(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 April 2011 (21.04.2011)

PCT

(10) International Publication Number
WO 2011/046700 A1

(51) **International Patent Classification**:
A61K 48/00 (2006.01)

(21) **International Application Number**:
PCT/US2010/048872

(22) **International Filing Date**:
15 September 2010 (15.09.2010)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/251,938 15 October 2009 (15.10.2009) US

(71) **Applicant** *(for all designated States except US)*: **CHILDREN'S MEMORIAL HOSPITAL** [US/US]; 2300 Children's Plaza, Chicago, IL 60614 (US).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: **HARRIS, Ann** [GB/US]; Children's Memorial Research Center, 2300 Children's Plaza #211, Chicago, IL 60614 (US). **OTT, Christopher, J.** [US/US]; Children's Memorial Research Center, 2300 Children's Plaza #211, Chicago, IL 60614 (US).

(74) **Agent**: **MRKSICH, K., Shannon**; Brinks Hofer Gilson & Lione, P.O. Box 10087, Chicago, IL 60610 (US).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2011/046700 A1

(54) **Title**: CFTR-ACTIVATING ENHANCER

(57) **Abstract**: Disclosed are novel compositions and methods for the control of the expression of a *CFTR* gene, in particular the human *CFTR* gene. The invention relates to a novel *CFTR* gene regulatory element capable of increasing the activity of the *CFTR* gene promoter, and to nucleic acid constructs comprising the element together with the *CFTR* gene coding region. The element and constructs containing it are useful in gene therapy for treating cystic fibrosis.

CFTR-ACTIVATING ENHANCER

GOVERNMENT INTERESTS

[0001] This invention was made with United States government support from the National Institutes of Health (NIH), grant number R01 HL094585. The United States government has certain rights in this invention.

FIELD OF THE INVENTION

[0002] The present invention relates generally to the field of the cystic fibrosis transmembrane conductance regulator (CFTR), and in particular to the field of the human *CFTR* gene.

BACKGROUND

[0003] Understanding the three-dimensional organization of individual loci within the human genome and how it relates to the regulation of gene expression is the focus of intense study. Many new technologies are being developed in order to locate and classify the functional elements of the human genome. These elements may contribute to the transcription of ubiquitously expressed genes and are also likely responsible for regulating genes with expression that is temporally and/or spatially controlled. *Cis*-acting regulatory elements located within non-coding regions of genomic DNA can influence the organization of chromosomes and the transcriptional activity of genes. These *cis* sequences include distal enhancers that may reside large distances from the gene promoters they control. Variations in these enhancer/promoter interactions and the nuclear localization of the genes they regulate are thought to be contributing factors in the diversity of transcriptional profiles between different cell types. Moreover, they are important in adjusting these profiles throughout

cellular differentiation and development (Ragoczy *et al.*, 2003, *Chromosome Res* 11: 513-525). These long-range associations are facilitated by the looping of chromatin whereby regulatory elements come together with requisite nuclear factors to function within 'transcriptional hubs' where transcriptional activity is coordinated (de Laat and Grosveld, 2003, *Chromosome Res.* 11: 447-459).

[0004] The regulated expression of large human genes can depend on long-range interactions to establish appropriate three-dimensional structures across the locus. For example, the cystic fibrosis (CF) transmembrane conductance regulator (*CFTR*) gene, which encompasses approximately 189 kb of genomic DNA, shows a complex pattern of expression with both spatial and temporal regulation. The flanking loci, *ASZ1* and *CTTNBP2*, show very different tissue-specific expression in comparison to *CFTR*.

[0005] Cystic fibrosis (CF) is one of the most common serious inherited diseases amongst Caucasians. One in 20-25 people harbor a mutation in the CF gene. As CF is an autosomal recessive syndrome this means that around 1 in 2,000 live births are affected. CF affects many organs including sweat glands (cystic fibrosis sufferers have salty sweat), the gut and the pancreas (85% of CF patients are pancreatic insufficient and require enzyme supplements in the diet). However, it is the effects on the lung which are most commonly life-threatening and lead to premature death. Specifically, cystic fibrosis patients accumulate mucus in the airways which is only relieved by regular physiotherapy. This mucus serves as a substrate for bacterial infection resulting in lung damage. The gene for cystic fibrosis was identified and sequenced in 1989 and subsequent studies showed that the gene product is a cAMP-activated chloride channel in the membrane of specialized epithelial cells and some other cell types.

[0006] In humans, the *CFTR* gene encompasses approximately 189 kb at human chromosome 7q31.2 and mutations within it cause the common

genetic disease cystic fibrosis (CF) (Rommens *et al.*, 1989, *Science* 245: 1059-1065). *CFTR* encodes a membrane-associated chloride ion channel that is expressed at the highest levels in chloride-secreting epithelial cells of the small intestine, pancreas, and male genital duct, and at lower levels in the respiratory epithelium and certain other sites.

[0007] The *CFTR* promoter has been characterized as a "housekeeping-like" promoter, and does not possess the regulatory elements responsible for the diverse expression profile of the gene (Chou *et al.*, 1991, *J. Biol. Chem.* 266: 24471-24476; Koh *et al.*, 1993, *J Biol. Chem.* 268:15912-15921; Yoshimura *et al.*, 1991, *J. Biol. Chem.* 266: 9140-9144). Classical methods of chromatin analysis were previously utilized to map potential regulatory elements in a number of cell lines that express the *CFTR* gene and identified several functionally important elements (Smith *et al.*, 1995, *Biochem. Biophys. Res. Commun.* 211: 274-281; Smith *et al.*, 1996, *J. Biol. Chem.* 271: 9947-9954; Nuthall *et al.*, 1999, *Biochem J* 341: 601-611; Nuthall *et al.*, 1999, *Eur. J. Biochem.* 266: 431-443; Smith *et al.*, 2000, *Genomics* 64: 90-96; Phylactides *et al.*, 2002, *Eur. J. Biochem.* 269: 553-559; Harris, U.S. Patent No. 6,573,073).

[0008] The mechanisms governing control of *CFTR* expression remain poorly understood, though they are known to involve intronic regulatory elements. Elucidating the regulatory mechanisms for the *CFTR* gene has been a significant challenge, due in part to the large size of the gene, the lack of tissue-specific control elements in the promoter, and the paucity of relevant cell types for analysis.

BRIEF SUMMARY

[0009] Systems for the regulation of the expression of a *CFTR* gene are disclosed. The systems preferably include an isolated nucleic acid element that is at least 95% identical to the nucleic acid sequence of SEQ ID NO:1, and a *CFTR* gene promoter that is at least 95% identical to the nucleic acid

sequence of SEQ ID NO:2. The systems preferably regulate enhanced expression of the *CFTR* gene. In some embodiments of the systems, the nucleic acid element may be recombinant. In other embodiments of the systems, the *CFTR* gene promoter may be recombinant. Yet in other embodiments of the systems, the *CFTR* gene promoter may be native. In some embodiments of the systems, the isolated nucleic acid element may be operably linked to the *CFTR* gene promoter. The systems may further include isolated nucleic acid elements comprising the *CFTR* gene enhancer sequence of intron 1.

[0010] Isolated nucleic acid elements that include functional sequences are disclosed. The functional sequences are contained within the sequence of SEQ ID NO:1. The nucleic acid elements regulate expression of the *CFTR* gene. Vectors containing these nucleic acid elements are also disclosed.

[0011] Isolated nucleic acid constructs that are up to 50 kbp (kilobase pairs) in length comprising the coding sequence of the *CFTR* gene together with an isolated nucleic acid element that is at least 95% identical to the nucleic acid sequence of SEQ ID NO:1, are disclosed. Also disclosed are vectors that comprise these isolated nucleic acid constructs.

[0012] Disclosed are methods of preparing delivery vectors for delivery of a *CFTR* gene to mammals, including humans. The methods include formulating the above nucleic acid constructs into delivery vectors suitable for administration to the mammals.

[0013] Methods for regulating expression of a *CFTR* gene in mammals are disclosed. The methods include delivering the delivery vectors described above to the mammals, thereby regulating the expression of the delivered *CFTR* gene in the mammals.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] Figure 1 (A, B) is a schematic diagram of the identification of DNase I hypersensitive sites (DHS) within the *CFTR* locus in cell types

relevant to CF pathology: (A) Averaged DNase-chip hybridization data from three (Caco2, skin fibroblasts), two (HT29, primary tracheal epithelial cells and NHBE cells), or a single (primary epididymis) experiment, analyzed with ACME statistical software; (B) *CFTR* mRNA levels measured by qRT-PCR (quantitative reverse transcription- polymerase chain reaction), where each value is relative to the amount of detected skin fibroblast transcript.

[0015] Figure 2 depicts graphs illustrating the long-range interactions between the CFTR promoter and specific DHS measured with q3C (quantitative chromosome conformation capture).

[0016] Figure 3 (A, B) depicts bar charts illustrating how intronic enhancers act cooperatively to augment the *CFTR* basal promoter: (A) Caco2 cells transfected with pGL3B luciferase reporter constructs containing the 787 bp *CFTR* basal promoter and fragments of the DHS regions found in introns 1, 10a,b, 11, or +15.6 cloned into the enhancer site of the vector in either forward or reverse orientations; (B) Caco2 cells transfected with the single enhancer luciferase reporter constructs (as in A), in addition to constructs containing multiple combinations of two or three DHS regions cloned into the enhancer site.

[0017] Figure 4 is a graph showing by chromatin immunoprecipitation (ChIP) how the transcription factor HNF1 and the transcriptional coactivator p300 associate with multiple DHS of the *CFTR* locus in Caco2 cells.

[0018] Figure 5 is a graph illustrating how EcoRI 3C (EcoRI chromatin conformation capture) detects interactions between the *CFTR* promoter and elements 3' of gene.

[0019] Figure 6 is a graph showing enrichment of Rad21 at -20.9, +6.8 and +48.9 kb DHS by ChIP.

DETAILED DESCRIPTION OF THE PRESENTLY PREFERRED EMBODIMENTS

[0020] In one aspect, the present invention relates to compositions and methods for the regulation of the expression of the *CFTR* gene, which encodes cystic fibrosis transmembrane conductance regulator (CFTR). In preferred embodiments, the *CFTR* gene is a human *CFTR* gene.

[0021] In one aspect, the present invention provides for the identification of one or more DNA sequences that interact with the promoter of a *CFTR* gene to regulate its expression. In particular, the present invention provides for the identification of one or more DNA sequences that interact with the promoter of a *CFTR* gene to enhance its expression.

[0022] In some embodiments, the present invention discloses nucleic acid elements as regulatory elements of the expression of the *CFTR* gene. The regulatory elements may be in their forward, 5' to 3' orientation. In other embodiments of the present invention, the regulatory elements may be in a reverse, 3' to 5' orientation.

[0023] "CFTR" means cystic fibrosis transmembrane conductance regulator. Preferably, the CFTR is a human CFTR.

[0024] "*CFTR* gene" refers to the gene that provides the polynucleotide sequence for making a protein called the cystic fibrosis transmembrane conductance regulator (CFTR). In preferred embodiments, *CFTR* gene refers to the human *CFTR* gene. Thus, the present invention specifically contemplates the human *CFTR* gene and the human CFTR protein.

[0025] Some aspects of the present invention are described in Ott *et al.*, 2009, *Proc. Natl. Acad. Sci.*, in press, which is herein incorporated by reference. As well, certain aspects of the present invention are disclosed in U.S. Patent No. 6,573,073 B2, which is herein incorporated by reference.

[0026] In one aspect, the present invention provides evidence that the cystic fibrosis transmembrane conductance regulator (*CFTR*) locus adopts a looped conformation to facilitate expression of the *CFTR* gene. The complex looped structure of the *CFTR* locus is present to varying degrees

in cells that express the *CFTR* gene, and is absent from cells in which the *CFTR* gene is inactive. The promoter region closely interacts with sequences in the middle of the gene about 100 kb from the promoter and with regions 3' to the locus that are about 200 kb away. These interacting regions correspond to prominent DNase I hypersensitive sites within the locus.

[0027] In one aspect the present invention provides regulatory elements outside the basal promoter region of the *CFTR* gene, which contribute to its diverse expression profile.

[0028] In another aspect, the present invention discloses the identification of one or more nucleic acid sequences that interact with the promoter of a *CFTR* gene to enhance its expression in a number of human primary cell types relevant to cystic fibrosis (CF) pathology. These nucleic acid sequences are preferably DNA sequences. Because these nucleic acid sequences enhance the expression of the *CFTR* gene, they may be viewed as *CFTR*-activating enhancers.

[0029] In one aspect, the present invention discloses the use of novel technologies to comprehensively map potential regulatory elements of the *CFTR* locus and to establish their mechanism of action.

[0030] The present invention contemplates the use of a variety of vectors that can be recombinantly engineered to express the nucleic acid elements that act as regulatory elements of the expression of the *CFTR* gene. A variety of vectors can be useful for practicing the invention. The term "vector" is intended to refer to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a plasmid, which refers to a circular double stranded DNA loop into which additional DNA segments may be ligated. Another type of vector is a viral vector, wherein additional DNA segments may be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other

vectors can be integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to as "recombinant expression vectors" (or simply, "expression vectors"). In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" may be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), also bacterial or human artificial chromosomes which serve equivalent functions.

[0031] In some embodiments of the present invention, the identified nucleic acid sequences can be used in reporter gene constructs. One utility of such gene constructs is their use to recruit proteins that modify chromatin structure. In preferred embodiments, the identified nucleic acid sequences can be used cooperatively in reporter gene constructs to recruit proteins that modify chromatin structure.

[0032] In one embodiment, the present invention discloses the use of high-resolution tiled microarrays for the detection of multiple intronic and extragenic DNase I hypersensitive sites (DHS), regions of open chromatin that are depleted of nucleosomes and are often associated with gene regulatory elements. It was discovered that several of these DHS regions bind both tissue-specific and general transcription factors and also possess cooperative enhancer activity *in vitro*. Moreover, *in vivo*, these enhancers interact directly with the *CFTR* promoter region. Thus, in some aspects of the invention, recent advances in methodology to evaluate regulatory elements in the human genome *in vivo*, combined with a biological approach to the expression and function of *CFTR,* have enabled

performing the first in-depth study of the organization of the entire *CFTR* locus.

[0033] In other embodiments of the invention, the properties of key regulatory elements for *CFTR* are demonstrated, as is demonstrated a novel example of a transcriptionally active human gene adopting distinct conformations in different cell types. It is also disclosed that the active *CFTR* locus exists in a complex conformation involving the direct interaction of several intronic enhancers with the promoter. These interactions are most evident in epithelial cells of intestinal lineage where *CFTR* is most highly expressed; however, they are also apparent in epithelial cells derived from the respiratory system, which generally express *CFTR* at lower levels. The very low levels of *CFTR* expression observed in airway cell types (such as those shown in Figure 1B) may be achieved primarily via promoter-mediated mechanisms or by inhibition of promoter-intronic enhancer contacts. In contrast, intestinal and genital duct cells appear to require the contribution of the intronic enhancers in order to maintain high *CFTR* expression. These promoter-enhancer interactions do not occur in skin fibroblasts, which lack *cis* regulatory elements detectable by DHS mapping in the *CFTR* locus.

[0034] In yet other aspects of the invention, disclosed is the use of various human non-epithelial cell types. These can be assayed for DHS using DNase-chip. Examples include, but are not limited to, GM06990 (B lymphoblastoid), primary H9 (undifferentiated embryonic stem cells), CD4+ T cells and K562 (erythroleukemia) (Xi *et al*., 2007, *PLoS Genet* 3: e136; data available at the Encyclopedia of DNA Elements (ENCODE) Consortium web site, University of California Santa Cruz, Santa Cruz, CA). None of these cell types exhibit the DHS within *CFTR* introns 1, 11, or at +15.6 kb that were detected in the various epithelial cell types utilized in this work.

[0035] The existence of an epithelial-specific looped conformation for *CFTR* is consistent with observations in other mammalian genes and gene

loci that achieve tightly regulated transcriptional activation via direct interactions of regulatory elements, including the T-helper 2 (T_H2) cytokine locus, the α- and β-*globin* loci, and others (Spilianakis *et al.*, 2004, *Nat. Immunol.* 5: 1017-1027; Cai *et al.*, 2006, *Nat. Genet.* 38: 1278-1288; Vernimmen *et al.*, 2007, *EMBO J.* 26: 2041-2051; Tolhuis *et al.*, 2002, *Mol. Cell* 10: 1453-1465; Palstra *et al.*, 2003, *Nat. Genet.* 35: 190-194). A novel aspect of the *CFTR* locus is that key interacting elements lie within introns in addition to those distal to the gene. In this context, the model for the active *CFTR* locus described herein predicts that the interaction of multiple enhancers and both tissue-specific and general transcription factors confers the complex regulation of *CFTR* expression. This is consistent with previous data obtained by the inventors, in which removal of the intron 1 DHS region from a human *CFTR* YAC resulted in a 60% decrease in *CFTR* expression levels in a human intestinal cell line and in the intestine (but not the lungs) of transgenic mice carrying the YAC (Rowntree *et al.*, 2001, *Hum. Mol. Genet.* 10: 1455-1464).

[0036] Not wanting to be bound by the following theory, it is possible that removal of multiple enhancer sites and the promoter would be required to completely extinguish *CFTR* expression in any epithelial cell type.

[0037] In one aspect of the invention, it is demonstrated herein that one of the tissue-specific transcription factors interacting with the enhancers in intestinal cells is HNF1, which the inventors previously showed to be involved in *CFTR* expression both *in vitro* and *in vivo*, through interactions at multiple intronic sites (Mouchel *et al.*, 2003, *Biochem. J.* 378: 909-908. HNF1 has been implicated in a similar mechanism of transcriptional regulation in the human *ADH* gene, where it binds a distal enhancer and augments gene expression (Su *et al.*, 2006, *J. Biol. Chem.* 281: 19809-19821).

[0038] In another aspect of the invention, it is demonstrated herein that one of the general transcription factors involved in the *CFTR* transcriptional hub is the widely expressed factor p300. This protein has histone acetyl transferase (HAT) activity and is known to be a marker of transcriptionally active elements including promoters and enhancers. In a recent study of p300-enriched sites using tiled microarrays (Heintzman *et al.*, 2009, *Nature* 459:108-112) it was discovered that several regions close to the intron 11 enhancer are associated with p300 binding in several cell *CFTR*-expressing cell types, corroborating the data from Caco2 cells. The presence of p300 in the *CFTR* transcriptional unit is consistent with the inventors' previous work demonstrating increased histone acetylation at certain regulatory elements within the locus (Blackledge *et al.*, 2007, *Biochem. J.* 408: 267-275).

[0039] Some major advances in the understanding of regulatory mechanisms for the *CFTR* gene that arise from this work were enabled by the application of relatively novel technologies including DNase-chip and 3C (chromatin conformation capture). A comparison of the power of these techniques with classical methods is warranted, particularly with respect to DHS mapping. Previously, a Southern blot-based method was used to screen for DHS across 400 kb spanning the *CFTR* locus in a number of cell lines, and thus a number of important regulatory elements within the locus were identified (Smith *et al.*, 2000, *Genomics* 64: 90-96; Phylactides *et al.*, 2002, *Eur. J. Biochem.* 269: 553-559). While the experiments using DNase-chip were able to reproducibly and reliably detect DHS of the locus that functionally associate with the *CFTR* promoter, it was not possible to detect several DHS that the inventors observed in earlier studies. These include the site at -20.9 kb from the *CFTR* translational start site (Smith *et al.*, 1995, *Biochem. Biophys. Res. Commun.* 211: 274-281; Nuthall *et al.*, 1999, *Eur. J. Biochem.* 266: 431-443) that was evident in many cell types and in human YAC-transgenic mice (Moulin *et al.*, 1999, *Mol. Med.* 5: 211-223); also, a cluster of four sites +6.8 to +7.4 kb distal to the translational

stop site of *CFTR* (Nuthall et al., 1999, *Biochem. J.* 341: 601-611). The 20.9 kb and +6.8 kb sites both possess enhancer-blocking insulator activity and associate with CCCTC-binding factor (CTCF) (Blackledge *et al*., 2007, *Biochem. J.* 408: 267-275; Blackledge *et al*., 2009, *Nucleic Acids Res.* 37: 1068-1094). The lack of detection of the -20.9 kb and +6.8 kb DHS in the current studies is likely due to their locations within or adjacent to repetitive regions, which are excluded from the tiled arrays; alternatively, the classical method of DHS mapping may be more sensitive than DNase-chip. Cell lines may also have evolved in culture and this process was accompanied by changes in the chromatin structure. Many groups have noted the loss of *CFTR* expression with prolonged culture of epithelial cell lines, which would be consistent with gradual changes in chromatin structure and modification. This phenomenon reinforces the approach of using primary human cell cultures in the definition of regulatory elements for *CFTR*.

[0040] The potential involvement in *CFTR* transcription of DHS flanking the *CFTR* gene warrants further comment, particularly since several sites bind CTCF and may contribute to the chromatin structure of the active locus. In primary genital duct cells, that exhibit the +6.8 kb DHS, corresponding to a functional CTCF-binding site, this region may mark the end of the active domain. In contrast, other cells, such as Caco2, which lack the +6.8 kb DHS and show very low occupancy of CTCF in this region may recruit the +15.6 kb DHS region and the end of the looped domain could then be provided by a more distal CTCF binding site at (+48.9 kb:116950000). The precise definition of elements at the 3' end of the locus that interact with the *CFTR* promoter cannot be provided by the 3C data presented here, since one *Hind* III fragment encompasses both the +6.8 kb region and the +15.6kb DHS. In further experiments using 3C libraries generated with EcoRI (Figure 5), which cleaves between +6.8 and +15.6 (at +11.2 kb) and also 3' to the +15.6 kb DHS (at +17.2 kb and 20.1 kb) all these fragments interact with the CFTR promoter bait,

demonstrating their close proximity in the active locus and suggesting that they could potentially contribute to a transcriptional hub. Consistent with these data is the observed binding of Rad21 (a subunit of the cohesin complex known to interact with CTCF on chromatin) at the -20.9 kb, +6.8 kb and +48.9 kb sites (Figure 6).

[0041] In some embodiments, the present invention provides as a new compound a regulatory nucleic acid element, which regulates the expression of the *CFTR* gene. The nucleic acid element is preferably located within intron 11 of the *CFTR* gene. In one preferred embodiment, the sequence of the regulatory nucleic acid element includes the 1,514 nucleotides located at positions 116,821,781-116,823,295 on human chromosome 7, described herein as the nucleic acid sequence of SEQ ID NO:1. In other preferred embodiments, the sequence of the nucleic acid element includes at least 75% of the 1,514 nucleotides of SEQ ID NO:1, preferably at least 85% of the 1,514 nucleotides of SEQ ID NO:1, and more preferably at least 95% of the 1,514 nucleotides of SEQ ID NO:1.

[0042] Yet in other preferred embodiments, one or more parts of the regulatory nucleic acid element are located in intron 11, whereas other part or parts of the sequence are located in an adjacent exon, i.e. in one or more of sequences flanking intron 11. Preferably the adjacent exon is exon 11 of the *CFTR* gene. Thus, for example, a preferred 1,514 nucleotides long nucleic acid element (such as the one shown in SEQ ID NO:1) may include 1,457 nucleotides from intron 11, and 57 nucleotides from exon 11. The nucleic acid element preferably interacts with the promoter of the *CFTR* gene, regulating the expression of the *CFTR* gene. More preferably, the nucleic acid element enhances the expression of the *CFTR* gene.

[0043] Since intron flanking sequences may also exert a subsidiary regulating function, the invention also includes: Nucleic acid elements containing up to 200 nucleotide residues including at least one defined oligonucleotide and having at least 70% homology with part or all of the sequence shown in SEQ ID NO:1; nucleic acid elements that are capable

of controlling expression of the human *CFTR* gene and which have at least 70% homology with part or all of the sequence shown in SEQ ID NO:1; nucleic acid elements that are capable of controlling expression of the *CFTR* gene and which include a functional sequence at least about 10 nucleotides in length, preferably about 50 nucleotides in length, and more preferably about 250 nucleotides in length and contained within one of the sequences set out above; nucleic acid elements that are capable of regulating expression of the *CFTR* gene and which are derived exclusively or mainly from intron 11 of the *CFTR* gene, in particular from the 1,514 bp intron 11 sequence as shown in SEQ ID NO:1.

[0044] In other embodiments, one or more of the above nucleic acid elements is used as a part of a system for the regulation of the expression of the *CFTR* gene. In some embodiments, one or more regulatory nucleic acid elements are used in conjunction with the promoter of the *CFTR* gene, to regulate the expression of the *CFTR* gene. The sequence of the *CFTR* promoter is shown as the 787 nucleotides long sequence of SEQ ID NO:2. In other embodiments, the sequence of the *CFTR* promoter includes at least 75% of the nucleic acid sequence of SEQ ID NO:2, preferably at least 85% of the nucleic acid sequence of SEQ ID NO:2, and more preferably at least 95% of the nucleic acid sequence of SEQ ID NO:2. In various aspects, the present invention contemplates combinations of various embodiments of the nucleic acid element disclosed above and various embodiments of the *CFTR* promoter.

[0045] In some embodiments the present invention provides novel DHS that were detected in the *CFTR* gene, both in cell lines that had been evaluated previously by classical methods and in primary tracheal epithelial cells that were investigated here for the first time. A novel DHS lies in intron 11. This enhancer lies almost exactly at the midpoint of the locus, and its association with the promoter results in the formation of two ~100 kb chromosome loops. One skilled in the art can further determine

the molecular basis of the enhancer activity in the intron 11 DHS element and additional transcription factors that directly bind to this region.

[0046] The inventors previously demonstrated that HNF1 binds directly to the tissue-specific enhancer in the intron 1 DHS (Ott *et al.*, 2009, *J. Cell. Mol. Med.* 13: 680-692) and this factor is also a strong candidate for involvement at the intron 11 DHS element since the ~1.5 kb region contains a number of predicted binding sites for HNF1 and is enriched by ChIP with an antibody specific for this factor. Not wanting to be bound by the following theory, it is possible that in certain cell types HNF1 recruits and coordinates the other nuclear factors necessary for *CFTR* transcription. Therefore, in some embodiments, the newly discovered enhancer sequence of intron 11 may be used in conjunction with the enhancer sequence of intron 1 (see e.g. Harris, U.S. Patent No. 6,573,073), and in conjunction with the *CFTR* promoter, to effectuate regulation of the expression of the *CFTR* gene.

[0047] In one aspect, the invention describes the analysis of elements associated with DHS that show enhancer activity and that interact directly with the *CFTR* promoter. However, a number of other DHS were identified in each of the cell types that do not encompass such elements. The inventors previously showed that some of these sites had modest enhancer activity (Phylactides *et al.*, 2002, *Eur. J. Biochem.* 269: 553-559), though others appear to be associated with cell type-specific *CFTR* expression by alternative mechanisms, as yet unidentified. It is also possible that some of these DHS may contain regulatory elements for other genes, particularly since many of the sites do not show a direct correlation with *CFTR* expression. The genes that flank *CFTR*, *ASZ1* on the 5' side and *CTTBP2* on the 3' side, have different expression profiles than *CFTR*. Moreover, while embodiments of the invention described herein are mainly focused on regulatory elements detected within the *CFTR* locus itself, it is possible that other *cis* and/or *trans* elements outside this region also coordinate regulation of the *CFTR* gene.

[0048] In one aspect, the present invention includes unexpected results. The data in Figure 3, discussed below, illustrate how according to the present invention the identified regulatory elements, i.e. intronic enhancers, act cooperatively to augment the *CFTR* basal promoter.

[0049] The diversity in DNase-chip data generated from the airway-derived cell types is of interest. Not wanting to be bound by the following theory, it may suggest different regulatory mechanisms in these cells. The NHBE cells, which are a mixture of bronchial and tracheal epithelial cells, provide evidence for similar CFTR regulatory mechanisms to those of intestinal epithelial cells. In contrast, the pHTE cultures of tracheal cells alone exhibit a different set of DHS and moreover, show only slight evidence for intra-locus looping. For example, each of these lung cell types can be investigated further, in order to dissect their *CFTR* regulatory pathways.

[0050] In one aspect of the invention, the model for CFTR gene expression that is revealed by the data presented herein provides a paradigm for other large genes with multiple regulatory elements lying within both introns and intergenic regions.

[0051] In one aspect of the invention, understanding the complex mechanisms by which large, disease-associated genes maintain their transcriptional activity *in vivo* provides novel avenues for therapy development. The observations disclosed herein can enable novel approaches to designing regulated transgenes for tissue-specific gene therapy protocols. For example, endogenous enhancers such as those identified herein may be incorporated into gene therapy vectors in order to provide stable expression, at appropriate levels, in the relevant cell types. Moreover, identification of the specific transcription factors acting upon genes such as *CFTR* will begin to reveal the cellular signaling pathways involved in regulating transcript levels, which might also be targeted for therapeutic benefit.

[0052] In some embodiments, the present invention contemplates gene therapy. The *CFTR* gene is about 189 kb in length. The coding sequence comprises 27 exons and is in total about 6.5 kb in length. Preferably, in gene therapy, most or all of the coding sequence of a gene is delivered into target cells. In some embodiments of the present invention, regulatory regions and/or introns or intron fragments may be delivered to cells in the context of the coding sequence. Thus, when a regulatory element (intronic enhancer) according to the present invention is used in conjunction with the *CFTR* gene, the expression of the *CFTR* gene is greatly enhanced. It is known that the CFTR gene promoter by itself is not very active.

[0053] A variety of gene delivery systems may be used for cystic fibrosis gene therapy, including, e.g., adenoviruses, liposomes, and retroviral vectors. Whatever the delivery system used, the idea of gene therapy is that the DNA in question is introduced into human cells, in an episomal (non-integrated) or integrated form, and will there express the CFTR protein which acts as a chloride ion channel. The DNA administered in the course of gene therapy will in principle be expressed in all human cells in which it becomes incorporated. This is because the viral and other promoters hitherto used do not have any cell specificity. But the desired effect of gene therapy is that the *CFTR* gene should be expressed in, and only in, those cells where it would be expressed in a normal healthy individual.

[0054] In another aspect, the invention provides a nucleic acid construct up to 15 kbp in length comprising the coding sequence of the *CFTR* gene together with a nucleic acid element as described above. The specified maximum length of 15 kbp is not critical, and longer constructs are envisaged which are, however, much shorter than the complete *CFTR* gene. Preferably the length of the construct is such that the construct can conveniently be incorporated in a viral or other vector for administration.

[0055] The nucleic acid regulator that is provided in this construct is preferably present between exon 11 and exon 12 of the coding sequence

of the *CFTR* gene. However, elements of this kind sometimes have powerful long-distance effects that are independent of their position or orientation in the gene. So it is envisaged that this element may alternatively be present, together with the promoter upstream of exon 1, or adjacent to some other pair of exons comprising the gene, or even at the downstream end of the coding sequence of the gene, provided that the regulator is positioned so as to be capable of modulating expression of the CFTR gene.

[0056] Preferably a *CFTR* gene promoter sequence is also present, operably linked to the *CFTR* gene and generally in a functional position at the upstream end of the coding sequence of the gene.

[0057] In some aspects, the invention provides a vector containing a nucleic acid construct as defined; and a method of treating by gene therapy a patient suffering from cystic fibrosis which method comprises administering to the patient this vector.

[0058] In yet other aspects, the invention further provides a method of preparing an agent for treating cystic fibrosis by gene therapy, which method comprises bringing a nucleic acid construct as defined into a form suitable for administration. For this purpose, the following steps are envisaged:1) Place the CFTR cDNA in a vector where it is driven by its own promoter rather than by a viral one, despite the relative weakness of the CFTR promoter; 2) Modify the CFTR cDNA by PCR-mutagenesis to insert the splice donor site at the end of exon 11 (within intron 11) and the splice acceptor site of exon 12 (also within intron 11). At the same time also insert about 1500 bp, preferably about 1000 bp, and more preferably about 600 bp of known intron 11 sequence, about 150 bp adjacent to exon 11 and about 150 bp adjacent to exon 12, and in the centre the regulatory element. This construct would then, due to the presence of the relevant splice sites, be capable of splicing out the inserted mini-intron 11 and generate a normal full length CFTR cDNA in human cells. The construct would be tested for this ability.

[0059] An alternative stratagem would be to insert the regulatory element directly into the vector backbone either upstream of the CFTR promoter or downstream of the CFTR cDNA.

[0060] Details of vectors which may be useful for gene therapy for CFTR, and of their construction, can be found in the published literature, e.g. Boucher *et al.*, 1994, *Human Gene Therapy* 5: 615-639 (adenovirus vectors); Zabner *et al.*, 1993, *Cell* 75: 205-216 (adenovirus vectors); Caplen *et al.*, 1994, *Gene Therapy* 1: 139-147 (liposome-mediated DNA transfer); Caplen *et al.*, 1995, *Nature Med.* 1: 39-46 (liposome-mediated DNA transfer). In accordance with the present invention, similar strategies may be used, but employing the regulatory element described herein and the *CFTR* gene promoter. The literature also describes transfer of the *CFTR* gene to target cells, both *in vitro* and *in vivo*, including suitable pharmaceutically acceptable formulations and protocols for administration.

[0061] Although the invention is described herein in relation to the human *CFTR* gene and its regulators, animal models are not excluded. The regulatory elements, nucleic acid constructs, vectors and uses of all of these may therefore be from and/or for use in non-human mammals.

EXAMPLES

[0062] Cell culture. The human colon carcinoma cell lines Caco2 and HT29, primary skin fibroblasts (ATCC# GM08333) were grown by standard methods. Primary human fetal male epididymis cells and primary tracheal epithelial cells were cultured as previously described (Harris *et al.*, 1991, *Development* 113: 305-310; Davis P *et al.*, 1990, *Am. J. Physiol.* 258: C71-C76). Normal Human Bronchial Epithelial (NHBE) cells were a mixture of primary human bronchial and tracheal epithelial cells (Lonza, CC-2541).

[0063] DNase-chip. DNase-chip was performed as previously described (Crawford *et al.*, 2006, *Nat. Methods* 3: 503-509), with modifications (Ott *et al.*, 2009, *J. Cell. Mol. Med.* 13: 680-692) and experiment analyzed with ACME statistical software (Scacheri *et al.*, 2006, *Meth. Enzymol.* 411: 270-282). DNase-chip data will be publicly available at the Encyclopedia of DNA Elements (ENCODE) Consortium web site, University of California Santa Cruz, Santa Cruz, CA (The ENCODE Project Consortium, 2007, *Nature* 447: 799-816) and will also be deposited at the Gene Expression Omnibus (GEO) database, National Center for Biotechnology Information (NCBI), U.S. National Library of Medicine, Bethesda, MD.

[0064] qRT-PCR. *CFTR* expression was assayed as described previously using Taqman® primer/probe set spanning *CFTR* exons 5 and 6 (TAQEX5/6) (Mouchel *et al.*, 2003, *Biochem. J.* 378: 909-908).

[0065] Quantitative chromosome conformation capture (q3C). 3C was performed as described previously (Hagège *et al.*, 2007, *Nat. Protocols* 2: 1722-1733), with minor modifications (Blackledge *et al.*, 2009, *Nucleic Acids Res.* 37: 1068-1094).

[0066] Transient promoter/enhancer reporter assays. Sequences encompassing the DHS in introns 10a,b, 11, and at +15.6, were amplified by PCR using *Pfu* DNA polymerase (Stratagene, La Jolla, CA). Primers are shown in Table 2. Reporter assays were performed by standard methods using a reporter gene construct driven by the 787 bp CFTR minimal promoter (Smith *et al.*, 1996, *J. Biol. Chem.* 271: 9947-9954; Phylactides *et al.*, 2002, *Eur. J. Biochem.* 269: 553-559).

[0067] Chromatin immunoprecipitation (ChIP). ChIP was performed as described previously (Ott *et al.*, 2009, *J. Cell. Mol. Med.* 13: 680-692).

Immunoprecipitations were performed with antibodies specific for HNF1 (Santa Cruz, sc-8986), p300 (Santa Cruz, sc-585X), RAD21 Abcam (ab992-50) or IgG (Santa Cruz, sc-2027).

[0068] Primer sequences. The primer sequences and locations used for RT-PCR, plasmid cloning, 3C, and ChIP are listed in Table 2.

[0069] Detection of DNase I hypersensitive sites across the *CFTR* locus. DNase-chip was used to identify DHS in a number of cell types relevant to *CFTR* expression including primary human tracheal and bronchial epithelial cells, primary human fetal epididymis epithelial cells, and the human colon carcinoma cell lines Caco2 and HT29, all of which express *CFTR*. Human skin fibroblasts that do not express *CFTR* were also evaluated, to provide an example of the chromatin structure of the transcriptionally inactive *CFTR* locus.

[0070] Figure 1 shows identification of DHS within the *CFTR* locus in cell types relevant to CF pathology. Figure 1(A) shows averaged DNase-chip hybridization data from three (Caco2, skin fibroblasts), two (HT29, primary tracheal epithelial cells and NHBE cells), or a single (primary epididymis) experiment was analyzed with ACME statistical software (Scacheri *et al.*, 2006, *Meth. Enzymol* 411: 270-282). A major DHS was identified at the *CFTR* promoter (Pr) in all cells that express the gene; several specific DHS of interest were detected including those in intron 1 (Int1) and intron 11 (Int11). A DHS at +15.6 kb 3' to the *CFTR* translational stop site (Nuthall *et al.*, 1999, *Biochem. J.* 341: 601-611; Blackledge *et al.*, 2007, *Biochem. J.* 408: 267-275) was seen in several cell types and at +48.9 kb a ubiquitous DHS (Ubiq) is marked in the last intron of the *CTTBP2* gene. The zero point of the x-axis represents the beginning of the first *CFTR* exon. The y-axis for each DHS track represents -log10(P-value) between 0-16 as determined by ACME. Figure 1(B) shows *CFTR* mRNA

levels measured by qRT-PCR; each value is relative to the amount of detected skin fibroblast transcript. Error bars represent SEM, n=3.

[0071] Three DNase-chip experiments were performed on independent cultures of skin fibroblasts and Caco2 cells, two experiments were carried out on primary tracheal and NHBE (bronchial and tracheal) epithelial cells and HT29 cells and one experiment evaluated primary epididymis cells (Figure 1A). The DNase hypersensitivity tracks represent averaged data where multiple experiments were performed. The data demonstrate that each cell type has a specific and unique pattern of DHS along the locus. The promoter region of the gene is hypersensitive in each *CFTR*-expressing cell type, but is DNase–resistant in the non-expressing skin fibroblasts. Outside the extended promoter region multiple DHS were identified, including those that are ubiquitous, common to several cell types, or cell-type specific.

[0072] No DHS were evident in the *CFTR* locus in skin fibroblasts, which correlates with the gene being transcriptionally inactive in these cells. In airway epithelial cells, that express low levels of CFTR, several DHS were seen, though these were not consistent between the different cell types. In primary human tracheal epithelial cells, several distal DHS were detected: in intron 18 (3600 + 10 kb, 3600 is the last base in exon 18 and the site maps 10 kb into the next intron), intron 19 (3849 + 12.5 kb), and intron 23 (4374 + 1.3 kb). The precise chromosome 7 coordinates of each DHS region on the hg 17 build are shown in Table 1. DHS were also detected 3' to the locus at +15.6 kb, +21.5 kb, +36.6 kb and +48.9 kb from the translational stop. In the NHBE cells, which are a mixture of primary bronchial and tracheal epithelial cells, strong DHS were detected in intron 11 (1811 + 0.8 kb) and at +15.6 kb, in addition to the +21.5 kb, 36.6 kb and 48.9 kb sites. The intron 11 DHS was also detected in Caco2, HT29, and primary epididymis cells. Both colon carcinoma cell lines, Caco2 and HT29, also showed the +15.6 kb DHS. A DHS in intron 1 (185 + 10 kb) that the inventors previously showed to contain an intestinal-specific enhancer

element *in vitro* and *in vivo* was detected in both Caco2 and HT29 cells. Additional cell-line specific DHS are evident in intron 10 (1716 + 13.2 kb) in Caco2 cells and intron 18 (3600 + 1.6 kb) in HT29. The intron 10 site encompasses two closely spaced DHS at 1716 + 13.2 kb and +13.7 kb, that were characterized in the inventors' previous work (Smith *et al.*, 2000, *Genomics* 64: 90-96; Phylactides *et al.*, 2002, *Eur. J. Biochem.* 269: 553-559; Mouchel *et al.*, 2003, *Biochem. J.* 378: 909-908). Also of interest are the DHS located -35 kb and -44 kb with respect to the *CFTR* translational start site in primary epididymis cells. These upstream DHS are closer to the neighboring gene *ASZ1* (ankyrin repeat, SAM and basic leucine zipper) and may be involved in its regulation.

[0073] The relative *CFTR* expression levels in each cell type were determined using RNA isolated at the same time they were harvested for DNase-chip (Figure 1B). With the exception of skin fibroblasts, which do not express *CFTR,* nor exhibit any DHS in the locus, there is no correlation between the frequency or the location of the DHS and the abundance of *CFTR* mRNA.

[0074] The *CFTR* locus is organized in a complex looped structure *in vivo.* Quantitative chromosome conformation capture (q3C), as described in Hagège *et al.*, 2007, *Nat. Protocols* 2: 1722-1733, was used to test the hypothesis that *cis*-acting elements located within DHS interact with each other and the *CFTR* promoter to regulate gene expression (Figure 2).

[0075] Figure 2 shows long-range interactions between the *CFTR* promoter and specific DHS measured with q3C. The organization of the *CFTR* locus is displayed above the graphs. Each assayed *Hind* III fragment is represented by a gray bar with the restriction sites and primer locations along the top of each graph. The 'bait' region of the *CFTR* promoter is schematically shown, which includes a primer and Taqman probe adjacent to the 5' *Hind* III site; this *Hind* III fragment spans the identified *CFTR* transcriptional start sites. The x-axis in each graph

represents position relative to the translational start site; the y-axis represents interaction frequency relative to the interaction frequency between two *Hind* III fragments within the ubiquitously expressed *ERCC3* gene. Below each graph the major DHS of each cell type (see Figure 1) are represented by bars. Data for each cell type are from a single representative 3C experiment (each experiment performed at least twice), error bars represent SEM of at least 2 PCR reactions for each fragment.

[0076] This enabled interrogation of the three-dimensional organization of the locus in the cell types in which DHS were mapped by DNase-chip. Formaldehyde-crosslinked nuclei from each cell type were subjected to *Hind* III digestion and subsequent q3C analysis. A fixed Taqman® probe and reverse primer were designed within a *Hind* III fragment at the *CFTR* promoter (bait), and multiple forward primers were generated within distal regions across the *CFTR* locus (Figure 2 and Table 2). The forward primers were located close to the 3' end of individual *Hind* III fragments that encompassed relevant DHS containing potential *cis*-acting regulatory elements and also other *Hind* III fragments spaced at regular intervals across the locus. Real-time PCR reactions using the reverse probe/primer and each of the forward primers enabled quantification of ligation events (subsequently referred to as 'interaction frequency') between the *CFTR* promoter and specific distal regions within each sample. Hence, the forward primers in *Hind* III fragments located near DHS enabled us to measure directly whether elements within these regions were physically associated with the *CFTR* promoter region.

[0077] In skin fibroblasts, no significant interactions were detected between the promoter and any other region of the locus. These cells show a characteristic pattern of interaction frequency that gradually decreases for fragments at increasing distances from the 'bait' region, signifying a decrease in random ligation events. In contrast, in the intestinal cell types (Caco2 and HT29), the primary epididymis cells, and the NHBE cells,

interactions between the promoter and *Hind* III fragment were detected close to and spanning the intron 11 DHS and regions 3' to the gene. These associations were strongest in Caco2 cells in which the three DHS in intron 10 also demonstrated a high interaction frequency with the promoter. The DHS in intron 10 may be directly interacting with the promoter or exhibiting a bystander effect due to the stronger association of the intron 11 DHS with the promoter as these DHS map within about 17 kb of each other. The interactions were moderate in HT29, primary epididymis, and NHBE cells. Several of these interactions were weakly evident in the primary tracheal epithelial cells which display a different pattern of DHS. Although other intronic DHS exist in these cells, none were found to interact with the promoter region.

[0078] An element within *CFTR* intron 11 is a strong transcriptional enhancer. The DNase-chip analysis revealed a strong DHS within *CFTR* intron 11 at 1811 + 0.8 kb in several cell types. Since the inventors previously demonstrated enhancer function associated with several intronic DHS in the *CFTR* gene (Smith *et al*., 1996, *J. Biol. Chem.* 271: 9947-9954; Phylactides *et al*., 2002, *Eur. J. Biochem.* 269: 553-559; Rowntree *et al*., 2001, *Hum. Mol. Genet.* 10: 1455-1464; Ott *et al*., 2009, *J. Cell. Mol. Med.* 13: 680-692), the inventors tested whether the intron 11 DHS exhibited similar properties.

[0079] Figure 3 shows how regulatory elements, i.e. intronic enhancers in this case, act cooperatively to augment the *CFTR* basal promoter. Figure 3A: Caco2 cells were transfected with pGL3B luciferase reporter constructs containing the 787 bp *CFTR* basal promoter and fragments of the DHS regions found in introns 1, 10a,b, 11, or +15.6 cloned into enhancer site of the vector in either forward or reverse orientations. Figure 3B: Caco2 cells were transfected with the single enhancer luciferase reporter constructs (as in Figure 3A), in addition to constructs containing multiple combinations of two or three DHS regions cloned into the

enhancer site. For both Figure 3A and Figure 3B data are shown relative to the *CFTR* basal promoter-alone vector; error bars represent standard errors of the mean (n=6). * P<0.01 in comparison to the CFTR promoter-only vector, by using unpaired *t* tests.

[0080] The entire region flanking this site was cloned, in forward and reverse orientations, into the enhancer site of a luciferase reporter vector (Phylactides *et al.*, 2002, *Eur. J. Biochem.* 269: 553-559) in which 787 bp of the *CFTR* minimal promoter drives luciferase expression. These constructs were evaluated for firefly luciferase expression following transient transfection into Caco2 cells (Figure 3A). Positive controls included the intron 1 DHS region enhancer in both orientations and negative controls were provided by the intron 10a,b DHS that was previously demonstrated to have no enhancer activity (Phylactides *et al.*, 2002, *Eur. J. Biochem.* 269: 553-559). The fragment encompassing the DHS in intron 11 enhanced *CFTR* promoter activity almost 20-fold in the forward and reverse orientation. The intron 1 enhancer (Ott *et al.*, 2009, *J. Cell. Mol. Med.* 13: 680-692) acts as a modest enhancer with about 5-fold effect on the *CFTR* promoter, while the +15.6 kb DHS region does not show enhancer activity.

[0081] Assays were performed to identify whether these intronic elements act cooperatively to influence the *CFTR* promoter. To achieve this, constructs were built in which combinations of two or three DHS regions were cloned into the enhancer site of the vector in linear order equivalent to their locations with the *CFTR* gene. The intron 10a,b DHS region had no effect on luciferase activity when combined with the other enhancer elements. However, in all other cases, the combination of two or three DHS elements in the same construct had a cooperative effect on the *CFTR* promoter (Figure 3B). The combination of the intron 1 DHS element with the intron 11 DHS enhancer doubled the *CFTR* promoter activity in comparison to the intron 11 element alone. Similarly, the combination of the intron 1 DHS element with the +15.6 kb DHS element had a

cooperative effect on luciferase expression, as did joining the intron 11 DHS element with the +15.6 kb DHS fragment, though alone the latter element does not act as an enhancer. Maximal *CFTR* promoter activity was seen when all three elements, the enhancers in introns 1 and 11 and the +15.6 kb element, were combined in the same construct. In contrast, when the middle fragment in the enhancer site was the intron 10a,b DHS, this additional enhancement was not evident and luciferase expression was similar to the constructs containing only the intron 1 DHS element. These data demonstrate that intronic enhancers within the *CFTR* locus act cooperatively to augment promoter activity in intestinal epithelial cells.

[0082] *In vivo* association of HNF1 and p300 with *CFTR* regulatory elements. The mechanism of interaction between these intronic enhancer sequences and the *CFTR* promoter is of significant interest. Not wanting to be bound by the following theory, it is likely to involve multiple protein complexes, including tissue-specific transcription factors, and general factors including chromatin remodeling enzymes and proteins associated with the nuclear scaffold (Ott *et al.*, 2009, *J. Cell. Mol. Med.* 13: 680-692; Mouchel *et al.*, 2003, *Biochem. J.* 378: 909-908; Blackledge *et al.*, 2007, *Biochem. J.* 408: 267-275; Blackledge et al., 2009, *Nucleic Acids Res.* 37: 1068-1094). The inventors previously showed that hepatocyte nuclear factor 1 (HNF1) binds directly to the core element within the intron 1 DHS and is necessary for enhancer activity (Ott *et al.*, 2009, *J. Cell. Mol. Med.* 13: 680-692). To provide further evidence for the close interaction between the intronic enhancer elements, ChIP experiments were performed with formaldehyde-crosslinked Caco2 chromatin and an HNF1-specific antibody to determine enrichment of enhancer regions (Figure 4).

[0083] Figure 4 shows how HNF1 and p300 associate with multiple DHS of the *CFTR* locus in Caco2 cells. Real-time PCR analysis of Caco2 chromatin immunoprecipitated with indicated antibodies at various locations along the locus. Each value shown is relative to enrichment

measured with isotype-matched IgG control (dotted line). Data for each antibody are from a single representative ChIP experiment (each experiment performed at least twice). All data points were calculated as percentage of input material and then normalized to background 18s rRNA levels; error bars represent SEM of at least 2 PCR reactions for each fragment. Black bars below graphs represent locations of Caco2 DHS (see Figure 1).

[0084] Strongest enrichment was observed at the intron 1 DHS enhancer and modest but significant enrichment was observed at each of the additional intronic DHS observed in Caco2 cells, which suggests either direct or indirect association with HNF1. These data were confirmed by ChIP-chip analysis. ChIP experiments were also performed with an antibody for the histone acetyltransferase p300, a chromatin-associated enzyme found at transcriptionally active elements (Heintzman *et al.*, 2007, *Nat. Genet.* 39: 311-318). In these experiments, p300 was significantly enriched only at each Caco2 intronic DHS, the +15.6 kb DHS, and the *CFTR* promoter. Thus, in Caco2 cells the direct interaction between the intronic DHS and *CFTR* promoter involves a protein scaffold that includes the tissue-specific transcription factor HNF1 and the general transcriptional activator p300 to coordinate the active *CFTR* locus.

[0085] Figure 5 is a graph illustrating how EcoRI 3C detects interactions between *CFTR* promoter and elements 3' of the gene. 3C was carried out as described except using EcoRI instead of HindIII. Data shown are from *CFTR*-expressing HT-29 cells. Primers used: Bait Rev (5' – TACATTCATAACCCCTCTTCCCTACA-3'; SEQ ID NO:3); Bait Probe (5'-TTCAAAGGAAAACATAAGATGCA-3'; SEQ ID NO:4); A1 Forw (5'-CCTGGCACCTGAGGGTACTTC-3'; SEQ ID NO:5); A2 Forw (5'-TTCTACAACTGCACATACAAACTTCAA-3'; SEQ ID NO:6); A3 Forw (5'-CGAAGGGAGGTATTATTCATATGGA-3'; SEQ ID NO:7); B1 Forw (5'-GAGATGTGCCTGATGAATTACAGAA-3'; SEQ ID NO:8); B2 Forw (5'-AAAAAGGCACTTTGCAAGTAAAGAA-3'; SEQ ID NO:9).

[0086] Figure 6 is a graph showing enrichment of Rad21 at -20.9, +6.8 and +48.9 kb DHS by ChIP. Experiments were carried out as described using Caco2 chromatin and rabbit anti-Rad21 antibody. All primer sets used are listed in Table1, except the two most 3' sets located at: +33.5 kb (+33.5 kb Forw; 5'-TTTTTTCAGCTTGTTCATGAGTGAT-3'; SEQ ID NO:10); (+33.5 kb Rev; 5'-GAATAATACCTCCCTTCGACTGAGAA-3'; SEQ ID NO:11); and +48.9 kb (+48.9 kb Forw; 5'-GGCATCAGCCAGTCAAGGTT-3'; SEQ ID NO:12); (+48.9 kb Rev; 5'-AGCAGAGGGCAAAGTGGTACTT-3'; SEQ ID NO:13).

Table 1. DHS locations

DHS	Location in CFTR locus	hg17 coordinates
-44 kb*	5'	chr7: 116,669,500-116,670,700
-35 kb*	5'	chr7: 116,678,400-116,680,000
185 + 10 kb	intron 1	chr7:116,723,600-116,724,700
1716 + 13.2/13.7 kb	intron 10	chr7: 116,806,300-116,807,600
1716 + 23 kb	intron 10	chr7: 116,816,300-116,817,000
1811 + 0.8 kb	intron 11	chr7: 116,822,000-116,823,400
3600 + 1.6 kb	intron 18	chr7: 116,850, 100-116,851,100
3600 + 10 kb	intron 18	chr7: 116,858,000-116,858,500
3849 + 12.5 kb	intron 19	chr7: 116,873,900-116,874,800
4374 + 1.3 kb	intron 23	chr7: 116,899,700-116,901,100
+15.6 kb †	3'	chr7 :116,916,600-116,918,100

+21.5 kb †	3'	chr7 :116,922, 100-116, 923,100
+36.6 kb †	3'	chr7 :116,937,100-116,938,300

[0087] In Table 1, * denotes distance upstream of the CFTR translational start site.

[0088] In Table 1, † denotes distance from last coding base (4574) in the CFTR transcript.

Table 2. Primer sequences

Probe name	**Sequence**	**SEQ ID NO:**
qRT-PCR		
TAQEX5 F	AGCTGTCAAGCCGTGTTCTAGATA	14
TAQEX6a R	ATGAGGAGTGCCACTTGCAAA	15
Ex5/6a Probe	FAM-CACACGAAATGTGCCAATGCAAGTCCTT-TAMRA	16
18s rRNA F	CGGCTACCACATCCAAGGAA	17
18s rRNA R	GCTGGAATTACCGCGGCT	18
18s rRNA Probe	FAM-TGCTGGCACCAGACTTGCCCTC-TAMRA	19
q3C		
-41.7	TGGACCTGGGTACATCTGTC	20
-36.8	TGTTATGTGCTGGTCGAAGA	21
-20.9	CCAAATACTGGACAAAAATGAACTTC	22
-3	TGCCTGGCACAGAGATGGA	23
S2	CATTGTGTCAGCAAAGAGCTTAGG	24

S3	GGTTTAAATATGTGCAAGGAGGATGTC	25
S4b	TTCTGGGATTCACTGTGGAAT	26
S6	GACTCTGCCCTTGACAAATTAACTAA	27
Int3	TTTTCCCTAGCCTGTCAATTGC	28
Int8	GCTTCCTCCTGATCAATCTTTAGG	29
Int10(1)	ACCCTATATCATTGCCCTTTGTATG	30
Int10(2)	AAGCATTTTCTCCTTTTCCTCAAC	31
Int11	TGACTCTGATGTCAAATGTTTCTCAA	32
Int14a	TGAGCCAGAAGTGGGCAACCT	33
Int18	GCTTGCTGACCCTTCTTCAG	34
Int19	GGGAGGAGGTGCATTGAAGTTA	35
Int20(2)	GCGGTTCATAAAGGGTCATA	36
Int21	TGTTGACTTTTCTCTAATGAAGATG	37
Ex24	GTGTTGGAGAAGAACTGAAATCATACTT	38
3'Ins	CTGTGGTTTGAATGTGTTCCCTAA	39
3'+20	CTGCTACAAGCCCACTATACAATATTG	40
Promoter R	GCAGTGTGGGTCTGATGCAT	41
Promoter Probe	FAM-CTCCGACACGCAAAGGAAGCGCT-TAMRA	42
Ercc F	CCCTGAGGTGAGTTTGTGGAAT	43
Ercc R	AGGATCTCTGTTTAATGGAAAAGCTT	44
Ercc Probe	FAM-CAGAGTCGGATCACGCCTTCATCCTT-TAMRA	45
B13 F	CGTGAGAGCATACTTCCTGGTTC	46
B13 R	ACACCAGAGAGGTCTTGCCCT	47
Plasmid Cloning		

SalI DHS10ab F	CGTCGACTGGGGTTTTGCCATGTTGGCCAA	48
SalI/NdeI DHS10ab R	CGTCGACATATGCTTTCAAACCATCCTCCACTGCC	49
SalI DHS11 F	CGTCGACTGGAGAAGGTGGAATCACACTG	50
SalI/NdeI DHS11 R	CGTCGACATATGGAAGACAGTATGCAAGAGCTAC AT	51
SalI/NdeI +15.6 F	CGTCGACATATGGCTGGTATTACAGGTGTACACC	52
SalI/NdeI +15.6 R	CGTCGACATATGAGCTTCACAGCAAAGATGGGAC	53
BamHI DHS11 F	CGGATCCTGGAGAAGGTGGAATCAC	54
BamHI DHS11 R	CGGATCCGAAGACAGTATGCAAGAGCTA	55
ChIP		
-20.9kb F	CCGGGATGTTGTTTGAAGCTT	56
-20.9kb R	TTTAAATAGTTGAATAGAGGACGAGATACTTT	57
-4kb F	TTCCACAGTACAAGGGCAACC	58
-4kb R	CAGATTAAGTTAGGGTCTCTCTACCTCAG	59
Promoter F	GTCCTCCCGCCGGTGG	60
Promoter R	CAGTCGCGGCCTCTCTTTAG	61
+5.7kb F (Int1[1])	CCTGTGTCTATTGTAGATCGATGACA	62
+5.7kb R	CTTGCTTTGACCCCTTTTTATCA	63
DHS1 F (Int1[2])	TCATTGTCAACTGTCAGGTAGCAA	64
DHS1 R	CAGAGTTAGGATTCCAGCCAGG	65
Int8 F	GCTTCCTCCTGATCAATCTTTAGG	66
Int8 R	AAAAATTCCTTGCCTCACTATTGC	67
Int10(1) F	TGCTTTATTGAATGGCATTACCTCTA	68
Int10(1) R	AGATGCTTGTGGTAAGGGAGGAG	69
Int10(2) F	GGAGAGTACTGTCTCTTATCAGCCATCT	70
Int10(2) R	CTGGTCTTTTCAACACTTTGAGTCA	71
Int11 F	TCCAAAAGCTGAGACAGGAAACT	72

Int11 R	ATTACATACACACAAAAGTACACACATGACT	73
Int14a F	TGAGCCAGAAGTGGGCAACCT	74
Int14a R	GCTGCAGACCTTCCCAGTGA	75
Int17a F	GGATAGTGCTGCTATTACTAAAGGTTTCT	76
Int17a R	ATGGCAGCTCCAACACATGA	77
+6.8kb F	TCTTCTTTCCCATTCACCTTTGTC	78
+6.8kb R	TTTTGGTTTCATTTATCAGCACATC	79
+15.6kb F	ATCCATTTTCTTCAAGTCTCTCTCCAT	80
+15.6kb R	GGAATGAGGATTGTTTATGATTTG	81

[0089] In Table 2, FAM denotes FAM labeling of the sequence. TAMRA denotes TAMRA labeling of the sequence.

[0090] It is to be understood that this invention is not limited to the particular devices, methodology, protocols, subjects, or reagents described, and as such may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention, which is limited only by the claims. Other suitable modifications and adaptations of a variety of conditions and parameters, obvious to those skilled in the art of molecular biology, genetic engineering, and medicine, are within the scope of this invention. All publications, patents, and patent applications cited herein are incorporated by reference in their entirety for all purposes.

CLAIMS

What is claimed is:

1. A system for regulation of the expression of a *CFTR* gene, comprising an isolated nucleic acid element that is at least 95% identical to the nucleic acid sequence of SEQ ID NO:1, and a *CFTR* gene promoter that is at least 95% identical to the nucleic acid sequence of SEQ ID NO:2.

2. The system of claim 1, which regulates enhanced expression of the *CFTR* gene.

3. The system of claim 1, wherein the nucleic acid element is recombinant.

4. The system of claim 1, wherein the *CFTR* gene promoter is recombinant.

5. The system of claim 1, wherein the *CFTR* gene promoter is native.

6. The system of claim 1, wherein the isolated nucleic acid element is operably linked to the *CFTR* gene promoter.

7. The system of claim 1, further comprising an isolated nucleic acid element comprising the *CFTR* gene enhancer sequence of intron 1.

8. An isolated nucleic acid element including a functional sequence, which functional sequence is contained within the sequence of SEQ ID NO:1, and which nucleic acid element regulates expression of the *CFTR* gene.

9. A vector containing the nucleic acid element according to claim 8.

10. An isolated nucleic acid construct up to 50 kbp in length comprising the coding sequence of the *CFTR* gene together with an isolated nucleic acid element that is at least 95% identical to the nucleic acid sequence of SEQ ID NO:1.

11. A vector comprising the isolated nucleic acid construct according to claim 10.

12. A method of preparing a delivery vector for delivery of a *CFTR* gene to a mammal, comprising formulating the nucleic acid construct according to claim 10 into a delivery vector suitable for administration to the mammal.

13. The method of claim 12, wherein the nucleic acid construct further comprises a *CFTR* gene promoter that is at least 95% identical to the nucleic acid sequence of SEQ ID NO:2.

14. A method for regulating expression of a *CFTR* gene in a mammal, the method comprising delivering the vector according to claim 10 to the mammal, thereby regulating the expression of the delivered *CFTR* gene in the mammal.

15. The method of claim 14, wherein the mammal is human.

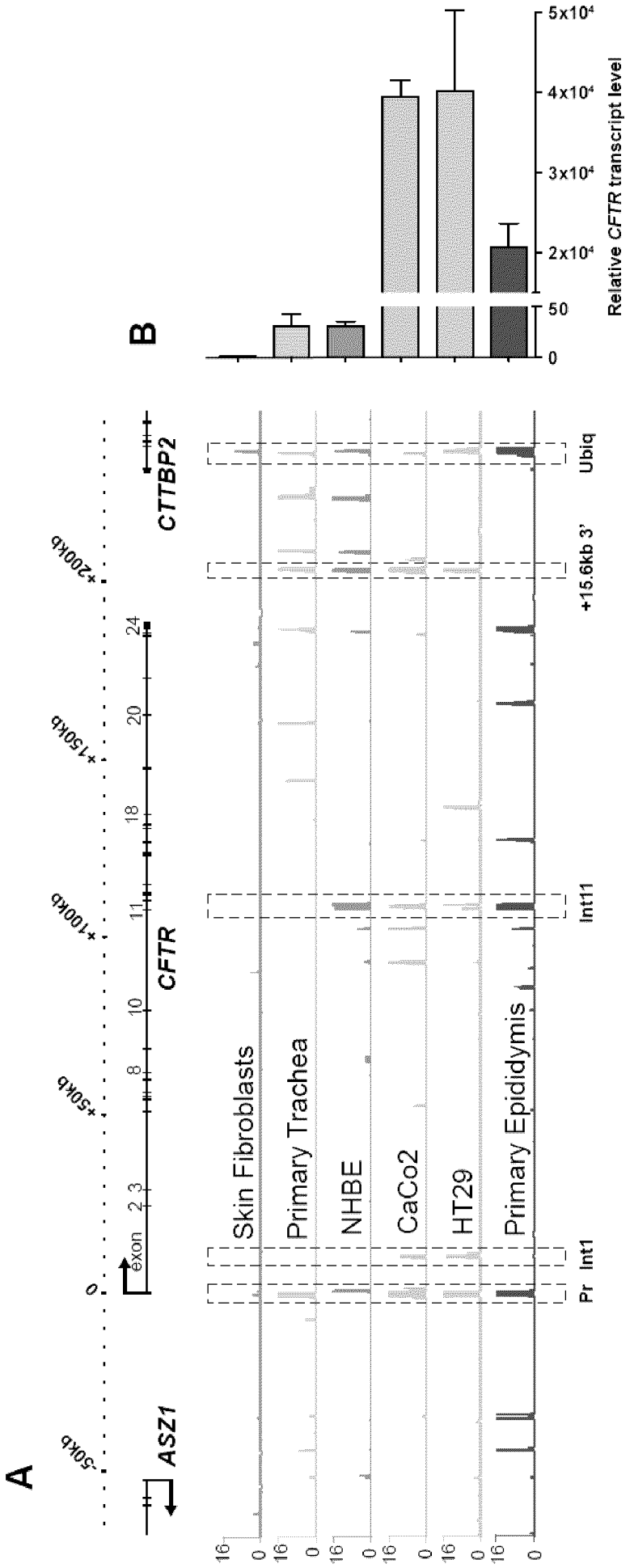
A
B
ASZ1
CFTR
CTTBP2
-50kb
0
+50kb
+100kb
+150kb
+200kb
exon
2 3
8
10
11
18
20
24
Pr Int1
Int11
+15.6kb 3'
Ubiq
Skin Fibroblasts
Primary Trachea
NHBE
CaCo2
HT29
Primary Epididymis
16
0
5x10⁴
4x10⁴
3x10⁴
2x10⁴
50
0
Relative CFTR transcript level

FIGURE 1

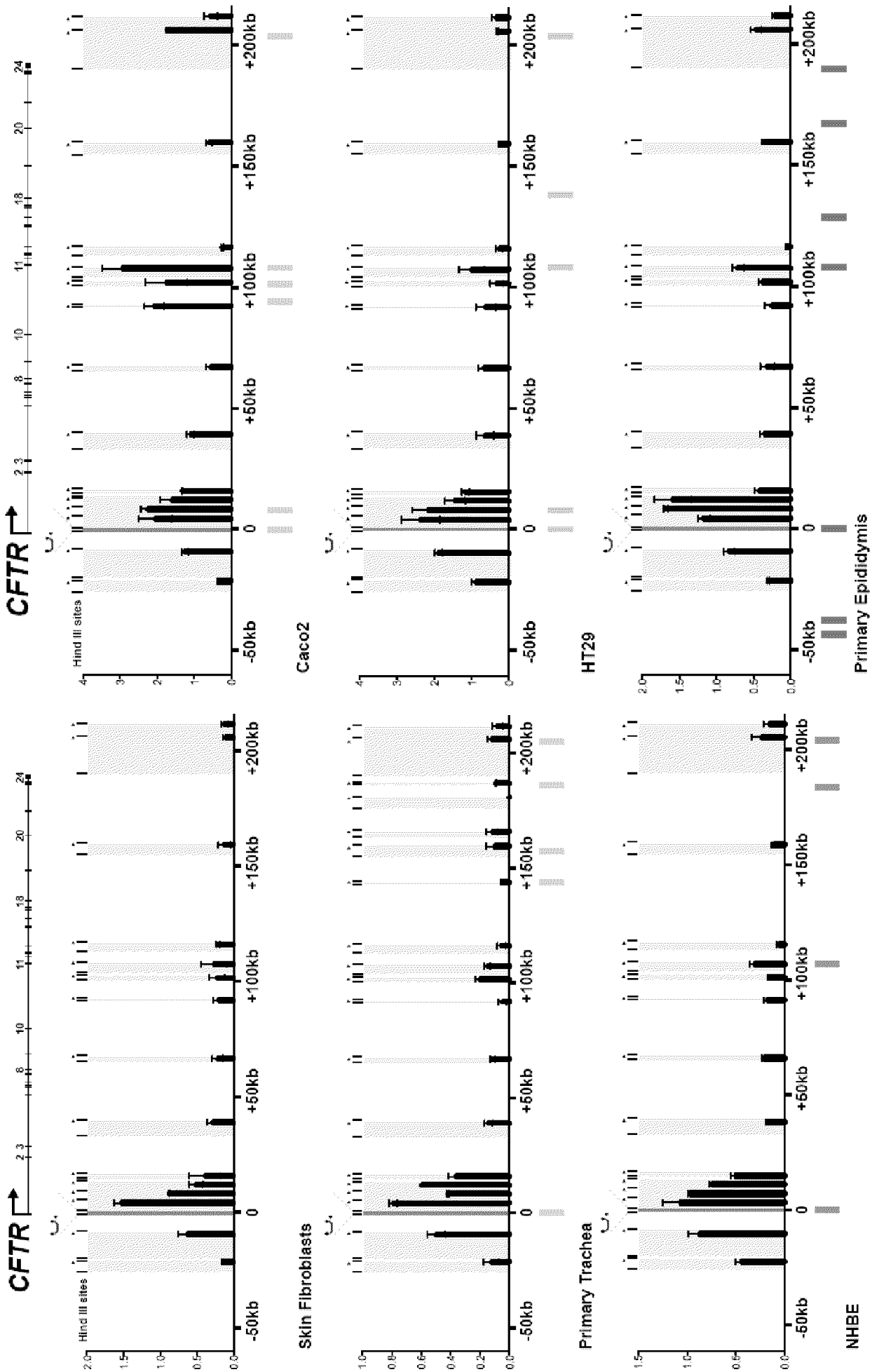
CFTR
Hind III sites
-50kb
0
+50kb
+100kb
+150kb
+200kb
Caco2
HT29
Primary Epididymis
Skin Fibroblasts
Primary Trachea
NHBE

FIGURE 2

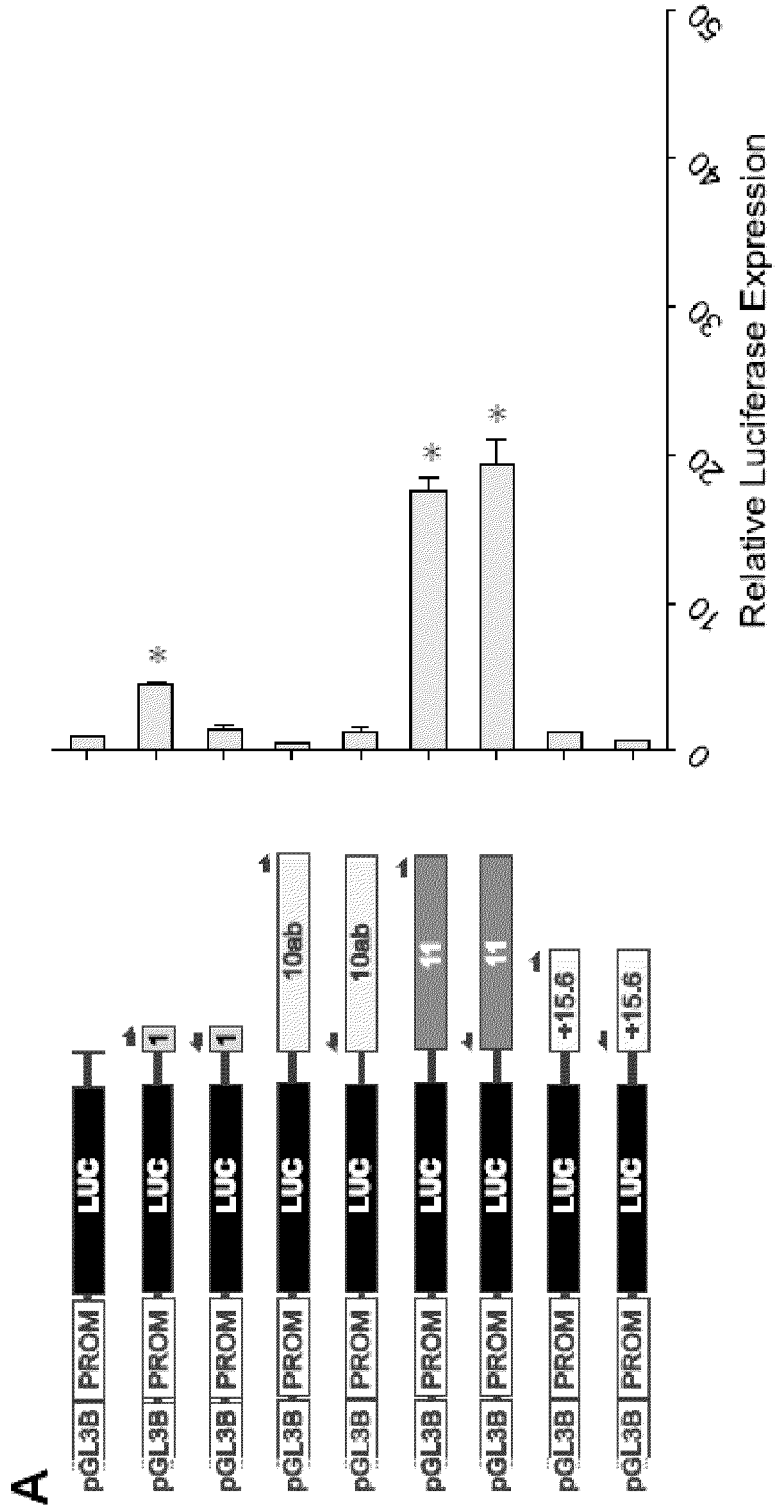
A
pGL3B
PROM
Luc
1
10ab
11
+15.6
Relative Luciferase Expression
0
10
20
30
40
50

FIGURE 3A

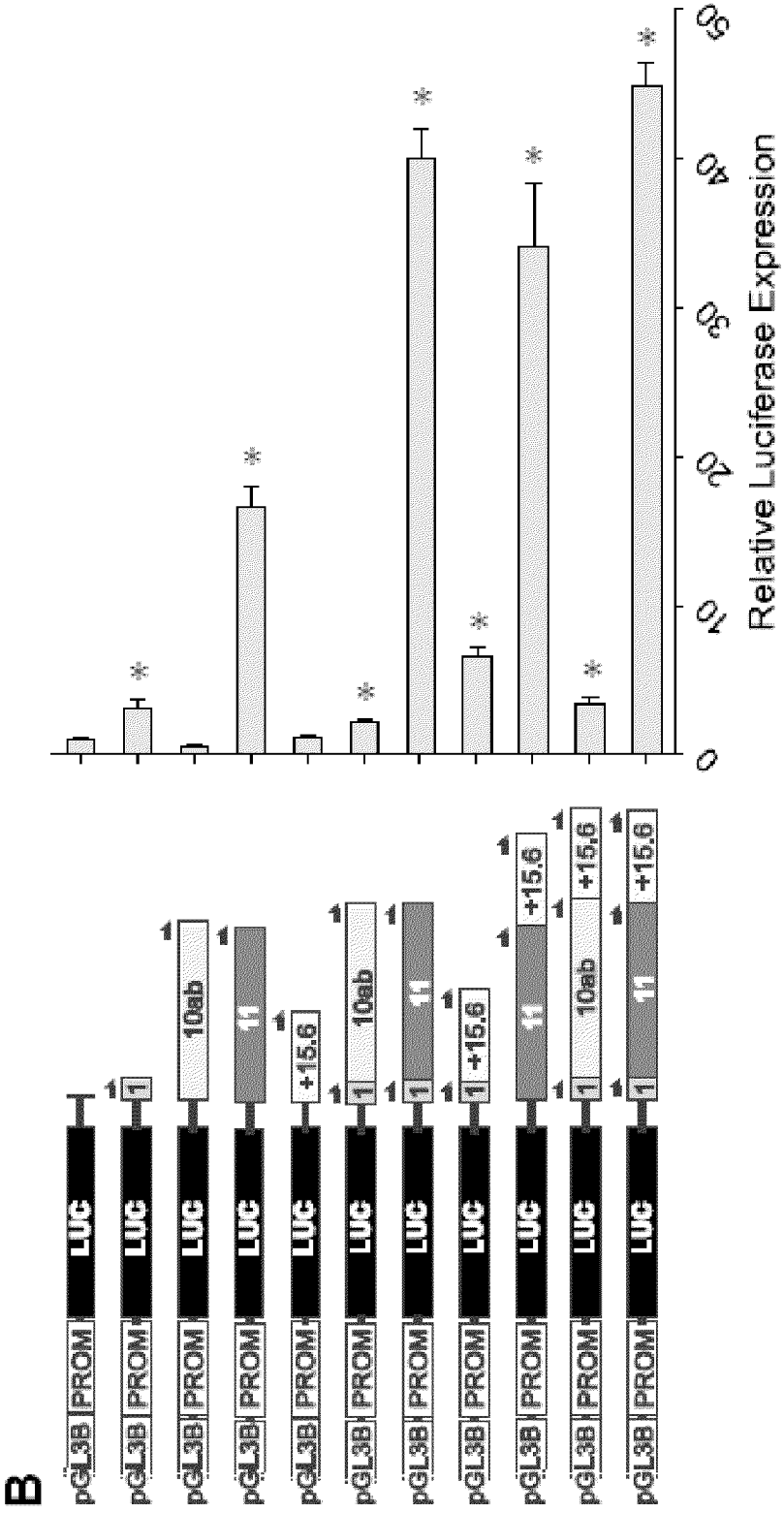
B
pGL3B
PROM
Luc
1
10ab
11
+15.6
Relative Luciferase Expression
0
20
40
60
80
*

FIGURE 3B

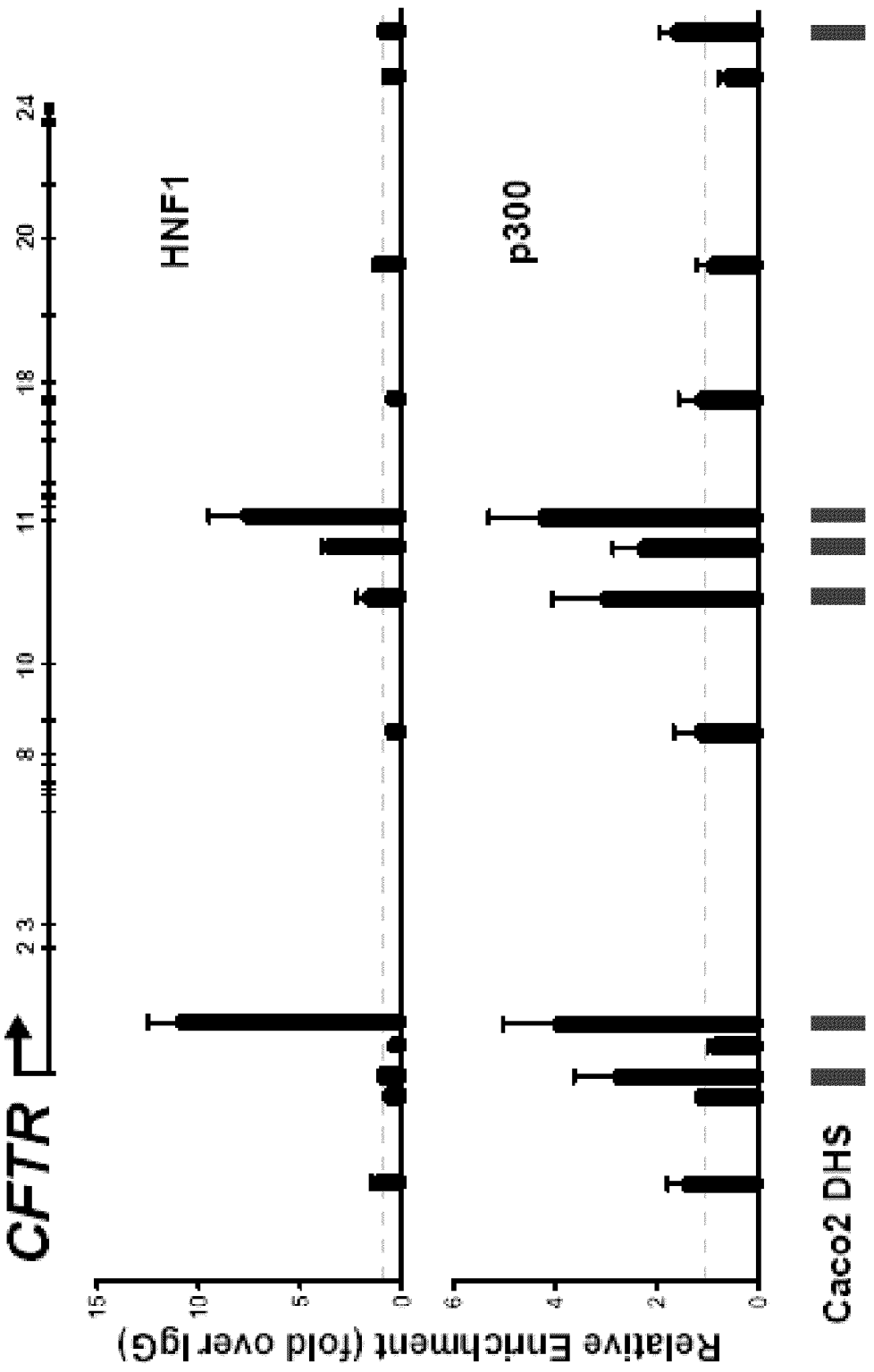
CFTR
2 3
8
10
11
18
20
24
HNF1
p300
Relative Enrichment (fold over IgG)
15
10
5
0
6
4
2
0
Caco2 DHS

FIGURE 4

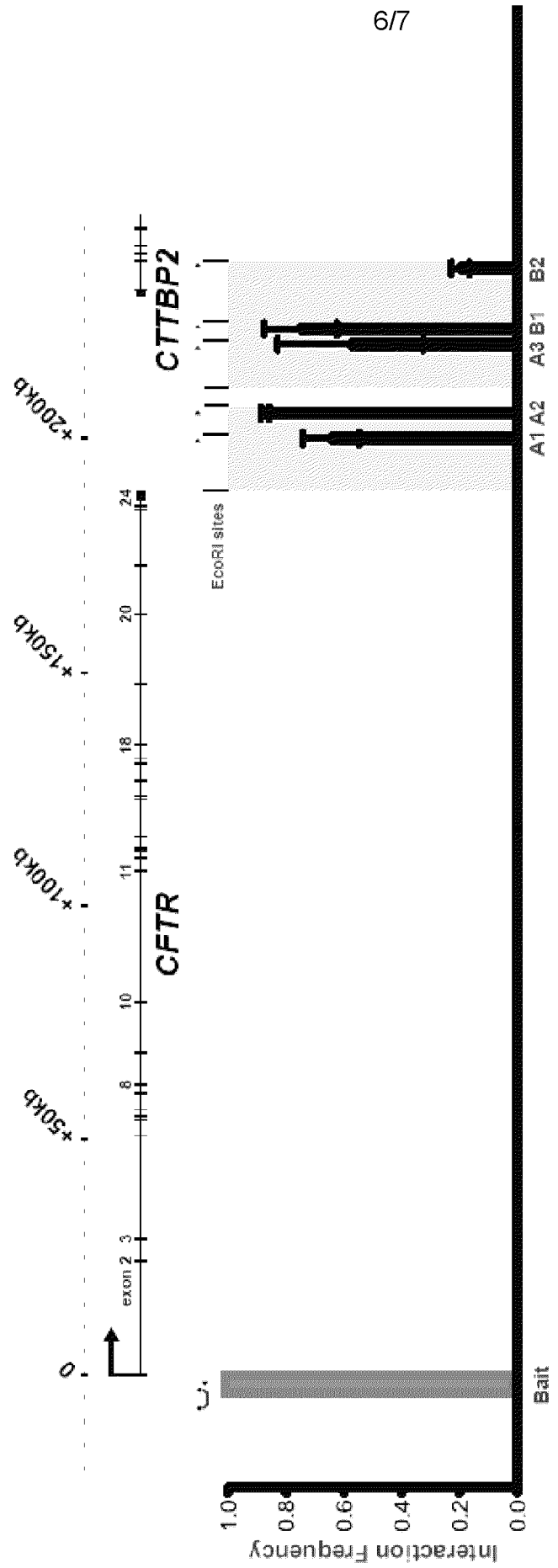
Interaction Frequency
1.0
0.8
0.6
0.4
0.2
0.0
Bait
CTTBP2
A1 A2
A3 B1
B2
EcoRI sites
CFTR
exon 2 3
8
10
11
18
20
24
0
+50Kb
+100Kb
+150Kb
+200Kb

FIGURE 5

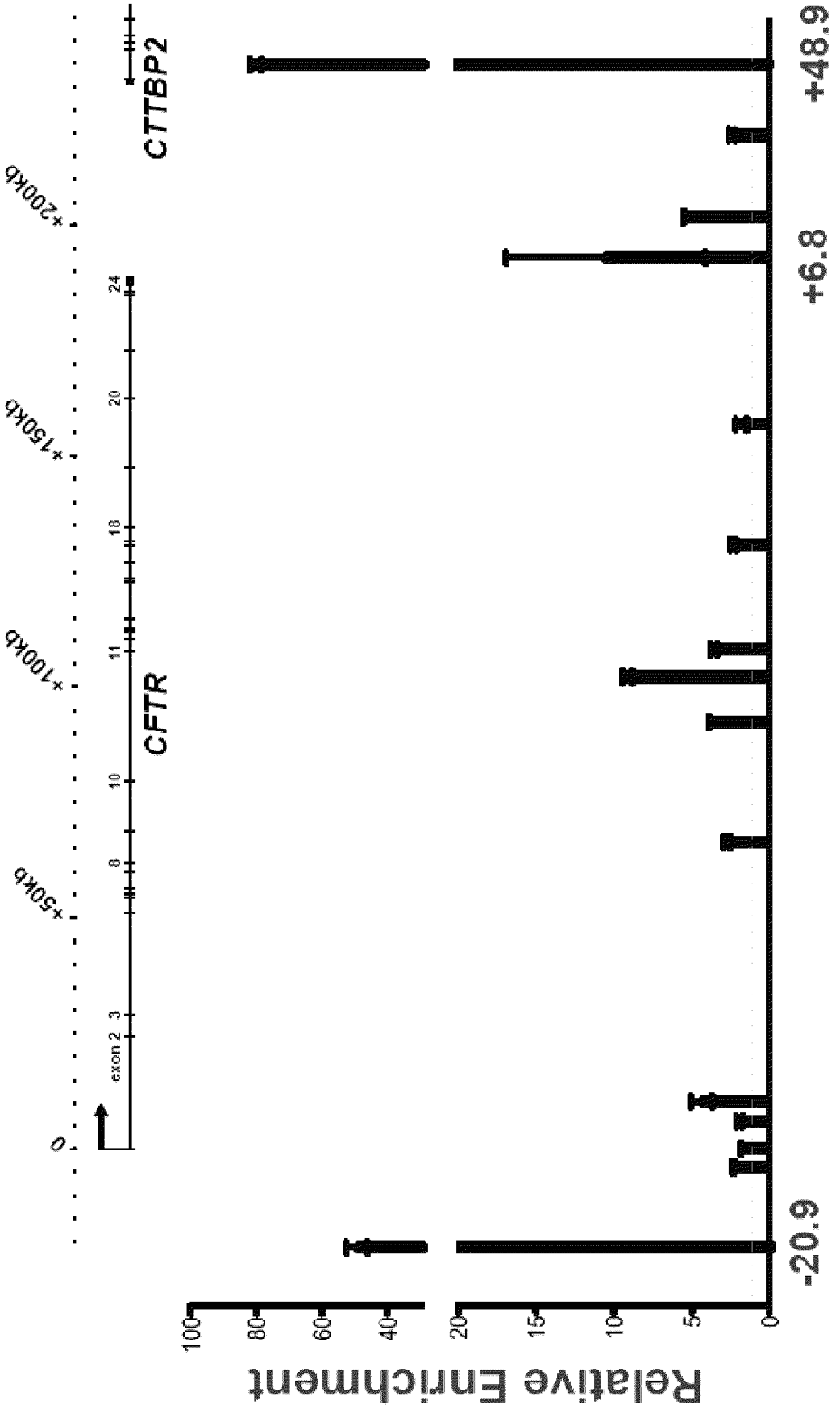
Relative Enrichment
100
80
60
40
20
15
10
5
0
-20.9
+6.8
+48.9
CTTBP2
CFTR
exon 2
0
50Kb
100Kb
150Kb
200Kb

FIGURE 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 10/48872

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - A61K 48/00 (2010.01)
USPC - 514/44A
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - A61K 48/00 (2010.01)
USPC - 514/44A

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC - 536/23.1; 435/320.1
(Text Search)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWEST (PGPB, USPT, USOC, EPAB, JPAB); Google Scholar; PubMed and GenCore 6.3
Search Terms: cystic fibrosis transmembrane conductance regulator, CFTR, enhancer, intron, intronic, 11, regulator, intronic enhancer, promoter, exon.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2002/0086836 A1 (HARRIS) 04 July 2002 (04.07.2002) claim 14, 15; para [0006], [0007], [0011], [0012], [0023]-[0029], [0036]-[0038], [0040]-[0045].	1-15
Y	HILLIER et al. Homo Sapiens BAC Clone CTB-68P20 from 7, Complete Sequence. GenBank Direct Submission Acession: AC000111, 01 March 2004 [online]. [Retrieved on 18 November 2010]. Retrieved from the Internet:<URL: http://www.ncbi.nlm.nih.gov/nuccore/1809237>, pg 3, 35.	1-15
Y	US 2007/0161029 A1 (LI et al.) 12 July 2007 (12.07.2007) table 1; SEQ ID NO: 11; para [0200].	1-7 and 13
A, P	OTT et al., Intronic Enhancers Coordinate Epithelial-Specific Looping of the Active CFTR Locus. PNAS. 24 November 2009, Vol 106, No 47, pages 19934-19939: abstract; pg 19936, col 2, para 2 to pg 19937, col 2, para 1.	1-15

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 November 2010 (18.11.2010)	14 DEC 2010
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (July 2009)