The Development of a Novel and Efficient HAC Vector Delivery System to Human Cells

A thesis submitted in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy
Trinity Term, 2007

Kirsty M Simpson
Green College
Faculty of Clinical Medicine
University of Oxford

DEDICATION

This thesis is dedicated to my family: Mum, Dad, Katy, Gran and Grandpa; and also to Daniela for all her help and support. Thank you.

ACKNOWLEDGEMENTS

My time as a post-grad student here in Oxford is finally drawing to a close. All those who know me will know that the last four years have been full of highs and lows for me in many ways, both academically and personally. In the course of finishing this DPhil I have done extra studies, spent many an hour dancing, been through heartache and had a lot of fun along the way. It has been an experience I will remember forever more. There are so many people who have become important to me along the way and I want to say thank you:

Firstly, to my family: Mum, Dad, Katy, Gran and Grandpa. Thank you for your love and support. Financially as well as emotionally!

To the department of Clinical Medicine and Green College for funding me.

To my supervisor Zoia for her support and guidance and to Daniela with whom I have been very lucky to work with. Without your help this would not have been possible and I owe you so much. There have been so many people I have worked with in the lab along the way and I only mention a few here. Thank you to Richard Wade-Martins for all your help with the amplicon preparations, to Doug Vernimmen for your help with the real-time PCR and to Antonio Velayos for your help with the western blots. To all those who I have worked with at the IMM: Anas, David, Kathryn, Alison, Chun, Wade, Louise, Andrew, Tom and Joyce to name just a few.

To those friends from Bristol and back home who have always been there for me: Kim, Andrew, Mim, Paul, Claire and Gary.

To all those who I met at 5 St Margarets, in particular to Becky for your support in my first year.

To Gav for saving me that year! Thank you for all of the memorable (and not so in some cases!) Wednesday nights.

To Richard (at least one of us remembered!). Thank you for your love and support over a great two years.

To all of my friends at OUDC. In particular to Helen for all the girlie chats and to Ian whose wonderful friendship and support I would have been lost without.

To Andy who has had the dubious honour of living with me over the last year and putting up with my stressing and to Andrew who helped me get over the stress by accompanying me around the world.

To the Featherwaits, who have played a massive role in my life and rescuing my sanity over the last two years both in this country and abroad! So, thank you to you all, past and present: Big Andrew, Justin, Carrie, Mel, Rhi, Small Andrew, Naomi, Rich the groupie, and of course to Bruce. In particular: to Tamasin and Alex. The phenomenal amount of work you two put into running “a small irrelevant dance team” has never gone unnoticed and I will be eternally grateful to you both for giving me the

opportunity to dance, for always being there for me and for being such fantastic friends.

To the Babywaits. Where would I be without you in reliving my undergrad years? Rosanna and Becky: the fun and carnage is engrained in my memory; Paul, I have never had such late night msn conversations; Sebastian and Geoffrey, thank you for your comfort and support. And who could forget the broom?

And of course to Paul. Much as I liked your suggestion as to what I should write here, I think I will stick with mine. I have never known anything like it: a rollercoaster from the start. The dancing and some of the happiest memories I have made the tears worthwhile.

So, thank you to you all for making my time here a memorable one. As this chapter of my life closes I look forward to what the future holds.

DECLARATION

The work presented in this thesis was conducted by the author unless otherwise stated. No part of this thesis has been submitted for another degree at this or any other university.

ABSTRACT

THE DEVELOPMENT OF A NOVEL AND EFFICIENT HAC VECTOR DELIVERY SYSTEM TO HUMAN CELLS

Kirsty M Simpson

Thesis presented to the University of Oxford for the degree of Doctor of Philosophy
Green College
Faculty of Clinical Medicine
Trinity Term 2007

Human Artificial Chromosomes (HACs) have been confirmed as viable gene expression vectors and a potential tool for gene therapy. However, standard lipid-based delivery methods pose a developmental barrier. The work presented in this thesis includes the development of a novel and efficient HAC vector system for gene delivery into human cells using Herpes Simplex Virus-1 (HSV-1) amplicon technology. The development of HSV-1 amplicons for HAC delivery is a major step forward in the HAC field.

In this study, utilising the technology allowed the generation of HACs at a high efficiency in a range of human cell types, which is a significant step in the development for HAC gene expression systems. Further work also showed a significant difference in HAC stability between cell lines. Real-time PCR analysis determined that Aurora B was over expressed in cell lines in which the HACs were unstable. This correlated with high levels of chromosomal instability and was confirmed by western blot analysis. Since Aurora B is a kinase involved in at least two cell cycle checkpoints, cellular phosphorylation levels were perturbed to mimic that observed in the unstable cells, using okadaic acid, which is both a protein phosphatase inhibitor and activates Aurora B. Treatment of cells showed an increase in both HAC and overall chromosomal instability and an increase in histone H3 Serine 10 and Serine 28 phosphorylation.

The project also focussed on the development of a gene expression system using HSV-1 amplicons. Two different strategies were explored. Firstly, one approach involved engineering the HPRT genomic locus into an HSV-HAC vector, by Red mediated recombination for complementing the HPRT deficiency in HPRT- HT1080 cells. As an alternative approach, co-infection of two different HSV-1 HAC amplicons for generating a single HAC gene vector was investigated. Initial experiments utilising the latter approach were the most successful and show promise for generating HAC containing genes via this strategy.

Contents

List of Figures

List of Tables

Abbreviations

α	alphoid/alpha DNA
A	Adenine
aa	amino acid
ABI	Applied Biosystems
ACA	Anti-Centromere Autoimmune Serum
ACE	Artificial Chromosome Engineering System
AIR-2	Aurora/Ipl1-related-2
Amp^R	Ampicillin Resistance
APC	Anaphase Promoting Complex
ARSs	Autonomously Replicating Sequences
ATP	Adenosine Tri-Phosphate
BAC	Bacterial Artificial Chromosome
BCIP/NBT	5-bromo-4-chloro-3-indolylphosphate/nitro blue tetrazolium
bp	base pairs
BUB	Budding Unhibited by Benzimidazoles
C	Cytosine
ºC	degrees Centigrade
C. elegans	*Caenorhabditis elegans*
CBMN	Cytokinesis Block Micronucleus
cdc	cell division control protein
CDE	Centromere Determining Element
Cdk	Cyclin dependent kinase
cDNA	complementary DNA
CENP	Centromere Protein
CFTR	Cystic Fibrosis Trans-Regulator
cTc	chlorotetracycline
CHO	Chinese Hamster Ovary
Cid	Centromere identifier
CLIP-170	Cytoplasmic Linker Protein-170
cm	centimetre
Cm^R	Chloramphenicol Resistance
Cnp1	Centromere protein 1
CREST	Calcinosis, Raynaud Syndrome, Oesophageal dysmotility, Scleroderma and Telangiectasia
Cse4	Chromosome segregation 4
C_T	Threshold Cycle
DABCO	1,4-Diazabicyclo(2,2,2)octane
DAPI	4,6-Diamidino-2-Phenylindole
D-MEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTPs	Deoxynucleotide Triphosphates
Drosophila	*Drosophila melanogaster*
dUTP	2'-Deoxyuridine 5'-Triphosphate
E. coli	*Escherichia coli*
EBNA-1	Epstein Barr Nuclear Antigen-1

EBV	Epstein Barr Virus
EGF	Epidermal Growth Factor
EGFP	Enhanced Green Fluorescent Protein
EPO	Erythropoietin
ES cells	Embryonic Stem cells
FAM	Fluorencin
FCS	Foetal Calf Serum
FISH	Fluorescent in situ Hybridisation
FITC	Fluorescein Isothiocyanate
FRET	Förster Resonance Energy Transfer
G	Guanine
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GCH1	Guanosine triphosphate Cyclohydrolase-1
GFP	Green Fluorescent Protein
H3K4	Histone 3 lysine 4
H3K9	Histone 3 lysine 9
H3S10	Histone 3 Serine 10
H3S28	Histone 3 Serine 28
HAC	Human Artificial Chromosome
hCF	human Chromosome Fragment
HCP	Holocentric Protein
HCV	Human Chromosomal Vector
HMG-1	High Mobility Group Protein-1
HP1	Heterochromatin Protein 1
HPRT	Hypoxanthine Phosphoribosyltransferase
HSV-1	Herpes Simplex Virus-1
Hzwint	Human ZW10-interacting protein
ICP-1	Inner Centromere Protein-1
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
INCENP	Inner Centromere Protein
Kan^R	Kanamycin Resistance
kb	kilo bases
kD	kilo Dalton
kMt	kinetochore Microtubule
kV	kilovolts
LB	Luria–Bertani Broth
LDLR	Low-density Lipoprotein Receptor
LMP	Low Melting Point
m	metres
M	Molar
MAD	Mitotic Arrest Deficient
MAP	Mitogen Activated Protein
Mb	Mega base
MCAK	Mitotic Centromere-Associated Kinesin
Me	Methylation
Mis	Mis-segregation
mg	milligrams
ml	millilitres

mM	millimolar
MMCT	Microcell-mediated Chromosome Transfer
MOI	Multiplicity Of Infection
MPF	Maturation Promoting Factor
mRNA	messenger RNA
mV	millivolts
N	Normality
NC-MiCs	Neocentromere-based Minichromosomes
NEB	New England Biolabs
Neo^R	Neomycin Resistance
ng	nanograms
nm	nanometres
nM	nanomolar
Ω	ohms
OA	Okadaic acid
OD	Optical Density
ORC	Origin Recognition Complex
OTC	Ornithine Transcarbamylase
PAC	P1 Artificial Chromosome
PARP	Poly (ADP-Ribose) Polymerase
PCR	Polymerase Chain Reaction
PEI	Polyethylenimine
PFGE	Pulse Field Gel Electrophoresis
pg	picogram
pmoles	picomoles
PP	Protein Phospatase
pRB	Retinoblastoma protein
PVDF	Polyvinylidene difluoride
Rag2	Recombination-activating gene 2
RFP	Red Fluorescent Protein
RLF	Replication Licensing Factor
RNA	Ribonucleic Acid
RNAi	RNA interference
ROD	Rough Deal
ROX	Passive Reference Dye
rpm	rotations per minute
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharomyces pombe*
SARs	Scaffold Attachment Regions
SATAC	Satellite DNA based Artificial Chromosome
SCID-X1	Severe Combined Immunodeficiency-X1
SCNT	Somatic Cell Nuclear Transfer
SMC	Structural Maintenance of Chromosomes
SPF	S phase Promoting Factor
SUMO	Small Ubiquitin Related Modifier
SU(VAR)3-9	Suppressors of Variegation 3-9
Swi6	Switching gene 6
T	Thymine
TACF	Telomere Associated Chromosome Fragmentation
TAMRA	Tetramethylrhodamine

TDT	Telomere Directed Truncation
T_M	Melting temperature
Topo II	Topoisomerase II
U	Unit
µF	microFarads
µg	micrograms
µl	microlitres
µM	micromolar
UV	Ultra violet
V	Volts
Xenopus	*Xenopus laevis*
YAC	Yeast Artificial Chromosome
ZW10	Zeste White 10

Chapter 1: Literature Review

1.1 The Mammalian Chromosome

The idea of chromosomes first appeared at the end of the nineteenth century with German zoologist Anton Schneider's description of the "chromatic nuclear figure" during nuclear division. The term chromosome was then proposed by the anatomy professor Waldeyer in 1888 (Zacharias, 2001). Very early studies indicated the involvement of chromosomes in inheritance. Once Mendel's laws came to light in 1900 it was quickly realised that the behaviour of chromosomes at cell division explained the distribution of genes to the daughter generation.

1.2 The Somatic Cell Cycle

The growth of a multicellular organism is achieved by mitosis, the process whereby chromosomes segregate evenly to daughter cells. The interphase stage between mitoses, where chromosomal DNA (Deoxyribo Nucleic Acid) is actively replicated, comprises the remainder of the somatic cell cycle (Figure 1.1). Mitosis must be tightly regulated to ensure that every daughter cell receives the correct set of chromosomes otherwise severe consequences can result: death, abnormality or cancer. This is achieved by a number of checkpoints in the cell cycle (Lukas et al., 2004). These ensure DNA replication has been completed, that the DNA is undamaged and that all the chromosomes are properly attached to the spindle at metaphase.

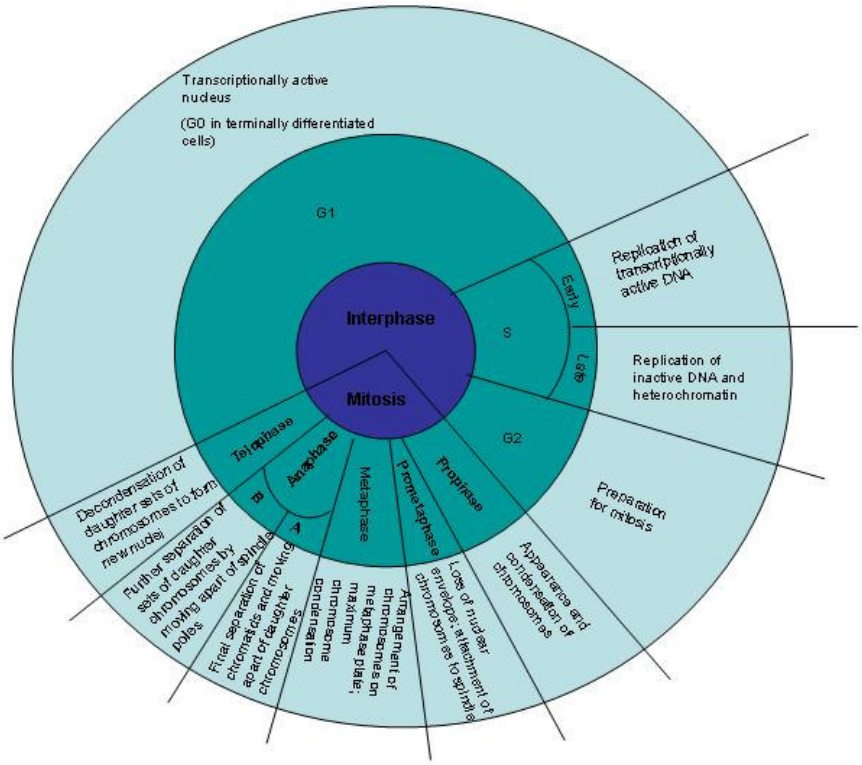


Figure 1.1: The Mitotic Cell Cycle
Adapted from (Sumner, 2003).
The mitotic cell cycle showing the activities that occur at each stage.

G1, during which cellular growth occurs, comprises most of the cell cycle. If cells arrest during G1, for example by limited nutrient supplies, they can enter what is referred to as G0, as is the case for terminally differentiated cells. The G1-S transition in mammals involves phosphorylation of the retinoblastoma protein, pRB. This results in its inactivation and it no longer acts as a transcriptional repressor. Phosphorylation of pRB is controlled by D and E-type cyclins, in turn controlled by a Cdk (Cyclin Dependent Kinase) inhibitor, $p27^{KIP1}$, levels of which decline in G1 in response to mitogen stimulation (Giacinti and Giordano, 2006; Sumner, 2003).

DNA replication occurs during S phase. In order to ensure genome stability from one generation to the next, mechanisms exist to ensure that the entire genome is replicated once only. Once any sequence has been replicated during a given cell cycle, synthesis cannot be initiated again until the next cell cycle. A central component of this system is Replication Licensing Factor (RLF) (Chong et al., 1996) which puts replication origins in an initiation-competent state. S phase Promoting Factor (SPF) induces these origins to replicate and thus removes RLF. Mechanisms exist to ensure that RLF and SPF cannot act on the DNA at the same time (Chong et al., 1996). Further studies reveal that RLF may in fact be part of a much larger complex involved in replication origin licensing (Nishitani and Lygerou, 2002). In addition to licensing, an S phase checkpoint, dependent on proteins at replication forks, prevents the cell proceeding to G2 and mitosis if replication is incomplete (Clarke and Gimenez-Abian, 2000). Once replication is complete, the cell is in G2 and is assessed for damage and the completeness of replication. A large number of conserved proteins are involved in this checkpoint (Clarke and Gimenez-Abian, 2000). Cells are stimulated to enter mitosis by M-phase kinase or Maturation Promoting Factor (MPF), a complex of Cdc (cell division control protein) 2 protein kinase with a cyclin B (Sumner, 2003).

For mitosis to proceed, decatenation of the newly replicated DNA is necessary. This requires the enzyme Topoisomerase II (Topo II) and a G2 checkpoint ensures that DNA catenations have been resolved successfully before cells enter mitosis (Clarke and Gimenez-Abian, 2000).

The first mitotic stage is prophase during which the chromosomes first become visible due to their condensation. The 2 metres (m) of DNA in each mammalian nucleus are

packed into a manageable form by combining the DNA with proteins in a number of stages, with the higher levels differing somewhat between interphase and metaphase. The first stage of compaction is the formation of the nucleosome fibre in which 147 base pairs (bp) of DNA, 1.7 turns, wraps around the histone octamer (containing two molecules each of histones H2A, H2B, H3 and H4; small basic proteins). This core particle forms the basic unit of chromatin (Kornberg, 1974; Luger et al., 1997). Each nucleosome is connected to the next by a stretch of approximately 20-60 bp of linker DNA and H1 holds together the DNA where it enters and leaves the core particle. Formation of the nucleosome fibre compacts the DNA sevenfold and produces a chromatin fibre approximately 10 nanometres (nm) in diameter. This 10 nm fibre forms a 30 nm solenoid structure in which the nucleosome fibre is wound in a regular helix, with about six nucleosomes per turn, resulting in a packing ratio of about 40-50 fold (Sumner, 2003). The compaction of this structure is thought to be modified by the linker histone (Robinson and Rhodes, 2006). The next level of packing involves an arrangement of loops that radiate out from a core or scaffold and in metaphase there appears to be another level beyond this. In interphase the 30 nm fibre appears to be arranged in loops but it is unclear how the mitotic scaffold is related to the nuclear matrix. The mitotic scaffold is a proteinaceous structure that runs along the centre of the chromatids and from which loops of the 30 nm fibres radiate (Stack and Anderson, 2001). In contrast, the nuclear matrix appears to be associated with the surface of the chromosomes (Sumner, 2003). However, there are some similarities. This includes protein components such as Topo II and the DNA sequences that attach to the scaffold seem to be the same as those that attach to the matrix. How the interphase chromatin is arranged with respect to the nuclear matrix is still a matter of much debate (Hancock, 2000).

Specific A (adenine) T (Thymine) -rich DNA sequences, often flanking genes, known as scaffold attachment regions (SARs), bind the chromosome loops to the metaphase chromosome scaffold. The scaffold itself consists of two major non-histone proteins, Topo IIα and ScII (a member of the SMC (Structural Maintenance of Chromosomes) family) and a number of minor proteins (Gassmann et al., 2004b). In addition, a number of centromere proteins are tightly associated with the scaffold.

The packing ratio of a fully condensed mitotic metaphase chromosome is about 10000 fold. The organisation of chromosomes into loops attached to a scaffold may be insufficient to achieve this and an extra level of coiling or condensation may be required. However, little is known about this at the present time. Two classes of proteins have been implicated in chromosome condensation: histones (H1 phosphorylation and H3 Serine 10 (H3S10) phosphorylation) and condensins (one possible function may be to induce supercoils in the DNA and stabilise them by binding to them); as well as Topo II, the role of which is unclear (Hirano, 2000).

At prometaphase the nuclear envelope breaks down and the mitotic spindle is formed. This consists of microtubules oriented with their negative ends towards the poles and their plus ends towards the chromosomes. In somatic cells the microtubules are assembled round the centrosomes, the microtubule organising centres, one at each pole of the cell, in the centre of which lie the centrioles (Sumner, 2003).

At metaphase chromosomes attach to the spindle via their kinetochores. This is essentially a random process until the chromosome is stabilised in the middle of the cell on the metaphase plate. The spindle checkpoint ensures that the cell cannot

proceed to anaphase until both kinetochores of every duplicated chromatid pair have attached correctly to spindle microtubules. This results in the microtubules being under tension. Several spindle checkpoint proteins have been identified (Clarke and Gimenez-Abian, 2000; Cleveland et al., 2003) including a phosphorylated protein recognised by the 3F3/2 antibody present specifically on unattached kinetochores, MAD (Mitotic Arrest Deficient) and BUB (Budding Unhibited by Benzimidazoles) proteins and Cdc20. The target of this checkpoint is the anaphase promoting complex (APC) which is a multisubunit E3 ubiquitin-ligase that specifies the destruction of specific proteins involved in sister chromatid cohesion to initiate anaphase (King et al., 1996).

Once all the chromosomes have become attached correctly to the spindle microtubules the cell can proceed to anaphase, the first step of which is the separation of the sister chromatids to form daughter chromosomes. It is thought that decatenation of centromeric DNA sequences by Topo II is necessary for this (Porter and Farr, 2004). Destruction of securin, which inhibits separase, which in turn digests cohesin, the protein complex holding the sister chromatids together, by the APC is also necessary (Nasmyth, 2001). Once the sister chromosomes have been separated, the spindle can pull the two groups of daughter chromosomes apart. This occurs in two stages: anaphase A where the daughter chromosomes move towards their respective poles by shortening of the microtubules at their chromosomal ends; and anaphase B where further separation is produced by the spindle poles moving further apart (Sumner, 2003).

The final stage of mitosis is telophase when the groups of daughter chromosomes acquire a new nuclear envelope and the chromosomes decondense. Cytokinesis follows this. Cytokinesis severs the cytoplasmic bridge, formed by an actomyosin contractile ring constricting the plasma membrane after anaphase onset, between the two daughter cells. A conserved core of about twenty proteins are involved (Glotzer, 2005). These include components of the central spindle, including Aurora B kinase, members of the RhoA GTPase pathway which direct contractile ring assembly, non-muscle myosin, actin and its regulators (the contractile ring is a network of myosin and actin filaments and myosin motor activity drives the constriction of the contractile ring) and proteins required for trafficking and fusion of membrane vesicles, necessary to complete cytokinesis (Glotzer, 2005).

1.3 Chromosome Classification

Chromosome shape is determined by the position of the centromere (Figure 1.2). Metacentric chromosomes have the centromere at, or close to, the middle of the chromosome thus having long and short arms of similar size. Acrocentric chromosomes have their centromeres close to one end and so they are characterised by a very short arm, often carrying satellites, and a long arm. Intermediates can be referred to as sub-metacentric or sub-acrocentric. Telocentric chromosomes have a terminal centromere and no appreciable short arm.

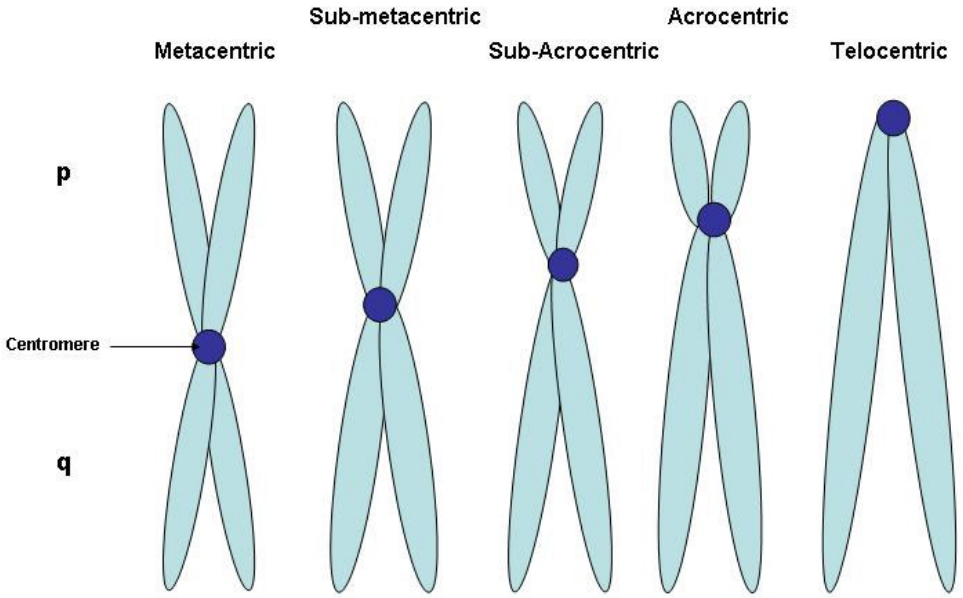


Figure 1.2: Chromosome Shape and Centromere Position
Adapted from (Sumner, 2003).
The short arm of the chromosome is designated the "p" arm and the long arm is known as the "q" arm.

Three elements essential for the stable maintenance and transmission of chromosomes in eukaryotes have been cloned and characterised in yeast: the centromere (Clarke and Carbon, 1980), telomeres (Szostak and Blackburn, 1982) and the origin of replication (Stinchcomb et al., 1979). However, systems for characterising the components of mammalian chromosomes have only recently been developed in mammalian cells.

1.4 Telomeres

Telomeres are the one cis-acting functional component of human chromosomes that have been cloned and shown to function when re-introduced into human cells (Barnett

et al., 1993; Farr et al., 1991). A telomere is functionally defined as a region of DNA at the molecular end of a linear chromosome that is required for replication and stability of the chromosome (Blackburn, 1984). Telomeres are essential for maintaining genome integrity. They protect the chromosomes from recombination, exonuclease degradation and end-to-end fusion and are replicated in germ cells using telomerase, consisting of a reverse transcriptase and RNA (Ribo Nucleic Acid) template complementary to the G (Guanine) -rich telomeric strand. For much of the cell cycle, telomeric DNA is maintained as a loop structure which protects the chromosome ends (Neidle and Parkinson, 2003). Telomeric sequences of eukaryotes consist of noncoding, short tandem repeats organised in arrays of variable length from 2-10 kilobases (kb) in humans. There is a strand bias in G+C (Cytosine) composition. The G-rich strand runs 5'→ 3' towards the chromosomal end. In mammals and other vertebrates this repeat sequence is TTAGGG (Moyzis et al., 1988). The repetitive sequence culminates in a short overhang of single-stranded sequence at the extreme 3' ends (approximately two hundred bases in man), which can fold into a variety of quadruplex structures. This loop can associate with various telomere proteins, making it inaccessible to nucleases (Shay, 1999). Access to telomerase is probably also prevented thus these proteins can play a role in regulating telomere length.

1.5 Origins of Replication

Replication origins in *Saccharomyces cerevisiae* (*S. cerevisiae*) have been clearly identified as Autonomously Replicating Sequences (ARSs) of about 150 bp that contain an essential 11 bp ARS consensus sequence and several other elements that contribute to initiation (Gilbert, 2001). The ARS acts as a binding site for the Origin

Recognition Complex (ORC). In multicellular eukaryotes, replication origins have been difficult to identify. In some systems, any DNA sequence can promote replication but other systems require specific DNA sequences (Gilbert, 2001). Therefore, the relative sequence and epigenetic contributions need to be further defined. In one study in human cells all DNA fragments greater than 15 kb promoted autonomous replication with equal efficiency, thus indicating that only a low level of sequence specificity is required for initiation of replication and that replication signals on average must exist at least every 10 kb in human DNA (Heinzel et al., 1991). The same group investigated the autonomous replication ability of human alphoid/alpha (α)-DNA. A single 2.7 kb module replicated with a very low efficiency. However, increasing the module number to six lead to very efficient replication. These results indicate that efficient autonomous replication can be mediated by sequence features which are not sufficient in small amounts at supporting replication (Krysan et al., 1993). The positions of origins cannot be randomly distributed. This would lead to incomplete replication due to S phase time constraints if origins were too far apart or duplication if origins were too close together. In *Xenopus* (*Xenopus laevis*) eggs replication origins are grouped into clusters (Blow et al., 2001). Within each cluster, most origins are spaced every 5 to 15 kb, the average being 8-9 kb. The average cluster covers ~50 kb. Origins within a cluster fire synchronously whilst each cluster fires at a different time during S phase. Replication is initiated at sites that appear random with respect to sequence but are regularly distributed. More recent research indicates that gene loci frequently contain several origins that are used at relatively low frequencies, meaning that a specific genomic region can be replicated by different origins. This is supported by the finding that both the chicken β-globin and the mouse HoxB loci contain several active origins (Norio, 2006).

1.6 The Centromere and Alpha Satellite DNA

In human chromosomes the centromere is the primary constriction of the chromosome, the site of kinetochore formation and of spindle attachment and is critical to chromosome inheritance. The kinetochore is the protein structure at the centromere, essential for attachment and movement along microtubules. Unattached kinetochores also generate signals for the mitotic checkpoint thus preventing aneuploidy during the cell cycle (Cleveland et al., 2003). The trilaminar organisation of the kinetochore implies the presence of different sub domains. The region closest to the centromeric heterochromatin is the inner plate, the only region to bind both DNA and proteins, where both CENP (Centromere Protein)-A and CENP-C localise (Saitoh et al., 1992). An electron-translucent space separates the inner plate from an electron dense outer plate (Sumner, 2003). The outermost fibrous corona regions of the kinetochore directly interact with the spindle microtubules. Proteins that localise here during mitosis include the motor proteins CENP-E and cytoplasmic dynein, and proteins such as BUB1 that are involved in the mitotic checkpoint.

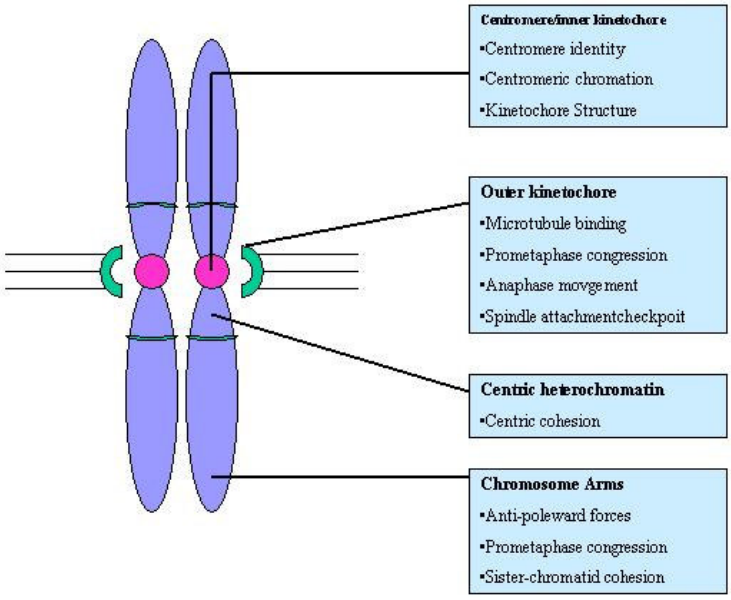


Figure 1.3: Structural and Functional Elements Of The Centromere Region
Adapted from (Sullivan et al., 2001).
This figure details the associated functions of the centromere region.

There are three basic centromere forms: puntiform, localised and holocentric. Puntiform centromeres are very small and bind one microtubule only. The best characterised is that of *S. cerevisiae*. Most eukaryotes have localised centromeres where the centromere forms on a large, but specific chromosomal region. Holocentric chromosomes, where the centromere forms along the length of the chromosome, can be found in some species, including nematodes.

The best defined centromere is that of *S. cerevisiae* (Hyman and Sorger, 1995). These 125 bp "point" centromeres consist of three conserved elements (CDEI, Centromere Determining Element I, an 8 bp imperfect palindrome; the central element CDEII, 80 bp of AT-rich DNA which isn't conserved in sequence, and CDEIII, a 25 bp

imperfect palindrome) packaged into a compact nuclease-resistant chromatin structure of ~250 bp (Cleveland et al., 2003). This sequence dependence for centromere function is absent in other eukaryotes. However, centromeres are generally surrounded by heterochromatin, a region of repetitive, gene-poor DNA.

Heterochromatin was first described as chromatin that remained compact throughout interphase. Two main classes have since been defined: facultative and constitutive heterochromatin (Grewal and Jia, 2007). Facultative heterochromatin occurs usually in only one of a pair of chromosomes, thus having the same DNA composition as its euchromatic homologue. It can be regarded as regions that are epigenetically repressed, possibly for only part of the life cycle, assembled when genes need to be silenced. A key feature of facultative heterochromatin is its ability to influence gene expression in a specific manner via propagation, a phenomenon known as epigenetic silencing (Grewal and Jia, 2007). In contrast, constitutive heterochromatin remains condensed throughout an organism's lifespan, occurs at the same site in both chromosome homologues and tends to have a substantially different DNA composition to that of euchromatin. It can also be defined by its staining properties. C-banding stains almost all large segments of constitutive heterochromatin that fail to decondense at metaphase. Heterochromatin is generally transcriptionally inactive and late replicating. Constitutive heterochromatin is found in the centromeric region and terminal regions of chromosomes. In eukaryotes it typically contains high amounts of satellite DNA and this generally has no common properties. Within a species, or even a single chromosome, there can be different types of DNA in constitutive heterochromatin and its properties cannot be solely attributed to DNA sequence. However, one consistent feature of heterochromatin that appears to be important for

condensation is cytosine methylation. Certain proteins or modifications are found in heterochromatin. For example, histone underacetylation, associated with transcriptional inactivity and HP1 (Heterochromatin Protein 1) which binds to heterochromatin via its chromodomain at H3 lysine 9 (H3K9) methylated sites and forms complexes with proteins such as SU(VAR)3-9 (Suppressors of Variegation 3-9), a histone methyltransferase, largely responsible for the H3K9 methylation (Grewal and Jia, 2007). A large degree of heterochromatin variability exists due to the associated protein factors and the order in which they are assembled (Craig, 2005; Grewal and Jia, 2007).

This heterochromatic state may help localise centromeric components (Henikoff et al., 2000) and the study of a *Drosophila* (*Drosophila melanogaster)* chromosome suggests that the alteration of heterochromatin can affect a chromosome's mitotic stability (Wines and Henikoff, 1992) and function (Allshire et al., 1995). Heterochromatin may also be important in chromosome segregation through recruitment of cohesin (essential for sister-chromatid cohesion) and the interaction of HP1 with Inner Centromere Protein, INCENP (Grewal and Jia, 2007).

In the fission yeast, *Schizosaccharomyces pombe* (*S. pombe*), a non-homologous, AT-rich, 4-5 kb central core sequence is flanked by conserved inverted inner and outer repeats (Baum et al., 1994; Nakaseko et al., 1987). In the nematode worm *Caenorhabditis elegans* (*C. elegans*) the chromosomes are holocentric and recruit centromeric proteins along their length without the requirement of specific DNA sequences (Pidoux and Allshire, 2000). The *Drosophila* centromere has been defined on a 1.3 Mega base (Mb) X-derived minichromosome Dp1187. A 420 kb region of

centric heterochromatin contains a fully functional centromere (Murphy and Karpen, 1995). It consists of a core of 5 bp satellites (AATAT and AAGAG), making up more than 85% of the 420 kb, and transposons comprising 10%, flanked by repetitive DNA (Sun et al., 1997). However, in the rest of the genome these satellites and transposons are neither unique to centromeres nor present at all centromeres.

The alphoid DNA found at human centromeres is a complex family of tandemly repeated DNA based on a monomer repeat length of 171 bp (Waye and Willard, 1986) that exhibit chromosome-specific variation in nucleotide sequence and higher order repeat arrangement (Choo et al., 1991). These monomers can exist in a tandemly arranged heterogeneous (monomers are ~70% identical) monomeric form or they can be arranged in a homogenous (monomers are typically 97-100% identical), higher order, multimeric form (Rudd and Willard, 2004). There are at least 33 alphoid subfamilies (Choo et al., 1991). Some are chromosome specific, some are not, some chromosomes possess only a single sub-family within their centromeres and others possess several different subfamilies. Higher order alpha satellite DNA arrays have been identified at the centromeric regions of each human chromosome, ranging from 500 kb to 5 Mb, altogether constituting some 5% of total human DNA. Higher order alpha satellite DNA is known to be important for centromere function (Schueler et al., 2001). Centromere proteins A, -B and –C have been shown to assemble on higher order alphoid DNA (but not monomeric alphoid DNA) that has integrated ectopically into host chromosomes following transfection. These integration sites often form secondary constrictions, and can break off to form minichromosomes (Nakano et al., 2003). This provides evidence that alphoid DNA is required for centromere function.

Alphoid DNA is considered the best candidate for the DNA requirement of centromere function. Several papers have estimated the minimal amount of α-DNA necessary for centromeric function to be in the range of 78 kb (Lo et al., 1999) and 65 kb (Yang et al., 2000). Analysis of rearranged Y chromosomes (Tyler-Smith et al., 1993), identified in patient cell lines, localised the sequence necessary for centromere function to ~500 kb. When the centromere of the Y chromosome was fragmented with telomeric DNA, it was shown that a derivative chromosome with only 140 kb of α-DNA was capable of centromere function (Brown et al., 1994). A familial rearranged chromosome 17 with only 20-30% of the normal alphoid array still retained full mitotic centromere activity supporting the idea that human centromeres are structurally and functionally repetitive (Wevrick et al., 1990). However, it is also clear that α-DNA is not always sufficient for centromeric activity. In human dicentric chromosomes, normal amounts of α-DNA are present at two distinct chromosomal sites but only one of these forms an active centromere (Earnshaw et al., 1989; Sullivan and Willard, 1998). However, reducing the distance between centromeres on dicentric chromosomes can result in both centromeres being capable of kinetochore formation, implying epigenetic events influence centromere function (Sullivan and Willard, 1998). It is also clear that in some situations α-DNA is not necessary for centromere formation. Centromeres can form on human chromosomes regions lacking alphoid repeats (Amor and Choo, 2002). These neocentromeres suggest an epigenetic component to kinetochore assembly.

1.7 Protein Components of the Kinetochore

1.7.1 CENP-A

CENP-A was originally identified as an autoantigen that was recognised by human CREST (Calcinosis, Raynaud Syndrome, Oesophageal dysmotility, Scleroderma and Telangiectasia) antisera (Earnshaw and Rothfield, 1985). CREST is an autoimmune disease and patients have circulating anti-centromere antibodies. CENP-A has been identified as a 17 kilo Dalton (kD) histone-H3 related protein localised at the inner kinetochore plate of mammalian mitotic chromosomes (Palmer et al., 1991). H3 is one of the four types of histone that form the histone octamer (containing two molecules each of histones H2A, H2B, H3 and H4) which 147 bp of DNA wraps around to form the basic unit of chromatin (Kornberg, 1974; Luger et al., 1997). CENP-A associates with all active centromeres, including neocentromeres (Warburton et al., 1997). This histone H3 variant is conserved in *S. cerevisiae* (Cse4, Chromosome segregation 4) (Meluh et al., 1998), *S. pombe* (Cnp1, Centromere protein 1) (Takahashi et al., 2000), *C. elegans* (HCP-3, Holocentric Protein-3) (Buchwitz et al., 1999) and *Drosophila* (Cid, Centromere identifier) (Henikoff et al., 2000) (Figure 1.4), displaying a histone H3-like histone fold domain, which specifies centromeric localisation, coupled to a unique amino-terminal domain (Sullivan et al., 1994). The histone fold domain consists of three α-helices separated by two β-sheet structures (Arents et al., 1991).

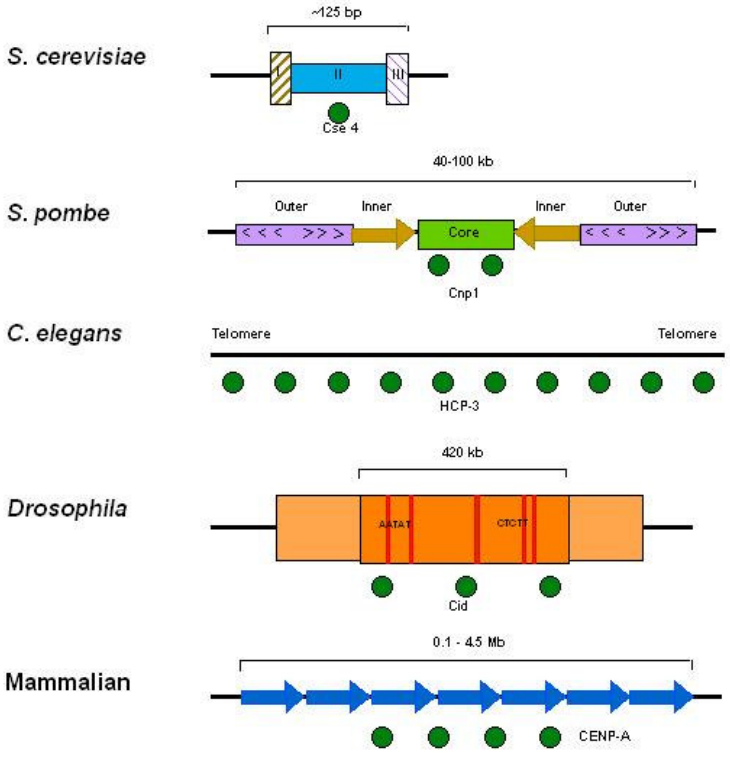


Figure 1.4: Centromeric DNAs and CENP-A Conservation

Adapted from (Sullivan et al., 2001).

This figure demonstrates the difference between eukaryotic centromeric DNA yet the high level of conservation of CENP-A and its homologues:

a) In *S. cerevisiae* Cse4 localises to the conserved element CDEII
b) In *S. pombe* Cnp1 localises to the central core, flanked by conserved inner and outer repeats
c) *C. elegans* kinetochores assemble along the length of the chromosome and HCP-3 assembles on the poleward facing side of the chromatids
d) In the defined *Drosophila* minichromosome Dp1187, cid is localised at a core of 5 bp satellites and transposons flanked by repetitive DNA
e) Human centromeres are composed of alphoid DNA higher order repeats to which CENP-A localises.

1.7.1.1 CENP-A Nucleosomes

Nucleosomes can be assembled *in vitro* from CENP-A and histones H2A, H2B and H4 (Yoda et al., 2000). CENP-A has also been recovered in mononucleosome fractions (Palmer et al., 1987) and locates only to active centromeres throughout the cell cycle (Warburton et al., 1997). In vivo, CENP-A nucleosomes are interspersed with H3 nucleosomes and the resulting chromatin is thought to be organised in a cylindrical structure so that CENP-A faces outwards for kinetochore assembly and interaction with spindle microtubules (Blower et al., 2002). Over expression of histone H3 is lethal in Cse4 mutants, suggesting that H3 competes with Cse4 (Glowczewski et al., 2000). The CENP-A centromere-targeting domain and H4 form more conformationally rigid interactions than H3 and H4 which may contribute to centromere identity (Black et al., 2004). It is presumed that specific regions of CENP-A are responsible for localisation (Palmer et al., 1991). Inactivating CENP-A in yeasts, worms, flies and mice disrupts mitosis and cell cycle progression and causes the mislocalisation of many kinetochore and centromere-region proteins, including CENP-C (Blower and Karpen, 2001; Howman et al., 2000; Moore and Roth, 2001). Disruption of Cse4 in *S. cerevisiae* causes chromosome missegregation and mitotic arrest (Stoler et al., 1995). This implies that CENP-A is at the top of a pathway driving kinetochore assembly. However, one kinetochore protein is unaffected by CENP-A RNAi (RNA interference) and that is Mis12 (Mis-segregation 12) (Goshima et al., 2003; Kline et al., 2006). This indicates a possible kinetochore assembly pathway independent from CENP-A. In addition, the presence of CENP-A is not sufficient to trigger kinetochore assembly. Over expression of CENP-A leads to its incorporation into chromosomal arms and to the localisation of CENP-C and some kinetochore proteins, but not to the assembly of a functional kinetochore (Van Hooser

et al., 2001). This evidence, the homology of CENP-A to histone H3 and its presence at the centromere throughout the cell cycle makes it a strong candidate for an epigenetic mark of kinetochore assembly by packaging centromeric DNA into a specialised chromatin structure.

1.7.1.2 CENP-A and the Cell Cycle

Recent evidence has implicated a role for CENP-A in spindle checkpoint signalling (Regnier et al., 2005). In chicken DT40 cells CENP-A knockouts resulted in missegregation after a mitotic delay, possibly due to the cells being unable to maintain kinetochore localisation of the checkpoint protein BubR1 (Regnier et al., 2005). A further possible role for CENP-A is in cytokinesis. CENP-A is phosphorylated in prophase after the completion of H3 phosphorylation, which is necessary for chromosome condensation, suggesting that they serve different functions (Zeitlin et al., 2001a). The passenger protein Aurora B has been shown to phosphorylate both CENP-A and H3 with phosphorylation mutants of CENP-A resulting in a delay in cytokinesis (Zeitlin et al., 2001b).

1.7.1.3 CENP-A and Centromere Identity and Inheritance

Using the cid promoter from *Drosophila*, H3 histones were deposited in euchromatin, native cid in centromeres and heterologous CENP-A homologues from yeast, worm and human in heterochromatin (Henikoff et al., 2000). This implies that the heterochromatic state facilitates the localisation of centromeric proteins. Centromere

flanking heterochromatin is a prerequisite for centromere maintenance and may be required to organise the higher order structure of centromeric chromatin (Henikoff et al., 2001). Selective methylation of histone H3 lysine 9 by SU(VAR)3-9 homologues (originally identified in *Drosophila* as a chromatin-binding translation initiation factor that suppresses position effect variegation) can create binding sites for heterochromatin proteins such as HP1 that are involved in marking chromatin states and chromosome domain organisation (Jenuwein, 2001), and thus propagate an epigenetic signal. They regulate epigenetic silencing by promoting and maintaining chromosome condensation and bind to methylated H3 via their chromodomain (Grewal and Jia, 2007; Lachner et al., 2001).

In addition to CENP-A, centromeric chromatin underlying the kinetochore has distinct H3 modifications (Sullivan and Karpen, 2004). Acetylation of H3, H4 and methylation (Me) of H3 on Lys4 (H3K4Me) are euchromatic marks. H3K4 di-methylation ($H3K4Me_2$) and H3K4 tri-methylation ($H3K4Me_3$) are associated with potentially active and transcriptionally active sites respectively. On the other hand, methylation of H3 on Lys9 (H3K9Me) is associated with transcriptionally silent chromatin and heterochromatin with di-methylation and tri-methylation marking facultative and constitutive heterochromatin, respectively, in mammals (Dunleavy et al., 2005; Sullivan and Karpen, 2004). Centromeric chromatin has turned out to have its own conserved modification pattern. H3 present within the centromeric chromatin is enriched for $H3K4Me_2$, normally associated with open but not active euchromatin. No H3K9Me, associated with heterochromatin, or H3 or H4 acetylation, associated with euchromatin, was observed. $H3K4Me_3$, associated with actively transcribed regions was also absent. It therefore appears that centromeric chromatin contains H3

and H4 marks in a pattern distinct from either heterochromatin or euchromatin and this may contribute to the structure and function of the centromere in ways such as differential replication timing, maintaining centromere size, co-ordinating the 3D centromeric chromatin structure and the epigenetic propagation of centromere identity (Sullivan and Karpen, 2004).

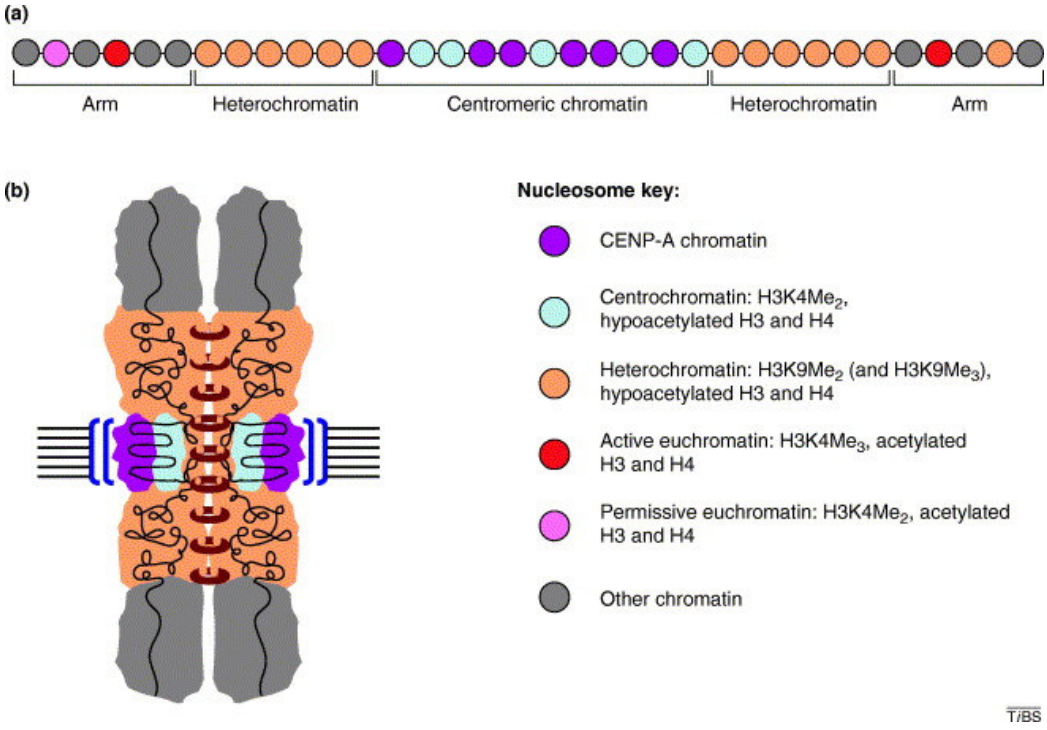


Figure 1.5: The Structure and Organisation of Centromeric Chromatin
Taken from Dunleavy et al., 2005.

a) Representation of an extended chromatin fibre. The centromeric domain is marked by CENP-A and H3 nucleosomes marked with $H3K4Me_2$, and is flanked by heterochromatin.
b) 3D organisation of human centromeric chromatin. CENP-A nucleosomes face outward and are underlined by H3 nucleosomes marked with $H3K4Me_2$.

Pericentromeric heterochromatin is a prerequisite for centromere maintenance (Henikoff et al., 2001). Murine and human SU(VAR)3-9 proteins are constitutive heterochromatin components involved in transcriptional silencing (Aagaard et al., 1999). SUV39H1 (human homologue) accumulates at the outer centromere during prometaphase and dissociates at the meta- to anaphase transition. It is also present at the active centromere of a dicentric chromosome and at a neocentromere (Aagaard et al., 2000). Histone methylation can be considered a long-term epigenetic mark due to the relatively low turnover of the methyl group. Methylation of H3K9 interferes with the phosphorylation of Serine-10 (Rea et al., 2000) that is dependent on Aurora B kinase (Hsu et al., 2000). Ectopic SUV39H1 disperses the pericentric localisation of Serine-10-phosphorylated H3, delays mitotic progression and induces chromosome lag at anaphase (Melcher et al., 2000). Thus, SUV39H1 dependent methylation of H3K9 may induce chromatin condensation. In addition, SU(VAR)3-9 enzymes associate with HP1 proteins (Melcher et al., 2000) and the heterochromatin association of HP1 is lost in the absence of SUV39-dependent methyltransferase activities (Lachner et al., 2001). This scenario is similar to that seen in *S. pombe* where Swi6 (Switching gene 6, a HP1 homologue) association to the outer repeats of the centromere is dependent on Rik1 and Clr4 (SU(VAR)3-9 homologues) (Ekwall et al., 1996). This association mediates the silencing of genes within the repetitive regions. Mutations in HP1 can also prevent the repression of genes lying in close proximity to centromeric heterochromatin and cause defective chromosome segregation (Kellum and Alberts, 1995). These results suggest that SUV39H1 can propagate an epigenetic signal which can be stabilised by HP1 protein interactions (Jenuwein, 2001). A possible regulatory mechanism has been revealed as H3S10

phosphorylation by Aurora B is necessary for the dissociation of HP1 from chromatin during mitosis (Fischle et al., 2005).

A number of other proteins acting at the DNA/chromatin level have been localised to the pericentric region of chromosomes (Choo, 2000). These include Topo II (involved in chromosome condensation and sister chromatid disjunction), HMG-1 (High Mobility Group Protein-1, involved in nucleosome positioning), HP1 and PARP (Poly (ADP-Ribose) Polymerase) which is involved in the maintenance of chromatin structure and DNA repair. PARP is able to bind to human α-satellite DNA and to human neocentromeres but not inactive centromeres on dicentric chromosomes (Earle et al., 2000). It appears to play a role in chromosome stability (Choo, 2000) and its specific binding to centromeres may indicate an epigenetic role in centromere function, possibly by the modification of its target proteins such as histones, topoisomerases, HMG proteins and DNA polymerases and ligases (Choo, 2000). It has been shown to interact with CENP-A, CENP-B and BUB3 but not CENP-C (Saxena et al., 2002a; Saxena et al., 2002b).

The *S. pombe* centromere, consisting of many Cnp1-containing nucleosomes, may form a stem loop structure in which the core is in the loop and the outer and inner repeats are in the stem (Steiner and Clarke, 1994). This model provides an epigenetic system where specific *de novo* folding of centromere components leads to a higher order chromatin structure. A repeat-subunit model has been proposed which could explain how this can be extrapolated to higher eukaryotes and holocentric organisms (Zinkowski et al., 1991).

It may be that centromere identity occurs through a post-replication, CENP-A specific loading factor (Sullivan et al., 2001). *S. pombe* Mis6 (Mis-segregation 6), an essential centromere connector protein acting during G_1-S phase, is required for Cnp1 localisation to the central core (Takahashi et al., 2000) and further proteins are upstream. Mis-16 and Mis-18 are thought to be the most upstream factors in kinetochore assembly (Hayashi et al., 2004). However, CENP-I which has similar sequence similarity to Mis6, is not required for CENP-A localisation (Nishihashi et al., 2002) and in *S. cerevisiae*, Ctf3p, the Mis6 homologue, is not required for Cse4 loading (Measday et al., 2002). The identification of further proteins that interact with CENP-A will lead to a better understanding of this theory.

Parental CENP-A nucleosomes inherited by replicated kinetochores could direct a chromatin assembly or remodelling factor to the kinetochore thus not relying on DNA sequence for centromere identity. Mammalian centromeres are known to replicate during the second half of S phase (Shelby et al., 2000). Histone H3 expression peaks in early to mid –S phase and CENP-A mRNA (messenger RNA) and protein levels are highest in late S phase and G2. Human CENP-A associated DNA is neither the first nor the last region to be replicated. Human centromeres replicate asynchronously in mid- to late S phase, simultaneously with non-centromeric DNA. CENP-A is actively assembled at centromeres in G2. Therefore, human centromere replication occurs before new CENP-A is assembled into chromatin. New CENP-A was also shown to be assembled into centromeres even when replication was blocked (Shelby et al., 2000). Unlike most genomic chromatin it can be concluded that histone synthesis and assembly are uncoupled from DNA replication at the kinetochore. *Drosophila* centromeres have also been found to replicate asynchronously in mid- to

late S phase, simultaneously with H3-containing chromatin (Sullivan and Karpen, 2001). These studies provide evidence that centromere identity is not linked to CENP-A expression and replication timing and that regulated chromatin assembly or remodelling may instead be important in centromere identity. Recent evidence has been revealed by over expressing CENP-A, which leads to the expansion of centromeric chromatin (Lam et al., 2006). This expansion occurred over both higher order alphoid DNA and active genes indicating that there is no sequence requirement for CENP-A-containing chromatin formation. Purification of a CENP-A pre-assembly complex from *Drosophila* yielded a simple complex with the only centromere-specific protein present being Cid itself (Furuyama et al., 2006). One possibility is that CENP-A can deposit promiscuously but can only gain access for assembly onto DNA at centromeres. This could account for *de novo* artificial chromosome and neocentromere formation (Dawe and Henikoff, 2006). A CENP-A driven self-assembly mechanism is further supported by the finding that substitution of the CENP-A targeting domain into histone H3 generates the same conformationally rigid nucleosome as does CENP-A (Black et al., 2007a) and that while CENP-A depletion is lethal, it can be rescued by H3 carrying the CENP-A targeting domain (Black et al., 2007b).

1.7.2 CENP-B

The highly conserved 80 kD CENP-B protein (Masumoto et al., 2004), found in the inner heterochromatin region beneath the kinetochore plates (Cooke et al., 1990), localises to human and mouse centromeres through direct binding to the 17 bp CENP-B box sequence (Masumoto et al., 1989) at its amino-terminal region (which contains

four alpha helices) and forms homodimers at its carboxy-terminal region (Yoda et al., 1992). This appears to form a distinct complex of a CENP-B dimer and two molecules of alphoid DNA. The CENP-B box is found at regular intervals both in human α-DNA and mouse minor satellite DNA. In humans it is found in every other alphoid repeat (Ikeno et al., 1994). However, it appears only in higher order alphoid DNA of the human autosomes and X chromosome (Masumoto et al., 1989). CENP-B is absent on active neocentromeres and the normal Y chromosome centromere, which has no capacity for *de novo* artificial chromosome formation (Mejía et al., 2002), as well as the centromeres of African Green Monkey cells (Goldberg et al., 1996). In mice, the Y chromosome lacks the CENP-B containing minor satellite DNA but in humans α-DNA is present yet seems to lack CENP-B boxes (Tyler-Smith and Brown, 1987). This protein is insufficient for centromerisation as it is found on both active and inactive centromeres (Sullivan and Schwartz, 1995) in pseudo-dicentric and –multicentric chromosomes. There is a lack of obvious mitotic and meiotic phenotype in Cenpb knockout mice (Fowler et al., 2004; Hudson et al., 1998). However, a decrease in body weight and testis size accompanied protein deficiency. Despite this, both the CENP-B box and higher alphoid DNA are required for *de novo* artificial chromosome formation and the assembly of functional centromere components such as CENP-A, CENP-C and CENP-E (Basu et al., 2005; Ohzeki et al., 2002); thus implying that CENP-B is important in centromere function. One theory is that CENP-B boxes promote CENP-A binding (Irvine et al., 2004). It may organise centromeric DNA by juxtaposing two CENP-B boxes in the alphoid DNA repeat through both its DNA-protein interactions and its protein-protein interactions (Yoda et al., 1992). The binding of CENP-B to the CENP-B box sequences causes it to bend by 59˚ (Tanaka et al., 2001). This kinked structure may be important for the centromere-specific

chromatin structure. CENP-B appears to cause nucleosome positioning between pairs of CENP-B boxes, thus forming a centromere-specific higher order structure (Yoda et al., 1998) (Figure 1.6). This has been supported further by recent evidence. CENP-B can still bind to nucleosomal DNA when the CENP-B box is wrapped within the nucleosome core particle and induce translational positioning of the nucleosome (Tanaka et al., 2005).

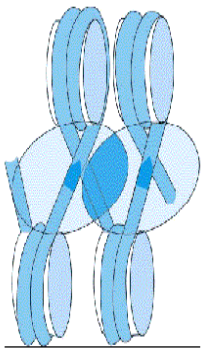


Figure 1.6: A Proposed Structure of the CENP-B/α-satellite DNA/Core Histone Complex

From (Yoda et al., 1998).
CENP-B forms dimers at the C-terminus (blue ovals) and the N-terminus binds to the adjacent CENP-B box (blue arrow). This links the nucleosomes in a compact structure.

1.7.3 CENP-C

CENP-C, a 140 kD constitutive centromere protein, directly binds DNA *in vitro* via a DNA-binding domain (Sugimoto et al., 1994). This domain, in the central region, overlaps with the region necessary to target the protein to the centromere (Yang et al., 1996). The domain proposed to bind DNA *in vitro* has been shown to be necessary for the recognition of centromeric DNA *in vivo* (Politi et al., 2002). CENP-C is localised at the inner kinetochore laminar on mitotic chromosomes (Saitoh et al., 1992) and is associated with active centromeres throughout the cell cycle. There is a peak in its expression in G1 (Knehr et al., 1996), hinting at a role in cell cycle control. It is only detected at active centromeres (Earnshaw et al., 1989; Sullivan and Schwartz, 1995). It appears that CENP-B and CENP-C associate with the same types of alpha-satellite arrays but in distinct non-overlapping centromere domains (Politi et al., 2002). CENP-C localisation is dependent on CENP-A and, in addition, on CENP-H and CENP-I in chicken DT40 cells (Fukagawa et al., 2001; Howman et al., 2000; Nishihashi et al., 2002). CENP-C alone is insufficient for centromerisation (Fukagawa et al., 1999) but it is essential for proper mitotic segregation and cell survival (Kalitsis et al., 1998), with disruption leading to mitotic arrest and irregular cellular and nuclear morphologies. Injection of cells with anti-CENP-C antibodies shows mitotic delay and the formation of fewer and shortened kinetochores (Tomkiel et al., 1994) but only if injected before prophase, thus suggesting that the antibodies interfere with the localisation of CENP-C at centromeres. Depletion in *C. elegans* results in a kinetochore null phenotype with complete failure of mitotic chromosome segregation, failure to recruit kinetochore components and an unstable spindle (Oegema et al., 2001). The failure of resolution of sister centromeres has also been reported (Moore and Roth, 2001). The carboxy terminal of CENP-C shares a region

of homology with the essential *S. cerevisiae* centromere protein Mif2 (Meluh and Koshland, 1995). Mutations in this result in defective chromosome segregation and delayed progression to mitosis (Brown et al., 1993). However, this Mif2 homology region of CENP-C is unnecessary for centromere targeting (Yang et al., 1996). An interaction between CENP-C and CENP-B has been detected and the CENP-C domains required overlap with three Mif2 homologous regions (Suzuki et al., 2004). Overproduction of CENP-B with no CENP-C interaction domains caused abnormal duplication of CENP-C domains at G2 and cell cycle delay at metaphase. This interaction may be important for the assembly of CENP-C on alphoid DNA yet CENP-C is able to assemble without CENP-B (Hudson et al., 1998). CENP-C may be involved in the epigenetic marking of the centromere through modifications by SUMO-1 (Small Ubiquitin Related Modifier-1) (Everett et al., 1999).

1.7.4 CENP-E

The 312 kD CENP-E protein was identified as a kinesin-like motor protein (the homology being at the amino-terminus) which accumulates in G2 (Yen et al., 1992) and is localised to active centromeres (Sullivan and Schwartz, 1995; Yen et al., 1992). It localises specifically during mitosis to the outermost corona region of the kinetochore and associates with the kinetochores during congression towards the midplate and relocates to the spindle midzone at anaphase. It is discarded at the end of mitosis (Yen et al., 1992). The kinetochore targeting domain lies near the carboxy-terminus (Chan et al., 1998) and through this domain CENP-E can interact with itself, CENP-F and hBUBR1. When depleted from mammalian kinetochores, mitotic arrest with a mixture of aligned and unaligned chromosomes results (McEwen et al., 2001).

Kinetochore Microtubule (kMt) binding is reduced at aligned chromosomes and is severely reduced at unaligned chromosomes. Centromere tension also results, thus increasing spindle fragmentation and the number of mono oriented chromosomes abnormally close to the spindle pole. However, typical patterns of congression and checkpoint protein localisation were observed. McEwen et al. propose that these results can be explained by a model in which redundant mechanisms, involving dynein, ZW10 (Zeste White 10) and ROD (Rough Deal), enable kinetochore microtubule binding and checkpoint monitoring when CENP-E is absent. CENP-E and dynein may form redundant pathways for kMt binding with the CENP-E pathway being monitored by hBUBR1 and the dynein pathway being monitored by ZW10 and ROD (McEwen et al., 2001). CENP-E has in fact been shown to activate the checkpoint protein kinase BubR1 (Mao et al., 2003) and is thought to act as a kinetochore attachment sensor that simultaneously binds microtubules and kinetochore bound checkpoint components. BubR1 appears to recruit CENP-E onto unattached kinetochores whereupon CENP-E activates the BubR1 kinase leading to a signalling cascade which prevents advance to anaphase (Mao et al., 2003). More recently, it has been shown that microtubule capture by CENP-E silences the BubR1 kinase activity. Thus, CENP-E is responsible for silencing BubR1-dependent mitotic checkpoint signalling through its capture at kinetochores of spindle microtubules (Mao et al., 2005). Human CENP-E has two microtubule-binding domains. One at the amino-terminus which is ATP (Adenosine Tri-Phosphate) dependent and one at the carboxy-terminus which is ATP independent and has four potential MAP (Mitogen Activated Protein) kinase and cdc2 (cell division control protein) phosphorylation sites (Liao et al., 1994). Thus, MAP kinase could be involved in the mediation of chromosome interactions with the microtubules (Zecevic et al., 1998).

1.7.5 CENP-F

CENP-F is a 367 kD facultative mammalian kinetochore protein of the nuclear matrix, originally identified by an autoimmune sera (Rattner et al., 1993). It peaks in G2 and mitosis and is rapidly degraded when mitosis is completed (Liao et al., 1995). During G2 it is detected at the pre-kinetochore complex and during mitosis CENP-F is associated with the kinetochores from pro-metaphase until early anaphase. It is then detected at the spindle midzone until telophase where it is concentrated at the intracellular bridge at either side of the mid-body (Liao et al., 1995). The 3210 amino acid (aa) protein is mainly coiled coil sequence and the kinetochore localisation region is located at the carboxy-terminus (Zhu et al., 1995). CENP-F is capable of interacting with itself, CENP-E and hBUB1. The kinetochore-targeting domain of CENP-F interacts with the kinase domain of BUB1 (Jablonski et al., 1998). In *C. elegans*, HCP-3 (CENP-A homologue) and HCP-4 (CENP-C homologue) are both required for the localisation of HCP-1 (CENP-F homologue) (Moore and Roth, 2001). Both CENP-E and CENP-F are farnesylated (Ashar et al., 2000). Farnesylation of CENP-E is required for microtubule interaction (Ashar et al., 2000) and farnesylation of CENP-F is required for its localisation to the nuclear envelope and kinetochores, progression through G2/M and degradation after mitosis (Hussein and Taylor, 2002). CENP-I also appears to be required for the localisation of CENP-F (Liu et al., 2003). Depleting CENP-F leads to misaligned chromosomes with a phenotype similar to that seen in CENP-E depleted cells (Yang et al., 2005). Premature chromosome condensation followed by mitotic cell death was also observed (Yang et al., 2005). A strong mitotic delay, thought to be due to continued activation of the mitotic checkpoint caused by unstable microtubule capture by CENP-F depleted kinetochores, has also been observed (Bomont et al., 2005; Feng et al., 2006). CENP-

F contains two microtubule-binding domains (Feng et al., 2006) and whilst CENP-F may not be essential for the mitotic checkpoint, kinetochores depleted of CENP-F may not be capable of generating the required level of signal for a mitotic block (Feng et al., 2006). Due to its interaction with CENP-E it may be that CENP-E and –F form a complex, possibly targeted to the centromere by CENP-C, which could be important for chromosome segregation.

1.7.6 CENP-G

CENP-G was identified in serum from a patient with gastric antral vascular ectasia disease (He et al., 1998). The 95 kD protein is a constitutive centromere protein, detected at the centromere throughout the cell cycle and is associated with the nuclear matrix. In mitosis it's restricted to the inner plate. It is specifically associated with higher order alphoid DNA. Its presence in Chinese Hamster Ovary (CHO) cells indicates that it is a conserved protein (He et al., 1998). CENP-G has been detected at inactive centromeres, suggesting it binds to alphoid DNA, at neocentromeres and on the Y chromosome which lacks higher order alphoid DNA, suggesting that it interacts with other centromeric proteins (Gimelli et al., 2000). This contrasts with CENP-A, -C, –E and -F which are found at all active centromeres including neocentromeres and with CENP-B which is not present on neocentromeres, or the Y chromosome centromere, yet it has been detected at inactive centromeres.

1.7.7 CENP-H

The 28 kD CENP-H colocalises with CENP-A and CENP-C at active centromeres in both interphase and metaphase (Sugata et al., 2000). It is also present outside centromeric heterochromatin where CENP-B is localised and inside the kinetochore corona where CENP-E is localised. It has been detected at neocentromeres but not at inactive centromeres. Binding assays suggest that CENP-H binds to itself and MCAK (Mitotic Centromere-Associated Kinesin). Results suggest that CENP-H multimers localise to the kinetochore plate (Sugata et al., 2000). Recently CENP-H has been shown to interact with the Nuf2 complex, thought to be a structural component of the centromere, possibly connecting the inner and outer kinetochores (Mikami et al., 2005). Mouse CENP-H has also been isolated (Sugata et al., 1999). This 33 kD CENP-H protein is specifically and constitutively localised to centromeres throughout the cell cycle. It contains a coiled coil structure and a clear nuclear localisation signal yet it has only 67% amino acid sequence identity and 75% amino acid similarity to human CENP-H (Sugata et al., 2000). No other CENP-H homologues have been found. CENP-H is required for centromere targeting of CENP-C (Fukagawa et al., 2001) and disruption of CENP-H leads to metaphase arrest (Fukagawa et al., 2001). A more recent study observed misaligned chromosomes, multipolar spindles, and reduced levels of both CENP-C and –E (Orthaus et al., 2006).

1.7.8 CENP-I

CENP-I is a constitutive 87 kD centromere protein that colocalises with CENP-A, -C and –H throughout the cell cycle (Nishihashi et al., 2002). It shows significant

sequence similarity with *S. pombe* Mis6 which is required for Cnp1 localisation to the central core (Takahashi et al., 2000). In the absence of CENP-I cells arrest at prometaphase with misaligned chromosomes. Eventually cells exit mitosis but undergo no cytokinesis. Both CENP-I and CENP-H are necessary for the localisation of CENP-C but not CENP-A (Nishihashi et al., 2002). CENP-I also appears to be necessary for the localisation of CENP-F, MAD1 and MAD2 to the kinetochores (Liu et al., 2003) as depletion leads to a lack of these proteins at the kinetochore.

1.7.9 Chromosomal Passengers

Chromosomal passengers tightly associate with chromosomes early in mitosis, mainly concentrating at the centromere in metaphase. They then abruptly disassociate at the metaphase/anaphase transition and finally concentrate at the midbody at cytokinesis (Mackay et al., 1998). They are involved in coordinating the chromosomal and cytoskeleton events of mitosis (Vagnarelli and Earnshaw, 2004). Known functions include chromatin modification, correction of kinetochore attachment errors, involvement in the spindle assembly checkpoint, spindle assembly and the completion of cytokinesis (Vagnarelli and Earnshaw, 2004).

INCENP is the founding member of the chromosomal passengers and functions as a targeting subunit for a second chromosomal passenger, Aurora B kinase (Kaitna et al., 2000). INCENP binds to noncentromeric sites on the chromosome arms and then localises to the inner centromere during metaphase and the spindle midzone at anaphase. The amino-terminal portion appears to direct transfer to the mitotic spindle at anaphase and the carboxy-terminus appears to be involved in interactions with the

cytoskeleton (Mackay et al., 1998; Mackay et al., 1993). A short conserved stretch of amino acids near the C-terminus bind Aurora B and INCENP is subsequently phosphorylated within this region which enhances Aurora B kinase activity (Vagnarelli and Earnshaw, 2004). Over expression of the amino-terminus causes incompletion in chromosome alignment and a block in cytokinesis by displacing the endogenous protein from centromeres (Mackay et al., 1998). Within the amino-terminus is a 68 aa region which directs the movement of the protein to the centromeres (13 aa) and subsequently, the spindle mid-zone (11 aa) (Ainsztein et al., 1998). Deletion of INCENP in *C. elegans* leads to segregation defects in anaphase and spindle defects in anaphase and telophase (Oegema et al., 2001). In another study, deletion of ICP-1 (Inner Centromere Protein-1), one of the *C. elegans* INCENP homologues, lead to chromosome segregation and cytokinesis defects (Kaitna et al., 2000), remarkably similar to depletion of AIR-2 (Aurora/Ipl1-related-2) an aurora kinase. Deletion in mice reveals embryonic lethality with numerous mitotic faults, including defective chromosome segregation (Cutts et al., 1999). It is also essential in yeast, *Drosophila* and human (Vagnarelli and Earnshaw, 2004). However, its localisation is independent of CENP-A. INCENP interacts with the hinge region of HP1 and appears to be involved in the transfer of INCENP to the anaphase spindle (Ainsztein et al., 1998). A recent paper has implied a functional link between INCENP and CENP-C as localisation of one appears to rely on the presence of the other (Faragher et al., 2007).

Aurora B is related to a serine/threonine protein kinase first identified in *Drosophila*. Mutants displayed monopolar spindles with duplicated centrosomes (Glover et al., 1995). The family is conserved from yeast to humans. There are three aurora

families in mammals (Adams et al., 2001a). Aurora A kinases are involved in spindle dynamics, B kinases in chromosomal events and cytokinesis and C kinases, the role of which is unclear although recent data indicates it as a possible chromosomal passenger protein (Chen et al., 2005). Aurora B has been implicated in many forms of human cancer, correlating with its role in chromosome segregation (Bischoff and Plowman, 1999). It has been shown that INCENP and Aurora B kinase bind to each other directly and that INCENP is required for targeting Aurora B to the centromeres at interphase and the spindle during mitosis (Adams et al., 2001b; Adams et al., 2000). This interaction is evolutionarily conserved. After the metaphase-anaphase transition, this complex recruits centralspindlin (a complex of a kinesin protein ZEN-4/MKLP, which is phosphorylated by Aurora B, and a Rho family GAP) to the midzone, where it bundles microtubules and allows the completion of cytokinesis (Guse et al., 2005). Cells depleted of Aurora B arrest in mitosis with unaligned metaphase chromosomes but have bipolar spindles, suggesting that unlike the *Drosophila* homologue, human Aurora B is not required for centrosome separation (Bischoff and Plowman, 1999). In *S. cerevisiae*, *C. elegans* and *Drosophila*, Aurora B has been shown to phosphorylate H3S10 (Adams et al., 2001b; Hsu et al., 2000) thus implying a role in chromatin structure. In humans, it also phosphorylates H3 Serine-28 (H3S28). It is thought to be necessary for the dissociation of HP1 from chromatin during mitosis (Fischle et al., 2005; Terada, 2006). Phosphorylation of Serine-10 in the amino-terminal tail of H3 is known to be required for condensation and chromosome segregation in *Tetrahymena thermophila* as well as mammalian cells but as yet the relevance of the modification is unclear (Wei et al., 1999). One possibility is the recruitment of condensin during chromosome condensation (Giet and Glover, 2001). Aurora B has been shown to phosphorylate CENP-A with phosphorylation mutants of CENP-A resulting in a delay

in cytokinesis (Zeitlin et al., 2001b). Aurora B may be responsible for maintaining the phosphorylated state of CENP-A during mitosis. Other targets of Aurora B include myosin II, Topo IIα, survivin and MCAK, implicating a role in kinetochore microtubule attachments (Vagnarelli and Earnshaw, 2004). New targets are being discovered rapidly (Faragher et al., 2007).

Survivin is also a chromosomal passenger (Skoufias et al., 2000). It is expressed mainly in G2 and mitosis and has been associated with the regulation of both mitosis and apoptosis as it is a member of the inhibitor of apoptosis protein family (Vagnarelli and Earnshaw, 2004). It first associates with the kinetochores, then the spindle midzone in anaphase and the midbody during cell cleavage, where it is proposed to have an active function (Skoufias et al., 2000). Null mouse embryos have disrupted microtubule formation and become polyploid (Uren et al., 2000). This phenotype and cellular localisation are similar to that of INCENP, indicating that they function in concert (Uren et al., 2000). Phosphorylation of survivin by Aurora B may be important for its function (Wheatley et al., 2004).

Borealin was identified as a fourth protein of the passenger complex and essentially all survivin in mitotic cells is associated with borealin. Its role is unknown but if it is depleted by RNAi, mitosis is delayed and a dramatic increase in kinetochore attachment errors and failures in cytokinesis are observed (Gassmann et al., 2004a).

A recent paper has shown that INCENP appears to stabilise the quaternary complex of INCENP, Aurora B, survivin and borealin, and that borealin acts to promote binding of survivin to INCENP and thus survivin acts as a mediator of centromere and

midbody docking of Aurora B (Vader et al., 2005). INCENP, survivin and borealin have been identified recently in a ternary complex (Klein et al., 2006) with borealin directing the complex to centromeric DNA.

1.7.10 Other Proteins Which Localise to the Centromere

The mammalian checkpoint proteins BUB1, BUBR1, BUB3, MAD1 and MAD2 localise to the kinetochores during mitosis (Pidoux and Allshire, 2000). These spindle checkpoint proteins function to determine that cells which do not have all of their chromosomes attached to the mitotic spindle do not progress to anaphase by sequestering Cdc20, the activator of the APC (Amon, 1999). It is essential for ensuring fidelity in chromosome transmission. Bub1 plays a key role in the assembly of checkpoint proteins at the kinetochore and localises to the kinetochore early in mitosis. It is required for the subsequent localisation of CENP-F, BubR1, CENP-E and MAD2 (Johnson et al., 2004).

The 90 kD microtubule-destabilising kinesin, MCAK is a homodimer which localises to the centromere, between the kinetochore plates, early in mitosis, when the kinetochore plate structure is formed, and persists until telophase (Wordeman and Mitchison, 1995). Functional deficiency of MCAK results in abnormal chromosome segregation (Maney et al., 1998). It may function in microtubule depolymerisation during anaphase. Like CENP-G it localises to normal centromeres, neocentromeres and inactive centromeres (Warburton, 2001). Its premature loss from inactive centromeres suggests that it may be stabilised by active kinetochore proteins such as CENP-H which binds MCAK (Sugata et al., 2000).

The motor protein cytoplasmic dynein (Pfarr et al., 1990; Steuer et al., 1990) and its associated multisubunit complex dynactin, required for spindle assembly and function, localise specifically during mitosis to the outermost corona region of the kinetochore and are likely to be a primary motor for congression (Cleveland et al., 2003). ROD, originally identified in *Drosophila*, and ZW10, associated with functional centromeres, are both required for dynein localisation to the kinetochore and checkpoint function (Chan et al., 2000; Starr et al., 1998). The dynamitin subunit of dynactin interacts with ZW10 in a yeast two-hybrid screen (Starr et al., 1998). Both ZW10 and ROD are localised to the kinetochores in prometaphase, move to the microtubules at metaphase, then back to the kinetochores at anaphase where they remain until the end of telophase (Scaerou et al., 1999; Williams et al., 1996), implying that they function together and are sensitive to tension across the kinetochore and thus are involved in the spindle checkpoint. It is thought that they contribute to checkpoint activation by promoting MAD2 recruitment, and to checkpoint inactivation by recruiting dynein/dynactin that subsequently removes MAD2 from attached kinetochores (Buffin et al., 2005; Kops et al., 2005). In addition, localisation of both ZW10 and dynein to the kinetochore is lost in cells mutant for ROD (Starr et al., 1998). Human ZW10-interacting protein (Hzwint) localises at the kinetochore of active centromeres in a complex which includes Mis12 (Kops et al., 2005) and targets ZW10 to the kinetochore at prometaphase (Starr et al., 2000).

Topo II catalyses the strand passing of double-stranded DNA in an ATP-dependent fashion and thus can remover positive or negative supercoils and can catenate or decatenate DNA duplexes. It has been implicated in DNA replication, transcription,

recombination, chromosome condensation and segregation (Porter and Farr, 2004). It is recruited to centromeres at the onset of heterochromatic condensation during the late S-G2 period where it remains until the completion of cell division (Rattner et al., 1996). Vertebrates have two very similar isoforms (α and β) encoded by distinct genes but only the α isoform is essential for viability and is associated with chromosomes throughout mitosis (Carpenter and Porter, 2004). The Topo II inhibitor ICRF-193 prevents the centromeric accumulation of Topo II and results in the disruption of the kinetochore structure and the disruption of chromatin condensation adjacent to the kinetochore (Rattner et al., 1996). However, an overt role in centromere function still remains unclear. Chromosome segregation is also seriously affected by loss of Topo II expression, the likely cause being residual catenations between sister chromatids (Porter and Farr, 2004). Recent evidence from a cell line in which endogenous Topo IIα genes are disrupted shows Topo IIα is essential for chromosome segregation but not for chromosome condensation (Carpenter and Porter, 2004). Also, recent data has linked SUMO-modified Topo II to the regulation of sister centromere cohesion and thus the separation of centromeres at anaphase (Porter and Farr, 2004). Evidence has recently shown Topo II cleavage activity within the centromeric alpha-satellite DNA arrays of human chromosomes and this has been mapped to specific regions (Spence et al., 2005).

CLIP-170 (Cytoplasmic Linker Protein-170) is required for *in vitro* binding of endocytic vesicles to microtubules. It also transiently associates with the kinetochore and colocalises with dynein and dynactin during mitosis but not on kinetochores already aligned at the metaphase plate. This implies a role in metaphase chromosome alignment (Dujardin et al., 1998). The carboxy-terminus appears to be essential for

kinetochore targeting, possibly mediated by dynein and dynactin (Dujardin et al., 1998). Recent work implies that CLIP-170 facilitates the formation of kinetochore-microtubule attachments, possibly through direct capture of microtubules at the kinetochore (Tanenbaum et al., 2006).

A phosphorylated 3F3/2 antigen transiently locates to the interzone between the inner and outer plates and is proposed to regulate the spindle assembly checkpoint by binding to kinetochores not under tension (Campbell and Gorbsky, 1995). The phosphorylation required for kinetochore localisation of 3F3/2 may be regulated by the MAP kinase pathway (Shapiro et al., 1998). 3F3/2 loading onto kinetochores appears to require the prior assembly of checkpoint proteins such as BUB1 and BUBR1 (Wong and Fang, 2006).

It must be noted that proteins mentioned above are by no means a comprehensive list of centromere-localised proteins. New proteins are being added to this list all the time including histone deacetylases, DNA methyltransferases, transcriptional repressors, chromatin remodellers and histone acetyltransferases (Craig et al., 2003; Liu et al., 2006). Several reviews detailing the possible 3D structure of the centromere and kinetochore can be found (Amor et al., 2004b; Liu et al., 2006).

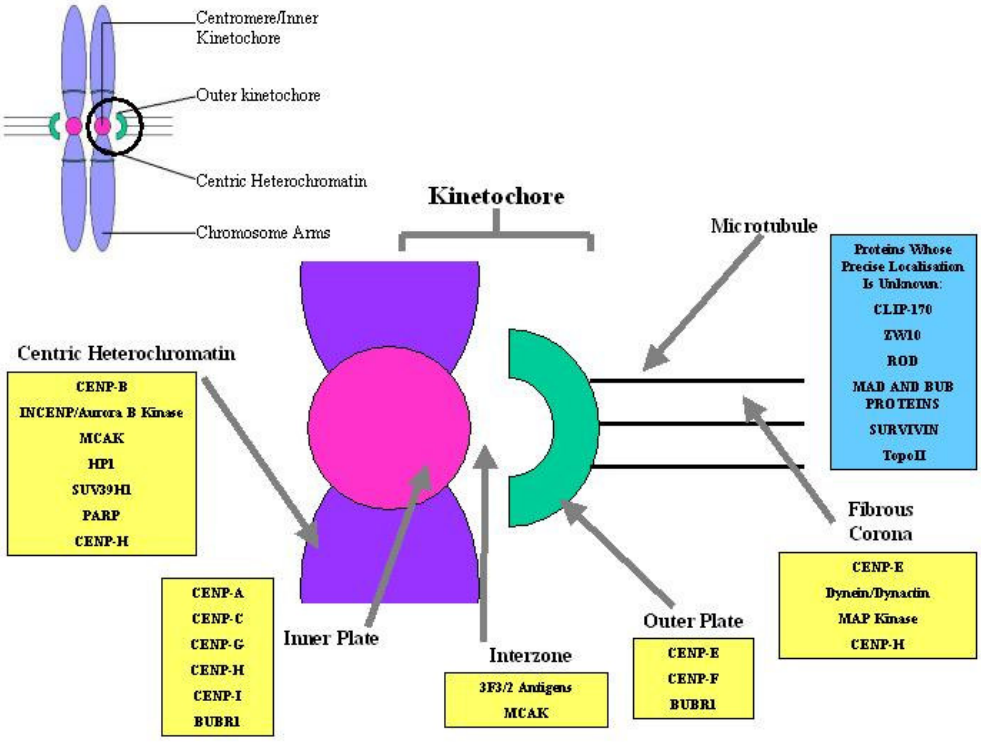


Figure 1.7: Schematic Representation Of Kinetochore Structure and the Location Of Associated Proteins

1.8 Neocentromeres

Neocentromeres are ectopic centromeres that originate occasionally from noncentromeric regions of chromosomes (Amor and Choo, 2002). This rare event was first observed on a marker chromosome of a boy with learning difficulties during routine karyotyping (Voullaire et al., 1993). The marker was designated mardel(10) and was derived from a *de novo* rearrangement of chromosome 10 that resulted in the loss of the original centromere. Although no α-DNA was present at the neocentromere, it was able to form a functional kinetochore and was mitotically stable. Over 60 different neocentromeres have now been reported, typically similar to

mardel(10) and ascertained in a similar way. They typically lack α-DNA, are C-band negative, contain a primary constriction and bind essential centromeric proteins (Saffery et al., 2000). The most common mechanism for neocentromere formation is the *de novo* inverted duplication of a distal chromosome segment, often creating an unbalanced karyotype. Neocentromeres have also been detected in human cancers and produced experimentally in *Drosophila*.

Neocentromeres can form either during mitosis or meiosis and then be transmitted (Depinet et al., 1997). The most notable case of meiotic transmission was of a Y-derived neocentric chromosome in three generations (Tyler-Smith et al., 1999). Neocentromeres most commonly form at distal, euchromatic regions yet heterochromatic proteins such as HP1 and SUV39H1 have been detected at neocentromeres formed at euchromatic regions (Saffery et al., 2000). This implicates a switch from a euchromatic to a heterochromatic state. Several hypotheses for neocentromere formation have been put forward (Amor and Choo, 2002). Neocentromeres may form randomly throughout the genome but only some changes are subsequently stable, or, there are some regions of the genome that favour neocentromere formation. For example, heterochromatin favours CENP-A binding (Henikoff et al., 2000). Another possibility is that neocentromeres share some sequence characteristics with α-DNA which, as discussed above, is a preferred site for centromere formation. Support for this theory comes from studies which have investigated the CENP-A binding domains (Lo et al., 2001a; Lo et al., 2001b). These studies found a high AT content (65% and 60.8%), compared with the genome average of 58%. AT-rich DNA has a narrower minor groove than GC-rich DNA so this feature may be important for the maintenance of centromere structure. Another

study has identified common structural features between the primary sequences of α-DNA, a human neocentromere and the centromere of *S. cervisiae* (Koch, 2000). Despite the absence of α-DNA, protein binding at the neocentromeres has been found to be indistinguishable to that of normal centromeres with the exception of the absence of CENP-B (Saffery et al., 2000). Neocentromeres have different replication timing to that of normal centromeres (Lo et al., 2001a). Mammalian centromeres are known to replicate during the second half of S-phase (Shelby et al., 2000). Lo et al. found that neocentromere formation was accompanied by a delay in replication timing, but not in the CENP-A binding domain, thus demonstrating that neocentromere formation alters replication timing.

Neocentromere formation has been induced experimentally in *Drosophila.* In one study a self-propagating neocentromere, derived from irradiation of a minichromosome derived from the tip of the X chromosome, is described (Williams et al., 1998). The minichromosome contained 290 kb of subtelomeric DNA next to the centromeric DNA. It was then shown that this 290 kb segment can be released from several sites within the *Drosophila* genome but neocentromere formation only occurred when it was released from sites adjacent to the centromeric chromatin (Maggert and Karpen, 2001). This suggests that centromere identity and activity can spread to an unrelated sequence and that neocentromeric activity is not sequence dependant. In a second study, chromosome fragments containing a distal block of heterochromatin from the tip of chromosome 2 had centromeric function (Platero et al., 1999). This argues against the cis-spreading described above and suggests that centromeric determinants are present at all times on this heterochromatin. These sites may normally be repressed by the presence of a more dominant centromere.

Neocentromeres have also been produced in animal cells. A 2.6 Mb minichromosome lacked human α-DNA and mouse minor or major satellite DNA. It consisted of sequences derived from the human Y chromosome and mouse chromosomes 12 and 15 and was stable in chicken DT40 cells (Shen et al., 2001a).

Neocentromere formation is usually detrimental to the individual and is a rare occurrence. Why then does it occur? It is possible that centromere relocation may occur via neocentromere formation as has been indicated in primate studies (Ventura et al., 2001). The X chromosome of three different species had the same order of genetic markers despite the different centromeric locations. The transmission of the neocentric Y chromosome through three generations also supports this theory (Tyler-Smith et al., 1999). It has been proposed that this centromere relocalisation could lead to speciation (Amor et al., 2004a; Henikoff et al., 2001). In one family, centromere repositioning occurred on chromosome 4 due to the inactivation of the original centromere. The formation of a neocentromere at a euchromatic site had little detrimental effect (Amor et al., 2004a).

1.9 Artificial Chromosome Formation

The first artificial chromosomes to be constructed were yeast artificial chromosomes (YACs) in the 1980s by adding the yeast signals for replication, segregation and maintenance of a linear chromosome to a segment of DNA (Murray and Szostak, 1983). The ability to clone large DNA fragments (Burke et al., 1987), up to 10 times greater than with cosmid technology, i.e. 300-1000 kb, into such YAC vectors has facilitated genome analysis. YACs are capable of carrying Mb-sized inserts.

However, they can display a relatively high frequency of rearrangement. The development of bacterial and P1 artificial chromosomes (BACs and PACs) has allowed fragments of several hundred kb to be cloned into single-copy *Escherichia coli* (*E. coli*) vectors which display negligible levels of rearrangement (Ioannou et al., 1994; Shizuya et al., 1992). The ability to clone centromeric DNA into these vectors has largely facilitated the generation of human artificial chromosomes.

In general, there are four approaches for generating human chromosome-based vectors (Basu and Willard, 2005; Grimes and Monaco, 2005; Irvine et al., 2005; Saffery and Choo, 2002):

1. Constructs containing selectable marker genes can integrate into pre-existing centromere regions. This causes chromosome breakage and the formation of large derivative satellite-rich chromosomes (Satellite DNA based Artificial Chromosomes (SATACs) and Artificial Chromosome Engineering System (ACEs)) (Csonka et al., 2000; Hadlaczky et al., 1991; Kereso et al., 1996) (Figure 1.8A).
2. The modification of small human marker chromosomes via irradiation or the transfection of selectable markers (Kuroiwa et al., 2000; Voet et al., 2001) (Figure1.8B).
3. The use of telomeric DNA to sequentially truncate human chromosomes and generate smaller derivative minichromosomes ("top down") (Barnett et al., 1993; Farr et al., 1991) (Figure1.8C). This is known as Telomere Associated Chromosome Fragmentation (TACF) or Telomere Directed Truncation (TDT).
4. The generation of *de novo* Human Artificial Chromosomes (HACs) by introducing cloned DNA into human cultured cells ("bottom up").

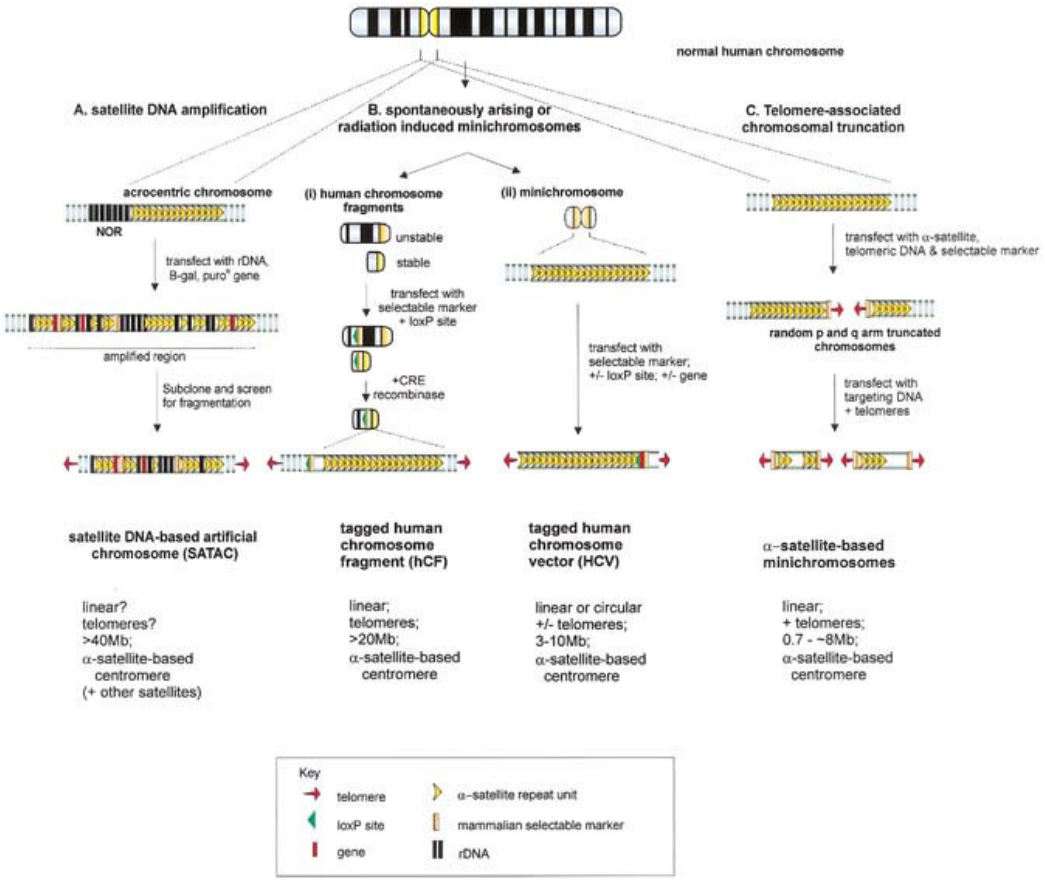


Figure 1.8: Modification of Normal Human Chromosomes and Endogenous Minichromosomes to Produce HACs

A. The integration of a selectable marker and r (recombinant) DNA into the pericentromeric NOR (nucleoli-organizing regions) of an acrocentric chromosome causes amplification. Cells lines in which fragmentation has occurred can be isolated. These will contain SATACs.

B. i) *loxP* sites and selectable markers can be inserted via homologous recombination to allow chromosome fragments to be transferred to small stable vectors via Cre-recombination (Kuroiwa et al., 2000).

ii) Minichromosomes can be modified by the incorporation of selectable markers and/or genes via a variety of approaches: random integration, homologous targeting or Cre-recombination after insertion of a *loxP* site (Voet et al., 2001).

C. Truncation of human chromosomes via the incorporation of telomeric DNA into the chromosomal arms. This seeds new telomeres and removes the distal portion of the chromosome. This is repeated until α-satellite-based linear chromosomes are obtained (Barnett et al., 1993; Farr et al., 1991).

Figure taken from (Saffery and Choo, 2002).

1.9.1 Modification of Normal Human Chromosomes and Endogenous Minichromosomes to Produce HACs

The targeting of heterologous DNA to the pericentromeric region of chromosomes can cause chromosome breakage and the formation of large derivative satellite-rich heterochromatic chromosomes, SATACs and ACEs, typically ten to hundreds of megabases in size, following the amplification of pericentromeric DNA and centromere duplication (Figure 1.8A) (Csonka et al., 2000; Hadlaczky et al., 1991; Kereso et al., 1996). The resulting chromosomes have been generated both in mouse and human cells, are typically mitotically stable (in one case a human SATAC showed a 95% retention rate after 50 generations with continued transgene expression (Csonka et al., 2000)), can be purified by flow sorting (deJong et al., 1999) and have shown to be transmitted successfully down the germline of mice (Co et al., 2000). Bacteriophage integrases which carry out recombination between sites on the phage and bacterial genomes, attP and attB, (Groth et al., 2000) have been used to insert genes into SATACs. The advantage of this system is that it is uni-directional compared with the Cre-recombinase reaction which is reversible. This has allowed the development of the ACE system (Lindenbaum et al., 2004). A platform ACE contains multiple attP integration sites which are targeted via a vector (containing an attB site) and an integrase. Using this strategy high levels of Erythropoietin (EPO) expression have been obtained (Lindenbaum et al., 2004). An intact ACE containing the Red Fluorescent Protein (RFP) gene has been transferred to human mesenchymal stem cells and demonstrated chromosomal maintenance and stable gene expression throughout differentiation (Vanderbyl et al., 2004) thus demonstrating therapeutic potential.

Artificial chromosomes can be produced by modifying spontaneously occurring, or radiation-induced, minichromosomes. Kuroiwa and colleagues modified a chromosome 22 fragment via targeted insertion of a *loxP* cassette. Introduction of this fragment into homologous recombination-proficient DT40 cells, followed by transfection of a Cre-encoding plasmid, catalysed the transfer of the chromosome 22 fragment (10 Mb) into a more stable vector from chromosome 14 (independently tagged with a *loxP* site) therefore generating a mitotically stable minichromosome (Kuroiwa et al., 2000) (Figure1.8B i). Chimeric mice generated from embryonic stem (ES) cells containing this >20 Mb human chromosome fragment (hCF), stably maintained it and the human Ig (immunoglobulin) λ gene on the chromosome 22 derived insert was functionally expressed. hCFs have been shown to be transmitted through the germline (Tomizuka et al., 1997) and have been used to produce transchromosomic mammals. Human/mouse ES cell hybrids containing episomally maintained hCFs including Ig genes were injected into murine embryos and the resulting chimeras exhibited correct expression of the Ig genes. One of these chromosomal fragments was propagated down four generations (Tomizuka et al., 1997). Mice whose endogenous IgH and Igκ loci were inactivated maintained two individual human chromosome fragments containing the IgH and Igκ loci and expressed the genes correctly in response to antigen stimulation (Tomizuka et al., 2000). Following this, a HAC containing the IgH and Igλ loci was incorporated into primary foetal fibroblasts to produce cloned cattle (Kuroiwa et al., 2002). Several calves maintained the artificial chromosome episomally and correctly expressed the genes allowing for the large-scale production of therapeutic human polyclonal antibodies.

Voet et al. targeted selectable markers (Figure1.8B ii) to a circular marker chromosome, isolated from a patient fibroblast line, and transferred it to hamster cells. The ring chromosome was found to be a mixture of alphoid DNA from chromosome 20 and the p arm of chromosome 1 (Voet et al., 2001). The Hypoxanthine Phosphoribosyltransferase (HPRT) gene was then introduced at the *loxP* site via Cre-recombination to produce a human chromosomal vector (HCV). Transfer to ES cells allowed the production of chimeric mice which transmitted the HCV through the germline (Voet et al., 2001). Subsequently a PAC containing 62.5 kb of genomic DNA was inserted into the *loxP* site. Efficient germline transmission and mitotic stability were shown (Voet et al., 2003) indicating no influence of the large insert on stability. In another study, a *loxP* site was inserted into a chromosome 9-derived minichromosome using a co-transfection strategy (Moralli et al., 2001). A histone H2B-GFP (Green Fluorescent Protein) reporter gene was introduced to the *loxP* site and was stably expressed.

Minichromosomes can arise as a result of chromosomal fragmentation by radiation. A 5.5 Mb minichromosome derived from chromosome 1 was obtained after irradiation of a CHO-human hybrid (Guiducci et al., 1999). It was tagged with the neo and IL (interleukin)-2 genes and allowed IL-2 dependent growth in mouse cells. Analysis of this minichromosome revealed vast alphoid DNA rearrangements (Auriche et al., 2001). This defined minichromosome has been used to clone the Cystic Fibrosis Trans-Regulator (CFTR) locus by targeting a 320 kb YAC containing the CFTR gene and upstream sequences to the pericentromeric region (Auriche et al., 2002). One of three positive clones generated showed regulated CFTR expression.

It has long been know that in *S. cerivisiae* transformation of telomeric sequences can cause chromosome fragmentation (Surosky and Tye, 1985). This has proven to be a vital tool for physical mapping in yeast (Vollrath et al., 1988) and in YACs (Pavan et al., 1990) and has allowed the chromosomal position to be determined in terms of the physical distance between a gene and each of the telomeric ends of a linear chromosomal molecule. It was shown in the early 1990s that cloned telomeric DNA can fragment mammalian chromosomes (Farr et al., 1991), the combined use of gene–targeting via homologous recombination and TACF can generate a defined human chromosome truncation (Itzhaki et al., 1992), and that cloned human telomeric DNA can integrate into mammalian chromosomes and seed the formation of new telomeres, therefore truncating the chromosome arm (Barnett et al., 1993) (Figure1.8C). This "top down" method is known as TACF or TDT. These techniques have been used to generate human X-derived minichromosomes ranging in size from 2.4 to ~8 Mb (Farr et al., 1995; Mills et al., 1999). Heller et al. have also used sequential truncations to generate minichromosomes from the human Y chromosome ranging in size from 4-9 Mb (Heller et al., 1996). Subsequently one of these minichromosomes was dissected further in DT40 cells and indicated that a functional centromere is maintained by an activity which functions in trans and is sensitive to the amount of alphoid DNA (Yang et al., 2000). A chromosome 21-derived minichromosome constructed in DT40 cells by TDT demonstrated Enhanced Green Fluorescent Protein (EGFP) expression of the EGFP gene targeted to an inserted *loxP* site (Katoh et al., 2004). Subsequently, the EPO expression cassette was targeted to this *loxP* site (Kakeda et al., 2005). The resulting minichromosome was introduced into human primary fibroblasts by Microcell-mediated Chromosome Transfer (MMCT) where it was mitotically stable and EPO expression was maintained for three months. This transgene expression in

primary cells from a minichromosome is an exciting development in the progress towards the use of artificial chromosomes as a therapeutic treatment. Furthermore, the 21-derived minichromosome containing an EGFP reporter gene and G418-resistance gene has now been introduced into human bone marrow and cord blood derived haematopoetic stem cells by MMCT. The minichromosome containing haematopoietic cells showed resistance to G418, expressed EGFP, and retained the ability to differentiate into various lineages where the minichromosome was maintained (Yamada et al., 2006).

Another approach to the generation of artificial chromosomes has been to modify a neocentromere. TACF has been used to remove the long and short arms of the mardel(10) chromosome (Voullaire et al., 1993) surrounding the neocentromere in HT1080 cells (Saffery et al., 2001). This resulted in neocentromere-based minichromosomes (NC-MiCs) either linear or circular in nature, ranging from 650 kb to 1.8 Mb that associated with essential centromere antigens and were mitotically stable. Further analysis revealed the 650 kb NC-MiC to be composed of a 330 kb stretch containing the neocentromere and 320 kb of flanking DNA (Wong et al., 2002).

1.9.2 *De Novo* HAC Formation

This bottom up approach (Figure 1.9) generates artificial chromosomes ranging in size from 1-10 Mb by introducing cloned centromeric DNA into human cultured cells. They are believed to be circular in structure, consist of multiple copies of the input DNA, which may rearrange, and they are typically present at 1-2 copies per cell. Generally they are analysed by checking for their input DNA components, a *de novo* centromere and their mitotic stability when grown in culture in the absence of selective pressure. Harrington et al. first used this method in 1997. They introduced synthetic concatemerised alpha satellite arrays into human cells with Polymerase Chain Reaction (PCR)-generated human telomeric DNA and randomly sheared human genomic DNA and were able to observe and characterise a mitotically stable *de novo* artificial chromosome in one transfected cell line (Harrington et al., 1997). The resulting HAC was between 6 and 10 Mb, much larger than the input DNA. How this amplification occurs is unclear. Subsequent studies have formed HACs following transfection of a YAC (Henning et al., 1999; Ikeno et al., 1998), BAC (Mejía et al., 2002) or PAC (Ebersole et al., 2000) containing alphoid DNA (for reviews see (Conese et al., 2004; Larin and Mejía, 2002; Saffery and Choo, 2002).

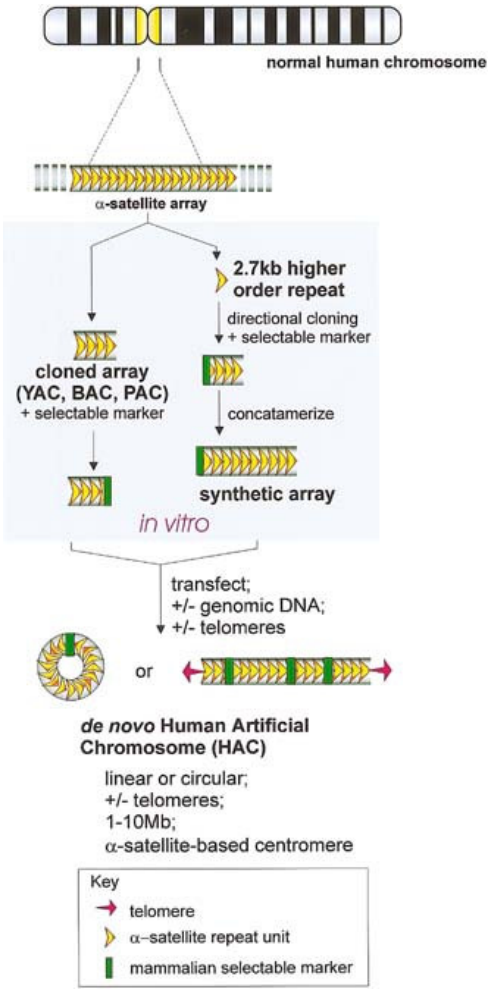


Figure 1.9: Formation of *De Novo* HACs
Alphoid arrays can be cloned in YACs, BACs or PACs (Ikeno et al., 1998; Mejía et al., 2002) or they can be synthesised as large concatemers (Harrington et al., 1997). These are then transfected into cells (with/without telomeres and genomic DNA). The resulting HACs can be linear, or more commonly circular.

Figure taken from (Saffery and Choo, 2002).

1.10 HAC-based studies

1.10.1 Are telomeres necessary?

Early studies engineered mammalian telomere sequences into HAC vectors (Henning et al., 1999; Ikeno et al., 1998). However, telomeres are not necessary for HAC formation. The construction of mitotically stable HACs has been demonstrated using circular α-DNA which does not contain telomere repeats (Ebersole et al., 2000; Grimes et al., 2002; Mejía et al., 2002; Ohzeki et al., 2002; Rudd et al., 2003; Schueler et al., 2001). HAC formation occurred at a high efficiency in HT1080 cells using PAC based vectors containing chromosome 21 alphoid DNA in either a linear or circular format. Circular DNA formed HACs with or without input telomeric DNA. However, the removal of telomeres from linear constructs was severely detrimental to HAC formation (Ebersole et al., 2000). Despite this, none of the HACs tested positive for telomeres using Fluorescent in situ Hybridisation (FISH), suggesting that the requirement for telomeres on the linear construct may only be necessary for initial stabilisation, or that the HACs contained very short, telomeres undetectable by FISH (Ebersole et al., 2000).

1.10.2 The effect of CENP-B boxes

YACs containing chromosome 21-derived alphoid DNA lacking CENP-B boxes are unable to seed *de novo* HAC formation (Ikeno et al., 1998). One study synthesised a series of alpha satellite DNA constructs from chromosome 21 in which the CENP-B boxes were either left intact or mutated (Ohzeki et al., 2002). HACs only formed from alpha satellite DNA containing intact CENP-B boxes. A further study

demonstrated that the addition of more CENP-B boxes into alphoid arrays (in 16/16 monomers as opposed to the naturally occurring 5/16 monomers) may improve the efficiency of HAC formation (Basu et al., 2005). This work has been supported by a further recent study in which a low density of CENP-B boxes did not support HAC formation (Okamoto et al., 2007). However, artificial chromosomes have been created using chromosome 22-derived alphoid DNA consisting of higher order repeat monomers devoid of CENP-B boxes (Kouprina et al., 2003), suggesting that CENP-B might enhance the generation of *de novo* HACs but is not absolutely required.

1.10.3 Variable efficiency of centromere formation

Human artificial chromosome formation, through the "bottom up" approach, is only associated with higher order alphoid sequences (Ebersole et al., 2000; Grimes et al., 2002; Harrington et al., 1997; Henning et al., 1999; Ikeno et al., 1998; Mejía et al., 2002; Ohzeki et al., 2002; Rudd et al., 2003; Schueler et al., 2001; Yoda et al., 1998) and the CENP-B (Centromere Protein-B) box appears only in type-I alphoid sequences of autosomes and X chromosomes (Ikeno et al., 1994; Masumoto et al., 1989).

The length of the input alphoid DNA influences the HAC formation process. Grimes et al. found that decreasing the 17 α-satellite input length from 80 kb to 35 kb lowered HAC formation. They also demonstrated that the selectable marker genes and /or vector sequences influence HAC formation (Grimes et al., 2002). In a recent study, HACs with 30 kb of alphoid DNA were as stable as HACs with longer alphoid DNA regions (Okamoto et al., 2007). Alphoid arrays smaller than this (10 kb) failed to

form functional centromeres and it was proposed that this is due to a preferential tendency of these arrays to form H3K9Me$_3$ heterochromatin. The study also demonstrated that CENP-A chromatin can spread into flanking vector DNA. If the length of alphoid DNA is insufficient, this may be required to support centromere function (Okamoto et al., 2007). A further study has suggested that the epigenetic assembly of heterochromatin is distinct from centromere chromatin assembly and is required for stable HAC formation (Nakashima et al., 2005).

The actual size of the resultant artificial chromosome may influence stability. YACs containing approximately 100 kb of chromosome 21-derived alphoid DNA retrofitted with human telomeric DNA and selectable markers were transfected into HT1080 cells (Ikeno et al., 1998). Chromosomes of 2.5 –5 Mb showed up to 99.5% stability whereas smaller 1-2 Mb HACs had over 1% loss rates per generation. This suggests a decrease in mitotic stability with a decrease in HAC size.

Different alphoid sequences have different HAC forming efficiencies. Constructs based on 17 α-satellite DNA form HACs at a much greater efficiency (in up to 79% of clones) than do Y-based α-satellite constructs (in less than 4% of clones) (Grimes et al., 2002; Mejía et al., 2002). 17 and Y α-satellite DNA are not identical. Firstly, they belong to different suprachromosomal families. Secondly, 17 α-DNA has a 16-monomer based higher-order repeat structure of 2.7 kb (Waye and Willard, 1986) whilst Y α-DNA has a 33-monomer based arrangement totalling 5.7 kb (Tyler-Smith and Brown, 1987). Thirdly, 17 α-DNA contains 6 CENP-B boxes per repeat (Warburton et al., 1993) whilst Y α-DNA has none.

1.10.4 *De Novo* HAC Analysis

In addition to the confirmation of CENP-A deposition on HACs (Grimes et al., 2002; Grimes et al., 2001; Ikeno et al., 2002; Ohzeki et al., 2002), indicating a functional centromere and substantiating its role in kinetochore formation; there is evidence that de novo HACs contain both heterochromatin (H3K9Me marks and the presence of HP1), essential for centromere function, and transcriptionally permissive euchromatin ($H3K4Me_2$ marks), a requirement for gene expression (Grimes et al., 2004). The level of heterochromatin varied, with it being enriched in artificial chromosomes estimated to be larger than 3 Mb in size, replicating late in S phase, and depleted on those smaller than 3 Mb, replicating earlier in S phase, characteristic of expressed sequences. This suggests that only a small amount of heterochromatin may be required for centromere function (Grimes et al., 2004). Further evidence that kinetochores can form within open chromatin regions comes from another study where $H3K4Me_2$ modifications were present within the kinetochore (Sullivan and Karpen, 2004), from the expression of genes located within the kinetochore of human neocentromeres and a rice centromere (Cooke, 2004; Nagaki et al., 2004; Saffery et al., 2003) and from the localisation of CENP-A to gene expressing regions (Lam et al., 2006). This study also revealed that $H3K4Me_2$ decreased in response to CENP-A over expression, possibly indicating a role for $H3K4Me_2$ as a regional boundary, separating centromeric chromatin from constitutive heterochromatin (Lam et al., 2006).

Multiple factors influence the mitotic stability of artificial chromosomes. For example: epigenetics, the cell/genetic background, the sequence composition and the chromosome size may exert an effect. Although artificial chromosomes may be stable

in culture, they exhibit more segregation errors than natural human chromosomes (Rudd et al., 2003). In instances where *de novo* HACs have been transferred into mouse cells using MMCT (Alazami et al., 2004; Ikeno et al., 2002), the stability of HACs in the absence of selection indicated moderate to high loss rates (1-5%/day).

1.10.5 Transgene expression from *de novo* HACs

Previous work in our lab has demonstrated the efficacy of *de novo* HACs for delivery and expression of human transgenes in human cell lines by complementation of the HPRT deficiency in human cells using a HAC vector containing chromosome 17 α DNA, telomeres and a large genomic segment from human Xq26.2 spanning the HPRT locus (Mejía and Larin, 2000; Mejía et al., 2001). HPRT encodes a ubiquitously expressed purine-salvage enzyme and mutations in this gene lead to a build up of uric acid and subsequently Lesch-Nyhan syndrome, a neurodevelopmental disorder (Caskey and Kruh, 1979). This was the first documented report of the complementation of a gene deficiency leading to a genetic disorder using *de novo* artificial chromosomes. Analysis of these HAC fibres showed irregular and alternating 17 α and HPRT DNA implying amplification of the input DNA. Another study has also reported expression of the HPRT gene using a co-transfection assay with one PAC containing 21 α-DNA and a second PAC containing the HPRT gene (Grimes et al., 2001). HACs were estimated to contain 1-2 copies of the HPRT gene. In a similar co-transfection study two BACs, one containing 21 α-DNA and one containing the GCH1 (Guanosine triphosphate Cyclohydrolase-1) genomic locus (involved in the synthesis of tetrahydrobiopterin, the essential co-factor for aromatic amino acid hydroxylases and nitric oxide synthase) generated HACs in HT1080 cells

with multiple copies of GCH1 (Ikeno et al., 2002). These HACs were sensitive to interferon (IFN)-γ induction, replicating the response of the gene expression from the endogenous chromosome genes. Recently, HPRT-expressing HACs have been generated from an α-DNA BAC linked to a BAC carrying the HPRT gene by Red homologous recombination (Kotzamanis et al., 2005).

1.11 The advantages of using HACs in Gene Therapy Studies

The use of viral vectors for gene therapy has had a turbulent history, despite their relatively high levels of gene expression and efficiency of transfer. The most well known set back to viral gene therapy occurred in 1999 with the death of patient Jesse Gelsinger caused by multi organ failure after administration of an adenovirus vector (Marshall, 1999; Thomas et al., 2003). The first successful gene therapy report was published in the year 2000. Three children were cured of the fatal immunodeficiency disease Severe Combined Immunodeficiency-X1 (SCID-X1) by *ex vivo* transduction of their haematopoietic stem cells with a retroviral vector (Cavazzana-Calvo et al., 2000). However, this therapy has subsequently caused a leukaemia-like disease in two of the eleven patients that were treated (Hacein-Bey-Abina et al., 2003). This was due to retroviral integration in proximity to the LMO2 proto-oncogene (a central regulator of haematopoiesis) promoter, leading to deregulated transcription and expression of LMO2.

Although these problems of insertional mutagenesis and immunogenicity are being addressed in new generation viral vectors (for reviews see (Thomas et al., 2003; Verma and Weitzman, 2004), HACs can avoid these problems commonly

encountered with viral systems and they offer a viable alternative to viral vectors, including episomal viral-based vectors which require the constitutive expression of EBNA-1 (Epstein Barr Nuclear Antigen-1), a transactivator protein from Epstein Barr Virus (EBV), or other viral proteins (Huertas et al., 2000; Wade-Martins et al., 2000). The stability of EBV vectors is lower than that of endogenous chromosomes, however, gene expression has been sustained in culture over several months (Wade-Martins et al., 1999; Wade-Martins et al., 2000) and EBV episomes containing the HPRT locus, β-globin locus and the β-globin and Factor IX minigenes have been successful in gene complementation (Basu and Willard, 2005). As HACs utilise human functional components for their replication and stable propagation, it is not necessary for them to integrate into the host genome. This ensures that the introduced genes are neither interfered with or affect the host cell genome. Therefore, there is no potential for insertional mutagenesis or transgene silencing. Virus vectors are limited by the space available for exogenous DNA. HACs are capable of carrying entire genes and their regulatory elements (Grimes et al., 2001; Mejía et al., 2001) due to their large cloning capacity, since, unlike viral vectors, they have no space limitations. For example, adeno-associated viruses only have a packaging capacity of 5 kb (Thomas et al., 2003). The extra regulatory elements which HACs are able to incorporate, may allow tissue specific expression and physiological gene expression. HACs derived entirely from human DNA pose no risk to the immune system as no viral elements are involved.

Despite the studies above which describe the value of HACs as gene transfer vectors, there remain several issues to be addressed. One such important issue is the efficient delivery of HAC vectors into target cells.

1.12 Developing an Efficient Delivery System – An Infectious Herpes Simplex Virus-1 System for HAC Vector Delivery

Herpes Simplex Virus-1 (HSV-1) infects mucosal surfaces and establishes latent infection in sensory neurons. The 152 kb genome consists of approximately 80 genes and only two non-coding regions are required to package the DNA into infectious virions (Hibbitt and Wade-Martins, 2006). Plasmids that contain these elements, the origin of replication, *ori*$_S$, and the packaging signal, *pac*, are incorporated into HSV virions as concatemers. Current transfection efficiencies of HAC constructs into HT1080 cells using cationic lipids are approximately 10^{-6}. This method also requires the condensation of large DNA to prevent shearing and degradation (Marschall et al., 1999). Therefore, our group began investigating a more efficient delivery system using a HSV-1 system.

The helper-virus free system has undergone several rounds of development (Sena-Esteves et al., 2000). Initially, the HSV-1 genome was provided in a set of five overlapping cosmids, but a low packaging efficiency was observed due to transfection difficulties (Cunningham and Davison, 1993). The HSV-1 genome was then localised on a single BAC lacking the *pac* signal to enable its packaging (Saeki et al., 1998). The most recently developed system was from Saeki et al. (Saeki et al., 2001). It uses an oversized, ICP27-deleted BAC (the ICP27 gene is essential for viral replication) to provide the necessary genes for packaging amplicon DNA, but does not itself get packaged. This was further developed by Wade-Martins et al. for BAC delivery (Wade-Martins et al., 2001).

This system allows the delivery of large genomic fragments (150 kb) into cells, without immunogenic effects, using HSV-1 amplicons in a helper virus free system (Figure 1.10). We have developed this system for HAC delivery due to the high efficiency (up to 100%) the system has shown in delivering BACs and PACs to human and mouse cells (Wade-Martins et al., 2001) and it could accommodate large sized alphoid arrays with relative ease. The system has also been used to deliver genomic BACs containing genes for complementing deficiencies, indicating another advantage for generating therapeutic HACs. The HPRT locus contained within a BAC was expressed when delivered by transduction to both a human HPRT-deficient MRC5-V2 cell line and a mouse primary hepatocyte culture derived from $Hprt^{-/-}$ mice (Wade-Martins et al., 2001). This system was also able to generate expression of the human Low-density Lipoprotein Receptor (LDLR) from a BAC delivered by HSV-1 amplicons, at physiologically appropriate levels in both CHO cells deficient in the LDLR receptor and in human fibroblasts derived from familial hypercholesterolemia patients who have a mutation in the LDLR receptor (Wade-Martins et al., 2003). The large genome of HSV-1, at 152 kb, allows transgene capacity large enough to accommodate approximately 95% of human genomic loci (Hibbitt and Wade-Martins, 2006). Further studies have highlighted the use of this system in delivering a genomic locus which retains its correct expression and splicing profile and the potential of the HSV-1 amplicons in the study of polygenic disorders due to the physiological expression from the delivered locus (Hibbitt and Wade-Martins, 2006).

Amplicons are generated by co-transfecting, into African Green Monkey Cells, (Vero 2-2) the desired BAC which has been modified by Cre-*loxP* recombination to contain the only two non-coding regions of DNA sequence required to facilitate the packaging of DNA into infectious virions: the HSV-1 packaging/cleavage signals (*pac*) and

origin of replication (*ori*$_S$) together with the helper BAC fHSVΔpacΔ270+ (containing the HSV-1 genome) and plasmid pEBHICP27 (Figure 1.10). All the enzymes and proteins required to replicate and package the desired BAC are provided in trans by fHSVΔpacΔ270+ and pEBHICP27. The HSV-1 genome (152 kb) is present as an over-sized BAC (178 kb) which lacks the *pac* signals and a functional copy of the immediate early gene ICP27 which is essential for viral replication, and instead contains "stuffer" sequences to increase its size to 178 kb (Saeki et al., 2001). These "stuffer" sequences increase the size of the HSV-1 BAC over the packaging capacity of the HSV-1 virions thus virtually eliminating the probability of replication-competent helper virus contamination (Saeki et al., 2001). This is obviously important for a gene therapy approach. HSV-1 recognises the *ori*$_S$ and *pac* signals and results in the incorporation of the plasmid containing such signals into HSV-1 virions in concatemers of up to 150 kb. After packaging, the amplicons are purified from the cells, concentrated and titred in glioma cells (G16-9). These cells are highly infectable and the number of transducing units per ml can be determined. Approximately 10^7 to 10^8 units/ml can be achieved. Results from various groups indicate that the immune response in the helper-virus free system is minimal (Costantini et al., 1999; Olschowka et al., 2003). This is vital for gene therapy. Other advantages of this system are the broad cell host range (Hibbitt and Wade-Martins, 2006) as illustrated by our group (Moralli et al., 2006), the ability to infect dividing and non-dividing cells and the transgene capacity of up to 150 kb (Sena-Esteves et al., 2000).

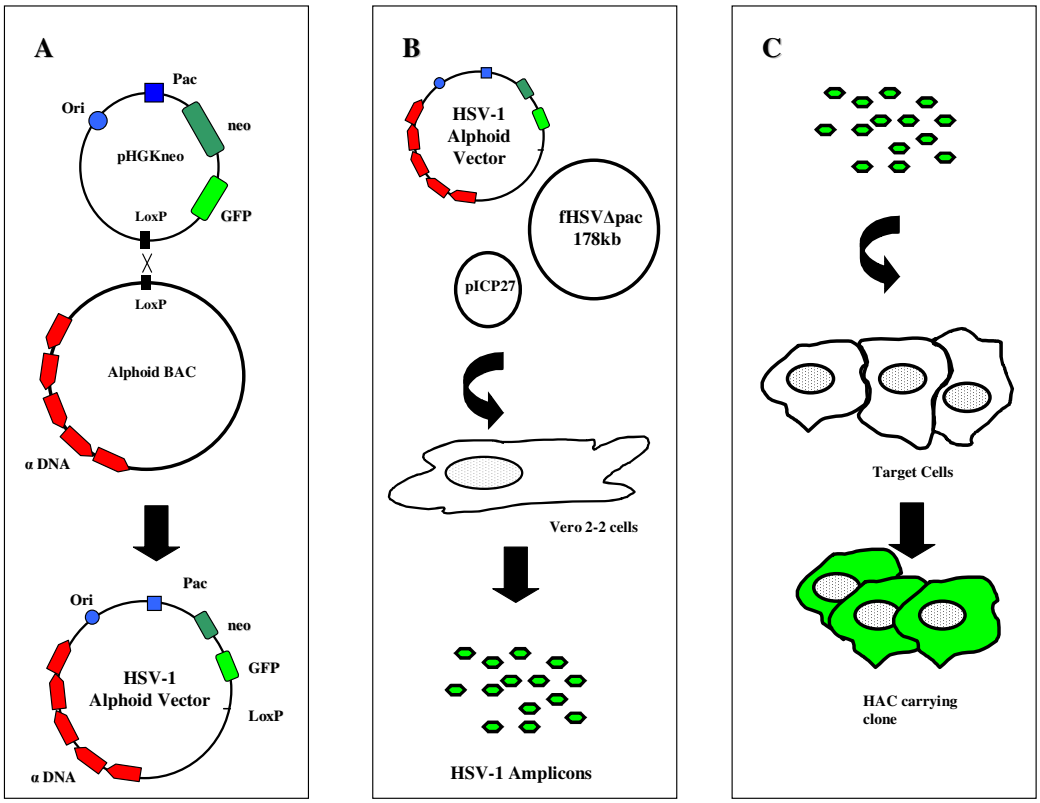


Figure 1.10: Using HSV-1 To Generate HACs In Mammalian Cells

A The desired alphoid BAC is modified to contain the HSV-1 *pac* and origin of replication by Cre-loxP recombination.

B This HSV-1 alphoid vector is co-transfected with the helper plasmids fHSVΔpacΔ270+ (the HSV-1 BAC) and pEBHICP27 into African Green Monkey Cells (Vero 2-2). Amplicons are generated, harvested and titred.

C Target cells are infected with the amplicons. Selection is applied and GFP positive, HAC containing, clones are isolated.

1.13 Project Outline

The overall aim of this project was to develop a novel HAC vector system for gene expression studies in human cells. Whilst HACs have been confirmed as viable gene expression vectors (Grimes et al., 2001; Ikeno et al., 2002; Kotzamanis et al., 2005; Mejía et al., 2001), current transfection efficiencies of HAC constructs into HT1080 cells using cationic lipids are poor at approximately 10^{-6}, and in addition, condensation of the DNA is required to prevent shearing and degradation (Marschall et al., 1999). HAC delivery has posed a barrier to the development of HACs as gene transfer vectors. The aim of the first stage of this work (discussed in Chapter 3) was to develop a novel and efficient a novel HAC vector system for gene delivery into human cells using HSV-1 amplicon technology. This allows the delivery of large genomic fragments (150 kb) into a range of cell types, without immunogenic effects, using HSV-1 amplicons in a helper virus free system.

HACs were generated with a reasonable efficiency in human cell types other than HT1080, which are commonly used for HAC studies, compared to previous delivery methods used. The transduction efficiency, determined by transient expression of the reporter gene, GFP, was several orders of magnitude higher than lipofection, illustrating that HSV-1 amplicons are a valuable system for HAC vector delivery. HACs in HT1080 and G16-9 cells were stable up to 90 days in the absence of selection. However, in MRC5-V2 and 293 cells, the HACs were unstable suggesting that the requirements for centromere function and HAC stability are different in each cell type.

The second stage of this study (discussed in Chapter 4) investigated this further. Real-time PCR analysis showed that Aurora B was over expressed in MRC5-V2 and 293 cells. High levels of chromosomal instability in MRC5-V2 and 293 cells correlated with high levels of Aurora B and Topo II as confirmed by western blot analysis. Since Aurora B is a kinase involved in at least two cell cycle checkpoints, cellular phosphorylation levels were perturbed to mimic what was observed in 293 and MRC5-V2 cells, using Okadaic acid (OA), which is both a protein phosphatase (PP) inhibitor and activates Aurora B. Treatment of cells showed an increase in both HAC and overall chromosomal instability and an increase in histone H3 Serine 10 and Serine 28 phosphorylation.

HACs have become an important gene vector transfer system for non-viral delivery into human cells and the third stage of this study (discussed in Chapter 5) involved the development of a gene expression system using HSV-1 amplicons. Firstly this involved engineering the HPRT genomic locus into an HSV-HAC vector, the 65 kb pHSV17α302Neo BAC, by Red mediated recombination for complementing the HPRT deficiency in HPRT- HT1080 cells. Secondly, a co-infection approach with the HSV-1 system, one vector containing alpha satellite DNA and the other the genomic locus of interest was studied. Preliminary experiments were successful and further work to develop HSV-HAC amplicons will hopefully enable this approach to generate artificial chromosomes in cells which can complement deficiencies in a similar way to previous co-transfection studies (Grimes et al., 2001; Ikeno et al., 2002).

Chapter 2: Materials and Methods

2.1 Suppliers

Unless otherwise stated, the chemicals and solvents used in experiments were obtained from Sigma, the tissue culture and western blotting reagents were from Invitrogen, all DNA restriction enzymes were from New England Biolabs (NEB), the bacterial media capsules were from BIO101, the quantitative real time PCR reagents were from Applied Biosystems (ABI) and the antibodies were from Abcam. Oligonucleotides were synthesised by MWG.

2.2 Solutions and Media

Dot Blot Buffer

25 millimolar (mM) Tris-HCl (pH 8.0), 1mM EDTA (pH 8.0), 150 mM NaCl.

KCM Buffer

120 mM KCl, 0.02 Molar (M) NaCl, 0.01 M Tris-HCl (pH 8.0), 0.5 M EDTA (pH 8.0), 0.1% Triton X-100.

6x Loading Dye

0.03% bromophenol blue, 0.4% orangeG, 15% Ficoll-400, 10 mM Tris-HCl (pH 7.5), 50 mM EDTA (pH 8).

Luria – Bertani Broth (LB)

LB medium and agar were both prepared using BIO101's capsules in accordance with the manufacturer's protocol.

10x Northern Gel Loading Dye

5% glycerol, 0.02% xylene cyanol, 0.2% bromophenol blue in 1x Northern Gel Running Buffer.

10x Northern Gel Running Buffer

0.2 M MOPS, 50 mM sodium acetate, 10 mM EDTA (pH 8).

SOC

SOB medium (Hanahan's Broth) (Sigma) was prepared according to the manufacturer's instructions and then glucose was added to a final concentration of 20 mM.

20x SSC

3 M NaCl, 0.3 M $C_6H_5Na_3O*2H_2O$ (trisodium citrate 2 hydrate).

STET Buffer

8% sucrose, 5% Triton X-100, 50 mM EDTA, 50 mM Tris-HCl (pH 8.0).

Alkaline SDS Buffer

1% SDS, 0.2 M NaOH.

50x TAE

40 mM Tris base, 20 mM glacial acetic acid, 1 mM EDTA (pH 8.0).

5x TBE

450 mM Tris base, 450 mM Orthoboric Acid, 10 mM EDTA (pH 8.0).

TBS

10 mM Tris-HCl (pH 8.0), 150 mM NaCl.

TE

10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0).

TE 50:10

50 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0).

TEEN buffer

1 mM Triethanolamine-HCl (pH 8.5), 0.2 mM EDTA, 25 mM NaCl.

10x TNE

100 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0), 2 M NaCl.

2.3 The Polymerase Chain Reaction

The PCR primers used in this study are listed in Table 2.1. The concentration of $MgCl_2$ used in each reaction was 1.5 mM and reactions were carried out in a total volume of 25 microlitres (µl) using 1 Unit (U) of Taq polymerase, 0.5 micromolar (µM) of each of the forward and reverse primers and 1x Amplitaq DNA polymerase Buffer II from ABI. Cycle conditions were:

95˚C 5 minutes

94˚C 30 seconds
59˚C 50 seconds
72˚C 1 minute
} 25x

72˚C 10 minutes

Reactions were carried out on a PTC-200 thermal cycler from MJ Research.

Table 2.1: Nucleotide Sequences and PCR Amplification Conditions for the Primers Utilised

Primer Name	Primer Sequence (5' to 3')	Annealing Temp (˚C)	Extension Time
Neo2 Forward	TTGCGTGGAGAGGCTATTCG	59	1 min
Neo2 Reverse	CCATTCGCCGCCAAGCTCTT		
CENP-B Box Forward	CGGGTATATCTTCACATGAC	50	1 min
CENP-B Box Reverse	CGAACGAAGGACACAGAGTG		

2.3.1 Harvesting Mammalian Cells for the Polymerase Chain Reaction (PCR)

Cells were harvested by trypsinisation, and pelleted by centrifugation at 1200 rotations per minute (rpm) and the pellet was washed in PBS. The cells were counted and resuspended in 200 µl/10^6 cells in lysis buffer (50 mM KCl, 10 mM Tris-HCl (pH 8.0), 0.5% Tween 20, 10 micrograms (µg)/ml Proteinase K). Samples were incubated at 55˚C for one hour followed by 95˚C for ten minutes before 1 µl was used in a PCR reaction.

2.4 Bacterial Strains

Table 2.2: Bacterial Strains Utilised in this Study

Strain	Genotype
DH5α	F- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (r_k-, m_k+) *pho*A *sup*E44 λ- *thi*-1 *gyr*A96 *rel*A1
DH5αpir	As DH5α but also containing the pir gene.
DH10B	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara, leu*)7697 *gal*U *gal*K λ- *rps*L *nup*G

2.5 Vectors

Table 2.3: Vectors Utilised in this Study

Vector	Features
pUC19	A small, high copy number *E. coli* plasmid cloning vector. 2.7 kb, pUC ori, ampicillin resistance (Amp^R), *lac*Z
pBeloBAC11	A mini-F factor based 7.5 kb vector constructed by (Kim et al., 1996), for generating bacterial artificial chromosomes. It contains the gene for chloramphenicol resistance (Cm^R), a bacterial origin of replication (Ori_S, RepE), copy number control genes (ParA, ParB), a *loxP* site, the *lacZ* gene containing unique BamHI and HindIII sites and T7/SP6 promoters.
227J24,15K01201B, 302N02, 292P20 and 302P22	BACs (Cm^R) based on pBeloBAC11 with 17 alphoid DNA inserts ranging from 50-150 kb (227J24, 130kb; 15K01201B, 150kb; 302N02, 90kb; 292P20, 110kb; 302P22, 50kb).
pWTR11.32	A 67 kb BAC vector constructed by (Ohzeki et al., 2002), containing kanamycin resistance (Kan^R), neomycin resistance (Neo^R) under the control of the SV40 promoter and a 60 kb 21 alphoid DNA array.
pCTP-T	A small plasmid vector developed for Cre-mediated recombination containing the pSC101 temperature-sensitive replication origin, tetracycline resistance gene, the *pir* gene and the Cre recombinase gene under the control of the tetracycline response element (Saeki et al., 2001; Wade-Martins et al., 2001).
pSG80A-HG	A 6 kb plasmid vector developed for generating HSV-1 amplicons, containing the HSV-1 amplicon elements (*ori*$_S$ and *pac*), the GFP gene, Amp^R, the R6Kγ bacterial replication origin and a *loxP* site (Wade-Martins et al., 2003).
fHSVΔpacΔ270+	A 178 kb PAC for generating HSV-1 amplicons, containing the HSV-1 genome without *ori*$_S$ and *pac,* Cm^R, Kan^R. It is deleted for ICP27 and has two additional copies of the ICP0 gene which act as a "stuffer" sequence (Saeki et al., 2001).
pEBHICP27	A 12.2 kb plasmid vector which expresses ICP27 with its own promoter and polyadenylation signal as well as EBNA1 and *oriP* from EBV, Amp^R, the R6Kγ bacterial replication origin and a *loxP* site, developed for HSV-1 amplicon generation (Saeki et al., 2001).

2.6 DNA Manipulations

2.6.1 DNA Purification from Protein Contaminated Samples

When necessary, DNA that was contaminated by protein was purified by extracting twice with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) and once with an equal volume of chloroform:isoamyl alcohol. DNA was precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol. The sample was left on dry ice for 15 minutes, spun at 13000 rpm for 10 minutes at 4ºC to pellet the DNA, air-dried and resuspended in an appropriate volume of TE buffer.

2.6.2 DNA Quantitation

DNA extracted from protein contaminated samples/bacterial cultures was quantitated on the Biophotometer from Eppendorf. DNA samples were diluted 1:50 in distilled water. Readings were taken at 230, 260 and 280 nm in comparison to an appropriate control (1x TE, diluted 1:50 in water). DNA concentrations and the DNA purity (A260/A280) were calculated by the spectrophotometer.

High molecular weight DNA (from mammalian cells or BACs) embedded in agarose was quantified with a DyNA Quant 200 fluorometer (Hoefer). A single agarose plug (100 µl) was melted at 68ºC, and then the liquid solution was equilibrated to 42ºC, before adding β-agaraseI and incubating at 37ºC for 1 hour. The DNA was diluted 1000-fold in an assay solution composed of 1 µg/ml Hoechst dye in 1x TNE buffer.

2.6.3 Purification of DNA from Agarose Gels

DNA was purified from 0.8% agarose gels using the QIAquick Gel Extraction kit (Qiagen). DNA adsorbs to the silica-membrane in the spin columns in the presence of high salt while contaminants pass through the column. Impurities are efficiently washed away with optimised buffers, and the pure DNA is eluted with Tris buffer or water. The manufacturer's protocol for using a microcentrifuge was followed with the exception of the addition of isopropanol at step 5 which was omitted. 30 µl of elution buffer was used, the column left to stand for 2 minutes and then centrifuged for 2 minutes at 13000 rpm.

2.6.4 Restriction Enzyme Digestion

DNA prepared for analysis was digested in an appropriate volume of water, with 1x digestion buffer and 1x BSA. 5 mM spermidine was added if the NaCl concentration in the buffer exceeded 50 mM. At other NaCl concentrations, 2 mM spermidine was added. BAC DNA prepared in agarose plugs was washed in TE for 6 lots of 20 minutes. Half a plug was then equilibrated in restriction enzyme buffer containing the appropriate amount of BSA and spermidine for two periods of 30 minutes on ice. This was then replaced with 150 µl of fresh restriction buffer/BSA/spermidine along with 10 U of enzyme, and incubated overnight at 37ºC. Plugs were washed for two periods of 30 minutes in 0.5x TBE before loading on Pulse Field Gel Electrophoresis (PFGE) apparatus (Bio-Rad).

2.6.5 DNA End Sequencing

Approximately 500 nanograms (ng) of BAC DNA was mixed with 4 µl BigDye Sequencing Mix (PE Applied Biosystems), 4 µl BigDye diluent (200 mM Tris-HCl, 5 mM $MgCl_2$) and 4 µl of the directional primer (M13; Promega). The volume was adjusted to 20 µl with water. The PCR programme was 96ºC for 15 seconds, then 25 cycles of 96ºC for 45 seconds, 50ºC for 45 seconds and 60ºC for 3 minutes 30 seconds. The reactions were ethanol precipitated and analysed on a 3100 Sequencer. The data was assessed using EditView 1.0.1 and sequence alignments were performed using NCBI nucleotide BLAST (BLASTN) searches on the human genome without filters for repetitive sequences.

2.7 Agarose Gel Electrophoresis

DNA fragments up to ~12 kb were generally separated by electrophoresis through 1% agarose containing 0.5 µg/ millilitre (ml) ethidium bromide, in 1x TAE buffer containing 0.5 µg/ml ethidium bromide. Samples were loaded in 1x loading dye. The DNA size markers used were the 1 kb and 100 bp ladders (NEB) reconstituted in 1x loading buffer. Electrophoresis was normally conducted between 80 and 120 volts (V) for 60-120 minutes. Low melting point (LMP) agarose gels were usually prepared with 0.8% agarose and run in 0.5x TBE buffer. Gels were visualised on a short wave ultra violet (UV) transilluminator (Bio-Rad) and photographed. LMP gels were usually visualised on a long wave (356 nm) UV transilluminator to minimise the formation of thymidine dimers during gel excision.

2.8 Pulsed-Field Gel Electrophoresis (PFGE)

PFGE was performed on a Chef Mapper Machine (Bio-Rad). Gels contained 1% SeaKem GTG agarose dissolved in 0.5x TBE and were run in 0.5x TBE at 14ºC at 6V/centimetre (cm) and an induced angle of 120º. The switch times and durations of the run varied according to the samples loaded and were automatically selected by the BioRad equipment algorithm once the expected smallest and largest fragment sizes were entered. The DNA markers used were the λ concatemer ladder (48.5 kb to 1018.5 kb) the midrange II ladder (24 kb to 291 kb) and the λ/HindIII ladder (125 bp to 23130 bp) (all NEB). Gels were stained for at least an hour in water containing 0.5 µg/ml ethidium bromide and visualised as described in section 2.7.

2.9 *E. coli* Culture

2.9.1 Growing *E. coli* on Plates and in Liquid Culture

E. coli clones were streaked out onto LB plates containing the appropriate antibiotic at the following concentrations: 100 µg/ml ampicillin; 33.3 µg/ml kanamycin; 20 µg/ml chloramphenicol; 12.5 µg/ml tetracyclin and grown overnight at 37ºC. To grow *E. coli* in liquid culture, 10 ml of LB containing the appropriate antibiotic was inoculated with a single bacterial colony and incubated at 37ºC with shaking at 300 rpm overnight. These cultures were either manipulated for preparing DNA as a miniprep or 1 ml was used to inoculate 200 ml LB (with the appropriate antibiotic) and the culture was incubated at 37ºC with shaking at 300 rpm overnight before a maxipreparation was performed. Glycerol stocks were made with 50% glycerol in LB and stored at -70ºC.

2.9.2 Minipreparation of plasmid DNA

Plasmid 10 ml starter cultures were processed using the QIAprep Miniprep Kit (Qiagen) according to the manufacturer's protocol for use with a microcentrifuge. The QIAprep miniprep procedure is based on alkaline lysis of bacterial cells followed by adsorption of DNA onto silica in the presence of high salt. The procedure consists of three basic steps: the preparation and clearing of a bacterial lysate, adsorption of DNA onto the QIAprep membrane and washing and elution of plasmid DNA.

2.9.3 Minipreparation of BAC DNA

BAC DNA was prepared as follows. 5 ml of an overnight culture was pelleted and resuspended in 200 µl of STET buffer by vortexing. 580 µl of alkaline lysis buffer was added while vortexing to lyse the cells, followed immediately by 430 µl of 7.5 M ammonium acetate. The cells were incubated on ice for 5 minutes and pelleted at 13000 rpm for 20 minutes at 4ºC. The supernatant was removed and mixed with 250 µl of isopropanol before further centrifugation at 10000 rpm for 6 minutes at room temperature. The supernatant was then aspirated and the pellet washed with 200 µl of 70% ethanol at 10000 rpm for 4 minutes. The supernatant was again aspirated and the pellet air-dried before resuspension in 50 µl of TE.

2.9.4 Maxipreparation of Plasmid and BAC DNA

Plasmid DNA was extracted from an overnight bacterial culture (prepared from 200 ml LB) using the Qiagen Maxiprep Kit. The plasmid purification protocol is based on

a modified alkaline lysis procedure, followed by binding of plasmid DNA to Qiagen Anion-Exchange Resin under appropriate low-salt and pH conditions. RNA, proteins, dyes, and low molecular weight impurities are removed by a medium-salt wash. Plasmid DNA is eluted in a high salt buffer and then concentrated and desalted by isopropanol precipitation. The following deviations from the manufacturer's protocol were observed. Instead of the first centrifuge spin, lysates were filtered through Whatman paper directly into the Qiagen-tip columns. BAC DNA was prepared in the same way with the same modification as described above and in addition, the elution buffer QF was heated to 65 ºC.

2.9.5 Maxipreparation of BAC DNA in Agarose Plugs

100 ml of LB medium containing the relevant antibiotics was inoculated with 1 ml of an overnight bacterial pre-culture and shaken at 37ºC for 6 hours. Cells were pelleted by centrifugation at 2800 rpm in a Beckman J-6B centrifuge and resuspended in 5 ml of TE 50:10. 10 ml of fresh alkaline lysis solution were added and mixed by gently inverting, followed immediately by 8 ml of 7.5 M ammonium acetate. Lysates from three such reactions were pooled and centrifuged at 13000 rpm for 20 minutes at 4ºC. Supernatants were filtered and the DNA precipitated by the addition of 35 ml of isopropanol followed by centrifugation at 13000 rpm for 15 minutes. The supernatant was aspirated thoroughly and the pellet resuspended in 1.5 ml of 25 mM EDTA pH 8.0, 0.05% N-lauroyl sarcosine with 75 µg of RNase. A 1.5% LMP agarose in TE 50:10 was equilibrated at 37ºC. 1.5 ml of this agarose solution was added to the DNA, mixed and pipetted into pre-chilled plug formers. The agarose was left to set for 20 minutes after which the plugs were pushed into digestion buffer containing 0.5

M EDTA pH 8.0 with 1% sarcosine and 2 milligrams (mg)/ml pronase. Plugs were incubated at 55ºC overnight, washed three times in TE 50:10 and stored at 4ºC.

2.10 DNA Cloning

2.10.1 Preparation of Electrocompetent Cells

A single bacterial colony from a fresh streak was used to seed an overnight 1.5 ml culture in LB plus antibiotic for 12-16 hours at 37ºC with shaking. The starter culture was then tipped into 100 ml of LB plus antibiotic and grown at 37ºC with shaking until an optical density (OD) $_{600}$ of between 0.4 and 0.5 was reached. The entire culture was then chilled on ice for 30 minutes prior to centrifugation at 3000 rpm for 10 minutes in a Beckman J6-B centrifuge at 4ºC. Whilst keeping the bacteria on ice, the pellet was resuspended in 100 ml of sterile, ice-cold 10% glycerol. This wash was repeated three times and then the cells were resuspended in 600 µl of 10% glycerol, aliquoted as necessary and snap-frozen in a dry ice/ethanol bath and stored at -70ºC.

2.10.2 Plasmid End Dephosphorylation and Ligation

50 ng of plasmid DNA linearised with the appropriate enzymes was dephosphorylated with 1 U of shrimp alkaline phosphatatse (Roche) in 1x dephosphorylation buffer for one hour at 37ºC before heat inactivation of the SAP at 70ºC for 20 minutes. 100 ng of the insert DNA, linearised with the appropriate enzyme, was ligated to the dephosporylated plasmid using T4 DNA ligase (Roche) according to the

manufacturer's protocols. As a negative control, a similar ligation reaction containing only the dephosphorylated plasmid was set up in parallel.

2.10.3 Bacterial Transformation

Electrocompetent cells were thawed on ice. 40 µl of cells were added to a chilled microfuge tube containing DNA. If necessary, Millipore filter discs (Catalogue Number: VSWP01300) were used according to the manufacturer's instructions to dialyse samples on distilled water prior to electroporation. The cell/DNA mixture was pipetted into a chilled 0.1 cm cuvette (Bio-Rad) and elecroporated at 2.0 kilovolts (kV), 200 ohms (Ω) and 25 microFarads (µF). 1 ml of SOC medium was added to the cuvette and the mixture transferred to a 15 ml Falcon tube and shaken at 300 rpm for 1 hour at 37ºC. The reactions were then plated at various dilutions onto LB plates containing the appropriate antibiotic and incubated overnight at 37ºC.

2.10.4 Cre-*loxP* - Mediated BAC Retrofitting

An aliquot (40 µl) of electrocompetent cells containing a *loxP*, Cm^R carrying BAC was transformed with 50 ng of pCTP-T and 100 ng of a plasmid carrying the R6γ origin of replication, a *loxP* site and Amp^R, using conditions as described in section 2.10.3. Cells were then incubated at 30ºC for 1 hour with 500 µl of SOC containing 20 µg/ml of heat inactivated chlorotetracycline (cTc) to induce Cre expression from pCPT-T. This solution was diluted 1:10 in LB containing the appropriate antibiotics and incubated at 30ºC (permissive temperature for pCTP-T) for 3 hours before being

plated at various dilutions and incubated overnight at 42ºC (restrictive temperature for the pCTP-T) on chloramphenicol/ampicillin LB plates.

2.11 Cultured Cell Lines

2.11.1 Tissue Culture Conditions

The human cell lines, HT1080 (fibrosarcoma; ECACC 85111505), HPRT deficient (HPRT-) HT1080, HT1.1 (a tetraploid HT1080 strain efficient at HAC formation (Grimes et al., 2004); a gift from Brenda Grimes), MRC5-V2 (lung fibroblast transformed with SV40 large T antigen; a gift from Elaine Levy), 293 (kidney transformed with adenovirus 5; ECACC 85120602), G16-9 (a derivative of the human glioma cell line Gli-36 (Kashima et al., 1995) which expresses the HSV-1 VP-16 gene), Hep3B (hepatocarcinoma; ECACC 86062703), HeLa (cervix epitheloid carcinoma; ECACC 92021013), D98 (a HeLa derivative; ECACC 85112701), MRC5 (primary foetal lung; ECACC 97113601) and Simian Vero 2-2 (a Vero derivative, expressing HSV-1 protein ICP27; a gift from Rozanne M. Sandri-Goldin) were cultured at 37°C in 5% CO_2 (vol/vol) in high glucose D-MEM (Dulbecco's Modified Eagle's Medium) Glutamax (Invitrogen) supplemented with 10% FCS (Foetal Calf Serum, PAA Laboratories), penicillin (100 U/ml) (Biowhittaker) and streptomycin (100 μg/ml) (Biowhittaker). BeFA (immortalised keratinocyte from a patient with Epidermolysis Bullosa; a gift from Alain Hovnanian) cells were cultured in Keratinocyte-SFM supplemented with Epidermal Growth Factor (EGF) and Bovine Pituitary Extract (according to manufacturer's instructions). The packaging cell line Vero 2-2 (a gift from Rozanne M. Sandri-Goldin) was supplemented with G418 500

μg/ml. G16-9 cells were cultured with 200 μg/ml hygromycin. Cells were cultured in 10 cm dishes and as all cells were adherent, split with 1x Trypsin in PBS. Aliquots of cells were frozen by centrifuging at 1200 rpm, resuspending the cell pellet in 1 ml of freezing medium (5% Dimethyl Sulfoxide (DMSO) in FCS) and pipetting into cryotubes (Nalgene). Cells were stored at -70ºC overnight before transferring them to liquid nitrogen.

2.11.2 Subcloning of cultured cells

Some HAC containing clones were subcloned to obtain a higher percentage of HAC-containing cells. Cells were harvested from subconfluent dishes using trypsin, counted using a haemocytometer, and diluted to (on average) a single cell per well. Cells were seeded in 96-well plates and wells containing single cells were identified. Clones were allowed to form and then cells were expanded for further analysis.

2.12 Lipofection of Cultured Mammalian Cells

Cells were grown to 50-70% confluence. BAC DNA was prepared in agarose plugs as described in section 2.9.5. 2 μg of DNA per 10 cm dish of cells was used. Plugs were washed 3 times, for 20 minutes each in TE. Plugs were then equilibrated twice in 1x TE containing 3 μM polyethylenimine (PEI) for 30 minutes each. Most of the PEI solution was then removed although a few drops were left to help prevent the LMP agarose from re-setting. The plugs were melted at 65ºC for 10 minutes and

placed in a 37ºC water bath. For every plug, 2 U of β-agaraseI were diluted in 25 µl of the 3 µM PEI in TE solution and added to the plugs. The solution was mixed gently and incubated in the water bath at 37ºC for 2 hours.

Once agarose digestion was complete the DNA was diluted in OptiMEM media using 2 ml per µg of DNA. For each µg of DNA, 6 µl of Lipofectamine 2000 (1 mg/ml) was diluted in 2 ml OptiMEM and incubated at room temperature for 5 minutes. The DNA and lipofectamine reagent solutions were then mixed at a ratio of 1 µg DNA (in 2 ml OptiMEM) to 6 µg Lipofectamine 2000 (in 2 ml OptiMEM). This mixture was incubated at room temperature for 20 minutes.

The growth medium from the cells was removed and the cells washed with OptiMEM. Cells were then incubated in the transfection mix overnight at 37ºC, 5% CO_2. The transfection mixture was removed the following morning and replaced with D-MEM containing 10% FCS. The cells were then split 1:4 and selection was applied 24 hours later (350 µg/ml for HT1080). Colonies were picked 10-14 days later into 96-well plates and expanded for analysis.

2.13 HSV-1 Amplicon Preparation

$1x10^6$ Vero 2-2 cells were seeded in 6 cm dishes 24 hours prior to transfection. For each dish, a mix of 2 µg of fHSVΔpacΔ270+ DNA, 1.8 µg of BAC DNA, 200 ng of plasmid pEBHICP27 and 10 µl of Plus Reagent (Invitrogen), diluted in 250 µl of OptiMEM was prepared. A second mix, composed of 250 µl of OptiMEM, and 23 µl of Lipofectamine reagent (Invitrogen) was added to the first and incubated for 30

minutes at room temperature. After adding OptiMEM to a final volume of 1.5 ml, the DNA/Lipofectamine Plus mixture was applied to the cells. The cells were incubated at 37°C for 4 hours and then the medium was removed and replaced with D-MEM containing 6% FCS and 25 mM HEPES. After 60 hours, the cells were harvested in D-MEM, in groups of 4 dishes, and frozen on dry ice. The cells were then sonicated and spun at 3500 rpm at 4ºC to remove cellular debris. The supernatant was layered onto a 25% sucrose cushion and spun at 20000 rpm, 4°C in a Beckman SW28 rotor for 3 hours to concentrate the amplicons. The pellet was resuspended in 250 µl of D-MEM containing 10% FCS. The titre of each amplicon preparation and the number of transducing units was estimated by infecting $2.5x10^5$ G16-9 cells seeded in a 24 well dish and counting the number of GFP positive cells 24 hours after infection, assuming that each fluorescent cell corresponded to 1 transducing unit. GFP positive cells were viewed using an Olympus IX51 microscope. Images were captured with a Jenoptik CCD camera and visualised using Improvision Openlab software. 10 µl aliquots of amplicons were stored at -80˚C for a maximum of one month.

2.14 Cell Infection With HSV-1 Amplicons

For each cell line, a confluent 24 well dish (approximately $1\text{-}5x10^4$ cells) was infected with HSV-1 amplicons at a MOI (Multiplicity Of Infection) ranging from 1 to 10, in 250 µl of D-MEM and 10% FCS medium. After 24 hours, the efficiency of infection was determined by counting the number of GFP positive cells in each well divided by the total number of cells. Cells were split into 4 wells. After 24 hours G418 selection was applied: 200 µg/ml for 293, 300 µg/ml for Hep3B, 350 µg/ml for HT1080, MRC5-V2, HeLa and D98 and 500 µg/ml for G16-9 to allow for the recovery of

stably transformed clones. Cells typically died under selection within 4-8 days and positive clones were identified after about 14-21 days, isolated and expanded.

2.15 Preparation of Cultured Cells for Cytogenetic Analysis

Actively dividing cells were treated with 30 ng/ml colcemid (Karyomax, Invitrogen). at 37ºC for 2-4 hours. The medium was collected; cells were washed with trypsin and the trypsin collected along with the detached cells. The cells were spun at 1200 rpm for 10 minutes. The pellet was resuspended in 10 ml of pre-warmed hypotonic solution (37ºC) (75 mM KCl for HT1080 and 56 mM KCl for all other cell lines) and incubated for 6 minutes at room temperature. Following another 1200 rpm centrifugation for 10 minutes the supernatant was carefully removed and the pellet was resuspended in 10 ml of ice-cold fixative (3 methanol:1 acetic acid) and left at room temperature for 30 minutes. Cells were spun again and the pellet resuspended in 10 ml of ice-cold fixative and incubated at room temperature for 10 minutes. Cells were spun once more and resuspended in 0.2-1 ml of fixative depending on the size of the pellet. 30 µl of the cell preparation was dropped on a clean slide and the slide dried on a 37ºC heat block. Cell preparations were stored in 1.5 ml of ice-cold fixative at -20ºC and slides were stored for short periods at room temperature. Slides were at least left overnight before proceeding to cytogenetic analysis.

2.16 Fluorescent in situ Hybridisation (FISH)

2.16.1 Preparation of Probes

FISH probes were labelled with biotin-16-dUTP or digoxigenin-11-dUTP (Roche) by nick translation using a commercial kit (Nick Translation System, Invitrogen) following the manufacturer's instructions. Probe length was determined by running a small aliquot on a 1% agarose gel. Probes were precipitated with a 10x excess of Cot I DNA (Roche), 3 M ammonium acetate and 2 volumes of ethanol for 15 minutes at room temperature. Probes were resuspended in hybridisation buffer (50% formamide, 10% dextran sulphate, 2x SSC) at a concentration of 10 ng/µl.

2.16.2 Dot Blot

To check for the level of labelled dUTP incorporation, a dot blot assay was used. A serial dilution of the labelled probe ranging from 1ng to 1 picogram (pg) (diluted in 20x SSC) was spotted onto a nitrocellulose membrane (Amersham Hybond N+) and allowed to dry. After a brief fixation at 60°C for 15 minutes, the membrane was pre-incubated in dot blot buffer containing 3% BSA for 30 minutes at room temperature. Anti-digoxin or Streptavidin conjugated to alkaline phosphatase (both Sigma) was added to this solution at a dilution of 1:5000 and incubated for 30 minutes at room temperature. The membrane was then washed for 3 lots of 5 minutes in dot blot buffer, followed by a final wash in dot blot buffer containing 5 mM $MgCl_2$. A BCIP/NBT (5-bromo-4-chloro-3-indolylphosphate/nitro blue tetrazolium) tablet (Sigma) dissolved in 10 ml water (to give a solution containing 0.15 mg/ml BCIP, 0.3 mg/ml NBT, 100 mM Tris and 5 mM $MgCl_2$, pH 9.5) was then added to the

membrane. The colorimetric reaction was allowed to proceed for a maximum of 30 minutes, then stopped with 0.1 M EDTA (pH 8.0). Probes were considered to be optimally labelled when the 5 pg spot was clearly identifiable.

2.16.3 FISH on Metaphase Chromosomes

Chromosomal DNA was denatured in 50 mM KCl, 10 mM Tris-HCl (pH 8.0) and 5% glycerol at 95°C for 10 minutes. Slides were washed briefly in 0.1x SSC followed by dehydration for 3 minutes each in 70%, 90% and 100% ethanol. Probes were denatured for 10 minutes at 85°C. Usually, 150 ng of labelled probe was applied to the slides and hybridisation was carried out at 42°C overnight, followed by three post-hybridisation washes in 0.1x SSC at 60°C. Biotin labelled probes were detected using Avidin-FITC (Fluorescein Isothiocyanate) (Molecular Probes, Invitrogen), followed by biotinylated anti-avidin antibody (Vector Laboratories), and a second layer of Avidin-FITC. Digoxigenin probes were detected with rhodamine conjugated sheep anti-digoxigenin antibody (Roche), followed by rhodamine anti-sheep antibody (Chemicon Europe). Antibodies were diluted in 4x SSC, 0.1% Tween 20, 3% BSA with an incubation period of 30 minutes at 37°C followed by three washes in 4x SSC, 0.1% Tween 20 at 42°C. Chromosomes were counterstained with DAPI (4',6-Diamidino-2-Phenylindole) and mounted in DABCO (1,4-Diazabicyclo(2,2,2)octane) (Sigma) antifade. Slides were analyzed with an Olympus BX60 microscope for epifluorescence equipped with a Sensys CCD camera (Photometrics). Images were collected and pseudocoloured using MacProbe 4.3 software.

2.16.4 ACA Detection

For ACA (Anti-Centromere Autoimmune Serum) detection, 50 µl of chromosome suspension fixed in methanol:acetic acid (3:1), were dropped onto clean slides and briefly (but not completely) allowed to air dry. After a 5 minute incubation in PBS, slides were washed in TEEN buffer containing 0.1% Triton X-100 and 0.1% BSA and incubated for 30 minutes at 37°C in ACA serum (a gift from William Earnshaw) diluted 1:5000 in the same buffer. After 3 washes in 10 mM Tris-HCl (pH 8.0) with 0.1% BSA, 10 ng/ml FITC-anti-human secondary antibody (Sigma) was applied to the slides. After 3 washes in 10 mM Tris-HCl (pH 8.0) with 0.1% BSA, cells were fixed in 4% paraformaldehyde (Science Services) diluted in PBS, for 15 minutes. Slides were washed in 2x SSC for 5 minutes and then 100-150 ng of FISH probe was applied under a coverslip. The probe and the slide DNA were simultaneously denatured at 80°C for 5 minutes on a PCR block. FISH was then carried out as described in section 2.16.3.

2.16.5 CENP-A Detection

For CENP-A detection, a subconfluent layer of cells was incubated for 3 hours in 0.03 µg/ml colcemid (Karyomax, Invitrogen). Cells were harvested and resuspended in cold 56 mM KCl at $2x10^5$ cells/ml. Approximately 200 µl of cell suspension was spun onto poly-lysine coated slides (BDH) at 800 rpm for 4 minutes, using a Shandon cytospin machine. The cells were then permeabilised in KCM buffer for 10 minutes and incubated for 30 minutes at 37°C in 2 ng/ml of anti-human murine CENP-A antibody (Abcam) diluted in PBS and 1% BSA. Slides were washed three times for 5

minutes in KCM and then 10 ng/ml of FITC conjugated anti-mouse secondary antibody (Chemicon) was applied to the slides. Slides were washed again three times for 5 minutes in KCM buffer and then the preparations were fixed in 2% formaldehyde (Science Services) diluted in KCM for 15 minutes. Slides were washed in 2x SSC for 5 minutes and then 100-150 ng of FISH probe was applied under a coverslip. The probe and the slide DNA were simultaneously denatured at 80°C for 5 minutes on a PCR block. FISH and image processing was then carried out as described in section 2.16.3.

2.16.6 Histone H3 Phosphorylated at Serine 28 Immunoflourescence

Cells were seeded onto a coverslip and allowed to reach subconfluency. The cells were then fixed in methanol for 5 minutes followed by a 5 minute wash in PBS containing 0.1% Triton. Slides were incubated with rabbit Anti-H3S28 (Abcam) diluted 1:200 in PBS, 0.1% Triton for 30 minutes at 37°C. Following 3x 5minute washes in PBS, 0.1% Triton at room temperature, the slides were incubated with a secondary anti-rabbit antibody conjugated to Texas Red (Chemicon) for 30 minutes at 37°C. After final washes, slides were stained with DAPI and mounted in DABCO antifade. Image processing was carried out as described above.

2.16.7 Cytokinesis Block Micronucleus Assay

Cytochalasin B (4 μg/ml) was added to subconfluent dishes of cells. After 48 hours the cells were harvested by trypsinisation and resuspended in PBS. Cells were

pelleted and resuspended in methanol:acetic acid 3:1 fixative containing 100 µl/ml formalin then washed twice in fixative only, and dropped onto clean slides. The cells were then hybridised as described above to BAC 227J24 and/or BAC pWTR11.32. To estimate the level of aneuploidy, binucleated cells were analysed for the presence of non-disjunction events (one nucleus showing one or more centromeric signal than the modal number, and the other one or more less) and the presence of bridges and micronuclei. Only binucleated cells with round nuclei, of similar size and shape were considered. Micronuclei were scored when they appeared to be 1/3-1/6 of the nucleus and of similar colour and density. Where both probes were used, the data for the two probes were then pooled together. It must be noted that BAC pWTR11.32 detects both chromosome 21 and chromosome 13 alpha satellite DNA, while BAC 227J24 only recognises chromosome 17 alpha satellite DNA.

2.16.8 Mitotic Stability

To investigate the mitotic stability of HACs, cells were passaged under non-selective conditions, in the absence of G418, and HACs were detected by FISH at 0, 30, 60 and 90 days in culture. HAC stability was calculated by the formula: $N_n=N_0 \times (1-R)^n$, where N_0 is the number of metaphase chromosome spreads showing HACs in the cells cultured under selection, N_n is the number of HAC-containing metaphase chromosome spreads after n days of culture in the absence of selection and R is the daily rate of loss.

2.17 Quantitative Real Time PCR

2.17.1 cDNA Preparation for Real-Time PCR

RNA was harvested from one confluent plate of cells using Qiagen's RNeasy Midikit, following the manufacturer's instructions. RNA was quantitated on the Biophotometer from Eppendorf. Samples were diluted 1:50 in distilled water. Readings were taken at 260 nm and the concentrations were calculated by the spectrophotometer. If deemed necessary, the RNA was checked on a northern gel (section 2.17.2). cDNA (complementary DNA) was prepared from 2 μg of RNA using Qiagen's QuantiTect Reverse Transcription Kit according to the manufacturer's instructions.

2.17.2 Northern Gel Analysis

A 1% agarose gel was prepared for each analysis. Agarose was melted in 1x northern gel running buffer and 6% formaldehyde was subsequently added. 500 ng of RNA was resuspended in a total volume of 20 μl in: 50% formamide, 6% formaldehyde, 1x gel buffer, 0.5 ng/μl ethidium bromide and 2 μl loading dye. The samples were heated at 65˚C for 10 minutes and then placed on ice prior to loading the gel. Gels were run in 1x running buffer at 30 millivolts (mV) for 3 hours and then visualised as described in section 2.7.

2.17.3 The Real-Time PCR Reaction

cDNA from the reverse transcription reaction was diluted 1:25 and 3 µl were used in each real time PCR reaction with 7.5 picomoles (pmoles) of each of forward and reverse primer, 2.5 pmoles of fluorogenic probe (labelled with FAM (fluorencin) at the 5' end and TAMRA (tetramethylrhodamine) at the 3' end) and 2x TaqMan Universal PCR Mastermix (ABI) in a total volume of 25 µl/well. The Universal Mastermix contains AmpiTaq Gold DNA Polymerase, dNTPs (Deoxynucleotide Triphosphates), optimised buffer components, $MgCl_2$, and Passive Reference Dye (ROX). ROX provides an internal reference and is used to normalise the real-time PCR reactions by normalising fluorescent fluctuations and compensating for well-to-well variations. Primers and probes were designed using Primer Express (ABI) and where possible where chosen to overlap exon boundaries to ensure no signal from contaminating genomic DNA. Real time reactions were carried out on an ABI PRISM 7700 machine using universal thermal cycling parameters (Table 2.4). The universal cycling conditions eliminate optimisation of the thermal cycling conditions and allow multiple assays to be run on the same plate.

Table 2.4: Universal Cycling Parameters for Real-Time Quantitative TaqMan Assays

Times and Temperatures			
Initial Steps		**Each of 40 cycles**	
		Melt	Anneal/Extend
HOLD	**HOLD**	**CYCLE**	
2 min* 50ºC	10 min** 95ºC	15 sec 95ºC	1 min 60ºC

* The 2 min hold at 50ºC is required for optimal AmpErase UNG activity.
** The 10 min hold at 95ºC is required for AmpliTaq Gold DNA Polymerase activation.

2.17.4 Real-Time PCR Primer and Probe Design

Primers and probes were designed using Primer Express (Table 2.5) following the TaqMan probe guidelines:

- amplicons of 50-150 bp are recommended as they promote high-efficiency assays
- the G/C content should be 20-80% to enable efficient denaturing and reduce non-specific interactions and runs of identical nucleotides (particularly of 4 or more Gs) are to be avoided
- the T_M (Melting Temperature) should be 58-60°C for primers and 68-70°C for probes to allow the use of universal cycling parameters
- there should be no G on the 5' end of the probe as this will have a quenching effect even after probe cleavage
- the 5 nucleotides at the 3' end of the primers should have no more than 2 G and/or C bases to prevent non-specific product formation
- the probe should have more C than G bases to prevent reduced normalised fluorescence values (otherwise the complementary strand should be selected).

Results were analysed using the comparative C_T (Threshold Cycle) method and the formula:

$$2^{-\Delta\Delta CT}$$

where ΔCT is the difference between the C_T of the target gene and the C_T of the control.

Table 2.5: Primers and Probes Used in Real-Time PCR Analysis

	Forward Primer	Probe	Reverse Primer
CENP-A	CCCTCCTTAGGCGCTTCCT Overlaps exons 1 and 2	CCATCAACACAGTCGGCGGAGAC	CTGAAGCTTTCGGATCTCCTTTAG
CENP-B*	GGCTTACTTTGCCATGGTCAA	TGACCTCCTTCCCCATTGATGACCG	TTGATGTCCAAGACCTCGAACTC
CENP-C	AGAAGTGACTTCAACTGTCACGAAA	CCACCAATCAGATGGACGCCTGGA	TGCTATAAACAGGACTCTCCTCTGATT Overlaps exons 9 and 10
CENP-E	ATCGAGATAGCAAGTTAACACGAATTC	TCCAAAGACACGTATTATCTGCACAATTACTCCA	AGTACTGGCAAACTGGAGAGCAGTA Overlaps exons 11 and 12
CENP-F	AGAAAGCAGTCATGAGTGGTATTCA	CCTGTCGGGATGTCAGCAAACCCTT Overlaps exons 18 and 19	CTTCGCAGGATATATGGGCTAGTCT
TopoI	CTGACAGCCCCGGATGAG Overlaps exons 17 and 18	TCCCAGCGAAGATCCTTTCTTATAACCGTG	AAAGAATTGCAACAGCTCGATTG
Kin17	TCACGTTTAATTTGAGTAAAGGAGCAT	TAGCTCATCCGGAGCAACATCTTCCAAG	GCACTCGGTCCCAGAGTACTTG Overlaps exons 7 and 8
Aurora B	CTGCGCAGAGAGATCGAAATC	AGGCCCACCTGCACCATCCC Overlaps exons 5 and 6	AAAAATAGTTGTAGAGACGCAGGATGT
GAPDH	CATCCATGACAACTTTGGTATCGT	AAGGACTCATGACCACAGTCCATGC Overlaps exons 7 and 8	CAGTCTTCTGGGTGGCAGTGA
β-2M	TGACTTTGTCACAGCCCAAGATA	TGATGCTGCTTACATGTCTCGATCCCA Overlaps exons 3 and 4	AATCCAAATGCGGCATCTTC

* CENP-B has no introns so it was not possible to design probes overlapping exon-intron junction

2.18 Western Blotting

2.18.1 Nuclear Protein Fraction Preparations

For each cell line, a nuclear protein fraction was prepared from a confluent 10 cm dish, either unsynchronized or blocked in metaphase by treatment with colcemid at 1 μg/ml for 24 hours. The harvested cells were lysed in PBS, 0.5% Triton X-100, 1x Protease Inhibitor (Roche) on ice for 10 minutes at a concentration of 10^7 cells/ml. Preparations were spun at 2000 rpm for 10 minutes at 4˚C and the pellet washed in half the volume of the above buffer. The resulting pellets were then extracted at 4˚C overnight in 0.2 Normality (N) HCl at a concentration of $4x10^7$ cells/ml. In some instances nuclear and cytosolic protein fractions were prepared using the Nuclear Extraction Kit from Panomics according to the supplier's instructions. Protein concentrations were measured using a Bradford assay with BSA as the standard. Spectrophotometer readings were taken at 595 nm on an Ultraspec 2100pro from Amersham Biosciences.

2.18.2 SDS-PAGE and Transfer

10 μg of each preparation was then run on 4-12% Bis-Tris polyacrilamide gels in MES-SDS buffer using the SeeBlue Plus2 Pre-Stained Standard (4 kD to 250 kD) (all from Invitrogen) following the manufacturer's instructions. For the detection of proteins greater than 150 kD, preparations were run on 3-8% Tris-Acetate polyacrilamide gels in Tris-Acetate-SDS buffer using the HiMark Pre-stained Protein Standard (40 kD to 500 kD) (all from Invitrogen) following the manufacturer's instructions. Gels were electroblotted onto PDVF (polyvinylidene difluoride)

membranes (Invitrogen) using the Invitrogen Xcell II Blot Module according to the manufacturer's instructions.

2.18.3 Protein Detection

Protein levels were checked via Ponceau (Sigma) staining and then membranes were blocked overnight at 4˚C using 5% milk in TBS/0.1% Tween 20. Membranes were probed with the primary antibody (see Table 2.6) for 1 hour at room temperature in 5% milk/TBS/0.1% Tween 20 followed by 3x 15 minute washes in TBS/0.1% Tween 20 and incubation with the secondary antibody (see Table 2.6) for 1 hour at room temperature. The washes were repeated before detection with the ECL plus system (Amersham) using Biorad ChemDoc XRS software. The densitometrical analysis of each band was carried out using the software QuantityOne.

If it was necessary to strip a membrane of previously bound antibody, this was done using a solution containing 62.5 mM Tris-HCl (pH 6.8), 2% SDS and 100 mM β-mercaptoethanol. The membranes were shaken in this solution for 20 minutes at 50˚C and then washed for two lots of 10 minutes in TBS/0.1% Tween 20 before being blocked as usual.

Table 2.6: Antibodies Used to Probe Western Blots in this Study.

Antibody	Company	Dilution	Size of Detected Band (kD)
Primary Antibodies			
Mouse Anti-Aurora B	BD Biosciences	1:250	41
Rabbit Anti-CENP-A	Upstate	1:1000	17
Rabbit Anti-GAPDH	Abcam	1:5000	40
Mouse Anti-H3S10	Upstate	1:1000	17
Rabbit Anti-H3S28	Abcam	1:1000	17
Rabbit Anti-H3triK9	Abcam	1:500	17
Anti-HP1 α	Upstate	1:500	26
Rabbit Anti-Kin17	Abcam	1:2500	45
Rabbit Anti-PP2a alpha	Abcam	1:5000	36
Rabbit Anti-SMC3	Abcam	1:5000	140
Rabbit Anti-Topo I	Abcam	1:1000	92
Anti-Topo II α	Oncogene	1:1000	170
Secondary Antibodies			
Goat Anti-rabbit-HRP	Biorad	1:5000	
Goat Anti-mouse-HRP	Biorad	1:5000	

Chapter 3:

Results I: Infectious Delivery of Human Artificial Chromosomes into Human Cells using HSV-1 Amplicon Vectors

3.1 Introduction

Whilst HACs have been confirmed as viable gene expression vectors (Grimes et al., 2001; Ikeno et al., 2002; Kotzamanis et al., 2005; Mejía et al., 2001), HAC delivery has posed a barrier to the development of HACs as gene transfer vectors, with HACs only being formed efficiently in HT1080 cells to date. Using lipofection techniques, transfection efficiencies are poor and to prevent shearing and degradation of the DNA, condensation is required (Marschall et al., 1999). To overcome these problems, the aim of this work was to develop a novel HAC vector system for gene delivery into human cells using the Herpes Simplex Virus-1 (HSV-1) amplicon technology. This system was developed by Saeki et al. (Saeki et al., 2001) and utilised for BAC delivery by Wade-Martins et al. (Wade-Martins et al., 2001). It allows the delivery of large genomic fragments (150 kb) into cells, without immunogenic effects, using HSV-1 amplicons in a helper virus free system.

This system was chosen for the development of HAC delivery due to the high efficiency (up to 100%) it has shown in delivering BACs and PACs to human and mouse cells (Wade-Martins et al., 2001) and it can accommodate large sized alphoid arrays with relative ease due to the large genome size of HSV-1. Results from various groups indicate that the immune response in the helper-virus free system is minimal (Costantini et al., 1999; Olschowka et al., 2003). This is vital for gene therapy. Other

advantages of this system are the broad host range and the ability to infect dividing and non-dividing cells. Previously, the system has also been used to deliver large genomic genes, carried on BACs, (the HPRT locus and LDLR) for complementing deficiencies. Because of the large size it can accommodate, the HSV-1 amplicon system presents another advantage for generating therapeutic HACs in that it allows the gene to be kept under control of its own promoters and enhancers, thus maintaining physiological regulation of the gene expression (Wade-Martins et al., 2003; Wade-Martins et al., 2001).

This study illustrates that HSV-1 is a valuable method for HAC vector delivery by showing the successful generation of HSV-1 HAC vectors and the efficient infection of several human cell types including human HT1080 cells which are commonly used for HAC studies. HACs can be generated at a higher efficiency in human cell types other than HT1080 compared to previous delivery methods used. The transduction efficiency, determined by transient expression of the reporter gene, GFP, was several orders of magnitude higher than lipofection. HACs in HT1080 and G16-9 cells were stable up to 3 months in the absence of selection. This is in contrast to MRC5-V2 and 293 cells, where the HACs were unstable in the presence and absence of selection possibly indicating that the HACs have not retained a fully functional centromere in these cells. Alternatively, it is also possible that a number of other factors including the genetic background, and/or epigenetic factors which influence chromosome segregation (Rudd et al., 2003) are different in the different cell types and may influence HAC stability.

3.2 The Generation of HSV-1 HAC Vectors

This study focussed on the analysis and modification of 17 and 21 α-BACs for HSV-1 delivery to human cells, as chromosome 17 and 21 α-DNA is know to be efficient at HAC formation (Grimes et al., 2002; Mejía et al., 2002). The BAC 227J24, containing a 130 kb insert of 17 α-DNA and BAC pWTR11.32, containing 60 kb of 21 α-DNA, a gift from Hiroshi Masumoto (Ohzeki et al., 2002) were already available in the laboratory for use in this study. In an initial analysis for a third BAC to utilise, four BACs, 15K01201B, 302N02, 292P20 and 302P22, with 17 alphoid DNA inserts ranging from 50-150 kb were analysed. A variety of BACs were studied due to the possible instability of repetitive DNA in a bacterial background. Therefore the isolation of a construct containing a stable insert was desirable. For each clone, four independent colonies were analysed by digesting the DNA with NotI to release the α-DNA insert in order to check the size on a PFGE gel. A stable colony was then restreaked on an agar plate and the process repeated a further twice using NotI digestion to check the stability of the α-DNA insert (Figure 3.1). BAC 292P20 was ruled out from further analysis as restriction digestion showed the presence of deleted forms (Figure 3.1). BACs 15K01201B and 302N02 were both stable and BAC 302P22 was chosen for further development due to its stability and insert size similarity (50 kb) with the 21 alphoid BAC pWTR11.32 (60 kb). Using these two BACs allowed a comparison of two different alphoid sequences whilst controlling for vector size. The third BAC, 227J24, containing a 130 kb insert of 17 α-DNA, allowed the examination of the effect of size between two 17 α vectors.

The homology of 302P22 and 227J24 to chromosome 17 centromere core α-DNA was confirmed by sequencing. 17 α-DNA has a 16-monomer based higher-order repeat structure of 2.7 kb (Waye and Willard, 1986), released by EcoRI digestion. The 2.7 kb EcoRI fragment from both 302P22 and 227J24 was subcloned into the multiple cloning site of pUC19 where it was sequenced using M13 primers. The obtained sequences were then compared to human chromosome core 17 α-DNA with NCBI BLASTN software and found to be 98% homologous for both BACs 302P22 and 227J24, confirming that they contain non-degenerate core 17 α-DNA. Figure 3.2 shows the BLAST sequence alignments.

The frequency of CENP-B boxes (found at regular intervals in human α-DNA and required for the binding of CENP-B) in the alphoid DNA of BACs 302P22 and 227J24 was determined due to the importance of CENP-B boxes in HAC formation (Basu et al., 2005; Ohzeki et al., 2002). The CENP-B Box primers, detailed in Table 2.1, amplify from the start of one CENP-B box to the next (Figure 3.3C). From Figure 3.3 it can be seen that there are bands visible at 900 and 320 bp. If these values are divided by the length of the α-DNA monomer (171 bp) (Waye and Willard, 1986), values of 5.3 and 1.9 respectively are obtained. Thus, CENP-B boxes are present either every other monomer repeat or every 6 monomers. A similar distribution has been found in the chromosome 17 centromere (Basu et al., 2005).

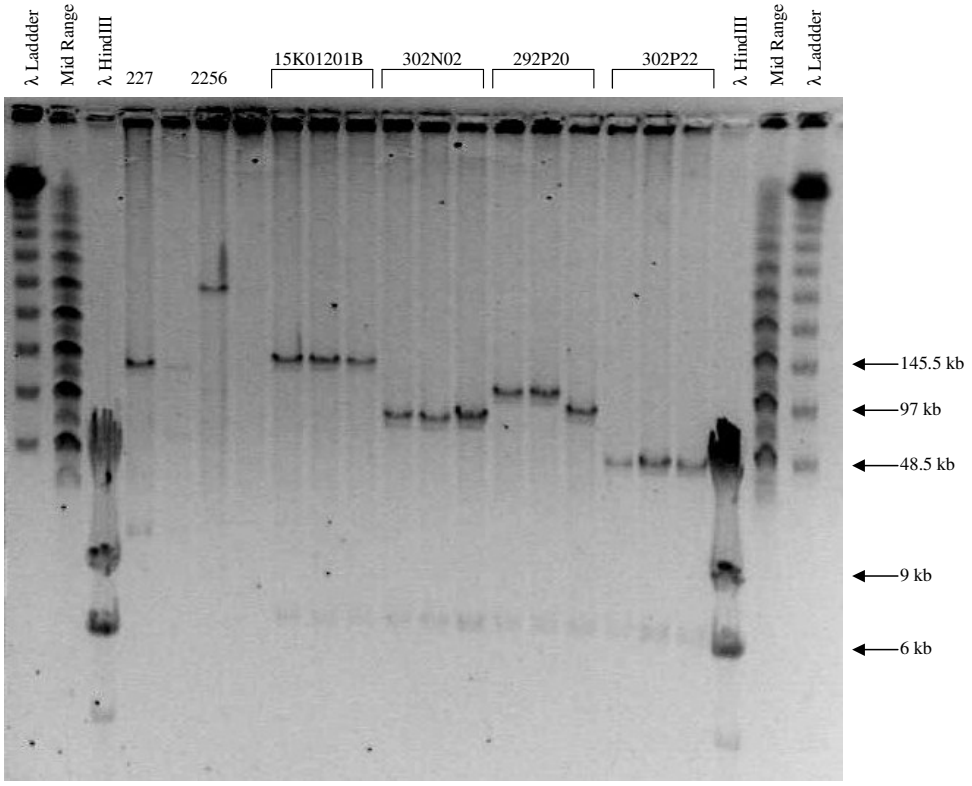


Figure 3.1: Pulsed Field Gel Electrophoresis Showing BAC Analysis
This gel illustrates the NotI digestion analysis of the four analysed BACs (15K01201B, 302N02, 292P20 and 302P22) compared with the BACs 227J24 and 2256 (Mejía et al., 2002). NotI digestion releases the α-DNA from the BAC vector. The instability of the 292P20 BAC can be observed. Regarding the 302P22 BAC, the 50 kb α fragment can be seen along with the 7 kb vector band.

A

```
Query  13   AAGCAGTCTCAGAAATCTTCTTTGTGATGTTTGTATTCAAATCCCCGAGTTGAACTTTCC  72
            ||||||||||||||| |||||||||||||||| ||||||||||||||||||||||||||
Sbjct  40   AAGCAGTCTCAGAA-TCTTCTTTGTGATGTTTGCATTCAAATCCCCGAGTTGAACTTTCC  98

Query  73   TTTCAAAGTTCACGTTTGAAACACTCTTTTTGCAGGATCTACAAGTGGATATTTGGACCA  132
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  99   TTTCAAAGTTCACGTTTGAAACACTCTTTTTGCAGGATCTACAAGTGGATATTTGGACCA  158

Query  133  CTCTGTGTCCTTCGTTCGAAACGGGTATATCTTCACATGACATCTAGACAGAAGCATCCT  192
            |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Sbjct  159  CTCTGTGTCCTTCGTTCGAAACGGGTATATCTTCACATGCCATCTAGACAGAAGCATCCT  218

Query  193  CAGAAGCTTCTCTGTGATGACTGCATTCAACTCACGGAGTTGAACACTCCTTTTGAGAGC  252
            |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Sbjct  219  CAGAAGCTTCTCTGTGATGACTGCATTCAACTCACGGAGTTGAACTCTCCTTTTGAGAGC  278

Query  253  GCAGTTTTGAAACTCTCTTTCTGTGGCATCTGCAAGGGGACATGTAGACCTCTTTGAAGA  312
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  279  GCAGTTTTGAAACTCTCTTTCTGTGGCATCTGCAAGGGGACATGTAGACCTCTTTGAAGA  338

Query  313  TTTCGTTGGAAACGGAATCATCTTCACATAAAGAACTACACAGATGCATTCTCAGGAACT  372
            |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Sbjct  339  TTTCGTTGGAAACGGAATCATCTTCACATAAA-AACTACACAGATGCATTCTCAGGAACT  397

Query  373  TTTTGGTGATGTTTGTATTCAACTCCCAGAGTTGAACTTTCCTTTGGAAAGAGCAGCTAT  432
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  398  TTTTGGTGATGTTTGTATTCAACTCCCAGAGTTGAACTTTCCTTTGGAAAGAGCAGCTAT  457

Query  433  GAAACACTGTTTTTCTAGAATCTGCAAGTGGACGTTTGGAGGGCTTTGTGGTTTGTGGTG  492
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  458  GAAACACTCTTTTTCTAGAATCTGCAAGTGGACGTTTGGAGGGCTTTGTGGTTTGTGGTG  517

Query  493  GAAAAGGAAATATCTTCACCTAAATACTAAATAGAACCATTCTCAGAAACTGC  545
            |||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Sbjct  518  GAAAAGGAAATATCTTCACCTAAATACTAGATAGAAGCATTCTCAGAAACTGC  570
```

B

```
Strand=Plus/Minus

Query  1     GCTCAGCTCTGTGAGTTCAACTCAATCATCCCAAAGAATTTTCTGAGAAAGCTTCTGTCT  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2638  GCTCAGCTCTGTGAGTTCAACTCAATCATCCCAAAGAATTTTCTGAGAAAGCTTCTGTCT  2579

Query  61    TCTTTTTATAGGAAGTTATTTCCTTTACTACGGTACTCCTCAAAGAGTGCAATTATCCCC  120
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Sbjct  2578  TCTTTTTATAGGAAGTTATTTCCTTTACTACGGTACTCCTCAAAGAGTGCAATGATCCCC  2519

Query  121   TTGCAGTTTCTACAGAAAGAGTGTTTCAAACCTGAACTATCAAAGAAAGGTTCCACACTG  180
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2518  TTGCAGTTTCTACAAAAAGAGTGTTTCAAACCTGAACTATCAAAGAAAGGTTCCACACTG  2459

Query  181   TGAGTTGAATGCAGACATCACGAAGAAGGTTCTGAGAATGCTTCTGTTTACTTCTGTGCA  240
             |||||||||||||||||||||||||||||||||||| ||||||||||| |||||||||
Sbjct  2458  TGAGTTGAATGCAGACATCACGAAGAAGGTTCTGAGAGTGCTTCTGTTTAGTTCTGTGCA  2399

Query  241   GTTTATCCCGTTTCCAACGAAATGCTCAGAGAGGACCAAATATCCACTTGCAGTTTCTAC  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2398  GTTTATCCCGTTTCCAACGAAATGCTCAGAGAGGACCAAATATCCACTTGCAGTTTCTAC  2339

Query  301   AAAAAGAGTGTTTCAAAGCTGAACTATCAAACAAAGGTTCAGCACTGTGAGTTGAATGCA  360
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  2338  AAAAAGAGTGTTTCAAAGCTGAACTATCAAAGAAAGGTTCAGCACTGTGAGTTGAATGCA  2279

Query  361   AACATTCACGAAGAGGGTTCTGAGAATGCTTCTGTTTTAGTTCTGTGCGGGTTATCCCGT  420
             |||| |||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Sbjct  2278  AACA-TCACGAAGAGGGTTCTGAGAATGCTTCTG-TTTAGTTCTGTGCGGGTTATCCCGT  2221

Query  421   TTCC-ACGAAATCCTCAGAGAGGTCCAAAATTTCTACTTGCAGTTTCCTACAGAAAGACC  479
             |||| |||||||||||||| ||||| |||| |||||||||||||| ||||||||||||
Sbjct  2220  TTCCAACGAAATCCTCAGAGCGGTCC-AAATATCTACTTGCAGTTT-CTACAGAAAGACC  2163

Query  480   GTTTCAAACCTGAAACTATCAAAGAAAGGGTTAACCACTGTGAGTTGAATGC  531
             |||||||||||| ||||||||||||||| |||| | ||||||||||||||||
Sbjct  2162  GTTTCAAACCTG-AACTATCAAAGAAA-GGTTCAACACTGTGAGTTGAATGC  2113
```

Figure 3.2: BLAST Alignments of Analysed BACs with Chromosome 17 Alphoid DNA

A 227J24
B 302P22
In both A and B the query is the sequenced alphoid DNA and the subject is human chromosome core 17 α-DNA. Both BACs show 98% homology to chromosome 17.

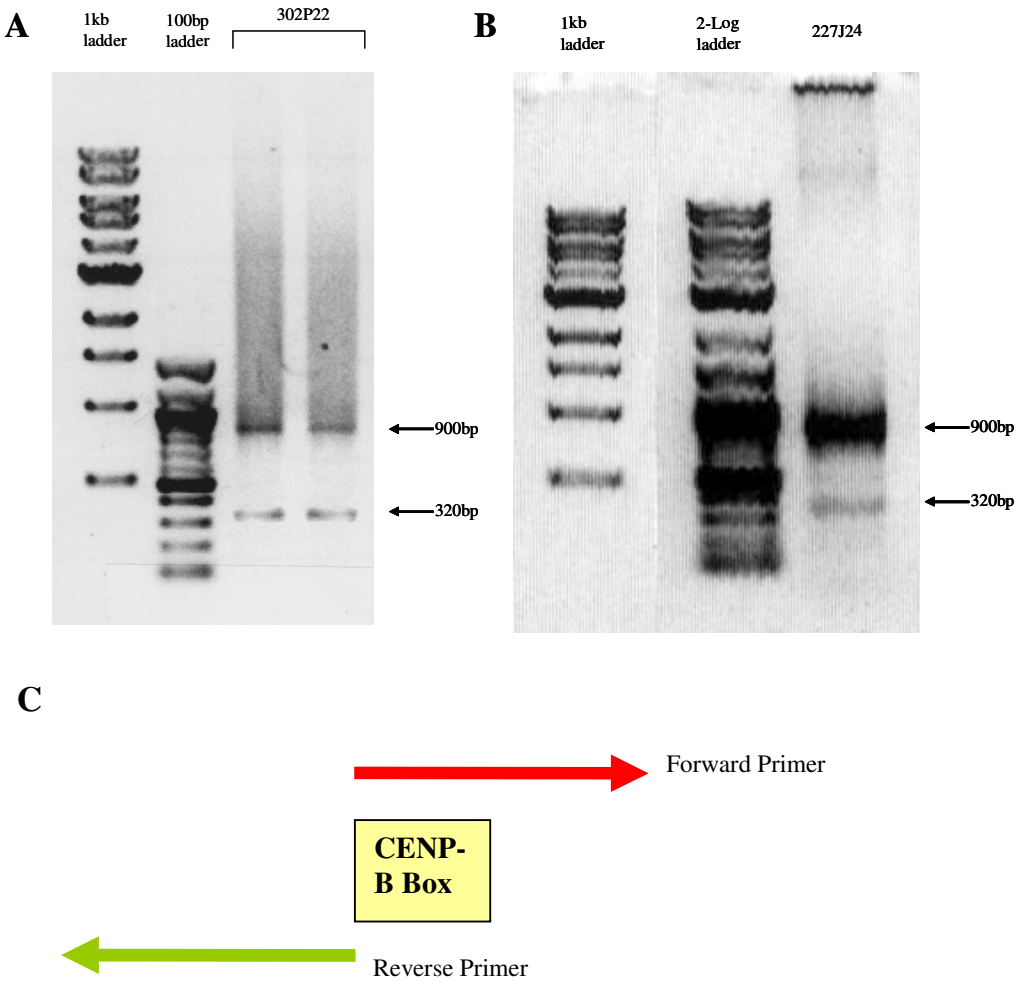


Figure 3.3: Agarose Gel Electrophoresis Showing CENP-B Box Analysis in BACs 302P22 and 227J24

A. This gel illustrates the PCR products from CENP-B box amplification in BAC 302P22.
B. This gel illustrates the PCR products from CENP-B box amplification in BAC 227J24.
Bands at 900 and 320bp can be observed in both gels indicating the presence of CENP-B boxes every other or every 6 monomers in the 17α 16-monomer based higher-order repeat structure.
C. The location of the forward and reverse primers indicating how they amplify from the start of one CENP-B box to the next.

In summary, three BAC vectors containing either chromosome 17 α-DNA (BAC 227J24, 138 kb and BAC 302P22, 57 kb) or chromosome 21 α-DNA (BAC pWTR11.32, 67kb, (Ohzeki et al., 2002)) were used in this study. The three BACs allowed a comparison of 17 and 21 alphoid templates (227J24 and 302P22 vs. pWTR11.32 respectively) and a comparison of different sized alphoid arrays (227J24 vs. 302P22 and pWTR11.32).

BAC pWTR11.32 contains a neomycin resistance gene (Neo^R) under the control of the SV40 promoter, and was modified by the vector pSG80A-HG (Saeki et al., 2001; Wade-Martins et al., 2003), containing the HSV-1 amplicon elements (ori_s and *pac*) and the GFP gene, by restriction enzyme digestion and ligation, because it lacks a *loxP* sequence. The plasmid vector pSG80A-HG, a kind gift from Yoshinaga Saeki, was linearised with SalI generating a 6 kb SalI fragment which was ligated to a 67 kb linearised SalI fragment containing the entire BAC pWTR11.32. This generated pHSV21αNeo, 73 kb, containing the HSV-1 elements and GFP gene (Figure 3.4).

As the 17α BACs 227J24 and 302P22 lacked a mammalian selectable marker, vector pSG80A-HG was modified to contain Neo^R (Moralli et al., 2006) under the control of the mouse pgk-1 promoter, generating pHGKNeo, prior to retrofitting BAC 227J24 and BAC 302P22 by Cre-*loxP* recombination (Wade-Martins et al., 2001). A NotI/Bsp120I fragment containing the Neo^R gene cassette from vector pJM1815 (Mejía and Larin, 2000) was inserted into the NotI site of pSG80A-HG. This generated pHGKNeo (8.3 kb) (Figure 3.4). pHSVα227Neo (146 kb) and pHSVα302Neo (65 kb) (Figure 3.4) were generated following Cre-*loxP* recombination of 227J24 and 302 P22 respectively with pHGKNeo.

Electrocompetent cells containing the chloramphenicol resistant BAC were transformed with pCTP-T and ampicillin resistant pHGKNeo. Cre expression was induced by incubating the cells at 30ºC for 1 hour in SOC containing heat inactivated chlorotetracycline (cTc). The cells were diluted in LB and incubated at 30ºC (permissive temperature for pCTP-T) for 3 hours before being plated at various dilutions on ampicillin-chloramphenicol plates and incubated overnight at 42ºC (restrictive temperature for the pCTP-T). Bacterial colonies were cultured, the DNA was harvested by minipreparation and checked by restriction digestion with PacI which linearises the recombined vector due to a unique site in pHGKNeo.

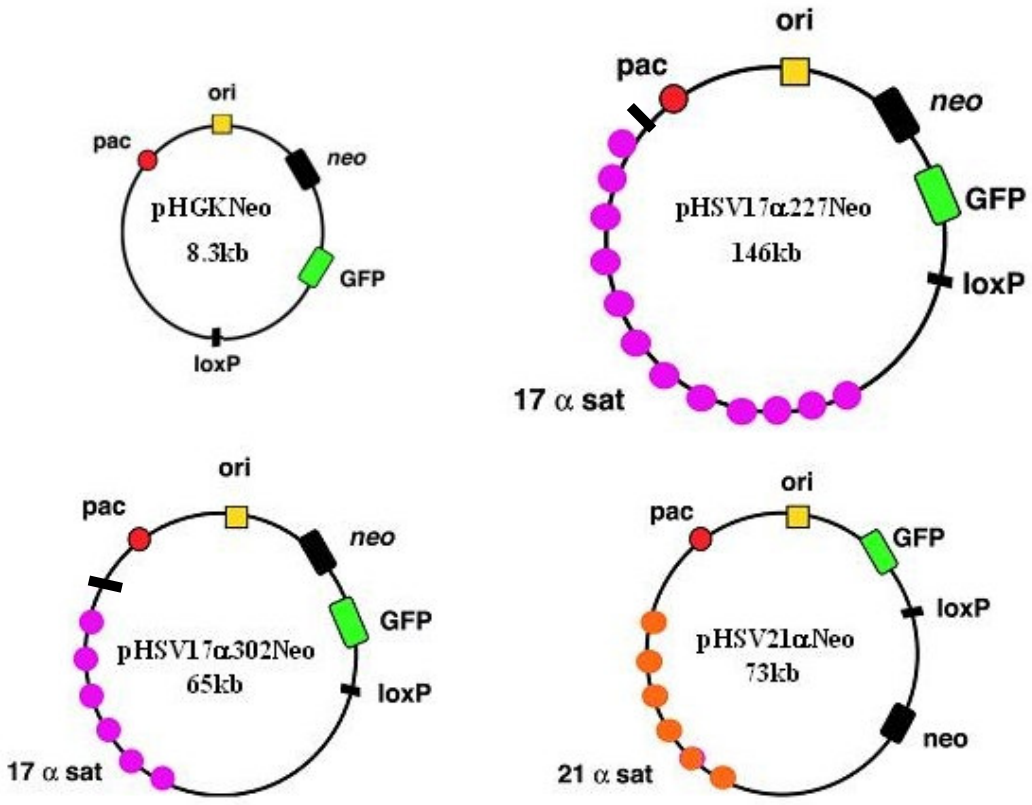


Figure 3.4: Vectors Used In Amplicon Generation
These include: pHGKNeo, containing the HSV elements *pac* (for packaging and cleavage) and ori_S (the origin of replication), the gene for GFP and the neomycin resistance (Neo^R) gene that confers resistance to G418; and the alphoid BACs pHSV17α227Neo, pHSV17α302Neo and pHSV21αNeo.

3.3 Amplicon Production and Infection of Human Cells

HSV-1 amplicons were generated containing pHSV17α302Neo and pHSV21αNeo (Figure 3.4) in Vero 2-2 cells and purified and concentrated as described in sections 1.12 and 2.13 (Wade-Martins et al., 2001). Vero 2-2 cells were transformed using LipofectAMINE Plus (Invitrogen) with fHSVΔpacΔ270+ DNA, either pHSV17α302Neo or pHSV21αNeo and pEBHICP27. The cells were incubated at 37°C for 4 hours and then the medium was removed and replaced with D-MEM containing 6% FCS. After 60 hours, the cells were harvested, sonicated, spun to remove cellular debris and concentrated on a sucrose gradient. Amplicons containing pHGKNeo and pHSV17α227Neo was work done in parallel by Daniela Moralli in the group (Moralli et al., 2006). Figure 3.5 illustrates the GFP expression following transfection of pHSV17α302Neo and pHSV21αNeo into Vero 2-2 cells after 60 hours incubation before the amplicons are harvested. The titre of the purified and concentrated preparations was estimated following the transduction of $2.5x10^5$ G16-9 (glioma) cells. The number of GFP transducing units was determined after 24 hours by counting the number of GFP positive cells. From various amplicon preparations (each preparation derived from $4x10^6$ Vero 2-2 cells) titres ranging from $5.3x10^6$-7.8 $x10^7$ units/ml were obtained.

In order to compare the efficiency of HAC formation of different alphoid templates (using pHSV17α227Neo and pHSV17α302Neo versus pHSV21αNeo) and different sized alphoid arrays (pHSV17α227Neo versus pHSV17α302Neo and pHSV21αNeo), a range of cell types were infected (HT1080 (fibrosarcoma), MRC5-V2 (lung fibroblast), 293 (kidney), MRC5 (primary foetal lung), Hep3B (hepatocarcinoma),

HeLa (cervix epithelioid carcinoma), and D98 (a HeLa derivative)) with pHSV17α302Neo and pHSV21αNeo at a range of MOIs from 1-10 (Table 3.1). HT1080, MRC5-V2, 293, BeFA (primary keratinocyte) and MRC5 cells were also transduced with pHGKNeo, pHSV17α227Neo and pHSV21αNeo amplicons in parallel experiments by Daniela Moralli (Moralli et al., 2006).

Table 3.1 shows the percentage transduction efficiency of pHSV17α302Neo amplicons monitored by GFP expression after 24hrs. As G16-9 cells were used to estimate the number of transducing units and titre of each amplicon preparation the GFP data related to this cell line is not shown in Table 3.1. GFP expression from all of the cell types discussed, transduced with the four amplicons (pHGKNeo, pHSV17α227Neo, pHSV21αNeo and pHSV17α302Neo) are shown in Figures 3.6 and 3.7.

Each amplicon was able to infect each of the cell lines studied. The efficiency of transduction ranged from 2% (BeFA MOI 1, pHGKNeo) to 100% (Hep3B MOI 5, pHSV21αNeo). On average, the number of GFP positive cells detected after infection at MOI 10 was higher than that observed at MOI 1. In some experiments the transduction efficiency was similar at MOI 1 and MOI 10 (pHSV17α302Neo amplicons in MRC5 and pHSV21αNeo amplicons in HT1080 and MRC5-V2). This could be because small volumes were used during transduction in these experiments and not all amplicons may have been able to target the cells. The HSV-1 amplicon system does not appear to be very sensitive to the size of the DNA delivered with no obvious inverse relationship between the efficiency of transduction, as monitored by transient GFP expression, and the size of the construct used (looking at

pHSV17α227Neo (146 kb) and pHSV17α302Neo (65 kb)). In HT1080, MRC5-V2 and 293 cells there appears to be a small effect visible at MOI 1. For example in HT1080 cells, the percentage of GFP expression at 24 hours, with pHSV17α227Neo is 3% and with pHSV17α302Neo is 16%. This effect is not visible at MOI 10; in HT1080 cells, the percentage of GFP expression at 24 hours, with either pHSV17α227Neo (146 kb) or pHSV17α302Neo (65 kb) at MOI 10 was 60% and 47% respectively, possibly because not all amplicons may have been able to target the cells due to the small volumes were used in these experiments. The efficiencies of transductions observed here with HSV-1 amplicons is at least 10000 fold greater than introducing large DNA as BACs and YACs by lipofection (Mejía et al., 2002).

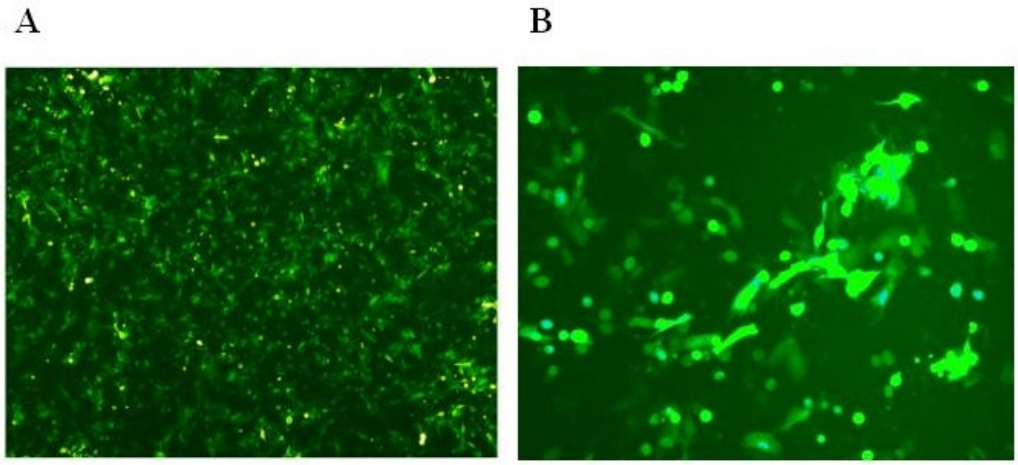


Figure 3.5: GFP Expression of pHSV17α302Neo and pHSV21αNeo in Vero 2-2 Cells

This figure illustrates the GFP expression following transfection of pHSV17α302Neo (A) (4x objective) and pHSV21αNeo (B) (10x objective) DNA into Vero 2-2 cells after 60 hours incubation before the amplicons are harvested.

Table 3.1: GFP Expression (%) of HAC-HSV Amplicons in Human Cell Lines 24 hours Post Transduction

Amplicon	HT1080		MRC5-V2		293		BeFA		MRC5		Hep3B		HeLa		D98	
	MOI 1	MOI 10	MOI 1	MOI 10	MOI 1	MOI 10	MOI 1	MOI 10	MOI 1	MOI 10	MOI 1	MOI 5	MOI 1	MOI 5	MOI 1	MOI 5
pHGKNeo	7%	64%	20%	40%	3%	40%	2%	30%	5%	40%	Nd	Nd	Nd	Nd	Nd	Nd
pHSV17α227Neo	3%	60%	21%	70%	10%	80%	8%	70%	10%	50%	Nd	Nd	Nd	Nd	Nd	Nd
pHSV17α302Neo	16%	47%	55%	98%	23%	65%	Nd	Nd	4%	7%	Nd	Nd	Nd	Nd	Nd	Nd
pHSV21αNeo	16%	20%	70%	70%	3%	30%	3%	20%	5%	10%	82%	100%	10%	41%	6%	66%

All cell lines in Table 3.1 were susceptible to amplicon infection with a range of efficiencies observed (2% (pHGKNeo, BeFA) to 100% (pHSV21αNeo, Hep3B)). The values in this table indicate the best results obtained.

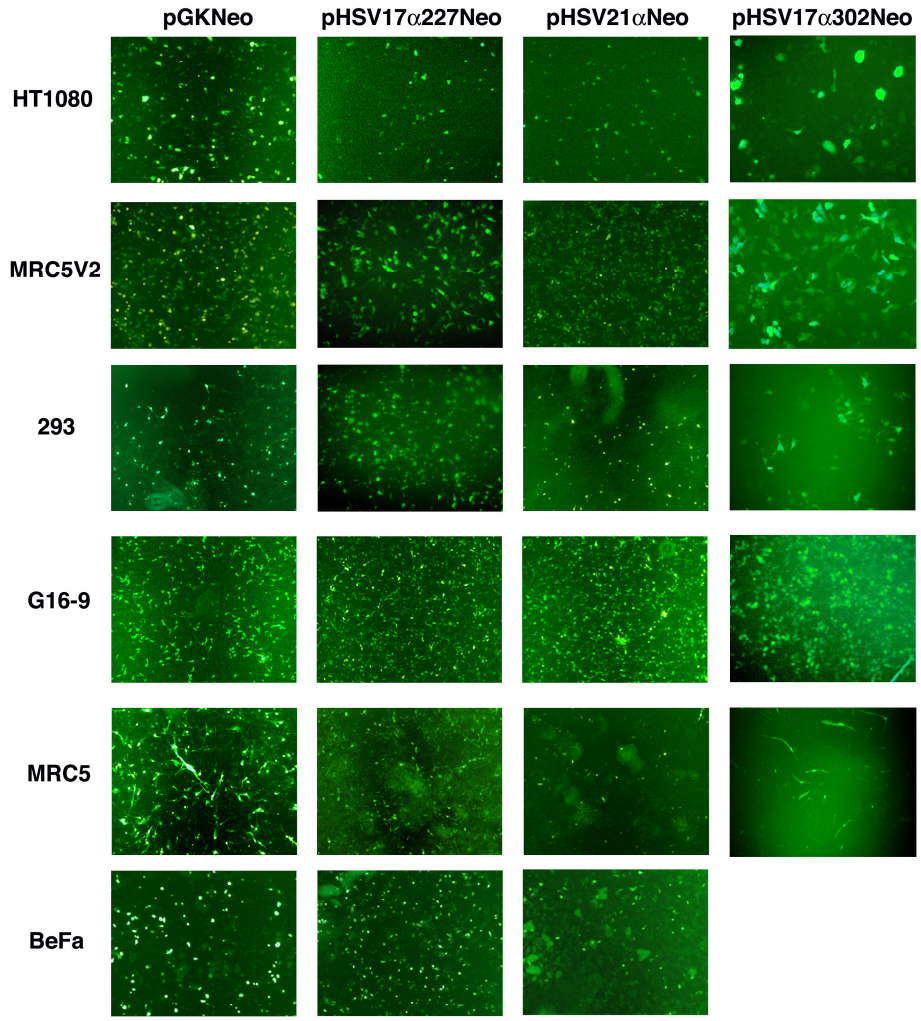


Figure 3.6: GFP Expression in Human Cells Following Transduction with HSV-1 HAC Amplicons

GFP expression in different human cell lines 24 hours after infection, following transduction with HSV-1 amplicons containing pHGKNeo, pHSV17α227Neo, pHSV21αNeo and pHSV17α302Neo. (HT1080, MRC5-V2 and 293 cells transduced with amplicons containing pHSV17α302Neo are shown at 10x objective. All other images are shown at 4x objective.)

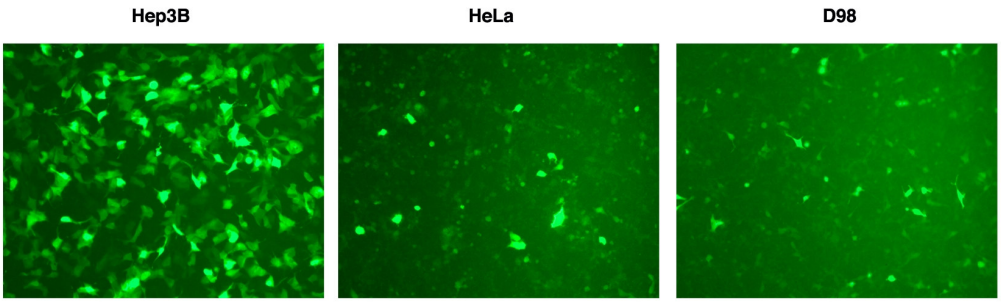


Figure 3.7: GFP Expression in Hep3B, HeLa and D98 Cells 24 hours after Infection, Following Transduction with Amplicons Containing pHSV21αNeo.

Images are shown at 4x objective.

3.4 Stable Clone Formation and FISH Analysis

Approximately 24 hours post transduction with HSV-1 amplicons, G418 selection was applied to the media to allow for the recovery of stably transformed clones. Colonies were isolated after 14-21 days from HT1080, MRC5-V2, 293 and G16-9. Following transduction with pHSV17α302Neo, twelve HT1080 clones, eight MRC5-V2 clones, twelve G16-9 clones and seven 293 clones were isolated and analysed (Table 3.2). Seven clones in G16-9 were isolated and analysed following transduction with pHSV21αNeo. No stable clones were obtained following transduction of Hep3B, HeLa and D98 cells. Possible reasons for this could have been a higher sensitivity to the infection process together with problems from bacterial contamination experienced during these experiments.

Colonies from pHSV17α227Neo and additional pHSV21αNeo transductions were isolated as described above in parallel experiments done by Daniela Moralli (Moralli et al., 2006). From pHSV17α227Neo transductions, eighteen clones were analysed from HT1080 cells, ten from MRC5-V2, thirty from G16-9 and three from 293 cells. From pHSV21αNeo transductions, twenty-one clones were characterised from HT1080 cells, twenty from MRC5-V2 and an additional nine clones from G16-9 cells (Table 3.2).

The clones were expanded and metaphase spreads were prepared from cells and dropped onto glass slides. Clones were screened by FISH using either a 17α or 21α probe to determine the presence of HACs and the results were confirmed using the

vector pHGKNeo as a probe. The analysis of positive clones and frequency of HAC formation resulting from the stable clones are shown in Table 3.2.

Table 3.2: Generation of HAC Clones Following HSV-1 Transduction with Amplicons Containing pHSV17α227Neo, pHSV17α302Neo and pHSV21αNeo.

Cell line	Amplicon	Clones Analysed	HAC (% of spreads)	Other Events
HT1080	pHSV17α227Neo	18	4 (6-26%)	14
	pHSV17α302Neo	12	7 (10-30%)*	5
	pHSV21αNeo	21	3 (7-15%)*	18
MRC5-V2	pHSV17α227Neo	10	9 (5-35%)	1
	pHSV17α302Neo	8	5 (5-35%)*	3
	pHSV21αNeo	20	4 (5-40%)*	16
G16-9	pHSV17α227Neo	30	5 (8-65%)*	25
	pHSV17α302Neo	24	14 (5-40%)*	10
	pHSV21αNeo	16	2 (16-50%)*	14
293	pHSV17α227Neo	3	2 (10-13%)	1
	pHSV17α302Neo	7	0	7
* some of these clones also contained an integration into a host chromosome				

Stably transformed clones were obtained after transduction with pHSV17α302Neo and pHSV17α227Neo in HT1080, MRC5-V2, G16-9 and 293 cells. HACs were present in a proportion of these clones and clones containing only integrations and no artificial chromosomes were obtained in each of the cell lines. Stably transformed clones were obtained after transduction with pHSV21αNeo in HT1080, MRC5-V2 and G16-9 cells.

For example, in HT1080 cells transduced with pHSV17α302Neo, 12 clones were analysed. 7 of these clones contained HACs varying from 10-30% in the metaphase spreads analysed, with some of these metaphases also containing integrations. 5 of these clones contained integrations only.

Twelve clones derived from HT1080, following transduction with pHSV17α302Neo amplicons, were analysed. Seven of these clones (58%) contained a HAC present in 10-30% of metaphase spreads analysed and three of these clones also contained an integration into a host chromosome. In addition, five separate clones contained integration events.

Five out of eight clones (63%) derived from MRC5-V2, following transduction with pHSV17α302Neo amplicons, contained a HAC present in 5-35% of metaphase spreads analysed and two of these clones also contained an integration into a host chromosome. Three separate clones contained integration events.

Twenty-four clones derived from G16-9, following transduction with pHSV17α302Neo amplicons, were analysed. Fourteen of these clones (58%) contained a HAC present in 5-40% of metaphase spreads analysed and eight of these clones also contained an integration into a host chromosome. In addition, ten separate clones contained integration events.

Seven clones derived from 293, following transduction with pHSV17α302Neo amplicons, were analysed. All of these clones contained integration events and no HACs. One possible reason for this could be that any HACs formed were unstable and lost by the time clones were expanded for analysis.

One clone out of seven, derived from G16-9, following transduction with pHSV21αNeo amplicons, was found to contain 50% HACs whilst the remaining six

contained integrations. In Table 3.2 these clones are combined with data obtained in parallel experiments.

The additional analysis of positive clones and the frequency of HAC formation from cell lines infected with pHSV17α227Neo and pHSV21αNeo in experiments by Daniela Moralli (Moralli et al., 2006) is shown in Table 3.2. Again, these clones were screened by FISH with either a 17α or 21α probe to determine the presence of HACs and this was confirmed with a pHGKNeo probe. Stably transformed clones were obtained after transduction with pHSV17α227Neo in HT1080, MRC5-V2, G16-9 and 293 cells. HACs were present in a proportion of these clones, ranging from 17-90%. In some of the clones, integrations were also observed. Clones containing only integrations and no artificial chromosomes were obtained in each of the cell lines. Stably transformed clones were obtained after transduction with pHSV21αNeo in HT1080, MRC5-V2 and G16-9 cells. The proportion of HAC containing clones ranged from 13-20% with integrations also observed in both HAC and non-HAC containing clones from all of these cell lines.

HACs were detected in HT1080, MRC5-V2, G16-9 and 293 cells (Figure 3.8a) at a range of frequencies (5-65%). The observed percentage of HACs was not necessarily as high as one would expect. There might have been smaller HACs that escaped detection by FISH, or negative cells which survived G418 selection due to bystander effects being present. The pHSVα302Neo vector generated HACs in HT1080, MRC5-V2 and G16-9 cells at a relatively high frequency. For example, 58%, 63% and 58% of the analysed clones respectively contained HACs in a percentage of spreads ranging from 5% to 40%. Fig 3.8a illustrates HACs formed from

pHSV17α302Neo amplicons in HT1080, MRC5-V2 and G16-9 cells. Only clones containing integration events were obtained in 293 cells following the transduction of pHSV17α302Neo.

Three of the G16-9 clones, GT1, GT9 and GT10, and one HT1080 clone, HT5.11, were subcloned to generate sublines containing a higher percentage of HACs (Table 3.3). The clones GT1 and GT10 were chosen due to the high percentage of HACs (40 and 25% respectively) detected in metaphase spreads. GT9 had a HAC frequency of only 10% but contained no DNA that had integrated into the host chromosomes. HT5.11 contained 30% HACs. Analysis of thirty-three subclones revealed HACs at different percentages with the subline GT9SC7 containing 80% HACs and the subline HT5.11SCD8 containing 70% HACs. Again, HACs were not observed in 100% of cells possibly due to non-detection by FISH, or surviving negative cells.

Table 3.3: Generation of Subclones from Clones Transduced with Amplicons Containing pHSV17α302Neo.

Parent Clone	Clones Analysed	HAC (% of spreads)	Integration (% of spreads)	Other Events
GT1	11	8 (10-40%)	2 (10-40%)	1 clone containing 10% HACs and 20% integrations
GT9	13	6 (10-80%)	6 (10-30%)	1 clone containing 20% HACs and 10% integrations
GT10	4	2 (10-40%)	2 (10%)	0
HT5.11	5	5 (20-70%)	0	0

Three G16-9 clones (GT1, GT9 and GT10) and one HT1080 clone (HT5.11), all stably transduced with pHSV17α302Neo, were subcloned to generate sublines. Stable subclones were generated from all four clones.

For example, in clones generated from the G16-9 clone GT1, 11 clones were analysed. 8 of these clones contained HACs varying from 10-40% in the metaphase spreads analysed. 2 of these clones contained integrations only (either in 10 or 40% of metaphases) and 1 clone contained both a HAC and an integration.

In HT1080, pHSV17α227Neo formed HACs in 22% of clones with frequencies ranging from 6-26% and pHSV21αNeo formed HACs in 14% of clones with frequencies ranging from 7-15%. pHSV17α227Neo was able to form HACs at a high efficiency in MRC5-V2 cells (90%) with the percentage of HACs in metaphase spreads ranging from 5-35% (Figure 3.8a). In contrast pHSV21αNeo formed HACs in 20% of clones with a similar range of frequencies observed in metaphase spreads (5-40%) (Figure 3.8a). In 293 cells, the frequency of HAC containing clones with

pHSV17α227Neo was 67% and HACs were present in 10-13% of spreads from each clone. Both pHSV17α227Neo and pHSV21αNeo formed HACs at similar efficiencies (17% and 13% respectively) in G16-9 cells. HACs ranged from 8-65% in positive clones. The clone 17α III.10 derived from pHSV17α227Neo in G16-9 cells, containing a contaminating sub-population with integrated DNA, was subcloned to generate sublines containing a higher percentage of HACs. One subline (17α III.10.1) was generated containing only HACs in 85% of metaphases analysed.

In summary, HAC containing clones were recovered after transduction with amplicons containing pHSV17α227Neo, pHSV17α302Neo and pHSV21αNeo from HT1080, MRC5-V2, and G16-9 cells. HAC containing clones were recovered from 293 cells following transduction with amplicons containing pHSV17α227Neo.

3.5 ACA and CENP-A Detection

The ability of HACs to bind centromeric proteins was firstly shown by immunofluorescence on metaphase chromosome spreads with anti-centromere autoimmune serum (ACA; a kind gift from William Earnshaw) since the antibodies present work well on methanol:acetic acid chromosome suspensions. The serum reacts with many of the major centromere proteins at a centromere, both necessary and non-necessary, so whenever possible, the HACs were also tested with anti-human CENP-A antibody which recognises the protein CENP-A, an essential centromere component (Figure 3.8b). Centromere specific histone-like protein was present on analysed MRC5-V2 and G16-9 clones generated with pHSVα302Neo (Figure 3.8b).

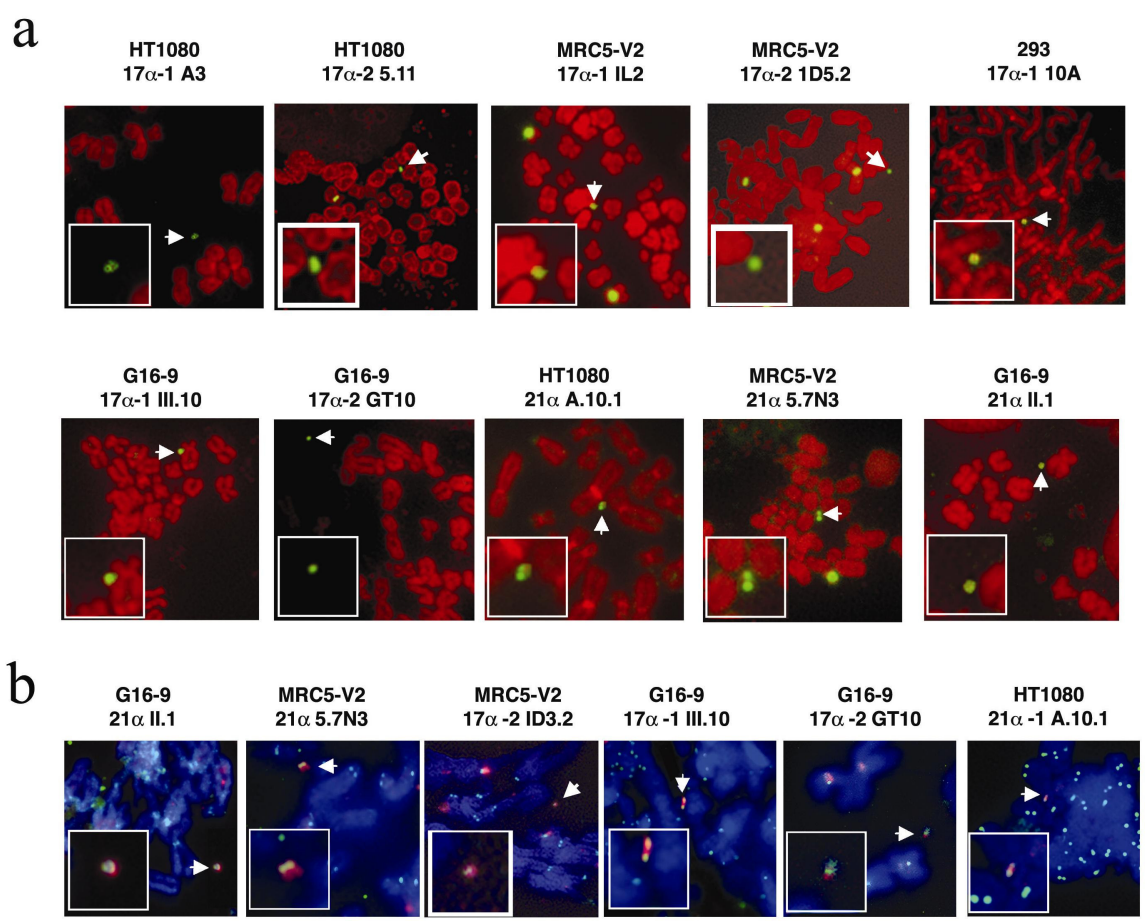
However, the signal was lower than on the other chromosomes in the MRC5-V2 HAC.

Co-localization of the 21α probe and ACA on the HACs in 21α.II.1 (G16-9) and 21α.57N3 (MRC5-V2) (Figure 3.8b) was observed. The ACA signal appeared to be localised to a distinct area, suggesting the presence of a discrete centromere. The rapid loss of the HACs in MRC5-V2, even in the presence of selection, made CENP-A analysis difficult. Two clones, 21α A.10.1 (HT1080) and 17α III.10 (G16-9) were analysed further. Using anti-human CENP-A antibodies, it was demonstrated that CENP-A is present on the HAC (Figure 3.8b) in both clones, as well as in G16-9 clones obtained from pHSV17α302Neo transduction. The signal intensity appeared to be at a similar level to the signal at endogenous centromeres, suggesting that the HAC formed a functional centromere.

Figure 3.8: Analysis of HAC Containing Cell Lines by FISH and Immunofluorescence

a. HAC clones generated in HT1080, MRC5-V2, G16-9 and 293 cells following infection with pHSV17α227Neo (17α-1), pHSV17α302Neo (17α-2) and pHSV21αNeo amplicons. Host chromosomes are in red, and the HACs were identified by FISH using either a 17α/21α or a pHGKNeo probe.

b. ACA and CENP-A staining of the HAC clones in HT1080, MRC5-V2 and G16-9 derived from pHSV17α227Neo, pHSV17α302Neo and pHSV21αNeo amplicons. The HACs in 21αII.1 (G16-9) and 21α5.7N3 (MRC5-V2) were identified by FISH with a 21α probe (red), and the signal co-localized with the ACA signal (green, white arrow; see inset). Chromosomes were counterstained with DAPI, (blue). The HACs in clones 17α-2 1D3.2 (MRC5-V2), 17α-1 III.10 (G16-9), 17α2 GT10 (G16-9) and 21αA.10.1 (HT1080) were identified by FISH with the 17α or 21α probes (red, white arrow) respectively, and the signal co-localized with CENP-A signal (green, white arrow; see inset). Chromosomes were counterstained with DAPI (blue).



3.6 Mitotic Stability

The mitotic stability of six HAC containing clones derived from pHSV17α302Neo was investigated: one from HT1080 (HT5.11), one from MRC5-V2 (1D3.2) and four from G16-9 cells (GT1, GT7, GT10 and GT2.1). Cells from each clone were passaged under non-selective conditions, in the absence of G418, and HACs were detected by FISH with 17α and pHGKNeo probes at 0, 30, 60 and 90 days in culture (Table 3.4). 50 metaphases from each clone were analysed. The mitotic stability of HACs derived from pHSV17α227Neo and pHSV21αNeo in either HT1080, MRC5-V2, 293 and G16-9 clones was also investigated in parallel experiments by Daniela Moralli (Table 3.4).

The HACs in HT5.11 (derived from pHSV17α302Neo in HT1080) and 21α A10.1 (derived from pHSV21αNeo in HT1080) were stable in the absence of selection. The percentage of HACs present in chromosome metaphase spreads in HT5.11 remained at 30% from days 0 to 90 and did not change from day 0 (40%) to day 60 (40%) in 21α A10.1, with only a slight decrease at day 90 (37.5%) (Table 3.4). The overall loss rate per day of 0.07% indicated that the HAC segregated efficiently in over 99% of metaphases. In comparison, the HAC in 21α 57N3 (derived from pHSV21αNeo in MRC5-V2) was lost within 30 days (Table 3.4), despite having bound ACA. The HAC in 1D3.2 (derived from pHSV17α303Neo in MRC5-V2) was also unstable with a daily loss rate of 6% after 30 days and it was completely lost after 60 days. HACs in other MRC5-V2 and 293 clones behaved similarly. In 17α GT1 (derived from pHSV17α302Neo in G16-9), it was observed that at day 0-90 HACs were present in 40% of chromosome spreads (Table 3.4) similar to the HAC frequency obtained in

pHSV21αNeo (HT1080) indicating that the HACs formed and maintained a fully functional centromere in G16-9 cells.

Table 3.4: Stability of HACs in the Absence of Selection

Clone	HAC (%) Off Selection (Days)				% Loss rate[a]
	0 Days	30 Days	60 Days	90 Days	
HT5.11 (HT1080)	30	30	30	30	0
1D3.2 (MRC5-V2)	35	5	0	0	6[b]
GT1 (G16-9)	40	40	40	40	0
GT7 (G16-9)	20	20	20	20	0
GT10 (G16-9)	25	20	20	20	0.74
GT2.1 (G16-9)	25	20	20	20	0.74
α21 A.10.1 (HT1080)	40	40	40	37.5	0.07
α21 57N3 (MRC5V2)	35	0	-	-	nd[c]
α17 G17 III.10 (G16-9)	65	20	15	15	0.9-3.8

a Calculated by the formula: $N_n=N_0 \times (1-R)^n$ (Mejía et al., 2001), where N_0 is the number of metaphase chromosome spreads showing HACs in the cells cultured under selection, N_n is the number of HAC-containing metaphase chromosome spreads after n days of culture in the absence of selection and R is the daily rate of loss.
b Calculated from 0-30 days.
c Not determined.

3.7 Investigation of HAC Frequency Following Transfection with the 17α BAC Vector pHSV17α302Neo

A series of lipofection experiments of pHSV17α302Neo into the HT1080 cell line were carried out to check the HAC forming frequency by standard techniques. Twelve G418 resistant clones were obtained in total. These clones were analysed by FISH as described above. The results indicated that 67% of the clones obtained contained HACs ranging from 5-25% of metaphase spreads analysed. In addition, 67% of the clones contained an integration of vector DNA into a host chromosome (ranging from 5-20% of metaphase spreads analysed), with 63% of these events being in clones that also contained a HAC. One clone designated HTL3 contained HACs in 25% of metaphase spreads and no integrations as identified with a 17α FISH probe (Figure 3.9A). The HAC in this clone was analysed further with ACA antisera (Figure 3.9B), which reacts with the major centromere proteins at centromeres and was found to be positive for ACA staining.

In other parallel experiments by Daniela Moralli, HACs were generated with pHSV17α227Neo and pHSV21αNeo in HT1080 cells by lipofection. No significant difference in HAC formation frequency was observed compared to HAC clones derived from HSV-1 delivery (Moralli et al., 2006).

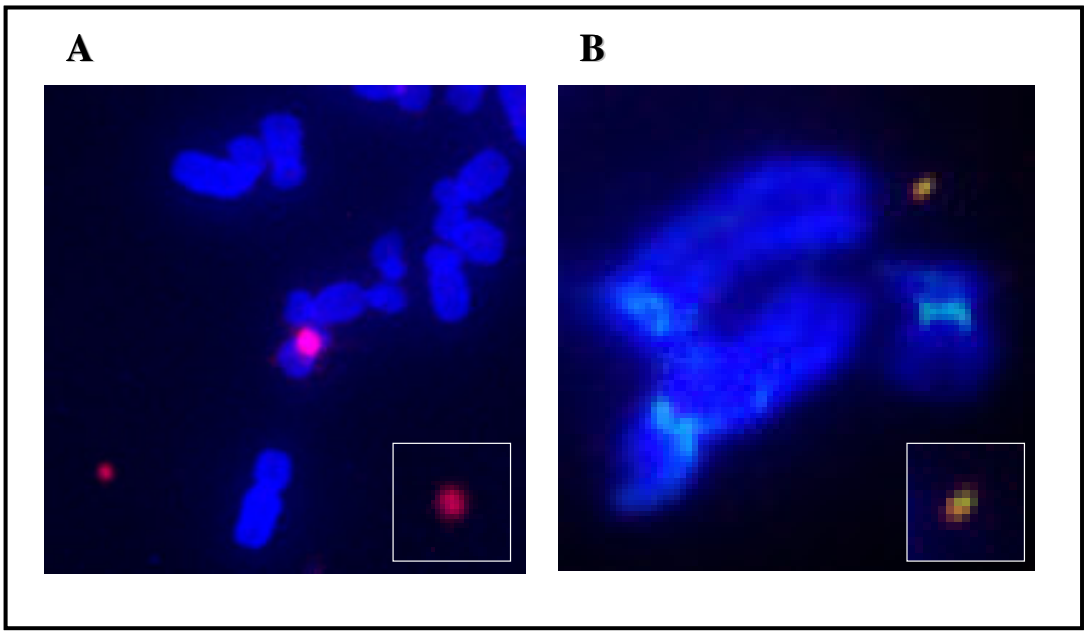


Figure 3.9: Analysis of HAC Containing Clone HTL3 by FISH and Immunofluorescence

A. HAC clone HTL3 (HT1080) derived following transfection with pHSV17α227Neo. Chromosomes are in blue, and the HACs were identified using a 17α probe, red.

B. ACA staining of the centromere of a HAC in HTL3. The HAC is identified by FISH with a 17α probe, in red, while the ACA signal is green. Chromosomes are counterstained in DAPI, blue.

3.8 Conclusions and Discussion

Our group has successfully generated HSV-1 amplicons containing HAC vector DNA and infected several human cell types including human HT1080 cells which are commonly used for HAC studies (Grimes et al., 2002; Harrington et al., 1997; Mejía et al., 2002). The efficiency, determined by transient expression of the reporter gene, GFP, generally depended on the cell type and did not always increase proportionally with an increase in MOI. This could be important for a gene therapy approach as a smaller number of transducing particles may be just as efficient as a large number. HAC containing cell lines were established in HT1080, MRC5-V2 and G16-9 cells following HSV-1 transduction, and the efficiency of generating stable clones ranged from 10^{-2}–10^{-5}. In contrast, the efficiency of clone formation in HT1080 following the introduction of the HAC vectors via lipofection is typically 10^{-5}-10^{-6}, similar to that observed in other studies (Mejía et al., 2002; Mejía et al., 2001).

HACs in HT1080 and G16-9 cells reacted to human ACA and anti-human CENP-A antibodies suggesting that they formed an active centromere. HACs in HT1080 cells were stable up to 90 days in the absence of selection and the HAC stability was comparable to HACs in HT1080 cell lines generated following lipofection, indicating the formation of a fully functional centromere. In G16-9 cells the HACs also appeared to be stable. However, in MRC5-V2 and 293 cells, the HACs were unstable in the presence and absence of selection indicating the possible inability of HACs to maintain a functional centromere in these cells despite ACA and CENP-A staining (Figure 3.8b). In the presence of selection, freely segregating HACs were lost, and the DNA integrated into the host chromosome thus allowing the cells to survive.

These results suggest that the requirements for centromere function and HAC stability are different in each cell type and this is investigated further in the following chapter.

HSV-1 is an effective method for HAC vector delivery. It offers a safe alternative to viral vector systems in which cytotoxic and cytopathic effects have been observed in the recent gene therapy trials (Check, 2003). No adverse reactions have been reported from the use of the disabled HSV-1 system (Olschowka et al., 2003). This study has shown that HACs can be generated with a reasonable efficiency in human cell types other than HT1080 compared to previous delivery methods used. Using this technology, differences in HAC stability between the cultured cell types have been identified and further work is required to identify the factors involved. This is the focus of the next results chapter.

Chapter 4:

Results II: Investigation of a Possible Candidate for HAC Instability

4.1 Introduction

As discussed in Chapter 1, a large range of proteins are involved in centromere maintenance and may thus influence HAC stability. CENP-A is a histone-H3 related protein localised at the inner kinetochore plate of mammalian mitotic chromosomes (Palmer et al., 1991). Associated with all active centromeres (Warburton et al., 1997), its inactivation disrupts mitosis and cell cycle progression and causes the mislocalisation of many kinetochore and centromere-region proteins, including CENP-C in yeasts, worms, flies and mice (Blower and Karpen, 2001; Howman et al., 2000; Moore and Roth, 2001) implying that CENP-A is at the top of a pathway driving kinetochore assembly and evidence suggests that it is a strong candidate for an epigenetic mark of kinetochore assembly. CENP-B (Masumoto et al., 2004) localises to human and mouse centromeres through direct binding to the 17 bp CENP-B box sequence (Masumoto et al., 1989). There is a lack of obvious mitotic and meiotic phenotype in Cenpb knockout mice (Fowler et al., 2004; Hudson et al., 1998) but despite this, both the CENP-B box and higher alphoid DNA are required for *de novo* artificial chromosome formation and the assembly of functional centromere components such as CENP-A, CENP-C and CENP-E (Basu et al., 2005; Ohzeki et al., 2002). CENP-C is only associated with active centromeres throughout the cell cycle and is essential for proper mitotic segregation and cell survival (Kalitsis et al., 1998). CENP-E protein is a kinesin-like motor protein which is localised to active

centromeres only (Sullivan and Schwartz, 1995; Yen et al., 1992). When depleted from mammalian kinetochores, mitotic arrest with a mixture of aligned and unaligned chromosomes results (McEwen et al., 2001). CENP-F is a facultative mammalian kinetochore protein and its depletion leads to misaligned chromosomes with a phenotype similar to that seen in CENP-E depleted cells (Yang et al., 2005). Aurora B is a serine/threonine protein kinase chromosomal passenger protein with a variety of roles in chromosome maintenance and a variety of phosphorylation targets (Vagnarelli and Earnshaw, 2004). DNA Topoisomerases localise to the centromere and are implicated in a range of functions including DNA replication, transcription, recombination, chromosome condensation, segregation and chromatid resolution (Corbett and Berger, 2004).

The results described in Chapter 3 suggest that the requirements for centromere function and HAC stability may be different in each cell type. The mitotic stability of HACs generated in HT1080, MRC5-V2, 293 and G16-9 clones were investigated in the absence of selection. Cells from nine clones were passaged under non-selective conditions, in the absence of G418, and HACs were detected by FISH with either a 17α or 21α and pHGKNeo probes at 0, 30, 60 and 90 days in culture (Table 3.4). HACs in HT1080 and G16-9 cells were stable for 90 days, but in MRC5-V2 and 293 cells the HACs were unstable although ACA/CENP-A binding to these HACs was observed (Table 3.4, Figure 3.8b). To investigate this further the level of gene expression of possible candidates involved in kinetochore/centromere function (CENP-A, -B, -C, -E and –F, sections 1.7.1 to 1.7.5), chromatid resolution (Topo I), replication (Kin17) and cell cycle checkpoints (Aurora B, section 1.7.9) were compared in the different cell types using quantitative PCR analysis of total cell

cDNA. Kin17 is a DNA-binding protein involved in replication that interacts with the SV40 T antigen and inhibits replication (Miccoli et al., 2002). As MRC5-V2 cells are transformed with the large T antigen, over-expression of Kin17 in this cell type could account for the observed HAC instability.

In summary, real-time PCR analysis, used as it allows for the accurate quantitation of gene expression, showed that Aurora B was over expressed in MRC5-V2 and 293 cells. Western blot analysis confirmed these results and high levels of chromosomal instability in MRC5-V2 and 293 cells correlated with the findings. To further confirm the effect of cellular phosphorylation levels on chromosome stability, cells were treated with Okadaic Acid (OA) which is both a protein phosphatase inhibitor and activates Aurora B (Sugiyama et al., 2002). It was hypothesised that keeping stable HACs under a constant low dosage of okadaic acid would result in a decrease in their stability. Treatment of cells showed an increase in both HAC and overall chromosomal instability and an increase in H3 phosphorylation at Serine 10 and 28. Among the targets of OA, PP2A was investigated as a possible candidate for the cause of the increase in instability and increase in H3 phosphorylation but results were inconclusive, with no clear differences observed, indicating that other factors may be involved.

4.2 Real-Time Quantitative PCR

Traditional PCR uses agarose gels to detect PCR amplification at the end-point of the reaction. The end-points may be variable from reaction to reaction and the size discrimination of the band intensity, on which the results are based, can only detect a

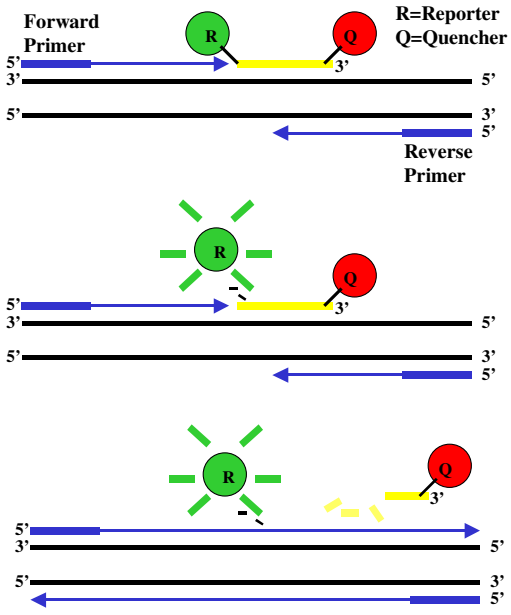
ten-fold difference. Real-time PCR can detect as little as a two-fold change. Real-time PCR has the ability to collect data in the exponential growth phase of the PCR reaction and monitor the process of the PCR as it occurs (i.e., in real time) thus providing accurate quantitation results. This allows for the accurate quantitation of gene expression using cDNA.

ABI have two sequence detection methods: SYBR Green Dye I chemistry and TaqMan Chemistry. The SYBR Green Dye I system detects all double-stranded DNA, including non-specific reaction products. For this reason the TaqMan system was chosen, particularly as any amplification from genomic DNA was undesirable.

In addition to the components of a typical PCR, in a TaqMan assay, an oligonucleotide probe with both a reporter fluorescent dye at the 5' end and a quencher dye at the 3' end is used in real-time PCR. While the probe is intact, the proximity of the quencher reduces the fluorescence emitted by the reporter dye by Förster Resonance Energy Transfer (FRET). During amplification, annealing of the probe to its target generates a substrate that is cleaved by the 5' fork-like, structure dependant, polymerisation-associated, nuclease activity of the Taq DNA polymerase when the enzyme extends from an upstream primer into the region of the probe. This cleavage separates the reporter dye from the quencher dye, thus increasing the reporter dye fluorescence intensity. Cleavage removes the probe from the target so it does not interfere with the PCR process. Additional reporter dye molecules are cleaved from their probes with each cycle. This gives an increase in fluorescence intensity proportional to the amount of amplicon produced (Figure 4.1). The combination of FRET and the 5' nuclease activity of the Taq polymerase enables data

to be collected in real-time. When the fluorescent signal from the reporter dye increases to a detectable level it can be captured and displayed as an amplification plot. Reactions are characterised by the point in time at which amplification is first detected above the background. This is known as the threshold and is set in the exponential phase of the reaction. The higher the starting copy of the target, the sooner a significant increase in fluorescence is observed and the sooner a reaction reaches the threshold level. The cycle at which the sample reaches this level is known as the C_T (Threshold Cycle) value, defined as the fractional cycle number at which the fluorescence passes the fixed threshold. The higher the amount of starting material is, the lower the C_T.

Using this TaqMan assay, the expression of several proteins involved in centromere/kinetochore maintenance have been analysed in several cell types.



Polymerisation and Strand Displacement

Two fluorescent dyes, a reporter (R) and a quencher (Q), are attached to the 5' and 3' ends of a probe. When both dyes are attached to the probe, reporter dye emission is quenched.

Cleavage

During each extension cycle, the *Taq* DNA polymerase cleaves the reporter dye from the probe.

Polymerisation Completed

Once separated from the quencher, the reporter dye emits its characteristic fluorescence.

Figure 4.1: The Fluorogenic 5' Nuclease Chemistry Involved in a Real-Time PCR Reaction

Stepwise representation of *Taq* DNA polymerase acting on a fluorogenic probe during one extension phase of PCR.

4.3 cDNA Preparation and Real-Time PCR

RNA was harvested from one confluent plate of each of the following cell lines: HT1080, HPRT deficient (HPRT-) HT1080, HT1.1 (a tetraploid HT1080 strain efficient at HAC formation (Grimes et al., 2004), a gift from Brenda Grimes), MRC5-V2, 293 and G16-9. The HPRT- and HT1.1 HT1080 strains were studied as a comparison to the HT1080 wild type cells as HACs are known to form in all three strains. Commercial MRC5 RNA was used due to difficulties in obtaining enough RNA from cultured cells. cDNA was prepared from 2 μg of RNA using Qiagen's QuantiTect Reverse Transcription Kit according to the manufacturer's instructions. The resulting cDNA was diluted 1:25 and 3 μl was used in each real-time PCR reaction with 7.5 pmoles of each of forward and reverse primer, 2.5 pmoles of fluorogenic probe (labelled with FAM at the 5' end and TAMRA at the 3' end) and 2x TaqMan Universal PCR Mastermix (ABI) in a total volume of 25 μl/well. Primers and probes were designed using Primer Express and where possible where chosen to overlap exon boundaries to ensure no amplification from contaminating genomic DNA (see section 2.17.4). Real time reactions were carried out on an ABI PRISM 7700 machine using universal thermal cycling parameters (see Table 2.4).

4.4 Real-Time PCR Data Analysis

A relative quantitation assay using the comparative C_T method was used to analyse changes in gene expression of the genes of interest relative to the reference gene, GAPDH (Glyceraldehyde-3-Phosphate Dehydrogenase).

Firstly, standard curves were generated for all sets of primers to ensure that they were efficient over a series of cDNA concentrations and also to ensure that the efficiency of the target amplification and GAPDH control were approximately equal. Standard curves were graphically represented as plot of C_T values vs. log of input nucleic acid. Slopes in the range of –3.1 to –3.3 were obtained where a slope of –3.32 indicates 100% efficiency (a 100% efficient reaction will yield a ten-fold increase in amplicon every 3.32 cycles during the exponential phase of amplification).

Three cDNA preparations from each cell line were obtained (with the exception of MRC5 as this obtained from commercial RNA) and twelve replicates of each cDNA sample were run.

Average C_T values were then obtained from each protein of interest and normalised to GAPDH, a control housekeeping gene involved in glycolysis, to give ΔC_T.

$$\Delta C_T = C_{T\ target} - C_{T\ GAPDH}$$

Normalisation to an endogenous control allows for the correction of results that can be skewed by differing amounts of nucleic acid template.

These values were then made comparative to the primary fibroblast cell line, MRC5. This results in $\Delta\Delta C_T$.

$$\Delta\Delta C_T = \Delta C_{T\ target} - \Delta C_{T\ MRC5}$$

The target protein gene expression level, relative to the MRC5 cell line is then determined by evaluating the expression $2^{-\Delta\Delta CT}$.

Data analysis is shown in Tables 4.3-4.10 in the appendix to this chapter and the correlating graphs in Figures 4.2-4.9. Each column is an average of the relative expression (of 3 different samples) of the protein in the cell type relative to that of MRC5. For example: the MRC5 expression of CENP-A is 1 and the HT1080 expression is 7. Therefore HT1080 has 7 times as much CENP-A as MRC5. The bars on each data column give the range of the gene expression as evaluated by $2^{-\Delta\Delta CT}$ with $\Delta\Delta C_T$+s and $\Delta\Delta C_T$-s, where s= the standard deviation of $\Delta\Delta C_T$. The lowest and highest values from each of the three samples are indicated.

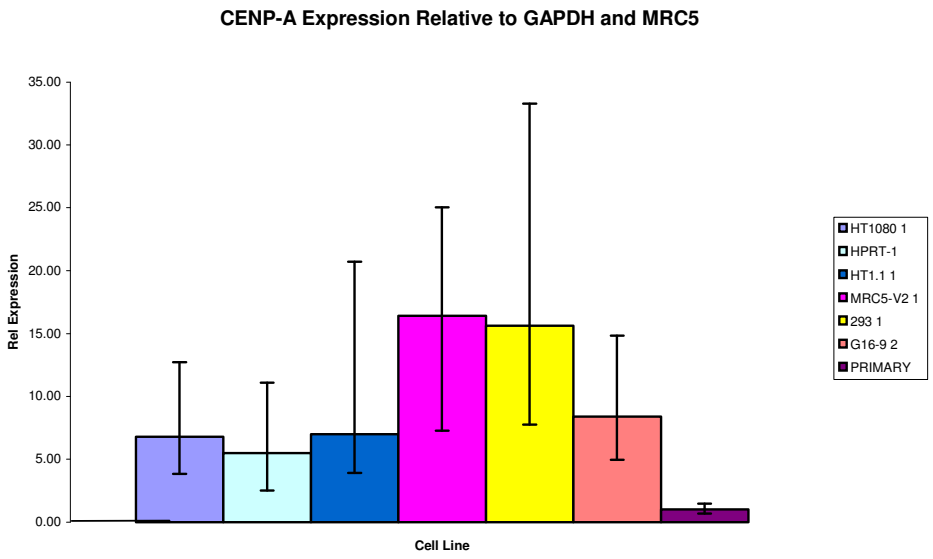


Figure 4.2: CENP-A Expression Relative to GAPDH and MRC5
The bars on each data column give the range of the gene expression as evaluated by $2^{-\Delta\Delta CT}$ with $\Delta\Delta C_T$+s and $\Delta\Delta C_T$-s, where s= the standard deviation of $\Delta\Delta C_T$. The lowest and highest values from each of the three samples are indicated.

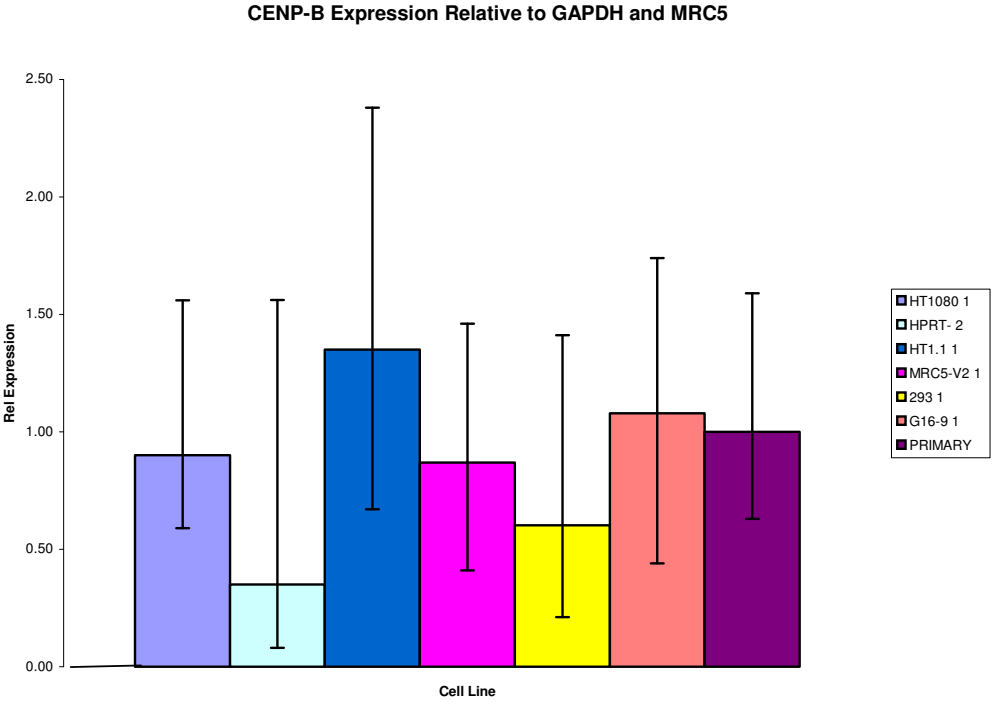


Figure 4.3: CENP-B Expression Relative to GAPDH and MRC5
The bars on each data column give the range of the gene expression as evaluated by $2^{-\Delta\Delta CT}$ with $\Delta\Delta C_T$+s and $\Delta\Delta C_T$-s, where s= the standard deviation of $\Delta\Delta C_T$. The lowest and highest values from each of the three samples are indicated.

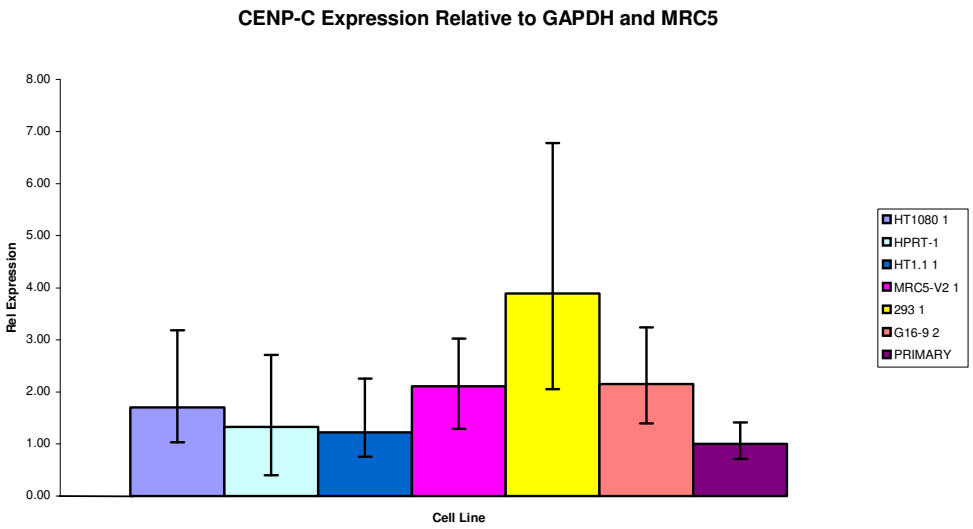


Figure 4.4: CENP-C Expression Relative to GAPDH and MRC5
The bars on each data column give the range of the gene expression as evaluated by $2^{-\Delta\Delta CT}$ with $\Delta\Delta C_T$+s and $\Delta\Delta C_T$-s, where s= the standard deviation of $\Delta\Delta C_T$. The lowest and highest values from each of the three samples are indicated.

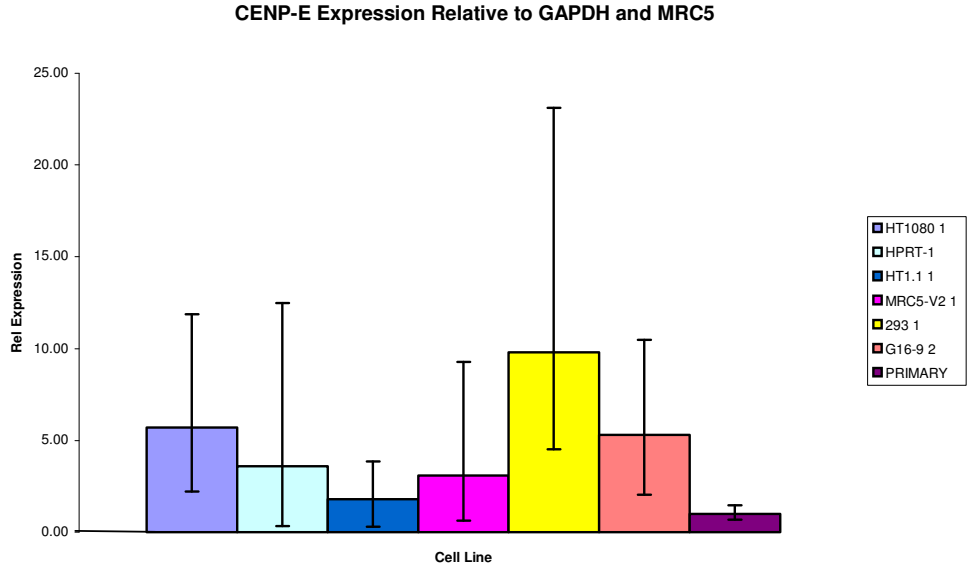


Figure 4.5: CENP-E Expression Relative to GAPDH and MRC5
The bars on each data column give the range of the gene expression as evaluated by $2^{-\Delta\Delta CT}$ with $\Delta\Delta C_T$+s and $\Delta\Delta C_T$-s, where s= the standard deviation of $\Delta\Delta C_T$. The lowest and highest values from each of the three samples are indicated.

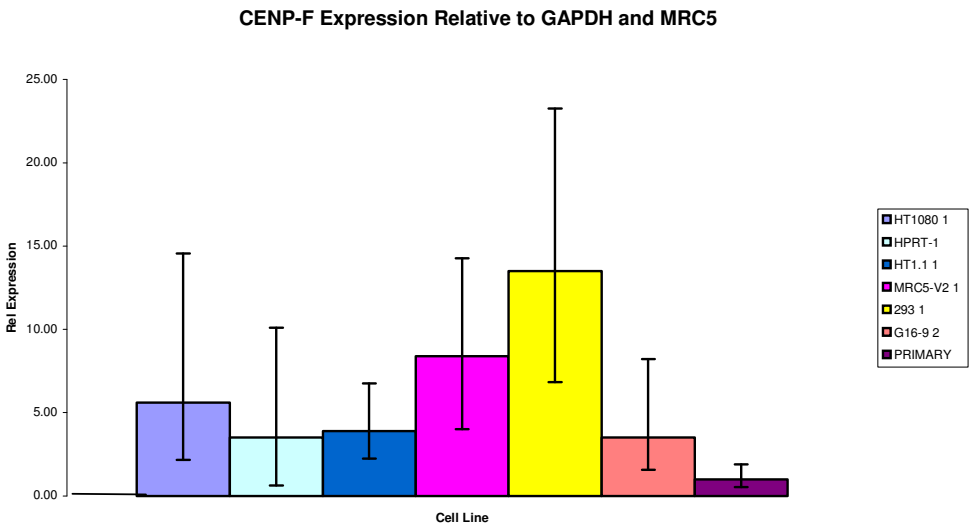


Figure 4.6: CENP-F Expression Relative to GAPDH and MRC5
The bars on each data column give the range of the gene expression as evaluated by $2^{-\Delta\Delta CT}$ with $\Delta\Delta C_T$+s and $\Delta\Delta C_T$-s, where s= the standard deviation of $\Delta\Delta C_T$. The lowest and highest values from each of the three samples are indicated.

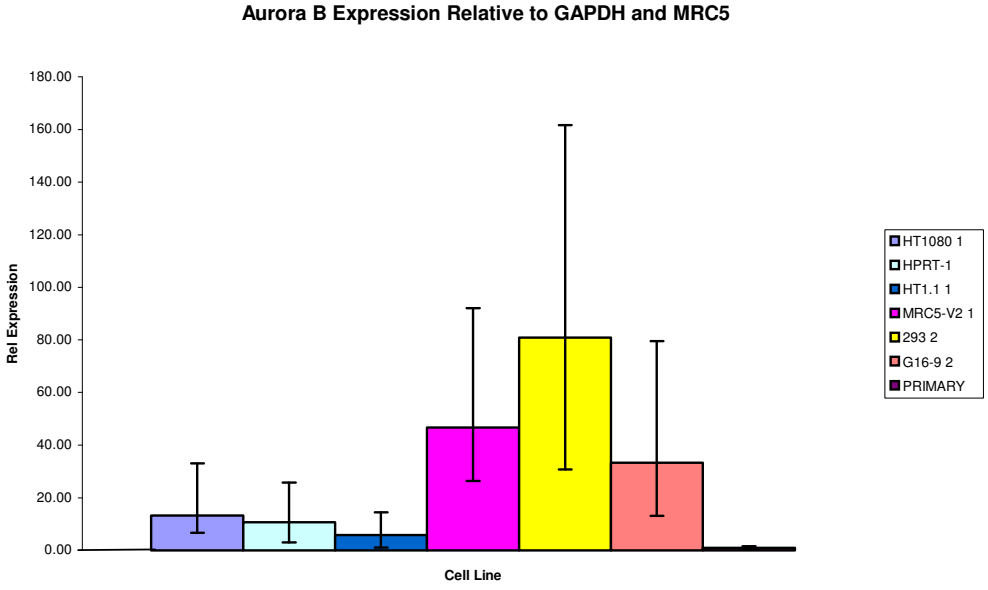


Figure 4.7: Aurora B Expression Relative to GAPDH and MRC5
The bars on each data column give the range of the gene expression as evaluated by $2^{-\Delta\Delta CT}$ with $\Delta\Delta C_T$+s and $\Delta\Delta C_T$-s, where s= the standard deviation of $\Delta\Delta C_T$. The lowest and highest values from each of the three samples are indicated.

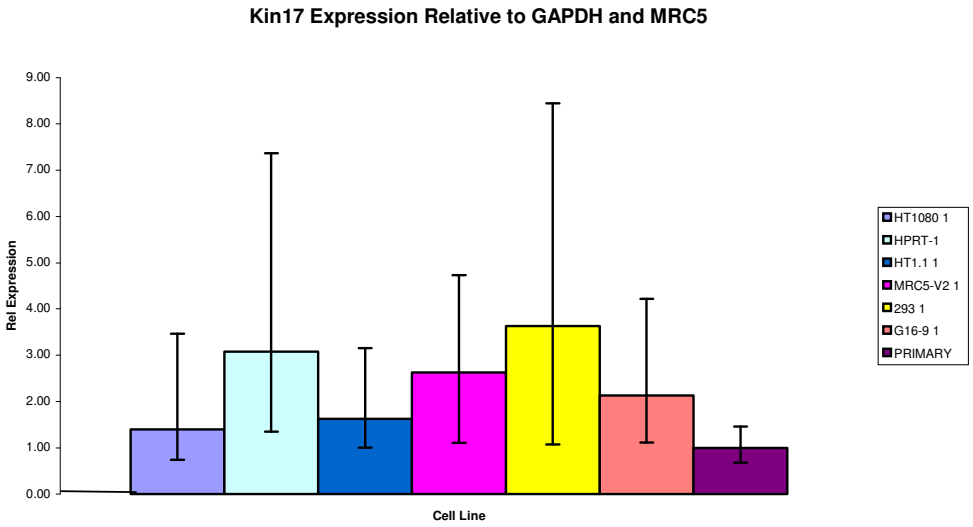


Figure 4.8: Kin17 Expression Relative to GAPDH and MRC5
The bars on each data column give the range of the gene expression as evaluated by $2^{-\Delta\Delta CT}$ with $\Delta\Delta C_T$+s and $\Delta\Delta C_T$-s, where s= the standard deviation of $\Delta\Delta C_T$. The lowest and highest values from each of the three samples are indicated.

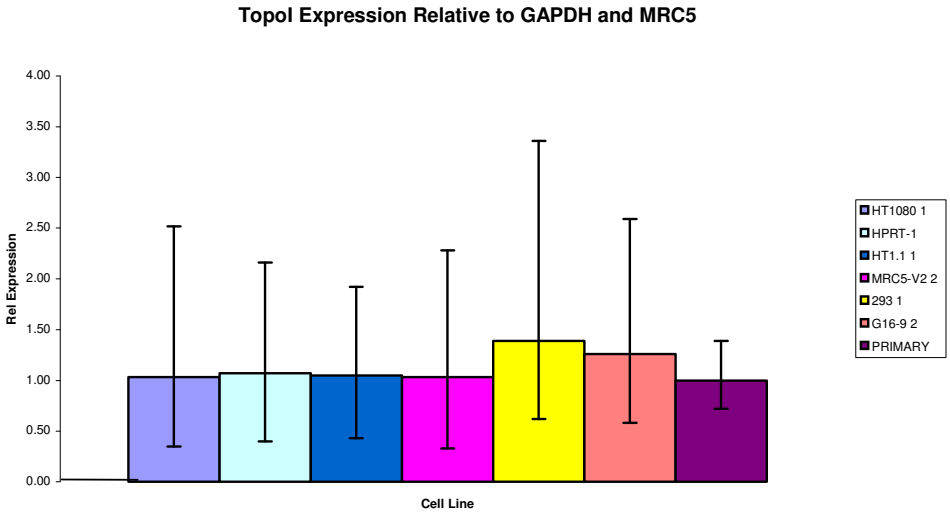


Figure 4.9: Topo I Expression Relative to GAPDH and MRC5
The bars on each data column give the range of the gene expression as evaluated by $2^{-\Delta\Delta CT}$ with $\Delta\Delta C_T$+s and $\Delta\Delta C_T$-s, where s= the standard deviation of $\Delta\Delta C_T$. The lowest and highest values from each of the three samples are indicated.

4.5 Conclusions from Real-Time PCR Analysis

From the above real-time PCR data it was observed that relative to MRC5 (Table 4.3, Figure 4.2) all the cultured cell types have an increased expression of CENP-A. The cultured cell types have an expression of CENP-A ranging between 4.60 (HPRT-HT1080) and 19.63 (MRC5-V2), whilst the MRC5 value is 1. CENP-A associates with all active centromeres (Warburton et al., 1997) and is thought to be at the top of a pathway driving kinetochore assembly and is therefore vital in a stable HAC. Using a t-test, HT1080, HPRT- HT1080, HT1.1 HT1080, MRC5-V2, 293 and G16-9 are all significantly over expressed with respect to MRC5 with statistical probabilities of *p*0.021, *p*0.017, *p*0.037, *p*0.024, *p*0.005 and *p*0.014 respectively. This data shows that relative to the primary type, CENP-A is over expressed in all cultured cells and thus should not hinder HAC stability in these cell types. Other statistically significant results were the over expression of CENP-E in 293 and G16-9 relative to MRC5 (*p*0.05 and *p*0.018 respectively) and the over expression of CENP-F in MRC5-V2 and 293 relative to MRC5 (*p*0.012 and *p*0.021 respectively). The most interesting observation however, is that Aurora B (Figure 4.7 and Table 4.8) is over expressed in both MRC5-V2 (an average expression of 46.5) and 293 (an average expression of 80.9) compared to MRC5 (an expression of 1) with statistical significances of *p*0.021 and *p*0.05 respectively. HACs in both MRC5-V2 and 293 were shown to be unstable in the presence and absence of selection (Table 3.4). In addition, using a paired t-test, a statistical difference in Aurora B levels was also observed when comparing the MRC-V2 cells where HACs are unstable, to HT1080 cells where they are stable (*p*0.033). This is in contrast to the same comparison of the CENP-F protein, which was insignificant.

High variability was observed between samples from the same cell type (Tables 4.3-4.10). Due to the incorporation of the standard deviation in the data analysis where the range of the gene expression is evaluated by $2^{-\Delta\Delta CT}$ with $\Delta\Delta C_T$+s and $\Delta\Delta C_T$-s, where s= the standard deviation of $\Delta\Delta C_T$, experiments looking for low-fold changes in target expression can only tolerate low variation among identical replicates (Bustin, 2000). This is the case in the data analysed above. One further explanation for the disparities between samples could be the presence of different sublines within the population, showing different levels in gene expression, either due to the cell cycle phase they are at, or due to different genetic backgrounds (i.e. tetraploid or aneuploid cells). A possible solution to this would be to isolate homogenous clones or to synchronise the cell populations.

4.6 Chromosomal Instability and Western Blot Analysis

To investigate the observed HAC instability further, the overall chromosomal stability of HT1080, HPRT- HT1080, G16-9, MRC5-V2, 293, Hep3B, HeLa and D98 cells was checked using the Cytokinesis Block Micronucleus (CBMN) assay which allows the analysis of the level of aneuploidy present. Cytochalasin B, which inhibits monomer addition to actin filaments at the fast-growing end of polymers, was added to sub confluent dishes of cells to arrest cell cytokinesis and cytodieresis after DNA replication and mitosis. Cells were harvested and hybridised using FISH to a 17α and 21α probe. Binucleated cells were analysed for the presence of non-disjunction events (one nucleus showing one or more centromeric signal than the modal number, and the other one or more less) and the presence of bridges and micronuclei. Only binucleated cells with round nuclei, of similar size and shape were considered.

Micronuclei were scored when they appeared to be 1/3-1/6 of the nucleus and of similar colour and density. Non-disjunction events were the most frequent aberration (Figure 4.10). Non-disjunction events were low in HT1080 (3.74%), HPRT- HT1080 (5.3%), G16-9 (1.15%) and D98 (4%), but high in MRC5-V2 (24%), 293 (38.7%), Hep3B (15.9%) and HeLa (10.2%) cells (Figure 4.10). A very low percentage of micronuclei and bridges were observed in the HT1080 (0.75%), G16-9 (0.38%) and MRC5-V2 (1.71%) cells with a higher level (13.98%) in 293 cells. No micronuclei or bridges were observed in the HPRT- HT1080, Hep3B, HeLa or D98 cells.

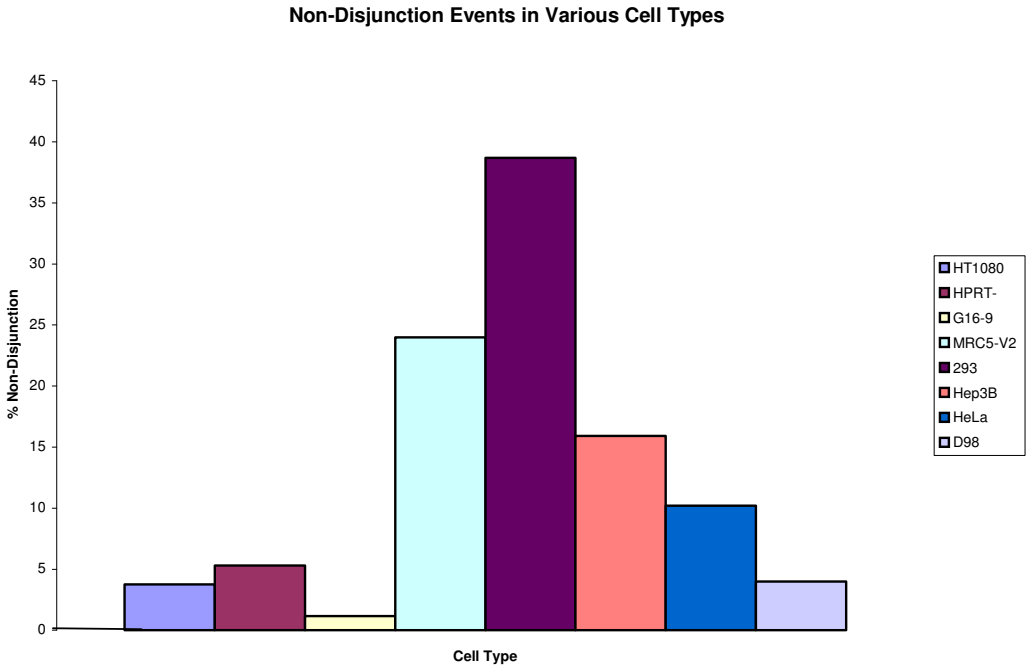


Figure 4.10: Percentage of Non-Disjunction Events as Analysed by the CBMN Assay in Bi-Nucleated Cells.
MRC5-V2 and 293 cells show a higher level of instability at 24% and 38.7% respectively.

Previous work in the laboratory (Moralli et al., 2006) investigated whether the high chromosomal instability in MRC5-V2 and 293 cells was associated with specific protein levels. Proteins involved in kinetochore/centromere function (CENP-A, section 1.7.1), chromatid cohesion and resolution (SMC3, section 1.2, Topo I and Topo II, section 1.7.10), replication (Kin17, section 4.1), pericentromere-specific chromatin modifications (HP1, $H3K9Me_3$, section 1.7.1 and Figure 1.5) and the cell cycle checkpoint (Aurora B, section 1.7.9) were investigated. No significant differences in protein levels were observed between cell types (HT1080, G16-9, MRC5-V2, 293 and MRC5) when western blots of nuclear fractions were reacted with antibodies for CENP-A, SMC3, Kin17, Topo I, HP1 and $H3K9Me_3$ (Moralli et al., 2006). However, Aurora B and Topo II were present in significantly higher amounts in MRC5-V2 and 293 cells. These results all correlate with the real-time PCR findings. Compared with HT1080 cells, the MRC5-V2 cells have 4.5 times more Aurora B and 1.2 times more Topo II and 293 cells have 7.75 times more Aurora B and 2 times more Topo II (Moralli et al., 2006) (Figure 4.11). Aurora B is required for the phosphorylation of H3S10 and H3S28 (Adams et al., 2001b; Goto et al., 2002; Hsu et al., 2000) before mitosis therefore the levels of these modifications were studied on nuclear protein fractions prepared from cells synchronised with colcemid, as described on page 98. The values obtained were adjusted to take into account the difference in mitotic index in each cell type which were as follows: 33.3% in HT1080, 19% in G16-9, 26% in MRC5-V2, 22% in 293 and 52% in MRC5 cells. The over expression of Aurora B corresponded to an increase in phosphorylation of H3S10/28 in MRC5-V2 cells but not in 293 cells (Figure 4.11) suggesting that another of Aurora B's functions other than H3S10/28 phosphorylation was responsible for the high level of chromosomal instability.

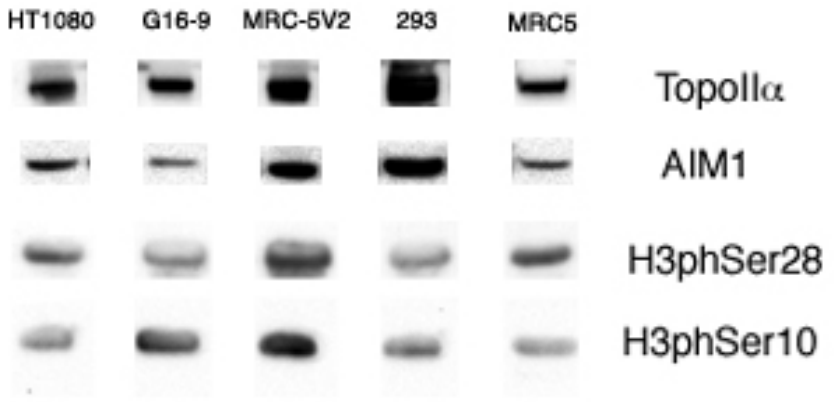


Figure 4.11: Western Blot of Nuclear Protein Fractions From Each Cell Type Reacted with Topo IIα, Aurora B (AIM-1), and Phosphorylated H3S28 and H3S10 Antibodies.

Significantly higher levels of Aurora B and Topo II were observed in MRC5-V2 and 293 cells. The over expression of Aurora B correlates to an increase in H3S10 and H3S28 phosphorylation in MRC5-V2 cells.

Aurora B is involved in chromosome condensation, in the checkpoint controlling the correct alignment of chromosomes on the metaphase spindle, and in the congression of chromosomes at metaphase (Hauf et al., 2003; Vagnarelli and Earnshaw, 2004). If the chromosomes are not correctly aligned on the metaphase spindle due to the over expression of Aurora B as observed in MRC5-V2 and 293 cells, this could account for the increase in chromosome instability resulting from non-disjunction that was evident in the CBMN assay. Increased levels of Aurora B in these cells may interfere with the spindle checkpoint mechanism, thus allowing the cells to proceed to anaphase when the chromosomes are not all correctly bi-oriented on the spindle and

therefore explain the high number of non-disjunction event observed in the two cell lines. In addition, the over expression of Aurora B has been found in numerous types of cancer, where it is usually associated with an increase in H3S10 phosphorylation and chromosomal instability (Ota et al., 2002). Over expression of Topo II is often found in human cancer cells (Jiao et al., 2005). It has been associated with high levels of genome instability, mainly due to centrosome amplification (Kronenwett et al., 2003). Therefore, the presence of excess Topo II possibly adds to overall chromosome instability levels. This overall increase in chromosome instability obviously has direct implications for the stability of any HAC present and offers an explanation as to why HACs are unstable in MRC5-V2 and 293 cells.

4.7 Okadaic Acid Treatment and Analysis

Aurora B is a serine/threonine protein kinase, and it has been demonstrated that raised levels of Aurora B, causing interference with phosphorylation levels, have an effect on chromosomal instability. The next step was to see if increasing phosphorylation levels via another system would have a similar effect.

Okadaic acid is a protein phosphatase (PP) inhibitor and is a potent inhibitor of PP1, PP2A and PP5. Human Aurora B forms complexes with PP1 and PP2A and is activated by okadaic acid as both PP1 and PP2A inactivate Aurora B in a dose-dependent manner (Sugiyama et al., 2002). Okadaic acid increases Aurora B kinase activity in a dose-dependent manner without altering the expression level of Aurora B. This suggests that Aurora B must be phosphorylated to exert its activity.

As high levels of Aurora B were observed in MRC5-V2 and 293 cells, where HACs were unstable, it was investigated whether keeping stable HACs under a constant low dosage of okadaic acid would result in an increase in their instability.

Two HAC containing cell lines were chosen for this study. LJ2-1 is an HT1080 derived clone containing a stable 17α HAC (Mejía et al., 2002). AG6-1 is an HPRT- HT1080 derived clone containing a stable HAC with 17α DNA and the entire HPRT genomic locus (Mejía et al., 2001).

An initial titration experiment was set up to determine the concentration of okadaic acid to use in the study. Okadaic acid was obtained from Sigma and dissolved in DMSO. At 10 nanomolar (nM) all cells in each of the two cell lines died. At 5 nM of OA cells in AG6-1 died while the LJ2-1 growth rate decreased by 10%. Neither of the cell lines showed notable growth disturbances at 1 nM (Table 4.1).

Table 4.1: Okadaic Acid Titration Results

	Concentration of Okadaic Acid		
Cell Line	**10 nM**	**5 nM**	**1 nM**
LJ2-1	All cells died	Growth decreased by 10%	No effect
AG6-1	All cells died	All cells died	No effect

1 nM and 5 nM concentrations were applied to two separate samples of cells and, in parallel, an equal volume of DMSO as the 1 nM concentration (2 μl) was added to a third set of cells to serve as a control. The cells were cultured for 30 days in the presence and absence of selection (125 μg/ml G418 and 0.25 μg/ml puromycin). After 30 days chromosome harvests were prepared. Slides were hybridised with a 17α probe to check for the presence of a HAC. A CBMN assay was also performed as described above. Results are shown in Table 4.2 and Figure 4.12.

Table 4.2: HAC Stability Analysis and CBMN Assay Results for Okadaic Acid Treated Cell Lines LJ2-1 and AG6-1.

Cell Line	HAC %	HAC % 30 Days Off Selection	% Loss Rate[a]	CBMN % Instability 30 Days Off Selection α17	CBMN % Instability 30 Days Off Selection α21	Average % Instability
LJ2-1	94 (Mejía et al., 2002)	88 (Mejía et al., 2002)	0.16			22.2[b]
LJ2-1 DMSO Control	94	70	0.98	36.3	37.93	37.12
LJ2-1 1nM OA	94	75	0.75	38.1	46.15	42.13
LJ2-1 5nM OA	94	50	2.08	57.69	50	53.85
AG6-1	96 (Mejía et al., 2001)	82 (Mejía et al., 2001)	0.49			23.6[b]
AG6-1 Control	96	60	1.55	25	22.73	23.87
AG6-1 1nM OA	96	40	2.88	50	48	49

a Calculated by the formula: $N_n=N_0 \times (1-R)^n$, where N_0 is the number of metaphase chromosome spreads showing HACs in the cells cultured under selection, N_n is the number of HAC-containing metaphase chromosome spreads after n days of culture in the absence of selection and R is the daily rate of loss.
b Calculated by Daniela Moralli from cells under selection using a 17α probe.

OA caused a decrease in stability of the HACs in both cell lines. However, in the LJ2-1 cells this effect was only observed at the higher concentration of 5 nM, with a drop from 70% in the control to 50% at 5 nM, a daily loss percentage of just over 2%. The DMSO appears to contribute slightly towards this result as a decrease in stability was observed in the LJ2-1 and AG6-1 controls compared to the original analysis. In the cytokinesis block assay, OA caused an increase in overall chromosomal instability in the LJ2-1 cells (37.12% in the control compared to 53.85% in the 5 nM OA treated cells) and AG6-1 cells (23.87% in the control compared to 49% in the 1 nM treated cells) (Figure 4.12).

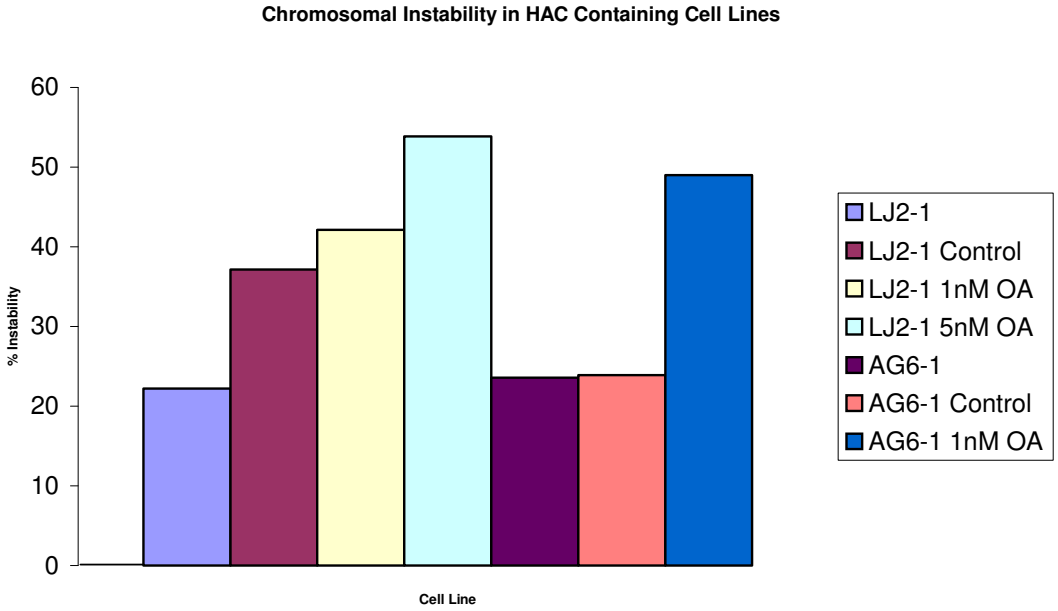
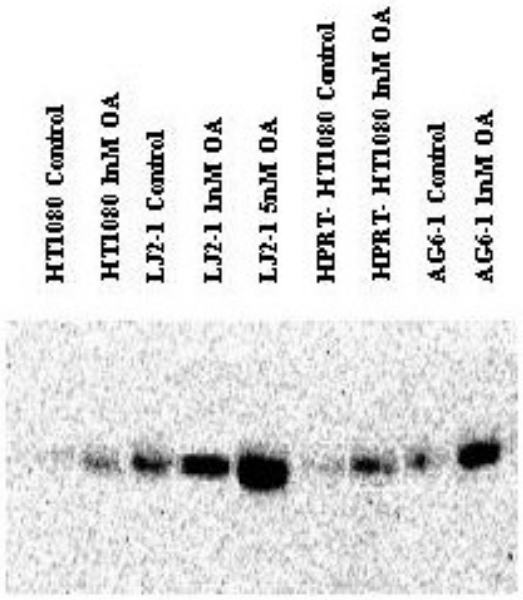


Figure 4.12: Percentage Chromosomal Instability as Analysed by the CBMN Assay in Okadaic Acid Treated, HAC-containing Cell Lines.
The graph shows a general increase in chromosomal instability with OA treatment.

As OA is both a protein phosphatase inhibitor and activates Aurora B, responsible for H3S10 and 28 phosphorylation, (Adams et al., 2001b; Goto et al., 2002; Hsu et al., 2000) the levels of H3 phosphorylation were checked in the treated cells. Nuclear fractions were prepared from the parental cell lines, HT1080 and HPRT- HT1080 as well as the HAC containing cell lines, LJ2-1 and AG6-1 after 30 days in the presence of either DMSO (the control) or OA. The results are shown in Figure 4.13.

A



B

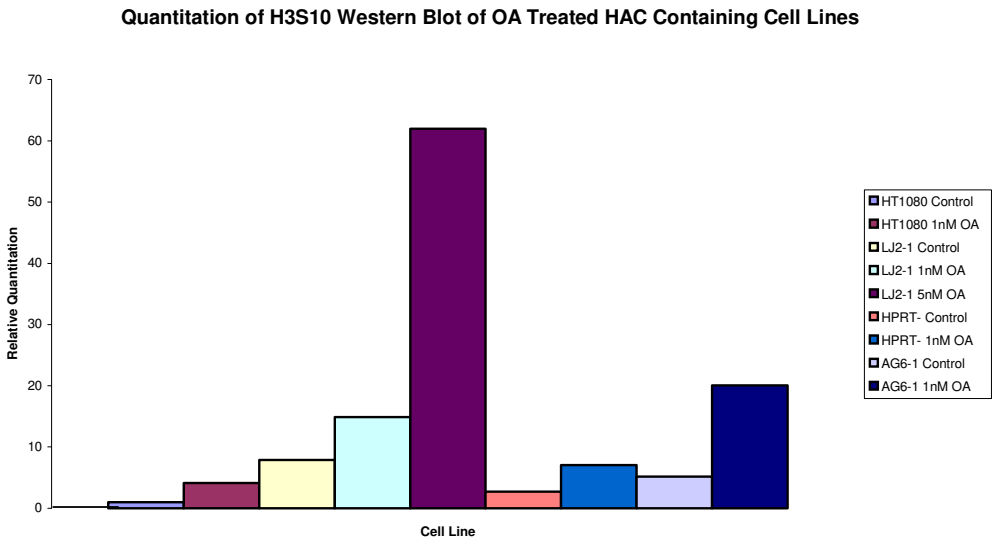


Figure 4.13: Analysis of H3S10 Levels in Okadaic Acid Treated HAC Containing Cell Lines

A Western blot of nuclear protein fractions from each cell line reacted with a phosphorylated H3S10 antibody.

B Quantification of western blot shown in A.

The western blot analysis shows that OA treatment results in an increase in H3S10 phosphorylation. HT1080 cells cultured with 1 nM OA had 4.1 times as much phosphorylation than then HT1080 control. The LJ2-1 cells at 1 nM and 5 nM had 14.5 and 61.9 times as much respectively. A similar trend was observed in the HPRT- parental cell line and in the AG6-1 1 nM OA treated cells.

In addition to the cell lines discussed above, OA was applied to other cell lines used in this study (MRC5-V2, 293 and G16-9) at 1 nM and 5 nM concentrations. Cells were cultured for 30 days and a CBMN assay was performed as well as checking the levels of H3S10 by western blot analysis. The only cell line to survive at 5 nM OA was MRC5-V2, and the 293 cells reached a crisis point and died under both 1 nM and 5 nM concentrations. HT1080, HPRT- HT1080, MRC5-V2 and G16-9 cells all survived at 1 nM OA. An increase in H3S10 phosphorylation was observed by western blot analysis with OA treatment. The results from the cytokinesis block are shown in Figure 4.14.

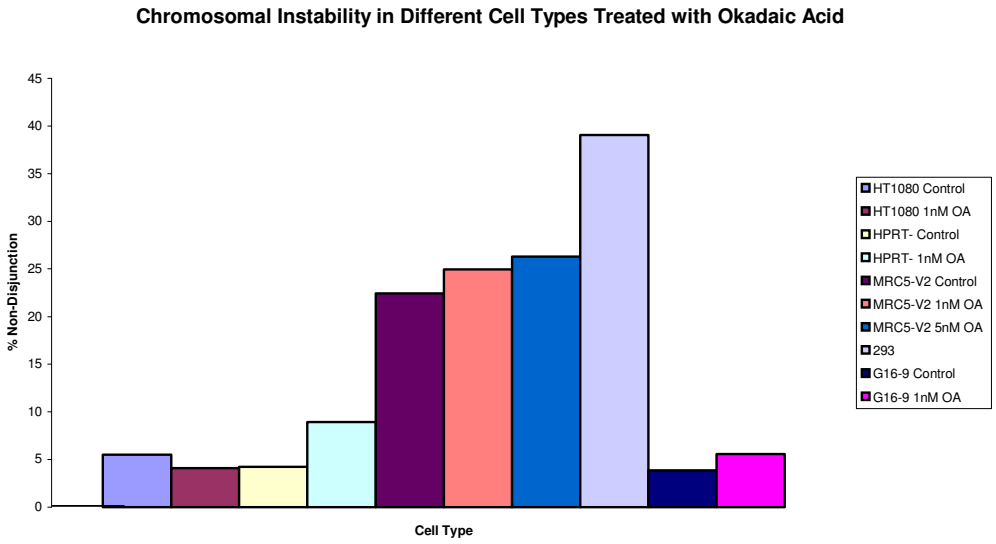


Figure 4.14: Chromosomal Instability as Analysed by the CBMN Assay in Okadaic Acid Treated Cell Types.

With the exception of HT1080, an increase in instability was observed in all of the OA treated cell types (4.23% in the control to 8.93% at 1 nM OA in HPRT- HT1080, 22.45% in the control, 24.95% at 1 nM and 26.3% at 5 nM OA in MRC5-V2 and 3.85% in the control to 5.57% at 1 nM OA in G16-9). The DMSO used in the control experiments also seems to have an a slight affect on the instability, causing an increase from 3.74% in untreated HT1080 cells (Figure 4.10) to 5.5%, 38.7% to 39.1% in 293 and 1.15% to 3.85% in G16-9.

H3 phosphorylation was also checked by immunofluorescence (section 2.16.6) to see if the over-phosphorylation of H3 caused mislocalisation of the histone modification. A phosphorylated H3S28 antibody was used on cells cultured on a slide and fixed in methanol. No differences in localisation were observed between the control cells and

the OA treated cells. H3S28 phosphorylation was only observed in mitotic cells and not interphases as expected (Figure 4.15).

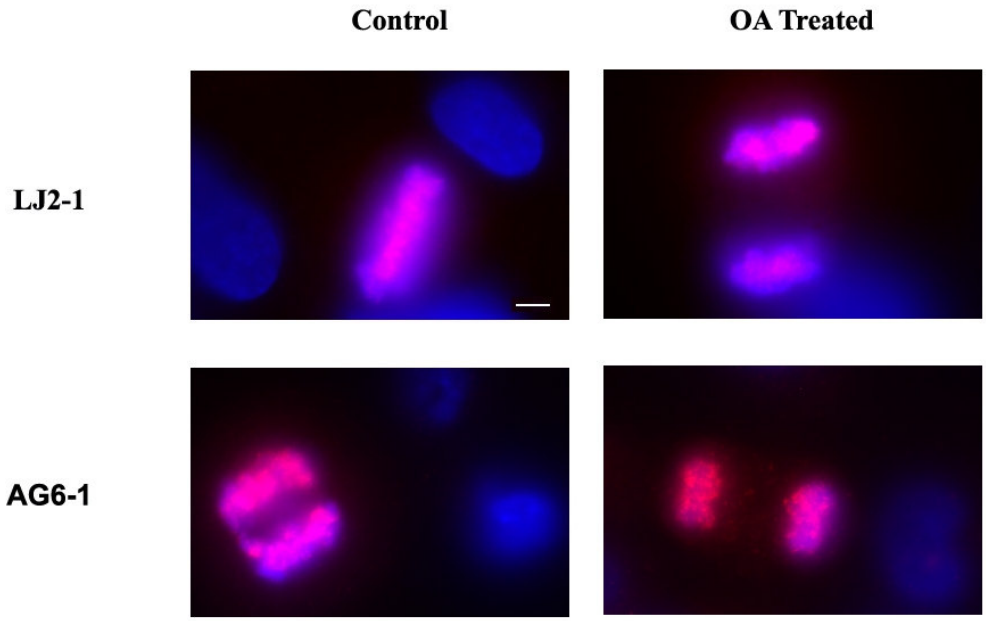


Figure 4.15: Phosphorylated H3S28 Localisation in Okadaic Acid Treated Cell Lines

Phosphorylated H3S28 is shown in red. Cells are countered stained with DAPI. LJ2-1 treated cells are shown at a concentration of 5 nM OA while AG6-1 is shown at a concentration of 1 nM OA. From left to right the images show:

- Phosphorylated H3S28 in metaphase
- Phosphorylated H3S28 in telophase
- Phosphorylated H3S28 in anaphase
- Phosphorylated H3S28 in telophase

In all cases, phosphorylation was only observed during mitosis. The scale bar is equivalent to 2μM.

These results indicate that the H3 phosphorylation in these cells is still occurring at the correct point in the cell cycle in these cells at 1 nM and 5 nM concentrations of OA. H3S10 and H3S28 are phosphorylated during mitosis by Aurora B (Goto et al.,

2002), peaking when chromosomes are in their most condensed state. This phosphorylation is required for proper chromosome condensation and segregation (Wei et al., 1999). After chromosome segregation, Aurora B dissociates from chromosomes (Adams et al., 2001a) and H3 phosphorylation at both sites decreases (Goto et al., 2002). The observation of no mislocalisation indicates that this is still functioning correctly. A low level of H3 phosphorylation is present in interphase cells but it is below the detectability of this experiment. Therefore there may be an increase in H3S10 in interphase that has not been detected. A possibility for the observed death at 5 nM of OA of the AG6-1 cells, and at all concentrations for the 293 cells, is that this mechanism may no longer be functioning.

4.7 Investigation of a Cause for the Increase in H3 Phosphorylation

A possible cause of the increase in instability and increase in H3 phosphorylation was investigated. Mitotic H3 phosphorylation is a balance of kinase and phosphatase activities (Hsu et al., 2000). Both PP1 and PP2A inactivate Aurora B in a dose-dependent manner (Sugiyama et al., 2002).

PP2A stabilises securin (which in turn inhibits separase, required for sister chromatid segregation (Nasmyth, 2001)), by counteracting its phosphorylation, which targets it for destruction by the SCF ubiquitin ligase complex (Gil-Bernabe et al., 2006). This leads to the release of cohesin which PP2A therefore prevents.

Several protein kinases are substrates for PP2A and PP2A appears to be the major kinase phosphatase in eukaryotic cells that down regulates activated protein kinases

(Millward et al., 1999). PP2A could be a possible candidate for the removal of phosphorylation of H3S10 and H3S28 after mitosis.

Via both of these pathways, OA, an inhibitor of PP2A could prevent both the activation of separase, resulting in poor chromosome segregation, and H3S10 and S28 dephosphorylation, leading to condensation and segregation errors. Therefore, the levels of PP2A were investigated in the OA treated cells.

Nuclear fractions were obtained after 30 days in culture for the OA treated cell lines, and checked by western blot for the levels of PP2A (Figure 4.16).

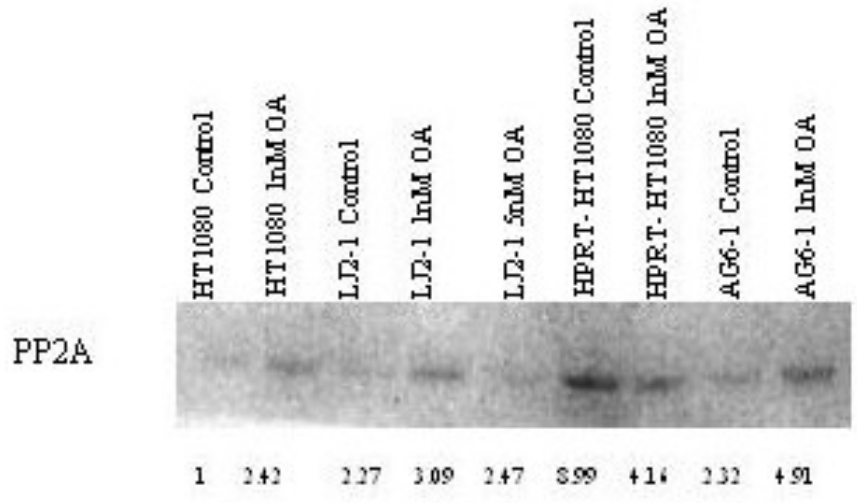


Figure 4.16: PP2A Levels in Okadaic Acid Treated Cell Line

From Figure 4.16 there is no clear difference in PP2A with okadaic acid treatment. This suggests that there may be another candidate responsible for the observed instability in these cell lines when treated with OA or, rather than a change in the total level of protein present, the effect is due to a difference in PP2A activity.

4.8 Conclusions and Discussion

HACs generated in HT1080 and G16-9 cells following transduction with HSV-1 HAC amplicons containing pHSV17α302Neo, pHSV17α 227Neo and pHSV21αNeo were stable for 90 days, but in MRC5-V2 and 293 cells the HACs were unstable in the presence and absence of selection. These results suggested that the requirements for centromere function and HAC stability might be different in each cell type. Real-time PCR allowed the accurate quantitation of gene expression of several possible candidates. The most interesting observation was that Aurora B is over expressed in both MRC5-V2 and 293 cells compared to MRC5 with statistical significances of $p0.021$ and $p0.05$ respectively. In addition, a statistical difference was also observed when comparing the MRC-V2 cells where HACs are unstable, to HT1080 cells where they are stable ($p0.033$). Few differences in expression were observed between the cell lines in the analysis of the other candidates studied, possibly due to different sub populations of cells within the samples. One possible explanation is differences in the cell cycle stages of the cell populations within the preparations. Synchronising the cell populations or isolating clones may remedy this problem.

High levels of chromosomal instability in MRC5-V2 and 293 cells correlated with high levels of Aurora B and Topo II. Numerous types of cancer show an increase in

the level of Aurora B, associated with an increase in H3S10 phosphorylation and chromosomal instability (Ota et al., 2002). However, an increase in H3S10 phosphorylation in all the cell lines analysed was not observed. Most notably 293, which had the highest level of chromosomal instability, had H3Ser10/28 phosphorylation similar to that of MRC5. Thus, the effect of Aurora B is likely to be due to another mechanism. Aurora B is involved in controlling the correct alignment of chromosomes on the metaphase spindle (Hauf et al., 2003; Vagnarelli and Earnshaw, 2004). The observed over expression in MRC5-V2 and 293 cells suggests that increased levels of Aurora B in these cells interferes with the spindle checkpoint mechanism, thus allowing the cells to proceed to anaphase when the chromosomes are not all correctly bi-oriented on the spindle. This could explain the high number of non-disjunction event observed in the two cell lines. The presence of excess Topo II may add to overall chromosome instability levels as its over expression is associated with centrosome amplification, leading to high levels of genome instability (Kronenwett et al., 2003).

OA is both a protein phosphatase inhibitor and activates Aurora B (Adams et al., 2001b; Goto et al., 2002; Hsu et al., 2000). OA caused a decrease in stability of HACs in both the LJ2-1 and AG6-1 cell lines. In the cytokinesis block assay, OA caused an increase in overall chromosomal instability in the LJ2-1 and AG6-1 cells. The DMSO used in the control experiments also seemed to have a slight affect on the instability for unknown reasons. Western blot analysis confirmed that OA treatment results in an increase in H3S10 phosphorylation. Thus, the OA effect observed does not exactly mimic what is seen in 293 cells, as could be expected, since OA acts on a large number of phosphatases and possibly kinases. 5 nM treatment of the AG6-1

cells caused cell death and 293 cells were not able to grow at all, even at very low OA concentrations. This could indicate that the cellular phosphorylation levels of these cell lines are already close to damagingly high levels. A mechanism must be in place to counteract the OA treatment that is successful at a lower dosage of OA, at least for AG6-1, but at a higher dosage a crisis point is reached and possibly induces apoptosis. It must also be taken into consideration that when applying OA, a cell population resistant to the pressure is selected and these cells may not display typical properties.

H3S28 phosphorylation was only observed in mitotic cells and not interphases via immunofluorescence (Figure 4.15), indicating that histone phosphorylation is still functioning correctly in these cells at 1 nM and 5 nM concentrations of OA. The observation of no mislocalisation indicates that H3 is still being phosphoryated by Aurora B (Goto et al., 2002) (required for proper chromosome condensation and segregation (Wei et al., 1999)), and then Aurora B dissociates from the chromosomes with a resulting decrease in H3 phosphorylation (Adams et al., 2001a). A possibility for the observed death at 5 nM of OA of the AG6-1 cells and the death of other cell lines including the 293 cells, is that this mechanism may no longer be functioning. However, mislocalisation of phosphorylated H3 to interphases may be below the detection of these experiments.

PP2A was investigated as a possible candidate for the cause of the increase in instability and increase in H3 phosphorylation. Mitotic H3 phosphorylation is a balance of kinase and phosphatase activities (Hsu et al., 2000). PP2A stabilises securin (which in turn inhibits separase, required for sister chromatid segregation (Nasmyth, 2001)), by counteracting its phosphorylation, which targets it for

destruction by the SCF complex (Gil-Bernabe et al., 2006). This leads to the release of cohesin that PP2A therefore prevents. PP2A also inactivates Aurora B in a dose-dependent manner (Sugiyama et al., 2002) and could be a possible candidate for the removal of phosphorylation of H3S10 and H3S28 after mitosis. Via both of these pathways, OA, an inhibitor of PP2A could prevent both the activation of separase and H3S10 and S28 dephosphorylation, leading to condensation and segregation errors.

The western blot analysis indicates that this is not the case and from Figure 4.16 there is no clear difference in PP2A with okadaic acid treatment. This indicates that there may be another candidate responsible for the observed instability in these cell lines when treated with OA. However, the function of PP2A was not assayed and while there may be no observed difference in expression, there may be functional differences present.

Appendix to Chapter 4: Results II

Table 4.3: Relative Quantitation of CENP-A (to MRC5) Using the Comparative C_T Method

Cell Line	CENP-A Average C_T	GAPDH Average C_T	ΔC_T CENP-A-GAPDH[a]	$\Delta\Delta C_T$ ΔC_T - $\Delta C_{T,MRC5}$[b]	CENP-A_N Rel. to MRC5[c]
HT1080	26.76±0.57	21.50±0.48	8.26±0.75	-2.93±0.75	7.61 (4.53-12.80)
	30.79±0.44	22.58±0.31	8.20±0.54	-2.98±0.54	7.90 (5.43-11.48)
	29.53±0.36	20.71±0.18	8.82±0.40	-2.36±0.40	5.15 (3.90-6.79)
HPRT-HT1080	33.06±0.37	24.60±0.33	8.46±0.50	-2.72±0.50	6.60 (4.67-9.33)
	35.05±0.29	26.07±0.74	8.98±0.79	-2.20±0.79	4.60 (2.66-7.96)
	32.01±0.65	23.21±0.86	8.80±1.08	-2.39±1.08	5.24 (2.48-11.08)
HT1.1 HT1080	29.74±0.27	21.25±0.15	8.49±0.31	-2.69±0.31	6.48 (5.22-8.03)
	29.47±0.29	20.68±0.30	8.80±0.42	-2.39±0.42	5.24 (3.92-7.02)
	29.38±1.07	21.41±0.45	7.98±1.16	-3.21±1.16	9.26 (4.14-20.69)
MRC5-V2	29.28±0.44	22.26±0.20	7.02±0.48	-4.17±0.48	17.94 (12.86-25.02)
	29.20±0.28	22.30±0.13	6.89±0.31	-4.29±0.31	19.63 (15.83-24.34)
	29.18±0.54	21.54±0.41	7.65±0.68	-3.54±0.68	11.61 (7.25-18.61)
293	30.51±0.55	23.11±0.57	7.41±0.79	-3.78±0.79	13.71 (7.93-23.71)
	30.62±0.79	23.44±0.69	7.18±1.05	-4.01±1.05	16.08 (7.77-33.30)
	28.85±0.35	21.76±0.44	7.09±0.56	-4.10±0.56	17.11 (11.60-25.22)
G16-9	29.63±0.27	21.30±0.47	8.33±0.54	-2.86±0.54	7.25 (4.98-10.54)
	31.78±0.31	23.94±0.44	7.83±0.54	-3.35±0.54	10.22 (7.03-14.86)
	29.78±0.46	21.57±0.36	8.21±0.58	-2.98±0.58	7.86 (5.26-11.76)
MRC5 PRIMARY	32.49±0.37	21.30±0.45	11.19±0.58	0±0.58	1 (0.67-1.49)

a. The ΔC_T value is determined by subtracting the average GAPDH C_T value from the average CENP-A C_T value. The standard deviation of the difference is calculated from the standard deviations of the CENP-A and GAPDH values using the formula $s=\sqrt{s_1^2+s_2^2}$ where s= standard deviation, s_1= standard deviation of CENP-A and s_2= standard deviation of GAPDH.
b. $\Delta\Delta C_T$ is formed by subtracting the calibrator value, in this case $\Delta C_{T,MRC5}$, from ΔC_T. As this is subtraction of an arbitrary value, the standard deviation remains unchanged from the standard deviation of ΔC_T.
c. The range given for CENP-A_N relative to MRC5 is determined by the expression $2^{-\Delta\Delta CT}$ with $\Delta\Delta C_T$+s and $\Delta\Delta C_T$-s, where s= the standard deviation of $\Delta\Delta C_T$.

Table 4.4: Relative Quantitation of CENP-B (to MRC5) Using the Comparative C_T Method

Cell Line	CENP-B Average C_T	GAPDH Average C_T	ΔC_T CENP-B-GAPDH[a]	$\Delta\Delta C_T$ ΔC_T - $\Delta C_{T,MRC5}$[b]	$CENP\text{-}B_N$ Rel. to MRC5[c]
HT1080	31.35±0.49	21.50±0.48	9.85±0.69	0.05±0.69	0.96 (0.60-1.56)
	32.68±0.35	22.58±0.31	10.09±0.47	0.30±0.47	0.81 (0.59-1.13)
	30.61±0.25	20.71±0.18	9.90±0.31	0.10±0.31	0.93 (0.75-1.15)
HPRT-HT1080	37.38±0.36	26.07±0.74	11.31±0.82	1.51±0.82	0.35 (0.20-0.62)
	33.92±1.29	23.21±0.86	10.71±1.55	0.91±1.55	0.53 (0.18-1.55)
	34.06±0.68	21.30±0.45	12.76±0.82	2.97±0.82	0.13 (0.07-0.23)
HT1.1 HT1080	30.39±0.19	21.25±0.15	9.14±0.24	-0.66±0.24	1.58 (1.33-1.86)
	30.66±0.27	20.68±0.30	9.99±0.40	0.19±0.40	0.88 (0.66-1.16)
	30.60±0.45	21.41±0.45	9.19±0.64	-0.60±0.64	1.52 (0.98-2.37)
MRC5-V2	32.00±0.40	22.26±0.20	9.74±0.45	-0.06±0.45	1.04 (0.76-1.42)
	32.46±0.82	22.30±0.13	10.16±0.83	0.36±0.83	0.78 (0.44-1.38)
	31.70±0.81	21.54±0.41	10.16±0.91	0.36±0.68	0.78 (0.41-1.46)
293	33.63±0.27	23.11±0.57	10.53±0.63	0.73±0.63	0.60 (0.39-0.93)
	34.12±1.20	23.44±0.69	10.68±1.38	0.88±1.38	0.54 (0.21-1.41)
	32.16±0.35	21.76±0.44	10.40±0.71	0.60±0.71	0.66 (0.40-1.08)
G16-9	32.31±0.21	22.89±0.38	9.42±0.43	-0.38±0.43	1.30 (0.96-1.75)
	33.62±0.45	23.94±0.44	9.67±0.63	-0.12±0.63	1.09 (0.70-1.68)
	31.56±0.87	21.57±0.36	9.99±0.94	0.20±0.94	0.87 (0.45-1.67)
MRC5	31.10±0.50	21.30±0.45	9.80±0.67	0±0.67	1 (0.63-1.59)

Table 4.5: Relative Quantitation of CENP-C (to MRC5) Using the Comparative C_T Method

Cell Line	CENP-C Average C_T	GAPDH Average C_T	ΔC_T CENP-C-GAPDH[a]	$\Delta\Delta C_T$ ΔC_T - $\Delta C_{T,MRC5}$[b]	CENP-C_N Rel. to MRC5[c]
HT1080	30.25±0.45	21.50±0.48	8.75±0.66	-1.02±0.66	2.02 (1.28-3.20)
	31.42±0.19	22.58±0.31	8.84±0.36	-0.93±0.36	1.90 (1.48-2.44)
	30.18±0.36	20.71±0.18	9.46±0.23	-0.30±0.23	1.23 (1.05-1.44)
HPRT-HT1080	33.31±0.22	24.60±0.33	8.72±0.40	-1.05±0.40	2.06 (1.56-2.72)
	35.44±0.40	26.07±0.74	9.37±0.84	-0.40±0.84	1.32 (0.74-2.36)
	31.80±0.65	21.30±0.45	10.51±0.56	0.74±0.56	0.60 (0.41-0.88)
HT1.1 HT1080	30.86±0.17	21.25±0.15	9.61±0.23	-0.15±0.23	1.11 (0.95-1.30)
	30.35±0.37	20.68±0.30	9.68±0.48	-0.09±0.48	1.06 (0.76-1.48)
	30.56±0.34	21.41±0.45	9.15±0.56	-0.62±0.56	1.53 (1.04-2.26)
MRC5-V2	30.70±0.16	22.26±0.20	8.44±0.26	-1.33±0.26	2.51 (2.10-3.01)
	31.16±0.41	22.30±0.13	8.85±0.43	-0.91±0.43	1.88 (1.39-2.53)
	30.37±0.41	21.54±0.41	8.83±0.58	-0.93±0.58	1.91 (1.28-2.85)
293	30.76±0.19	23.11±0.57	7.66±0.60	-2.11±0.60	4.31 (2.84-6.54)
	31.30±0.52	23.44±0.69	7.86±0.86	-1.90±0.86	3.74 (2.06-6.79)
	29.66±0.45	21.76±0.44	7.89±0.63	-1.87±0.63	3.66 (2.36-5.66)
G16-9	30.00±0.21	21.30±0.47	8.70±0.51	-1.06±0.51	2.09 (1.47-2.98)
	32.35±0.40	23.94±0.44	8.41±0.59	-1.06±0.59	2.09 (1.39-3.15)
	30.16±0.37	21.57±0.36	8.59±0.52	-1.18±0.52	2.26 (1.58-3.24)
MRC5	31.06±0.22	21.30±0.45	9.76±0.50	0±0.50	1 (0.71-1.41)

Table 4.6: Relative Quantitation of CENP-E (to MRC5) Using the Comparative C_T Method

Cell Line	CENP-E Average C_T	GAPDH Average C_T	ΔC_T CENP-E-GAPDH[a]	$\Delta\Delta C_T$ ΔC_T - $\Delta C_{T,MRC5}$[b]	CENP-E_N Rel. to MRC5[c]
HT1080	31.23±0.46	21.50±0.48	9.73±0.67	-2.90±0.67	7.48 (4.70-11.90)
	32.62±0.33	22.58±0.31	10.04±0.45	-2.60±0.45	6.05 (4.43-8.26)
	31.47±0.68	20.71±0.18	10.76±0.70	-1.87±0.70	3.66 (2.25-5.94)
HPRT-HT1080	34.04±0.32	24.60±0.33	9.45±0.46	-3.19±0.46	9.11 (6.62-12.53)
	39.13±0.60	26.07±0.74	13.06±0.95	0.43±0.95	0.74 (0.38-1.44)
	33.85±0.94	21.30±0.45	12.55±1.04	-0.08±1.04	1.06 (0.51-2.17)
HT1.1 HT1080	32.18±0.19	21.25±0.15	10.49±0.24	-1.70±0.24	3.24 (2.74-3.83)
	36.90±0.54	23.13±0.36	13.77±0.65	1.13±0.65	0.46 (0.29-0.71)
	32.60±0.58	20.68±0.30	11.93±0.65	-0.71±0.65	1.63 (1.04-2.56)
MRC5-V2	32.02±0.29	22.26±0.20	9.76±0.35	-2.87±0.35	7.32 (5.74-9.33)
	35.75±0.28	23.35±0.50	12.40±0.57	-0.24±0.57	1.18 (0.79-1.75)
	35.02±0.43	22.30±0.13	12.72±0.45	0.09±0.45	0.94 (0.69-1.29)
293	31.98±0.52	23.11±0.57	8.88±0.77	-3.76±0.77	13.53 (7.94-23.08)
	32.87±0.72	23.44±0.69	9.43±1.00	-3.20±1.00	9.20 (4.60-18.40)
	31.68±0.33	21.76±0.44	9.92±0.55	-2.71±0.55	6.56 (4.48-9.60)
G16-9	31.26±0.55	21.30±0.47	9.97±0.72	-2.67±0.72	6.63 (3.86-10.47)
	34.38±1.09	23.94±0.44	10.43±1.18	-2.20±1.18	4.59 (2.03-10.41)
	31.98±0.25	21.57±0.36	10.42±0.44	-2.22±0.44	4.65 (3.43-6.31)
MRC5	33.93±0.34	21.30±0.45	12.63±0.56	0±0.56	1 (0.68-1.47)

Table 4.7: Relative Quantitation of CENP-F (to MRC5) Using the Comparative C_T Method

Cell Line	CENP-F Average C_T	GAPDH Average C_T	ΔC_T CENP-F-GAPDH[a]	$\Delta\Delta C_T$ ΔC_T - $\Delta C_{T,MRC5}$[b]	CENP-F_N Rel. to MRC5[c]
HT1080	26.52±0.70	21.50±0.48	5.02±0.85	-3.00±0.85	8.02 (4.45-14.45)
	28.15±0.11	22.58±0.31	5.56±0.33	-2.46±0.33	5.50 (4.38-6.92)
	27.51±0.51	20.71±0.18	6.44±0.54	-1.59±0.54	3.01 (2.07-4.37)
HPRT-HT1080	29.75±0.34	24.60±0.33	5.16±0.47	-2.87±0.47	7.29 (5.27-10.10)
	33.68±0.78	26.07±0.74	7.61±1.08	-0.42±1.08	1.34 (0.63-2.83)
	28.45±0.59	21.30±0.45	7.15±0.74	-0.87±0.74	1.83 (1.10-3.06)
HT1.1 HT1080	27.19±0.24	21.25±0.15	5.94±0.28	-2.09±0.28	4.24 (3.49-5.15)
	27.08±0.37	20.68±0.30	6.40±0.48	-1.63±0.48	3.08 (2.21-4.30)
	27.32±0.45	21.41±0.45	5.92±0.64	-2.11±0.64	4.32 (2.77-6.73)
MRC5-V2	26.97±0.17	22.26±0.20	4.71±0.26	-3.32±0.26	9.98 (8.33-11.94)
	27.41±0.19	22.30±0.13	5.10±0.23	-2.92±0.23	7.57 (6.46-8.88)
	26.65±0.82	21.54±0.41	5.11±0.92	-2.91±0.92	7.52 (3.97-14.22)
293	27.18±0.15	23.11±0.57	4.08±0.59	-3.95±0.59	15.42 (10.24-23.21)
	29.38±0.35	25.27±0.35	4.11±0.49	-3.92±0.49	15.10 (10.75-21.21)
	26.49±0.30	21.76±0.44	4.73±0.53	-3.30±0.53	9.82 (6.80-14.17)
G16-9	27.71±0.36	21.30±0.47	6.41±0.59	-1.61±0.59	3.06 (2.03-4.61)
	29.60±0.50	23.94±0.44	5.65±0.67	-2.37±0.67	5.18 (3.26-8.25)
	28.36±0.42	21.57±0.36	6.79±0.55	-1.23±0.55	2.35 (1.60-3.44)
MRC5	29.33±0.77	21.30±0.45	8.03±0.92	0±0.92	1 (0.53-1.89)

Table 4.8: Relative Quantitation of Aurora B (to MRC5) Using the Comparative C_T Method

Cell Line	Aurora B Average C_T	GAPDH Average C_T	ΔC_T Aurora B - GAPDH[a]	$\Delta\Delta C_T$ ΔC_T - $\Delta C_{T,MRC5}$[b]	Aurora B_N Rel. to MRC5[c]
HT1080	26.96±0.25	21.50±0.48	5.45±0.54	-3.61±0.54	12.21 (8.40-17.75)
	27.78±1.13	22.58±0.31	5.20±1.17	-3.87±1.17	14.57 (6.48-32.79)
	26.15±0.48	20.71±0.18	5.44±0.51	-3.62±0.51	12.33 (8.66-17.56)
HPRT-HT1080	29.99±0.33	24.60±0.33	5.40±0.47	-3.67±0.47	12.68 (9.16-17.57)
	31.29±0.42	26.07±0.74	5.22±0.85	-3.85±0.85	14.39 (7.98-25.93)
	27.85±0.71	21.30±0.45	6.55±0.84	-2.51±0.84	5.71 (3.19-10.21)
HT1.1 HT1080	26.90±0.43	21.25±0.15	5.65±0.46	-3.42±0.46	10.67 (7.75-14.67)
	31.15±0.58	23.13±0.36	8.01±0.68	-1.05±0.68	2.07 (1.29-3.31)
	27.40±0.29	20.68±0.30	6.72±0.42	-2.34±0.42	5.07 (3.79-6.79)
MRC5-V2	26.27±0.22	22.26±0.20	4.00±0.30	-5.06±0.30	33.32 (27.06-41.02)
	25.57±0.71	22.30±0.13	3.26±0.72	-5.80±0.72	55.78 (33.86-91.88)
	24.95±0.85	21.54±0.41	3.41±0.94	-5.65±0.94	50.27 (26.20-96.45)
293	27.47±0.32	25.27±0.35	2.20±0.47	-6.87±0.47	116.70 (84.25-161.64)
	26.85±0.16	23.44±0.69	3.41±0.71	-5.66±0.71	50.42 (30.82-82.47)
	24.59±0.58	21.76±0.44	2.83±0.73	-6.24±0.73	75.50 (45.52-125.22)
G16-9	25.46±0.42	21.30±0.47	4.17±0.63	-4.90±0.63	29.79 (19.25-46.10)
	27.42±0.58	23.94±0.44	3.48±0.73	-5.59±0.73	48.03 (28.96-79.66)
	26.15±0.66	21.57±0.36	4.58±0.75	-4.48±0.75	22.34 (13.28-37.57)
MRC5	30.36±0.49	21.30±0.45	9.06±0.67	0±0.67	1 (0.63-1.59)

Table 4.9: Relative Quantitation of Kin17 (to MRC5) Using the Comparative C_T Method

Cell Line	Kin17 Average C_T	GAPDH Average C_T	ΔC_T Kin17-GAPDH[a]	$\Delta\Delta C_T$ ΔC_T - $\Delta C_{T,MRC5}$[b]	$Kin17_N$ Rel. to MRC5[c]
HT1080	28.42±0.13	21.50±0.48	6.92±0.50	-0.48±0.50	1.40 (0.99-1.98)
	30.96±0.57	24.42±0.72	6.54±0.92	-0.87±0.92	1.83 (0.97-3.46)
	28.19±0.34	20.71±0.18	7.48±0.38	0.07±0.38	0.95 (0.73-1.24)
HPRT-HT1080	30.37±0.52	24.60±0.33	5.77±0.60	-1.63±0.60	3.11 (2.05-4.71)
	31.90±0.51	26.07±0.74	5.83±0.90	-1.58±0.90	2.99 (1.60-5.58)
	28.96±0.86	23.21±0.86	5.74±1.22	-1.66±1.22	3.17 (1.36-7.38)
HT1.1 HT1080	28.27±0.34	21.25±0.15	7.02±0.37	-0.39±0.37	1.31 (1.01-1.69)
	29.59±0.61	23.13±0.36	6.46±0.71	-0.95±0.71	1.93 (1.18-3.16)
	28.09±0.51	21.41±0.45	6.68±0.68	-0.72±0.68	1.65 (1.03-2.65)
MRC5-V2	27.99±0.38	22.26±0.20	5.73±0.43	-1.68±0.43	3.20 (2.38-4.32)
	28.21±0.73	22.30±0.13	5.91±0.74	-1.50±0.74	2.83 (1.69-4.72)
	28.08±0.59	21.54±0.41	6.54±0.72	-0.86±0.72	1.82 (1.10-2.99)
293	28.71±0.10	23.11±0.57	5.61±0.58	-1.80±0.58	3.48 (2.33-5.20)
	28.57±0.40	23.44±0.69	5.13±0.80	-2.28±0.80	4.85 (2.78-8.44)
	27.81±0.43	21.76±0.44	6.05±0.62	-1.36±0.62	2.56 (1.67-3.93)
G16-9	28.76±0.39	22.89±0.38	5.87±0.54	-1.53±0.54	2.89 (1.99-4.21)
	27.74±0.34	21.30±0.47	6.45±0.58	-0.96±0.58	1.94 (1.30-2.90)
	28.36±0.28	21.57±0.36	6.79±0.46	-0.61±0.46	1.53 (1.11-2.10)
MRC5	28.71±0.31	21.30±0.45	7.41±0.55	0±0.55	1 (0.68-1.46)

Table 4.10: Relative Quantitation of Topo I (to MRC5) Using the Comparative C_T Method

Cell Line	Topo I Average C_T	GAPDH Average C_T	ΔC_T TopoI - GAPDH[a]	$\Delta\Delta C_T$ ΔC_T - $\Delta C_{T,MRC5}$[b]	$TopoI_N$ Rel. to MRC5[c]
HT1080	27.52±1.09	21.50±0.48	6.02±1.42	0.09±1.42	0.94 (0.35-2.52)
	28.14±0.51	22.58±0.31	5.56±0.60	-0.37±0.60	1.30 (0.85-1.96)
	26.90±0.38	20.71±0.18	6.19±0.42	0.26±0.42	0.84 (0.62-1.12)
HPRT-HT1080	29.87±0.28	24.60±0.33	5.27±0.43	-0.66±0.43	1.58 (1.18-2.13)
	32.01±0.80	26.07±0.74	5.94±1.09	0.00±1.09	1.00 (0.47-2.13)
	28.10±0.35	21.30±0.45	6.80±0.57	0.87±0.57	0.55 (0.37-0.81)
HT1.1 HT1080	26.55±0.26	21.25±0.15	5.30±0.30	-0.64±0.30	1.55 (1.26-1.91)
	27.35±0.44	20.68±0.30	6.67±0.53	0.73±0.53	0.60 (0.42-0.87)
	27.41±0.58	21.41±0.45	6.00±0.73	0.06±0.73	0.96 (0.58-1.59)
MRC5-V2	30.32±0.23	23.35±0.50	6.97±0.55	1.03±0.55	0.49 (0.33-0.72)
	27.90±0.42	22.30±0.13	5.60±0.44	-0.34±0.44	1.26 (0.93-1.72)
	27.03±0.63	21.54±0.41	5.50±0.75	-0.44±0.75	1.35 (0.80-2.28)
293	27.96±0.36	23.11±0.57	4.86±0.67	-1.08±0.67	2.11 (1.33-3.36)
	29.18±0.55	23.44±0.69	5.74±0.88	-0.20±0.88	1.14 (0.62-2.11)
	27.82±0.23	21.76±0.44	6.06±0.50	0.13±0.50	0.92 (0.65-1.30)
G16-9	26.94±0.43	21.30±0.47	5.64±0.64	-0.29±0.64	1.23 (0.79-1.91)
	29.06±0.34	23.94±0.44	5.12±0.56	-0.82±0.56	1.76 (1.20-2.60)
	27.78±0.32	21.57±0.36	6.21±0.58	0.28±0.48	0.83 (0.59-1.15)
MRC5	27.23±0.14	21.30±0.45	5.93±0.47	0±0.47	1 (0.72-1.39)

Chapter 5:

Results III: Development of a Gene Expression System Using HSV-1 Amplicons

5.1 Introduction

HACs have become an important gene vector transfer system for non-viral gene delivery into human cells. HAC vectors have been effective in both gene expression and complementation studies in human HT1080 cells. In our lab, a HAC vector containing chromosome 17 α-DNA, telomeres and a large genomic segment from human Xq26.2 spanning the HPRT locus, complemented the HPRT deficiency in human cells thus demonstrating the efficacy of *de novo* HACs for delivery and expression of human transgenes in human cell lines (Mejía and Larin, 2000; Mejía et al., 2001). Another study reported expression of the HPRT gene from an artificial chromosome in HPRT deficient HT1080 cells using a co-transfection assay (Grimes et al., 2001). Two PACs, one containing 70 kb of 21 α-DNA and a second PAC containing a 140 kb genomic insert spanning the HPRT gene were transfected into HT1080 cells, resulting in the formation of HPRT expressing HACs. In a similar co-transfection study two BACs, one containing 21 α-DNA and one containing the GCH1 genomic locus generated HACs in HT1080 cells with multiple copies of GCH1 (Ikeno et al., 2002). These HACs were sensitive to IFN-γ induction, replicating the response of the gene expression from the endogenous chromosomal gene.

The HSV-1 system has been used to deliver genomic BACs containing genes for complementing deficiencies. HPRT-deficient MRC5-V2 cells and a mouse primary

hepatocyte culture derived from $Hprt^{-/-}$ mice were both complemented following transduction with a BAC containing the HPRT locus (Wade-Martins et al., 2001). Expression of the human LDLR, at physiologically appropriate levels, was generated from a BAC delivered by HSV-1 amplicons in both CHO cells deficient in the LDLR receptor and in human fibroblasts derived from familial hypercholesterolemia patients who have a mutation in the LDLR receptor (Wade-Martins et al., 2003).

Highly efficient homologous recombination systems have been developed in *E. coli* which allow the modification of large BACs without the use of restriction enzymes and ligation (Copeland et al., 2001; Muyrers et al., 2001). Red mediated recombination genes (*gam* and *red*) from bacteriophage lambda have been introduced into the *E. coli* genome (Lee et al., 2001). Expression of these genes is under the control of a temperature sensitive lambda repressor, inactive at 42ºC. This system has been used to recombine two BACs (Kotzamanis and Huxley, 2004). Each BAC is modified with a linking vector containing a region of homology for recombination. One BAC is linearised and introduced into the *E. coli* strain EL350, which contains both the recombination genes and the second BAC. Recombination then occurs between the homologous regions on each BAC (Kotzamanis and Huxley, 2004). This method has been used to generate a BAC containing the entire CFTR gene (Kotzamanis and Huxley, 2004) and to construct a HAC vector containing 70 kb of alphoid DNA on a 156 kb BAC carrying the human HPRT gene (Kotzamanis et al., 2005).

The work in this chapter has focussed on creating a vector to obtain gene expression from a genomic locus carrying HAC using the HSV-1 system due to the ability of this

system to form HACs in a variety of cells lines (as detailed in Chapter 3), a vital feature of a successful gene therapy approach. Conventional subcloning based on restriction digest and ligation is relatively inefficient for incorporating large genomic regions into vectors due to the difficulty in manipulating such large DNA fragments. Thus, more reliable systems such as Red mediated recombination have been developed. The first section of this work investigates the possibility of engineering the HPRT genomic locus into the 65 kb pHSV17α302Neo vector by Red mediated recombination for complementing the HPRT deficiency in HPRT- HT1080 cells using the HSV-1 system.

The second section of this work investigates whether a strategy using co-infection of two HSV-HAC amplicons, one containing alpha satellite DNA and the other the genomic locus of interest, is possible for generating HAC carrying genes in cells. Pilot experiments using BAC vector DNA were successful and suggest that this is a feasible approach.

5.2 PAC Subcloning by Red/ET Recombination to Engineer the HPRT Genomic Locus into the pHSV17α302Neo Vector

The final HAC gene vector construct for HSV-1 delivery is limited to 150 kb due to the packaging capacity of the HSV-1 virions. The HPRT genomic locus is present on a large, 175 kb, PAC construct, hPAC71G4 (Schindelhauer and Cooke, 1997). Therefore, to develop a smaller construct suitable for recombination with the 17α BAC pHSV17α302Neo, 65 kb, I aimed to modify the PAC construct by Red/ET recombination using Gene Bridge's BAC Subcloning Kit (Figure 5.1).

The experimental steps of this technique are as follows: oligonucleotides are designed containing, at the 5' end, stretches of homology to the regions flanking the genomic fragment which is to be subcloned. At their 3' ends, these oligos also contain primer sequences for amplification of a BAC or PAC minimal vector. Using these oligos a linear vector backbone with flanking homology arms is constructed in a PCR reaction. The cells containing the genomic locus are transformed with the expression plasmid pRed/ET. The expression of genes mediating recombination is induced by the addition of L-arabinose. The linear PCR product is electroporated into these cells. Recombination will take place and the clones carrying the subcloned fragment are identified by selection for the appropriate antibiotic resistance. The recombined BAC would then be recombined with pHSV17α302Neo using the protocol described by (Kotzamanis and Huxley, 2004) for linking two overlapping BACs into a single clone. The desired region of one BAC is cloned into a specifically designed BAC linking vector via homologous recombination. This is then linearised via a site in the BAC-linking vector and recombined with the second BAC. The regions of homology used are the overlapping regions of the BAC backbone and the chloramphenicol resistance gene in the linking vector (Kotzamanis and Huxley, 2004).

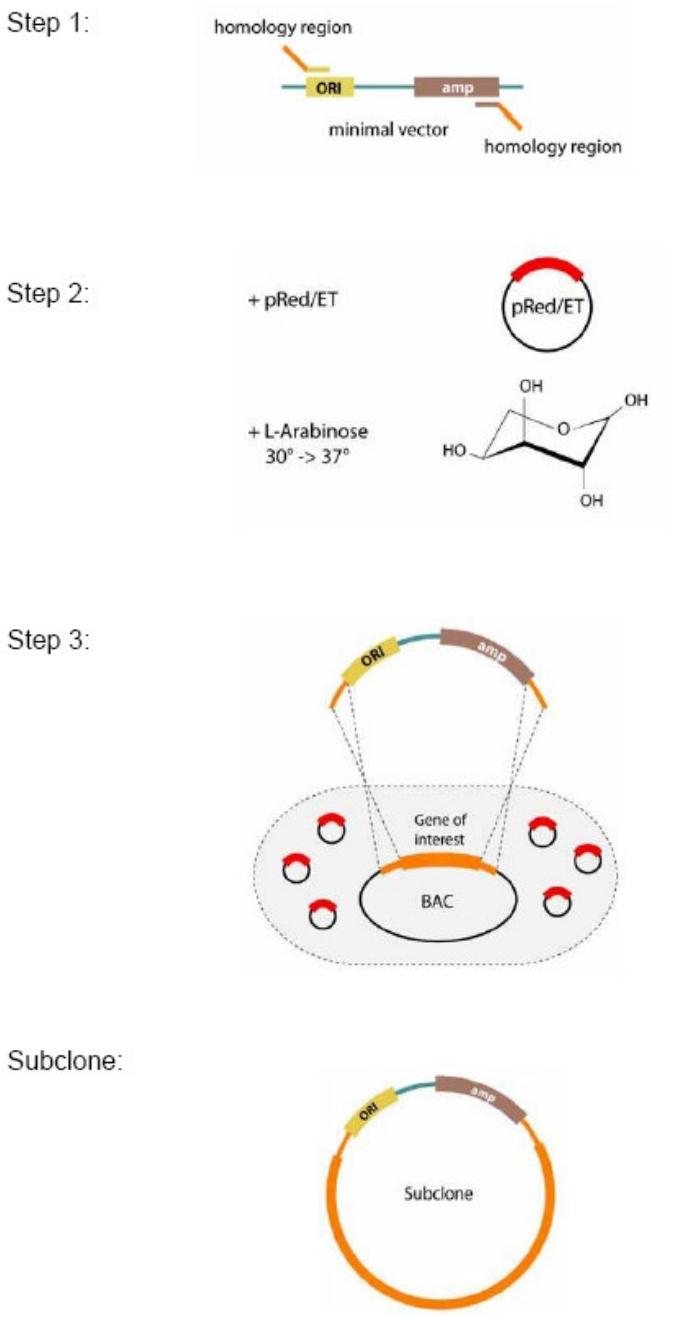


Figure 5.1: General Subcloning Strategy

Step 1: Oligos are designed containing stretches of homology to the fragment of the BAC/PAC which is to be subcloned. At their 3' ends, these oligos also contain primer sequences for amplification of the minimal vector. Using these oligos a linear minimal vector with flanking homology arms is constructed in a PCR reaction.

Step 2: The *E. coli* strain carrying the BAC/PAC, which is to be modified, is transformed with the expression plasmid pRed/ET. The expression of genes mediating Red/ET is induced by the addition of L-arabinose.

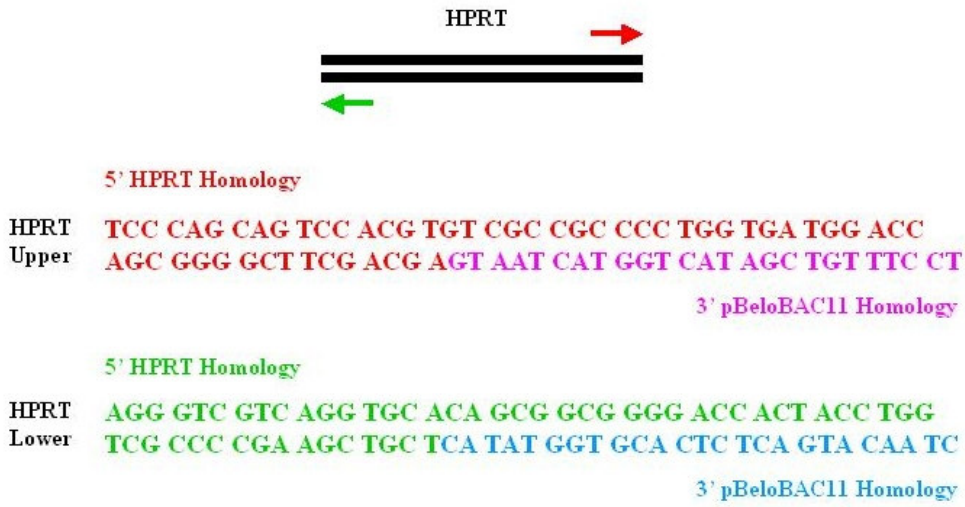
Step 3: The linear vector (PCR product with the added homology arms) is electroporated into the cells. Recombination will take place and the clones carrying the subcloned fragment are identified by selection for the appropriate antibiotic resistance. Only colonies with a circularised (recombined) vector will survive.

(From Gene Bridges BAC Subcloning Protocol Manual)

Since the HPRT genomic locus is carried by a PAC, (hPAC71G4) it was decided to move the 40 kb region containing the HPRT gene to pBeloBAC11. The first step was to design PCR primers which serve to amplify the pBeloBAC11 vector. These contain homology to the regions flanking the HPRT genomic locus at the 5' end, which were chosen to ensure that only the region containing the HPRT gene and its regulatory sequences was moved from the PAC to the BAC, by positioning them 50 bp immediately upstream of the unwanted section (Figure 5.2A). At their 3' ends, the primers contain sequences for amplification of pBeloBAC11 (Figure 5.2A). Using these oligos allows a linear pBeloBAC11 with flanking homology arms to HPRT to be constructed (Figure 5.2B). Primers were HPLC purified (Figure 5.2A).

pBeloBAC11 was linearised with BamHI and the fragment purified using the QIAquick Gel Extraction kit (Qiagen) according to the manufacturer's instructions.

A



B

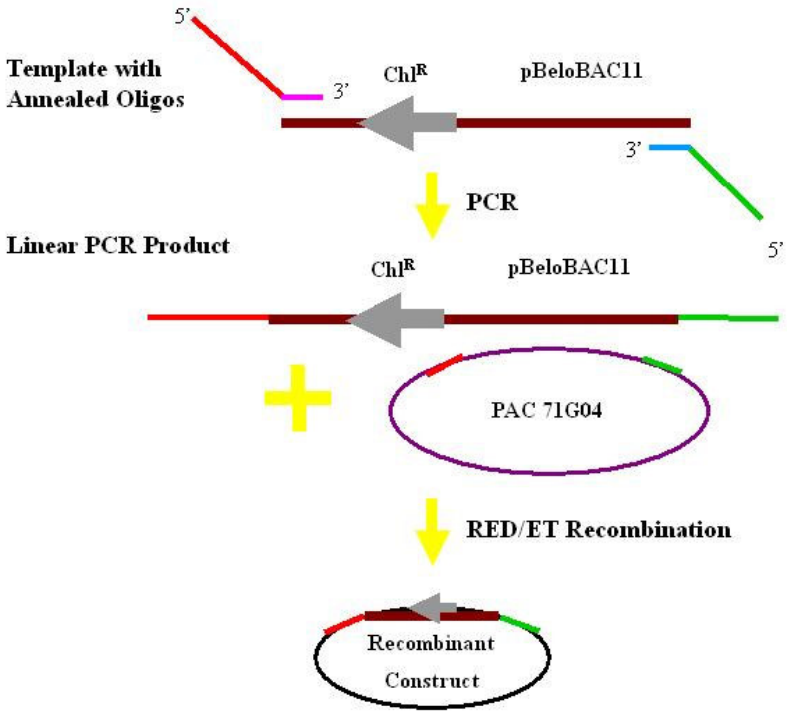


Figure 5.2: Production of the Recombinant BAC Containing the HPRT Genomic Locus

A PCR Primers

B Practical Steps Involved in the Recombination Process

PCR reactions were carried out in a total volume of 25 µl with 1x enzyme buffer, 1.75 mM $MgCl_2$, 500 µM dNTPs, 15 ng of pBeloBAC11 BamHI linearised template, 2.5 pmoles of each primer and 1 U of BIO-X-ACT Long DNA polymerase (Bioline).

PCR conditions were generally as follows:

1. 95˚C for 15 minutes
2. 95˚C for 30 seconds
3. 55˚C for 1 minute
4. 72˚C for 12 minutes
5. Go to 2 and repeat for 34 cycles
6. 72˚C for 9 minutes
7. 10˚C for ever

However, these conditions yielded no results. Doubling the primer or template concentration, using different dNTPs, preparing fresh DNA template and using circular template, utilising different PCR machines and increasing the extension time all had no effect. Eventually, DNA extraction using phenol-chloroform purification of the template led to a visible 7 kb product.

A temperature gradient from 50-70˚C was run to determine the best annealing temperature (Figure 5.3). For ease of preparation, 6x 25 µl reactions were set up and then each halved to give 12 reactions. 4 µl of each was run on a 1% agarose gel. 55˚C was determined as the best annealing temperature due to the large amount of product and small amount of by-product observed.

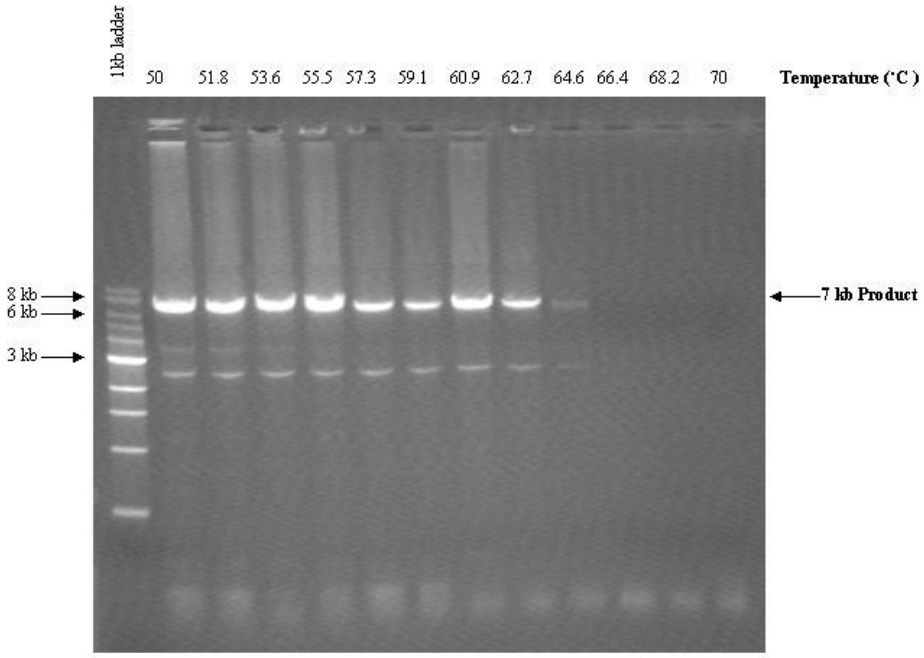


Figure 5.3: Temperature Gradient PCR from 50-70˚C to Determine the Annealing Temperature for the HPRT Primers.
55˚C was the most appropriate temperature with a large product band and little by-product.

The necessity of preparing large quantities of pBeloBAC11 template was time consuming and laborious, particularly with the PCR only working sporadically. New primers were ordered and aliquoted to avoid repeated freezing and thawing. Different commercial enzymes were tested including PfuTurbo Hotstart DNA Polymerase (Stratagene), 1.25 U in a 25 µl reaction; DyNAzyme EXT PCR Kit (NEB), following the manufacturer's instructions, with and without the use of 5% DMSO; and the Expand Long Template PCR System (Roche), where each of the three provided buffers were tested and the elongation time was increased for 20 seconds/cycle. None

of these approaches gave consistent results. In particular, much time was lost as the NEB control reaction failed to yield positive results and a new kit had to be obtained.

Eventually 6 successful reactions were obtained. 3 each were pooled and purified using the QIAquick PCR purification kit (Qiagen). Samples were eluted twice with water to give a total volume of 85 µl. This was digested with DpnI in 1x buffer and 1x BSA to remove any remaining template. The reaction product was purified again and eluted in 20 µl water before being precipitated.

hPAC71G4 electrocompetent cells were made as detailed in section 2.10.1. 10 ng of pSC101-BAD-gbaA-tetra/pRed/ET (BAC Subcloning Kit, Gene Bridges) was transformed into these competent cells by electroporation at 2.0 kV, 200 Ω and 25 µF.

pSC101-BAD-gbaA-tetra is the Red/ET recombination protein expression plasmid. The plasmid carries the redα, β, and γ genes of the λ phage along with the recA gene in a polycistronic operon under the control of an inducible promoter. The time period for recombination is therefore limited by the induced expression, limiting the risk of unwanted rearrangement. The plasmid is a derivative of a temperature sensitive pSC101 replicon which is a low copy number plasmid depending on the oriR101. pSC101 encodes the RepA protein which is necessary for plasmid DNA replication (Miller et al., 1995). Cells have to be cultured at 30˚C as at 37°C the Red/ET plasmid will be lost. The redαβγ operon is expressed under the control of the arabinose-inducible BAD promoter (Guzman et al., 1995). pSC101-BAD-gbaA-tetra confers tetracycline resistance. The pBAD promoter is regulated by AraC (Schleif, 1992) which forms a complex with L-arabinose. Arabinose binds to AraC and allows

transcription to begin. In the absence of arabinose an AraC dimmer blocks transcription.

After electroporation cells were cultured at 30˚C for 75 minutes in 1 ml SOC and plated on agar plates containing kanamycin (33.3 μg/ml), to select for hPAC71G4, and tetracycline (3 μg/ml), to select for pRed/ET. Eight colonies were picked and grown in the presence of kanamycin at 37˚C overnight. Cultures were miniprepped as per section 2.9.3 and the resulting DNA was digested with NotI to ensure that the hPAC71G4 was still intact. Digestion with NotI releases a 16 kb fragment (Figure 5.4).

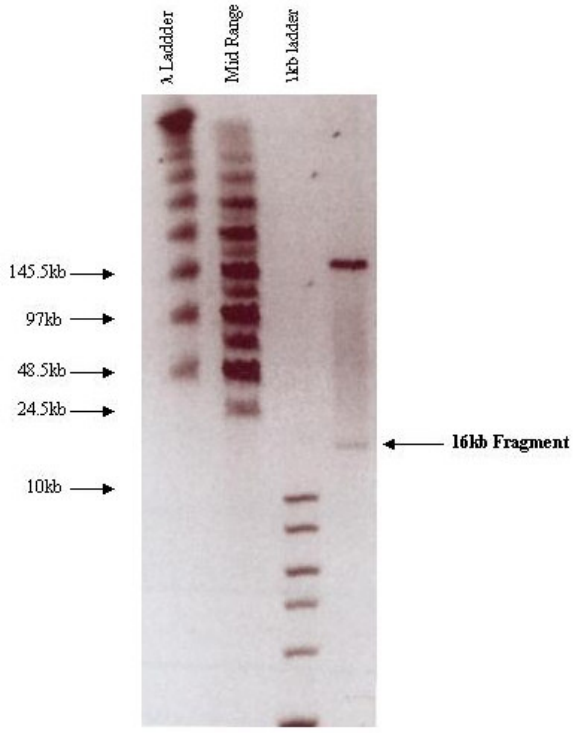


Figure 5.4: NotI Digestion of hPAC71G4 after Transformation of Electrocompetent Cells with pRed/ET.
Digestion with NotI releases a 16 kb fragment.

The same colonies were grown in the presence of kanamycin and tetracycline at 30˚C overnight. Cultures were miniprepped using the QIAprep Miniprep Kit according to the manufacturer's instructions and the resulting DNA was digested with EcoRI to ensure that the pRed/ET plasmid was still intact. Digestion with EcoRI releases 6.6, 1.5 and 1.2 kb fragments as detailed in the Gene Bridges Protocol. Correct colonies were grown in 1.5 ml LB with kanamycin (33.3 μg/ml) and tetracycline (3 μg/ml) overnight at 30˚C before being transferred to a 100 ml culture. Cells were grown to an OD of 0.15 and then induced with 1.5 ml of 10% L-arabinose. Electrocompetent cells were prepared as normal.

10 μl of the purified PCR product was transformed into these cells by electroporation at 2.0 kV, 200 Ω and 25 μF. After electroporation cells were cultured at 37˚C for 75 minutes in 1 ml SOC and plated on chloramphenicol agar plates. This selects for the chloramphenicol resistance conferred by pBeloBAC11 which should have recombined with the HPRT genomic locus of the PAC leading to a circular molecule. No colonies were obtained and due to the PCR difficulties discussed this recombination approach was no longer pursued.

5.3 Development of a Dual Infection System Using HSV-1 Amplicons

The second section of this work looks at using a co-infection approach with the HSV-1 system. Simultaneously delivering an HSV-1 vector carrying alpha satellite and a second HSV-1 vector containing a genomic gene with its regulatory sequences would allow the development of a genomic locus carrying HAC to study gene expression. Expression has been achieved previously via co-transfection experiments where

HPRT and GCH1 expressing HACs have been obtained (Grimes et al., 2001; Ikeno et al., 2002). Initial experiments look promising and suggest that using the HSV-1 system for dual infections is a feasible approach to generating gene complementation in cells.

To monitor dual infection events in mammalian cells, it's necessary to use two HSV-1 amplicon constructs carrying different selectable markers, and different fluorescent reporter genes. We decided to use pHGKneo, described in Chapter 3, which carries the GFP and the Neo^R gene. The second vector had to be engineered starting from plasmid pHR (a kind gift from Richard Wade-Martins) (Figure 5.5) which is essentially pSG80A-HG, as described in Table 2.3, but with the GFP gene replaced with RFP gene.

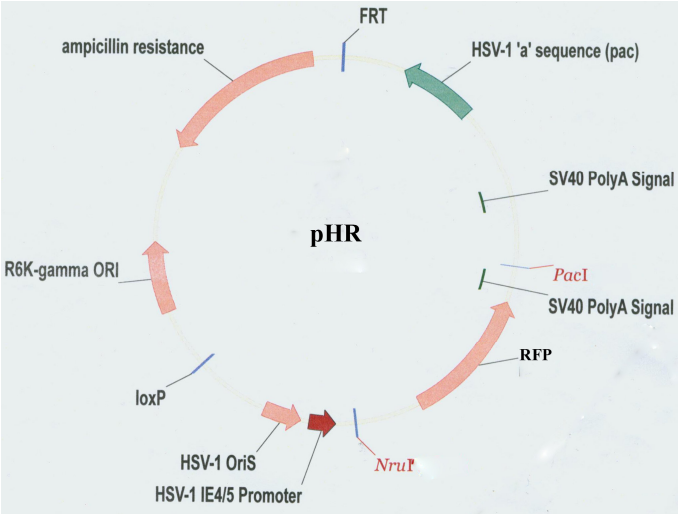
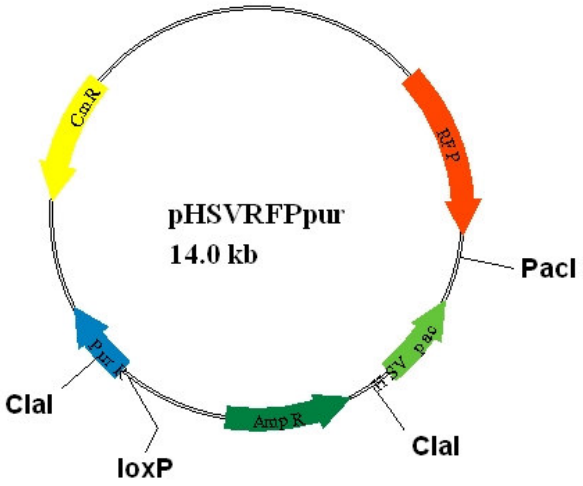


Figure 5.5: pHR Plasmid Map

The paucity of suitable single site restriction enzymes on pHR prevented the insertion of a selectable marker cassette by standard cloning. Previous work in the lab generated a pBeloBAC11 vector with the puromycin resistance cassette inserted into the unique ApaLI site at position 87 (Alazami, unpublished), pBelopur7. This vector was chosen for development for two reasons. Firstly it has a *loxP* site and so can easily be recombined with pHR via Cre- *loxP* - mediated recombination. Secondly, the puromycin cassette has been utilised before in the lab so the appropriate level of selective pressure was known.

pBelopur7 electrocompetent cells were made and transformed with 100 ng of pHR and 50 ng of pCTP-T as described in sections 2.10.4. By Cre-*loxP* - mediated recombination this generated pHSVRFPpur (Figure 5.6A). The construct was checked by PacI and ClaI digestion (Figure 5.6B). There is no PacI site in pBeloBAC11 and one unique PacI site in pHR, thus a linear fragment approximately 14 kb is expected. There is one ClaI site in pHR and one in the puromycin cassette of pBeloBAC11, thus two fragments are obtained (approximately 12 kb and 2.2 kb).

A



B

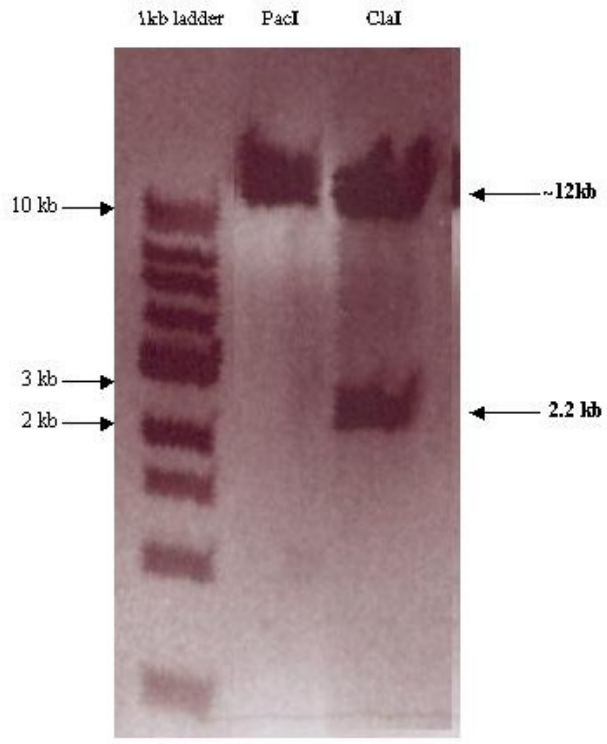


Figure 5.6: pHSVRFPpur Analysis
A Vector map indicating the restriction sites used in analysis (the fragment sizes and enzyme sites are approximate).
B Restriction digest gel showing that PacI linearises the vector and ClaI releases a 2.2kb fragment.

HSV-1 amplicons were generated containing pHGKNeo (GFP and neomycin resistance, Figure 3.4) and pHSVRFPpur (RFP and puromycin resistance, Figure 5.6) in Vero 2-2 cells and purified and concentrated as described in sections 1.12 and 2.13 (Wade-Martins et al., 2001). Figure 5.7 illustrates the GFP and RFP expression following transfection of pHGKNeo and pHSVRFPpur into Vero 2-2 cells after 60 hours incubation before the amplicons are harvested.

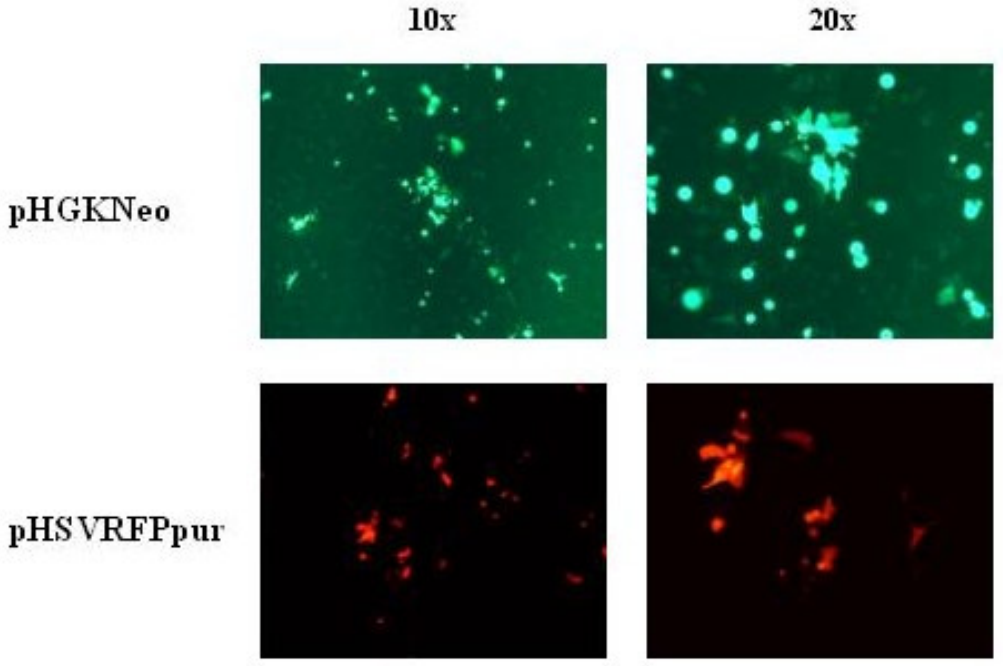


Figure 5.7: GFP Expression of pHGKNeo and RFP Expression of pHSVRFPpur in Vero 2-2 Cells

This figure illustrates the GFP and RFP expression following transduction of pHGKNeo and pHSVRFPpur respectively at both 10x and 20x objective in Vero 2-2 cells after 60 hours incubation before the amplicons are harvested.

The titre was estimated following the transduction of $2.5x10^5$ G16-9 cells. The number of transducing units was determined after 24 hours by counting the number of GFP/RFP positive cells. A titre of $1x10^6$ units/ml was obtained.

HT1080, MRC5-V2, 293 and G16-9 cells were infected with:

a) pHGKNeo

b) pHSVRFPpur

c) pHGKNeo and pHSVRFPpur

at MOI 1. Figures 5.8 and 5.9 show the GFP and RFP expression as well as the merged images to indicate dually transformed cells in MRC5-V2 and G16-9 cells respectively. All the cell lines used were susceptible to infection by each amplicon but the efficiency of transduction was very poor, probably due to the low amplicon titre. Efficiencies ranged in the single infections from 0.6% with pHGKNeo in 293 cells and pHSVRFPpur in HT1080 cells, to 33% with pHGKNeo in G16-9 cells. The pHGKNeo levels are very low compared with the values obtained previously (Table 3.1). For example 7% in HT1080 compared with 0.8% observed here. As was to be expected, the transduction efficiencies of the dual infections were lower, ranging from 0.4% (pHGKNeo in HT1080) to 24.3% (pHSVRFPpur in G16-9). The frequency of yellow cells, as calculated from the merged images was much lower than expected (at 0.03% in both MRC5-V2 and G16-9), particularly in G16-9. There are several possibilities to explain this result. The frequency of dually infected cells may in fact have been higher but they were not visible due to one of the two fluorescent genes not yet being expressed. The simultaneous expression of the two fluorescent proteins may be toxic to the cells. This could be overcome by the use of different markers. Another possibility is that one of the reporter fluorescent signals was stronger than the

other, and many cells were considered either green or red only when they were in fact yellow. Or, there could be something preventing the dual infection. To investigate these factors further, amplicon preparations with a very good titre (10^7) should be used.

Table 5.1: GFP/RFP Expression (%) of HSV Amplicons in Human Cell Lines 24hrs Post Transduction.

	HT1080	MRC5-V2	293	G16-9
pHGKNeo	0.8%	6.6%	0.6%	33%
pHSVRFPpur	0.6%	11%	1%	30%
pHGKNeo (Dual)	0.4%	5.2%	0.6%	18%
pHSVRFPpur (Dual)	0.6%	6.6%	0.6%	24%
Yellow Cells[a] (Dual)	0	0.03%	0	0.03%

a As estimated from the merged images from the pHGKNeo and pHSVRFPpur dual infections. The frequency was too low to observe in HT1080 and 293 cells.
All cell lines in Table 5.1 were susceptible to amplicon infection with a range of efficiencies observed (0.4% (pHGKNeo, dual infection, HT1080) to 33% (pHGKNeo, G16-9)).

Approximately 24 hours post transduction with HSV-1 amplicons, selection was applied to the media to allow for the recovery of stably transformed clones (Table 5.2). Both G418 and puromycin have been used as a selective agent in the laboratory before thus the selective levels were known.

Table 5.2: Selection Applied to Cells Transduced with pHGKNeo and pHSVRFPpur HSV-1 Amplicons.

Cell Type	pHGKNeo	pHSVRFPpur	Double Infection	
	G418 (µg/ml)	Puromycin (µg/ml)	G418 (µg/ml)	Puromycin (µg/ml)
HT1080	350	0.5	125	0.25
MRC5-V2	350	0.5	125	0.25
293	200	0.3	70	0.17
G16-9	500	0.75	180	0.35

Unfortunately no stable clones were obtained. Possible reasons for this could have been: the low transduction efficiency, probably due to the low amplicon titre, together with problems from bacterial contamination experienced during these experiments.

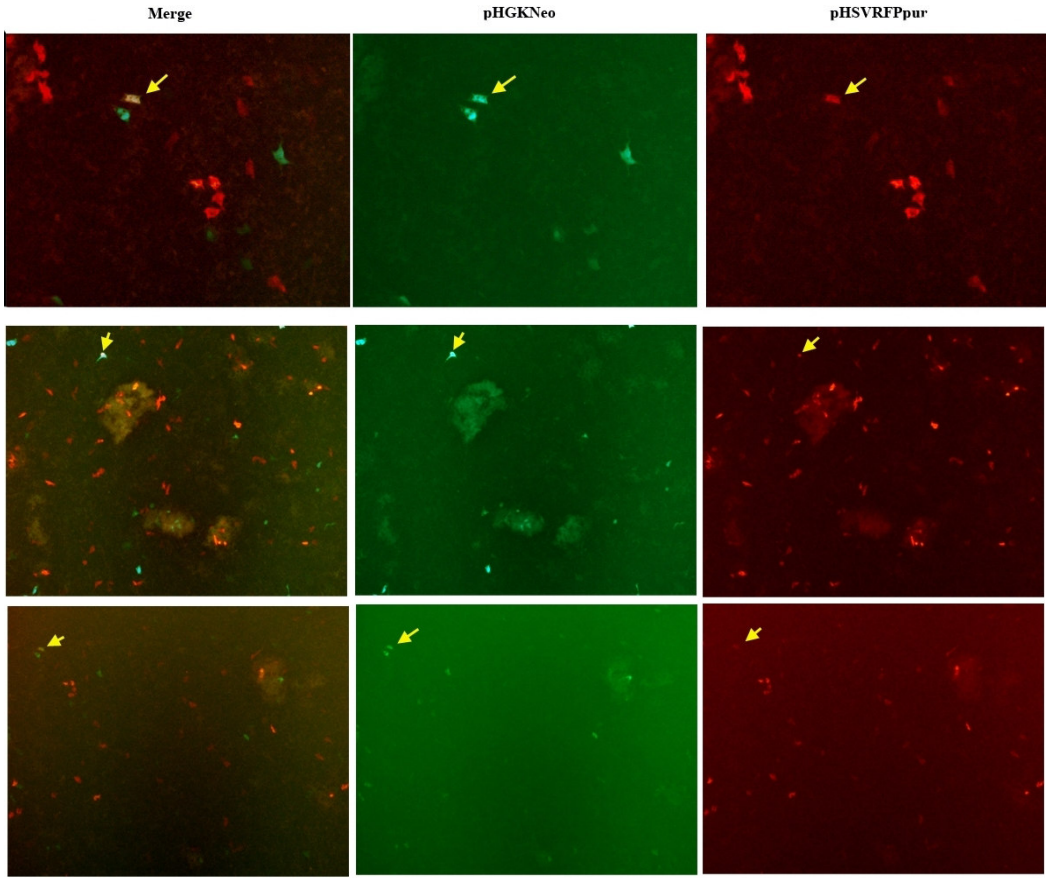


Figure 5.8: GFP and RFP Expression in MRC5-V2 Cells Following Transduction with HSV-1 HAC Amplicons

GFP expression in MRC5-V2 cells 24 hours after infection, following transduction with pHGKNeo and RFP expression following transduction with pHSVRFPpur. The two images were merged and dually transformed cells are indicated with a yellow arrow. All images are at 4x objective.

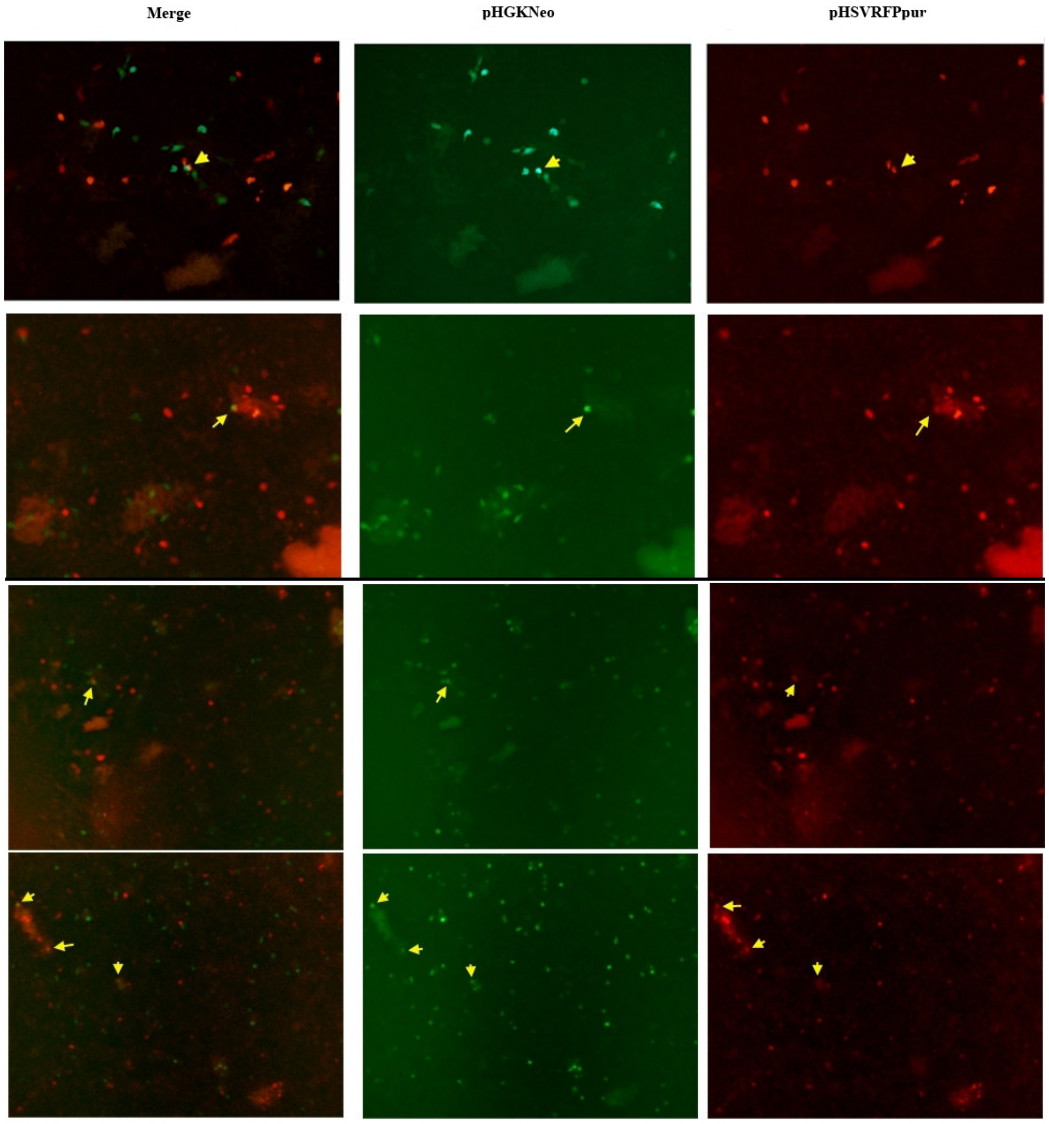


Figure 5.9: GFP and RFP Expression in G16-9 Cells Following Transduction with HSV-1 HAC Amplicons

GFP expression in G16-9 cells 24 hours after infection, following transduction with pHGKNeo and RFP expression following transduction with pHSVRFPpur. The two images were merged and dually transformed cells are indicated with a yellow arrow. All images are at 4x objective.

5.4 Conclusions and Discussion

The aim of the work in this chapter was to develop a vector to obtain gene expression from a genomic locus carrying HAC using the HSV-1 system. As previously discussed, HACs have become an important gene vector transfer system for non-viral gene delivery into human cells (Grimes et al., 2001; Ikeno et al., 2002; Mejía and Larin, 2000; Mejía et al., 2001) and the HSV-1 system has been used to deliver genomic BACs containing genes for complementing deficiencies (Wade-Martins et al., 2003; Wade-Martins et al., 2001).

The first section of this work investigated the possibility of engineering the HPRT genomic locus into the 65 kb pHSV17α302Neo vector by Red mediated recombination for complementing the HPRT deficiency in HPRT- HT1080 cells using the HSV-1 system. The final HAC gene vector construct for HSV-1 delivery is constrained to150 kb due to the packaging capacity of the HSV-1 amplicons. The HPRT genomic locus was present on a large, 175 kb, PAC construct, hPAC71G4 (Schindelhauer and Cooke, 1997) and therefore this construct needed to be modified by Red/ET recombination to develop a smaller construct suitable for recombination with the 17α BAC pHSV17α302Neo, 65 kb. This approach was hampered by numerous PCR technical problems. The PCR only worked sporadically and a number of control experiments were set up including varying the primer and template concentration, using different dNTPs, preparing fresh DNA template and using circular template, utilising different PCR machines and increasing the extension time. None of these had an effect. Finally, phenol-chloroform purification of the template led to success and a visible 7 kb product. However, the reaction remained unreliable. New primers were ordered and aliquoted to avoid repeated freezing and thawing.

Four different commercial enzymes were tested (designed specifically for long range PCR) and the elongation time was increased for 20 seconds/cycle. However, none of these approaches gave consistent results. Finally, enough successful product was generated to transform the arabinose induced electrocompetent cells containing the HPRT PAC and the pRed/ET plasmid. In this step the HPRT genomic locus of the PAC to be subcloned should recombine into the linear vector leading to a circular molecule. No colonies were obtained, probably due to the poor quality of the PCR product.

However, Red mediated recombination has been successful previously and has been used to generate a BAC containing the entire CFTR gene (Kotzamanis and Huxley, 2004) and to construct a HAC vector containing 70 kb of alphoid DNA on a 156 kb BAC carrying the human HPRT gene (Kotzamanis et al., 2005). One further possibility to pursue would be an alternative method which has utilised Tn5-based transposons to construct HAC vectors (Basu et al., 2005) to enable a BAC with alpha satellite DNA, selectable markers and transposase recognition elements to be recombined with a second BAC containing a large genomic segment.

The second section of this work investigated whether using a co-infection approach with the HSV-1 system is possible. pHSVRFPpur (Figure 5.6A) was generated with relative ease by Cre-*loxP* - mediated recombination, transforming pBelopur7 electrocompetent cells with pHR. HSV-1 amplicons were generated and a dual infection with pHGKNeo and pHSVRFPpur yielded cells expressing both GFP and RFP (Figures 5.8 and 5.9). The frequency of yellow cells, as calculated from the merged images was much lower than expected. Contributory factors could include

expression differences between the two fluorescent genes, toxicity to the cells, or, there could be something preventing the dual infection. Using amplicon preparations with a high titre (10^7) would enable further study. Unfortunately no stable clones were obtained after the application of selection, a possible reason being the low transduction efficiency, probably due to the low amplicon titre.

Despite the lack of stable clones, this initial experiment has shown that transducing cells with two different amplicons is possible. The transduction efficiency in the dually infected cells was lower than that of the cells transduced with one amplicon only, however, transduction with two amplicons was not so inhibitory as to prevent the observation of fluorescent cells even with a low amplicon titre. Further work could focus on applying this approach to generate artificial chromosomes in cells which can complement deficiencies in a similar way to previous co-transfection studies (Grimes et al., 2001; Ikeno et al., 2002).

Chapter Six: Conclusions and Discussion

6.1 Project Summary

This study was carried out in three separate stages. The first stage (Chapter 3) involved the development of an efficient HSV-1 system for delivering HAC vectors to cells. We successfully engineered HSV-1 amplicons containing HAC vector DNA and infected several human cell types including human HT1080 cells which are commonly used for HAC studies. The transduction efficiency, determined by transient expression of the reporter gene, GFP, was several orders of magnitude higher than lipofection. HACs were generated with a reasonable efficiency in human cell types other than HT1080. This is a noteworthy improvement on standard lipid-based delivery methods which only appear to be successful in HAC formation in HT1080 cells. All HACs were able to bind specific centromeric proteins, as demonstrated by immunostaining with specific antibodies. HACs in HT1080 and G16-9 cells were stable up to 90 days in the absence of selection. However, in MRC5-V2 and 293 cells, the HACs were unstable suggesting that the requirements for centromere function and HAC stability are different in each cell type.

The second stage of the study (Chapter 4) investigated this further. To determine if the HACs were the only chromosomes displaying stability problems, the basal chromosomal instability levels of all cell lines were characterised by CBMN assays. In MRC5-V2 and 293 cells very high levels of non-disjunction events were detected, involving all the tested chromosomes. Several candidate genes were analysed by real-time PCR and Aurora B was found to be over expressed in MRC5-V2 and 293 cells, compared to both MRC5 and HT1080 cells. High levels of Aurora B and Topo II

proteins were also evidenced by western blot analysis, and the protein levels showed statistically significant correlation with the chromosomal instability levels. Since Aurora B is a kinase involved in several processes, such as chromosome condensation and correct chromosome alignment on the mitotic spindle, we perturbed the cellular phosphorylation levels of cell lines containing stable HACs using okadaic acid, with the aim of reproducing conditions similar to those observed in 293 and MRC5-V2 cells. OA acts as a protein phosphatase inhibitor and activator of Aurora B. Treatment of cells showed an increase in both HAC and overall chromosomal instability and an increase in H3 S10/28 phosphorylation. Among the various OA targets, apart from Aurora B, PP2A was investigated as a possible candidate for the increase in instability and H3 phosphorylation but results did not give a clear-cut answer.

The third stage (Chapter 5) involved developing a novel gene expression system utilising HSV-1 amplicons. Investigation of one approach involved engineering the HPRT genomic locus into a 17α HSV-HAC vector (pHSV17α302Neo, 65kb) by Red mediated recombination for complementing the HPRT deficiency in HPRT- HT1080 cells. Another approach included investigating if a co-infection approach with two different HSV-1 amplicons was feasible.

6.2 Conclusions to Chapter 3: The Infectious Delivery of Human Artificial Chromosomes into Human Cells using HSV-1 Amplicon Vectors

The work described in Chapter 3 aimed to overcome the current poor transfection efficiencies of HAC constructs into mammalian cells using cationic lipids. A novel HAC vector system for gene delivery into human cells was developed using the more efficient HSV-1 system. This allows the delivery of large genomic fragments (150 kb) into cells, without immunogenic effects, using HSV-1 amplicons in a helper virus free system. This system was chosen for the development of HAC delivery due to the high efficiency (up to 100%) it has shown in delivering BACs and PACs to human and mouse cells (Wade-Martins et al., 2001) and for complementing genetic deficiencies (Wade-Martins et al., 2003).

The result of this study was the generation of HACs in several cell types with viable efficiencies. The efficiency was determined by transient expression of the reporter gene, GFP. What is of interest with regards to a gene therapy approach is that the efficiency, whilst generally dependent on the cell type, did not always increase proportionally with an increase in MOI. In a therapeutic setting, this could mean that a smaller number of transducing particles may be just as efficient as a large number. HACs were generated in MRC5-V2, 293 and G16-9 cells for the first time. No HACs have been obtained in these cells using standard transfection techniques. Perhaps not enough DNA entered the cells using this method or the DNA was damaged in the process, possibly due to shearing. In addition, the HAC DNA delivered to cells using the HSV-1 system is packaged by a mammalian cell line. This may result in epigenetic marking such as methylation which would be lacking in the BAC vectors extracted from bacterial cells and delivered via lipofection.

The generation of HACs in a variety of cell lines allowed the detection of chromosome stability differences. The findings indicated that the requirements for centromere function and HAC stability are different in each cell type and may depend on a number of factors including the genetic background, and/or epigenetic factors which influence chromosome segregation (Rudd et al., 2003). HACs in HT1080 and G16-9 cells were stable up to 90 days in the absence of selection and reacted to human ACA and anti-human CENP-A antibodies, indicating that they bound CENP-A and formed an active centromere. In contrast, the HACs in MRC5-V2 and 293 cells were unstable in the presence and absence of selection. Possible candidates for this instability were studied in Chapter 4. A further development to this study would be to use the HACs as a segregation marker in the investigation of the effect of substances which influence chromosome stability, or in the study of knock-out/over expression of possible candidate genes. The HAC-HSV system makes this possible as the HACs express GFP. Thus, when cells lose the HAC, they lose the GFP expression, making it very easy to follow HAC segregation. This would reduce the need for time-consuming karyotype analysis.

No HACs were obtained from the MRC5 cells. These primary cells only have a limited lifetime and senescence begins to occur once the cells have been passaged approximatly twenty times. Despite the cells being obtained from the supplier at the lowest possible passage, once cell stocks had been frozen, and then seeded and transduced with amplicons, they always entered senescence, either as an effect of the infection, or simply because they reached their terminal lifespan before the clones were large enough to be analysed. These difficulties could be overcome with the use

of staminal cells which are capable of a prolonged period of culture without differentiating.

This work shows the value of the HSV-1 delivery method in the generation of HACs. It is relatively easy to transform HAC vectors such as 302P22 into HAC vectors such as pHSV17α302Neo due to the easy manipulation of pHGKNeo either by Cre-*loxP* recombination or restriction digestion and ligation. Further advantages of the HSV-1 system are: its ease of use; the broad host range, indicated by the generation of HACs in cell types other than HT1080 at reasonable efficiencies; the ability to infect dividing and non-dividing cells; and the absence of cytotoxic and cytopathic effects (Olschowka et al., 2003). This makes amplicon delivery of genes via a disabled HSV-1 system a safe alternative to viral vector systems where immunogenic effects have been observed (Check, 2003).

6.3 Conclusions to Chapter 4: Investigation of a Possible Candidate for HAC Instability

Real-time PCR allowed the accurate quantitation of gene expression of several possible candidates responsible for the observed chromosome instability in MRC5-V2 and 293 cells. Aurora B, was found to be significantly over expressed in both MRC5-V2 and 293 cells in comparison with MRC5 cells and was also over expressed with relation to the HT1080 lines and G16-9 cells (where HACs are stable). It is possible that in future studies, synchronising the cell populations may highlight more differences in expression between other candidate genes, due to the cell cycle dependent expression of some of the genes studied.

High levels of chromosomal instability in MRC5-V2 and 293 cells correlated with high levels of Aurora B and Topo II as confirmed by western blot analysis. Aurora B is required for the phosphorylation of H3S10 and H3S28 (Adams et al., 2001b; Goto et al., 2002; Hsu et al., 2000) before mitosis and is involved in chromosome condensation, in the checkpoint controlling the correct alignment of chromosomes on the metaphase spindle, and in the congression of chromosomes at metaphase (Hauf et al., 2003; Vagnarelli and Earnshaw, 2004). In cancer research, the main effect of Aurora B over expression is thought to be due to increased phosphorylation of H3S10/S28 and consequently condensation errors and chromosomal instability (Ota et al., 2002). This does not appear to be the case from the experiments outlined above. 293 cells have the highest instability levels, yet the levels of H3S10/28 phosphorylation were not significantly increased. This implies that either there is another pathway where the main effect takes place, for example, interference with the spindle checkpoint mechanism, thus allowing the cells to proceed to anaphase when the chromosomes are not all correctly bi-oriented on the spindle. This could explain the high number of non-disjunction event observed in the two cell lines, or there is something antagonising the Aurora B kinase activity on H3S10/S28 in 293 cells. To check this hypothesis, the overall levels of protein phosphorylation in 293 cells would need to be checked to determine if there is a general decrease in the levels of phosphorylation or if it was only affecting H3 phosphorylation. Interestingly, 293 cells are extremely sensitive to treatment with okadaic acid, and die at very low concentrations. This implies that the overall phosphorylation level of the proteins in these cells is low and little inhibition is necessary to tip the cells into non-viability.

OA, a protein phosphatase inhibitor and activator of Aurora B, caused an overall increase in chromosomal instability and therefore a decrease in stability of HACs in the cell lines LJ2-1 and AG6-1. Western blot analysis confirmed that OA treatment resulted in an increase in H3S10 phosphorylation. H3S10 and H3S28 are phosphorylated prior to mitosis by Aurora B (Goto et al., 2002), peaking when chromosomes are in their most condensed state. After chromosome segregation, Aurora B dissociates from chromosomes (Adams et al., 2001a) and H3 phosphorylation at both sites decreases (Goto et al., 2002). Because of the increased phosphorylation levels induced by the OA treatment, we could have expected errors in the strict progression from non-phosphorylated, uncondensed chromosomes to phosphorylated mitotic chromosomes, and back to non-phosphorylated, decondensed interphase chromosomes. The observation of no mislocalisation of H3S28 in interphase cells indicates that this mechanism, required for correct chromosome segregation, is still functioning correctly at the concentrations of OA utilised. Since this and other phosphorylation pathways are likely to be important for cell viability, it could account for the observed cell death at higher levels of OA. However, we cannot rule out that the mislocalisation of phosphorylated H3 to interphases may be below the detection of these experiments. Further work could involve a functional assay of Aurora B activity in the various cell types as well as the study of other kinase/phosphatase proteins. Among these, the phosphatase PP2A, as it inactivates Aurora B in a dose-dependent manner (Sugiyama et al., 2002) and could be a possible candidate for the removal of phosphorylation of H3S10 and H3S28 after mitosis. Western blot analysis however showed no clear difference in PP2A levels with okadaic acid treatment. To determine if PP2A is involved in the chromosome

instability, the amount of protein present in the various cell lines used in this study will be characterised in further studies.

6.4 Conclusions to Chapter 5: Development of a Gene Expression System Using HSV-1 Amplicons

The work of Chapter 5 focussed on obtaining gene expression from genomic loci carried by HSV-1 HACs. The HSV-1 system is limited to a final HAC vector size of 150 kb due to the packaging capacity of the amplicons. Therefore, the first section of this work looked at engineering the HPRT genomic locus, present on a 175 kb PAC, into a smaller construct by Red mediated recombination, which would be suitable for recombination with the 65 kb pHSV17α302Neo vector. This would allow the complementation of the HPRT deficiency in HPRT- HT1080 cells using the HSV-1 system. The HPRT locus was chosen as it is well characterised, easy to select for and HPRT- HT1080 cells were readily available.

This approach was greeted with many PCR technical problems. The PCR production of a linear pBeloBAC11 vector only worked sporadically despite varying conditions many times and pursuing this further would have been beyond the time constraints of this study. Because of the difficulties associated with the long PCR reaction, one possibility to pursue would be to try and use a smaller vector such as pBAC108L which is 6.9 kb as opposed to pBeloBAC11, 7.4 kb.

An alternative approach to pursue would be to utilise Tn5-based transposons to construct HAC vectors (Basu et al., 2005) to enable a BAC with alpha satellite DNA,

selectable markers and transposase recognition elements to be recombined with a second BAC containing a large genomic segment.

One way to overcome the HSV size limitation is by co-infection. A co-infection approach with the HSV-1 system was shown to be successful. pHSVRFPpur was generated successfully by Cre-*loxP* – mediated recombination. HSV-1 amplicons were generated and a dual infection with pHGKNeo and pHSVRFPpur resulted in yellow cells expressing both GFP and RFP. However, the frequency of dual transduction events, as calculated from the merged images was much lower than expected. There are several possible explanations for this. Maybe the expression of the two fluorescent reporters was not simultaneous and so one of the two fluorescent proteins was not yet present at the time of cell counting. Alternatively, one of the reporter fluorescent signals could have been stronger than the other resulting in cells being considered either green or red only when they were in fact yellow. Lastly, the simultaneous expression of the two fluorescent proteins may be toxic to the cells and so different markers might be required. However, it cannot be ruled out that there could be some cellular system preventing the dual infection. Using amplicon preparations with a very good titre (10^7/ml) would enable these questions to be addressed more fully. Although no stable clones were obtained after the application of selection, probably due to the low amplicon titre, the results of initial experiments look promising. This work suggests that co-infection using the HSV-1 system could be a feasible approach to generating artificial chromosomes in cells for complementing deficiencies in a similar way to previous co-transfection studies (Grimes et al., 2001; Ikeno et al., 2002).

6.5 The Future of HACs as a Therapeutic Approach

6.5.1 Advantages Over Viral Gene Therapy

Human Artificial Chromosomes offer an alternative strategy to viral-based gene therapy. The use of viral vectors for gene therapy has had varying degrees of success. In 2000, the first successful gene therapy was reported. *Ex vivo* transduction of haematopoietic stem cells with a retroviral vector, resulted in the cure of three children from the fatal immunodeficiency disease SCID-X1 (Cavazzana-Calvo et al., 2000). However, three years later two of the eleven patients developed leukaemia-like disease due to retroviral integration in proximity to the LMO2 proto-oncogene (a central regulator of haematopoiesis) promoter. This lead to deregulated transcription and expression of LMO2 (Hacein-Bey-Abina et al., 2003). In addition to the report above, retrovirally-transduced bone marrow cells resulted in leukaemia in mice (Li et al., 2002). Also, there have been reports that random integration can lead to position effects, resulting in loss of expression through DNA modifications such as methylation (Challita and Kohn, 1994).

In 1999, Jesse Gelsinger died as a direct cause of the administration of an adenovirus vector being used in a viral gene therapy clinical trial to treat an Ornithine Transcarbamylase (OTC) deficiency, which causes an accumulation of ammonia in the bloodstream. The dissemination of the vector from the liver to the blood circulation triggered a massive inflammatory response which ultimately led to multi organ failure (Marshall, 1999; Thomas et al., 2003). Previous exposure to the virus may have sensitised his immune response (Bostanci, 2002). This was a serious set back to gene therapy.

The problems described above, such as insertional mutagenesis and immunogenicity are being tackled in various new generation viral vectors (Thomas et al., 2003; Verma and Weitzman, 2004). However, HACs offer a viable alternative. They can avoid such problems encountered with viral systems.

6.5.2 Criteria for Successful Gene Therapy

The ultimate goal of gene therapy is "to achieve the authentic expression of transgenes or nucleic acid sequences of therapeutic value in a specific cell type in the absence of undesired interactions with the host's genome and without the risk of cellular transformation or stimulating the host's immune system" (Lipps et al., 2003).

6.5.2.1 Authentic Expression

Virus vectors are limited by the space available for exogenous DNA. Replication defective adenoviruses only have a packaging capacity of 8 kb as do retro and lentiviruses while adeno-associated viruses can only incorporate 5 kb (Thomas et al., 2003). HACs are capable of carrying entire genes and their regulatory elements (Grimes et al., 2001; Mejía et al., 2001) due to their large cloning capacity. These regulatory elements may allow physiologically relevant and appropriately regulated tissue specific expression and allow the faithful mimicry of normal gene expression. In addition, HACs provide a chromatin context that is comprised of both euchromatin and chromatin (Grimes et al., 2004), thus providing a suitable environment for gene

expression. The large cloning capacity of HACs may overcome the current size limitations that conventional viral vectors impose.

Recently HACs have been reported to be stable in human primary fibroblasts (Kakeda et al., 2005) , expressing human erythropoietin for twelve weeks. This is another property required for a gene therapy vector: sustained expression of the transgene. Work in our laboratory with an HSV-1 HAC has shown prolonged expression from a gene not subject to selective pressure (Moralli et al., 2006). In addition to *de novo* HACs, transfer of a mammalian SATAC into human mesenchymal stem cells showed correct segregation and stable expression of an exogenous gene throughout differentiation (Vanderbyl et al., 2004). Voet et al. produced a human chromosomal vector which was transmitted through the germline of chimeric mice (Voet et al., 2001) illustrating the potential of HACs as gene therapy vectors in stem cells. Several studies have also illustrated the potential of HACs to transmit through the germline (Co et al., 2000; Kuroiwa et al., 2002). In the latter study, transchromosomic calves produced human immunoglobulins from a HAC vector, further indicating the importance of HACs as therapeutic vectors. A recent study has followed GFP-tagged *de novo* HACs and revealed that they are properly aligned at the spindle mid-zone in cells and segregate with the same timing as the host chromosomes (Tsuduki et al., 2006) thus illustrating the stability of the HACs.

However, the efficient delivery of HACs/minichromosomes into human cells is a major obstacle and has hindered progress in developing gene expression systems. Microcell-mediated chromosome transfer is widely used to shuttle HACs and minichromosomes between cell lines and has been used to transfer both into mouse

embryonic stem cells (Kuroiwa et al., 2000; Tomizuka et al., 2000; Voet et al., 2001) and somatic cells (Alazami et al., 2004; Ikeno et al., 2002) and the chicken DT40 cell line (Shen et al., 2001b). However, it is inefficient, as is transfection as discussed previously. An approach using both cationic lipids and high-frequency ultrasound has been used to transfer previously flow-purified minichromosomes into cells with efficiencies ranging from 12-53% (Oberle et al., 2004). However, flow cytometry is not readily applicable to most artificial chromosomes (those in his study were large AT-rich satellite-repeat based). Microinjection of artificial chromosomes, such as that seen with the introduction of a 60 Mb SATAC into fertilised murine (Co et al., 2000) and bovine embryos (Wang et al., 2001) is a possibility. However, it also requires the purification of minichromosomes by flow cytometry.

The HSV-1 system described in this study is a valuable method for HAC vector delivery, allowing efficient delivery of large genomic fragments (150 kb) into cells. The use of a disabled HSV-1 system to deliver genes by amplicons is a safe alternative method to other viral vectors systems in which reactions have been observed. There have been no such effects reported in response to HSV-1 amplicons *in vivo* (Costantini et al., 1999; Olschowka et al., 2003). This method could particularly be applied *ex vivo*.

6.5.2.2 Absence of Undesired Interactions and Immune Responses

The extrachromosomal nature, usually as single-copy entities, ensures that the introduced genes do not interfere or are affected by the host cell genome. HACs have no need to integrate into the host genome for replication due to their use of human

functional components for their replication and stable propagation, thus there is no potential for insertional mutagenesis or transgene silencing. As no viral elements are involved, there is no risk of immune reactions or recombination with wild type virus from a HAC derived entirely from human DNA. This is a vast contrast to the viral gene therapy trials described above where retroviral integration led to leukaemia (Hacein-Bey-Abina et al., 2003) and an immune response to an adenovirus vector led to the death of a patient (Marshall, 1999; Thomas et al., 2003). However, *de novo* HACs have proven difficult to characterise, with the mode of formation giving no control over gene copy number/structural integrity, and the observed integration events must be overcome. In an *ex vivo* gene therapy approach however, populations of cells with no integrations can be selected.

6.5.2.3 The Future of HACs in Stem Cells

Recently there have been huge advances in stem cell technology. Pluripotent embryonic stem cells have first been derived from human blastocysts (Cowan et al., 2004; Thomson et al., 1998) and now have been derived from blastocysts cloned by Somatic Cell Nuclear Transfer (SCNT) with patient-specific embryonic stem cell lines being generated (Vanikar et al., 2007). Directed differentiation will allow the formation from embryonic stem cells of fully differentiated cells. Using extracellular signals, significant progress is being made in this field. Various studies have described the potential of human embryonic stem cells to differentiate into multiple lineages such as neuronal progenitors (Reubinoff et al., 2001), haematopoietic precursors (Kaufman et al., 2001) and insulin-secreting cells (Assady et al., 2001). Several groups have also combined embryonic stem cells and gene therapy. In one

study, embryonic stem cells were modified by lentiviral vectors and maintained GFP expression throughout differentiation (Gropp et al., 2003). The next step in the HSV-1 system would be to use stem cells as a target for the HSV amplicons to demonstrate that HACs can be formed in primary cells.

Further research demonstrated that nuclear transfer can be combined with gene therapy to treat a genetic disorder (Rideout et al., 2002). Tail-tip cells were derived from a mouse with a deficiency in the recombination-activating gene 2 (Rag2). The nuclei were transferred into enucleated oocytes and embryonic stem cells were derived from the resulting blastocysts. The Rag2 gene was repaired by homologous recombination and the stem cells were induced to differentiate into embryoid bodies and then haematopoietic precursors. These could then be used to treat the Rag2 deficiency (Rideout et al., 2002). These exciting developments raise the possibility of creating patient-specific stem cells which contain a stable, transgene-expressing artificial chromosome, possibly introduced through the HSV-1 amplicon system. Stem cells could then undergo directed differentiation before introduction to the patient. These differentiated cells would express the therapeutic gene in a tissue-specific manner with no insertional or immunogenic effect. Prior to this, animal cell studies, notably murine, would be required. However, various groups including our own (Alazami et al., 2004), have run into difficulties generating *de novo* HACs in murine cells at a reasonable frequency. The HSV-1 system could offer an alternative approach and the possibility of generating physiologically realistic transgene expression to enable the further development of this exciting field.

References

Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, A., Lebersorger, A., Singh, P. B. et al. (1999). Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. *Embo J* **18**, 1923-38.

Aagaard, L., Schmid, M., Warburton, P. and Jenuwein, T. (2000). Mitotic phosphorylation of SUV39H1, a novel component of active centromeres, coincides with transient accumulation at mammalian centromeres. *J Cell Sci* **113 (Pt 5)**, 817-29.

Adams, R. R., Carmena, M. and Earnshaw, W. C. (2001a). Chromosomal passengers and the (aurora) ABCs of mitosis. *Trends Cell Biol* **11**, 49-54.

Adams, R. R., Maiato, H., Earnshaw, W. C. and Carmena, M. (2001b). Essential roles of Drosophila inner centromere protein (INCENP) and aurora B in histone H3 phosphorylation, metaphase chromosome alignment, kinetochore disjunction, and chromosome segregation. *J Cell Biol* **153**, 865-80.

Adams, R. R., Wheatley, S. P., Gouldsworthy, A. M., Kandels-Lewis, S. E., Carmena, M., Smythe, C., Gerloff, D. L. and Earnshaw, W. C. (2000). INCENP binds the Aurora-related kinase AIRK2 and is required to target it to chromosomes, the central spindle and cleavage furrow. *Curr Biol* **10**, 1075-8.

Ainsztein, A. M., Kandels-Lewis, S. E., Mackay, A. M. and Earnshaw, W. C. (1998). INCENP centromere and spindle targeting: identification of essential conserved motifs and involvement of heterochromatin protein HP1. *J Cell Biol* **143**, 1763-74.

Alazami, A. M., Mejía, J. E. and Monaco, Z. L. (2004). Human artificial chromosomes containing chromosome 17 alphoid DNA maintain an active centromere in murine cells but are not stable. *Genomics* **83**, 844-51.

Allshire, R. C., Nimmo, E. R., Ekwall, K., Javerzat, J. P. and Cranston, G. (1995). Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. *Genes Dev* **9**, 218-33.

Amon, A. (1999). The spindle checkpoint. *Curr Opin Genet Dev* **9**, 69-75.

Amor, D. J., Bentley, K., Ryan, J., Perry, J., Wong, L., Slater, H. and Choo, K. H. (2004a). Human centromere repositioning "in progress". *Proc Natl Acad Sci U S A* **101**, 6542-7.

Amor, D. J. and Choo, K. H. (2002). Neocentromeres: role in human disease, evolution, and centromere study. *Am J Hum Genet* **71**, 695-714.

Amor, D. J., Kalitsis, P., Sumer, H. and Choo, K. H. (2004b). Building the centromere: from foundation proteins to 3D organization. *Trends Cell Biol* **14**, 359-68.

Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E. and Moudrianakis, E. N. (1991). The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. *Proc Natl Acad Sci U S A* **88**, 10148-52.

Ashar, H. R., James, L., Gray, K., Carr, D., Black, S., Armstrong, L., Bishop, W. R. and Kirschmeier, P. (2000). Farnesyl transferase inhibitors block the farnesylation of CENP-E and CENP-F and alter the association of CENP-E with the microtubules. *J Biol Chem* **275**, 30451-7.

Assady, S., Maor, G., Amit, M., Itskovitz-Eldor, J., Skorecki, K. L. and Tzukerman, M. (2001). Insulin production by human embryonic stem cells. *Diabetes* **50**, 1691-7.

Auriche, C., Carpani, D., Conese, M., Caci, E., Zegarra-Moran, O., Donini, P. and Ascenzioni, F. (2002). Functional human CFTR produced by a stable minichromosome. *EMBO Rep* **3**, 862-8.

Auriche, C., Donini, P. and Ascenzioni, F. (2001). Molecular and cytological analysis of a 5.5 Mb minichromosome. *EMBO Rep* **2**, 102-7.

Barnett, M. A., Buckle, V. J., Evans, E. P., Porter, A. C., Rout, D., Smith, A. G. and Brown, W. R. (1993). Telomere directed fragmentation of mammalian chromosomes. *Nucleic Acids Res* **21**, 27-36.

Basu, J., Stromberg, G., Compitello, G., Willard, H. F. and Van Bokkelen, G. (2005). Rapid creation of BAC-based human artificial chromosome vectors by transposition with synthetic alpha-satellite arrays. *Nucleic Acids Res* **33**, 587-96.

Basu, J. and Willard, H. F. (2005). Artificial and engineered chromosomes: non-integrating vectors for gene therapy. *Trends Mol Med* **11**, 251-8.

Baum, M., Ngan, V. K. and Clarke, L. (1994). The centromeric K-type repeat and the central core are together sufficient to establish a functional Schizosaccharomyces pombe centromere. *Mol Biol Cell* **5**, 747-61.

Bischoff, J. R. and Plowman, G. D. (1999). The Aurora/Ipl1p kinase family: regulators of chromosome segregation and cytokinesis. *Trends Cell Biol* **9**, 454-9.

Black, B. E., Brock, M. A., Bedard, S., Woods, V. L., Jr. and Cleveland, D. W. (2007a). An epigenetic mark generated by the incorporation of CENP-A into centromeric nucleosomes. *Proc Natl Acad Sci U S A* **104**, 5008-13.

Black, B. E., Foltz, D. R., Chakravarthy, S., Luger, K., Woods, V. L., Jr. and Cleveland, D. W. (2004). Structural determinants for generating centromeric chromatin. *Nature* **430**, 578-82.

Black, B. E., Jansen, L. E., Maddox, P. S., Foltz, D. R., Desai, A. B., Shah, J. V. and Cleveland, D. W. (2007b). Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. *Mol Cell* **25**, 309-22.

Blackburn, E. H. (1984). The molecular structure of centromeres and telomeres. *Annu Rev Biochem* **53**, 163-94.

Blow, J. J., Gillespie, P. J., Francis, D. and Jackson, D. A. (2001). Replication origins in Xenopus egg extract Are 5-15 kilobases apart and are activated in clusters that fire at different times. *J Cell Biol* **152**, 15-25.

Blower, M. D. and Karpen, G. H. (2001). The role of Drosophila CID in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nat Cell Biol* **3**, 730-9.

Blower, M. D., Sullivan, B. A. and Karpen, G. H. (2002). Conserved organization of centromeric chromatin in flies and humans. *Dev Cell* **2**, 319-30.

Bomont, P., Maddox, P., Shah, J. V., Desai, A. B. and Cleveland, D. W. (2005). Unstable microtubule capture at kinetochores depleted of the centromere-associated protein CENP-F. *Embo J* **24**, 3927-39.

Bostanci, A. (2002). Gene therapy. Blood test flags agent in death of Penn subject. *Science* **295**, 604-5.

Brown, K. E., Barnett, M. A., Burgtorf, C., Shaw, P., Buckle, V. J. and Brown, W. R. (1994). Dissecting the centromere of the human Y chromosome with cloned telomeric DNA. *Hum Mol Genet* **3**, 1227-37.

Brown, M. T., Goetsch, L. and Hartwell, L. H. (1993). MIF2 is required for mitotic spindle integrity during anaphase spindle elongation in Saccharomyces cerevisiae. *J Cell Biol* **123**, 387-403.

Buchwitz, B. J., Ahmad, K., Moore, L. L., Roth, M. B. and Henikoff, S. (1999). A histone-H3-like protein in C. elegans. *Nature* **401**, 547-8.

Buffin, E., Lefebvre, C., Huang, J., Gagou, M. E. and Karess, R. E. (2005). Recruitment of Mad2 to the kinetochore requires the Rod/Zw10 complex. *Curr Biol* **15**, 856-61.

Burke, D. T., Carle, G. F. and Olson, M. V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**, 806-12.

Bustin, S. A. (2000). Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* **25**, 169-93.

Campbell, M. S. and Gorbsky, G. J. (1995). Microinjection of mitotic cells with the 3F3/2 anti-phosphoepitope antibody delays the onset of anaphase. *J Cell Biol* **129**, 1195-204.

Carpenter, A. J. and Porter, A. C. (2004). Construction, characterization, and complementation of a conditional-lethal DNA topoisomerase IIalpha mutant human cell line. *Mol Biol Cell* **15**, 5700-11.

Caskey, C. T. and Kruh, G. D. (1979). The HPRT locus. *Cell* **16**, 1-9.

Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J. L. et al. (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288**, 669-72.

Challita, P. M. and Kohn, D. B. (1994). Lack of expression from a retroviral vector after transduction of murine hematopoietic stem cells is associated with methylation in vivo. *Proc Natl Acad Sci U S A* **91**, 2567-71.

Chan, G. K., Jablonski, S. A., Starr, D. A., Goldberg, M. L. and Yen, T. J. (2000). Human Zw10 and ROD are mitotic checkpoint proteins that bind to kinetochores. *Nat Cell Biol* **2**, 944-7.

Chan, G. K., Schaar, B. T. and Yen, T. J. (1998). Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and hBUBR1. *J Cell Biol* **143**, 49-63.

Check, E. (2003). Cancer risk prompts US to curb gene therapy. *Nature* **422**, 7.

Chen, H. L., Tang, C. J., Chen, C. Y. and Tang, T. K. (2005). Overexpression of an Aurora-C kinase-deficient mutant disrupts the Aurora-B/INCENP complex and induces polyploidy. *J Biomed Sci* **12**, 297-310.

Chong, J. P., Thommes, P. and Blow, J. J. (1996). The role of MCM/P1 proteins in the licensing of DNA replication. *Trends Biochem Sci* **21**, 102-6.

Choo, K. H. (2000). Centromerization. *Trends Cell Biol* **10**, 182-8.

Choo, K. H., Vissel, B., Nagy, A., Earle, E. and Kalitsis, P. (1991). A survey of the genomic distribution of alpha satellite DNA on all the human chromosomes, and derivation of a new consensus sequence. *Nucleic Acids Res* **19**, 1179-82.

Clarke, D. J. and Gimenez-Abian, J. F. (2000). Checkpoints controlling mitosis. *Bioessays* **22**, 351-63.

Clarke, L. and Carbon, J. (1980). Isolation of the centromere-linked CDC10 gene by complementation in yeast. *Proc Natl Acad Sci U S A* **77**, 2173-7.

Cleveland, D. W., Mao, Y. and Sullivan, K. F. (2003). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. *Cell* **112**, 407-21.

Co, D. O., Borowski, A. H., Leung, J. D., van der Kaa, J., Hengst, S., Platenburg, G. J., Pieper, F. R., Perez, C. F., Jirik, F. R. and Drayer, J. I. (2000). Generation of transgenic mice and germline transmission of a mammalian artificial chromosome introduced into embryos by pronuclear microinjection. *Chromosome Res* **8**, 183-91.

Conese, M., Auriche, C. and Ascenzioni, F. (2004). Gene therapy progress and prospects: episomally maintained self-replicating systems. *Gene Ther* **11**, 1735-41.

Cooke, C. A., Bernat, R. L. and Earnshaw, W. C. (1990). CENP-B: a major human centromere protein located beneath the kinetochore. *J Cell Biol* **110**, 1475-88.

Cooke, H. J. (2004). Silence of the centromeres--not. *Trends Biotechnol* **22**, 319-21.

Copeland, N. G., Jenkins, N. A. and Court, D. L. (2001). Recombineering: a powerful new tool for mouse functional genomics. *Nat Rev Genet* **2**, 769-79.

Corbett, K. D. and Berger, J. M. (2004). Structure, molecular mechanisms, and evolutionary relationships in DNA topoisomerases. *Annu Rev Biophys Biomol Struct* **33**, 95-118.

Costantini, L. C., Jacoby, D. R., Wang, S., Fraefel, C., Breakefield, X. O. and Isacson, O. (1999). Gene transfer to the nigrostriatal system by hybrid herpes simplex virus/adeno-associated virus amplicon vectors. *Hum Gene Ther* **10**, 2481-94.

Cowan, C. A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J. P., Wang, S., Morton, C. C., McMahon, A. P., Powers, D. et al. (2004). Derivation of embryonic stem-cell lines from human blastocysts. *N Engl J Med* **350**, 1353-6.

Craig, J. M. (2005). Heterochromatin--many flavours, common themes. *Bioessays* **27**, 17-28.

Craig, J. M., Earle, E., Canham, P., Wong, L. H., Anderson, M. and Choo, K. H. (2003). Analysis of mammalian proteins involved in chromatin modification reveals new metaphase centromeric proteins and distinct chromosomal distribution patterns. *Hum Mol Genet* **12**, 3109-21.

Csonka, E., Cserpan, I., Fodor, K., Hollo, G., Katona, R., Kereso, J., Praznovszky, T., Szakal, B., Telenius, A., deJong, G. et al. (2000). Novel generation of human satellite DNA-based artificial chromosomes in mammalian cells. *J Cell Sci* **113 (Pt 18)**, 3207-16.

Cunningham, C. and Davison, A. J. (1993). A cosmid-based system for constructing mutants of herpes simplex virus type 1. *Virology* **197**, 116-24.

Cutts, S. M., Fowler, K. J., Kile, B. T., Hii, L. L., O'Dowd, R. A., Hudson, D. F., Saffery, R., Kalitsis, P., Earle, E. and Choo, K. H. (1999). Defective chromosome segregation, microtubule bundling and nuclear bridging in inner centromere protein gene (Incenp)-disrupted mice. *Hum Mol Genet* **8**, 1145-55.

Dawe, R. K. and Henikoff, S. (2006). Centromeres put epigenetics in the driver's seat. *Trends Biochem Sci* **31**, 662-9.

deJong, G., Telenius, A. H., Telenius, H., Perez, C. F., Drayer, J. I. and Hadlaczky, G. (1999). Mammalian artificial chromosome pilot production facility: large-scale isolation of functional satellite DNA-based artificial chromosomes. *Cytometry* **35**, 129-33.

Depinet, T. W., Zackowski, J. L., Earnshaw, W. C., Kaffe, S., Sekhon, G. S., Stallard, R., Sullivan, B. A., Vance, G. H., Van Dyke, D. L., Willard, H. F. et al. (1997). Characterization of neo-centromeres in marker chromosomes lacking detectable alpha-satellite DNA. *Hum Mol Genet* **6**, 1195-204.

Dujardin, D., Wacker, U. I., Moreau, A., Schroer, T. A., Rickard, J. E. and De Mey, J. R. (1998). Evidence for a role of CLIP-170 in the establishment of metaphase chromosome alignment. *J Cell Biol* **141**, 849-62.

Dunleavy, E., Pidoux, A. and Allshire, R. (2005). Centromeric chromatin makes its mark. *Trends Biochem Sci* **30**, 172-5.

Earle, E., Saxena, A., MacDonald, A., Hudson, D. F., Shaffer, L. G., Saffery, R., Cancilla, M. R., Cutts, S. M., Howman, E. and Choo, K. H. (2000). Poly(ADP-ribose) polymerase at active centromeres and neocentromeres at metaphase. *Hum Mol Genet* **9**, 187-94.

Earnshaw, W. C., Ratrie, H., 3rd and Stetten, G. (1989). Visualization of centromere proteins CENP-B and CENP-C on a stable dicentric chromosome in cytological spreads. *Chromosoma* **98**, 1-12.

Earnshaw, W. C. and Rothfield, N. (1985). Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* **91**, 313-21.

Ebersole, T. A., Ross, A., Clark, E., McGill, N., Schindelhauer, D., Cooke, H. and Grimes, B. (2000). Mammalian artificial chromosome formation from circular alphoid input DNA does not require telomere repeats. *Hum Mol Genet* **9**, 1623-31.

Ekwall, K., Nimmo, E. R., Javerzat, J. P., Borgstrom, B., Egel, R., Cranston, G. and Allshire, R. (1996). Mutations in the fission yeast silencing factors clr4+ and rik1+ disrupt the localisation of the chromo domain protein Swi6p and impair centromere function. *J Cell Sci* **109 (Pt 11)**, 2637-48.

Everett, R. D., Earnshaw, W. C., Findlay, J. and Lomonte, P. (1999). Specific destruction of kinetochore protein CENP-C and disruption of cell division by herpes simplex virus immediate-early protein Vmw110. *Embo J* **18**, 1526-38.

Faragher, A. J., Sun, X. M., Butterworth, M., Harper, N., Mulheran, M., Ruchaud, S., Earnshaw, W. C. and Cohen, G. M. (2007). Death Receptor-induced Apoptosis Reveals a Novel Interplay between the Chromosomal Passenger Complex and CENP-C during Interphase. *Mol Biol Cell*.

Farr, C., Fantes, J., Goodfellow, P. and Cooke, H. (1991). Functional reintroduction of human telomeres into mammalian cells. *Proc Natl Acad Sci U S A* **88**, 7006-10.

Farr, C. J., Bayne, R. A., Kipling, D., Mills, W., Critcher, R. and Cooke, H. J. (1995). Generation of a human X-derived minichromosome using telomere-associated chromosome fragmentation. *Embo J* **14**, 5444-54.

Feng, J., Huang, H. and Yen, T. J. (2006). CENP-F is a novel microtubule-binding protein that is essential for kinetochore attachments and affects the duration of the mitotic checkpoint delay. *Chromosoma* **115**, 320-9.

Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., Hunt, D. F., Funabiki, H. and Allis, C. D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature*.

Fowler, K. J., Wong, L. H., Griffiths, B. K., Sibson, M. C., Reed, S. and Choo, K. H. (2004). Centromere protein b-null mice display decreasing reproductive performance through successive generations of breeding due to diminishing endometrial glands. *Reproduction* **127**, 367-77.

Fukagawa, T., Mikami, Y., Nishihashi, A., Regnier, V., Haraguchi, T., Hiraoka, Y., Sugata, N., Todokoro, K., Brown, W. and Ikemura, T. (2001). CENP-H, a constitutive centromere component, is required for centromere targeting of CENP-C in vertebrate cells. *Embo J* **20**, 4603-17.

Fukagawa, T., Pendon, C., Morris, J. and Brown, W. (1999). CENP-C is necessary but not sufficient to induce formation of a functional centromere. *Embo J* **18**, 4196-209.

Furuyama, T., Dalal, Y. and Henikoff, S. (2006). Chaperone-mediated assembly of centromeric chromatin in vitro. *Proc Natl Acad Sci U S A* **103**, 6172-7.

Gassmann, R., Carvalho, A., Henzing, A. J., Ruchaud, S., Hudson, D. F., Honda, R., Nigg, E. A., Gerloff, D. L. and Earnshaw, W. C. (2004a). Borealin: a novel chromosomal passenger required for stability of the bipolar mitotic spindle. *J Cell Biol* **166**, 179-91.

Gassmann, R., Vagnarelli, P., Hudson, D. and Earnshaw, W. C. (2004b). Mitotic chromosome formation and the condensin paradox. *Exp Cell Res* **296**, 35-42.

Giacinti, C. and Giordano, A. (2006). RB and cell cycle progression. *Oncogene* **25**, 5220-7.

Giet, R. and Glover, D. M. (2001). Drosophila aurora B kinase is required for histone H3 phosphorylation and condensin recruitment during chromosome condensation and to organize the central spindle during cytokinesis. *J Cell Biol* **152**, 669-82.

Gil-Bernabe, A. M., Romero, F., Limon-Mortes, M. C. and Tortolero, M. (2006). Protein phosphatase 2A stabilizes human securin, whose phosphorylated forms are degraded via the SCF ubiquitin ligase. *Mol Cell Biol* **26**, 4017-27.

Gilbert, D. M. (2001). Making sense of eukaryotic DNA replication origins. *Science* **294**, 96-100.

Gimelli, G., Zuffardi, O., Giglio, S., Zeng, C. and He, D. (2000). CENP-G in neocentromeres and inactive centromeres. *Chromosoma* **109**, 328-33.

Glotzer, M. (2005). The molecular requirements for cytokinesis. *Science* **307**, 1735-9

Glover, D. M., Leibowitz, M. H., McLean, D. A. and Parry, H. (1995). Mutations in aurora prevent centrosome separation leading to the formation of monopolar spindles. *Cell* **81**, 95-105.

Glowczewski, L., Yang, P., Kalashnikova, T., Santisteban, M. S. and Smith, M. M. (2000). Histone-histone interactions and centromere function. *Mol Cell Biol* **20**, 5700-11.

Goldberg, I. G., Sawhney, H., Pluta, A. F., Warburton, P. E. and Earnshaw, W. C. (1996). Surprising deficiency of CENP-B binding sites in African green monkey alpha-satellite DNA: implications for CENP-B function at centromeres. *Mol Cell Biol* **16**, 5156-68.

Goshima, G., Kiyomitsu, T., Yoda, K. and Yanagida, M. (2003). Human centromere chromatin protein hMis12, essential for equal segregation, is independent of CENP-A loading pathway. *J Cell Biol* **160**, 25-39.

Goto, H., Yasui, Y., Nigg, E. A. and Inagaki, M. (2002). Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation. *Genes Cells* **7**, 11-7.

Grewal, S. I. and Jia, S. (2007). Heterochromatin revisited. *Nat Rev Genet* **8**, 35-46.

Grimes, B. R., Babcock, J., Rudd, M. K., Chadwick, B. and Willard, H. F. (2004). Assembly and characterization of heterochromatin and euchromatin on human artificial chromosomes. *Genome Biol* **5**, R89.

Grimes, B. R. and Monaco, Z. L. (2005). Artificial and engineered chromosomes: developments and prospects for gene therapy. *Chromosoma* **114**, 230-41.

Grimes, B. R., Rhoades, A. A. and Willard, H. F. (2002). Alpha-satellite DNA and vector composition influence rates of human artificial chromosome formation. *Mol Ther* **5**, 798-805.

Grimes, B. R., Schindelhauer, D., McGill, N. I., Ross, A., Ebersole, T. A. and Cooke, H. J. (2001). Stable gene expression from a mammalian artificial chromosome. *EMBO Rep* **2**, 910-4.

Gropp, M., Itsykson, P., Singer, O., Ben-Hur, T., Reinhartz, E., Galun, E. and Reubinoff, B. E. (2003). Stable genetic modification of human embryonic stem cells by lentiviral vectors. *Mol Ther* **7**, 281-7.

Groth, A. C., Olivares, E. C., Thyagarajan, B. and Calos, M. P. (2000). A phage integrase directs efficient site-specific integration in human cells. *Proc Natl Acad Sci U S A* **97**, 5995-6000.

Guiducci, C., Ascenzioni, F., Auriche, C., Piccolella, E., Guerrini, A. M. and Donini, P. (1999). Use of a human minichromosome as a cloning and expression vector for mammalian cells. *Hum Mol Genet* **8**, 1417-24.

Guse, A., Mishima, M. and Glotzer, M. (2005). Phosphorylation of ZEN-4/MKLP1 by aurora B regulates completion of cytokinesis. *Curr Biol* **15**, 778-86.

Guzman, L. M., Belin, D., Carson, M. J. and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**, 4121-30.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C. S., Pawliuk, R., Morillon, E. et al. (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**, 415-9.

Hadlaczky, G., Praznovszky, T., Cserpan, I., Kereso, J., Peterfy, M., Kelemen, I., Atalay, E., Szeles, A., Szelei, J., Tubak, V. et al. (1991). Centromere formation in mouse cells cotransformed with human DNA and a dominant marker gene. *Proc Natl Acad Sci U S A* **88**, 8106-10.

Hancock, R. (2000). A new look at the nuclear matrix. *Chromosoma* **109**, 219-25.

Harrington, J. J., Van Bokkelen, G., Mays, R. W., Gustashaw, K. and Willard, H. F. (1997). Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat Genet* **15**, 345-55.

Hauf, S., Cole, R. W., LaTerra, S., Zimmer, C., Schnapp, G., Walter, R., Heckel, A., van Meel, J., Rieder, C. L. and Peters, J. M. (2003). The small molecule Hesperadin reveals a role for Aurora B in correcting kinetochore-microtubule attachment and in maintaining the spindle assembly checkpoint. *J Cell Biol* **161**, 281-94.

Hayashi, T., Fujita, Y., Iwasaki, O., Adachi, Y., Takahashi, K. and Yanagida, M. (2004). Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell* **118**, 715-29.

He, D., Zeng, C., Woods, K., Zhong, L., Turner, D., Busch, R. K., Brinkley, B. R. and Busch, H. (1998). CENP-G: a new centromeric protein that is associated with the alpha-1 satellite DNA subfamily. *Chromosoma* **107**, 189-97.

Heinzel, S. S., Krysan, P. J., Tran, C. T. and Calos, M. P. (1991). Autonomous DNA replication in human cells is affected by the size and the source of the DNA. *Mol Cell Biol* **11**, 2263-72.

Heller, R., Brown, K. E., Burgtorf, C. and Brown, W. R. (1996). Mini-chromosomes derived from the human Y chromosome by telomere directed chromosome breakage. *Proc Natl Acad Sci U S A* **93**, 7125-30.

Henikoff, S., Ahmad, K. and Malik, H. S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* **293**, 1098-102.

Henikoff, S., Ahmad, K., Platero, J. S. and van Steensel, B. (2000). Heterochromatic deposition of centromeric histone H3-like proteins. *Proc Natl Acad Sci U S A* **97**, 716-21.

Henning, K. A., Novotny, E. A., Compton, S. T., Guan, X. Y., Liu, P. P. and Ashlock, M. A. (1999). Human artificial chromosomes generated by modification of a yeast artificial chromosome containing both human alpha satellite and single-copy DNA sequences. *Proc Natl Acad Sci U S A* **96**, 592-7.

Hibbitt, O. C. and Wade-Martins, R. (2006). Delivery of large genomic DNA inserts >100 kb using HSV-1 amplicons. *Curr Gene Ther* **6**, 325-36.

Hirano, T. (2000). Chromosome cohesion, condensation, and separation. *Annu Rev Biochem* **69**, 115-44.

Howman, E. V., Fowler, K. J., Newson, A. J., Redward, S., MacDonald, A. C., Kalitsis, P. and Choo, K. H. (2000). Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. *Proc Natl Acad Sci U S A* **97**, 1148-53.

Hsu, J. Y., Sun, Z. W., Li, X., Reuben, M., Tatchell, K., Bishop, D. K., Grushcow, J. M., Brame, C. J., Caldwell, J. A., Hunt, D. F. et al. (2000). Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell* **102**, 279-91.

Hudson, D. F., Fowler, K. J., Earle, E., Saffery, R., Kalitsis, P., Trowell, H., Hill, J., Wreford, N. G., de Kretser, D. M., Cancilla, M. R. et al. (1998). Centromere protein B null mice are mitotically and meiotically normal but have lower body and testis weights. *J Cell Biol* **141**, 309-19.

Huertas, D., Howe, S., McGuigan, A. and Huxley, C. (2000). Expression of the human CFTR gene from episomal oriP-EBNA1-YACs in mouse cells. *Hum Mol Genet* **9**, 617-29.

Hussein, D. and Taylor, S. S. (2002). Farnesylation of Cenp-F is required for G2/M progression and degradation after mitosis. *J Cell Sci* **115**, 3403-14.

Hyman, A. A. and Sorger, P. K. (1995). Structure and function of kinetochores in budding yeast. *Annu Rev Cell Dev Biol* **11**, 471-95.

Ikeno, M., Grimes, B., Okazaki, T., Nakano, M., Saitoh, K., Hoshino, H., McGill, N. I., Cooke, H. and Masumoto, H. (1998). Construction of YAC-based mammalian artificial chromosomes. *Nat Biotechnol* **16**, 431-9.

Ikeno, M., Inagaki, H., Nagata, K., Morita, M., Ichinose, H. and Okazaki, T. (2002). Generation of human artificial chromosomes expressing naturally controlled guanosine triphosphate cyclohydrolase I gene. *Genes Cells* **7**, 1021-32.

Ikeno, M., Masumoto, H. and Okazaki, T. (1994). Distribution of CENP-B boxes reflected in CREST centromere antigenic sites on long-range alpha-satellite DNA arrays of human chromosome 21. *Hum Mol Genet* **3**, 1245-57.

Ioannou, P. A., Amemiya, C. T., Garnes, J., Kroisel, P. M., Shizuya, H., Chen, C., Batzer, M. A. and de Jong, P. J. (1994). A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat Genet* **6**, 84-9.

Irvine, D. V., Amor, D. J., Perry, J., Sirvent, N., Pedeutour, F., Choo, K. H. and Saffery, R. (2004). Chromosome size and origin as determinants of the level of CENP-A incorporation into human centromeres. *Chromosome Res* **12**, 805-15.

Irvine, D. V., Shaw, M. L., Choo, K. H. and Saffery, R. (2005). Engineering chromosomes for delivery of therapeutic genes. *Trends Biotechnol* **23**, 575-83.

Itzhaki, J. E., Barnett, M. A., MacCarthy, A. B., Buckle, V. J., Brown, W. R. and Porter, A. C. (1992). Targeted breakage of a human chromosome mediated by cloned human telomeric DNA. *Nat Genet* **2**, 283-7.

Jablonski, S. A., Chan, G. K., Cooke, C. A., Earnshaw, W. C. and Yen, T. J. (1998). The hBUB1 and hBUBR1 kinases sequentially assemble onto kinetochores during prophase with hBUBR1 concentrating at the kinetochore plates in mitosis. *Chromosoma* **107**, 386-96.

Jenuwein, T. (2001). Re-SET-ting heterochromatin by histone methyltransferases. *Trends Cell Biol* **11**, 266-73.

Jiao, W., Lin, H. M., Timmons, J., Nagaich, A. K., Ng, S. W., Misteli, T. and Rane, S. G. (2005). E2F-dependent repression of topoisomerase II regulates heterochromatin formation and apoptosis in cells with melanoma-prone mutation. *Cancer Res* **65**, 4067-77.

Johnson, V. L., Scott, M. I., Holt, S. V., Hussein, D. and Taylor, S. S. (2004). Bub1 is required for kinetochore localization of BubR1, Cenp-E, Cenp-F and Mad2, and chromosome congression. *J Cell Sci* **117**, 1577-89.

Kaitna, S., Mendoza, M., Jantsch-Plunger, V. and Glotzer, M. (2000). Incenp and an aurora-like kinase form a complex essential for chromosome segregation and efficient completion of cytokinesis. *Curr Biol* **10**, 1172-81.

Kakeda, M., Hiratsuka, M., Nagata, K., Kuroiwa, Y., Kakitani, M., Katoh, M., Oshimura, M. and Tomizuka, K. (2005). Human artificial chromosome (HAC) vector provides long-term therapeutic transgene expression in normal human primary fibroblasts. *Gene Ther* **12**, 852-6.

Kalitsis, P., Fowler, K. J., Earle, E., Hill, J. and Choo, K. H. (1998). Targeted disruption of mouse centromere protein C gene leads to mitotic disarray and early embryo death. *Proc Natl Acad Sci U S A* **95**, 1136-41.

Kashima, T., Vinters, H. V. and Campagnoni, A. T. (1995). Unexpected expression of intermediate filament protein genes in human oligodendroglioma cell lines. *J Neuropathol Exp Neurol* **54**, 23-31.

Katoh, M., Ayabe, F., Norikane, S., Okada, T., Masumoto, H., Horike, S., Shirayoshi, Y. and Oshimura, M. (2004). Construction of a novel human artificial chromosome vector for gene delivery. *Biochem Biophys Res Commun* **321**, 280-90.

Kaufman, D. S., Hanson, E. T., Lewis, R. L., Auerbach, R. and Thomson, J. A. (2001). Hematopoietic colony-forming cells derived from human embryonic stem cells. *Proc Natl Acad Sci U S A* **98**, 10716-21.

Kellum, R. and Alberts, B. M. (1995). Heterochromatin protein 1 is required for correct chromosome segregation in Drosophila embryos. *J Cell Sci* **108 (Pt 4)**, 1419-31.

Kereso, J., Praznovszky, T., Cserpan, I., Fodor, K., Katona, R., Csonka, E., Fatyol, K., Hollo, G., Szeles, A., Ross, A. R. et al. (1996). De novo chromosome formations by large-scale amplification of the centromeric region of mouse chromosomes. *Chromosome Res* **4**, 226-39.

Kim, U. J., Birren, B. W., Slepak, T., Mancino, V., Boysen, C., Kang, H. L., Simon, M. I. and Shizuya, H. (1996). Construction and characterization of a human bacterial artificial chromosome library. *Genomics* **34**, 213-8.

King, R. W., Deshaies, R. J., Peters, J. M. and Kirschner, M. W. (1996). How proteolysis drives the cell cycle. *Science* **274**, 1652-9.

Klein, U. R., Nigg, E. A. and Gruneberg, U. (2006). Centromere targeting of the chromosomal passenger complex requires a ternary subcomplex of Borealin, Survivin, and the N-terminal domain of INCENP. *Mol Biol Cell* **17**, 2547-58.

Kline, S. L., Cheeseman, I. M., Hori, T., Fukagawa, T. and Desai, A. (2006). The human Mis12 complex is required for kinetochore assembly and proper chromosome segregation. *J Cell Biol* **173**, 9-17.

Knehr, M., Poppe, M., Schroeter, D., Eickelbaum, W., Finze, E. M., Kiesewetter, U. L., Enulescu, M., Arand, M. and Paweletz, N. (1996). Cellular expression of human centromere protein C demonstrates a cyclic behavior with highest abundance in the G1 phase. *Proc Natl Acad Sci U S A* **93**, 10234-9.

Koch, J. (2000). Neocentromeres and alpha satellite: a proposed structural code for functional human centromere DNA. *Hum Mol Genet* **9**, 149-54.

Kops, G. J., Kim, Y., Weaver, B. A., Mao, Y., McLeod, I., Yates, J. R., 3rd, Tagaya, M. and Cleveland, D. W. (2005). ZW10 links mitotic checkpoint signaling to the structural kinetochore. *J Cell Biol* **169**, 49-60.

Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* **184**, 868-71.

Kotzamanis, G., Cheung, W., Abdulrazzak, H., Perez-Luz, S., Howe, S., Cooke, H. and Huxley, C. (2005). Construction of human artificial chromosome vectors by recombineering. *Gene* **351**, 29-38.

Kotzamanis, G. and Huxley, C. (2004). Recombining overlapping BACs into a single larger BAC. *BMC Biotechnol* **4**, 1.

Kouprina, N., Ebersole, T., Koriabine, M., Pak, E., Rogozin, I. B., Katoh, M., Oshimura, M., Ogi, K., Peredelchuk, M., Solomon, G. et al. (2003). Cloning of human centromeres by transformation-associated recombination in yeast and generation of functional human artificial chromosomes. *Nucleic Acids Res* **31**, 922-34.

Kronenwett, U., Castro, J., Roblick, U. J., Fujioka, K., Ostring, C., Faridmoghaddam, F., Laytragoon-Lewin, N., Tribukait, B. and Auer, G. (2003). Expression of cyclins A, E and topoisomerase II alpha correlates with centrosome amplification and genomic instability and influences the reliability of cytometric S-phase determination. *BMC Cell Biol* **4**, 8.

Krysan, P. J., Smith, J. G. and Calos, M. P. (1993). Autonomous replication in human cells of multimers of specific human and bacterial DNA sequences. *Mol Cell Biol* **13**, 2688-96.

Kuroiwa, Y., Kasinathan, P., Choi, Y. J., Naeem, R., Tomizuka, K., Sullivan, E. J., Knott, J. G., Duteau, A., Goldsby, R. A., Osborne, B. A. et al. (2002). Cloned transchromosomic calves producing human immunoglobulin. *Nat Biotechnol* **20**, 889-94.

Kuroiwa, Y., Tomizuka, K., Shinohara, T., Kazuki, Y., Yoshida, H., Ohguma, A., Yamamoto, T., Tanaka, S., Oshimura, M. and Ishida, I. (2000). Manipulation of human minichromosomes to carry greater than megabase-sized chromosome inserts. *Nat Biotechnol* **18**, 1086-90.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-20.

Lam, A. L., Boivin, C. D., Bonney, C. F., Rudd, M. K. and Sullivan, B. A. (2006). Human centromeric chromatin is a dynamic chromosomal domain that can spread over noncentromeric DNA. *Proc Natl Acad Sci U S A* **103**, 4186-91.

Larin, Z. and Mejía, J. E. (2002). Advances in human artificial chromosome technology. *Trends Genet* **18**, 313-9.

Lee, E. C., Yu, D., Martinez de Velasco, J., Tessarollo, L., Swing, D. A., Court, D. L., Jenkins, N. A. and Copeland, N. G. (2001). A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* **73**, 56-65.

Li, Z., Dullmann, J., Schiedlmeier, B., Schmidt, M., von Kalle, C., Meyer, J., Forster, M., Stocking, C., Wahlers, A., Frank, O. et al. (2002). Murine leukemia induced by retroviral gene marking. *Science* **296**, 497.

Liao, H., Li, G. and Yen, T. J. (1994). Mitotic regulation of microtubule cross-linking activity of CENP-E kinetochore protein. *Science* **265**, 394-8.

Liao, H., Winkfein, R. J., Mack, G., Rattner, J. B. and Yen, T. J. (1995). CENP-F is a protein of the nuclear matrix that assembles onto kinetochores at late G2 and is rapidly degraded after mitosis. *J Cell Biol* **130**, 507-18.

Lindenbaum, M., Perkins, E., Csonka, E., Fleming, E., Garcia, L., Greene, A., Gung, L., Hadlaczky, G., Lee, E., Leung, J. et al. (2004). A mammalian artificial chromosome engineering system (ACE System) applicable to biopharmaceutical protein production, transgenesis and gene-based cell therapy. *Nucleic Acids Res* **32**, e172.

Lipps, H. J., Jenke, A. C., Nehlsen, K., Scinteie, M. F., Stehle, I. M. and Bode, J. (2003). Chromosome-based vectors for gene therapy. *Gene* **304**, 23-33.

Liu, S. T., Hittle, J. C., Jablonski, S. A., Campbell, M. S., Yoda, K. and Yen, T. J. (2003). Human CENP-I specifies localization of CENP-F, MAD1 and MAD2 to kinetochores and is essential for mitosis. *Nat Cell Biol* **5**, 341-5.

Liu, S. T., Rattner, J. B., Jablonski, S. A. and Yen, T. J. (2006). Mapping the assembly pathways that specify formation of the trilaminar kinetochore plates in human cells. *J Cell Biol* **175**, 41-53.

Lo, A. W., Craig, J. M., Saffery, R., Kalitsis, P., Irvine, D. V., Earle, E., Magliano, D. J. and Choo, K. H. (2001a). A 330 kb CENP-A binding domain and altered replication timing at a human neocentromere. *Embo J* **20**, 2087-96.

Lo, A. W., Liao, G. C., Rocchi, M. and Choo, K. H. (1999). Extreme reduction of chromosome-specific alpha-satellite array is unusually common in human chromosome 21. *Genome Res* **9**, 895-908.

Lo, A. W., Magliano, D. J., Sibson, M. C., Kalitsis, P., Craig, J. M. and Choo, K. H. (2001b). A novel chromatin immunoprecipitation and array (CIA) analysis identifies a 460-kb CENP-A-binding neocentromere DNA. *Genome Res* **11**, 448-57.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-60.

Lukas, J., Lukas, C. and Bartek, J. (2004). Mammalian cell cycle checkpoints: signalling pathways and their organization in space and time. *DNA Repair (Amst)* **3**, 997-1007.

Mackay, A. M., Ainsztein, A. M., Eckley, D. M. and Earnshaw, W. C. (1998). A dominant mutant of inner centromere protein (INCENP), a chromosomal protein, disrupts prometaphase congression and cytokinesis. *J Cell Biol* **140**, 991-1002.

Mackay, A. M., Eckley, D. M., Chue, C. and Earnshaw, W. C. (1993). Molecular analysis of the INCENPs (inner centromere proteins): separate domains are required for association with microtubules during interphase and with the central spindle during anaphase. *J Cell Biol* **123**, 373-85.

Maggert, K. A. and Karpen, G. H. (2001). The activation of a neocentromere in Drosophila requires proximity to an endogenous centromere. *Genetics* **158**, 1615-28.

Maney, T., Hunter, A. W., Wagenbach, M. and Wordeman, L. (1998). Mitotic centromere-associated kinesin is important for anaphase chromosome segregation. *J Cell Biol* **142**, 787-801.

Mao, Y., Abrieu, A. and Cleveland, D. W. (2003). Activating and silencing the mitotic checkpoint through CENP-E-dependent activation/inactivation of BubR1. *Cell* **114**, 87-98.

Mao, Y., Desai, A. and Cleveland, D. W. (2005). Microtubule capture by CENP-E silences BubR1-dependent mitotic checkpoint signaling. *J Cell Biol* **170**, 873-80.

Marschall, P., Malik, N. and Larin, Z. (1999). Transfer of YACs up to 2.3 Mb intact into human cells with polyethylenimine. *Gene Ther* **6**, 1634-7.

Marshall, E. (1999). Gene therapy death prompts review of adenovirus vector. *Science* **286**, 2244-5.

Masumoto, H., Masukata, H., Muro, Y., Nozaki, N. and Okazaki, T. (1989). A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J Cell Biol* **109**, 1963-73.

Masumoto, H., Nakano, M. and Ohzeki, J. (2004). The role of CENP-B and alpha-satellite DNA: de novo assembly and epigenetic maintenance of human centromeres. *Chromosome Res* **12**, 543-56.

McEwen, B. F., Chan, G. K., Zubrowski, B., Savoian, M. S., Sauer, M. T. and Yen, T. J. (2001). CENP-E is essential for reliable bioriented spindle attachment, but chromosome alignment can be achieved via redundant mechanisms in mammalian cells. *Mol Biol Cell* **12**, 2776-89.

Measday, V., Hailey, D. W., Pot, I., Givan, S. A., Hyland, K. M., Cagney, G., Fields, S., Davis, T. N. and Hieter, P. (2002). Ctf3p, the Mis6 budding yeast homolog, interacts with Mcm22p and Mcm16p at the yeast outer kinetochore. *Genes Dev* **16**, 101-13.

Mejía, J. E., Alazami, A., Willmott, A., Marschall, P., Levy, E., Earnshaw, W. C. and Larin, Z. (2002). Efficiency of de novo centromere formation in human artificial chromosomes. *Genomics* **79**, 297-304.

Mejía, J. E. and Larin, Z. (2000). The assembly of large BACs by in vivo recombination. *Genomics* **70**, 165-70.

Mejía, J. E., Willmott, A., Levy, E., Earnshaw, W. C. and Larin, Z. (2001). Functional complementation of a genetic deficiency with human artificial chromosomes. *Am J Hum Genet* **69**, 315-26.

Melcher, M., Schmid, M., Aagaard, L., Selenko, P., Laible, G. and Jenuwein, T. (2000). Structure-function analysis of SUV39H1 reveals a dominant role in heterochromatin organization, chromosome segregation, and mitotic progression. *Mol Cell Biol* **20**, 3728-41.

Meluh, P. B. and Koshland, D. (1995). Evidence that the MIF2 gene of Saccharomyces cerevisiae encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Mol Biol Cell* **6**, 793-807.

Meluh, P. B., Yang, P., Glowczewski, L., Koshland, D. and Smith, M. M. (1998). Cse4p is a component of the core centromere of Saccharomyces cerevisiae. *Cell* **94**, 607-13.

Miccoli, L., Biard, D. S., Creminon, C. and Angulo, J. F. (2002). Human kin17 protein directly interacts with the simian virus 40 large T antigen and inhibits DNA replication. *Cancer Res* **62**, 5425-35.

Mikami, Y., Hori, T., Kimura, H. and Fukagawa, T. (2005). The functional region of CENP-H interacts with the Nuf2 complex that localizes to centromere during mitosis. *Mol Cell Biol* **25**, 1958-70.

Miller, C. A., Ingmer, H. and Cohen, S. N. (1995). Boundaries of the pSC101 minimal replicon are conditional. *J Bacteriol* **177**, 4865-71.

Mills, W., Critcher, R., Lee, C. and Farr, C. J. (1999). Generation of an approximately 2.4 Mb human X centromere-based minichromosome by targeted telomere-associated chromosome fragmentation in DT40. *Hum Mol Genet* **8**, 751-61.

Millward, T. A., Zolnierowicz, S. and Hemmings, B. A. (1999). Regulation of protein kinase cascades by protein phosphatase 2A. *Trends Biochem Sci* **24**, 186-91.

Moore, L. L. and Roth, M. B. (2001). HCP-4, a CENP-C-like protein in Caenorhabditis elegans, is required for resolution of sister centromeres. *J Cell Biol* **153**, 1199-208.

Moralli, D., Simpson, K. M., Wade-Martins, R. and Monaco, Z. L. (2006). A novel human artificial chromosome gene expression system using herpes simplex virus type 1 vectors. *EMBO Rep* **7**, 911-8.

Moralli, D., Vagnarelli, P., Bensi, M., De Carli, L. and Raimondi, E. (2001). Insertion of a loxP site in a size-reduced human accessory chromosome. *Cytogenet Cell Genet* **94**, 113-20.

Moyzis, R. K., Buckingham, J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., Meyne, J., Ratliff, R. L. and Wu, J. R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc Natl Acad Sci U S A* **85**, 6622-6.

Murphy, T. D. and Karpen, G. H. (1995). Localization of centromere function in a Drosophila minichromosome. *Cell* **82**, 599-609.

Murray, A. W. and Szostak, J. W. (1983). Construction of artificial chromosomes in yeast. *Nature* **305**, 189-93.

Muyrers, J. P., Zhang, Y. and Stewart, A. F. (2001). Techniques: Recombinogenic engineering--new options for cloning and manipulating DNA. *Trends Biochem Sci* **26**, 325-31.

Nagaki, K., Cheng, Z., Ouyang, S., Talbert, P. B., Kim, M., Jones, K. M., Henikoff, S., Buell, C. R. and Jiang, J. (2004). Sequencing of a rice centromere uncovers active genes. *Nat Genet* **36**, 138-45.

Nakano, M., Okamoto, Y., Ohzeki, J. and Masumoto, H. (2003). Epigenetic assembly of centromeric chromatin at ectopic alpha-satellite sites on human chromosomes. *J Cell Sci* **116**, 4021-34.

Nakaseko, Y., Kinoshita, N. and Yanagida, M. (1987). A novel sequence common to the centromere regions of Schizosaccharomyces pombe chromosomes. *Nucleic Acids Res* **15**, 4705-15.

Nakashima, H., Nakano, M., Ohnishi, R., Hiraoka, Y., Kaneda, Y., Sugino, A. and Masumoto, H. (2005). Assembly of additional heterochromatin distinct from centromere-kinetochore chromatin is required for de novo formation of human artificial chromosome. *J Cell Sci* **118**, 5885-98.

Nasmyth, K. (2001). Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annu Rev Genet* **35**, 673-745.

Neidle, S. and Parkinson, G. N. (2003). The structure of telomeric DNA. *Curr Opin Struct Biol* **13**, 275-83.

Nishihashi, A., Haraguchi, T., Hiraoka, Y., Ikemura, T., Regnier, V., Dodson, H., Earnshaw, W. C. and Fukagawa, T. (2002). CENP-I is essential for centromere function in vertebrate cells. *Dev Cell* **2**, 463-76.

Nishitani, H. and Lygerou, Z. (2002). Control of DNA replication licensing in a cell cycle. *Genes Cells* **7**, 523-34.

Norio, P. (2006). DNA replication: the unbearable lightness of origins. *EMBO Rep* **7**, 779-81.

Oberle, V., de Jong, G., Drayer, J. I. and Hoekstra, D. (2004). Efficient transfer of chromosome-based DNA constructs into mammalian cells. *Biochim Biophys Acta* **1676**, 223-30.

Oegema, K., Desai, A., Rybina, S., Kirkham, M. and Hyman, A. A. (2001). Functional analysis of kinetochore assembly in Caenorhabditis elegans. *J Cell Biol* **153**, 1209-26.

Ohzeki, J., Nakano, M., Okada, T. and Masumoto, H. (2002). CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. *J Cell Biol* **159**, 765-75.

Okamoto, Y., Nakano, M., Ohzeki, J., Larionov, V. and Masumoto, H. (2007). A minimal CENP-A core is required for nucleation and maintenance of a functional human centromere. *Embo J* **26**, 1279-91.

Olschowka, J. A., Bowers, W. J., Hurley, S. D., Mastrangelo, M. A. and Federoff, H. J. (2003). Helper-free HSV-1 amplicons elicit a markedly less robust innate immune response in the CNS. *Mol Ther* **7**, 218-27.

Orthaus, S., Ohndorf, S. and Diekmann, S. (2006). RNAi knockdown of human kinetochore protein CENP-H. *Biochem Biophys Res Commun* **348**, 36-46.

Ota, T., Suto, S., Katayama, H., Han, Z. B., Suzuki, F., Maeda, M., Tanino, M., Terada, Y. and Tatsuka, M. (2002). Increased mitotic phosphorylation of histone H3 attributable to AIM-1/Aurora-B overexpression contributes to chromosome number instability. *Cancer Res* **62**, 5168-77.

Palmer, D. K., O'Day, K., Trong, H. L., Charbonneau, H. and Margolis, R. L. (1991). Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. *Proc Natl Acad Sci U S A* **88**, 3734-8.

Palmer, D. K., O'Day, K., Wener, M. H., Andrews, B. S. and Margolis, R. L. (1987). A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. *J Cell Biol* **104**, 805-15.

Pavan, W. J., Hieter, P. and Reeves, R. H. (1990). Generation of deletion derivatives by targeted transformation of human-derived yeast artificial chromosomes. *Proc Natl Acad Sci U S A* **87**, 1300-4.

Pfarr, C. M., Coue, M., Grissom, P. M., Hays, T. S., Porter, M. E. and McIntosh, J. R. (1990). Cytoplasmic dynein is localized to kinetochores during mitosis. *Nature* **345**, 263-5.

Pidoux, A. L. and Allshire, R. C. (2000). Centromeres: getting a grip of chromosomes. *Curr Opin Cell Biol* **12**, 308-19.

Platero, J. S., Ahmad, K. and Henikoff, S. (1999). A distal heterochromatic block displays centromeric activity when detached from a natural centromere. *Mol Cell* **4**, 995-1004.

Politi, V., Perini, G., Trazzi, S., Pliss, A., Raska, I., Earnshaw, W. C. and Della Valle, G. (2002). CENP-C binds the alpha-satellite DNA in vivo at specific centromere domains. *J Cell Sci* **115**, 2317-27.

Porter, A. C. and Farr, C. J. (2004). Topoisomerase II: untangling its contribution at the centromere. *Chromosome Res* **12**, 569-83.

Rattner, J. B., Hendzel, M. J., Furbee, C. S., Muller, M. T. and Bazett-Jones, D. P. (1996). Topoisomerase II alpha is associated with the mammalian centromere in a cell cycle- and species-specific manner and is required for proper centromere/kinetochore structure. *J Cell Biol* **134**, 1097-107.

Rattner, J. B., Rao, A., Fritzler, M. J., Valencia, D. W. and Yen, T. J. (1993). CENP-F is a .ca 400 kDa kinetochore protein that exhibits a cell-cycle dependent localization. *Cell Motil Cytoskeleton* **26**, 214-26.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D. et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593-9.

Regnier, V., Vagnarelli, P., Fukagawa, T., Zerjal, T., Burns, E., Trouche, D., Earnshaw, W. and Brown, W. (2005). CENP-A is required for accurate chromosome segregation and sustained kinetochore association of BubR1. *Mol Cell Biol* **25**, 3967-81.

Reubinoff, B. E., Itsykson, P., Turetsky, T., Pera, M. F., Reinhartz, E., Itzik, A. and Ben-Hur, T. (2001). Neural progenitors from human embryonic stem cells. *Nat Biotechnol* **19**, 1134-40.

Rideout, W. M., 3rd, Hochedlinger, K., Kyba, M., Daley, G. Q. and Jaenisch, R. (2002). Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy. *Cell* **109**, 17-27.

Robinson, P. J. and Rhodes, D. (2006). Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr Opin Struct Biol* **16**, 336-43.

Rudd, M. K., Mays, R. W., Schwartz, S. and Willard, H. F. (2003). Human artificial chromosomes with alpha satellite-based de novo centromeres show increased frequency of nondisjunction and anaphase lag. *Mol Cell Biol* **23**, 7689-97.

Rudd, M. K. and Willard, H. F. (2004). Analysis of the centromeric regions of the human genome assembly. *Trends Genet* **20**, 529-33.

Saeki, Y., Fraefel, C., Ichikawa, T., Breakefield, X. O. and Chiocca, E. A. (2001). Improved helper virus-free packaging system for HSV amplicon vectors using an ICP27-deleted, oversized HSV-1 DNA in a bacterial artificial chromosome. *Mol Ther* **3**, 591-601.

Saeki, Y., Ichikawa, T., Saeki, A., Chiocca, E. A., Tobler, K., Ackermann, M., Breakefield, X. O. and Fraefel, C. (1998). Herpes simplex virus type 1 DNA amplified as bacterial artificial chromosome in Escherichia coli: rescue of replication-competent virus progeny and packaging of amplicon vectors. *Hum Gene Ther* **9**, 2787-94.

Saffery, R. and Choo, K. H. (2002). Strategies for engineering human chromosomes with therapeutic potential. *J Gene Med* **4**, 5-13.

Saffery, R., Irvine, D. V., Griffiths, B., Kalitsis, P., Wordeman, L. and Choo, K. H. (2000). Human centromeres and neocentromeres show identical distribution patterns of >20 functionally important kinetochore-associated proteins. *Hum Mol Genet* **9**, 175-85.

Saffery, R., Sumer, H., Hassan, S., Wong, L. H., Craig, J. M., Todokoro, K., Anderson, M., Stafford, A. and Choo, K. H. (2003). Transcription within a functional human centromere. *Mol Cell* **12**, 509-16.

Saffery, R., Wong, L. H., Irvine, D. V., Bateman, M. A., Griffiths, B., Cutts, S. M., Cancilla, M. R., Cendron, A. C., Stafford, A. J. and Choo, K. H. (2001). Construction of neocentromere-based human minichromosomes by telomere-associated chromosomal truncation. *Proc Natl Acad Sci U S A* **98**, 5705-10.

Saitoh, H., Tomkiel, J., Cooke, C. A., Ratrie, H., 3rd, Maurer, M., Rothfield, N. F. and Earnshaw, W. C. (1992). CENP-C, an autoantigen in scleroderma, is a component of the human inner kinetochore plate. *Cell* **70**, 115-25.

Saxena, A., Saffery, R., Wong, L. H., Kalitsis, P. and Choo, K. H. (2002a). Centromere proteins Cenpa, Cenpb, and Bub3 interact with poly(ADP-ribose) polymerase-1 protein and are poly(ADP-ribosyl)ated. *J Biol Chem* **277**, 26921-6.

Saxena, A., Wong, L. H., Kalitsis, P., Earle, E., Shaffer, L. G. and Choo, K. H. (2002b). Poly(ADP-ribose) polymerase 2 localizes to mammalian active centromeres and interacts with PARP-1, Cenpa, Cenpb and Bub3, but not Cenpc. *Hum Mol Genet* **11**, 2319-29.

Scaerou, F., Aguilera, I., Saunders, R., Kane, N., Blottiere, L. and Karess, R. (1999). The rough deal protein is a new kinetochore component required for accurate chromosome segregation in Drosophila. *J Cell Sci* **112 (Pt 21)**, 3757-68.

Schindelhauer, D. and Cooke, H. J. (1997). Efficient combination of large DNA in vitro: in gel site specific recombination (IGSSR) of PAC fragments containing alpha satellite DNA and the human HPRT gene locus. *Nucleic Acids Res* **25**, 2241-3.

Schleif, R. (1992). DNA looping. *Annu Rev Biochem* **61**, 199-223.

Schueler, M. G., Higgins, A. W., Rudd, M. K., Gustashaw, K. and Willard, H. F. (2001). Genomic and genetic definition of a functional human centromere. *Science* **294**, 109-15.

Sena-Esteves, M., Saeki, Y., Fraefel, C. and Breakefield, X. O. (2000). HSV-1 amplicon vectors--simplicity and versatility. *Mol Ther* **2**, 9-15.

Shapiro, P. S., Vaisberg, E., Hunt, A. J., Tolwinski, N. S., Whalen, A. M., McIntosh, J. R. and Ahn, N. G. (1998). Activation of the MKK/ERK pathway during somatic cell mitosis: direct interactions of active ERK with kinetochores and regulation of the mitotic 3F3/2 phosphoantigen. *J Cell Biol* **142**, 1533-45.

Shay, J. W. (1999). At the end of the millennium, a view of the end. *Nat Genet* **23**, 382-3.

Shelby, R. D., Monier, K. and Sullivan, K. F. (2000). Chromatin assembly at kinetochores is uncoupled from DNA replication. *J Cell Biol* **151**, 1113-8.

Shen, M. H., Ross, A., Yang, J., de las Heras, J. I. and Cooke, H. (2001a). Neo-centromere formation on a 2.6 Mb mini-chromosome in DT40 cells. *Chromosoma* **110**, 421-9.

Shen, M. H., Yang, J. W., Yang, J., Pendon, C. and Brown, W. R. (2001b). The accuracy of segregation of human mini-chromosomes varies in different vertebrate cell lines, correlates with the extent of centromere formation and provides evidence for a trans-acting centromere maintenance activity. *Chromosoma* **109**, 524-35.

Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y. and Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. *Proc Natl Acad Sci U S A* **89**, 8794-7.

Skoufias, D. A., Mollinari, C., Lacroix, F. B. and Margolis, R. L. (2000). Human survivin is a kinetochore-associated passenger protein. *J Cell Biol* **151**, 1575-82.

Spence, J. M., Fournier, R. E., Oshimura, M., Regnier, V. and Farr, C. J. (2005). Topoisomerase II cleavage activity within the human D11Z1 and DXZ1 alpha-satellite arrays. *Chromosome Res* **13**, 637-48.

Stack, S. M. and Anderson, L. K. (2001). A model for chromosome structure during the mitotic and meiotic cell cycles. *Chromosome Res* **9**, 175-98.

Starr, D. A., Saffery, R., Li, Z., Simpson, A. E., Choo, K. H., Yen, T. J. and Goldberg, M. L. (2000). HZwint-1, a novel human kinetochore component that interacts with HZW10. *J Cell Sci* **113 (Pt 11)**, 1939-50.

Starr, D. A., Williams, B. C., Hays, T. S. and Goldberg, M. L. (1998). ZW10 helps recruit dynactin and dynein to the kinetochore. *J Cell Biol* **142**, 763-74.

Steiner, N. C. and Clarke, L. (1994). A novel epigenetic effect can alter centromere function in fission yeast. *Cell* **79**, 865-74.

Steuer, E. R., Wordeman, L., Schroer, T. A. and Sheetz, M. P. (1990). Localization of cytoplasmic dynein to mitotic spindles and kinetochores. *Nature* **345**, 266-8.

Stinchcomb, D. T., Struhl, K. and Davis, R. W. (1979). Isolation and characterisation of a yeast chromosomal replicator. *Nature* **282**, 39-43.

Stoler, S., Keith, K. C., Curnick, K. E. and Fitzgerald-Hayes, M. (1995). A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. *Genes Dev* **9**, 573-86.

Sugata, N., Li, S., Earnshaw, W. C., Yen, T. J., Yoda, K., Masumoto, H., Munekata, E., Warburton, P. E. and Todokoro, K. (2000). Human CENP-H multimers colocalize with CENP-A and CENP-C at active centromere--kinetochore complexes. *Hum Mol Genet* **9**, 2919-26.

Sugata, N., Munekata, E. and Todokoro, K. (1999). Characterization of a novel kinetochore protein, CENP-H. *J Biol Chem* **274**, 27343-6.

Sugimoto, K., Yata, H., Muro, Y. and Himeno, M. (1994). Human centromere protein C (CENP-C) is a DNA-binding protein which possesses a novel DNA-binding motif. *J Biochem (Tokyo)* **116**, 877-81.

Sugiyama, K., Sugiura, K., Hara, T., Sugimoto, K., Shima, H., Honda, K., Furukawa, K., Yamashita, S. and Urano, T. (2002). Aurora-B associated protein phosphatases as negative regulators of kinase activation. *Oncogene* **21**, 3103-11.

Sullivan, B. and Karpen, G. (2001). Centromere identity in Drosophila is not determined in vivo by replication timing. *J Cell Biol* **154**, 683-90.

Sullivan, B. A., Blower, M. D. and Karpen, G. H. (2001). Determining centromere identity: cyclical stories and forking paths. *Nat Rev Genet* **2**, 584-96.

Sullivan, B. A. and Karpen, G. H. (2004). Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat Struct Mol Biol*.

Sullivan, B. A. and Schwartz, S. (1995). Identification of centromeric antigens in dicentric Robertsonian translocations: CENP-C and CENP-E are necessary components of functional centromeres. *Hum Mol Genet* **4**, 2189-97.

Sullivan, B. A. and Willard, H. F. (1998). Stable dicentric X chromosomes with two functional centromeres. *Nat Genet* **20**, 227-8.

Sullivan, K. F., Hechenberger, M. and Masri, K. (1994). Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *J Cell Biol* **127**, 581-92.

Sumner, A. T. (2003). Chromosomes : organization and function. Malden, Mass. ; Oxford: Blackwell.

Sun, X., Wahlstrom, J. and Karpen, G. (1997). Molecular structure of a functional Drosophila centromere. *Cell* **91**, 1007-19.

Surosky, R. T. and Tye, B. K. (1985). Construction of telocentric chromosomes in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **82**, 2106-10.

Suzuki, N., Nakano, M., Nozaki, N., Egashira, S., Okazaki, T. and Masumoto, H. (2004). CENP-B interacts with CENP-C domains containing Mif2 regions responsible for centromere localization. *J Biol Chem* **279**, 5934-46.

Szostak, J. W. and Blackburn, E. H. (1982). Cloning yeast telomeres on linear plasmid vectors. *Cell* **29**, 245-55.

Takahashi, K., Chen, E. S. and Yanagida, M. (2000). Requirement of Mis6 centromere connector for localizing a CENP-A-like protein in fission yeast. *Science* **288**, 2215-9.

Tanaka, Y., Nureki, O., Kurumizaka, H., Fukai, S., Kawaguchi, S., Ikuta, M., Iwahara, J., Okazaki, T. and Yokoyama, S. (2001). Crystal structure of the CENP-B protein-DNA complex: the DNA-binding domains of CENP-B induce kinks in the CENP-B box DNA. *Embo J* **20**, 6612-8.

Tanaka, Y., Tachiwana, H., Yoda, K., Masumoto, H., Okazaki, T., Kurumizaka, H. and Yokoyama, S. (2005). Human centromere protein B induces translational positioning of nucleosomes on alpha-satellite sequences. *J Biol Chem* **280**, 41609-18.

Tanenbaum, M. E., Galjart, N., van Vugt, M. A. and Medema, R. H. (2006). CLIP-170 facilitates the formation of kinetochore-microtubule attachments. *Embo J* **25**, 45-57.

Terada, Y. (2006). Aurora-B/AIM-1 regulates the dynamic behavior of HP1alpha at the G2-M transition. *Mol Biol Cell* **17**, 3232-41.

Thomas, C. E., Ehrhardt, A. and Kay, M. A. (2003). Progress and problems with the use of viral vectors for gene therapy. *Nat Rev Genet* **4**, 346-58.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S. and Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145-7.

Tomizuka, K., Shinohara, T., Yoshida, H., Uejima, H., Ohguma, A., Tanaka, S., Sato, K., Oshimura, M. and Ishida, I. (2000). Double trans-chromosomic mice: maintenance of two individual human chromosome fragments containing Ig heavy and kappa loci and expression of fully human antibodies. *Proc Natl Acad Sci U S A* **97**, 722-7.

Tomizuka, K., Yoshida, H., Uejima, H., Kugoh, H., Sato, K., Ohguma, A., Hayasaka, M., Hanaoka, K., Oshimura, M. and Ishida, I. (1997). Functional expression and germline transmission of a human chromosome fragment in chimaeric mice. *Nat Genet* **16**, 133-43.

Tomkiel, J., Cooke, C. A., Saitoh, H., Bernat, R. L. and Earnshaw, W. C. (1994). CENP-C is required for maintaining proper kinetochore size and for a timely transition to anaphase. *J Cell Biol* **125**, 531-45.

Tsuduki, T., Nakano, M., Yasuoka, N., Yamazaki, S., Okada, T., Okamoto, Y. and Masumoto, H. (2006). An artificially constructed de novo human chromosome behaves almost identically to its natural counterpart during metaphase and anaphase in living cells. *Mol Cell Biol* **26**, 7682-95.

Tyler-Smith, C. and Brown, W. R. (1987). Structure of the major block of alphoid satellite DNA on the human Y chromosome. *J Mol Biol* **195**, 457-70.

Tyler-Smith, C., Gimelli, G., Giglio, S., Floridia, G., Pandya, A., Terzoli, G., Warburton, P. E., Earnshaw, W. C. and Zuffardi, O. (1999). Transmission of a fully functional human neocentromere through three generations. *Am J Hum Genet* **64**, 1440-4.

Tyler-Smith, C., Oakey, R. J., Larin, Z., Fisher, R. B., Crocker, M., Affara, N. A., Ferguson-Smith, M. A., Muenke, M., Zuffardi, O. and Jobling, M. A. (1993). Localization of DNA sequences required for human centromere function through an analysis of rearranged Y chromosomes. *Nat Genet* **5**, 368-75.

Uren, A. G., Wong, L., Pakusch, M., Fowler, K. J., Burrows, F. J., Vaux, D. L. and Choo, K. H. (2000). Survivin and the inner centromere protein INCENP show similar cell-cycle localization and gene knockout phenotype. *Curr Biol* **10**, 1319-28.

Vader, G., Kauw, J. J., Medema, R. H. and Lens, S. M. (2005). Survivin mediates targeting of the chromosomal passenger complex to the centromere and midbody. *EMBO Rep*.

Vagnarelli, P. and Earnshaw, W. C. (2004). Chromosomal passengers: the four-dimensional regulation of mitotic events. *Chromosoma*.

Van Hooser, A. A., Ouspenski, II, Gregson, H. C., Starr, D. A., Yen, T. J., Goldberg, M. L., Yokomori, K., Earnshaw, W. C., Sullivan, K. F. and Brinkley, B. R. (2001). Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. *J Cell Sci* **114**, 3529-42.

Vanderbyl, S., MacDonald, G. N., Sidhu, S., Gung, L., Telenius, A., Perez, C. and Perkins, E. (2004). Transfer and stable transgene expression of a mammalian artificial chromosome into bone marrow-derived human mesenchymal stem cells. *Stem Cells* **22**, 324-33.

Vanikar, A. V., Mishra, V. V., Firoz, A., Shah, V. R., Dave, S. D., Patel, R. D., Kanodia, K. V., Patel, J. V., Patel, C. N. and Trivedi, H. L. (2007). Successful generation of donor specific hematopoietic stem cell lines from co-cultured bone marrow with human embryonic stem cell line: a new methodology. *Transplant Proc* **39**, 658-61.

Ventura, M., Archidiacono, N. and Rocchi, M. (2001). Centromere emergence in evolution. *Genome Res* **11**, 595-9.

Verma, I. M. and Weitzman, M. D. (2004). Gene Therapy: Twenty-First Century Medicine. *Annu Rev Biochem*.

Voet, T., Schoenmakers, E., Carpentier, S., Labaere, C. and Marynen, P. (2003). Controlled transgene dosage and PAC-mediated transgenesis in mice using a chromosomal vector. *Genomics* **82**, 596-605.

Voet, T., Vermeesch, J., Carens, A., Durr, J., Labaere, C., Duhamel, H., David, G. and Marynen, P. (2001). Efficient male and female germline transmission of a human chromosomal vector in mice. *Genome Res* **11**, 124-36.

Vollrath, D., Davis, R. W., Connelly, C. and Hieter, P. (1988). Physical mapping of large DNA by chromosome fragmentation. *Proc Natl Acad Sci U S A* **85**, 6027-31.

Voullaire, L. E., Slater, H. R., Petrovic, V. and Choo, K. H. (1993). A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? *Am J Hum Genet* **52**, 1153-63.

Wade-Martins, R., Frampton, J. and James, M. R. (1999). Long-term stability of large insert genomic DNA episomal shuttle vectors in human cells. *Nucleic Acids Res* **27**, 1674-82.

Wade-Martins, R., Saeki, Y. and Chiocca, E. A. (2003). Infectious delivery of a 135-kb LDLR genomic locus leads to regulated complementation of low-density lipoprotein receptor deficiency in human cells. *Mol Ther* **7**, 604-12.

Wade-Martins, R., Smith, E. R., Tyminski, E., Chiocca, E. A. and Saeki, Y. (2001). An infectious transfer and expression system for genomic DNA loci in human and mouse cells. *Nat Biotechnol* **19**, 1067-70.

Wade-Martins, R., White, R. E., Kimura, H., Cook, P. R. and James, M. R. (2000). Stable correction of a genetic deficiency in human cells by an episome carrying a 115 kb genomic transgene. *Nat Biotechnol* **18**, 1311-4.

Wang, B., Lazaris, A., Lindenbaum, M., Stewart, S., Co, D., Perez, C., Drayer, J. and Karatzas, C. N. (2001). Expression of a reporter gene after microinjection of mammalian artificial chromosomes into pronuclei of bovine zygotes. *Mol Reprod Dev* **60**, 433-8.

Warburton, P. E. (2001). Epigenetic analysis of kinetochore assembly on variant human centromeres. *Trends Genet* **17**, 243-7.

Warburton, P. E., Cooke, C. A., Bourassa, S., Vafa, O., Sullivan, B. A., Stetten, G., Gimelli, G., Warburton, D., Tyler-Smith, C., Sullivan, K. F. et al. (1997). Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. *Curr Biol* **7**, 901-4.

Warburton, P. E., Waye, J. S. and Willard, H. F. (1993). Nonrandom localization of recombination events in human alpha satellite repeat unit variants: implications for higher-order structural characteristics within centromeric heterochromatin. *Mol Cell Biol* **13**, 6520-9.

Waye, J. S. and Willard, H. F. (1986). Structure, organization, and sequence of alpha satellite DNA from human chromosome 17: evidence for evolution by unequal crossing-over and an ancestral pentamer repeat shared with the human X chromosome. *Mol Cell Biol* **6**, 3156-65.

Wei, Y., Yu, L., Bowen, J., Gorovsky, M. A. and Allis, C. D. (1999). Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. *Cell* **97**, 99-109.

Wevrick, R., Earnshaw, W. C., Howard-Peebles, P. N. and Willard, H. F. (1990). Partial deletion of alpha satellite DNA associated with reduced amounts of the centromere protein CENP-B in a mitotically stable human chromosome rearrangement. *Mol Cell Biol* **10**, 6374-80.

Wheatley, S. P., Henzing, A. J., Dodson, H., Khaled, W. and Earnshaw, W. C. (2004). Aurora-B phosphorylation in vitro identifies a residue of survivin that is essential for its localization and binding to inner centromere protein (INCENP) in vivo. *J Biol Chem* **279**, 5655-60.

Williams, B. C., Gatti, M. and Goldberg, M. L. (1996). Bipolar spindle attachments affect redistributions of ZW10, a Drosophila centromere/kinetochore component required for accurate chromosome segregation. *J Cell Biol* **134**, 1127-40.

Williams, B. C., Murphy, T. D., Goldberg, M. L. and Karpen, G. H. (1998). Neocentromere activity of structurally acentric mini-chromosomes in Drosophila. *Nat Genet* **18**, 30-7.

Wines, D. R. and Henikoff, S. (1992). Somatic instability of a Drosophila chromosome. *Genetics* **131**, 683-91.

Wong, L. H., Saffery, R. and Choo, K. H. (2002). Construction of neocentromere-based human minichromosomes for gene delivery and centromere studies. *Gene Ther* **9**, 724-6.

Wong, O. K. and Fang, G. (2006). Loading of the 3F3/2 antigen onto kinetochores is dependent on the ordered assembly of the spindle checkpoint proteins. *Mol Biol Cell* **17**, 4390-9.

Wordeman, L. and Mitchison, T. J. (1995). Identification and partial characterization of mitotic centromere-associated kinesin, a kinesin-related protein that associates with centromeres during mitosis. *J Cell Biol* **128**, 95-104.

Yamada, H., Kunisato, A., Kawahara, M., Tahimic, C. G., Ren, X., Ueda, H., Nagamune, T., Katoh, M., Inoue, T., Nishikawa, M. et al. (2006). Exogenous gene expression and growth regulation of hematopoietic cells via a novel human artificial chromosome. *J Hum Genet* **51**, 147-50.

Yang, C. H., Tomkiel, J., Saitoh, H., Johnson, D. H. and Earnshaw, W. C. (1996). Identification of overlapping DNA-binding and centromere-targeting domains in the human kinetochore protein CENP-C. *Mol Cell Biol* **16**, 3576-86.

Yang, J. W., Pendon, C., Yang, J., Haywood, N., Chand, A. and Brown, W. R. (2000). Human mini-chromosomes with minimal centromeres. *Hum Mol Genet* **9**, 1891-902.

Yang, Z., Guo, J., Chen, Q., Ding, C., Du, J. and Zhu, X. (2005). Silencing mitosin induces misaligned chromosomes, premature chromosome decondensation before anaphase onset, and mitotic cell death. *Mol Cell Biol* **25**, 4062-74.

Yen, T. J., Li, G., Schaar, B. T., Szilak, I. and Cleveland, D. W. (1992). CENP-E is a putative kinetochore motor that accumulates just before mitosis. *Nature* **359**, 536-9.

Yoda, K., Ando, S., Morishita, S., Houmura, K., Hashimoto, K., Takeyasu, K. and Okazaki, T. (2000). Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. *Proc Natl Acad Sci U S A* **97**, 7266-71.

Yoda, K., Ando, S., Okuda, A., Kikuchi, A. and Okazaki, T. (1998). In vitro assembly of the CENP-B/alpha-satellite DNA/core histone complex: CENP-B causes nucleosome positioning. *Genes Cells* **3**, 533-48.

Yoda, K., Kitagawa, K., Masumoto, H., Muro, Y. and Okazaki, T. (1992). A human centromere protein, CENP-B, has a DNA binding domain containing four potential alpha helices at the NH2 terminus, which is separable from dimerizing activity. *J Cell Biol* **119**, 1413-27.

Zacharias, H. (2001). Key word: chromosome. *Chromosome Res* **9**, 345-55.

Zecevic, M., Catling, A. D., Eblen, S. T., Renzi, L., Hittle, J. C., Yen, T. J., Gorbsky, G. J. and Weber, M. J. (1998). Active MAP kinase in mitosis: localization at kinetochores and association with the motor protein CENP-E. *J Cell Biol* **142**, 1547-58.

Zeitlin, S. G., Barber, C. M., Allis, C. D., Sullivan, K. F. and Sullivan, K. (2001a). Differential regulation of CENP-A and histone H3 phosphorylation in G2/M. *J Cell Sci* **114**, 653-61.

Zeitlin, S. G., Shelby, R. D. and Sullivan, K. F. (2001b). CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. *J Cell Biol* **155**, 1147-57.

Zhu, X., Chang, K. H., He, D., Mancini, M. A., Brinkley, W. R. and Lee, W. H. (1995). The C terminus of mitosin is essential for its nuclear localization, centromere/kinetochore targeting, and dimerization. *J Biol Chem* **270**, 19545-50.

Zinkowski, R. P., Meyne, J. and Brinkley, B. R. (1991). The centromere-kinetochore complex: a repeat subunit model. *J Cell Biol* **113**, 1091-110.