of whole chromosome sets or defined portions of genomes is not uncommon and selective gene amplification has been shown to be part of normal development in both protozoa and higher organisms. Chromosomal rearrangements are now a well-documented feature of normal development of, for example, B- and T-lymphocytes and trypanosomes. Transposable elements, whose mobility may not be part of normal developmental processes, can have marked effects on development if their transposition takes them to developmentally important genes. This article reviews some of the structural changes that occur during normal development, and discusses some of the consequences for development when the mechanisms which bring about these rearrangements go wrong.
INTRODUCTION

During the 1940's and '50's there arose the firm idea of constancy of nuclear DNA content and with it came the concept of the C-value (Boivin, Vendrely & Vendrely, 1948; Swift, 1950). At about the same time cytogeneticists were busy trying to establish the concept of the karyotype; again one of constancy but in this case it is of number, shape and size of mitotic chromosomes. These structurally static visions of a constant genotype fitted in with the emerging information about how genes could be regulated and it became possible to consider that differentiation is controlled simply by turning on and off genes. Gurdon's elegant demonstration of totipotency of a single somatic nucleus of a frog also concentrated enquiring minds to the quest for mechanisms of gene regulation. The scene was set for the full-scale investigation of differential gene control, without entertaining the possibility of differential gene alteration.

In following this trail developmental biologists had temporarily pushed aside several long-standing observations on a number of organisms, which quite clearly established the principle that major genomic alterations can accompany differentiation, at least of germline and somatic cells. In the nineteenth century Zur Strassen and Bovieri had described chromosome elimination from the somatic cells of Ascaris, and since that time a number of similar phenomena have been found in insects and crustaceans (see review by Beams & Kessel, 1974).

In the last decade interest in the role that genome rearrangements might play in normal and abnormal development has been reawakened by intense molecular and cytogenetic studies on a number of disparate biological systems. I will consider some of the characteristics of rearrangements which are associated with developmental changes, and, as far as it is known, the effects that these rearrangements have on the cell or organism.

WHAT IS MEANT BY GENOME CHANGE AND DEVELOPMENTAL CHANGE?

First we should be clear what is meant by a genome change or a developmental change. I have been deliberately broad in what I have considered to be a developmental change, but I have classified them into two groups. The first group is a developmental switch that happens at a definable stage and as an essential part of the normal development of an organism. Amplification of the ribosomal genes during oogenesis would be grouped in this category, as would the mating-type switch in yeast. The second group of developmental changes are those that are not predictable within the normal developmental programme of the organism, but nevertheless they manifest themselves as developmental effects on individual cells, and ultimately on the organism as a whole. The chromosomal changes associated with specific forms of cancer would fall into this group.

Then we should be clear about the types of changes that can occur to the genome. Again I have chosen to make two basic groups, based on the distinction
Chromosome rearrangements

<table>
<thead>
<tr>
<th>Table 1. Types of changes associated with recognizable differentiated states</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rearrangements which do not necessarily involve gain or loss of DNA with:-</td>
</tr>
<tr>
<td>1. Predictable developmental effects</td>
</tr>
<tr>
<td>e.g. Yeast mating-type locus</td>
</tr>
<tr>
<td>Variant surface glycoproteins of trypanosomes</td>
</tr>
<tr>
<td>Immunoglobulin genes of mammals</td>
</tr>
<tr>
<td>2. Unpredictable developmental effects</td>
</tr>
<tr>
<td>e.g. Translocation in tumours</td>
</tr>
<tr>
<td>Rearrangements necessarily involving quantitative change</td>
</tr>
<tr>
<td>1. DNA sequence amplification</td>
</tr>
<tr>
<td>e.g. rDNA in oogenesis</td>
</tr>
<tr>
<td>Macronuclear development in ciliates</td>
</tr>
<tr>
<td>Chorion protein genes in <em>Drosophila</em></td>
</tr>
<tr>
<td>Gene amplification in mammalian cells</td>
</tr>
<tr>
<td>2. DNA sequence diminution</td>
</tr>
<tr>
<td>e.g. Satellite DNA in ciliates, nematodes and polytene chromosomes of insects</td>
</tr>
</tbody>
</table>

between those changes which involve rearrangement of pre-existing genomic segments without significant gain or loss of material, and those alterations which involve significant quantitative changes. With the latter group one runs into the difficult semantic problem of whether the change represents an amplification or a diminution. In some cases it is clear, but in others it is far from clear. It seems to me that often it does not matter what it is called so long as the essential point is not lost; namely that the relative amounts of different sequences within the genome has changed in association with a recognizable developmental switch.

Table 1 lists the changes that I will discuss, in the groups and in the order that they appear. To some readers the absence of transposable elements from the list might be surprising. Their omission is mainly for reasons of space, since it would take a chapter on its own to cover them adequately. It should, however, be noted that the insertion or excision of a mobile element can have dramatic developmental effects, and also that the mobilization of particular elements can occur at defined developmental stages. On the other hand, they have not yet been shown to be instrumental in bringing about a normal developmental change as a consequence of their regular and predicted transposition (see review by Federoff, 1983).

**Rearrangements not involving gain or loss of DNA**

The basic feature in common to this class of rearrangements is that a gene, or a copy of it, is activated by translocation to an active site or is rearranged in a way which brings together its component parts to form a structure that can be expressed. The following three cases are particularly good examples of genomic
rearrangements which are instrumental in bringing about a developmental switch.

**The mating type locus of yeast**

Cells of *Saccharomyces* are of one of two mating types, \(a\) or \(\alpha\), and mating only occurs between an \(a\) cell and an \(\alpha\) cell. One might expect all the individual cells of a clone grown from a single cell to be either \(a\) or \(\alpha\), and thus to be self-infertile. When this happens the yeast is said to be heterothallic, but more often than not it is found to be self-fertile. This means that cells of the opposite mating type must have arisen during clonal growth. These yeasts are called homothallic and their ability to switch between mating types is brought about by a genomic rearrangement.

The alleles which determine the \(a\) and \(\alpha\) mating types are called MAT\(a\) or MAT\(\alpha\), respectively. A homothallic yeast carries a single unexpressed copy of both \(a\) and \(\alpha\) in regions on either side (HML\(_{\text{Left}}\) and HMR\(_{\text{Right}}\)) of the expressed MAT allele (see Fig. 1). HML and HMR can be occupied by complete sequences of either \(a\) or \(\alpha\), but they are normally not expressed at these sites because of the negative regulation of two other genes, call MAR and SIR. The mating type

![Diagram of DNA rearrangements resulting in the mating-type switch in yeast. Chromosome 3 of yeast is represented on the top line, with HML\(_a\) to the left of the centromere and MAT\(a\) and HML\(_a\) to the right of the centromere. In this configuration the yeast cell would be of \(\alpha\) mating type because \(a\) sequences are transcribed at the MAT locus. The first switch would replace the \(a\) sequence at MAT with \(a\) sequences from HMR\(_a\); the second switch would replace \(a\) sequence at MAT with \(a\) sequences from HML\(_a\), and so on. HML and MAT have extensive sequence homology at W, X Z1 and Z2, whereas HMR and MAT share only X and Z1 in common. The signal for transposition of \(a\) or \(\alpha\) sequences is a double-stranded cut made at the \(3'\) end of the Ya or Ya segment in MAT. The numbers at the top represent the sizes, in kilobase pairs, of the various regions. Based on data in Strathern *et al.* (1982) and Nasmyth (1983).](image)
Chromosome rearrangements

which is expressed is determined by the type of sequences present in the MAT locus (reviewed by Nasmyth, 1982).

Interconversion between a and α at MAT can occur every second budding cycle by movement and substitution of an unexpressed a or α from the storage sites in HML or HMR. This transposition is duplicative since the a and α sequences in HML and HMR are retained. This is the cassette mechanism of mating-type switch (Hicks, Strathern & Herskowitz, 1977). The switch is initiated by a double-stranded cut made by an endonuclease at the end of the a or α sequences located in the MAT locus (Strathern et al. 1982). The endonuclease is probably coded for, or regulated by, the HO gene, which determines the homothallic (HO) or heterothallic (ho) status of the yeast (Nasmyth, 1983). Very similar molecular exchanges occur during the mating-type switch in Schizosaccharomyces (Beach, 1983).

The variant surface glycoprotein genes of trypanosomes

The antigenic properties of a trypanosome are determined at any one time by a single surface glycoprotein, called the variant surface glycoprotein (VSG). There are over 100 different VSGs, each of which is coded for by a different gene. The trypanosome avoids elimination by the host’s immune system by switching to a different VSG when an immune response is mounted to the pre-existing VSG.

By using complementary DNAs (cDNAs) made from messenger RNAs (mRNAs) coding for VSGs, it was established that a single copy of each of the VSG genes is always present in the genome of the trypanosome (Hoeijimakers et al. 1980). These are the basic copies (BC; see Fig. 2). Additionally it was found that the expressed VSG gene was often present in two copies – the extra copy being in a different chromosomal environment. This is the expression linked copy (ELC). This suggested that a BC had been moved to a different site before it could be expressed; an idea confirmed by the fact that the BCs often lack a 35 nucleotide sequence which is found at the 5' end of the mature VSG mRNA. It turns out that this short sequence is present as a small exon in the DNA which flanks the ELC, and it is from this copy that it is transcribed and spliced to the main portion of the VSG mRNA (Michels, Bernard, Van der Ploeg & Borst, 1983; Liu, Van der Ploeg, Rijsewijk & Borst, 1983). At the 5' end of the ELC gene are found variable portions – between 0.4 and 2.0 kb – of sequences also present 5' to the BC gene (Van der Ploeg, Liu & Borst, 1982). At the 3' end of the ELC gene, however, the breakpoint is usually found to be at a variable position within, but close to, the 3' end of the BC gene (Borst & Cross, 1982). Thus activation may also involve the addition of essential 3' sequences to the BC gene.

Further outside the 3' end of the ELC there are two peculiarities. There is a stretch devoid of any recognition sites for restriction enzymes, the so-called barren region, which is followed by a region of an apparent clustering of recognition


Fig. 2. Mapping expression linked copies and the basic copy of variant surface glycoprotein genes in *Trypanosoma brucei*. At the top is a restriction map of the basic copy of the VSG gene 118. In each independently isolated clone of trypanosome expressing the 118 VSG an additional copy can be detected – the ELC. Four such examples are shown. Comparison of the maps show that each ELC and the BC have in common the coding region of the gene and a small region 5' (to the left) of the coding region. Apart from this ELCs and BC share no homology. The ELCs do have homology. They share the same map 5' to the coding region, although there is variation in the point of recombination in the different variants. They also share the common property of a barren region 3' to the coding region which ends at a telomere. The variation in the map immediately 3' to the coding region shows that these variants must have arisen by independent rearrangements. Based mainly on the data of Michels *et al.* (1983).

sites of many restriction enzymes. Since the latter can also be degraded by the exonuclease Bal 31 it has been suggested to be the natural end to the chromosome (De Lange & Borst, 1982).

The length of the barren region varies between different ELCs (Bernards *et al.* 1981; Van der Ploeg, Bernards, Rijsewijk & Borst, 1982; Michels *et al.* 1982) as well as in different clonal isolates of trypanosomes making the same VSG. This variation is probably due to the fact that chromosome ends appear to increase in length during serial growth. The telomere grows steadily at an estimated rate of between six (Van der Ploeg *et al.* 1984) and about ten (Bernards, Michels, Lincke & Borst, 1983) nucleotides per cell generation, with abrupt and large decreases in size being induced by *in vitro* heat shocks. Since the telomere is a series of tandem repeats of the sequence 5'-CCCTAA-3' (Blackburn & Chalonner, 1984; Van der Ploeg *et al.* 1984), it suggests that the telomere may grow by one repeat per cell generation. Continual changes in the length of telomeric sequences may be a common property of the ends of linear chromosomes, and would explain the length heterogeneity of ribosomal DNA in *Tetrahymena*
Chromosome rearrangements

(Blackburn & Gall, 1978); Physarum (Johnson, 1980) and Dictyostelium (Emery & Weiner, 1981).

Not all VSG genes are activated by duplication and movement to an expression site. Those that are not appear to be always next to a chromosome end (e.g. Williams, Young & Majiwa, 1982; Borst & Cross, 1982). Activation may still involve exchange between the chromosome carrying the pre-existing ELC and the telomERICally located BC of the VSG gene which is to be activated (Borst et al. 1983).

Whatever the precise mechanisms are that bring about the various rearrangements to the VSG genes, and whatever the functional requirements for the rearrangements might be, it is clear that here is a case in which defined rearrangements to the DNA which flanks a VSG gene bring about the expression of that VSG gene. As far as is known, most of the variation of DNA sequence between the various VSG genes already exists in the BCs, and is not generated during the process of rearrangement. The genomic rearrangement is therefore a developmental switch between existing variability, which contrasts with the creation of diversity as an integral part of rearrangement in the immunoglobulin genes.

Immunoglobulin genes of mammals

There are three separate families of genes that participate in expression of immunoglobulins (Ig); in humans the heavy chain genes are on chromosome 14, and the kappa and lambda light chain genes are on chromosomes 2 and 22 respectively. The structures of Ig genes are different in B-lymphocytes from that which is present in all non-immunoglobulin producing somatic cells or germline cells. This is because the Ig genes are rearranged as an integral part of B-cell differentiation, to create the diversity of specificity of antibodies.

To illustrate the principles of Ig gene rearrangement we will consider only the heavy chain genes (Fig. 3). The first stage of rearrangement brings together one of the several hundred variable genes (VH), one of the ‘diversity’ segments (DH) and one of the junction segments (JH), see review by Marcu (1982). Each VH gene has a promoter, but they all remain inactive unless juxtaposed to a CH gene by fusion with a DH and a JH segment (Mather & Perry, 1981). Between the JH segments and the class-switching sequence (S) is an enhancer (E) of transcription (Sakano et al. 1980), so the activating effect of the rearrangement is to bring together a promoter and an enhancer.

The joining of VH, DH and JH genes is mediated by short sequences which occur in opposite orientation on either side of each of the ‘internal’ DH segments, as well as in opposite orientations at the 3’ side of the VH gene and the 5’ side of the JH genes (e.g. Sakano, Kurosawa, Weigert & Tonegawa, 1981). Not all rearrangements result in productive expression of an immunoglobulin heavy chain, but if successful, and if followed by a similarly successful rearrangement in either kappa or lambda light chain genes, the pre-B cell will begin to synthesize IgM as a cell surface receptor.
Fig. 3. DNA rearrangements in immunoglobulin heavy chain genes. At the top is a diagram of an immunoglobulin molecule with its variable, hypervariable junction and constant regions. These are coded for by the variable genes (VH), diversity (DH) and junction (JH) segments and constant genes (CH). There are multiple copies of each of these components and it is the joining of one each of the VH, DH and JH segments taken at random which generates the wide potential variation in the specificity of the IgM molecule. VH genes each carry a promoter (p), but these are only active (P) when rearrangement brings them next to the enhancer sequence (E) between the 3' JH segment and the switch sites (S) of the C gene. After the VDJ rearrangement transcription from P results in an RNA which can be processed to either IgM or IgD. Subsequently, the class of Ig synthesized by the mature B cell is established by deletions within the CH gene cluster. In the example shown, rearrangement occurs between the switch site of Cλ and Cy, with deletion of the intervening chromosome segment. Transcription from P results in IgG1 mRNA. Based on Marcu (1982) and references cited therein.

IgM class of immunoglobulins is specified by the heavy chain gene C. This is the first CH gene downstream of the rearranged V-D-J complex. Transcription from the VH promoter results in an RNA which can be alternatively spliced into either membrane-bound or secreted forms of IgM or IgD (Moore et al. 1981; Uaki et al. 1981).

A further rearrangement within the Ig heavy chain gene family accompanies maturation of the B-cell and is responsible for what is known as the class switch. The B cell continues to produce and secrete IgM or changes to one of the alternative classes of Ig – IgA, IgG or IgE – while still retaining the specificity for the antigen. This rearrangement involves the deletion of sequences between the switch site 3' to JH and the switch site 5' to the selected CH gene (e.g.
Chromosome rearrangements

Burrows, Beck-Engeser & Wabl, 1983). The switch sites contain recognition sequences composed of blocks of tandemly repeated short sequences (Marcu, Lang, Stanton & Harris, 1982; Nikaido, Nakai & Honjo, 1981). Each C_H gene has its own switch site 5' to its coding sequences (Davis, Kim & Hood, 1980; Tyler & Adams, 1980; Kataoka, Miyata & Honjo, 1981). The continued synthesis of secreted IgM by a mature B-cell is ensured by the deletion of a segment 5' to the C_H gene, which includes the entire switch-site-recognition sequence (Cory, Adams & Kemp, 1980).

Thus DNA rearrangements are central to B-cell differentiation. They mediate the fusion of a V_H gene with one each of the D_H and J_H segments to establish the unique specificity of the antibody. Subsequently they are instrumental in bringing about the class switch. Rearrangements which increase the potential for antibody diversity, also occur in the light chain genes although there are no light chain D segments and the V_L or V_H genes are joined directly with one of their respective J genes. Apart from establishing antibody diversity and heavy chain class switching, rearrangements probably have two additional functions. One of these is to increase the potential for diversity over and above that already present in the genome by introducing 'flexibility' into the precision with which the V-D-J junctions are made. Very small differences in the exact points of recombination can affect whether a particular codon is present or changed, and so alter the final amino acid composition of the hypervariable region (e.g. Gough & Bernard, 1981).

The other additional function may be to ensure that only one light chain and one heavy chain gene is ever active in any one B-cell. This is the phenomenon of allelic exclusion. It has also been suggested that DNA rearrangements might determine which class of light chain is expressed, or, in the very early stage of lymphocyte development, whether differentiation is to a B-cell or a T-cell (Coleclough, 1983).

Chromosomal rearrangements associated with the tumour phenotype

It may well be that there are a number of developmentally important rearrangements occurring in somatic cells, but that they go undetected because we do not have informative probes. There are, however, a class of DNA rearrangements, which by their very nature, are self-selective and relatively easily detected. These are rearrangements which result in the development of the tumour phenotype. The clearest and consistent correlations between specific human tumours and the chromosomal rearrangements they carry are listed in Table 2, but there are many others which are listed and discussed elsewhere (see for example Human Gene Mapping 7, 1983; Mitelman, 1983; Honey & Shows, 1983; Forman & Rowley, 1982; and for the mouse, Klein, 1981).

One of the clearest associations is in Burkitt's lymphoma, where translocations between chromosome 8 and chromosomes 2, 14 or 22 are often found. The significance of these particular chromosomes lies in the fact that the cellular
Table 2. *Human tumours associated with specific chromosome rearrangements*

<table>
<thead>
<tr>
<th>Tumour</th>
<th>Type of rearrangement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carcinoma of:</td>
<td></td>
</tr>
<tr>
<td>kidney</td>
<td>t(3;8)(p1;q24)</td>
</tr>
<tr>
<td>lung</td>
<td>del(3)(p14-p23)</td>
</tr>
<tr>
<td>ovary</td>
<td>t(6;14)(q21;q24)</td>
</tr>
<tr>
<td>Leukaemia</td>
<td></td>
</tr>
<tr>
<td>Acute prolymphocytic</td>
<td>t(6;12)(q15;p13)</td>
</tr>
<tr>
<td>Acute promyelocytic</td>
<td>t(15;17)(q26;q21)</td>
</tr>
<tr>
<td>Chronic myelocytic</td>
<td>t(9;22)(q34;q11)</td>
</tr>
<tr>
<td>Lymphoma</td>
<td></td>
</tr>
<tr>
<td>Burkitt's</td>
<td>t(2;8)(q12;q24)</td>
</tr>
<tr>
<td></td>
<td>t(8;14)(q24;q32)</td>
</tr>
<tr>
<td></td>
<td>t(8;22)(q24;q11)</td>
</tr>
<tr>
<td>Retinoblastoma-1 (Rb-1)</td>
<td>del(13)(q14)</td>
</tr>
<tr>
<td>Wilms' tumour</td>
<td>del(11)(p13)</td>
</tr>
</tbody>
</table>

* Compiled from data listed in Human Gene Mapping 7 (1983) and Honey & Shows (1983).

Oncogene *c-myc* is located at the observed breakpoint on chromosome 8, and one or other of the immunoglobulin gene clusters is localized at the breakpoints on chromosomes 2 (κ-light chain), 14 (heavy chain) and 22 (λ-light chain). It is therefore tempting to think that the rearrangement in some way abnormally activates the *c-myc* gene, thereby contributing to the establishment tumourigenicity. There is, however, precious little evidence to show enhanced expression of *c-myc*, or structural mutations that result from the translocation (see review by Rabbitts, 1983). It may be that the change in control is rather subtle, such as the loss of correct cell cycle regulation of *c-myc* (Taub et al. 1984).

I should stress that I do not consider this type of rearrangement to be of any significance to the normal development of the organism. Rather it is a by-product which is a mistake of the normal rearrangement processes going on during B-cell maturation, but which nevertheless results in a developmental change in the B-lymphocyte. But rearrangements which lead to tumours of different kinds are found in other somatic cells. Are these also mistakes which are the by-products of as yet unknown developmentally important rearrangements?

Another kind of somatic chromosomal change gives rise to the malignant phenotype of retinoblastoma (Rb-1). The gene, which is responsible for the inherited form of the disease, has an autosomally dominant pattern of inheritance, and is located in band q14 of chromosome 13. Using DNA restriction enzyme site polymorphisms to define and follow the two homologues of chromosome 13, Cavenee et al. (1983) were able to show that the Rb-1 phenotype was only expressed when the region at 13q14 had become homozygous. They were further able to show that homozygosity resulted either from
Chromosome rearrangements

mitotic non-disjunction or from mitotic recombination. Thus Rb-1 is, in fact, a recessive mutation at the cellular level when in a fully heterozygous genotype. It appears to be inherited as an autosomal dominant mutation because the establishment of homozygosity at 13q14 is so frequent in retinoblasts that it is inevitable that one homozygous cell will be produced in both eyes during the life of the person carrying the Rb-1 mutation. This high error rate suggests that some form of chromosome rearrangements may be occurring as part of normal development in retinoblasts.

REARRANGEMENTS INVOLVING QUANTITATIVE CHANGES

Amplification of ribosomal RNA genes during oogenesis

The classic and most investigated example of ribosomal DNA (rDNA) amplification is during oogenesis in *Xenopus laevis*. It proceeds in two distinct phases. The first occurs about a month after fertilization as the primordial germ cells begin to proliferate as gonial cells. The second phase starts after about three months of development, but only in female larvae as the gonial cells enter meiosis to differentiate into oocytes. In the first phase rDNA is amplified about 40-fold (e.g. Gall & Pardue, 1969), but it is not at all clear what role this amplification step serves; especially in view of the fact that the amplified copies are promptly lost during meiosis in spermatocytes. The second amplification boosts the oocyte rDNA content to 2500 times the normal haploid value (Gall, 1969). This second amplification is presumably an essential developmental response to the special demands of the oocyte for large amounts of rRNA (and ultimately proteins) to supply the cleaving embryo after fertilization. Ribosomal DNA amplification has been recently and comprehensively reviewed by Bird (1980) and Macgregor (1982).

Amplification during both premeiotic and meiotic phases results in the additional copies of rDNA being in the form of extrachromosomal circles (Bird, 1978; Miller, 1964). These circles have contour lengths which are simple multiples of the rDNA repeat length (Hourcade, Dressler & Wolfson, 1973). The identification of replication intermediates in both primordial (Bird, 1978) and ovarian (Gilbert & Dressler, 1968; Rochaix, Bird & Bakken, 1974) extrachromosomal rDNA shows that they are capable of autonomous replication and do not need to be continuously excised from the chromosome.

Ribosomal DNA is present in chromosomes as tandem arrays, but with a heterogeneous distribution of lengths of repeat units, due to variation in the length of the non-transcribed spacer. The heterogeneity of repeat length even in whole ovary extrachromosomal rDNA is much less (Wellauer, Reeder, Dawid & Brown, 1976) and only a very small subset of possible repeat length variants is amplified within a single oocyte (Bird, 1978). However, different repeat length variants are amplified in different oocytes. This means that amplification draws
on only a few copies of rDNA within a single oocyte, but that there do not appear to be any special copies set aside for amplification.

Amplification of rDNA is not confined to *Xenopus*, but is widespread in the amphibia and is well documented in a number of insects (see Bird, 1980; Macgregor, 1982). In all cases, the process would seem to serve the developmental requirement for the accumulation of rRNA and protein prior to fertilization.

**Genomic happenings in ciliates**

Unicellar ciliates go through complicated sets of nuclear changes following conjugation. Although the details differ between the different species, the end result is always to re-establish the presence of a macronucleus and a micronucleus within the cell. From our point of view the micronucleus, which is transcriptionally inactive during vegetative growth, is the 'germline' nucleus, and the transcriptionally active macronucleus is the equivalent of the 'somatic' nucleus. The formation of the macronucleus following fertilization is the development change.

**Ribosomal DNA and genome rearrangements in *Tetrahymena***

The macronucleus of *Tetrahymena* carries only a single copy of the rDNA unit, which is integrated into chromosomal DNA (Yao & Gall, 1977). During macronuclear development this single copy becomes excised and amplified several hundred fold as a small extrachromosomal linear molecule. During this process a 3 kb segment of DNA, which is immediately 3' to the micronuclear rDNA, is completely lost from the macronucleus (Yao, Blackburn & Gall, 1979). It is not a simple excision because each of the amplified rDNA molecules carries two copies of the rDNA unit joined together head-to-head to form a huge inverted repeat (Engberg, Anderson, Leick & Collins, 1976; Karrer & Gall, 1976). Also the ends of the palindrome contain between 20 and 70 tandem copies of the sequence 5'-CCCCAA-3', which seems to define a structure capable of functioning as a telomere in yeast (Szostak & Blackburn, 1982) as well as *Tetrahymena*. In the micronuclear DNA there is only one copy of the C₄A₂ sequence and it is located at the 3' end of the rDNA (King & Yao, 1982). This amplification of rDNA not only produces several hundred copies of the gene, but it also involves loss of some micronuclear DNA and the tandem repetition of the C₄A₂ sequence. Whether the latter are co-opted from blocks of C₄A₂ sequences elsewhere in the genome, or whether they 'grow' in situ is not known (but see the growth of trypanosome telomeres).

Alterations to the rDNA genes have been the best characterized, but it is clear that extensive rearrangements and deletions occur elsewhere in the genome as the macronucleus is formed. Some 10 to 20 per cent of micronuclear DNA sequences are eliminated (Yao & Gorovsky, 1974), resulting from the deletion of defined segments from specific sites in the macronucleus. Sequences on either side of the deleted segments become linked together to form contiguous
Chromosome rearrangements

stretches of macronuclear chromosome (Yao et al. 1984). Comparison of structure of a specific α-tubulin gene in macro- and micronuclear DNA of *Tetrahymena* shows that several deletions occur in the flanking sequences during macronuclear differentiation. Those sequences remaining in the macronucleus are joined together in the same relative order that they occur in the micronucleus (Callahan, Shalke & Gorovsky, 1984). Such rearrangements occur widely in the genome, perhaps involving up to 5000 sites. The precise function they serve is unknown at present, but some role in the transcriptional activation of sequences required for vegetative growth would seem a likely candidate.

Gene amplification and diminution in hypotrichous ciliates

Just as with *Tetrahymena*, macronuclear formation in the hypotrichous ciliates (e.g. *Stylonychia*, *Oxytricha*, *Euplotes*) involves loss of sequences, structural rearrangement of DNA and gene amplification, but on a far more bizarre scale (reviewed in Prescott, Murti & Bostock, 1973; Lawn, Heumann, Herrick & Prescott, 1978; Swanton, Heumann & Prescott, 1980). Initially the chromosomes within the developing macronucleus undergo several rounds of polytenization, but, as in the insects (see below), not all DNA sequences are replicated equally. The highly repeated, satellite-like sequences are under replicated; this is the first stage of diminution. At the height of polytenization the chromosomes are transected and the chromatin within each polytene band is enclosed in a vesicle. Selective degradation within the vesicles results in the elimination of any remaining repeated sequences as well as 95% of the single copy sequences. This is the second stage of diminution. Finally, the remaining 5% or so of unique sequences are released as small chromatin fragments from the vesicles and replicated several times to form the genomic content of the mature macronucleus. It will now contain about 1000 copies of each of 24000 different linear DNA molecules which range in size from a few hundred to several thousand base pairs in size. This is the amplification stage.

The effect of cutting up the DNA is to create thousands of mini-chromosomes, each of which will require a stable end. It is not surprising therefore that the linear molecules of the macronucleus contain at their ends inverted blocks of the short sequence 5'-CCCCAAAAA-3' repeated several times (cf. the C₄A₂ sequence of *Tetrahymena rDNA*). This poses the interesting question of where the C₄A₄ blocks come from. Are they pre-existing in the micronuclear chromosomes, perhaps serving as recognition signals for excision and preservation? Do single copies of their C₄A₄ sequence lie on either side of sequences to be preserved from micronuclear DNA, being tandemly duplicated after excision? Or are the ends derived from chromosomal 'stores' of C₄A₄ and attached to the macronuclear fragments after the vesicle stage? Using DNA:DNA hybridization as a test, Boswell, Klobutcher & Prescott (1982) were able to rule out the first possibility, since they could not find the C₄A₄ unit in micronuclear DNA within the flanking regions of the macronuclear sequences. The methods
used could not detect a single copy of the C₄A₄ unit so the second possibility remains to be tested. Large blocks of the sequence do occur in micronuclear DNA, but these are thought to be at the natural telomeres of micronuclear chromosomes (Dawson & Herrick, 1984). They could nevertheless serve as stores of the C₄A₄ sequence.

**Differential gene amplification in polytene tissues of insects**

*Chorion protein genes in Drosophila*

Synthesis of the chorion occurs during a short period of time and requires a very high rate of synthesis of individual chorion proteins. Insects have met this demand in different ways. The silkmoth, for example, carries multiple copies of the genes for chorion proteins (Jones & Kafatos, 1980), whereas *Drosophila* carries only single copies of each of the chorion protein genes, but amplifies them between 15 and 60 fold to meet demand for eggshell production (Spradling & Mahowald, 1980; Spradling, 1981).

In this case amplification seems to be by several additional rounds of replication from a specific replication origin within the chorion protein gene clusters in the already polytenized chromosomes of the follicle cells. Amplification extends outwards from the centre of a region that covers 100 kb; the sequences in the centre being amplified several-fold more than those at the periphery (Spradling, 1981). Evidence that the specificity for amplification lies in the replication origin comes from the fact that the ability to amplify can be transferred by a chromosomal inversion to sequences which normally never amplify and which have absolutely no functional relationship to the chorion protein genes. The chorion protein genes which normally amplify, but which are separated from the amplification origin by the inversion, no longer undergo amplification (Spradling & Mahowald, 1981). It is as though the chromosomal regions which contain the chorion protein genes are programmed to go through four to six additional rounds of polytenization, giving rise to an 'onion-skin' structure of partially replicated bubbles. This notion is supported by the electron microscopic visualization of complex branching structures (Osheim & Miller, 1983).

**DNA ‘puffs’**

The amplification of chorion protein genes in *Drosophila* may simply be a specific example of a ‘DNA puff’; a well-documented phenomenon in late larval instars of sciarid flies, when certain polytene bands go through additional rounds of DNA replication (Rudkin & Corlette, 1957; Crouse & Keyl, 1968). Contemporary cloning experiments have shown that the DNA puffs of a related fly, *Rynchosciara*, do, in fact, represent the amplification of genes which are expressed during the late larval stages (Glover *et al.* 1982). Once again, amplification here would seem to represent a developmentally programmed chromosomal change to meet the demands for high levels of specific gene products.
Chromosome rearrangements

Gene amplification in mammalian cells

In vitro selection

Gene amplification occurs in a number of mammalian cells which have been selected in vitro for high levels of resistance to particular metabolic inhibitors. There have been several excellent reviews of this subject recently (see Cowell, 1982; Schimke, 1982; Hamlin, Milbrandt, Heintz & Azizkhan, 1984; Stark & Wahl, 1984), so I will mention only some features which I believe summarize the field. These are listed in Table 3.

Together the first two points are really just a restatement of the fact that, when undergoing selection for drug resistance, cells can respond in a developmental way by overproducing a gene product. The basis for this developmental change is a change at the chromosomal level. The overproduced protein may be the direct target for the drug (e.g. dihydrofolate reductase in response to Methotrexate), but not necessarily so. Cells which show multiple drug cross-resistances (to vincristine, actinomycin D and driamycin, amongst others) over-produce a common small acidic soluble cytoplasmic protein and a phosphorylated glycoprotein of relative molecular mass 170×10^3, neither of which appears to be a target for these drugs.

The third point in Table 3 is important because it establishes that, at least in the case of CAD and DHFR gene amplifications, the mechanisms which give rise to amplification occur in the absence of the selective agent (Kempe, Swyryd

Table 3. Features of in vitro selected gene amplification in mammalian cells

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Increase in amount of one or more proteins.</td>
</tr>
<tr>
<td>2.</td>
<td>Increase in number of genes coding for overproduced proteins.</td>
</tr>
<tr>
<td>3.</td>
<td>Frequency of spontaneous occurrence within the range 10^-4 to 10^-7.</td>
</tr>
<tr>
<td>4.</td>
<td>Treatments which disrupt DNA synthesis can dramatically enhance the spontaneous rate of occurrence.</td>
</tr>
<tr>
<td>5.</td>
<td>Amplified genes may or may not be regulated normally.</td>
</tr>
<tr>
<td>6.</td>
<td>Amplification is often associated with a chromosome translocation and DNA rearrangements.</td>
</tr>
<tr>
<td>7.</td>
<td>Chromosomally integrated copies of amplified genes are not necessarily located at the wild-type position.</td>
</tr>
<tr>
<td>8.</td>
<td>Amplification is often associated with the presence of abnormal chromosome structures, dm or hsr.</td>
</tr>
<tr>
<td>9.</td>
<td>Large amounts of DNA are co-amplified with the selected gene.</td>
</tr>
</tbody>
</table>
Thus amplification of chromosomal segments is occurring spontaneously in cells; the effect of the selective agent is simply to select for those cells which have already amplified the gene of relevance to the specific selective agent. Nevertheless the spontaneous rate of amplification can be greatly increased by treatments which interfere with DNA synthesis (e.g. Mariani & Schimke, 1984).

The selected gene is usually found to be amplified as part of a much larger unit of DNA; often 500 to 1000 kb in size, but in one case up to 3000 kb. At the cytogenetic level the amplified material can be seen as small acentric chromosome fragments called double minutes (dm) or as large abnormal homogeneously staining regions (hsr) (see Fig. 4). In building up these large structures it is clear that many rearrangements occur, and that sequences which can be juxtaposed in the amplified DNA may have been drawn from non-contiguous regions of wild-type chromosomes. Rearrangement of DNA sequences by translocation has been implicated in the initial stages of amplification of both the DHFR gene (Worton, Duff & Flintoff 1981; Flintoff et al. 1982) and the asparagine synthetase gene.
Chromosome rearrangements

(Andrulis et al. 1983) in Chinese hamster ovary cells. The variable location of the hsr on chromosome 2 in several independently selected MTX-resistant Chinese hamster cells (Biedler, Melera & Spengler, 1980) also implicates chromosomal rearrangements in the amplification process.

Non-selected tumour cells

As a result of studies on the in vitro model systems a good correlation has been established between, at the molecular level, the presence of amplified DNA sequences and, at the cytogenetic level, the presence of either dm or hsr. Of course, simply the presence of the latter cannot be taken as proof of gene amplification, but it is a good indicator. It is therefore interesting that many spontaneously arising tumours contain the abnormal chromosome entities. This suggests that the tumours may contain amplified gene sequences, and that their presence may have played a role in the development of the tumour phenotype. This is especially so if amplification is a way of inducing elevated expression of normal cellular oncogenes, which would in turn result in tumourigenesis (see Bishop, 1983). The correlation between size of hsr and degree of tumourigenicity supports this idea (Gilbert et al. 1983).

In a few specific cases we do not need to assume that the presence of dm or hsr equals gene amplification, since molecular studies have identified amplification of DNA sequences including cellular oncogenes (see Table 3 of Stark & Wahl, 1984). A number of different human neuroblastomas contain amplified copies of a variant, called N-myc, of the c-myc oncogene. The amplified sequences are contained in hsr, but these are at different sites within different tumours (Schwab et al. 1983). In one of these, IMR 32, the hsr is on chromosome 1, making it very large and ideal for sorting in a FACS (Kanda et al. 1983). Cloned DNA probes prepared from the hsr of chromosome 1 of IMR 32 were found to have been derived from sequences normally on the short arm of chromosome 2 (Kohl et al. 1983). Thus amplification of the sequences on chromosome 2 to form the hsr on chromosome 1 involved both rearrangements and translocation.

DNA sequence diminution or elimination

We have already mentioned examples of DNA diminution when considering macronuclear development in the ciliates. There are, however, other clear examples in which the loss of chromosomal material is an integral part of the developmental process. I would consider the elimination of the whole genome during maturation of red blood cells as a necessary step in the development of the concave structure of the mature mammalian red blood cell. Whole chromosome sets are also eliminated in a developmentally programmed way during spermatogenesis in Sciara (Abbott, Hess & Gerbi, 1981) and in the establishment of haploidy in the mite, Metaseiulus (Nelson-Rees, Hoy & Roush, 1980). The elimination of whole mitotic chromosome sets also occurs in a tissue-specific manner in hybrids of barley (Finch, 1983).
This last section brings us back to Bovieri’s first observation on chromosome changes associated with development, namely the elimination of large segments of chromosomes during early development in *Ascaris*. Fig. 5 shows the
Chromosome rearrangements

Chromosomes of each of the blastomeres of a 4-cell embryo of *Parascaris equorum univalens* as they enter mitosis. The progenitor of the germ cells retains its two chromosomes intact, whereas the three presomatic blastomeres contain fragmented chromosomes. The four large heterochromatic blocks, one from each telomere, lie to the periphery of the cells, do not attach to the spindle, and are lost from the daughter nuclei. As in the case of the hypotrichous ciliates, the loss of heterochromatin is paralleled by the loss of almost the entire germline complement of highly repeated tandemly arranged sequences (e.g. Roth, 1979). These sequences of *Ascaris suum* fall into two related families (Roth & Moritz, 1981), both of which have been sequenced (Streeck, Moritz & Beer, 1982). Their properties look very much like other satellite DNAs and give no clue as to why they should be eliminated. The fact that a few copies are retained in somatic cells (Roth & Moritz, 1981) is reminiscent of the C4A4 repeats in ciliates (see above). Could it be that the large blocks of satellite function as hugh telomeres in the germline chromosome, whereas individual copies, or small numbers, of the repeated sequence can function as telomeres for the fragmented somatic chromosomes?

Although less well characterized at the molecular level, similar developmental changes occur in presomatic cells during early development of several species of Copepods (Beermann, 1977). Up to half the chromatin can be eliminated, being derived from both interstitial and telomeric blocks of heterochromatin. Following diminution the somatic chromosomes retain structural continuity, which suggest that Copepods have retained mechanisms for bringing together and rejoining chromosome ends (Beermann & Meyer, 1980).

Heterochromatin and satellite DNA are also under-represented in polytene tissues of *Drosophila* larvae (e.g. Gall, Cohen & Atherton, 1973; Endow & Gall, 1975), and in some species the ribosomal RNA genes are also under-replicated during polytenization (Hennig & Meyer, 1971). The developmental significance of this under-replication of satellite DNA is not at all clear, unless it is related in some way to the formation of the chromocentre, or, perhaps, to the maximization of synthetic capacity by increasing the relative ‘concentration’ of expressed genes within defined constraints of the nucleocytoplasmic ratio.

Fig. 5. Elimination of heterochromatin and satellite DNA from somatic tissues of *Parascaris equorum univalens*. By the four blastomere stage (top) there is clear differentiation between germ-line and pre-somatic blastomeres. The former retains the two large chromosomes intact (middle left) and its DNA contains both main band (density of 1·700 g.cm⁻³) as well as large amounts of two satellite DNAs (densities of 1·692 g.cm⁻³ and 1·697 g.cm⁻³). In pre-somatic blastomeres the chromosomes fragment and the large heterochromatic regions are lost by not attaching to the spindle (middle right). DNA extracted from somatic tissue contains essentially only the main band density component. The micrographs were kindly supplied by Dr. Karl Moritz; see text for details and sources from which this figure was compiled.
CONCLUDING REMARKS

I have described a number of cases in which there is a clear link between a structural alteration of some kind at the level of the genome and a change of some kind in the developmental state of the cell or organism. In many of these there is a strong case for thinking that the chromosomal change is the initiator of the developmental change. In others a causative link has not been established (although it may exist) and the chromosomal change may simply follow as a direct result of the altered development. Taken together they comprise a significant body of evidence to show that programmed chromosomal rearrangement has been maintained throughout evolution as one of several mechanisms for effecting steps in differentiation. They also show that somatic genomic rearrangements are hard to detect without the appropriate molecular probes. Now that just about any piece of DNA can be cloned it will be interesting to see how many other ‘developmental switches’ turn out to be mediated by rearrangements at the DNA level.

REFERENCES


Chromosome rearrangements


