POLYNUCLEOTIDE

FIELD OF THE INVENTION

The present invention relates to a polynucleotide comprising a ubiquitous chromatin opening element (UCOE) that does not occur in nature. The present invention also relates to a vector comprising the polynucleotide sequence, a host cell comprising the vector and use of the polynucleotide, vector or host cell in therapy, or for in vitro protein expression applications.

BACKGROUND OF THE INVENTION

The current model of chromatin structure in higher eukaryotes postulates that genes are organised in "domains"[1,2]. Chromatin domains are envisaged to exist in either a condensed, "closed", transcriptionally silent state, or in a de-condensed, "open" and transcriptionally competent configuration. The establishment of an open chromatin structure characterised by increased DNaseI sensitivity, DNA hypomethylation and histone hyperacetylation, is considered a pre-requisite to the commencement of gene expression.

The open and closed nature of chromatin regions is reflected in the behaviour of transgenes that are randomly integrated into the host cell genome. Identical constructs give different patterns of tissue-specific and development stage-specific expression when integrated at different locations in the mouse genome[3-5]. A variegated expression pattern within a given transgenic mouse tissue, known as position effect variegation (PEV), is also frequently observed[6]. When exogenous genes are integrated into the chromosome of mammalian cells cultures *in vitro*, many of the integration events result in rapid silencing of the transgene and the remainder give large variability in expression levels[7,8]. These position effects render transgene expression inefficient, with implication for both basic research and biotechnology applications.

The chromatin domain model of gene organisation suggests that genetic control elements that are able to establish and maintain a transcriptionally competent open chromatin structure should be associated with active regions of the genome.

Locus Control Regions (LCRs) are a class of transcriptional regulatory elements with long-range chromatin remodelling capability. LCRs are functionally defined in transgenic mice by their ability to confer site-of-integration independent, transgene copy number-dependent, physiological levels of expression on a gene linked in *cis*,, especially single copy transgenes[9,10]. Crucially, such expression is tissue-specific. LCRs are able to obstruct the spread of heterochromatin, prevent PEV[6] and consist of a series of Dnase I hypersensitive (HS) sites which can be located either 5' or 3' of the genes that they regulate[10].

LCRs appear to be comprised of two separate, although not necessary independent components. First, the establishment of an open chromatin domain, and second a dominant transcriptional activation capacity to confer transgene copy number dependent expression[9]. The molecular mechanisms by which LCRs exert their function remain a point of contention[2,11-13].

The generation of cultured mammalian cell lines producing high levels of a therapeutic protein product is a major developing industry. Chromatin position effects make it a difficult, time consuming and expensive process. The most commonly used approach to the production of such mammalian "cell factories" relies on gene amplification induced by a combination of a drug resistance gene (e.g., DHFR, glutamine synthetase (Kaufman, 1990, Methods Enzymol., 185:537-566, which is incorporated herein by reference)) and the mainatenance of stringent selective pressure. The use of vectors containing LCRs from highlyexpressed gene domains, using cells derived from the appropriate tissue, greatly simplifies the procedure (Needham *et al.*, 1992, Nucleic Acids Res., 20:997-1003; Needham *et al.*, 1995, Protein Expr. Purif., 6:124-131, each of which is incorporated herein by reference).

However, the tissue-specificity of LCRs, although useful in some circumstances, is also a major limitation for many applications, for instance where no LCR is known for the tissue in which expression is required, or where expression in many, or all, tissues is required.

Our co-pending patent application PCT/GB99/02357 (WO 00/05393), incorporated by reference herein, describes elements that are responsible for establishing an open chromatin structure across a locus that consists exclusively of ubiquitously expressed, housekeeping genes. These elements are not derived from an LCR. The invention provides a polynucleotide comprising a ubiquitous chromatin opening element (UCOE) which opens chromatin or maintains chromatin in an open state and facilitates reproducible expression of an operably-linked gene in cells of at least two different tissue types, wherein the polynucleotide is not derived from a locus control region.

Methylation-free CpG islands are well-known in the art (Bird *et al* (1985) Cell 40: 91-99, Tazi and Bird (1990) Cell 60: 909-920) and may be defined as CpG-rich regions of DNA with above average (>60%) content of CpG di-nucleotides where the cytosine residues are not methylated and which extend over the 5' ends of two closely spaced (0.1-3 kb) divergently transcribed genes. These regions of DNA remain unmethylated in all tissues throughout development (Wise and Pravtcheva (1999) Genomics 60: 258-271). They are associated with the 5' ends of all ubiquitously expressed genes, as well as an estimated 40% of genes showing a tissue restricted expression profile[16,17] and are known to be localised regions of active chromatin[18].

An 'extended' methylation-free CpG island is a methylation-free CpG island that extends across a region encompassing more than one transcriptional start site and/or extends for more than 300bp and preferably more than 500bp. The borders of the extended methylation-free CpG island are functionally defined through the use of PCR over the region in combination with restriction endonuclease enzymes whose

ability to digest (cut) DNA at their recognition sequence is sensitive to the methylation status of any CpG residues that are present. One such enzyme is HpaII, which recognises and digests at the site CCGG, which is commonly found within CpG islands, but only if the central CG residues are *not* methylated. Therefore, PCR conducted with HpaII-digested DNA and over a region harbouring HpaII sites, does *not* give an amplification product due to HpaII digestion if the DNA is *unmethylated*. The PCR will *only* give an amplified product if the DNA is *methylated*. Therefore, beyond the methylation-free region HpaII will not digest the DNA a PCR amplified product will be observed thereby defining the boundaries of the "extended methylation-free CpG island".

We have demonstrated that regions spanning methylation-free CpG islands encompassing dual, divergently transcribed promoters from the human TATA binding protein (*TBP*)/proteosome component-B1 (*PSMBI*) and heterogenous nuclear ribonucleoprotein A2/B1 (*hnRNPA2*)/heterochromatin protein 1Hsγ (*HP1 $^{Hs\gamma}$*) gene loci give reproducible, physiological levels of gene expression and that they are able to prevent a variegated expression pattern and silencing that normally occurs with transgene integration within centromeric heterochomatin.

We have shown that methylation-free CpG islands associated with actively transcribing promoters possess the ability to remodel chromatin and are thus thought to be a prime determinant in establishing and maintaining an open domain at housekeeping gene loci.

UCOE's confer an increased proportion of productive gene delivery events with improvements in the level and stability of transgene expression. This has important research and biotechnological applications including the generation of transgenic animals and recombinant protein products in cultured cells. We have shown beneficial effects of UCOEs on expression of the CMV-EGFP reporter construct and with the secreted, pharmaceutically valuable protein erythropoietin. The properties of UCOEs also suggest utility in gene therapy, the effectiveness of which is often

limited by a low frequency of productive gene delivery events and an inadequate level and duration of expression[45].

Given these significant implications and wide ranging applications, there is a desire to further optimise transgene expression levels and achieve improved stability of gene expression over a prolonged period of culture.

One particular need is to overcome the directional bias observed in some naturally-occurring UCOEs. Although UCOEs confer position-independent transcriptional enhancement on operably-linked promoters, this is, to some extent, orientation-dependent (i.e. the UCOE is significantly more effective in one orientation than the other). In some circumstances, such as an expression vector comprising two expression units transcribed divergently with a UCOE situated between them, there is an advantage in being able to obtain balanced, high-level expression from both promoters, which may not be possible with a natural UCOE. There is therefore a need for artificially-constructed UCOEs that are effective in both orientations.

STATEMENTS OF THE INVENTION

According to the present invention there is provided a polynucleotide comprising a UCOE, which opens chromatin or maintains chromatin in an open state and facilitates reproducible expression of an operably-linked gene in cells of at least two different tissue types, wherein the nucleotide sequence of the UCOE does not occur in nature.

Genomic regions comprising regulatory sequences from at least two genes were combined to form a chimeric UCOE which significantly enhanced gene expression over a prolonged period of culture. Such a chimeric UCOE constitutes a nucleotide sequence that does not occur in nature. Accordingly the phrase "does not occur in nature" refers to a situation wherein the nucleotide sequence of the element constituting the UCOE does not naturally exist as such and is man-made or

artificially constructed, being a combination of naturally-occurring and/or artificially-generated sequences.

As used herein, the terms "artificial", "artificially-constructed", "chimeric", and the like, in reference to a UCOE, are used interchangeably throughout to mean that the UCOE or the element constituting the UCOE does not naturally exist; i.e., "does not occur in nature." Where the UCOE is a combination of naturally-occurring sequences, it is their arrangement or organization into the UCOE that is non-natural. By way of non-limiting example, an artificial UCOE could be comprised of two naturally-occurring sequences that are normally disparate (from different regions of a chromosome, from different chromosomes, from different organisms, etc.) and that have been brought together in a non-natural organization, to create a chimeric or artificial UCOE.

As used herein, the terms "artificially synthesized" and "artificially-generated" in reference to sequences in a UCOE, refer to sequences that are non-natural; i.e., sequences that are not naturally-occurring and are wholly synthetic. Such "artificially synthesized" and "artificially-generated" sequences can also be combined with naturally-occurring sequences to make up or create an artificial UCOE.

According to an alternative aspect of the invention there is provided a nucleic acid molecule comprising a DNA sequence selected from :

i) the DNA sequence as represented in Figure 2 or Figure 7;

ii) DNA sequences which hybridise to the sequence presented in Figure 2 or Figure 7 which encode a polypeptide according to the invention.

In a preferred embodiment of the invention there is provided an isolated nucleic acid molecule which anneals under stringent hybridisation conditions to the sequence presented in Figure 2 or Figure 7.

Stringent hybridisation/washing conditions are well known in the art. For example, nucleic acid hybrids that are stable after washing in 0.1 x SSC ,0.1% SDS at 60°C. It is well known in the art that optimal hybridisation conditions can be calculated if the sequence of the nucleic acid is known. For example, hybridisation conditions can be determined by the GC content of the nucleic acid subject to hybridisation. Please see Sambrook *et al* (1989), Molecular Cloning; A Laboratory Approach. A common formula for calculating the stringency conditions required to achieve hybridisation between nucleic acid molecules of a specified homology is:

$T_m = 81.5^0\ C + 16.6\ Log\ [Na^+] + 0.41[\ \%\ G + C] - 0.63$ (%formamide)

Preferably the polynucleotide of the present invention facilitates reproducible expression of an operably-linked gene non-tissue specifically.

Preferably the polynucleotide of the present invention facilitates reproducible expression of an operably-linked gene in all tissue types where active gene expression occurs.

Preferably the polynucleotide of the present invention facilitates expression of an operably-linked gene at a physiological level.

Preferably the polynucleotide of the present invention comprises an extended methylation-free, CpG-island.

Preferably the polynucleotide of the present invention comprises one or more naturally-occurring sequences associated with the control of gene expression.

Preferably the polynucleotide of the present invention comprises one or more naturally-occurring promoters.

Preferably the polynucleotide of the present invention comprises dual or bi-directional promoters that transcribe divergently.

Preferably the polynucleotide of the present invention comprises the human β-actin CpG island/ promoter region, or fragment thereof.

Preferably the polynucleotide of the present invention comprises a DNA fragment within the range of 100bp to 3 kb spanning the human β-actin CpG island/ promoter region or a fragment thereof.

Preferably the polynucleotide of the present invention comprises the human PDCD2 CpG island/ promoter region or a fragment thereof.

Preferably the polynucleotide of the present invention comprises a DNA fragment within the range from 100bp to 3.0 kb spanning the human PDCD2 CpG island/ promoter region, or a fragment thereof.

Preferably the polynucleotide of the present invention comprises a DNA fragment within the range from 100bp to 3.0kb spanning the human β-actin CpG island/ promoter region and a DNA fragment within the range from 100bp to 3.0 kb spanning the human PDCD2 CpG island/ promoter region. Preferably said fragments are directly adjacent with their promoters oriented divergently.

Preferably the polynucleotide of the present invention comprises a 2kb DNA fragment spanning the human β-actin CpG island/ promoter region and a 1.8kb DNA fragment spanning the human PDCD2 CpG island/ promoter region. Preferably said fragments are directly adjacent with their promoters oriented divergently.

In further preferred embodiment the polynucleotide comprises the sequence of Figure 2 or a fragment thereof, in either orientation.

Preferably the polynucleotide comprises the human RNP CpG island / promoter region or a fragment thereof.

Preferably the polynucleotide comprises a 4kb DNA fragment spanning the human RNP CpG island / promoter region.

Preferably the polynucleotide comprises an extended methylation-free CpG island containing bidivergent promoters adjacent to at least one further sequence comprising a methylation-free CpG island.

Preferably the polynucleotide comprises the human RNP CpG island /promoter region and a DNA fragment in the range 100bp to 3.0kb spanning the human ß-actin CpG island / promoter region.

In a further preferred embodiment the polynucleotide comprises the sequence of Figure 7 or a fragment thereof in either orientation. Preferably the polynucleotide comprises one or more promoter sequences.

It is known in the art that initiation of transcription may, under some circumstances, be inhibited by read-through transcripts from upstream promoters (Youssoufian and Lodish (1993) Transcriptional inhibition of the murine erythropoietin receptor gene by an upstream repetitive element, Mol Cell Biol 13 (1) 98–104). Therefore, one embodiment of the invention comprises the polynucleotide of the present invention wherein one or more of the promoter sequences are mutated in such a way that they are incapable of initiating transcription.

Preferably the promoter is selected from CMV, EF-1α, RSV LTR, or HIV2 LTR or combinations of sequences derived therefrom. More preferably the promoter is a CMV promoter. Most preferably it is the mouse CMV promoter.

Preferably the polynucleotide of the present invention comprises at least one sequence which is artificially synthesised.

The present invention also provides a vector comprising the polynucleotide of the present invention.

Preferably said vector is an expression vector adapted for eukaryotic gene expression.

Typically said adaptation includes, by example and not by way of limitation, the provision of transcription control sequences (promoter sequences) which mediate cell/tissue specific expression.

Promoter and enhancer are terms well-known in the art and include the following features which are provided by example only, and not by way of limitation. Promoters are 5', *cis*-acting regulatory sequences directly linked to the initiation of transcription. Promoter elements include so-called TATA box and RNA polymerase initiation selection (RIS) sequences which function to select a site of transcription initiation. These sequences also bind polypeptides which function, *inter alia*, to facilitate transcription initiation selection by RNA polymerase.
Enhancer elements are *cis* acting nucleic acid sequences often found 5' to the transcription initiation site of a gene (enhancers can also be found 3' to a gene sequence or even located in intronic sequences and is therefore position independent). Enhancers function to increase the rate of transcription of the gene to which the enhancer is linked. Enhancer activity is responsive to *trans* acting transcription factors (polypeptides) which have been shown to bind specifically to enhancer elements. The binding/activity of transcription factors is responsive to a number of environmental cues which include, by way of example and not by way of limitation, intermediary metabolites (eg glucose), environmental effectors (eg heat,). (See Eukaryotic Transcription Factors, by David S. Latchman, Academic Press Ltd, San Diego)

Adaptations also include the provision of selectable markers and autonomous replication sequences which both facilitate the maintenance of said vector in either the eukaryotic cell or prokaryotic host. Vectors which are maintained autonomously

are referred to as episomal vectors. Episomal vectors are desirable since they are self-replicating and so persist without the need for integration. Episomal vectors of this type are described in WO98/07876.

Adaptations which facilitate the expression of vector encoded genes include the provision of transcription termination/polyadenylation sequences. This also includes the provision of internal ribosome entry sites (IRES) which function to maximise expression of vector encoded genes arranged in bicistronic or multi-cistronic expression cassettes.

These adaptations are well-known in the art. There is a significant amount of published literature with respect to expression vector construction and recombinant DNA techniques in general. Please see, Sambrook et al (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY and references therein; Marston, F (1987) DNA Cloning Techniques: A Practical Approach Vol III IRL Press, Oxford UK; DNA Cloning: F M Ausubel et al, Current Protocols in Molecular Biology, John Wiley & Sons, Inc.(1994).

Preferably the vector comprises an expressible gene operably-linked to a promoter and the polynucleotide.

Preferably the vector is an episomal or integrating vector.

In a preferred embodiment, the vector of the present invention is a plasmid.

Alternatively, the vector may be a virus, such as an adenovirus, adeno-associated virus, a herpesvirus, vaccinia virus, lentivirus or other retrovirus.

Preferably the operably-linked gene is a therapeutic nucleic acid sequence.

Preferably the vector comprises two sites for insertion of open reading frames to be expressed, each transcribed from a distinct promoter, said promoters being arranged so as to transcribe divergently and both promoters being operably-linked to an artifically-constructed UCOE situated between them, and wherein said UCOE has been so constructed as to be effective in both orientations. This is particularly useful for the production of proteins that comprise two or more polypeptide chains, including, but not limited to, immunoglobulins. The insertion sites in the vector may be inserted with nucleic acid encoding different polypeptides of interest, including, but not limited to an open reading frame encoding an immunoglobulin heavy and an immunoglobulin light chain.

Preferably the vector comprises the sequence of Figure 2, the CMV promoter, a multiple cloning site, a polyadenylation sequence and genes encoding selectable markers under suitable control elements. It will be apparent to one of skill in the art that the UCOE can be inserted in both orientations.

Preferably the vector comprising the artificial UCOE is CET 500 as shown schematically in Figure 3a.

In an alternative embodiment, the vector is CET 501 as shown in Figure 3b.

Alternatively, the vector of any of claims comprises the sequence of Figure 7, the CMV promoter, a multiple cloning site, a polyadenylation sequence and genes encoding selectable markers under suitable control elements. Similarly, it will be apparent to one of skill in the art that the artificial UCOE can be inserted in both orientations.

In an preferred embodiment of this alternative, the vector comprising the artificial UCOE is known as CET 600 as shown schematically in Figure 8a.

In a further alternative embodiment, the vector is CET 601 as shown schematically in Figure 8b.

The present invention also provides a host cell that is transfected with the vector of the present invention.

The present invention also provides the polynucleotide, vector or the host cell for use in therapy.

The present invention also provides use of the polynucleotide, vector or host cell in the manufacture of a composition for use in gene therapy.

The present invention also provides a method of treatment, comprising administering to a patient in need of such treatment a pharmaceutically effective amount of the polynucleotide, vector or host cell of the present invention. Preferably the patient is suffering from a disease treatable by gene therapy.

The present invention also provides a pharmaceutical composition comprising the polynucleotide and/or the vector and/or host cell, optionally in admixture with a pharmaceutically acceptable carrier or diluent, for therapy to treat a disease or provide the cells of a particular tissue with an advantageous protein or function.

The polynucleotide, vector or host cell of the invention or the pharmaceutical composition may be administered via a route which includes systemic intramuscular, intravenous, aerosol, oral (solid or liquid form), topical, ocular, rectal, intraperitoneal and/or intrathecal and local direct injection.

The exact dosage regime will, of course, need to be determined by individual clinicians for individual patients and this, in turn, will be controlled by the exact nature of the protein expressed by the gene of interest and the type of tissue that is being targeted for treatment.

The dosage also will depend upon the disease indication and the route of administration. The number of doses will depend upon the disease, and the efficacy data from clinical trials.

The amount of polynucleotide or vector DNA delivered for effective gene therapy according to the invention will preferably be in the range of between 50 ng -1000 μg of vector DNA/kg body weight; and more preferably in the range of between about 1-100 μg vector DNA/kg.

Although it is preferred according to the invention to administer the polynucleotide, vector or host cell to a mammal for *in vivo* cell uptake, an *ex vivo* approach may be utilised whereby cells are removed from an animal, transduced with the polynucleotide or vector, and then re-implanted into the animal. The liver, for example, can be accessed by an *ex vivo* approach by removing hepatocytes from an animal, transducing the hepatocytes *in vitro* and re-implanting the transduced hepatocytes into the animal (e.g., as described for rabbits by Chowdhury *et al.*, *Science* 254:1802-1805, 1991, or in humans by Wilson, *Hum. Gene Ther.* 3:179-222, 1992). Such methods also may be effective for delivery to various populations of cells in the circulatory or lymphatic systems, such as erythrocytes, T cells, B cells and haematopoietic stem cells.

In another embodiment of the invention, there is provided a mammalian model for determining the efficacy of gene therapy using the polynucleotide, vector or host cell of the invention. The mammalian model comprises a transgenic animal whose cells contain the vector of the present invention. Methods of making transgenic mice (Gordon *et al., Proc. Natl. Acad. Sci. USA* 77:7380 (1980); Harbers *et al., Nature*293:540 (1981); Wagner *et al., Proc. Natl. Acad. Sci. USA* 78:5016 (1981); and Wagner *et al., Proc. Natl. Acad. Sci. USA* 78:6376 (1981), sheep pigs, chickens (see Hammer *et al., Nature* 315:680 (1985)), etc., are well-known in the art and are

contemplated for use according to the invention. Such animals permit testing prior to clinical trials in humans.

Transgenic animals containing the polynucleotide of the invention also may be used for long-term production of a protein of interest.

The present invention also provides for use of the polynucleotide and/or vector and/or host cell in a cell culture system in order to obtain a desired gene product. Suitable cell culture systems are well known in the art and are fully described in the body of literature known to those skilled in the art.

The present invention also provides for use of the polynucleotide to increase the expression of an endogenous gene comprising inserting the polynucleotide into the genome of a cell in a position operably associated with the endogenous gene thereby increasing the level of expression of the gene.

The present invention also provides the use of the polynucleotide of the present invention in producing transgenic plants.

The generation of transgenic plants which have increased yield, resistance, etc. are well known to those skilled in the art. The present invention also provides for transgenic plant containing cells which contain the polynucleotide of the present invention. Some or all of the cells comprising the artificial UCOE may originate from plants.

The present invention also provides for a transgenic non-human animal containing cells which contain the polynucleotide.

The present invention also relates to the use of polynucleotide of the present invention in functional genomics applications. Functional genomics relates principally to the sequencing of genes specifically expressed in particular cell types

or disease states and now provides thousands of novel gene sequences of potential interest for drug discovery or gene therapy purposes. The major problem in using this information for the development of novel therapies lies in how to determine the functions of these genes. UCOEs can be used in a number of functional genomic applications in order to determine the function of gene sequences. The functional genomic applications of the present invention include, but are not limited to:

(1) Using the polynucleotide of the present invention to achieve sustained expression of anti-sense versions of the gene sequences or ribozyme knockdown libraries, thereby determining the effects of inactivating the gene on cell phenotype.

(2) Using the polynucleotide of the present invention to prepare expression libraries for the gene sequences, such that delivery into cells will result in reliable, reproducible, sustained expression of the gene sequences. The resulting cells, expressing the gene sequences can be used in a variety of approaches to function determination and drug discovery. For example, raising neutralising antibodies to the gene product; rapid purification of the protein product of the gene itself for use in structural, functional or drug screening studies; or in cell-based drug screening.

(3) Using the polynucleotide of the present invention in approaches involving mouse embryonic stem (ES) cells and transgenic mice. One of the most powerful functional genomics approaches involves random insertion into genes in mouse ES cells of constructs which only allow drug selection following insertion into expressed genes, and which can readily be rescued for sequencing (G.Hicks *et al.*, Nature Genetics, **16**, 338-334). Transgenic mice with knockout mutations in genes with novel sequences can then readily be made to probe their function. At present this technology works well for the 10% of mouse genes which are well expressed in mouse ES cells.

Incorporation of UCOEs into the integrating constructs will enable this technique to be extended to identify all genes expressed in mice.

In an alternative embodiment of the invention, there is provided a method for the production of the polypeptide according to the invention comprising:

i) providing a cell transformed/transfected with a nucleic acid molecule according to the invention;

ii) growing said cell in conditions conducive to the manufacture of said polypeptide; and

iii) purifying said polypeptide from said cell, or its growth environment.

In a preferred embodiment of the invention said nucleic acid molecule is the vector according to the invention.

In a preferred method of the invention said vector encodes, and thus said polypeptide is provided with, a secretion signal to facilitate purification of said polypeptide.

Alternatively, other preferred embodiments may include further refinements to facilitate purification of expressed recombinant protein, such as affinity tags or epitopes, or enzymatic cleavage sites.

DETAILED DESCRIPTION OF THE INVENTION

The invention will now be described by way of example only and with reference to the accompanying figures wherein;

Figure 1 is a linear restriction map of the artificial UCOE *Nco* I fragment. The PDCD2 and actin transcripts are shown as shaded arrows, with the direction of the arrow depicting the direction of transcription from their respective promoters.

Figure 2 shows the nucleotide sequence of the PDCD2/Actin artificial UCOE *Nco* I fragment according to the invention.

Figures 3a and 3b depict plasmid maps of the PDCD2/Actin artificial UCOE-containing expression vectors.

Figure 3a shows CET 500, comprising the PDCD2/Actin artificial UCOE upstream of the CMV promoter and a multiple cloning site suitable for insertion of the open reading frame to be expressed. In this particular embodiment the plasmid backbone is from pEGFPN-1 and carries a kanamycin/ neomycin resistance gene.

Figure 3b shows CET 501, comprising the PDCD2/actin artificial UCOE in the reverse orientation.

Figure 4 shows levels of expression of a reporter transgene (EGFP) in pools of stably-transfected CHO cells maintained under G418 selection over a period of 85 days. Expression was driven by the CMV promoter alone (CMV), or in combination with the 8kb RNP/HP-1 UCOE (CET 220) or artificial PDCD2/actin UCOE CET 510) and was measured as median fluorescence by FACS .

Figure 5 shows the proportion of cells expressing the transgene in the above experiment expressed as % positive cells on FACS analysis over a period of 85 days.

Figure 6 is a linear restriction map of the RNP/HP-1/actin artificial UCOE fragment. The methylation-free islands are shown as shaded arrows, with the direction of the arrow depicting the direction of transcription from their respective promoters.

Figure 7 shows the nucleotide sequence of the RNP/HP-1/actin artificial UCOE fragment according to the invention.

Figure 8 depicts plasmid maps of RNP/HP-1/actin artificial UCOE-containing expression vectors.

Figure 8a shows CET 600, comprising the RNP/HP-1/actin artificial UCOE upstream of the CMV promoter and a multiple cloning site suitable for insertion of the open reading frame to be expressed. In this particular embodiment the plasmid backbone is from pEGFPN-1 and carries a kanamycin/ neomycin resistance gene.

Figure 8b shows CET 601, comprising the RNP/HP-1/actin artificial UCOE in the reverse orientation.

Figure 9 shows levels of expression of a reporter transgene (EGFP) in pools of stably-transfected CHO cells maintained under G418 selection over a period of 127 days. Expression was driven by the CMV promoter alone (CMV), or in combination with the 8kb RNP UCOE in either forward (CET 220) or reverse (CET 221) or by RNP/HP-1/actin artificial UCOE in either forward (CET 610) or reverse (CET 611) orientation and was measured as median fluorescence by FACS. It will be understood that CET 610 is therefore the equivalent of CET 600, but containing EGFP as the inserted gene, and that CET 611 is the corresponding EGFP-containing equivalent of CET 601.

Figure 10 shows the proportion of cells expressing the transgene in the above experiment expressed as % positive cells on FACS analysis over a period of 127 days.

Figures 11a - 11d depict plasmid maps of the bi-directional UCOE vectors for the expression of immunoglobulins.

Example 1

Heterologous combinations of UCOE fragments

We have found that the substitution of the *A2/HP1* fragments in the 4.0CMV and 8.0CMV constructs with the comparable region from the *TBP/B1* locus gave

substantial increases in the level and stability of EGFP expression from the hCMV promoter. These findings in tissue culture cells provide evidence that regions of DNA encompassing an extended CpG island and divergently transcribed promoters are responsible for conferring an increase in the proportion of viable integration events, significantly improved transgene expression and a resistance to transcriptional silencing even from within centromeric heterochromatin.

Genomic regions encompassing CpG islands from housekeeping genes associated with a single promoter were combined. The construct comprised the 2kb CpG island from the 5' end of the human *β-actin* gene[37] joined to a 3.2kb fragment containing the CpG island from the 5' end of the human *PDCD2* gene. The promoters were divergently transcribed and their transcriptional start sites separated by 1.9kb.

The entire 5.2kb combination was then linked to an EGFP reporter gene driven by the CMV promoter (PDCD2/ACTIN). As a comparison, the *β-actin* CpG *island/promoter region alone was also inserted upstream of the CMV-EGFP* expression vector (ACTIN).

Materials and Methods

With particular reference to Figure 2, the CMV promoter was removed from pEGFP-N1 by digestion with *AseI* and *NheI* followed by blunting with T4 DNA polymerase and religation to produce pΔ-EGFP-N1. A β-Actin promoter EGFP fusion was constructed by digesting MA39 (HSACCYBB) with *Bsm*I and *Nco*I, isolating the appropriate fragment and blunting it with T4 DNA polymerase. This fragment was ligated into pΔ-EGFP-N1 that had been digested with *Age*I, blunted with T4 DNA polymerase and then digested with SmaI. This produced an in frame fusion of the first codon of the β-Actin gene with EGFP, expression of which was driven off the β-Actin promoter/MFI (methylation free island). The construct, pActin-EGFP, was shown to express EGFP in stably transfected CHOK1 cells.

To construct a vector with bi-directional promoters and an extended methylation free island, the PDCD2 gene and some of its promoter was initially removed from pCP2-TNN (approximately 160kb of the TBP locus in pCYPAC-2) by digestion with *Swa*I and *Bsp*EI and sub cloned into pBluescriptKS+ that had been digested with *Eco*RV and *Xma*I (pPDCD2-KS). As this vector did not contain the whole of the PDCD2 methylation free island, the remaining 5' region of the PDCD2 methylation free island, which also contained the promoter, was obtained by PCR from pCP2-TNN using the following priers 5'GCGGTACCAAGGGCATTCTGAAGTTAACC-3', 5-AGCTCCACAGGCCTGG-3'. The PCR product was then digested with *Kpn*I (site generated with PCR primers) and *Stu*I (internal site), pActin-EGFP was digested with *Sal*I and *Kpn*I, and the PDCD2 gene was removed from pPDCD2-KS by digestion with *Sal*I and *Stu*I. All three fragments were ligated together to create pPDCD2-Actin-EGFP .

The region containing the methylation free island of pF-PDCD2-Actin-EGFP was removed as an *Nco* I fragment (approximately 5.2kb), this was blunted with T4 DNA polymerase and then ligated, in both orientations, into pEGFPN-1 that had been digested with *Ase* I and then blunted. These vectors were called CET 510 (UCOE in forward orientation) and CET511 (UCOE in reverse orientation). The corresponding empty expression vectors with no transgene inserted into the multiple cloning sites were termed CET 500 and CET 501, respectively (see Figure 3).

Constructs were linearized with *Pvu* I, transfected into CHO-K1 cells and selected under G418 selection (0.6mg/ml) for both experiments.

It will be understood that one of skill in the art may adapt these procedures for preparation and testing of other polynucleotides of the invention.

Results

As shown in Figures 4 and 5, the artificial UCOE construct gave levels of reporter gene expression that were comparable with those achieved with the 8kb RNP/HP-1

UCOE in terms of median fluorescence, and slightly better in terms of the proportion of cells expressing over the time course of the experiment.

Example 2

Heterologous combinations of UCOE and methylation-free CpG island fragments

Materials and Methods

The actin methylation-free island was removed from pActin-EGFP by digestion with *Nco* I, blunted followed by digestion with *Kpn* I. The RNP 4kb fragment was removed from CET 20 by digestion with *Kpn* I and *Hind* III. These two fragments were then ligated into pBKS which had been digested with *Cla* I, blunted and then cut with *Hind* III. This gave the artificial UCOE in pBKS.

The artificial UCOE was then removed from pBKS by digestion with *Sal* I and *Hind* III, blunted and ligated into pEGFP-N1 that had been digested with *Ase* I and blunted. The UCOE was inserted in both orientations to create CET 610 and CET 611 respectively. The corresponding expression vectors with no transgene inserted into multiple cloning sites were termed CET 600 and CET 601, respectively (see Figure 8).

Constructs were linearised and transfected into CHO-K1 cells and selected under G418 selection (0.6mg/ml) for the duration of the experiments.

Results

The effects of the artificial UCOE (in both orientations) on levels of transgene expression were assessed by comparison with results obtained with CMV promoter alone, and the 8kb RNP/HP-1 UCOE (in both orientations). Figure 9 shows that, in terms of median fluorescence on FACS analysis, the CMV promoter alone gave very low levels of expression. Both the RNP UCOE (CET 220) and artificial UCOE (CET 600) gave greatly increased (30-fold) expression in the forward orientation and were comparable. However, a directional bias seen with the RNP UCOE was less

marked with the artificial UCOE, i.e. in the reverse orientation the artificial UCOE was superior, although still slightly less good than either UCOE in the forward orientation. This was true for both overall level of expression and ability to maintain expression over a prolonged period (more than 120 days).

When the same experiment was analysed interms of proportion of cells maintaining expression over time, both UCOEs, in both orientations, gave consistently better results than CMV alone (approximately twice as many cells expressing in the early part of the experiment). This difference became more marked over longer periods, with the CMV only population progressively losing expressing cells, until by 120 days only approximately 10% cells were still expressing. This contrasts with maintenance at approximately 90% for both UCOEs in the forward orientation, and only slightly lower levels (75–85%) for the reverse orientations by the end of the experiment.

Example 3

Co-expression of two genes on the same vector

Efficient functional antibody production requires appropriately balanced expression of the heavy and light chains. Transfection of the two chains on separate plasmids makes the maintenance of an equal copy number difficult and provides the potential for transcriptional interference between the genes if the vectors integrate close to one another in the genome.

A series of new vectors for the co-expression of two genes on the same vector have been constructed to compare neo versus puro as resistance markers and hCMV, beta actin or mCMV promoters to drive light or heavy chain expression (figure 11).

Material and Methods

The two *Sfi* I sites of pORT1 (Cobra) were changed to *Mfe* I sites by introduction of adapter molecules comprised of annealed oligos Mfe.F, 5'-AACAATTGGCGGC and Mfe.R, 5'-GCCAATTGTTGCC.

The HSV TK polyA site was amplified from pVgRXR (Invitrogen) with primers TK.F, 5'ACGCGTCGACGGAAGGAGACAATACCGGAAG and TK.R, 5'-CCGCTCGAGTTGGGGTGGGGAAAAGGAA. The *Sal* I to *Xho* I fragment was inserted into the *Sal* I site. The murine PGK polyA site was amplified from male BALB/c genomic DNA (Clontech) using primers mPGK.F, 5'-CGGGATCCGCCTGAGAAAGGAAGTGAGCTG and mPGK.R, 5'-GAAGATCTGGAGGAATGAGCTGGCCCTTA, and the *Bam*H I to *Bgl* II fragment was cloned into the *Bam*H I site.

The *Ase* I to *Sal* I fragment of pcDNA3.1 containing the *neo* expression cassette was treated with T4 DNA polymerase, ligated to *Spe* I linkers (5'-GACTAGTC). The *Spe* I fragment was then cloned into the *Spe* I site to give pORTneoF or the *Eco*R I to *Not* I fragment of CET700 (Cobra) carrying the puromycin resistance cassette was treated with T4 DNA polymerase, ligated to *Xba* I linkers, and the *Xba* I fragment was cloned into the *Xba* I site to give pORTpuroF.

The *Hind* III to *Bam*H I murine CMV promoter fragment from pCMVEGFPN-1 (Cobra) was subcloned into the *Hind* III to *Bam*H I sites of the Hybrid UCOE in BKS+ (Cobra). The human CMV promoter was then amplified from plasmid pIRESneo (Clontech) using primers hCMVF, 5'-CTCGAGTTATTAATAGTAATCAATTACGGGGTCAT and hCMVR, 5'-GTCGACGATCTGACGGTTCACTAAACCAGCTCT and the *Xho* I to *Sal* I fragment was cloned into the *Sal* I site. The *Bam*H I to *Sal* I fragment was then cloned into the *Bam*H I to *Sal* I sites of pORTneoF to give pBDUneo100, or into pORTpuroF to give pBDUpuro300.

The two ATG codons upstream of the *Sal* I cloning site in the Hybrid UCOE in BKS+ were altered by site-directed mutagenesis and the *Bam*H I to *Sal* I fragment was cloned into the *Bam*H I to *Sal* I sites of pORTneoF to give pBDUneo200, or into pORTpuroF to give pBDUpuro400.

Human antibody light chain coding sequences were cloned into either the *Bam*H I or *Sal* I sites of all four bi-directional UCOE vectors (pBDUneo100, pBDUneo200, pBDUpuro300 and pBDUpuro400), followed by immunoglobulin heavy chain coding sequence at the remaining *Bam*H I or *Sal* I cloning site to give pBDUneo112, pBDUneo121, pBDUneo212, pBDUneo221, pBDUpuro112, pBDUpuro121, pBDUpuro212 and pBDUpuro221. All eight antibody expression constructs were transfected into CHO-K1 cells using Lipofectamine (Invitrogen) following the manufacturer's instructions, and selected with 500 μg/ml G418 (*neo* vectors) or 12.5 μg/ml puromycin (*puro* vectors).

Pools were selected and antibody production rates compared between the different constructs to determine the optimal promoter and selectable marker combination for antibody expression in CHO cells.

Results

Table 1 demonstrates that vectors containing the light chain expressed from the murine CMV promoter gave the best expression of the antibody. No significant difference was observed between production rates obtained with vectors containing the G418 or puromycin-resistance cassettes. The production rate from the pool of a co-transfection experiment performed separately is compared. Clones from this pool were isolated with production rates of 3-18pg/cell/day. However, clones above 5pg/cell/day were unstable and rapidly decreased in expression or stopped producing. Clones expressing approx. 5pg/cell/day were used for initial fermentation experiments. These preliminary indications are very encouraging that higher production rates will be observed in clones isolated from the bi-directional UCOE vector transfectants.

Table 1: Expression of hAb1 (IgG4) from Bi-directional UCOE vectors (CHO-K1 pools)

CHO-K1 POOLS

Vector (pg/cell/day)	H3 Promoter	K1 Promoter	Production Rate
pBDUneo112	murine CMV	human CMV	0.3
pBDUneo121	human CMV	murine CMV	1.5
pBDUneo212	murine CMV	human beta-actin	0.06

pBDUneo221	human beta-actin	murine CMV	1.3
pBDUpuro312	murine CMV	human CMV	0.5
pBDUpuro321	human CMV	murine CMV	1.4
pBDUpuro412	murine CMV	human beta-actin	0.05
pBDUpuro421	human beta-actin	murine CMV	2.3
Cotransfection**	human CMV	human CMV	0.7

**Contransfection was carried out previously using the same antibody genes each driven from 4kb UCOE CMV vectors (hydro and neo selection).

REFERENCES

1. Dillon, N. & Grosveld, F. Chromatin domains as potential units of eukaryotic gene function. *Curr. Opin. Genet. Dev. 4,* 260-264 (1994).

2. Higgs, D.R. Do LCRs open chromatin domains? *Cell* 95, 299-302 (1998).

3. Palmiter, R.D. & Brinster, R.L. Germline transformation of mice. *Ann. Ref. Genet.* 20, 465-499 (1986).

4. Allen, N.D. *et al.* Transgenes as probes for active chromosomal domains in mouse development. *Nature* 333, 852-855 (1988).

5. Bonnerot, C., Grimber, G., Briand, P. & Nicolas, J.F. Patterns of expression of position-dependent integrated transgenes in mouse embryo. *Proc. Natl. Acad. Sci. USA* 87:6331-6335 (1990).

6. Kioussis, D. & Festenstein, R. Locus control regions: overcoming heterochromatin-induced gene inactivation in mammals. *Curr. Opin. Genet. Dev.* 7, 614-619 (1997).

7. Pikaart, , M.J., Recillas-Targa, F. & Felsenfield, G. Loss of transciptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev.* 12, 2852-2862 (1998).

8. Fussenegger, M., Bailey, J.E., Hauser, H. & Mueller, P.P Genetic optimization of recombinant glycoprotein production by mammalian cells. *Trends Biotech.* 17, 35-42 (1999).

9. Fraser, P. & Grosveld, F. Locus control regions; chromatin activation and transcription. *Curr. Opin. Cell Biol.* 10, 361-365 (1998).

10. Li, Q., Harju, S. & Peterson, K.R. Locus Control Regions: coming of age at a decade plus. *Trends Genet.* 15: 403-408 (1999).

11. Bulger, M. & Groudine, M. Looping versus linking: toward a model for long-distance gene activation. *Genes Dev.* 13, 2465-2477 (1999).

12. Grosveld, F. Activation by locus control regions? *Curr. Opin. Genet. Dev. 9* 152-157 (1999).

13. Bender, M.A., Bulger, M., Close, J. & Groudine, M. ß-globin Gene Switching and Dnase I Sensitivity of the Endogenous ß-globin Locus in Mice Do Not Require the Locus Control Region. *Mol. Cell* 5, 387-393 (2000).

14. Ortiz, B.D., Cado, D., Chen, V., Diaz, P.W. & Winoto, A. Adjacent DNA elements dominantly restrict the ubiquitous activity of a novel chromatin-opening region to specific tissues. *EMBO J.* 16, 5037-5045 (1997).

15. Ortiz, B.D., Cado, D. & Winoto, A. A new element within the T-cell receptor alpha locus required for tissue-specific locus control region activity. *Mol. Cell. Biol.* 19, 1901-1909 (1999).

16. Antequera, F. & Bird, A. Number of CpG islands and genes in human and mouse. *Proc. Natl. Acad. Sci. USA* 90, 1195-11999 (1993).

17. Cross, S.H. & Bird, A.P. CpG islands and genes. *Curr. Opin, Genet. Dev. 5,* 309-314 (1995).

18. Tazi, J. & Bird, A. Alternative chromatin structure at CpG islands. *Cell* 60, 909-920 (1990).

19. Imbert, G., Trottier, Y., Bechmann, J., & Mandel, J.L. The gene for the TATA binding protein (TBP) that contains a highly polymorphic protein coding CAG repeat maps to 6q27. *Genomics* 21: 667-668 (1994).

20. Purrello, M. *et al.* Physical mapping at 6q27 of the locus for the TATA-box binding protein, the DNA bining subunit of TFIID and a component of SL1 and TFIIIB, strongly suggests that it is single copy in the human genome. *Genomics* 22, 94-100 (1994).

21. Trachtulec, Z *et al.* Linkage of TATA-binding protein and proteasome subunit C5 genes in mice and humans reveals synteny conserved between mammals and invertebrates. *Genomics* 44: 1-7 (1997).

22. Owens, G.P., Hahan, W.E., & Cohen, J.J. Identification of mRNAs associated with programmed cell death in immature thymocytes. *Mol. Cell. Biol.* 11, 4177-4188 (1991).

23. Chalut, C., Gallois, Y., Poterszman, A., Moncollin, V. & Egly, J.-M. Genomic structure of the human TATA-box-binding protein (TBP). *Gene* 161, 277-282 (1995).

24. Schmidt, E.E. & Schibler, U. High accumulation of components of the RNA polymerase II transcription machinery in rodent spermatids. *Development* 121, 2373-2383 (1995).

25. Biamonti, G., Ruggiu, M., Saccone, S., Della Valle, G. & Riva, S. Two homologous genes, originated by duplication, encode the human hnRNP proteins A2 and A1. *Nucl. Acids Res.* 22, 1996-2002 (1994).

26. Kozu, T., Henrich, B. & Schafer, K.P. Structure and expression of the gene (HNRPA2B1) encoding the human hnRNP protein A2/B1. *Genomics* 25, 365-371 (1995).

27. Ye, Q. & Worman, H.J. Interaction between an integral protein of the nuclear envelope inner membrane and human chromodomain proteins homologous to Drosophila HP1. *J. Biol. Chem.* 271, 14653-14656 (1996).

28. James, T.C. & Elgin, S.C.R. Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene. *Mol. Cell. Biol.* 6, 3861-3872 (1986).

29. Singh, P.B., *et al.* A sequence motif found in a drosophila heterochromatin protein is conserved in animals and plants. *Nucl. Acids Res.* 19, 789-794 (1991).

30. Gilliand, G., Perrin,m S., Blanchard, K. & Bunn, H.F. Analysis of cytokine Mrna and DNA: detection and quantitation by competitive polymerase chain reaction. *Proc. Natl. Acad. Sci. USA* 87, 2725-2729 (1990).

31. Furth, P.A. Hennighausen, L., Baker, C., Beatty, B. & Woychick, R. The viarability in activity of the universally expressed human cytomegalovirus immediate early gene 1 promoter/enhancer in transgenic mice. *Nucl. Acids Res.* 19 6205-6208 (1991).

32. Ray, P. *et al.* Ectopic expression of a *c-kit* ^{W42} minigene in transgenic mice: recapitulation of *W* phenotypes and evidence for *c-kit* function in melanoblast progenitors. *Genes Dev. 5, 2265-2273 (1991).*

33. Yamashita, T. *et al* High level expression of human *a*-fetoprotein in transgenic mice. *Biochem, Biophys. Res. Comm.* 191, 715-720 (1993).

34. Milot, E. *et al.* Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell 87, 104-114 (1996).*

35. Festenstein, R. *et al.* Locus control region function and heterochromatin-induced position effect variegation. *Science* 271, 1123-1125 (1996).

36. Sabbattini, P., Georgiou, A., Sinclaire, C. & Dillon, N. Analysis of mice with single and multiple copies of transgenes reveals a novel arrangement for the λ5-V-pre*BI* locus control region. *Mol. Cell. Biol.* 19, 671-679 (1999).

37. Ng, S.Y. *et al.* Evolution of the functional human ß-actin gene and its multi-pseudogene family: conservation of noncoding regions and chromosomal dispersion of pseudogenes. *Mol. Cell. Biol.* 5, 2720-2732 (1985).

38. Pravtcheva, D.D., Wise, T.L., Ensor, N.J. & Ruddle, F.H. Mosaic expression of an HPRT transgene integrated in a region of Y heterochromatin. *J. Exp. Zool.* 268, 452-468 (1994).

39. Bell, A.C. & Felsenfield, G. Stopped at the border: boundaries and insulators. *Curr. Opin. Genet. Dev.* 9, 191-198 (1999).

40. Winston, J.H., Hong, L., Datta, S.K. & Kellems, R.E. An intron 1 regulatory region from the murine adenosine deaminase gene can activate heterologous promoters for ubiquitous expression in transgenic mice. *Somat. Cell Mol. Genet.* 22, 261-278 (1996).

41. Hart, C.M. & Laemmli, U.K. Facilitation of chromatin dynamics by SARs. *Curr. Opin. Genet. Develop.* 8,519-525 (1998).

42. Klehr, D., Maass, K. & Bode, J. Scaffold-attached regions from the human interferon ß domain can be used to enhance the stable expression of genes under the control of various promoters. *Biochemistry* 30, 1264-1270 (1991).

43. McKnight, R.A., Shamay, L., Sankaran, L., Wall, R.J. & Henninghausen, L. Matrix-attachment regions can impart position-independent regulation of a tissue-specific gene in transgenic mice. *Proc. Natl. Acad. Sci. USA* 89, 6943-6947 (1992).

44. Van Drunen, C.M., et al. A bipartite sequence element associated with matrix/scaffold attachment regions. *Nucl. Acids Res.* 27, 2924-2930 (1999).

45. Verma, I.M. & Somia, N. Gene Therapy – promises, problems and prospects. *Nature* 389: 239-242 (1997).

46. Brown, S.A. & Kingston, R.E. Disruption of downstream chromatin directed by a transcriptional activator. *Genes Dev.* 11.3 116-3121 (1997).

47. Orphanides, G., LeRoy, G., Chang, C.H., Luse, D.S. & Reinberg, D. FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell*92, 105-116 (1998).

48. Schnitzler, G., Sif, S. & Kingston, R.E. Human SWI/SNF interconverts a nucleosome between its base state and a stable remodelled state. *Cell* 94, 17-27 (1998).

49. Travers, A. Chromatin modification by DNA tracking. *Proc. Natl. Acad. Sci. USA* 96, 13634-13637 (1999).

50. Ashe, H.L., Monks, J., Wijgerde, M., Fraser, P. & Proudfoot, N.J. Intergenic transcription and transinduction of the human ß-globin locus. *Genes Dev* 11, 2494-2509 (1997).

51. Rogan, D.F., Cousins, D.J. & Staynov, D.Z. Intergenic transcription occurs throughout the human IL-4/IL-13 gene cluster. *Biochem, Biophys, Res. Commun.* 255, 556-561 (1999).

52. Gribnau, J., Diderich, K., Pruzina, S., Calzorlari, R. & Fraser P. Intergenic Transcription and Developmental Remodeling of Chromatin Subdomains in the Human ß-globin Locus. *Mol. Cell* 5, 377-386 (2000).

53. Vyas, P. *et al.* Cis-acting sequences regulating expression of the human α-globin cluster lie within constitutively open chromatin. *Cell* 69, 781-793 (1992).

54. Ioannou, P.A. *et al.* A new bacteriophage P1-derrived vector for the propagation of large human DNA fragments. *Nature Gene.* 6, 84-89 (1994).

55. Raguz, S. *et al.* Muscle-specific locus control region activity associated with the human desmin gene. *Dev. Biol.* 201, 26-42 (1998).

56. Dillon, N. & Grosveld, F. Transcriptional analysis using transgenic animals, in *Gene Transcription: A practical approach* (eds Hames, B. D. & Higgins, S.J.) 153-188 (IRL Press, Oxford, 1993).

57. Morgenstern, J.P. & Land, H. 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, 3587-3595 (1990).

58. Horz, W. & Altenburger, W. Nucleotide sequence of mouse satellite DNA. *Nucl. Acids Res.* 9, 683-696 (1981).

59. Monfouilloux, S. *et al.* Recent human-specific spreading of a subtelomeric domain. *Genomics* 51, 165-176 (1998).

CLAIMS:

1. A polynucleotide comprising a ubiquitous chromatin opening element (UCOE) comprising an extended methylation free CpG-island, which opens chromatin or maintains chromatin in an open state and facilitates reproducible expression of an operably-linked gene in cells of at least two different tissue types, wherein the nucleotide sequence of the UCOE does not occur in nature.

2. The polynucleotide of claim 1, which facilitates reproducible expression of an operably-linked gene non-tissue specifically.

3. The polynucleotide of claim 1, which facilitates reproducible expression of an operably-linked gene in all tissue types where active gene expression occurs.

4. The polynucleotide of claim 1, which facilitates expression of an operably-linked gene at a physiological level.

5. The polynucleotide of claim 1, wherein the UCOE comprises an extended methylation-free, CpG-island.

6. The polynucleotide of claim 1, wherein the UCOE comprises at least one naturally-occurring sequence associated with the control of gene expression.

7. The polynucleotide of claim 1, wherein the UCOE comprises at least one naturally-occurring promoter.

8. The polynucleotide of claim 1, wherein the UCOE comprises dual or bi-directional promoters that transcribe divergently.

9. The polynucleotide of claim 1, wherein the UCOE comprises the human β-actin CpG island/promoter region or a fragment thereof.

10. The polynucleotide of claim 9, wherein the UCOE comprises a DNA fragment within the range of 100bp to 3.0 kb spanning the human β-actin CpG island/promoter region or a fragment thereof.

11. The polynucleotide of claim 1, wherein the UCOE comprises the human PCDC2 CpG island/promoter region or a fragment thereof.

12. The polynucleotide of claim 11, wherein the UCOE comprises a DNA fragment within the range of 100bp to 3.0 kb spanning the human PCDC2 CpG island/promoter region or a fragment thereof.

13. The polynucleotide of claim 1, wherein the UCOE comprises a DNA fragment within the range of 100bp to 3.0 kb spanning the human β-actin CpG island/promoter region and a DNA fragment within the range of 100bp to 3.0 kb spanning the human PDCD2 CpG island/promoter region.

14. The polynucleotide of claim 1, wherein the UCOE comprises a 2.0 kb DNA fragment spanning the human β-actin CpG island/promoter region and a 1.8 kb DNA fragment spanning the human PDCD2 CpG island/promoter region.

15. The polynucleotide of claim 13 or claim 14 wherein the promoters are orientated divergently.

16. The polynucleotide of claim 1, wherein the UCOE comprises SEQ ID

NO:1 or a fragment thereof in either orientation.

17. The polynucleotide of claim 1, wherein the UCOE comprises the human RNP CpG island/promoter region or a fragment thereof.

18. The polynucleotide of claim 17, wherein the UCOE comprises a 4kb DNA fragment spanning the human RNP CpG island/promoter region.

19. The polynucleotide of claim 1, comprising an extended methylation-free CpG island containing bidivergent promoters adjacent to at least one further sequence comprising a methylation-free CpG island.

20. The polynucleotide of claim 1, wherein the UCOE comprises the human RNP CpG island/promoter region and a DNA fragment in the range 100bp to 3.0kb spanning the human β-actin CpG island/promoter region.

21. The polynucleotide of claim 20, wherein the UCOE comprises SEQ ID NO:2 or fragment thereof in either orientation.

22. The polynucleotide of claim 1, further comprising a promoter.

23. The polynucleotide of claim 1, wherein the UCOE comprises at least one promoter sequence mutated to be incapable of initiating transcription.

24. The polynucleotide of claim 22 or claim 23 wherein the promoter is the CMV promoter.

25. The polynucleotide of claim 24 wherein the promoter is the mouse CMV promoter.

26. The polynucleotide of claim 1, wherein the UCOE comprises at least one sequence which is artificially synthesized.

27. A vector comprising the polynucleotide of claim 1.

28. The vector of claim 27, further comprising an expressible gene operably-linked to a promoter and the polynucleotide.

29. The vector of claim 27 or claim 28 wherein the vector is an episomal or integrating vector.

30. The vector according to claim 27, wherein the vector is a plasmid.

31. The vector according to claim 27, wherein the vector is a virus.

32. The vector according to claim 28, wherein the operably-linked gene is a therapeutic nucleic acid sequence.

33. The vector of claim 27, wherein the UCOE is effective in both orientations, said vector further comprising

a) a first and a second promoter, said arranged to transcribe divergently; and

b) a first and a second site for insertion of open reading frames to be expressed,

wherein each of said first and second promoters is operably-linked to the UCOE,

wherein the first site for insertion is under the control of the first promoter, and the second site for insertion is under the control of the second promoter,

and further wherein the UCOE is situated between the first and second promoter.

34. The vector of claim 33, further comprising a nucleic acid encoding an immunoglobulin heavy chain inserted into the first site for insertion, and a nucleic acid encoding an immunoglobulin light chain inserted into the second site for insertion.

35. The vector of claim 27, comprising SEQ ID NO:1, the CMV promoter, a multiple cloning site, a polyadenylation sequence and genes encoding selectable markers under suitable control elements.

36. The vector of claim 27, comprising SEQ ID NO:2, the CMV promoter, a multiple cloning site, a polyadenylation sequence and genes encoding selectable markers under suitable control elements.

37. The vector of claim 35 or claim 36, wherein the orientation of the UCOE is reversed.

38. The vector CET 500.

39. The vector CET 501.

40. The vector CET 600.

41. The vector CET 601.

42. A host cell transfected with the vector of any one of claims 27 or 38–41.

43. A composition comprising the polynucleotide of claim 1.

44. A composition comprising the host cell of claim 42, in combination with a pharmaceutically acceptable excipient.

45. A method of obtaining a desired gene product comprising using the polynucleotide of claim 1.

46. A method of obtaining a desired gene product comprising using the host cell of claim 42 in a cell culture system in order to obtain a desired gene product.

47. A nucleic acid molecule comprising a DNA sequence selected from the group consisting of SEQ ID NO:1 and SEQ ID NO:2.

48. A method for preparing a polypeptide comprising:

providing a cell transformed/transfected with the nucleic acid molecule according to claim 47;

growing said cell in conditions conducive to the production of said polypeptide; and

purifying said polypeptide from said cell, or its growth environment.

49. A composition comprising the vector of any one of claims 27 or 38–41, in combination with a pharmaceutically acceptable excipient.

50. A method of obtaining a desired gene product comprising using the vector of any one of claims 27 or 38–41 in a cell culture system in order to obtain a desired gene product.