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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 April 2016 (21.04.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/061073 A1

(51) International Patent Classification:
C12N 15/85 (2006.01) *C12N 15/113* (2010.01)
C12N 15/861 (2006.01) *C12N 15/63* (2006.01)

(21) International Application Number:
PCT/US2015/055286

(22) International Filing Date:
13 October 2015 (13.10.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/063,728 14 October 2014 (14.10.2014) US

(71) Applicant: **MEMORIAL SLOAN-KETTERING CANCER CENTER** [US/US]; Office Of Technology Development, 1275 York Avenue, Box 524, New York, NY 10065 (US).

(72) Inventors: **VENTURA, Andrea**; 430 East 63rd Street, Apt. 71, New York, NY 10065 (US). **MADDALO, Danilo**; 475 Main Street, Apt. 11J, New York, NY 10044 (US).

(74) Agent: **YE, Michael**; Andrews Kurth, LLP, 1350 I Street, N.W., Suite 1100, Washington, DC 20005 (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, BN, 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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

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

(54) Title: COMPOSITION AND METHOD FOR *IN VIVO* ENGINEERING OF CHROMOSOMAL REARRANGEMENTS

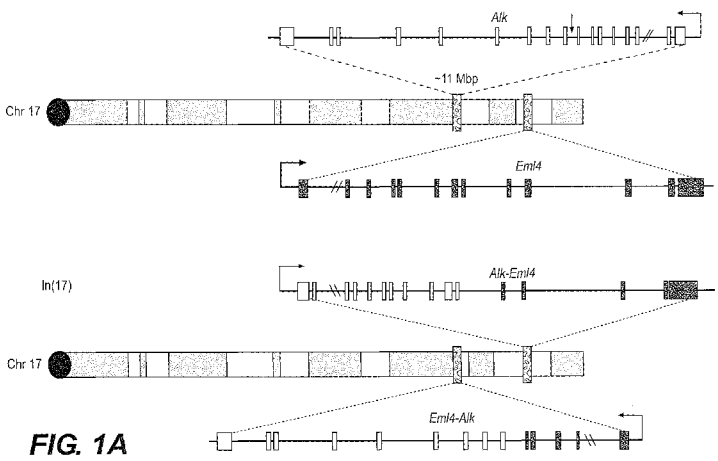


FIG. 1A

(57) Abstract: A method for inducing specific chromosomal rearrangements *in vivo* using the CRISPR/Cas system is disclosed. The method may be used to introduce chromosomal rearrangement *in vivo* and to generate animal models of human diseases, such as a mouse model of *Eml4-Alk*-driven non- small cell lung cancer.

WO 2016/061073 A1

COMPOSITION AND METHOD FOR *IN VIVO* ENGINEERING OF CHROMOSOMAL REARRANGEMENTS

[0001] This application claims priority from U.S. Provisional Application No. 62/063,728, filed October 14, 2014. The entirety of the aforementioned application is incorporated by reference.

FIELD

[0002] The present application relates generally to the field of genetic engineering and in particular to compositions and methods for inducing chromosomal rearrangements *in vivo*.

BACKGROUND

[0003] Chromosomal rearrangements have a central role in the pathogenesis of human cancers and often result in the expression of therapeutically actionable gene fusions. A recently discovered example is a fusion between the genes echinoderm microtubule-associated protein like 4 (*Eml4*) and anaplastic lymphoma kinase (*Alk*), generated by an inversion on the short arm of chromosome 2: inv(2)(p21p23). The *Eml4–Alk* oncogene is detected in a subset of human non-small cell lung cancers (NSCLC) and is clinically relevant because it confers sensitivity to *Alk* inhibitors. Despite their importance, modelling such genetic events in mice has proven challenging and requires complex manipulation of the germline.

SUMMARY

[0004] One aspect of the present application relates to a method for inducing chromosomal rearrangements *in vivo*. The method comprises administering into a subject an effective amount of an expression vector comprising a polynucleotide encoding a Cas gene and two or more distinct single guide RNAs (sgRNAs), wherein the expression vector expresses the Cas gene and the two or more sgRNAs *in vivo* and causes chromosomal rearrangements between two chromosomal loci targeted by the two or more sgRNAs.

[0005] Another aspect of the present application relates to a mouse produced by the method described above, wherein the mouse has non-small cell lung cancer (NSCLC) cells harboring an *Eml4-Alk* inversion and expressing an *Eml4-Alk* fusion gene.

[0006] Another aspect of the present application relates to a composition for inducing specific chromosomal rearrangements *in vivo*. The composition comprises an expression vector comprising a polynucleotide encoding a Cas gene and two or more distinct single guide RNAs (sgRNAs); wherein the expression vector is capable of expressing the Cas gene

and the two or more sgRNAs *in vivo* and causing chromosomal rearrangements between two chromosomal loci targeted by the two or more sgRNAs.

[0007] Another aspect of the present application relates to a kit for inducing specific chromosomal rearrangements *in vivo*. The kit comprises the composition described above and a reagent for introducing the expression vector into a cell *in vivo*.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The above and other objects and advantages of the application will be apparent upon consideration of the following detailed description, taken in conjunction with the accompanying figures.

[0009] **Fig. 1** is a composite showing induction of *Eml4–Alk* rearrangement in murine cells using the CRISPR-Cas9 system. Panel (a): Schematic of the In(17) involving the *Eml4* and *Alk* loci. The smaller arrows indicate the sites recognized by the sgRNAs. Panel (b): A schematic of the loci before and after the inversion with the location of the primers used (top). PCRs were performed on genomic DNA extracted from NIH/3T3 cells transfected with the indicated pX330 constructs (middle). The PCR bands were sub-cloned and the sequences of four independent clones and a representative chromatogram are shown (*Eml4-Alk* Predicted, GGGTCTACTATGTAAGGCTA (SEQ ID NO. 23); *Alk-Eml4* Predicted, ATGTCTATCTGGAAGGAGCC (SEQ ID NO. 24); *Eml4-Alk* Clone 1, GGGTCTACTATGTAAGGCTA (SEQ ID NO. 25); *Alk-Eml4* Clone 1, ATGTCTATCTGGAAGGAGCC (SEQ ID NO. 26); *Eml4-Alk* Clone 2, GGGTCTACTATGTAAGGCTA (SEQ ID NO. 27); *Alk-Eml4* Clone 2, ATGTCTA---GGAAGGAGCC (SEQ ID NO. 28); *Eml4-Alk* Clone 3, GGGTCTACT--GTAAGGCTA (SEQ ID NO. 29); *Alk-Eml4* Clone 3, ATGTCTA---GGAAGGAGCC (SEQ ID NO. 30); *Eml4-Alk* Clone 4, GGGTCTACT--GTAAGGCTA (SEQ ID NO. 31); *Alk-Eml4* Clone 4, ATG----------GAGCC (SEQ ID NO. 32)) (bottom). Panel (c): Schematic of the *Eml4–Alk* fusion transcript (top). Detection of the *Eml4–Alk* fusion transcript by RT–PCR on total RNAs extracted from NIH/3T3 cells transfected with the indicated pX330 constructs (bottom left). Sequence of the PCR product showing the correct *Eml4–Alk* junction is shown (*Eml4-Alk* Predicted, GGACCTAAAGTGTACCGTCG (SEQ ID NO. 33); *Eml4-Alk* Observed, GGACCTAAAGTGTACCGTCG (SEQ ID NO. 34); (bottom right).

[0010] **Fig. 2** is a schematic of the break-apart interphase FISH strategy. In cells with the *Eml4–Alk* inversion, the 5'*Alk* and 3'*Alk* probes become separated, and the 3'*Alk* and 5' *Eml4* probes become juxtaposed. .

[0011] **Fig. 3** shows induction of the Npm1–*Alk* translocation in NIH/3T3 cells. Panel (a): Schematic of the Npm1-*Alk* translocation. Smaller arrows indicate the sites recognized by the sgRNAs. Panel (b): Sequences recognized by the sgRNAs and location of primers used to detect the Npm1–*Alk* and *Alk*–Npm1 rearrangement (top). PCR on genomic DNA extracted from NIH/3T3 co-transfected with pX330 constructs expressing the indicated sgRNAs (middle). Sequences of four independent subclones obtained from the PCR products and representative chromatogram are shown (*Npm1-Alk* Predicted, AATGTGTTTTGTAAGGCTA (SEQ ID NO. 35); *Alk-Npm1* Predicted, CATGTCTATCTGGTTGGGTT (SEQ ID NO. 36); *Npm1-Alk* Clone 1, AATGTGTTTTGTAAGGCTA (SEQ ID NO. 37); *Alk-Npm1* Clone 1, CATGTCTATCTGGTTGGGTT (SEQ ID NO. 38); *Npm1-Alk* Clone 2, AATGTGTTTTGTAAGGCTA (SEQ ID NO. 39); *Alk-Npm1* Clone 2, CATGTCTATCTGGTTGGGTT (SEQ ID NO. 40); *Npm1-Alk* Clone 3, AATGTGTTTTGTAAGGCTA (SEQ ID NO. 41); *Alk-Npm1* Clone 3, CATGTCTATCTGGTTGGGTT (SEQ ID NO. 42); *Npm1-Alk* Clone 4, AATGTGTTTTGTAAGGCTA (SEQ ID NO. 43); *Alk-Npm1* Clone 4, CATGTCTATCTGGTTGGGTT (SEQ ID NO. 44)) (bottom). Panel (c): Detection of the Npm1–*Alk* fusion transcript by RT–PCR on total RNAs extracted from NIH/3T3 cells co-transfected with the indicated pX330 constructs (left). The PCR band was extracted and sequenced to confirm the presence of the correct Npm1–*Alk* junction (bottom-right). Representative results from two independent experiments are shown (*Npm1-Alk* Predicted, CATCTAGTAGTGTACCGTCG (SEQ ID NO. 45); *Npm1-Alk* Observed, CATCTAGTAGTGTACCGTCG (SEQ ID NO. 46).

[0012] **Fig. 4** shows comparison of dual and single sgRNA-expressing plasmids. Panel (a): Schematic of pX330 (A) and its derivatives (B–E) used in these experiments. NIH/3T3 were transfected with these constructs and lysed to extract total RNA and genomic DNA. Panel (b): RNAs were analysed by northern blotting with probes against the *Alk* (left) or *Eml4* (right) sgRNAs. Panels (c) and (d): The DNA samples were subjected to surveyor assays (Panel (c)), or amplified by PCR to detect the *Eml4–Alk* inversion (Panel (d)).

[0013] **Fig. 5** shows induction of the *Eml4–Alk* inversion in primary MEFs using an adenoviral vector expressing FLAG–Cas and tandem sgRNAs. Panel (a): Schematic of the Adenoviral vectors. Panel (b): Immunoblot using an anti-FLAG antibody on lysates from MEFs infected with the indicated adenoviruses. Panel (c): Small-RNA northerns using probes against sg*Eml4* and sg*Alk* on total RNAs from cells infected with Ad-Cas9 or Ad-EA.

Panel (d): PCR-mediated detection of the *Eml4–Alk* inversion in MEFs infected with Ad-Cas9 or Ad-EA for the indicated number of days. Panel (e): Standard curve generated performing quantitative PCR analysis on genomic DNA containing a known fraction of *Eml4–Alk* alleles. Average of two independent experiments. Panel (f): Quantification of the fraction of MEFs harbouring the *Eml4–Alk* inversion at the indicated time points after infection with Ad-EA or Ad-Cas9. Values are mean of three independent infections ± s.d.

[0014] **Fig. 6** is a composite showing that lung tumors induced by Ad-EA infection harbor the *Eml4–Alk* inversion. Panel (a): Detection of the wild type *Eml4* locus and *Eml4-Alk* inversion in micro-dissected tumors from Ad-EA-infected mice using a three-primer PCR strategy. Panel (b): RNAs extracted from the same tumors shown in Panel (a) were reverse-transcribed and amplified using a three-primer strategy to detect the *Eml4* and *Eml4–Alk* transcripts. Panel (c): RT–PCR detection (left) of the full length *Eml4–Alk* cDNA (~3.2 Kb) in the tumors shown in Panel (a). The full-length PCR products were sequenced on both strands. A chromatogram of the *Eml4–Alk* junction is shown (right). Observed sequence: GGACCTAAAGTGTACCGTCG (SEQ ID NO. 47); Predicted sequences GGACCTAAAGTGTACCGTCG (SEQ ID NO. 48). Panel (d): Representative immunohistochemistry of Ad-EA-induced lung tumors decorated with antibodies against the indicated phospho-proteins. A bar-plot of staining intensity for the indicated phospho-proteins is also shown. Tumors from two mice for each group were scored.

[0015] **Fig. 7** is a composite showing Ad-EA-induced lung tumors respond to crizotinib treatment. Panel (a): Schematic of the experiment. Panel (b): Representative μCT of the lungs of mice treated with crizotinib or vehicle at day 0 and after 2 weeks of treatment. Lung tumors are indicated by arrows. Asterisks mark the hearts. Panel (c): Macroscopic appearance of the lungs after 2 weeks of treatment. Panel (d): Low magnification of lung sections from two crizotinib-treated and 2 vehicle-treated mice (haematoxylin and eosin). Panel (e): Higher magnification of representative haematoxylin and eosin stained lung sections from crizotinib-treated mice showing residual atrofic foci of tumor cells (left) or necrotic-inflammatory debris (right).

DETAILED DESCRIPTION

[0016] Some modes for carrying out the present invention are presented in terms of its exemplary embodiments, herein discussed below. However, the present invention is not limited to the described embodiment and a person skilled in the art will appreciate that many other embodiments of the present invention are possible without deviating from the basic concept of the present invention, and that any such work around will also fall under scope of

this application. It is envisioned that other styles and configurations of the present invention can be easily incorporated into the teachings of the present invention, and only one particular configuration shall be shown and described for purposes of clarity and disclosure and not by way of limitation of scope.

[0017] Headings used herein are for organizational purposes only and are not meant to be used to limit the scope of the description or the claims. As used throughout this application, the word "may" is used in a permissive sense (i.e., meaning having the potential to), rather than the mandatory sense (i.e., meaning must). The terms "a" and "an" herein do not denote a limitation of quantity, but rather denote the presence of at least one of the referenced items.

Definitions

[0018] The term "CRISPR" as used herein, refers to Clustered Regularly Interspaced Short Palindromic Repeats, which can also be known as SPIDRs (Spacer Interspersed Direct Repeats. CRISPR are a family of DNA loci that are normally associated with a particular bacterial species. The CRISPR locus represents a particular class of interspersed short sequence repeats (SSRs) that were recognized in *E. coli*. Similar interspersed SSRs have been identified in *Haloferax mediterranei*, *Streptococcus pyogenes*, *Anabaena*, and *Mycobacterium tuberculosis*. The CRISPR loci usually differs from other SSRs by the structure of the repeats, which have been termed short regularly spaced repeats (SRSRs). Generally, the repeats are short elements that appear in regularly spaced clusters with unique intervening sequences of a substantially constant length. The repeat sequences are highly conserved between strains, but the number of interspersed repeats and the sequences of the spacer regions usually differs from strain to strain.

[0019] As used herein, "CRISPR system" refers collectively to transcripts and other elements involved in the expression of or directing the activity of CRISPR-associated ("Cas") genes, including sequences encoding a Cas gene, a guide sequence (also referred to as a "spacer" in the context of an endogenous CRISPR system), or other sequences and transcripts from a CRISPR locus. In some embodiments, one or more elements of a CRISPR system is derived from a type I, type II, or type III CRISPR system. In some embodiments, one or more elements of a CRISPR system is derived from a particular organism comprising an endogenous CRISPR system, such as *Streptococcus pyogenes*. In general, a CRISPR system is characterized by elements that promote the formation of a CRISPR complex at the site of a target sequence (also referred to as a protospacer in the context of an endogenous CRISPR system). In the context of formation of a CRISPR complex, "target sequence" refers to a

sequence to which a guide sequence is designed to have complementarity, where hybridization between a target sequence and a guide sequence promotes the formation of a CRISPR complex. Full complementarity is not necessarily required, provided there is sufficient complementarity to cause hybridization and promote formation of a CRISPR complex. A target sequence may comprise any polynucleotide, such as DNA or RNA polynucleotides. In some embodiments, a target sequence is located in the nucleus or cytoplasm of a cell. In some embodiments, the target sequence may be within an organelle of a eukaryotic cell, for example, mitochondrion or chloroplast.

[0020] As used herein, the terms "sgRNA", "guide RNA", "single guide RNA" and "synthetic guide RNA" are interchangeable and refer to the polynucleotide sequence comprising the guide sequence. The term "guide sequence" refers to the about 20 bp sequence within the guide RNA that specifies the target site and may be used interchangeably with the terms "guide" or "spacer."

[0021] As used herein, the terms "polynucleotide", "nucleotide", "nucleotide sequence", "nucleic acid" and "oligonucleotide" are used interchangeably. They refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides may have any three dimensional structure, and may perform any function, known or unknown. The following are non-limiting examples of polynucleotides: coding or non-coding regions of a gene or gene fragment, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, short interfering RNA (siRNA), short-hairpin RNA (shRNA), micro-RNA (miRNA), ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise one or more modified nucleotides, such as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component.

[0022] As used herein, "complementarity" refers to the ability of a nucleic acid to form hydrogen bond(s) with another nucleic acid sequence by either traditional Watson-Crick base pairing or other non-traditional types. A percent complementarity indicates the percentage of residues in a nucleic acid molecule which can form hydrogen bonds (e.g., Watson-Crick base pairing) with a second nucleic acid sequence (*e.g.*, 5, 6, 7, 8, 9, 10 out of 10 being 50%, 60%, 70%, 80%, 90%, and 100% complementary).

[0023] As used herein, "hybridization" refers to a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson Crick base pairing, Hoogstein binding, or in any other sequence specific manner. The complex may comprise two strands forming a duplex structure, three or more strands forming a multi stranded complex, a single self-hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of PCR, or the cleavage of a polynucleotide by an enzyme. A sequence capable of hybridizing with a given sequence is referred to as the "complement" of the given sequence.

[0024] As used herein, the terms “engineering” or “engineered” are used interchangeably and indicate the involvement of the hand of man.

[0025] As used herein, the terms “subject,” “individual,” and “animal” are used interchangeably herein to refer to a vertebrate, preferably a mammal. The term “mammal” or “mammalian” includes, but is not limited to, murines, simians, humans, farm animals, sport animals, and pets. Tissues, cells and their progeny of a biological entity obtained *in vivo* or cultured *in vitro* are also encompassed.

[0026] As used herein, “administered” includes delivery by viral or non-viral techniques. Viral delivery mechanisms include but are not limited to adenoviral vectors, adeno-associated viral (AAV) vectors, herpes viral vectors, retroviral vectors, lentiviral vectors, etc as described above. Non-viral delivery mechanisms include lipid mediated transfection, liposomes, immunoliposomes, lipofectin, cationic facial amphiphiles (CFAs) and combinations thereof.

Composition for inducing *in vivo* chromosomal rearrangements

[0027] One aspect of the present application is a composition for inducing specific chromosomal rearrangements *in vivo*. The composition comprises an expression vector comprising a polynucleotide encoding a Cas gene and two or more distinct single guide RNAs (sgRNAs). The expression vector, when introduced into a mammalian cell, is capable of expressing the Cas gene and the two or more distinct single guide RNAs (sgRNAs) *in vivo* and causing chromosomal rearrangement of a DNA sequence between two chromosomal loci targeted by the two or more sgRNAs (target loci). In some embodiment, the two target loci are separated by a genomic sequence of greater than 10kb, 30kb, 100kb, 300kb, 1Mb, 3 Mb or 10 Mb. In some embodiment, the two target loci are separated by a genomic sequence of 10kb-30kb, 10kb-100kb, 10kb-300kb, 10kb-1Mb, 10kb-3 Mb, 10kb-10 Mb, 10kb-30 Mb, 10kb-100Mb, 30kb-100kb, 30kb-300kb, 30kb-1Mb, 30kb-3 Mb, 30kb-10 Mb, 30kb-30 Mb,

30kb-100Mb, 100kb-300kb, 100kb-1Mb, 100kb-3 Mb, 100kb-10 Mb, 100kb-30 Mb, 100kb-100Mb, 300kb-1Mb, 300kb-3 Mb, 300kb-10 Mb, 300kb-30 Mb, 300kb-100Mb, 1Mb-3 Mb, 3Mb -10 Mb, 10Mb -30 Mb, 10Mb-100Mb or 30Mb-100Mb.

[0028] In some embodiments, the two or more distinct sgRNAs comprise sgRNAs targeting *Eml4* and *Alk* sites. In other embodiments, the two or more distinct sgRNAs comprise sgRNAs targeting Npml and *Alk* sites.

[0029] As used herein, the term "chromosomal rearrangement" refers to all possible allele combination of the sequences between two target loci on a chromosome, including but are not limited to, indels, inversions and deletions.

The Cas gene

[0030] The Cas gene encodes a RNA-guided DNA endonuclease enzyme that forms a CRISPR complex (comprising a guide sequence hybridized to a target sequence and complexed with one or more Cas proteins), which results in cleavage of one or both strands in or near (e.g. within 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or more base pairs from) the target sequence. Non-limiting examples of Cas proteins include Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homologs thereof, or modified versions thereof. These enzymes are known in the art, and may be found in publicly available databases of protein sequences. In some embodiments, the unmodified CRISPR enzyme has DNA cleavage activity, such as Cas9. In some embodiments the CRISPR enzyme is Cas9, and may be Cas9 from *S. pyogenes* or *S. pneumoniae*. In some embodiments, the CRISPR enzyme directs cleavage of one or both strands at the location of a target sequence, such as within the target sequence and/or within the complement of the target sequence. In some embodiments, the CRISPR enzyme directs cleavage of one or both strands within about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500, or more base pairs from the first or last nucleotide of a target sequence.

Guide Sequences

[0031] A guide sequence is any polynucleotide sequence with sufficient complementarity to a target polynucleotide sequence to hybridize with the target sequence and direct sequence-specific binding of a CRISPR complex to the target sequence. In some embodiments, the degree of complementarity, when optimally aligned using a suitable alignment algorithm, between a guide sequence and its corresponding target sequence is about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more.

Optimal alignment may be determined with the use of any suitable algorithm for aligning sequences, such as known to those of skill in the art. In some embodiments, a guide sequence is about or more than about 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75, or more nucleotides in length. In some embodiments, a guide sequence is less than about 75, 50, 45, 40, 35, 30, 25, 20, 15, 12, or fewer nucleotides in length.

[0032] Any appropriate assay as known to those of skill in the art may be used to assess the ability of a guide sequence to direct sequence-specific binding of a CRISPR complex to a target sequence. A guide sequence may be selected to target any target sequence. In some embodiments, the target sequence is a sequence within a genome of a cell. Exemplary target sequences include those that are unique in the target genome.

Expression Vectors

[0033] The expression vector can be a viral vector (such as a recombinant adenovirus) or a non-viral vector (such as a plasmid). In some embodiments, the expression vector comprises a Cas expression cassette comprising a regulatory element operationally linked to the Cas gene, and a sgRNA expression cassette comprising a regulatory element operationally linked to the DNA sequences encoding the two or more sgRNAs. In some embodiments, the expression vector comprises a Cas expression cassette comprising a regulatory element operationally linked to the Cas gene, and sgRNA expression cassettes for each of the two or more sgRNAs. Each sgRNA expression cassette comprises regulatory element operationally linked to a DNA sequence encoding a sgRNA. The regulatory element may comprise a promoter, an enhancer or both. The promoter can be a constitutive promoter, a tissue specific promoter or an inducible promoter. In some embodiments, the Cas gene and the sgRNA in the expression vector are under the control of a regulatable expression system.

[0034] As used herein, the term "operationally linked to" refers to the joining of an encoding nucleic acid sequence to a regulatory element which results in the biological production of the desired polynucleotide or polypeptide.

[0035] Examples of constitutive promoters include, but are not limited to, CMV promoter, SV40 promoter, CBh (Chicken Beta Action Short), Pgk promoter, H1 promoter, promoter and U6 promoter.

[0036] Examples of tissue specific promoters include, but are not limited to, promoters that direct expression primarily in a desired tissue of interest, such as muscle,

neuron, bone, skin, blood, specific organs (e.g. liver, pancreas), or particular cell types (e.g. lymphocytes).

[0037] Examples of regulatable expression system include, but are not limited to Tet-on/off system, Ecdysone system, progesterone system and rapamycin system. The Tet-on/off system is based on two regulatory elements derived from the tetracycline-resistance operon of the E. coli Tn10 transposon: the tet repressor protein (TetR) and the Tet operator DNA sequence (tetO) to which TetR binds. The system consists of two components, a "regulator" and a "reporter" plasmid. The "regulator" plasmid encodes a hybrid protein containing a mutated Tet repressor (rtetR) fused to the VP16 activation domain of herpes simplex virus. The "reporter" plasmid contains a tet-responsive element (TRE), which controls the "reporter" gene of choice. The rtetR-VP16 fusion protein can only bind to the TRE, therefore activates the transcription of the "reporter" gene, in the presence of tetracycline.

[0038] The ecdysone system is based on the molting induction system found in *Drosophila*, but modified for inducible expression in mammalian cells. The system uses an analog of the drosophila steroid hormone ecdysone, muristerone A, to activate expression of the gene of interest *via* a heterodimeric nuclear receptor. Expression levels have been reported to exceed 200-fold over basal levels with no effect on mammalian cell physiology.

[0039] The progesterone receptor is normally stimulated to bind to a specific DNA sequence and to activate transcription through an interaction with its hormone ligand. Conversely, the progesterone antagonist mifepristone (RU486) is able to block hormone-induced nuclear transport and subsequent DNA binding. A mutant form of the progesterone receptor that can be stimulated to bind through an interaction with RU486 has been generated. To generate a specific, regulatable transcription factor, the RU486-binding domain of the progesterone receptor has been fused to the DNA-binding domain of the yeast transcription factor GAL4 and the transactivation domain of the HSV protein VP16. The chimeric factor is inactive in the absence of RU486. The addition of hormone, however, induces a conformational change in the chimeric protein, and this change allows binding to a GAL4-binding site and the activation of transcription from promoters containing the GAL4-binding site.

[0040] The rapamycin system uses an immunosuppressive agent an inducer. The binding of rapamycin to FK506-binding protein (FKBP) results in its heterodimerization with another rapamycin binding protein FRAP, which can be reversed by removal of the drug. The ability to bring two proteins together by addition of a drug potentiates the regulation of a number of biological processes, including transcription. A chimeric DNA-binding domain

has been fused to the FKBP, which enables binding of the fusion protein to a specific DNA-binding sequence. A transcriptional activation domain also has been fused to FRAP. When these two fusion proteins are co-expressed in the same cell, a fully functional transcription factor can be formed by heterodimerization mediated by addition of rapamycin. The dimerized chimeric transcription factor can then bind to a synthetic promoter sequence containing copies of the synthetic DNA-binding sequence.

[0041] In some embodiments, the expression vector comprises three expression cassettes: a Cas expression cassette comprising a Cas9 gene under the control of a CBh promote and two sgRNA expression cassettes, each comprising a U6 promoter and the coding sequences of a distinct sgRNA. In some embodiments, the two sgRNA expression cassettes comprise coding sequences for sgRNAs targeting the *Alk* and *Eml4* gene. In other embodiments, the two sgRNA expression cassettes comprise coding sequences for sgRNAs targeting the *Alk* and *Npm1* gene.

[0042] In some embodiments, the expression vector also comprises an enhancer element, such as WPRE; CMV enhancers and the R-U5′ segment in LTR of HTLV-I.

Viral Vectors

[0043] In some embodiments, the expression vector of the present application is a viral vector. Viruses are logical tools for gene delivery. They replicate inside cells and therefore have evolved mechanisms to enter the cells and use the cellular machinery to express their genes. The concept of virus-based gene delivery is to engineer the virus so that it can express the gene of interest. Depending on the specific application and the type of virus, most viral vectors contain mutations that hamper their ability to replicate freely as wild-type viruses in the host.

[0044] Viruses from several different families have been modified to generate viral vectors for gene delivery. These viruses include retroviruses, lentivirus, adenoviruses, adeno-associated viruses, herpes simplex viruses, picornaviruses, and alphaviruses.

[0045] RNA or DNA viral based systems targeting a virus to specific cells in the body as known to one of skill in the art can be used for the delivery of nucleic acids and trafficking the viral payload to the nucleus. Viral vectors can be administered directly to subjects (*in vivo*) or they can be used to treat cells *in vitro*, and the modified cells may optionally be administered to subjects (*ex vivo*). Conventional viral based systems could include retroviral, lentivirus, adenoviral, adeno-associated and herpes simplex virus vectors for gene transfer. Integration in the host genome is possible with the retrovirus, lentivirus, and adeno-associated virus gene transfer methods.

[0046] In some embodiments, the expression vector of the present application is an adenoviral vector. The adenovirus is a double-stranded, linear DNA virus that does not go through an RNA intermediate. There are over 50 different human serotypes of adenovirus divided into 6 subgroups based on the genetic sequence homology all of which exhibit comparable genetic organization. Human adenovirus group C serotypes 2 and 5 (with 95% sequence homology) are most commonly used in adenoviral vector systems and are normally associated with upper respiratory tract infections in the young.

[0047] Adenoviral based vectors are capable of very high transduction efficiency in many cell types and do not require cell division. The adenoviruses/adenoviral vectors in aspects of the present application may be of human or animal origin. As regards the adenoviruses of human origin, preferred adenoviruses are those classified in group C, in particular the adenoviruses of type 2 (Ad2), 5 (Ad5), 7 (Ad7) or 12 (Ad12). More preferably, it is an Ad2 or Ad5 adenovirus. Among the various adenoviruses of animal origin, canine adenovirus, mouse adenovirus or an avian adenovirus such as CELO virus (Cotton et al., 1993, J Virol 67:3777-3785) may be used. With respect to animal adenoviruses it is preferred to use adenoviruses of canine origin, and especially the strains of the CAV2 adenoviruses [manhattan strain or A26/61 (ATCC VR-800) for example].

[0048] The organization of the adenovirus genome is similar in all of the adenovirus groups and specific functions are generally positioned at identical locations for each serotype studied. The genome of adenoviruses comprises an inverted terminal repeat (ITR) at each end, an encapsidation sequence (Psi), early genes and late genes. The main early genes have been classified into an array of intermediate early (E1a), delayed early (E1b, E2a, E2b, E3 and E4), and intermediate regions. In some embodiments, the adenoviral vector comprises a deletion in the E1, E2, E3 and/or E4 region in the viral genome.

[0049] In some embodiments, a heat-sensitive point mutation has been introduced into the ts125 mutant, making it possible to inactivate the 72 kDa DNA-binding protein (DBP). In some embodiments, the adenoviral vector comprises a deletion in the E1 region of its genome. In some embodiments, the adenoviral vector comprises a deletion in the E1a and E1b regions.

[0050] In some embodiments, the adenoviral vector comprise a deletion in the E4 region. The E4 region is involved in the regulation of the expression of the late genes, in the stability of the late nuclear RNAs, in decreasing host cell protein expression and in the efficiency of the replication of the viral DNA. In some embodiments, the adenoviral vector comprise deletions in both the E1 and the E4 region. Adenoviral vectors in which the E1 and

E4 regions are deleted therefore possess very reduced viral gene expression and transcriptional background noise. In some embodiments, the adenoviral vector is a "gutless" adenoviral vector in which all the viral genes are deleted.

[0051] The adenoviral vectors of the present application may possess other alterations in their genome. In particular, other regions may be deleted to increase the capacity of the virus and reduce its side effects linked to the expression of viral genes. Thus, all or part of the E3 or IVa2 region in particular may be deleted. As regards the E3 region, it may however be particularly preferred to conserve the part encoding the gp19K protein. This protein indeed makes it possible to prevent the adenoviral vector from becoming the subject of an immune reaction which (i) would limit its action and (ii) could have undesirable side effects. According to a specific mode, the E3 region is deleted and the sequence encoding the gp19K protein is reintroduced under the control of a heterologous promoter.

[0052] The cas gene and the coding sequences for sgRNAs can be inserted into various sites of the recombinant viral genome. It can be inserted into the E1, E3 or E4 region, as a replacement for the deleted or surplus sequences. It can also be inserted into any other site, outside the sequences necessary in cis for the production of the viruses (ITR sequences and encapsidation sequence).

[0053] In some embodiments, the adenoviral vector is replication deficient. In other embodiments, the adenoviral vector is a replication-competent adenoviral vector.

[0054] In other embodiments, the expression vector is a retroviral vector. Retroviruses are RNA viruses that replicate through an integrated DNA intermediate. Retroviral particles encapsidate two copies of the full-length viral RNA, each copy containing the complete genetic information needed for virus replication. Retroviruses possess a lipid envelope and use interactions between the virally encoded envelope protein that is embedded in the membrane and a cellular receptor to enter the host cells. Using the virally encoded enzyme reverse transcriptase, which is present in the virion, viral RNA is reverse transcribed into a DNA copy. This DNA copy is integrated into the host genome by integrase, another virally encoded enzyme. The integrated viral DNA is referred to as a provirus and becomes a permanent part of the host genome. The cellular transcriptional and translational machinery carries out expression of the viral genes. The host RNA polymerase II transcribes the provirus to generate RNA, and other cellular processes modify and transport the RNA out of the nucleus. A fraction of viral RNAs are spliced to allow expression of some genes whereas other viral RNAs remain full-length. The host translational machinery synthesizes and modifies the viral proteins. The newly synthesized viral proteins and the

newly synthesized full-length viral RNAs are assembled together to form new viruses that bud out of the host cells.

[0055] Retroviruses can also be classified into oncoviruses, lentiviruses, and spumaviruses. Most oncoviruses are simple retroviruses. Lentiviruses, spumaviruses, and some oncoviruses are complex retroviruses. Murine leukemia virus (MLV) is example of an oncovirus, human immunodeficiency virus 1 (HIV-1) is an example of a lentivirus, and human foamy virus is an example of a spumavirus. In some embodiments, the viral vector is a lentiviral vector.

[0056] A retrovirus can be altered by incorporating foreign envelope proteins, which can expand the potential target population of cells in a subject. Selection of a retroviral gene transfer system would therefore depend on the target tissue. Retroviral vectors are comprised of *cis*-acting long terminal repeats with packaging capacity for up to 6-10 kb of foreign sequence. The minimum *cis*-acting LTRs are sufficient for replication and packaging of the vectors. Widely used retroviral vectors include those based upon murine leukemia virus (MuLV), gibbon ape leukemia virus (GaLV), Simian Immunodeficiency virus (SIV), human immunodeficiency virus (HIV), and combinations thereof.

[0057] In some embodiments, the retroviral vector is a self-inactivating retroviral vector. After gene delivery, the vector will delete some of the cis-acting elements needed to complete another round of replication. Therefore, even in the presence of a replication-competent virus, these vectors cannot be transferred to other target cells efficiently.

[0100] Other viral vectors which may be used in aspects of the present application include Herpes Simplex Virus (HSV) vector, adeno-associated virus (AAV) vectors, vesicular stomatitis virus vectors, vaccinia virus vectors and SV-40-based viral vectors.

Non-viral vectors

[0101] In some embodiments, the expression vector of the present application is a non-viral expression vector, such as a plasmid comprising one or more expression cassettes. The non-viral expression vector may be administered using a delivery agent/method such as lipofection, nucleofection, microinjection, biolistics, virosomes, liposomes, immunoliposomes, polycation or lipid:nucleic acid conjugates, naked DNA, artificial virions, and agent-enhanced uptake of DNA. Lipofection is described in *e.g.*, U.S. Pat. Nos. 5,049,386, 4,946,787; and 4,897,355) and lipofection reagents are sold commercially (*e.g.*, Transfectam™ and Lipofectin™). Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides include those of Felgner, WO 91/17424;

WO 91/16024. Delivery can be to cells *(e.g. in vitro* or *ex vivo* administration) or target tissues (*e.g. in vivo* administration).

[0102] The practice of aspects of the present application employs, unless otherwise indicated, conventional techniques of immunology, biochemistry, chemistry, molecular biology, microbiology, cell biology, genomics and recombinant DNA, which are within the skill of the art. *See* Sambrook, Fritsch and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel, et al. eds., (1987)); the series METHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR 2: A PRACTICAL APPROACH (M. J. MacPherson, B. D. Hames and G. R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) ANTIBODIES, A LABORATORY MANUAL, and ANIMAL CELL CULTURE (R. I. Freshney, ed. (1987)).

Kit for inducing *in vivo* chromosomal rearrangements

[0103] Another aspect of the present application is a kit for inducing specific chromosomal rearrangements *in vivo*. The kit comprises: an expression vector comprising a polynucleotide encoding a Cas gene and two or more distinct single guide RNAs (sgRNAs); and one or more reagents for delivering the expression vector into a mammalian cell, wherein the expression vector is capable of expressing the Cas gene and the two or more sgRNAs inside the cell and causing chromosomal rearrangement between two chromosomal loci targeted by the two or more sgRNAs. In some embodiments, the expression vector is an adenoviral viral vector. In some embodiments, the kit further comprises instructions for how to use the kit.

Methods for inducing *in vivo* chromosomal rearrangements

[0104] Another aspect of the present application relates to a method for inducing specific chromosomal rearrangements *in vivo*. The method comprises the steps of administering into a subject an effective amount of an expression vector comprising a polynucleotide encoding a Cas gene and two or more distinct single guide RNAs (sgRNAs), wherein the expression vector expresses the Cas gene and the two or more sgRNAs inside a target cell and causes chromosomal rearrangement between two chromosomal loci targeted by the two or more sgRNAs.

[0105] In some embodiments, the two or more distinct sgRNAs consist of sgRNAs targeting *Eml4* and *Alk* sites. In other embodiments, the two or more distinct sgRNAs consist of sgRNAs targeting Npml and *Alk* sites.

[0106] In certain embodiments, the expression vector is a viral vector. In a particular embodiment, the viral vector is an adenoviral vector. In some embodiments, the expression vector is a non-viral vector.

[0107] In some embodiments, the method is utilized to correct a genomic deficiency in certain somatic or germ line cells. In other embodiments, the method is used to create animal models with conditions resulting from chromosomal rearrangement.

[0108] In some embodiments, the present application relates to a method for inducing specific chromosomal rearrangements *in vivo* to create an animal model of human non-small cell lung cancer (NSCLC). The method comprises infecting somatic cells in a subject with a viral vector capable of expressing the Cas9 gene and sgRNAs targeting *Eml4* and *Alk* sites, wherein the infection results in development of tumors harboring the *Eml4-Alk* inversion in the subject.

[0109] The CRISPR-based strategy described herein offers several advantages over germline engineering via transgenesis or homologous recombination. By inducing the rearrangement in only a subset of somatic cells, the resulting lesions more closely recapitulate the stochastic nature of tumor formation in humans. In addition, by modifying the endogenous loci, expression of the resulting fusion genes is subjected to physiologic transcriptional and post-transcriptional regulation, accurately modelling the reduced dosage of the wild-type alleles and the expression of the reciprocal product of the translocation/inversion. Finally, because the method requires only the generation of an appropriate viral vector and no germline manipulations, it can be readily adapted to model chromosomal rearrangements in other species, including non-human primates, and as such will facilitate the study of species-specific differences in tumor progression and therapy response *in vivo*.

Animal models

[0110] Another aspect of the present application relates to an animal produced by the methods of inducing an *in vivo* chromosomal rearrangement described herein, wherein the animal has NSCLCs that harbor the *Eml4-Alk* inversion, express the *Eml4-Alk* fusion gene, display histopathologic and molecular features typical of *Alk*+ human NSCLCs, and respond to treatment with *Alk*-inhibitors. In some embodiments, the animal is a mouse, a rat, a chimpanzee, a monkey, a sheep, a rabbit, a guinea pig, a frog, a zebrafish or any laboratory test animal as would be known to one of ordinary skill in the art.

[0111] The animal produced by the methods of inducing an *in vivo* chromosomal rearrangement described herein provides unique opportunities to dissect the molecular

mechanisms of diseases related to the chromosomal rearrangement, to test the efficacy of targeted therapies, and to investigate the mechanisms of drug resistance *in vivo*.

[0112] In some embodiments, one or more expression vectors described herein are used to produce a non-human transgenic animal. In some embodiments, the animal is a mammal, such as a mouse, rat, or rabbit. Methods for producing transgenic animals are known in the art, and generally begin with a method of cell transfection, such as is known in the art. The animal may be useful in applications outside of providing a disease model. In this regard, transgenic animals, especially mammals such as livestock (cows, sheep, goats and pigs), but also poultry and edible insects, are preferred.

[0113] Below are disclosed methods, devices, and systems for targeting the delivery of agents to specific cell types in mammals, in particular the lungs. Further aspects and advantages of the application will appear from the following description taken together with the accompanying drawings.

EXAMPLES

EXAMPLE 1: ADAPTING THE CRISPR SYSTEM TO ENGINEER CHROMOSOMAL REARRANGEMENTS IN EUKARYOTIC CELLS

[0114] In the mouse genome, *Eml4* and *Alk* are located on chromosome 17, approximately 11 megabases (Mb) apart, in a region that is syntenic to human chromosome 2(p21–p23) (Fig. 1, panel (a)). The most common *EML4–ALK* variant in human NSCLCs is modeled by introducing concomitant double-strand DNA breaks at intron 14 of *Eml4* (which corresponds to intron 13 of *EML4*) and at intron 19 of *Alk* (Fig. 1, panels (a) and (b) and Fig. 2). To induce the DNA breaks the CRISPR system was chosen because it only requires co-expression of Cas9 and an appropriately designed single guide RNA molecule (sgRNA).

[0115] The sgRNA sequences targeting the *Eml4* and *Alk* sites are cloned into the Cas9-expressing plasmid pX330 and the resulting constructs are-cotransfected in NIH/3T3. PCR analysis demonstrated the induction of the *Eml4–Alk* inversion and of a large deletion of the region between the two cut sites in the transfected cell population (Fig. 1, panel (b)). The presence of the desired *Eml4–Alk* inversion was confirmed by sequencing the corresponding *Eml4–Alk* fusion transcript (Fig. 1, panel (c)) and directly visualized by interphase FISH (Fig. 1, panels (d) and (e)). Using a similar strategy, this also modelled the *Npm1–Alk* rearrangement, a reciprocal chromosomal translocation commonly observed in anaplastic large cell lymphomas (Fig. 3). These results confirm that the CRISPR system can be adapted to engineer large deletions, inversions, and chromosomal translocations in eukaryotic cells.

[0116] The pX330 vector expressing Cas9 (Addgene plasmid 42230) was digested with BbsI and ligated to annealed and phosphorylated sgRNA oligos targeting *Eml4*, *Alk*, and Npm1. For cloning of tandem U6-sgRNA-Cas9 constructs, the second U6-sgRNA cassette was amplified using primers containing the XbaI and KpnI sites and cloned into the pX330 construct containing the appropriate sgRNA. For Adeno-*Eml4-Alk* cloning, pX330-*Alk-Eml4* vector was modified by adding an XhoI site upstream the first U6 promoter. An EcoRI/XhoI fragment containing the double U6-sgRNA cassette and the Flag-tagged Cas9 was then ligated the EcoRI/XhoI-digested pacAd5 shuttle vector. NIH/3T3 cells were transfected in 6-well plates with 3 μg of total plasmid DNA per well using lipofectamine 2000 (Invitrogen) following the manufacturer's instructions. To enrich for transfected cells, transfections included 1 μg of a plasmid expressing the Puro-resistance gene (pSico) and cells were incubated with 2 μg ml−1 puromycin for 2 days. Recombinant adenoviruses were generated by Viraquest (Ad-EA and Ad-Cas9) or purchased from the University of Iowa (Ad-Cre). MEFs infections were performed by adding adenovirus (3×10^6 p.f.u.) to each well of a 6-well plate.

PCR and RT–PCR analysis

[0117] For PCR analysis of genomic DNA, cells were collected in lysis buffer (100 nM Tris-HCl pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl supplemented with fresh proteinase K at final concentration of 100 ng ml−1). Genomic DNA was extracted with phenol-chloroform-isoamylic alcohol and precipitated in ethanol. The DNA pellet was dried and resuspended in double-distilled water. For RT–PCR, total RNAs were extracted with TRIzol (Life Technologies) following manufacturer's instructions. cDNAs were prepared using the Superscript III kit, following the manufacturer's instructions. The primers and the primer pairs used in the various PCR reactions are provided in Table 1 (which lists the name and sequence of each DNA oligonucleotide used in this study).

TABLE 1

SEQ ID NO	Name	Sequence
1	*Alk*_cDNA-rev	GGTCATGATGGTCGAGGTCC
2	*Alk*_Exon29_rev	GCTAGTGGAGTACAGGGCTC
3	*Alk*_gDNA-for (primer D Fig 1b)	GCAGCGGGGCTTCCGAAGGGGC
4	*Alk*_gDNA-rev (primer C Fig1b)	GTTTTACTGTGTCAGAAAGGG
5	*Alk*-rev	CAAGGCAGTGAGAACCTGAA
6	*Eml4*_cDNA-for	TGGAGTGGCAACTCACTAACAA
7	*Eml4*_cDNA-rev	GCAACTGCTCTAATGGTGCC
8	*Eml4*_Exon1_for	TAGAACTCGAGGCAAGATGGACGG TTTCGC
9	*Eml4*_gDNA-for (primer A Fig1b)	GCTCAAGAGGTGGGTTGTGT
10	*Eml4*_gDNA-rev (primer B Fig1b)	CAGGGCTGTGCCTAGATGAC
11	*Eml4*-for	GAGCCTTGTTGATACATCGTTC
12	*Eml4*-rev	TAGGAGGCAGTTTGGGCTAC
13	GAPDH_cDNA-for	ACCACAGTCCATGCCATCACTGCC
14	GAPDH_cDNA-rev	GTCTCGCTCCTGGAAGATGG
15	miR17-92_gDNA-for	TCGAGTATCTGACAATGTGG
16	miR17-92_gDNA-rev	TAGCCAGAAGTTCCAAATTGG
17	Npm1_cDNA-for	ACTACCTTTTCGGCTGTGAACT
18	Npm1_gDNA-for (primer A)	GTCTCTTGCGTCATTTGGGG
19	Npm1_gDNA-rev (primer B)	CTCCAGGAGCAGATCGCTTT

Quantification of inversion efficiency in MEFs

[0118] An NIH/3T3 subclone was first isolated carrying a mono-allelic *Eml4–Alk* inversion validated by interphase FISH. Genomic DNA extracted from this clone was mixed with increasing amounts of genomic DNA from parental NIH/3T3 cells to generate a series of standards containing known percentage of *Eml4–Alk* alleles. The standards and the test samples were then subjected to quantitative PCR (Applied Biosystem) using primers amplifying the *Eml4–Alk* junction (*Eml4*-for and *Alk*-rev, see Table 1) or a control gene (miR-17~92-gDNA-for and miR-17~92-gDNA-rev) and the fraction of *Eml4-Alk* alleles in

the test was calculated by plotting the ΔΔCt values on the standard curve. qPCR analysis was performed using SYBR Green (Life Technology).

Cell lines

[0119] MEFs were generated from E14.5 wild-type embryos following standard procedures. NIH/3T3 were purchased from ATCC.

EXAMPLE 2: DELIVERY OF sgRNA/Cas9 USING ADENOVIRAL VECTORS

[0120] Although appropriate for cell-based experiments, expression of two sgRNAs from separate constructs would be impractical *in vivo*. Therefore, plasmids were engineered to simultaneously express Cas9 and two distinct sgRNAs from tandem U6 promoters (Fig. 4, panel (a)). Their transfection in NIH/3T3 cells resulted in comparable levels of the two sgRNAs, efficient cleavage at the targeted sites, and accumulation of the *Eml4–Alk* inversion (Fig. 4, panels (b)–(d)).

[0121] To deliver Cas9 and sgRNAs targeting the *Alk* and *Eml4* loci to the lungs of adult mice, next the dual sgRNA/Cas9 cassette was transferred into an adenoviral shuttle vector (Fig. 5, panel (a)), which was then used to produce recombinant adenoviruses (hereafter referred to as 'Ad-EA'). Adenoviral vectors are ideal because they efficiently infect the lung epithelium of adult mice and do not integrate into the host genome. Infection of mouse embryo fibroblasts (MEFs) with Ad-EA led to the expression of Cas9 and both sgRNAs, and to the rapid generation of the desired *Eml4–Alk* inversion (Fig. 5, panels (b) and (c)). The *Eml4–Alk* inversion was estimated to occur in approximately 3–4% of infected MEFs (Fig. 5, panels (d)-(e)).

Mouse husbandry and adenoviral infection

[0122] Mice were purchased from The Jackson Laboratory (C57BL/6J) or from Charles River (CD1) and housed in the SPF Memorial Sloan-Kettering Cancer Center Institutional Animal Care and Use Committee, where the health status of the colony is constantly monitored by the veterinary staff and by a sentinel program. For adenoviral infection, 6–10-week-old mice were anaesthetized by intra peritoneal injection of ketamine (80 mg per kg) and xylazine (10 mg per kg) and treated by intratracheal instillation of 1.5×10^8 PFU adenovirus/mouse, as previously described (DuPage *et al.*, Nature Protocols, 2009, 4:1064-1072). Investigators were not blinded with respect to which adenovirus was injected. All studies and procedures were approved by the MSKCC Institutional Animal Care and Use Committee.

Surveyor assay

[0123] The genomic region flanking the CRISPR/Cas9 target site was first amplified by PCR. After a cycle of melting and re-annealing to allow heteroduplex formation, the amplicon was digested with the surveyor nuclease (Transgenomic) for 1 h at 42 °C according to the manufacturer's directions and the digestion products were separated on a 2% agarose gel.

Northern blot analysis

[0124] 10 μg of RNA previously extracted with TRIzol (Life Technologies) were run on a 15% denaturing polyacrylamide gel and blotted on a nitrocellulose membrane for 1 h at 100 V at room temperature. The membranes were then hybridized to radiolabelled oligonucleotides complementary to the *Alk* (5′-TACAGATAGACATGCCAGGAC) (SEQ ID NO: 20), *Eml4* (5′-TCCTAGTAGACCCCGACAAAC) (SEQ ID NO: 21) sgRNAs, or mU6 (5′-GCAGGGGCCATGCTAATCTTCTCTGTATCG) (SEQ ID NO: 22) dissolved in ExpressHyb (Clontech) at 42 °C overnight. Washes were performed at room temperature in 2X SSC and 0.2 SSC.

EXAMPLE 3: INDUCING CHROMOSOMAL REARRANGEMENTS IN MICE

[0125] To induce the *Eml4–Alk* rearrangement *in vivo* a cohort of adult CD1 and C57BL/6J (B6) mice were next infected by intratracheal instillation of Ad-EA (n = 52: 22 B6, 30 CD1) or control adenoviruses expressing either the Cre recombinase (Ad-Cre, n = 15: 6 B6, 9 CD1) or Cas9 alone (Ad-Cas9, n = 19: 9 B6, 10 CD1).

[0126] At two days, and at one week post-infection the lungs appeared histologically normal with no obvious signs of cytoxicity except for the presence of occasional inflammatory infiltrates (data not shown). However, one month after Ad-EA infection, the lungs of mice of both strains presented multiple small lesions that, upon histopathological examination, appeared to be papillary intrabronchiolar epithelial hyperplasia, atypical adenomatous hyperplasia (AAH) or early well-differentiated adenocarcinomas. By 6–8 weeks post-infection, larger tumors were easily detectable by micro-computed tomography (μCT) and macroscopically visible at necropsy. At 12–14 weeks post-infection, the lungs of Ad-EA-infected mice invariably contained multiple large lesions histologically classified as lung adenocarcinomas.

[0127] In Ad-EA-infected animals, multiple bilateral lung tumors were frequently detected by 4–7 weeks post-infection (n = 23/26 mice), and invariably after 8 weeks post-infection (n = 34). In contrast, Ad-Cre-infected mice remained tumor-free at all time points examined (n = 14 mice, range 4–18 weeks), with the exception of two CD1 mice in each of which was observed a single small adenoma. Analogously, even at the latest time point

examined (9 weeks post-infection), none of the Ad-Cas9 infected mice presented lung tumors (n = 8 mice), whereas at same time point all Ad-EA infected mice had developed multiple tumors (P value < 0.0001, Fisher's exact test). These results indicate that intra-tracheal delivery of Ad-EA can initiate lung tumorigenesis with high penetrance and low latency, and that this effect cannot be attributed to adenoviral infection or Cas9 expression alone.

[0128] All tumors examined were positive for the pneumocyte marker Nkx-2.1/TTF1 and negative for p63 and Sox2, in agreement with the diagnosis of lung adenocarcinoma. The tumors were also strongly positive for the alveolar type II marker surfactant protein C (SpC), whereas the Clara cell marker CCSP/CC10 was undetectable. The adenocarcinomas had a papillary or, less frequently, acinar architecture. Most of these tumors were in close proximity to bronchi and bronchioles showing papillary epithelial hyperplasia, and areas of AAH were frequently observed, especially at earlier time points. The majority of tumor cells appeared low-grade, with occasional instances of intermediate nuclear atypia with enlarged nuclei and prominent nucleoli. Approximately 20% of tumors contained cells with a large cytoplasmic vacuole and a peripherally located nucleus. These cells are reminiscent of signet ring cells, which are commonly observed in human *ALK*+ NSCLC. Approximately 30% of adenocarcinomas displayed areas of intense positivity at the periodic acid-Schiff (PAS) staining.

EXAMPLE 4: DEMONSTRATION OF *IN VIVO* CHROMOSOMAL REARRANGEMENTS

[0129] Interphase FISH analysis demonstrated the presence of a mono- or bi-allelic *Eml4–Alk* inversion in every Ad-EA-induced tumor examined (n = 4 animals), but not in control K-RasG12D-driven tumors (data not shown). The presence of the inversion and expression of the full-length *Eml4–Alk* transcript by PCR and PCR with reverse transcription (RT–PCR) analysis of micro-dissected tumors followed by sequencing (Fig. 6, panels (a)–(c)) was further confirmed.

[0130] Activation of the human *ALK* oncogene via deregulation, translocation or amplification has been shown to lead to constitutive phosphorylation of ERK, STAT3, and AKT. At 12–14 weeks post-infection, all lung tumors derived from Ad-EA-injected mice showed phosphorylation and nuclear localization of Stat3. Phosphorylation of Akt and Erk1/2 were also frequently, but not invariably, observed (Fig. 6, panel (d)).

Interphase fluorescent *in situ* hybridization

[0131] Interphase FISH experiments were performed and interpreted by the MSKCC cytogenetic core using a 3-colour probe mix designed to detect and discriminate between

Alk–Eml4 fusion and other rearrangements of *Alk*. The probe mix comprised mouse BAC clones mapping to: 3′ *Alk* (17qE1.3, RP23-306H20, RP23-397M18 labelled with Green dUTP), 5′ *Alk* (17qE1.3, RP23-12H17, RP23-403F20 labelled with Red dUTP), and 5′ *Eml4* (17qE4, RP23-193B15 labelled with Orange dUTP). Probe labelling, hybridization, washing, and fluorescence detection were done according to standard procedures. Cell line harvest and metaphase spreads were prepared according to standard cytogenetics procedures. For NIH/3T3, FISH signals were enumerated in a minimum of 20 metaphases to determine locus specificity, and 100 interphase cells to determine *Alk–Eml4* fusion status. Each paraffin section was first scanned under ×100 objective to assess signal pattern and select representative regions for analysis. At least three images per representative region were captured (each image was a compressed stack of 12 z-sections at 0.5 micron intervals). Signal counts were performed on the captured images and a minimum of 50 interphase nuclei was analysed to determine the *Alk-Eml4* fusion status. Based on the observed distance between the green (3′ *Alk*), red (5′ *Alk*), and orange (5′ *Eml4*) signal in the negative controls (parental cell line and Ad-Cre-infected cells), interphase cells were classified as normal, *Eml4–Alk* positive, or other.

EXAMPLE 5: SENSITIVITY OF TUMORS INDUCED BY *IN VIVO* CHROMOSOMAL REARRANGEMENTS TO CRIZOTINIB

[0132] Finally, the sensitivity of Ad-EA-induced lung tumors to crizotinib was examined, a dual *ALK*/MET inhibitor used in the clinic to treat patients affected by *ALK*+ NSCLCs. Ten Ad-EA infected CD1 mice were monitored by μCT scans starting at 9 weeks post-infection until the appearance of multiple large lung tumors, at which point the animals were randomly assigned to receive a daily dose of crizotinib (100mg/kg; n = 7) or vehicle (n = 3) (Fig. 7, panel (a)). After two weeks of treatment the animals in the crizotinib group displayed complete (6/7) or partial (1/7) tumor regression, as indicated by μCT scans and confirmed at necropsy, whereas all control animals showed signs of disease progression (Fig. 7, panels (b) and (c)). Histological analysis showed that in the crizotinib group the tumors had undergone marked atrophy or were replaced by areas of intense inflammatory necrosis (Fig. 7, panels (d) and (e)).

Lung processing and antibodies for immunohistochemistry

[0133] Lungs were inflated by intratracheal injection of 4% paraformaldehyde (PFA), incubated for 18–24 h in 4% PFA, and then transferred to 70% ethanol for at least 24 h before further processing. The following antibodies were used: phospho-Stat3 (Tyr705, Cell Signaling Technology #9135, 0.1 $\mu g\ ml^{-1}$); phospho-Erk1/2 (Thr202/Tyr204, Cell Signaling

Technology #4370 1 μg ml^{-1}); phospho-Akt (Ser473, Cell Signaling Technology #4060 1 μg ml^{-1}); Nkx-2.1 (Epitomics, EP1584Y 1:1,200); FLAG (Sigma, M2 1:1,000); P63 (Santa Cruz (H-137) sc8343, 1:1,000); Sox2 (Cell Signaling Technology, C70B1 #3728, 1:1,000); CC10/CCSP (Millipore, 07-623, 1:2,000); SpC (Millipore, AB3786, 1:1,000).

μCT imaging

[0134] μCT Scans were performed on the Mediso Nano SPECT/CT System covering only the lung fields of each mouse. Each scan averaged approximately 5 min using 240 projections with an exposure time of 1,000 ms set at a pitch of 1 degree. The tube energy of the X-ray was 55 kVp and 145 μA. The in-plane voxel sizes chosen were small and thin creating a voxel size of 73 × 73 × 73 μm. The final reconstructed image consisted of 368 × 368 × 1,897 voxels. Scans were analysed with the Osirix software.

Crizotinib treatment

[0135] Mice were randomized to receive either control vehicle (water) or crizotinib at 100 mg/kg p.o. daily for at least 14 consecutive days. Mice were monitored daily for weight loss and clinical signs. Investigators were not blind with respect to treatment.

[0136] The foregoing descriptions of specific embodiments of the present application have been presented for purposes of illustration and description. They are not intended to be exhaustive or to limit the application and method of use to the precise forms disclosed. Obviously many modifications and variations are possible in light of the above teaching. It is understood that various omissions or substitutions of equivalents are contemplated as circumstance may suggest or render expedient, but is intended to cover the application or implementation without departing from the spirit or scope of the claims of the present application.

What is claimed is:

1. A method for inducing chromosomal rearrangements *in vivo*, comprising:

administering into a subject an effective amount of an expression vector comprising a polynucleotide encoding a Cas gene and two or more distinct single guide RNAs (sgRNAs),

wherein the expression vector expresses the Cas gene and the two or more sgRNAs *in vivo* and causes chromosomal rearrangements between two chromosomal loci targeted by the two or more sgRNAs

2. The method of Claim 1, wherein the expression vector is a viral vector.

3. The method of Claim 2, wherein the viral vector is an adenoviral vector.

4. The method of Claim 1, wherein the expression vector is an non-viral vector.

5. The method of Claim 1, wherein the Cas gene is a Cas9 gene.

6. The method of Claim 1, wherein the expression vector is the only expression vector administered to the subject.

7. The method of Claim 1, wherein the Cas gene and each of the two or more distinct sgRNAs are controlled by separate regulatory element.

8. The method of Claim 1, wherein the subject is a mouse and wherein the two or more distinct sgRNAs consist of sgRNAs targeting *Eml4* and *Alk* sites.

9. The method of Claim 1, wherein the subject is a mouse and wherein the two or more distinct sgRNAs consist of sgRNAs targeting Npml and *Alk* sites.

10. A mouse produced by the method of Claim 7, wherein the mouse has non-small cell lung cancer (NSCLC) cells harboring an *Eml4-Alk* inversion and expressing an *Eml4-Alk* fusion gene.

11. The mouse of Claim 10, wherein the non-small cell lung cancer (NSCLC) cells display histopathologic and molecular features typical of *Alk*+ human NSCLCs, and respond to treatment with *Alk*-inhibitors.

12. A composition for inducing specific chromosomal rearrangements *in vivo*, comprising an expression vector comprising a polynucleotide encoding a Cas gene and two or more distinct single guide RNAs (sgRNAs); wherein the expression vector is capable of expressing the Cas gene and the two or more sgRNAs *in vivo* and causing chromosomal rearrangements between two chromosomal loci targeted by the two or more sgRNAs.

13. The composition of Claim 12, wherein the expression vector is a viral vector.

14. The composition of Claim 13, wherein the viral vector is an adenoviral vector.

15. The composition of Claim 12, wherein the expression vector is an non-viral vector.

16. The composition of Claim 12, wherein the Cas gene is a Cas9 gene.

17. The method of Claim 12, wherein the Cas gene and each of the two or more distinct sgRNAs are controlled by separate regulatory element.

18. A kit for inducing specific chromosomal rearrangements *in vivo*, comprising:

the composition of Claim 12; and

a reagent for introducing the expression vector into a cell *in vivo*.

19. The kit of Claim 18, wherein the cell is a somatic cell.

20. The kit of Claim 18, further comprising instructions for introducing the expression vector into the cell.

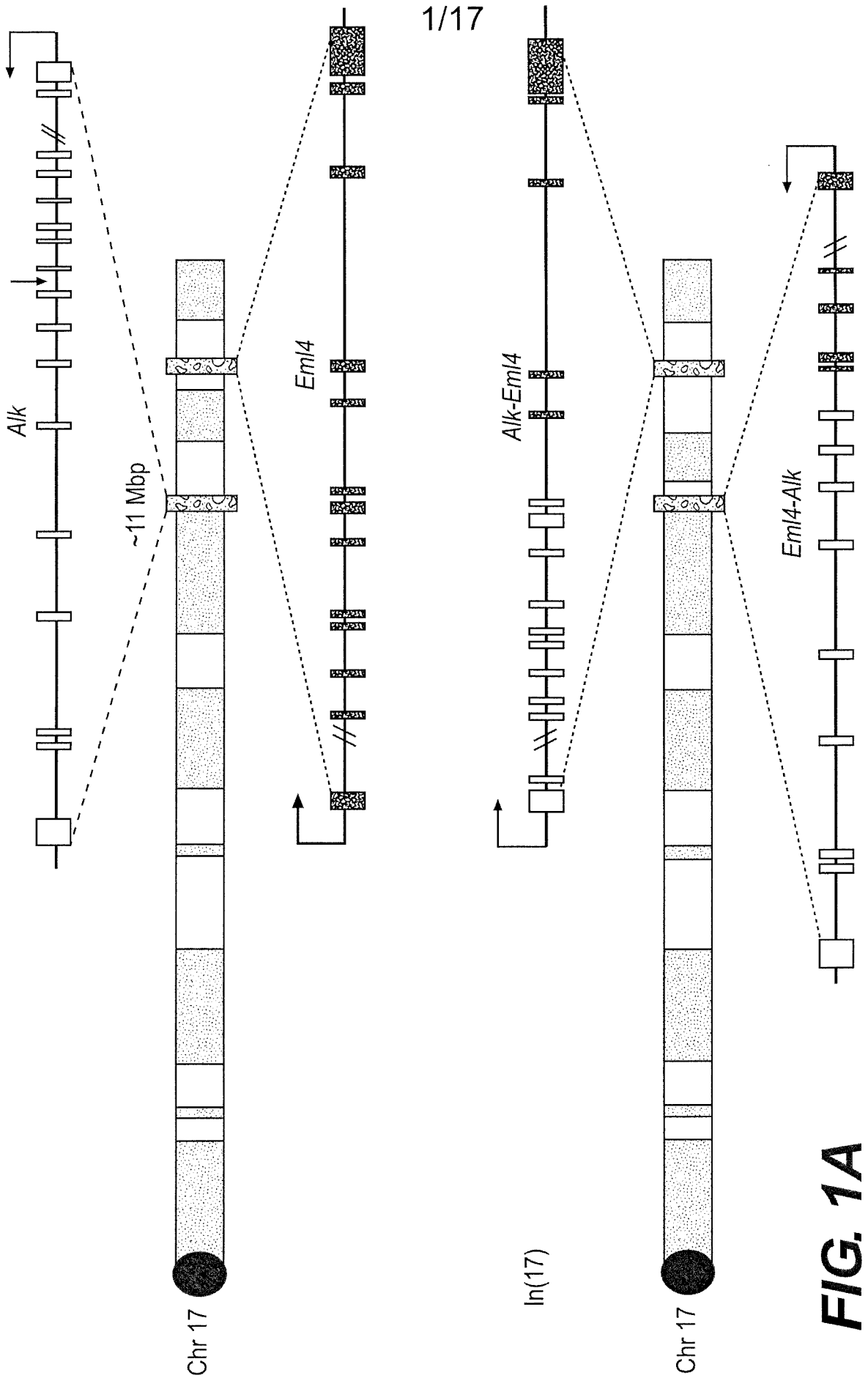


FIG. 1A

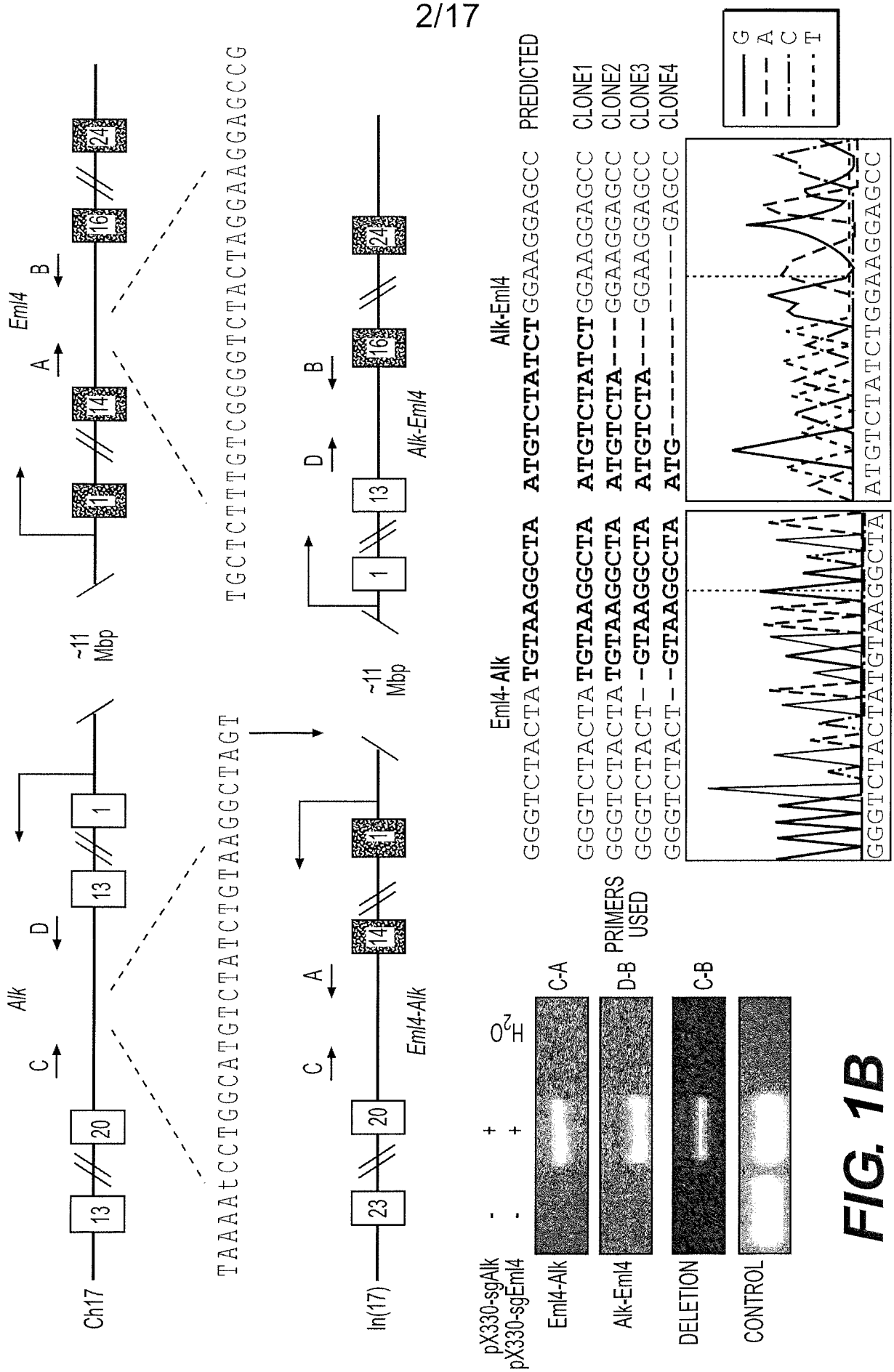

Ch17
Alk
Eml4
~11 Mbp
TAAAAtCCTGGCATGTCTATCTGTAAGGCTAGT
TGCTCTTTGTCGGGGTCTACTAGGAAGGAGCCG
In(17)
Eml4-Alk
Alk-Eml4
Eml4-Alk
Alk-Eml4
DELETION
CONTROL
pX330-sgAlk
pX330-sgEml4
PRIMERS USED
C-A
D-B
C-B
Eml4- Alk
Alk-Eml4
PREDICTED
CLONE1
CLONE2
CLONE3
CLONE4
GGGTCTACTATGTAAGGCTA
GGAAGGAGCC
G
A
C
T

FIG. 1B

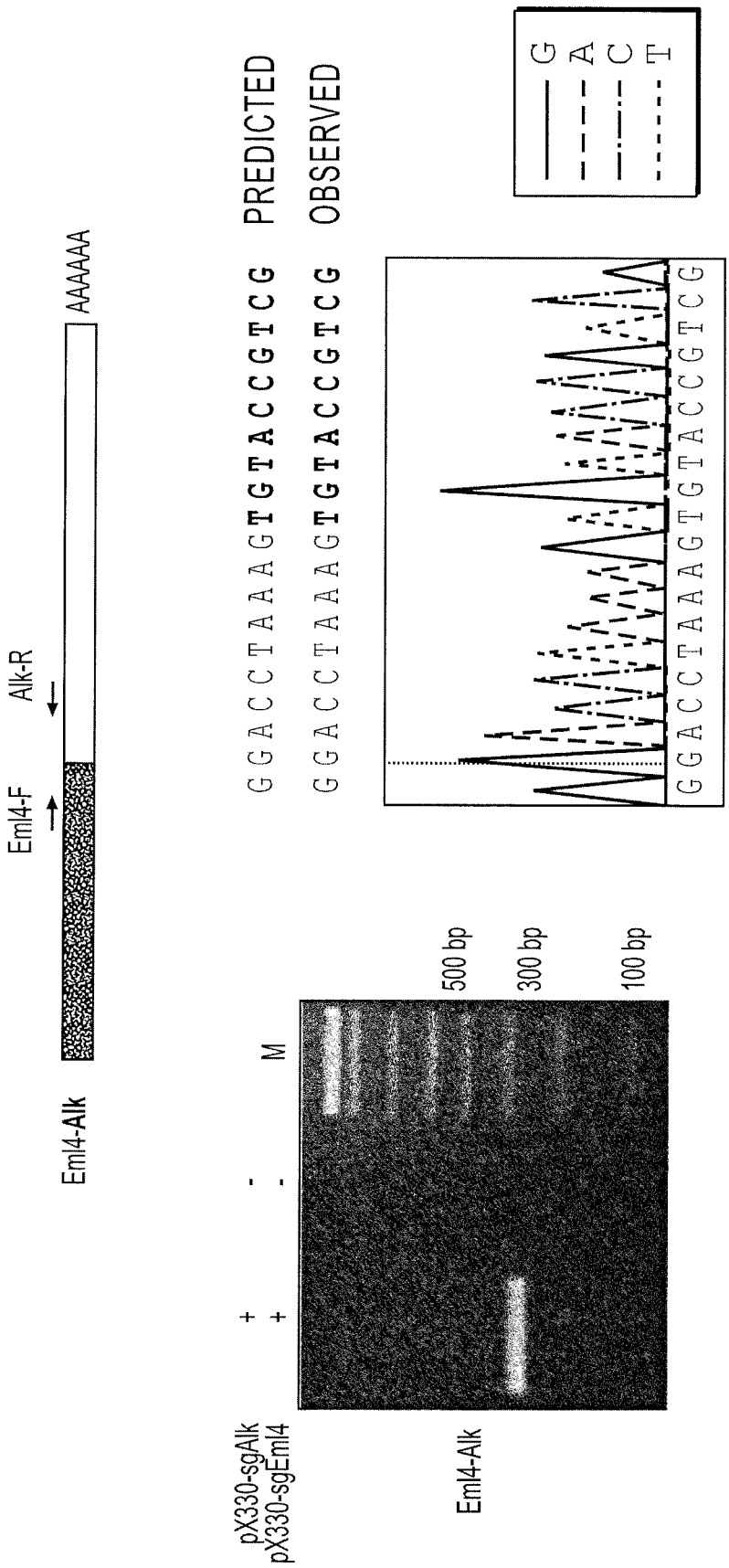

Eml4-F
Alk-R
AAAAAA
Eml4-Alk
M
500 bp
300 bp
100 bp
pX330-sgAlk
pX330-sgEml4
Eml4-Alk
PREDICTED
OBSERVED
G G A C C T A A A G T G T A C C G T C G
G G A C C T A A A G T G T A C C G T C G
G
A
C
T

FIG. 1C

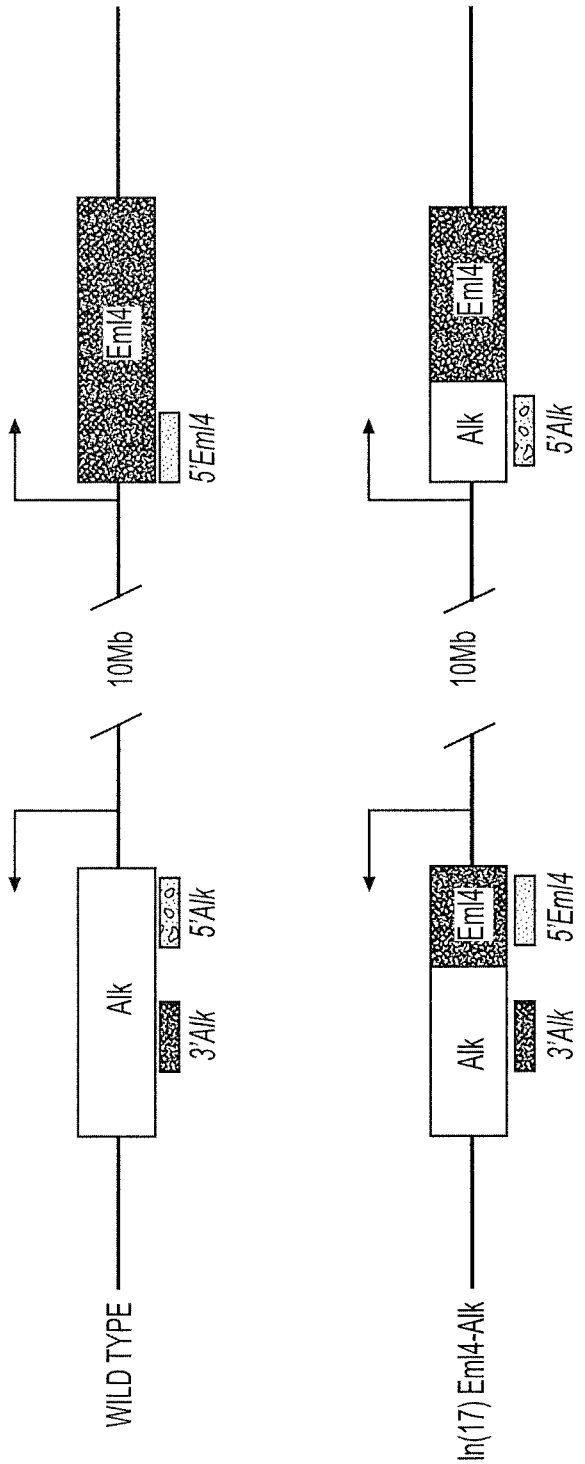

WILD TYPE
Alk
3'Alk
5'Alk
10Mb
Eml4
5'Eml4
In(17) Eml4-Alk
Eml4
Alk
3'Alk
5'Eml4
10Mb
Alk
5'Alk

FIG. 2

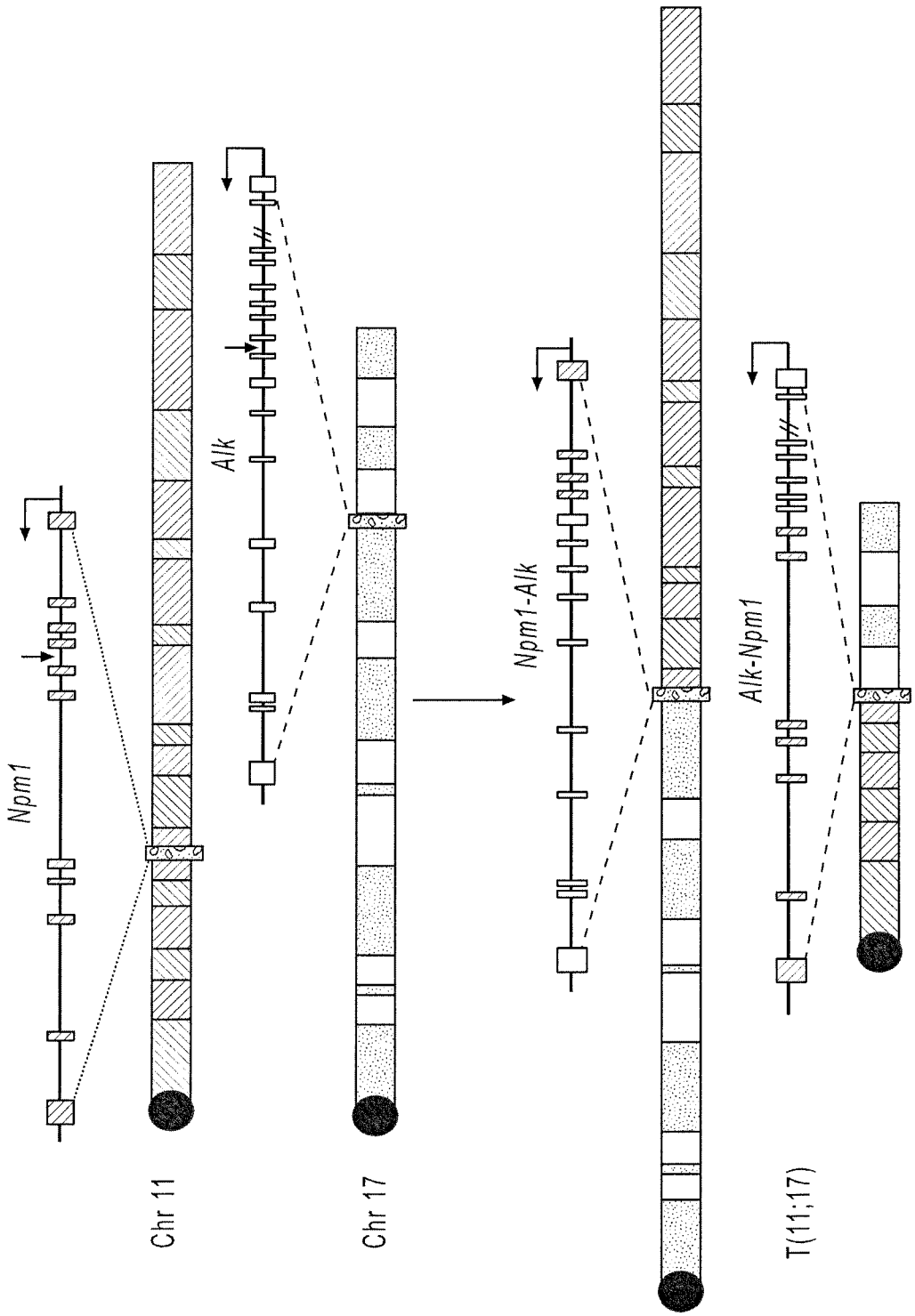


FIG. 3A

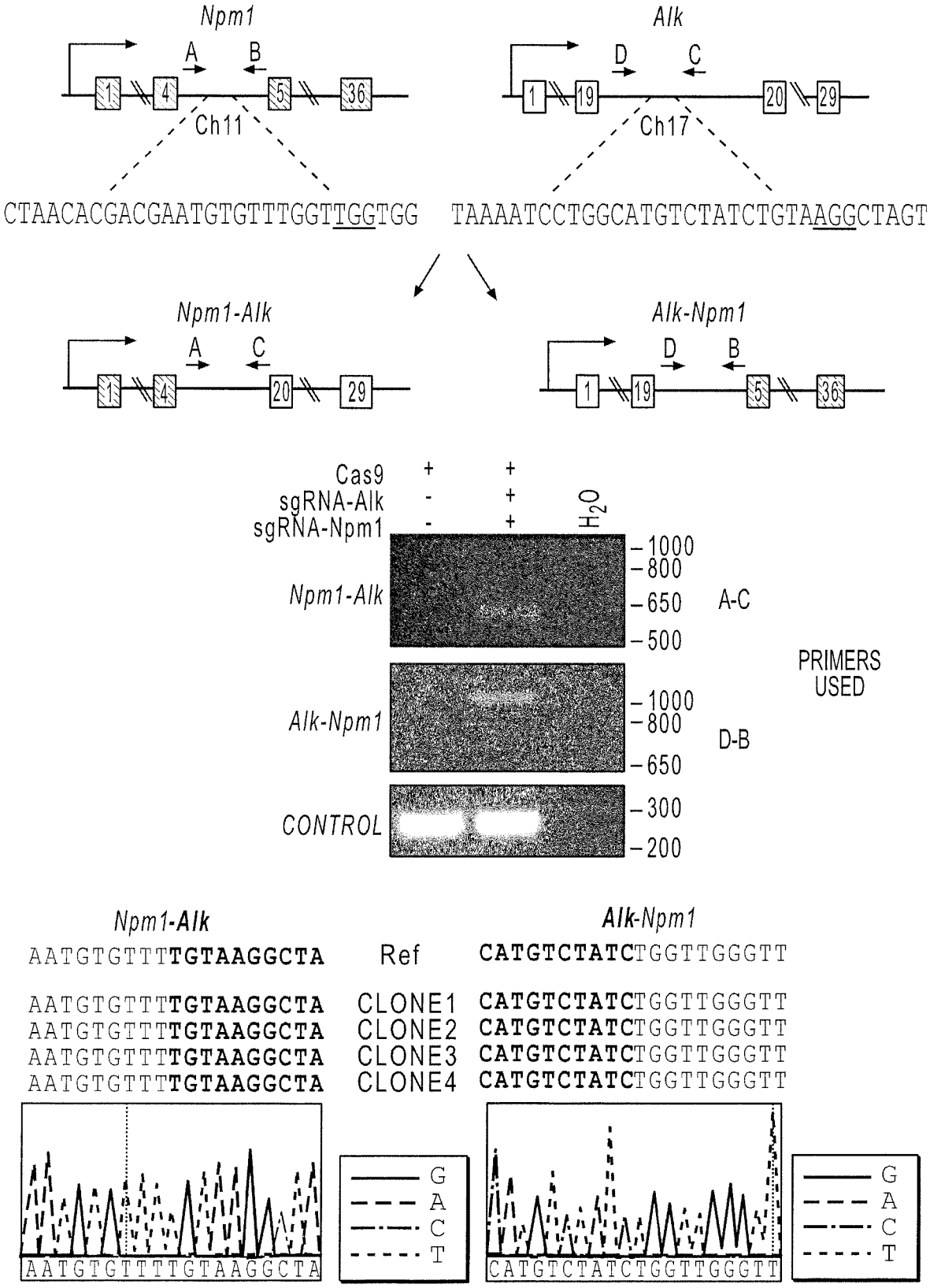
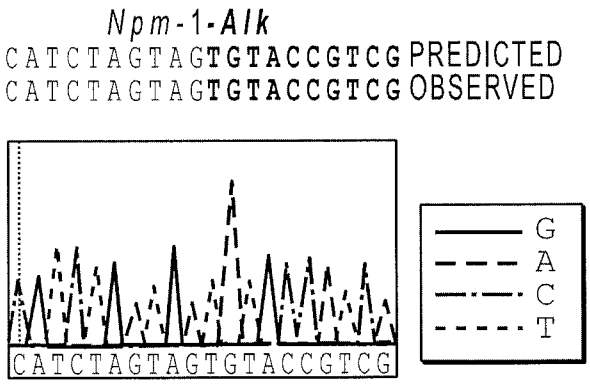

Npm1
A
B
1
4
5
36
Ch11
CTAACACGACGAATGTGTTTGGTTGGTGG
Alk
D
C
1
19
20
29
Ch17
TAAAATCCTGGCATGTCTATCTGTAAGGCTAGT
Npm1-Alk
A
C
1
4
20
29
Alk-Npm1
D
B
1
19
5
36
Cas9 + +
sgRNA-Alk - + H2O
sgRNA-Npm1 - +
Npm1-Alk
1000
800
650
500
A-C
Alk-Npm1
1000
800
650
D-B
CONTROL
300
200
PRIMERS USED
Npm1-Alk
AATGTGTTTTGTAAGGCTA
Ref
Alk-Npm1
CATGTCTATCTGGTTGGGTT
AATGTGTTTTGTAAGGCTA CLONE1 CATGTCTATCTGGTTGGGTT
AATGTGTTTTGTAAGGCTA CLONE2 CATGTCTATCTGGTTGGGTT
AATGTGTTTTGTAAGGCTA CLONE3 CATGTCTATCTGGTTGGGTT
AATGTGTTTTGTAAGGCTA CLONE4 CATGTCTATCTGGTTGGGTT
G
A
C
T
AATGTGTTTTGTAAGGCTA
CATGTCTATCTGGTTGGGTT

FIG. 3B


Npm1-F
Alk-R
Npm1-Alk
AAAAAA


Npm-1-Alk
CATCTAGTAGTGTACCGTCG PREDICTED
CATCTAGTAGTGTACCGTCG OBSERVED
CATCTAGTAGTGTACCGTCG
G
A
C
T

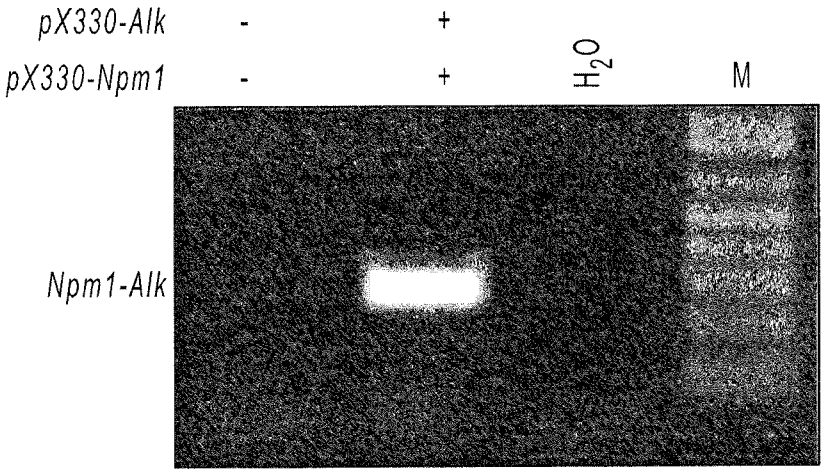

pX330-Alk
pX330-Npm1
-
+
-
+
H2O
M
Npm1-Alk

FIG. 3C

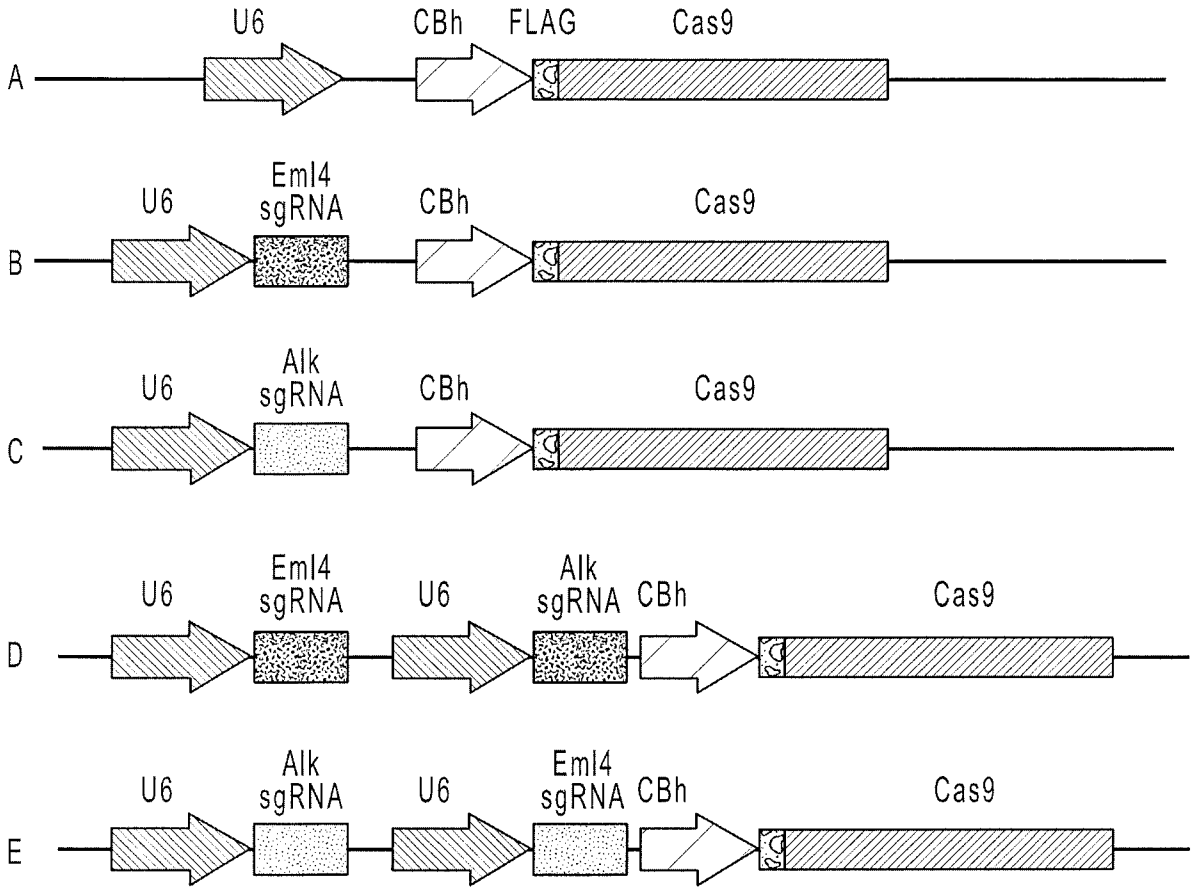

A
U6
CBh
FLAG
Cas9
B
U6
Eml4 sgRNA
CBh
Cas9
C
U6
Alk sgRNA
CBh
Cas9
D
U6
Eml4 sgRNA
U6
Alk sgRNA
CBh
Cas9
E
U6
Alk sgRNA
U6
Eml4 sgRNA
CBh
Cas9

FIG. 4A

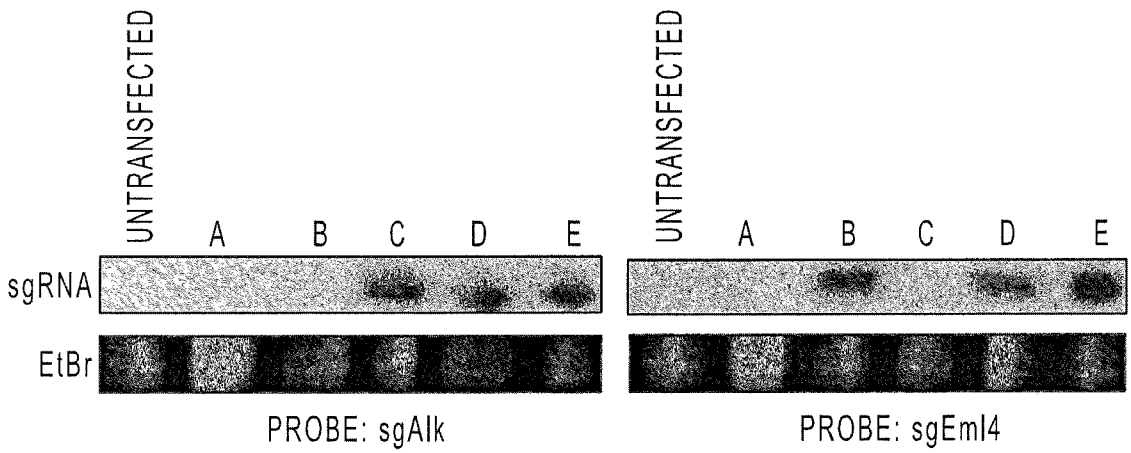

UNTRANSFECTED
A
B
C
D
E
UNTRANSFECTED
A
B
C
D
E
sgRNA
EtBr
PROBE: sgAlk
PROBE: sgEml4

FIG. 4B

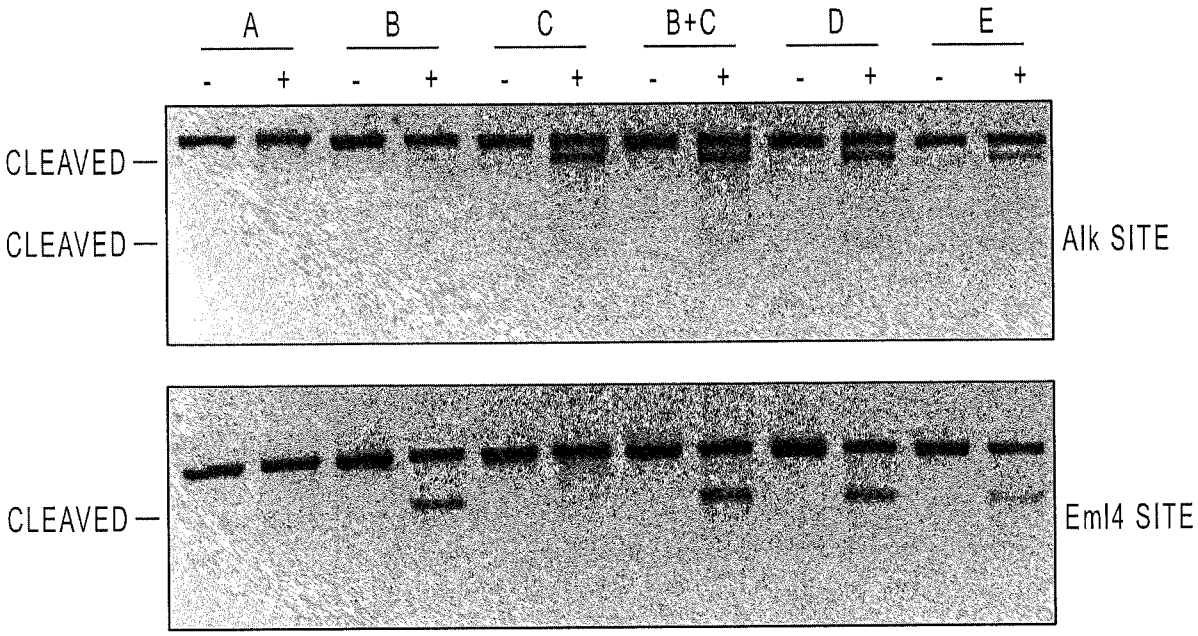

A
B
C
B+C
D
E
- + - + - + - + - + - +
CLEAVED
CLEAVED
Alk SITE
CLEAVED
Eml4 SITE

FIG. 4C

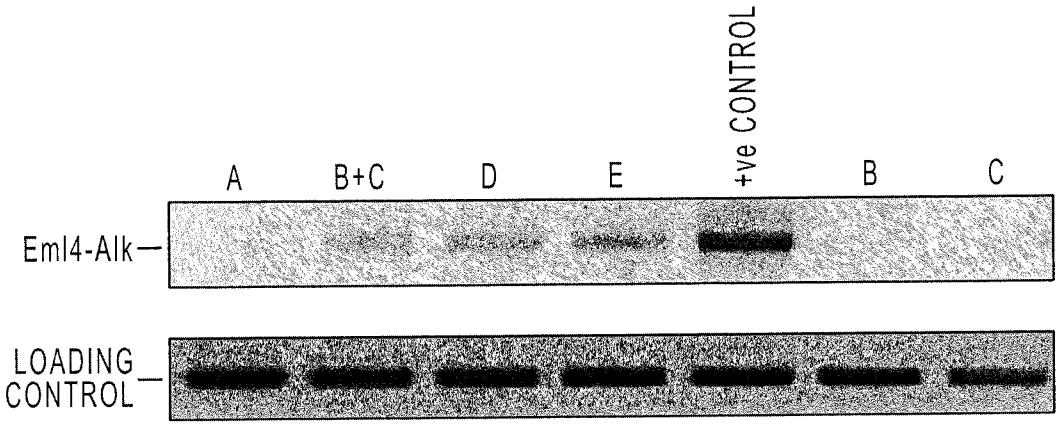

+ve CONTROL
A
B+C
D
E
B
C
Eml4-Alk
LOADING CONTROL

FIG. 4D

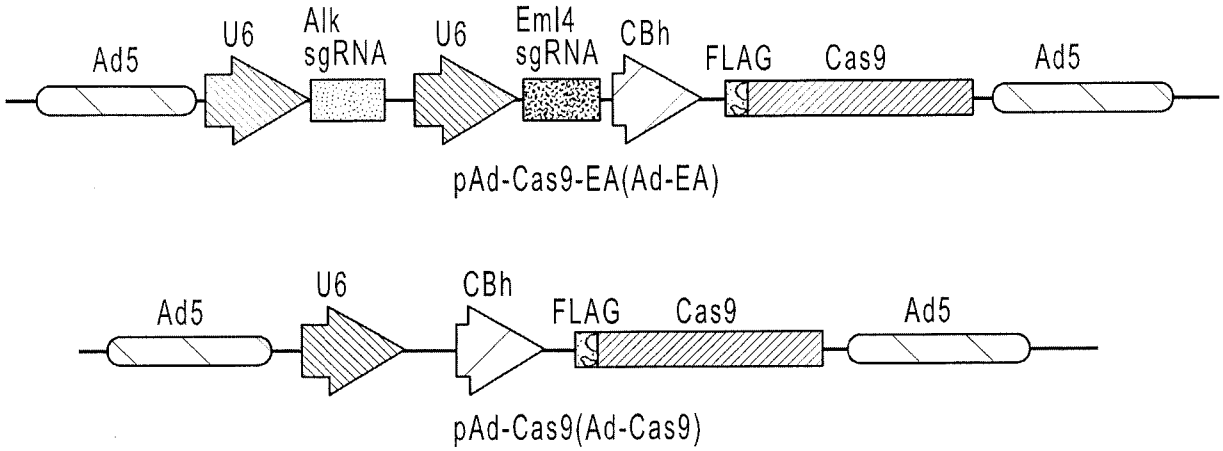


FIG. 5A

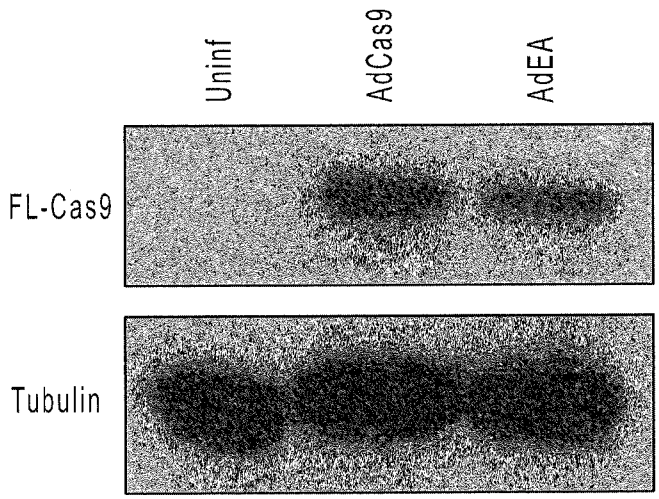


FIG. 5B

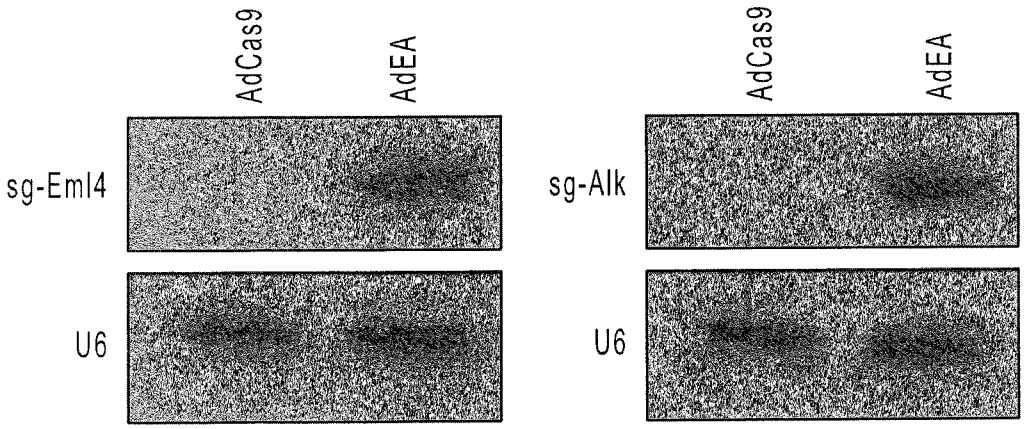


FIG. 5C

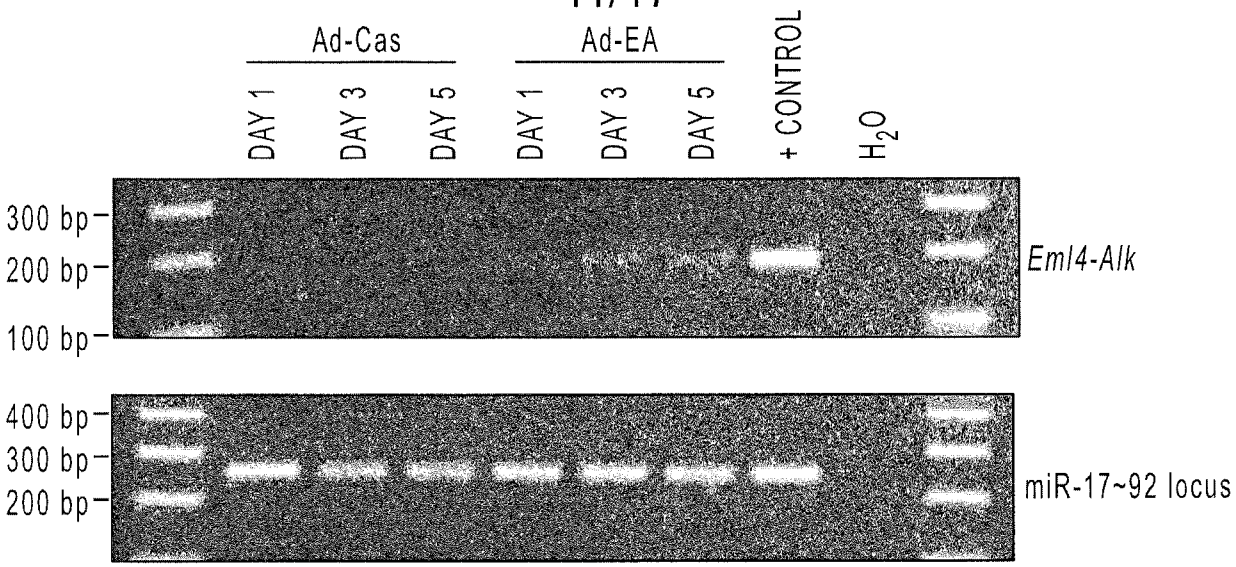


FIG. 5D

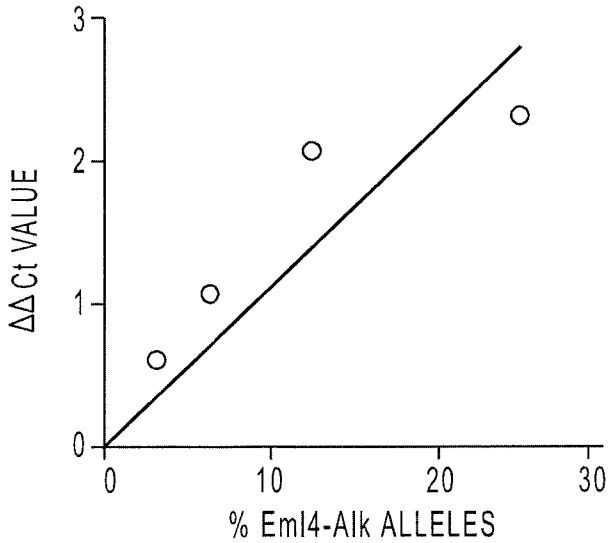


FIG. 5E

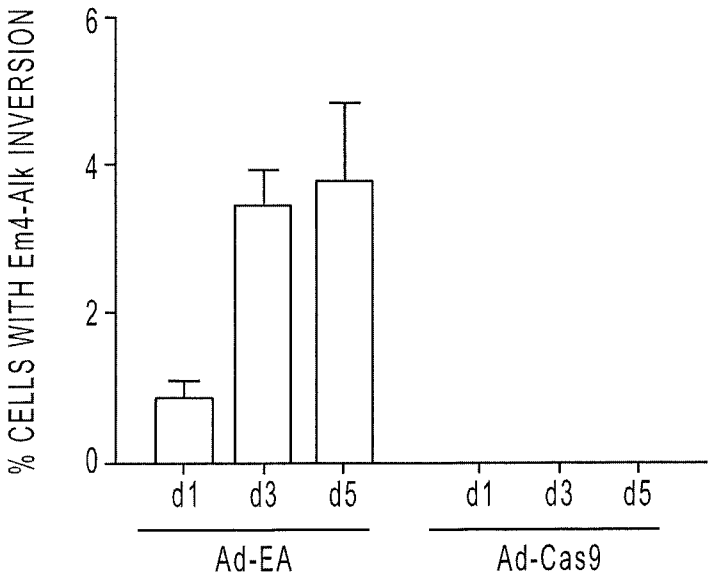


FIG. 5F

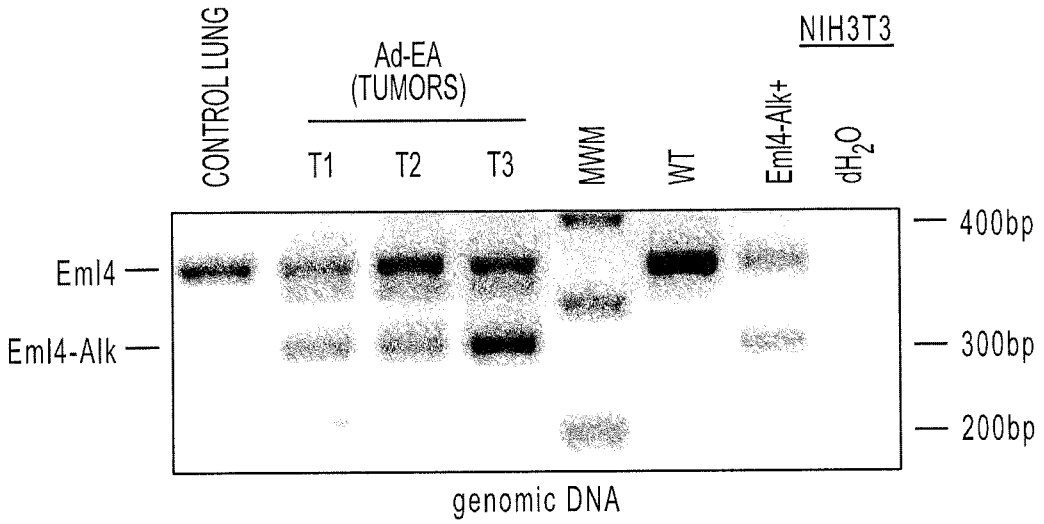

CONTROL LUNG
Ad-EA (TUMORS)
T1
T2
T3
MWM
WT
NIH3T3
Eml4-Alk+
dH2O
Eml4
Eml4-Alk
400bp
300bp
200bp
genomic DNA

FIG. 6A

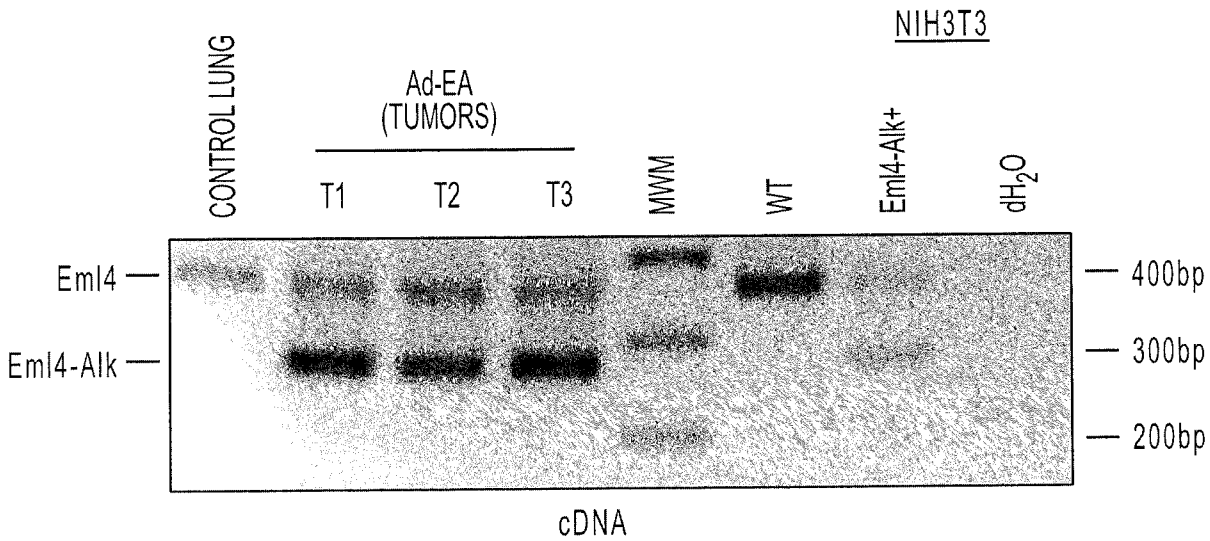
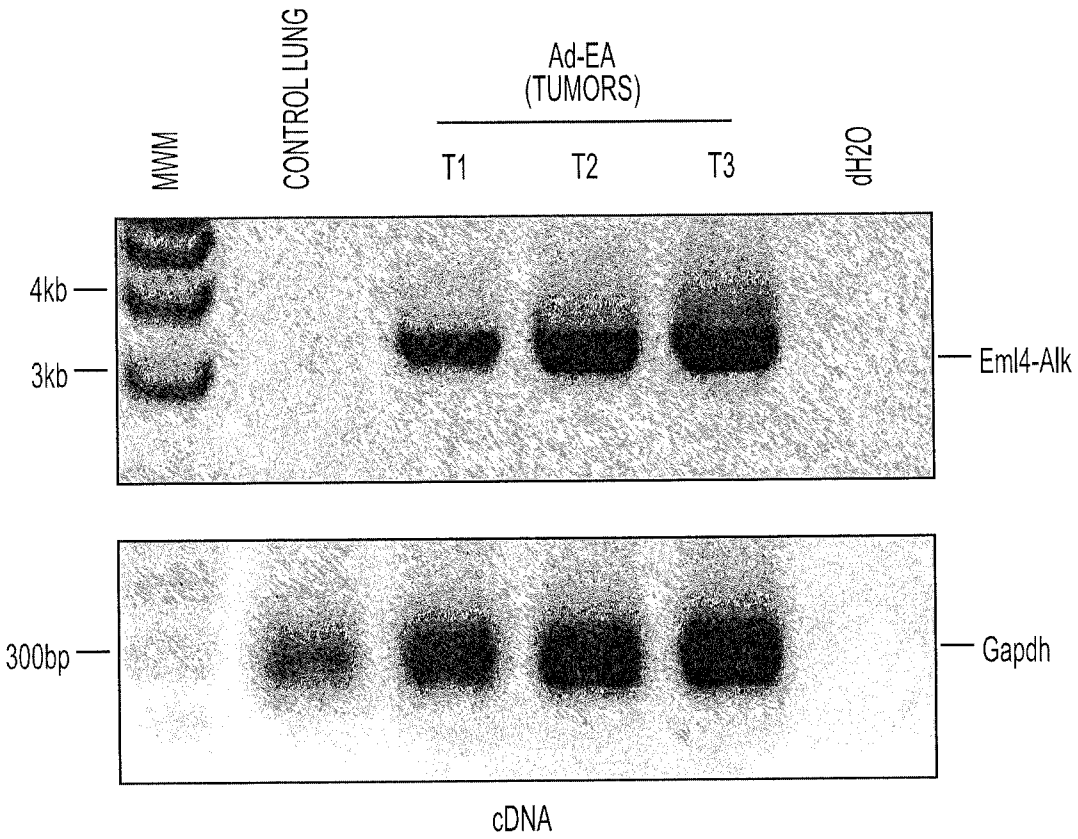

CONTROL LUNG
Ad-EA (TUMORS)
T1
T2
T3
MWM
WT
NIH3T3
Eml4-Alk+
dH2O
Eml4
Eml4-Alk
400bp
300bp
200bp
cDNA

FIG. 6B


MWM
CONTROL LUNG
Ad-EA
(TUMORS)
T1
T2
T3
dH2O
4kb
3kb
Eml4-Alk
300bp
Gapdh
cDNA

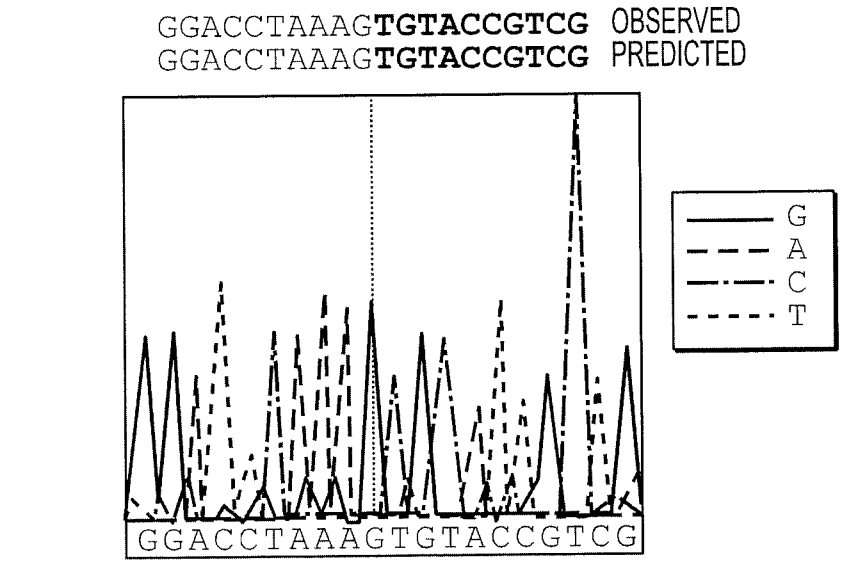

GGACCTAAAGTGTACCGTCG OBSERVED
GGACCTAAAGTGTACCGTCG PREDICTED
G
A
C
T
GGACCTAAAGTGTACCGTCG

FIG. 6C

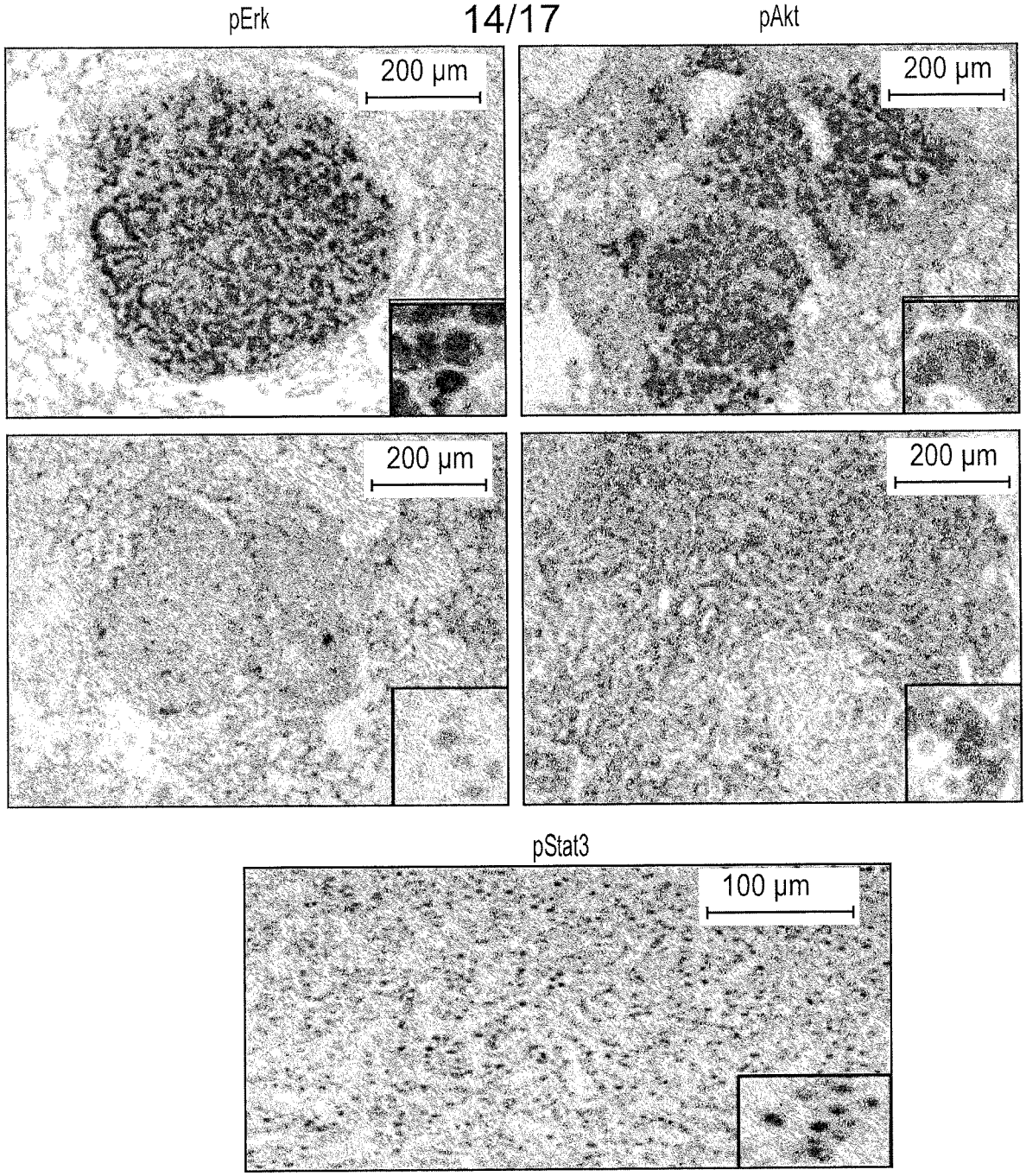

pErk
200 µm
pAkt
200 µm
200 µm
200 µm
pStat3
100 µm
n=32
n=24
n=29
% OF TUMORS
100
90
80
70
60
50
40
30
20
10
0
pErk
pStat3
pAkt
NEGATIVE
PARTIALLY POSITIVE
UNIFORMLY POSITIVE

FIG. 6D

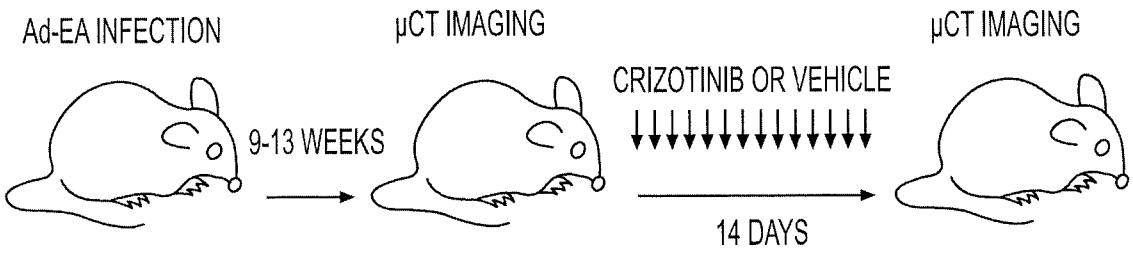

Ad-EA INFECTION
9-13 WEEKS
μCT IMAGING
CRIZOTINIB OR VEHICLE
14 DAYS
μCT IMAGING

FIG. 7A

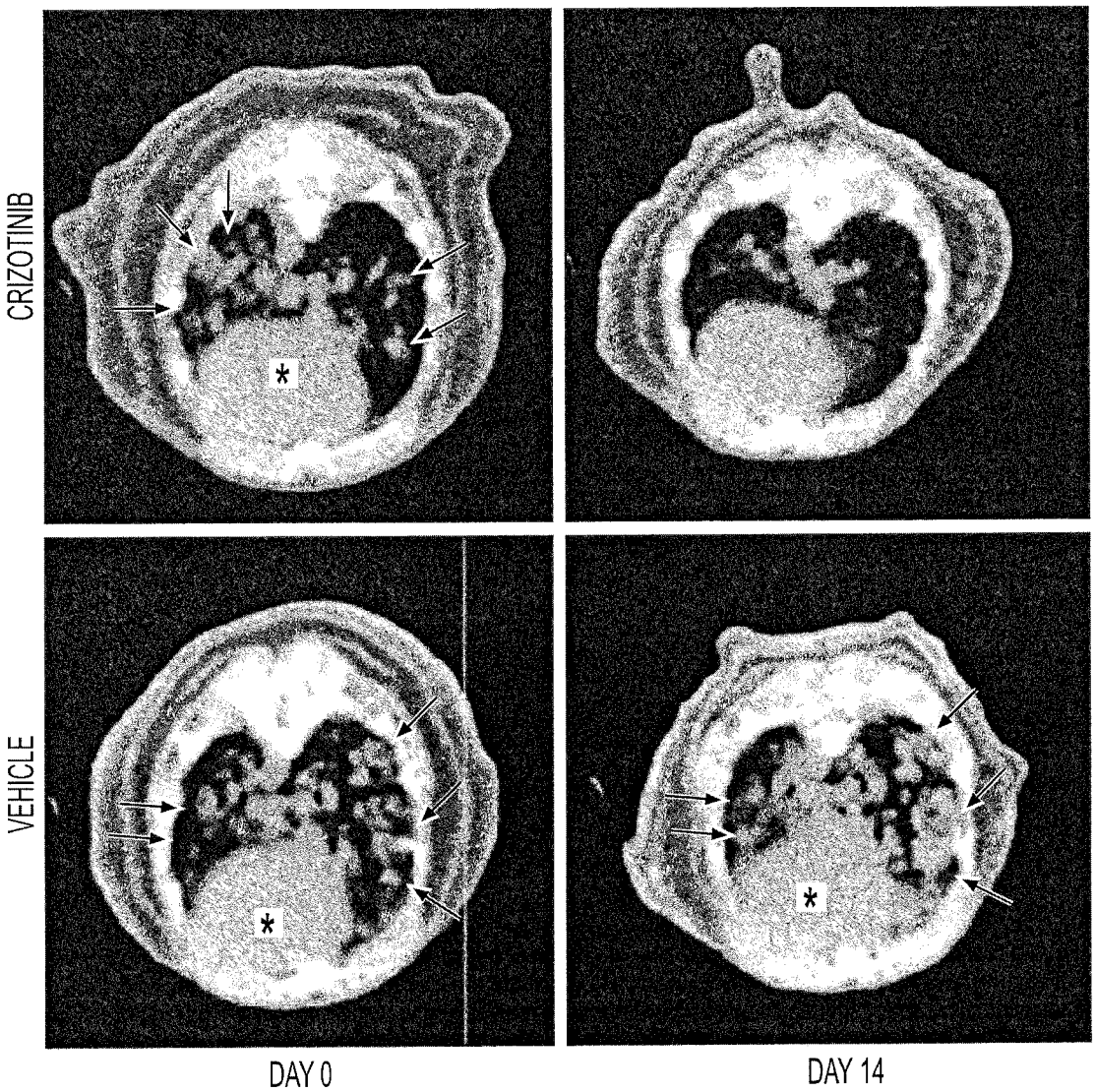

CRIZOTINIB
VEHICLE
DAY 0
DAY 14

FIG. 7B

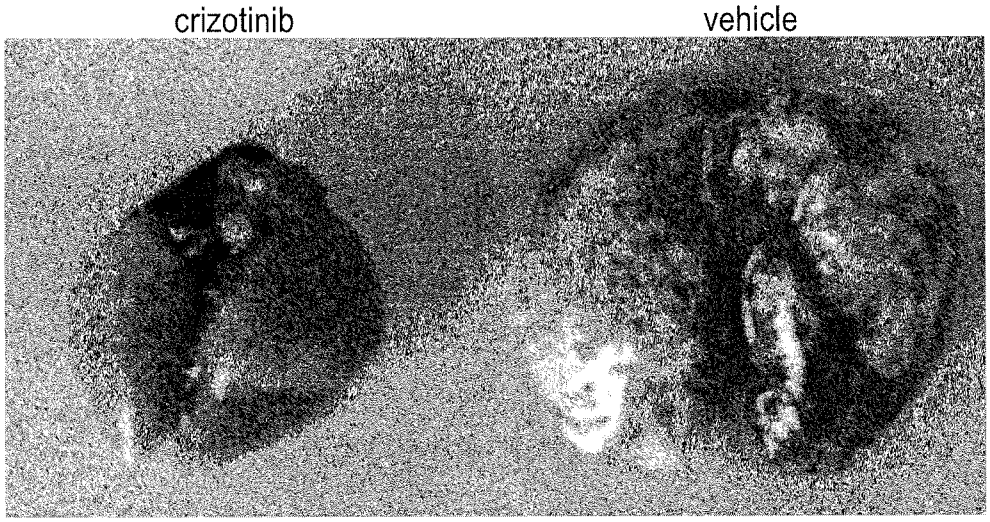

crizotinib
vehicle

FIG. 7C

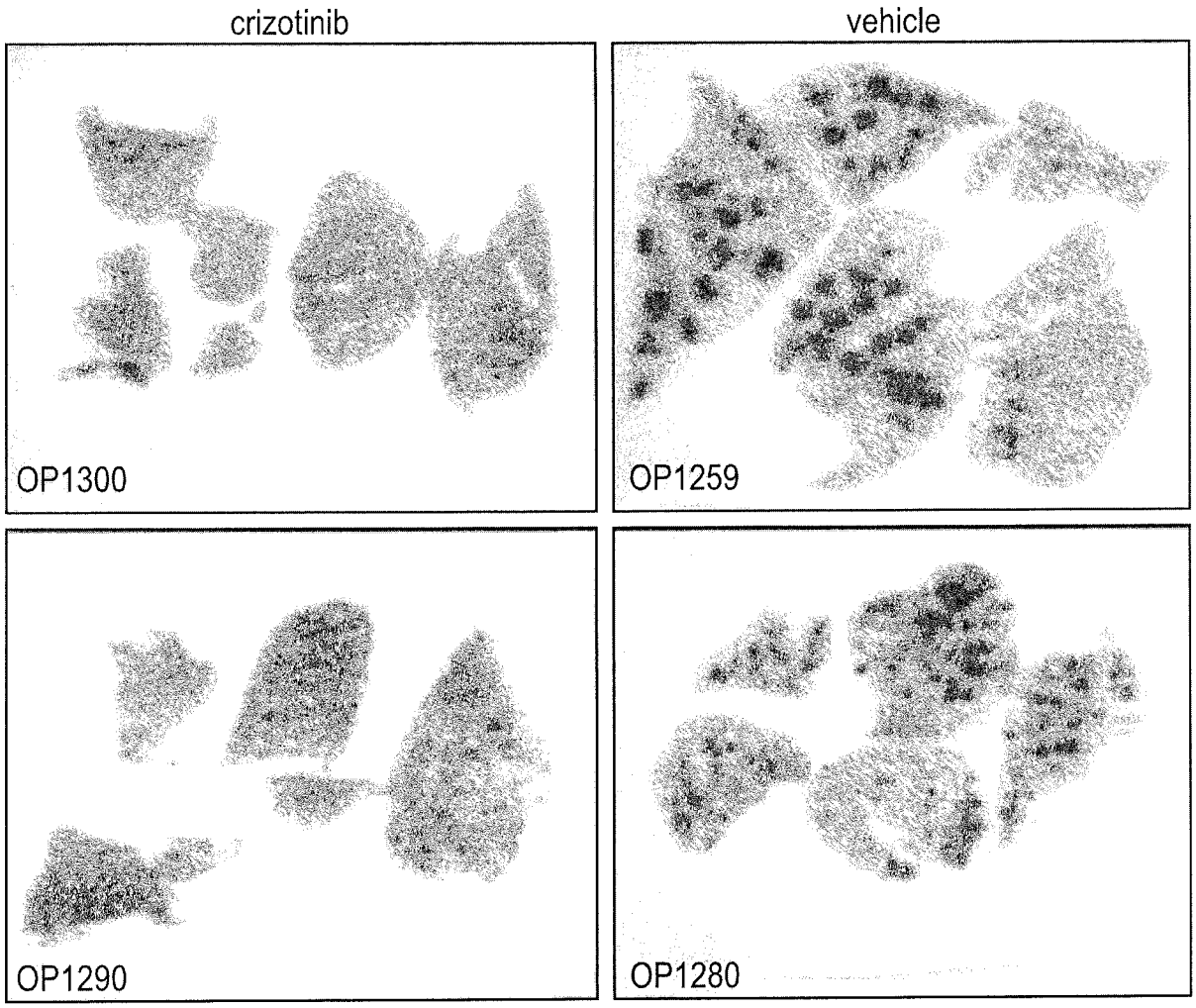

crizotinib
vehicle
OP1300
OP1259
OP1290
OP1280

FIG. 7D

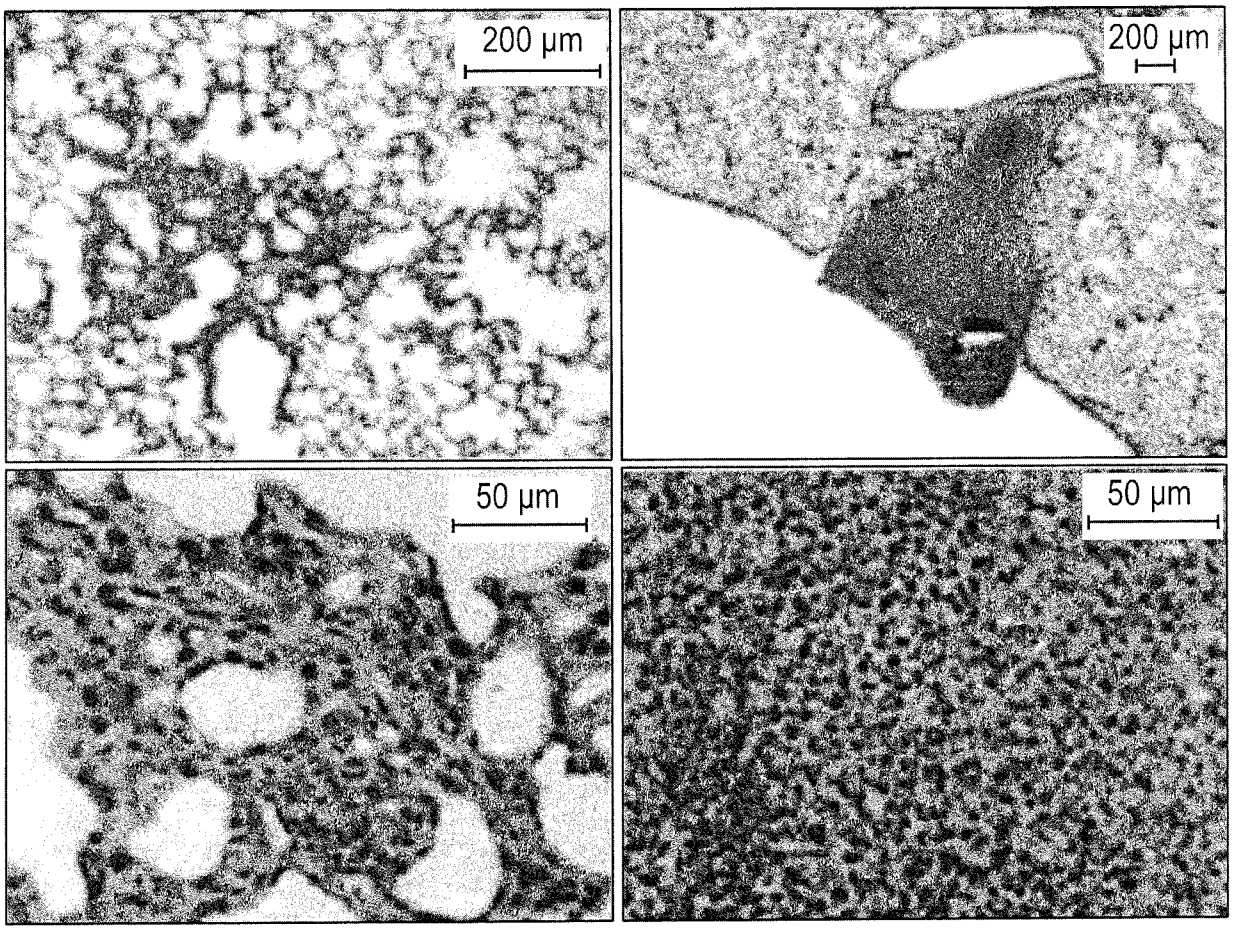


FIG. 7E

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US15/55286

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12N 15/85, 15/861, 15/113, 15/63 (2015.01)
CPC - C12N 15/907, 9/22, 15/63, 15/8509, 15/113, 15/102
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): C12N 15/85, 15/861, 15/113, 15/63 (2015.01)
CPC: C12N 15/907, 9/22, 15/63, 15/8509, 15/113, 15/102

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PatSeer (US, EP, WO, JP, DE, GB, CN, FR, KR, ES, AU, IN, CA, RU, AT, CH, TH, BR, PH, Other Countries (INPADOC)); Google; Google Scholar; PubMed; EBSCO; 'Cas9,' 'CRISPR,' 'sgRNA,' 'EML4-ALK,' translocation, inversion, fusion

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CHOI, PS et al. Targeted Genomic Rearrangements Using CRISPR/Cas Technology. Nat. Commun. 24 April 2014, Vol. 5, Article No. 3728. DOI: 10.1038/ncomms4728; pages 1-9.	1-20
Y	WO 2014/093595 A1 (THE BROAD INSTITUTE, INC.) June 19, 2014; paragraphs [0007], [0008], [0013], [0015], [00127], [00139]-[00143], [00146], [0150]-[00154].	1-20
Y	MARTINELLI, P et al. The Lymphoma-Associated NPM-ALK Oncogene Elicits A p16INK4a/pRb-Dependent Tumor-Suppressive Pathway. Blood. 16 June 2011, Vol. 117, No. 24, pages 6617-6626.	9
Y	SODA, M et al. A Mouse Model for EML4-ALK-Positive Lung Cancer. Proc. Natl. Acad. Sci. USA. 16 December 2008, Vol. 105, No. 50, pages 19893-7. DOI: 10.1073/pnas.0805381105.	10, 11
A	XUE, W et al. CRISPR-Mediated Direct Mutation of Cancer Genes in the Mouse Liver. Nature. 16 October 2014 [Published online 06 August 2014], Vol. 514, page 380-384.	1-20
A	US 2014/0273226 A1 (SYSTEM BIOSCIENCES, LLC) September 18, 2014; entire document	1-20
A	MAGGIO, I et al. Adenoviral Vector Delivery of RNA-guided CRISPR/Cas9 Nuclease Complexes Induces Targeted Mutagenesis In A Diverse Array Of Human Cells. Sci. Rep. 2014, Vol. 5, No. 5105, pages 1-11. DOI:10.1038/srep05105.	1-20

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* 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
05 January 2016 (05.01.2016)	02 FEB 2016
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer Shane Thomas PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)