Karyotype Analysis and Chromosome Banding

Wendy A Bickmore, *MRC Human Genetics Unit, Edinburgh, Scotland, UK*

A series of reproducible bands across metaphase chromosomes can be revealed by some treatments. These chromosome bands not only allow the identification of normal and abnormal chromosomes but they also tell us about fundamental aspects of the chromatin structure and compartmentalization of the genome.

Secondary article

Article Contents

Methods of Chromosome Banding

Nearly all methods of chromosome banding rely on harvesting chromosomes in mitosis. This is usually achieved by treating cells with tubulin inhibitors, such as colchicine or demecolcine (colcemid), that depolymerize the mitotic spindle and so arrest the cell at this stage. Excessively long incubations with Colcemid result in overcondensed chromosomes that band poorly and moreover some cell types, especially those from the mouse, eventually escape the Colcemid block and proceed through the cell cycle.

Chromosome banding methods are either based on staining chromosomes with a dye or on assaying for a particular function. The most common methods of dye-based chromosome banding are G- (Giemsa), R- (reverse), C- (centromere) and Q- (quinacrine) banding. Bands that show strong staining are referred to as positive bands; weakly staining bands are negative bands. However the staining patterns are not black and white, different bands stain to different intensities (Francke, 1994). G-positive bands are usually just called G-bands and likewise for R-positive (R-) bands. Positive C-bands contain constitutive heterochromatin. Q-bands are considered equivalent to G-bands.

The most widely used function-based banding method is replication banding and is based on the fact that different bands replicate their DNA at different times during S phase of the cell cycle. Generally, R-band DNA is replicated earlier than G-bands (Dutrillaux *et al*., 1976). G-bands also correspond to the condensed chromomeres of meiotic chromosomes and R-bands to the interchromomeric regions.

Uses of Chromosome Banding

G- and R-banding are the most commonly used techniques for chromosome identification (karyotyping) and for identifying abnormalities of chromosome number, translocations of material from one chromosome to another, and deletions, inversions or amplifications of chromosome segments. This has had an invaluable impact on human genetics and medicine and the power of this approach has been augmented by combining cytogenetics with fluorescence *in situ* hybridization (FISH). The detection of chromosome deletions associated with disorders, very often contiguous gene syndromes, provided some of the first disease gene localizations in humans. Similarly, translocations have been important in pinpointing the location of disease-associated genes and the characteristic translocations associated with some leukaemias is important, not only for understanding the molecular basis of these cancers, but also for their diagnosis and prognosis. One of the best examples of this is the translocation between human chromosomes 9 and 22 – t(9;22)(q34:q11) – or the Philadelphia chromosome diagnostic of chronic myelogenous leukaemia (CML).

Comparisons of chromosome banding patterns can confirm evolutionary relationships between species and also reveal changes in karyotype that may have been important in speciation. The banding patterns of human, gorilla and chimpanzee chromosomes are almost identical, though human chromosome 2 is the result of a fusion between two great ape chromosomes. There are also extensive similarities between human chromosome bands and those of lower primates.

Number and Size of Bands

Idealized diagrams (ideograms) of G-banded chromosomes are published as standard reference points for chromosome banding. The G-bands are usually portrayed in black and the R-bands in white. Bands are numbered consecutively away from the centromere on both the short (p) and long (q) arms (**Figure 1**). The total number of bands or 'resolution' in the human karyotype depends on how condensed the chromosomes are, and at what stage of mitosis they are in. A 350-band resolution corresponds to chromosomes late in metaphase. High-resolution ideograms (approximately 1250–2000 bands) have also been produced for human chromosomes in mid-prophase

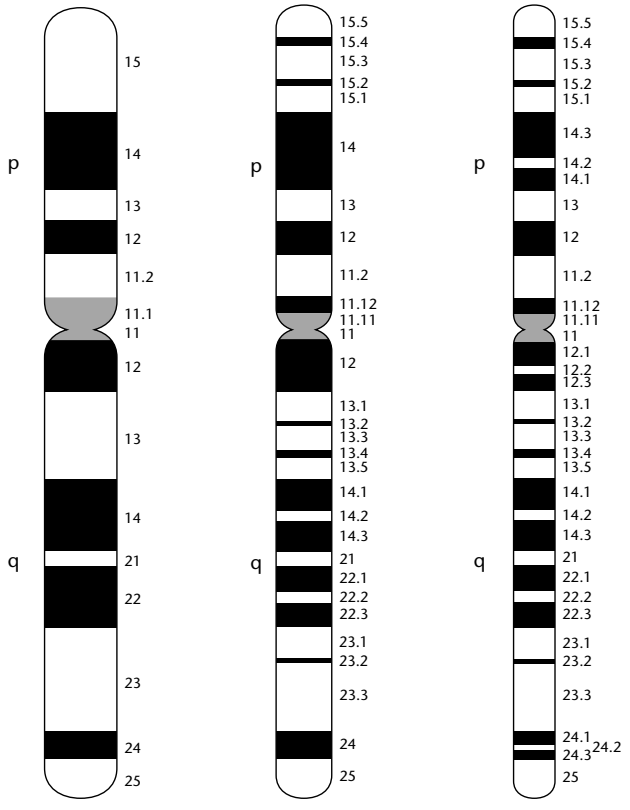

Figure 1 G-band ideograms of human chromosome 11 at (from left to right) 350, 550 and 850 band resolution.

(Yunis, 1981). When a low-resolution band is subdivided, the number of each subband is placed behind a decimal point following the first band designation. For example the most distal low-resolution band on the short arm of human chromosome 11 (11p15) can be subdivided into bands 11p15.1, 11p15.2, 11p15.3, 11p15.4 and 11p15.5 at higher resolution (**Figure 1**). A 2000-band resolution chromosome band may contain ~ 1.5 Mb of DNA, while a 300-band resolution band will contain 7–10 Mb of DNA. A skilled cytogeneticist may be able to spot a deletion of 5–10 Mb of DNA depending on its location, but at a molecular level the human genome probably comprises > 2000 'bands'.

Basis for G-/R-banding

G-banding involves staining protease-treated chromosomes with Giemsa dye and is thought to result from interactions of both DNA and protein with the thiazine and eosin components of the stain. The most common R-banding method involves heat denaturing chromosomes in hot acidic saline followed by Giemsa staining. This method is thought to preferentially denature AT-rich DNA and to stain the under-denatured GC-rich regions. T-banding identifies a subset of R-bands – the most intensely staining ones – by employing either a more severe heat treatment than R-banding. It is thought to identify the GC-richest R-bands, of which approximately half occur at telomeres in the human genome, hence the name.

The need to combine chromosome banding with fluorescence *in situ* hybridization has meant that banding techniques using fluorescent dyes has become more popular. Q-banding involves staining with quinacrine which reacts specifically with certain bases. Quinacrine intercalates into chromosomal DNA irrespective of sequence, but fluoresces brighter in regions of AT-rich DNA. There are a number of other molecules whose fluorescence is influenced by the base composition of the DNA to which they are bound. In addition to quinacrine, other commonly used fluorochromes with a specificity for AT-rich DNA include Hoechst 33258, DAPI (4′-6-diamidino-2-phenylindole) and daunomycin. The fluorescence of Hoechst and DAPI is not quenched by guanine and so they give less distinct bands than those produced by quinacrine; however, daunomycin fluorescence is greatly quenched by DNA with a GC content of $> 32\%$. DAPI staining has the advantage that it is very resistant to fading and that its excitation and emission spectra are compatible with reporter molecules and filters commonly used in FISH (**Figure 2**).

Fluorochromes with a preference for GC-rich DNA include chromomycin and 7-amino actinomycin D. These dyes give an R-band-like pattern.

Bands Reflect the Domain Organization of the Genome

Techniques used to reveal chromosome bands enhance an inherent pattern of chromosome organization. A chromosome band is a manifestation of a chromatin domain with functional and structural characteristics that are homogeneous and distinctive over a long enough stretch to be seen down the microscope. G- and R-banding reflect differences in chromatin structure and base composition between different regions of the genome. Fluorochrome banding also reflects variation in base composition along chromosome length.

A rather different sort of banding is that seen after staining of polytenized (interphase) chromosomes from some tissues of Dipteran insects. The basis for this banding is not well understood but arises through the alignment, in register, of many thousands of copies of the chromosome.

Patterns in DNA sequence

Banding patterns can arise as a consequence of differences in the DNA sequence along chromosomes. R- and

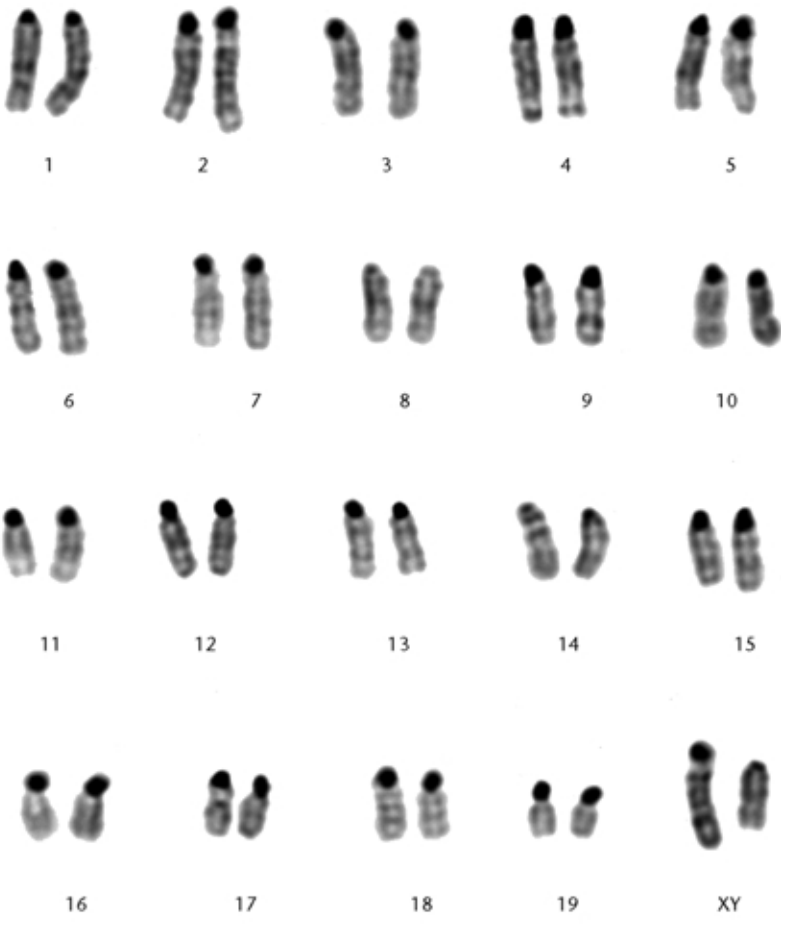

Figure 2 DAPI-stained chromosomes from mouse embryonic stem cells (male). The staining approximates to a G-band pattern, with the C- and G-bands stained more intensely by DAPI than the R-bands. The constitutive heterochromatin around the centromeres of these acrocentric chromosomes stains very intensely with DAPI.

G-banding patterns are revealed on human chromosomes by FISH with Alu and LI interspersed repeats, respectively (Korenberg and Rykowski, 1988). A similar distribution has also been reported for SINEs (short interspersed elements) and LINEs (long interspersed elements) on mouse chromosomes. However, molecular studies of the human genome including sequencing show that both SINE and LINE repeats can be found in close proximity to each other. To reconcile these differences it is suggested that the SINEs located in R-bands are those that have retroposed most recently and hence are closest in sequence to the progenitor copy. These will therefore hybridize better to the SINE probes that are based on the consensus repeat sequence than to those that are more diverged, and hence will produce stronger FISH signals than their G-band counterparts.

Fractions of DNA of differing base composition (isochores) produce banding patterns on human metaphase chromosomes with the most GC-rich fractions highlighting T-bands (Saccone *et al*., 1993). Molecular analyses on a finer scale reveal that, although there is a general tendency for G-bands to be quite AT-rich, R-bands can contain both AT-rich isochores and GC-rich isochores.

A fraction of DNA (CpG islands) that contains the 5′ ends of approximately 50% of mammalian genes also produces a T- and R-banding pattern on chromosomes telling us that genes are not uniformly distributed along the chromosome's length. Most islands appear to reside in T- and R-bands (Craig and Bickmore, 1994) (**Figure 3**). A similar clustering of CpG islands into early-replicating R-band compartments is seen in rodent genomes (Cross *et al*., 1997). Chromosomes of a nonmammalian vertebrate (the chicken) show a striking concentration of CpG islands in distinctive parts of the karyotype – the microchromosomes – rather than on the macrochromosomes.

Examination of genome databases suggests that CpG island clustering can be extrapolated to gene density itself.

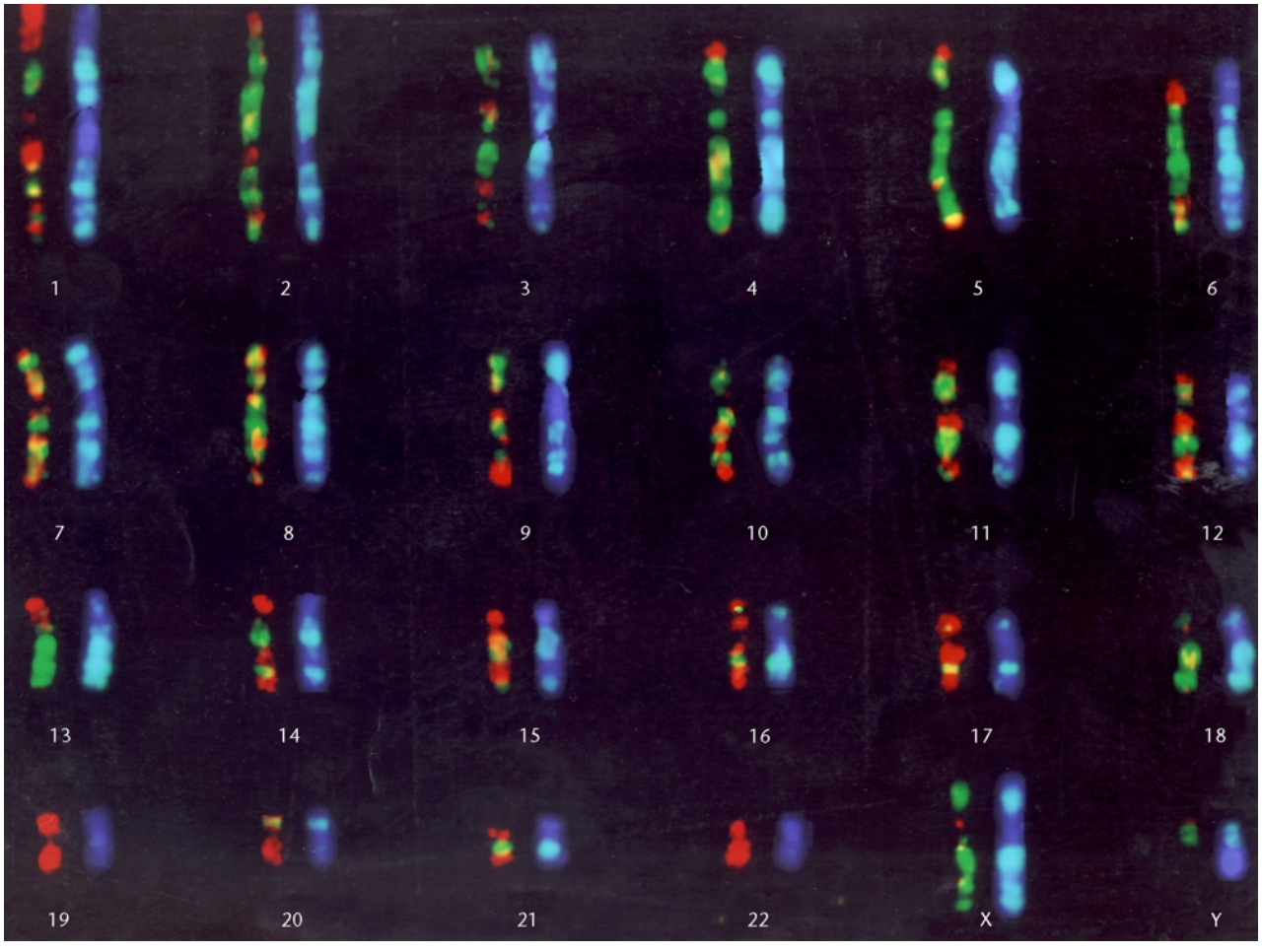

Figure 3 The location of CpG islands in the human karyotype. Left-hand chromosome of each set: CpG islands (red) painted by FISH onto human metaphase chromosomes that have been replication banded so that the G (late replicating)-bands are green (sites of BrdU incorporation). Right-hand chromosome of each set DAPI-stained (blue) chromosomes with a late-replication banding pattern (green).

Hence the highest densities of genes in our genomes are located in the T- and R-bands. This explains why human chromosomes with a high G-band content, e.g. 13, 18 and 21, are seen as viable trisomies in the population, whereas trisomies for small but T-/R-band-rich chromosomes (e.g. chromosome 22) are lost early in embryonic development. Whole genome sequencing has confirmed the low gene density of human chromosome 21 compared to that of 22 (Dunham *et al*., 1999; The Chromosome 21 Mapping and Sequencing Consortium, 2000).

There has been a conservation of chromosome organization at the chromosome band level in terms of relative gene density, time of DNA replication and banding characteristics over 100 million years of mammalian chromosome evolution. This suggests either the influence of a strong selective pressure to maintain these characteristics together in particular chromosome bands or the action of common mechanisms linking the properties of gene density, replication time and band type.

Patterns in DNA replication

Different regions of the genome replicate at different times during S phase. The relationship between timing of replication and chromosome banding is usually studied by incorporating pulses of the thymidine analogue 5-bromo-2′-deoxyuridine (BrdU) into cells during defined stages of S phase and then examining chromosomes in the subsequent metaphase. Sites of BrdU incorporation can be detected with antibodies that detect the presence of BrdU in denatured DNA (**Figure 3**). T-bands replicate on average earlier than ordinary R-bands, and DNA in G-bands is replicated even later (Dutrillaux *et al*., 1976).

Aspects of the primary DNA sequence or chromatin structure in different types of band could influence their replication time. Also differences in replication time could influence some characteristics of chromosome bands, e.g. base composition or chromatin structure. Sites of transcription at the G_1/S boundary may seed the assembly of the first replication factories in early S phase and hence the most transcriptionally active regions of the genome, and the regions with the highest concentrations of genes, would tend to be the ones to be replicated first.

Patterns in chromatin structure

Several banding techniques, especially G-banding, suggest that there are both qualitative and quantitative variations in the interaction of DNA and proteins along the length of metaphase chromosomes. The chromatin of active genes is generally considered to be more accessible to nuclease attack than is inactive chromatin. Consistent with this nucleases preferentially digest R-bands and T-bands of intact mitotic chromosomes, with G-bands and C-bands refractive to digestion. The extent of chromatin packaging in the interphase nucleus also differs between chromosome bands. C-band positive heterochromatin remains visibly condensed through interphase. FISH has shown that over the 150 kb to 1 Mb size range G-band chromatin is more tightly packaged than that of R-bands.

The N-terminal tails of core histones H3 and H4 can be modified by acetylation of lysine residues. A consequence of this acetylation may be to facilitate access of proteins such as transcription factors to the DNA. Histones can be deacetylated or acetylated throughout the cell cycle by nuclear acetyltransferase and deacetylase enzymes. Immunofluorescence of mammalian metaphase chromosomes with antibodies raised against each of the acetylated forms of H4 has shown that the differently modified forms are found preferentially in different regions of the chromosome. Mono-acetylated (Lys16) H4 is found throughout euchromatin, whereas acetylation at Lys8 and Lys12 occurs mainly in R-bands. Acetylation at Lys5 is found in the most highly modified (tri- and tetra-acetylated) forms of histone H4. Antibodies to this H4 isoform produce a good banding pattern on metaphase chromosomes, especially from cells that have been briefly exposed to histone deacetylase inhibitors. Bright immunofluorescence is seen over T-/R-band regions of the karyotype and only faint staining is seen in G-bands (Jeppesen and Turner, 1993). Hence the distribution of histone acetylation on mammalian metaphase chromosomes mirrors that of genes.

The radial loop/scaffold model of chromosome structure proposes that higher order chromosome packaging arises through the arrangement of the 30-nm chromatin fibre into loops tethered to a proteinaceous chromosome scaffold that runs centrally down the length of the chromosome. The two major protein components of the mitotic scaffold are topoisomerase IIα Sc I (170 kDa) and Sc II (135 kDa). Sc II is a member of the structural maintenance of chromosomes (SMC) family of proteins.

Immunofluorescence analysis shows that topoisomerase IIα is not uniform along the chromosome length. It is particularly concentrated in regions of centromeric heterochromatin and there is stronger staining of G-bands than R-bands. This might indicate that chromosome loop anchoring sites to the scaffold are the least frequent in R-bands. The pattern of immunofluorescence with anti-ScII antibodies on vertebrate mitotic chromosomes is very similar to that of topoisomerase II – the chromosome axes are lit up and centromeres are particularly brightly stained.

Specific sites along the DNA associate with the chromosome scaffold. Such sequences are generally referred to as SARs (scaffold attachment regions). They contain oligo(dA)-oligo(dT) tracts and fluoresce brightly with daunomycin. The path of aligned SARs along the scaffold at the core of metaphase chromosomes, as defined by daunomycin staining is G-band like and colocalizes with topoisomerase II immunofluorescence (Saitoh and Laemmli, 1994). This might mean that G-bands have smaller loops, and hence more frequent scaffold attachments than R-bands and chromosome painting with SARs suggests that there are indeed more of these in G-bands than in R-bands (Craig *et al*., 1997). However, these staining patterns could also result from differences in the path of the SARs between different types of band, with the chromosome scaffold being more tightly coiled within G-bands and straighter within R-bands.

Variations in the density of meiotic chiasmata along mammalian chromosomes are apparent when physical and genetic maps are compared. T-bands have the highest rates of exchange, followed by 'mundane' R-bands, then G-bands. Heterochromatin (C-bands) shows the lowest rates of recombination. R-bands are both the sites of synaptic initiation and the preferred sites of crossing-over in mammals and other vertebrates.

Evolution of chromosome bands

Whereas Q-, G- and R-banding patterns have only been observed in some eukaryotes, replication banding is almost universal among living organisms possessing chromosomes large enough to see by microscopy, suggesting that it is a fundamental consequence of, or requirement for, the compartmentalization of complex genomes.

Chromosomes from most mammals and birds can be G- and R-banded. In addition, most reptilian chromosomes band with G- and R-banding techniques to some extent. With amphibia, fish and plants, some species band whereas others do not. The lowest vertebrates with reported good G-banding are the bony fish.

Evolutionary analysis of chromosome banding patterns suggests that the first cytogenetically detectable compartmentalization that arose in the genomes of eukaryotes was the temporal control of replication and differences in chromatin packaging and the segregation of some chromosomal domains into heterochromatin. Ability to be G-banded (and we will assume here that this is a reflection of differences in chromatin structure on mitotic chromosomes) followed later. Fluorochrome banding seems to have appeared on the scene last of all.

Extreme Chromosome Bands

All groups of eukaryotes that have chromosomes large enough to be visualized contain a proportion of heterochromatin, visualized by C-banding, and even simple eukaryotes with microscopic chromosomes have heterochromatin at a molecular level. It seems unlikely that there is a completely strict dividing line between heterochromatin and euchromatin – we might consider the very gene-rich T-bands and gene-free constitutive heterochromatin as the opposite extremes of a continuum of chromatin flavours.

In humans, the main C-bands are on the long arm of the Y-chromosome and close to the centromeres of chromosomes 1, 9 and 16 (pericentric). The size of these C-bands differs between different individuals. Smaller C-bands are found at the centromere of each chromosome and on the p arms of the five acrocentric chromosomes (13, 14, 15, 21, 22). In the mouse, the main blocks of visible heterochromatin are found close to the centromere of each chromosome (**Figure 2**). The amount of heterochromatin at these sites also varies between different strains of *Mus musculus*. In *Drosophila* the main sites of constitutive heterochromatin are found at the chromocentre, telomeres and on the fourth/Y-chromosome. At the molecular level all of these sites are characterized by the presence of middle and highly repetitive sequences, often in tandem arrays of satellites. Classical satellites (I, II and III) locate to the pericentromeric C-bands of human chromosomes 1, 9, 16 and the long arm of the Y; α-satellite is located at the centromeres. It is the minor satellite repeat that is found at mouse centromeres, with major satellite being found beyond this, between the centromere and telomere. In *Drosophila*, satellite DNA at the chromocentre is interspersed with stretches of more complex DNAs (both unique and moderately repetitive DNA, including transposable elements).

Heterochromatin is normally associated with repression of transcription and recombination, and with late replication. In *Drosophila*, the constitutive heterochromatin is underrepresented in polytene chromosomes. No endogenous genes or CpG islands have yet been mapped to regions of C-banded constitutive heterochromatin in the human or mouse genomes, and transgenes that integrate close to constitutive heterochromatin are frequently silenced. A few endogeneous genes in *Drosophila* are found in a heterochromatic location and moreover they cease to function when moved to euchromatic locations.

C-band-positive constitutive heterochromatin has a distinctive chromatin structure. It is the site of most methylation in human and mouse chromosomes and so the ^{5Me}CpG-binding protein MECP2 is highly concentrated over constitutive heterochromatin in vertebrates in most cell types. Constitutive heterochromatin is also associated with very low levels of histone acetylation; with the exception that the chromocentre of *Drosophila* is enriched in histones acetylated at Lys12, and that in mammalian embryonic stem cells prior to differentiation histone H4 within the pericentromeric heterochromatin is acetylated. A set of distinctive chromosomal proteins can also be found concentrated in heterochromatin. Many of these proteins have motifs, such as the chromobox, that are characteristic of proteins involved in the formation of multiprotein heterochromatin complexes. Antibodies against M31, a chromobox containing protein that is a homologue of *Drosophila* heterochromatin protein HP1, highlight the pericentric heterochromatin of mouse and human chromosomes.

References

Craig JM and Bickmore WA (1994) The distribution of CpG islands in mammalian chromosomes. *Nature Genetics* **7**: 376–382.

Craig JM, Boyle S, Perry P and Bickmore WA (1997) Scaffold attachments within the human genome. *Journal of Cell Science* **110**: 2673–2682.

Cross SH, Lee M, Clark VH *et al*. (1997) The chromosomal distribution of CpG islands in the mouse: evidence for genome scrambling in the rodent lineage. *Genomics* **40**: 454–461.

Dunham I, Shimizu N, Roe BA *et al*. (1999) The DNA sequence of human chromosome 22. *Nature* **402**: 489–495.

Dutrillaux B, Couturier J, Richer C-L and Viegas-Péguinot E (1976) Sequence of DNA replication in 277 R- and Q-bands of human chromosomes using a BrdU treatment. *Chromosoma* **58**: 51–61.

Francke U (1994) Digitized and differentially shaded human chromosome ideograms for genomic applications. *Cytogenetics and Cell Genetics* **65**: 206–219.

Jeppesen P and Turner BM (1993) The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* **74**: 281–289.

Korenberg JR and Rykowski MC (1988) Human genome organisation. Alu, Lines, and the molecular structure of metaphase chromosome bands. *Cell* **53**: 391–400.

Saccone S, De Sario A, Wiegant J *et al*. (1993) Correlations between isochores and chromosomal bands in the human genome. *Proceedings of the National Academy of Sciences of the USA* **90**: 11929–11933.

Saitoh Y and Laemmli UK (1994) Metaphase chromosome structure: bands arise from a differential folding path of the highly AT-rich scaffold. *Cell* **76**: 609–622.

The Chromosome 21 Mapping and Sequencing Consortium (2000) The DNA sequence of human chromosome 21. *Nature* **405**: 311–319.

Yunis JJ (1981) Mid-prophase human chromosomes. The attainment of 2000 bands. *Human Genetics* **56**: 293–298.

Further Reading

Craig JM and Bickmore WA (1993) Chromosome bands – flavours to savour. *BioEssays* **15**: 349–354.

Craig JM and Bickmore WA (1997) *Chromosome Bands: Patterns in the Genome*. Heidelberg: Springer-Verlag.

Gosden JR (1994) Chromosome analysis protocols. In: Walker JM (ed.) *Methods in Molecular Biology*, vol. 29. Totawa: Humana Press.

Mitelman F (ed.) (1995) *An International System for Human Cytogenetic Nomenclature*. Basel: S Karger.

Sumner AT (1990) *Chromosome Banding*. London: Unwin Hyman.