Generation of Prototype Human Satellite DNA-based Artificial Chromosomes

Ph.D. Thesis

by

Erika Csonka, MD

Medical Microbiology Department

Faculty of Medicine

Albert Szent-Györgyi Medical and Pharmaceutical Center

University of Szeged

Supervisor: Gyula Hadlaczky, Ph.D., D.Sci. Scientific Advisor

Institute of Genetics, Biological Research Center

Hungarian Academy of Sciences

2001

CONTENTS

Introduction 4
1. What is gene therapy and why gene therapy? 4
1.1. Germline gene therapy 5
1.2. Somatic gene therapy 5
1.2.1. *Ex vivo* delivery 5
1.2.2. *In vivo* delivery 6
2. Delivery systems, vectors 6
2.1. Viral systems 6
2.1.1. Retroviral vectors 6
2.1.2. Adenoviral vectors 8
2.1.3. Adeno-associated viral vectors 9
2.1.4. Herpesvirus-based vectors 10
2.1.5. Other viral vectors and hybrid vector systems 11
2.2. Nonviral systems 11
2.2.1. Chemical delivery systems 11
2.2.2. Physical delivery systems 12
2.2.3. Receptor-mediated gene transfer 12
2.3. The "ideal" vector 13
3. Building a MAC 14
3.1. Synthetic (*in vitro*) approach 14
3.2. *In vivo* chromosome fragmentation 14
3.3. *In vivo* "inductive" approach 14
4. The SATAC technology 15
4.1. Generation of satellite DNA-based artificial chromosomes (SATACs) 15
4.2. Generation of human satellite DNA-based artificial chromosomes (hSATACs) ... 17
Materials and methods 18
Culture of cell lines and transfection of mammalian cells 18
DNA and RNA manipulations, plasmid constructs 18
Immunoblotting (Western-blot) 19
Chromosome banding 19

PCR amplification of DNA probes..... 19
Immunostaining of chromosomes and *in situ* hybridization..... 21
Flow cytometry..... 21
Results..... 22
Generation of *de novo* chromosomes..... 22
Analysis of primary clones..... 25
Isolation of stable subclones..... 27
Sequence composition of hSATACs..... 29
Expression of the marker genes from heterochromatic hSATACs..... 32
Separation of hSATACs..... 33
Discussion..... 35
Acknowledgement..... 39
Summary..... 40
References..... 41
List of publications..... 50

INTRODUCTION

My Ph.D. work was to generate human artificial chromosome(s). Significance of the successful development of mammalian artificial chromosomes (MACs) is twofold. MACs would provide a unique model system for studying the structure and function of higher eukaryotic chromosomes. Further more, MACs would represent a non-integrating, optimal vector system with large carrying capacity for different areas of biotechnology, such as cellular protein production, transgenic animal technology including protein production and xenotransplantation, and gene therapy. Successful construction of MAC vectors would have a substantial impact on the further development of gene therapy. Considering this potential of human artificial chromosomes, below there is a brief summary about the basics of gene therapy [1].

1. What is gene therapy and why gene therapy?

Despite advances in clinical practice, many diseases still have no cure and require expensive treatments to prolong the lifespan of patients. The idea that therapeutic gene transfer may permanently cure some diseases has sparked great interest in the field of gene therapy.

The basic concept of gene therapy can be described as the intracellular delivery of genetic material to generate a therapeutic effect by correcting an existing abnormality or providing cells with a new function.

Initially, inherited genetic disorders were the main focus of gene therapy, but now a wide range of diseases, including cancer, vascular diseases, neurodegenerative disorders and other acquired diseases are being considered as targets. At present, the suitability of a disease considered to be treated by gene therapy relies on several prerequisites:

- The disease must be life threatening, thus making its treatment ethically acceptable to the risk of side effects.
- The effects of disease must be potentially reversible by the treatment.
- The gene must be fully characterized.
- Delivery to the affected site must be feasible.
- Well-defined short-term diagnostic parameters are required to demonstrate the physiological effect and benefit of the inserted gene. (For example, in the case of cystic fibrosis the electrical conductance change over the nasal epithelium can be monitored.)

On the basis of the target cells distinction can be made between therapy on the germ cells and the somatic cells of an individual.

1.1. Germline gene therapy

Germline engineering, in practice, today means altering the fertilized eggs – the first cell of the will-be-embryo – so that the genetic changes will be copied into every cell of the future adult, including his or her reproductive cells. Such changes would be passed to the offspring.

Germline gene therapy is currently considered to be ethically unacceptable, even though it has the potential to eradicate many disorders. Until such time as gene therapy for many illnesses is commonplace, with all the difficulties to routine use of gene therapy overcome and side-effects identified and dealt with, germline gene therapy will not be attempted. Ultimately, the question of germline gene therapy will have to be re-evaluated and we may see many of inherited diseases disappear from the gene pool.

> *"It seems obvious that germline therapy will be much more successful than somatic. If we wait for the success of somatic therapy, we'll wait until the sun burns out. We might as well do what we finally can to take the threat of Alzheimer's away from a family or breast cancer away from a family. The biggest ethical problem we have is not using our knowledge,...people not having the guts to go ahead and try and help someone. We are always going to have to take chances."* (James D. Watson)**[2]**

1.2. Somatic gene therapy

The somatic gene therapy involves delivery of genes into diploid cells of an individual where the "implanted" genetic material is not passed onto its progeny. Somatic interventions do not reach beyond the patient being treated.

1.2.1. *Ex vivo* delivery

In this system, genetic material is inserted after explantation, cultivation and manipulation of cells *in vitro*, followed by a subsequent re-implantation. For this purpose either differentiated or stem cells can be used.

The *ex vivo* route is attractive mainly due to the reduced immunological problems and the enhanced efficiency of vector delivery *in vitro*.

So far, the only successful gene therapy clinical trial was carried out by applying the *ex vivo* delivery approach [3].

1.2.2. *In vivo* delivery

The *in vivo* delivery involves two major approaches as the *in situ* delivery, and the systemic administration of the exogenous genetic material.

The *in situ* delivery means the administration of the genetic material directly to the desired tissue. This approach offers effective targeting, but the continuing problem is the low efficiency of the delivery of therapeutic genetic information.

Administration of genetic material systemically is probably the least advanced strategy at present, but potentially the most useful and desirable. The main hurdles of systemic delivery approach are the insufficient targeting of the current vectors to the correct tissue sites, the strong immune response stimulated by most of the vectors used to date, and that the therapeutic lipid-DNA complexes are rapidly eliminated after systemic administration.

2. Delivery systems, vectors

All gene therapy applications depend on the fact that genetic material needs to be delivered across the cell membrane and preferably to the cell nucleus.

2.1. Viral systems

Viruses have throughout evolution developed highly effective methods of entering cells and delivering their genetic information into the cells.

Many viral vector systems now exist for use in gene delivery studies. Common to all viral vectors is the fact that they are genetically modified viruses, containing "foreign" genes and deleted areas, which render them replication incompetent. The most widely used virus-based delivery systems are summarized below.

2.1.1. Retroviral vectors (RV)

Most current retroviral vectors for gene therapy purposes are based on the retroviral oncogenic subgroup Moloney murine leukemia virus (MoMuLV)[4]. The nonpathogenic retroviruses, human spumaviruses are also being examined as potential gene delivery vehicles as they are able to infect quiescent cells. Currently retroviral vectors derived from the lentivirus group are being developed because they provide a promising tool to facilitate retroviral mediated gene therapy [5,6].

The removal of practically all of the viral genes provides an approximately 9-10kb of space for transgene incorporation. In this way infectious replication-incompetent virus particles can be generated.

Advantages offered by replication-defective retroviruses:

- Efficient and stable integration of the transgene into the host genome, enabling the long-term expression of the transgene **[7]**.
- The targeting of retroviruses to specific cell populations by engineering heterologous protein domains into the envelope glycoproteins that modifies their tropism **[8]**.

There are drawbacks limiting the application of retrovirus-based vector system:

- Inability to produce high titers of purified virus particles.
- The possible contamination of retroviral stocks with replication-competent retroviruses or recombination of the replication-defective viruses with wild-type (retro-) viruses.
- Size limitations on the amount of DNA, which can be stably incorporated into the retroviral cassettes to allow packaging into the retroviral particles. This limits the application of larger genes or large regulatory elements.
- The risk of insertional mutagenesis of essential genes or activation of cellular oncogenes upon random integration into the host genome.
- The conventional MoMuLV-based vectors, targeting only mitotically active cells, have limited gene delivery into targets such as muscle, neurons, hepatocytes and hematopoietic stem cells **[5,9]**.
- Sensitivity of murine retroviruses to inactivation by the human complement-mediated lysis pathway of the immune system **[10]**.
- The vector promoter silencing (promoter "shut-off") in stably transduced cells *in vivo* **[11]**.

At present, the retrovirus-based vectors are the most widely used delivery vehicles in clinical trials accounting for approximately 40% of studies. A landmark gene therapy trial for X-linked SCID was successfully completed in two infants **[3]**, based on the use of complementary DNA containing a defective Moloney retrovirus-derived vector and *ex vivo* infection of $CD34^+$ stem cells. After a 10-month follow-up period, gammac transgene-expressing T and NK cells were detected in the two patients. The T, B and NK cells counts and function, including antigen-specific responses, were comparable to those of age-matched controls. This clinical trial can be regarded as the first success of the gene therapy because

gene therapy was able to provide full correction of disease phenotype and, hence, clinical benefit.

2.1.2. Adenoviral vectors (Ad)

The adenovirus-based vectors are also commonly used for gene therapy applications. Adenoviruses types 2 and 5 are thought to be suitable as vectors for gene therapy purposes and recombinant adenovirus vaccines, since they are not implicated in serious illness **[12]**.

The first generation adenovirus vectors are deleted in the E1 region - resulting in approximately 7.8kb **[13]** space to incorporate exogenous DNA sequences -, theoretically rendering the adenovirus unable to replicate in non-E1 complementing cell lines **[14,15]**. Although deleted for E1 function, first generation adenovirus-based vectors were found to result in low level leaky expression of the remaining viral genes initiated by the host E1-like transcription factors **[16]**.

Second generation adenovirus vectors are being designed to be less immunogenic to the infected host, aiming at complete ablation of expression of the immunodominant viral proteins facilitating persistent long-term expression. As well as minimizing the leaky expression of highly immunogenic adenoviral proteins, these deletions increased the cloning capacity of the vectors. On the other hand, deletion of more than one essential regulatory element reduces the risk of the emergence of replication competent adenovirus from genetic recombination with adenovirus sequences in *trans*-complementing cell line or from endemic wild-type adenovirus infections.

More recently a new generation of the adenoviral system has been developed in which practically all of the viral genome has been deleted. These so-called "gutless" or "pseudo" adenoviruses (PAV) have a transgene capacity up to 37kb.

The major advantages of the adenovirus-based vectors:

- Their limited pathogenicity.
- Wide host range and tropism for most of the human cells.
- High infectivity *in vitro* and *in vivo*.
- The ability to infect quiescent as well as dividing cells **[17]**.
- The adenovirus genome rarely integrates into the host chromosome and therefore replicates in an extra-chromosomal state, minimizing the genotoxicity associated with insertional mutagenesis **[18]**.
- The ability to produce high numbers of purified particles.

Disadvantages of the adenoviral system:

- Current adenovirus vectors have a relatively tight restriction on the size of DNA, which can be packaged into virions **[14]**.
- Since adenovirus-delivered genes are episomally maintained and lost due to genetic instability, they are expressed transiently. Hence repeated administrations are necessary to maintain stable expression of the transgene.
- High immunogenic feature of the virus particles and their gene products, which limits the number of repeated applications *in vivo*. Host immune responses result in transient gene expression *in vivo* **[16,19,20]**, ultimately leading to the loss of transgene expression due to eradication of successfully transduced cells **[20]**.

The adenoviral vehicle system is particularly attractive to the cystic fibrosis therapists and also could be used in cancer gene therapy, where enhanced cell death due to immune recognition is desirable.

The first victim of the gene therapy trials was the 18-year-old patient, Jesse Gelsinger, a participant in the experimental gene therapy trial for ornithine transcarbamilase (OTC) deficiency. He died four days after being injected with corrective genetic material using an adenoviral-based vector system. The cause of his death was that the modified "cold" adenovirus-based vector initiated an unusual and deadly immune-system response that led to multiple organ failure and death **[21]**.

2.1.3. Adeno-associated viral vectors (AAV)

The AAV is naturally replication incompetent and requires additional genes from a helper virus infection which in nature are generally complemented by adenovirus or herpes simplex virus coinfection **[22]**. The wild-type AAV has been demonstrated to preferentially integrate into human chromosome #19 at site q13.4 directed by the *rep* gene **[23,24]**. (This chromosomal region is thought to be involved in malignant transformation events of cells.)

AAV-based vectors generally involve replacement of the *rep* and *cap* genes with a transgene of interest **[24]**.

The AAV-based vector system offered advantages:

- High transduction frequencies.
- The ability to infect all cultured cells tested **[24]**, and *in vivo* (systematic) applications have resulted in the transduction of the majority of tissues.
- The ability of preparation of high titers.

The *rep* deleted AAV vectors has a number of limitations at present:

- Complicated process of vector production.
- The purified stocks easily contaminated with wild-type AAV and helper viruses **[24]**.
- As the *rep* functions are deleted, the assembled virions lose the ability to selectively integrate into the host genome and tend to integrate randomly, introducing the same risk of insertional mutagenesis as retroviruses **[23]**.
- The limited transgene capacity of the particles, which is approximately 4.7kb **[24]**.

2.1.4. Herpesvirus-based vectors

The current herpesvirus vectors are based on the herpes simplex virus (HSV) type 1. The HSV based systems include the development of the so-called disabled infectious single copy (DISC) virus which comprises a glycoprotein H defective mutant HSV genome. Viral particles generated in complementing cells can infect any other cells, replicate their own genome but will not produce further infectious particles.

Current HSV-based vectors involve replacement of one or more of the five immediate early (IE) genes whose functions are *trans*-complemented by packaging cell lines **[25,26]**. The major interest comes from the possibility of a "gutless" HSV vector, which has the potential of accommodating up to 150kb of insert DNA **[27]**.

Advantageous features:

- HSV has a wide tropism, infecting virtually any human cell, and is capable of nuclear delivery, infecting dividing as well as quiescent cells **[25]**.
- Herpesviruses infect and persist in cells of the nervous system, hence herpesvirus-based vectors may provide a unique strategy for gene transfer to cells of the nervous system **[28]**.

Limitations:

- As the HSV genome is episomally maintained in the nucleus of the infected host, they provide only transient gene expression.
- The insert capacity of the current HSV based vectors is approximately 20kb.
- Immunogenic features due to the viral proteins expressed in the infected host cells.
- The risk of the emergence of replication competent herpes virus from genetic recombination with herpesvirus sequences in *trans*-complementing cell line or from wild-type herpesvirus infections.

2.1.5. Other viral vectors and hybrid vector systems

Although the RV, Ad, AAV and HSV are currently the most advanced viral vector systems applied in gene therapy trials, a large number of other viruses have potential: poliovirus, vaccinia virus, cytomegalovirus, measles virus, pseudorabies virus, baculovirus (insect virus) and parvovirus **[29]** are being developed for gene therapy. Hybrid vectors are also being developed which combine advantageous properties of current viral vector systems into single systems, generating novel viruses with unique combinations of functions **[30]**.

2.2. Nonviral systems

Alternatives to the viral strategies are the application of chemically synthesized vehicles, delivery by physical methods or receptor mediated gene transfer.

2.2.1. Chemical delivery systems

The vast majority of current chemical delivery systems are based on polycationic entities which cause compaction of negatively charged nucleic acids accompanied by the formation of nanometric complexes: lipoplexes or polyplexes **[31]**. These polycationic entities belong to one of two main categories which are either cationic liposome/micelle (DOPE, DOTMA **[32]**, DOTAP **[33]** etc.) or cationic polymer-based (poly-L-lysine **[34]**, polyamidoamine **[35]**, polyethylenimine **[36,37]** etc.). These nanometric complexes are generally stable enough to protect bound nucleic acids from degradation and are able to enter cells, usually by endocytosis. Inside the cells, a proportion of the bound nucleic acids are able to dissociate and escape from early endosomes into the cytoplasm either to perform a therapeutic function there or traffic into the nucleus.

There are several problems with the chemical delivery strategies:

- Lipoplexes or polyplexes frequently show instability towards aggregation even before entering a cell, especially under conditions of physiological salt levels. They are also unstable in the presence of body fluid components such as serum proteins.
- Neither lipoplexes nor polyplexes are cell type selective.
- None of them can enter cells very rapidly.
- Inside the cell, the vast majority of lipoplex and polyplex particles aggregate within late-endosome compartments producing macroparticles from which nucleic acid can hardly dissociate and escape. Moreover, if nucleic acids manage to escape the endosome

compartments after delivery, access to the nucleus represents another barrier to nucleic acid function.

2.2.2. Physical delivery systems

The most important physical delivery options are electroporation, microinjection, biolistics or "gene gun" and magnetofection.

Electroporation is a routine laboratory method for transfection of cultured mammalian and bacterial cells. Electrodes inserted into tissue emit controlled electrical pulses that are able to create transmembrane potentials sufficient to allow locally administered naked DNA to enter cells, probably through the transient pores **[38-40]**.

Microinjection involves the direct injection of the therapeutic gene into the desired tissue *in vivo*. Remaining problems with this strategy are to solve the long term expression of the introduced gene and to reach sufficient number of the cells of the chosen tissue.

In the case of biolistic or "gene gun" gold or tungsten particles are coated with DNA and accelerated to high speeds by electronic or helium pressure discharge, enabling the coated particles to penetrate the target tissues **[41-45]**. This method causes serious cellular damage.

In magnetofection magnetic iron-oxide nanoparticles are coated with cationic or anionic polymers and naked DNA or recombinant adenoviral vector. Complexes of vectors and magnetic particles can transfect cell lines in culture. Under the influence of the magnetic fields, transfection/transduction is enhanced up to several orders of magnitude and can be directed to a specific area within the cell population in the culture dish **[46]**.

2.2.3. Receptor-mediated gene transfer

In this strategy, the ligand specific to a receptor of a target cell is covalently bound to a polycationic molecule such as poly-L-lysine. The polycationic component of this complex is capable of making linkage with DNA molecule via electrostatic interactions. After a specific binding between the receptor and its ligand this complex can get into the cell by natural endocytosis.

Problematic features of this strategy:

- The endocytotic vesicles are transported into the lysosomes where they regularly degrade. Efficiency of the receptor-mediated gene delivery system can be increased exploiting advantage of a parallel adenoviral infection as the adenoviruses are able to crack the endosomes subsequently that they managed to get into the cell.

- Introduced genes are expressed transiently, though in some cases integration of the foreign DNA into the host genome was observed, which can result in significantly prolonged expression **[47]**.
- Theoretically, any size of foreign DNA can be used, however, a too long DNA molecule can break.
- Though this method enables targeting of particular cells and cell types, the safe and efficient targeting is still unsolved.

2.3. The "ideal" vector

A key issue in the successful development of gene therapy is the development of appropriate vector and delivery systems that allow efficient gene transfer and safe, controlled gene expression. Considering the benefits and drawbacks to the current vector systems the requirements of a "safe" or "low-risk" vector useable for gene therapy purposes can be established **[48]**.

- It should be free of elements that enable rearrangements in the host genome (recombination, mutation) and the vector must not have negative influence to the function of the cell, tissue or organism containing it.
- It should be stable in the host nucleus without integration into the host genome or designed for site-specific integration so that not interfere with the host genomic functions.
- The copy number of therapeutic gene(s) carried by the vector should be controlled and stable. The vector should provide independent and controlled function of the therapeutic gene(s).
- It should accept large (up to Mb size) inserts and ensure the functional stability of the insert.

None of the available DNA vectors meet all of these requirements, except the natural "vector" that is the eukaryotic chromosome. The perfection of eukaryotic chromosomes makes them the most sophisticated natural gene transfer and operation system. Successful modelling this natural system by construction of artificial chromosomes would provide a new vector system for gene therapy.

3. Building a MAC

In recent years, basically three main approaches were developed for construction of MACs:

3.1. Synthetic (*in vitro*) approach

This strategy involves combination of isolated and cloned structural and functional elements of chromosomes *in vitro* and introduction of them into cells. The main blocks of the chromosomes that are necessary to their stable maintenance and replication can be defined as centromere, telomeres, autonomously replicating sequence (ARS) and "filling" DNA providing appropriate physical size. In addition, normally a chromosome contains genes, which in the case of an artificial chromosome also involves a selectable marker gene **[48]**.

This methodology was successfully used for construction of artificial chromosomes in lower eukaryotes like yeasts *Saccharomyces cerevisiae* **[49]** and *Schizosaccharomyces pombe* **[50]**. *In vitro* assembly of MACs from cloned elements is hindered by the lack of an isolated, functional mammalian centromere, all the other elements of mammalian chromosomes are identified, cloned and available.

3.2. *In vivo* chromosome fragmentation

This approach utilizes the telomere-directed fragmentation of existing natural chromosomes **[51-59]**. Different human minichromosomes were generated in this way, they show different stability and their purification has not been solved yet. However, such artificial human minichromosomes hold the promise of becoming an engineerable vector system **[60]** for human gene therapy. Similarly, naturally occurring centric chromosome fragments **[61-64]** may have the same potential.

3.3. *In vivo* "inductive" approach

This method is based on an inducible intrinsic mechanism, which can facilitate large-scale amplifications and formation of *de novo* centromeres and chromosomes **[65,66]** in mammalian cells, upon the integration of exogenous DNA sequences into the specific region of the chromosomes. Induced amplification of the pericentric/centromeric region of mammalian chromosomes leads to the formation of new chromosomes (Satellite DNA-based Artificial Chromosomes = SATACs) that are composed of coamplified satellite DNA and exogenous DNA sequences. SATACs can be regarded as artificially generated accessory chromosomes composed of predictable DNA sequences, and contain defined genetic information.

4. The SATAC technology

4.1. Generation of satellite DNA-based artificial chromosomes (SATACs)

In our laboratory, results of transformation experiments revealed that in certain cases, integration of exogenous DNA into mouse chromosomes led to the formation of new centromeres, dicentric chromosomes, and ultimately, after breakage of dicentric chromosomes, *de novo* chromosomes **[65]**. Detailed analysis of the chromosomal structures in transformed mammalian cells indicated that *de novo* chromosomes were generated via a large-scale amplification process **[66]**. It became clear that large-scale amplification and *de novo* chromosome formation could be induced reproducibly, and independent from the sequence of exogenous DNA used in transformation **[67]**. Analysis of the replication and the amplicon structure of *de novo* formed chromosomes resulted in the identification of a higher order replication unit (megareplicon) and a primary replication initiation site (megareplicator) at the pericentromeric region of mouse chromosomes. These results suggested that large-scale amplification of the centromeric regions as well as the formation of new centromeres, chromosome segments, and stable chromosomes could be attributed to a replication-directed mechanism **[68]**. There are several lines of direct and indirect evidences suggesting that the ribosomal-DNA (rDNA) containing chromosomal sites can be responsible for the large-scale amplification events and the rDNA itself may have importance in the *de novo* chromosome formations [Fodor, K. *et al.* unpublished]. When the exogenous DNA integrates **(Fig. 1A,B)** into close proximity of a megareplicator, it could lead to a replication error, which initiates large-scale amplification **(Fig. 1C)** of surrounding sequences and brings about the formation of large inverted repeats (amplicons). These amplicons are composed of coamplified endogenous and exogenous DNA sequences, and they are the building blocks of the new chromosomal segments. Amplified centromeric regions can eventually form active centromere that leads to the formation of a dicentric chromosome **(Fig. 1D)**. The presence of two active centromeres on the same chromosome causes specific breakage between the two centromeres **(Fig. 1E)**, which ultimately brings about the existence of a new chromosome **(Fig. 1F)**. During the recent years, a number of mouse SATACs have been generated by this technology **[65-67,** Praznovszky, T. *et al.* unpublished].

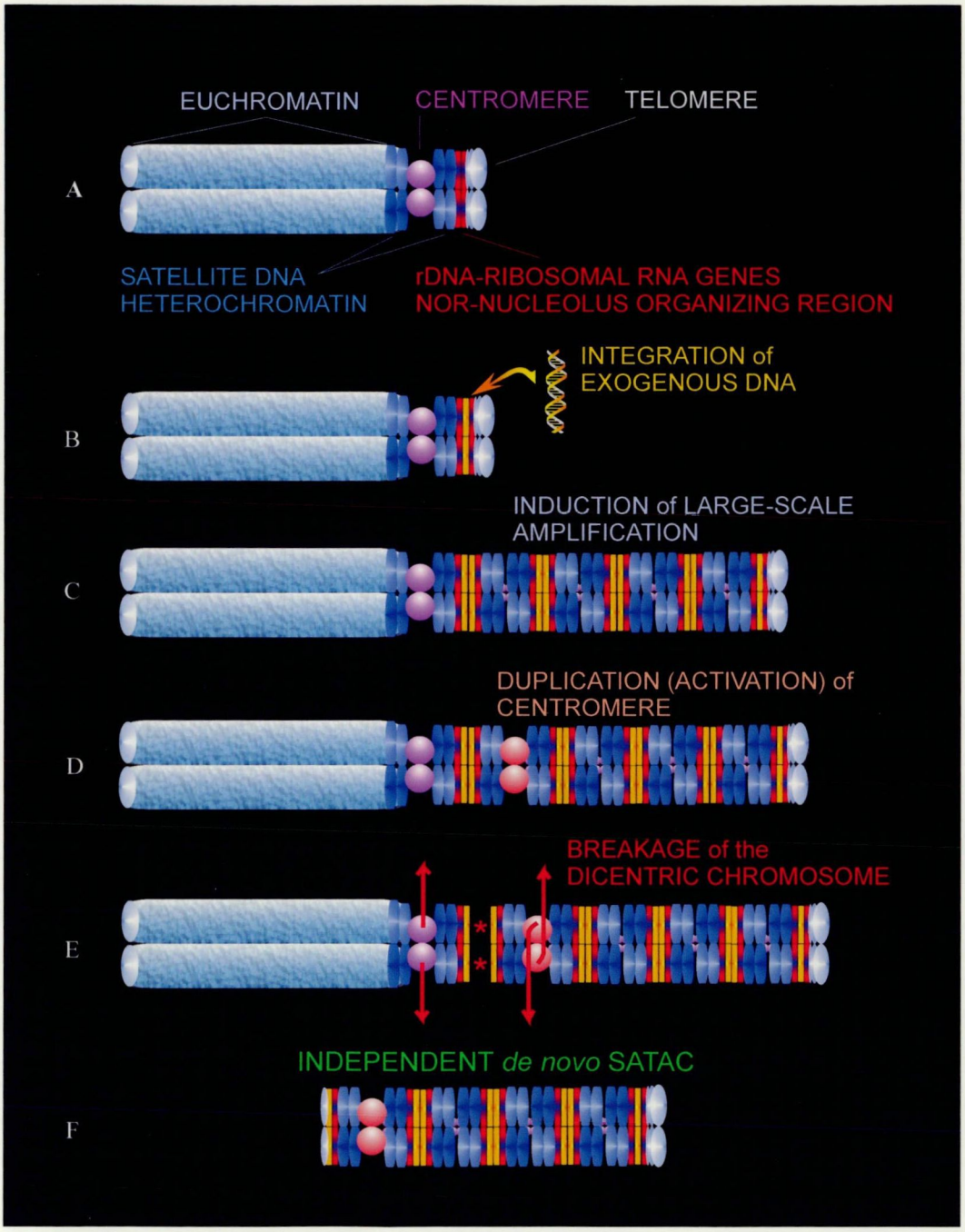

Figure 1. Generation of MAC via *in vivo* inductive approach. A. Mammalian NOR chromosome. **B.** Integration of exogenous DNA into close proximity to a megareplicator leads to a replication error, which initiates large-scale amplification (**C**) of surrounding sequences and brings about the formation of large inverted repeats (amplicons). **C.** These amplicons are composed of coamplified endogenous and exogenous DNA sequences, and they are the building blocks of the new chromosomal segments. **D.** Amplified centromeric regions can eventually form active centromere that leads to the formation of a dicentric chromosome. **E.** The presence of two active centromeres on the same chromosome causes specific breakage between the two centromeres, which ultimately brings about the existence of a new chromosome (**F.**).

4.2. Generation of human satellite DNA-based artificial chromosomes (hSATACs)

The question was raised, weather this *in vivo* inductive approach for generation of mammalian artificial chromosomes observed in rodent cells can be applied to other mammalian species or it is restricted to rodent cells. To answer this question we attempted to generate a human satellite DNA-based artificial chromosome on human chromosome #15.

This human acrocentric chromosome was chosen, because the NOR (NOR = Nucleolus Organizing Region) is fairly well characterized on the short arm of this chromosome. Clinical cytogenetics data indicate that polymorphism (spontaneous amplification) of the short arm of the human acrocentric chromosomes are the most frequent structural changes without phenotypic effect. In higher eukaryotes, ribosomal RNA genes are amplified and organized in the form of tandem repeats, their copy number shows significant variance in individuals. Based on these facts, it seems unlikely that induced amplification of rDNA or the surrounding satellite DNA sequences identified at these chromosomal regions would cause undesirable phenotypic effects. To our knowledge, no unsuspected or unwanted coding sequences are localized in these chromosomal regions, the only genes found in these chromosomal sites are the rRNA genes. Sequences of these regions are suitable basic material to build artificial chromosomes. Therefore, the short arm of human acrocentric chromosomes bearing amplified ribosomal RNA genes (rRNA, rDNA) and different non-coding satellite DNA sequences, seems to be an optimal chromosomal region for inducing *de novo* chromosome formation.

We hypothesized, that the basic steps of *de novo* chromosome formation **(Fig. 1)** are similar in mouse and in human cells. Induction of large-scale amplification in the short arm region of human chromosome #15 theoretically would lead to the formation of a *de novo* chromosome that is composed of amplified endogenous sequences of the centromeric/short arm region of the human chromosome, and the multiple copies of coamplified exogenous DNA sequences. Human SATACs generated on this way would possess all the elements necessary to be a stable chromosome: centromere, telomeres, and different noncoding satellite DNA sequences that are safe "filling" sequences. Because of the high satellite DNA content and the presence of amplified non-satellite sequences, these SATACs would be heterochromatic.

MATERIALS AND METHODS

Culture of cell lines and transfection of mammalian cells

Cell line 94-3 (NIGMS, GM 10664) was produced by fusion of lymphoblasts of a patient with a t(X;15) (q25 or q25;q26) and HPRT deficient Chinese-hamster RJK88 cells **[69]**, and was obtained from NIGMS Human Genetic Mutant Cell Repository, Coriell Cell Repositories, Coriell Institute for Medical Research, Camden, New Jersey, USA. Cells were cultured in Dulbecco's Modified Eagle Medium (high glucose) (GibcoBRL) supplemented with 10% heat inactivated foetal bovine serum (GibcoBRL) (DMEM-10FBS). The EJ30 human bladder carcinoma cell line was maintained in Nutrient Mixture F-12 HAM medium (Sigma) containing 10% heat inactivated foetal bovine serum (F12-10FBS). The mouse Satellite DNA-based Artificial Chromosome (mSATAC) carrying cell lines (produced in our laboratory), the mM2C1pCEP8 was cultured in F12-10FBS supplemented with 168μg/ml HygromycinB (Merck) and 10μg/ml Puromycin (Sigma), the E420-19 was maintained in Iscove's Modified Dulbecco's Medium (GibcoBRL) containing 15% heat inactivated foetal bovine serum (IMEM-15FBS) supplemented with 200μg/ml HygromycinB.

Cotransfection of plasmid DNAs was carried out by calcium-phosphate DNA coprecipitation method **[70]** using 2μg of pCH110 linearized with *BamHI*, 7μg of pBabePuro opened with *EcoRI*, and 100μg of carrier rDNA (pK161) linearized with *ClaI* restriction endonuclease per semiconfluent 100mm petri dish ($5x10^6$ cells) of 94-3 cell line. Selection of transformants was started 48 hours after transfection on 10μg/ml Puromycin in F12-10FBS medium. Stable subclones of transformants were maintained in IMEM-10FBS containing 10μg/ml Puromycin.

DNA and RNA manipulations, plasmid constructs

All general DNA and RNA manipulations were performed by standard procedures **[71]**; pBabePuro plasmid **[72]** carrying a mammalian selectable marker gene - the Puromycin acetyl-transferase gene –was a kind gift from Dr L. Székely (MTC, Karolinska Institutet, Stockholm); pCH110 carrying β–galactosidase gene was obtained from Pharmacia; pK161 carrying a 9kb coding sequence of mouse ribosomal-DNA in pWE15 (Stratagene) was cloned in our laboratory [Fodor, K. *et al.* unpublished]. ACR-1, a plasmid clone (ATCC 61650) of a 4.3kb unique sequence distal to the rDNA cluster on the short arm of all 5 human acrocentric

(NOR) chromosomes [73] was obtained from the American Type Culture Collection (ATCC), Rockville, USA.

Immunoblotting (Western-blot)

Western-blotting was performed using standard procedures [71]. The β-galactosidase enzyme was detected using 2000x diluted monoclonal anti- β-galactosidase antibody (mouse, Promega) and 1000x dilution of HRPO-conjugated anti-mouse antibody (rabbit, DAKO).

Chromosome banding

Constitutive heterochromatin was detected by Giemsa/Barium-hydroxide C-banding staining [74] with the modification that after the 1 hour depurination treatment in 0.2N HCl the slides were incubated in prewarmed saturated Barium-hydroxide octahydrate diluted 1:1 with distilled water at 50°C for 3-5 minutes.

PCR amplification of DNA probes

To avoid cross hybridization of vector sequences, DNA probes were generated by Polymerase Chain Reaction. Primers were synthesized on a Pharmacia Gene Assembler. List of forward (F) and reverse (R) primers:

β-galactosidase gene (pCH110; Pharmacia): LacZ-F (β-galactosidase 121-143) 5'-GAAGAG GCCCGCACCGATCGCCC; LacZ-R (β-galactosidase 772-750) 5'-TTACCCGTAGGTAGT CACGCAAC.

Puromycin acetyl-transferase gene (pBabePuro): Puro-F: 5'-ATGACCGAGTACAAGCCCA CGGTGCGC; Puro-R: 5'-TCAGGCACCGGGCTTGCGGGTCATGCA.

Human rDNA 5'-external transcribed spacer [75]: HETS-F (1-24 of U13369) 5'-GCTGACA CGCTGTCCTCTGGCGAC; HETS-R (U13369: 1520-1497) 5'-GTCCTCTGCGAGCGGGT CGCTACG.

rDNA [76] coding sequence pK161-M7 carrying part of conserved 18S RNA sequence: pK161-M7-F: 5'-AGGTGCGTCTGCGGGTTGGGGCTCGTC; pK161-M7-R: 5'-AAGGAT CCTCGTTAAAGGATTTAAAGTGGAC.

Chromosome-specific α-satellites: α-satellite primer 1 (pS12, [Cserpán, I. *et al.* unpublished], X60716: 3920-3898): 5'-CCTGAAAGCGCTCCAAATGTCCA; α-satellite primer 2 (pS12, X60716: 3582-3605): 5'-CCTAAGGTGAAAAAGGAAATATCT.

Centromeric *NotI* repeat DNA **[77]**: Notrep F primer (U53226: 50-75): 5'-GGGTTTAAATA GCCTCGGGCGCAGC; Notrep R primer (U53226: 571-545): 5'-TAGTAGATTGGATTAT CTGGAGCCACA.

β-satellite (distal) hybridizing to the p13 region of acrocentric chromosomes **[78]**: β-satdist F primer (M81228: 1-26): 5'-ATAAGCTTAGGCAAGAGTTGCATCACCT; β-satdist R primer (M81228: 955-930): 5'- TGAAGCTTTGCCTACAGGGGATTGTGAC.

β-satellite (proximal) hybridizing to the p11 and p13 regions of acrocentric chromosomes **[78]**: β-satprox F primer (M81226: 38-65): 5'- CAAAGCTTAGACAAGAGTTACATCACC T; β-satprox R primer (M81227: 200-176): 5'-TGAAGCTTTCCTAGAGGCACATTGGGA C.

Satellite III DNA **[79]**: Satellite III F primer (M21305: 1-25): 5'-TGGGAATTCAATAGAAT GGAATGGTAT; Satellite III R primer (M21305: 1608-1582): 5'-TGCAATAGAATGGAAT GGAATCAACT C.

147 bp repeat sequence located at the proximal junction of the human rDNA cluster on all acrocentric chromosomes **[80]**: 147rep22 F primer (D31961: 156-177): 5'-CGGGTTGGGAC CAGTTAGGT; 147rep22 R primer (D31961: 358-379): 5'-TGGAGTCTGTGAGTCTGTAG GT.

All PCR products listed above were isolated by agarose gel electrophoresis, and identity of the fragments was verified by sequencing from the both ends with the amplification primers on an ABI Prism automated DNA sequencer.

The PCR generated telomeric probe was produced by the template-free method **[81]**, and labelled either directly by PCR or nick translation (DIG-Nick Translation Kit, Boehringer) with DIG-11-dUTP (Boehringer) then purified on Nick Spin Column (Pharmacia).

DOP-PCR (Degenerate Oligonucleotide-Primed PCR) of flow sorted human Satellite DNA-based Artificial Chromosomes (hSATAC) was carried out as described **[82]**, with the modification that TaKaRa PCR Buffer, dNTP-mix and Taq Polymerase (5U/ 50μl reaction volume) were used in the presence of 8μM 6MW single degenerate primer. PCR was performed on MiniCycler (JM Research) thermal cycler. Probes generated by DOP-PCR were labelled by nick-translation for reverse painting **[83]**.

Immunostaining of chromosomes and *in situ* hybridization

Indirect immunofluorescence staining of chromosomes with human LU851 anti-centromere serum was performed as described previously **[67,84]**. Immunostaining with fluorescein conjugated anti BrdU monoclonal antibody (Boehringer) was performed according to the manufacturer's suggestions.

For *in situ* hybridizations biotin (BIO) or digoxigenin (DIG) labelled human genomic painting probe (total genomic DNA from EJ30 cell line), a 1kb synthetic telomere [Cserpán, I. *et al.* submitted], PCR generated telomere probe **[81]**, 147bp human rDNA flanking sequence probe, human 9kb repeat, ACR-1, LacZ PCR, Puro PCR, β-satellite, and satellite III probes were prepared by using a nick translation kit (Gibco BRL) according to the manufacturer's suggestions. BIO and DIG labelled human 15, 22, and X chromosome specific painting probes, D15Z, D22Z1 and DXZ1 α-satellite centromeric probes were obtained and used according to the suggestions of the supplier (Oncor). Standard *in situ* hybridizations were performed as described earlier **[65]**. To detect human chr #15 centromere a BIO or DIG labelled monomeric α-satellite variant was also used that was isolated in our laboratory [Cserpán, I. *et al.* unpublished]. The specificity of all human probes used was verified by *in situ* hybridization on human lymphocyte preparations made from the peripheral blood of a healthy individual. Microscopy was made with an Olympus AH-2 photomicroscope equipped with Quips XL Genetics Workstation system including a Photometrics KAF 1400-G2 CCD camera (Vysis).

Flow cytometry

Metaphase chromosomes were isolated from colchicine-blocked cells and stained with Hoechst 33258 and Chromomycin A3 using standard procedures **[85]**.

Flow analysis and purification of the hSATACs were performed on a FACS Vantage (BDIS, San Jose, CA) equipped with a turbosort option and two Inova 305 lasers (Coherent, Palo Alto, CA). Condensing agents were added to the sheath buffer to maintain condensed chromosomes after sorting. The final buffer contains 15 mM Tris HCl, 0.1 mM EDTA, 20 mM NaCl, 1% hexylene glycol, 100 mM glycine, 20 μM spermine and 50 μM spermidine, pH:7.6. The sorted SATACs were collected in 1.5 ml screw-capped Eppendorf tubes at 4°C, at a sort concentration of approximately $1x10^6$ chromosomes / ml, which were then stored at 4°C until needed. For FISH analysis of sorted fractions, aliquots of approximately 10,000

chromosomes from each sorted region were placed on glass slides after addition of 0.2% formaldehyde for 5 minutes. Slides were then left to dry and dehydrated before being hybridized to labelled probes (pBabe Puro plasmid and PCR-generated telomere) using standard conditions. Following FISH detection, at least 20 chromosomes were scored for each sorted region.

RESULTS

Generation of *de novo* chromosomes

A Chinese-hamster cell line (94-3) **[69]** was chosen as the host cell, because it carries human acrocentric chromosomes including the chromosome #15. It offers efficient targeting of the human ribosomal region and since hamster chromosomes carry practically no large constitutive heterochromatic segments, amplified human heterochromatic regions can be detected easily by Giemsa staining technique (C-banding) on Chinese-hamster background.

Initial characterization of the 94-3 cell line revealed that in 100% of the cells, human t(X;15) chromosome (referred herein as t(X;15q25)) and human chromosome #22 were retained. In addition, 15% of the cells had human/hamster translocated chromosomes, and 20% of the cells carried human chromosome fragments (NIGMS Human Genetic Mutant Cell Repository).

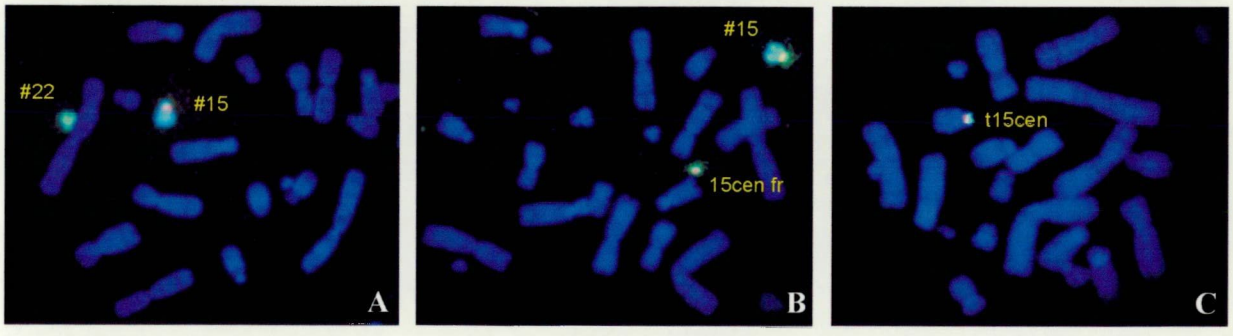

Figure 2. Human chromosomes in the host 94-3 cell line. Two color FISH on metaphases of 94-3 cells with biotin-labelled human genomic painting probe (green signal), and digoxigenin (DIG)-labelled D15Z probe (pink-white signals). **A.** human chromosomes t(X;15q25) and #22. **B.** Human chr t(X;15q25) and a small human chromosome fragment with chr #15 centromere. **C.** Acrocentric hamster chromosome with translocated human chr #15 centromere and short arm region.

First we verified the presence of human chromosomes in 94-3 cells by *in situ* hybridizations with biotin labelled human genomic DNA probe, and also with biotin labelled human α-satellite DNA sequences, and confirmed that 96.8% of the cells carried human chromosomal material. Two hundred and ninety-six metaphases of 94-3 cells were analyzed in detail by *in situ* hybridization using a human genomic painting probe (EJ30), human chromosome-specific painting probes (chr #15, #22, and X), and human chromosome-specific α-satellite centromeric probes (D15Z, D22Z1, and DXZ1). Details of the FISH analysis of the 94-3 cell line is presented in **Figure 2** and **Table 1**.

Table 1. Analysis of 94-3 cell line by FISH

Human chromosome (intact or fragment)	FISH probes	Percentage in 296 metaphases analyzed
Human chromosome material	Genomic painting	96.8
Chromosome t(X;15q25)	15 painting, D15Z	56.0
Chromosome X	X painting	56.0
Chromosome # 22	22 painting, D22Z1	15.5
Chr #15 centric fragment	D15Z	3.4
Chr # 22 centric fragment	D22Z1	34.1
Chr X centromere	DXZ1	0.0
Translocated human/hamster chromosome*	Genomic painting, D15Z, and D22Z1	24.7

* These metaphases contained an acrocentric hamster chromosome with a centromere and short arm composed of human chromosomal material. The majority (89%) of these translocated chromosomes showed centromeric hybridization with D22Z1, 7% of the centromeres were labelled with D15Z (Fig. 2C), and in 4% of the centromeres no human alphoid sequences were detected.

Our aim was to integrate marker gene(s) into the ribosomal DNA region (rDNA, NOR=Nucleolus Organising Region) of a human chromosome via site-specific integration, and induce large-scale amplification and *de novo* chromosome formation. Large amounts of target sequences used as a carrier in cotransfection, can promote site-specific integration of the exogenous DNA sequences **[67,86]**. The ribosomal RNA coding sequences are highly conserved in eukaryotes (from yeast to human), therefore mouse rDNA sequences are suitable for targeting the rDNA region of human (or hamster) chromosomes. In the case of human

acrocentric chromosomes, besides the rDNA, a number of neighbouring endogenous sequences have been identified in the NOR regions **(Fig. 3)**. These sequences are: the proximal 147bp repeat **[80]** proximal and distal β-satellite **[87]**, ACR-1, a unique flanking sequence distal to the rDNA cluster **[73]**, a methylated human 9-kb repetitive sequence **[77]** (*Not I* repeat), telomeric sequences on the distal ends, D15Z1 satellite III **[79]**, and D15Z α-satellite DNA sequences towards the centromere. Except for ACR-1 sequences, all these sequence elements were detectable on the human chr t(X;15q25) in the 94-3 cell line by *in situ* hybridization with the corresponding probes.

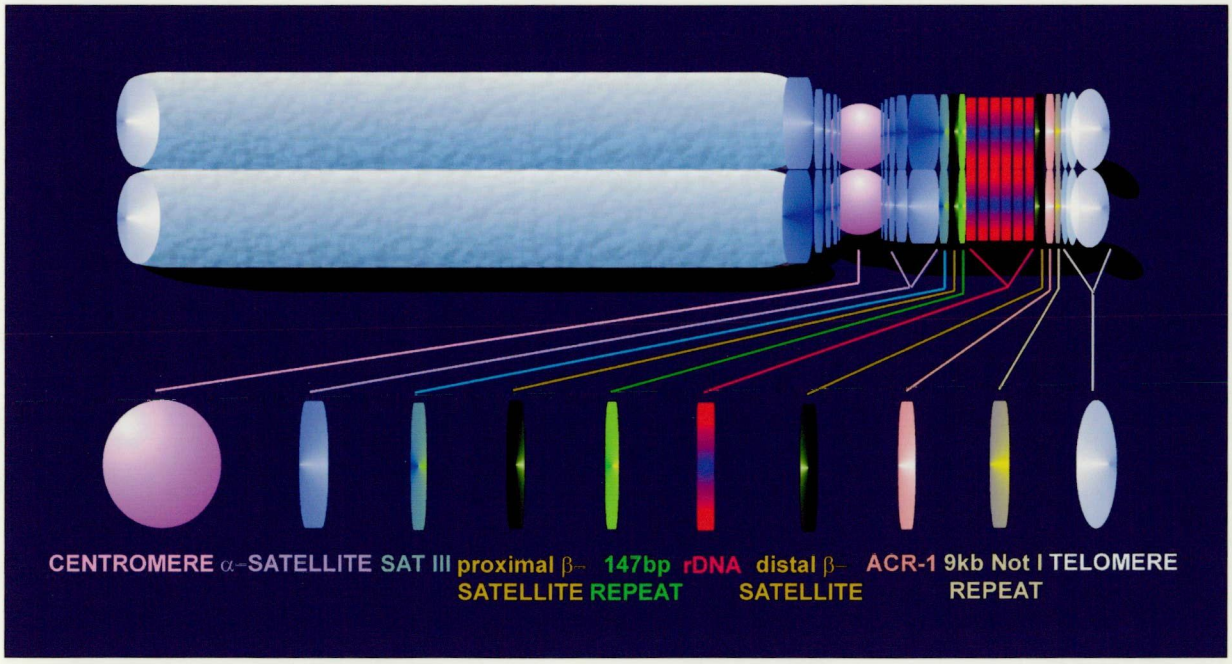

Figure 3. DNA composition of the short arm of the human chromosome #15. In the case of human acrocentric chromosomes, besides the rDNA, a number of neighbouring endogenous sequences have been identified in the NOR. These sequences from the centromere towards the telomere are: D15Z α-satellite DNA sequences, D15Z1 satellite III, proximal β-satellite, 147bp repeat, rDNA cluster, distal β-satellite, ACR-1 specific to the short arm of human acrocantric chromosomes, a methylated human 9-kb repetitive sequence (*Not I* repeat), and telomeric sequences on the distal ends.

To test whether large-scale amplification can be induced on human chromosomes in a host cell of another species, a semiconfluent culture of 94-3 cells was cotransfected with a mixture of plasmid DNAs of pBabePuro linearized with *EcoRI*, pCH110 linearized with *BamHI*, and carrier rDNA (pK161) linearized with *ClaI* restriction endonucleases as

described in the methods section. After 10-16 days of puromycin selection, 68 individual puromycin resistant colonies were rescued, and propagated for an additional 18-25 days. Initial cytological analyses of the clones were performed on day 28-42 after transfection, and at the same time the primary screening for amplification was carried out by Southern blot hybridization with a plasmid probe on *EcoRI* digested DNAs purified from the individual clones.

The generation time of the primary clones was estimated by BrdU incorporation **[68]**, and was found to be 14-16 hours. Therefore, these primary clones represented cultures of 45th-70th generations after transfection.

Analysis of primary clones

Over 50% of the 68 rescued clones showed "high copy number" hybridization signals with plasmid probe indicating either the integration of multiple copies of the bacterial constructs or the amplification of the integrated exogenous DNA sequences. Representative examples of the hybridizations on the primary clones are shown in **Figure 4.** Twenty-one primary clones showed β-galactosidase expression detected with the standard LacZ staining technique.

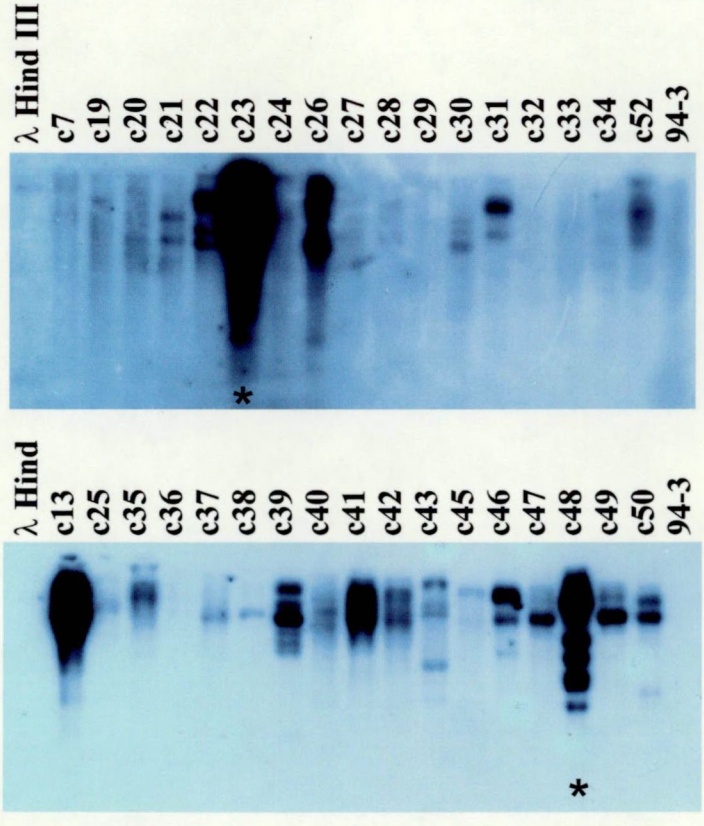

Figure 4. Southern blot hybridization on the primary clones. Representative example of Southern blot hybridization on ~10 μg *Eco RI* digested DNA of 34 primary clones with pBabePuro plasmid probe. Lanes representing clones C23 and C48 are marked with asterisks.

The first cytological analysis was made by C-banding **[74]**. Two clones (C23 and C48) were found to carry large *de novo* formed heterochromatic segments. In clone C23, 71% of the cells contained a typical "sausage" chromosome with a heterochromatic arm of 30-150MB **(Fig. 5A)**. The sizes of the *de novo* structures were estimated by comparing them to the 250-300Mb sized hamster chromosome #1. Fifteen per cent of the C48 metaphases also contained heterochromatic "sausage" chromosomes **(Fig. 5D)**. In addition to the amplified sausage chromosomes, free *de novo* heterochromatic chromosomes were observed in 5 out of 347 metaphases of the C23 clone **(Fig. 5B)**, and heterochromatic chromosomes of different size (~50-250MB) were found in 107 out of 219 metaphases of the C48 clone **(Fig. 5C)**. The heterochromatic segments in both clones were derived from the short arm of an acrocentric chromosome. Fluorescent *in situ* hybridization revealed one integration site of pBabePuro sequence in the short arm of an acrocentric chromosome in clone C23, and there were two integration sites in clone C48: one in the centromeric region of an acrocentric chromosome, and the other in a tiny metacentric chromosome. Initiation of the amplification detected by *in situ* hybridization with the D15Z probe correlated with these integration sites, and enlarged human chr #15 α-satellite (D15Z) regions were detected on the acrocentrics, and on a centric fragment of human chr #15 (not shown).

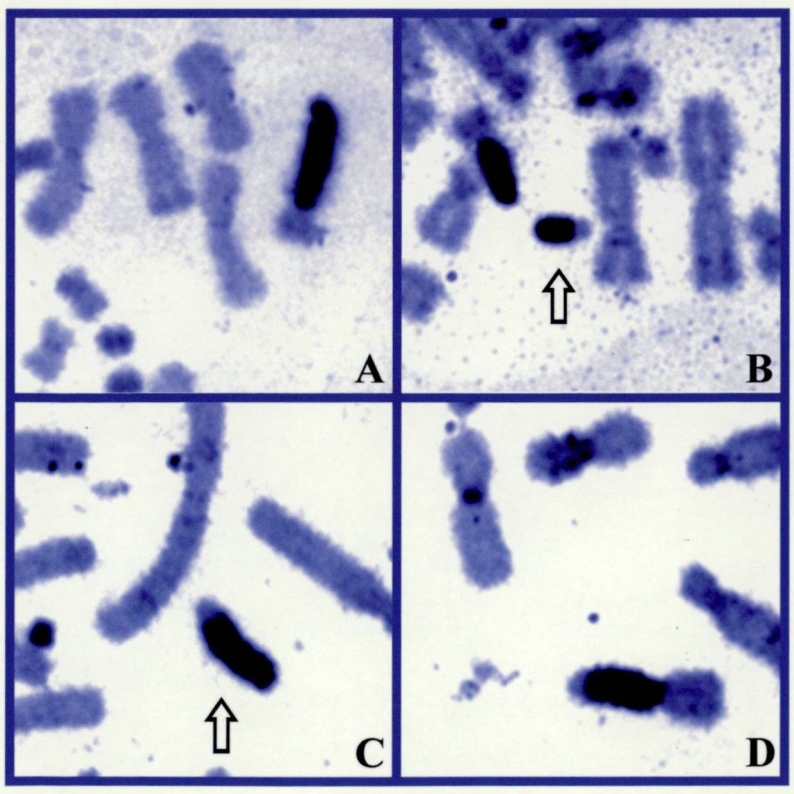

Figure 5. Heterochromatic staining on C23 and C48 primary clones. A. "Sausage" chromosome in C23. **B.** Independent heterochromatic chromosome in C23 (arrow). **C.** Independent heterochromatic chromosome in C48 metaphase (arrow). **D.** Heterochromatic "sausage" chromosome in C48.

Concurrent with the C-banding, results of *in situ* hybridizations (FISH) with human genomic, human chr #15 and chr #22 painting, and centromere specific probes showed that the large amplified regions were of human origin and contained human D15Z sequences. In clone C23, the short arm of human chr t(X;15q25) hosted the new amplified arm, while in clone C48, the "sausage" chromosome derived from the acrocentric hamster chromosome with translocated human chr #15cen/short arm region (not shown). *In situ* hybridization on C23 and C48 metaphases with chr #15cen/PuroPCR probes and immunostaining with LU851 human anti -centromere serum **[65]** confirmed that these *de novo* formed heterochromatic chromosomes carried amplified exogenous DNA **(Fig. 6D,F,H)** and functional centromeres **(Fig. 6C,E,G)**. Standard immunostaining on hSATACs with anti-centromere serum resulted in a typical centromeric label showing terminally located kinetochores. In addition, faint signals along the amplified chromosomal segments were observed **(Fig. 6C,E,G)**.

The morphology of the transitional structures found in the primary clones confirmed that the basic steps of *de novo* chromosome formations are similar to that of were observed on mouse chromosomes: i/ the integration(s) of exogenous DNA into centromeric regions **(Fig. 6B)** that induces large-scale amplifications; ii/ the formation of heterochromatic arms with an active centromere ("sausage" chromosome formation) **(Fig. 6C,D)**, iii/ breakage of the dicentric "sausage" chromosome separated the heterochromatic arm **(Fig. 6E,F)**, which finally became an independent new chromosome **(Fig. 6G,H)**.

Isolation of stable subclones

The occurrence of hSATACs in primary clones indicated that they were mitotically stable, and might allow selection of stable subclones. By single-cell subcloning we managed to establish subclones carrying stable hSATACs of different size and morphology from both primary clones **(Fig. 7)**. We regarded subclones as stable when at least 98 % of cells retained the hSATAC under selective conditions, and more than 95% of cells retained cytologically unchanged SATAC under non-selective conditions cultured for 33 days (~50 generations). In the two most extensively used subclones (C23-Z41/15 and C48-5B4-D2), under normal selective growth hSATACs remained cytologically unchanged after 274 and 731generations, respectively.

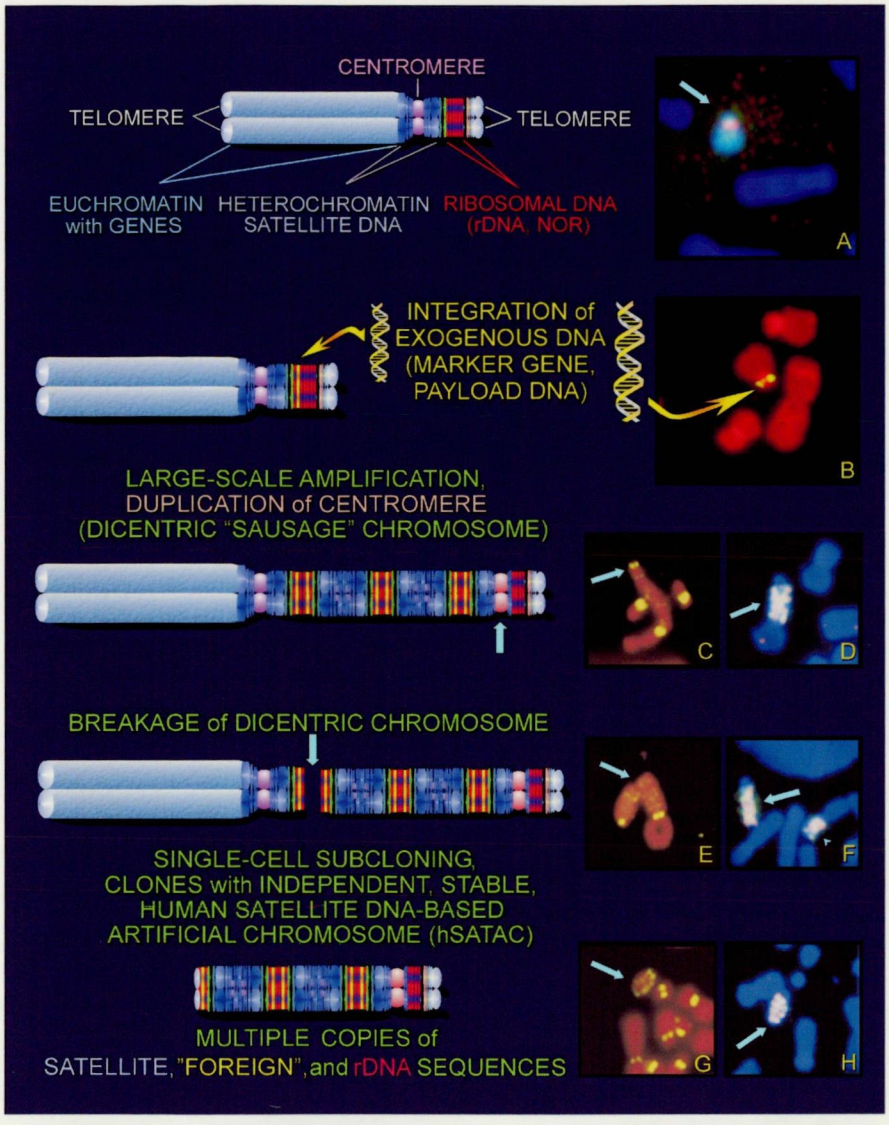

Figure 6. Generation of human satellite DNA-based artificial chromosomes *in vivo*. Subsequent steps of induced *de novo* chromosome formations are shown on **B-H**. Chromosomes on **A,D,F,H** were counterstained with DAPI (blue), overlapping green and red signals of the probes appear as pink-white. Chromosomes on **B, C,E,G** were counterstained with propidium iodide (red). **A.** *In situ* hybridization on human chromosome #15 (arrow) in a human/hamster monochromosomal hybrid cell with human chromosome specific painting probe (green) and human chromosome #15 centromere-specific alpha satellite DNA (red) probes. **B.** Yellow *in situ* hybridization signal with plasmid probe (arrow) demonstrates the integration of exogenous DNA into the centromeric/short arm region of chromosome #15 in a primary transformant cell. **C.** Large-scale amplification of the centromeric/short arm region of chromosome #15 result in the formation of a new chromosome arm with a *de novo* centromere. Arrow points to the newly formed centromere of the dicentric "sausage" chromosome. Yellow signals correspond to the centromeres visualized by indirect immunofluorescence with LU851 human anti-centromere antibody **D.** Two color *in situ* hybridization on the "sausage" chromosome with chromosome #15 α-satellite (D15Z1) (red) and exogenous DNA (green) probes. **E-F.** Breakage of the dicentric chromosome (arrows). The arrowhead on **F** indicates the formerly dicentric chromosome. **G-H.** Independent *de novo* SATACs (arrows) with anti centromere staining (**G**), and with *in situ* hybridization. Probes are the same as on **C**, **E** and **D**, **F**. (Note the faint double signals on the amplified chromosome arms (**C**, **E**) and on the SATAC (**G**) that may correspond to multiple inactive centromeres.)

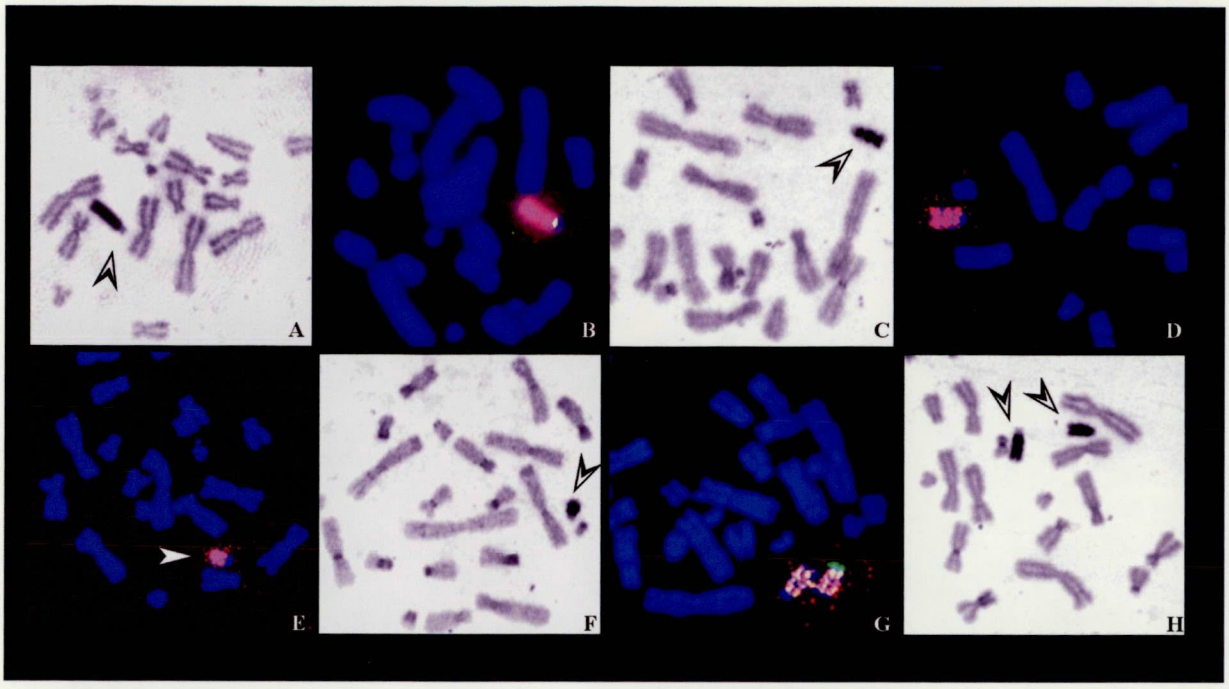

Figure 7. Chromosomes of stable hSATAC subclones with C-banding and FISH. A. Heterochromatic staining of clone C23-Z41/15. **B.** Two-color FISH with DIG-labelled human chr #15 α–satellite D15Z (red) and biotin labelled (green) human satellite III probes. Intense colocalized signals appear pink/white. **C.** Clone C48-6B5-D1 by C-banding **D.** FISH with DIG labelled (red) D15Z α-satellite probe. **E.** FISH on clone C48-6D6 with human D15Z α-satellite (red) probe. **F.** C-band staining on clone C48-6D6. **G.** FISH on clone C48-5B4-D2 with D15Z α-satellite (red), biotin labelled pBabePuro plasmid (pink/white) probes, and with biotin labelled β-satellite probe (green) that localized exclusively on the tip of one hSATAC. **H.** C-banding of clone C48-5B4-D2. Arrowheads on pictures point to the hSATACs. Chromosomes on **B,D,E,G** were counterstained with DAPI (blue).

Sequence composition of hSATACs

To identify DNA sequences present on the hSATACs other than the alpha-satellite (D15Z) and exogeneous sequences, *in situ* hybridizations were also performed with a number of different probes of known human DNA sequences specific to the short arm regions of acrocentric chromosomes **(Table 2)**. Analysis of sequence composition of hSATAC of clone C23-Z41/15 by FISH is demonstrated on **Figure 8**. All known sequences specific to the short arm of human chr #15 **(Fig. 3)** from the centromere **(**α-satellite and satellite III see **Fig. 7B, Fig. 8E)** to the telomere **(**ACR-1 and telomeric repeats see **Fig. 8K,L)** were detected on the hSATAC. Results of *in situ* hybridizations revealed that the most abundant sequences on the

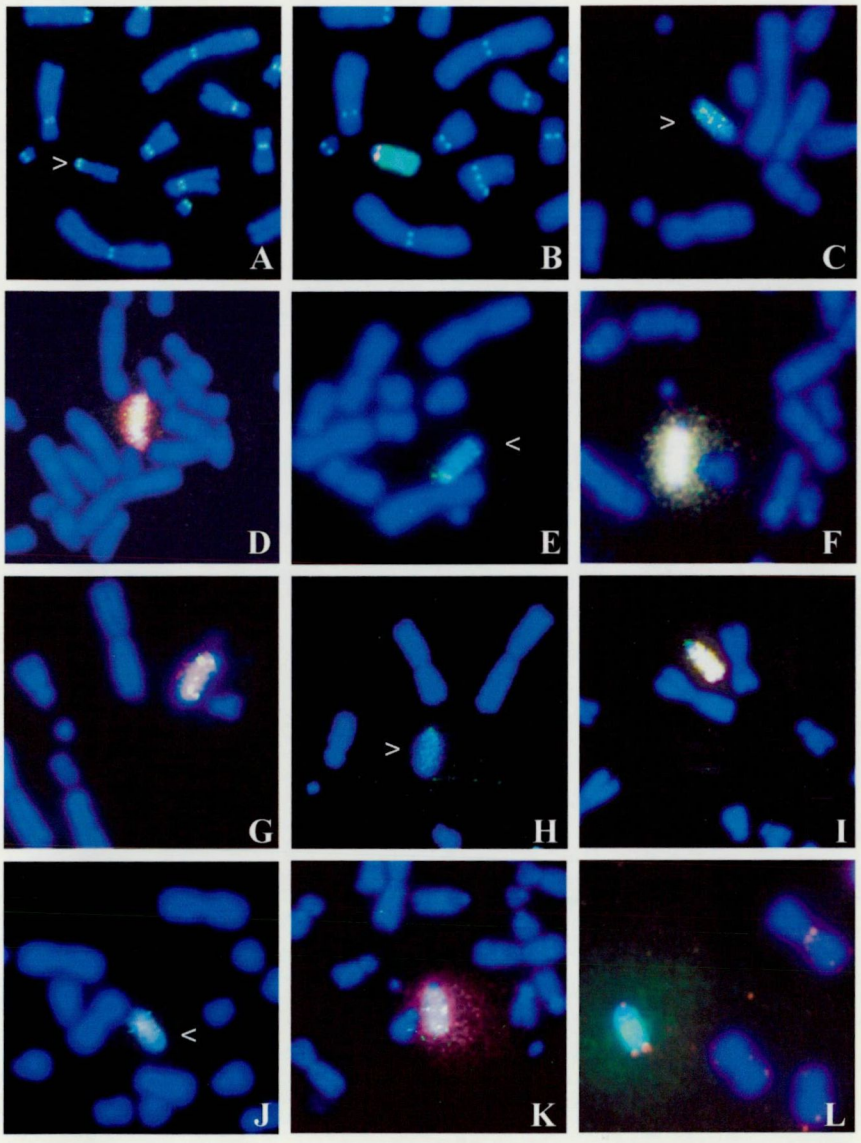

Figure 8. Sequence analysis of hSATAC of clone C23-Z41/15 by FISH. Chromosomes are counterstained with DAPI (blue), hSATAC is indicated with arrowheads. **A.** Indirect immunofluorescence staining of hSATAC with LU-851 anti-centromere serum showing terminally located centromere. **B.** FISH on the same metaphase with human chr #15 monomeric α-satellite probe (red) localized exclusively in the centromere and with plasmide probe (pBabePuro) (green) showing interspersed signal on the hSATAC. **C.** Interspersed LacZ (green) signal on the hSATAC. **D.** Colocalized (pink-white) Puro PCR (green) and D15Z (red) signals. **E.** Centromeric and interspersed localization of satellite III (D15Z1) sequences (green).**F.** Colocalized M7 rDNA (green) and D15Z (red) signals. **G.** Human rDNA externally transcribed spacer (HETS) (green) and D15Z (red) signals. **H.** Interspersed signal (green) with *Not I* centromeric repeat probe. **I.** Colocalized DIG labelled β-satellite (red) and biotin labelled D15Z (green) signals. **J.** Biotin (green) 147bp repeat probe. **K.** Two color FISH with D15Z (red) and ACR-1 (green) probes. **L.** Two color FISH with biotin labelled D15Z (green) and DIG labelled PCR generated telomere (red) probes on hSATAC, the presence of centromeric telomere sequences are characteristic to some hamster chromosomes.

hSATACs were: chr #15 α-satellite (D15Z) **(Fig. 8D,F,G,I,K,L)**, human rDNA **(Fig. 8F,G)**, satellite III (D15Z1) **(Fig. 7B, Fig. 8E)**, β-satellite (both distal and proximal) **(Fig. 7G, Fig. 8I)**, *Not I* repeat **(Fig. 8H)**, 147bp repeat **(Fig. 8J)**, and exogenous DNA **(Fig. 8B-D)**. These observations suggest that the heterochromatic nature of the newly formed chromosomes can be attributed to the presence of these amplified DNA sequences. The presence of interspersed

telomeric sequences on hSATACs (**Fig. 11B**: R1&2) indicates that these sequences also coamplified with the other sequences specific to the short arm of human acrocentric chromosomes. Following the chromosome breakage that occur during the SATAC formation, these interspersed telomeric repeats may contribute to the stability of the *de novo* chromosomes.

Table 2. Sequence analysis of hSATACs by FISH

Probes	hSATAC clones				
	C23-Z41/15	C48-6D6	C48-5B4-D2*		C48-6B5-D1
22cen D22Z1	-	-	-	-	-
15cen D15Z	+	+	+	+	+
Puro PCR	+	+	+	+	+
LacZ PCR	+	-	-	-	-
rDNA (M7)	+	+	+	+	+
rDNA HETS	+	-	-	+	-
Telomere	+	+	+	+	+
147bp repeat	+	+	+	+	+
9kb *Not I* repeat	+	+	-	+	-
ACR	+	-	-	-	-
β-satellite	+	-	-	+	-
Satellite-III	+	-	-	+	-

* C48-5B4-D2 cell line contains two hSATACs with similar morphology but different sequence composition (for an example see Fig. 7G).

Human SATACs were separated by FACS (see below) from clones of C23 and of C48 as well. Separated hSATACs were used as template in DOP-PCR **[82]**, and the product was labelled for FISH and were used in reverse painting **[83]** on hamster and human chromosomes. Results of reverse painting experiments revealed that the hSATACs of C23 clones carry no detectable amount of hamster specific sequences. SATACs of C48 clone, which derived from the acrocentric hamster chromosome with translocated human chr#15cen/short arm region hamster centromeric sequence was found **(Fig. 9)**. This sequence was cloned, sequenced and identified as hamster chr #5 centromeric repeat [Szakál, B. *et al.* unpublished].

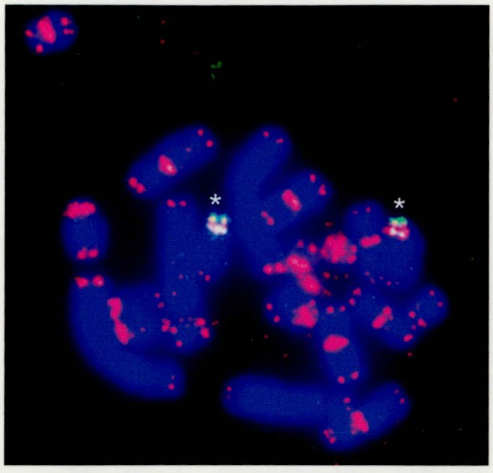

Figure 9. Sequence analysis of hSATAC of clone C48-5B4-D2 by reverse painting on Chinese-hamster spleen lymphocytes. Two color FISH with biotin labelled hSATAC specific painting probe generated by DOP-PCR from purified hSATACs (green) (asterisks), and DIG labelled PCR generated telomere (red) probes on Chinese hamster chromosomes. Chromosomes are counterstained with DAPI (blue).

Expression of the marker genes from heterochromatic hSATACs

Twenty-one out of 68 primary clones showed β–galactosidase expression by conventional lacZ staining. Clone C23 is one of the 21 β-galactosidase positive primary clones. Expression of the non-selected marker gene in hSATAC carrying subclones (C23-Z41/15) of C23 primary clone was confirmed by conventional LacZ staining **(Fig. 10B)** and by immunoblotting using monoclonal β-galactosidase antibody **(Fig. 10A)**.

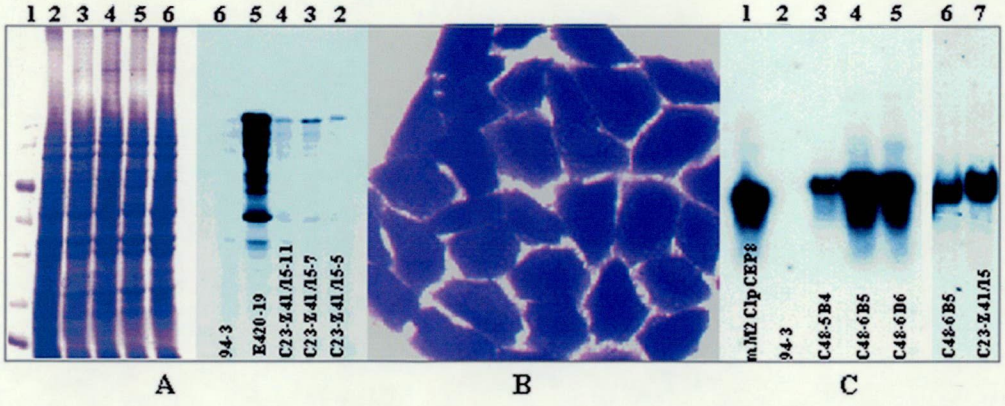

Figure 10. Expression of the marker genes from heterochromatic hSATACs. Detection of the expression of the non-selected marker gene (β-galactosidase): **A.** Immunoblotting with anti-β-galactosidase antibody. Protein marker (1), hSATAC subclones (2-4), proteins of a mouse SATAC carrying cell line containing 0.1% β-galactosidase enzyme (5), host cell line 94-3 (6). **B.** LacZ staining on hSATAC carrying cells (C23-Z41/15). Detection of the expression of the selected marker gene (puromycin acetyl-transferase): **C.** Northern blot hybridization with puromycin PCR probe. Mouse SATAC carrying cell line expressing puromycin acetyl-transferase (1), host cell line 94-3 (2), hSATAC subclones (3-7).

Expression of selectable marker gene (puromycin acetyl-transferase) was detected by resistance to 10μg/ml puromycin that is two fold of lethal concentration established for the host cell line (94-3). Z41/15 cells were also tested at high concentration of antibiotics, and up to 150μg/ml puromycin no killing effect or stunted growth were observed. Some subclones of C23 cells can tolerate 300μg/ml Puromycin concentration as well. In addition, Northern hybridizations were also used to detect the mRNA of the selectable marker gene in subclones of both C48 and C23 primary clones **(Fig. 10C)**.

Separation of hSATACs

Based on the difference between the human SATAC and the "natural" hamster chromosomes with respect to their DNA content, hSATACs can be purified efficiently by flow cytometry. The cell line containing two hSATACs, C48-5B4-D2, was used to demonstrate the separation of hSATACs by flow cytometry. Metaphase chromosomes were isolated from mitotic cells, stained with two fluorescent dyes, and passed through two laser beams: by plotting the fluorescence of each dye and chromosome, a "flow karyotype" was obtained from both the parental cell line (94-3) **(Fig. 11A)** and the hSATAC cell line **(Fig. 11B)**. The two novel clusters (R1 and R2) in the hSATAC line were assumed to represent the two hSATACs. Chromosomes from these two clusters were physically purified by the cytometer and placed on slides for FISH analysis. Two other regions (R3 and R4) were also isolated as controls. The FISH experiment using either plasmid probes or PCR generated telomere probes showed conclusively that the chromosomes isolated in regions 1 and 2 contained interspersed plasmid/telomere DNA consistent with hSATACs, while the chromosomes in regions 3 and 4 contained no plasmid DNA, and the telomere DNA signal was consistent with hamster chromosomes **(Fig. 11B**: R3-4). From R1 and R2, 62 chromosomes were analysed by *in situ* hybridization and 60 chromosomes proved to be hSATACs. Based on these results, the purity of separated hSATACs was >96%.

Purified SATACs can be used in quality control by FISH as templates to generate hybridization probes with degenerate oligonucleotide primed PCR (DOP-PCR) **[82]**. Fluorescent *in situ* hybridization with these probes to the chromosomes of host cell and SATACs can serve in monitoring the purity of separated SATACs, and to analyze the sequence composition of SATACs at the level of FISH detection **[83]**.

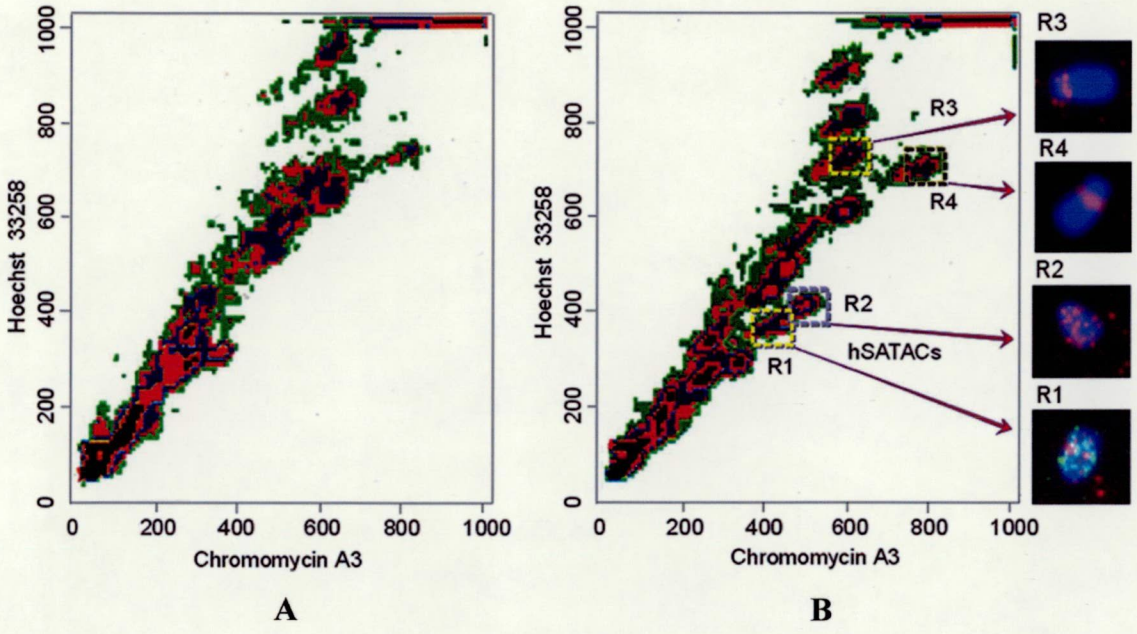

Figure 11. Separation of hSATACs with fluorescence activated cell sorter (FACS). A. FACS karyotype of the host 94-3 cell line. **B.** FACS karyotype of the C48-5B4-D2 clone containing two hSATACs. **R1-R4** represent the separated chromosome fractions, on the right side of the panel the verification of the separated chromosomes are shown by *in situ* hybridization with biotin-labelled pBabePuro probe (green) (**R1**), and DIG labelled PCR generated telomere probe (red) (**R1-R4**) on DAPI counterstained chromosomes (blue).

A stable cell line with a functional artificial chromosome has been established via direct chromosome uptake of isolated hSATAC by cultured human cells [Csonka, E. *et al.* unpublished] **(Fig. 12)**. This indicates the feasibility of direct transfer of purified hSATACs by "wrapping" them with various membrane-active compounds or by other physical and chemical methods.

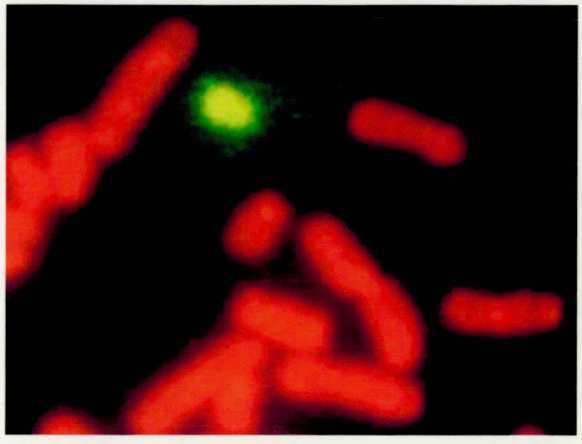

Figure 12. Chromosomes of a human fibrosarcoma cell line established by direct uptake of an isolated hSATAC. *In situ* hybridization with hSATAC specific painting probe (green) generated with degenerate oligonucleotide primed PCR from FACS purified hSATACs. Chromosomes are counterstained with propidium iodide (red).

DISCUSSION

We described the successful generation of the human counterpart of mouse satellite DNA-based artificial chromosomes **[67]**. We demonstrated that human satellite DNA-based artificial chromosomes can now be generated via induced large-scale amplification on the centromeric/short arm region of the human acrocentric chromosomes. In independent experiments, different hSATACs have been generated from the short arm of the human chromosome #15, in a human/hamster monochromosomal hybrid cell line (94-3) **[88]** and in HT-1080 human fibrosarcoma cells [Csonka, E. *et al.* unpublished] **(Fig. 13)**.

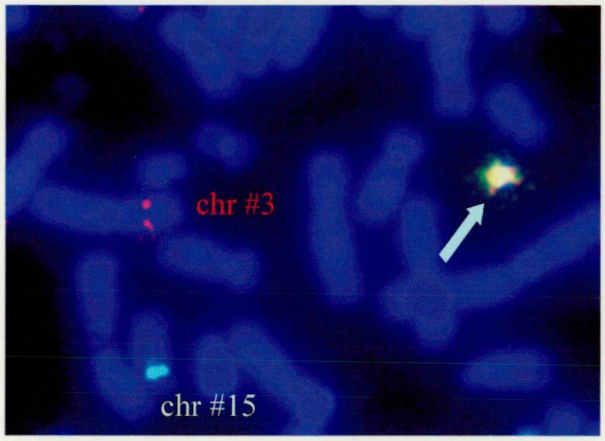

Figure 13. Human SATACs generated from sequences of the centromeric/short-arm region of human chromosome #15 in human fibrosarcoma cell. Two color *in situ* hybridization on hSATAC (arrow) carrying multiple copies of an ~120kb sequence of the short arm region of human chromosome #3. Green signals correspond to the human chromosome #15 α-satellite DNA, red signals label the ~120kb human sequence of chromosome #3. For an estimation of the amount of payload DNA carried by the SATAC, compare the intensity of the single copy hybridization signals (red) on chromosome #3 to the intensity on the SATAC (pink-white). Chromosomes are counterstained with DAPI (blue).

These results confirm that the methodology we initially used in mouse cells to create mouse satellite DNA-based artificial chromosomes is reproducible, and can be successfully applied to other species, and also in other host cells using different foreign DNA sequences. The ability to target amplifiable sites, as we did here using rDNA sequences, greatly accelerates the generation of SATACs. These mammalian artificial chromosomes are formed *in vivo* and acquire all the structural and functional elements necessary for chromosome formation from endogenous sequences. Consequently, the successful generation of *de novo* chromosomes, at the same time, is the ultimate test of the functionality of these components together with the integrated exogenous genetic material.

Human-hamster hybrid cells used in these experiments proved to be an efficient platform for *de novo* chromosome formation, and for establishing stable cell lines with

hSATACs. The short arms of the human acrocentric chromosomes are specific chromosomal regions (NOR) with well characterized DNA sequences. Amplification of these endogenous sequences of the short arms offers several advantages, which make these sites ideal for the construction of human artificial chromosomes *in vivo*: i/ Coamplified telomeric sequences can provide functioning telomeres for the *de novo* formed chromosome. ii/ The non-coding centromeric satellite DNAs lack transcription units for undesired and unknown genes and hence are safe "filling" sequences. iii/ Despite of the heterochromatic nature of the hSATACs, coamplified rDNA sequences may provide a suitable chromatin environment **[89,90]** for the expression of the integrated "foreign" gene(s). iv/ Non-coding repeated sequences specific to the short arm of the human acrocentric chromosomes can serve as useful markers to characterize newly formed chromosomes. v/ With respect to the possible use of hSATACs as vectors in human gene therapy, it is important that apart from the rRNA genes, no unsuspected or unwanted coding sequences are localized in these chromosomal regions, and it would seem unlikely that the amplification of rDNA or the known sequences identified at these chromosomal regions would cause phenotypic effects. In contrast, at least in tissue culture, amplification of rDNA presented selective advantage **[91]**. The spontaneous amplification of the short arm of the human acrocentric chromosomes is the most frequent structural polymorphism without phenotypic effect **[92]**. Follow-up studies on an unselected group of children with small supernumerary chromosomes revealed that in all 12 cases out of 14, where the inherited or *de novo* extra chromosomes were derived from the short arm of a NOR chromosome, the affected children developed normally **[93]**. Additionally there is a report of 8 healthy members of a three-generation family carrying a small supernumerary chromosome derived from the short arm of an acrocentric chromosome **[94]**. These examples indicate the existence of intrinsic cellular mechanisms that give rise to stable *de novo* chromosomes under natural circumstances without any dire consequences. The *in vivo* generation of SATACs is based upon the induction of these intrinsic mechanisms, and it is conceivable that satellite DNA-based artificial chromosomes formed by induced large-scale amplifications on the short arm of human chromosomes may become a "safe" or "low risk" vector in gene therapy.

The cell population of primary clones represents 45-70 generations of transfected cells, and the presence of hSATACs in these clones indicated that the *de novo* chromosomes appeared early after the transfection. We showed that stable subclones could be established

with different hSATACs from the primary clones, and that these *de novo* chromosomes were retained in the majority of the cells. These results indicate that hSATACs are mitotically stable, and demonstrate the feasibility and relative ease of establishing stable cell lines with hSATACs by prolonged culturing and repeated subclonings.

We have shown that mouse SATACs can be efficiently transferred by microcell mediated chromosome transfer to mammalian cells derived from different species **[95]** preserving their structural integrity, stability, and function. Further engineering of hSATACs can be performed with targeted integration of the required exogenous genetic material either in the host cells or in a homologous recombination system such as chicken DT-40 cells **[96]**. We demonstrated that hSATACs could be purified efficiently from the host cell chromosomes by flow cytometry. Purified hSATACs may allow their extensive molecular characterization and may be used for direct chromosome transfer. Recent results of direct transfer experiments with mouse SATACs indicated that the FACS procedure did not abolish the biological activity of purified artificial chromosomes. Furthermore, by pronuclear microinjection of FACS purified artificial chromosomes, successful generation of transgenic mice and germline transmission of mouse SATAC was demonstrated **[97]**.

The generation of prototype hSATACs demonstrates that the inductive approach represents an effective method for construction of human artificial chromosomes *in vivo*. In one transfection experiment, and by rounds of single-cell subcloning, different stable clones could be produced with different hSATACs. We showed that DNA sequences specific to the centromeric/short arm regions of human chromosomes are appropriate tags, and useful markers in characterizing the hSATACs generated in the NOR region of human chromosomes. For possible human gene therapy applications, construction of hSATACs on a human NOR chromosome of choice in stable monochromosomal human-rodent hybrids **[98]**, may serve as an appropriate platform.

SATACs have already complied some of the milestones considered necessary on the way of becoming a vector. These milestones are:

1. Generation of artificial chromosomes on a reproducible manner, from predictable DNA sequences **[88]**.
2. Stable transfer of SATACs into different cells and embryos preserving the structural integrity and function **[95,97]**.
3. Large-scale purification of SATACs **[85]**.

4. Generation of transgenic animals with purified SATACs and germline transmission of these mammalian artificial chromosomes **[97]**.
5. Tissue specific expression of therapeutic gene in transgenic animals (unpublished).

Considering the diversity of the requirements of different gene therapy applications it is conceivable that there will be no prevailing or ideal vector. Each application must have its optimal or perfect vector that complies best with the requirements of safety and serves the therapeutic goal at the possible highest level. For the success of gene therapy, different vector systems should be complementary rather then competitors. Artificial chromosomes will be suitable vectors in those applications where the stable, persistent or controlled gene expression can only be secured by the use of large (>100kb) therapeutic gene(s) or with gene complexes with appropriate "built-in" controlling elements. The present limitation, which applies to all existing vector systems, is the efficient delivery of the vectors into the target cells or tissues. Significant improvement of efficiency and specificity of delivery methods is one of the major tasks for the basic science that has to be accomplished in the future. Until then, in somatic gene therapy, the use of artificial chromosomes will be limited to *ex vivo* applications, particularly in combination with stem cell technology.

In general, the most imminent tasks to realize the practical applications of artificial chromosomes in *ex vivo* somatic gene therapy can be summarized below **[99]**.

1. Construction of therapeutic artificial chromosomes preferably with appropriate control of the expression of therapeutic gene(s).
2. Improved and efficient delivery of artificial chromosomes, preferably on a cell-specific manner.
3. Demonstration of therapeutic effect(s) in appropriate cellular or/and animal model systems.
4. Accomplishment of pre-clinical animal systems to test the toxicity, immunogeneicity, stability, and the long term-effect of artificial chromosomes.
5. Establishment of artificial chromosome vector-specific criteria for safety and quality control.

The successful generation of transgenic animals by pronuclear injection of purified artificial chromosomes made the germline gene therapy technically feasible, and refuelled public debate on the engineering of the human germline. Obviously, there are serious ethical and social issues related to germline gene therapy that have to be discussed. For the present,

engineering the human germline is a distant possibility of the future, and the debate on this is directed towards the principle of germline therapy. To put the principle of germline gene therapy into practice will require considerable time and substantial basic and clinical research. Impartial discussions on the implementation of germline gene therapy will be possible only when solid scientific facts will be available on the safety of a proposed protocol, and sufficient data will be accumulated from animal models on the potential risks and benefits of germline engineering. When the time comes, impartial discussions should also mean to give voice to the affected families and individuals facing incurable diseases.

ACKNOWLEDGEMENT

I am grateful to my colleagues at the Chromosome Group of the Institute of Genetics BRC, Szeged and at the Research and Development Laboratory of Chromos Molecular Systems Inc., Burnaby, Canada for their unrestricted help and support. Unpublished data cited here represents ongoing work in our laboratory, and ongoing collaborations with the Microbiology and Tumorbiology Center of Karolinska Institutet, Stockholm, Sweden. My work was supported partly by research funding from Chromos Molecular Systems Inc.

SUMMARY

We demonstrated that:

- **The "*in vivo*" construction of artificial chromosomes is an efficient and reproducible method, and this technology can be successfully applied for generation of human artificial chromosomes.**

- **Stable heterochromatic human SATACs can be generated in different mammalian cell lines from predictable endogenous DNA sequences.**

- **Human SATACs can carry >Mb size "payload" DNA.**

- **Based on the difference of the DNA content between the human SATACs and the natural chromosomes, they can be purified efficiently by flow cytometry.**

- **Human SATACs purified with dual laser beam cell sorter are suitable for quality control of artificial chromosomes by FISH (DOP-PCR, reverse painting).**

- **Purification of human SATACs makes possible their direct transfer into target cells.**

- **Human SATACs are heterochromatic, however, they provide suitable chromosomal environment for the stable expression of the integrated exogenous genetic material.**

- **Human SATACs are composed of genetically "neutral" DNA sequences.**

- **Based on the characteristics of prototype human SATACs they can serve as a start-point in developing a "safe" or "low risk" artificial chromosome-based vector system for gene therapy.**

REFERENCES

1. *Understanding Gene Therapy.* Lemoine, N.R. (ed.). BIOS Scientific Publishers Limited, Oxford, UK. (1999): pp. 1-69.
2. *Engineering the Human Germline.* Stock, G., Campbell, J. (eds.). Oxford University Press, New York. (2000): p.79.
3. Cavazzano-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J.L., Bousso, P., Deist, F.L., Fischer, A. (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science.* **288** (5466), 669-672.
4. Vile, R. (1991). The retroviral life cycle and the molecular construction of retroviral vectors. In: *Methods in Molecular Biology*, Vol. 8: *Practical Molecular Virology: Viral Vectors for Gene Expression.* (Collins, M. ed.). The Humana Press, Inc., Clifton, NJ: pp. 1-15.
5. Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L., Trono, D. (1997). Multiply attenuated lentiviral vector achieves efficient gene delivery *in vivo. Nat. Biotech.* **15,** 871-875.
6. Johnston, J., Power, C. (1999). Productive infection of human peripheral blood mononuclear cells by feline immunodeficiency virus: Implications for vector development. *J. Virol.* **73,** 2491-2498.
7. Miller, A.D., Bender, M.A., Harris, E.A.S., Kaleko, M., Gelinas, R.E. (1988). Design of retrovirus vectors for gene transfer and expression of the human β-globin gene. *J. Virol.* **62,** 4337-4345.
8. Cosset, F.-L., Russell, S.J. (1996). Targeting retrovirus entry. *Gene Therapy.* **3,** 946-956.
9. Miller, D.B., Adam, M.A., Miller, A.D. (1990). Gene transfer by retrovirus vectors occurs only in cells that are actively replicating at the time of infection. *Mol. Cell Biol.* **10,** 4239-4242.
10. Welsh, R.M., Cooper, N.R., Jensen, F.C., Oldstone, M.B.A. (1975). Human serum lyses RNA tumour viruses. *Nature.* **257,** 612-614.
11. Jahner, D., Stuhlmann, H., Stewart, C.L., Harbers, K., Lohler, J., Simon, I., Jaenisch, R. (1982). *De novo* methylation and expression of retroviral genomes during mouse embryogenesis. *Nature.* **298,** 623-628.

12. Zhang, W.W. (1997). Review: Adenovirus vectors: development and application. *Exp. Opin. Invest. Drugs.* **6,** 1419-1457.

13. Bett, A.J., Haddara, W., Prevec, L., Graham, F.L. (1994). An efficient and flexible system for construction of adenovirus vectors with inserts or deletions in early regions 1 and 3. *Proc. Natl. Acad. Sci. USA.* **91,** 8802-8806.

14. Bett, A.J., Prevec, L., Graham, F.L. (1993). Packaging capacity and stability of human adenovirus type 5 vectors. *J. Virol.* **67,** 5911-5921.

15. Graham, F.L., Smiley, J., Russell, W.C., Nairn, R. (1977). Characterization of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* **36,** 59-74.

16. Yang, Y., Nunes, F.A., Berencsi, K., Gönczöl, É., Engelhardt, J.F., Wilson, J.M. (1994). Inactivation of E2A in recombinant adenoviruses improves the prospect for gene therapy in cystic fibrosis. *Nat. Genet.* **7,** 362-369.

17. Berkner, K.L. (1988). Development of adenovirus vectors for the expression of heterologous genes. *BioTechniques.* **6,** 616-629.

18. Shenk, T. (1996). Adenoviridae: The viruses and their replication. In: *Fields Virology.* Vol. 2. 3rd Edn. (Fields, B.N., Knipe, D.M. and Howley, P.M. eds.). Lippincott-Raven, Philadelphia, PA. pp. 2111-2148.

19. Grubb, B.R., Pickles, R.J., Ye, H., Yankaskas, J.R., Vick, Rn., Engelhardt, J.F., Wilson, J.M., Johnson, J.G., Boucher, R.C. (1994). Inefficient gene transfer by adenovirus vector to cystic fibrosis airway epithelia of mice and humans. *Nature.* **371,** 802-806.

20. Yang, Y., Jooss, K.U., Su, Q., Ertl, H.C.J., Wilson, J.M. (1996). Immune response to viral antigens versus transgene product in the elimination of recombinant adenovirus-infected hepatocytes *in vivo*. *Gene Ther.* **3,** 137-144.

21. Preliminary Findings Reported on the Death of Jesse Gelsinger. (December 2, 1999.) *The Institute for Human Gene Therapy.* University of Pennsylvania Health System. (IHGT Home Page)

22. Verma, I.M., Somia, N. (1997). Gene therapy-promises, problems and prospects. *Nature.* **389,** 239-242.

23. Muzyczka, N. (1992). Use of adeno-associated virus as a general transduction vector for mammalian cells. *Curr. Top. Microbiol. Immunol.* **158,** 97-129.

24. Kremer, E.J., Perricaudet, M. (1995). Adenovirus and adeno-associated virus-mediated gene transfer. *British Med. Bull.* **51,** 31-44.

25. Huard, J., Krisky, D., Oligino, T., Marconi, P., Day, C.S., Watkins, S.C., Glorioso, J.C. (1997). Gene transfer to muscle using herpes simplex virus-based vectors. *Neuromusc. Disord.* **7,** 299-313.

26. Brehm, M., Samaniego, L.A., Bonneau, R.H., DeLuca, N.A., Tevethia, S.S. (1999). Immunogenicity if herpes simplex virus type 1mutants containing deletions in one or more of the α-genes: ICP4, ICP27, ICP22, and ICP0. *Virology.* **256,** 258-269.

27. Frenkel, N., Singer, O., Kwong, A.D. (1994). Minireview: The herpes simplex virus amplicon – a versatile defective virus vector. *Gene Ther.* **1,** S40-S46.

28. Hermens, W.T., Verhaagen, J. (1998). Viral vectors, tools for gene transfer in the nervous system. *Progress Neurobiol.* **55,** 399-432.

29. Ponnazhagan, S., Weigel, K.A., Raikwar, S.P., Mukherjee, P., Yoder, M.C., Srivastava, A. (1998). Recombinant human parvovirus B19 vectors: erythroid cell-specific delivery and expression of transduced genes. *J. Virol.* **72,** 5224-5230.

30. Wang, S., Vos, J.M. (1996). A hybrid herpesvirus infection vector based on Epstein-Barr virus and herpes simplex virus type I for gene transfer into human cells *in vitro* and *in vivo*. *J. Virol.* **70,** 8422-8430.

31. Felgner, P.L., Barenholz, Y., Behr, J.-P., Cheng, S.H., Cullis, P., Huang, L., Jessee, J.A., Seymour, L., Szoka, F., Thierry, A.R., Wagner, E., Wu, G. (1997). Nomenclature for synthetic gene delivery systems. *Hum. Gene Ther.* **8,** 511-512.

32. Felgner, P.L., Gadek, T.R., Holm, M., Roman, R., Chan, H.W., Wenz, M., Northrop, J.P., Ringold, G.M., Danielsen, M. (1987). Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc. Natl. Acad. Sci. USA.* **84,** 7413-7417.

33. Leventis, R., Silvius, J.R. (1990). Interactions of mammalian cells with lipid dispersions containing novel metabolizable cationic amphiphiles. *Biochim. Biophys. Acta.* **1023,** 124-132.

34. Wolfert, M.A., Seymour, L.W. (1996). Atomic force microscopic analysis of the influence of the molecular weight of poly(L)lysine on the size of polyelectrolyte complexes formed with DNA. *Gene Ther.* **3,** 269-273.

35. Haensler, J., Szoka, F.C. Jr. (1993). Polyamidoamine cascade polymers mediate efficient transfection of cells in culture. *Bioconjugate Chem.* **4,** 372-379.

36. Boussif, O., Lezoualch, F., Zanta, M.A., Mergny, M.D., Scherman, D., Demeneix, B., Behr, J.-P. (1995). A versatile vector for gene and oligonucleotide transfer into cells in culture and *in vivo*: polyethylenimine. *Proc. Natl. Acad. Sci. USA.* **92,** 7297-7301.

37. Abdallah, B., Hassan, A., Benoist, C., Goula, D., Behr, J.-P., Demeneix, B.A. (1996). A powerful nonviral vector for *in vivo* gene transfer into the adult mammalian brain: polyethylenimine. *Hum. Gene Ther.* **7,** 1947-1954.

38. Mathiesen, I. (1999). Electropermeabilization of skeletal muscle enhances gene transfer *in vivo. Gene Ther.* **6,** 508- 514.

39. Rizzuto, G., Cappalletti, M., Maione, D. *et al.* (1999). Efficient and regulated erythropoietin production by naked DNA injection and muscle electroporation. *Proc. Natl. Acad. Sci. USA.* **96,** 6417-6422.

40. Potter, H. (1988). Electroporation in biology: methods, applications and instrumentation. *Anal. Biochem.* **174,** 361-373.

41. Sanford, J.C., Klein, T.M., Wolf, E.D., Allen, N. (1987). Delivery of substances into cells and tissues using a particle bombardment process. *Part. Sci. Technol.* **5,** 27-37.

42. Williams, R.S., Johnston, S.A., Riedy, M., DeVit, M.J., McElligott, S.G., Sanford, J.C. (1991). Introduction of foreign genes into tissues of living mice by DNA-coated microprojectiles. *Proc. Natl. Acad. Sci. USA.* **88,** 2726-2730.

43. Shark, K.B., Smith, F.D., Harpending, P.R., Rasmussen, J.L., Sanford, J.C. (1991). Biolistic transformation of a prokaryote, *Bacillus megaterium. Appl. Environ. Microbiol.* **57,** 480-485.

44. Johnston, S.A., Anziano, P.Q., Shark, K., Sanford, J.C., Butow, R.A. (1988). Mitochondrial transformation in yeast by bombardment with microprojectiles. *Science.* **240,** 1538-1541.

45. Daniell, H., Krishnan, M., McFadden, B.F. (1991). Transient expression of beta-glucuronidase in different cellular compartments following Biolistic delivery of foreign DNA into wheat leaves and calli. *Plant Cell Rep.* **9,** 615-619.

46. Plank, C., Scherer, F., Schillinger, U., Anton, M. (2000). Magnetofection: Enhancement and localization of gene delivery with magnetic particles under the influence of a magnetic field. *J. Gene Med.* **2,**(5 Suppl.) 24. Or13.

47. Mitani, K., Caskey, C.T. (1993). Delivering therapeutic genes-matching approach and application. *Trends Biotech.* **11,** 162-166.

48. Hadlaczky, Gy. (1991). The mammalian minichromosome. *D.Sci. Thesis.* Institute of Genetics, Biological Research Center, Hungarian Academy of Sciences. p.23-24.

49. Murray, A.W., Szostak, J.W. (1983). Construction of artificial chromosomes in yeast. *Nature.* **305,**(5931) 189-193.

50. Hahnenberger, K.M., Baum, M.P., Polizzi, C.M., Carbon, J., Clarke, L. (1989). Construction of functional artificial minichromosomes in the fission yeast *Schizosaccharomyces pombe. Proc. Natl. Acad. Sci. USA.* **86,**(2) 577-581.

51. Farr, C., Fantes, J., Goodfellow, P., Cooke, H. (1991). Functional reintroduction of human telomeres into mammalian cells. *Proc. Natl. Acad. Sci. USA.* **88,** 7006-7010.

52. Farr, C. J., Stevanovic, M., Thomson, E. J., Goodfellow, P. N., Cooke, H. J. (1992). Telomere-associated chromosome fragmentation: applications in genome manipulation and analysis. *Nat. Genet.* **2,** 275-282.

53. Itzhaki, J.E., Barnett, M.A., MacCarthy, A.B., Buckle, V.J., Brown, W.R., Porter, A.C. (1992). Targeted breakage of a human chromosome mediated by cloned human telomeric DNA. *Nat. Genet.* **2,** 283-287.

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

55. Brown, K.E., Barnett, M.A., Burgtorf, C., Shaw, P., Buckle, V.J., Brown, W.R. (1994). Dissecting the centromere of the human Y chromosome with cloned telomeric DNA. *Hum. Mol. Genet.* **3,**(8):1227-1237.

56. Farr, C.J., Bayne, R.A., Kipling, D., Mills, W., Critcher, R., Cooke, H.J. (1995). Generation of a human X-derived minichromosome using telomere-associated chromosome fragmentation. *EMBO J.* **14,** 5444-5454.

57. Heller, R., Brown, K.E., Burgtorf, C., Brown, W.R. (1996). Mini-chromosomes derived from the human Y chromosome by telomere directed chromosome breakage. *Proc. Natl. Acad. Sci. USA.* **93,** 7125-7130.

58. Kuroiwa, Y., Shinohara, T., Notsu, T., Tomizuka, K., Yoshida, H., Takeda, S., Oshimura, M., Ishida, I. (1998). Efficient modification of a human chromosome by telomere-directed truncation in high homologous recombination-proficient chicken DT40 cells. *Nucleic Acids Res.* **26,** 3447-3448.

59. Mills, W., Critcher, R., Lee, C., 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,** 751-761.

60. Kuroiwa, Y., Tomizuka, K., Shinohara, T., Kazuki, Y., Yoshida, H., Ohguma, A., Yamamoto, T., Tanaka, S., Oshimura, M., Ishida, I. (2000). Manipulation of human minichromosomes to carry greater than megabase-sized chromosome inserts. *Nature Biotech.* **18,** 1086-1090.

61. Carine, K., Solus, J., Waltzer, E., Manch-Citron, J., Hamkalo, B.A., Scheffler, I.E. (1986). Chinese hamster cells with a minichromosome containing the centromere region of human chromosome 1. *Somat. Cell Mol. Genet.* **12,** 479-491.

62. Raimondi, E., Ferretti, L., Young, B.D., Sgaramella, V., DeCarli, L. (1991). The origin of a morphologically unidentifiable human supernumerary minichromosome traced through sorting, molecular cloning, and *in situ* hybridisation. *J. Med. Genet.* **28,** 92-96.

63. Haaf, T., Sumner, A.T., Kohler, J., Willard, H.F., Schmid, M. (1992). A microchromosome derived from chromosome 11 in a patient with the CREST syndrome of scleroderma. *Cytogenet. Cell Genet.* **60,** 12-17.

64. Guiducci, C., Ascenzioni, F., Auriche, C., Piccolella, E., Guerrini, A.M., Donini, P. (1999). Use of a human minichromosome as a cloning and expression vector for mammalian cells. *Hum. Mol. Genet*. **8,** 1417-1424.

65. Hadlaczky, Gy., Praznovszky, T., Cserpán, I., Kereső, J., Péterfy, M., Kelemen, I., Atalay, E., Szeles, A., Szelei, J., Tubak, V. (1991). Centromere formation in mouse cells cotransformed with human DNA and a dominant marker gene. *Proc. Natl. Acad. Sci. USA.* **88,** 8106-8110.

66. Praznovszky, T. Kereső, J. Tubak, V., Cserpán, I., Fátyol, K., Hadlaczky, Gy. (1991). *De novo* chromosome formation in rodent cells. *Proc. Natl. Acad. Sci. USA.* **88,** 11042-11046.

67. Kereső, J., Praznovszky, T., Cserpán, I., Fodor, K., Katona, R., Csonka, E., Fátyol, K., Holló, Gy., Szeles, A., Ross, A.R., Sumner, A.T., Szalay, A.A., Hadlaczky, Gy. (1996). *De novo* chromosome formations by large-scale amplification of the centromeric region of mouse chromosomes. *Chromosome Res.* **4,** 226-239.

68. Holló, Gy., Kereső, J., Praznovszky, T., Cserpán, I., Fodor, K., Katona, R., Csonka, E., Fátyol, K., Szeles, A., Szalay, A.A., Hadlaczky, Gy. (1996). Evidence for a megareplicon covering megabases of centromeric chromosome segments. *Chromosome Res.* **4,** 240-247.

69. Ledbetter, S.A., Schwartz, C.E., Davies, K.E., Ledbetter, D.H. (1991). New somatic cell hybrids for physical mapping in distal Xq and the fragile X region. *Am. J. Med. Genet.* **38,** 418-420.

70. Chen, C, Okayama, H. (1987). High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7,** 2745-2752.

71. Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). Molecular cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

72. Morgenstern, J.P., Land, H. (1990). Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucl. Acids Res.* **18,** 3587-3596.

73. Worton, R.G., Sutherland, J., Sylvester, J.E., Willard, H.F., Bodrug, S., Dube, I., Duff, C., Kean, V., Ray, P.N., Schmickel, R.D. (1988). Human ribosomal RNA genes: orientation of the tandem array and conservation of the 5' end. *Science.* **239,** 64-68.

74. Sumner, A.T. (1972). A simple technique for demonstrating centromeric heterochromatin. *Exp. Cell Res.* **75,** 304-306.

75. Gonzalez, I.L., Sylvester, J.E. (1995). Complete sequence of the 43-kb human ribosomal DNA repeat: analysis of the intergenic spacer. *Genomics.* **27,** 320-328.

76. Gogel, E., Langst, G., Grummt, I., Knuckle, E., Grummt, F. (1996). Mapping of replication initiation sites in the mouse ribosomal gene cluster. *Chromosoma.* **104,** 511-518.

77. Thoraval, D., Asakawa, J., Kodaira, M., Chang, C., Radany, E., Kuick, R., Lamb, B., Richardson, B., Neel, J.V., Glover, T., Hanash, S. (1996). A methylated human 9-kb repetitive sequence on acrocentric chromosomes is homologous to a subtelomeric repeat in chimpanzees. *Proc Natl. Acad. Sci. USA.* **93,** 4442-4447.

78. Greig, G.M., Willard, H.F. (1992). α-satellite DNA: Characterization and localization of two subfamilies from the distal and proximal short arms of the human acrocentric chromosomes. *Genomics.* **12,** 573-580.

79. Vissel, B., Nagy, A., Choo, K.H.A. (1992). A satellite III sequence shared by human chromosomes 13, 14, and 21 that is contiguous with alpha satellite DNA. *Cytogenet. Cell Genet.* **61,** 81-86.

80. Sakai, K., Ohta, T., Minoshima, S., Kudoh, J., Wang, Y., de Jong, P.J., Shimizu, N. (1995). Human ribosomal RNA gene cluster: identification of the proximal end containing a novel tandem repeat sequence. *Genomics.* **26,** 521-526.

81. Ijdo, J.W., Wells, R.A., Baldini, A., Reeders, S. T. (1991). Improved telomere detection using a telomere repeat probe $(TTAGGG)_n$ generated by PCR. *Nucl. Acids Res.* **19,** 4780.

82. Telenius, H., Carter, N.P., Bebb, C.E., Nordenskjöld, M., Ponder, B.A.J., Tunnacliffe, A. (1992). Degenerate oligonucleotide-primed PCR: General amplification of target DNA by a single degenerate primer. *Genomics.* **13,** 718-725.

83. Telenius, H., Pelmear, A.H., Tunnacliffe, A., Carter, N.P., Behmel, A., Ferguson-Smith, M.A., Nordenskjöld, M., Pfragner, R., Ponder, B.A. (1992). Cytogenetic analysis by chromosome painting using DOP-PCR amplified flow-sorted chromosomes. *Genes Chromosomes Cancer.* **4,** 257-263.

84. Nicol, I., Jeppesen, P. (1994). Human autoimmune sera recognize a conserved 26 kD protein associated with mammalian heterochromatin that is homologous to heterochromatin protein 1 of Drosophila. *Chromosome Res.* **2,** 245-253.

85. De Jong, G., Telenius, A.H., Telenius, H., Perez, C.F., Drayer, J.I., Hadlaczky, Gy. (1999). Mammalian artificial chromosome pilot facility: Large-scale isolation of functional satellite DNA-based artificial chromosomes. *Cytometry.* **35,** 129-133.

86. Raimondi, E., Balzaretti, M., Moralli, D., Vagnarelli, P., Tredici, F., Bensi, M., De Carli, L. (1996). Gene targeting to the centromeric DNA of a human minichromosome. *Hum. Gene Ther.* **7,** 1103-1109.

87. Waye, J.S., Willard, H.F. (1989). Human beta satellite DNA: genomic organization and sequence definition of a class of highly repetitive tandem DNA. *Proc. Natl. Acad. Sci. USA.* **86,** 6250-6254.

88. Csonka, E., Cserpán, I., Fodor, K., Holló, Gy., Katona, R., Kereső, J., Praznovszky, T., Szakál, B., Telenius, A., deJong, G., Udvardy, A., Hadlaczky, Gy. (2000). Novel Generation of Human Satellite DNA-based Artificial Chromosomes in Mammalian Cells. *J Cell Sci.* **113,** 3207-3216.

89. Lucchini, R., Sogo, J.M. (1992). Different chromatin structures along the spacers flanking active and inactive *Xenopus* rRNA genes. *Mol. Cell. Biol.* **12,** 4288-4296.

90. Karpen, G.H., Schaefer, J.E., Laird, C.D. (1988). A *Drosophila* rRNA gene located in euchromatin is active in transcription and nucleolus formation. *Genes Dev.* **2,** 1745-1763.

91. Roberts, C. Brasch, J., Tattersall, M.H. (1987). Ribosomal RNA gene amplification: a selective advantage in tissue culture. *Cancer Genet. Cytogenet.* **29,** 119-127.

92. Conte, R.A., Kleyman, S M., Laundon, C., Verma, R. S. (1997). Characterization of two extreme variants involving the short arm of chromosome 22: are they identical. *Ann. Genet.* **40,** 145-149.

93. Gravholt, C.H., Friedrich, U. (1995). Molecular cytogenetic study of supernumerary marker chromosomes in an unselected group of children. *Am. J. Med. Genet.***13,** 106-111.

94. Fu, S., Fu, H., Xiao, H., Song, X., Chen, J., Gao, C., Qiu, H., Cheng, Z. (1992). Molecular cytogenetic study of an extra small chromosome. *I Chuan Hsueh Pao.* **19,** 294-297.

95. Telenius, H., Szeles, A., Kereső, J., Csonka, E., Praznovszky, T., Imreh, S., Maxwell, A., Perez, C.F., Drayer, J.I., Hadlaczky, Gy. (1999). Stability of a functional murine satellite DNA-based artificial chromosome across mammalian species. *Chromosome Res.* **7,** 3-7.

96. Dieken, E.S., Epner, E.M., Fiering, S., Fournier, R.E.K., Groudine, M. (1996). Efficient modification of human chromosomal alleles using recombination-proficient chicken/human microcell hybrids. *Nature Genet.* **12,** 174-182.

97. Co, D.O., Borowski, A.H., Leung, J.D., van der Kaa, J., Hengst, S., Platenburg, G., Pieper, F.R., Perez, C.F., Jirik, F.R., 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,** 183-191.

98. Cuthbert, A P., Trott, D.A., Ekong, R.M., Jezzard, S., England, N.L., Themis, M., Todd, C.M., Newbold, R.F. (1995). Construction and characterization of a highly stable human:rodent monochromosomal hybrid panel for genetic complementation and genome mapping studies. *Cytogen. Cell Genet.* **71,** 68-76.

99. Hadlaczky, Gy. (2001). Satellite DNA-based Artificial Chromosomes for the Use in Gene Therapy. *Current Opinion in Molecular Therapeutics.* **3**(2), 125-132.

LIST OF PUBLICATIONS

Articles:

Kereső, J., Praznovszky, T., Cserpán, I., Fodor, K., Katona, R., **Csonka, E.**, Fátyol, K., Holló, Gy., Szeles, A., Ross, A.R., Sumner, A.T., Szalay, A.A., Hadlaczky, Gy. (1996). *De novo* chromosome formations by large-scale amplification of the centromeric region of mouse chromosomes. *Chromosome Res.* **4,** 226-239.

Holló, Gy., Kereső, J., Praznovszky, T., Cserpán, I., Fodor, K., Katona, R., **Csonka, E.**, Fátyol, K., Szeles, A., Szalay, A.A., Hadlaczky, Gy. (1996). Evidence for a megareplicon covering megabases of centromeric chromosome segments. *Chromosome Res.* **4,** 240-247.

Telenius, H., Szeles, A., Kereső, J., **Csonka, E.**, Praznovszky, T., Imreh, S., Maxwell, A., Perez, C.F., Drayer, J.I., Hadlaczky, Gy. (1999). Stability of a functional murine satellite DNA-based artificial chromosome across mammalian species. *Chromosome Res.* **7,** 3-7.

Csonka, E., Cserpán, I., Fodor, K., Holló, Gy., Katona, R., Kereső, J., Praznovszky, T., Szakál, B., Telenius, A., deJong, G., Udvardy, A., Hadlaczky, Gy. (2000). Novel Generation of Human Satellite DNA-based Artificial Chromosomes in Mammalian Cells. *J Cell Sci.* **113,** 3207-3216.

Abstracts:

Csonka, E., Cserpán, I., Fodor, K., Holló, Gy., Katona, R., Kereső, J., Praznovszky, T., Szakál, B., Telenius, A., deJong, G., Udvardy, A., Hadlaczky, Gy. (2000). Human Satellite DNA-based Artificial Chromosomes. *Eur. J. Hum. Genet.* **8,**(Suppl.) 40.

Csonka, E., Cserpán, I., Fodor, K., Holló, Gy., Katona, R., Kereső, J., Praznovszky, T., Szakál, B., Telenius, A., deJong, G., Udvardy, A., Hadlaczky, Gy. (2000). Human Satellite DNA-based Artificial Chromosomes. *J. Gene Medicine.* **2,**(5,Suppl.) 116.

Katona, R., Cserpán, I., **Csonka, E.**, Hadlaczky, Gy. (2000). Anti-HIV Ribozyme Transgenic Mice. *J. Gene Medicine.* **2,**(5,Suppl.) 89.