



- (51) **International Patent Classification:**
C12Q 1/68 (2006.01) **C12N 15/10** (2006.01)
C12N 15/63 (2006.01)
- (21) **International Application Number:**
PCT/US2015/000194
- (22) **International Filing Date:**
23 December 2015 (23.12.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
62/098,285 30 December 2014 (30.12.2014) US
- (71) **Applicants:** **THE BROAD INSTITUTE INC.** [US/US];
415 Main Street, Cambridge, MA 02142 (US). **MAS-**
SACHUSETTS INSTITUTE OF TECHNOLOGY
[US/US]; 77 Massachusetts Ave., Cambridge, MA 02142
(US).
- (72) **Inventors:** **CHEN, Sidi**; C/o Massachusetts Institute Of
Technology, 77 Massachusetts Ave., Cambridge, MA
02142 (US). **PLATT, Randall, Jeffrey**; 72 Sixth Street,
Apt. 1, Cambridge, MA 02141 (US). **SANJANA, Neville,**
Espi; 2a Union Terrace, Cambridge, MA 02142 (US).
SHARP, Phillip, A.; C/o Massachusetts Institute Of Tech-
nology, 77 Massachusetts Ave., Cambridge, MA 02142
(US). **ZHANG, Feng**; 100 Pacific Street, Apt. 11, Cam-
bridge, MA 02139 (US).
- (74) **Agents:** **KOWALSKI, Thomas, J.** et al.; Vedder Price
P.C., 1633 Broadway, New York, NY 10019 (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,

[Continued on next page]

- (54) **Title:** CRISPR MEDIATED IN VIVO MODELING AND GENETIC SCREENING OF TUMOR GROWTH AND META-
STASIS

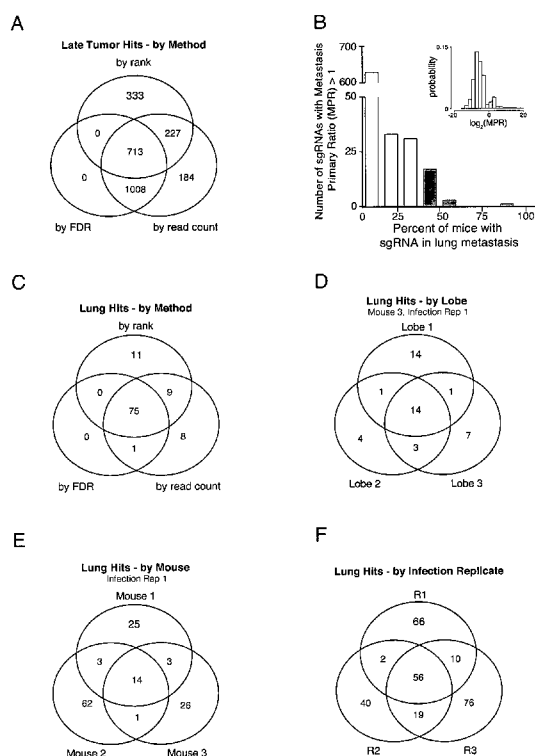


Fig. 11A-11F

(57) **Abstract:** The present invention relates to in vivo methods for modeling tumor formation and/or tumor evolution comprising the use of eukaryotic cells in which one or more genetic target locus has been altered by the CRISPR/Cas system, and which cells are transplanted in non-human eukaryote as a model system for tumor formation and tumor evolution. In particular in vivo genetic screening methods for identifying genes involved in tumorigenesis and metastasis are disclosed. The invention further relates to kits and components for practicing the methods, as well as materials obtainable by the methods, in particular tumor and metastasis samples and cells or cell lines derived therefrom. The invention also relates to diagnostic and therapeutic methods derived from the information obtained in the modeling methods.



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.

SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(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,

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

CRISPR MEDIATED IN VIVO MODELING AND GENETIC SCREENING OF TUMOR GROWTH AND METASTASIS

RELATED APPLICATIONS AND/OR INCORPORATION BY REFERENCE

[0001] This application claims benefit of and priority to US provisional application Serial No. 62/098,285, filed December 30, 2014.

[0002] All documents cited or referenced herein ("herein cited documents"), and all documents cited or referenced in herein cited documents, together with any manufacturer's instructions, descriptions, product specifications, and product sheets for any products mentioned herein or in any document incorporated by reference herein, are hereby incorporated herein by reference, and may be employed in the practice of the invention. More specifically, all referenced documents are incorporated by reference to the same extent as if each individual document was specifically and individually indicated to be incorporated by reference.

STATEMENT OF FEDERALLY SPONSORED RESEARCH

[0003] This invention was made with government support under grant numbers MH100706, CA133404, CA151884 and CA14051 awarded by the National Institutes of Health. The government has certain rights in the invention.

FIELD OF THE INVENTION

[0004] The present invention relates to modeling of tumorigenesis and metastasis. The present invention provides hereto in vivo systems for identifying genes involved in tumorigenesis and metastasis, as well as associated methodology and kits.

BACKGROUND OF THE INVENTION

[0005] Cancer genomes have complex landscape of mutations and diverse types of aberrations (Lawrence et al., 2013; Weinberg, 2007). A major challenge in understanding the cancer genome is to disentangle alterations that are driving the processes of tumor evolution (Garraway and Lander, 2013). Genetic screens are powerful tools for identifying causal genes in various hallmarks of cancer progression (Hanahan and Weinberg, 2011). Loss-of-function genetic screens have been applied in identifying tumor suppressors in liver cancer and skin

cancer models (Schramek et al., 2014; Shao et al., 2014; Zender et al., 2008). Current genetic screens in cancer primarily utilize either RNA interference (RNAi) reagents, which act through knock-down of the mRNAs of the target genes, or overexpression of open reading frames.

[0006] Metastasis is a major lethal factor of cancer (Valastyan and Weinberg, 2011). Clinical observation suggested that the probability of detecting metastases in a patient positively correlates with the size of a primary tumor (Weinberg, 2007). Metastasis of cancer cells involves intravasation, survival and traveling in circulation, extravasation and clonal growth at a distant site (Valastyan and Weinberg, 2011). This is a multi-step process regulated by many genes in complex pathways. Many studies using various approaches had been carried out to search for, and to characterize, genes regulating metastases (Kang et al., 2003; Nguyen et al., 2009; Valiente et al., 2014; Winslow et al., 2011).

[0007] CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) Cas9-mediated gene disruption has been widely used in generating loss-of-function mutations in diverse organisms including mammals (Cong et al., 2013; Mali et al., 2013) (reviewed in (Hsu et al., 2014)). Cas9-based knockout screens have been applied in identifying essential genes and genes involved in drug resistance in various cell lines (Koike-Yusa et al., 2014; Shalem et al., 2014; Wang et al., 2014). However, *in vivo* studies using pooled libraries have been challenging, due to the complexity of the library and the representation requirement in a successful screen, as well as the complex dynamics of cellular behavior in animals.

[0008] Citation or identification of any document in this application is not an admission that such document is available as prior art to the present invention.

SUMMARY OF THE INVENTION

[0009] There exists a pressing need for alternative and robust systems and techniques for experimental *in vivo* modeling of cancer mutations, especially cancer mutations affecting not only tumorigenesis, but also tumor progression and tumor evolution, in particular tumor metastasis.

[0010] Aspects of this invention may address this need and provide related advantages. Aspects of the present invention may address herein discussed challenges and need in the art by advantageously providing methods, systems, compositions and models for *ex vivo* and *in vivo*

modeling of multiple genetic, e.g., cancer or tissue-specific mutations. Aspects of the invention provide methods for enabling rapid and direct *in vivo* and *ex vivo* modeling of the dynamics of multiple genetic, e.g., cancer or tissue-specific mutations. Aspects of the present invention involve sequence targeting, such as genome perturbation or induction of multiple mutations using the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) system or components thereof. The invention provides systematic reverse engineering of causal genetic variations, including through selective perturbation of individual, and moreover, multiple genetic elements. For instance, in non-human eukaryote, e.g., animal, such as fish, e.g., zebra fish, mammal, e.g., primate, e.g., ape, chimpanzee, macaque, rodent, e.g., mouse, rabbit, rat, canine or dog, livestock (cow / bovine, sheep / ovine, goat or pig), fowl or poultry, e.g., chicken, insect, arthropod or plant, e.g., dicot (e.g., nightshade such as tobacco, tuber such as potato) or monocot (e.g., corn) models that constitutively or through induction or through administration or delivery, have cells that contain Cas9. The invention provides tools for studying genetic interaction between multiple individual genetic elements by allowing selective perturbation of e.g., one or more cancer-associated or correlated gene(s)/genetic element(s). In an aspect, the invention provides methods for using one or more elements/components of a CRISPR-Cas system *via* a vector and/or particle and/or nanoparticle delivery formulation or system as a means to modify a target polynucleotide. In preferred embodiments, the delivery is via a viral vector (e.g., AAV, adenovirus, lentivirus). The CRISPR complex of the invention provides an effective means for modifying a target polynucleotide. The CRISPR complex of the invention has a wide variety of utilities including modifying (e.g., deleting, inserting, translocating, inactivating, activating) a target polynucleotide in a multiplicity of cell types in various tissues and organs, e.g., endothelial cells, skin, heart, muscle or lung. As such the CRISPR complex of the invention has a broad spectrum of applications in modeling of multiple genetic, e.g., cancer or tissue-specific mutations, and hence gene therapy, drug discovery, drug screening, disease diagnosis, and prognosis.

[0011] The present invention generally relates to methods for modeling tumor formation and/or tumor evolution. In certain aspects, the invention relates to methods for identifying genes which are involved in tumor formation and tumor evolution. More particularly methods are provided to identify genes whose loss-of function confers a proliferative or metastatic phenotype.

The methods according to the invention foresee hereto the introduction of manipulated eukaryotic cells in a non-human eukaryote. The cells are in particular manipulated by the CRISPR/Cas system. Manipulation of the cells therefore entails the introduction or delivery of the components of the CRISPR/Cas system in the eukaryotic cells. Through the introduction or expression of Cas in combination with specific RNAs which are capable of guiding Cas to a genomic target locus, specific alterations of that target locus are accomplished, such as in particular gene inactivating mutations, but also epigenetic modifications or alterations of gene transcription. Cells so manipulated are subsequently introduced in a non-human eukaryote, after which in vivo tumor formation and/or evolution can be monitored, analyzed, or otherwise modeled.

[0012] Advantageously, the methods according to the present invention not only allow the identification of genes involved in primary tumor formation and evolution, but also the identification of genes involved in secondary tumor formation and evolution, i.e. metastases. In particular, the methods according to the invention allow for the identification of tumor suppressor genes and/or metastasis suppressor genes.

[0013] In a particularly preferred embodiment of the invention, the methods as described herein comprise administering a plurality of eukaryotic cells treated with the CRISPR/Cas system as described above, wherein each of the cells comprises or expresses a different RNA capable of guiding Cas to a genetic target locus of said eukaryote. Advantageously, a library of such RNAs may be used to this effect. The plurality of cells may be transplanted in a non-human eukaryote and allows genetic screening for genes involved in tumor formation and/or evolution. Specifically targeted genes which are involved in tumor formation and/or evolution become enriched over time. Moreover, analysis of the targeted genes in metastases allows identification of genes specifically involved in the development and/or progression of metastasis.

[0014] In related aspects, the invention provides for kits comprising one or more of the components for practicing the methods as described above. Such kits may thus comprise eukaryotic cells; Cas polypeptide, mRNA, or a polynucleotide such as an expression and/or delivery vector encoding Cas; and/or one or more RNA(s) capable of guiding Cas to a genomic target locus in the eukaryotic cell, or expression and/or delivery vector(s) encoding RNA(s)

capable of guiding Cas to (a) genomic target locus/loci in the eukaryotic cell. Such kits may also include a non-human eukaryote into which the cells manipulated with the CRISPR/Cas system are to be introduced.

[0015] It will be appreciated that in the present methods, where the non-human transgenic Cas9 organism is multicellular, e.g., an animal or plant, the modification may occur *ex vivo* or *in vitro*, for instance in a cell culture and in some instances not *in vivo*. In other embodiments, it may occur *in vivo*. In an aspect, the invention provides a method of modifying an organism or a non-human organism by manipulation of a target sequence in a genomic locus of interest comprising: Delivering, e.g., via particle(s) or nanoparticle(s) or vector(s) (e.g., viral vector, e.g., AAV, adenovirus, lentivirus) a non-naturally occurring or engineered composition. The composition can comprise: A) I. RNA(s) having polynucleotide sequence(s), e.g., a CRISPR-Cas system chimeric RNA (chiRNA) having polynucleotide a sequence, wherein the polynucleotide sequence comprises: (a) a guide sequence capable of hybridizing to a target sequence in a eukaryotic cell, (b) a tracr mate sequence, and (c) a tracr sequence; wherein (a), (b) and (c) are arranged in a 5' to 3' orientation. The composition can also comprise A) II. a polynucleotide sequence encoding a CRISPR enzyme advantageously comprising at least one or more or two or more nuclear localization sequences. When transcribed, the tracr mate sequence hybridizes to the tracr sequence and the guide sequence directs sequence-specific binding of a CRISPR complex to the target sequence, wherein the CRISPR complex comprises the CRISPR enzyme complexed with (1) the guide sequence that is hybridizable to the target sequence, and (2) the tracr mate sequence that is hybridizable to the tracr sequence. The polynucleotide sequence encoding a CRISPR enzyme is DNA or RNA. When the Cas9 is already present in the cell, e.g., through the cell having already been provided (A) II. or through the cell expressing Cas9, e.g., through the cell having been transformed to express Cas9, e.g., Cas9 is expressed constitutively or conditionally or inducibly — for instance when the cell is part of or from a non-human transgenic eukaryote, e.g., animal, mammal, primate, rodent, etc as herein discussed — then (A) I. is provided as the CRISPR complex is formed *in situ* or *in vivo*. In an aspect, the invention provides a method of modifying an organism or a non-human organism by manipulation of a target sequence in a genomic locus of interest comprising: Delivering, e.g., via particle(s) or nanoparticle(s) or vector(s) (e.g., viral vector, e.g., AAV, adenovirus, lentivirus) a non-naturally

occurring or engineered composition. The composition can comprise (B) I. polynucleotides comprising: (a) a guide sequence capable of hybridizing to a target sequence in a eukaryotic cell, and (b) at least one or more tracr mate sequences. The composition can also comprise (B) II. a polynucleotide sequence comprising a tracr sequence. The composition can also comprise (B) III. a polynucleotide sequence encoding a CRISPR enzyme advantageously comprising at least one or more or two or more nuclear localization sequences. When transcribed, the tracr mate sequence hybridizes to the tracr sequence and the guide sequence directs sequence-specific binding of a CRISPR complex to the target sequence. The CRISPR complex comprises the CRISPR enzyme complexed with (1) the guide sequence that is hybridizable to the target sequence, and (2) the tracr mate sequence that is hybridizable to the tracr sequence, and the polynucleotide sequence encoding a CRISPR enzyme is DNA or RNA. When the Cas9 is already present in the cell, e.g., through the cell having already been provided (B) III. or through the cell expressing Cas9, e.g., through the cell having been transformed to express Cas9, e.g., Cas9 is expressed constitutively or conditionally or inducibly — for instance when the cell is part of or from a non-human transgenic eukaryote, e.g., animal, mammal, primate, rodent, etc as herein discussed — then (B) I. and (B) II. are provided as the CRISPR complex is formed in situ or in vivo. Accordingly, components I and II or I, II and III or the foregoing embodiments can be delivered separately; for instance, in embodiments involving components I, II and III, components I and II can be delivered together, while component II can be delivered separately, e.g., prior to components I and II, so that the cell or eukaryote expresses Cas9. It will be further appreciated that heretofore it could not be expected that multiple, specific mutations, especially in the numbers herein discussed, e.g., 3-50 or more, or 3, 16, 32, 48 or 50 or more, could be able to be achieved. In undertaking embodiments of the invention, the Applicants have indeed divined that such multiple mutations are present in significant numbers of cells of the non-human eukaryote. For instance, it was surprising and unexpected that cells and tumors of the non-human transgenic organisms of the invention could have multiple mutations of the delivered RNA(s) (sgRNAs); indeed, all of the multiple mutations.

[0016] In some embodiments the invention comprehends delivering a CRISPR enzyme comprising delivering to a cell mRNA encoding the CRISPR enzyme, e.g., via nanoparticle complex(es). In some of these methods the CRISPR enzyme is a Cas9. In certain preferred

embodiments the Cas9 enzyme is constitutively present, e.g., through knock-in. Thus, in a preferred embodiment of the invention, the Cas9 enzyme is constitutively present in vivo (e.g., a non-human transgenic eukaryote, animal, mammal, primate, rodent, etc) or ex vivo (cells comprising a vector containing nucleic acid molecule(s) for in vivo expression of the Cas9). The CRISPR enzyme is a type I or III CRISPR enzyme, preferably a type II CRISPR enzyme. This type II CRISPR enzyme may be any Cas enzyme. A preferred Cas enzyme may be identified as Cas9 as this can refer to the general class of enzymes that share homology to the biggest nuclease with multiple nuclease domains from the type II CRISPR system. Most preferably, the Cas9 enzyme is from, or is derived from, SpCas9 or SaCas9. It will be appreciated that SpCas9 or SaCas9 are those from or derived from *S. pyogenes* or *S. aureus* Cas9. By derived, Applicants mean that the derived enzyme is largely based, in the sense of having a high degree of sequence homology with, a wildtype enzyme, but that it has been mutated (modified) in some way as described herein. It will be appreciated that the terms Cas and CRISPR enzyme are generally used herein interchangeably, unless otherwise apparent. The Cas enzyme can be for instance any naturally-occurring bacterial Cas9 as well as any chimaeras, mutants, homologs or orthologs. Many of the residue numberings used herein refer to the Cas9 enzyme from the type II CRISPR locus in *Streptococcus pyogenes* (annotated alternatively as SpCas9 or spCas9). However, it will be appreciated that this invention includes many more Cas9s from other species of microbes, e.g., orthologs of SpCas9, or Cas9s derived from microbes in addition to *S. pyogenes*, e.g., SaCas9 derived from *S. aureus*, St1Cas9 derived from *S. thermophilus* and so forth. The skilled person will be able to determine appropriate corresponding residues in Cas9 enzymes other than SpCas9 by comparison of the relevant amino acid sequences. Thus, where a specific amino acid replacement is referred to using the SpCas9 numbering, then, unless the context makes it apparent this is not intended to refer to other Cas9 enzymes, the disclosure is intended to encompass corresponding modifications in other Cas9 enzymes. Cas9 orthologs typically share the general organization of 3-4 RuvC domains and a HNH domain. The 5' most RuvC domain cleaves the non-complementary strand, and the HNH domain cleaves the complementary strand. All notations are in reference to the guide sequence. The catalytic residue in the 5' RuvC domain is identified through homology comparison of the Cas9 of interest with other Cas9 orthologs (from *S. pyogenes* type II CRISPR locus, *S. thermophilus* CRISPR locus 1, *S. thermophilus*

CRISPR locus 3, and *Franciscilla novicida* type II CRISPR locus), and the conserved Asp residue (D10) is mutated to alanine to convert Cas9 into a complementary-strand nicking enzyme. Similarly, the conserved His and Asn residues in the HNH domains are mutated to Alanine to convert Cas9 into a non-complementary-strand nicking enzyme. In some embodiments, both sets of mutations may be made, to convert Cas9 into a non-cutting enzyme. Thus, the Cas9 may comprise one or more mutations and may be used as a generic DNA binding protein with or without fusion to a functional domain. The mutations may be artificially introduced mutations or gain- or loss-of-function mutations. The mutations may include but are not limited to mutations in one of the catalytic domains (e.g., D10 and H840) in the RuvC and HNH catalytic domains respectively; or the CRISPR enzyme can comprise one or more mutations selected from the group consisting of D10A, E762A, H840A, N854A, N863A or D986A and/or one or more mutations in a RuvC1 or HNH domain of the CRISPR enzyme or has a mutation as otherwise as discussed herein. In one aspect of the invention, the Cas9 enzyme may be fused to a protein, e.g., a TAG, and/or an inducible/controllable domain such as a chemically inducible/controllable domain. The Cas9 in the invention may be a chimeric Cas9 proteins; e.g., a Cas9 having enhanced function by being a chimera. Chimeric Cas9 proteins may be new Cas9 containing fragments from more than one naturally occurring Cas9. These may comprise fusions of N-terminal fragment(s) of one Cas9 homolog with C-terminal fragment(s) of another Cas9 homolog. The Cas9 can be delivered into the cell in the form of mRNA. The expression of Cas9 can be under the control of an inducible promoter.

[0017] The tracrRNA and direct repeat sequences can be mutant sequences or the invention can encompass RNA of the CRISPR-Cas system that includes mutant chimeric guide sequences that allow for enhancing performance of these RNAs in cells. A suitable promoter, such as the Pol III promoter, such as a U6 promoter, can be added onto the guide RNA that is advantageously delivered via AAV or particle or nanoparticle. Aspects of the invention also relate to the guide RNA being transcribed in vitro or ordered from a synthesis company and directly transfected. Expression of RNA(s), e.g., guide RNAs or sgRNA under the control of the T7 promoter driven by the expression of T7 polymerase in the cell is also envisioned. In an advantageous embodiment, the cell is a eukaryotic cell. In a preferred embodiment the eukaryotic cell is a human cell. In a more preferred embodiment the cell is a patient-specific cell,

e.g., a cell in which 3-50 or more mutations associated or correlated with a patient's genetic disease, e.g., cancer, are expressed in the cell, e.g., via Cas9 being present in the cell and RNA(s) for such mutations delivered to the cell (e.g., any whole number between 3 and 50 of mutations, with it noted that in some embodiments there can be up to 16 different RNA(s), e.g., sgRNAs each having its own a promoter, in a vector, such as AAV, and that when each sgRNA does not have its own promoter, there can be twice to thrice that amount of different RNA(s), e.g., sgRNAs, e.g., 32 or even 48 different guides delivered by one vector), whereby the CRISPR-Cas complexes result in the cells having the mutations and the cells and the eukaryote, e.g., animal, containing the cells being a model for the patient's genetic disease. In an advantageous embodiment, AAV may package U6 tandem sgRNA targeting up to about 50 genes.

[0018] A codon optimized sequence can be a sequence optimized for a eukaryote, or for specific organs such as the lung. It will be appreciated that where reference is made to a polynucleotide, where that polynucleotide is RNA and is said to 'comprise' a feature such as a tracr mate sequence, the RNA sequence includes the feature. Where the polynucleotide is DNA and is said to comprise a feature such as a tracr mate sequence, the DNA sequence is or can be transcribed into the RNA that comprises the feature at issue. Where the feature is a protein, such as the CRISPR enzyme, the DNA or RNA sequence referred to is, or can be, translated (and in the case of DNA transcribed first). Furthermore, in cases where an RNA encoding the CRISPR enzyme is provided to a cell, it is understood that the RNA is capable of being translated by the cell into which it is delivered. By manipulation of a target sequence, Applicants mean the alteration of the target sequence, which may include the epigenetic manipulation of a target sequence. This epigenetic manipulation may be of the chromatin state of a target sequence, such as by modification of the methylation state of the target sequence (i.e. addition or removal of methylation or methylation patterns or CpG islands), histone modification, increasing or reducing accessibility to the target sequence, or by promoting 3D folding. It will be appreciated that where reference is made to a method of modifying an organism or mammal including human or a non-human mammal or organism by manipulation of a target sequence in a genomic locus of interest, this may apply to the organism (or mammal) as a whole or just a single cell or population of cells from that organism (if the organism is multicellular). Applicants envisage, inter alia, a single cell or a population of cells and these may preferably be modified ex vivo and

then re-introduced, e.g., transplanted to make transgenic organisms that express Cas9 in certain cells. The invention in some embodiments comprehends a method of modifying a eukaryote, such as a Cas9 transgenic eukaryote comprising delivering, e.g., via vector(s) and/or particle(s) and/or nanoparticles a non-naturally occurring or engineered composition. The composition comprises: I. a first regulatory element operably linked to (a) a first guide sequence capable of hybridizing to a first target sequence, and (b) at least one or more tracr mate sequences, II. a second regulatory element operably linked to (a) a second guide sequence capable of hybridizing to a second target sequence, and (b) at least one or more tracr mate sequences, III. a third regulatory element operably linked to (a) a third guide sequence capable of hybridizing to a third target sequence, and (b) at least one or more tracr mate sequences, and IV. a fourth regulatory element operably linked to a tracr sequence. There can be additional regulatory element(s) operably linked to additional guide sequence(s). Optionally, the composition can involve V. a fifth regulatory element operably linked to an enzyme-coding sequence encoding a CRISPR enzyme (e.g., for establishing the Cas9 transgenic eukaryote). Components I, II, III and IV (as well as any other regulatory element(s) linked to additional guide sequence(s)) are located on the same or different vectors and/or particles and/or nanoparticles of the system. When transcribed, the tracr mate sequence hybridizes to the tracr sequence and the first, second and the third guide sequences direct sequence-specific binding of a first, second and a third CRISPR complexes to the first, second and third target sequences respectively, wherein the first CRISPR complex comprises the CRISPR enzyme complexed with (1) the first guide sequence that is hybridizable to the first target sequence, and (2) the tracr mate sequence that is hybridizable to the tracr sequence, wherein the second CRISPR complex comprises the CRISPR enzyme complexed with (1) the second guide sequence that is hybridizable to the second target sequence, and (2) the tracr mate sequence that is hybridizable to the tracr sequence, and wherein the third CRISPR complex comprises the CRISPR enzyme complexed with (1) the third guide sequence that is hybridizable to the third target sequence, and (2) the tracr mate sequence that is hybridizable to the tracr sequence, whereby in a Cas9 transgenic eukaryote or cell thereof, at least three (3) mutations may be induced, and advantageously the mutations are correlated or associated with a genetic disease condition, whereby the eukaryote or cell becomes a model of the disease, e.g., cancer. The invention also provides a vector system as described herein. The system may comprise one,

two, three or four different vectors; and the system may comprise one, two, three or four different nanoparticle complex(es) delivering the component(s) of the system. Components I, II, III and IV may thus be located on one, two, three or four different vectors, and may be delivered by one, two, three or four different particle or nanoparticle complex(es) or AAVs or components I, II, III and IV can be located on same or different vector(s) / particle(s) / nanoparticle(s), with all combinations of locations envisaged. And complexes that target lung or lung tissue or cells are advantageous.

[0019] In some methods of the invention any or all of the polynucleotide sequence encoding the CRISPR enzyme, the first and the second guide sequence, the first and the second tracr mate sequence or the first and the second tracr sequence, is/are RNA; and advantageously delivered via nanoparticle complex(es). Embodiments of the invention also comprehend the guide RNAs comprising a guide sequence fused to a tracr mate sequence and a tracr sequence. In an aspect of the invention the Cas protein is codon optimized for expression in a eukaryotic cell, preferably a mammalian cell or a human cell. The invention also comprehends an engineered, non-naturally occurring vector system. The system comprises one or more vectors comprising: (a) a first regulatory element operably linked to each of two or more e.g., three, CRISPR-Cas system guide RNAs that target a first target sequence, a second target sequence and a third target sequence respectively of a double stranded DNA molecule, wherein either strand of the double stranded DNA molecule may be targeted by each CRISPR-Cas system guide RNA. The system can also comprise (b) a second regulatory element operably linked to a Cas protein. Components (a) and (b) are located on same or different vectors of the system, but advantageously separate vectors as it is preferred that cells receiving (a) contain Cas9, e.g., via the cells being those of a transgenic Cas9 eukaryote (whereby (b) may have been administered to cells that gave rise to the eukaryote). The guide RNAs target DNA and at least three mutations are induced in the cells, e.g., mutations correlated to or associated with a genetic disorder such as cancer.

[0020] In one aspect, the invention provides a method of modifying a target polynucleotide in a eukaryotic cell. In some embodiments, the method comprises allowing a CRISPR complex to bind to the target polynucleotide to effect cleavage of said target polynucleotide thereby modifying the target polynucleotide, wherein the CRISPR complex comprises a CRISPR enzyme

complexed with a guide sequence hybridized to a target sequence within said target polynucleotide, wherein said guide sequence is linked to a tracr mate sequence which in turn hybridizes to a tracr sequence; and advantageously the complex or a component thereof has been delivered via nanoparticle complex(es). In some embodiments, said cleavage comprises cleaving one or two strands at the location of the target sequence by said CRISPR enzyme. In some embodiments, said cleavage results in decreased transcription of a target gene. In some embodiments, the method further comprises repairing said cleaved target polynucleotide by homologous recombination with an exogenous template polynucleotide, wherein said repair results in a mutation comprising an insertion, deletion, or substitution of one or more nucleotides of said target polynucleotide. The mutation can be a mutation correlated to or associated with a genetic disease condition, such as cancer. In some embodiments, said mutation results in one or more amino acid changes in a protein expressed from a gene comprising the target sequence. In some embodiments, vectors are delivered to the eukaryotic cell in a transgenic Cas9 eukaryote. In some embodiments, said modifying takes place in said eukaryotic cell in a cell culture. In one aspect, the invention provides a method of generating a model eukaryotic cell or a model Cas9 transgenic eukaryote comprising mutated disease gene(s), e.g., having 3-50 mutations correlated to or associated with a genetic disease such as cancer (e.g., any whole number between 3 and 50 of mutations, with it noted that in some embodiments there can be up to 16 different RNA(s), e.g., sgRNAs each having its own a promoter, in a vector, such as AAV, and that when each sgRNA does not have its own promoter, there can be twice to thrice that amount of different RNA(s), e.g., sgRNAs, e.g., 32 or even 48 different guides delivered by one vector). In some embodiments, a disease gene is any gene associated with an increase in the risk of having or developing a disease.

[0021] In some methods, a target polynucleotide can be inactivated to effect the modification of the expression in a cell. For example, upon the binding of a CRISPR complex to a target sequence in a cell, the target polynucleotide is inactivated such that the sequence is not transcribed, the coded protein is not produced, or the sequence does not function as the wild-type sequence does. For example, a protein or microRNA coding sequence may be inactivated such that the protein or microRNA or pre-microRNA transcript is not produced. In certain embodiments, the target sequence is flanked or followed, at its 3' end, by a PAM suitable for the

CRISPR enzyme, typically a Cas and in particular a Cas9. For example, a suitable PAM is 5'-NRG or 5'-NNGRR for SpCas9 or SaCas9 enzymes (or derived enzymes), respectively.

[0022] Delivery can be in the form of a vector which may be a plasmid or other nucleic acid molecule form, especially when the delivery is via a nanoparticle complex; and the vector also can be viral vector, such as a herpes, e.g., herpes simplex virus, lenti- or baculo- or adeno-viral or adeno-associated viral vectors, but other means of delivery are known (such as yeast systems, microvesicles, gene guns/means of attaching vectors to gold nanoparticles) and are provided, especially as to those aspects of the complex not delivered via a nanoparticle complex. A vector may mean not only a viral or yeast system (for instance, where the nucleic acids of interest may be operably linked to and under the control of (in terms of expression, such as to ultimately provide a processed RNA) a promoter), but also direct delivery of nucleic acids into a host cell; and advantageously the complex or a component thereof is delivered via nanoparticle complex(es). Also envisaged is a method of delivering the present CRISPR enzyme comprising delivering to a cell mRNA encoding the CRISPR enzyme; and advantageously the complex or a component thereof has been delivered via nanoparticle complex(es). It will be appreciated that in certain embodiments the CRISPR enzyme is truncated, and/or comprised of less than one thousand amino acids or less than four thousand amino acids, and/or is a nuclease or nickase, and/or is codon-optimized, and/or comprises one or more mutations, and/or comprises a chimeric CRISPR enzyme, and/or the other options as herein discussed. When not delivering via a nanoparticle complex, AAV is a preferred vector. In certain embodiments, multiple RNA(s) or guide RNAs or sgRNAs formulated in one or more delivery vehicles (e.g., where some guide RNAs are provided in a vector and others are formulated in nanoparticles); and these may be provided alone (e.g., when Cas9 is already in a cell) or with a Cas9 delivery system. In certain embodiments, the Cas9 is also delivered in a nanoparticle formulation. In certain instances the RNA(s) or guide RNA or sgRNA-vector and/or particle and/or nanoparticle formulation(s) and the Cas9 vector and/or particle and/or nanoparticle formulation(s) may be delivered separately or may be delivered substantially contemporaneously (i.e., co-delivery). Sequential delivery could be done at separate points in time, separated by days, weeks or even months. And as Cas9 is advantageously present in a transgenic organism in the practice of the invention, e.g., through being constitutively or conditionally or inducibly present, sequential delivery can include

initially administering or delivering the Cas9 vector and/or particle and/or nanoparticle formulation(s) to cells that give rise to the non-human Cas9 transgenic eukaryote, and thereafter, at a suitable time in the life of the transgenic eukaryote, administering the the RNA(s) or guide RNA or sgRNA-vector and/or particle and/or nanoparticle formulation(s), e.g., so as to give rise to one or more, advantageously 3-50 mutations in the transgenic eukaryote (e.g., any whole number between 3 and 50 of mutations, with it noted that in some embodiments there can be up to 16 different RNA(s), e.g., sgRNAs each having its own a promoter, in a vector, such as AAV, and that when each sgRNA does not have its own promoter, there can be twice to thrice that amount of different RNA(s), e.g., sgRNAs, e.g., 32 or even 48 different guides delivered by one vector), and advantageously the mutations are associated or correlated with a genetic disease, whereby the transgenic eukaryote is a model of the genetic disease, e.g., cancer. Multiple mutations may thus be introduced using any number of sgRNAs, e.g. the vector may comprise at least 3 sgRNAs, at least 8 sgRNAs, at least 16 sgRNAs, at least 32 sgRNAs, at least 48 sgRNAs, or at least 50 sgRNAs. Alternatively, the vector may comprise 1-2 sgRNAs, 1-3 sgRNAs, 1-4 sgRNAs, 1-5 sgRNAs, 3-6 sgRNAs, 3-7 sgRNAs, 3-8 sgRNAs, 3-9 sgRNAs, 3-10 sgRNAs, 3-16 sgRNAs, 3-30 sgRNAs, 3-32 sgRNAs, 3-48 sgRNAs or 3-50 sgRNAs. In certain embodiments, vector (e.g., AAV, adenovirus, lentivirus) and/or particle and/or nanoparticle formulations comprising one or more RNA(s) e.g. guide RNAs or sgRNA are adapted for delivery in vitro, ex vivo or in vivo in the context of the CRISPR-Cas system, e.g., so as to form CRISPR-Cas complexes in vitro, ex vivo or in vivo, to different target genes, different target cells or different target different tissues/organs, with different target genes and/or cells of the lung. Multiplexed gene targeting using nanoparticle formulations comprising one or more guide RNAs are also envisioned. In an embodiment, a nanoparticle formulation comprising one or more components of the CRISPR-Cas system is provided. In an embodiment, a RNA(s) or gRNA or sgRNA-nanoparticle formulation comprising one or more guide RNAs or sgRNA is provided. In certain embodiments, a composition comprising a nanoparticle formulation comprising one or more components of the CRISPR-Cas system is provided. In certain embodiments, a composition, e.g., a pharmaceutical or veterinary composition, comprising a vector (e.g., AAV, adenovirus, lentivirus) and/or particle and/or nanoparticle formulation comprising one or more components of the CRISPR-Cas system and/or nucleic acid molecule(s)

coding therefor, advantageously with such nucleic acid molecule(s) operably linked to promoter(s) is provided. Accordingly, in certain embodiments, it may be useful to deliver the RNA(s) or guide RNA or sgRNA, e.g., vector and/or particle and/or nanoparticle formulations separately from the Cas9 or nucleic acid molecule(s) coding therefor. A dual-delivery system is envisaged such that the Cas 9 may be delivered via a vector and the RNA(s), e.g., guide RNAs or sgRNA are / is provided in a particle or nanoparticle formulation, for example, first Cas9 vector is delivered via a vector system followed by delivery of sgRNA-nanoparticle formulation. Vectors may be considered in the broadest light as simply any means of delivery, rather than specifically viral vectors.

[0023] In one aspect, the present invention provides a Cas9 transgenic eukaryote, e.g., mouse. In certain preferred embodiments, the Cas 9 transgenic eukaryote, e.g., mouse comprises a Cas9 transgene knocked into the Rosa26 locus. In one aspect, the present invention provides a Cas9 transgenic eukaryote, e.g., mouse wherein Cas9 transgene is driven by the ubiquitous CAG promoter thereby providing for constitutive expression of Cas9 in all tissues/cells/cell types of the mouse. In one aspect, the present invention provides a Cas9 transgenic eukaryote, e.g., mouse wherein the Cas9 transgene driven by the ubiquitous CAG promoter further comprises a Lox-Stop-polyA-Lox (LSL) cassette (Rosa26-LSL-Cas9 mouse) thereby rendering Cas9 expression inducible by the Cre recombinase. In one aspect, the present invention provides a constitutive Cas9 expressing eukaryote, e.g., mouse line obtained by crossing of the Rosa26-LSL-Cas9 mouse with a beta-actin-Cre eukaryote, e.g., mouse line. In certain embodiments, progeny (or progenies) derived from said Cas9 expressing eukaryote, e.g., mouse line, may be successfully bred over at least five generations without exhibiting increased levels of genome instability or cellular toxicity. In one aspect, the present invention provides a modular viral vector construct comprising a plurality of sgRNAs driven by a single RNA polymerase III promoter (e.g., U6), wherein the sgRNAs are in tandem, or where each of the sgRNAs is driven by one RNA polymerase III promoter. In one aspect, the present invention provides a modular viral vector construct comprising one or more cassettes expressing Cre recombinase, a plurality of sgRNAs to guide Cas9 cutting, and optionally including a Homology Directed Repair (HDR) template to model the dynamics of a complex pathological disease or disorder involving two or more genetic elements simultaneously using a single vector construct (in embodiments where the HDR

template is not involved, the cutting creates the mutation to model the dynamics of a complex pathological disease or disorder (e.g., loss of function or gain of function). In certain non-limiting embodiments, the complex pathological disease or disorder is cancer. It can be appreciated that any kind of cancer of any tissue type are within the scope of the present invention. In a preferred embodiment, the present invention provides for modeling of lung cancer. In one aspect, the present invention provides a modular viral vector construct to model the dynamics of multiple cancer lesions simultaneously using a single vector. In certain embodiments, the modular viral vector construct comprises one or more cassettes expressing Cre recombinase, a plurality of sgRNAs to guide Cas9 cutting, and optionally one or more Homology Directed Repair (HDR) template(s) to introduce specific gain-of-function mutations or precise sequence substitution in target loci. In one aspect, the present invention provides a method for simultaneously introducing multiple mutations ex vivo in a tissue, organ or a cell line, or in vivo in the same animal comprising delivering a single viral vector construct, wherein the viral vector construct comprises one or more cassettes expressing Cre recombinase, a plurality of sgRNAs to guide Cas9 cutting, and optionally HDR template(s) for achieving targeted insertion or precise sequence substitution at specific target loci of interest. In one aspect, the present invention provides a method for generating loss-of-function mutations in two or more tumor suppressor genes and gain-of-function mutations in one or more proto-oncogenes using a modular viral vector construct which comprises one or more cassettes expressing Cre recombinase, a plurality of sgRNAs to guide Cas9 cutting, and optionally one or more HDR template(s). In certain non-limiting embodiments, tumor suppressor genes may include p53 and Lkb1 (serine/threonine kinase 11). In certain non-limiting embodiments, the proto-oncogene may include Kras. Heretofore, it had not been expected that multiple mutations wherein some mutations are loss of function (knock out) and some mutations are gain of function (knock in) could be achieved; and it had not been expected to achieve mutations in p53, Lkb1 and Kras and nor had there heretofore been any direction to select these particular genes for being mutated together. It can be readily appreciated that mutations in any cancer-associated gene is within the scope of the present invention. In one aspect, the present invention provides a method for delivering ex vivo or in vivo of any of the modular viral constructs disclosed herein using an AAV. Selection of the AAV serotype is based on its suitability and specificity for a tissue type. In certain preferred

embodiments, AAV9 is used for delivery to lung tissue. In one aspect, the present invention provides a method for ex vivo and/or in vivo genome editing comprising delivering any of the above modular viral vector constructs, which comprise one or more cassettes expressing Cre recombinase, a plurality of sgRNAs to guide Cas9 cutting, and optionally one or more HDR template(s), into a Cas9 transgenic mouse (e.g., Rosa26-LSL-Cas9). In certain embodiments, the viral vector is AAV9. In one aspect, the present invention provides a Cas9 transgenic non-human eukaryote, e.g., animal model for lung cancer said model having loss-of-function mutations in p53 and/or Lkb1 and gain-of-function mutation in Kras. It can be appreciated that using the novel CRISPR-Cas9 tools disclosed herein, Cas9 transgenic non-human eukaryote, e.g., animal model with multiple mutations in any number of loci can be envisioned and are within the scope of the present invention. It will be appreciated that such a transgenic non-human eukaryote, e.g., animal model provides a valuable tool for research purposes, e.g., to delineate specific roles/contribution of individual mutations to cancer progression, to recapitulate specific combinations of mutations in a given cancer type, and opens the door for developing and testing new therapeutic interventions for cancers involving mutations at multiple loci. Such uses are within the scope of the present invention. In one aspect, the present invention provides a method of treating or inhibiting the development of a genetic disease in a subject in need thereof, comprising providing individualized or personalized treatment (or an individualized or personalized model or patient specific-modeling) comprising: delivering RNA(s), e.g., sgRNA, that targets a genetic locus correlated or associated with the genetic disease, e.g., cancer, to a Cas9 non-human transgenic eukaryote (e.g., animal, mammal, primate, rodent, fish etc as herein discussed), e.g., via vector such as AAV, adenovirus, lentivirus, or particle(s) or nanoparticle(s), whereby mutation(s), advantageously a plurality, e.g., 3-50 mutations (e.g., any whole number between 3 and 50 of mutations, with it noted that in some embodiments there can be up to 16 different RNA(s), e.g., sgRNAs each having its own a promoter, in a vector, such as AAV, and that when each sgRNA does not have its own promoter, there can be twice to thrice that amount of different RNA(s), e.g., sgRNAs, e.g., 32 or even 48 different guides delivered by one vector), are induced in the eukaryote and the eukaryote is a model for the disease; and obtaining and/or extrapolating data from the Cas9 non-human transgenic eukaryote to humans to provide individualized or personalized treatment. The obtaining and/or extrapolating data can be

subjecting the eukaryote to putative treatment(s) and/or therapy(ies), e.g., gene therapy, ascertaining whether such putative treatment(s) and/or therapy(ies) give rise to remission or treatment or alleviation or mitigation or stasis of the disease, and if so, then administering in dosing scaled to a 70 kg individual or subject, the putative treatment(s) and/or therapy(ies). The invention thus allows for one to ascertain whether a particular treatment and/or therapy may be effective as to a particular individual's disease.

[0024] In certain aspects the invention provides vector(s), particle(s) or nanoparticle(s) containing nucleic acid molecule(s), whereby in vivo in a eukaryotic cell containing or conditionally or inducibly expressing Cas9: the vector(s) express(es) a plurality of RNAs to guide the Cas9 and optionally delivers donor templates (e.g., HDR templates, and in certain embodiments advantageously includes and delivers such donor templates), and optionally in the event Cas9 is conditionally or inducibly expressed in the cell that which induces Cas9, e.g., Cre recombinase; whereby a plurality of specific mutations or precise sequence substitutions in a plurality of target loci are introduced. The vector(s) can be a viral vector such as lentivirus, adenovirus, or adeno-associated virus (AAV), e.g., AAV6 or AAV9. The Cas9 can be from *S. thermophiles*, *S. aureus*, or *S. pyogenes*. The eukaryotic cell can comprise a Cas9 transgene is functionally linked to a constitutive promoter, or a tissue specific promoter, or an inducible promoter; and, the eukaryotic cell can be part of a non-human transgenic eukaryote, e.g., a non-human mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect or arthropod; advantageously a mouse. The isolated eukaryotic cell or the non-human transgenic eukaryote can express an additional protein or enzyme, such as Cre; and, the expression of Cre can be driven by coding therefor functionally or operatively linked to a constitutive promoter, or a tissue specific promoter, or an inducible promoter.

[0025] The RNAs to guide Cas9 can comprise CRISPR RNA and transactivating (tracr) RNA. The tracr mate and the tracr sequence can be connected to form a transactivating (tracer) sequence. The tracr mate and the tracr sequence can optionally be designed to form a single guide RNA (sgRNA). Indeed, it is advantageous that the RNAs to guide Cas9 can comprise chimeric single guide RNA (sgRNA). The tracr sequence and tracr mate sequence along the

length of the shorter of the two when optimally aligned can be about or more than about 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97.5%, 99%, or higher. The tracr sequence can be about or more than about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, or more nucleotides in length. The degree of complementarity between a guide sequence and its corresponding target sequence can be about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or 100%. A guide or RNA or sgRNA can be 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. A guide or RNA or sgRNA can be less than about 75, 50, 45, 40, 35, 30, 25, 20, 15, 12, or fewer nucleotides in length.

[0026] The vector(s) can include the regulatory element(s), e.g., promoter(s). The vector(s) can comprise at least 3 or 8 or 16 or 32 or 48 or 50 RNA(s) (e.g., sgRNAs), such as 1-2, 1-3, 1-4 1-5, 3-6, 3-7, 3-8, 3-9, 3-10, 3-8, 3-16, 3-30, 3-32, 3-48, 3-50 RNA(s) (e.g., sgRNAs). In a single vector there can be a promoter for each RNA (e.g., sgRNA), advantageously when there are up to about 16 RNA(s) (e.g., sgRNAs); and, when a single vector provides for more than 16 RNA(s) (e.g., sgRNAs), one or more promoter(s) can drive expression of more than one of the RNA(s) (e.g., sgRNAs), e.g., when there are 32 RNA(s) (e.g., sgRNAs), each promoter can drive expression of two RNA(s) (e.g., sgRNAs), and when there are 48 RNA(s) (e.g., sgRNAs), each promoter can drive expression of three RNA(s) (e.g., sgRNAs). By simple arithmetic and well established cloning protocols and the teachings in this disclosure one skilled in the art can readily practice the invention as to the RNA(s) (e.g., sgRNA(s) for a suitable exemplary vector such as AAV, and a suitable promoter such as the U6 promoter, e.g., U6-sgRNAs. For example, the packaging limit of AAV is ~4.7 kb. The length of a single U6-sgRNA (plus restriction sites for cloning) is 361 bp. Therefore, the skilled person can readily fit about 12-16, e.g., 13 U6-sgRNA cassettes in a single vector. This can be assembled by any suitable means, such as a golden gate strategy used for TALE assembly (<http://www.genome-engineering.org/talectors/>). The skilled person can also use a tandem guide strategy to increase the number of U6-sgRNAs by approximately 1.5 times, e.g., to increase from 12-16, e.g., 13 to approximately 18-24, e.g., about 19 U6-sgRNAs. Therefore, one skilled in the art can readily reach approximately 18-24, e.g., about 19 promoter-RNAs, e.g., U6-sgRNAs in a single vector, e.g., an AAV vector. A further means for increasing the number of promoters and RNAs, e.g., sgRNA(s) in a vector is to use a

single promoter (e.g., U6) to express an array of RNAs, e.g., sgRNAs separated by cleavable sequences. And an even further means for increasing the number of promoter-RNAs, e.g., sgRNAs in a vector, is to express an array of promoter-RNAs, e.g., sgRNAs separated by cleavable sequences in the intron of a coding sequence or gene; and, in this instance it is advantageous to use a polymerase II promoter, which can have increased expression and enable the transcription of long RNA in a tissue specific manner. (see, e.g., <http://nar.oxfordjournals.org/content/34/7/e53.short>, <http://www.nature.com/mt/journal/v16/n9/abs/mt2008144a.html>). In an advantageous embodiment, AAV may package U6 tandem sgRNA targeting up to about 50 genes. Accordingly, from the knowledge in the art and the teachings in this disclosure the skilled person can readily make and use vector(s), e.g., a single vector, expressing multiple RNAs or guides or sgRNAs under the control or operatively or functionally linked to one or more promoters—especially as to the numbers of RNAs or guides or sgRNAs discussed herein, without any undue experimentation.

[0027] The RNA(s), e.g., sgRNA(s), can be functionally or operatively linked to regulatory element(s) and hence the regulatory element(s) drive expression. The promoter(s) can be constitutive promoter(s) and/or inducible promoter(s) and/or tissue specific promoter(s). The promoter can be selected from the group consisting of RNA polymerases, pol I, pol II, pol III, T7, U6, H1, retroviral Rous sarcoma virus (RSV) LTR promoter, the cytomegalovirus (CMV) promoter, the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1 α promoter. An advantageous promoter is the promoter is U6.

[0028] Advantageously, each RNA (e.g., sgRNA) is specific to a different target sequence. Each different target sequence can be associated with or correlated to a form of cancer. Each RNA or sgRNA can be specific to a different target sequence but these RNA(s) or sgRNA(s) target specific gene sequences associated with or correlated to cancer, advantageously a particular type or form of cancer; for instance each RNA or sgRNA can be specific to a different target sequence but the target sequences of the RNA(s) or sgRNA(s) are associated with or correlated to the same type or form of cancer. The cancer selected from the group consisting of

Lung cancer, Lung adenocarcinoma, Lung squamous cell carcinoma, Acute myeloid leukemia, Basal cell carcinoma (skin), Bladder cancer, Breast cancer, Carcinoid, Chronic lymphocytic leukemia, Colorectal cancer, Diffuse large B-cell lymphoma, Endometrial, Esophageal cancer, Esophageal adenocarcinoma, Glioblastoma multiforme, Glioma, Head and neck cancer, Kidney clear cell cancer, Medulloblastoma, Melanoma, Multiple myeloma, Nasopharyngeal, Neuroblastoma, Ovarian cancer, Prostate cancer, Rhabdoid tumor, Testicular germ cell tumor, Thyroid cancer, and Urinary bladder cancer.

[0029] Advantageously, each sgRNA can be driven by an independent U6 promoter. The vector can be an AAV, e.g., an AAV9. Each of the sgRNAs can target a different genetic locus associated with a multigenic disease or disorder, e.g., a cancer, such as lung cancer. The sgRNAs can target one or more tumor suppressor genes so as to introduce loss-of-function mutations. The tumor suppressor genes can be, without limitation, p53 and/or Lkb1. The sgRNAs can target one or more proto-oncogenes or oncogenes so as to introduce a loss-of-function mutation. The proto-oncogenes or oncogenes can be Kras. The sgRNAs can target one or more of p53, Lkb1, or Kras loci. From 3 to 50 specific mutations or precise sequence substitutions in from 3 to 50 target loci can be introduced. More generally, the sgRNAs can target one or more, e.g., 3 to 50 loci specific to a cancer.

[0030] In an aspect, the invention provides a method for modeling tumor formation and/or tumor evolution. In an embodiment of the invention, the method comprises the steps of providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas, introducing into said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, or introducing into said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, introducing said Cas transgenic eukaryotic cell into a non-human eukaryote, and analyzing tumor formation and/or tumor evolution in said non-human eukaryote.

[0031] In one embodiment, the method comprises introducing a plurality of Cas transgenic eukaryotic cells in a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus of

said eukaryote; or each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus of said eukaryote.

[0032] In an embodiment of the invention, the method comprises identifying one or more genes which are involved in tumor formation and/or evolution. In certain embodiments, the method comprises analyzing tumor metastasis. In certain embodiments, the method comprises analyzing gene expression. In certain embodiments, the method comprises analyzing gene expression in the tumor. In certain such embodiments, the method comprises identifying one or more of said RNAs introduced into the transgenic cell.

[0033] In an aspect, the invention provides a method for identifying genes which are involved in tumor formation and/or tumor metastasis. The method comprises the steps of providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas, introducing into the Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, or introducing into said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, introducing a plurality of said Cas transgenic eukaryotic cells into a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus, and identifying genes which are involved in tumor formation and/or tumor metastasis in said non-human eukaryote based on the identification in said tumor or tumor metastasis of one or more of said RNA capable of guiding Cas to one or more genetic target locus.

[0034] In certain embodiments, the genes involved in tumor formation, tumor evolution, and/or tumor metastasis are tumor suppressor genes and/or metastasis suppressor genes. In certain embodiments, the method comprises introducing one or more genomic mutations into said Cas transgenic eukaryotic cell.

[0035] In an embodiment of the invention, the RNA capable of guiding Cas comprises CRISPR RNA and transactivating (tracr) RNA. In an embodiment of the invention, the RNA capable of guiding Cas comprises a single guide (sg) RNA.

[0036] In an embodiment of the invention, the RNA capable of guiding Cas is introduced into said Cas transgenic eukaryotic cell by means of transduction. In an embodiment of the invention, the RNA capable of guiding Cas is introduced in said Cas transgenic eukaryotic cell by means of lentiviral transduction.

[0037] In certain embodiments, the one or more polynucleotide encoding one or more RNA capable of guiding Cas is capable of constitutively expressing said RNA or alternatively inducibly and/or conditionally expressing said RNA.

[0038] In an embodiment of the invention, the Cas transgenic eukaryotic cell is a tumor cell. In an embodiment of the invention, the Cas transgenic eukaryotic cell is a non-metastasizing tumor cell.

[0039] In an embodiment of the invention, the Cas transgenic eukaryotic cell is characterized by oncogenic expression of Kras (KrasG12D), homozygous p53 loss (p53 $-/-$), and/or heterozygous Dicer1 loss (Dicer $+/-$), preferably all.

[0040] In an embodiment of the invention, the Cas is a type II Cas. In certain embodiments, the Cas is Cas9. In certain embodiments, the Cas is a Cas originating from *Streptococcus pyogenes*, *Streptococcus thermophiles*, or *Staphylococcus aureus*.

[0041] In an embodiment of the invention, the Cas is a mutated Cas having an altered catalytic activity. In one such embodiment, the Cas is a catalytically inactive Cas. In another embodiment, the Cas is fused to an enzyme which modifies genomic DNA or genomic DNA architecture. In another embodiment, the Cas is fused to a polypeptide which alters gene transcription.

[0042] In an embodiment, the Cas transgenic eukaryotic cell is a cell from an animal. In certain such embodiments, the Cas transgenic eukaryotic cell is a cell from a eukaryote selected from the group consisting of mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod. In certain

such embodiments, the Cas transgenic eukaryotic cell is a mammalian cell, preferably a mouse cell. In certain embodiments, the eukaryote into which is inserted a Cas transgenic eukaryotic cell is an animal, preferably a mammal. In certain embodiments, the eukaryote is an immunocompromised animal, preferably an immunocompromised mammal. In certain embodiments, the eukaryote is selected from mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod. In an embodiment of the invention the eukaryote is a mouse.

[0043] In an embodiment of the invention, the method comprises introducing Cas transgenic eukaryotic cells subcutaneously.

[0044] In an embodiment of the invention, the selected tumor comprises, without limitation, lung cancer, lung adenocarcinoma, lung squamous cell carcinoma, acute myeloid leukemia, basal cell (skin) carcinoma, bladder cancer, breast cancer, carcinoid cancer, chronic lymphocytic leukemia, colorectal cancer, lymphoma, diffuse large B-cell lymphoma, endometrial cancer, esophageal cancer, esophageal adenocarcinoma, glioblastoma multiforme, glioma, head and neck cancer, kidney cell cancer, medulloblastoma, melanoma, multiple myeloma, nasopharyngeal cancer, neuroblastoma, ovarian cancer; prostate cancer, rhabdoid tumor, thyroid cancer, or urinary bladder cancer.

[0045] In an aspect, the invention provides a tumor sample or a tumor metastasis sample isolated from a non-human eukaryote after introduction of a Cas transgenic eukaryotic cell. In an non-limiting embodiment of the invention, the tumor is selected from lung cancer, lung adenocarcinoma, lung squamous cell carcinoma, acute myeloid leukemia, basal cell (skin) carcinoma, bladder cancer, breast cancer, carcinoid cancer, chronic lymphocytic leukemia, colorectal cancer, lymphoma, diffuse large B-cell lymphoma, endometrial cancer, esophageal cancer, esophageal adenocarcinoma, glioblastoma multiforme, glioma, head and neck cancer, kidney cell cancer, medulloblastoma, melanoma, multiple myeloma, nasopharyngeal cancer, neuroblastoma, ovarian cancer; prostate cancer, rhabdoid tumor, thyroid cancer, and urinary bladder cancer.

[0046] In an aspect, the invention provides a cell from the tumor sample or the tumor metastasis sample from a non-human eukaryote after introduction of a Cas transgenic eukaryotic

cell. In an aspect, the invention provides a cell line obtained or obtainable from the tumor sample or the tumor metastasis sample.

[0047] In an aspect, the invention provides a method for generating a non-human eukaryote model for tumor formation and/or tumor evolution, comprising the steps of providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas, introducing in said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, or introducing in said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, introducing a plurality of the Cas transgenic eukaryotic cell after step (b) in a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus;

[0048] In an aspect, the invention provides a method for generating a non-human eukaryote model for identifying genes involved in formation and/or tumor evolution, comprising the steps of providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas, introducing in said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote or or introducing in said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, introducing a plurality of the Cas transgenic eukaryotic cells in a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus. In an aspect, the invention provides a non-human eukaryote obtained by

[0049] In an aspect, the invention provides a non-human eukaryote obtained or obtainable from the eukaryotic model. In an embodiment the animal is a mammal. In another embodiment, the eukaryote is an immunocompromised animal or mammal. In another embodiment, the eukaryote is selected from a mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow,

bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, or arthropod. In yet another embodiment, the eukaryote is a mouse.

[0050] In an aspect, the invention provides a method for determining whether a compound is capable of suppressing tumor formation and/or tumor evolution comprising providing a non-human eukaryote model of tumor formation and/or tumor evolution, administering the compound to the animal model, and determining whether or not the compound is capable of suppressing tumor formation and/or tumor evolution in said animal model.

[0051] In an aspect, the invention provides a kit comprising an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas, one or more RNA capable of guiding Cas to one or more genetic target locus, or one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus. In an embodiment the kit comprises a non-human eukaryote. In an embodiment, the kit comprises instructions for performing a method for modeling tumor formation and/or tumor evolution, a method for identifying genes which are involved in tumor formation and/or tumor metastasis, or a method for generating a non-human eukaryote model for identifying genes involved in tumor formation and or tumor evolution, or a method for determining whether a compound is capable of suppressing tumor formation. In an embodiment, the kit comprises one or more RNA capable of guiding Cas to one or more genetic target locus, or one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus are comprised in a vector, preferably a lentiviral vector. In an embodiment, the kit comprises a plurality of RNAs capable of guiding Cas to a genetic target locus; or a plurality of polynucleotides encoding an RNA capable of guiding Cas to a genetic target locus. In an embodiment, the kit comprises an RNA capable of guiding Cas comprises CRISPR RNA and transactivating (tracr) RNA. In an embodiment, the kit comprises an RNA capable of guiding Cas comprises a single guide (sg) RNA. In another embodiment, the kit comprises a polynucleotide encoding one or more RNA capable of guiding Cas is capable of constitutively expressing said RNA or alternatively inducibly and/or conditionally expressing said RNA. In an embodiment, the Cas transgenic eukaryotic cell is a tumor cell.

[0052] In an embodiment, the Cas transgenic eukaryotic cell of the kit is a non-metastasizing tumor cell. In an embodiment, the Cas is a type II Cas. 61. In an embodiment, the Cas is Cas9. In an embodiment, the Cas9 originated from *Streptococcus pyogenes*, *Streptococcus thermophiles*, or *Staphylococcus aureus*. In another embodiment, the Cas is a mutated Cas having an altered catalytic activity. In one such embodiment, the Cas is a catalytically inactive Cas. In another embodiment, the Cas is fused to an enzyme which modifies genomic DNA or genomic DNA architecture. In an embodiment, the Cas is fused to a polypeptide which alters gene transcription. In an embodiment, the Cas transgenic eukaryotic cell is a cell from an animal. In certain embodiments, the Cas transgenic cell is selected from mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, or arthropod. In an embodiment, the kit comprises a mouse cell.

[0053] In certain aspects, the invention provides a method for diagnosing tumorigenesis and/or tumor metastasis, a method for prognosing tumorigenesis and/or tumor metastasis, and a method of determining the likelihood of developing tumor metastasis,

[0054] In certain embodiments of the invention, analyzing expression of one or more genes comprises comparing expression level with a predetermined expression level. In certain embodiments of the invention, the gene comprises *Nf2*, *Pten*, *Trim72*, *Cdkn2a*, *miR-152*, or *miR-345*. In an embodiment, the tumor metastasis is lung metastasis. In an embodiment, the tumor or tumor metastasis is selected from lung cancer, lung adenocarcinoma, lung squamous cell carcinoma, acute myeloid leukemia, basal cell (skin) carcinoma, bladder cancer, breast cancer, carcinoid cancer, chronic lymphocytic leukemia, colorectal cancer, lymphoma, diffuse large B-cell lymphoma, endometrial cancer, esophageal cancer, esophageal adenocarcinoma, glioblastoma multiforme, glioma, head and neck cancer, kidney cell cancer, medulloblastoma, melanoma, multiple myeloma, nasopharyngeal cancer, neuroblastoma, ovarian cancer; prostate cancer, rhabdoid tumor, thyroid cancer, or urinary bladder cancer.

[0055] The invention also relates to a non-human eukaryote to which the methods as described above have been applied, i.e. a non-human eukaryote having transplanted eukaryotic cells which have been subjected to the CRISPR/Cas system as described herein.

[0056] The invention further relates to tumors, tumor samples, and cells derived therefrom, or cell lines derived from such tumors, tumor samples, or cells derived therefrom, which tumors originate from a non-human eukaryote to which the methods as described above have been applied. Such tumors, tumor samples, and cells derived therefrom, or cell lines derived from such tumors, tumor samples, or cells derived therefrom comprise one or more genetic modification, epigenetic modification, or gene transcription alteration. Such tumors, tumor samples, and cells derived therefrom, or cell lines derived from such tumors, tumor samples, or cells derived therefrom may also be provided in a kit.

[0057] The invention further relates to tumor metastases, tumor metastase samples, and cells derived therefrom, or cell lines derived from such tumor metastases, tumor metastase samples, or cells derived therefrom, which tumor metastases originate from a non-human eukaryote to which the methods as described above have been applied. Such tumor metastases, tumor metastase samples, and cells derived therefrom, or cell lines derived from such tumor metastases, tumor metastase samples, or cells derived therefrom comprise one or more genetic modification, epigenetic modification, or gene transcription alteration. Such tumor metastases, tumor metastase samples, and cells derived therefrom, or cell lines derived from such tumor metastases, tumor metastase samples, or cells derived therefrom may also be provided in a kit.

[0058] In further aspects, the invention relates to methods for diagnosing and/or prognosing cancer, being it the formation and/or evolution of primary tumors or the formation and/or evolution of secondary tumors, i.e. metastases. Such methods include analysis of one or more of the genes which are identified in the methods according to the invention as described herein. Genetic analysis and/or expression analysis of such genes may provide diagnostic and/or prognostic information of tumor and/or metastasis development and/or evolution, as well as disease progression and/or outcome. Genetic or epigenetic as well as transcriptional alterations may be correlated with disease type, progression, evolution, prognosis, etc.

[0059] In related aspects, the invention also foresees in methods for treating cancer, being it a primary tumor or secondary tumor, i.e. metastasis. The non-human eukaryote to which the methods as described herein are applied may be used as a test model for therapeutic applications. Therapeutic applications obtained as such may then be applied to a subject in need thereof.

[0060] Accordingly, it is an object of the invention to not encompass within the invention any previously known product, process of making the product, or method of using the product such that Applicants reserve the right and hereby disclose a disclaimer of any previously known product, process, or method. It is further noted that the invention does not intend to encompass within the scope of the invention any product, process, or making of the product or method of using the product, which does not meet the written description and enablement requirements of the USPTO (35 U.S.C. §112, first paragraph) or the EPO (Article 83 of the EPC), such that Applicants reserve the right and hereby disclose a disclaimer of any previously described product, process of making the product, or method of using the product.

[0061] It is noted that in this disclosure and particularly in the claims and/or paragraphs, terms such as “comprises”, “comprised”, “comprising” and the like can have the meaning attributed to it in U.S. Patent law; e.g., they can mean “includes”, “included”, “including”, and the like; and that terms such as “consisting essentially of” and “consists essentially of” have the meaning ascribed to them in U.S. Patent law, e.g., they allow for elements not explicitly recited, but exclude elements that are found in the prior art or that affect a basic or novel characteristic of the invention. It may be advantageous in the practice of the invention to be in compliance with Art. 53(c) EPC and Rule 28(b) and (c) EPC. Nothing herein is intended as a promise.

[0062] It is noted that in this disclosure and particularly in the claims and/or paragraphs, terms such as “comprises”, “comprised”, “comprising” and the like can have the meaning attributed to it in U.S. Patent law; e.g., they can mean “includes”, “included”, “including”, and the like; and that terms such as “consisting essentially of” and “consists essentially of” have the meaning ascribed to them in U.S. Patent law, e.g., they allow for elements not explicitly recited, but exclude elements that are found in the prior art or that affect a basic or novel characteristic of the invention. Nothing herein is intended as a promise.

[0063] These and other embodiments are disclosed or are obvious from and encompassed by, the following Detailed Description.

BRIEF DESCRIPTION OF THE FIGURES

[0064] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be

obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0065] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[0066] **Figure 1A-F:** Schematic representation of the loss-of-function metastasis screen using the mouse genome-scale CRISPR/Cas9 knock-out (mGeCKOa) library. (B) Representative hematoxylin and eosin (H&E) stains of primary tumor and various organs of Nu/Nu mice subcutaneously transplanted with a KrasG12D/+;p53-/-;Dicer1+/- NSCLC cell line, which was transduced with lentivirus carrying a Cas9 vector (Cas9 only), or a Cas9 vector plus the mGeCKOa library (Cas9 + mGeCKOa). (C) Primary tumor growth curve of Nu/Nu mice transplanted with Cas9 vector cells (n = 3 mice) or Cas9 + mGeCKOa cells (n = 9 mice). (D) Representative EGFP fluorescence images of lung lobes of animals transplanted with Cas9 vector or Cas9 + mGeCKOa cells. Yellow arrows indicate lung metastases. (E) Number of lobes with metastases visible after dissection of Nu/Nu mice transplanted with Cas9 vector cells (n = 3 mice) or Cas9 + mGeCKOa cells with three independent infection replicate experiments (R1, R2 and R3, n = 3 mice per replicate). (F) Representative H&E stains from various organs of Nu/Nu mice subcutaneously transplanted with Cas9 only or Cas9 + mGeCKOa cells.

[0067] **Figure 2A-D:** (A) Pearson correlation coefficient of the normalized sgRNA read counts from the mGeCKOa plasmid pool, cells before transplantation (day 7 after spinfection), early tumors (~ 2 weeks after transplantation), and lungs (~ 6 weeks after transplantation). For each biological sample type, 3 independent infection replicates (R1, R2 and R3) are shown. n = 1 mouse per infection replicate for early primary tumors and n=3 mice per infection replicate for late primary tumors and lung samples. (B) Number of unique sgRNAs in the plasmid, cells before transplantation, early and late primary tumors and lung metastases as in (A); (C) Boxplot of the sgRNA normalized read counts for the mGeCKOa plasmid pool, cells before transplantation, early and late primary tumors and lung metastases as in (A). Gray dots overlaid on each boxplot indicate read counts for the 1,000 control (non-targeting) sgRNAs in the

mGeCKOa library. (D) Cumulative probability distribution of library sgRNAs in the plasmid, cells before transplantation, early and late primary tumors and lung metastases as in (A). Distributions for each sample type are averaged across individual mice and infection replications.

[0068] Figure 3A-C: (A) Pie charts of the most abundant sgRNAs in the primary tumors (at ~6 weeks post-transplantation) of three representative mice (one from each replicate mGeCKOa infection). The area for each sgRNA corresponds to the fraction of total reads from the lobe for the sgRNA. All sgRNAs with $\geq 2\%$ of total reads are plotted individually. (B) Number of genes with 0, 1, 2 or 3 significantly enriched (FDR < 0.2% for at least one mouse) mGeCKOa sgRNAs targeting that gene. For genes/miRNAs with 2 or more enriched sgRNAs, genes/miRNAs are categorized by how many sgRNAs targeting that gene/miRNAs are enriched as indicated in the colored bubbles adjacent to each bar. (C) Inset: Waterfall plot of sgRNAs where multiple sgRNAs targeting the same gene are significantly enriched in primary tumors. Each sgRNA is ranked by the number of mice in which it is enriched. Only sgRNAs enriched in 2 or more mice are shown. Main panel: Enlargement and gene labels for sgRNAs at the top of the list from the inset (boxed region).

[0069] Figure 4A-H: (A) Pie charts of the most abundant sgRNAs in three individual lobes of the lungs of two representative animals transplanted with Cas9 + mGeCKOa infected cells. The area for each sgRNA corresponds to the fraction of total reads from the lobe for the sgRNA. All sgRNAs with $\geq 2\%$ of total reads are plotted individually. (B) Pie charts of the most abundant sgRNAs in the lung (averaged across three individual lobes) for the two animals shown in (A). All sgRNAs with $\geq 2\%$ of average reads are plotted individually. (C) *Left:* Percentage of late tumor reads for the significantly enriched (FDR<0.2%) mGeCKOa sgRNAs found in the lung metastases (averaged across three dissected lobes). *Right:* In purple, the percentage of late tumor reads for the significantly enriched (FDR<0.2%) mGeCKOa sgRNAs found in the lung metastases (average across all mice, $n=9$ mice). In grey, the percentage of late tumor reads for random, size-matched samples of sgRNAs present in the late tumor ($n=100$ samples). (D) *Inset:* All sgRNAs found in individual lung lobes ordered by the percent of lobes in which a particular sgRNA was amongst the significantly enriched (FDR<0.2%) sgRNAs for that lobe. *Main panel:*

Enlargement and gene labels for sgRNAs at the top of the list from the inset (boxed region). (E) *Inset*: All sgRNAs found in individual mouse (averaged across three dissected lobes) ordered by the percent of mice in which a particular sgRNA was amongst the significantly enriched ($FDR < 0.2\%$) sgRNAs for that mouse. *Main panel*: Enlargement and gene labels for sgRNAs at the top of the list from the inset (boxed region). (F) *Bottom*: Metastasis Primary Ratio (MPR) for the sgRNAs in mGeCKOa with enrichment in metastases over late tumor ($MPR > 1$) observed in at least 3 mice. The sgRNAs are sorted by the number of mice in which the MPR for the sgRNA is greater than 1. *Bottom*: Number of mice in which the MPR for this sgRNA is greater than 1. In both panels, individual sgRNAs are labeled by gene target. (G) Number of genes with 0, 1, 2 or 3 significantly enriched ($FDR < 0.2\%$ for at least one mouse) mGeCKOa sgRNAs in the lung metastases. For genes with 2 enriched sgRNAs, gene names are indicated in the colored bubble adjacent to the bar. (H) Number of mice and percentage of mice in which each sgRNA was enriched in the lung metastases for all genes with multiple enriched sgRNAs.

[0070] Figure 5A-F: (A) Schematic representation of lentiviral transduction of Cas9-EGFP KPD cells with single sgRNAs designed to target one gene or miRNA. After puromycin selection, the cell population is transplanted into Nu/Nu mice and also deep sequenced to examine the distribution of indels at the target site. After 6 weeks, the primary tumor and lungs are examined. (B) Histograms of indel sizes at the genomic locus targeted by a representative sgRNA for each gene/miRNA after 3 days of puromycin selection. (C) Representative H&E staining of lung lobes from uninjected animals ($n = 3$ animals), animals transplanted with cells transduced with Cas9 only ($n = 5$), and animals transplanted with cells containing Cas9 and a single sgRNA ($n = 6$). Single sgRNAs are either control/non-targeting sgRNAs ($n = 6$ mice for control sgRNAs, 3 distinct control sgRNAs with 2 mice each) or targeting sgRNAs ($n = 6$ mice for each gene/miRNA target, 3 sgRNAs per target with 2 mice each). Blue arrows indicate lung metastases. Scale bar: 10 μm . (D) Percent of lung lobes with metastases after 6 weeks for the mice in (C). (E) Primary tumor growth curve of Nu/Nu mice transplanted with NSCLC cells transduced with Cas9 only ($n = 5$) or single sgRNAs ($n = 6$ mice per gene/miRNA target, 3 sgRNAs per target with 2 mice each; $n = 6$ mice for control sgRNAs, 3 control sgRNAs with 2 mice each). (F) Correlation between primary tumor volume and percent of lobes with metastases for each gene in (D) and (E). Error bars indicate s.e.m.

[0071] Figure 6A-E: (A) Schematic representation of the loss-of-function metastasis minipool screen 624 sgRNAs. Briefly, Cas9-EGFP KPD cells were transduced with either validation or control (non-targeting) pools of 624 sgRNAs. After puromycin selection, the cell pools are transplanted into Nu/Nu mice. After 6 weeks, minipool sgRNAs are sequenced from primary tumor and lung samples. (B) Primary tumor growth curve of Nu/Nu mice transplanted with Cas9 vector + validation pool cells (n = 5 mice) or Cas9 + control pool cells (n = 5 mice). (C) Percent of lung lobes with metastases after 6 weeks for the mice in (B). C = control pool. V = validation pool. (D) Boxplot of the sgRNA normalized read counts for the plasmid library, cells before transplantation, primary tumor and lung metastases using the validation minipool. (E) Cumulative probability distribution of library sgRNAs in the validation plasmid pool, cells before transplantation, early tumor and lung metastases. Tumor and metastases distributions are averaged across mice.

[0072] Figure 7A-E: (A) Pie charts of the most abundant sgRNAs in the primary tumor and the whole lung of two representative mice transplanted with Cas9 + validation pool cells. The area for each sgRNA corresponds to the fraction of total reads from the tissue (primary tumor or lung metastases) for the sgRNA. All sgRNAs with $\geq 2\%$ of total reads are plotted individually. (B) Scatterplot of normalized sgRNA read counts in primary tumor and lung metastases for all sgRNAs in the validation minipool for each mouse (different color dots indicate different mice). (C) Base-2 logarithm of the ratio of sgRNA abundance in the lung metastases over the primary tumor (Metastasis-Primary Ratio, or MPR) plotted against the abundance in the lung metastases (n = 5 mice per sgRNA). Green dots are the 100 control sgRNAs. Dots with black outlines are non-control sgRNAs that target genes or miRNAs. Red dots indicate non-control sgRNAs for which more than one sgRNA targeting the same gene/miRNA is enriched in the lung metastases over the primary tumor (i.e. $\log_2(\text{MPR}) > 0$) and are labeled with the gene/miRNA targeted. The lung-primary ratio is calculated for individual mice and these quantities are averaged across mice. (D) Number of genes with 0 to 10 significantly enriched (FDR<0.2% in at least one mouse) validation minipool sgRNAs in the lung metastases. For genes/miRNAs with 2 or more enriched sgRNAs, genes/miRNAs are categorized by how many sgRNAs targeting that gene/miRNAs are enriched as indicated in the colored bubbles adjacent to each bar. (E) Schematic illustration of tumor growth and metastasis in the library-transduced NSCLC

transplant model. The initially diverse set of loss-of-function mutations in the subcutaneously transplanted pool is selected over time for mutations that promote growth of the primary tumor. These mutants are also found in higher numbers in the blood and, subsequently, in lung metastases where a subset of mutations from the primary tumor are able to proliferate.

[0073] Figure 8A-D (S1): (A) Brightfield image of the parental *Kras*G12D/+;*p53*-/-;*Dicer1*+/- NSCLC cell line and EGFP fluorescence images of clonally derived lines after transduction with a NLS-3xFLAG-Cas9-EGFP lentivirus. (B) Western blot analysis of Cas9 expression in clonally derived lines from (A) using anti-FLAG and anti-Cas9 antibodies. Both antibodies detect an ~190 kDa band. (C) Micro-CT snapshots of the lungs of representative mice transplanted with control (Cas9 only) and mGeCKOa-transduced (Cas9 + mGeCKOa) cell pools. Yellow arrows indicate lung metastases. (D) Dissecting scope snapshots of the lungs of representative mice transplanted with control (Cas9 only) and mGeCKOa-transduced (Cas9 + mGeCKOa) cell pools. Yellow arrows indicate lung metastases.

[0074] Figure 9A-F (S2): (A) Pearson correlation of sgRNA normalized read counts between two technical replicates of the mGeCKOa-transduced cell pool from infection replicate 1 (R1) before transplantation. Technical replicates have separate PCR readout and Illumina sequencing. (B) Pearson correlation of sgRNA normalized read counts between two infection replicates (R1 and R2) of the transduced cell pool before transplantation. (C) Pearson correlation of sgRNA normalized read counts between the mGeCKOa plasmid and the transduced cell pools (average of all 3 infection replicates). (D) Pearson correlation of change in rank between mGeCKOa plasmid and transduced cell pools (average of all 3 infection replicates) for all pairs of sgRNAs targeting the same gene. (E) Pearson correlation of normalized sgRNA read counts and dendrogram based on complete linkage (farthest neighbor) clustering for all samples in mGeCKOa initial screen. Plasmid denotes the mGeCKOa plasmid before lentivirus production. Cell denotes the mGeCKOa-transduced cell pool at 7 days post-transduction. Early tumor denotes samples harvested from the primary tumor at ~ 2 weeks after transplantation of the transduced cell pool. Late tumor and lobe denote samples harvested from the primary tumor and lungs, respectively, at ~ 6 weeks after transplantation of the transduced cell pool. Rn denotes independent infection replicates and mn denotes different mice for that infection replicate. (F)

Gene Set Enrichment Analysis (GSEA) of the 10 most depleted gene sets in the cells before transplantation compared to the plasmid pool (FDR $q < 10^{-3}$). Gray lines indicate the rank position of a single gene from the gene set indicated. The green bar indicates the cross-over between enriched and depleted genes. The red line indicates the position of the median gene rank for the genes in the set. Numbers next to the gene set name are the Normalized Enrichment Score.

[0075] Figure 10A-G (S3): (A) Venn diagram of sgRNAs from the mGeCKOa library detected in the 3 infection replicates of the mGeCKOa-transduced cell pool at 7 days post-transduction. (B) Venn diagram of sgRNAs from the mGeCKOa library detected in the 3 infection replicates of the primary tumor at ~ 2 weeks after transplantation of the transduced cell pool. (C) Number of unique sgRNAs from the mGeCKOa library detected in the plasmid pool, cells before transplantation, early and late primary tumors, and lung metastases. **Plasmid** denotes the mGeCKOa plasmid before lentivirus production. **Cell** denotes the mGeCKOa-transduced cell pool at 7 days post-transduction. **Early tumor** denotes samples harvested from the primary tumor at ~ 2 weeks after transplantation of the transduced cell pool. **Late tumor** and **lobe** denote samples harvested from the primary tumor and lungs, respectively, at ~ 6 weeks after transplantation of the transduced cell pool. **R*n*** denotes independent infection replicates and **m*n*** denotes different mice for that infection replicate. Technical replicates of PCR readout and sequencing for **Cell** are denoted by *a* or *b* after the infection replicate (R1, R2, R3). (D) Boxplot of the sgRNA normalized read counts for the mGeCKOa plasmid pool, cells before transplantation, early and late primary tumors, and lung metastases at lobe level with labels as given in Figure S4. Green dots indicate 1,000 control (non-targeting) sgRNAs in the mGeCKOa library. (E) Venn diagram of sgRNAs from the mGeCKOa library detected in lung metastases and primary tumors across all 3 infection replicates ($n = 3$ mice per replicate, 9 mice total) at ~ 6 weeks after transplantation of the transduced cell pool. (F) Scatterplot of the number of sgRNAs detected in the primary tumor and the lung metastases for each individual mouse transplanted with mGeCKOa cells, showing correlation of sgRNA diversity between lung and late primary tumors for each mouse ($n=9$). (G) Pearson correlation of mGeCKOa sgRNA abundance in late primary tumor and lung metastases in the same mouse (taken from one representative mouse with a median correlation among the 9 mice).

[0076] Figure 11A-F (S4): (A) Venn diagram of enriched sgRNAs in late primary tumor, using three different methods, by rank, by read count cutoff, or by FDR. (B) Histogram of all 713 sgRNAs with a Metastasis-Primary Ratio (MPR: ratio of normalized reads in metastases [averaged over lobes] to normalized reads in late tumor) greater than 1 in at least one mouse. sgRNAs are binned by how many mice for which the MPR>1. Red-filled bars indicate the sgRNAs plotted in Figure 4F (MPR>1 in 3 or more mice). *Inset:* Distribution of MPRs for all sgRNAs in the mGeCKOa library detected in either late tumor or metastases. Values are averaged over all mice ($n=9$ mice). (C) Venn diagram of enriched sgRNAs in lung metastases, using three different methods, by rank, by read count cutoff, or by FDR. (D) Venn diagram of sgRNAs from the mGeCKOa library enriched in 3 lobes of the lung from a single mouse at ~ 6 weeks after transplantation of the transduced cell pool. (E) Venn diagram of sgRNAs from the mGeCKOa library enriched in 3 different mice at ~ 6 weeks after transplantation of the transduced cell pool from a single infection replicate. The union of the sgRNAs from 3 lung lobes is used for each mouse. (F) Venn diagram of sgRNAs from the mGeCKOa library enriched in lung metastases across all mice from 3 independent infection replicates ($n = 3$ mice per replicate) at ~ 6 weeks after transplantation of the transduced cell pool.

[0077] Figure 12A-C (S5): (A) Barplot of mutation frequency of the late primary multi-hit gene set (24 multiple enriched sgRNA - targeting genes in late primary tumor) in human cancer sequencing datasets. (B) Co-mutation plot of the late primary multi-hit gene set in TCGA LUAD patients. (C) Co-mutation plot of the late primary multi-hit gene set in TCGA LUSC patients.

[0078] Figure 13A-G (S6): (A) Western blot protein analysis of individual sgRNAs targeting *Nf2* and *Pten*. For subsequent *in vivo* experiments, *Nf2* sgRNAs 1, 2 and 4 and *Pten* sgRNAs 1, 3 and 4 were used. (B) Distribution of cell size (diameter) of the Cas9-GFP transduced mouse *Kras*^{G12D/+};*p53*^{-/-};*Dicer1*^{+/-} NSCLC cell line using Cellometer Auto T4. (C) (upper) Schematics of CTC capturing device. (lower) Brightfield and EGFP images of NLS-3xFLAG-Cas9-EGFP transduced *Kras*^{G12D/+};*p53*^{-/-};*Dicer1*^{+/-} NSCLC cells trapped in PDMS cell capture device. Scale bar: 50 μ m. (D) Quantification of circulating tumor cells (CTCs) from the blood of mice injected with *sg-Nf2*, *sg-Pten*, *sg-Cdkn2a*, *sg-Trim72* or control (pooled Cas9-only and non-targeting sgRNAs) cells at ~ 5 weeks after transplantation. CTCs from $n = 2\sim4$ mice

were captured for each condition. *Sg-miR152* has only one replicate thus is not plotted in this panel. (E) Correlation between primary tumor size and CTC concentration for genes in (D). (F) Correlation between CTC concentration and lung metastasis rate for genes in (D).

[0079] **Figure 14A-I (S7):** (A) Pearson correlation of sgRNA normalized read counts between two technical replicates of the Validation minipool transduced cells before transplantation (~7 days post-transduction). Technical replicates have separate PCR readout and Illumina sequencing. (B) Pearson correlation of sgRNA normalized read counts between the Validation minipool plasmid and the transduced cell pool before transplantation (~7 days post-transduction). (C) Number of unique sgRNAs from the mGeCKOa library detected in the plasmid pool, cells before transplantation, primary tumors, and lung metastases. Different mice are denoted by m1-m5. (D) Pearson correlation of normalized sgRNA read counts and dendrogram based on complete linkage (farthest neighbor) clustering for all samples in validation minipool screen. TR, technical replicate. (E) Venn diagram of sgRNAs from the mGeCKOa library detected in primary tumors of 5 different mice at ~ 6 weeks after transplantation of the validation minipool transduced cell pool from a single infection replicate. (F) Venn diagram of sgRNAs from the mGeCKOa library detected in lung metastases of 5 different mice at ~ 6 weeks after transplantation of the validation minipool transduced cell pool from a single infection replicate. (G) All sgRNAs from the Validation minipool found in the lungs of individual mice ordered by the percent of mice where this sgRNA was amongst the top 10 most enriched sgRNAs for that lobe. In cases where more than 1 sgRNA for the gene was enriched, distinct sgRNAs are labeled. NTCs denote non-targeting, control sgRNAs from the validation minipool. (H) Pie charts of the top 10 most abundant sgRNAs in the primary tumors of 5 mice transplanted with Cas9 + Validation minipool infected cells. The area for each sgRNA corresponds to the fraction of total reads from the tumor for the sgRNA. The 10 most abundant sgRNAs are plotted individually. (I) Pie charts of the top 10 most abundant sgRNAs in the lung metastases of 5 mice transplanted with Cas9 + Validation minipool infected cells. The area for each sgRNA corresponds to the fraction of total reads from the tumor for the sgRNA. The 10 most abundant sgRNAs are plotted individually.

[0080] **Figure 15A-C (Table S1)** Comparison of the overall distributions of sgRNAs from early primary tumors with those in the cell populations. Identification of the sgRNAs which are depleted in cells and in early tumors.

[0081] **Figure 16A-B** Tiny pool sgRNA representation in cell pool, late tumor and lung metastasis.

[0082] **Figure 17** Boxplot of the sgRNA representation in the human xenograft model. Tu= human cell line H23 mouse xenograft sgRNA representation, ETP = human cell line H23 in vitro culture sgRNA representation, log=log2 normalized reads.

[0083] **Figure 18A-CF** Supplemental tables.

DETAILED DESCRIPTION OF THE INVENTION

[0084] Before the present methods of the invention are described, it is to be understood that this invention is not limited to particular methods, components, products or combinations described, as such methods, components, products and combinations may, of course, vary. It is also to be understood that the terminology used herein is not intended to be limiting, since the scope of the present invention will be limited only by the appended claims.

[0085] As used herein, the singular forms “a”, “an”, and “the” include both singular and plural referents unless the context clearly dictates otherwise.

[0086] The terms “comprising”, “comprises” and “comprised of” as used herein are synonymous with “including”, “includes” or “containing”, “contains”, and are inclusive or open-ended and do not exclude additional, non-recited members, elements or method steps. It will be appreciated that the terms “comprising”, “comprises” and “comprised of” as used herein comprise the terms “consisting of”, “consists” and “consists of”, as well as the terms “consisting essentially of”, “consists essentially” and “consists essentially of”.

[0087] The recitation of numerical ranges by endpoints includes all numbers and fractions subsumed within the respective ranges, as well as the recited endpoints.

[0088] The term “about” or “approximately” as used herein when referring to a measurable value such as a parameter, an amount, a temporal duration, and the like, is meant to encompass

variations of +/-20% or less, preferably +/-10% or less, more preferably +/-5% or less, and still more preferably +/-1% or less of and from the specified value, insofar such variations are appropriate to perform in the disclosed invention. It is to be understood that the value to which the modifier “about” or “approximately” refers is itself also specifically, and preferably, disclosed.

[0089] Whereas the terms “one or more” or “at least one”, such as one or more or at least one member(s) of a group of members, is clear per se, by means of further exemplification, the term encompasses inter alia a reference to any one of said members, or to any two or more of said members, such as, e.g., any ≥ 3 , ≥ 4 , ≥ 5 , ≥ 6 or ≥ 7 etc. of said members, and up to all said members.

[0090] All references cited in the present specification are hereby incorporated by reference in their entirety. In particular, the teachings of all references herein specifically referred to are incorporated by reference.

[0091] Unless otherwise defined, all terms used in disclosing the invention, including technical and scientific terms, have the meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. By means of further guidance, term definitions are included to better appreciate the teaching of the present invention.

[0092] In the following passages, different aspects of the invention are defined in more detail. Each aspect so defined may be combined with any other aspect or aspects unless clearly indicated to the contrary. In particular, any feature indicated as being preferred or advantageous may be combined with any other feature or features indicated as being preferred or advantageous.

[0093] Standard reference works setting forth the general principles of recombinant DNA technology include Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, ed. Sambrook et al., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989; Current Protocols in Molecular Biology, ed. Ausubel et al., Greene Publishing and Wiley-Interscience, New York, 1992 (with periodic updates) (“Ausubel et al. 1992”); the series Methods in Enzymology (Academic Press, Inc.); Innis et al., PCR Protocols: A Guide to Methods and Applications, Academic Press: San Diego, 1990; 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). General principles of microbiology are set forth, for example, in Davis, B. D. et al., Microbiology, 3rd edition, Harper & Row, publishers, Philadelphia, Pa. (1980).

[0094] Reference throughout this specification to “one embodiment” or “an embodiment” means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, appearances of the phrases “in one embodiment” or “in an embodiment” in various places throughout this specification are not necessarily all referring to the same embodiment, but may. Furthermore, the particular features, structures or characteristics may be combined in any suitable manner, as would be apparent to a person skilled in the art from this disclosure, in one or more embodiments. Furthermore, while some embodiments described herein include some but not other features included in other embodiments, combinations of features of different embodiments are meant to be within the scope of the invention, and form different embodiments, as would be understood by those in the art. For example, in the appended claims, any of the claimed embodiments can be used in any combination.

[0095] In the following detailed description of the invention, reference is made to the accompanying drawings that form a part hereof, and in which are shown by way of illustration only of specific embodiments in which the invention may be practiced. It is to be understood that other embodiments may be utilised and structural or logical changes may be made without departing from the scope of the present invention. The following detailed description, therefore, is not to be taken in a limiting sense, and the scope of the present invention is defined by the appended claims.

[0096] It is an object of the invention to not encompass within the invention any previously known product, process of making the product, or method of using the product such that Applicants reserve the right and hereby disclose a disclaimer of any previously known product, process, or method. It is further noted that the invention does not intend to encompass within the scope of the invention any product, process, or making of the product or method of using the product, which does not meet the written description and enablement requirements of the

USPTO (35 U.S.C. §112, first paragraph) or the EPO (Article 83 of the EPC), such that Applicants reserve the right and hereby disclose a disclaimer of any previously described product, process of making the product, or method of using the product.

[0097] Preferred statements (features) and embodiments of this invention are set herein below. Each statements and embodiments of the invention so defined may be combined with any other statement and/or embodiments unless clearly indicated to the contrary. In particular, any feature indicated as being preferred or advantageous may be combined with any other feature or features or statements indicated as being preferred or advantageous. Hereto, the present invention is in particular captured by any one or any combination of one or more of the below numbered statements and embodiments 1 to 75, with any other statement and/or embodiments.

[0098] 1. A method for modeling tumor formation and/or tumor evolution comprising the steps of:

[0099] (a) providing an isolated Cas (transient or stable) transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas;

[00100] (b) introducing in said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; or

[00101] introducing in said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote;

[00102] (c) introducing said Cas transgenic eukaryotic cell after step (b) in a non-human eukaryote;

[00103] (d) analyzing tumor formation and/or tumor evolution in said non-human eukaryote.

[00104] 2. The method according to statement 1, 78, or 79, wherein step (c) comprises introducing a plurality of eukaryotic cells in a non-human eukaryote, each cell of said plurality of eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus of said eukaryote; or

each cell of said plurality of eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus of said eukaryote.

[00105] 3. The method according to any of statements 1 to 2 or 78 to 79, wherein step (d) comprises identifying one or more genes which are involved in tumor formation and/or evolution.

[00106] 4. The method according to any of statements 1 to 3 or 78 to 79, wherein step (d) comprises analyzing tumor metastasis.

[00107] 5. The method according to any of statements 1 to 4 or 78 to 79, wherein step (d) comprises analyzing gene expression.

[00108] 6. The method according to any of statements 1 to 5 or 78 to 79, wherein step (d) comprises analyzing gene expression in said tumor.

[00109] 7. The method according to any of statements 1 to 5 or 78 to 79, wherein step (d) comprises identifying one or more of said RNA of step (b) in said tumor.

[00110] 8. The method according to any of statements 4 to 5 or 78 to 79, wherein step (d) comprises analyzing gene expression in said tumor metastasis.

[00111] 9. The method according to any of statements 4 to 5 or 78 to 79, wherein step (d) comprises identifying one or more of said RNA of step (b) in said tumor metastasis.

[00112] 10. A method for identifying genes which are involved in tumor formation and/or tumor metastasis, comprising the steps of:

(a) providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas;

(b) introducing in said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; or

introducing in said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote;

(c) introducing a plurality of said Cas transgenic eukaryotic cell after step (b) in a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different

RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus;

(d) identifying genes which are involved in tumor formation and/or tumor metastasis in said non-human eukaryote based on the identification in said tumor or tumor metastasis of one or more of said RNA capable of guiding Cas to one or more genetic target locus.

[00113] 11. The method according to any of statements 3 to 10, 42-43, 47 or 78 to 81, wherein said genes involved in tumor formation, tumor evolution, and/or tumor metastasis are tumor suppressor genes and/or metastasis suppressor genes.

[00114] 12. The method according to any of statements 1 to 11, 42-43, 47 or 78 to 81, wherein said method after step (b) comprises introducing one or more genomic mutations in said eukaryotic cell.

[00115] 13. The method according to any of statements 1 to 12, 42-43, 47 or 78 to 81, wherein said RNA capable of guiding Cas comprises CRISPR RNA and transactivating (tracr) RNA.

[00116] 14. The method according to any of statements 1 to 13, 42-43, 47 or 78 to 81, wherein said RNA capable of guiding Cas comprises a single guide (sg) RNA.

[00117] 15. The method according to any of statements 1 to 14, 42-43, 47 or 78 to 81, wherein said RNA capable of guiding Cas is introduced in said eukaryotic cell by means of transduction.

[00118] 16. The method according to any of statements 1 to 15, 42-43, 47 or 78 to 81, wherein said RNA capable of guiding Cas is introduced in said eukaryotic cell by means of lentiviral transduction.

[00119] 17. The method according to any of statements 1 to 16, 42-43, 47 or 78 to 81, wherein said polynucleotide encoding one or more RNA capable of guiding Cas is capable of constitutively expressing said RNA or alternatively inducibly and/or conditionally expressing said RNA.

[00120] 18. The method according to any of statements 1 to 17, 42-43, 47 or 78 to 81, wherein said eukaryotic cell of step (a) is a tumor cell.

- [00121] 19. The method according to any of statements 1 to 18, 42-43, 47 or 78 to 81, wherein said eukaryotic cell of step (a) is a non-metastasizing tumor cell.
- [00122] 20. The method according to any of statements 1 to 18, 42-43, 47 or 78 to 81, wherein said eukaryotic cell of step (a) is characterized by oncogenic expression of Kras (KrasG12D), homozygous p53 loss (p53 -/-), and/or heterozygous Dicer1 loss (Dicer +/-), preferably all.
- [00123] 21. The method according to any of statements 1 to 20, 42-43, 47 or 78 to 81, wherein said Cas is a type II Cas.
- [00124] 22. The method according to any of statements 1 to 21, 42-43, 47 or 78 to 81, wherein said Cas is Cas9.
- [00125] 23. The method according to any of statements 1 to 22, 42-43, 47 or 78 to 81, wherein said Cas is a Cas originating from *Streptococcus pyogenes*, *Streptococcus thermophiles*, or *Staphylococcus aureus*.
- [00126] 24. The method according to any of statements 1 to 23, 42-43, 47 or 78 to 81, wherein said Cas is a mutated Cas having an altered catalytic activity.
- [00127] 25. The method according to any of statements 1 to 24, 42-43, 47 or 78 to 81, wherein said Cas is a catalytically inactive Cas.
- [00128] 26. The method according to any of statements 1 to 25, 42-43, 47 or 78 to 81, wherein said Cas is fused to an enzyme which modifies genomic DNA or genomic DNA architecture.
- [00129] 27. The method according to any of statements 1 to 26, 42-43, 47 or 78 to 81, wherein said Cas is fused to a polypeptide which alters gene transcription.
- [00130] 28. The method according to any of statements 1 to 27, 42-43, 47 or 78 to 81, wherein said eukaryotic cell is a cell from an animal.
- [00131] 29. The method according to any of statements 1 to 28, 42-43, 47 or 78 to 81, wherein said eukaryotic cell is a cell from a eukaryote selected from the group consisting of mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod.

[00132] 30. The method according to any of statements 1 to 29, 42-43, 47 or 78 to 81, wherein said eukaryotic cell is a mammalian cell, preferably a mouse cell.

[00133] 31. The method according to any of statements 1 to 30, 42-43, 47 or 78 to 81, wherein said eukaryote in step (c) is an animal, preferably a mammal.

[00134] 32. The method according to any of statements 1 to 31, 42-43, 47 or 78 to 81, wherein said eukaryote in step (c) is an immunocompromised animal, preferably an immunocompromised mammal.

[00135] 33. The method according to any of statements 1 to 32, 42-43, 47 or 78 to 81, wherein said eukaryote in step (c) is selected from the group consisting of mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod.

[00136] 34. The method according to any of statements 1 to 33, 42-43, 47 or 78 to 81, wherein said eukaryote in step (c) is a mouse.

[00137] 35. The method according to any of statements 1 to 34, 42-43, 47 or 78 to 81, wherein step (c) comprises introducing said cells subcutaneously.

[00138] 36. The method according to any of statements 1 to 35, 42-43, 47 or 78 to 81, wherein said tumor is selected from the group consisting of lung cancer, lung adenocarcinoma, lung squamous cell carcinoma, acute myeloid leukemia, basal cell (skin) carcinoma, bladder cancer, breast cancer, carcinoid cancer, chronic lymphocytic leukemia, colorectal cancer, lymphoma, diffuse large B-cell lymphoma, endometrial cancer, esophageal cancer, esophageal adenocarcinoma, glioblastoma multiforme, glioma, head and neck cancer, kidney cell cancer, medulloblastoma, melanoma, multiple myeloma, nasopharyngeal cancer, neuroblastoma, ovarian cancer; prostate cancer, rhabdoid tumor, thyroid cancer, and urinary bladder cancer.

[00139] 37. A tumor sample or a tumor metastasis sample isolated from said non-human eukaryote after introduction of said Cas transgenic eukaryotic cell of step (c) in statement 1.

[00140] 38. The tumor sample or the tumor metastasis sample according to statement 37, wherein said tumor is selected from the group consisting of lung cancer, lung adenocarcinoma, lung squamous cell carcinoma, acute myeloid leukemia, basal cell (skin) carcinoma, bladder

cancer, breast cancer, carcinoid cancer, chronic lymphocytic leukemia, colorectal cancer, lymphoma, diffuse large B-cell lymphoma, endometrial cancer, esophageal cancer, esophageal adenocarcinoma, glioblastoma multiforme, glioma, head and neck cancer, kidney cell cancer, medulloblastoma, melanoma, multiple myeloma, nasopharyngeal cancer, neuroblastoma, ovarian cancer; prostate cancer, rhabdoid tumor, thyroid cancer, and urinary bladder cancer.

[00141] 39. A cell isolated from said tumor sample or said tumor metastasis sample according to any of statements 37 to 38.

[00142] 40. A cell line obtained or obtainable from said tumor sample or said tumor metastasis sample according to any of statements 37 to 38.

[00143] 41. A cell line obtained or obtainable from said cell of statement 39.

[00144] 42. A method for generating a non-human eukaryote model for tumor formation and/or tumor evolution, comprising the steps of:

- (a) providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas;
- (b) introducing in said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; or introducing in said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote;
- (c) introducing a plurality of said Cas transgenic eukaryotic cell after step (b) in a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus;

[00145] 43. A method for generating a non-human eukaryote model for identifying genes involved in formation and/or tumor evolution, comprising the steps of:

- (a) providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas;

(b) introducing in said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; or introducing in said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote;

(c) introducing a plurality of said Cas transgenic eukaryotic cell after step (b) in a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus;

[00146] 44. A non-human eukaryote obtained or obtainable by statement 42 or 43.

[00147] 45. The non-human eukaryote according to statement 44, wherein said eukaryote is an animal, preferably a mammal.

[00148] 46. The non-human eukaryote according to any of statements 44 to 45, wherein said eukaryote is an immunocompromised animal, preferably an immunocompromised mammal.

[00149] 47. The non-human eukaryote according to any of statements 44 to 46, wherein said eukaryote is selected from the group consisting of mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod.

[00150] 48. The non-human eukaryote according to any of statements 44 to 47, wherein said eukaryote is a mouse.

[00151] 49. A method for determining whether a compound is capable of suppressing tumor formation and/or tumor evolution comprising the steps of:

- (a) providing an animal model according to the method of statement 42,
- (b) administering said compound to said animal model
- (c) determining the whether or not said compound is capable of suppressing tumor formation and/or tumor evolution in said animal model.

[00152] 50. A kit comprising:

- an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas;
- one or more RNA capable of guiding Cas to one or more genetic target locus, or one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus

[00153] 51. The kit according to statement 50, further comprising a non-human eukaryote.

[00154] 52. The kit according to any of statements 50 to 51, further comprising instructions for performing the method according to any of statements 1 to 36.

[00155] 53. The kit according to any of statements 50 to 52, wherein said one or more RNA capable of guiding Cas to one or more genetic target locus, or one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus are comprised in a vector, preferably a lentiviral vector.

[00156] 54. The kit according to any of statements 50 to 53, comprising a plurality of RNAs capable of guiding Cas to a genetic target locus; or a plurality of polynucleotides encoding an RNA capable of guiding Cas to a genetic target locus.

[00157] 55. The kit according to any of statements 50 to 54, wherein said RNA capable of guiding Cas comprises CRISPR RNA and transactivating (tracr) RNA.

[00158] 56. The kit according to any of statements 50 to 55, wherein said RNA capable of guiding Cas comprises a single guide (sg) RNA.

[00159] 57. The kit according to any of statements 50 to 56, wherein said polynucleotide encoding one or more RNA capable of guiding Cas is capable of constitutively expressing said RNA or alternatively inducibly and/or conditionally expressing said RNA.

[00160] 58. The kit according to any of statements 50 to 57, wherein said Cas transgenic eukaryotic cell is a tumor cell.

[00161] 59. The kit according to any of statements 50 to 58, wherein said Cas transgenic eukaryotic cell is a non-metastasizing tumor cell.

[00162] 60. The kit according to any of statements 50 to 59, wherein said Cas is a type II Cas.

- [00163] 61. The kit according to any of statements 50 to 60, wherein said Cas is Cas9.
- [00164] 62. The kit according to any of statements 50 to 61, wherein said Cas is a Cas originating from *Streptococcus pyogenes*, *Streptococcus thermophiles*, or *Staphylococcus aureus*.
- [00165] 63. The kit according to any of statements 50 to 62, wherein said Cas is a mutated Cas having an altered catalytic activity.
- [00166] 64. The kit according to any of statements 50 to 63, wherein said Cas is a catalytically inactive Cas.
- [00167] 65. The kit according to any of statements 50 to 64, wherein said Cas is fused to an enzyme which modifies genomic DNA or genomic DNA architecture.
- [00168] 66. The kit according to any of statements 50 to 65, wherein said Cas is fused to a polypeptide which alters gene transcription.
- [00169] 67. The kit according to any of statements 50 to 66, wherein said Cas transgenic eukaryotic cell is a cell from an animal.
- [00170] 68. The kit according to any of statements 50 to 67, wherein said Cas transgenic eukaryotic cell is a cell from a eukaryote selected from the group consisting of mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod.
- [00171] 69. The kit according to any of statements 50 to 68, wherein said Cas transgenic eukaryotic cell is a mammalian cell, preferably a mouse cell.
- [00172] 70. The kit according to any of statements 51 to 69, wherein said eukaryote is an animal, preferably a mammal.
- [00173] 71. The kit according to any of statements 51 to 70, wherein said eukaryote is an immunocompromised animal, preferably an immunocompromised mammal.
- [00174] 72. The kit according to any of statements 51 to 71, wherein said eukaryote is selected from the group consisting of mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod.

[00175] 73. The kit according to any of statements 52 to 72, wherein said eukaryote is a mouse.

[00176] 74. A method for diagnosing tumorigenesis and/or tumor metastasis, comprising the step of genetically analyzing or analyzing expression of one or more genes identified according to the method according to any of statements 10 to 36 or 83 to 84.

[00177] 75. A method for prognosing tumorigenesis and/or tumor metastasis, comprising the step of genetically analyzing or analyzing expression of one or more genes identified according to the method according to any of statements 10 to 36 or 83 to 84.

[00178] 76. A method for determining the likelihood of developing tumor metastasis, comprising the step of genetically analyzing or analyzing expression of one or more genes identified according to the method according to any of statements 10 to 36 or 83 to 84.

[00179] 77. The method according to any of statements 74 to 76, wherein analyzing expression of said one or more genes comprises comparing expression level with a predetermined expression level.

[00180] 78. The method according to any of statements 74 to 77, wherein said one or more gene is selected from the group consisting of Nf2, Pten, Trim72, and Ube2g2.

[00181] 79. The method according to statement 78, wherein said tumor metastasis is lung metastasis.

[00182] 80. The method according to any of statements 74 to 79, wherein said tumor or tumor metastasis is selected from the group consisting of lung cancer, lung adenocarcinoma, lung squamous cell carcinoma, acute myeloid leukemia, basal cell (skin) carcinoma, bladder cancer, breast cancer, carcinoid cancer, chronic lymphocytic leukemia, colorectal cancer, lymphoma, diffuse large B-cell lymphoma, endometrial cancer, esophageal cancer, esophageal adenocarcinoma, glioblastoma multiforme, glioma, head and neck cancer, kidney cell cancer, medulloblastoma, melanoma, multiple myeloma, nasopharyngeal cancer, neuroblastoma, ovarian cancer; prostate cancer, rhabdoid tumor, thyroid cancer, and urinary bladder cancer.

[00183] 81. A method for modeling tumor formation and/or tumor evolution comprising the steps of:

(a) introducing in a eukaryotic cell a Cas polypeptide; or stably or transiently introducing in a eukaryotic cell a polynucleotide encoding a Cas polypeptide, said polynucleotide being capable of constitutively expressing said Cas or alternatively inducibly and/or conditionally expressing said Cas;

(b) introducing in said eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; or

introducing in said eukaryotic cell stably or transiently one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, said polynucleotide being capable of constitutively expressing said RNA or alternatively inducibly and/or conditionally expressing said RNA;

wherein steps (a) and (b) are performed simultaneously or subsequently, in either order;

(c) subsequently introducing said eukaryotic cell in a non-human eukaryote;

(d) analyzing tumor formation and/or tumor evolution in said non-human eukaryote.

[00184] 82. A method for modeling tumor formation and/or tumor evolution comprising the steps of:

(a) providing an isolated eukaryotic cell,

said cell comprising a Cas polypeptide, or stably or transiently expressing Cas or capable of inducibly and/or conditionally expressing Cas; and

said cell comprising one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, or stably or transiently expressing one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote or capable of inducibly and/or conditionally expressing one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote;

(b) subsequently introducing said eukaryotic cell in a non-human eukaryote;

(c) analyzing tumor formation and/or tumor evolution in said non-human eukaryote.

[00185] 83. A method for identifying genes which are involved in tumor formation and/or tumor metastasis, comprising the steps of:

(a) introducing in a eukaryotic cell a Cas polypeptide; or stably or transiently introducing in a eukaryotic cell a polynucleotide encoding a Cas polypeptide, said polynucleotide being capable of constitutively expressing said Cas or alternatively inducibly and/or conditionally expressing said Cas;

(b) introducing in said eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; or

introducing in said eukaryotic cell stably or transiently one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, said polynucleotide being capable of constitutively expressing said RNA or alternatively inducibly and/or conditionally expressing said RNA;

wherein steps (a) and (b) are performed simultaneously or subsequently, in either order;

(c) subsequently introducing a plurality of said eukaryotic cells in a non-human eukaryote, each cell of said plurality of eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus;

(d) identifying genes which are involved in tumor formation and/or tumor metastasis in said non-human eukaryote based on the identification in said tumor or tumor metastasis of one or more of said RNA capable of guiding Cas to one or more genetic target locus.

[00186] 84. A method for identifying genes which are involved in tumor formation and/or tumor metastasis, comprising the steps of:

(a) providing an isolated eukaryotic cell

said cell comprising a Cas polypeptide, or stably or transiently expressing Cas or capable of inducibly and/or conditionally expressing Cas; and

said cell comprising one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, or stably or transiently expressing one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote or capable of inducibly and/or

conditionally expressing one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote;

(b) subsequently introducing a plurality of said eukaryotic cells in a non-human eukaryote, each cell of said plurality of eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus;

(c) identifying genes which are involved in tumor formation and/or tumor metastasis in said non-human eukaryote based on the identification in said tumor or tumor metastasis of one or more of said RNA capable of guiding Cas to one or more genetic target locus.

[00187] As used herein, the term “modeling tumor formation and/or tumor evolution” and “modeling tumor metastasis formation and/or tumor evolution” generally encompasses the analysis of tumor (metastasis) formation and/or tumor (metastasis) evolution. Such analyses may encompass structural and functional analyses, as is known in the art. Also spatio-temporal analyses of tumor (metastasis) formation and/or evolution are envisaged. Non-limiting examples of such analyses include, but are not limited to (epi)genetic and/or morphological characterization of the tumor (metastases), analysis of the type or stage and growth characteristics of the tumor (metastases), including changes over time. Also transcriptome and/or proteome analyses are included, as well as analyses of individual genes and/or expression products.

[00188] As used herein, the term “gene(s) which is (are) involved in tumor formation and/or evolution” refers to genes – any type of gene, i.e. including for instance microRNA genes – the mutation of which, or the altered expression of which is causally correlated with tumor formation and/or evolution. In certain embodiments, the method according to the invention allow identification of tumor suppressor genes and/or metastasis suppressor genes. Tumor suppressor genes are those genes which prevent tumor formation and/or development. Metastasis suppressors are those genes which prevent metastasis formation and/or development. A metastasizing tumor cell is a cell capable of leading to tumor metastasis, i.e. a cell capable of undergoing the typical hallmarks of metastasis, i.e. local invasion into the surrounding tissue, intravasation into the blood stream from the primary tumor site, circulation in the blood stream,

and extravasation from the blood stream into a secondary site, followed by settling/colonization and metastasis growth and development. In this context, a non-metastasising tumor cell is thus a cell which does not undergo, or is not capable of undergoing the above cascade of events.

[00189] As used herein, the term “genetic target locus” refers to a specific physical site in the genome of the eukaryotic cell which is targeted by the CRISPR/Cas system as referred to herein elsewhere. Such genetic target locus is preferably but need not be a specific sequence of a gene, including intron, exon, promoter, and other regulatory sequences including terminators, translational regulatory sequences such as ribosome binding sites and internal ribosome entry sites, enhancers, silencers, insulators, boundary elements, replication origins, matrix attachment sites and locus control regions, in as far as targeting the locus results in genomic alteration and/or altered gene expression.

[00190] An altered expression of one or more genome sequences can be determined by assaying for a difference in the mRNA levels of the corresponding genes between the test model cell and a control cell. Alternatively, the differential expression of sequences is determined by detecting a difference in the level of the encoded polypeptide or gene product.

[00191] As used herein, the term “eukaryotic cell” may refer to a cell or a plurality of cells derived from a eukaryotic organism. In preferred embodiments, such eukaryotic cells are derived from an animal, such as mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, or arthropod, preferably a mammal, such as a rodent, in particular a mouse. In certain embodiments, such eukaryotic cells are non-human eukaryotic cells. The cell type and cell origin are not particularly limiting according to embodiments of the invention. Eukaryotic cells may be primary cells or cell lines. Eukaryotic cells may be dividing cells (e.g. stem cells) or partially or terminally differentiated cells. Eukaryotic cells may in certain embodiments be tumor cells, which may or may not be capable of metastasis or which may or may not be derived from a metastatic tumor. Eukaryotic cells may also be in vitro transformed eukaryotic cells, e.g. in order to render them tumorigenic, whether or not with metastatic potential. Exemplary eukaryotic cell lines include, but are not limited to C8161 , CCRF-CEM, MOLT, mI CD-3, NHDF, HeLa-S3, Huh 1, Huh4, Huh7, HUVEC, HASMC, HEKn, HEKa, MiaPaCell, Panel , PC-3, TFI, CTLL-2, CiR, Rat6, CV1 , RPTE, A10,

T24, .182, A375, ARH-77, Calul, SW480, SW620, S OV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI-231, HB56, TIB55, Jurkat, J45.0L LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, BS-C-1 monkey kidney epithelial, BALB/ 3T3 mouse embryo fibroblast, 3T3 Swiss, 3T3-L1, 132-d5 human fetal fibroblasts; 10.1 mouse fibroblasts, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780eis, A172, A20, A253, A431, A-549, ALC, B16, B35, BCP-1 cells, BEAS- 2B, bEnd.3, ! -2 1 . BR 293, BxPC3, C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CeO-IR, CHO-K1, CHO-K2, CHQ-T, CHO Dhfr -/-, COR-L23, COR-L23/CPR, COR-L23/5010, CGR- L23/R23, COS-7, COV-434, CM L TI, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Flepal cl c7, HL-60, HMEC, HT-29, Jurkat, JY ceils, 562 cells, Ku812, KCL22, KG 1, KAO1, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-IOA, MDA-MB-231, MDA-MB-468, MDA-MB- 435, MDCK II, MDCK II, MOR/0.2R, MONO-MAC 6, MTD-1A, My End, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, \ H i-3T3. N ALM- 1, NW-145, OPCN / OPCT cell lines, Peer, PNT-1A / PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2 cells, Sf-9, SkBr3, T2, T-47D, T84, TI-IPI cell line, U373, U87, U937, VCaP, Vero cells, WM39, WT-49, X63, YAC-1, YAR, and transgenic varieties thereof. Cell lines are available from a variety of sources known to those with skill in the art (see, e.g., the American Type Culture Collection (ATCC) (Manassus, Va.)). In some embodiments, a cell transfected with one or more vectors described herein is used to establish a new cell line comprising one or more vector-derived sequences. In some embodiments, a cell transiently transfected with the components of a CRISPR/Cas system as described herein (such as by transient transfection of one or more vectors, or transfection with RNA), and modified through the activity of a CRISPR complex, is used to establish a new cell line comprising cells containing the modification but lacking any other exogenous sequence.

[00192] As used herein, the term “non-human eukaryote” refers to a eukaryotic organism different than *Homo sapiens*. In preferred embodiments, such eukaryote is a non-human animal, such as non-human mammal, non-human primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, or arthropod, preferably a mammal, such as a rodent, in particular a mouse. The skilled person will appreciate that the eukaryotic cells which are transplanted or introduced in a non-human eukaryote according to the

methods as referred to herein are preferably derived from or originate from the same species as the eukaryote to which they are transplanted. For example, a mouse cell is transplanted in a mouse in certain embodiment according to the methods of the invention as described herein. In certain embodiments, the eukaryote is an immunocompromized eukaryote, i.e. a eukaryote in which the immune system is partially or completely shut down. For instance, immunocompromized mice may be used in the methods according to the invention as described herein. Examples of immunocompromized mice include, but are not limited to Nude mice, RAG -/- mice, SCID (severe compromised immunodeficiency) mice, SCID-Beige mice, NOD (non-obese diabetic)-SCID mice, NOG or NSG mice, etc.

[00193] As used herein “introducing a eukaryotic cell in a non-human eukaryote” generally refers to transplanting or grafting eukaryotic cells in a non-human eukaryote. In particular embodiments, the eukaryotic cell is introduced into a non-human eukaryote of a different species, such as, but not limited to a human cell into a rodent.

[00194] Typically, cells in a suitable carrier or medium are injected in the animal at a desired site, such as without limitation subdermal, intradermal, transdermal, intracavernous, intravitreal, intra-articular, transscleral, intracerebral, intrathecal, epidural, intramuscular, intravenous, intracardiac, intraosseous, intraperitoneal, etc. the amount and concentration of cells to be injected may vary, but typically the amount of injected cells will be between 10^2 to 10^{10} or between 10^2 to 10^9 , or between 10^3 to 10^{10} or between 10^3 to 10^9 , or between 10^4 to 10^{10} or between 10^4 to 10^9 , such as between 10^4 and 10^8 , or between 10^5 and 10^7 , e.g., about 1×10^5 , about 5×10^5 , about 1×10^6 , about 5×10^6 , about 1×10^7 , about 5×10^7 , about 1×10^8 , about 5×10^8 , about 1×10^9 , about 2×10^9 , about 3×10^9 , about 4×10^9 , about 5×10^9 , about 6×10^9 , about 7×10^9 , about 8×10^9 , about 9×10^9 or about 1×10^{10} cells per injection site. For example, such number of cells may particularly refer to the total number of cells to be administered to a non-human eukaryote, which administration may be suitably distributed over one or more doses (e.g., distributed over 2, 3, 4, 5, 6, 7, 8 9 or 10 or more doses) administered over one or more days (e.g., over 1, 2, 3, 4 or 5 or more days). Suitably, in a composition to be administered, cells may be present at a concentration between about 10^4 /ml to about 10^8 /ml, preferably between about

10^5 /ml and about 10^7 /ml, yet more preferably between about 1×10^6 /ml and about 1×10^7 /ml, such as, e.g., about 5×10^6 /ml.

[00195] In certain preferred embodiments, one RNA capable of guiding Cas to a genomic target locus may be used in the methods according to the invention. Introduction of a single RNA into a pool of cells will result in a variety of different genetic modifications at the same locus.

[00196] In certain preferred embodiments, a plurality of RNAs capable of guiding Cas to a genomic target locus may be used in the methods according to the invention. Advantageously, a genome-wide library, such as a knock-out library, may be used, such as for instance provided in Sanjana et al. (2014); Shalem et al. (2014); and as described in WO 2014/093701, which is incorporated herein by reference in its entirety. A genome wide library may comprise a plurality of CRISPR-Cas system guide RNAs that may comprise guide sequences that are capable of targeting a plurality of target sequences in a plurality of genomic loci, wherein said targeting preferably results in a knockout of gene function. This library may potentially comprise guide RNAs that target each and every gene in the genome of an organism. In some embodiments of the invention the organism or subject is a eukaryote (including mammal including human) or a non-human eukaryote or a non-human animal or a non-human mammal. In some embodiments, the organism or subject is a non-human animal, and may be an arthropod, for example, an insect, or may be a nematode. In some methods of the invention the organism or subject is a plant. In some methods of the invention the organism or subject is a mammal or a non-human mammal. A non-human mammal may be for example a rodent (preferably a mouse or a rat), an ungulate, or a primate. In some methods of the invention the organism or subject is algae, including microalgae, or is a fungus. A eukaryotic cell population can thus be established in which each cell has a different RNA of the library, such that the entire library or substantially the entire library is represented in the cell population. A gene knock-out cell library may thus be generated, in which each cell has a single gene knocked out. The skilled person will understand that the cell population may represent the library multiple times.

[00197] As used herein, "hybridization" or "hybridizing" 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, Hoogsteen 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 PGR, 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.

[00198] As used herein, "expression" refers to the process by which a polynucleotide is transcribed from a DNA template (such as into mRNA or other RNA transcript) and/or the process by which a transcribed mRNA is subsequently translated into peptides, polypeptides, or proteins. Transcripts and encoded polypeptides may be collectively referred to as "gene product." If the polynucleotide is derived from genomic DNA, expression may include splicing of the mRNA in a eukaryotic cell. As used herein "expression" of a gene or nucleic acid encompasses not only cellular gene expression, but also the transcription and translation of nucleic acid(s) in cloning systems and in any other context.

[00199] The terms "polypeptide", "peptide" and "protein" are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non amino acids. The terms also encompass an amino acid polymer that has been modified; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation, such as conjugation with a labeling component. As used herein the term "amino acid" includes natural and/or unnatural or synthetic amino acids, including glycine and both the D or L, optical isomers, and amino acid analogs and peptidomimetics.

[00200] The terms "subject," "individual," and "patient" are used interchangeably herein to refer to a vertebrate, preferably a mammal, more preferably a human. Mammals include, but are 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.

[00201] The terms "therapeutic agent", "therapeutic capable agent" or "treatment agent" are used interchangeably and refer to a molecule or compound that confers some beneficial effect

upon administration to a subject. The beneficial effect includes enablement of diagnostic determinations; amelioration of a disease, symptom, disorder, or pathological condition; reducing or preventing the onset of a disease, symptom, disorder or condition; and generally counteracting a disease, symptom, disorder or pathological condition.

[00202] As used herein, "treatment" or "treating," or "palliating" or "ameliorating" are used interchangeably. These terms refer to an approach for obtaining beneficial or desired results including but not limited to a therapeutic benefit and/or a prophylactic benefit. By therapeutic benefit is meant any therapeutically relevant improvement in or effect on one or more diseases, conditions, or symptoms under treatment. For prophylactic benefit, the compositions may be administered to a subject at risk of developing a particular disease, condition, or symptom, or to a subject reporting one or more of the physiological symptoms of a disease, even though the disease, condition, or symptom may not have yet been manifested.

[00203] As used herein, "diagnosis" or "diagnosing" in the context of tumor (metastasis) formation refers to the determination as to whether a subject has a tumor or a tumor metastasis. As used herein, "prognosis" or "prognosing" in the context of tumor (metastasis) refers to establishing or predicting the progression of a tumor or tumor metastasis or clinical outcome of a subject having a tumor or tumor metastasis. As used herein, "determining the likeliness of developing tumor metastasis" refers to establishing the chance or risk that a tumor will evolve to produce metastasis. These methods generally involve the determination of a genetic event identified as being involved in tumor (metastasis) formation and/or evolution according to the *in vivo* methods according to the invention as described herein. In certain embodiments, the invention relates to such diagnostic and/or prognostic and/or predictive methods wherein the genetic status or expression status of *Nf2* (neurofibromin 2), *Pten* (phosphatase and tensin homolog), *Trim72* (tripartite motif-containing protein 72), and/or *Ube2g2* (ubiquitin-conjugating enzyme E2 G2) is determined. These genes have been identified with the *in vivo* methods according to the invention as described herein as being involved in tumor metastasis, i.e. these genes are knocked out in tumor metastasis, in particular in lung metastasis.

[00204] As used herein, the terms "chimeric RNA", "chimeric guide RNA", "guide RNA", "single guide RNA" and "synthetic guide RNA" are used interchangeably and refer to the

polynucleotide sequence comprising the guide sequence, the tracr sequence and the tracr mate sequence. The term "guide sequence" refers to the about 20bp sequence within the guide RNA that specifies the target site and may be used interchangeably with the terms "guide" or "spacer". The term "tracr mate sequence" may also be used interchangeably with the term "direct repeat(s)".

[00205] With respect to general information on CRISPR-Cas Systems, components thereof, and delivery of such components, including methods, materials, delivery vehicles, vectors, particles, AAV, and making and using thereof, including as to amounts and formulations, all useful in the practice of the instant invention, reference is made to: US Patents Nos. 8,697,359, 8,771,945, 8,795,965, 8,865,406, 8,871,445, 8,889,356, 8,889,418, 8,895,308, 8,906,616, 8,932,814, 8,945,839, 8,993,233 and 8,999,641; US Patent Publication US 2015-0031134 (US App. Ser. No. 14/497,627), which is allowed; US Patent Publications US 2014-0310830 (US App. Ser. No. 14/105,031), US 2014-0287938 A1 (U.S. App. Ser. No. 14/213,991), US 2014-0273234 A1 (U.S. App. Ser. No. 14/293,674), US 2014-0273232 A1 (U.S. App. Ser. No. 14/290,575), US 2014-0273231 (U.S. App. Ser. No. 14/259,420), US 2014-0256046 A1 (U.S. App. Ser. No. 14/226,274), US 2014-0248702 A1 (U.S. App. Ser. No. 14/258,458), US 2014-0242700 A1 (U.S. App. Ser. No. 14/222,930), US 2014-0242699 A1 (U.S. App. Ser. No. 14/183,512), US 2014-0242664 A1 (U.S. App. Ser. No. 14/104,990), US 2014-0234972 A1 (U.S. App. Ser. No. 14/183,471), US 2014-0227787 A1 (U.S. App. Ser. No. 14/256,912), US 2014-0189896 A1 (U.S. App. Ser. No. 14/105,035), US 2014-0186958 (U.S. App. Ser. No. 14/105,017), US 2014-0186919 A1 (U.S. App. Ser. No. 14/104,977), US 2014-0186843 A1 (U.S. App. Ser. No. 14/104,900), US 2014-0179770 A1 (U.S. App. Ser. No. 14/104,837) and US 2014-0179006 A1 (U.S. App. Ser. No. 14/183,486), US 2014-0170753 (US App Ser No 14/183,429); US 2015-0184139 (U.S. App. Ser. No. 14/324,960); 14/054,414 European Patent Applications EP 2 771 468 (EP13818570.7), EP 2 764 103 (EP13824232.6), and EP 2 784 162 (EP14170383.5); and PCT Patent Publications WO 2014/093661 (PCT/US2013/074743), WO 2014/093694 (PCT/US2013/074790), WO 2014/093595 (PCT/US2013/074611), WO 2014/093718 (PCT/US2013/074825), WO 2014/093709 (PCT/US2013/074812), WO 2014/093622 (PCT/US2013/074667), WO 2014/093635 (PCT/US2013/074691), WO 2014/093655 (PCT/US2013/074736), WO 2014/093712 (PCT/US2013/074819), WO

2014/093701 (PCT/US2013/074800), WO 2014/018423 (PCT/US2013/051418), WO 2014/204723 (PCT/US2014/041790), WO 2014/204724 (PCT/US2014/041800), WO 2014/204725 (PCT/US2014/041803), WO 2014/204726 (PCT/US2014/041804), WO 2014/204727 (PCT/US2014/041806), WO 2014/204728 (PCT/US2014/041808), WO 2014/204729 (PCT/US2014/041809), WO 2015/089351 (PCT/US2014/069897), WO 2015/089354 (PCT/US2014/069902), WO 2015/089364 (PCT/US2014/069925), WO 2015/089427 (PCT/US2014/070068), WO 2015/089462 (PCT/US2014/070127), WO 2015/089419 (PCT/US2014/070057), WO 2015/089465 (PCT/US2014/070135), WO 2015/089486 (PCT/US2014/070175), PCT/US2015/051691, PCT/US2015/051830. Reference is also made to US provisional patent applications 61/758,468; 61/802,174; 61/806,375; 61/814,263; 61/819,803 and 61/828,130, filed on January 30, 2013; March 15, 2013; March 28, 2013; April 20, 2013; May 6, 2013 and May 28, 2013 respectively. Reference is also made to US provisional patent application 61/836,123, filed on June 17, 2013. Reference is additionally made to US provisional patent applications 61/835,931, 61/835,936, 61/835,973, 61/836,080, 61/836,101, and 61/836,127, each filed June 17, 2013. Further reference is made to US provisional patent applications 61/862,468 and 61/862,355 filed on August 5, 2013; 61/871,301 filed on August 28, 2013; 61/960,777 filed on September 25, 2013 and 61/961,980 filed on October 28, 2013. Reference is yet further made to: PCT/US2014/62558 filed October 28, 2014, and US Provisional Patent Applications Serial Nos.: 61/915,148, 61/915,150, 61/915,153, 61/915,203, 61/915,251, 61/915,301, 61/915,267, 61/915,260, and 61/915,397, each filed December 12, 2013; 61/757,972 and 61/768,959, filed on January 29, 2013 and February 25, 2013; 62/010,888 and 62/010,879, both filed June 11, 2014; 62/010,329, 62/010,439 and 62/010,441, each filed June 10, 2014; 61/939,228 and 61/939,242, each filed February 12, 2014; 61/980,012, filed April 15, 2014; 62/038,358, filed August 17, 2014; 62/055,484, 62/055,460 and 62/055,487, each filed September 25, 2014; and 62/069,243, filed October 27, 2014. Reference is made to PCT application designating, inter alia, the United States, application No. PCT/US14/41806, filed June 10, 2014. Reference is made to US provisional patent application 61/930,214 filed on January 22, 2014. Reference is made to PCT application designating, inter alia, the United States, application No. PCT/US14/41806, filed June 10, 2014.

[00206] Mention is also made of US application 62/180,709, 17-Jun-15, PROTECTED GUIDE RNAS (PGRNAS); US application 62/091,455, filed, 12-Dec-14, PROTECTED GUIDE RNAS (PGRNAS); US application 62/096,708, 24-Dec-14, PROTECTED GUIDE RNAS (PGRNAS); US applications 62/091,462, 12-Dec-14, 62/096,324, 23-Dec-14, 62/180,681, 17-Jun-2015, and 62/237,496, 5-Oct-2015, DEAD GUIDES FOR CRISPR TRANSCRIPTION FACTORS; US application 62/091,456, 12-Dec-14 and 62/180,692, 17-Jun-2015, ESCORTED AND FUNCTIONALIZED GUIDES FOR CRISPR-CAS SYSTEMS; US application 62/091,461, 12-Dec-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR GENOME EDITING AS TO HEMATOPOETIC STEM CELLS (HSCs); US application 62/094,903, 19-Dec-14, UNBIASED IDENTIFICATION OF DOUBLE-STRAND BREAKS AND GENOMIC REARRANGEMENT BY GENOME-WISE INSERT CAPTURE SEQUENCING; US application 62/096,761, 24-Dec-14, ENGINEERING OF SYSTEMS, METHODS AND OPTIMIZED ENZYME AND GUIDE SCAFFOLDS FOR SEQUENCE MANIPULATION; US application 62/098,059, 30-Dec-14, 62/181,641, 18-Jun-2015, and 62/181,667, 18-Jun-2015, RNA-TARGETING SYSTEM; US application 62/096,656, 24-Dec-14 and 62/181,151, 17-Jun-2015, CRISPR HAVING OR ASSOCIATED WITH DESTABILIZATION DOMAINS; US application 62/096,697, 24-Dec-14, CRISPR HAVING OR ASSOCIATED WITH AAV; US application 62/098,158, 30-Dec-14, ENGINEERED CRISPR COMPLEX INSERTIONAL TARGETING SYSTEMS; US application 62/151,052, 22-Apr-15, CELLULAR TARGETING FOR EXTRACELLULAR EXOSOMAL REPORTING; US application 62/054,490, 24-Sep-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR TARGETING DISORDERS AND DISEASES USING PARTICLE DELIVERY COMPONENTS; US application 61/939,154, 12-FEB-14, SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS; US application 62/055,484, 25-Sep-14, SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS; US application 62/087,537, 4-Dec-14, SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS; US application 62/054,651, 24-Sep-14, DELIVERY,

USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR MODELING COMPETITION OF MULTIPLE CANCER MUTATIONS IN VIVO; US application 62/067,886, 23-Oct-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR MODELING COMPETITION OF MULTIPLE CANCER MUTATIONS IN VIVO; US applications 62/054,675, 24-Sep-14 and 62/181,002, 17-Jun-2015, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS IN NEURONAL CELLS/TISSUES; US application 62/054,528, 24-Sep-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS IN IMMUNE DISEASES OR DISORDERS; US application 62/055,454, 25-Sep-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR TARGETING DISORDERS AND DISEASES USING CELL PENETRATION PEPTIDES (CPP); US application 62/055,460, 25-Sep-14, MULTIFUNCTIONAL-CRISPR COMPLEXES AND/OR OPTIMIZED ENZYME LINKED FUNCTIONAL-CRISPR COMPLEXES; US application 62/087,475, 4-Dec-14 and 62/181,690, 18-Jun-2015, FUNCTIONAL SCREENING WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS; US application 62/055,487, 25-Sep-14, FUNCTIONAL SCREENING WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS; US application 62/087,546, 4-Dec-14 and 62/181,687, 18-Jun-2015, MULTIFUNCTIONAL CRISPR COMPLEXES AND/OR OPTIMIZED ENZYME LINKED FUNCTIONAL-CRISPR COMPLEXES; and US application 62/098,285, 30-Dec-14, CRISPR MEDIATED IN VIVO MODELING AND GENETIC SCREENING OF TUMOR GROWTH AND METASTASIS.

[00207] Mention is made of US applications 62/181,659, 18-Jun-2015 and 62/207,318, 19-Aug-2015, ENGINEERING AND OPTIMIZATION OF SYSTEMS, METHODS, ENZYME AND GUIDE SCAFFOLDS OF CAS9 ORTHOLOGS AND VARIANTS FOR SEQUENCE MANIPULATION. Mention is made of US applications 62/181,663, 18-Jun-2015 and 62/245,264, 22-Oct-2015, NOVEL CRISPR ENZYMES AND SYSTEMS, US applications 62/181,675, 18-Jun-2015, and Attorney Docket No. 46783.01.2128, filed 22-Oct-2015, NOVEL CRISPR ENZYMES AND SYSTEMS, US application 62/232,067, 24-Sep-2015, US application 62/205,733, 16-Aug-2015, US application 62/201,542, 5-Aug-2015, US application 62/193,507,

16-Jul-2015, and US application 62/181,739, 18-Jun-2015, each entitled NOVEL CRISPR ENZYMES AND SYSTEMS and of US application 62/245,270, 22-Oct-2015, NOVEL CRISPR ENZYMES AND SYSTEMS. Mention is also made of US application 61/939,256, 12-Feb-2014, and WO 2015/089473 (PCT/US2014/070152), 12-Dec-2014, each entitled ENGINEERING OF SYSTEMS, METHODS AND OPTIMIZED GUIDE COMPOSITIONS WITH NEW ARCHITECTURES FOR SEQUENCE MANIPULATION. Mention is also made of PCT/US2015/045504, 15-Aug-2015, US application 62/180,699, 17-Jun-2015, and US application 62/038,358, 17-Aug-2014, each entitled GENOME EDITING USING CAS9 NICKASES.

[00208] Each of these patents, patent publications, and applications, and all documents cited therein or during their prosecution (“appln cited documents”) and all documents cited or referenced in the appln cited documents, together with any instructions, descriptions, product specifications, and product sheets for any products mentioned therein or in any document therein and incorporated by reference herein, are hereby incorporated herein by reference, and may be employed in the practice of the invention. All documents (e.g., these patents, patent publications and applications and the appln cited documents) are incorporated herein by reference to the same extent as if each individual document was specifically and individually indicated to be incorporated by reference.

[00209] Also with respect to general information on CRISPR-Cas Systems, mention is made of the following (also hereby incorporated herein by reference):

- *Multiplex genome engineering using CRISPR/Cas systems.* Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., & Zhang, F. *Science* Feb 15;339(6121):819-23 (2013);
- *RNA-guided editing of bacterial genomes using CRISPR-Cas systems.* Jiang W., Bikard D., Cox D., Zhang F, Marraffini LA. *Nat Biotechnol* Mar;31(3):233-9 (2013);
- *One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering.* Wang H., Yang H., Shivalila CS., Dawlaty MM., Cheng AW., Zhang F., Jaenisch R. *Cell* May 9;153(4):910-8 (2013);
- *Optical control of mammalian endogenous transcription and epigenetic states.* Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, Platt RJ, Scott DA, Church GM, Zhang F. *Nature*. 2013 Aug 22;500(7463):472-6. doi: 10.1038/Nature12466. Epub 2013 Aug 23;

- *Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity.* Ran, FA., Hsu, PD., Lin, CY., Gootenberg, JS., Konermann, S., Trevino, AE., Scott, DA., Inoue, A., Matoba, S., Zhang, Y., & Zhang, F. *Cell* Aug 28. pii: S0092-8674(13)01015-5. (2013);
- *DNA targeting specificity of RNA-guided Cas9 nucleases.* Hsu, P., Scott, D., Weinstein, J., Ran, FA., Konermann, S., Agarwala, V., Li, Y., Fine, E., Wu, X., Shalem, O., Cradick, TJ., Marraffini, LA., Bao, G., & Zhang, F. *Nat Biotechnol* doi:10.1038/nbt.2647 (2013);
- *Genome engineering using the CRISPR-Cas9 system.* Ran, FA., Hsu, PD., Wright, J., Agarwala, V., Scott, DA., Zhang, F. *Nature Protocols* Nov;8(11):2281-308. (2013);
- *Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells.* Shalem, O., Sanjana, NE., Hartenian, E., Shi, X., Scott, DA., Mikkelsen, T., Heckl, D., Ebert, BL., Root, DE., Doench, JG., Zhang, F. *Science* Dec 12. (2013). [Epub ahead of print];
- *Crystal structure of cas9 in complex with guide RNA and target DNA.* Nishimasu, H., Ran, FA., Hsu, PD., Konermann, S., Shehata, SI., Dohmae, N., Ishitani, R., Zhang, F., Nureki, O. *Cell* Feb 27. (2014). 156(5):935-49;
- *Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells.* Wu X., Scott DA., Kriz AJ., Chiu AC., Hsu PD., Dadon DB., Cheng AW., Trevino AE., Konermann S., Chen S., Jaenisch R., Zhang F., Sharp PA. *Nat Biotechnol.* (2014) Apr 20. doi: 10.1038/nbt.2889,
- *CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling,* Platt et al., *Cell* 159(2): 440-455 (2014) DOI: 10.1016/j.cell.2014.09.014,
- *Development and Applications of CRISPR-Cas9 for Genome Engineering,* Hsu et al, *Cell* 157, 1262-1278 (June 5, 2014) (Hsu 2014),
- *Genetic screens in human cells using the CRISPR/Cas9 system,* Wang et al., *Science*. 2014 January 3; 343(6166): 80–84. doi:10.1126/science.1246981,
- *Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation,* Doench et al., *Nature Biotechnology* published online 3 September 2014; doi:10.1038/nbt.3026, and
- *In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9,* Swiech et al, *Nature Biotechnology* ; published online 19 October 2014; doi:10.1038/nbt.3055.
- *Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex,* Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu

- PD, Habib N, Gootenberg JS, Nishimasu H, Nureki O, Zhang F., Nature. Jan 29;517(7536):583-8 (2015).
- *A split-Cas9 architecture for inducible genome editing and transcription modulation*, Zetsche B, Volz SE, Zhang F., (published online 02 February 2015) Nat Biotechnol. Feb;33(2):139-42 (2015);
 - *Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis*, Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, Scott DA, Song J, Pan JQ, Weissleder R, Lee H, Zhang F, Sharp PA. Cell 160, 1246–1260, March 12, 2015 (multiplex screen in mouse), and
 - *In vivo genome editing using Staphylococcus aureus Cas9*, Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV, Sharp PA, Zhang F., (published online 01 April 2015), Nature. Apr 9;520(7546):186-91 (2015).
 - *High-throughput functional genomics using CRISPR-Cas9*, Shalem et al., Nature Reviews Genetics 16, 299-311 (May 2015).
 - *Sequence determinants of improved CRISPR sgRNA design*, Xu et al., Genome Research 25, 1147-1157 (August 2015).
 - *A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks*, Parnas et al., Cell 162, 675-686 (July 30, 2015).
 - *CRISPR/Cas9 cleavage of viral DNA efficiently suppresses hepatitis B virus*, Ramanan et al., Scientific Reports 5:10833. doi: 10.1038/srep10833 (June 2, 2015).
 - *Crystal Structure of Staphylococcus aureus Cas9*, Nishimasu et al., Cell 162, 1113-1126 (Aug. 27, 2015).
 - *BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis*, Canver et al., Nature 527(7577):192-7 (Nov. 12, 2015) doi: 10.1038/nature15521. Epub 2015 Sep 16.
 - *Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System*, Zetsche et al., Cell 163, 759-71 (Sep 25, 2015).

- *Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems*, Shmakov et al., Molecular Cell, 60(3), 385–397 doi: 10.1016/j.molcel.2015.10.008 Epub October 22, 2015.
- *Rationally engineered Cas9 nucleases with improved specificity*, Slaymaker et al., Science 2015 Dec 1. pii: aad5227. [Epub ahead of print]

each of which is incorporated herein by reference, and discussed briefly below:

- Cong *et al.* engineered type II CRISPR/Cas systems for use in eukaryotic cells based on both *Streptococcus thermophilus* Cas9 and also *Streptococcus pyogenes* Cas9 and demonstrated that Cas9 nucleases can be directed by short RNAs to induce precise cleavage of DNA in human and mouse cells. Their study further showed that Cas9 as converted into a nicking enzyme can be used to facilitate homology-directed repair in eukaryotic cells with minimal mutagenic activity. Additionally, their study demonstrated that multiple guide sequences can be encoded into a single CRISPR array to enable simultaneous editing of several at endogenous genomic loci sites within the mammalian genome, demonstrating easy programmability and wide applicability of the RNA-guided nuclease technology. This ability to use RNA to program sequence specific DNA cleavage in cells defined a new class of genome engineering tools. These studies further showed that other CRISPR loci are likely to be transplantable into mammalian cells and can also mediate mammalian genome cleavage. Importantly, it can be envisaged that several aspects of the CRISPR/Cas system can be further improved to increase its efficiency and versatility.
- Jiang *et al.* used the clustered, regularly interspaced, short palindromic repeats (CRISPR)–associated Cas9 endonuclease complexed with dual-RNAs to introduce precise mutations in the genomes of *Streptococcus pneumoniae* and *Escherichia coli*. The approach relied on dual-RNA:Cas9-directed cleavage at the targeted genomic site to kill unmutated cells and circumvents the need for selectable markers or counter-selection systems. The study reported reprogramming dual-RNA:Cas9 specificity by changing the sequence of short CRISPR RNA (crRNA) to make single- and multinucleotide changes carried on editing templates. The study showed that simultaneous use of two crRNAs

enabled multiplex mutagenesis. Furthermore, when the approach was used in combination with recombineering, in *S. pneumoniae*, nearly 100% of cells that were recovered using the described approach contained the desired mutation, and in *E. coli*, 65% that were recovered contained the mutation.

- Wang *et al.* (2013) used the CRISPR/Cas system for the one-step generation of mice carrying mutations in multiple genes which were traditionally generated in multiple steps by sequential recombination in embryonic stem cells and/or time-consuming intercrossing of mice with a single mutation. The CRISPR/Cas system will greatly accelerate the *in vivo* study of functionally redundant genes and of epistatic gene interactions.
- Konermann *et al.* addressed the need in the art for versatile and robust technologies that enable optical and chemical modulation of DNA-binding domains based CRISPR Cas9 enzyme and also Transcriptional Activator Like Effectors.
- Ran *et al.* (2013-A) described an approach that combined a Cas9 nickase mutant with paired guide RNAs to introduce targeted double-strand breaks. This addresses the issue of the Cas9 nuclease from the microbial CRISPR-Cas system being targeted to specific genomic loci by a guide sequence, which can tolerate certain mismatches to the DNA target and thereby promote undesired off-target mutagenesis. Because individual nicks in the genome are repaired with high fidelity, simultaneous nicking via appropriately offset guide RNAs is required for double-stranded breaks and extends the number of specifically recognized bases for target cleavage. The authors demonstrated that using paired nicking can reduce off-target activity by 50- to 1,500-fold in cell lines and to facilitate gene knockout in mouse zygotes without sacrificing on-target cleavage efficiency. This versatile strategy enables a wide variety of genome editing applications that require high specificity.
- Hsu *et al.* (2013) characterized SpCas9 targeting specificity in human cells to inform the selection of target sites and avoid off-target effects. The study evaluated >700 guide RNA variants and SpCas9-induced indel mutation levels at >100 predicted genomic off-target loci in 293T and 293FT cells. The authors that SpCas9 tolerates mismatches between guide RNA and target DNA at different positions in a sequence-dependent manner,

sensitive to the number, position and distribution of mismatches. The authors further showed that SpCas9-mediated cleavage is unaffected by DNA methylation and that the dosage of SpCas9 and sgRNA can be titrated to minimize off-target modification. Additionally, to facilitate mammalian genome engineering applications, the authors reported providing a web-based software tool to guide the selection and validation of target sequences as well as off-target analyses.

- Ran *et al.* (2013-B) described a set of tools for Cas9-mediated genome editing *via* non-homologous end joining (NHEJ) or homology-directed repair (HDR) in mammalian cells, as well as generation of modified cell lines for downstream functional studies. To minimize off-target cleavage, the authors further described a double-nicking strategy using the Cas9 nickase mutant with paired guide RNAs. The protocol provided by the authors experimentally derived guidelines for the selection of target sites, evaluation of cleavage efficiency and analysis of off-target activity. The studies showed that beginning with target design, gene modifications can be achieved within as little as 1–2 weeks, and modified clonal cell lines can be derived within 2–3 weeks.
- Shalem *et al.* described a new way to interrogate gene function on a genome-wide scale. Their studies showed that delivery of a genome-scale CRISPR-Cas9 knockout (GeCKO) library targeted 18,080 genes with 64,751 unique guide sequences enabled both negative and positive selection screening in human cells. First, the authors showed use of the GeCKO library to identify genes essential for cell viability in cancer and pluripotent stem cells. Next, in a melanoma model, the authors screened for genes whose loss is involved in resistance to vemurafenib, a therapeutic that inhibits mutant protein kinase BRAF. Their studies showed that the highest-ranking candidates included previously validated genes NF1 and MED12 as well as novel hits NF2, CUL3, TADA2B, and TADA1. The authors observed a high level of consistency between independent guide RNAs targeting the same gene and a high rate of hit confirmation, and thus demonstrated the promise of genome-scale screening with Cas9.
- Nishimasu *et al.* reported the crystal structure of *Streptococcus pyogenes* Cas9 in complex with sgRNA and its target DNA at 2.5 Å resolution. The structure revealed a bilobed architecture composed of target recognition and nuclease lobes, accommodating

the sgRNA:DNA heteroduplex in a positively charged groove at their interface. Whereas the recognition lobe is essential for binding sgRNA and DNA, the nuclease lobe contains the HNH and RuvC nuclease domains, which are properly positioned for cleavage of the complementary and non-complementary strands of the target DNA, respectively. The nuclease lobe also contains a carboxyl-terminal domain responsible for the interaction with the protospacer adjacent motif (PAM). This high-resolution structure and accompanying functional analyses have revealed the molecular mechanism of RNA-guided DNA targeting by Cas9, thus paving the way for the rational design of new, versatile genome-editing technologies.

- Wu *et al.* mapped genome-wide binding sites of a catalytically inactive Cas9 (dCas9) from *Streptococcus pyogenes* loaded with single guide RNAs (sgRNAs) in mouse embryonic stem cells (mESCs). The authors showed that each of the four sgRNAs tested targets dCas9 to between tens and thousands of genomic sites, frequently characterized by a 5-nucleotide seed region in the sgRNA and an NGG protospacer adjacent motif (PAM). Chromatin inaccessibility decreases dCas9 binding to other sites with matching seed sequences; thus 70% of off-target sites are associated with genes. The authors showed that targeted sequencing of 295 dCas9 binding sites in mESCs transfected with catalytically active Cas9 identified only one site mutated above background levels. The authors proposed a two-state model for Cas9 binding and cleavage, in which a seed match triggers binding but extensive pairing with target DNA is required for cleavage.
- Platt *et al.* established a Cre-dependent Cas9 knockin mouse. The authors demonstrated *in vivo* as well as *ex vivo* genome editing using adeno-associated virus (AAV)-, lentivirus-, or particle-mediated delivery of guide RNA in neurons, immune cells, and endothelial cells.
- Hsu *et al.* (2014) is a review article that discusses generally CRISPR-Cas9 history from yogurt to genome editing, including genetic screening of cells.
- Wang *et al.* (2014) relates to a pooled, loss-of-function genetic screening approach suitable for both positive and negative selection that uses a genome-scale lentiviral single guide RNA (sgRNA) library.

- Doench *et al.* created a pool of sgRNAs, tiling across all possible target sites of a panel of six endogenous mouse and three endogenous human genes and quantitatively assessed their ability to produce null alleles of their target gene by antibody staining and flow cytometry. The authors showed that optimization of the PAM improved activity and also provided an on-line tool for designing sgRNAs.
- Swiech *et al.* demonstrate that AAV-mediated SpCas9 genome editing can enable reverse genetic studies of gene function in the brain.
- Konermann *et al.* (2015) discusses the ability to attach multiple effector domains, e.g., transcriptional activator, functional and epigenomic regulators at appropriate positions on the guide such as stem or tetraloop with and without linkers.
- Zetsche *et al.* demonstrates that the Cas9 enzyme can be split into two and hence the assembly of Cas9 for activation can be controlled.
- Chen *et al.* relates to multiplex screening by demonstrating that a genome-wide *in vivo* CRISPR-Cas9 screen in mice reveals genes regulating lung metastasis.
- Ran *et al.* (2015) relates to SaCas9 and its ability to edit genomes and demonstrates that one cannot extrapolate from biochemical assays. Shalem *et al.* (2015) described ways in which catalytically inactive Cas9 (dCas9) fusions are used to synthetically repress (CRISPRi) or activate (CRISPRa) expression, showing advances using Cas9 for genome-scale screens, including arrayed and pooled screens, knockout approaches that inactivate genomic loci and strategies that modulate transcriptional activity.
- Shalem *et al.* (2015) described ways in which catalytically inactive Cas9 (dCas9) fusions are used to synthetically repress (CRISPRi) or activate (CRISPRa) expression, showing advances using Cas9 for genome-scale screens, including arrayed and pooled screens, knockout approaches that inactivate genomic loci and strategies that modulate transcriptional activity.
- Xu *et al.* (2015) assessed the DNA sequence features that contribute to single guide RNA (sgRNA) efficiency in CRISPR-based screens. The authors explored efficiency of CRISPR/Cas9 knockout and nucleotide preference at the cleavage site. The authors also found that the sequence preference for CRISPRi/a is substantially different from that for CRISPR/Cas9 knockout.

- Parnas *et al.* (2015) introduced genome-wide pooled CRISPR-Cas9 libraries into dendritic cells (DCs) to identify genes that control the induction of tumor necrosis factor (Tnf) by bacterial lipopolysaccharide (LPS). Known regulators of Tlr4 signaling and previously unknown candidates were identified and classified into three functional modules with distinct effects on the canonical responses to LPS.
- Ramanan *et al.* (2015) demonstrated cleavage of viral episomal DNA (cccDNA) in infected cells. The HBV genome exists in the nuclei of infected hepatocytes as a 3.2kb double-stranded episomal DNA species called covalently closed circular DNA (cccDNA), which is a key component in the HBV life cycle whose replication is not inhibited by current therapies. The authors showed that sgRNAs specifically targeting highly conserved regions of HBV robustly suppresses viral replication and depleted cccDNA.
- Nishimasu *et al.* (2015) reported the crystal structures of SaCas9 in complex with a single guide RNA (sgRNA) and its double-stranded DNA targets, containing the 5'-TTGAAT-3' PAM and the 5'-TTGGGT-3' PAM. A structural comparison of SaCas9 with SpCas9 highlighted both structural conservation and divergence, explaining their distinct PAM specificities and orthologous sgRNA recognition.
- Canver *et al.* (2015) demonstrated a CRISPR-Cas9-based functional investigation of non-coding genomic elements. The authors we developed pooled CRISPR-Cas9 guide RNA libraries to perform *in situ* saturating mutagenesis of the human and mouse BCL11A enhancers which revealed critical features of the enhancers.
- Zetsche *et al.* (2015) reported characterization of Cpf1, a class 2 CRISPR nuclease from *Francisella novicida* U112 having features distinct from Cas9. Cpf1 is a single RNA-guided endonuclease lacking tracrRNA, utilizes a T-rich protospacer-adjacent motif, and cleaves DNA via a staggered DNA double-stranded break.
- Shmakov *et al.* (2015) reported three distinct Class 2 CRISPR-Cas systems. Two system CRISPR enzymes (C2c1 and C2c3) contain RuvC-like endonuclease domains distantly related to Cpf1. Unlike Cpf1, C2c1 depends on both crRNA and tracrRNA for DNA cleavage. The third enzyme (C2c2) contains two predicted HEPN RNase domains and is tracrRNA independent.

- Slaymaker et al (2015) reported the use of structure-guided protein engineering to improve the specificity of *Streptococcus pyogenes* Cas9 (SpCas9). The authors developed "enhanced specificity" SpCas9 (eSpCas9) variants which maintained robust on-target cleavage with reduced off-target effects.

[00210] Mention is also made of Tsai et al, "Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing," *Nature Biotechnology* 32(6): 569-77 (2014) which is not believed to be prior art to the instant invention or application, but which may be considered in the practice of the instant invention. Mention is also made of Konermann et al., "Genome-scale transcription activation by an engineered CRISPR-Cas9 complex," doi:10.1038/nature14136, incorporated herein by reference.

[00211] In general, the CRISPR-Cas or CRISPR system is as used in the foregoing documents, such as WO 2014/093622 (PCT/US2013/074667) and refers collectively to transcripts and other elements involved in the expression of or directing the activity of CRISPR-associated ("Cas") enzyme, including sequences encoding or delivering a Cas enzyme, a tracr (trans-activating CRISPR) sequence (e.g. tracrRNA or an active partial tracrRNA), a tracr-mate sequence (encompassing a "direct repeat" and a tracrRNA-processed partial direct repeat in the context of an endogenous CRISPR system), a guide sequence (also referred to as a "spacer" in the context of an endogenous CRISPR system), or "RNA(s)" as that term is herein used (e.g., RNA(s) to guide Cas, such as Cas9, e.g. CRISPR RNA and transactivating (tracr) RNA or a single guide RNA (sgRNA) (chimeric RNA)) or other sequences and transcripts from a CRISPR locus. 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. 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, direct repeats may be identified *in silico* by searching for repetitive motifs that fulfill any or all of the following criteria: 1. found in a 2Kb window of genomic sequence flanking the type II CRISPR locus; 2. span from 20 to 50 bp; and 3.

interspaced by 20 to 50 bp. In some embodiments, 2 of these criteria may be used, for instance 1 and 2, 2 and 3, or 1 and 3. In some embodiments, all 3 criteria may be used.

[00212] In embodiments of the invention the terms guide sequence and guide RNA, i.e. RNA capable of guiding Cas to a target genomic locus, are used interchangeably as in foregoing cited documents such as WO 2014/093622 (PCT/US2013/074667). In general, a guide sequence is any polynucleotide sequence having sufficient complementarity with 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 between a guide sequence and its corresponding target sequence, when optimally aligned using a suitable alignment algorithm, 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, non-limiting example of which include the Smith-Waterman algorithm, the Needleman-Wunsch algorithm, algorithms based on the Burrows-Wheeler Transform (e.g. the Burrows Wheeler Aligner), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies; available at www.novocraft.com), ELAND (Illumina, San Diego, CA), SOAP (available at soap.genomics.org.cn), and Maq (available at maq.sourceforge.net). 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. Preferably the guide sequence is 10-30 nucleotides long. The ability of a guide sequence to direct sequence-specific binding of a CRISPR complex to a target sequence may be assessed by any suitable assay. For example, the components of a CRISPR system sufficient to form a CRISPR complex, including the guide sequence to be tested, may be provided to a host cell having the corresponding target sequence, such as by transfection with vectors encoding the components of the CRISPR sequence, followed by an assessment of preferential cleavage within the target sequence, such as by Surveyor assay as described herein. Similarly, cleavage of a target polynucleotide sequence may be evaluated in a test tube by providing the target sequence, components of a CRISPR complex, including the guide sequence to be tested and a control guide sequence different from the test guide sequence, and comparing

binding or rate of cleavage at the target sequence between the test and control guide sequence reactions. Other assays are possible, and will occur to those skilled in the art.

[00213] A guide sequence, i.e. an RNA capable of guiding Cas to a genomic target locus, 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. For example, for the *S. pyogenes* Cas9, a unique target sequence in a genome may include a Cas9 target site of the form MMMMMMMMNNNNNNNNNNNNNNXGG where NNNNNNNNNNNNNXGG (N is A, G, T, or C; and X can be anything) has a single occurrence in the genome. A unique target sequence in a genome may include an *S. pyogenes* Cas9 target site of the form MMMMMMMMNNNNNNNNNNNNNNXGG where NNNNNNNNNNNNNXGG (N is A, G, T, or C; and X can be anything) has a single occurrence in the genome. For the *S. thermophilus* CRISPR1 Cas9, a unique target sequence in a genome may include a Cas9 target site of the form MMMMMMMMNNNNNNNNNNNNNNXXAGAAW where NNNNNNNNNNNNNXXAGAAW (N is A, G, T, or C; X can be anything; and W is A or T) has a single occurrence in the genome. A unique target sequence in a genome may include an *S. thermophilus* CRISPR1 Cas9 target site of the form MMMMMMMMNNNNNNNNNNNNNNXXAGAAW where NNNNNNNNNNNNNXXAGAAW (N is A, G, T, or C; X can be anything; and W is A or T) has a single occurrence in the genome. For the *S. pyogenes* Cas9, a unique target sequence in a genome may include a Cas9 target site of the form MMMMMMMMNNNNNNNNNNNNNNXGGXG where NNNNNNNNNNNNNXGGXG (N is A, G, T, or C; and X can be anything) has a single occurrence in the genome. A unique target sequence in a genome may include an *S. pyogenes* Cas9 target site of the form MMMMMMMMNNNNNNNNNNNNNNXGGXG where NNNNNNNNNNNNNXGGXG (N is A, G, T, or C; and X can be anything) has a single occurrence in the genome. In each of these sequences "M" may be A, G, T, or C, and need not be considered in identifying a sequence as unique. In some embodiments, a guide sequence is selected to reduce the degree secondary structure within the guide sequence. In some embodiments, about or less than about 75%, 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 1%, or fewer of the nucleotides of the guide sequence participate in self-complementary base pairing when optimally folded. Optimal folding may be determined by any suitable polynucleotide folding algorithm. Some programs are based on

calculating the minimal Gibbs free energy. An example of one such algorithm is mFold, as described by Zuker and Stiegler (Nucleic Acids Res. 9 (1981), 133-148). Another example folding algorithm is the online webserver RNAfold, developed at Institute for Theoretical Chemistry at the University of Vienna, using the centroid structure prediction algorithm (see e.g. A.R. Gruber et al., 2008, Cell 106(1): 23-24; and PA Carr and GM Church, 2009, Nature Biotechnology 27(12): 1151-62).

[00214] In general, a tracr mate sequence includes any sequence that has sufficient complementarity with a tracr sequence to promote one or more of: (1) excision of a guide sequence flanked by tracr mate sequences in a cell containing the corresponding tracr sequence; and (2) formation of a CRISPR complex at a target sequence, wherein the CRISPR complex comprises the tracr mate sequence hybridized to the tracr sequence. In general, degree of complementarity is with reference to the optimal alignment of the tracr mate sequence and tracr sequence, along the length of the shorter of the two sequences. Optimal alignment may be determined by any suitable alignment algorithm, and may further account for secondary structures, such as self-complementarity within either the tracr sequence or tracr mate sequence. In some embodiments, the degree of complementarity between the tracr sequence and tracr mate sequence along the length of the shorter of the two when optimally aligned is about or more than about 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97.5%, 99%, or higher. In some embodiments, the tracr sequence is about or more than about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, or more nucleotides in length. In some embodiments, the tracr sequence and tracr mate sequence are contained within a single transcript, such that hybridization between the two produces a transcript having a secondary structure, such as a hairpin. In an embodiment of the invention, the transcript or transcribed polynucleotide sequence has at least two or more hairpins. In preferred embodiments, the transcript has two, three, four or five hairpins. In a further embodiment of the invention, the transcript has at most five hairpins. In a hairpin structure the portion of the sequence 5' of the final "N" and upstream of the loop corresponds to the tracr mate sequence, and the portion of the sequence 3' of the loop corresponds to the tracr sequence. Further non-limiting examples of single polynucleotides comprising a guide sequence, a tracr mate sequence, and a tracr sequence are as follows (listed 5' to 3'), where "N" represents a base of a guide sequence, the first block of lower case letters

represent the tracr mate sequence, and the second block of lower case letters represent the tracr sequence, and the final poly-T sequence represents the transcription terminator: (1) NNNNNNNNNNNNNNNNNNNNNggtttgtactctcaagatttaGAAAtaaatctgcagaagctacaaagataaggcttcatgccgaaatcaacaccctgtcattttatggcagggtgtttcgtatttaaTTTTTT (SEQ ID NO: X); (2) NNNNNNNNNNNNNNNNNNNNNggtttgtactctcaGAAAtgcagaagctacaaagataaggcttcatgccgaaatcaacaccctgtcattttatggcagggtgtttcgtatttaaTTTTTT; (SEQ ID NO: X) (3) NNNNNNNNNNNNNNNNNNNNNggtttgtactctcaGAAAtgcagaagctacaaagataaggcttcatgccgaaatcaacaccctgtcattttatggcagggtgtTTTTTT; (SEQ ID NO: X) (4) NNNNNNNNNNNNNNNNNNNNNggttttagagctaGAAAtagcaagttaaaataaggctagtcggttatcaactgaaaaagtggcaccgagtcggtgcTTTTTT; (SEQ ID NO: X) (5) NNNNNNNNNNNNNNNNNNNNNggttttagagctaGAAATAGcaagttaaaataaggctagtcggttatcaactgaaaagtgtTTTTTT; and (SEQ ID NO: X) (6) NNNNNNNNNNNNNNNNNNNNNggttttagagctagAAATAGcaagttaaaataaggctagtcggttatcaTTTTTTT (SEQ ID NO: X). In some embodiments, sequences (1) to (3) are used in combination with Cas9 from *S. thermophilus* CRISPR1. In some embodiments, sequences (4) to (6) are used in combination with Cas9 from *S. pyogenes*. In some embodiments, the tracr sequence is a separate transcript from a transcript comprising the tracr mate sequence.

[00215] In some embodiments, candidate tracrRNA may be subsequently predicted by sequences that fulfill any or all of the following criteria: 1. sequence homology to direct repeats (motif search in Geneious with up to 18-bp mismatches); 2. presence of a predicted Rho-independent transcriptional terminator in direction of transcription; and 3. stable hairpin secondary structure between tracrRNA and direct repeat. In some embodiments, 2 of these criteria may be used, for instance 1 and 2, 2 and 3, or 1 and 3. In some embodiments, all 3 criteria may be used.

[00216] In some embodiments, chimeric synthetic guide RNAs (sgRNAs) designs may incorporate at least 12 bp of duplex structure between the direct repeat and tracrRNA.

[00217] The RNAs to guide Cas, such as Cas9, can comprise CRISPR RNA and transactivating (tracr) RNA. The tracr mate and the tracr sequence can be connected to form a transactivating (tracr) sequence. The tracr mate and the tracr sequence can optionally be

designed to form a single guide RNA (sgRNA). Indeed, it is advantageous that the RNAs to guide Cas can comprise chimeric single guide RNA (sgRNA). The tracr sequence and tracr mate sequence along the length of the shorter of the two when optimally aligned can be about or more than about 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97.5%, 99%, or higher. The tracr sequence can be about or more than about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, or more nucleotides in length. The degree of complementarity between a guide sequence and its corresponding target sequence can be about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or 100%. A guide or RNA or sgRNA can be 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. A guide or RNA or sgRNA can be less than about 75, 50, 45, 40, 35, 30, 25, 20, 15, 12, or fewer nucleotides in length.

[00218] In particularly preferred embodiments according to the invention, the RNA capable of guiding Cas to a genomic target locus may comprise (1) a guide suequence capable of hybridizing to a genomic target locus in the eukaryotic cell; (2) a tracr sequence; and (3) a tracr mate sequence. All (1) to (3) may reside in a single RNA, i.e. an sgRNA (arranged in a 5' to 3' orientation), or the tracr RNA may be a different RNA than the RNA containing the guide and tracr sequence. The tracr hybridizes to the tracr mate sequence and directs the CRISPR/Cas complex to the target sequence.

[00219] The methods according to the invention as described herein comprehend inducing one or more mutations in a eukaryotic cell (in vitro, i.e. in an isolated eukaryotic cell) as herein discussed comprising delivering to cell a vector as herein discussed. The mutation(s) can include the introduction, deletion, or substitution of one or more nucleotides at each target sequence of cell(s) via the guide(s) or RNA(s) or sgRNA(s) resulting in loss-of-function and/or gain of function of the target(s). The mutations can include the introduction, deletion, or substitution of 1-75 nucleotides at each target sequence of said cell(s) via the guide(s) or or RNA(s) sgRNA(s) resulting in loss-of-function or gain of function of the target(s). The mutations can include the introduction, deletion, or substitution of 1, 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, or 75 nucleotides at each target sequence of said cell(s) via the guide(s) or RNA(s) or sgRNA(s) resulting in loss-of-function or gain of function of the

target(s). The mutations can include the introduction, deletion, or substitution of 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, or 75 nucleotides at each target sequence of said cell(s) via the guide(s) or RNA(s) or sgRNA(s) resulting in loss-of-function or gain of function of the target(s). The mutations include the introduction, deletion, or substitution of 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, or 75 nucleotides at each target sequence of said cell(s) via the guide(s) or RNA(s) or sgRNA(s) resulting in loss-of-function or gain of function of the target(s). The mutations can include the introduction, deletion, or substitution of 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, or 75 nucleotides at each target sequence of said cell(s) via the guide(s) or RNA(s) or sgRNA(s) resulting in loss-of-function or gain of function of the target(s). The mutations can include the introduction, deletion, or substitution of 40, 45, 50, 75, 100, 200, 300, 400 or 500 nucleotides at each target sequence of said cell(s) via the guide(s) or RNA(s) or sgRNA(s) resulting in loss-of-function or gain of function of the target. The method or mutation(s) thereof, when cell treated as such are implanted in a eukaryote can induce a cancer which may or may not metastasize, e.g., Lung cancer, Lung adenocarcinoma, Lung squamous cell carcinoma, Acute myeloid leukemia, Basal cell carcinoma (skin), Bladder cancer, Breast cancer, Carcinoid, Chronic lymphocytic leukemia, Colorectal cancer, Diffuse large B-cell lymphoma, Endometrial, Esophageal cancer, Esophageal adenocarcinoma, Glioblastoma multiforme, Glioma, Head and neck cancer, Kidney clear cell cancer, Medulloblastoma, Melanoma, Multiple myeloma, Nasopharyngeal, Neuroblastoma, Ovarian cancer, Prostate cancer, Rhabdoid tumor, Testicular germ cell tumor, Thyroid cancer, or Urinary bladder cancer. The methods as described herein, may thus also be used to identify genes or allow identification of genes which are involved in cancer and/or metastasis.

[00220] In certain embodiments, each RNA (e.g., sgRNA) used in the methods as described herein is specific to a different target sequence. Each different target sequence can be associated with or correlated to a form of cancer. Each RNA or sgRNA can be specific to a different target sequence but these RNA(s) or sgRNA(s) target specific gene sequences associated with or correlated to cancer, advantageously a particular type or form of cancer; for instance each RNA or sgRNA can be specific to a different target sequence but the target sequences of the RNA(s) or sgRNA(s) are associated with or correlated to the same type or form of cancer. The cancer

selected from the group consisting of Lung cancer, Lung adenocarcinoma, Lung squamous cell carcinoma, Acute myeloid leukemia, Basal cell carcinoma (skin), Bladder cancer, Breast cancer, Carcinoid, Chronic lymphocytic leukemia, Colorectal cancer, Diffuse large B-cell lymphoma, Endometrial, Esophageal cancer, Esophageal adenocarcinoma, Glioblastoma multiforme, Glioma, Head and neck cancer, Kidney clear cell cancer, Medulloblastoma, Melanoma, Multiple myeloma, Nasopharyngeal, Neuroblastoma, Ovarian cancer, Prostate cancer, Rhabdoid tumor, Testicular germ cell tumor, Thyroid cancer, and Urinary bladder cancer.

[00221] For minimization of toxicity and off-target effect, it will be important to control the concentration of Cas mRNA and guide RNA delivered. Optimal concentrations of Cas mRNA and guide RNA can be determined by testing different concentrations in a cellular or non-human eukaryote animal model and using deep sequencing to analyze the extent of modification at potential off-target genomic loci. For example, for the guide sequence targeting 5'-GAGTCCGAGCAGAAGAAGAA-3' (SEQ ID NO: X) in the EMX1 gene of the human genome, deep sequencing can be used to assess the level of modification at the following two off-target loci, 1: 5'-GAGTCCTAGCAGGAGAAGAA-3' (SEQ ID NO: X) and 2: 5'-GAGTCTAAGCAGAAGAAGAA-3' (SEQ ID NO: X). The concentration that gives the highest level of on-target modification while minimizing the level of off-target modification should be chosen for in vivo delivery. Alternatively, to minimize the level of toxicity and off-target effect, Cas nickase mRNA (for example *S. pyogenes* Cas9 with the D10A mutation) can be delivered with a pair of guide RNAs targeting a site of interest. The two guide RNAs need to be spaced as follows. Guide sequences and strategies to minimize toxicity and off-target effects can be as in WO 2014/093622 (PCT/US2013/074667).

[00222] In some embodiments, the CRISPR system is derived advantageously from a type II 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 preferred embodiments of the invention, the CRISPR system is a type II CRISPR system and the Cas enzyme is Cas9, which catalyzes DNA cleavage. 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, homologues thereof, or modified versions thereof. A preferred Cas enzyme may be identified as Cas9 as this can refer to the general class of enzymes that share homology to the biggest nuclease with multiple nuclease domains from the type II CRISPR system. Most preferably, the Cas9 enzyme is from, or is derived from, SpCas9 or SaCas9. It will be appreciated that SpCas9 or SaCas9 are those from or derived from *S. pyogenes* or *S. aureus* Cas9. By derived, Applicants mean that the derived enzyme is largely based, in the sense of having a high degree of sequence homology with, a wildtype enzyme, but that it has been mutated (modified) in some way as described herein. It will be appreciated that the terms Cas and CRISPR enzyme are generally used herein interchangeably, unless otherwise apparent. The Cas enzyme can be for instance any naturally-occurring bacterial Cas9 as well as any chimaeras, mutants, homologs or orthologs. Many of the residue numberings used herein refer to the Cas9 enzyme from the type II CRISPR locus in *Streptococcus pyogenes* (annotated alternatively as SpCas9 or spCas9). However, it will be appreciated that this invention includes many more Cas9s from other species of microbes, e.g., orthologs of SpCas9, or Cas9s derived from microbes in addition to *S. pyogenes*, e.g., SaCas9 derived from *S. aureus*, St1Cas9 derived from *S. thermophilus* and so forth. The skilled person will be able to determine appropriate corresponding residues in Cas9 enzymes other than SpCas9 by comparison of the relevant amino acid sequences. Thus, where a specific amino acid replacement is referred to using the SpCas9 numbering, then, unless the context makes it apparent this is not intended to refer to other Cas9 enzymes, the disclosure is intended to encompass corresponding modifications in other Cas9 enzymes.

[00223] In some embodiments, the unmodified Cas has DNA cleavage activity, such as Cas9. In some embodiments, the Cas 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 Cas 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. In some embodiments, a vector encodes a Cas that is mutated to with respect to a corresponding wild-type enzyme such that the mutated Cas lacks the ability to cleave one or both strands of a target polynucleotide containing a target sequence. For example,

an aspartate-to-alanine substitution (D10A) in the RuvC I catalytic domain of Cas9 from *S. pyogenes* converts Cas9 from a nuclease that cleaves both strands to a nickase (cleaves a single strand). Other examples of mutations that render Cas9 a nickase include, without limitation, H840A, N854A, and N863A. As a further example, two or more catalytic domains of Cas9 (RuvC I, RuvC II, and RuvC III or the HNH domain) may be mutated to produce a mutated Cas9 substantially lacking all DNA cleavage activity. In some embodiments, a D10A mutation is combined with one or more of H840A, N854A, or N863A mutations to produce a Cas9 enzyme substantially lacking all DNA cleavage activity. In some embodiments, a Cas is considered to substantially lack all DNA cleavage activity when the DNA cleavage activity of the mutated enzyme is about no more than 25%, 10%, 5%, 1%, 0.1%, 0.01%, or less of the DNA cleavage activity of the non-mutated form of the enzyme; an example can be when the DNA cleavage activity of the mutated form is nil or negligible as compared with the non-mutated form. Thus, the Cas may comprise one or more mutations and may be used as a generic DNA binding protein with or without fusion to a functional domain. The mutations may be artificially introduced mutations or gain- or loss-of-function mutations. The mutations may include but are not limited to mutations in one of the catalytic domains (e.g., D10 and H840) in the RuvC and HNH catalytic domains respectively; or the CRISPR enzyme can comprise one or more mutations selected from the group consisting of D10A, E762A, H840A, N854A, N863A or D986A and/or one or more mutations in a RuvC1 or HNH domain of the Cas or has a mutation as otherwise as discussed herein. In one aspect of the invention, the Cas enzyme may be fused to a protein, e.g., a TAG, and/or an inducible/controllable domain such as a chemically inducible/controllable domain. The Cas in the invention may be a chimeric Cas proteins; e.g., a Cas having enhanced function by being a chimera. Chimeric Cas proteins may be new Cas containing fragments from more than one naturally occurring Cas. These may comprise fusions of N-terminal fragment(s) of one Cas9 homolog with C-terminal fragment(s) of another Cas homolog. The Cas can be delivered into the cell in the form of mRNA. The expression of Cas can be under the control of an inducible promoter.

[00224] Where the enzyme is not SpCas9, mutations may be made at any or all residues corresponding to positions 10, 762, 840, 854, 863 and/or 986 of SpCas9 (which may be ascertained for instance by standard sequence comparison tools). In particular, any or all of the

following mutations are preferred in SpCas9: D10A, E762A, H840A, N854A, N863A and/or D986A; as well as conservative substitution for any of the replacement amino acids is also envisaged. The same (or conservative substitutions of these mutations) at corresponding positions in other Cas9s are also preferred. Particularly preferred are D10 and H840 in SpCas9. However, in other Cas9s, residues corresponding to SpCas9 D10 and H840 are also preferred. Orthologs of SpCas9 can be used in the practice of the invention. A Cas enzyme may be identified Cas9 as this can refer to the general class of enzymes that share homology to the biggest nuclease with multiple nuclease domains from the type II CRISPR system. Most preferably, the Cas9 enzyme is from, or is derived from, spCas9 (*S. pyogenes* Cas9) or saCas9 (*S. aureus* Cas9). StCas9” refers to wild type Cas9 from *S. thermophilus*, the protein sequence of which is given in the SwissProt database under accession number G3ECR1. Similarly, *S. pyogenes* Cas9 or spCas9 is included in SwissProt under accession number Q99ZW2. By derived, Applicants mean that the derived enzyme is largely based, in the sense of having a high degree of sequence homology with, a wildtype enzyme, but that it has been mutated (modified) in some way as described herein. It will be appreciated that the terms Cas and CRISPR enzyme are generally used herein interchangeably, unless otherwise apparent. As mentioned above, many of the residue numberings used herein refer to the Cas9 enzyme from the type II CRISPR locus in *Streptococcus pyogenes*. However, it will be appreciated that this invention includes many more Cas9s from other species of microbes, such as SpCas9, SaCa9, St1Cas9 and so forth. Enzymatic action by Cas9 derived from *Streptococcus pyogenes* or any closely related Cas9 generates double stranded breaks at target site sequences which hybridize to 20 nucleotides of the guide sequence and that have a protospacer-adjacent motif (PAM) sequence (examples include NGG/NRG or a PAM that can be determined as described herein) following the 20 nucleotides of the target sequence. CRISPR activity through Cas9 for site-specific DNA recognition and cleavage is defined by the guide sequence, the tracr sequence that hybridizes in part to the guide sequence and the PAM sequence. More aspects of the CRISPR system are described in Karginov and Hannon, *The CRISPR system: small RNA-guided defence in bacteria and archaea*, *Mol Cell* 2010, January 15; 37(1): 7. The type II CRISPR locus from *Streptococcus pyogenes* SF370, which contains a cluster of four genes Cas9, Cas1, Cas2, and Csn1, as well as two non-coding RNA elements, tracrRNA and a characteristic array of repetitive

sequences (direct repeats) interspaced by short stretches of non-repetitive sequences (spacers, about 30bp each). In this system, targeted DNA double-strand break (DSB) is generated in four sequential steps. First, two non-coding RNAs, the pre-crRNA array and tracrRNA, are transcribed from the CRISPR locus. Second, tracrRNA hybridizes to the direct repeats of pre-crRNA, which is then processed into mature crRNAs containing individual spacer sequences. Third, the mature crRNA:tracrRNA complex directs Cas9 to the DNA target consisting of the protospacer and the corresponding PAM via heteroduplex formation between the spacer region of the crRNA and the protospacer DNA. Finally, Cas9 mediates cleavage of target DNA upstream of PAM to create a DSB within the protospacer. A pre-crRNA array consisting of a single spacer flanked by two direct repeats (DRs) is also encompassed by the term “tracr-mate sequences”). In certain embodiments, Cas9 may be constitutively present or inducibly present or conditionally present or administered or delivered. Cas9 optimization may be used to enhance function or to develop new functions, one can generate chimeric Cas9 proteins. And Cas9 may be used as a generic DNA binding protein.

[00225] Typically, in the context of an endogenous CRISPR system, formation of a CRISPR complex (comprising a guide sequence hybridized to a target sequence and complexed with one or more Cas proteins) 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. Without wishing to be bound by theory, the tracr sequence, which may comprise or consist of all or a portion of a wild-type tracr sequence (e.g. about or more than about 20, 26, 32, 45, 48, 54, 63, 67, 85, or more nucleotides of a wild-type tracr sequence), may also form part of a CRISPR complex, such as by hybridization along at least a portion of the tracr sequence to all or a portion of a tracr mate sequence that is operably linked to the guide sequence.

[00226] In some embodiments, the Cas as referred to herein is a codon optimized Cas. An example of a codon optimized sequence, is in this instance a sequence optimized for expression in a eukaryote, e.g., humans (i.e. being optimized for expression in humans), or for another eukaryote, animal or mammal as herein discussed; see, e.g., SaCas9 human codon optimized sequence in WO 2014/093622 (PCT/US2013/074667). Whilst this is preferred, it will be appreciated that other examples are possible and codon optimization for a host species other than

human, or for codon optimization for specific organs is known. In some embodiments, an enzyme coding sequence encoding a Cas is codon optimized for expression in particular cells, such as eukaryotic cells. The eukaryotic cells may be those of or derived from a particular organism, such as a mammal, including but not limited to human, or non-human eukaryote or animal or mammal as herein discussed, e.g., mouse, rat, rabbit, dog, livestock, or non-human mammal or primate. In some embodiments, processes for modifying the germ line genetic identity of human beings and/or processes for modifying the genetic identity of animals which are likely to cause them suffering without any substantial medical benefit to man or animal, and also animals resulting from such processes, may be excluded. In general, codon optimization refers to a process of modifying a nucleic acid sequence for enhanced expression in the host cells of interest by replacing at least one codon (e.g. about or more than about 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more codons) of the native sequence with codons that are more frequently or most frequently used in the genes of that host cell while maintaining the native amino acid sequence. Various species exhibit particular bias for certain codons of a particular amino acid. Codon bias (differences in codon usage between organisms) often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, among other things, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization. Codon usage tables are readily available, for example, at the “Codon Usage Database” available at www.kazusa.or.jp/codon/ and these tables can be adapted in a number of ways. See Nakamura, Y., et al. “Codon usage tabulated from the international DNA sequence databases: status for the year 2000” *Nucl. Acids Res.* 28:292 (2000). Computer algorithms for codon optimizing a particular sequence for expression in a particular host cell are also available, such as Gene Forge (Aptagen; Jacobus, PA), are also available. In some embodiments, one or more codons (e.g. 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more, or all codons) in a sequence encoding a Cas correspond to the most frequently used codon for a particular amino acid.

[00227] As used herein, the term “Cas transgenic eukaryotic cell” refers to a eukaryotic cell in which a Cas gene has been genomically integrated. The nature, type, or origin of the eukaryotic

cell are not particularly limiting according to the present invention. Also the way how the Cas transgene is introduced in the eukaryotic cell is may vary and can be any method as is known in the art. In certain embodiments, the Cas transgenic eukaryotic cell is obtained by introducing the Cas transgene in an isolated eukaryotic cell. In certain other embodiments, the Cas transgenic eukaryotic cell is obtained by isolating cells from a Cas transgenic eukaryote. By means of example, and without limitation, the Cas transgenic eukaryotic cell as referred to herein may be derived from a Cas transgenic eukaryote, such as a Cas knock-in eukaryote. Reference is made to WO 2014/093622 (PCT/US13/74667), incorporated herein by reference. Methods of US Patent Publication Nos. 20120017290 and 20110265198 assigned to Sangamo BioSciences, Inc. directed to targeting the Rosa locus may be modified to utilize the CRISPR Cas system of the present invention. Methods of US Patent Publication No. 20130236946 assigned to Cellectis directed to targeting the Rosa locus may also be modified to utilize the CRISPR Cas system of the present invention. By means of further example reference is made to Platt et. al. (Cell; 159(2):440-455 (2014)), describing a Cas9 knock-in mouse, which is incorporated herein by reference. The Cas transgene can further comprise a Lox-Stop-polyA-Lox(LSL) cassette thereby rendering Cas expression inducible by Cre recombinase. Alternatively, the Cas transgenic eukaryotic cell may be obtained by introducing the Cas transgene in an isolated eukaryotic cell. Delivery systems for transgenes are well known in the art. By means of example, the Cas transgene may be delivered in the eukaryotic cell by means of vector (e.g., AAV, adenovirus, lentivirus) and/or particle and/or nanoparticle delivery, as also described herein elsewhere.

[00228] It will be understood by the skilled person that the eukaryotic cell, such as the Cas transgenic eukaryotic cell, as referred to herein may comprise further genomic alterations besides having an integrated Cas gene or the mutations arising from the sequence specific action of Cas when complexed with RNA capable of guiding Cas to a genomic target locus, such as for instance one or more oncogenic mutations, as for instance and without limitation described in Platt et al. (2014), Chen et al., (2014) or Kumar et al.. (2009). In certain embodiments, the eukaryotic cell as used herein has a homozygous inactivation of p53 and/or a heterozygous inactivation of Dicer 1 and/or an oncogenic mutation in Kras G12D.

[00229] In some embodiments, the Cas sequence is fused to one or more nuclear localization sequences (NLSs), such as about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs. In some embodiments, the Cas comprises about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs at or near the amino-terminus, about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs at or near the carboxy-terminus, or a combination of these (e.g. zero or at least one or more NLS at the amino-terminus and zero or at one or more NLS at the carboxy terminus). When more than one NLS is present, each may be selected independently of the others, such that a single NLS may be present in more than one copy and/or in combination with one or more other NLSs present in one or more copies. In a preferred embodiment of the invention, the Cas comprises at most 6 NLSs. In some embodiments, an NLS is considered near the N- or C-terminus when the nearest amino acid of the NLS is within about 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50, or more amino acids along the polypeptide chain from the N- or C-terminus. Non-limiting examples of NLSs include an NLS sequence derived from: the NLS of the SV40 virus large T-antigen, having the amino acid sequence PKKKRKV (SEQ ID NO: X); the NLS from nucleoplasmin (e.g. the nucleoplasmin bipartite NLS with the sequence KRPAATKKAGQAKKKK) (SEQ ID NO: X); the c-myc NLS having the amino acid sequence PAAKRVKLD (SEQ ID NO: X) or RQRRNELKRSP (SEQ ID NO: X); the hRNPA1 M9 NLS having the sequence NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: X); the sequence RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: X) of the IBB domain from importin-alpha; the sequences VSRKRPRP (SEQ ID NO: X) and PPKKARED (SEQ ID NO: X) of the myoma T protein; the sequence POPKKKPL (SEQ ID NO: X) of human p53; the sequence SALIKKKKKMAP (SEQ ID NO: X) of mouse c-abl IV; the sequences DRLRR (SEQ ID NO: X) and PKQKKRK (SEQ ID NO: X) of the influenza virus NS1; the sequence RKLKKKIKKL (SEQ ID NO: X) of the Hepatitis virus delta antigen; the sequence REKKKFLKRR (SEQ ID NO: X) of the mouse Mx1 protein; the sequence KRKGDEV DGVDEVAKKKSKK (SEQ ID NO: X) of the human poly(ADP-ribose) polymerase; and the sequence RKCLQAGMNLEARKTKK (SEQ ID NO: X) of the steroid hormone receptors (human) glucocorticoid. In general, the one or more NLSs are of sufficient strength to drive accumulation of the Cas in a detectable amount in the nucleus of a eukaryotic cell. In general, strength of nuclear localization activity may derive from the number of NLSs in

the Cas, the particular NLS(s) used, or a combination of these factors. Detection of accumulation in the nucleus may be performed by any suitable technique. For example, a detectable marker may be fused to the Cas, such that location within a cell may be visualized, such as in combination with a means for detecting the location of the nucleus (e.g. a stain specific for the nucleus such as DAPI). Cell nuclei may also be isolated from cells, the contents of which may then be analyzed by any suitable process for detecting protein, such as immunohistochemistry, Western blot, or enzyme activity assay. Accumulation in the nucleus may also be determined indirectly, such as by an assay for the effect of CRISPR complex formation (e.g. assay for DNA cleavage or mutation at the target sequence, or assay for altered gene expression activity affected by CRISPR complex formation and/or Cas enzyme activity), as compared to a control not exposed to the Cas or complex, or exposed to a Cas lacking the one or more NLSs.

[00230] In certain aspects the invention involves vectors, e.g. for delivering or introducing in a eukaryotic cell Cas and/or RNA capable of guiding Cas to a genomic target locus, but also for propagating these components (e.g. in prokaryotic cells). As used herein, a “vector” is a tool that allows or facilitates the transfer of an entity from one environment to another. It is a replicon, such as a plasmid, phage, or cosmid, into which another DNA segment may be inserted so as to bring about the replication of the inserted segment. Generally, a vector is capable of replication when associated with the proper control elements. In general, the term “vector” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. Vectors include, but are not limited to, nucleic acid molecules that are single-stranded, double-stranded, or partially double-stranded; nucleic acid molecules that comprise one or more free ends, no free ends (e.g. circular); nucleic acid molecules that comprise DNA, RNA, or both; and other varieties of polynucleotides known in the art. One type of vector is a “plasmid,” which refers to a circular double stranded DNA loop into which additional DNA segments can be inserted, such as by standard molecular cloning techniques. Another type of vector is a viral vector, wherein virally-derived DNA or RNA sequences are present in the vector for packaging into a virus (e.g. retroviruses, replication defective retroviruses, adenoviruses, replication defective adenoviruses, and adeno-associated viruses (AAVs)). Viral vectors also include polynucleotides carried by a virus for transfection into a host cell. 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 (e.g., non-episomal mammalian vectors) are 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 herein as “expression vectors.” Common expression vectors of utility in recombinant DNA techniques are often in the form of plasmids.

[00231] Recombinant expression vectors can comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory elements, which may be selected on the basis of the host cells to be used for expression, that is operatively-linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, “operably linked” is intended to mean that the nucleotide sequence of interest is linked to the regulatory element(s) in a manner that allows for expression of the nucleotide sequence (e.g. in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell). With regards to recombination and cloning methods, mention is made of U.S. patent application 10/815,730, published September 2, 2004 as US 2004-0171156 A1, the contents of which are herein incorporated by reference in their entirety.

[00232] The vector(s) can include the regulatory element(s), e.g., promoter(s). The vector(s) can comprise Cas, and/or a single, but possibly also can comprise at least 3 or 8 or 16 or 32 or 48 or 50 RNA(s) (e.g., sgRNAs), such as 1-2, 1-3, 1-4 1-5, 3-6, 3-7, 3-8, 3-9, 3-10, 3-8, 3-16, 3-30, 3-32, 3-48, 3-50 RNA(s) (e.g., sgRNAs). In a single vector there can be a promoter for each RNA (e.g., sgRNA), advantageously when there are up to about 16 RNA(s) (e.g., sgRNAs); and, when a single vector provides for more than 16 RNA(s) (e.g., sgRNAs), one or more promoter(s) can drive expression of more than one of the RNA(s) (e.g., sgRNAs), e.g., when there are 32 RNA(s) (e.g., sgRNAs), each promoter can drive expression of two RNA(s) (e.g., sgRNAs), and when there are 48 RNA(s) (e.g., sgRNAs), each promoter can drive expression of three RNA(s) (e.g., sgRNAs). By simple arithmetic and well established cloning protocols and the teachings in this disclosure one skilled in the art can readily practice the invention as to the RNA(s) (e.g., sgRNA(s)) for a suitable exemplary vector such as AAV, and a suitable promoter such as the U6

promoter, e.g., U6-sgRNAs. For example, the packaging limit of AAV is ~4.7 kb. The length of a single U6-sgRNA (plus restriction sites for cloning) is 361 bp. Therefore, the skilled person can readily fit about 12-16, e.g., 13 U6-sgRNA cassettes in a single vector. This can be assembled by any suitable means, such as a golden gate strategy used for TALE assembly (<http://www.genome-engineering.org/taleffectors/>). The skilled person can also use a tandem guide strategy to increase the number of U6-sgRNAs by approximately 1.5 times, e.g., to increase from 12-16, e.g., 13 to approximately 18-24, e.g., about 19 U6-sgRNAs. Therefore, one skilled in the art can readily reach approximately 18-24, e.g., about 19 promoter-RNAs, e.g., U6-sgRNAs in a single vector, e.g., an AAV vector. A further means for increasing the number of promoters and RNAs, e.g., sgRNA(s) in a vector is to use a single promoter (e.g., U6) to express an array of RNAs, e.g., sgRNAs separated by cleavable sequences. And an even further means for increasing the number of promoter-RNAs, e.g., sgRNAs in a vector, is to express an array of promoter-RNAs, e.g., sgRNAs separated by cleavable sequences in the intron of a coding sequence or gene; and, in this instance it is advantageous to use a polymerase II promoter, which can have increased expression and enable the transcription of long RNA in a tissue specific manner. (see, e.g., <http://nar.oxfordjournals.org/content/34/7/e53.short>, <http://www.nature.com/mt/journal/v16/n9/abs/mt2008144a.html>). In an advantageous embodiment, AAV may package U6 tandem sgRNA targeting up to about 50 genes. Accordingly, from the knowledge in the art and the teachings in this disclosure the skilled person can readily make and use vector(s), e.g., a single vector, expressing multiple RNAs or guides or sgRNAs under the control or operatively or functionally linked to one or more promoters—especially as to the numbers of RNAs or guides or sgRNAs discussed herein, without any undue experimentation.

[00233] The RNA(s), e.g., sgRNA(s), can be functionally or operatively linked to regulatory element(s) and hence the regulatory element(s) drive expression. The promoter(s) can be constitutive promoter(s) and/or inducible promoter(s) and/or tissue specific promoter(s). The promoter can be selected from the group consisting of RNA polymerases, pol I, pol II, pol III, T7, U6, H1, retroviral Rous sarcoma virus (RSV) LTR promoter, the cytomegalovirus (CMV) promoter, the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the

phosphoglycerol kinase (PGK) promoter, and the EF1 α promoter. An advantageous promoter is the promoter is U6.

[00234] Aspects of the invention also relate to bicistronic vectors for chimeric RNA and Cas. Bicistronic expression vectors for chimeric RNA and Cas are preferred. In general and particularly in this embodiment Cas is preferably driven by the CBh promoter. The chimeric RNA may preferably be driven by a Pol III promoter, such as a U6 promoter. Ideally the two are combined. The chimeric guide RNA typically consists of a 20bp guide sequence (Ns) and this may be joined to the tracr sequence (running from the first “U” of the lower strand to the end of the transcript). The tracr sequence may be truncated at various positions as indicated. The guide and tracr sequences are separated by the tracr-mate sequence, which may be GUUUUAGAGCUA (SEQ ID NO:X). This may be followed by the loop sequence GAAA. Both of these are preferred examples. ChiRNAs are indicated by their “+n” designation, and crRNA refers to a hybrid RNA where guide and tracr sequences are expressed as separate transcripts. Throughout this application, chimeric RNA may also be called single guide, or synthetic guide RNA (sgRNA). The loop is preferably GAAA, but it is not limited to this sequence or indeed to being only 4bp in length. Indeed, preferred loop forming sequences for use in hairpin structures are four nucleotides in length, and most preferably have the sequence GAAA. However, longer or shorter loop sequences may be used, as may alternative sequences. The sequences preferably include a nucleotide triplet (for example, AAA), and an additional nucleotide (for example C or G). Examples of loop forming sequences include CAAA and AAAG. In practicing any of the methods disclosed herein, a suitable vector can be introduced to a cell or an embryo via one or more methods known in the art, including without limitation, microinjection, electroporation, sonoporation, biolistics, calcium phosphate-mediated transfection, cationic transfection, liposome transfection, dendrimer transfection, heat shock transfection, nucleofection transfection, magnetofection, lipofection, impalefection, optical transfection, proprietary agent-enhanced uptake of nucleic acids, and delivery via liposomes, immunoliposomes, virosomes, or artificial virions. In some methods, the vector is introduced into an embryo by microinjection. The vector or vectors may be microinjected into the nucleus or the cytoplasm of the embryo. In some methods, the vector or vectors may be introduced into a cell by nucleofection.

[00235] The term “regulatory element” is intended to include promoters, enhancers, internal ribosomal entry sites (IRES), and other expression control elements (e.g. transcription termination signals, such as polyadenylation signals and poly-U sequences). Such regulatory elements are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory elements include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). A tissue-specific promoter may 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). Regulatory elements may also direct expression in a temporal-dependent manner, such as in a cell-cycle dependent or developmental stage-dependent manner, which may or may not also be tissue or cell-type specific. In some embodiments, a vector comprises one or more pol III promoter (e.g. 1, 2, 3, 4, 5, or more pol III promoters), one or more pol II promoters (e.g. 1, 2, 3, 4, 5, or more pol II promoters), one or more pol I promoters (e.g. 1, 2, 3, 4, 5, or more pol I promoters), or combinations thereof. Examples of pol III promoters include, but are not limited to, U6 and H1 promoters. Examples of pol II promoters include, but are not limited to, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) [see, e.g., Boshart et al, Cell, 41:521-530 (1985)], the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1 α promoter. Also encompassed by the term “regulatory element” are enhancer elements, such as WPRE; CMV enhancers; the R-U5' segment in LTR of HTLV-I (Mol. Cell. Biol., Vol. 8(1), p. 466-472, 1988); SV40 enhancer; and the intron sequence between exons 2 and 3 of rabbit β -globin (Proc. Natl. Acad. Sci. USA., Vol. 78(3), p. 1527-31, 1981). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression desired, etc. A vector can be introduced into host cells to thereby produce transcripts, proteins, or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., clustered regularly interspersed short palindromic repeats (CRISPR) transcripts, proteins, enzymes, mutant forms thereof, fusion proteins thereof, etc.). With regards

to regulatory sequences, mention is made of U.S. patent application 10/491,026, the contents of which are incorporated by reference herein in their entirety. With regards to promoters, mention is made of PCT publication WO 2011/028929 and U.S. application 12/511,940, the contents of which are incorporated by reference herein in their entirety.

[00236] Vectors can be designed for expression of CRISPR transcripts (e.g. nucleic acid transcripts, proteins, or enzymes) in prokaryotic or eukaryotic cells. For example, CRISPR transcripts can be expressed in bacterial cells such as *Escherichia coli*, insect cells (using baculovirus expression vectors), yeast cells, or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated in vitro, for example using T7 promoter regulatory sequences and T7 polymerase.

[00237] Vectors may be introduced and propagated in a prokaryote or prokaryotic cell. In some embodiments, a prokaryote is used to amplify copies of a vector to be introduced into a eukaryotic cell or as an intermediate vector in the production of a vector to be introduced into a eukaryotic cell (e.g. amplifying a plasmid as part of a viral vector packaging system). In some embodiments, a prokaryote is used to amplify copies of a vector and express one or more nucleic acids, such as to provide a source of one or more proteins for delivery to a host cell or host organism. Expression of proteins in prokaryotes is most often carried out in *Escherichia coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, such as to the amino terminus of the recombinant protein. Such fusion vectors may serve one or more purposes, such as: (i) to increase expression of recombinant protein; (ii) to increase the solubility of the recombinant protein; and (iii) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Example fusion expression vectors include pGEX

(Pharmacia Biotech Inc; Smith and Johnson, 1988. *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann et al., (1988) *Gene* 69:301-315) and pET 11d (Studier et al., *GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY* 185, Academic Press, San Diego, Calif. (1990) 60-89). In some embodiments, a vector is a yeast expression vector. Examples of vectors for expression in yeast *Saccharomyces cerevisiae* include pYepSec1 (Baldari, et al., 1987. *EMBO J.* 6: 229-234), pMFa (Kuijan and Herskowitz, 1982. *Cell* 30: 933-943), pJRY88 (Schultz et al., 1987. *Gene* 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (Invitrogen Corp, San Diego, Calif.). In some embodiments, a vector drives protein expression in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., SF9 cells) include the pAc series (Smith, et al., 1983. *Mol. Cell. Biol.* 3: 2156-2165) and the pVL series (Lucklow and Summers, 1989. *Virology* 170: 31-39).

[00238] In some embodiments, a vector is capable of driving expression of one or more sequences in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, 1987. *Nature* 329: 840) and pMT2PC (Kaufman, et al., 1987. *EMBO J.* 6: 187-195). When used in mammalian cells, the expression vector's control functions are typically provided by one or more regulatory elements. For example, commonly used promoters are derived from polyoma, adenovirus 2, cytomegalovirus, simian virus 40, and others disclosed herein and known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells see, e.g., Chapters 16 and 17 of Sambrook, et al., *MOLECULAR CLONING: A LABORATORY MANUAL*. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

[00239] In some embodiments, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters

include the albumin promoter (liver-specific; Pinkert, et al., 1987. *Genes Dev.* 1: 268-277), lymphoid-specific promoters (Calame and Eaton, 1988. *Adv. Immunol.* 43: 235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989. *EMBO J.* 8: 729-733) and immunoglobulins (Baneiji, et al., 1983. *Cell* 33: 729-740; Queen and Baltimore, 1983. *Cell* 33: 741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle, 1989. *Proc. Natl. Acad. Sci. USA* 86: 5473-5477), pancreas-specific promoters (Edlund, et al., 1985. *Science* 230: 912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, e.g., the murine hox promoters (Kessel and Gruss, 1990. *Science* 249: 374-379) and the α -fetoprotein promoter (Campes and Tilghman, 1989. *Genes Dev.* 3: 537-546). With regards to these prokaryotic and eukaryotic vectors, mention is made of U.S. Patent 6,750,059, the contents of which are incorporated by reference herein in their entirety. Other embodiments of the invention may relate to the use of viral vectors, with regards to which mention is made of U.S. Patent application 13/092,085, the contents of which are incorporated by reference herein in their entirety. Tissue-specific regulatory elements are known in the art and in this regard, mention is made of U.S. Patent 7,776,321, the contents of which are incorporated by reference herein in their entirety. In some embodiments, a regulatory element is operably linked to one or more elements of a CRISPR system so as to drive expression of the one or more elements of the CRISPR system. In general, CRISPRs (Clustered Regularly Interspaced Short Palindromic Repeats), also known as SPIDRs (SPacer Interspersed Direct Repeats), constitute a family of DNA loci that are usually specific to a particular bacterial species. The CRISPR locus comprises a distinct class of interspersed short sequence repeats (SSRs) that were recognized in *E. coli* (Ishino et al., *J. Bacteriol.*, 169:5429-5433 [1987]; and Nakata et al., *J. Bacteriol.*, 171:3553-3556 [1989]), and associated genes. Similar interspersed SSRs have been identified in *Haloferax mediterranei*, *Streptococcus pyogenes*, *Anabaena*, and *Mycobacterium tuberculosis* (See, Groenen et al., *Mol. Microbiol.*, 10:1057-1065 [1993]; Hoe et al., *Emerg. Infect. Dis.*, 5:254-263 [1999]; Masepohl et al., *Biochim. Biophys. Acta* 1307:26-30 [1996]; and Mojica et al., *Mol. Microbiol.*, 17:85-93 [1995]). The CRISPR loci typically differ from other SSRs by the structure of the repeats, which have been termed short regularly spaced repeats (SRSRs) (Janssen et al., *OMICS J. Integ. Biol.*, 6:23-33 [2002]; and Mojica et al., *Mol.*

Microbiol., 36:244-246 [2000]). In general, the repeats are short elements that occur in clusters that are regularly spaced by unique intervening sequences with a substantially constant length (Mojica et al., [2000], supra). Although the repeat sequences are highly conserved between strains, the number of interspersed repeats and the sequences of the spacer regions typically differ from strain to strain (van Embden et al., J. Bacteriol., 182:2393-2401 [2000]). CRISPR loci have been identified in more than 40 prokaryotes (See e.g., Jansen et al., Mol. Microbiol., 43:1565-1575 [2002]; and Mojica et al., [2005]) including, but not limited to *Aeropyrum*, *Pyrobaculum*, *Sulfolobus*, *Archaeoglobus*, *Halocarcula*, *Methanobacterium*, *Methanococcus*, *Methanosarcina*, *Methanopyrus*, *Pyrococcus*, *Picrophilus*, *Thermoplasma*, *Corynebacterium*, *Mycobacterium*, *Streptomyces*, *Aquifex*, *Porphyromonas*, *Chlorobium*, *Thermus*, *Bacillus*, *Listeria*, *Staphylococcus*, *Clostridium*, *Thermoanaerobacter*, *Mycoplasma*, *Fusobacterium*, *Azarcus*, *Chromobacterium*, *Neisseria*, *Nitrosomonas*, *Desulfovibrio*, *Geobacter*, *Myxococcus*, *Campylobacter*, *Wolinella*, *Acinetobacter*, *Erwinia*, *Escherichia*, *Legionella*, *Methylococcus*, *Pasteurella*, *Photobacterium*, *Salmonella*, *Xanthomonas*, *Yersinia*, *Treponema*, and *Thermotoga*. In some embodiments, one or more vectors driving expression of one or more elements of a CRISPR system (such as the Cas and or the RNA guiding the Cas to a genomic target locus in a eukaryotic cell as referred to herein elsewhere) are introduced into a host cell such that expression of the elements of the CRISPR system direct formation of a CRISPR complex at one or more target sites. For example, a Cas enzyme, a guide sequence linked to a tracr-mate sequence, and a tracr sequence could each be operably linked to separate regulatory elements on separate vectors. Alternatively, two or more of the elements expressed from the same or different regulatory elements, may be combined in a single vector, with one or more additional vectors providing any components of the CRISPR system not included in the first vector. CRISPR system elements that are combined in a single vector may be arranged in any suitable orientation, such as one element located 5' with respect to ("upstream" of) or 3' with respect to ("downstream" of) a second element. The coding sequence of one element may be located on the same or opposite strand of the coding sequence of a second element, and oriented in the same or opposite direction. In some embodiments, a single promoter drives expression of a transcript encoding a Cas and one or more of the guide sequence, tracr mate sequence (optionally operably linked to the guide sequence), and a tracr sequence embedded within one or more intron

sequences (e.g. each in a different intron, two or more in at least one intron, or all in a single intron). In some embodiments, the Cas, guide sequence, tracr mate sequence, and tracr sequence are operably linked to and expressed from the same promoter. Delivery vehicles, vectors, particles, nanoparticles, formulations and components thereof for expression of one or more elements of a CRISPR system are as used in the foregoing documents, such as WO 2014/093622 (PCT/US2013/074667). In some embodiments, a vector comprises one or more insertion sites, such as a restriction endonuclease recognition sequence (also referred to as a “cloning site”). In some embodiments, one or more insertion sites (e.g. about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more insertion sites) are located upstream and/or downstream of one or more sequence elements of one or more vectors. In some embodiments, a vector comprises an insertion site upstream of a tracr mate sequence, and optionally downstream of a regulatory element operably linked to the tracr mate sequence, such that following insertion of a guide sequence into the insertion site and upon expression the guide sequence directs sequence-specific binding of a CRISPR complex to a target sequence in a eukaryotic cell. In some embodiments, a vector comprises two or more insertion sites, each insertion site being located between two tracr mate sequences so as to allow insertion of a guide sequence at each site. In such an arrangement, the two or more guide sequences may comprise two or more copies of a single guide sequence, two or more different guide sequences, or combinations of these. When multiple different guide sequences are used, a single expression construct may be used to target CRISPR activity to multiple different, corresponding target sequences within a cell. For example, a single vector may comprise about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or more guide sequences. In some embodiments, about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more such guide-sequence-containing vectors may be provided, and optionally delivered to a cell. In some embodiments, a vector comprises a regulatory element operably linked to an enzyme-coding sequence encoding a Cas protein. Cas protein or Cas mRNA or CRISPR guide RNA or RNA(s) can be delivered separately; and advantageously at least one of these is delivered via a particle complex. Cas mRNA can be delivered prior to the guide RNA to give time for Cas to be expressed. Cas mRNA might be administered 1-12 hours (preferably around 2-6 hours) prior to the administration of guide RNA. Alternatively, Cas mRNA and guide RNA can be administered together. Advantageously, a second booster dose of guide RNA can be administered 1-12 hours

(preferably around 2-6 hours) after the initial administration of Cas mRNA + guide RNA. Additional administrations of Cas mRNA and/or guide RNA might be useful to achieve the most efficient levels of genome modification.

[00240] In certain embodiments, this invention provides a method of modifying expression of a polynucleotide in a eukaryotic cell. The method comprises increasing or decreasing expression of a target polynucleotide by using a CRISPR complex that binds to the polynucleotide. In some methods, a target polynucleotide can be inactivated to effect the modification of the expression in a cell. For example, upon the binding of a CRISPR complex to a target sequence in a cell, the target polynucleotide is inactivated such that the sequence is not transcribed, the coded protein is not produced, or the sequence does not function as the wild-type sequence does. For example, a protein or microRNA coding sequence may be inactivated such that the protein or microRNA or pre-microRNA transcript is not produced. In some methods, a control sequence can be inactivated such that it no longer functions as a control sequence. As used herein, “control sequence” refers to any nucleic acid sequence that effects the transcription, translation, or accessibility of a nucleic acid sequence. Examples of a control sequence include, a promoter, a transcription terminator, and an enhancer are control sequences. The target polynucleotide of a CRISPR complex can be any polynucleotide endogenous or exogenous to the eukaryotic cell. For example, the target polynucleotide can be a polynucleotide residing in the nucleus of the eukaryotic cell. The target polynucleotide can be a sequence coding a gene product (e.g., a protein) or a non-coding sequence (e.g., a regulatory polynucleotide or a junk DNA). Examples of target polynucleotides include a sequence associated with a signaling biochemical pathway, e.g., a signaling biochemical pathway-associated gene or polynucleotide. Examples of target polynucleotides include a disease associated gene or polynucleotide. A “disease-associated” gene or polynucleotide refers to any gene or polynucleotide which is yielding transcription or translation products at an abnormal level or in an abnormal form in cells derived from a disease-affected tissues compared with tissues or cells of a non disease control, such as oncogenes or tumor suppressor genes or metastasis suppressor genes. It may be a gene that becomes expressed at an abnormally high level; it may be a gene that becomes expressed at an abnormally low level, where the altered expression correlates with the occurrence and/or progression of the disease. A disease-associated gene also refers to a gene possessing mutation(s) or genetic variation that is

directly responsible or is in linkage disequilibrium with a gene(s) that is responsible for the etiology of a disease. The transcribed or translated products may be known or unknown, and may be at a normal or abnormal level. The target polynucleotide of a CRISPR complex can be any polynucleotide endogenous or exogenous to the eukaryotic cell. For example, the target polynucleotide can be a polynucleotide residing in the nucleus of the eukaryotic cell. The target polynucleotide can be a sequence coding a gene product (e.g., a protein) or a non-coding sequence (e.g., a regulatory polynucleotide or a junk DNA). Without wishing to be bound by theory, it is believed that the target sequence should be associated with a PAM (protospacer adjacent motif); that is, a short sequence recognized by the CRISPR complex. The precise sequence and length requirements for the PAM differ depending on the Cas used, but PAMs are typically 2-5 base pair sequences adjacent the protospacer (that is, the target sequence. In some embodiments, the method comprises allowing a CRISPR complex to bind to the target polynucleotide to effect cleavage of said target polynucleotide thereby modifying the target polynucleotide, wherein the CRISPR complex comprises a Cas complexed with a guide sequence hybridized to a target sequence within said target polynucleotide, wherein said guide sequence is linked to a tracr mate sequence which in turn hybridizes to a tracr sequence. In one aspect, the invention provides a method of modifying expression of a polynucleotide in a eukaryotic cell. In some embodiments, the method comprises allowing a CRISPR complex to bind to the polynucleotide such that said binding results in increased or decreased expression of said polynucleotide; wherein the CRISPR complex comprises a Cas complexed with a guide sequence hybridized to a target sequence within said polynucleotide, wherein said guide sequence is linked to a tracr mate sequence which in turn hybridizes to a tracr sequence. Similar considerations and conditions apply as above for methods of modifying a target polynucleotide. In fact, these sampling, culturing and re-introduction options apply across the aspects of the present invention. In one aspect, the invention provides for methods of modifying a target polynucleotide in a eukaryotic cell, which may be in vivo, ex vivo or in vitro. In some embodiments, the method comprises sampling a cell or population of cells from a human or non-human animal, and modifying the cell or cells. Culturing may occur at any stage ex vivo. The cell or cells may even be re-introduced into the non-human animal or plant. For re-introduced cells it is particularly preferred that the cells are stem cells.

[00241] In certain embodiments, a Cas and/or an RNA capable of guiding the Cas to a target locus as described herein elsewhere is delivered to or introduced in a eukaryotic cell. Conventional viral and non-viral based gene transfer methods can be used to introduce nucleic acids in cells. Non-viral vector delivery systems include DNA plasmids, RNA (e.g. a transcript of a vector described herein), naked nucleic acid, and nucleic acid complexed with a delivery vehicle, such as a liposome. Methods of non-viral delivery of nucleic acids include lipofection, 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). The preparation of lipid:nucleic acid complexes, including targeted liposomes such as immunolipid complexes, is well known to one of skill in the art (see, e.g., Crystal, Science 270:404-410 (1995); Blaese et al., Cancer Gene Ther. 2:291-297 (1995); Behr et al., Bioconjugate Chem. 5:382-389 (1994); Remy et al., Bioconjugate Chem. 5:647-654 (1994); Gao et al., Gene Therapy 2:710-722 (1995); Ahmad et al., Cancer Res. 52:4817-4820 (1992); U.S. Pat. Nos. 4,186,183, 4,217,344, 4,235,871, 4,261,975, 4,485,054, 4,501,728, 4,774,085, 4,837,028, and 4,946,787). Viral vector delivery systems include DNA and RNA viruses, which have either episomal or integrated genomes after delivery to the cell. Viral vectors can be used to treat cells in vitro, and the modified cells can then be administered to a eukaryote, such as a non-human eukaryote. 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, often resulting in long term expression of the inserted transgene. Additionally, high transduction efficiencies have been observed in many different cell types and target tissues. The tropism of a retrovirus can be altered by incorporating foreign envelope proteins, expanding the potential target population of target cells. Lentiviral vectors are retroviral vectors (and hence both lentiviral and retroviral vectors may be used in the practice of the invention). Moreover, lentiviral vectors are preferred

as they are able to transduce or infect non-dividing cells and typically produce high viral titers. Selection of a retroviral gene transfer system may 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, which are then used to integrate the therapeutic gene into the target cell to provide permanent transgene expression. Widely used retroviral vectors that may be used in the practice of the invention include those based upon murine leukemia virus (MuLV), gibbon ape leukemia virus (GaLV), Simian Immuno deficiency virus (SIV), human immuno deficiency virus (HIV), and combinations thereof (see, e.g., Buchscher et al., *J. Virol.* 66:2731-2739 (1992); Johann et al., *J. Virol.* 66:1635-1640 (1992); Sommerfelt et al., *Virol.* 176:58-59 (1990); Wilson et al., *J. Virol.* 63:2374-2378 (1989); Miller et al., *J. Virol.* 65:2220-2224 (1991); PCT/US94/05700). Zou et al. administered about 10 μ l of a recombinant lentivirus having a titer of 1×10^9 transducing units (TU)/ml by an intrathecal catheter. These sorts of dosages can be adapted or extrapolated to use of a retroviral or lentiviral vector in the present invention.

[00242] If a Cas transgenic eukaryotic cell provided for herein is used, then only delivery of guide(s) is necessary, i.e. RNA capable of guiding Cas to a genomic target locus. In some embodiments, one or more vectors described herein are used to produce a non-human transgenic Cas9 eukaryote, e.g., animal, mammal, primate, rodent, mouse, rat, rabbit. In some embodiments, the transgenic animal is a mammal, such as a mouse, rat, or rabbit. Guides or RNA(s) can be delivered via the same vector types as Cas9. When both guides or RNA(s) and Cas9 are being delivered a dual-vector system where the Cas9 is delivered via in vivo expression from an AAV vector and the guide(s) are delivered by a separate AAV vector. This can be done substantially contemporaneously (i.e., co-delivery), but it could also be done at separate points in time, separated even by weeks or months. Of course, the ultimate separation is where the transgenic Cas9 eukaryote is generated and thereafter the guide(s) or RNA(s) are delivered. Alternatively a first round of CRISPR-Cas9 systems can be delivered, and subsequently further guides or RNA(s) are delivered as the original Cas9 is still functional in the target cells may be re-used. If the Cas9 is under the control of an inducible promoter, then induction of transcription of new Cas9 in the target cells is preferred.

[00243] Adeno-associated virus ("AAV") vectors may also be used to transduce cells with target nucleic acids, e.g., in the in vitro production of nucleic acids and peptides, and for in vivo and ex vivo gene therapy procedures (see, e.g., West et al., *Virology* 160:38-47 (1987); U.S. Pat. No. 4,797,368; WO 93/24641; Kotin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1351 (1994). Construction of recombinant AAV vectors is described in a number of publications, including U.S. Pat. No. 5,173,414; Tratschin et al., *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin, et al., *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat & Muzyczka, *PNAS* 81:6466-6470 (1984); and Samulski et al., *J. Virol.* 63:03822-3828 (1989). Packaging cells are typically used to form virus particles that are capable of infecting a host cell. Such cells include 293 cells, which package adenovirus, and ψ 2 cells or PA317 cells, which package retrovirus. Viral vectors used in gene therapy are usually generated by producer a cell line that packages a nucleic acid vector into a viral particle. The vectors typically contain the minimal viral sequences required for packaging and subsequent integration into a host, other viral sequences being replaced by an expression cassette for the polynucleotide(s) to be expressed. The missing viral functions are typically supplied in trans by the packaging cell line. For example, AAV vectors used in gene therapy typically only possess ITR sequences from the AAV genome which are required for packaging and integration into the host genome. Viral DNA is packaged in a cell line, which contains a helper plasmid encoding the other AAV genes, namely rep and cap, but lacking ITR sequences. The cell line may also be infected with adenovirus as a helper. The helper virus promotes replication of the AAV vector and expression of AAV genes from the helper plasmid. The helper plasmid is not packaged in significant amounts due to a lack of ITR sequences. Contamination with adenovirus can be reduced by, e.g., heat treatment to which adenovirus is more sensitive than AAV. Accordingly, AAV is considered an ideal candidate for use as a transducing vector. Such AAV transducing vectors can comprise sufficient cis-acting functions to replicate in the presence of adenovirus or herpesvirus or poxvirus (e.g., vaccinia virus) helper functions provided in trans. Recombinant AAV (rAAV) can be used to carry exogenous genes into cells of a variety of lineages. In these vectors, the AAV cap and/or rep genes are deleted from the viral genome and replaced with a DNA segment of choice. Current AAV vectors may accommodate up to 4300 bases of inserted DNA. There are a number of ways to produce rAAV, and the invention provides rAAV and methods for preparing rAAV. For

example, plasmid(s) containing or consisting essentially of the desired viral construct are transfected into AAV-infected cells. In addition, a second or additional helper plasmid is cotransfected into these cells to provide the AAV rep and/or cap genes which are obligatory for replication and packaging of the recombinant viral construct. Under these conditions, the rep and/or cap proteins of AAV act in trans to stimulate replication and packaging of the rAAV construct. Two to Three days after transfection, rAAV is harvested. Traditionally rAAV is harvested from the cells along with adenovirus. The contaminating adenovirus is then inactivated by heat treatment. In the instant invention, rAAV is advantageously harvested not from the cells themselves, but from cell supernatant. Accordingly, in an initial aspect the invention provides for preparing rAAV, and in addition to the foregoing, rAAV can be prepared by a method that comprises or consists essentially of: infecting susceptible cells with a rAAV containing exogenous DNA including DNA for expression, and helper virus (e.g., adenovirus, herpesvirus, poxvirus such as vaccinia virus) wherein the rAAV lacks functioning cap and/or rep (and the helper virus (e.g., adenovirus, herpesvirus, poxvirus such as vaccinia virus) provides the cap and/or rev function that the rAAV lacks); or infecting susceptible cells with a rAAV containing exogenous DNA including DNA for expression, wherein the recombinant lacks functioning cap and/or rep, and transfecting said cells with a plasmid supplying cap and/or rep function that the rAAV lacks; or infecting susceptible cells with a rAAV containing exogenous DNA including DNA for expression, wherein the recombinant lacks functioning cap and/or rep, wherein said cells supply cap and/or rep function that the recombinant lacks; or transfecting the susceptible cells with an AAV lacking functioning cap and/or rep and plasmids for inserting exogenous DNA into the recombinant so that the exogenous DNA is expressed by the recombinant and for supplying rep and/or cap functions whereby transfection results in an rAAV containing the exogenous DNA including DNA for expression that lacks functioning cap and/or rep. The rAAV can be from an AAV as herein described, and advantageously can be an rAAV1, rAAV2, AAV5 or rAAV having hybrid or capsid which may comprise AAV1, AAV2, AAV5 or any combination thereof. One can select the AAV of the rAAV with regard to the cells to be targeted by the rAAV. In addition to 293 cells, other cells that can be used in the practice of the invention and the relative infectivity of certain AAV serotypes in vitro as to these cells; see Grimm, D. et al, J. Virol. 82: 5887-5911 (2008) The invention provides rAAV that contains or consists

essentially of an exogenous nucleic acid molecule encoding a CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) system or component(s) or coding therefor, e.g., a plurality of cassettes comprising or consisting a first cassette comprising or consisting essentially of a promoter, a nucleic acid molecule encoding a CRISPR-associated (Cas) protein (putative nuclease or helicase proteins), e.g., Cas9 and a terminator, and a two, or more, advantageously up to the packaging size limit of the vector, e.g., in total (including the first cassette) five, cassettes comprising or consisting essentially of a promoter, nucleic acid molecule encoding guide RNA (gRNA) and a terminator (e.g., each cassette schematically represented as Promoter-gRNA1-terminator, Promoter-gRNA2-terminator Promoter-gRNA(N)-terminator (where N is a number that can be inserted that is at an upper limit of the packaging size limit of the vector), or two or more individual rAAVs, each containing one or more than one cassette of a CRISPR system, e.g., a first rAAV containing the first cassette comprising or consisting essentially of a promoter, a nucleic acid molecule encoding Cas, e.g., Cas9 and a terminator, and a second rAAV containing a plurality, four, cassettes comprising or consisting essentially of a promoter, nucleic acid molecule encoding guide RNA (gRNA) and a terminator (e.g., each cassette schematically represented as Promoter-gRNA1-terminator, Promoter-gRNA2-terminator Promoter-gRNA(N)-terminator (where N is a number that can be inserted that is at an upper limit of the packaging size limit of the vector). As rAAV is a DNA virus, the nucleic acid molecules in the herein discussion concerning AAV or rAAV are advantageously DNA. The promoter is in some embodiments advantageously human Synapsin I promoter (hSyn).

[00244] AAV has a packaging limit of 4.5 or 4.75 Kb. This means that Cas9 as well as a promoter and transcription terminator have to be all fit into the same viral vector. Constructs larger than 4.5 or 4.75 Kb will lead to significantly reduced virus production. SpCas9 is quite large, the gene itself is over 4.1 Kb, which makes it difficult for packing into AAV. Therefore embodiments of the invention include utilizing homologs of Cas9 that are shorter. For example:

Species	Cas9 Size
<i>Corynebacter diphtheriae</i>	3252
<i>Eubacterium ventriosum</i>	3321
<i>Streptococcus pasteurianus</i>	3390

Lactobacillus farciminis	3378
Sphaerochaeta globus	3537
Azospirillum B510	3504
Gluconacetobacter diazotrophicus	3150
Neisseria cinerea	3246
Roseburia intestinalis	3420
Parvibaculum lavamentivorans	3111
Staphylococcus aureus	3159
Nitratifactor salsuginis DSM 16511	3396
Campylobacter lari CF89-12	3009
Streptococcus thermophilus LMD-9	3396

[00245] The invention also can be practiced with an adenovirus vector, e.g., an E1-, partial E3-, E4-deleted adenoviral vector may be used in the practice of the invention. Such vectors are safe as twenty-eight patients with advanced neovascular age-related macular degeneration (AMD) were given a single intravitreal injection of an E1-, partial E3-, E4-deleted adenoviral vector expressing human pigment epithelium-derived factor (AdPEDF.II) (see, e.g., Campochiaro et al., Human Gene Therapy 17:167-176 (February 2006)); and previous adenovirus doses ranging from 10⁶ to 10^{9.5} particle units (PU) can be adapted to or employed in the practice of the instant invention (see, e.g., Campochiaro et al., Human Gene Therapy 17:167-176 (February 2006)). Adenoviral vector-mediated RNA transfer appears to be a viable approach for delivery of RNA(S). For adenoviral vector injections into a rat, 2 x 10⁹ infectious particles were injected in 3 ml of normal saline solution (NSS). This can be adapted to or extrapolated from in the practice of the present invention. For siRNA, a rat was injected into the great saphenous vein with 12.5 µg of a siRNA and a primate was injected into the great saphenous vein with 750 µg of a siRNA. This can be adapted to or extrapolated from in the practice of the present invention.

[00246] In certain embodiments the Cas and/or RNA capable of guiding Cas to a genomic target locus may be delivered by lentiviral delivery systems. Lentiviruses are complex retroviruses that have the ability to infect and express their genes in both mitotic and post-mitotic

cells. The most commonly known lentivirus is the human immunodeficiency virus (HIV), which uses the envelope glycoproteins of other viruses to target a broad range of cell types.

[00247] Lentiviruses may be prepared as follows. After cloning pCasES10 (which contains a lentiviral transfer plasmid backbone), HEK293FT at low passage (p=5) were seeded in a T-75 flask to 50% confluence the day before transfection in DMEM with 10% fetal bovine serum and without antibiotics. After 20 hours, media was changed to OptiMEM (serum-free) media and transfection was done 4 hours later. Cells were transfected with 10 µg of lentiviral transfer plasmid (pCasES10) and the following packaging plasmids: 5 µg of pMD2.G (VSV-g pseudotype), and 7.5ug of psPAX2 (gag/pol/rev/tat). Transfection was done in 4mL OptiMEM with a cationic lipid delivery agent (50uL Lipofectamine 2000 and 100ul Plus reagent). After 6 hours, the media was changed to antibiotic-free DMEM with 10% fetal bovine serum. These methods use serum during cell culture, but serum-free methods are preferred.

[00248] Lentivirus may be purified as follows. Viral supernatants were harvested after 48 hours. Supernatants were first cleared of debris and filtered through a 0.45um low protein binding (PVDF) filter. They were then spun in a ultracentrifuge for 2 hours at 24,000 rpm. Viral pellets were resuspended in 50ul of DMEM overnight at 4C. They were then aliquotted and immediately frozen at -80°C.

[00249] Also useful in the practice of the invention is a minimal non-primate lentiviral vector, such as a lentiviral vector based on the equine infectious anemia virus (EIAV) (see, e.g., Balagaan, J Gene Med 2006; 8: 275 285, Published online 21 November 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jgm.845). The vectors may have cytomegalovirus (CMV) promoter driving expression of the target gene. Intracameral, subretinal, intraocular and intravitreal injections are all within the ambit of the instant invention (see, e.g., Balagaan, J Gene Med 2006; 8: 275 285, Published online 21 November 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jgm.845). In this regard, mention is made of RetinoStat®, an equine infectious anemia virus-based lentiviral gene therapy vector that expresses angiostatic proteins endostatin and angiostatin that is delivered via a subretinal injection for the treatment of the wet form of age-related macular degeneration is also contemplated (see, e.g., Binley et al., HUMAN GENE THERAPY 23:980–991 (September

2012)). Such a vector may be modified for practice of the present invention. Dosing of RetinoStat® (e.g., 1.1×10^5 transducing units per eye (TU/eye) in a total volume of 100 μ l) can be applied or extrapolated from in practicing the present invention with a lentivirus.

[00250] In certain embodiments, use is made of self-inactivating lentiviral vectors with an siRNA targeting a common exon shared by HIV tat/rev, a nucleolar-localizing TAR decoy, and an anti-CCR5-specific hammerhead ribozyme (see, e.g., DiGiusto et al. (2010) *Sci Transl Med* 2:36ra43) may be used/and or adapted to the CRISPR-Cas system of the present invention. A minimum of 2.5×10^6 CD34+ cells per kilogram patient weight may be collected and prestimulated for 16 to 20 hours in X-VIVO 15 medium (Lonza) containing 2 μ mol/L-glutamine, stem cell factor (100 ng/ml), Flt-3 ligand (Flt-3L) (100 ng/ml), and thrombopoietin (10 ng/ml) (CellGenix) at a density of 2×10^6 cells/ml. Prestimulated cells may be transduced with lentiviral at a multiplicity of infection of 5 for 16 to 24 hours in 75-cm² tissue culture flasks coated with fibronectin (25 mg/cm²) (RetroNectin, Takara Bio Inc.).

[00251] Accordingly, the invention contemplates amongst vector(s) useful in the practice of the invention: viral vectors, including retroviral vectors, lentiviral vectors, adenovirus vectors, or AAV vectors.

[00252] The Cas, for instance a Cas9, and/or any of the present RNAs, for instance a guide RNA, can also be delivered in the form of RNA. Cas mRNA can be generated using in vitro transcription. For example, Cas mRNA can be synthesized using a PCR cassette containing the following elements: T7_promoter-kozak sequence (GCCACC)-Cas9-3' UTR from beta globin-polyA tail (a string of 120 or more adenines). The cassette can be used for transcription by T7 polymerase. Guide RNAs can also be transcribed using in vitro transcription from a cassette containing T7_promoter-GG-guide RNA sequence.

[00253] Several types of particle and nanoparticle delivery systems and/or formulations are known to be useful in a diverse spectrum of biomedical applications; and particle and nanoparticle delivery systems in the practice of the instant invention can be as in WO 2014/093622 (PCT/US13/74667). In general, a particle is defined as a small object that behaves as a whole unit with respect to its transport and properties. Particles are further classified according to diameter. Coarse particles cover a range between 2,500 and 10,000 nanometers.

Fine particles are sized between 100 and 2,500 nanometers. Ultrafine particles, or nanoparticles, are generally between 1 and 100 nanometers in size. The basis of the 100-nm limit is the fact that novel properties that differentiate particles from the bulk material typically develop at a critical length scale of under 100 nm. As used herein, a particle delivery system/formulation is defined as any biological delivery system/formulation which includes a particle in accordance with the present invention. A particle in accordance with the present invention is any entity having a greatest dimension (e.g. diameter) of less than 100 microns (μm). In some embodiments, inventive particles have a greatest dimension of less than 10 μm . In some embodiments, inventive particles have a greatest dimension of less than 2000 nanometers (nm). In some embodiments, inventive particles have a greatest dimension of less than 1000 nanometers (nm). In some embodiments, inventive particles have a greatest dimension of less than 900 nm, 800 nm, 700 nm, 600 nm, 500 nm, 400 nm, 300 nm, 200 nm, or 100 nm. Typically, inventive particles have a greatest dimension (e.g., diameter) of 500 nm or less. In some embodiments, inventive particles have a greatest dimension (e.g., diameter) of 250 nm or less. In some embodiments, inventive particles have a greatest dimension (e.g., diameter) of 200 nm or less. In some embodiments, inventive particles have a greatest dimension (e.g., diameter) of 150 nm or less. In some embodiments, inventive particles have a greatest dimension (e.g., diameter) of 100 nm or less. Smaller particles, e.g., having a greatest dimension of 50 nm or less are used in some embodiments of the invention. In some embodiments, inventive particles have a greatest dimension ranging between 25 nm and 200 nm. Particle characterization (including e.g., characterizing morphology, dimension, etc.) is done using a variety of different techniques. Common techniques are electron microscopy (TEM, SEM), atomic force microscopy (AFM), dynamic light scattering (DLS), X-ray photoelectron spectroscopy (XPS), powder X-ray diffraction (XRD), Fourier transform infrared spectroscopy (FTIR), matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF), ultraviolet-visible spectroscopy, dual polarisation interferometry and nuclear magnetic resonance (NMR). Characterization (dimension measurements) may be made as to native particles (i.e., preloading) or after loading of the cargo (herein cargo refers to e.g., one or more components of CRISPR-Cas system e.g., Cas enzyme or mRNA or guide RNA, or any combination thereof, and may include additional carriers and/or excipients) to provide particles of an optimal size for delivery

according to certain embodiments of the present invention. In certain preferred embodiments, particle dimension (e.g., diameter) characterization is based on measurements using dynamic laser scattering (DLS). Particles delivery systems within the scope of the present invention may be provided in any form, including but not limited to solid, semi-solid, emulsion, or colloidal particles. As such any of the delivery systems described herein, including but not limited to, e.g., lipid-based systems, liposomes, micelles, microvesicles, exosomes, or gene gun may be provided as particle delivery systems within the scope of the present invention.

[00254] In general, a “nanoparticle” refers to any particle having a diameter of less than 1000 nm. In certain preferred embodiments, nanoparticles have a greatest dimension (e.g., diameter) of 500 nm or less. In other preferred embodiments, nanoparticles have a greatest dimension ranging between 25 nm and 200 nm. In other preferred embodiments, nanoparticles have a greatest dimension of 100 nm or less. In other preferred embodiments, nanoparticles have a greatest dimension ranging between 35 nm and 60 nm. Nanoparticles encompassed in the present invention may be provided in different forms, e.g., as solid nanoparticles (e.g., metal such as silver, gold, iron, titanium), non-metal, lipid-based solids, polymers), suspensions of nanoparticles, or combinations thereof. Metal, dielectric, and semiconductor nanoparticles may be prepared, as well as hybrid structures (e.g., core-shell nanoparticles). Nanoparticles made of semiconducting material may also be labeled quantum dots if they are small enough (typically sub 10 nm) that quantization of electronic energy levels occurs. Such nanoscale particles are used in biomedical applications as drug carriers or imaging agents and may be adapted for similar purposes in the present invention.

[00255] With regard to particles that can deliver RNA, see, e.g., Alabi et al., *Proc Natl Acad Sci U S A*. 2013 Aug 6;110(32):12881-6; Zhang et al., *Adv Mater*. 2013 Sep 6;25(33):4641-5; Jiang et al., *Nano Lett*. 2013 Mar 13;13(3):1059-64; Karagiannis et al., *ACS Nano*. 2012 Oct 23;6(10):8484-7; Whitehead et al., *ACS Nano*. 2012 Aug 28;6(8):6922-9 and Lee et al., *Nat Nanotechnol*. 2012 Jun 3;7(6):389-93. Lipid particles, Spherical Nucleic Acid (SNATM) constructs, nanoplexes and other particles (particularly gold particles) are also contemplated as a means for delivery of CRISPR/Cas system or component(s) thereof or vector(s) to intended targets. Particles, nanoparticles, and the like and vectors are advantageous for delivering the

RNA(s) of the CRISPR-Cas system and particles and nanoparticles and the like may be advantageous for delivery of vector containing nucleic acid(s) encoding or comprising RNA(s) of the invention. In certain instances, e.g., where Cas is constitutively or inducibly or conditionally expressed by an organism or cells thereof, it is useful to deliver the RNA(s) (also herein sometimes termed “guides”) of the CRISPR-Cas system separately from the Cas9. It is considered as advantageous that the Cas may be delivered via a viral vector or be constitutively or inducibly or conditionally expressed and that guides specific to genomic targets are delivered separately. A recent publication, entitled “In vivo endothelial siRNA delivery using polymeric nanoparticles with low molecular weight” by James E. Dahlman and Carmen Barnes et al. *Nature Nanotechnology* (2014) published online 11 May 2014, doi:10.1038/nnano.2014.84, incorporated herein in its entirety, showed that polymeric particles made of low-molecular-weight polyamines and lipids can deliver siRNA to endothelial cells with high efficiency, thereby facilitating the simultaneous silencing of multiple endothelial genes in vivo. The authors reported that unlike lipid or lipid-like particles, the particle formulation they used (termed 7C1), differed from traditional lipid-based particle formulations because it can deliver siRNA to lung endothelial cells at low doses without substantially reducing gene expression in pulmonary immune cells, hepatocytes or peritoneal immune cells. The study further demonstrated that that 7C1-mediated endothelial gene silencing affects function in vivo, by using the nanoformulation to modify mouse models of vascular permeability, emphysema, lung tumor growth and lung metastasis.

[00256] In some embodiments, the Cas is part of a fusion protein (i.e. chimeric protein) comprising one or more heterologous protein domains (e.g. about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more domains in addition to the Cas). A Cas fusion protein may comprise any additional protein sequence, and optionally a linker sequence between any two domains. Examples of protein domains that may be fused to a Cas include, without limitation, epitope tags, reporter gene sequences, and protein domains having one or more of the following activities: methylase activity, demethylase activity, transcription activation activity, transcription repression activity, transcription release factor activity, histone modification activity, RNA cleavage activity and nucleic acid binding activity. In such embodiments, it is preferred that Cas itself is catalytically inactive, or partially catalytically inactive. Non-limiting examples of epitope tags

include histidine (His) tags, V5 tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Examples of reporter genes include, but are not limited to, glutathione-S-transferase (GST), horseradish peroxidase (HRP), chloramphenicol acetyltransferase (CAT) beta-galactosidase, beta-glucuronidase, luciferase, green fluorescent protein (GFP), HcRed, DsRed, cyan fluorescent protein (CFP), yellow fluorescent protein (YFP), and autofluorescent proteins including blue fluorescent protein (BFP). A Cas may be fused to a gene sequence encoding a protein or a fragment of a protein that bind DNA molecules or bind other cellular molecules, including but not limited to maltose binding protein (MBP), S-tag, Lex A DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. Additional domains that may form part of a fusion protein comprising a Cas are described in US20110059502, incorporated herein by reference. In some embodiments, a tagged Cas is used to identify the location of a target sequence. In certain embodiments, Cas is fused to a heterologous protein capable of manipulating a target sequence. By manipulation of a target sequence, Applicants mean the alteration of the target sequence, which may include the epigenetic manipulation of a target sequence. This epigenetic manipulation may be of the chromatin state of a target sequence, such as by modification of the methylation state of the target sequence (i.e. addition or removal of methylation or methylation patterns or CpG islands), histone modification, increasing or reducing accessibility to the target sequence, or by promoting 3D folding.

[00257] In some embodiments, a Cas *f* sequence and/or RNA capable of guiding Cas to a genomic target locus may form a component of an inducible system. The inducible nature of the system would allow for spatiotemporal control of gene editing or gene expression using a form of energy. The form of energy may include but is not limited to electromagnetic radiation, sound energy, chemical energy and thermal energy. Examples of inducible system include tetracycline inducible promoters (Tet-On or Tet-Off), small molecule two-hybrid transcription activations systems (FKBP, ABA, etc), or light inducible systems (Phytochrome, LOV domains, or cryptochrome). In one embodiment, the Cas sequence and/or RNA capable of guiding Cas to a genomic target locus may be a part of a Light Inducible Transcriptional Effector (LITE) to direct changes in transcriptional activity in a sequence-specific manner. The components of a LITE may include a Cas sequence and/or RNA capable of guiding Cas to a genomic target locus, a

light-responsive cytochrome heterodimer (e.g. from *Arabidopsis thaliana*), and a transcriptional activation/repression domain. Further examples of inducible DNA binding proteins and methods for their use are provided in US 61/736465 and US 61/721,283, which is hereby incorporated by reference in its entirety.

[00258] In certain embodiments, one or more of the components of the CRISPR-Cas system may be conditionally (e.g. tissue or cell type specific) and/or inducibly (e.g. chemically inducible) expressed in the cell. Inducible and conditional expression systems are described herein elsewhere. In particular embodiments, one or more of the guide RNA(s) may be conditionally and/or inducibly expressed in the cell. In particular preferred embodiments, the Cas may be conditionally and/or inducibly expressed in the cell.

In an aspect the invention also features a non-naturally occurring or engineered inducible Cas9 CRISPR-Cas system, comprising: a first Cas9 fusion construct attached to a first half of an inducible dimer and a second Cas9 fusion construct attached to a second half of the inducible dimer, wherein the first Cas9 fusion construct is operably linked to one or more nuclear localization signals, wherein the second Cas9 fusion construct is operably linked to one or more nuclear export signals, wherein contact with an inducer energy source brings the first and second halves of the inducible dimer together, wherein bringing the first and second halves of the inducible dimer together allows the first and second Cas9 fusion constructs to constitute a functional Cas9 CRISPR-Cas system, wherein the Cas9 CRISPR-Cas system comprises a guide RNA (gRNA) comprising a guide sequence capable of hybridizing to a target sequence in a genomic locus of interest in a cell, and wherein the functional Cas9 CRISPR-Cas system binds to the target sequence and, optionally, edits the genomic locus to alter gene expression.

[00259] In an aspect of the invention in the inducible Cas9 CRISPR-Cas system, the inducible dimer is or comprises or consists essentially of or consists of an inducible heterodimer. In an aspect, in inducible Cas9 CRISPR-Cas system, the first half or a first portion or a first fragment of the inducible heterodimer is or comprises or consists of or consists essentially of an FKBP, optionally FKBP12. In an aspect of the invention, in the inducible Cas9 CRISPR-Cas system, the second half or a second portion or a second fragment of the inducible heterodimer is or comprises or consists of or consists essentially of FRB. In an aspect of the invention, in the inducible Cas9 CRISPR-Cas system, the arrangement of the first Cas9 fusion construct is or

comprises or consists of or consists essentially of N' terminal Cas9 part-FRB-NES. In an aspect of the invention, in the inducible Cas9 CRISPR-Cas system, the arrangement of the first Cas9 fusion construct is or comprises or consists of or consists essentially of NES-N' terminal Cas9 part-FRB-NES. In an aspect of the invention, in the inducible Cas9 CRISPR-Cas system, the arrangement of the second Cas9 fusion construct is or comprises or consists essentially of or consists of C' terminal Cas9 part-FKBP-NLS. In an aspect the invention provides in the inducible Cas9 CRISPR-Cas system, the arrangement of the second Cas9 fusion construct is or comprises or consists of or consists essentially of NLS-C' terminal Cas9 part-FKBP-NLS. In an aspect, in inducible Cas9 CRISPR-Cas system there can be a linker that separates the Cas9 part from the half or portion or fragment of the inducible dimer. In an aspect, in the inducible Cas9 CRISPR-Cas system, the inducer energy source is or comprises or consists essentially of or consists of rapamycin. In an aspect, in inducible Cas9 CRISPR-Cas system, the inducible dimer is an inducible homodimer. In an aspect, in inducible Cas9 CRISPR-Cas system, the Cas9 is FnCas9. In an aspect, in the inducible Cas9 CRISPR-Cas system, one or more functional domains are associated with one or both parts of the Cas9, e.g., the functional domains optionally including a transcriptional activator, a transcriptional or a nuclease such as a FokI nuclease. In an aspect, in the inducible Cas9 CRISPR-Cas system, the functional Cas9 CRISPR-Cas system binds to the target sequence and the enzyme is a dead-Cas9, optionally having a diminished nuclease activity of at least 97%, or 100% (or no more than 3% and advantageously 0% nuclease activity) as compared with the Cas9 not having the at least one mutation. The invention further comprehends and an aspect of the invention provides, a polynucleotide encoding the inducible Cas9 CRISPR-Cas system as herein discussed.

[00260] In an aspect, the invention provides a vector for delivery of the first Cas9 fusion construct, attached to a first half or portion or fragment of an inducible dimer and operably linked to one or more nuclear localization signals, according as herein discussed. In an aspect, the invention provides a vector for delivery of the second Cas9 fusion construct, attached to a second half or portion or fragment of an inducible dimer and operably linked to one or more nuclear export signals.

[00261] In an aspect, the invention provides a vector for delivery of both: the first Cas9 fusion construct, attached to a first half or portion or fragment of an inducible dimer and operably

linked to one or more nuclear localization signals, as herein discussed; and the second Cas9 fusion construct, attached to a second half or portion or fragment of an inducible dimer and operably linked to one or more nuclear export signals, as herein discussed.

[00262] In an aspect, the vector can be single plasmid or expression cassette.

[00263] The invention, in an aspect, provides a eukaryotic host cell or cell line transformed with any of the vectors herein discussed or expressing the inducible Cas9 CRISPR-Cas system as herein discussed.

[00264] The invention, in an aspect provides, a transgenic organism transformed with any of the vectors herein discussed or expressing the inducible Cas9 CRISPR-Cas system herein discussed, or the progeny thereof. In an aspect, the invention provides a model organism which constitutively expresses the inducible Cas9 CRISPR-Cas system as herein discussed.

In an aspect, the invention provides non-naturally occurring or engineered inducible Cas9 CRISPR-Cas system, comprising: a first Cas9 fusion construct attached to a first half of an inducible heterodimer and a second Cas9 fusion construct attached to a second half of the inducible heterodimer, wherein the first Cas9 fusion construct is operably linked to one or more nuclear localization signals, wherein the second Cas9 fusion construct is operably linked to a nuclear export signal, wherein contact with an inducer energy source brings the first and second halves of the inducible heterodimer together, wherein bringing the first and second halves of the inducible heterodimer together allows the first and second Cas9 fusion constructs to constitute a functional Cas9 CRISPR-Cas system, wherein the Cas9 CRISPR-Cas system comprises a guide RNA (gRNA) comprising a guide sequence capable of hybridizing to a target sequence in a genomic locus of interest in a cell, and wherein the functional Cas9 CRISPR-Cas system edits the genomic locus to alter gene expression.

[00265] In an aspect, the invention provides a method of treating a subject in need thereof, comprising inducing gene editing by transforming the subject with the polynucleotide as herein discussed or any of the vectors herein discussed and administering an inducer energy source to the subject. The invention comprehends uses of such a polynucleotide or vector in the manufacture of a medicament, e.g., such a medicament for treating a subject or for such a method of treating a subject. The invention comprehends the polynucleotide as herein discussed or any of the vectors herein discussed for use in a method of treating a subject in need thereof

comprising inducing gene editing, wherein the method further comprises administering an inducer energy source to the subject. In an aspect, in the method, a repair template is also provided, for example delivered by a vector comprising said repair template.

[00266] The invention also provides a method of treating a subject in need thereof, comprising inducing transcriptional activation or repression by transforming the subject with the polynucleotide herein discussed or any of the vectors herein discussed, wherein said polynucleotide or vector encodes or comprises the catalytically inactive Cas9 and one or more associated functional domains as herein discussed; the method further comprising administering an inducer energy source to the subject. The invention also provides the polynucleotide herein discussed or any of the vectors herein discussed for use in a method of treating a subject in need thereof comprising inducing transcriptional activation or repression, wherein the method further comprises administering an inducer energy source to the subject.

[00267] Accordingly, the invention comprehends *inter alia* homodimers as well as heterodimers, dead-Cas9 or Cas9 having essentially no nuclease activity, e.g., through mutation, systems or complexes wherein there is one or more NLS and/or one or more NES; functional domain(s) linked to split Cas9; methods, including methods of treatment, and uses.

[00268] It will be appreciated that where reference is made herein to Cas9, Cas9 protein or Cas9 enzyme, this includes the present split Cas9. In one aspect, the invention provides a method for altering or modifying expression of a gene product. The said method may comprise introducing into a cell containing and expressing a DNA molecule encoding the gene product an engineered, non-naturally occurring Cas9 CRISPR-Cas system comprising a Cas9 protein and guide RNA that targets the DNA molecule, whereby the guide RNA targets the DNA molecule encoding the gene product and the Cas9 protein cleaves the DNA molecule encoding the gene product, whereby expression of the gene product is altered; and, wherein the Cas9 protein and the guide RNA do not naturally occur together. The invention comprehends the guide RNA comprising a guide sequence linked to a direct repeat (DR) sequence. The invention further comprehends the Cas9 protein being codon optimized for expression in a eukaryotic cell. In a preferred embodiment the eukaryotic cell is a mammalian cell and in a more preferred embodiment the mammalian cell is a human cell. In a further embodiment of the invention, the expression of the gene product is decreased.

[00269] In one aspect, the invention provides an engineered, non-naturally occurring Cas9 CRISPR-Cas system comprising a Cas9 protein and a guide RNA that targets a DNA molecule encoding a gene product in a cell, whereby the guide RNA targets the DNA molecule encoding the gene product and the Cas9 protein cleaves the DNA molecule encoding the gene product, whereby expression of the gene product is altered; and, wherein the Cas9 protein and the guide RNA do not naturally occur together; this including the present split Cas9. The invention comprehends the guide RNA comprising a guide sequence linked to a DR sequence. The invention further comprehends the Cas9 protein being codon optimized for expression in a eukaryotic cell. In a preferred embodiment the eukaryotic cell is a mammalian cell and in a more preferred embodiment the mammalian cell is a human cell. In a further embodiment of the invention, the expression of the gene product is decreased.

[00270] In another aspect, the invention provides an engineered, non-naturally occurring vector system comprising one or more vectors comprising a first regulatory element operably linked to a Cas9 CRISPR-Cas system guide RNA that targets a DNA molecule encoding a gene product and a second regulatory element operably linked to a Cas9 protein; this includes the present split Cas9. Components (a) and (b) may be located on same or different vectors of the system. The guide RNA targets the DNA molecule encoding the gene product in a cell and the Cas9 protein cleaves the DNA molecule encoding the gene product, whereby expression of the gene product is altered; and, wherein the Cas9 protein and the guide RNA do not naturally occur together. The invention comprehends the guide RNA comprising a guide sequence linked to a DR sequence. The invention further comprehends the Cas9 protein being codon optimized for expression in a eukaryotic cell. In a preferred embodiment the eukaryotic cell is a mammalian cell and in a more preferred embodiment the mammalian cell is a human cell. In a further embodiment of the invention, the expression of the gene product is decreased.

[00271] In one aspect, the invention provides a vector system comprising one or more vectors. In some embodiments, the system comprises: (a) a first regulatory element operably linked to a DR sequence and one or more insertion sites for inserting one or more guide sequences downstream of the DR sequence, wherein when expressed, the guide sequence directs sequence-specific binding of a Cas9 CRISPR-Cas complex to a target sequence in a eukaryotic cell, wherein the Cas9 CRISPR-Cas complex comprises Cas9 complexed with (1) the guide sequence

that is hybridized to the target sequence, and (2) the DR sequence; and (b) a second regulatory element operably linked to an enzyme-coding sequence encoding said Cas9 enzyme comprising a nuclear localization sequence; wherein components (a) and (b) are located on the same or different vectors of the system; this includes the present split Cas9. In some embodiments, component (a) further comprises two or more guide sequences operably linked to the first regulatory element, wherein when expressed, each of the two or more guide sequences direct sequence specific binding of a Cas9 CRISPR-Cas complex to a different target sequence in a eukaryotic cell.

[00272] In some embodiments, the Cas9 CRISPR-Cas complex comprises one or more nuclear localization sequences of sufficient strength to drive accumulation of said Cas9 CRISPR-Cas complex in a detectable amount in the nucleus of a eukaryotic cell. Without wishing to be bound by theory, it is believed that a nuclear localization sequence is not necessary for Cas9 CRISPR-Cas complex activity in eukaryotes, but that including such sequences enhances activity of the system, especially as to targeting nucleic acid molecules in the nucleus.

[00273] In some embodiments, the Cas9 enzyme is Cas9 of a bacterial species selected from the group consisting of *Francisella tularensis* 1, *Francisella tularensis* subsp. *novicida*, *Prevotella albensis*, *Lachnospiraceae* bacterium MC2017 1, *Butyrivibrio proteoclasticus*, *Peregrinibacteria* bacterium GW2011_GWA2_33_10, *Parcubacteria* bacterium GW2011_GWC2_44_17, *Smithella* sp. SCADC, *Acidaminococcus* sp. BV3L6, *Lachnospiraceae* bacterium MA2020, *Candidatus Methanoplasma termitum*, *Eubacterium eligens*, *Moraxella bovoculi* 237, *Leptospira inadai*, *Lachnospiraceae* bacterium ND2006, *Porphyromonas crevioricanis* 3, *Prevotella disiens*, and *Porphyromonas macacae*, and may include mutated Cas9 derived from these organisms. The enzyme may be a Cas9 homolog or ortholog. In some embodiments, the Cas9 is codon-optimized for expression in a eukaryotic cell. In some embodiments, the Cas9 directs cleavage of one or two strands at the location of the target sequence. In a preferred embodiment, the strand break is a staggered cut with a 5' overhang. In some embodiments, the first regulatory element is a polymerase III promoter. In some embodiments, the second regulatory element is a polymerase II promoter. In some embodiments, the direct repeat has a minimum length of 16 nts and a single stem loop. In further embodiments the direct repeat has a length longer than 16 nts, preferably more than 17 nts, and has more than one stem loop or optimized secondary structures.

[00274] In one aspect, the invention provides a eukaryotic host cell comprising (a) a first regulatory element operably linked to a direct repeat sequence and one or more insertion sites for inserting one or more guide sequences downstream of the DR sequence, wherein when expressed, the guide sequence directs sequence-specific binding of a Cas9 CRISPR-Cas complex to a target sequence in a eukaryotic cell, wherein the Cas9 CRISPR-Cas complex comprises Cas9 complexed with (1) the guide sequence that is hybridized to the target sequence, and (2) the DR sequence; and/or (b) a second regulatory element operably linked to an enzyme-coding sequence encoding said Cas9 enzyme comprising a nuclear localization sequence. In some embodiments, the host cell comprises components (a) and (b); this includes the present split Cas9. In some embodiments, component (a), component (b), or components (a) and (b) are stably integrated into a genome of the host eukaryotic cell. In some embodiments, component (a) further comprises two or more guide sequences operably linked to the first regulatory element, wherein when expressed, each of the two or more guide sequences direct sequence specific binding of a Cas9 CRISPR-Cas complex to a different target sequence in a eukaryotic cell. In some embodiments, the Cas9 is codon-optimized for expression in a eukaryotic cell. In some embodiments, the Cas9 directs cleavage of one or two strands at the location of the target sequence. In a preferred embodiment, the strand break is a staggered cut with a 5' overhang. In some embodiments, the Cas9 lacks DNA strand cleavage activity. In some embodiments, the first regulatory element is a polymerase III promoter. In some embodiments, the direct repeat has a minimum length of 16 nts and a single stem loop. In further embodiments the direct repeat has a length longer than 16 nts, preferably more than 17 nts, and has more than one stem loop or optimized secondary structures. In an aspect, the invention provides a non-human eukaryotic organism; preferably a multicellular eukaryotic organism, comprising a eukaryotic host cell according to any of the described embodiments. In other aspects, the invention provides a eukaryotic organism; preferably a multicellular eukaryotic organism, comprising a eukaryotic host cell according to any of the described embodiments. The organism in some embodiments of these aspects may be an animal; for example a mammal. Also, the organism may be an arthropod such as an insect. The organism also may be a plant. Further, the organism may be a fungus.

[00275] In one aspect, the invention provides a kit comprising one or more of the components described herein. In some embodiments, the kit comprises a vector system and instructions for

using the kit. In some embodiments, the vector system comprises (a) a first regulatory element operably linked to a direct repeat sequence and one or more insertion sites for inserting one or more guide sequences downstream of the DR sequence, wherein when expressed, the guide sequence directs sequence-specific binding of a Cas9 CRISPR-Cas complex to a target sequence in a eukaryotic cell, wherein the Cas9 CRISPR-Cas complex comprises Cas9 complexed with (1) the guide sequence that is hybridized to the target sequence, and (2) the DR sequence; and/or (b) a second regulatory element operably linked to an enzyme-coding sequence encoding said Cas9 enzyme comprising a nuclear localization sequence and advantageously this includes the present split Cas9. In some embodiments, the kit comprises components (a) and (b) located on the same or different vectors of the system. In some embodiments, component (a) further comprises two or more guide sequences operably linked to the first regulatory element, wherein when expressed, each of the two or more guide sequences direct sequence specific binding of a Cas9 CRISPR-Cas complex to a different target sequence in a eukaryotic cell. In some embodiments, the Cas9 comprises one or more nuclear localization sequences of sufficient strength to drive accumulation of said Cas9 in a detectable amount in the nucleus of a eukaryotic cell. In some embodiments, the Cas9 enzyme is Cas9 of a bacterial species selected from the group consisting of *Francisella tularensis* 1, *Francisella tularensis* subsp. *novicida*, *Prevotella albensis*, *Lachnospiraceae* bacterium MC2017 1, *Butyrivibrio proteoclasticus*, *Peregrinibacteria* bacterium GW2011_GWA2_33_10, *Parcubacteria* bacterium GW2011_GWC2_44_17, *Smithella* sp. SCADC, *Acidaminococcus* sp. BV3L6, *Lachnospiraceae* bacterium MA2020, *Candidatus Methanoplasma termitum*, *Eubacterium eligens*, *Moraxella bovoculi* 237, *Leptospira inadai*, *Lachnospiraceae* bacterium ND2006, *Porphyromonas crevioricanis* 3, *Prevotella disiens*, and *Porphyromonas macacae*, and may include mutated Cas9 derived from these organisms. The enzyme may be a Cas9 homolog or ortholog. In some embodiments, the Cas9 is codon-optimized for expression in a eukaryotic cell. In some embodiments, the Cas9 directs cleavage of one or two strands at the location of the target sequence. In a preferred embodiment, the strand break is a staggered cut with a 5' overhang. In some embodiments, the CRISPR enzyme lacks DNA strand cleavage activity. In some embodiments, the direct repeat has a minimum length of 16 nts and a single stem loop. In further embodiments the direct repeat has a length longer than 16 nts, preferably more than 17 nts, and has more than one stem loop or optimized secondary structures.

[00276] In one aspect, the invention provides a method of modifying a target polynucleotide in a eukaryotic cell. In some embodiments, the method comprises allowing a Cas9 CRISPR-Cas complex to bind to the target polynucleotide to effect cleavage of said target polynucleotide thereby modifying the target polynucleotide, wherein the Cas9 CRISPR-Cas complex comprises Cas9 complexed with a guide sequence hybridized to a target sequence within said target polynucleotide, wherein said guide sequence is linked to a direct repeat sequence. In some embodiments, said cleavage comprises cleaving one or two strands at the location of the target sequence by said Cas9; this includes the present split Cas9. In some embodiments, said cleavage results in decreased transcription of a target gene. In some embodiments, the method further comprises repairing said cleaved target polynucleotide by homologous recombination with an exogenous template polynucleotide, wherein said repair results in a mutation comprising an insertion, deletion, or substitution of one or more nucleotides of said target polynucleotide. In some embodiments, said mutation results in one or more amino acid changes in a protein expressed from a gene comprising the target sequence. In some embodiments, the method further comprises delivering one or more vectors to said eukaryotic cell, wherein the one or more vectors drive expression of one or more of: the Cas9, and the guide sequence linked to the DR sequence. In some embodiments, said vectors are delivered to the eukaryotic cell in a subject. In some embodiments, said modifying takes place in said eukaryotic cell in a cell culture. In some embodiments, the method further comprises isolating said eukaryotic cell from a subject prior to said modifying. In some embodiments, the method further comprises returning said eukaryotic cell and/or cells derived therefrom to said subject.

[00277] In one aspect, the invention provides a method of modifying expression of a polynucleotide in a eukaryotic cell. In some embodiments, the method comprises allowing a Cas9 CRISPR-Cas complex to bind to the polynucleotide such that said binding results in increased or decreased expression of said polynucleotide; wherein the Cas9 CRISPR-Cas complex comprises Cas9 complexed with a guide sequence hybridized to a target sequence within said polynucleotide, wherein said guide sequence is linked to a direct repeat sequence; this includes the present split Cas9. In some embodiments, the method further comprises delivering one or more vectors to said eukaryotic cells, wherein the one or more vectors drive expression of one or more of: the Cas9, and the guide sequence linked to the DR sequence.

[00278] In one aspect, the invention provides a method of generating a model eukaryotic cell comprising a mutated disease gene. In some embodiments, a disease gene is any gene associated an increase in the risk of having or developing a disease. In some embodiments, the method comprises (a) introducing one or more vectors into a eukaryotic cell, wherein the one or more vectors drive expression of one or more of: Cas9, and a guide sequence linked to a direct repeat sequence; and (b) allowing a Cas9 CRISPR-Cas complex to bind to a target polynucleotide to effect cleavage of the target polynucleotide within said disease gene, wherein the Cas9 CRISPR-Cas complex comprises the Cas9 complexed with (1) the guide sequence that is hybridized to the target sequence within the target polynucleotide, and (2) the DR sequence, thereby generating a model eukaryotic cell comprising a mutated disease gene; this includes the present split Cas9. In some embodiments, said cleavage comprises cleaving one or two strands at the location of the target sequence by said Cas9. In a preferred embodiment, the strand break is a staggered cut with a 5' overhang. In some embodiments, said cleavage results in decreased transcription of a target gene. In some embodiments, the method further comprises repairing said cleaved target polynucleotide by homologous recombination with an exogenous template polynucleotide, wherein said repair results in a mutation comprising an insertion, deletion, or substitution of one or more nucleotides of said target polynucleotide. In some embodiments, said mutation results in one or more amino acid changes in a protein expression from a gene comprising the target sequence.

[00279] In one aspect, the invention provides a method for developing a biologically active agent that modulates a cell signaling event associated with a disease gene. In some embodiments, a disease gene is any gene associated an increase in the risk of having or developing a disease. In some embodiments, the method comprises (a) contacting a test compound with a model cell of any one of the described embodiments; and (b) detecting a change in a readout that is indicative of a reduction or an augmentation of a cell signaling event associated with said mutation in said disease gene, thereby developing said biologically active agent that modulates said cell signaling event associated with said disease gene.

[00280] In one aspect, the invention provides a recombinant polynucleotide comprising a guide sequence downstream of a direct repeat sequence, wherein the guide sequence when expressed directs sequence-specific binding of a Cas9 CRISPR-Cas complex to a corresponding

target sequence present in a eukaryotic cell. In some embodiments, the target sequence is a viral sequence present in a eukaryotic cell. In some embodiments, the target sequence is a proto-oncogene or an oncogene.

[00281] In one aspect the invention provides for a method of selecting one or more cell(s) by introducing one or more mutations in a gene in the one or more cell (s), the method comprising: introducing one or more vectors into the cell (s), wherein the one or more vectors drive expression of one or more of: Cas9, a guide sequence linked to a direct repeat sequence, and an editing template; wherein the editing template comprises the one or more mutations that abolish Cas9 cleavage; allowing homologous recombination of the editing template with the target polynucleotide in the cell(s) to be selected; allowing a Cas9 CRISPR-Cas complex to bind to a target polynucleotide to effect cleavage of the target polynucleotide within said gene, wherein the Cas9 CRISPR-Cas complex comprises the Cas9 complexed with (1) the guide sequence that is hybridized to the target sequence within the target polynucleotide, and (2) the direct repeat sequence, wherein binding of the Cas9 CRISPR-Cas complex to the target polynucleotide induces cell death, thereby allowing one or more cell(s) in which one or more mutations have been introduced to be selected; this includes the present split Cas9. In another preferred embodiment of the invention the cell to be selected may be a eukaryotic cell. Aspects of the invention allow for selection of specific cells without requiring a selection marker or a two-step process that may include a counter-selection system.

[00282] Herein there is the phrase “this includes the present split Cas9” or similar text; and, this is to indicate that Cas9 in embodiments herein can be a split Cas9 as herein discussed.

[00283] In an aspect the invention involves a non-naturally occurring or engineered inducible Cas9 CRISPR-Cas system, comprising a first Cas9 fusion construct attached to a first half of an inducible heterodimer and a second Cas9 fusion construct attached to a second half of the inducible heterodimer, wherein the first Cas9 fusion construct is operably linked to one or more nuclear localization signals, wherein the second Cas9 fusion construct is operably linked to a nuclear export signal, wherein contact with an inducer energy source brings the first and second halves of the inducible heterodimer together, wherein bringing the first and second halves of the inducible heterodimer together allows the first and second Cas9 fusion constructs to constitute a functional Cas9 CRISPR-Cas system, wherein the Cas9 CRISPR-Cas system comprises a guide

RNA (gRNA) comprising a guide sequence capable of hybridizing to a target sequence in a genomic locus of interest in a cell, and wherein the functional Cas9 CRISPR-Cas system edits the genomic locus to alter gene expression. In an embodiment of the invention the first half of the inducible heterodimer is FKBP12 and the second half of the inducible heterodimer is FRB. In another embodiment of the invention the inducer energy source is rapamycin.

[00284] An inducer energy source may be considered to be simply an inducer or a dimerizing agent. The term 'inducer energy source' is used herein throughout for consistency. The inducer energy source (or inducer) acts to reconstitute the Cas9. In some embodiments, the inducer energy source brings the two parts of the Cas9 together through the action of the two halves of the inducible dimer. The two halves of the inducible dimer therefore are brought together in the presence of the inducer energy source. The two halves of the dimer will not form into the dimer (dimerize) without the inducer energy source.

[00285] Thus, the two halves of the inducible dimer cooperate with the inducer energy source to dimerize the dimer. This in turn reconstitutes the Cas9 by bringing the first and second parts of the Cas9 together.

[00286] The CRISPR enzyme fusion constructs each comprise one part of the split Cas9. These are fused, preferably via a linker such as a GlySer linker described herein, to one of the two halves of the dimer. The two halves of the dimer may be substantially the same two monomers that together form the homodimer, or they may be different monomers that together form the heterodimer. As such, the two monomers can be thought of as one half of the full dimer.

[00287] The Cas9 is split in the sense that the two parts of the Cas9 enzyme substantially comprise a functioning Cas9. That Cas9 may function as a genome editing enzyme (when forming a complex with the target DNA and the guide), such as a nickase or a nuclease (cleaving both strands of the DNA), or it may be a dead-Cas9 which is essentially a DNA-binding protein with very little or no catalytic activity, due to typically mutation(s) in its catalytic domains.

[00288] The two parts of the split Cas9 can be thought of as the N' terminal part and the C' terminal part of the split Cas9. The fusion is typically at the split point of the Cas9. In other words, the C' terminal of the N' terminal part of the split Cas9 is fused to one of the dimer halves, whilst the N' terminal of the C' terminal part is fused to the other dimer half.

[00289] The Cas9 does not have to be split in the sense that the break is newly created. The split point is typically designed in silico and cloned into the constructs. Together, the two parts of the split Cas9, the N' terminal and C' terminal parts, form a full Cas9, comprising preferably at least 70% or more of the wildtype amino acids (or nucleotides encoding them), preferably at least 80% or more, preferably at least 90% or more, preferably at least 95% or more, and most preferably at least 99% or more of the wildtype amino acids (or nucleotides encoding them). Some trimming may be possible, and mutants are envisaged. Non-functional domains may be removed entirely. What is important is that the two parts may be brought together and that the desired Cas9 function is restored or reconstituted.

[00290] The dimer may be a homodimer or a heterodimer.

[00291] One or more, preferably two, NLSs may be used in operable linkage to the first Cas9 construct. One or more, preferably two, NESs may be used in operable linkage to the first Cas9 construct. The NLSs and/or the NESs preferably flank the split Cas9-dimer (i.e., half dimer) fusion, i.e., one NLS may be positioned at the N' terminal of the first Cas9 construct and one NLS may be at the C' terminal of the first Cas9 construct. Similarly, one NES may be positioned at the N' terminal of the second Cas9 construct and one NES may be at the C' terminal of the second Cas9 construct. Where reference is made to N' or C' terminals, it will be appreciated that these correspond to 5' and 3' ends in the corresponding nucleotide sequence.

[00292] A preferred arrangement is that the first Cas9 construct is arranged 5'-NLS-(N' terminal Cas9 part)-linker-(first half of the dimer)-NLS-3'. A preferred arrangement is that the second Cas9 construct is arranged 5'-NES--(second half of the dimer)-linker-(C' terminal Cas9 part)-NES-3'. A suitable promoter is preferably upstream of each of these constructs. The two constructs may be delivered separately or together.

[00293] In some embodiments, one or all of the NES(s) in operable linkage to the second Cas9 construct may be swapped out for an NLS. However, this may be typically not preferred and, in other embodiments, the localization signal in operable linkage to the second Cas9 construct is one or more NES(s).

[00294] It will also be appreciated that the NES may be operably linked to the N' terminal fragment of the split Cas9 and that the NLS may be operably linked to the C' terminal fragment of the split Cas9. However, the arrangement where the NLS is operably linked to the N' terminal

fragment of the split Cas9 and that the NES is operably linked to the C' terminal fragment of the split Cas9 may be preferred.

[00295] The NES functions to localize the second Cas9 fusion construct outside of the nucleus, at least until the inducer energy source is provided (e.g., at least until an energy source is provided to the inducer to perform its function). The presence of the inducer stimulates dimerization of the two Cas9 fusions within the cytoplasm and makes it thermodynamically worthwhile for the dimerized, first and second, Cas9 fusions to localize to the nucleus. Without being bound by theory, Applicants believe that the NES sequesters the second Cas9 fusion to the cytoplasm (i.e., outside of the nucleus). The NLS on the first Cas9 fusion localizes it to the nucleus. In both cases, Applicants use the NES or NLS to shift an equilibrium (the equilibrium of nuclear transport) to a desired direction. The dimerization typically occurs outside of the nucleus (a very small fraction might happen in the nucleus) and the NLSs on the dimerized complex shift the equilibrium of nuclear transport to nuclear localization, so the dimerized and hence reconstituted Cas9 enters the nucleus.

[00296] Beneficially, Applicants are able to reconstitute function in the split Cas9. Transient transfection is used to prove the concept and dimerization occurs in the background in the presence of the inducer energy source. No activity is seen with separate fragments of the Cas9. Stable expression through lentiviral delivery is then used to develop this and show that a split Cas9 approach can be used.

[00297] This present split Cas9 approach is beneficial as it allows the Cas9 activity to be inducible, thus allowing for temporal control. Furthermore, different localization sequences may be used (i.e., the NES and NLS as preferred) to reduce background activity from auto-assembled complexes. Tissue specific promoters, for example one for each of the first and second Cas9 fusion constructs, may also be used for tissue-specific targeting, thus providing spatial control. Two different tissue specific promoters may be used to exert a finer degree of control if required. The same approach may be used in respect of stage-specific promoters or there may a mixture of stage and tissue specific promoters, where one of the first and second Cas9 fusion constructs is under the control of (i.e. operably linked to or comprises) a tissue-specific promoter, whilst the other of the first and second Cas9 fusion constructs is under the control of (i.e. operably linked to or comprises) a stage-specific promoter.

[00298] The inducible Cas9 CRISPR-Cas system comprises one or more nuclear localization sequences (NLSs), as described herein, for example as operably linked to the first Cas9 fusion construct. These nuclear localization sequences are ideally of sufficient strength to drive accumulation of said first Cas9 fusion construct in a detectable amount in the nucleus of a eukaryotic cell. Without wishing to be bound by theory, it is believed that a nuclear localization sequence is not necessary for Cas9 CRISPR-Cas complex activity in eukaryotes, but that including such sequences enhances activity of the system, especially as to targeting nucleic acid molecules in the nucleus, and assists with the operation of the present 2-part system.

[00299] Equally, the second Cas9 fusion construct is operably linked to a nuclear export sequence (NES). Indeed, it may be linked to one or more nuclear export sequences. In other words, the number of export sequences used with the second Cas9 fusion construct is preferably 1 or 2 or 3. Typically 2 is preferred, but 1 is enough and so is preferred in some embodiments. Suitable examples of NLS and NES are known in the art. For example, a preferred nuclear export signal (NES) is human protein tyrosin kinase 2. Preferred signals will be species specific.

[00300] Where the FRB and FKBP system are used, the FKBP is preferably flanked by nuclear localization sequences (NLSs). Where the FRB and FKBP system are used, the preferred arrangement is N' terminal Cas9 – FRB – NES : C' terminal Cas9-FKBP-NLS. Thus, the first Cas9 fusion construct would comprise the C' terminal Cas9 part and the second Cas9 fusion construct would comprise the N' terminal Cas9 part.

[00301] Another beneficial aspect to the present invention is that it may be turned on quickly, i.e. that it has a rapid response. It is believed, without being bound by theory, that Cas9 activity can be induced through dimerization of existing (already present) fusion constructs (through contact with the inducer energy source) more rapidly than through the expression (especially translation) of new fusion constructs. As such, the first and second Cas9 fusion constructs may be expressed in the target cell ahead of time, i.e. before Cas9 activity is required. Cas9 activity can then be temporally controlled and then quickly constituted through addition of the inducer energy source, which ideally acts more quickly (to dimerize the heterodimer and thereby provide Cas9 activity) than through expression (including induction of transcription) of Cas9 delivered by a vector, for example.

[00302] The terms Cas9 or Cas9 enzyme and CRISPR enzyme are used interchangeably herein unless otherwise apparent.

[00303] Applicants demonstrate that Cas9 can be split into two components, which reconstitute a functional nuclease when brought back together. Employing rapamycin sensitive dimerization domains, Applicants generate a chemically inducible Cas9 for temporal control of Cas9-mediated genome editing and transcription modulation. Put another way, Applicants demonstrate that Cas9 can be rendered chemically inducible by being split into two fragments and that rapamycin-sensitive dimerization domains may be used for controlled reassembly of the Cas9. Applicants show that the re-assembled Cas9 may be used to mediate genome editing (through nuclease/nickase activity) as well as transcription modulation (as a DNA-binding domain, the so-called “dead Cas9”).

[00304] As such, the use of rapamycin-sensitive dimerization domains is preferred. Reassembly of the Cas9 is preferred. Reassembly can be determined by restoration of binding activity. Where the Cas9 is a nickase or induces a double-strand break, suitable comparison percentages compared to a wildtype are described herein.

[00305] Rapamycin treatments can last 12 days. The dose can be 200nM. This temporal and/or molar dosage is an example of an appropriate dose for Human embryonic kidney 293FT (HEK293FT) cell lines and this may also be used in other cell lines. This figure can be extrapolated out for therapeutic use *in vivo* into, for example, mg/kg. However, it is also envisaged that the standard dosage for administering rapamycin to a subject is used here as well. By the “standard dosage”, it is meant the dosage under rapamycin’s normal therapeutic use or primary indication (i.e. the dose used when rapamycin is administered for use to prevent organ rejection).

[00306] It is noteworthy that the preferred arrangement of Cas9-FRB/FKBP pieces are separate and inactive until rapamycin-induced dimerization of FRB and FKBP results in reassembly of a functional full-length Cas9 nuclease. Thus, it is preferred that first Cas9 fusion construct attached to a first half of an inducible heterodimer is delivered separately and/or is localized separately from the second Cas9 fusion construct attached to a first half of an inducible heterodimer.

[00307] To sequester the Cas9(N)-FRB fragment in the cytoplasm, where it is less likely to dimerize with the nuclear-localized Cas9(C)-FKBP fragment, it is preferable to use on Cas9(N)-FRB a single nuclear export sequence (NES) from the human protein tyrosin kinase 2 (Cas9(N)-FRB-NES). In the presence of rapamycin, Cas9(N)-FRB-NES dimerizes with Cas9(C)-FKBP-2xNLS to reconstitute a complete Cas9 protein, which shifts the balance of nuclear trafficking toward nuclear import and allows DNA targeting.

[00308] High dosage of Cas9 can exacerbate indel frequencies at off-target (OT) sequences which exhibit few mismatches to the guide strand. Such sequences are especially susceptible, if mismatches are non-consecutive and/or outside of the seed region of the guide. Accordingly, temporal control of Cas9 activity could be used to reduce dosage in long-term expression experiments and therefore result in reduced off-target indels compared to constitutively active Cas9.

[00309] Viral delivery is preferred. In particular, a lentiviral or AAV delivery vector is envisaged. Applicants generate a split-Cas9 lentivirus construct, similar to the lentiCRISPR plasmid. The split pieces should be small enough to fit the ~4.7kb size limitation of AAV.

[00310] Applicants demonstrate that stable, low copy expression of split Cas9 can be used to induce substantial indels at a targeted locus without significant mutation at off-target sites. Applicants clone Cas9 fragments (2 parts based on split 5, described herein).

[00311] A dead Cas9 may also be used, comprising a VP64 transactivation domain, for example added to Cas9(C)-FKBP-2xNLS (dead-Cas9(C)-FKBP-2xNLS-VP64). These fragments reconstitute a catalytically inactive Cas9-VP64 fusion (dead-Cas9-VP64). Transcriptional activation is induced by VP64 in the presence of rapamycin to induce the dimerization of the Cas9(C)-FKBP fusion and the Cas9(N)-FRB fusion. In other words, Applicants test the inducibility of split dead-Cas9-VP64 and show that transcriptional activation is induced by split dead-Cas9-VP64 in the presence of rapamycin. As such, the present inducible Cas9 may be associated with one or more functional domain, such as a transcriptional activator or repressor or a nuclease (such as FokI). A functional domain may be bound to or fused with one part of the split Cas9.

[00312] A preferred arrangement is that the first Cas9 construct is arranged 5'-First Localization Signal-(N' terminal Cas9 part)-linker-(first half of the dimer)-First Localization

Signal-3' and the second Cas9 construct is arranged 5'- Second Localization Signal--(second half of the dimer)-linker-(C' terminal Cas9 part)-Second Localization Signal-Functional Domain-3'. Here, a functional domain is placed at the 3' end of the second Cas9 construct. Alternatively, a functional domain may be placed at the 5' end of the first Cas9 construct. One or more functional domains may be used at the 3' end or the 5' end or at both ends. A suitable promoter is preferably upstream of each of these constructs. The two constructs may be delivered separately or together. The Localization Signals may be an NLS or an NES, so long as they are not inter-mixed on each construct.

[00313] In an aspect the invention provides an inducible Cas9 CRISPR-Cas system wherein the Cas9 has a diminished nuclease activity of at least 97%, or 100% as compared with the Cas9 enzyme not having the at least one mutation.

[00314] Accordingly, it is also preferred that the Cas9 is a dead-Cas9. Ideally, the split should always be so that the catalytic domain(s) are unaffected. For the dead-Cas9 the intention is that DNA binding occurs, but not cleavage or nickase activity is shown.

[00315] In an aspect the invention provides an inducible Cas9 CRISPR-Cas system as herein discussed wherein one or more functional domains is associated with the Cas9. This functional domain may be associated with (i.e. bound to or fused with) one part of the split Cas9 or both. There may be one associated with each of the two parts of the split Cas9. These may therefore be typically provided as part of the first and/or second Cas9 fusion constructs, as fusions within that construct. The functional domains are typically fused via a linker, such as GlySer linker, as discussed herein. The one or more functional domains may be transcriptional activation domain or a repressor domain. Although they may be different domains it is preferred that all the functional domains are either activator or repressor and that a mixture of the two is not used.

[00316] The transcriptional activation domain may comprise VP64, p65, MyoD1, HSF1, RTA or SET7/9.

[00317] In an aspect, the invention provides an inducible Cas9 CRISPR-Cas system as herein discussed wherein the one or more functional domains associated with the Cas9 is a transcriptional repressor domain.

[00318] In an aspect, the invention provides an inducible Cas9 CRISPR-Cas system as herein discussed wherein the transcriptional repressor domain is a KRAB domain.

[00319] In an aspect, the invention provides an inducible Cas9 CRISPR-Cas system as herein discussed wherein the transcriptional repressor domain is a NuE domain, NcoR domain, SID domain or a SID4X domain.

[00320] In an aspect the invention provides an inducible Cas9 CRISPR-Cas system as herein discussed wherein the one or more functional domains associated with the adaptor protein have one or more activities comprising methylase activity, demethylase activity, transcription activation activity, transcription repression activity, transcription release factor activity, histone modification activity, RNA cleavage activity, DNA cleavage activity, DNA integration activity or nucleic acid binding activity.

[00321] Histone modifying domains are also preferred in some embodiments. Exemplary histone modifying domains are discussed below. Transposase domains, HR (Homologous Recombination) machinery domains, recombinase domains, and/or integrase domains are also preferred as the present functional domains. In some embodiments, DNA integration activity includes HR machinery domains, integrase domains, recombinase domains and/or transposase domains.

[00322] In an aspect the invention provides an inducible Cas9 CRISPR-Cas system as herein discussed wherein the DNA cleavage activity is due to a nuclease.

[00323] In an aspect the invention provides an inducible Cas9 CRISPR-Cas system as herein discussed wherein the nuclease comprises a FokI nuclease.

[00324] The use of such functional domains, which are preferred with the present split Cas9 system, is also discussed in detail in Konermann et al. ("Genome-scale transcriptional activation with an engineered CRISPR-Cas9 complex" Nature published 11 Dec 2014).

[00325] The present system may be used with any guide.

[00326] Modified guides may be used in certain embodiments. Particularly preferred are guides embodying the teachings of Konermann Nature 11 Dec 2014 paper mentioned above. These guides are modified so that protein-binding RNA portions (such as aptamers) are added. Such portion(s) may replace a portion of the guide. Corresponding RNA-binding protein domains can be used to then recognise the RNA and recruit functional domains, such as those described herein, to the guide. This is primarily for use with dead-Cas9 leading to transcriptional activation or repression or DNA cleavage through nucleases such as FokI. The use of such

guides in combination with dead-Cas9 is powerful, and it is especially powerful if the Cas9 itself is also associated with its own functional domain, as discussed herein. When a dead-Cas9 (with or without its own associated functional domain) is induced to reconstitute in accordance with the present invention, i.e. is a split Cas9, then the tool is especially useful.

[00327] A guide RNA (gRNA), also preferred for use in the present invention, can comprise a guide sequence capable of hybridizing to a target sequence in a genomic locus of interest in a cell, wherein the gRNA is modified by the insertion of distinct RNA sequence(s) that bind to one or more adaptor proteins, and wherein the adaptor protein is associated with one or more functional domains. The Cas9 may comprise at least one mutation, such that the Cas9 enzyme has no more than 5% of the nuclease activity of the Cas9 enzyme not having the at least one mutation; and/or at least one or more nuclear localization sequences. Also provided is a non-naturally occurring or engineered composition comprising: one or more guide RNA (gRNA) comprising a guide sequence capable of hybridizing to a target sequence in a genomic locus of interest in a cell, a Cas9 enzyme comprising at least one or more nuclear localization sequences, wherein the Cas9 enzyme comprises at least one mutation, such that the Cas9 enzyme has no more than 5% of the nuclease activity of the Cas9 enzyme not having the at least one mutation, wherein the at least one gRNA is modified by the insertion of distinct RNA sequence(s) that bind to one or more adaptor proteins, and wherein the adaptor protein is associated with one or more functional domains.

[00328] The gRNA that is preferably modified by the insertion of distinct RNA sequence(s) that bind to one or more adaptor proteins. The insertion of distinct RNA sequence(s) that bind to one or more adaptor proteins is preferably an aptamer sequence or two or more aptamer sequences specific to the same or different adaptor protein(s). The adaptor protein preferably comprises MS2, PP7, Q β , F2, GA, fr, JP501, M12, R17, BZ13, JP34, JP500, KU1, M11, MX1, TW18, VK, SP, FI, ID2, NL95, TW19, AP205, ϕ Cb5, ϕ Cb8r, ϕ Cb12r, ϕ Cb23r, 7s, PRR1. Cell lines stably expressing inter alia split dead-Cas9 can be useful.

[00329] Applicants demonstrate that Cas9 can be split into two distinct fragments, which reconstitute a functional full-length Cas9 nuclease when brought back together using chemical induction. The split Cas9 architecture will be useful for a variety of applications. For example, split Cas9 may enable genetic strategies for restricting Cas9 activity to intersectional cell

populations by putting each fragment under a different tissue specific promoter. Additionally, different chemically inducible dimerization domains such as APA and gibberellin may also be employed.

[00330] The inducer energy source is preferably chemical induction.

[00331] The split position or location is the point at which the first part of the Cas9 enzyme is separated from the second part. In some embodiments, the first part will comprise or encode amino acids 1 to X, whilst the second part will comprise or encode amino acids X+1 to the end. In this example, the numbering is contiguous, but this may not always be necessary as amino acids (or the nucleotides encoding them) could be trimmed from the end of either of the split ends, provided that sufficient DNA binding activity and, if required, DNA nickase or cleavage activity is retained, for example at least 40%, 50%, 60%, 70%, 80%, 90% or 95% activity compared to wildtype Cas9.

[00332] The exemplary numbering provided herein may be in reference to the wildtype protein, preferably the wildtype FnCas9. However, it is envisaged that mutants of the wildtype Cas9 such as of FnCas9 protein can be used. The numbering may also not follow exactly the FnCas9 numbering as, for instance, some N' or C' terminal truncations or deletions may be used, but this can be addressed using standard sequence alignment tools. Orthologs are also preferred as a sequence alignment tool.

[00333] Thus, the split position may be selected using ordinary skill in the art, for instance based on crystal data and/or computational structure predictions.

[00334] For example, computational analysis of the primary structure of Cas9 nucleases reveals three distinct regions (Fig. 1). First a C-terminal RuvC like domain, which is the only functional characterized domain. Second a N-terminal alpha-helical region and third a mixed alpha and beta region, located between the RuvC like domain and the alpha-helical region. Several small stretches of unstructured regions are predicted within the Cas9 primary structure. Unstructured regions, which are exposed to the solvent and not conserved within different Cas9 orthologs, may represent preferred sites for splits (Fig. 2 and Fig. 3).

[00335] The following table presents non-limiting potential split regions within As and LbCas9. A split site within such a region may be opportune.

Split region	AsCas9	LbCas9
1	575-588	566-571
2	631-645	754-757
3	653-664	-
4	818-844	-

[00336] For Fn, As and Lb Cas9 mutants, it should be readily apparent what the corresponding position for a potential split site is, for example, based on a sequence alignment. For non-Fn, As and Lb enzymes one can use the crystal structure of an ortholog if a relatively high degree of homology exists between the ortholog and the intended Cas9, or one can use computational prediction.

[00337] Ideally, the split position should be located within a region or loop. Preferably, the split position occurs where an interruption of the amino acid sequence does not result in the partial or full destruction of a structural feature (e.g. alpha-helices or beta-sheets). Unstructured regions (regions that do not show up in the crystal structure because these regions are not structured enough to be “frozen” in a crystal) are often preferred options. Applicants can for example make splits in unstructured regions that are exposed on the surface of Cas9.

[00338] In one aspect the invention involves a method for generating a non-human eukaryote model for tumor formation and/or tumor evolution by introducing into a non-human animal a plurality of eukaryotic Cas transgenic cells, each cell comprising one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote. The Applicants have found that when the plurality of cells originates from an immortal (cancer) cell line which does not have the capacity to metastasize upon injection into a non-human eukaryote (or at least not until the primary tumor is too large), upon introduction into such cells of one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote, will cause the cells to form macroscopic metastases in the non-human eukaryotic animal. The applicants have moreover found that the model can be used to analyse human tumors in a non-human animal. In principle virtually any cell lines can be used as xenografts, including human cell lines.

[00339] In particular embodiments, the methods comprise the steps of: (a) providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas; (b) introducing in said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; or introducing in said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; (c) introducing a plurality of said Cas transgenic eukaryotic cell after step (b) in a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus. In particular embodiments, the one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote is a CRISPR library with multiple sgRNAs targeting each gene that is likely to generate loss of function mutation. In further particular embodiments, a customized library is used, whereby genes are selected based on genomic analysis, pathways, clinical relevance, literature or customized informatics for focused libraries for screens. Indeed, this allows the possibility to reduce the size of the library. In particular embodiments, a targeted sub-pool is used to confirm a primary screen. This can optionally including libraries comprising multiple RNAs guiding to the same target genes.

[00340] It will be understood that even where only one CAS guiding RNA is used (i.e. to one target DNA sequence), introduction thereof into a plurality of eukaryotic cells may nevertheless result in a variety of phenotypes resulting from the different possibilities that the repair mechanism will address the DNA cleavage. However, in particular embodiments, more than one CAS guiding RNA is used.

[00341] Accordingly, also provided herein is a concept of multiple levels of screening which allows further validation of the identified targets. More particularly, a first level may include the introduction of a genome-wide library targeting each gene that is likely to generate loss of function mutation. In a further level the library may aim to target a sub-pool of genes which have been identified in a first level screen to affect tumor formation and evolution. The identification of the sub-pool of genes which most strongly affect tumor formation and evolution may be

performed based either on relative abundance of the sgRNAs, i.e. by taking sgRNAs above certain threshold, based on rank, i.e. by taking top ranked sgRNAs in each mouse, or based on false discovery rate (FDR), i.e. by taking sgRNAs enriched compared to the distribution of the 1,000 non-targeting sgRNAs at 0.2% FDR. In particular embodiments, the genes identified in two out of three or in all three methods are selected as a sub-pool. In particular embodiments, the targeting of the subpool may comprise using multiple RNAs, such as 10 RNAs per gene. Typically, the methods will also comprise a validation the individual genes. In particular embodiments this is ensured by cloning multiple sgRNAs targeting one gene into the same lentiviral vector, transducing them into a cell line, introducing them into a non-human eukaryote and considering the tumor formation and evolution of said cell line. Additionally or alternatively, the methods may also take into account the different bottlenecks that can affect CAS guiding RNA representation in the library. In one level, the method may involve introducing individual RNAs into cells and cultivating cells prior to mixing them and introducing them into the non-human eukaryote animal. For instance the method can comprise maintaining separate populations of cells comprising different RNAs for guiding Cas, whereby these are combined only upon injection into the non-human animal. Representation of each of the RNAs within the RNA population can be investigated at the plasmid, cell pool, primary tumor and metastasis level. This targeted subpool strategy provides a higher-throughput method for validating the top candidates and assaying competing mutants in parallel. This strategy is of interest for any type of genome-scale screens and more particularly loss-of-function screens such as the one described herein or RNAi-based screens.

[00342] In an aspect, the invention involves a non-human eukaryote, such as an animal, preferably a mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, or arthropod, etc or cell thereof or tissue thereof that may be used as a disease model for cancer, such as a tumor isolated therefrom, in particular a primary tumor or a metastasis, or a cell isolated from such tumor or metastasis, or a cell line derived therefrom.

[00343] For example, a method of the invention may be used to create a non-human eukaryote, e.g., an animal, mammal, primate, rodent or cell that comprises a cell or a plurality of

cells having one or more modification, e.g., 1-50 modifications, such as in one or more nucleic acid sequences associated or correlated with cancer. It is to be understood that such modifications result from the methods as described herein. Such a mutated nucleic acid sequence be associated or correlated with cancer (including metastasis) and may encode a cancer or metastasis associated protein sequence or may be a cancer or metastasis associated or correlated control sequence. The cell may be in vivo or ex vivo in the cases of multicellular organisms. In the instance where the cell is in culture (i.e. in vitro), a cell line may be established if appropriate culturing conditions are met and preferably if the cell is suitably adapted for this purpose (for instance a stem cell). Hence, cell lines are also envisaged. In some methods, the disease model can be used to study the effects of mutations on the animal or cell and development and/or progression of cancer and/or metastasis using measures commonly used in the study of cancer. Alternatively, such a cancer model is useful for studying the effect of a putatively pharmaceutically active compound or gene therapy on the cancer or metastasis. A cancer- or metastasis-associated gene or polynucleotide can be modified to give rise to cancer or metastasis in the model, and then putatively pharmaceutically active compound and/or gene therapy can be administered so as to observe whether cancer or metastasis development and/or progression is inhibited or reduced. In particular, the method comprises modifying so as to produce, one or more, such as a plurality cancer or metastasis-associated or correlated gene(s) or polynucleotide(s). the plurality of cancer or metastasis associated or correlated gene(s) or polynucleotide(s) may in certain embodiments also be targeted by introduction in a plurality of cells of a library of RNAs capable of guiding Cas to a genomic target locus, preferably in such way that in each cell a different genomic locus is targeted (or different regions of a genomic locus). Accordingly, in some methods, a genetically modified animal may be compared with an animal predisposed to development of the disease, such that administering putative gene therapy, or pharmaceutically acceptable compound(s), or any combination thereof can be performed to assess how such putative therapy(ies) or treatment(s) may perform in a human. The invention can involve administering to a eukaryote obtained by a method as described herein a test compound or contacting a test compound with a tumor or metastasis tissue, cell or cell line derived therefrom, wherein the tissue, or cell comprises - or wherein the eukaryote comprises a cancerous or metastatic tissue or cell that comprises - one or more, e.g., 1-50 or more mutations

from the CRISPR-Cas system, e.g., in an animal comprising cells expressing Cas and to which RNA(s) generating the mutations has / have been administered; and detecting a reduction or an augmentation of a cell signaling event associated with the mutation(s) or lack thereof. Screening of such putative pharmaceutically active compound(s) and/or gene therapy(ies) can be by cellular function change and/or intracellular signaling or extracellular signaling change. Such screening can involve evaluating for dosages or dose curves, as well as combinations of potential drugs and/or therapies. An altered expression of one or more genome sequences associated with a signaling biochemical pathway can be determined by assaying for a difference in the mRNA levels of the corresponding genes between the disease model eukaryote or animal or cell or tissue thereof and a normal eukaryote, animal, tissue or cell, and to ascertain whether when the disease model is administered or contacted with a candidate chemical agent or gene therapy it reverts to or towards normal. An assay can be for mutation(s)-induced alteration in the level of mRNA transcripts or corresponding polynucleotides in comparison with such level(s) in a normal eukaryote or animal and whether such level(s) are placed towards or to normal when a therapy or treatment or agent is employed.

[00344] In an aspect the invention involves cells, e.g., non-human eukaryotic, e.g., animal, such as mammal, e.g., primate, rodent, mouse, rat, rabbit, etc. as described herein elsewhere, or even human cells, containing Cas polypeptide or transformed to constitutively express or alternatively inducibly and/or conditionally express Cas, e.g., such cells as to which a vector that contains nucleic acid molecule(s) encoding a Cas, e.g., with nucleic acid(s) encoding a promoter and preferably at least one NLS, advantageously two or more NLSs, or such cells that have had their genome altered, e.g., through the vector being an integrating virus or through such cells being stem cells or cells that give rise to a cell line or a living organism (but wherein such an organism is advantageously non-human), that contains and expresses nucleic acid molecule(s) encoding Cas. To these cells is then administered (simultaneously or subsequently) RNA(s) or vector(s), e.g., AAV, adenovirus, lentivirus containing or providing RNA(s) that guide Cas to a genomic target locus, e.g., under the control of a promoter such as a U6 promoter and/or particle(s) and/or nanoparticle(s) containing the RNA(s) and/or vector(s), whereby the RNA(s) direct the Cas in the cells to provide a mutation, or a plurality of mutation(s) such as from 3 to 50 mutations, advantageously mutation(s) associated or correlated with cancer. Such cells are then

transplanted into or onto a eukaryote, such as an animal suitable for being a cancer model, e.g., a rodent such as a mouse (see, e.g., literature on mouse transplantation cancer models, generally discussed at the NIH website; see <http://emice.nci.nih.gov/aam/mouse/transplantation-mouse-models-1>), chickens or chicken embryo or chicken embryo membrane (Kuzminien et al, "Evaluation of the Chicken Embryochorioallantoic membrane Model for Laryngeal Tumor Transplantation," *Papers on Anthropology* XX, 2011, pp. 229–240), zebra fish (see, e.g., Haldi et al, "Human melanoma cells transplanted into zebrafish proliferate, migrate, produce melanin, form masses and stimulate angiogenesis in zebrafish," *Angiogenesis*. 2006;9(3):139-51. Epub 2006 Oct 19)). The cells proliferate on or in the non-human eukaryote, e.g., animal model and tumor formation and/or evolution can be studied.

[00345] The non-human eukaryote, e.g., animal model can then be used for testing, e.g., as to potential therapy and/or putative treatment via a possibly pharmaceutically active compound. The administering of such compound can be at or to or for body delivery to the proliferated heterologous transplanted cells, e.g., direct injection at or near such proliferated heterologous transplanted cells, or injection or other administration in such a way that the compound is delivered into the heterologous transplanted cells, e.g., injection into the bloodstream whereby bodily functions transport to the proliferated heterologous transplanted cells. In an aspect of the invention, barcoding techniques of WO/2013/138585 A1 can be adapted or integrated into the practice of the invention.

[00346] The invention also comprehends a method for the identification of a treatment, e.g., chemical or gene therapy treatment, for cancer comprising applying, administering or delivering one or more treatments to the non-human transgenic eukaryote obtained by the methods as herein discussed and identifying whether the cancer has improved; and, the cancer can be Lung cancer, Lung adenocarcinoma, Lung squamous cell carcinoma, Acute myeloid leukemia, Basal cell carcinoma (skin), Bladder cancer, Breast cancer, Carcinoid, Chronic lymphocytic leukemia, Colorectal cancer, Diffuse large B-cell lymphoma, Endometrial, Esophageal cancer, Esophageal adenocarcinoma, Glioblastoma multiforme, Glioma, Head and neck cancer, Kidney clear cell cancer, Medulloblastoma, Melanoma, Multiple myeloma, Nasopharyngeal, Neuroblastoma, Ovarian cancer, Prostate cancer, Rhabdoid tumor, Testicular germ cell tumor, Thyroid cancer, or

Urinary bladder cancer. The method can comprise applying, administering or delivering different doses of the treatment and/or employing different routes of administration and/or different carriers or excipients and/or applying, administering or delivering at different time intervals, e.g., applying, administering or delivering different doses at different time intervals.

[00347] The invention thus also envisions an individualized or personalized treatment of cancer, being it a primary tumor or metastasis, in a subject in need of such treatment comprising: (a) performing any of the methods as described herein; (b) testing treatment(s) for the developed or developing cancer and/or metastasis; and (c) treating the subject based on results from the testing of treatment(s) of step (b).

[00348] In an aspect, the invention provides kits containing any one or more of the elements discussed herein. Elements may be provided individually or in combinations, and may be provided in any suitable container, such as a vial, a bottle, or a tube. In some embodiments, the kit includes instructions in one or more languages, for example in more than one language. In some embodiments, a kit comprises one or more reagents for use in a process utilizing one or more of the elements described herein. Reagents may be provided in any suitable container. For example, a kit may provide one or more reaction or storage buffers. Reagents may be provided in a form that is usable in a particular assay, or in a form that requires addition of one or more other components before use (e.g. in concentrate or lyophilized form). A buffer can be any buffer, including but not limited to a sodium carbonate buffer, a sodium bicarbonate buffer, a borate buffer, a Tris buffer, a MOPS buffer, a HEPES buffer, and combinations thereof. In some embodiments, the buffer is alkaline. In some embodiments, the buffer has a pH from about 7 to about 10. In some embodiments, the kit comprises Cas (or a vector encoding Cas, or a Cas transgenic eukaryotic cell) and/or one or more, such as a library, oligonucleotides corresponding to a guide sequence for insertion into a vector (or vectors having already inserted such oligonucleotides) so as to operably link the guide sequence and a regulatory element. In some embodiments, the kit comprises a homologous recombination template polynucleotide. In some embodiments, the kit comprises one or more of the vectors and/or one or more of the polynucleotides described herein. The kit may advantageously allow the provision of all elements of the systems of the invention. Kits can involve vector(s) and/or particle(s) and/or

nanoparticle(s) containing or encoding RNA(s) to be administered to a eukaryotic cell, e.g., animal, mammal, primate, rodent, etc., with such a kit including instructions for administering to such a cell; and such a kit can optionally include a non-human eukaryote.

[00349] The invention uses nucleic acids to bind target DNA sequences. This is advantageous as nucleic acids are much easier and cheaper to produce than proteins, and the specificity can be varied according to the length of the stretch where homology is sought. Complex 3-D positioning of multiple fingers, for example is not required. 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. The term also encompasses nucleic-acid-like structures with synthetic backbones, see, e.g., Eckstein, 1991; Baserga et al., 1992; Milligan, 1993; WO 97/03211; WO 96/39154; Mata, 1997; Strauss-Soukup, 1997; and Samstag, 1996. 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.

[00350] It will be appreciated that where reference is made to a polynucleotide, where that polynucleotide is RNA and is said to ‘comprise’ a feature such as a target sequence, the RNA sequence includes the feature. Where the polynucleotide is DNA and is said to comprise a feature such as a target sequence, the DNA sequence is or can be transcribed into the RNA that comprises the feature at issue. Where the feature is a protein, such as the CRISPR enzyme, the DNA or RNA sequence referred to is, or can be, translated (and in the case of DNA

transcribed first). Furthermore, in cases where an RNA encoding the CRISPR enzyme is provided to a cell, it is understood that the RNA is capable of being translated by the cell into which it is delivered.

[00351] As used herein the term “wild type” is a term of the art understood by skilled persons and means the typical form of an organism, strain, gene or characteristic as it occurs in nature as distinguished from mutant or variant forms. A “wild type” can be a base line. As used herein the term “variant” should be taken to mean the exhibition of qualities that have a pattern that deviates from what occurs in nature. The terms “non-naturally occurring” or “engineered” are used interchangeably and indicate the involvement of the hand of man. The terms, when referring to nucleic acid molecules or polypeptides mean that the nucleic acid molecule or the polypeptide is at least substantially free from at least one other component with which they are naturally associated in nature and as found in nature. “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). “Perfectly complementary” means that all the contiguous residues of a nucleic acid sequence will hydrogen bond with the same number of contiguous residues in a second nucleic acid sequence. “Substantially complementary” as used herein refers to a degree of complementarity that is at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100% over a region of 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, or more nucleotides, or refers to two nucleic acids that hybridize under stringent conditions. As used herein, “stringent conditions” for hybridization refer to conditions under which a nucleic acid having complementarity to a target sequence predominantly hybridizes with the target sequence, and substantially does not hybridize to non-target sequences. Stringent conditions are generally sequence-dependent, and vary depending on a number of factors. In general, the longer the sequence, the higher the temperature at which the sequence specifically hybridizes to its target sequence. Non-limiting examples of stringent conditions are described in detail in Tijssen (1993), *Laboratory Techniques In Biochemistry And Molecular Biology-Hybridization With*

Nucleic Acid Probes Part I, Second Chapter "Overview of principles of hybridization and the strategy of nucleic acid probe assay", Elsevier, N.Y. Where reference is made to a polynucleotide sequence, then complementary or partially complementary sequences are also envisaged. These are preferably capable of hybridising to the reference sequence under highly stringent conditions. Generally, in order to maximize the hybridization rate, relatively low-stringency hybridization conditions are selected: about 20 to 25° C lower than the thermal melting point (T_m). The T_m is the temperature at which 50% of specific target sequence hybridizes to a perfectly complementary probe in solution at a defined ionic strength and pH. Generally, in order to require at least about 85% nucleotide complementarity of hybridized sequences, highly stringent washing conditions are selected to be about 5 to 15° C lower than the T_m . In order to require at least about 70% nucleotide complementarity of hybridized sequences, moderately-stringent washing conditions are selected to be about 15 to 30° C lower than the T_m . Highly permissive (very low stringency) washing conditions may be as low as 50° C below the T_m , allowing a high level of mis-matching between hybridized sequences. Those skilled in the art will recognize that other physical and chemical parameters in the hybridization and wash stages can also be altered to affect the outcome of a detectable hybridization signal from a specific level of homology between target and probe sequences. Preferred highly stringent conditions comprise incubation in 50% formamide, 5×SSC, and 1% SDS at 42° C, or incubation in 5×SSC and 1% SDS at 65° C, with wash in 0.2×SSC and 0.1% SDS at 65° C. "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, Hoogsteen 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. As used herein, the term "genomic locus" or "locus" (plural loci) is the specific location of a gene or DNA sequence on a chromosome. A "gene" refers to stretches of DNA or RNA that encode a polypeptide or an RNA chain that has functional role to play in an organism

and hence is the molecular unit of heredity in living organisms. For the purpose of this invention it may be considered that genes include regions which regulate the production of the gene product, whether or not such regulatory sequences are adjacent to coding and/or transcribed sequences. Accordingly, a gene includes, but is not necessarily limited to, promoter sequences, terminators, translational regulatory sequences such as ribosome binding sites and internal ribosome entry sites, enhancers, silencers, insulators, boundary elements, replication origins, matrix attachment sites and locus control regions. As used herein, "expression of a genomic locus" or "gene expression" is the process by which information from a gene is used in the synthesis of a functional gene product. The products of gene expression are often proteins, but in non-protein coding genes such as rRNA genes or tRNA genes, the product is functional RNA. The process of gene expression is used by all known life - eukaryotes (including multicellular organisms), prokaryotes (bacteria and archaea) and viruses to generate functional products to survive. As used herein "expression" of a gene or nucleic acid encompasses not only cellular gene expression, but also the transcription and translation of nucleic acid(s) in cloning systems and in any other context. As used herein, "expression" also refers to the process by which a polynucleotide is transcribed from a DNA template (such as into mRNA or other RNA transcript) and/or the process by which a transcribed mRNA is subsequently translated into peptides, polypeptides, or proteins. Transcripts and encoded polypeptides may be collectively referred to as "gene product." If the polynucleotide is derived from genomic DNA, expression may include splicing of the mRNA in a eukaryotic cell. The terms "polypeptide", "peptide" and "protein" are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non amino acids. The terms also encompass an amino acid polymer that has been modified; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation, such as conjugation with a labeling component. As used herein the term "amino acid" includes natural and/or unnatural or synthetic amino acids, including glycine and both the D or L optical isomers, and amino acid analogs and peptidomimetics. As used herein, the term "domain" or "protein domain" refers to a part of a protein sequence that may exist and function independently of the rest of the protein chain. As described in aspects of the invention, sequence identity is related to sequence homology.

Homology comparisons may be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. These commercially available computer programs may calculate percent (%) homology between two or more sequences and may also calculate the sequence identity shared by two or more amino acid or nucleic acid sequences. In some preferred embodiments, the capping region of the dTALEs described herein have sequences that are at least 95% identical or share identity to the capping region amino acid sequences provided herein. Sequence homologies may be generated by any of a number of computer programs known in the art, for example BLAST or FASTA, etc. A suitable computer program for carrying out such an alignment is the GCG Wisconsin Bestfit package (University of Wisconsin, U.S.A; Devereux et al., 1984, *Nucleic Acids Research* 12:387). Examples of other software that may perform sequence comparisons include, but are not limited to, the BLAST package (see Ausubel et al., 1999 *ibid* – Chapter 18), FASTA (Atschul et al., 1990, *J. Mol. Biol.*, 403-410) and the GENWORKS suite of comparison tools. Both BLAST and FASTA are available for offline and online searching (see Ausubel et al., 1999 *ibid*, pages 7-58 to 7-60). However it is preferred to use the GCG Bestfit program. Percentage (%) sequence homology may be calculated over contiguous sequences, i.e., one sequence is aligned with the other sequence and each amino acid or nucleotide in one sequence is directly compared with the corresponding amino acid or nucleotide in the other sequence, one residue at a time. This is called an “ungapped” alignment. Typically, such ungapped alignments are performed only over a relatively short number of residues. Although this is a very simple and consistent method, it fails to take into consideration that, for example, in an otherwise identical pair of sequences, one insertion or deletion may cause the following amino acid residues to be put out of alignment, thus potentially resulting in a large reduction in % homology when a global alignment is performed. Consequently, most sequence comparison methods are designed to produce optimal alignments that take into consideration possible insertions and deletions without unduly penalizing the overall homology or identity score. This is achieved by inserting “gaps” in the sequence alignment to try to maximize local homology or identity. However, these more complex methods assign “gap penalties” to each gap that occurs in the alignment so that, for the same number of identical amino acids, a sequence alignment with as few gaps as possible - reflecting higher relatedness between the two compared sequences - may achieve a higher score than one with many gaps. “Affinity gap costs” are

typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties may, of course, produce optimized alignments with fewer gaps. Most alignment programs allow the gap penalties to be modified. However, it is preferred to use the default values when using such software for sequence comparisons. For example, when using the GCG Wisconsin Bestfit package the default gap penalty for amino acid sequences is -12 for a gap and -4 for each extension. Calculation of maximum % homology therefore first requires the production of an optimal alignment, taking into consideration gap penalties. A suitable computer program for carrying out such an alignment is the GCG Wisconsin Bestfit package (Devereux et al., 1984 *Nuc. Acids Research* 12 p387). Examples of other software that may perform sequence comparisons include, but are not limited to, the BLAST package (see Ausubel et al., 1999 *Short Protocols in Molecular Biology*, 4th Ed. – Chapter 18), FASTA (Altschul et al., 1990 *J. Mol. Biol.* 403-410) and the GENEWORKS suite of comparison tools. Both BLAST and FASTA are available for offline and online searching (see Ausubel et al., 1999, *Short Protocols in Molecular Biology*, pages 7-58 to 7-60). However, for some applications, it is preferred to use the GCG Bestfit program. A new tool, called BLAST 2 Sequences is also available for comparing protein and nucleotide sequences (see *FEMS Microbiol Lett.* 1999 174(2): 247-50; *FEMS Microbiol Lett.* 1999 177(1): 187-8 and the website of the National Center for Biotechnology information at the website of the National Institutes for Health). Although the final % homology may be measured in terms of identity, the alignment process itself is typically not based on an all-or-nothing pair comparison. Instead, a scaled similarity score matrix is generally used that assigns scores to each pair-wise comparison based on chemical similarity or evolutionary distance. An example of such a matrix commonly used is the BLOSUM62 matrix - the default matrix for the BLAST suite of programs. GCG Wisconsin programs generally use either the public default values or a custom symbol comparison table, if supplied (see user manual for further details). For some applications, it is preferred to use the public default values for the GCG package, or in the case of other software, the default matrix, such as BLOSUM62. Alternatively, percentage homologies may be calculated using the multiple alignment feature in DNASISTM (Hitachi Software), based on an algorithm, analogous to CLUSTAL (Higgins DG & Sharp PM (1988), *Gene* 73(1); 237-244). Once the software has produced an optimal alignment, it is possible to

calculate % homology, preferably % sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result. The sequences may also have deletions, insertions or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent substance. Deliberate amino acid substitutions may be made on the basis of similarity in amino acid properties (such as polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues) and it is therefore useful to group amino acids together in functional groups. Amino acids may be grouped together based on the properties of their side chains alone. However, it is more useful to include mutation data as well. The sets of amino acids thus derived are likely to be conserved for structural reasons. These sets may be described in the form of a Venn diagram (Livingstone C.D. and Barton G.J. (1993) "Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation" *Comput. Appl. Biosci.* 9: 745-756) (Taylor W.R. (1986) "The classification of amino acid conservation" *J. Theor. Biol.* 119; 205-218). Conservative substitutions may be made, for example according to the table below which describes a generally accepted Venn diagram grouping of amino acids.

Set		Sub-set	
Hydrophobic	F W Y H K M I L V A G C	Aromatic	F W Y H
		Aliphatic	I L V
Polar	W Y H K R E D C S T N Q	Charged	H K R E D
		Positively charged	H K R
		Negatively charged	E D
Small	V C A G S P T N D	Tiny	A G S

[00352] Embodiments of the invention include sequences (both polynucleotide or polypeptide) which may comprise homologous substitution (substitution and replacement are both used herein to mean the interchange of an existing amino acid residue or nucleotide, with an

alternative residue or nucleotide) that may occur i.e., like-for-like substitution in the case of amino acids such as basic for basic, acidic for acidic, polar for polar, etc. Non-homologous substitution may also occur i.e., from one class of residue to another or alternatively involving the inclusion of unnatural amino acids such as ornithine (hereinafter referred to as Z), diaminobutyric acid ornithine (hereinafter referred to as B), norleucine ornithine (hereinafter referred to as O), pyriylalanine, thienylalanine, naphthylalanine and phenylglycine. Variant amino acid sequences may include suitable spacer groups that may be inserted between any two amino acid residues of the sequence including alkyl groups such as methyl, ethyl or propyl groups in addition to amino acid spacers such as glycine or β -alanine residues. A further form of variation, which involves the presence of one or more amino acid residues in peptoid form, may be well understood by those skilled in the art. For the avoidance of doubt, “the peptoid form” is used to refer to variant amino acid residues wherein the α -carbon substituent group is on the residue’s nitrogen atom rather than the α -carbon. Processes for preparing peptides in the peptoid form are known in the art, for example Simon RJ et al., PNAS (1992) 89(20), 9367-9371 and Horwell DC, Trends Biotechnol. (1995) 13(4), 132-134.

[00353] For purpose of this invention, amplification means any method employing a primer and a polymerase capable of replicating a target sequence with reasonable fidelity. Amplification may be carried out by natural or recombinant DNA polymerases such as TaqGold™, T7 DNA polymerase, Klenow fragment of E.coli DNA polymerase, and reverse transcriptase. A preferred amplification method is PCR.

Crystallization of CRISPR-Cas9 and Characterization of Crystal Structure

[00354] The crystals of the Cas9 can be obtained by techniques of protein crystallography, including batch, liquid bridge, dialysis, vapor diffusion and hanging drop methods. Generally, the crystals of the invention are grown by dissolving substantially pure CRISPR-Cas9 and a nucleic acid molecule to which it binds in an aqueous buffer containing a precipitant at a concentration just below that necessary to precipitate. Water is removed by controlled evaporation to produce precipitating conditions, which are maintained until crystal growth ceases. The crystal structure information is described in U.S. provisional applications 61/915,251 filed December 12, 2013, 61/930,214 filed on January 22, 2014, 61/980,012 filed April 15, 2014;

and Nishimasu et al, "Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA," Cell 156(5):935–949, DOI: <http://dx.doi.org/10.1016/j.cell.2014.02.001> (2014), and PCT/US14/70175; each and all of which are incorporated herein by reference, and together are "herein cited materials" concerning the Cas9 crystal structure. Uses of the Crystals, Crystal Structure and Atomic Structure Co-Ordinates of the herein cited materials: The crystals of the Cas9, and particularly the atomic structure co-ordinates obtained therefrom, have a wide variety of uses. The crystals and structure co-ordinates are particularly useful for identifying compounds (nucleic acid molecules) that bind to CRISPR-Cas9, and CRISPR-Cas9s that can bind to particular compounds (nucleic acid molecules). Thus, the structure co-ordinates described herein can be used as phasing models in determining the crystal structures of additional synthetic or mutated CRISPR-Cas9s, Cas9s, nickases, binding domains. The provision of the crystal structure of CRISPR-Cas9 complexed with a nucleic acid molecule as applied in conjunction with the herein teachings provides the skilled artisan with a detailed insight into the mechanisms of action of CRISPR-Cas9. This insight provides a means to design modified CRISPR-Cas9s, such as by attaching thereto a functional group, such as a repressor or activator. While one can attach a functional group such as a repressor or activator to the N or C terminal of CRISPR-Cas9, the crystal structure demonstrates that the N terminal seems obscured or hidden, whereas the C terminal is more available for a functional group such as repressor or activator. Moreover, the crystal structure demonstrates that there is a flexible loop between approximately CRISPR-Cas9 (*S. pyogenes*) residues 534-676 which is suitable for attachment of a functional group such as an activator or repressor. Attachment can be via a linker, e.g., a flexible glycine-serine (GlyGlyGlySer) or (GGGS)₃ or a rigid alpha-helical linker such as (Ala(GluAlaAlaAlaLys)Ala). In addition to the flexible loop there is also a nuclease or H3 region, an H2 region and a helical region. By "helix" or "helical", is meant a helix as known in the art, including, but not limited to an alpha-helix. Additionally, the term helix or helical may also be used to indicate a c-terminal helical element with an N-terminal turn. The provision of the crystal structure of CRISPR-Cas9 complexed with a nucleic acid molecule allows a novel approach for drug or compound discovery, identification, and design for compounds that can bind to CRISPR-Cas9 and thus the invention provides tools useful in diagnosis, treatment, or prevention of conditions or diseases of multicellular organisms, e.g., algae, plants, invertebrates, fish, amphibians, reptiles, avians,

mammals; for example domesticated plants, animals (e.g., production animals such as swine, bovine, chicken; companion animal such as felines, canines, rodents (rabbit, gerbil, hamster); laboratory animals such as mouse, rat), and humans. The invention further involves, in place of or in addition to “in silico” methods, other “wet” methods, including high throughput screening of a binder (e.g., target nucleic acid molecule) and a candidate CRISPR-Cas9 system (e.g., *S. pyogenes* Cas9), or a candidate binder (e.g., target nucleic acid molecule) and a CRISPR-Cas9 system (e.g., *S. pyogenes* Cas9), or a candidate binder (e.g., target nucleic acid molecule) and a candidate CRISPR-Cas9 system (e.g., *S. pyogenes* Cas9) (the foregoing CRISPR-Cas9 system(s) with or without one or more functional group(s)), to select compounds with binding activity. Those pairs of binder and CRISPR-Cas9 system which show binding activity may be selected and further crystallized with the CRISPR-Cas9 crystal having a structure herein, e.g., by co-crystallization or by soaking, for X-ray analysis. The resulting X-ray structure may be compared with that of the Cas9 Crystal Structure for a variety of purposes, e.g., for areas of overlap. Having designed, identified, or selected possible pairs of binder and CRISPR-Cas9 system by determining those which have favorable fitting properties, e.g., predicted strong attraction based on the pairs of binder and CRISPR-Cas9 crystal structure data herein, these possible pairs can then be screened by “wet” methods for activity. Consequently, in an aspect the invention can involve: obtaining or synthesizing the possible pairs; and contacting a binder (e.g., target nucleic acid molecule) and a candidate CRISPR-Cas9 system (e.g., *S. pyogenes* Cas9), or a candidate binder (e.g., target nucleic acid molecule) and a CRISPR-Cas9 system (e.g., *S. pyogenes* Cas9), or a candidate binder (e.g., target nucleic acid molecule) and a candidate CRISPR-Cas9 system (e.g., *S. pyogenes* Cas9) (the foregoing CRISPR-Cas9 system(s) with or without one or more functional group(s)) to determine ability to bind. In the latter step, the contacting is advantageously under conditions to determine function. Instead of, or in addition to, performing such an assay, the invention may comprise: obtaining or synthesizing complex(es) from said contacting and analyzing the complex(es), e.g., by X-ray diffraction or NMR or other means, to determine the ability to bind or interact. Detailed structural information can then be obtained about the binding, and in light of this information, adjustments can be made to the structure or functionality of a candidate CRISPR-Cas9 system or components thereof. These steps may be repeated and re-repeated as necessary. Alternatively or additionally, potential CRISPR-Cas9

systems from or in the foregoing methods can be with nucleic acid molecules in vivo, including without limitation by way of administration to an organism (including non-human animal and human) to ascertain or confirm function, including whether a desired outcome (e.g., reduction of symptoms, treatment) results therefrom. If X-ray crystallographic or NMR spectroscopic data are provided for a CRISPR-cas system or complex of unknown crystal structure, the structure of a CRISPR-Cas9 complex as defined of the herein cited materials may be used to interpret that data to provide a likely structure for the unknown system or complex by such techniques as by phase modeling in the case of X-ray crystallography. Thus, from this disclosure and the knowledge in the art one can perform a method comprising: aligning a representation of the CRISPR-cas system or complex having an unknown crystal structure with an analogous representation of the CRISPR-cas(9) system and complex of the crystal structure herein to match homologous or analogous regions (e.g., homologous or analogous sequences); modeling the structure of the matched homologous or analogous regions (e.g., sequences) of the CRISPR-cas system or complex of unknown crystal structure based on the structure of the Cas9 Crystal Structure of the corresponding regions (e.g., sequences); and, determining a conformation (e.g. taking into consideration favorable interactions should be formed so that a low energy conformation is formed) for the unknown crystal structure which substantially preserves the structure of said matched homologous regions. "Homologous regions" describes, for example as to amino acids, amino acid residues in two sequences that are identical or have similar, e.g., aliphatic, aromatic, polar, negatively charged, or positively charged, side-chain chemical groups. Homologous regions as of nucleic acid molecules can include at least 85% or 86% or 87% or 88% or 89% or 90% or 91% or 92% or 93% or 94% or 95% or 96% or 97% or 98% or 99% homology or identity. Identical and similar regions are sometimes described as being respectively "invariant" and "conserved" by those skilled in the art. Homology modeling is a technique that is well known to those skilled in the art (see, e.g., Greer, Science vol. 228 (1985) 1055, and Blundell et al. Eur J Biochem vol 172 (1988), 513). The computer representation of the conserved regions of the CRISPR-Cas9 crystal structure and those of a CRISPR-cas system of unknown crystal structure aid in the prediction and determination of the crystal structure of the CRISPR-cas system of unknown crystal structure. The CRISPR-Cas9 crystal structure in silico may be equally applied to new CRISPR-cas crystal structures divined by using the teachings herein and

the herein cited materials concerning the CRISPR-Cas9 crystal structure. In this fashion, a library of CRISPR-Cas crystal structures can be obtained. Rational CRISPR-Cas system design is thus provided by the instant invention. For instance, having obtained a CRISPR-Cas system or complex using teachings herein and determining a conformation or crystal structure of a CRISPR-Cas system or complex, such a conformation may be used in a computer-based methods herein for determining the conformation or crystal structure of other CRISPR-Cas systems or complexes whose crystal structures are yet unknown. Data from all of these crystal structures can be in a database, and the herein methods can be more robust by having herein comparisons involving the herein crystal structure or portions thereof be with respect to one or more crystal structures in the library. Thus, there is the provision of systems, such as computer systems, intended to generate structures and/or perform rational design of a CRISPR-Cas system or complex. The system can contain: atomic co-ordinate data according to the herein cited materials concerning the Cas9 Crystal Structure or be derived therefrom e.g., by modeling, said data defining the three-dimensional structure of a CRISPR-Cas system or complex or at least one domain or sub-domain thereof, or structure factor data therefor, said structure factor data being derivable from the atomic co-ordinate data of the herein-referenced Crystal Structure. The invention also involves computer readable media with: atomic co-ordinate data according to the teachings herein e.g., by or from homology modeling, said data defining the three-dimensional structure of a CRISPR-Cas system or complex or at least one domain or sub-domain thereof, or structure factor data therefor, said structure factor data being derivable from the atomic co-ordinate data of the herein-referenced Crystal Structure. "Computer readable media" refers to any media which can be read and accessed directly by a computer, and includes, but is not limited to: magnetic storage media; optical storage media; electrical storage media; cloud storage and hybrids of these categories. By providing such computer readable media, the atomic co-ordinate data can be routinely accessed for modeling or other "in silico" methods. The invention further comprehends methods of doing business by providing access to such computer readable media, for instance on a subscription basis, via the Internet or a global communication/computer network; or, the computer system can be available to a user, on a subscription basis. A "computer system" refers to the hardware means, software means and data storage means used to analyze the atomic co-ordinate data of the present invention. The minimum hardware means of

computer-based systems of the invention may comprise a central processing unit (CPU), input means, output means, and data storage means. Desirably, a display or monitor is provided to visualize structure data. The invention further comprehends methods of transmitting information obtained in any method or step thereof described herein or any information described herein, e.g., via telecommunications, telephone, mass communications, mass media, presentations, internet, email, etc. The crystal structures of the invention can be analyzed to generate Fourier electron density map(s) of CRISPR-Cas systems or complexes; advantageously, the three-dimensional structure being as defined by the atomic co-ordinate data according to the herein-referenced Crystal Structure. Fourier electron density maps can be calculated based on X-ray diffraction patterns. These maps can then be used to determine aspects of binding or other interactions. Electron density maps can be calculated using known programs such as those from the CCP4 computer package (Collaborative Computing Project, No. 4. The CCP4 Suite: Programs for Protein Crystallography, Acta Crystallographica, D50, 1994, 760-763). For map visualization and model building programs such as "QUANTA" (1994, San Diego, Calif.: Molecular Simulations, Jones et al., Acta Crystallography A47 (1991), 110-119) can be used.

[00355] The herein cited materials concerning the Crystal Structure gives atomic co-ordinate data for a CRISPR-Cas9 (*S. pyogenes*), and lists each atom by a unique number; the chemical element and its position for each amino acid residue (as determined by electron density maps and antibody sequence comparisons), the amino acid residue in which the element is located, the chain identifier, the number of the residue, co-ordinates (e.g., X, Y, Z) which define with respect to the crystallographic axes the atomic position (in angstroms) of the respective atom, the occupancy of the atom in the respective position, "B", isotropic displacement parameter (in angstroms²) which accounts for movement of the atom around its atomic center, and atomic number. The conformational variations in the crystal structures of the CRISPR-Cas9 system or of components of the CRISPR-Cas9 may provide important and critical information about the flexibility or movement of protein structure regions relative to nucleotide (RNA or DNA) structure regions that may be important for CRISPR-Cas system function. The structural information provided for Cas9 (e.g. *S. pyogenes* Cas9) as the CRISPR enzyme in the present application may be used to further engineer and optimize the CRISPR-Cas system and this may be extrapolated to interrogate structure-function relationships in other CRISPR enzyme systems

as well. The herein cited materials relate to the crystal structure of *S. pyogenes* Cas9 in complex with sgRNA and its target DNA at 2.4 Å resolution. The structure revealed a bilobed architecture composed of target recognition and nuclease lobes, accommodating a sgRNA:DNA duplex in a positively-charged groove at their interface. The recognition lobe is essential for sgRNA and DNA binding and the nuclease lobe contains the HNH and RuvC nuclease domains, which are properly positioned for the cleavage of complementary and non-complementary strands of the target DNA, respectively. This high-resolution structure and the functional analyses provided herein elucidate the molecular mechanism of RNA-guided DNA targeting by Cas9, and provides an abundance of information for generating optimized CRISPR-Cas systems and components thereof. The crystal structure in conjunction with herein teachings may provide steps towards understanding the molecular mechanism of RNA-guided DNA targeting by Cas9. The structural and functional analyses of herein teachings and herein cited materials provide a useful scaffold for rational engineering of Cas9-based genome modulating technologies and may provide guidance as to Cas9-mediated recognition of PAM sequences on the target DNA or mismatch tolerance between the sgRNA:DNA duplex. Aspects of the invention may also relate to truncation mutants, e.g. an *S. pyogenes* Cas9 truncation mutant may facilitate packaging of Cas9 into size-constrained viral vectors for in vivo and therapeutic applications. Similarly, engineering of the PAM Interacting (PI) domain may allow programming of PAM specificity, improve target site recognition fidelity, and increase the versatility of the Cas9 genome engineering platform.

[00356] The invention comprehends optimized functional CRISPR-Cas enzyme systems. In particular the CRISPR enzyme comprises one or more mutations that converts it to a DNA binding protein to which ligase domains may be recruited or appended or inserted or attached. In certain embodiments, the CRISPR enzyme comprises one or more mutations which include but are not limited to D10A, E762A, H840A, N854A, N863A or D986A (based on the amino acid position numbering of a *S. pyogenes* Cas9) and/or the one or more mutations is in a RuvC1 or HNH domain of the CRISPR enzyme or is a mutation as otherwise as discussed herein. In some embodiments, the CRISPR enzyme has one or more mutations in a catalytic domain, wherein when transcribed, the tracr mate sequence hybridizes to the tracr sequence and the guide sequence directs sequence-specific binding of a CRISPR complex to the target sequence.

[00357] The teachings herein and structural information of herein cited materials provided herein allows for interrogation of sgRNA (or chimeric RNA) interaction with the target DNA and the CRISPR enzyme (e.g. Cas9) permitting engineering or alteration of sgRNA structure to optimize functionality of the entire CRISPR-Cas system. For example, loops of the sgRNA may be extended, without colliding with the Cas9 protein by the insertion of distinct RNA loop(s) or distinct sequence(s) that may recruit adaptor proteins that can bind to the distinct RNA loop(s) or distinct sequence(s). The adaptor proteins may include but are not limited to orthogonal RNA-binding protein / aptamer combinations that exist within the diversity of bacteriophage coat proteins. A list of such coat proteins includes, but is not limited to: Q β , F2, GA, fr, JP501, M12, R17, BZ13, JP34, JP500, KU1, M11, MX1, TW18, VK, SP, FI, ID2, NL95, TW19, AP205, ϕ Cb5, ϕ Cb8r, ϕ Cb12r, ϕ Cb23r, 7s and PRR1. These adaptor proteins or orthogonal RNA binding proteins can further recruit effector proteins or fusions which comprise one or more ligase domains.

[00358] Although the present invention and its advantages have been described in detail, it should be understood that various changes, substitutions and alterations can be made herein without departing from the spirit and scope of the invention as defined in the appended claims.

[00359] The present invention will be further illustrated in the following Examples which are given for illustration purposes only and are not intended to limit the invention in any way.

Examples

Example 1: Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis.

Materials and methods

Generation of Cas9-GFP expression vector and transduction of mouse NSCLC cell lines

[00360] A lentiviral vector, lenti-NLS-FLAG-Cas9-GFP, was generated by subcloning Cas9 from pX330 vector into a lentiviral vector (Cong et al., 2013). The plasmid was co-transfected into HEK293FT cells (Invitrogen R700-07) with lentiviral packaging plasmids psPAX2 and pMD2.G (Addgene 12260 and 12259). HEK293FT cells were cultured in D10 media and seeded in a 15cm culture dish the day before transfection such that they would be 80-90% confluent at the time of transfection. Two hours before transfection, the media was replaced with 15 ml of

prewarmed OptiMEM (Invitrogen 51985-091). The transfection mix was aspirated after 12 hours and replaced with fresh prewarmed DMEM + 10% FBS. Viral particles were harvested 48 hours after this media change and frozen at -80C.

Cas9-GFP expressing mouse NSCLC cell lines

[00361] A mouse NSCLC cell line with genotype *Kras*G12D/+;*p53*-/-;*Dicer1*+/- (KPD cell line) was derived from a mouse model (Kumar et al., 2009) and described in (Chen et al., 2014) and cultured in high-glucose DMEM (Invitrogen 10566-024) + 10% FBS (Hyclone). This cell line was transduced with a lentiviral vector encoding NLS-Cas9-GFP (lentiCas9-EGFP) at low MOI (<0.01). GFP-positive cells were sorted as single cells into 96-well plates and cultured as clonal cell lines. Multiple clonal lines were established and genotyped by PCR. Lines with 100% cell GFP-positive were kept and those with segregating GFP expression were discarded. FLAG-Cas9 expression was confirmed by antibodies against FLAG (Sigma) or Cas9 (Diagenode).

Pooled guide-only library cloning and viral production

[00362] A genome-wide mouse CRISPR knockout guide-only library (Sanjana et al., 2014) containing 67,405 sgRNAs (mGeCKOa, Addgene 1000000053) was used for the transduction of the clonal lentiCas9-EGFP-expressing NSCLC cells. Briefly, the library of sgRNA oligos was cloned into an sgRNA-expressing vector that also expresses puromycin resistance (lentiGuide-Puro) using Gibson assembly, transformed into *E. coli* electrocompetent cells (Lucigen), and spread onto ten 245 mm x 245 mm plates (Corning) agar plates containing 100 µg/ml carbenicillin. The number of colonies was estimated to be $> 2.5 \times 10^6$ ($> 35 \times$ coverage). All colonies were scraped off and pooled for plasmid purification using EndoFree Maxi Kit (Qiagen) with 0.5g per maxi reaction.

[00363] After plasmid purification, Applicants amplified the libraries using the following primers in to ensure adequate representation: *Library Amp F*: AATGATACGGCGACCAACGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC TtAAGTAGAGTcttgtggaaggacgaaacaccg (SEQ ID NO:X).; *Library Amp R*: CAAGCAGAAGACGGCATAACGAGATAAGTAGAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTtTCTACTATTCTTTCCCCTCACTGT (SEQ ID NO:X).. With ~15M reads, Applicants identified 94% of the designed sgRNAs and there was a <10-fold difference in

representation between the 10th and 90th percentile of sgRNA read counts. Two technical replicates of the PCR amplification and sequencing readout of the plasmid library (15M reads per replicate) were well correlated ($r^2=0.98$).

[00364] To produce virus, the mGeCKOa pooled plasmid was co-transfected into HEK293FT cells (Invitrogen R700-07) with lentiviral packaging plasmids psPAX2 and pMD2.G (Addgene 12260 and 12259). HEK293FT cells were cultured in DMEM (Invitrogen 10566-024) + 10% FBS (Hyclone) and seeded in T-225 flasks the day before transfection such that they would be 80-90% confluent at the time of transfection. Two hours before transfection, the media was replaced with 13 ml of prewarmed OptiMEM (Invitrogen 51985-091). For transfection of each T-225 flask, 200ul of Plus reagent (Invitrogen 11514-015) was diluted into 4mL of OptiMEM and then the following DNA was added: 20 ug mGeCKOa, 15 ug psPAX2, 10 ug pMD2.G. Separately, 100ul of Lipofectamine 2000 was diluted into 4mL of OptiMEM, briefly vortexed, and incubated at room temperature for 5 minutes. After incubation, the Plus+DNA and Lipofectamine 2000 (Invitrogen 11668019) mixtures were combined, briefly vortexed, and incubated at room temperature for 20 minutes. The mixture was then very gently added to the T-225 flask with 13 ml OptiMEM. All media was aspirated after 6 hours and replaced with fresh prewarmed DMEM + 10% FBS. Viral particles were harvested 48 hours after this media change and frozen at -80°C.

Pooled library transduction into mouse NSCLC cell line

[00365] The virus was titred by spinfection of 3×10^6 mouse NSCLC cells per well in a 12-well plate with different dilutions of the virus (and no virus control) in each well. After adding virus, cells were spun at 2000 rpm for 2 hours at 37 °C and then placed into the incubator overnight. The next day 1.5×10^5 cells from each viral concentration were plated into two replicate wells of a 6-well plate. At 24 hours post-transduction, 2 µg/ml puromycin (Sigma) was added to one of the replicate wells. Cells not transduced with the library (no virus control) did not survive past 24 hours with puromycin at 2 µg/ml. After 72 hours, cells were counted in all wells to determine the viral volume that resulted in 20-40% of cells surviving in puromycin. Assuming infection events occur independently, this corresponds to a multiplicity of infection

(MOI) of 0.2-0.5 and a single-infection percentage (SIP) of 77% (at 40% puromycin survival) to 89% (at 20% puromycin survival).

[00366] The SIP is calculated directly from the puromycin survival ($p_{survival}$), as shown below. The probability of a cell being infected by n viral particles if the MOI is m is given by the Poisson distribution,

$$P(n) = \frac{m^n \cdot e^{-m}}{n!}$$

and the observed percent of cells surviving puromycin selection is

$$p_{survival} = P(n > 0) = 1 - P(n = 0) = 1 - e^{-m}$$

Solving for the SIP as a function of $p_{survival}$, we get:

$$\begin{aligned} SIP &= \frac{P(n = 1)}{P(n \geq 1)} = \frac{P(n = 1)}{P(n > 0)} \\ &= -\frac{(1 - p_{survival}) \ln(1 - p_{survival})}{p_{survival}} \end{aligned}$$

[00367] For each experimental replicate of the pooled screen, a total of 1.2×10^8 cells were infected at MOI ~ 0.4 , and selected with puromycin at 2 ug/ml for 7 days. MOI was calculated with an in-line control using a similar procedure as given in the previous paragraph. Infected cells were expanded under puromycin selection for 7 days and split every 2-3 days. After 7 days, 3×10^7 cells were spun down and frozen for genomic DNA extraction. At the same time, 1.7×10^8 cells were washed twice in sterile PBS and resuspended at 5×10^7 cells/ml in PBS for tumor transplantation.

Animal work statement

[00368] All animal work were performed under the guidelines of Division of Comparative Medicine (DCM), with protocols (0411-040-14, 0414-024-17, 0911-098-11, 0911-098-14 and 0914-091-17) approved by Massachusetts Institute of Technology Committee for Animal Care

(CAC), and were consistent with the Guide for Care and Use of Laboratory Animals, National Research Council, 1996 (institutional animal welfare assurance no. A-3125-01).

Mice, tumor transplant and metastasis analysis

[00369] NSCLC cells infected with NLS-Cas9-GFP vector only or NLS-Cas9-GFP and mGeCKOa library were injected subcutaneously into the right side flank of Nu/Nu mice at 3×10^7 cells per mouse. Transplanted primary tumor sizes were measured by caliper. At 6 weeks post transplantation, mice were sacrificed and several organs (liver, lung, kidney and spleen) were dissected for examination of metastases under a fluorescent stereoscope.

Mouse tissue collection

[00370] Primary tumors and other organs were dissected manually. For molecular biology, tissues were flash frozen with liquid nitrogen, ground in 24 Well Polyethylene Vials with metal beads in a GenoGrinder machine (OPS diagnostics). Homogenized tissues were used for DNA/RNA/protein extractions using standard molecular biology protocols. Tissues for histology were then fixed in 4% formaldehyde or 10% formalin overnight, embedded in paraffin, and sectioned at 6 μ m with a microtome as described previously (Chen et al., 2014). Slices were subjected to hematoxylin and eosin (H&E) staining as described previously (Chen et al., 2014).

MicroCT

[00371] MicroCT (μ CT) were performed as described previously (Platt et al., 2014). Briefly, μ CT imaging was performed using standard imaging protocol with a μ CT machine (GE Healthcare). Briefly, animals were anesthetized using isoflurane, and setup in the imaging bed with a nosecone providing constant isoflurane. A total of 720 views were acquired for each mouse using a soft-tissue-fast-scan setting. Raw image stacks were processed for lung reconstruction using the standard ROI tool (MicroView). Rendering and quantification were performed using render volume tool and measurement tool in MicroView.

Genomic DNA extraction from cells and mouse tissues

[00372] Genomic DNA from cells and tissues (primary tumors and lungs) was extracted using a homemade modified salting out precipitation method similar to the Puregene (Qiagen/Gentra) procedure. After a hands-on comparison of several commercial DNA extractions kits, such as the

QIAamp Blood Midi/Max (Qiagen), the Quick-gDNA MidiPrep (Zymo) and Puregene (Qiagen), Applicants found that Applicants' homemade version provided consistent, high-quality yields with a low-cost, simple protocol.

[00373] For gDNA extraction from either 100-200 mg of frozen ground tissue or 3×10^7 - 5×10^7 frozen cells, the same procedure was used. For different amounts of tissue or cells, the quantities below can be scaled as needed. In a 15 ml conical tube, Applicants added 6 ml of NK Lysis Buffer (50 mM Tris, 50 mM EDTA, 1% SDS, pH 8) and 30 μ l of 20 mg/ml Proteinase K (Qiagen 19131) to the tissue/cell sample and incubated at 55 °C overnight. The next day, Applicants added 30 μ l of 10 mg/ml RNase A (Qiagen 19101, diluted in NK Lysis Buffer to 10 mg/ml) to the lysed sample, inverted 25 times and incubated at 37 °C for 30 minutes. Samples were cooled on ice before addition of 2 ml of pre-chilled 7.5M ammonium acetate (Sigma A1542) to precipitate proteins. Stock solutions of 7.5M ammonium acetate should be made in sterile H₂O and kept at 4 °C until use. After adding ammonium acetate, the samples were vortexed at high speed for 20 seconds and then centrifuged at $\geq 4,000 \times g$ for 10 minutes. After the spin, a tight pellet was visible and the supernatant was carefully decanted into a new 15 ml conical tube. Applicants then added 6 ml 100% isopropanol to the tube, invert 50 times and centrifuged at $\geq 4,000 \times g$ for 10 minutes. Genomic DNA were visible as a small white pellet. Applicants discarded the supernatant and add 6 ml of freshly prepared 70% ethanol, inverted the tube 10 times, and centrifuged at $\geq 4,000 \times g$ for 1 minute. Applicants then discarded the supernatant by pouring, briefly spun again, and used a P200 pipette to remove any remaining ethanol. After air drying for 10-30 minutes, the DNA changed appearance from a milky white pellet to slightly translucent. At this stage, Applicants added 500 μ l of 1x TE buffer (Sigma T9285) and incubated at 65 °C for 1 hour and at room temperature overnight to fully resuspend the DNA. The next day, Applicants vortexed the samples briefly and then used a Nanodrop (Thermo Scientific) to measure the DNA concentration.

SgRNA library readout by deep sequencing

[00374] The sgRNA library for each sample (plasmid, genomic DNA from cells and tissues) was amplified and prepared for Illumina sequencing using a two-step PCR procedure. For

PCR#1, we amplified using primers specific to the sgRNA-expression vector (lentiGuide-PCR1-F (pHKO25-F3), and lentiGuide -PCR1-R (pHKO25-R1)):

[00375] pHKO25-F3 CCCGAGGGGACCCAGAGAG (SEQ ID NO:X).

[00376] pHKO25-R1 GCGCACCGTGGGCTTGTAC (SEQ ID NO:X).

[00377] All PCR was performed using Phusion Flash High Fidelity Master Mix (Thermo). For PCR#1, the thermocycling parameters were: 98 °C for 30s, 18-24 cycles of (98 °C for 1s, 62 °C for 5s, 72 °C for 35 s), and 72 °C for 1 minute. In each PCR#1 reaction, Applicants used 3 µg of gDNA. For each sample, the appropriate number of PCR#1 reactions was used to capture the full representation of the screen. For example, at ~400x coverage of Applicants' 67,405 mGeCKOa sgRNA library, Applicants used gDNA from 3×10^7 cells. Assuming 6.6 pg of gDNA per cell, Applicants would use ~200 µg of gDNA per sample. Since Applicants used 3 µg of gDNA per 100 µl PCR#1 reaction, each biological sample required 67 PCR#1 reactions.

[00378] PCR#1 products for each biological sample were pooled and then amplified with barcoded second PCR primers (Figure 18 -Supplemental Tables). For each sample, Applicants performed at least 7 PCR#2 reactions (One 100ul reaction per 10^4 constructs in the library) using 10 µl of the pooled PCR#1 product per PCR#2 reaction.

[00379] Second PCR products were pooled and then normalized for each biological sample before combining uniquely barcoded separate biological samples. The pooled product was then gel purified from a 2% E-gel EX (Life Technologies) using the QiaQuick kit (Qiagen). The purified pooled library was then quantified with dsDNA High-Sensitivity Qubit (Life Technologies) and/or a gel-based method using the Low-Range Quantitative Ladder Life Technologies). Diluted libraries with 5-20% PhiX were sequenced with MiSeq or HiSeq 2500 (Illumina).

sgRNA deep sequencing data processing

[00380] Deep sequencing data were processed for sgRNA representation using custom scripts. Briefly, sequencing reads were demultiplexed using the 8-basepair barcodes in the reverse primer, and then demultiplexed using the 8-basepair barcodes in the forward primer. Demultiplexed reads were trimmed using cutadapt (Martin, 2011), leaving only the 20bp spacer

(guide) sequences. The spacer sequences were then mapped to the spacers of designed sgRNA library using bowtie (Langmead et al., 2009). For mapping, a maximum of one mismatch was allowed in the 20 bp spacer sequence. Mapped sgRNA spacers were then quantified by counting the total number of reads. In each biological sample, any sgRNA spacer with only a single read was filtered out. The total numbers of reads for all sgRNAs in each sample were normalized. Figures were generated using the normalized read counts in R and RStudio (R project, Revolution Analytics), Matlab (Mathworks), and the Gene Set Enrichment Analysis tool (Broad Institute).

Validation and control minipool synthesis, viral transduction and in vivo transplantation

[00381] Validation and control minipools were synthesized in a single oligonucleotide pool using a semiconductor-based electrochemical detritylation synthesis (CustomArray) and SAFC Proligo reagents (Sigma). Each minipool was separately PCR amplified from the pooled synthesis, gel purified and cloned into the lentiGuide-Puro vector using Gibson assembly, as previously described (Sanjana et al., 2014; Shalem et al., 2014). Cloned minipool libraries were deep-sequenced to verify representation (Illumina) and lentivirus was produced in the same manner as for the mGeCKOa library. As before, validation and control minipool library viruses were titred by spinfection of 3×10^6 mouse Cas9-EGFP NSCLC cells per well in a 12-well plate with different dilutions of the virus (and no virus control) in each well to find what viral volume was need to yield a 20-40% survival after treatment for 48 hours with 2 ug/ml puromycin (Sigma).

[00382] Using these viral volumes, the same clonal Cas9-EGFP NSCLC mouse cell line as used for the mGeCKOa screen was transduced via spinfection with either control or validation minipool virus. Representation for both screens (validation and control minipool) was ~1,000 fold. Cells were maintained in D10 media with 2 ug/ml puromycin and split 1:4 every 2-3 days.

[00383] After 7 days in culture, Cas8-EGFP NSCLC cells infected with validation minipool or control minipool were injected subcutaneously into the right side flank of *Nu/Nu* mice at 3×10^7 cells per mouse. For each minipool, cells were transplanted into 5 mice. After five weeks, mice were sacrificed and primary tumors and lungs were dissected. Histology samples and genomic DNA were collected in a similar manner as described above for the mGeCKOa screen.

Single sgRNA design, lentiviral transduction and in vivo transplantation

[00384] Six sgRNAs per protein coding gene and four per microRNA gene were chosen for single agent validation. For protein coding genes, we cloned both the 3 sgRNAs from the mGeCKOa library and 3 additional sgRNAs to target each gene. For miRNAs, we used all 4 sgRNAs from the mGeCKOa library. SgRNAs targeting individual genes/miRNAs were cloned into the lentiGuide-Puro vector (Addgene). Single sgRNA viruses were generated by transfection of HEK293FT using the same procedure as described for the mGeCKOa library virus. After harvest, viruses were functionally tittered (percent puromycin survival) and used to transduce Cas9-EGFP NSCLC cells at a MOI $\ll 1$ for all transductions. Cells were maintained in D10 media (with 2 ug/ml puromycin added at 24 hours post-transduction) and split 1:4 every 2-3 days.

[00385] After 3 days in culture, a portion of the cells infected with each single sgRNA virus was collected in QuickExtract buffer (Epicentre) and subjected to amplicon-sequencing (Illumina) for indels, as described previously (Hsu et al., 2013). After 7 days in culture, another portion of the cells was collected in RIPA buffer for western blot of protein levels using antibodies against Nf2 or Pten (CST). SgRNAs that were efficient in generating indels or in reducing protein levels were chosen for in vivo experiments. After 7 days in culture, cells were injected subcutaneously into the right side flank of *Nu/Nu* mice at 5×10^6 cells per mouse. Each gene or microRNA was targeted with three independent sgRNAs that went into validation experiments, with two mice injected per sgRNA. A total of 6 mice were used to validate each gene or miRNA. After five weeks, mice were sacrificed and primary tumors and lungs were dissected. Histology samples were collected and analyzed in a similar manner as described above for the mGeCKOa screen.

CTC chip design and capture

[00386] Applicants designed a microfluidic device for CTC capture similar to one previously described (Chung et al., 2013). This device has two functions: magnetic depletion of leukocytes, and capture CTCs based on their size criteria (5~30 μm). The microfluidic system were fabricated with standard soft lithography and bonded with glass substrate. Each CTC chip contains a large number of capturing sites ($>10,000$). Peripheral blood was collected from mice after

injected with cells for five weeks, using terminal cardiac puncture method. The blood samples were depleted of red blood cells using red blood cell lysis buffer (BD and/or Miltenyl), fixed with 2% formaldehyde and subject to CTC capture. For each sample, 250~500 µl of the fixed cells was run through the chip at flow rate of 2-5 mL/hour. The chips with captured cells were imaged under a fluorescence microscope with 15~17 images tiling the whole chip for each sample. GFP-positive cells were counted across all images for each chip using custom Matlab scripts.

Standard molecular biology

[00387] Routine DNA cloning, nucleic acid purification western blot were performed using standard molecular biology protocols with commercially available kits.

Results

CRISPR/Cas9 library-mediated mutagenesis accelerates metastasis

[00388] The Applicants derived and cloned a cancer cell line (Chen et al., 2014) from a mouse non-small cell lung cancer (NSCLC) (Kumar et al., 2009). This cell line possesses an oncogenic *Kras* in conjunction with homozygous *p53* loss and heterozygous *Dicer1* loss (*Kras*^{G12D}/+;*p53*^{-/-};*Dicer1*^{+/-}, or KPD), and is capable of inducing tumors when transplanted into immunocompromised mice (Chen et al., 2014; Kumar et al., 2009). This KPD cell line was transduced with a lentivirus carrying a Cas9 transgene fused to a nuclear localization sequence (NLS), a triple FLAG epitope (3xFLAG), and a green fluorescent protein (GFP). The transduction was performed at low multiplicity of infection (MOI < 0.01) to ensure that most cells receive only one lentiviral transgene. Clonal cell lines expressing Cas9-GFP were established (termed Cas9-GFP KPD cell lines thereafter), as shown by fluorescence imaging and western blot using antibodies against Cas9 or FLAG (Figure 8 A-B). A clonal Cas9-GFP KPD cell line was selected to provide cellular homogeneity for the subsequent screen.

[00389] For the knockout screens, the Applicants utilized a pooled genome-wide mouse sgRNA library cloned into lentiGuide-Puro, a bicistronic lentiviral vector that expresses both the sgRNA and a puromycin resistance gene (Sanjana et al., 2014). This library (termed mouse Genome-scale CRISPR knockout library A, or mGeCKOa thereafter) has 67,405 sgRNAs

targeting 20,611 annotated protein-coding and 1,175 microRNA precursor genes in the mouse genome. The library also contains 1,000 control sgRNAs (termed non-targeting sgRNAs) designed to have minimal homology to sequences in the mouse genome (Sanjana et al., 2014; Shalem et al., 2014). The Applicants transduced the Cas9-GFP KPD cell line with the mGeCKOa lentiviral library with three independent infection replicates with greater than 400x coverage for each replicate (Figure 1 A). In these three infection replicates, the MOI was on average 0.4 ± 0.02 , which ensures that over 80% of the cells surviving puromycin selection received only one sgRNA-expressing lentiviral integrant.

[00390] After culture *in vitro* for a week, the Applicants subcutaneously transplanted 3×10^7 cells into the flanks of immunocompromised *Nu/Nu* mice (Figure 1 A). With the cells from each infection replicate, Applicants transplanted 4 mice, using one mouse for early tumor sequencing and three for later stage (see below)(Figure 1A). Both mGeCKOa transduced and untransduced Cas9-GFP KPD cells form tumors at the injection site (Figure 1 B). Like most subcutaneously transplanted tumors, these tumors were poorly differentiated. The primary tumors induced by mGeCKOa transduced cells grew slightly faster than tumors from the untransduced cells at an early stage (Figure 1 C) (two weeks post-transplantation) (paired two-tailed *t*-test, $p = 0.05$), but at late stages all tumors were similar in size (paired two-tailed *t*-test, $p = 0.18$ for four weeks, and $p = 0.6$ for six week) (Figure 1 C).

[00391] At 6 weeks post-transplantation, Applicants imaged the mice using micro-computed tomography (μ CT), and found tumors in the lungs of the mice transplanted with mGeCKOa transduced Cas9-GFP KPD cells (mGeCKOa mice), but not in the mice transplanted with untransduced Cas9-GFP KPD cells (control mice) (Figure 1 D, Figure 8 (S1) C). Mice were sacrificed and examined for metastases in various organs. Under a fluorescent stereoscope at 6x magnification, metastases were visually detected in the lung in 90% (8/9) of the mGeCKOa mice (Figure 8 (S1) D). In three independent experiments, mGeCKOa mice on average had 80% of their lobes positive for metastases (Figure 1 E). In contrast, none (0/3) of the Cas9-only control animals developed detectable metastases (Figure 1 E). At this time, metastases were not detected in the liver, kidney or spleen in either group (Figure 1 F). These data suggested that mGeCKOa

library transduction accelerate the ability of the Cas9-GFP KPD cells to form metastases in the lung.

Dynamic evolution of sgRNA library representation during tumor growth and metastasis

[00392] To investigate the sgRNA representation through different stages of tumor evolution and to infer genes whose loss-of function confers a preliferative or metastatic phenotype, Applicants used deep sequencing to read out the sgRNA representation. At six-week post transplantation, Applicants harvested and sequences the late stage primary tumor and three random lobes from the lung of each of the nine mGeCKOa mice (Figure 1 A). In parallel, Applicants also sequenced the mGeCKOa plasmid library, the pre-transplantation mGeCKOa transduced Cas9-GFP KPD cells, as well as early stage primary tumors (two-week post transplantation, one animal from each replicate experiment). In the cell samples, the sgRNA representations showed high concordance between technical replicates (correlation, $\rho = 0.95$ on average) and biological infection replicates (correlation, $\rho = 0.84$ on average) (Figure 2 A, Figure 9 (S2) A, B, F). Cell pools on average highly correlate with the plasmid representation (correlation, $\rho = 0.93$) (Figure 2 A, Figure 9 (S2) C, F). Furthermore, different sgRNAs that target the same gene are correlated in terms of rank change (correlation $\rho = 0.49$ on average)(Figure 9(S2) D). The sgRNAs with significantly decreased abundance in cells compared to plasmid are enriched for multiple categories of essential genes, such as ribosomal proteins, translation factors, RNA splicing factors and RNA processing factors, indicating selection against the loss of these genes after one-week in culture (Figure 9 (S2) F).

[00393] To investigate the sgRNA library dynamics in different sample types (plasmid, cell, early primary tumor, late primary tumor and lung metastases), Applicants compared the overall distributions of sgRNAs from all samples sequenced. Cell samples tightly clustered with each other and the plasmid, forming a cell-plasmid clade in the overall clustering dendrogram (Figure 2 A, Figure 9(S2) E). Early primary tumor samples also clustered with each other and then with the cell-plasmid clade, whereas late tumors and lung metastases clustered with each other as a clade (Figure 2 A, Figure 9(S2) E). The detected sgRNAs in different replicates of infected cell populations overlap by more than 95% with each other (Figure 10 (S3) A). The detected sgRNAs in the three biological replicates of early tumor overlap 63~76% with each other (Figure 10 (S3)

B). Early primary tumors retained less than half (31.5 ~ 49.4%) of the sgRNAs as compared to the cell populations (Figure 2 B-C, Figure 10 (S3) C-D). Compared to the cell pool, sgRNAs whose targets are essential genes are further depleted in early tumors (Figure 15 - Table S1).

[00394] Interestingly, only a small fraction of sgRNAs (less than 4% of all sgRNAs, or less than 8% of the sgRNAs in early primary tumor of corresponding replicate) was detected in the late stage primary tumor samples (Figure 2 B-C, Figure 10 (S3) C-D). The sgRNA diversity (i.e. number of different sgRNAs detected) further decreases in samples from lung metastases (Figure 2 B-C, Figure 10 (S3) C-D). The lung samples retained at most 0.4% of all sgRNAs, or at most 1.1% of the sgRNAs in early primary tumor of corresponding replicate were found in the lung samples, with a subset of highly enriched ones (Figure 2 B-C, Figure 10 (S3) C-D). The global patterns of sgRNA distributions in different sample types are distinct, as is evident in the strong shifts in respective cumulative distribution functions (Kolmogorov-Smirnov (KS) test, $p < 10^{-15}$ for all pairwise comparisons) (Figure 2 D), reflective of cancer evolution at different stages.

Enriched sgRNAs in primary tumors

[00395] Late primary tumors retain few sgRNAs (on average 813 +/- 264 sgRNAs, n=9 mice), with even fewer at high frequencies (4 +/- 1 sg RNAs with less than 5% of total reads) in each mouse (Figure 2 B-C, Figure 9 (S2) C-D, Figure 3 A). Applicants used three methods to identify enriched sgRNAs in late primary tumor. The first method was based on relative abundance, i.e. by taking sgRNAs above certain threshold; the second method was based on ranks, i.e. by taking top ranked sgRNAs in each mouse; the third method was based on false discovery rate (FDR), i.e. by taking sgRNAs enriched compared to the distribution of the 1,000 non-targeting sgRNAs at 0.2% FDR. All three methods generated similar results (Figure 11 (S4) A). Taking the results from the FDR method as an example, a total of 935 sgRNAs (targeting 909 genes) are enriched over the non-targeting controls (FDR cut-off=0.2%) in the late primary tumor of one or more mice (Figure 3 B-C). These sgRNAs target genes that are highly enriched in apoptosis pathway (Figure 18 - Supplemental tables), with many of them being pro-apoptotic, such as *BH3 interacting-domain death agonist (Bid)*, *Phosphatase and Tensin Homolog (Pten)*, *Cyclin-dependent kinase inhibitor 2a (Cdkn2a)*, and *O-6-Methylguanine-DNA Methyltransferase (Mgmt)*, suggesting strong selection for mutations inactivating apoptosis in primary tumor cells.

[00396] There are 24 genes targeted by two or more independent sgRNAs enriched in late primary tumors (Figure 3 B-C). These genes were found to be mutated in patients in many previously reported cancer sequencing studies curated by cBioPortal (Cerami et al. 2013; Gao et al. 2013). For example, in somatic mutations identified by The Cancer Genome Atlas (TCGA) for non-small cell lung cancer (NSCLC), including adenocarcinoma (LUAD) (TCGA-Network, 2014b) and lung squamous cell carcinoma (LUSC) (TCGA-Network, 2012), 36% (107/407) of the patients have one or more of these 24 genes mutated (Figure 12 (S5) B-C). Several of them are well-known tumor suppressors, such as *Pten*, *Cyclin-dependent kinase inhibitor 2b* (*Cdkn2b*), *Neurofibromin 2* (*Nf2*, *Merlin*), *Alpha-type platelet-derived growth factor receptor* (*Pdgfra*), and *Integrin alpha X* (*Itgax*). Several of these genes are associated with cancer, but their functions in tumor growth are poorly understood. For example, *Mgmt*, a gene with two enriched sgRNAs, is required for DNA repair and thus crucial for genome stability (Kaina et al., 2007; Tano et al., 1990; Teo et al., 2001). Mutation, silencing or promoter methylation of *MGMT* is associated with primary glioblastomas (Hegi et al., 2005; Jesien-Lewandowicz et al., 2009; Molenaar et al., 2014). *Med16* encodes the 16th subunit of the mediator complex of transcription regulation (Herbig et al., 2010), which has been recently implicated in cancer (Huang et al., 2012; Schiano et al., 2014).

Enriched sgRNAs in metastases

[00397] Applicants also sequenced the sgRNA distributions from three lung lobes for each mouse transplanted with mGeCKOa-transduced Cas9-GFP KPD cell. In each lobe the sgRNA representation is dominated by one or a few sgRNAs (Figure 4A, Figure 10(S3)D). In each mouse, the lung sgRNA representation (average of normalized sgRNA representations from three lobes) is also dominated by a small number of sgRNAs (on average 3.4 +/- 0.4 sgRNAs with > 5% of total reads)(Figure 4 B), suggesting those select sgRNAs-integrated cells seeded metastases and grew to dominance over this time scale. Non-targeting sgRNAs were occasionally detected in the metastases, but never observed at high frequency (every Non-targeting sgRNA occupies <0.1% of total reads in any lobe) (Figure 2 C, Figure 4 A-B). These observations are consistent with the fact that untransduced tumors are not metastatic (Figure 1E), suggesting specific sgRNA-mediated mutations led to metastasis.

[00398] The sgRNA representations in the lung metastases are similar to those in the late stage primary tumors in several ways. First, the detected sgRNAs in lung samples significantly overlap with those in late tumor samples (Chi-square test, $p < 10^{-15}$) (Figure 10 (S3) E). Second, the number of sgRNA detected in lung samples correlates, albeit weakly, with the number of sgRNAs detected in late primary tumor samples (correlation, $\rho = 0.42$, F-test, $p = 0.097$) (Figure 10 (S3) F). Third, as a genome-wide measurement, the abundance of sgRNAs in the lung positively correlates with that in the late primary tumors of the same animal (correlation, $\rho = 0.18$ on average, F-test, $p < 0.001$, $n = 9$) (Figure 10 (S3) G). Fourth, in most mice (8/9), the lung metastasis enriched sgRNAs occupy a large fraction of reads in the late primary tumor of the same animal (Figure 4 E, left panel), significantly larger than random samplings of the same numbers of sgRNAs (Figure 4 E, right panel). These data indicate that mutants with preferential ability to proliferate in late primary tumors are more likely to dominate the metastases.

[00399] The three methods (by threshold, by rank, or by FDR) of finding enriched sgRNAs in the lung metastases yield similar results, overlapping with each other by 87~99% in pairwise comparison (Figure 11 (S4) B). Taking the analysis by FDR for example, the detected sgRNAs in different lobes of the same animal overlap with each other by more than 40% (Figure 11 (S4) C), while different animals show substantial variability (overlap by 10~20%) (Figure 11 (S4) D). The detected sgRNAs in different biological / infection replicate experiments, with a union set of all animals pooled from each replicate, overlap with each other by more than 55% (Figure 11 (S4) E), suggesting pooling animals in the same experiment facilitate capturing the identification of shared strong hits. Overall, a small fraction of sgRNAs was found enriched multiple times (in one or more lobes, or in one or more animals) (Figure 4 C-D). These include sgRNAs targeting *Nf2*, *Pten*, *Tripartite motif-containing protein 72 (Trim72)*, *Bid*, *Cyclin-dependent kinase inhibitor 2a (Cdkn2a)*, *Zinc Finger, FYVE Domain Containing 28 (Zfyve28)*, *Reproductive homeobox 13 (Rhox13)*, *BRISC and BRCA1 A complex member 1 (Babam1)*, as well as microRNA genes *miR-152* and *miR-345*. Intriguingly, a few sgRNAs targeting the PolII subunits and olfactory receptor are also enriched in the lung, likely due to unknown or off-target effects. In four genes, *Nf2*, *Pten*, *Trim72*, and *Zfyve28*, two independent sgRNAs targeting different regions of the same gene were observed as enriched (Figure 4 H). One of the *Zfyve28* sgRNA, however, is enriched in only one animal; whereas *Nf2*, *Pten* or *Trim72* all have two sgRNAs

enriched in multiple animals. These three genes, several representative genes with one sgRNA with high enrichment in more than one sample (*Cdkn2a*, *Fga*, *Cryba4*), and the top two scoring microRNAs (*miR-152* and *miR-345*) were chosen to assay individually for primary tumor and metastases formation. The most frequently enriched sgRNA is an sgRNA targeting *Nf2*, appearing in 90% of the lobes (Figure 4 D), or 100% of the animals (Figure 4 E), suggesting *Nf2* loss-of-function mutants are the best competitors in the library for metastases in this setting.

[00400] For most sgRNAs detected in metastases, the relative abundance in metastasis is lower than in primary tumor of the same mouse, with a metastasis-primary ratio (MPR) less than 1 (Figure 11 (S4) F), likely due to more skewed distributions of sgRNAs in the metastasis compared to those in the late primary tumors. A small subset of sgRNAs, however, are more abundant in metastasis than in primary tumor (MPR > 1) in multiple mice, such as sgRNAs targeting *Nf2*, *Trim72*, *Prostaglandin E Synthase 2 (Ptgse2)* or *Ubiquitin-conjugating enzyme E2G 2 (Ube2g2)* (Figure 4 F, Figure 11 (S4) F). The strongest hits (i.e. genes targeted by the most abundant sgRNAs, by multiple enriched sgRNAs, and/or by occurrence in multiple animals) in late primary tumor also partly overlap with those in the lung metastases (described later in text) (Compare Figures 3 and 4).

[00401] In four genes, *Nf2*, *Pten*, *Trim72*, and *Zfyve28*, two independent sgRNAs targeting different regions of the same gene were enriched in lung metastases (Figure 4 G). One of the *Zfyve28*-targeting sgRNAs, however, is enriched in only one animal; whereas *Nf2*, *Pten* or *Trim72* all have two sgRNAs enriched in multiple animals (Figure 4 H). These three genes, several representative genes with one sgRNA having high enrichment in more than one mice (*Cdkn2a*, *Fga*, *Cryba4*), and the top two scoring microRNAs (*miR-152* and *miR-345*) were chosen to assay individually for primary tumor growth and metastases formation.

Validation of genes using single sgRNAs

[00402] For these targets (*Nf2*, *Pten*, *Trim72*, *Cdkn2a*, *Fga*, *Cryba4*, *miR-152* and *miR-345*), Applicants cloned multiple sgRNAs targeting each of them into the same lentiviral vector (lentiGuid-Puro), and transduced them into the Cas9-GFP KPD cell line (Figure 5 A). As expected, these sgRNAs generated a broad distribution of NHEJ - mediated indels at the target site when examined three days post-transduction, with a bias toward deletions (Figure 5 B). For

protein coding genes, the majority (> 80%) of indels are out-of-frame, which potentially disrupt the protein functions. SgRNAs targeting either the loop or mature 5p- or 3p- microRNAs for *miR-152* and *miR-345* generated deletions (average size = -7 bp), overlapping with the loop or mature micro-RNA sequences in the hairpins, which are structures required for the maturation of microRNAs. For protein-coding genes where specific antibodies are available (*Nf2* and *Pten*), the majority of the protein products were significantly reduced one week after lentiviral sgRNA infection (Figure 13 (S6) A).

[00403] When these single-sgRNA-transduced cells were transplanted into the flanks of immunocompromised mice, they all formed tumors *in situ*. With two animals injected per sgRNA and three sgRNAs per gene, all genes tested showed increased lung metastases formation compared to controls (untransduced and non-targeting sgRNAs), with the most significant ones being *Nf2*, *Pten* and *Cdkn2a* (One tailed t test, $p < 0.05$) (Figure 5 C-D). *Trim72* and *Fga* also have effects but were marginally significant (One tailed t test, $p = 0.056$). *Cryba4* is not statistically different from controls ($p = 0.11$). Both of the microRNAs targeted significantly increased metastasis ($p < 0.05$). These data suggest that loss-of-function mutations in any of *Nf2*, *Pten*, *Cdkn2a*, *Trim72*, *Fga*, *miR345* or *miR-152* alone are sufficient to accelerate the rate of metastasis formation in this genetic background. Among these genes *Nf2*, has the strongest effect, with metastases detected in all mice and in over 80% of the lobes examined.

[00404] To examine if primary tumor shedding could explain which gene targets were more likely to form lung metastases, Applicants analyzed blood samples for the presence of circulating tumor cells (CTCs). A microfluidic device was generated based on the physical size of the Cas9-GFP KPD cells to capture CTCs (Figure 13 (S6) B-C). CTC capture was performed with terminal peripheral blood samples from animals injected with *sg-Nf2*, *sg-Pten*, *sg-Trim72*, *sg-Cdkn2a*, *sg-miR-152* and control cells (Figure 13 (S6) C). Animals transplanted with either *sg-Nf2*, *sg-Pten*, *sg-Trim72*, *sg-Cdkn2a* or *sg-miR-152* cells had a higher concentration of CTCs as compared to controls (Figure 13 (S6) D), consistent with the higher rate of lung metastasis formation. The absolute number of CTCs captured by this device, however, is only weakly correlated with either primary tumor size or metastasis formation rate in a gene-by-gene analysis (Figure 13 (S6) E-F), suggesting CTC count alone can not explain all of the variation in

metastasis under this setting, as metastasis is also influenced by multiple other factors, such as extravasation or clonal growth at destination site. This is not surprising since metastasis is also influenced by multiple other factors such as extravasation or clonal growth at destination site.

[00405] Most genes targeted with single sgRNAs also showed effects in accelerating primary tumor growth compared to controls (Figure 5 E). *Nf2* and *Pten* loss-of-function dramatically enhanced tumor growth (KS test, $p < 0.001$) (Figure 5 E); *Cdkn2a*-, *Trim72*- and *Fga*- targeting sgRNAs slightly accelerate primary tumor growth (KS test, $p = 0.003 \sim 0.01$); *Cryba4* has a marginal effect (KS test, $p = 0.08$); neither *miR-152*- nor *miR-345*- targeting sgRNAs promote primary tumor growth (KS test, $p > 0.1$). Overall, for the targets Applicants examined using individual sgRNAs, the rate of development of lung metastases strongly correlates with the terminal volume of the late primary tumor (or average primary tumor growth rate)(correlation, $\rho = 0.83$, F-test, $p < 0.01$) (Figure 5 F), indicating at a single gene level that mutant cells with a stronger ability to promote primary tumor growth generate metastases faster.

Competitive dynamics of top hits in a targeted sub-library

[00406] To better understand the relative metastatic potential of multiple genes from the genome-wide screen, Applicants designed a targeted pooled screen with a much smaller library. This small library (termed validation pool hereafter) contains 624 sgRNAs targeting 53 genes (10 sgRNAs per gene for most genes) plus 100 non-targeting sgRNAs. Applicants also created a size-matched library containing 624 non-targeting sgRNAs (termed control pool hereafter) (Figure 6 A). Lentiviruses from these two pools were used to transduce the Cas9-GFP KPD cells, which were then cultured *in vitro* for one week, and then transplanted into *Nu/Nu* mice (Figure 6 A). Both validation pool and control pool transduced cells induced primary tumor growth at a similar rate (Figure 6 B). However, the validation pool had a dramatically increased the rate of lung metastases formation (Figure 6 C).

[00407] Applicants sequenced the validation pool plasmid library, the transduced cells pre-transplantation, as well as the late stage primary tumors and whole lungs of the animals at 5 weeks post-transplantation. The sgRNA representations correlate strongly between technical replicates of the transduced cell pool, late primary tumor or and lung metastases (correlation, $\rho = 0.55$, F-test, $p < 0.01$, $n=5$)(Figure 14 (S7) B). The sgRNA distribution in the cell sample

strongly correlated with the plasmid (correlation, $\rho = 0.91$) (Figure 14 (S7) B,D). Almost all (99.4%) sgRNAs were recovered in the plasmid and the cell population (Figure 14 (S7) C). The sgRNA distributions in lung metastases clustered with each other and the primary tumors, away from plasmid or cell (Figure 14 (S7) D). The late primary tumors retained less than half of the sgRNAs, while the metastases in the whole lung only retained a small fraction (2~7%) of all sgRNAs (Figure 14 (S7) C). Similar to the case of full library, in the minipool plasmid and the cell samples showed tight distributions of sgRNAs, whereas the late primary tumors and lung metastases contained a largely bi-modal distribution, with the majority being zero and a small fraction of sgRNAs spanning a large range of non-zero read counts (Figure 6 D). Intriguingly, two animals retained relatively high sgRNA diversity in primary tumors (Figure 6 D), likely due to dormant or slowly proliferating cells that remained in low numbers during tumor growth. Similar to the genome-wide screen, there is a large shift in the distribution of sgRNA read counts in the tumor and lung metastases (Kolmogorov-Smirnov (KS) test, $p < 10^{-15}$ for pairwise comparisons between the cell, primary tumor and lung metastases, $p = 0.02$ between plasmid and cell) (Figure 6 E). As in the genome-wide library, the sgRNAs detected in the late primary tumors or the lungs of five different mice significantly overlap with each other (Figure 14 (S7) E-F).

[00408] The late stage primary tumors and lung metastases with the validation pool are dominated by a few sgRNAs (Figure 7 A, Figure 14 (S7) G-I), suggesting these sgRNAs outcompete the others relatively during tumor growth and metastasis. The gene with the strongest effect is again *Nf2*, with all ten sgRNAs significantly enriched in both late primary tumors and lung metastases in one or more mice, accounting for 71 - 98% of the total reads in each mouse (Figure 7 A, Figure 14 (S7) G-I), indicating that *Nf2* loss-of-function is the main driver of tumor growth and metastases in this experimental setting. In this small library, the sgRNA read counts are correlated between late stage primary tumors and lung metastases (correlation, $\rho = 0.6$ on average, F test, $p < 0.001$, $n = 5$) (Figure 7 B). The late stage primary tumors and lung metastases have dozens of sgRNAs targeting genes at moderate to high frequencies (Figure 7 B-C). Several genes have multiple independent sgRNAs that are enriched in the lung over the primary tumor (MPR > 1), such as *Nf2* (8 sgRNAs), *Pten* (4 sgRNAs), *Trim72* (3 sgRNAs), *Ube2g2* (3 sgRNAs), *Ptges2* (2 sgRNAs) and *ATP-dependent DNA Ligase*

IV (Lig4) (2 sgRNAs) (Figure 7 C). Two *Cdkn2a* sgRNAs were present in both late primary tumors and lung metastases in two mice, but with MPR < 1. *Fga*-, *Cryba4*-, *miR-152*- or *miR-345*- targeting sgRNAs were not found at high frequency in either late primary tumors or lung metastases, which suggests that they are outcompeted by other loss-of-function mutations (such as *Nf2*), also in agreement with the relatively reduced metastases formation in single sgRNA validation. These results further validated several of the top hits from the primary screen and, in contrast to screen validation using single sgRNAs, provides a higher-throughput method for validating the top candidates. This validation pool also furthers our understanding of the dynamics of multiple competing mutants within a heterogeneous tumor, and indicates that mutants with strong pro-growth effects tend to enhance metastases in this model.

Further validation with a tiny pool

[00409] Applicants designed a further targeted pooled screen with an even smaller library. This very small library (termed tiny pool hereafter) contains 23 sgRNAs targeting 6 genes (3 sgRNAs per gene) and 4 non-targeting sgRNAs (Ctrl) (Figure 16). The six selected genes were *Nf2*, *Pten*, *Trim72*, *Cdkn2a*, *miR-152*, and *miR-345*. Lentiviruses from these pools were used to transduce the Cas9-GFP KPD cells, which were then cultured *in vitro* for one week, and then transplanted into *Nu/Nu* mice. Both tiny pool and control pool transduced cells induced primary tumor growth at a similar rate, while non-transduced cells did not induce tumor growth (Figure 16B). This targeted subpool strategy provides a higher-throughput method for validating the top candidates and assaying competing mutants in parallel.

Generation of a xenograft model

[00410] The human cell line H23 was transduced with a lentivirus carrying a Cas9 transgene fused to a nuclear localization sequence (NLS) a triple FLAG epitope (3xFLAG), and a green fluorescent protein (GFP). The transduction was performed at a very low multiplicity of infection (MOI < 0.01) to ensure that most cells receive only one lentiviral transgene. Clonal cell lines expressing Cas9-GFP were established. A clonal Cas9-H23 cell line was selected for introduction of the genome wide knockout library.

A human genome-scale CRISPR knockout library hGeCKOa was made as described in Shalem et al. 2014 and used to transduce the human Cas9-H23 cell line as described above.

Human H23 cells infected with NLS-Cas9-GFP vector only or NLS-Cas9-GFP and hGeCKO library were injected subcutaneously into the right side flank of Nu/Nu mice at 3×10^7 cells per mouse.

The in vitro culture sgRNA and mouse xenograft sgRNA representation are provided in Figure 17.

References

- [00411] Aceto, N., Bardia, A., Miyamoto, D.T., Donaldson, M.C., Wittner, B.S., Spencer, J.A., Yu, M., Pely, A., Engstrom, A., Zhu, H., et al. (2014). Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* 158, 1110-1122.
- [00412] Bell, C.L., Vandenberghe, L.H., Bell, P., Limberis, M.P., Gao, G.P., Van Vliet, K., Agbandje-McKenna, M., and Wilson, J.M. (2011). The AAV9 receptor and its modification to improve in vivo lung gene transfer in mice. *The Journal of clinical investigation* 121, 2427-2435.
- Chen, S., Xue, Y., Wu, X., Le, C., Bhutkar, A., Bell, E.L., Zhang, F., Langer, R., and Sharp, P.A. (2014). Global microRNA depletion suppresses tumor angiogenesis. *Genes & development* 28, 1054-1067.
- [00413] Bibikova, M., Beumer, K., Trautman, J.K., and Carroll, D. (2003). Enhancing gene targeting with designed zinc finger nucleases. *Science* 300, 764.
- [00414] Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery* 2, 401-404.
- [00415] Chen, S., Xue, Y., Wu, X., Le, C., Bhutkar, A., Bell, E.L., Zhang, F., Langer, R., and Sharp, P.A. (2014). Global microRNA depletion suppresses tumor angiogenesis. *Genes & development* 28, 1054-1067.

- [00416] Cheng, Z., Ma, R., Tan, W., and Zhang, L. (2014). MiR-152 suppresses the proliferation and invasion of NSCLC cells by inhibiting FGF2. *Experimental & molecular medicine* 46, e112.
- [00417] Chung, J., Issadore, D., Ullal, A., Lee, K., Weissleder, R., and Lee, H. (2013). Rare cell isolation and profiling on a hybrid magnetic/size-sorting chip. *Biomicrofluidics* 7, 54107.
- [00418] Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823.
- [00419] Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602-607.
- [00420] Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D., and Naldini, L. (1998). A third-generation lentivirus vector with a conditional packaging system. *Journal of virology* 72, 8463-8471.
- [00421] DuPage, M., Dooley, A.L., and Jacks, T. (2009). Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nature protocols* 4, 1064-1072.
- [00422] Francia, G., Cruz-Munoz, W., Man, S., Xu, P., and Kerbel, R.S. (2011). Mouse models of advanced spontaneous metastasis for experimental therapeutics. *Nature reviews Cancer* 11, 135-141.
- [00423] Frese, K.K., and Tuveson, D.A. (2007). Maximizing mouse cancer models. *Nature reviews Cancer* 7, 645-658.
- [00424] Garraway, L.A., and Lander, E.S. (2013). Lessons from the cancer genome. *Cell* 153, 17-37.
- [00425] Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria.

Proceedings of the National Academy of Sciences of the United States of America 109, E2579-2586.

[00426] Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling* 6, p11.

[00427] Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*.

[00428] Govindan, R., Ding, L., Griffith, M., Subramanian, J., Dees, N.D., Kanchi, K.L., Maher, C.A., Fulton, R., Fulton, L., Wallis, J., et al. (2012). Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* 150, 1121-1134.

[00429] Halbert, C.L., Allen, J.M., and Miller, A.D. (2002). Efficient mouse airway transduction following recombination between AAV vectors carrying parts of a larger gene. *Nature biotechnology* 20, 697-701.

[00430] Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646-674.

[00431] Hegi, M.E., Diserens, A.C., Gorlia, T., Hamou, M.F., de Tribolet, N., Weller, M., Kros, J.M., Hainfellner, J.A., Mason, W., Mariani, L., et al. (2005). MGMT gene silencing and benefit from temozolomide in glioblastoma. *The New England journal of medicine* 352, 997-1003.

[00432] Heimann, R., and Hellman, S. (1998). Aging, progression, and phenotype in breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 16, 2686-2692.

[00433] Herbig, E., Warfield, L., Fish, L., Fishburn, J., Knutson, B.A., Moorefield, B., Pacheco, D., and Hahn, S. (2010). Mechanism of Mediator recruitment by tandem Gcn4 activation domains and three Gal11 activator-binding domains. *Molecular and cellular biology* 30, 2376-2390.

- [00434] Herbst, R.S., Heymach, J.V., and Lippman, S.M. (2008). Lung cancer. *The New England journal of medicine* 359, 1367-1380.
- [00435] Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology* 31, 827-832.
- [00436] Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262-1278.
- [00437] Huang, S., Holzel, M., Knijnenburg, T., Schlicker, A., Roepman, P., McDermott, U., Garnett, M., Grenrum, W., Sun, C., Prahallad, A., et al. (2012). MED12 controls the response to multiple cancer drugs through regulation of TGF-beta receptor signaling. *Cell* 151, 937-950.
- [00438] Ioannidis John, P.A., Castaldi P., Evangelou E. A compendium of genome-wide associations for cancer: critical synopsis and reappraisal. *J. Natl Cancer Inst* 2010; 102: 846-858
- [00439] Iwasaki, M., Homma, S., Hishiya, A., Dolezal, S.J., Reed, J.C., and Takayama, S. (2007). BAG3 regulates motility and adhesion of epithelial cancer cells. *Cancer research* 67, 10252-10259.
- Jasin, M., de Villiers, J., Weber, F., and Schaffner, W. (1985). High frequency of homologous recombination in mammalian cells between endogenous and introduced SV40 genomes. *Cell* 43, 695-703.
- [00440] Jackson, E.L., Willis, N., Mercer, K., Bronson, R.T., Crowley, D., Montoya, R., Jacks, T., and Tuveson, D.A. (2001). Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes & development* 15, 3243-3248.
- [00441] Jesien-Lewandowicz, E., Jesionek-Kupnicka, D., Zawlik, I., Szybka, M., Kulczycka-Wojdala, D., Rieske, P., Sieruta, M., Jaskolski, D., Och, W., Skowronski, W., et al. (2009). High incidence of MGMT promoter methylation in primary glioblastomas without correlation with TP53 gene mutations. *Cancer genetics and cytogenetics* 188, 77-82.

- [00442] Ji, H., Ramsey, M.R., Hayes, D.N., Fan, C., McNamara, K., Kozlowski, P., Torrice, C., Wu, M.C., Shimamura, T., Perera, S.A., et al. (2007). LKB1 modulates lung cancer differentiation and metastasis. *Nature* 448, 807-810.
- [00443] Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821.
- [00444] Johnson, L., Mercer, K., Greenbaum, D., Bronson, R.T., Crowley, D., Tuveson, D.A., and Jacks, T. (2001). Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature* 410, 1111-1116.
- [00445] Kaczmarczyk, G., Lewandowski, R., Trautsolt, W., Ziolkowski, A., and Kozielski, J. (2012). Cytological examination of pleural cavity lavage accompanied by the study of gene promoter hypermethylation of p16 and O6-methylguanine-DNA-methyltransferase genes in diagnostics of non-small cell lung cancer metastatic changes into pleura. *Contemporary oncology* 16, 322-327.
- [00446] Kaina, B., Christmann, M., Naumann, S., and Roos, W.P. (2007). MGMT: key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents. *DNA repair* 6, 1079-1099.
- [00447] Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., et al. (2013). Mutational landscape and significance across 12 major cancer types. *Nature* 502, 333-339.
- [00448] Kang, Y., Siegel, P.M., Shu, W., Drobnjak, M., Kakonen, S.M., Cordon-Cardo, C., Guise, T.A., and Massague, J. (2003). A multigenic program mediating breast cancer metastasis to bone. *Cancer cell* 3, 537-549.
- [00449] Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome research* 22, 568-576.

- [00450] Koike-Yusa, H., Li, Y., Tan, E.P., Velasco-Herrera Mdel, C., and Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nature biotechnology* 32, 267-273.
- [00451] Kumar, M.S., Pester, R.E., Chen, C.Y., Lane, K., Chin, C., Lu, J., Kirsch, D.G., Golub, T.R., and Jacks, T. (2009). *Dicer1* functions as a haploinsufficient tumor suppressor. *Genes & development* 23, 2700-2704.
- [00452] Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* 10, R25.
- [00453] Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214-218.
- [00454] Lawrence, M.S. Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R, Meyerson, M., Gabriel, S.B., and Lander E.S. Discovery and stuation analysis of cancer genes across 21 tumour types. *Nature* 23 January 2014; vol 505, 495 – 355 and supplementary tables 1 – 6.
- [00455] Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589-595.
- [00456] Limberis, M.P., and Wilson, J.M. (2006). Adeno-associated virus serotype 9 vectors transduce murine alveolar and nasal epithelia and can be readministered. *Proceedings of the National Academy of Sciences of the United States of America* 103, 12993-12998.
- [00457] Mali, P., Yang, L.H., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-Guided Human Genome Engineering via Cas9. *Science* 339, 823-826.
- [00458] Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet* 17, 1.

- [00459] McClatchey, A.I., Saotome, I., Mercer, K., Crowley, D., Gusella, J.F., Bronson, R.T., and Jacks, T. (1998). Mice heterozygous for a mutation at the Nf2 tumor suppressor locus develop a range of highly metastatic tumors. *Genes & development* 12, 1121-1133.
- [00460] McFadden, D.G., Papagiannakopoulos, T., Taylor-Weiner, A., Stewart, C., Carter, S.L., Cibulskis, K., Bhutkar, A., McKenna, A., Dooley, A., Vernon, A., et al. (2014). Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. *Cell* 156, 1298-1311.
- [00461] Mitzner, W., Brown, R., and Lee, W. (2001). In vivo measurement of lung volumes in mice. *Physiological genomics* 4, 215-221.
- [00462] Molenaar, R.J., Verbaan, D., Lamba, S., Zanon, C., Jeuken, J.W., Boots-Sprenger, S.H., Wesseling, P., Hulsebos, T.J., Troost, D., van Tilborg, A.A., et al. (2014). The combination of IDH1 mutations and MGMT methylation status predicts survival in glioblastoma better than either IDH1 or MGMT alone. *Neuro-oncology* 16, 1263-1273.
- [00463] Naba, A., Clauser, K.R., Lamar, J.M., Carr, S.A., and Hynes, R.O. (2014a). Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters. *eLife* 3, e01308.
- [00464] Naba, A., Clauser, K.R., Whittaker, C.A., Carr, S.A., Tanabe, K.K., and Hynes, R.O. (2014b). Extracellular matrix signatures of human primary metastatic colon cancers and their metastases to liver. *BMC cancer* 14, 518.
- [00465] Nguyen, D.X., Chiang, A.C., Zhang, X.H., Kim, J.Y., Kris, M.G., Ladanyi, M., Gerald, W.L., and Massague, J. (2009). WNT/TCF signaling through LEF1 and HOXB9 mediates lung adenocarcinoma metastasis. *Cell* 138, 51-62.
- [00466] Nissen, L.J., Cao, R., Hedlund, E.M., Wang, Z., Zhao, X., Wetterskog, D., Funa, K., Brakenhielm, E., and Cao, Y. (2007). Angiogenic factors FGF2 and PDGF-BB synergistically promote murine tumor neovascularization and metastasis. *The Journal of clinical investigation* 117, 2766-2777.
- [00467] Park, W.Y., Kim, M.H., Shin, D.H., Lee, J.H., Choi, K.U., Kim, J.Y., Park do, Y., Lee, C.H., and Sol, M.Y. (2012). Ciliated adenocarcinomas of the lung: a tumor of non-terminal

respiratory unit origin. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 25, 1265-1274.

[00468] Platt, R.J., Chen, S., Zhou, Y., Yim, M.J., Swiech, L., Kempton, H.R., Dahlman, J.E., Parnas, O., Eisenhaure, T.M., Jovanovic, M., et al. (2014). CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. *Cell* 159, 440-455.

[00469] Pylayeva-Gupta, Y., Grabocka, E., and Bar-Sagi, D. (2011). RAS oncogenes: weaving a tumorigenic web. *Nature reviews Cancer* 11, 761-774.

[00470] Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nature Methods* 11, 783-784.

[00471] Schiano, C., Casamassimi, A., Rienzo, M., de Nigris, F., Sommese, L., and Napoli, C. (2014). Involvement of Mediator complex in malignancy. *Biochimica et biophysica acta* 1845, 66-83.

[00472] Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature methods* 9, 671-675.

[00473] Shackelford, D.B., and Shaw, R.J. (2009). The LKB1-AMPK pathway: metabolism and growth control in tumour suppression. *Nature reviews Cancer* 9, 563-575.

[00474] Schramek, D., Sendoel, A., Segal, J.P., Beronja, S., Heller, E., Oristian, D., Reva, B., and Fuchs, E. (2014). Direct in vivo RNAi screen unveils myosin IIa as a tumor suppressor of squamous cell carcinomas. *Science* 343, 309-313.

[00475] Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343, 84-87.

[00476] Shao, D.D., Xue, W., Krall, E.B., Bhutkar, A., Piccioni, F., Wang, X., Schinzel, A.C., Sood, S., Rosenbluh, J., Kim, J.W., et al. (2014). KRAS and YAP1 converge to regulate EMT and tumor survival. *Cell* 158, 171-184.

- [00477] Tang, J.T., Wang, J.L., Du, W., Hong, J., Zhao, S.L., Wang, Y.C., Xiong, H., Chen, H.M., and Fang, J.Y. (2011). MicroRNA 345, a methylation-sensitive microRNA is involved in cell proliferation and invasion in human colorectal cancer. *Carcinogenesis* 32, 1207-1215.
- [00478] Tano, K., Shiota, S., Collier, J., Foote, R.S., and Mitra, S. (1990). Isolation and structural characterization of a cDNA clone encoding the human DNA repair protein for O6-alkylguanine. *Proceedings of the National Academy of Sciences of the United States of America* 87, 686-690.
- [00479] TCGA-Network (2012). Comprehensive genomic characterization of squamous cell lung cancers. *Nature* 489(7417):519-25.
- [00480] TCGA-Network (2014a). Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature* 507, 315-322.
- [00481] TCGA-Network (2014b). Comprehensive molecular profiling of lung adenocarcinoma. *Nature* Published online 09 July 2014.
- [00482] Teo, A.K., Oh, H.K., Ali, R.B., and Li, B.F. (2001). The modified human DNA repair enzyme O(6)-methylguanine-DNA methyltransferase is a negative regulator of estrogen receptor-mediated transcription upon alkylation DNA damage. *Molecular and cellular biology* 21, 7105-7114.
- [00483] Valastyan, S., and Weinberg, R.A. (2011). Tumor metastasis: molecular insights and evolving paradigms. *Cell* 147, 275-292.
- [00484] Valiente, M., Obenaus, A.C., Jin, X., Chen, Q., Zhang, X.H., Lee, D.J., Chaff, J.E., Kris, M.G., Huse, J.T., Brogi, E., et al. (2014). Serpins promote cancer cell survival and vascular co-option in brain metastasis. *Cell* 156, 1002-1016.
- [00485] Vanharanta, S., and Massague, J. (2013). Origins of metastatic traits. *Cancer cell* 24, 410-421.
- [00486] Waghorne, C., Thomas, M., Lagarde, A., Kerbel, R.S., and Breitman, M.L. (1988). Genetic evidence for progressive selection and overgrowth of primary tumors by metastatic cell subpopulations. *Cancer research* 48, 6109-6114.

- [00487] Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science* 343, 80-84.
- [00488] Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153, 910-918.
- [00489] Weinberg, R.A. (2007). *The biology of cancer* (New York, Garland Science).
- [00490] Whittaker, S.R., Theurillat, J.P., Van Allen, E., Wagle, N., Hsiao, J., Cowley, G.S., Schadendorf, D., Root, D.E., and Garraway, L.A. (2013). A genome-scale RNA interference screen implicates NF1 loss in resistance to RAF inhibition. *Cancer discovery* 3, 350-362.
- [00491] Winslow, M.M., Dayton, T.L., Verhaak, R.G., Kim-Kiselak, C., Snyder, E.L., Feldser, D.M., Hubbard, D.D., DuPage, M.J., Whittaker, C.A., Hoersch, S., et al. (2011). Suppression of lung adenocarcinoma progression by Nkx2-1. *Nature* 473, 101-104.
- [00492] Xue, W., Chen, S., Yin, H., Tammela, T., Papagiannakopoulos, T., Joshi, N.S., Cai, W., Yang, G., Bronson, R., Crowley, D.G., et al. (2014). CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature*.
- [00493] Yokota, J., Nishioka, M., Tani, M., and Kohno, T. (2003). Genetic alterations responsible for metastatic phenotypes of lung cancer cells. *Clinical & experimental metastasis* 20, 189-193.
- [00494] Zender, L., Xue, W., Zuber, J., Semighini, C.P., Krasnitz, A., Ma, B., Zender, P., Kubicka, S., Luk, J.M., Schirmacher, P., et al. (2008). An oncogenomics-based in vivo RNAi screen identifies tumor suppressors in liver cancer. *Cell* 135, 852-864.

What is claimed is:

1. A method for modeling tumor formation and/or tumor evolution comprising the steps of:
 - (a) providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas;
 - (b) introducing in said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; or
introducing in said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote;
 - (c) introducing said Cas transgenic eukaryotic cell after step (b) in a non-human eukaryote;
 - (d) analyzing tumor formation and/or tumor evolution in said non-human eukaryote.
2. The method according to claim 1, wherein step (c) comprises introducing a plurality of Cas transgenic eukaryotic cells in a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus of said eukaryote; or
each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus of said eukaryote.
3. The method according to any of claims 1 to 2, wherein step (d) comprises identifying one or more genes which are involved in tumor formation and/or evolution.
4. The method according to any of claims 1 to 3, wherein step (d) comprises analyzing tumor metastasis.
5. The method according to any of claims 1 to 4, wherein step (d) comprises analyzing gene expression.
6. The method according to any of claims 1 to 5, wherein step (d) comprises analyzing gene expression in said tumor.

7. The method according to any of claims 1 to 5, wherein step (d) comprises identifying one or more of said RNA of step (b) in said tumor.
8. The method according to any of claims 4 to 5, wherein step (d) comprises analyzing gene expression in said tumor metastasis.
9. The method according to any of claims 4 to 5, wherein step (d) comprises identifying one or more of said RNA of step (b) in said tumor metastasis.
10. A method for identifying genes which are involved in tumor formation and/or tumor metastasis, comprising the steps of:
 - (a) providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas;
 - (b) introducing in said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; or
introducing in said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote;
 - (c) introducing a plurality of said Cas transgenic eukaryotic cell after step (b) in a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus;
 - (d) identifying genes which are involved in tumor formation and/or tumor metastasis in said non-human eukaryote based on the identification in said tumor or tumor metastasis of one or more of said RNA capable of guiding Cas to one or more genetic target locus.
11. The method according to any of claims 3 to 10, wherein said genes involved in tumor formation, tumor evolution, and/or tumor metastasis are tumor suppressor genes and/or metastasis suppressor genes.
12. The method according to any of claims 1 to 11, wherein said method after step (b) comprises introducing one or more genomic mutations in said Cas transgenic eukaryotic cell.

13. The method according to any of claims 1 to 12, wherein said RNA capable of guiding Cas comprises CRISPR RNA and transactivating (tracr) RNA.
14. The method according to any of claims 1 to 13, wherein said RNA capable of guiding Cas comprises a single guide (sg) RNA.
15. The method according to any of claims 1 to 14, wherein said RNA capable of guiding Cas is introduced in said Cas transgenic eukaryotic cell by means of transduction.
16. The method according to any of claims 1 to 15, wherein said RNA capable of guiding Cas is introduced in said Cas transgenic eukaryotic cell by means of lentiviral transduction.
17. The method according to any of claims 1 to 16, wherein said polynucleotide encoding one or more RNA capable of guiding Cas is capable of constitutively expressing said RNA or alternatively inducibly and/or conditionally expressing said RNA.
18. The method according to any of claims 1 to 17, wherein said Cas transgenic eukaryotic cell of step (a) is a tumor cell.
19. The method according to any of claims 1 to 18, wherein said Cas transgenic eukaryotic cell of step (a) is a non-metastasizing tumor cell.
20. The method according to any of claims 1 to 18, wherein said Cas transgenic eukaryotic cell of step (a) is characterized by oncogenic expression of *Kras* (*KrasG12D*), homozygous *p53* loss (*p53* $-/-$), and/or heterozygous *Dicer1* loss (*Dicer* $+/-$), preferably all.
21. The method according to any of claims 1 to 20, wherein said Cas comprises a type II Cas.
22. The method according to any of claims 1 to 21, wherein said Cas comprises Cas9.
23. The method according to any of claims 1 to 22, wherein said Cas comprises a Cas originating from *Streptococcus pyogenes*, *Streptococcus thermophiles*, or *Staphylococcus aureus*.
24. The method according to any of claims 1 to 23, wherein said Cas comprises a mutated Cas having an altered catalytic activity.
25. The method according to any of claims 1 to 24, wherein said Cas comprises a catalytically inactive Cas.

26. The method according to any of claims 1 to 25, wherein said Cas is fused to an enzyme which modifies genomic DNA or genomic DNA architecture.
27. The method according to any of claims 1 to 26, wherein said Cas is fused to a polypeptide which alters gene transcription.
28. The method according to any of claims 1 to 27, wherein said Cas transgenic eukaryotic cell is a cell from an animal.
29. The method according to any of claims 1 to 28, wherein said Cas transgenic eukaryotic cell is a cell from a eukaryote selected from the group consisting of mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod.
30. The method according to any of claims 1 to 29, wherein said Cas transgenic eukaryotic cell is a mammalian cell, preferably a mouse cell.
31. The method according to any of claims 1 to 30, wherein said eukaryote in step (c) is an animal, preferably a mammal.
32. The method according to any of claims 1 to 31, wherein said eukaryote in step (c) is an immunocompromised animal, preferably an immunocompromised mammal.
33. The method according to any of claims 1 to 32, wherein said eukaryote in step (c) is selected from mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod.
34. The method according to any of claims 1 to 33, wherein said eukaryote in step (c) is a mouse.
35. The method according to any of claims 1 to 34, wherein step (c) comprises introducing said cells subcutaneously.
36. The method according to any of claims 1 to 35, wherein said tumor is selected from lung cancer, lung adenocarcinoma, lung squamous cell carcinoma, acute myeloid leukemia, basal cell (skin) carcinoma, bladder cancer, breast cancer, carcinoid cancer, chronic lymphocytic leukemia, colorectal cancer, lymphoma, diffuse large B-cell lymphoma, endometrial cancer, esophageal

cancer, esophageal adenocarcinoma, glioblastoma multiforme, glioma, head and neck cancer, kidney cell cancer, medulloblastoma, melanoma, multiple myeloma, nasopharyngeal cancer, neuroblastoma, ovarian cancer; prostate cancer, rhabdoid tumor, thyroid cancer, and urinary bladder cancer.

37. A tumor sample or a tumor metastasis sample isolated from said non-human eukaryote after introduction of said Cas transgenic eukaryotic cell of step (c) in claim 1.

38. The tumor sample or the tumor metastasis sample according to claim 37, wherein said tumor is selected from the group consisting of lung cancer, lung adenocarcinoma, lung squamous cell carcinoma, acute myeloid leukemia, basal cell (skin) carcinoma, bladder cancer, breast cancer, carcinoid cancer, chronic lymphocytic leukemia, colorectal cancer, lymphoma, diffuse large B-cell lymphoma, endometrial cancer, esophageal cancer, esophageal adenocarcinoma, glioblastoma multiforme, glioma, head and neck cancer, kidney cell cancer, medulloblastoma, melanoma, multiple myeloma, nasopharyngeal cancer, neuroblastoma, ovarian cancer; prostate cancer, rhabdoid tumor, thyroid cancer, and urinary bladder cancer.

39. A cell isolated from said tumor sample or said tumor metastasis sample according to any of claims 37 to 38.

40. A cell line obtained or obtainable from said tumor sample or said tumor metastasis sample according to any of claims 37 to 38.

41. A cell line obtained or obtainable from said cell of claim 39.

42. A method for generating a non-human eukaryote model for tumor formation and/or tumor evolution, comprising the steps of:

(a) providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas;

(b) introducing in said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; or

introducing in said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote;

(c) introducing a plurality of said Cas transgenic eukaryotic cell after step (b) in a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus;

43. A method for generating a non-human eukaryote model for identifying genes involved in formation and/or tumor evolution, comprising the steps of:

(a) providing an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas;

(b) introducing in said Cas transgenic eukaryotic cell one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote; or

introducing in said Cas transgenic eukaryotic cell one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus of a eukaryote;

(c) introducing a plurality of said Cas transgenic eukaryotic cell after step (b) in a non-human eukaryote, each cell of said plurality of Cas transgenic eukaryotic cells comprising a different RNA capable of guiding Cas to a genetic target locus, or each cell of said plurality of Cas transgenic eukaryotic cells comprising a different polynucleotide encoding an RNA capable of guiding Cas to a genetic target locus;

44. A non-human eukaryote obtained or obtainable by the method of claim 42 or 43.

45. The non-human eukaryote according to claim 44, wherein said eukaryote is an animal, preferably a mammal.

46. The non-human eukaryote according to any of claims 44 to 45, wherein said eukaryote is an immunocompromised animal, preferably an immunocompromised mammal.

47. The non-human eukaryote according to any of claims 44 to 46, wherein said eukaryote is selected from the group consisting of mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod.

48. The non-human eukaryote according to any of claims 44 to 47, wherein said eukaryote is a mouse.

49. A method for determining whether a compound is capable of suppressing tumor formation and/or tumor evolution comprising the steps of:

- (a) providing an animal model according to the method of claim 42,
- (b) administering said compound to said animal model
- (c) determining the whether or not said compound is capable of suppressing tumor formation and/or tumor evolution in said animal model.

50. A kit comprising:

- an isolated Cas transgenic eukaryotic cell expressing Cas or capable of inducibly and/or conditionally expressing Cas;
- one or more RNA capable of guiding Cas to one or more genetic target locus, or one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus

51. The kit according to claim 50, further comprising a non-human eukaryote.

52. The kit according to any of claims 50 to 51, further comprising instructions for performing the method according to any of claims 1 to 36.

53. The kit according to any of claims 50 to 52, wherein said one or more RNA capable of guiding Cas to one or more genetic target locus, or one or more polynucleotide encoding one or more RNA capable of guiding Cas to one or more genetic target locus are comprised in a vector, preferably a lentiviral vector.

54. The kit according to any of claims 50 to 53, comprising a plurality of RNAs capable of guiding Cas to a genetic target locus; or a plurality of polynucleotides encoding an RNA capable of guiding Cas to a genetic target locus.

55. The kit according to any of claims 50 to 54, wherein said RNA capable of guiding Cas comprises CRISPR RNA and transactivating (tracr) RNA.

56. The kit according to any of claims 50 to 55, wherein said RNA capable of guiding Cas comprises a single guide (sg) RNA.
57. The kit according to any of claims 50 to 56, wherein said polynucleotide encoding one or more RNA capable of guiding Cas is capable of constitutively expressing said RNA or alternatively inducibly and/or conditionally expressing said RNA.
58. The kit according to any of claims 50 to 57, wherein said Cas transgenic eukaryotic cell is a tumor cell.
59. The kit according to any of claims 50 to 58, wherein said Cas transgenic eukaryotic cell is a non-metastasizing tumor cell.
60. The kit according to any of claims 50 to 59, wherein said Cas is a type II Cas.
61. The kit according to any of claims 50 to 60, wherein said Cas is Cas9.
62. The kit according to any of claims 50 to 61, wherein said Cas is a Cas originating from *Streptococcus pyogenes*, *Streptococcus thermophiles*, or *Staphylococcus aureus*.
63. The kit according to any of claims 50 to 62, wherein said Cas is a mutated Cas having an altered catalytic activity.
64. The kit according to any of claims 50 to 63, wherein said Cas is a catalytically inactive Cas.
65. The kit according to any of claims 50 to 64, wherein said Cas is fused to an enzyme which modifies genomic DNA or genomic DNA architecture.
66. The kit according to any of claims 50 to 65, wherein said Cas is fused to a polypeptide which alters gene transcription.
67. The kit according to any of claims 50 to 66, wherein said Cas transgenic eukaryotic cell is a cell from an animal.
68. The kit according to any of claims 50 to 67, wherein said Cas transgenic eukaryotic cell is a cell from a eukaryote selected from the group consisting of mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod.

69. The kit according to any of claims 50 to 68, wherein said Cas transgenic eukaryotic cell is a mammalian cell, preferably a mouse cell.
70. The kit according to any of claims 51 to 69, wherein said eukaryote is an animal, preferably a mammal.
71. The kit according to any of claims 51 to 70, wherein said eukaryote is an immunocompromised animal, preferably an immunocompromised mammal.
72. The kit according to any of claims 51 to 71, wherein said eukaryote is selected from the group consisting of mammal, primate, rodent, mouse, rat, rabbit, canine, dog, cow, bovine, sheep, ovine, goat, pig, fowl, poultry, chicken, fish, insect, and arthropod.
73. The kit according to any of claims 51 to 72, wherein said eukaryote is a mouse.
74. A method for diagnosing tumorigenesis and/or tumor metastasis, comprising the step of genetically analyzing or analyzing expression of one or more genes identified according to the method according to any of claims 10 to 36.
75. A method for prognosing tumorigenesis and/or tumor metastasis, comprising the step of genetically analyzing or analyzing expression of one or more genes identified according to the method according to any of claims 10 to 36.
76. A method for determining the likelihood of developing tumor metastasis, comprising the step of genetically analyzing or analyzing expression of one or more genes identified according to the method according to any of claims 10 to 36.
77. The method according to any of claims 74 to 76, wherein analyzing expression of said one or more genes comprises comparing expression level with a predetermined expression level.
78. The method according to any of claims 74 to 77, wherein said one or more gene is selected from the group consisting of *Nf2*, *Pten*, *Trim72*, *Cdkn2a*, *miR-152*, and *miR-345*.
79. The method according to claim 78, wherein said tumor metastasis is lung metastasis.
80. The method according to any of claims 74 to 78, wherein said tumor or tumor metastasis is selected from the group consisting of lung cancer, lung adenocarcinoma, lung squamous cell carcinoma, acute myeloid leukemia, basal cell (skin) carcinoma, bladder cancer, breast cancer,

carcinoid cancer, chronic lymphocytic leukemia, colorectal cancer, lymphoma, diffuse large B-cell lymphoma, endometrial cancer, esophageal cancer, esophageal adenocarcinoma, glioblastoma multiforme, glioma, head and neck cancer, kidney cell cancer, medulloblastoma, melanoma, multiple myeloma, nasopharyngeal cancer, neuroblastoma, ovarian cancer; prostate cancer, rhabdoid tumor, thyroid cancer, and urinary bladder cancer.

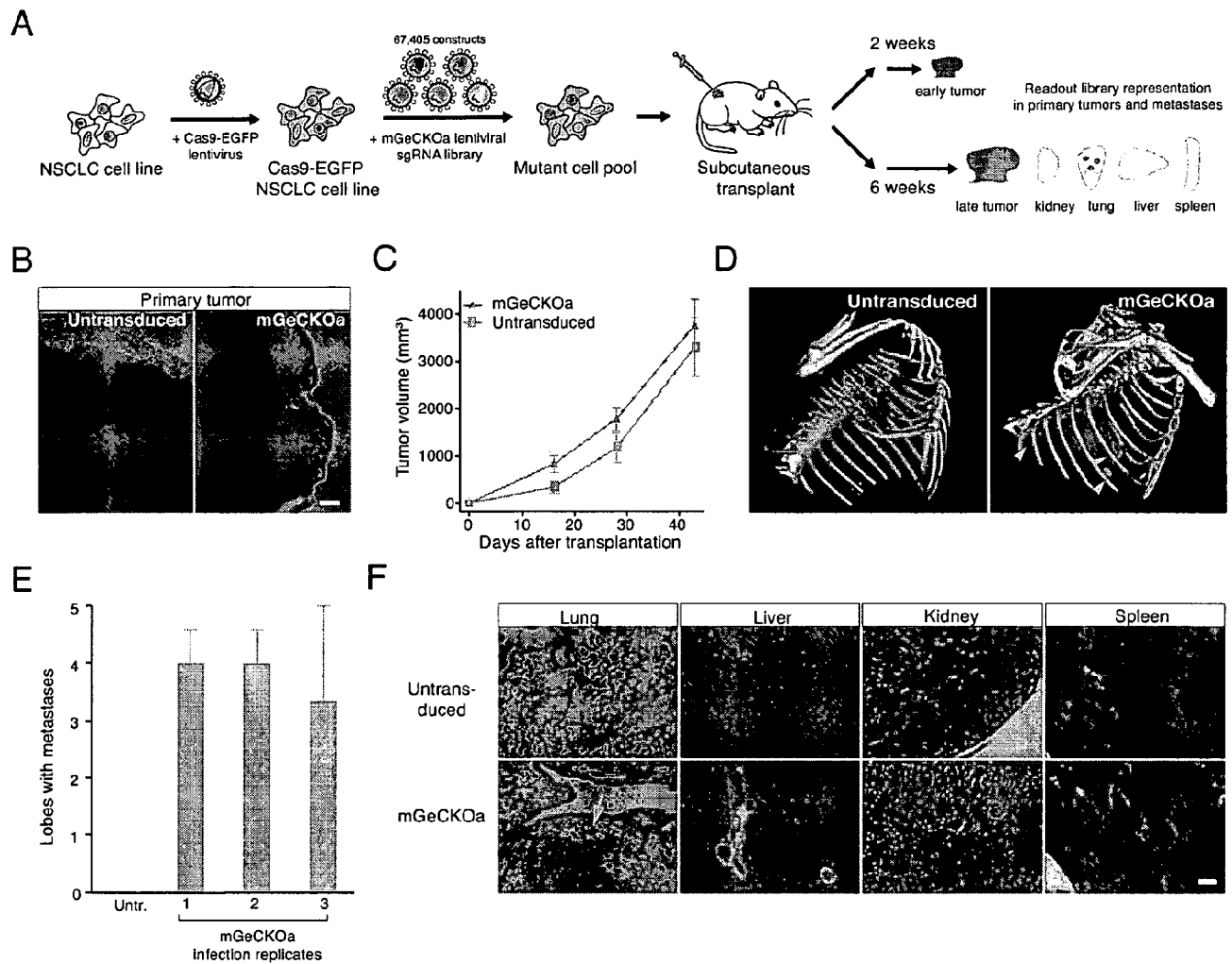


Fig. 1A-1F

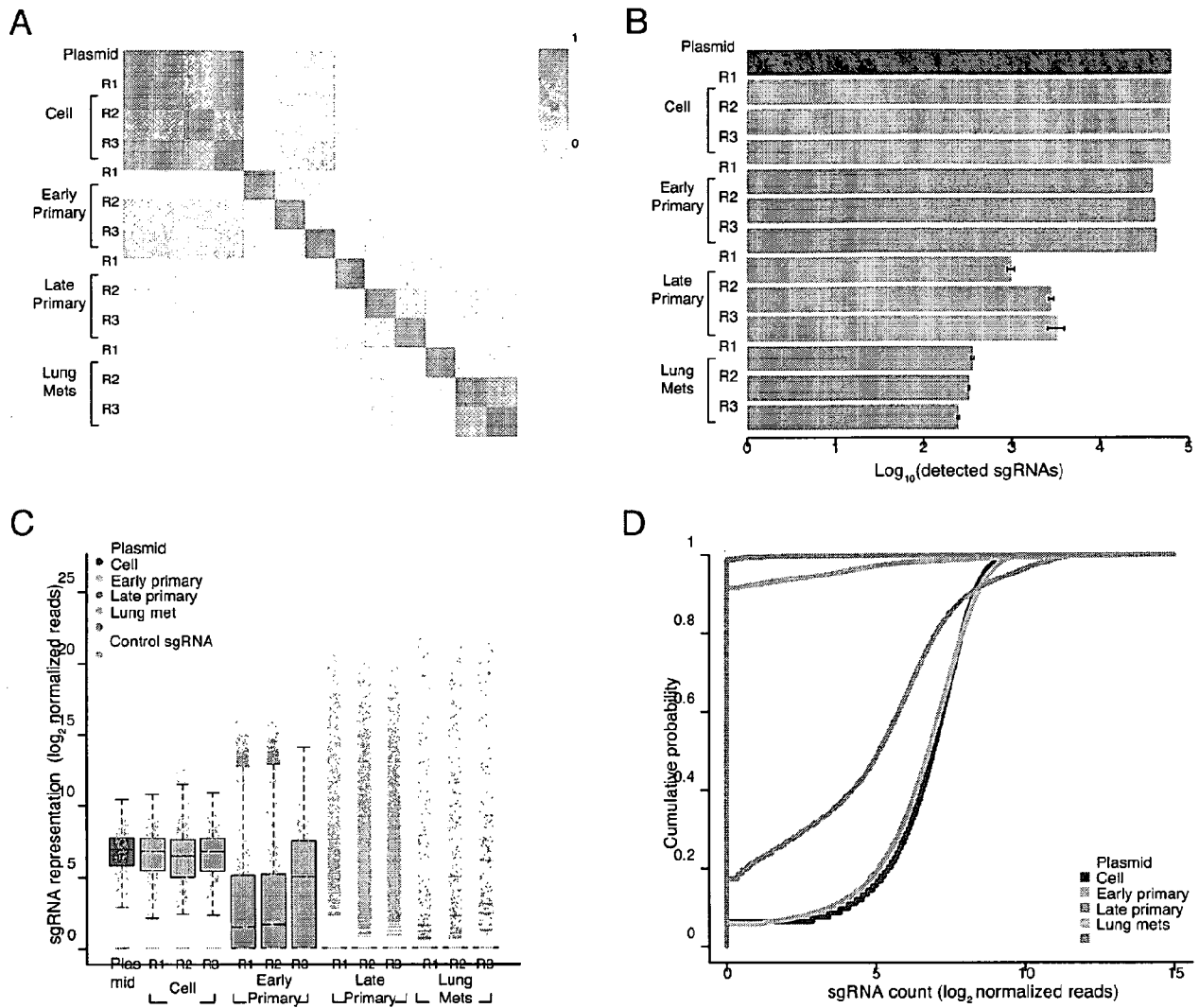


Fig. 2A-2D

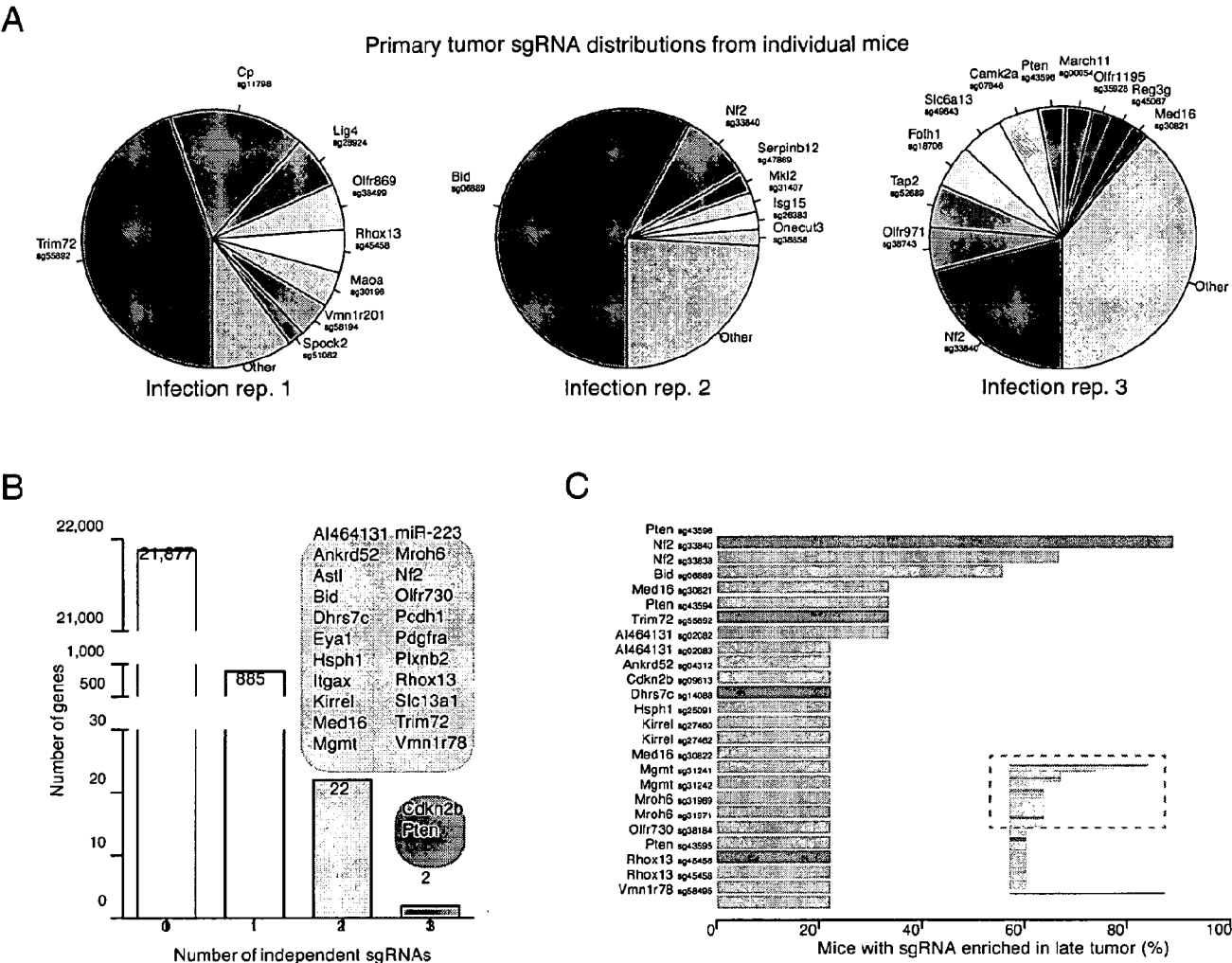


Fig. 3A-3C

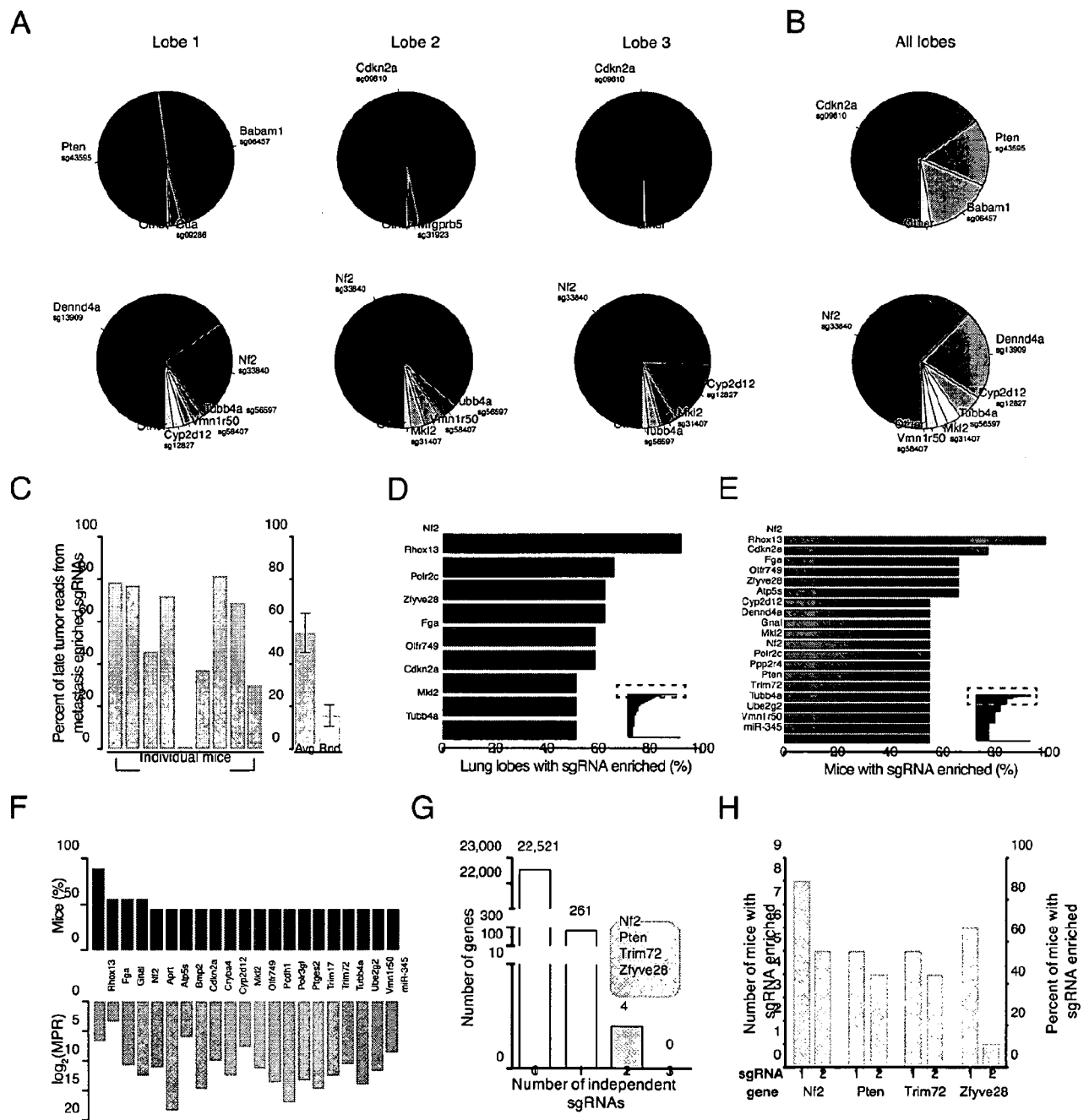


Fig. 4A-4H

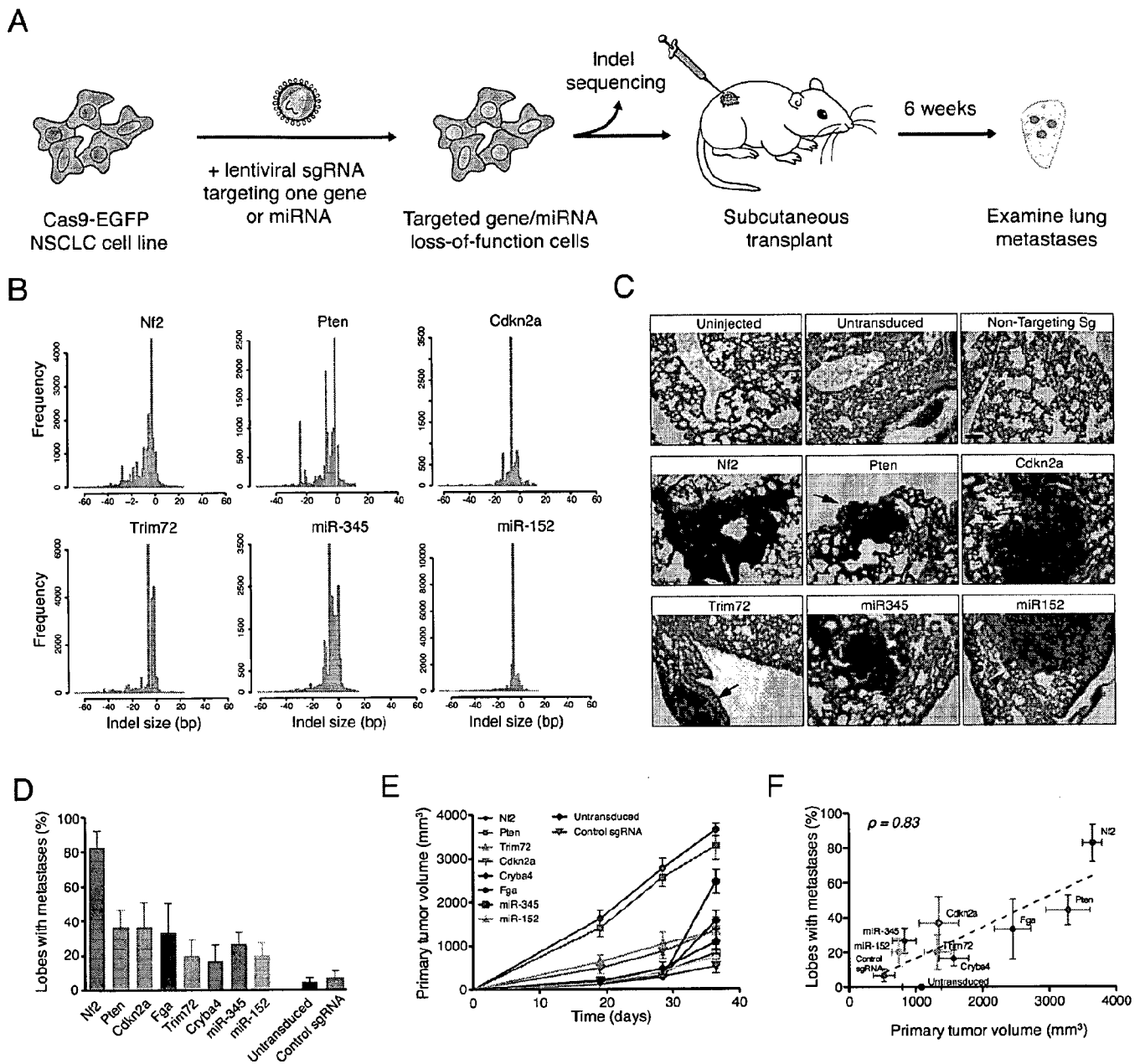
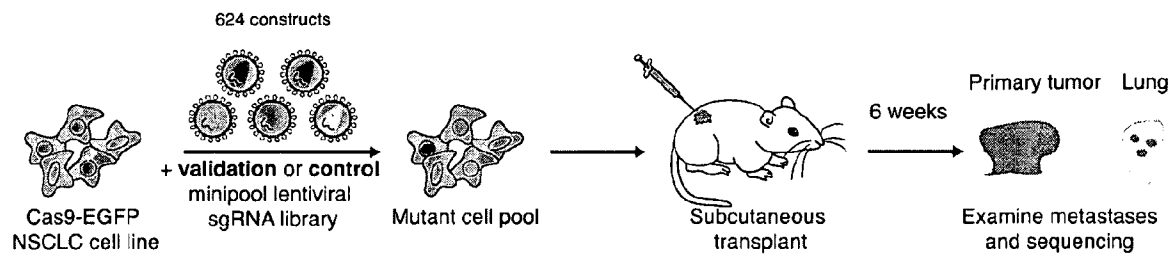
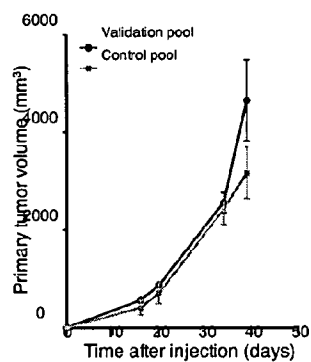


Fig. 5A-5F

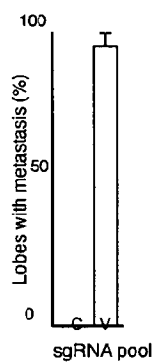
A



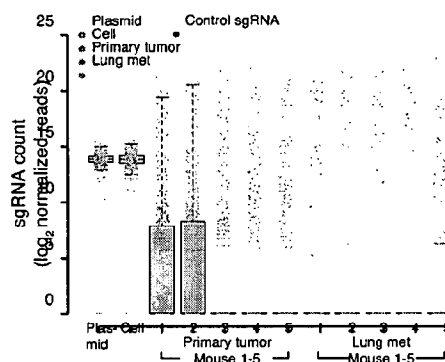
B



C



D



E

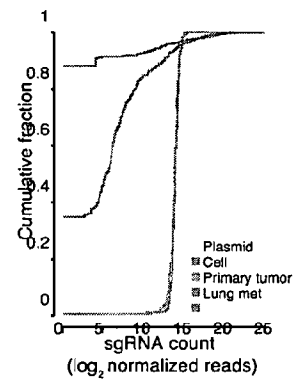


Fig. 6A-6E

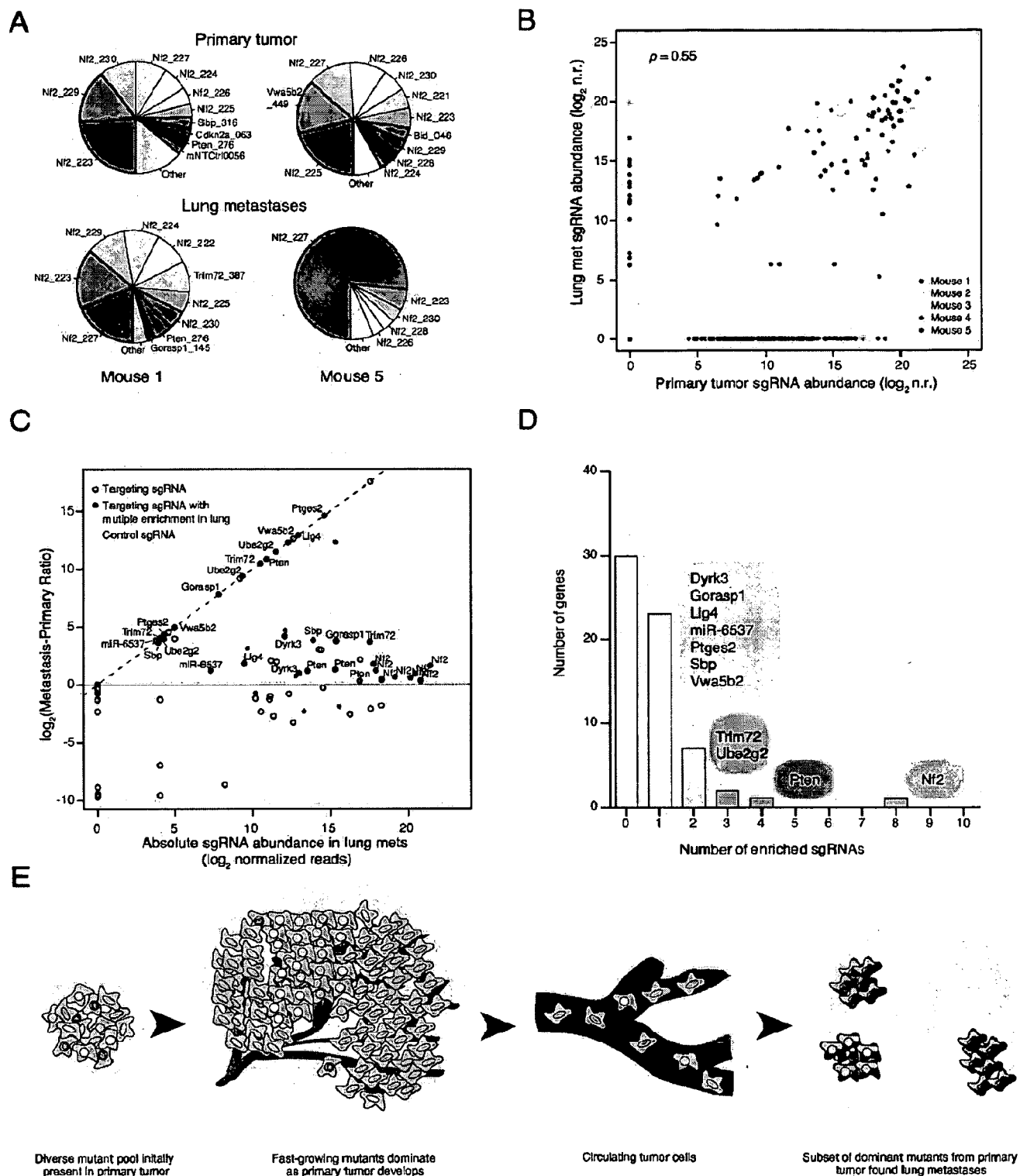


FIG. 7A-E

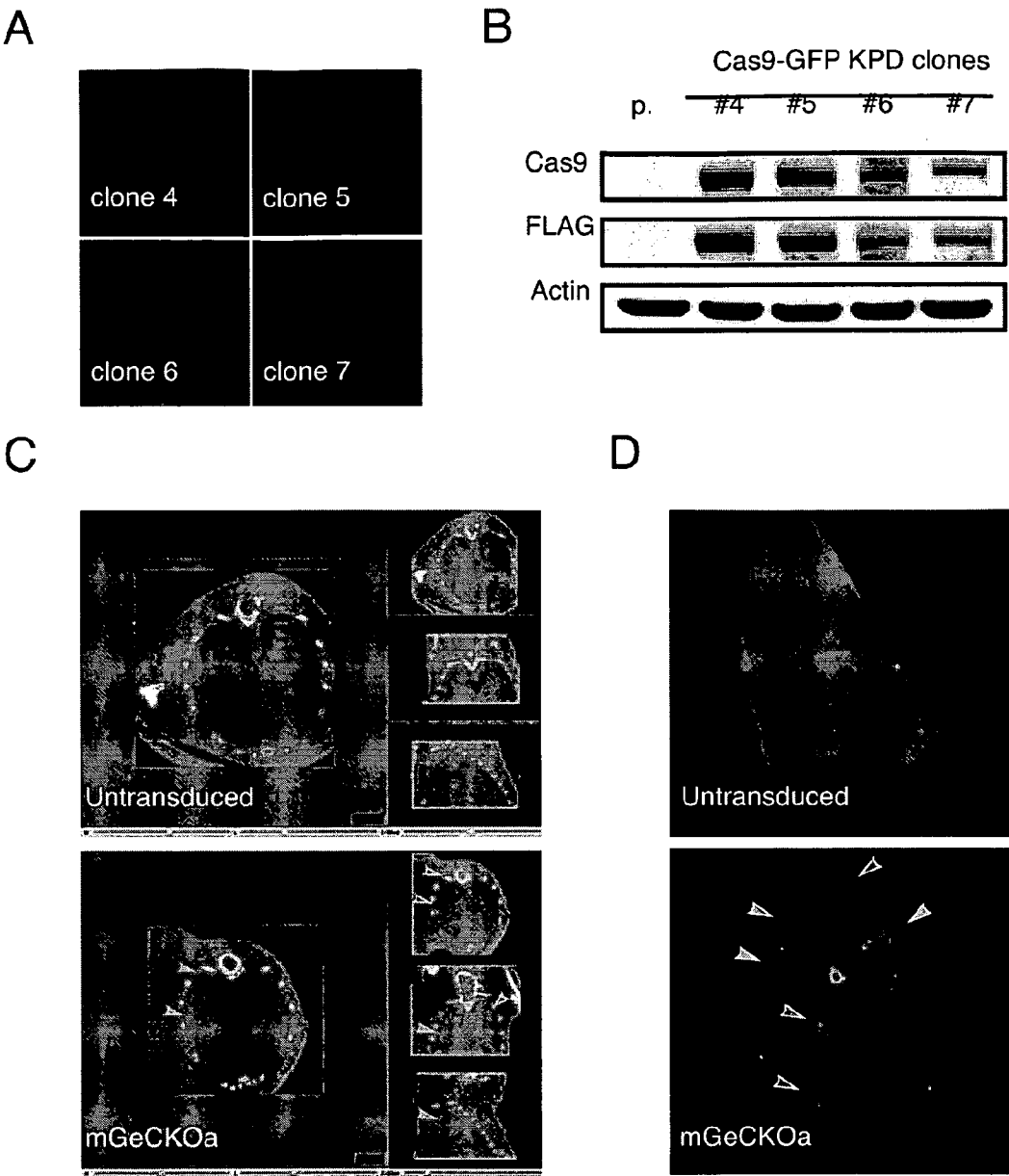


Fig. 8A-8D

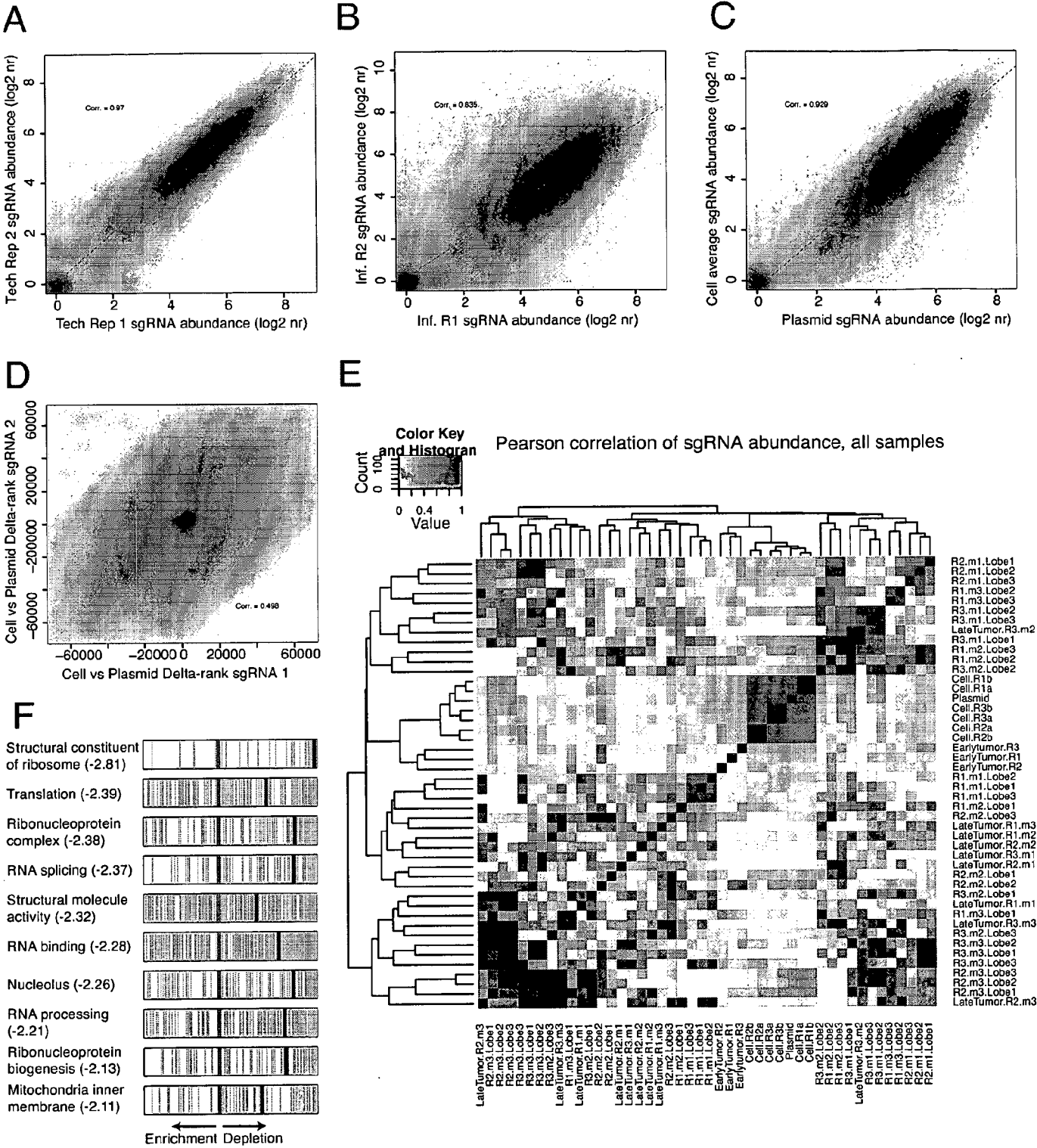


Fig. 9A-9F

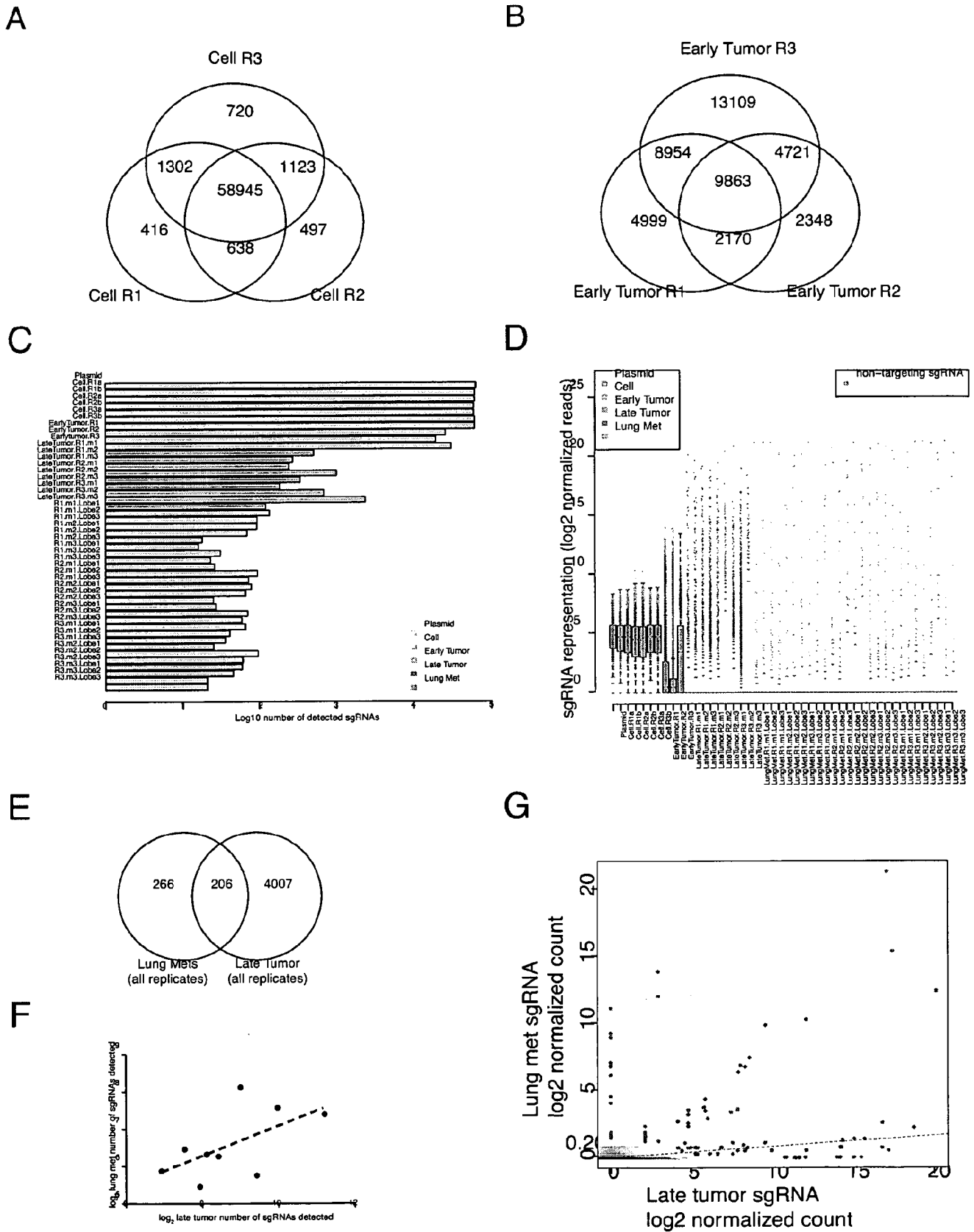


Fig. 10A-10G

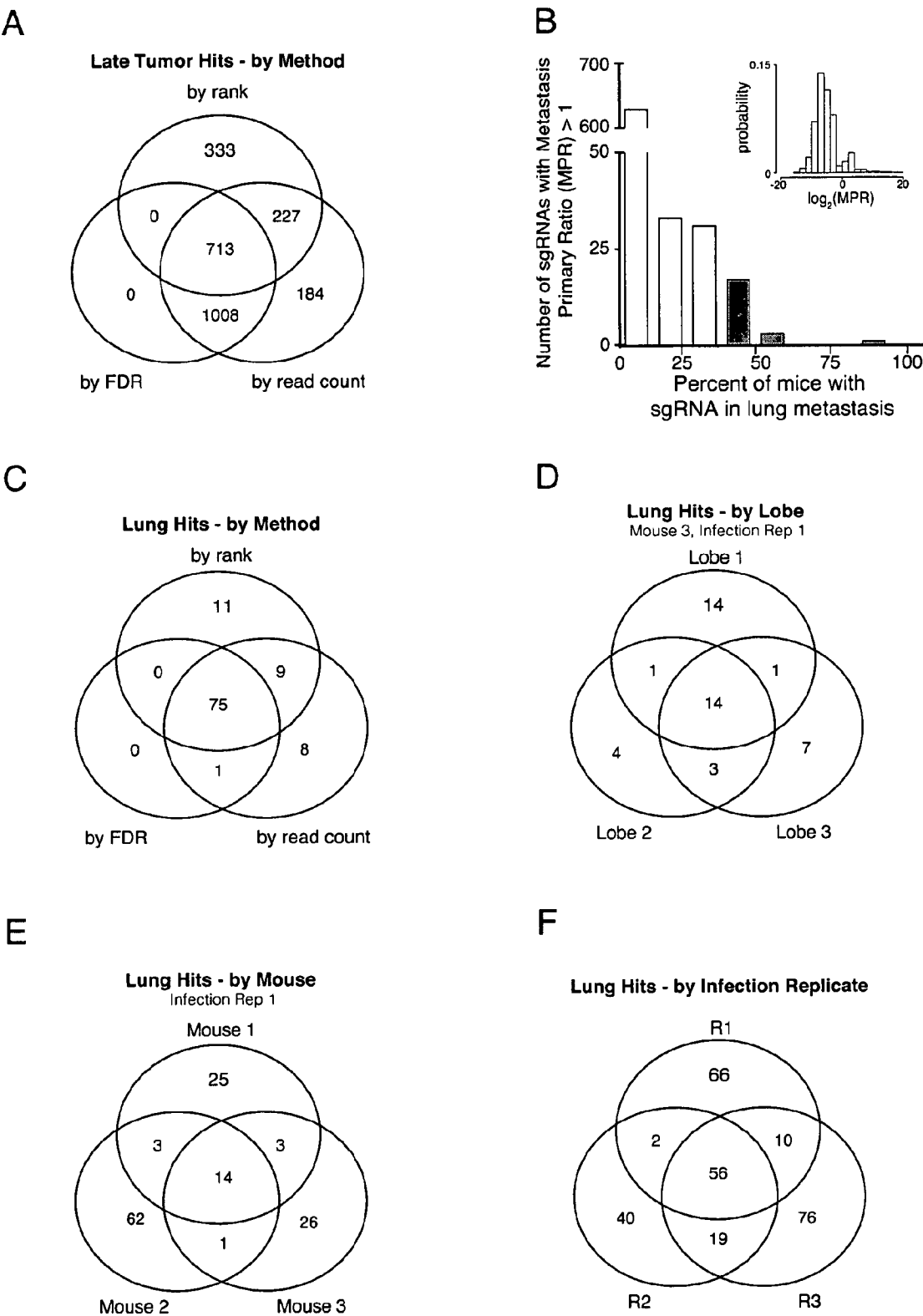


Fig. 11A-11F

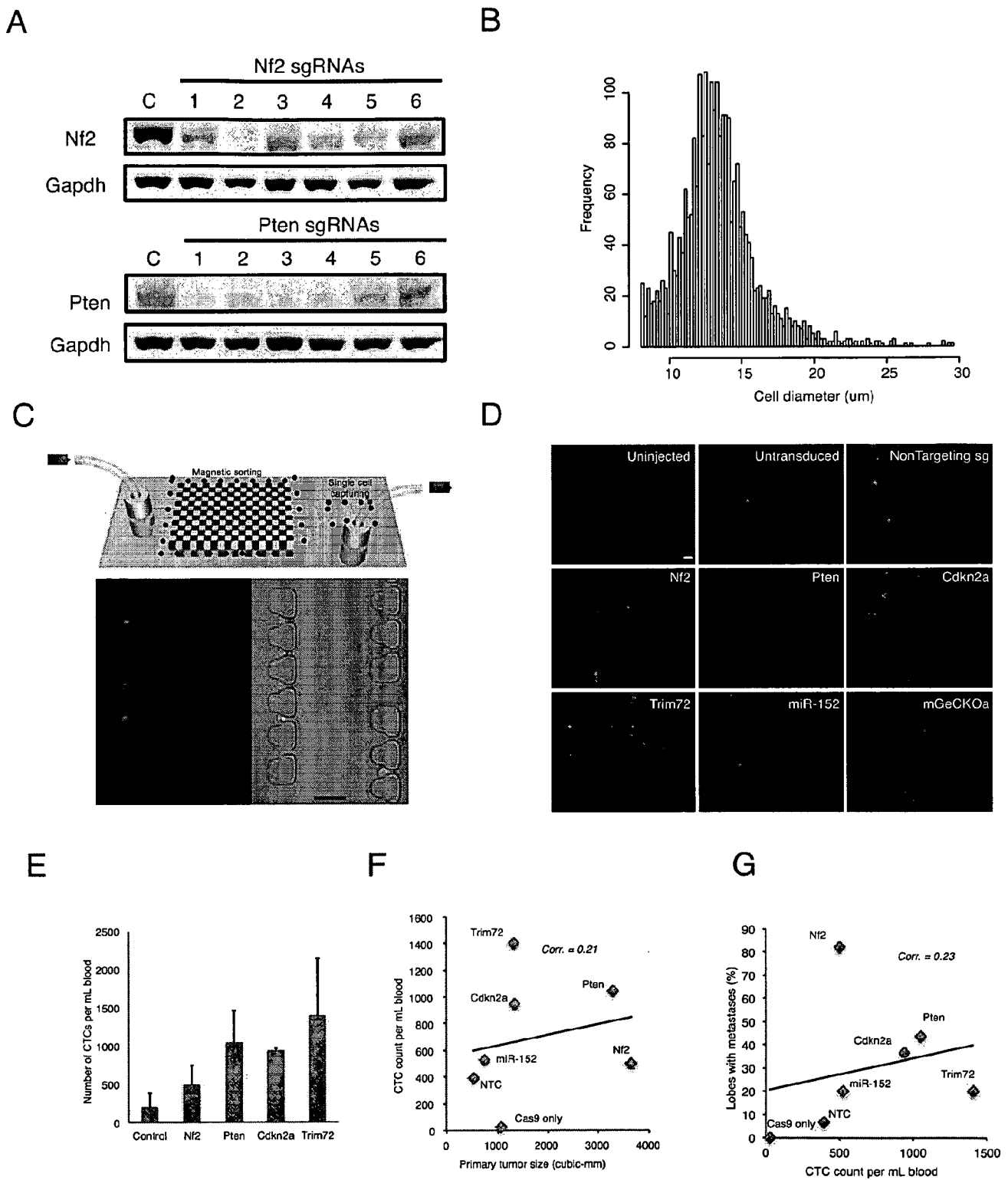


Fig. 13A-13G

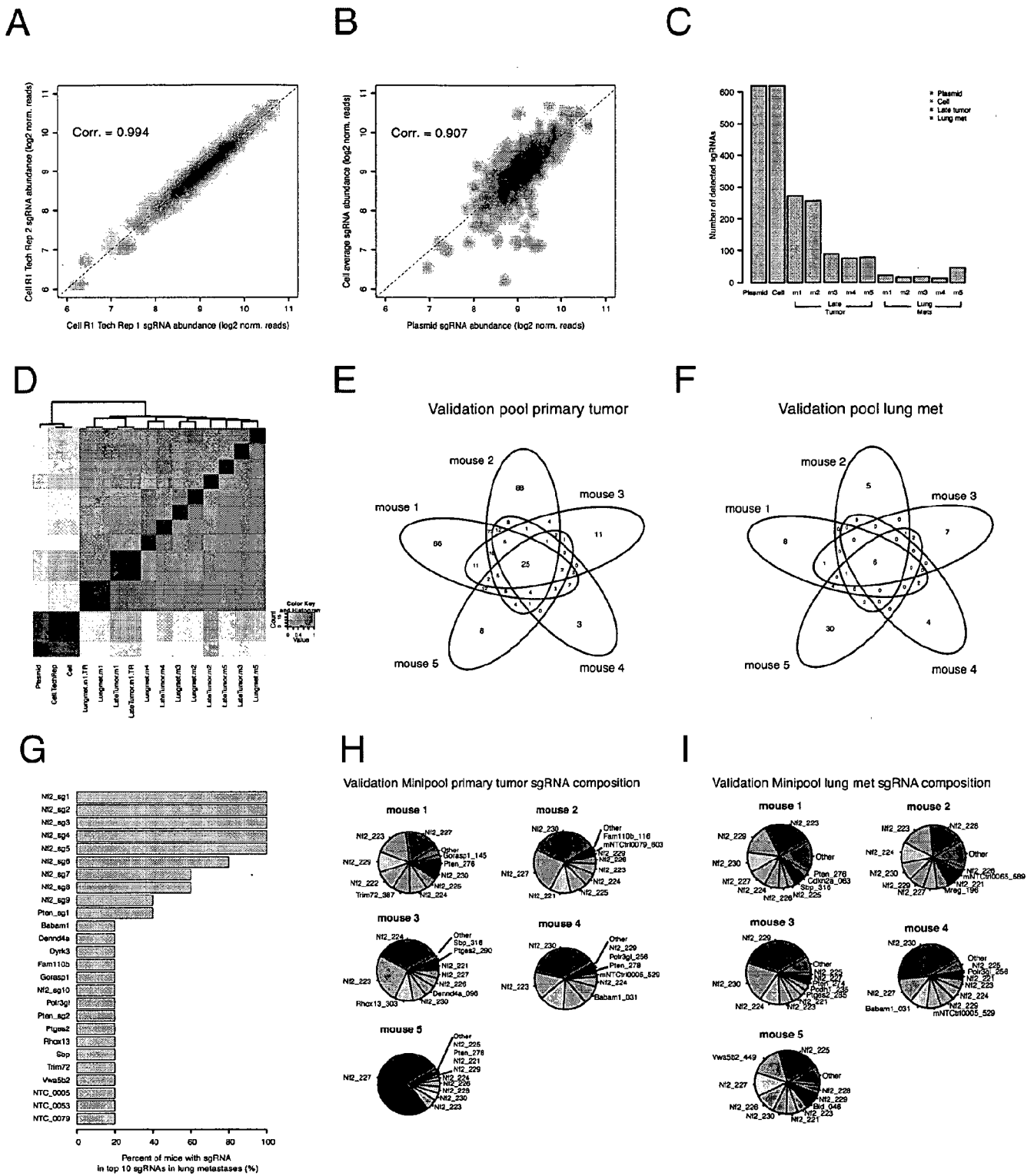


Fig. 14A-14I

Term	PValue
GO:0034660~ncRNA metabolic process	5,34E-07
GO:0043039~tRNA aminoacylation	3,11E-06
GO:0043038~amino acid activation	3,11E-06
GO:0006418~tRNA aminoacylation for protein translation	3,11E-06
GO:0006399~tRNA metabolic process	2,06E-05
GO:0006396~RNA processing	5,52E-05
GO:0006412~translation	3,06E-04
GO:0034660~ncRNA metabolic process	5,34E-07
GO:0034470~ncRNA processing	0,001201556
GO:0006432~phenylalanyl-tRNA aminoacylation	0,001489208
GO:0006397~mRNA processing	0,003053718
GO:0006260~DNA replication	0,007731845
GO:0016071~mRNA metabolic process	0,008403937
GO:0043123~positive regulation of I-kappaB kinase/NF-kappaB cascade	0,00886185
GO:0006259~DNA metabolic process	0,011189394
GO:0032268~regulation of cellular protein metabolic process	0,014163531
GO:0006417~regulation of translation	0,014199449
GO:0006364~rRNA processing	0,018535588
GO:0007605~sensory perception of sound	0,019167945
GO:0048663~neuron fate commitment	0,019850216
GO:0016072~rRNA metabolic process	0,020331833
GO:0008380~RNA splicing	0,022282055
GO:0010608~posttranscriptional regulation of gene expression	0,023399291
GO:0043122~regulation of I-kappaB kinase/NF-kappaB cascade	0,024794682
GO:0008033~tRNA processing	0,030900708
GO:0050885~neuromuscular process controlling balance	0,035221747
GO:0006281~DNA repair	0,036031873
GO:0050954~sensory perception of mechanical stimulus	0,036204544
GO:0060548~negative regulation of cell death	0,036592552
GO:0015031~protein transport	0,038747958
GO:0008654~phospholipid biosynthetic process	0,039571922

Table S1 (B- GOPB Cell dropout)

FIG. 15A

GO:0006338~chromatin remodeling	0,039968437
GO:0015671~oxygen transport	0,041644621
GO:0045184~establishment of protein localization	0,04511632
GO:0043066~negative regulation of apoptosis	0,045963495
GO:0043069~negative regulation of programmed cell death	0,056198261
GO:0007179~transforming growth factor beta receptor signaling pathway	0,059282224
GO:0015669~gas transport	0,064733689
GO:0042058~regulation of epidermal growth factor receptor signaling pathway	0,064733689
GO:0051438~regulation of ubiquitin-protein ligase activity	0,066170603
GO:0032318~regulation of Ras GTPase activity	0,067036136
GO:0045664~regulation of neuron differentiation	0,069824626
GO:0043524~negative regulation of neuron apoptosis	0,071016952
GO:0006084~acetyl-CoA metabolic process	0,072958657
GO:0006974~response to DNA damage stimulus	0,074384011
GO:0060284~regulation of cell development	0,075062878
GO:0008104~protein localization	0,07671382
GO:0022613~ribonucleoprotein complex biogenesis	0,077202885
GO:0043087~regulation of GTPase activity	0,077849976
GO:0009967~positive regulation of signal transduction	0,080490559
GO:0048741~skeletal muscle fiber development	0,081496332
GO:0051028~mRNA transport	0,082284352
GO:0007528~neuromuscular junction development	0,083608105
GO:0000080~G1 phase of mitotic cell cycle	0,088232581
GO:0051340~regulation of ligase activity	0,088232581
GO:0006302~double-strand break repair	0,09211767
GO:0031123~RNA 3'-end processing	0,094779917
GO:0010552~positive regulation of specific transcription from RNA polymerase II promoter	0,094779917
GO:0080135~regulation of cellular response to stress	0,097167815
GO:0016568~chromatin modification	0,099227623

Table S1 (B- GOPB Cell dropout- continued)

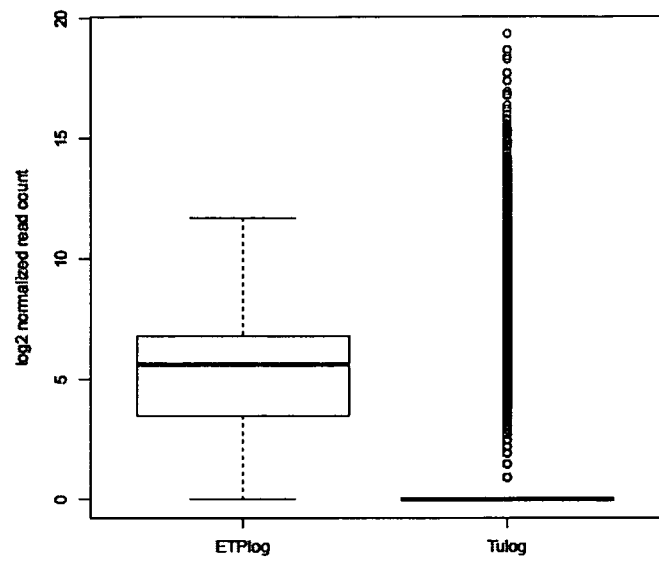
FIG. 15B

Term	PValue
GO:0008380~RNA splicing	0,01637065
GO:0048593~camera-type eye morphogenesis	0,02560709
GO:0034728~nucleosome organization	0,02629497
GO:0006333~chromatin assembly or disassembly	0,03394217
GO:0019370~leukotriene biosynthetic process	0,03448161
GO:0016070~RNA metabolic process	0,04179123
GO:0006334~nucleosome assembly	0,04609538
GO:0048592~eye morphogenesis	0,04609538
GO:0060040~retinal bipolar neuron differentiation	0,04859248
GO:0000082~G1/S transition of mitotic cell cycle	0,05530344
GO:0044242~cellular lipid catabolic process	0,05713935
GO:0065004~protein-DNA complex assembly	0,05800605
GO:0015031~protein transport	0,05826797
GO:0006397~mRNA processing	0,05925072
GO:0006396~RNA processing	0,06029652
GO:0048568~embryonic organ development	0,06345821
GO:0032940~secretion by cell	0,06727065
GO:0008645~hexose transport	0,0709633
GO:0016071~mRNA metabolic process	0,0721051
GO:0051329~interphase of mitotic cell cycle	0,0725353
GO:0007423~sensory organ development	0,07519978
GO:0015749~monosaccharide transport	0,07961359
GO:0051325~interphase	0,08608706
GO:0032647~regulation of interferon-alpha production	0,09230152
GO:0048706~embryonic skeletal system development	0,09306471
GO:0048562~embryonic organ morphogenesis	0,09804005

Table S1 (B- GOPB early tumor dropout)

FIG. 15C

A



B

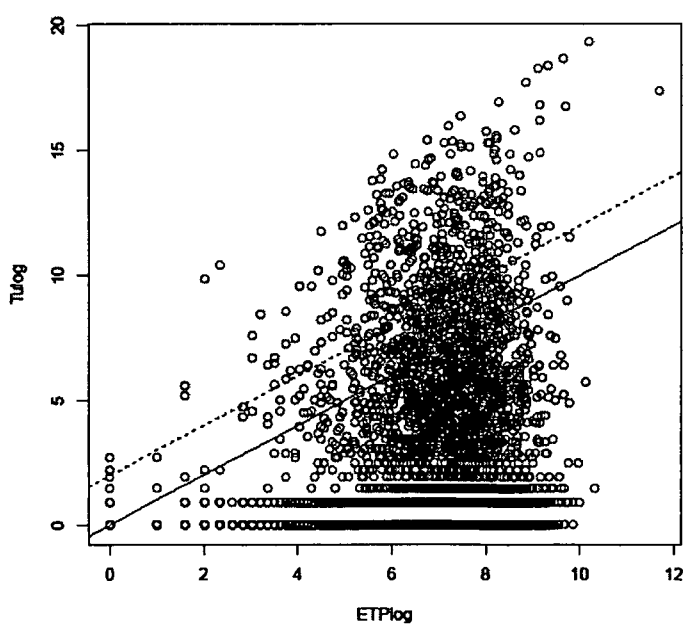


FIG. 16A-16B

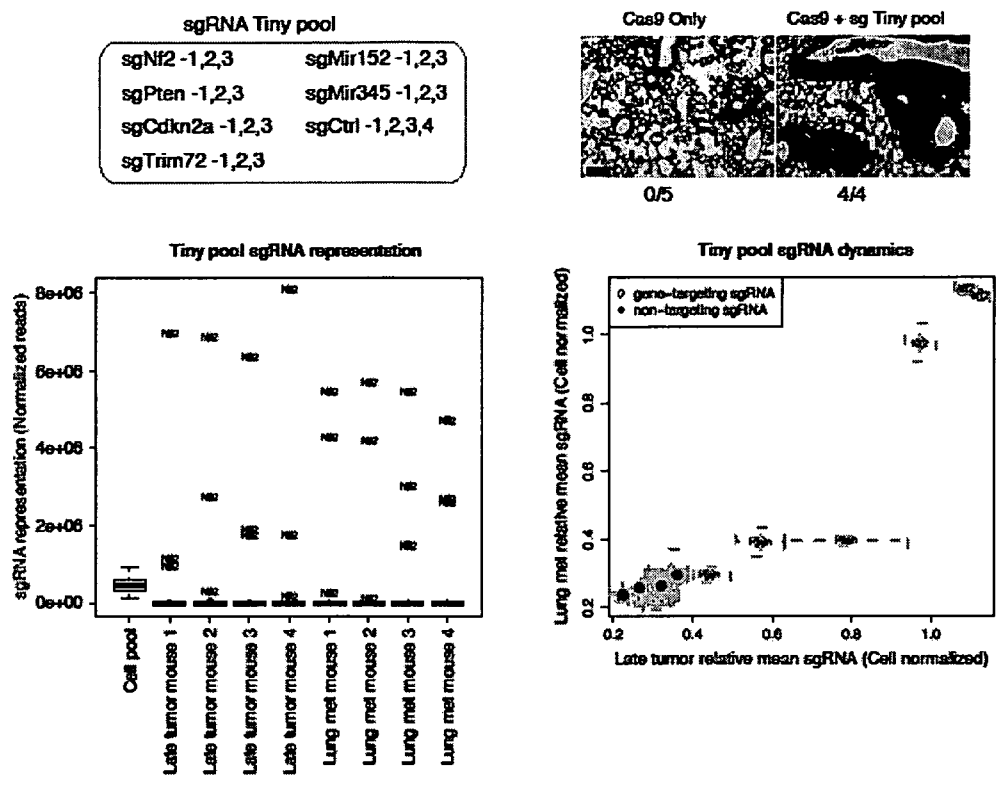


FIG. 17

Early tumor dropout

Term	Count	%	PValue	Genes	FDR
GO:0034660-ncRNA metabolic process	37	2,840970735	5,34E-07	ELAC2, TARS2, NARS, FARS2, UTP6, UTP15, IARS2, FDXACB1, VARS, LIN28B, DMT1, DCAF13, TCP11, CDKN2A, CLP1, TRMT5, RPP30, PIWIL4, NSUN2, EXOSC7, AARS, EXOSC1, LARS2, NOP10, NOP14, TARS, DDX56, RARS, TRMT12, KHSRP, HARS, FARSB, FARSA, PES1, DUS4L, AARS2, DDX51	9,47E-04
GO:0043039-tRNA aminoacylation	15	1,070663812	3,11E-06	TARS2, NARS, FARS2, AARS, IARS2, LARS2, FDXACB1, VARS, TARS, RARS, FARSB, HARS, KHSRP, FARSA, AARS2	0,005513874
GO:0043038-amino acid activation	15	1,070663812	3,11E-06	TARS2, NARS, FARS2, AARS, IARS2, LARS2, FDXACB1, VARS, TARS, RARS, FARSB, HARS, KHSRP, FARSA, AARS2	0,005513874
GO:0006418-tRNA aminoacylation for protein translation	15	1,070663812	3,11E-06	TARS2, NARS, FARS2, AARS, IARS2, LARS2, FDXACB1, VARS, TARS, RARS, FARSB, HARS, KHSRP, FARSA, AARS2	0,005513874
GO:0006399-tRNA metabolic process	23	1,841684511	2,06E-05	TARS2, ELAC2, NARS, FARS2, AARS, IARS2, LARS2, VARS,	0,038462543

Figure 18 (supplemental tables)

FIG. 18A

				FDXACB1, TARS, TCP11, CLP1, RARS, TRMT12, TRMT5, FARSB, KHSRP, HARS, RPP30, FARSA, DUS4L, NSUN2, AARS2	
GO:0006396~RNA processing	56	3,997144897	6.52E-05	U2AF2, SNRPD3, LSM7, UTP15, FDXACB1, YBX1, SMNDC1, DCAF13, APP, CDKN2A, CLP1, TRMT5, RPP30, DBR1, LSM3, ZCCHC8, EXOSC7, AARS, EXOSC1, TTF2, SNRNP48, TRMT12, SNRPB, KHSRP, CPSF6, RBM39, CPSF4, THOC3, XRN2, ELAC2, FARS2, PPIL1, ADAD1, UTP6, LIN28B, DMT1, TCP11, ZFP36L3, PRPF8, GEMIN8, GEMIN7, NSUN2, CSTF3, SNW1, PRPF18, NOP10, NOP14, DDX58, PAPOLA, SFPQ, PAPOLG, DUS4L, PES1, SNRNP27, CSTF1, DDX51	0.097870478
GO:0006412~translation	42	2,997858872	3.06E-04	RPL17, RPL19, EEF1B2, TARS2, NARS, FARS2, HBS1L, EIF5B, RPL35, RPS15A, IARS2, FDXACB1, VARS, MRPL10, EIF3B, MRPL15, RPL7, EIF3H, MRPL36, MRPL55, FLT3L, RPS21, MRPL33	0.542255487

Figure 18 (supplemental tables continued)

FIG. 18B

				RPS27A, EIF2B5, MRPL2, MRPL4, AARS, MRPS8, DENR, LARS2, RPS8, TARS, MRPL21, RARS, HARS, KHSRP, RPL3L, FARSB, PELO, FARSA, AARS2	
GO:0034470-ncRNA processing	24	1,713062099	0,00120156	ELAC2, EXOSC7, FARS2, AARS, UTP6, UTP15, EXOSC1, FDXACB1, NOP10, LIN28B, DMT1, NOP14, DCAF13, DDX58, TCP11, CDKN2A, CLP1, TRMT12, TRMT5, RPP30, PES1, DUS4L, NSUN2, DDX51	2,110398883
GO:0006432-phenylalanyl-tRNA aminoacylation	4	0,28551035	0,00148921	FARS2, FARSB, FARSA, FDXACB1	2,60935578
GO:0006397-mRNA processing	33	2,355460385	0,00305372	PPIL1, SNRPD3, U2AF2, LSM7, YBX1, SMNDC1, ZFP36L3, APP, TCP11, CLP1, PRPF8, DBR1, GEMIN8, LSM3, GEMIN7, ZCCHC8, CSTF3, SNW1, PRPF18, TTF2, PAPOLA, SNRNP48, SFPO, SNRPB, KHSRP, CPSF6, RBM39, CPSF4, PAPOLG, SNRNP27, CSTF1, THOC3, XRN2	5,281374654
GO:0006260-DNA replication	21	1,498929338	0,00773185	ING5, UPF1, PDGFB, PHB, GINS4, POLE, PAPD7, RNASEH1, LIG4, WRN, RBBP7, TK2, RAD51, RFC5, RFC4, HELB, TOP3A, TEP1, CHAF1B, NFIB, RBMS1	12,86445645
GO:0016071-mRNA metabolic	35	2,49821556	0,00840394	PPIL1, SNRPD3,	13,80561571

Figure 18 (supplemental tables continued)

FIG. 18C

process				U2AF2, LSM7, MAPKAPK2, YBX1, SMNDC1, ZFP38L3, APP, TCP11, PRPF8, CLP1, DBR1, GEMIN8, LSM3, GEMIN7, ZCCHC8, CSTF3, UPF1, SNW1, PRPF18, TTF2, PAPOLA, SNRNP48, SFPQ, SNRPB, KHSRP, CPSF6, CPSF4, RBM39, PAPOLG, SNRNP27, CSTF1, THOC3, XRN2	
GO:0043123~positive regulation of I-kappaB kinase/NF-kappaB cascade	8	0,5710207	0,00886185	UBE2N, MAP3K7, CARD11, LST1, MYD88, NOD1, NUP62, LTB	14,60824868
GO:0006259~DNA metabolic process	45	3,211891435	0,01118939	ING5, KIF22, HMGB2, APEX2, PDGFB, FAM175A, MRE11A, RNASEH1, BNIP3, SMUG1, TK2, SETX, MUTYH, PIWIL4, CCNO, RBMS1, SGIP1, UPF1, PHB, EME2, MSH4, POLE, GINS4, NEIL1, PAPD7, LIG4, WRN, GTF2H5, RAD52, RBBP7, EEPD1, RAD51, RFC5, UBE2N, XPA, RFC4, HELB, ENDOG, SFPQ, TOP3A, TEP1, RUVBL1, CHAF1B, XRN2, NFIB	18,09881674
GO:0032268~regulation of cellular protein metabolic process	32	2,284082788	0,01416353	SRP14, ATG10, PDGFB, IMPACT, ADAD1, MKNK2, MKNK1, IGF2BP3, IL11, IL31RA, PRKAR2B, APP, CDKN2A, HBB-B1, PIWIL4, EIF2B5,	22,35900433

Figure 18 (supplemental tables continued)

FIG. 18D

				FCER1A, UPF1, FBXO2, BARHL2, IREB2, SMAD4, 4931440F15R8K, CDK4, RNF180, UBE2N, CD59B, NGDN, IGFBP3, NCOR1, SRP9, CBS	
GO:0006417~regulation of translation	15	1,070663812	0,01419945	SRP14, UPF1, IMPACT, BARHL2, ADAD1, MKNK2, IREB2, MKNK1, IGF2BP3, CDK4, APP, NGDN, PIWIL4, SRP9, EIF2B5	22,40817475
GO:0006364~rRNA processing	12	0,856531049	0,01853559	NOP14, DMT1, DCAF13, DDX58, CDKN2A, EXOSC7, UTP6, UTP15, EXOSC1, PES1, NOP10, DDX51	28,24621148
GO:0007605~sensory perception of sound	13	0,927908637	0,01916795	MYO1A, THRB, HEXA, HEXB, SOBP, OTOS, GPR88, ESPN, GJB2, MARVELD2, CLIC5, GTF3C2, AXIN1	29,08199548
GO:0048663~neuron fate commitment	9	0,642398267	0,01985022	NOTCH3, DLX1, ISL2, LBX1, HES5, FOXA1, SMAD4, NKO2-1, TLX3	29,9323531
GO:0016072~rRNA metabolic process	12	0,856531049	0,02033183	NOP14, DMT1, DCAF13, DDX58, CDKN2A, EXOSC7, UTP6, UTP15, EXOSC1, PES1, NOP10, DDX51	30,54066283
GO:0008380~RNA splicing	24	1,713062099	0,02228205	SNRPD3, UZAF2, PPIL1, LSM7, SNW1, PRPF18, YBX1, TTF2, SMNDC1, TCP11, SNRNP48, CLP1, PRPF8, SFPO, SNRPB, KHSRP, DBR1, GEMIN8, LSM3, RBM39, SNRNP27, GEMIN7, THOC3,	32,95332856

Figure 18 (supplemental tables continued)

FIG. 18E

GO:0010608~posttranscriptional regulation of gene expression	19	1,356174161	0,02339929	ZCCHC8 SRP14, UPF1, IMPACT, BARHL2, ADAD1, IREB2, MKNK2, MKNK1, IGF2BP3, MAPKAPK2, CDK4, LIN28B, APP, CDKN2A, CLP1, AGDN, PIWIL4, SRP9, EIF2B5	34,29963175
GO:0043122~regulation of I-kappaB kinase/NF-kappaB cascade	8	0,5710207	0,02479468	UBE2N, MAP3K7, CARD11, LST1, MYD88, NOD1, NUP62, LTB	35,94530447
GO:0008033~tRNA processing	11	0,765153462	0,03090071	TCP11, ELAC2, FARS2, CLP1, TRMT12, AARS, TRMT5, RPP30, DUS4L, FDXACB1, NSUN2	42,69971538
GO:0050885~neuromuscular process controlling balance	7	0,499643112	0,03522175	APP, TPP1, HEXA, CLIC5, ADCY5, HEXB, AARS	47,06716372
GO:0006281~DNA repair	25	1,784439686	0,03603187	KIF22, HMGB2, APEX2, FAM175A, MRE11A, SMUG1, SETX, MUTYH, CCNO, UPF1, EME2, POLE, MSH4, NEIL1, GTF2H5, WRN, LIG4, RAD52, EEPD1, RAD51, UBE2N, XPA, SFPQ, CHAF1B, XRN2	47,85020284
GO:0050954~sensory perception of mechanical stimulus	13	0,927908837	0,03820454	MYO1A, THRB, HEXA, HEXB, SOBP, OTOS, GPR98, ESPN, GJB2, MARVELD2, CLIC5, GTF3C2, AXIN1	48,01568068
GO:0060548~negative regulation of cell death	27	1,927194861	0,03659255	STIL, STAT5A, SNCA, BNP3, SGMS1, BCL2L1, CITED2, MAP3K7, PAK7, G2E3, CHST11, DNAJC8, FAS, PTCRA, SPP1, PIK3CG, AARS, SKP2	48,3857205

Figure 18 (supplemental tables continued)

FIG. 18F

				LIG4, BCL2L10, NME5, DLX1, CD59B, NUP62, BCL2A1A, AGTR1A, NGF	
GO:0015031~protein transport	61	4,354032834	0,03874796	SRP14, SEC24B, SEC31B, RAB5B, HPS1, VPS52, CCHCR1, COPB2, TIMM9, RANBP2, SAR1B, RAB21, SCAMP1, ZDHHC3, VPS45, VPS41, NUPL2, ATG4C, AGTR1A, RAB17, USO1, TOMM22, SRP9, ARFGAP1, RAB3B, ATG10, SNX1, ARF8, SNX3, TIMM13, RAB42, RAB43, TOMM7, TOMM6, CSE1L, DDX19A, AP3M1, EXOC4, SEC22A, PEX14, PEX13, SNX24, NUP54, TNPO1, TNPO3, ASPSCR1, RAB8B, RABIF, IPO11, TOMM40, NUP155, AP4S1, PREB, ARF1, NUP62, CADPS2, TOM1L2, SEC13, HGS, SNX30, ARAP1	50,39647209
O:0008654~phospholipid biosynthetic process	12	0,856531049	0,03957192	PIGK, AGPAT6, SERINC5, PGS1, CDIPT, PIGM, AGPAT5, HEXB, DPM2, PCYT1B, SGMS1, CDS1	51,14542295
GO:0006338~chromatin remodeling	7	0,499643112	0,03996844	SATB2, BAZ1B, SMARCB1, BNIP3, INO80, HMG20B, RBBP7	51,50202418
GO:0015671~oxygen transport	4	0,28551035	0,04164462	HBB-BH1, HBB-B1, HBA-X, MB	52,98252166
GO:0045184~establishment of	61	4,354032834	0,04511632	SRP14, SEC24B,	55,91435741

Figure 18 (supplemental tables continued)

FIG. 18G

protein localization				SEC31B, RAB5B, HPS1, VPS52, CCHCR1, COPB2, TIMM9, RANBP2, SAR1B, RAB21, SCAMP1, ZDHHC3, VPS45, VPS41, NUPL2, ATG4C, AGTR1A, RAB17, USO1, TOMM22, SRP9, ARFGAP1, RAB38, ATG10, SNX1, ARF6, SNX3, TIMM13, RAB42, RAB43, TOMM7, TOMM8, CSE1L, DDX19A, AP3M1, EXOC4, SEC22A, PEX14, PEX13, SNX24, NUP54, TNPO1, TNPO3, ASPSCR1, RAB8B, RABIF, IPO11, TOMM40, NUP155, AP4S1, PREB, ARF1, NUP62, CADPS2, TOM1L2, SEC13, HGS, SNX30, ARAP1	
GO:0043066~negative regulation of apoptosis	26	1,855817273	0,04596349	STIL, STAT5A, SNCA, BNIP3, SGMS1, BCL2L1, CITED2, MAP3K7, PAK7, G2E3, CHST11, DNAJC5, FAS, PTCRA, SPP1, PIK3CG, AARS, SKP2, LIG4, BCL2L10, NME5, DLX1, NUP62, BCL2A1A, AGTR1A, NGF	56,60313774
GO:0043068~negative regulation of programmed cell death	26	1,855817273	0,05619828	STIL, STAT5A, SNCA, BNIP3, SGMS1, BCL2L1, CITED2, MAP3K7, PAK7, G2E3, CHST11, DNAJC5, FAS, PTCRA, SPP1, PIK3CG, AARS, SKP2	64,16110385

Figure 18 (supplemental tables continued)

FIG. 18H

				LIG4, BCL2L10, NME5, DLX1, NUP62, BCL2A1A, AGTR1A, NGF	
GO:0007179~transforming growth factor beta receptor signaling pathway	8	0,5710207	0,05928222	MAP3K7, PDGFB, LTBP4, SMAD6, USP9X, SMAD4, TAB1, ARAP1	66,18286002
GO:0015669~gas transport	4	0,28551035	0,06473389	HBB-BH1, HBB-B1, HBA-X, MB	69,4959843
GO:0042058~regulation of epidermal growth factor receptor signaling pathway	4	0,28551035	0,06473369	APP, NUP62, SOS1, ZFYVE28	69,4959843
GO:0051438~regulation of ubiquitin-protein ligase activity	3	0,214132762	0,0661706	UBE2N, CDKN2A, 4931440F15RIK	70,31682673
GO:0032318~regulation of Ras GTPase activity	11	0,785153462	0,06703614	ARFGAP1, ALS2, ADAP2, TBC1D12, RALBP1, TBC1D10A, ASAP3, TBC1D20, ARAP1, TBCK, TBC1D9B	70,80115861
GO:0045664~regulation of neuron differentiation	13	0,927908637	0,06982463	RTN4, LBX1, SOCS2, BARHL2, FOXA1, REST, NOTCH3, ISL2, RNF8, HES5, ROBO2, TLX3, NGF	72,31131655
GO:0043524~negative regulation of neuron apoptosis	8	0,5710207	0,07101695	DLX1, AGTR1A, SNCA, AARS, DNAJC5, LIG4, BCL2L1, NGF	72,93427912
GO:0006084~acetyl-CoA metabolic process	6	0,428265525	0,07295866	SDHA, MLYCD, SDHD, IDH2, DLAT, FH1	73,92055117
GO:0006974~response to DNA damage stimulus	29	2,068950036	0,07438401	KIF22, HMGB2, APEX2, FAM175A, MRE11A, SMUG1, SETX, MUTYH, CCNO, CDK1, UPF1, EME2, POLE, MSH4, NEIL1, LIG4, GTF2H5, WRN, RAD52, EEPD1, RAD51, UBE2N, MAPK1, XPA, BAZ1B, SFPQ, MAPK3, CHAF1B, XRN2	74,62285477

Figure 18 (supplemental tables continued)

FIG. 18I

GO:0060284~regulation of cell development	18	1.284796574	0.07506288	RTN4, LBX1, SOCS2, FOXA1, BARHL2, SMAD4, LIG4, REST, NOTCH3, DLX1, ISL2, RNF6, CDKN2A, HES5, ROBO2, TLX3, IGFBP3, NGF	74.95103753
GO:0008104~protein localization	67	4.762298358	0.07671382	ALS2, SRP14, SEC24B, SEC31B, RAB5B, HPS1, VPS52, CCHCR1, COPB2, TIMM9, RANBP2, SAR1B, RAB21, SCAMP1, ZDHHC3, VPS45, VPS41, NUPL2, ATG4C, AGTR1A, CLIC5, RAB17, USO1, TOMM22, SRP9, ARFGAP1, CAV3, RAB3B, ATG10, SNX1, ARF6, SNX3, TIMM13, SUFU, RAB42, RAB43, TOMM7, TOMM6, CSE1L, DDX18A, AP3M1, EXOC4, SEC22A, PEX14, PEX13, NUP54, SNX24, TNPO1, TNPO3, ASPSCR1, RAB8B, RABIF, IPO11, TOMM40, NUP155, AP4S1, PREB, SYNE2, ARF1, NUP62, CADPS2, TOM1L2, SEC13, HGS, FAF1, SNX30, ARAP1	75.73251455
GO:0022613~ribonucleoprotein complex biogenesis	16	1.142041399	0.07720288	EXOSC7, UTP6, UTP15, EXOSC1, NOP10, SMNDC1, DMT1, NOP14, DCAF13, MRPL10, DDX56, CDKN2A, CLP1, GEMIN8, PES1, DDX51	75.95956013

Figure 18 (supplemental tables continued)

FIG. 18J

GO:0043087~regulation of GTPase activity	12	0,856531049	0,07784998	ARFGAP1, ALS2, ADAP2, TBC1D12, RASGRP3, RALBP1, TBC1D10A, ASAP3, TBC1D20, ARAP1, TBCK, TBC1D9B	76,25688935
GO:0009967~positive regulation of signal transduction	19	1,358174161	0,08048056	FCER1A, LST1, PDGFB, SKP2, SMAD4, IL11, IL31RA, CITED2, MAP3K7, UBE2N, CARD11, GPC3, MYD88, CDKN2A, NOD1, NUP62, SOS1, LTB, AXIN1	77,43480657
GO:0048741~skeletal muscle fiber development	6	0,428265525	0,08148633	ALS2, APP, ACTA1, TNC, CACNA1S, SNTA1	77,86851507
GO:0051028~mRNA transport	9	0,642398287	0,08228435	ENY2, NUP62, DDX19A, KHSPR, NUP54, RANBP2, NUP155, NUPL2, THOC3	78,20296721
GO:0007528~neuromuscular junction development	5	0,356887937	0,0836081	ALS2, APP, TNC, CACNA1S, SNTA1	78,75408388
GO:0000080~G1 phase of mitotic cell cycle	3	0,214132762	0,08823258	PPP2R3A, CDC123, MTBP	80,5779525
GO:0051340~regulation of ligase activity	3	0,214132762	0,08823258	UBE2N, CDKN2A, 4931440F15RK	80,5779525
GO:0006302~double-strand break repair	7	0,499643112	0,09211767	UBE2N, FAM175A, MRE11A, LIG4, RAD52, SETX, RAD51	81,89497006
GO:0031123~RNA 3'-end processing	5	0,356887937	0,09477992	ZFP36L3, APP, PAPOLA, PAPOLG, LIN28B	82,90901821
GO:0010552~positive regulation of specific transcription from RNA polymerase II promoter	5	0,356887937	0,09477992	HMG82, HES5, NFKB1, IRF4, CITED2	82,90901821
GO:0080135~regulation of cellular response to stress	10	0,713775874	0,09716782	UBE2N, FCER1A, MAP3K7, CDKN2A, MYD88, NOD1, FAM175A, HBB-B1, AXIN1, CBS	83,69145467
GO:0016568~chromatin	24	1,713062099	0,09922762	ENY2, ING5, DPF3,	84,33811046

Figure 18 (supplemental tables continued)

FIG. 18K

modification				FAM175A, INO80, BNIP3, DMAP1, CHD9, SMARCB1, PET117, ASF1B, SATB2, CSRP2BP, L3MBTL2, HMG20B, PAD14, RBBP7, UBE2N, NPTXR, BAZ1B, SMARCC1, RUVBL1, IRF4, RBM14, NCOR1	
--------------	--	--	--	---	--

Late tumor hits

Category	Term	Count	%	PValue	Genes	FDR
GOTERM_BP_FAT	GO:0042981~regulation of apoptosis	30	5.464480874	0.00136618	BID, BCLAF1, HTATIP2, FGFR3, SNCB, MGMT, PRKDC, SOX9, PTEN, IKBIP, IRAK3, CDKN2A, LHX3, DEDD2, SPN, PLAGL2, CCAR1, MUC2, NTF3, MUC20, GRIN1, CIDEA, LIG4, PIM2, PEA15A, TAX1BP1, EYA1, GRM2, HIPK1, PTPRV	2.298984748
GOTERM_BP_FAT	GO:0043067~regulation of programmed cell death	30	5.464480874	0.001649572	BID, BCLAF1, HTATIP2, FGFR3, SNCB, MGMT, PRKDC, SOX9, PTEN, IKBIP, IRAK3, CDKN2A, LHX3, DEDD2, SPN, PLAGL2, CCAR1, MUC2, NTF3, MUC20, GRIN1, CIDEA, LIG4, PIM2, PEA15A,	2.769598817

Figure 18 (supplemental tables continued)

FIG. 18L

					TAX1BP1, EYA1, GRM2, HIPK1, PTPRV	
GOTERM_BP_FAT	GO:0010941~regulation of cell death	30	5.464480874	0.001783665	BID, BCLAF1, HTATIP2, FGFR3, SNCB, MGMT, PRKDC, SOX9, PTEN, IKBIP, IRAK3, CDKN2A, LHX3, DEDD2, SPN, PLAGL2, CCAR1, MUC2, NTF3, MUC20, GRIN1, CIDEA, LIG4, PIM2, PEA15A, TAX1BP1, EYA1, GRM2, HIPK1, PTPRV	2,991537368
GOTERM_BP_FAT	GO:0043086~negative regulation of catalytic activity	10	1.821493625	0.002501466	IRAK3, CDKN2A, NF2, PKIG, ZFYVE28, GABBR1, IL1B, PPP2R4, NOS3, OAS1D	4,171482788
GOTERM_BP_FAT	GO:0048585~negative regulation of response to stimulus	8	1.4571949	0.002583844	IRAK3, KLK8, GRIN1, CD276, IL1B, MTOR, GRIN3A, SPN	4,273382449
GOTERM_BP_FAT	GO:0006816~calcium ion transport	11	2.003642987	0.002574547	ATP2B2, PLN, GRIN1, CACNG6, CACNB3, GRIN3A, CACNA1D, CACNA2D2, CAMK2A, TTPR1, TPCN2	4,290856022
GOTERM_BP_FAT	GO:0032102~negative regulation of response to external stimulus	6	1.092898175	0.003476997	KLK8, GRIN1, CD276, MTOR, GRIN3A, SPN	5,753453937
GOTERM_BP_FAT	GO:0051348~negative regulation of transferase activity	7	1.275045537	0.00400413	IRAK3, CDKN2A, NF2, PKIG, ZFYVE28, IL1B, OAS1D	6,598017324
GOTERM_BP_FAT	GO:0044092~negative	11	2.003642987	0.004812453	IRAK3, CDKN2A	7,879272441

Figure 18 (supplemental tables continued)

FIG. 18M

	regulation of molecular function				THRA, NF2, PKIG, ZFYVE28, GABBR1, IL1B, PPP2R4, NOS3, OAS1D	
GOTERM_BP_FAT	GO:0008285~negative regulation of cell proliferation	15	2.732240437	0.005391769	MUC2, FGFR3, NF2, JARID2, STK11, CD276, CXADR, PTEN, CDKN2A, EIF2AK1, CDKN2B, PTPRV, CDKN2C, NOS3, SPN	8.787337203
GOTERM_BP_FAT	GO:0044057~regulation of system process	14	2.550091075	0.005452268	EDN3, KLK8, THRA, NTF3, GRIN1, BMPR2, CACNA2D2, ATP2B2, GRM2, P2RX3, PLN, IL1B, NOS3, CAMK2A	8.881728042
GOTERM_BP_FAT	GO:0043068~positive regulation of programmed cell death	16	2.9143888	0.005873626	BID, MUC2, BCLAF1, HTATIP2, FGFR3, GRIN1, MUC20, PRKDC, PTEN, IKBP, CDKN2A, HIPK1, PTPRV, SPN, CCAR1, PLAGL2	9.538213098
GOTERM_BP_FAT	GO:0010642~positive regulation of cell death	16	2.9143888	0.008310514	BID, MUC2, BCLAF1, HTATIP2, FGFR3, GRIN1, MUC20, PRKDC, PTEN, IKBP, CDKN2A, HIPK1, PTPRV, SPN, CCAR1, PLAGL2	10.21018358
GOTERM_BP_FAT	GO:0015674~di-, tri-valent inorganic cation transport	12	2.18579235	0.008893582	ATP2B2, PLN, GRIN1, CACNG6, CACNB3, GRIN3A, TCN2, CACNA1D,	11.10233985

Figure 18 (supplemental tables continued)

FIG. 18N

					CACNA2D2, CAMK2A, ITPR1, TPCN2	
GOTERM_BP_FAT	GO:0046164~alcohol catabolic process	7	1,275045537	0,010813172	GNPDA1, MAOA, PKLR, HK1, ENO4, DLAT, PDHB	16,88645728
GOTERM_BP_FAT	GO:0043065~positive regulation of apoptosis	15	2,732240437	0,012670397	BID, MUC2, BCLAF1, FGFR3, GRIN1, MUC20, PRKDC, PTEN, IKBIP, CDKN2A, HIPK1, PITPRV, SPN, CCAR1, PLAGL2	19,50171132
GOTERM_BP_FAT	GO:0050767~regulation of neurogenesis	10	1,821493825	0,014040692	KLK8, EYA1, NTF3, SEMA4F, PLXNB2, SIX1, GRIN1, LIG4, NTN1, BMPR1A	21,38142364
GOTERM_BP_FAT	GO:0033673~negative regulation of kinase activity	6	1,092896175	0,015270158	IRAK3, CDKN2A, NF2, PKIG, ZFYVE28, IL1B	23,03271818
GOTERM_BP_FAT	GO:0006469~negative regulation of protein kinase activity	6	1,092896175	0,015270158	IRAK3, CDKN2A, NF2, PKIG, ZFYVE28, IL1B	23,03271818
GOTERM_BP_FAT	GO:0060284~regulation of cell development	11	2,003642987	0,016715538	KLK8, EYA1, CDKN2A, NTF3, SEMA4F, PLXNB2, SIX1, GRIN1, LIG4, NTN1, BMPR1A	24,93224899
GOTERM_BP_FAT	GO:0030334~regulation of cell migration	8	1,4571949	0,016735505	MUC2, EDN3, IRS2, PDPN, RRAS2, UNC5C, PTEN, NTN1	24,95817935
GOTERM_BP_FAT	GO:0042110~T cell activation	9	1,638344262	0,018795072	ITGAX, THEMIS, RPL22, PRKDC, LIG4, NCOR1, LCP1, SPN, CD1D2	27,58792023
GOTERM_BP_FAT	GO:0046365~monosaccharide catabolic process	6	1,092896175	0,019178777	GNPDA1, PKLR, HK1, ENO4, DLAT, PDHB	28,06818296

Figure 18 (supplemental tables continued)

FIG. 180

GOTERM_BP_FAT	GO:0006366~transcription from RNA polymerase II promoter	7	1,275045537	0,019584752	SCAF1, TAF13, MED16, THRAP3, MED24, POLR2D, POLR2C	28,57301545
GOTERM_BP_FAT	GO:0080135~regulation of cellular response to stress	7	1,275045537	0,020789881	KLK8, EYA1, EIF2AK1, CDKN2A, IL1B, MTOR, TLR6	30,05208194
GOTERM_BP_FAT	GO:0030217~T cell differentiation	7	1,275045537	0,022044262	THEMIS, RPL22, PRKDC, LIG4, NCOR1, SPN, CD1D2	31,56095604
GOTERM_BP_FAT	GO:0007229~integrin-mediated signaling pathway	7	1,275045537	0,022044262	VAV3, ITGAX, ITGB8, ADAMTS13, ADAM23, ITGAE, ITGA4	31,56095604
GOTERM_BP_FAT	GO:0031344~regulation of cell projection organization	6	1,092898175	0,023682897	KLK8, SEMA4F, PLXNB2, GRIN1, MTOR, NTN1	33,49762517
GOTERM_BP_FAT	GO:0008351~transcription, DNA-dependent	9	1,639344282	0,024607005	SCAF1, TAF13, CDKN2A, MED16, THRAP3, MED24, TCEA2, POLR2D, POLR2C	34,54890995
GOTERM_BP_FAT	GO:0050770~regulation of axonogenesis	5	0,910746812	0,025125899	KLK8, SEMA4F, PLXNB2, GRIN1, NTN1	35,1387561
GOTERM_BP_FAT	GO:0051241~negative regulation of multicellular organismal process	8	1,4571949	0,025220931	IRAK3, KLK8, GRM2, PLN, BMPR2, NOS3, CXADR, SOX9	35,24623903
GOTERM_BP_FAT	GO:0051960~regulation of nervous system development	10	1,821493825	0,027314722	KLK8, EYA1, NTF3, SEMA4F, PLXNB2, SIX1, GRN1, LIG4, NTN1, BMPR1A	37,57224802
GOTERM_BP_FAT	GO:0007507~heart development	13	2,387941712	0,027755837	MYL2, PRKDC, ITGA4, SOX9, CXADR, PTEN, ECE2, HAND2, PLN, VCAN, MKL2, ENG, BMPR1A	38,05215168

Figure 18 (supplemental tables continued)

FIG. 18P

GOTERM_BP_FAT	GO:0048514-blood vessel morphogenesis	12	2,18579235	0,027999749	SEMA5A, HTATIP2, HAND2, IL1B, NOS3, ADRA2B, ITGA4, MKL2, PTEN, ENG, PNPLA8, CITED1	38,31601749
GOTERM_BP_FAT	GO:0044275-cellular carbohydrate catabolic process	6	1,092896175	0,02884128	GNPDA1, PKLR, HK1, ENO4, DLAT, PDHB	39,21830561
GOTERM_BP_FAT	GO:0051338-regulation of transferase activity	12	2,18579235	0,028946587	IRAK3, CDKN2A, VAV3, NF2, WRAP53, PKIG, ZFYVE28, MUC20, IL1B, MTOR, TLR6, OAS1D	39,33030447
GOTERM_BP_FAT	GO:0018052-carbohydrate catabolic process	7	1,275045537	0,029082118	GNPDA1, LYG2, PKLR, HK1, ENO4, DLAT, PDHB	39,47422129
GOTERM_BP_FAT	GO:0032774-RNA biosynthetic process	9	1,639344262	0,029117697	SCAF1, TAF13, CDKN2A, MED16, THRAP3, MED24, TCEA2, POLR2D, POLR2C	39,51194554
GOTERM_BP_FAT	GO:0030001-metal ion transport	21	3,825136612	0,030842264	SLC9A7, SCN3A, SLC39A10, SLC20A2, CACNG8, GRIN1, CACNB3, GRIN3A, TCN2, KCTD5, CACNA2D2, ITPR1, CNGA2, TPCN2, KCNJ13, ATP2B2, PLN, SLC13A1, CP, CACNA1D, CAMK2A	41,31409014
GOTERM_BP_FAT	GO:0016486-peptide hormone processing	3	0,548448087	0,031899318	PCSK2, ECE2, CHST8	42,39358432
GOTERM_BP_FAT	GO:0051270-regulation of cell motion	8	1,4571949	0,034736172	MUC2, EDN3, IRS2, PDPN, RRAS2, UNC5C,	45,18901867

Figure 18 (supplemental tables continued)

FIG. 18Q

GOTERM_BP_FAT	GO:0051989-regulation of transmission of nerve impulse	8	1.4571949	0.034736172	PTEN, NTN1, ATP2B2, KLK8, GRM2, NTF3, P2RX3, GRIN1, CACNA2D2, CAMK2A	45.19901867
GOTERM_BP_FAT	GO:0007155-cell adhesion	25	4.553734062	0.034933395	CLDN17, THRA, CLDN8, ITGAE, ASTN1, SOX9, CXADR, PCDH1, WISP2, ITGAX, ITGB8, KLRAB, ADAM23, PDPN, CPXM2, ITGA4, BTBD9, NCAM2, SIGLEC5, PKP4, CLDN1, DSC3, VCAN, ENG, CDH11	45.38919728
GOTERM_BP_FAT	GO:0022610-biological adhesion	25	4.553734062	0.035968976	CLDN17, THRA, CLDN8, ITGAE, ASTN1, SOX9, CXADR, PCDH1, WISP2, ITGAX, ITGB8, KLRAB, ADAM23, PDPN, CPXM2, ITGA4, BTBD9, NCAM2, SIGLEC5, PKP4, CLDN1, DSC3, VCAN, ENG, CDH11	46.37763707
GOTERM_BP_FAT	GO:0006096-glycolysis	5	0.910746812	0.037114558	PKLR, HK1, ENO4, DLAT, PDHB	47.4514436
GOTERM_BP_FAT	GO:0042127-regulation of cell proliferation	24	4.371584689	0.038127944	EDN3, MUC2, IRS2, FGFR3, NF2, STK11, JARID2, CD276, LIG4, FGF21, SOX9, CXADR, PTEN, NTN1, POLD4, CDKN2A, EIF2AK1, CDKN2B, HIPK1	48.38443487

Figure 18 (supplemental tables continued)

FIG. 18R

					PTPRV, CDKN2C, NOS3, SPN, BMPR1A	
GOTERM_BP_FAT	GO:0001525~angiogenesis	9	1,639344262	0,038326179	SEMA5A, HTATIP2, HAND2, IL1B, NOS3, ADRA2B, PTEN, ENG, PNPLA6	48,56511004
GOTERM_BP_FAT	GO:0040012~regulation of locomotion	8	1,4571949	0,039456659	MUC2, EDN3, IRS2, PDPN, RRAS2, UNC5C, PTEN, NTN1	49,58403921
GOTERM_BP_FAT	GO:0050680~negative regulation of epithelial cell proliferation	4	0,72859745	0,041501972	CDKN2A, FGFR3, STK11, PTEN	51,37966407
GOTERM_BP_FAT	GO:0006812~cation transport	23	4,189435337	0,04279551	SLC9A7, SCN3A, SLC39A10, SLC20A2, CACNG8, GRIN1, CACNB3, GRIN3A, TCN2, KCTD5, CACNA2D2, ITPR1, CNGA2, TPCN2, KCNJ13, ATP6V1C1, ATP2B2, PLN, P2RX3, SLC13A1, CP, CACNA1D, CAMK2A	52,48397615
GOTERM_BP_FAT	GO:0031644~regulation of neurological system process	8	1,4571949	0,044578002	ATP2B2, KLRK8, GRM2, NTF3, P2RX3, GRIN1, CACNA2D2, CAMK2A	53,96708119
GOTERM_BP_FAT	GO:0051094~positive regulation of developmental process	12	2,18579235	0,04509441	LPL, CDKN2A, FGFR3, NTF3, PLXNB2, CD276, IL1B, NOS3, LIG4, MKL2, NTN1, CD1D2	54,38654496
GOTERM_BP_FAT	GO:0014031~mesenchymal cell development	5	0,910746812	0,045639308	EDN3, NRTN, HAND2, SOX9, ALX1	54,82932021

Figure 18 (supplemental tables continued)

FIG. 18S

GOTERM_BP_FAT	GO:0010875--regulation of neuron projection development	5	0.910746812	0.045639308	KLK8, SEMA4F, PLXNB2, GRIN1, NTN1	54,82932021
GOTERM_BP_FAT	GO:0048762--mesenchymal cell differentiation	5	0.910746812	0.0518785	EDN3, NRTN, HAND2, SOX9, ALX1	59,59870644
GOTERM_BP_FAT	GO:0002237--response to molecule of bacterial origin	5	0.910746812	0.0518785	IRAK3, TAP2, MAPK3, IL1B, 1100001G20RIK	59,59870844
GOTERM_BP_FAT	GO:0019220--regulation of phosphate metabolic process	15	2.732240437	0.052596226	FGFR3, VAV3, NF2, MUC20, PKIG, BMPR2, TLR8, IRAK3, CDKN2A, EIF2AK1, PPP1R2, ZFYVE28, IL1B, PPP2R4, MTOR	60,11587427
GOTERM_BP_FAT	GO:0051174--regulation of phosphorus metabolic process	15	2.732240437	0.052596226	FGFR3, VAV3, NF2, MUC20, PKIG, BMPR2, TLR8, IRAK3, CDKN2A, EIF2AK1, PPP1R2, ZFYVE28, IL1B, PPP2R4, MTOR	60,11587427
GOTERM_BP_FAT	GO:0010769--regulation of cell morphogenesis involved in differentiation	5	0.910746812	0.055163919	KLK8, SEMA4F, PLXNB2, GRIN1, NTN1	61,91549855
GOTERM_BP_FAT	GO:0008894--regulation of catabolic process	5	0.910746812	0.055163919	IRAK3, IL1B, CIDEA, MTOR, PIM2	61,91549855
GOTERM_BP_FAT	GO:0060485--mesenchyme development	5	0.910746812	0.055163919	EDN3, NRTN, HAND2, SOX9, ALX1	61,91549855
GOTERM_BP_FAT	GO:0019318--hexose metabolic process	10	1.821483625	0.055828001	PTGES3, PPP1R2, GNPDA1, PPP1R1A, PFKFB2, PKLR, HK1, ENO4, DLAT, PDHB	62,23249293
GOTERM_BP_FAT	GO:0001944--vasculature	13	2.367941712	0.05740593	HTATIP2, PDPN	63,42408542

Figure 18 (supplemental tables continued)

FIG. 18T

	development				ITGA4, PTEN, CITED1, PNPLA8, SEMA5A, HAND2, IL1B, NOS3, ADRA2B, MKL2, ENG	
GOTERM_BP_FAT	GO:0008892~prostanoid metabolic process	3	0.546448087	0.059828326	PTGES3, PTGES2, PDPN	64,99074558
GOTERM_BP_FAT	GO:0008893~prostaglandin metabolic process	3	0.546448087	0.059828326	PTGES3, PTGES2, PDPN	64,99074558
GOTERM_BP_FAT	GO:0008219~cell death	22	4.007285974	0.060294584	BID, MUC2, KIK8, HTATIP2, PRKDC, CIDEA, GAN, LIG4, PIM2, PEA15A, PTEN, NTN1, TAX1BP1, IRAK3, TNFRSF11B, CDKN2A, UNC5A, HAND2, UNC5C, PDCD5, DEDD2, CCAR1	65,28495056
GOTERM_BP_FAT	GO:0022604~regulation of cell morphogenesis	7	1.275045537	0.0608751	KLK8, PDPN, SEMA4F, PLXNB2, EPB4.2, GRIN1, NTN1	65,64789983
GOTERM_BP_FAT	GO:0019320~hexose catabolic process	5	0.910748812	0.062083417	PKLR, HK1, ENO4, DLAT, PDHB	66,38004654
GOTERM_BP_FAT	GO:0006007~glucose catabolic process	5	0.910748812	0.062083417	PKLR, HK1, ENO4, DLAT, PDHB	66,38004654
GOTERM_BP_FAT	GO:0033077~T cell differentiation in the thymus	4	0.72859745	0.063423542	PRKDC, LIG4, SPN, CD1D2	67,19990596
GOTERM_BP_FAT	GO:0006928~cell motion	17	3.096539182	0.084708426	EDN3, NRTN, SMOK2B, VAV3, NTF3, PDPN, ASTN1, PRKDC, ITGA4, PTEN, NTN1, TNP2, SEMA5A, SEMA4F, LHX3, SIK1, IL1B	67,95709491
GOTERM_BP_FAT	GO:0007389~pattern	14	2.550091075	0.068616779	BMPR2, DLL3,	69,05139243

Figure 18 (supplemental tables continued)

FIG. 18U

	specification process				PRKDC, SEMA5A, CTNNBIP1, EYA1, HOXD9, HIPK1, SFRP1, LHX3, SIX1, ENG, ALX1, BMPR1A	
GOTERM_BP_FAT	GO:0007517~muscle organ development	10	1.821493825	0.0681874	MYL2, PLN, SIX1, MKL2, CXADR, TRIM72, PTEN, CACNA2D2, ENG, TAGLN3	69.9310274
GOTERM_BP_FAT	GO:0014033~neural crest cell differentiation	4	0.72859745	0.068336688	EDN3, NRTN, HAND2, SOX9	70.00740394
GOTERM_BP_FAT	GO:0014032~neural crest cell development	4	0.72859745	0.068336688	EDN3, NRTN, HAND2, SOX9	70.00740394
GOTERM_BP_FAT	GO:0050804~regulation of synaptic transmission	7	1.275045537	0.068486872	ATP2B2, GRM2, NTF3, P2RX3, GRIN1, CACNA2D2, CAMK2A	70.08955099
GOTERM_BP_FAT	GO:0006890~icosanoid metabolic process	4	0.72859745	0.073417103	PTGES3, GGT5, PTGES2, PDPN	72.87162064
GOTERM_BP_FAT	GO:0016265~death	22	4.007285974	0.073478734	BID, MUC2, KLK8, HTATIP2, PRKDC, CIDEA, GAN, LIG4, PIM2, PEA15A, PTEN, NTN1, TAX1BP1, IRAK3, TNFRSF11B, CDKN2A, UNC5A, HAND2, UNC5C, PDCD5, DEDD2, CCAR1	72.7025284
GOTERM_BP_FAT	GO:0045684~regulation of neuron differentiation	7	1.275045537	0.073854582	KLK8, EYA1, SEMA4F, PLXNB2, SIX1, GRIN1, NTN1	72.8903048
GOTERM_BP_FAT	GO:0009952~anterior/posterior pattern formation	9	1.639344262	0.074691198	CTNNBIP1, HOXD8, SFRP1, HIPK1, DLL3, BMPR2, PRKDC, BMPR1A, ALX1	73.30392897
GOTERM_BP_FAT	GO:0050771~negative	3	0.546448087	0.076001207	KLK8, SEMA4F	73.9396568

Figure 18 (supplemental tables continued)

FIG. 18V

	regulation of axonogenesis				NTN1	
GOTERM_BP_FAT	GO:0010627~regulation of protein kinase cascade	9	1,839344282	0,079185774	FGFR3, NF2, STK11, IL1B, FGF21, PIM2, TLR6, PTEN, NCOR1	75,42626531
GOTERM_BP_FAT	GO:0006915~apoptosis	20	3,64288725	0,080855939	BID, MUC2, HTATIP2, PRKDC, CIDEA, LIG4, PEA15A, PIM2, PTEN, NTN1, TAX1BP1, IRAK3, TNFRSF11B, CDKN2A, UNC5A, HAND2, UNC5C, PDCD5, DEDD2, CCAR1	76,08527945
GOTERM_BP_FAT	GO:0048738~cardiac muscle tissue development	5	0,910746812	0,081187313	MYL2, PLN, MKL2, CXADR, PTEN	76,31935213
GOTERM_BP_FAT	GO:0009617~response to bacterium	9	1,639344282	0,083838455	MBL1, IRAK3, DEFB15, TAP2, MAPK3, IL1B, WFDC15A, 1100001G20RK, SPN	77,45532448
GOTERM_BP_FAT	GO:0002568~somatic diversification of T cell receptor genes	2	0,364298725	0,083871921	PRKDC, LIG4	77,46933088
GOTERM_BP_FAT	GO:0033153~T cell receptor V(D)J recombination	2	0,364298725	0,083871921	PRKDC, LIG4	77,46933088
GOTERM_BP_FAT	GO:0002681~somatic recombination of T cell receptor gene segments	2	0,364298725	0,083871921	PRKDC, LIG4	77,46933088
GOTERM_BP_FAT	GO:0033559~unsaturated fatty acid metabolic process	4	0,72859745	0,084062516	PTGES3, GGT5, PTGES2, PDPN	77,54894247
GOTERM_BP_FAT	GO:0018149~peptide cross-linking	3	0,546448087	0,084546704	SPOCK2, F13A1, EPB4.2	77,74999986
GOTERM_BP_FAT	GO:0051180~vitamin transport	3	0,546448087	0,084546704	PDPN, SLC2A3, TCN2	77,74999986
GOTERM_BP_FAT	GO:0010648~negative regulation of cell communication	10	1,821493625	0,089019308	CTNNBIP1, IRAK3, GRM2, NF2, ZFYVE28, IL1B, CIDEA	79,52877711

Figure 18 (supplemental tables continued)

FIG. 18W

					MTOR, PTEN, NCOR1	
GOTERM_BP_FAT	GO:0007498~mesoderm development	5	0.910746812	0.089562248	AMH, NF2, DLL3, BMPR2, BMPR1A	79,73535655
GOTERM_BP_FAT	GO:0012501~programmed cell death	20	3,64298725	0.090105783	BID, MUC2, HTATIP2, PRKDC, CIDEA, LIG4, PEA15A, PIM2, PTEN, NTN1, TAX1BP1, IRAK3, TNFRSF11B, CDKN2A, UNC5A, HAND2, UNC5C, PDCD5, DEDD2, CCAR1	79,94020063
GOTERM_BP_FAT	GO:0002709~regulation of T cell mediated immunity	3	0.546448087	0.093368037	TAP2, SPN, CD1D2	81,12854892
GOTERM_BP_FAT	GO:0009895~negative regulation of catabolic process	3	0.546448087	0.093368037	IRAK3, CIDEA, MTOR	81,12854892
GOTERM_BP_FAT	GO:0001568~blood vessel development	12	2.18579235	0.09381057	SEMA5A, HTATIP2, HAND2, IL1B, NOS3, ADRA2B, ITGA4, MKL2, PTEN, ENG, PNPLA8, CITED1	81,28534875
GOTERM_BP_FAT	GO:0050768~negative regulation of neurogenesis	4	0.72859745	0.095323828	KLK8, SEMA4F, NTN1, BMPR1A	81,80997544
GOTERM_BP_FAT	GO:0032496~response to lipopolysaccharide	4	0.72859745	0.095323828	IRAK3, MAPK3, IL1B, 1100001G20RIK	81,80997544
GOTERM_BP_FAT	GO:0060537~muscle tissue development	8	1.4571949	0.097754734	MYL2, PLN, SIX1, MKL2, CXADR, PTEN, CACNA2D2, ENG	82,62385976

Figure 18 (supplemental tables continued)

FIG. 18X

primers

Sequence Name	Bases	Sequence
HK112	91	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTAAGTAGAGTCTTGTGAAAGGACGAAACA CCG
HK113	92	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTATACACGATCTCTTGTGAAAGGACGAAAC ACCG
HK114	93	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGATCGCGCGGTTCTTGTGAAAGGACGAAA CACCG
HK115	94	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCGATCATGATCGTCTTGTGAAAGGACGAA ACACCG
HK116	95	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCGATCGTTACCATCTTGTGAAAGGACGAA AACACCG
HK117	96	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTATCGATTCTTGGTTCTTGTGAAAGGACG AAACACCG
HK118	97	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGATCGATAACGCATTCTTGTGAAAGGAC GAAACACCG
HK119	98	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCGATCGATACAGGTATTCTTGTGAAAGGA CGAAACACCG
HK120	99	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACGATCGATAGGTAAGGTCTTGTGAAAGG ACGAAACACCG
HK121	91	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTAACAATGGTCTTGTGAAAGGACGAAACA CCG
HK122	92	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTATACTGTATCTCTTGTGAAAGGACGAAACA CCG
HK123	93	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGATAGGTGCGATCTTGTGAAAGGACGAAA CACCG
HK086	92	CAAGCAGAAGACGGCATACGAGATAAGTAGAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTTCTACTATTCTTTCCCTGCA CTGT
HK087	93	CAAGCAGAAGACGGCATACGAGATACGATCGTGAAGTTCAGACGTGTGCTCTTCCGATCTATTCTACTATTCTTTCCCTGCA ACTGT
HK088	94	CAAGCAGAAGACGGCATACGAGATCGCGCGGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATTCTACTATTCTTTCCCTG CACTGT
HK089	95	CAAGCAGAAGACGGCATACGAGATCATGATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGATTCTACTATTCTTTCCCT GCACTGT
HK090	96	CAAGCAGAAGACGGCATACGAGATCGTTACCAAGTGAAGTTCAGACGTGTGCTCTTCCGATCTCGATCTCTACTATTCTTTCCCT TGCACTGT
HK091	92	CAAGCAGAAGACGGCATACGAGATTCTTGGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTTCTACTATTCTTTCCCTGCA TGT
HK112	91	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTAAGTAGAGTCTTGTGAAAGGACGAAACA CCG
HK113	92	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTATACACGATCTCTTGTGAAAGGACGAAAC

Figure 18 (supplemental tables continued)

FIG. 18Y

		ACCG
HK092	93	CAAGCAGAAGACGGCATACGAGATAACGCATTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATTCTACTATTCTTTCCCTGCACTGT
HK093	94	CAAGCAGAAGACGGCATACGAGATACAGGTATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATTCTACTATTCTTTCCCTGCACTGT
HK094	95	CAAGCAGAAGACGGCATACGAGATAGGTAAGGGTACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGATTCTACTATTCTTTCCCTTGCACACTGT
HK095	96	CAAGCAGAAGACGGCATACGAGATAACAATGGGTACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGATCTCTACTATTCTTTCCCTTGCACACTGT
HK096	92	CAAGCAGAAGACGGCATACGAGATACTGTATCGTACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
HK097	93	CAAGCAGAAGACGGCATACGAGATAGGTCGAGTACTGGAGTTCAGACGTGTGCTCTTCCGATCTATTCTACTATTCTTTCCCTGCACTGT

Single SgRNAs

Spacer No.	Spacer 20bp	PAM	Target
S001	GATCCGCACCGTGAATGTCT	NGG	Nf2
S002	GAAGCTCATGCCGAGAAGCGA	NGG	Nf2
S003	CTTGGCGTCATATGCTGTCC	NGG	Nf2
S004	CACTGGGGCTTCGGGAAACC	NGG	Nf2
S005	TGGACTGCAGTATACAATCA	NGG	Nf2
S006	GTTGGATCATGATGTTTCCA	NGG	Nf2
S007	ACCGCCAAATTTAACTGCAG	NGG	Pten
S008	GCAGCAATTCACGTAAAGC	NGG	Pten
S009	TGTCATCTTCACTAGCCAT	NGG	Pten
S010	ACAATATTGATGATGTAGTA	NGG	Pten
S011	CATACCTCTGCAGTTAAAT	NGG	Pten
S012	AATCCCATAGCAATAATTT	NGG	Pten
S013	GCTTGACAGCTCCCGACGCA	NGG	Trim72
S014	ATGCTGGGGTTGCCTTGCGT	NGG	Trim72
S015	ATGTTCCGGGCTCTGATGCC	NGG	Trim72
S016	GCGCAAGGAGAAGACTGTAG	NGG	Trim72
S017	CCTGTCAGAGTCACCAACAC	NGG	Trim72
S018	TTGCTCACCTGCTGGTTACC	NGG	Trim72
S019	GTGCGATATTGCGTTCCGC	NGG	Cdkn2a
S020	CCCAACGCCCGAACTCTTT	NGG	Cdkn2a
S021	GGGGTACGACCGAAAGAGTT	NGG	Cdkn2a
S022	GGCTGGATGTGCGCGATGCC	NGG	Cdkn2a
S023	TCGTGCGATCCCGAGACCC	NGG	Cdkn2a
S024	CCGCACCGGAATCCTGGACC	NGG	Cdkn2a

Figure 18 (supplemental tables continued)

FIG. 18Z

S025	AGAGGGGCGATTACCGGGC	NGG	Cryba4
S026	TGTAGAACCACCGCGACTCA	NGG	Cryba4
S027	AAGGAGGTGAGCCTCTCCGC	NGG	Cryba4
S028	CTTACGCCCCAGATTGAACA	NGG	Cryba4
S029	GACAGACTCACGCTCCGCTC	NGG	Cryba4
S030	AAGAAGGCTTCCAGGGCCGA	NGG	Cryba4
S031	GTTGATTGACGAAGCCAACC	NGG	Fga
S032	GTTGATAGACATGAAGCGCC	NGG	Fga
S033	AGTACGTGGCCCAAGAGTTG	NGG	Fga
S034	GGCGCAGAAATTGAGATCCTG	NGG	Fga
S035	CACCTGCCTCATCTTGAGCG	NGG	Fga
S036	ACTACAGATACCGAAGATAA	NGG	Fga
S037	GCTACTGGGCCCTGAAGTAG	NGG	mmu-mir-345
S038	CCAGTGCTTGTGGTGCTAC	NGG	mmu-mir-345
S039	GACCCCTAGTCCAGTGCTTG	NGG	mmu-mir-345
S040	TGAAGCCGACCCCAAGTCC	NGG	mmu-mir-345
S041	AGTCAGTGCATGACAGAACT	NGG	mmu-mir-152
S042	ATACACTCCGACTCGGGCTC	NGG	mmu-mir-152
S043	TTCTGTGATACACTCCGACT	NGG	mmu-mir-152
S044	CGCCGCTGTTCGCCGGGCCT	NGG	mmu-mir-152
S045	GCGAGGTATTCGGCTCCGCG	NGG	NonTargetingControlGuideForMouse_0001
S046	GCTTTCACGGAGGTTTCGACG	NGG	NonTargetingControlGuideForMouse_0002
S047	ATGTTGCAGTTCCGGCTCGAT	NGG	NonTargetingControlGuideForMouse_0003
S048	ACGTGTAAGGCGAAGGCCTT	NGG	NonTargetingControlGuideForMouse_0004

Figure 18 (supplemental tables continued)

Mouse Valid guides

long_name	uid	gene	seq	Pool name	spacer num	display
MouseValid_MetsMIR_001_Adra2b	MetsMIR_001	Adra2b	TACCAGCGCATTG CCGAAAA	Mouse Valid	001	Adra2b_001
MouseValid_MetsMIR_002_Adra2b	MetsMIR_002	Adra2b	GAACTCCCCCGG CGGC CGGA	Mouse Valid	002	Adra2b_002
MouseValid_MetsMIR_003_Adra2b	MetsMIR_003	Adra2b	ACTTCCGCCGTGC CTTTCGA	Mouse Valid	003	Adra2b_003
MouseValid_MetsMIR_004_Adra2b	MetsMIR_004	Adra2b	TGGCAATTACATA GATTCGC	Mouse Valid	004	Adra2b_004
MouseValid_MetsMIR_005_Adra2b	MetsMIR_005	Adra2b	AGAAGAGCACGTC TAGCGCC	Mouse Valid	005	Adra2b_005
MouseValid_MetsMIR_006_Adra2b	MetsMIR_006	Adra2b	CCGGCAAAGGAT CCTTCGAA	Mouse Valid	006	Adra2b_006

FIG. 18AA

MouseValid_MetsMIR_007_Adra2b	MetsMi R_007	Adra2b	CCAATAGCCCAGC AGCTCGT	Mouse Valid	007	Adra2b_007
MouseValid_MetsMIR_008_Adra2b	MetsMi R_008	Adra2b	AATGACCACCGCC AGCACAA	Mouse Valid	008	Adra2b_008
MouseValid_MetsMIR_009_Adra2b	MetsMi R_009	Adra2b	GGACATGTCCGG CCCCGCCA	Mouse Valid	009	Adra2b_009
MouseValid_MetsMIR_010_Adra2b	MetsMi R_010	Adra2b	AGCAAAAAAAGAT CCGATGC	Mouse Valid	010	Adra2b_010
MouseValid_MetsMIR_011_Aprt	MetsMi R_011	Aprt	GTGATCTTGCCG CTGTGCG	Mouse Valid	011	Aprt_011
MouseValid_MetsMIR_012_Aprt	MetsMi R_012	Aprt	CCGAGCTTCCATC CGCCTCT	Mouse Valid	012	Aprt_012
MouseValid_MetsMIR_013_Aprt	MetsMi R_013	Aprt	CCCTAACAGGTCT AGACTCC	Mouse Valid	013	Aprt_013
MouseValid_MetsMIR_014_Aprt	MetsMi R_014	Aprt	AGCTGACCTCGCT GAAGGGC	Mouse Valid	014	Aprt_014
MouseValid_MetsMIR_015_Aprt	MetsMi R_015	Aprt	CCGGCAGCTTCC CCTGTTTC	Mouse Valid	015	Aprt_015
MouseValid_MetsMIR_016_Aprt	MetsMi R_016	Aprt	CACACTCCACCAC TTCAGCC	Mouse Valid	016	Aprt_016
MouseValid_MetsMIR_017_Aprt	MetsMi R_017	Aprt	CGGGCAGAGAGT GGTCATG	Mouse Valid	017	Aprt_017
MouseValid_MetsMIR_018_Aprt	MetsMi R_018	Aprt	CACAATGACCACT CTCTGCC	Mouse Valid	018	Aprt_018
MouseValid_MetsMIR_019_Aprt	MetsMi R_019	Aprt	CTCGCCCTCTTG AAAGACC	Mouse Valid	019	Aprt_019
MouseValid_MetsMIR_020_Aprt	MetsMi R_020	Aprt	GCGTGGGCTGTG TGCTCATC	Mouse Valid	020	Aprt_020
MouseValid_MetsMIR_021_Atp5s	MetsMi R_021	Atp5s	GGAGCCAAAGTAC GCTACTG	Mouse Valid	021	Atp5s_021
MouseValid_MetsMIR_022_Atp5s	MetsMi R_022	Atp5s	TGCCACCGATTCT TGTATCA	Mouse Valid	022	Atp5s_022
MouseValid_MetsMIR_023_Atp5s	MetsMi R_023	Atp5s	TCATTGCTCTGCG ACATTTC	Mouse Valid	023	Atp5s_023
MouseValid_MetsMIR_024_Atp5s	MetsMi R_024	Atp5s	AAATTTACGAAAA AGCCTAT	Mouse Valid	024	Atp5s_024
MouseValid_MetsMIR_025_Atp5s	MetsMi R_025	Atp5s	TCTAGATCAATGA TGATGTT	Mouse Valid	025	Atp5s_025
MouseValid_MetsMIR_026_Atp5s	MetsMi R_026	Atp5s	CAACCATCCCCAG AAGTATC	Mouse Valid	026	Atp5s_026
MouseValid_MetsMIR_027_Atp5s	MetsMi R_027	Atp5s	CAGTTAAGAACAA ATACTTG	Mouse Valid	027	Atp5s_027
MouseValid_MetsMIR_028_Atp5s	MetsMi R_028	Atp5s	CAGTGCTGTCTTA AAGACTT	Mouse Valid	028	Atp5s_028

Figure 18 (supplemental tables continued)

FIG. 18AB

MouseValid_MetsMIR_029_Atp5s	MetsMi R_029	Atp5s	GAGTAGATTATGA ACGCCTC	Mouse Valid	029	Atp5s_029
MouseValid_MetsMIR_030_Atp5s	MetsMi R_030	Atp5s	TATAGTCGTGTAG CCACTTC	Mouse Valid	030	Atp5s_030
MouseValid_MetsMIR_031_Babam1	MetsMi R_031	Babam1	TAGTGAATGACGA CTCCGCC	Mouse Valid	031	Babam1_031
MouseValid_MetsMIR_032_Babam1	MetsMi R_032	Babam1	ACGCCGTCTCCAG GTCATAC	Mouse Valid	032	Babam1_032
MouseValid_MetsMIR_033_Babam1	MetsMi R_033	Babam1	CAAGAGCCACGAA TTTGAC	Mouse Valid	033	Babam1_033
MouseValid_MetsMIR_034_Babam1	MetsMi R_034	Babam1	TATGTTTCAGATC ATCTGTC	Mouse Valid	034	Babam1_034
MouseValid_MetsMIR_035_Babam1	MetsMi R_035	Babam1	GCAGAGCTCACGT GGGTCAG	Mouse Valid	035	Babam1_035
MouseValid_MetsMIR_036_Babam1	MetsMi R_036	Babam1	GGAGATGTCTGTG CCAAAGC	Mouse Valid	036	Babam1_036
MouseValid_MetsMIR_037_Babam1	MetsMi R_037	Babam1	GCACTCACATGAG GCTGAAG	Mouse Valid	037	Babam1_037
MouseValid_MetsMIR_038_Babam1	MetsMi R_038	Babam1	CTACAACCAAGTGC AAATTCG	Mouse Valid	038	Babam1_038
MouseValid_MetsMIR_039_Babam1	MetsMi R_039	Babam1	CTGCAGCTGCCTG TATGACC	Mouse Valid	039	Babam1_039
MouseValid_MetsMIR_040_Babam1	MetsMi R_040	Babam1	CCATTAAAGGACT CCAGCTT	Mouse Valid	040	Babam1_040
MouseValid_MetsMIR_041_Bid	MetsMi R_041	Bid	GGTCCATCTCATC GCCTATT	Mouse Valid	041	Bid_041
MouseValid_MetsMIR_042_Bid	MetsMi R_042	Bid	GCCAGACATCTCG CCCAAT	Mouse Valid	042	Bid_042
MouseValid_MetsMIR_043_Bid	MetsMi R_043	Bid	CCTTAGGTACGCA ACGGTTC	Mouse Valid	043	Bid_043
MouseValid_MetsMIR_044_Bid	MetsMi R_044	Bid	CTGTGATGTGCTC GGCCCCC	Mouse Valid	044	Bid_044
MouseValid_MetsMIR_045_Bid	MetsMi R_045	Bid	TAATAAAGTTGAC AGTCGTG	Mouse Valid	045	Bid_045
MouseValid_MetsMIR_046_Bid	MetsMi R_046	Bid	TCTAACCAAGTTC CTCATAT	Mouse Valid	046	Bid_046
MouseValid_MetsMIR_047_Bid	MetsMi R_047	Bid	CTTCTTAGGAGA TGGACTG	Mouse Valid	047	Bid_047
MouseValid_MetsMIR_048_Bid	MetsMi R_048	Bid	CTGTCTGCAGCTC GTCTTCG	Mouse Valid	048	Bid_048
MouseValid_MetsMIR_049_Bid	MetsMi R_049	Bid	CTAGCCGCACAGT TCATGAA	Mouse Valid	049	Bid_049
MouseValid_MetsMIR_050_Bid	MetsMi R_050	Bid	TCATACTTACTTC CTCCGAC	Mouse Valid	050	Bid_050

Figure 18 (supplemental tables continued)

FIG. 18AC

MouseValid_MetsMIR_051_Bmp2	MetsMIR_051	Bmp2	ATGCCGCGGCGA ACTTCTTG	Mouse Valid	051	Bmp2_051
MouseValid_MetsMIR_052_Bmp2	MetsMIR_052	Bmp2	ATATGCTAGATCT GTACCGC	Mouse Valid	052	Bmp2_052
MouseValid_MetsMIR_053_Bmp2	MetsMIR_053	Bmp2	GATCAAACCTAGAA GCCGTGG	Mouse Valid	053	Bmp2_053
MouseValid_MetsMIR_054_Bmp2	MetsMIR_054	Bmp2	TTAAATTGAAGAA GAAGCGC	Mouse Valid	054	Bmp2_054
MouseValid_MetsMIR_055_Bmp2	MetsMIR_055	Bmp2	GGCACCACGAC GTCCTTGC	Mouse Valid	055	Bmp2_055
MouseValid_MetsMIR_056_Bmp2	MetsMIR_056	Bmp2	GTCTTCCGAAGGC CGGGACA	Mouse Valid	056	Bmp2_056
MouseValid_MetsMIR_057_Bmp2	MetsMIR_057	Bmp2	GCGGTACAGATCT AGCATAT	Mouse Valid	057	Bmp2_057
MouseValid_MetsMIR_058_Bmp2	MetsMIR_058	Bmp2	TGGTCYGGGGCG GGCGCTCC	Mouse Valid	058	Bmp2_058
MouseValid_MetsMIR_059_Bmp2	MetsMIR_059	Bmp2	GTCCGAGCTTCC ATCACGA	Mouse Valid	059	Bmp2_059
MouseValid_MetsMIR_060_Bmp2	MetsMIR_060	Bmp2	TTCGTGATGGAAG CTGCCGA	Mouse Valid	060	Bmp2_060
MouseValid_MetsMIR_061_Cdkn2a	MetsMIR_061	Cdkn2a	GTCCGATATTTC GTCCGC	Mouse Valid	061	Cdkn2a_061
MouseValid_MetsMIR_062_Cdkn2a	MetsMIR_062	Cdkn2a	CCCAACGCCCCG AACTCTT	Mouse Valid	062	Cdkn2a_062
MouseValid_MetsMIR_063_Cdkn2a	MetsMIR_063	Cdkn2a	GGGTACGACCG AAAGAGTT	Mouse Valid	063	Cdkn2a_063
MouseValid_MetsMIR_064_Cdkn2a	MetsMIR_064	Cdkn2a	GGCTGGATGTGC GCGATGCC	Mouse Valid	064	Cdkn2a_064
MouseValid_MetsMIR_065_Cdkn2a	MetsMIR_065	Cdkn2a	TGGTGGATCCCG GACACCC	Mouse Valid	065	Cdkn2a_065
MouseValid_MetsMIR_066_Cdkn2a	MetsMIR_066	Cdkn2a	CCGACCCGAAT CCTGGACC	Mouse Valid	066	Cdkn2a_066
MouseValid_MetsMIR_067_Cdkn2a	MetsMIR_067	Cdkn2a	GCCGGATTAGCT CTGCTCT	Mouse Valid	067	Cdkn2a_067
MouseValid_MetsMIR_068_Cdkn2a	MetsMIR_068	Cdkn2a	TGCAGGGCCCTG GAACCTCG	Mouse Valid	068	Cdkn2a_068
MouseValid_MetsMIR_069_Cdkn2a	MetsMIR_069	Cdkn2a	GGGAACCTCGCC CAGACCGA	Mouse Valid	069	Cdkn2a_069
MouseValid_MetsMIR_070_Cdkn2a	MetsMIR_070	Cdkn2a	AGGGCCGTGTGC ATGACGTG	Mouse Valid	070	Cdkn2a_070
MouseValid_MetsMIR_071_Cryba4	MetsMIR_071	Cryba4	AGAGGGGCGATT ACCCGGGC	Mouse Valid	071	Cryba4_071
MouseValid_MetsMIR_072_Cryba4	MetsMIR_072	Cryba4	TGTAGAACCACCG CGACTCA	Mouse Valid	072	Cryba4_072

Figure 18 (supplemental tables continued)

FIG. 18AD

MouseValid_MetsMIR_073_Cryba4	MetsMi R_073	Cryba4	AAGGAGGTGAGC CTCTCCGC	Mouse Valid	073	Cryba4_073
MouseValid_MetsMIR_074_Cryba4	MetsMi R_074	Cryba4	CTTACGCCCCAGA TTGAACA	Mouse Valid	074	Cryba4_074
MouseValid_MetsMIR_075_Cryba4	MetsMi R_075	Cryba4	GACAGACTCACGC TCCGCTC	Mouse Valid	075	Cryba4_075
MouseValid_MetsMIR_076_Cryba4	MetsMi R_076	Cryba4	AAGAAGGCTTCCA GGGCCGA	Mouse Valid	076	Cryba4_076
MouseValid_MetsMIR_077_Cryba4	MetsMi R_077	Cryba4	TTCTGGCCACACT CTGCACC	Mouse Valid	077	Cryba4_077
MouseValid_MetsMIR_078_Cryba4	MetsMi R_078	Cryba4	AGTGCTTGTAGTC ACCTGAG	Mouse Valid	078	Cryba4_078
MouseValid_MetsMIR_079_Cryba4	MetsMi R_079	Cryba4	CCAAGGACAGCAA TATGTGC	Mouse Valid	079	Cryba4_079
MouseValid_MetsMIR_080_Cryba4	MetsMi R_080	Cryba4	AGATGGTCAGCCT TGAGTCG	Mouse Valid	080	Cryba4_080
MouseValid_MetsMIR_081_Cyp2d12	MetsMi R_081	Cyp2d12	TTGTCTTCTCTGC GCAATTT	Mouse Valid	081	Cyp2d12_081
MouseValid_MetsMIR_082_Cyp2d12	MetsMi R_082	Cyp2d12	CCCACCAGCCTGA GTAGTGA	Mouse Valid	082	Cyp2d12_082
MouseValid_MetsMIR_083_Cyp2d12	MetsMi R_083	Cyp2d12	GTGCCCATCTTTG AGCATCT	Mouse Valid	083	Cyp2d12_083
MouseValid_MetsMIR_084_Cyp2d12	MetsMi R_084	Cyp2d12	GGTTGTGATCAAC AGAATGA	Mouse Valid	084	Cyp2d12_084
MouseValid_MetsMIR_085_Cyp2d12	MetsMi R_085	Cyp2d12	ATAACCTGTTGAC TGAAAC	Mouse Valid	085	Cyp2d12_085
MouseValid_MetsMIR_086_Cyp2d12	MetsMi R_086	Cyp2d12	TGAAGGACTTCTG ACTTTGC	Mouse Valid	086	Cyp2d12_086
MouseValid_MetsMIR_087_Cyp2d12	MetsMi R_087	Cyp2d12	GATAGTTTGACAG AACTCTC	Mouse Valid	087	Cyp2d12_087
MouseValid_MetsMIR_088_Cyp2d12	MetsMi R_088	Cyp2d12	GTACTTTTCAGCAT CCTGATG	Mouse Valid	088	Cyp2d12_088
MouseValid_MetsMIR_089_Cyp2d12	MetsMi R_089	Cyp2d12	GGTATGATCTTTG CACCTTA	Mouse Valid	089	Cyp2d12_089
MouseValid_MetsMIR_090_Cyp2d12	MetsMi R_090	Cyp2d12	CTGTGATGCCTTC ACTACTC	Mouse Valid	090	Cyp2d12_090
MouseValid_MetsMIR_091_Dennd4a	MetsMi R_091	Dennd4a	GTTTAACTCTCT GGTAGAT	Mouse Valid	091	Dennd4a_091
MouseValid_MetsMIR_092_Dennd4a	MetsMi R_092	Dennd4a	AATATTGGCAGGT CGTCCAT	Mouse Valid	092	Dennd4a_092
MouseValid_MetsMIR_093_Dennd4a	MetsMi R_093	Dennd4a	CTGTATCGTTATT CCATCTC	Mouse Valid	093	Dennd4a_093
MouseValid_MetsMIR_094_Dennd4a	MetsMi R_094	Dennd4a	AAACACGACTCAG AATACGC	Mouse Valid	094	Dennd4a_094

Figure 18 (supplemental tables continued)

FIG. 18AE

MouseValid_MetsMIR_095_Dennd4a	MetsMi R_095	Dennd4a	AACGACTCGTAAT CTTCTTG	Mouse Valid	095	Dennd4a_095
MouseValid_MetsMIR_096_Dennd4a	MetsMi R_096	Dennd4a	CGTACGCACCTTC TCAGCTG	Mouse Valid	096	Dennd4a_096
MouseValid_MetsMIR_097_Dennd4a	MetsMi R_097	Dennd4a	CTTATGCTATAAA AAGTCAG	Mouse Valid	097	Dennd4a_097
MouseValid_MetsMIR_098_Dennd4a	MetsMi R_098	Dennd4a	TCTCTCCTAGTGG GGATCTG	Mouse Valid	098	Dennd4a_098
MouseValid_MetsMIR_099_Dennd4a	MetsMi R_099	Dennd4a	CATTCAATAGTTG CCCCCAT	Mouse Valid	099	Dennd4a_099
MouseValid_MetsMIR_100_Dennd4a	MetsMi R_100	Dennd4a	AGAGGGTATTTGC TATCTGG	Mouse Valid	100	Dennd4a_100
MouseValid_MetsMIR_101_Dyrk3	MetsMi R_101	Dyrk3	ATCGTCGTAGCCA CCGTAT	Mouse Valid	101	Dyrk3_101
MouseValid_MetsMIR_102_Dyrk3	MetsMi R_102	Dyrk3	GGTTGCACCTTCAC CGCGCTC	Mouse Valid	102	Dyrk3_102
MouseValid_MetsMIR_103_Dyrk3	MetsMi R_103	Dyrk3	CGCCGTTACCTCC GCTGCTG	Mouse Valid	103	Dyrk3_103
MouseValid_MetsMIR_104_Dyrk3	MetsMi R_104	Dyrk3	CGTCCTTCCTCCC GCGATCG	Mouse Valid	104	Dyrk3_104
MouseValid_MetsMIR_105_Dyrk3	MetsMi R_105	Dyrk3	GTTGCTGAATGTG TCCGAAT	Mouse Valid	105	Dyrk3_105
MouseValid_MetsMIR_106_Dyrk3	MetsMi R_106	Dyrk3	CCTACTTACATTTA GCCTTC	Mouse Valid	106	Dyrk3_106
MouseValid_MetsMIR_107_Dyrk3	MetsMi R_107	Dyrk3	GGGCGAACCTCC GAACCAAC	Mouse Valid	107	Dyrk3_107
MouseValid_MetsMIR_108_Dyrk3	MetsMi R_108	Dyrk3	GCACCACCTCCC GTCCGTC	Mouse Valid	108	Dyrk3_108
MouseValid_MetsMIR_109_Dyrk3	MetsMi R_109	Dyrk3	ACAGCACCCCGAT CGACATA	Mouse Valid	109	Dyrk3_109
MouseValid_MetsMIR_110_Dyrk3	MetsMi R_110	Dyrk3	TCTCCTCGGCCGC CTGGCGA	Mouse Valid	110	Dyrk3_110
MouseValid_MetsMIR_111_Fam110b	MetsMi R_111	Fam110b	CTCACGTTGCACA CCGCTCT	Mouse Valid	111	Fam110b_111
MouseValid_MetsMIR_112_Fam110b	MetsMi R_112	Fam110b	AGTAGTCGGYTCC TTTGTTT	Mouse Valid	112	Fam110b_112
MouseValid_MetsMIR_113_Fam110b	MetsMi R_113	Fam110b	CAGGGGCACCGC CGACGTGA	Mouse Valid	113	Fam110b_113
MouseValid_MetsMIR_114_Fam110b	MetsMi R_114	Fam110b	ACCCGACTACTTC CGCAGGC	Mouse Valid	114	Fam110b_114
MouseValid_MetsMIR_115_Fam110b	MetsMi R_115	Fam110b	CTCCTGGCTCTTG ACGTATT	Mouse Valid	115	Fam110b_115
MouseValid_MetsMIR_116_Fam110b	MetsMi R_116	Fam110b	GGCTTTAAGGGG CTTCACAC	Mouse Valid	116	Fam110b_116

Figure 18 (supplemental tables continued)

FIG. 18AF

MouseValid_MetsMIR_117_Fam110b	MetsMi R_117	Fam110b	ACAAGCATAGCTC CGAAAC	Mouse Valid	117	Fam110b_117
MouseValid_MetsMIR_118_Fam110b	MetsMi R_118	Fam110b	CAACTACTGCCGA CTTGACC	Mouse Valid	118	Fam110b_118
MouseValid_MetsMIR_119_Fam110b	MetsMi R_119	Fam110b	AGAGCAGCTCCCA CGTGAGC	Mouse Valid	119	Fam110b_119
MouseValid_MetsMIR_120_Fam110b	MetsMi R_120	Fam110b	AGCTGATCATGCT TGCCTGT	Mouse Valid	120	Fam110b_120
MouseValid_MetsMIR_121_Fga	MetsMi R_121	Fga	GTTGATTGACGAA GCCAACC	Mouse Valid	121	Fga_121
MouseValid_MetsMIR_122_Fga	MetsMi R_122	Fga	GTTGATAGACATG AAGCGCC	Mouse Valid	122	Fga_122
MouseValid_MetsMIR_123_Fga	MetsMi R_123	Fga	AGTACGTGGCCCA AGAGTTG	Mouse Valid	123	Fga_123
MouseValid_MetsMIR_124_Fga	MetsMi R_124	Fga	GGCGCAGAATTGA GATCCTG	Mouse Valid	124	Fga_124
MouseValid_MetsMIR_125_Fga	MetsMi R_125	Fga	CACCTGCCTCATC TTGAGCG	Mouse Valid	125	Fga_125
MouseValid_MetsMIR_126_Fga	MetsMi R_126	Fga	ACTACAGATACCG AAGATAA	Mouse Valid	126	Fga_126
MouseValid_MetsMIR_127_Fga	MetsMi R_127	Fga	AGCACATACCCAG ACTGTGC	Mouse Valid	127	Fga_127
MouseValid_MetsMIR_128_Fga	MetsMi R_128	Fga	CAAATACTCCATG ATATTCC	Mouse Valid	128	Fga_128
MouseValid_MetsMIR_129_Fga	MetsMi R_129	Fga	CCTTCTGCTCTGA TGATGAC	Mouse Valid	129	Fga_129
MouseValid_MetsMIR_130_Fga	MetsMi R_130	Fga	AGGATTCTAACTC ACTGACC	Mouse Valid	130	Fga_130
MouseValid_MetsMIR_131_Gnal	MetsMi R_131	Gnal	AAATGYCAAAGAT GCGATCG	Mouse Valid	131	Gnal_131
MouseValid_MetsMIR_132_Gnal	MetsMi R_132	Gnal	ATTGTGCACAGTC GATCAGC	Mouse Valid	132	Gnal_132
MouseValid_MetsMIR_133_Gnal	MetsMi R_133	Gnal	GGTTGAAGCCATT GACGTGC	Mouse Valid	133	Gnal_133
MouseValid_MetsMIR_134_Gnal	MetsMi R_134	Gnal	CTCGTTGGATCTC TCAAAGC	Mouse Valid	134	Gnal_134
MouseValid_MetsMIR_135_Gnal	MetsMi R_135	Gnal	CCAACCCCTGAGAA CCAGTTC	Mouse Valid	135	Gnal_135
MouseValid_MetsMIR_136_Gnal	MetsMi R_136	Gnal	CACTGACTTTGAA TATTCCC	Mouse Valid	136	Gnal_136
MouseValid_MetsMIR_137_Gnal	MetsMi R_137	Gnal	TTACAGGGGGCTG GTGAGTC	Mouse Valid	137	Gnal_137
MouseValid_MetsMIR_138_Gnal	MetsMi R_138	Gnal	GGAAAAGAAGCA GAAAATTTC	Mouse Valid	138	Gnal_138

Figure 18 (supplemental tables continued)

FIG. 18AG

MouseValid_MetsMIR_139_Gnal	MetsMi R_139	Gnal	TGCTCTTGATATA ATCTGAC	Mouse Valid	139	Gnal_139
MouseValid_MetsMIR_140_Gnal	MetsMi R_140	Gnal	GAATATTCAAAGT CAGTGAT	Mouse Valid	140	Gnal_140
MouseValid_MetsMIR_141_Gorasp1	MetsMi R_141	Gorasp1	CACCAGCCTCGA GTGCCCCGA	Mouse Valid	141	Gorasp1_141
MouseValid_MetsMIR_142_Gorasp1	MetsMi R_142	Gorasp1	CACATGCCACACG TGTTCCG	Mouse Valid	142	Gorasp1_142
MouseValid_MetsMIR_143_Gorasp1	MetsMi R_143	Gorasp1	CTGGCCGCCCCA CATGTTGC	Mouse Valid	143	Gorasp1_143
MouseValid_MetsMIR_144_Gorasp1	MetsMi R_144	Gorasp1	ACAGGTACAAGAG AACTCCG	Mouse Valid	144	Gorasp1_144
MouseValid_MetsMIR_145_Gorasp1	MetsMi R_145	Gorasp1	TTATACACCATCA GCTTCAG	Mouse Valid	145	Gorasp1_145
MouseValid_MetsMIR_146_Gorasp1	MetsMi R_146	Gorasp1	CCAACATGTAGT CTGTGTA	Mouse Valid	146	Gorasp1_146
MouseValid_MetsMIR_147_Gorasp1	MetsMi R_147	Gorasp1	TACTCTCATTGAG TCCCATG	Mouse Valid	147	Gorasp1_147
MouseValid_MetsMIR_148_Gorasp1	MetsMi R_148	Gorasp1	CCTACACAGACT ACATAGT	Mouse Valid	148	Gorasp1_148
MouseValid_MetsMIR_149_Gorasp1	MetsMi R_149	Gorasp1	GAAGCTACAGAAG CGCACGC	Mouse Valid	149	Gorasp1_149
MouseValid_MetsMIR_150_Gorasp1	MetsMi R_150	Gorasp1	TGCGCTTCTGTAG CTTCCGC	Mouse Valid	150	Gorasp1_150
MouseValid_MetsMIR_151_Hdgr	MetsMi R_151	Hdgr	TCTATCACCGTC ACCCCTCA	Mouse Valid	151	Hdgr_151
MouseValid_MetsMIR_152_Hdgr	MetsMi R_152	Hdgr	CGAGAACAACCCCT ACAGTCA	Mouse Valid	152	Hdgr_152
MouseValid_MetsMIR_153_Hdgr	MetsMi R_153	Hdgr	CTCCTATCCAGGG CATTCCCT	Mouse Valid	153	Hdgr_153
MouseValid_MetsMIR_154_Hdgr	MetsMi R_154	Hdgr	TGAGAGGCCCCCT GCCTGTAG	Mouse Valid	154	Hdgr_154
MouseValid_MetsMIR_155_Hdgr	MetsMi R_155	Hdgr	GGTGGAGCCCGA AGCCCATG	Mouse Valid	155	Hdgr_155
MouseValid_MetsMIR_156_Hdgr	MetsMi R_156	Hdgr	TCTCATGATCTCT GACGCCC	Mouse Valid	156	Hdgr_156
MouseValid_MetsMIR_157_Hdgr	MetsMi R_157	Hdgr	GGGTCCCAAAAA AAAGACT	Mouse Valid	157	Hdgr_157
MouseValid_MetsMIR_158_Hdgr	MetsMi R_158	Hdgr	CCAGATTGATGAG ATGCCTG	Mouse Valid	158	Hdgr_158
MouseValid_MetsMIR_159_Hdgr	MetsMi R_159	Hdgr	CAGCGACGAAGA AGGGAAAC	Mouse Valid	159	Hdgr_159
MouseValid_MetsMIR_160_Hdgr	MetsMi R_160	Hdgr	CTGTTCTTCTCCA CCTCTAC	Mouse Valid	160	Hdgr_160

Figure 18 (supplemental tables continued)

FIG. 18AH

MouseValid_MetsMIR_161_Jrkl	MetsMi R_161	Jrkl	CTTAAGTACTCTC GAACCTG	Mouse Valid	161	Jrkl_161
MouseValid_MetsMIR_162_Jrkl	MetsMi R_162	Jrkl	TTACGTATTACCA GTCTGTC	Mouse Valid	162	Jrkl_162
MouseValid_MetsMIR_163_Jrkl	MetsMi R_163	Jrkl	CTGGACCAATTTG TGCCAAA	Mouse Valid	163	Jrkl_163
MouseValid_MetsMIR_164_Jrkl	MetsMi R_164	Jrkl	ATGTGTTGCACCA ATGCAAC	Mouse Valid	164	Jrkl_164
MouseValid_MetsMIR_165_Jrkl	MetsMi R_165	Jrkl	AGGCTCATCCAAG CAATTGG	Mouse Valid	165	Jrkl_165
MouseValid_MetsMIR_166_Jrkl	MetsMi R_166	Jrkl	AAAAAAGAACTCT GCCCGTT	Mouse Valid	166	Jrkl_166
MouseValid_MetsMIR_167_Jrkl	MetsMi R_167	Jrkl	ATGAAAAGAACT ATCGTGC	Mouse Valid	167	Jrkl_167
MouseValid_MetsMIR_168_Jrkl	MetsMi R_168	Jrkl	CAAGCTCTGCTAA TAGTAAC	Mouse Valid	168	Jrkl_168
MouseValid_MetsMIR_169_Jrkl	MetsMi R_169	Jrkl	TTCAATCAGCAGC GAGCGAA	Mouse Valid	169	Jrkl_169
MouseValid_MetsMIR_170_Jrkl	MetsMi R_170	Jrkl	CGTGTTAACCAAC CAGCAGA	Mouse Valid	170	Jrkl_170
MouseValid_MetsMIR_171_Lig4	MetsMi R_171	Lig4	AGCATCCGTGCGA GCTCCAC	Mouse Valid	171	Lig4_171
MouseValid_MetsMIR_172_Lig4	MetsMi R_172	Lig4	ATGATGTTGTCAA GCCCAG	Mouse Valid	172	Lig4_172
MouseValid_MetsMIR_173_Lig4	MetsMi R_173	Lig4	CCATTATTACACA ATGCCGT	Mouse Valid	173	Lig4_173
MouseValid_MetsMIR_174_Lig4	MetsMi R_174	Lig4	ATTAGTCCACTGA CGTACTC	Mouse Valid	174	Lig4_174
MouseValid_MetsMIR_175_Lig4	MetsMi R_175	Lig4	TACAACTATACCG ACCAGTT	Mouse Valid	175	Lig4_175
MouseValid_MetsMIR_176_Lig4	MetsMi R_176	Lig4	ACTTAGAATGTCG TTATTCC	Mouse Valid	176	Lig4_176
MouseValid_MetsMIR_177_Lig4	MetsMi R_177	Lig4	AGACAAAAACATG CGTGCCG	Mouse Valid	177	Lig4_177
MouseValid_MetsMIR_178_Lig4	MetsMi R_178	Lig4	TAATCTCGTGGCT TCAATTC	Mouse Valid	178	Lig4_178
MouseValid_MetsMIR_179_Lig4	MetsMi R_179	Lig4	ATTAAACCAGAGT ACGTCAG	Mouse Valid	179	Lig4_179
MouseValid_MetsMIR_180_Lig4	MetsMi R_180	Lig4	TGAATTAGACGTC CTAATTG	Mouse Valid	180	Lig4_180
MouseValid_MetsMIR_181_Mkl2	MetsMi R_181	Mkl2	TCTTACGTGGCAT GATGCC	Mouse Valid	181	Mkl2_181
MouseValid_MetsMIR_182_Mkl2	MetsMi R_182	Mkl2	GCCTCATTTGCCA CTAGTCT	Mouse Valid	182	Mkl2_182

Figure 18 (supplemental tables continued)

FIG. 18A1

MouseValid_MetsMIR_183_Mk2	MetsMi R_183	Mk2	CCCAAGACTAATG GCAATG	Mouse Valid	183	Mk2_183
MouseValid_MetsMIR_184_Mk2	MetsMi R_184	Mk2	TTAACAGGAAGTA TAGTCTT	Mouse Valid	184	Mk2_184
MouseValid_MetsMIR_185_Mk2	MetsMi R_185	Mk2	TCCTTTGTTGCAG CCTCAGC	Mouse Valid	185	Mk2_185
MouseValid_MetsMIR_186_Mk2	MetsMi R_186	Mk2	CCTGTGCTCACAG ACTGAGC	Mouse Valid	186	Mk2_186
MouseValid_MetsMIR_187_Mk2	MetsMi R_187	Mk2	GAAGCCTTGTTGGC TGCTTCT	Mouse Valid	187	Mk2_187
MouseValid_MetsMIR_188_Mk2	MetsMi R_188	Mk2	TCCTCAGCAGTTA AAATTC	Mouse Valid	188	Mk2_188
MouseValid_MetsMIR_189_Mk2	MetsMi R_189	Mk2	GGCTGCATTAGA ACCTGCC	Mouse Valid	189	Mk2_189
MouseValid_MetsMIR_190_Mk2	MetsMi R_190	Mk2	AGGTCTAAATGC AGCCTGC	Mouse Valid	190	Mk2_190
MouseValid_MetsMIR_191_Mreg	MetsMi R_191	Mreg	CCTCTCCAGAGTC GCTCCAA	Mouse Valid	191	Mreg_191
MouseValid_MetsMIR_192_Mreg	MetsMi R_192	Mreg	GCTCTCCGAGAG GTACCTCT	Mouse Valid	192	Mreg_192
MouseValid_MetsMIR_193_Mreg	MetsMi R_193	Mreg	CTTCCCTGCGGAT CTGCCGC	Mouse Valid	193	Mreg_193
MouseValid_MetsMIR_194_Mreg	MetsMi R_194	Mreg	CCTCTAAGATTCT TCTCCAT	Mouse Valid	194	Mreg_194
MouseValid_MetsMIR_195_Mreg	MetsMi R_195	Mreg	AACAATCCGTATT CCTCCTT	Mouse Valid	195	Mreg_195
MouseValid_MetsMIR_196_Mreg	MetsMi R_196	Mreg	CCTCTATCTTCCC CGCCAGC	Mouse Valid	196	Mreg_196
MouseValid_MetsMIR_197_Mreg	MetsMi R_197	Mreg	TCCACCAGGCATG GGACCCC	Mouse Valid	197	Mreg_197
MouseValid_MetsMIR_198_Mreg	MetsMi R_198	Mreg	CAAATGTACCCCA AGAAACC	Mouse Valid	198	Mreg_198
MouseValid_MetsMIR_199_Mreg	MetsMi R_199	Mreg	TATCGTCATCCGC CTCTGTG	Mouse Valid	199	Mreg_199
MouseValid_MetsMIR_200_Mreg	MetsMi R_200	Mreg	CCGGGAGATGCT GTAAAGC	Mouse Valid	200	Mreg_200
MouseValid_MetsMIR_201_Mroh6	MetsMi R_201	Mroh6	CTACAGGCCACTC GCGCTCT	Mouse Valid	201	Mroh6_201
MouseValid_MetsMIR_202_Mroh6	MetsMi R_202	Mroh6	CATCTCTCCAAGA GCGCGAG	Mouse Valid	202	Mroh6_202
MouseValid_MetsMIR_203_Mroh6	MetsMi R_203	Mroh6	GCCCCTTTACTCG CTGATTG	Mouse Valid	203	Mroh6_203
MouseValid_MetsMIR_204_Mroh6	MetsMi R_204	Mroh6	CCGGTCCGGGGG TAGCGAAC	Mouse Valid	204	Mroh6_204

Figure 18 (supplemental tables continued)

FIG. 18AJ

MouseValid_MetsMIR_205_Mroh6	MetsMI R_205	Mroh6	TATACAACGCAAG GTCTGCA	Mouse Valid	205	Mroh6_205
MouseValid_MetsMIR_206_Mroh6	MetsMI R_206	Mroh6	TGTGGCGGCTCC CGAGTCTG	Mouse Valid	206	Mroh6_206
MouseValid_MetsMIR_207_Mroh6	MetsMI R_207	Mroh6	GCAGCCCACGCA CCAGACCG	Mouse Valid	207	Mroh6_207
MouseValid_MetsMIR_208_Mroh6	MetsMI R_208	Mroh6	GGCTGGCTATGC CGGGACCC	Mouse Valid	208	Mroh6_208
MouseValid_MetsMIR_209_Mroh6	MetsMI R_209	Mroh6	TTTGGGAGCCTTC CCACCGA	Mouse Valid	209	Mroh6_209
MouseValid_MetsMIR_210_Mroh6	MetsMI R_210	Mroh6	AGCCGGAATCAG CGAGTAA	Mouse Valid	210	Mroh6_210
MouseValid_MetsMIR_211_Mxra7	MetsMI R_211	Mxra7	CCACGGCGCGAT GGAGTCGC	Mouse Valid	211	Mxra7_211
MouseValid_MetsMIR_212_Mxra7	MetsMI R_212	Mxra7	CCCGTCTGACATC TCGCCAA	Mouse Valid	212	Mxra7_212
MouseValid_MetsMIR_213_Mxra7	MetsMI R_213	Mxra7	ACTCCGACTCTAT GACATGT	Mouse Valid	213	Mxra7_213
MouseValid_MetsMIR_214_Mxra7	MetsMI R_214	Mxra7	GCCACGCGCTCG GACTGGCT	Mouse Valid	214	Mxra7_214
MouseValid_MetsMIR_215_Mxra7	MetsMI R_215	Mxra7	GCCACCCACTGAA GAACCTG	Mouse Valid	215	Mxra7_215
MouseValid_MetsMIR_216_Mxra7	MetsMI R_216	Mxra7	TGCTCACCCTCTT CTTCCTC	Mouse Valid	216	Mxra7_216
MouseValid_MetsMIR_217_Mxra7	MetsMI R_217	Mxra7	TGGCTTCCCCTCA GCTGCC	Mouse Valid	217	Mxra7_217
MouseValid_MetsMIR_218_Mxra7	MetsMI R_218	Mxra7	GATGATGACCAA GAAGAAC	Mouse Valid	218	Mxra7_218
MouseValid_MetsMIR_219_Mxra7	MetsMI R_219	Mxra7	CGCCAGCGCCGT GACCAGCG	Mouse Valid	219	Mxra7_219
MouseValid_MetsMIR_220_Mxra7	MetsMI R_220	Mxra7	ACCGAGCCAGTC CGAGCGCG	Mouse Valid	220	Mxra7_220
MouseValid_MetsMIR_221_Nf2	MetsMI R_221	Nf2	GATCCGCACCGT GAATGTCT	Mouse Valid	221	Nf2_221
MouseValid_MetsMIR_222_Nf2	MetsMI R_222	Nf2	GAAGCTCATGCGA GAAGCGA	Mouse Valid	222	Nf2_222
MouseValid_MetsMIR_223_Nf2	MetsMI R_223	Nf2	CTTGCGTCATAT GCTGTCC	Mouse Valid	223	Nf2_223
MouseValid_MetsMIR_224_Nf2	MetsMI R_224	Nf2	CACTGGGGCTTC GGGAAACC	Mouse Valid	224	Nf2_224
MouseValid_MetsMIR_225_Nf2	MetsMI R_225	Nf2	TGGACTGCAGTAT ACAATCA	Mouse Valid	225	Nf2_225
MouseValid_MetsMIR_226_Nf2	MetsMI R_226	Nf2	GTTGGATCATGAT GTTTCGA	Mouse Valid	226	Nf2_226

Figure 18 (supplemental tables continued)

FIG. 18AK

MouseValid_MetsMIR_227_Nf2	MetsMi R_227	Nf2	AAAGGTCTACTGC CCTCCCG	Mouse Valid	227	Nf2_227
MouseValid_MetsMIR_228_Nf2	MetsMi R_228	Nf2	ACCAGTTACCTTT CACTTCC	Mouse Valid	228	Nf2_228
MouseValid_MetsMIR_229_Nf2	MetsMi R_229	Nf2	CGTCACCATGGAC GCCGAGA	Mouse Valid	229	Nf2_229
MouseValid_MetsMIR_230_Nf2	MetsMi R_230	Nf2	AGCCCAAGACATT CACGGTG	Mouse Valid	230	Nf2_230
MouseValid_MetsMIR_231_Pcdh1	MetsMi R_231	Pcdh1	CTGGGGCGTGCT AAGCCGAA	Mouse Valid	231	Pcdh1_231
MouseValid_MetsMIR_232_Pcdh1	MetsMi R_232	Pcdh1	GTTACACCGCCTC GTGGTGA	Mouse Valid	232	Pcdh1_232
MouseValid_MetsMIR_233_Pcdh1	MetsMi R_233	Pcdh1	GTAGGTAGGCCG CTCGAATT	Mouse Valid	233	Pcdh1_233
MouseValid_MetsMIR_234_Pcdh1	MetsMi R_234	Pcdh1	CGCCATTGGGGC CGCGGTCA	Mouse Valid	234	Pcdh1_234
MouseValid_MetsMIR_235_Pcdh1	MetsMi R_235	Pcdh1	GCCCCGACCATC GAGATCCG	Mouse Valid	235	Pcdh1_235
MouseValid_MetsMIR_236_Pcdh1	MetsMi R_236	Pcdh1	CGTGCGCAAGCCC CCTCGCTA	Mouse Valid	236	Pcdh1_236
MouseValid_MetsMIR_237_Pcdh1	MetsMi R_237	Pcdh1	CAGTACTAGTGGC CTACTGC	Mouse Valid	237	Pcdh1_237
MouseValid_MetsMIR_238_Pcdh1	MetsMi R_238	Pcdh1	CGTGTAGTCAATC TCCGCAT	Mouse Valid	238	Pcdh1_238
MouseValid_MetsMIR_239_Pcdh1	MetsMi R_239	Pcdh1	GCGTGTCCAGGC TATGGACG	Mouse Valid	239	Pcdh1_239
MouseValid_MetsMIR_240_Pcdh1	MetsMi R_240	Pcdh1	CCCCTGGATATTG ACATCGC	Mouse Valid	240	Pcdh1_240
MouseValid_MetsMIR_241_Potr2c	MetsMi R_241	Potr2c	ATCGCAGAGGTG CCCAAT	Mouse Valid	241	Potr2c_241
MouseValid_MetsMIR_242_Potr2c	MetsMi R_242	Potr2c	AGACACTAACCGG AATAACC	Mouse Valid	242	Potr2c_242
MouseValid_MetsMIR_243_Potr2c	MetsMi R_243	Potr2c	GTAGGGTGGCCA ATTCCATT	Mouse Valid	243	Potr2c_243
MouseValid_MetsMIR_244_Potr2c	MetsMi R_244	Potr2c	GCGCGATGAATTC ATCATGA	Mouse Valid	244	Potr2c_244
MouseValid_MetsMIR_245_Potr2c	MetsMi R_245	Potr2c	GTCAATCTGAACC CAGTCGA	Mouse Valid	245	Potr2c_245
MouseValid_MetsMIR_246_Potr2c	MetsMi R_246	Potr2c	CCACAATGTCATC ACTAGTG	Mouse Valid	246	Potr2c_246
MouseValid_MetsMIR_247_Potr2c	MetsMi R_247	Potr2c	CCTCACTAGTGAT GACATTG	Mouse Valid	247	Potr2c_247
MouseValid_MetsMIR_248_Potr2c	MetsMi R_248	Potr2c	GTTCTGCCCTGAG TGTCTCG	Mouse Valid	248	Potr2c_248

Figure 18 (supplemental tables continued)

FIG. 18AL

MouseValid MetsMIR 249 Polr2c	MetsMi R 249	Polr2c	TGCGATGAAGACC CTCCGAA	Mouse Valid	249	Polr2c 249
MouseValid MetsMIR 250 Polr2c	MetsMi R 250	Polr2c	TCAGATTGACGCC AACTCCT	Mouse Valid	250	Polr2c 250
MouseValid MetsMIR 251 Polr3gl	MetsMi R 251	Polr3gl	CTATTGACAACGC CATCGAT	Mouse Valid	251	Polr3gl 251
MouseValid MetsMIR 252 Polr3gl	MetsMi R 252	Polr3gl	CTTACCAGGGTTC CAATCGA	Mouse Valid	252	Polr3gl 252
MouseValid MetsMIR 253 Polr3gl	MetsMi R 253	Polr3gl	GTGAGCTCAAGAT TCGAGTG	Mouse Valid	253	Polr3gl 253
MouseValid MetsMIR 254 Polr3gl	MetsMi R 254	Polr3gl	CCCTAAGAGCACA GACGATA	Mouse Valid	254	Polr3gl 254
MouseValid MetsMIR 255 Polr3gl	MetsMi R 255	Polr3gl	CACCTCGAATCTTG AGCTCAC	Mouse Valid	255	Polr3gl 255
MouseValid MetsMIR 256 Polr3gl	MetsMi R 256	Polr3gl	CCACCATTATCCT TCCCAAG	Mouse Valid	256	Polr3gl 256
MouseValid MetsMIR 257 Polr3gl	MetsMi R 257	Polr3gl	GGCAGCTCCCCTA CTTTCATC	Mouse Valid	257	Polr3gl 257
MouseValid MetsMIR 258 Polr3gl	MetsMi R 258	Polr3gl	AGAGGAGGGGGA GTATGTCC	Mouse Valid	258	Polr3gl 258
MouseValid MetsMIR 259 Polr3gl	MetsMi R 259	Polr3gl	GGAGACGATACA GAAACTAG	Mouse Valid	259	Polr3gl 259
MouseValid MetsMIR 260 Polr3gl	MetsMi R 260	Polr3gl	TCCTTATCGTCTG TGCTCTT	Mouse Valid	260	Polr3gl 260
MouseValid MetsMIR 261 Ppp2r4	MetsMi R 261	Ppp2r4	CACCTCTTGATAC GCTGGAT	Mouse Valid	261	Ppp2r4 261
MouseValid MetsMIR 262 Ppp2r4	MetsMi R 262	Ppp2r4	TGTGCCATAGTCA ATTCTGT	Mouse Valid	262	Ppp2r4 262
MouseValid MetsMIR 263 Ppp2r4	MetsMi R 263	Ppp2r4	CTCAGAGACTTTG TAGTCGA	Mouse Valid	263	Ppp2r4 263
MouseValid MetsMIR 264 Ppp2r4	MetsMi R 264	Ppp2r4	CGCTTCCATTGTC CCATATC	Mouse Valid	264	Ppp2r4 264
MouseValid MetsMIR 265 Ppp2r4	MetsMi R 265	Ppp2r4	GATTGATGAAACC CCGCCAG	Mouse Valid	265	Ppp2r4 265
MouseValid MetsMIR 266 Ppp2r4	MetsMi R 266	Ppp2r4	AACTCCACACGAA TTGACTA	Mouse Valid	266	Ppp2r4 266
MouseValid MetsMIR 267 Ppp2r4	MetsMi R 267	Ppp2r4	TAGGCATATGCTG ACTACAT	Mouse Valid	267	Ppp2r4 267
MouseValid MetsMIR 268 Ppp2r4	MetsMi R 268	Ppp2r4	ATGATAAAGTCT GAGTAGT	Mouse Valid	268	Ppp2r4 268
MouseValid MetsMIR 269 Ppp2r4	MetsMi R 269	Ppp2r4	TTCACAGGCCATC GAGAAGC	Mouse Valid	269	Ppp2r4 269
MouseValid MetsMIR 270 Ppp2r4	MetsMi R 270	Ppp2r4	CAGTCTCACCTGA TCAAGTT	Mouse Valid	270	Ppp2r4 270

Figure 18 (supplemental tables continued)

FIG. 18AM

MouseValid_MetsMIR_271_Pten	MetsMIR_271	Pten	ACCGCCAAATTTA ACTGCAG	Mouse Valid	271	Pten_271
MouseValid_MetsMIR_272_Pten	MetsMIR_272	Pten	GCAGCAATTCAC GTAAAGC	Mouse Valid	272	Pten_272
MouseValid_MetsMIR_273_Pten	MetsMIR_273	Pten	TGTCATCTTCACT TAGCCAT	Mouse Valid	273	Pten_273
MouseValid_MetsMIR_274_Pten	MetsMIR_274	Pten	ACAATATTGATGA TGAGTA	Mouse Valid	274	Pten_274
MouseValid_MetsMIR_275_Pten	MetsMIR_275	Pten	CATACCTCTGCAG TTAAATT	Mouse Valid	275	Pten_275
MouseValid_MetsMIR_276_Pten	MetsMIR_276	Pten	AATCCCATAGCAA TAATATT	Mouse Valid	276	Pten_276
MouseValid_MetsMIR_277_Pten	MetsMIR_277	Pten	ACAGATTGTATAT CTTGTA	Mouse Valid	277	Pten_277
MouseValid_MetsMIR_278_Pten	MetsMIR_278	Pten	GGTTTGATAAGTT CTAGCTG	Mouse Valid	278	Pten_278
MouseValid_MetsMIR_279_Pten	MetsMIR_279	Pten	TCTGTGAAGATCT TGACCAA	Mouse Valid	279	Pten_279
MouseValid_MetsMIR_280_Pten	MetsMIR_280	Pten	TTATCCAAATATTA TTGCTA	Mouse Valid	280	Pten_280
MouseValid_MetsMIR_281_Ptes2	MetsMIR_281	Ptes2	CGTGTACCGAACA CCCGCTG	Mouse Valid	281	Ptes2_281
MouseValid_MetsMIR_282_Ptes2	MetsMIR_282	Ptes2	CTTCGACTACATT GTCCGTG	Mouse Valid	282	Ptes2_282
MouseValid_MetsMIR_283_Ptes2	MetsMIR_283	Ptes2	CACCTATTACCCA CCCATGA	Mouse Valid	283	Ptes2_283
MouseValid_MetsMIR_284_Ptes2	MetsMIR_284	Ptes2	TAGTGCCCTCAAG ACCTACC	Mouse Valid	284	Ptes2_284
MouseValid_MetsMIR_285_Ptes2	MetsMIR_285	Ptes2	AGAAGGGACATGT CTTGTA	Mouse Valid	285	Ptes2_285
MouseValid_MetsMIR_286_Ptes2	MetsMIR_286	Ptes2	CCAGTACTTGTTG CAAACT	Mouse Valid	286	Ptes2_286
MouseValid_MetsMIR_287_Ptes2	MetsMIR_287	Ptes2	GAGTGACTTACCC TGAAACC	Mouse Valid	287	Ptes2_287
MouseValid_MetsMIR_288_Ptes2	MetsMIR_288	Ptes2	TGGAGGTGAATCC CGTGACA	Mouse Valid	288	Ptes2_288
MouseValid_MetsMIR_289_Ptes2	MetsMIR_289	Ptes2	AGCGGGTGTTTCG GTACACGT	Mouse Valid	289	Ptes2_289
MouseValid_MetsMIR_290_Ptes2	MetsMIR_290	Ptes2	GGCAGCCGCGCC CACATACT	Mouse Valid	290	Ptes2_290
MouseValid_MetsMIR_291_Rhoh	MetsMIR_291	Rhoh	ACCCGTATTCTCG TACACCG	Mouse Valid	291	Rhoh_291
MouseValid_MetsMIR_292_Rhoh	MetsMIR_292	Rhoh	ACCCACGGGTGAC GAGAATA	Mouse Valid	292	Rhoh_292

Figure 18 (supplemental tables continued)

FIG. 18AN

MouseValid_MetsMIR_293_Rhoh	MetsMl R_293	Rhoh	CCCTTCTATGGCA TTGATGC	Mouse Valid	293	Rhoh_293
MouseValid_MetsMIR_294_Rhoh	MetsMl R_294	Rhoh	ACGACGCCCTCAG AAGTATC	Mouse Valid	294	Rhoh_294
MouseValid_MetsMIR_295_Rhoh	MetsMl R_295	Rhoh	GAGCAACCTACCC TGTACCC	Mouse Valid	295	Rhoh_295
MouseValid_MetsMIR_296_Rhoh	MetsMl R_296	Rhoh	CAGGGGCCGGAT ACTTCTGA	Mouse Valid	296	Rhoh_296
MouseValid_MetsMIR_297_Rhoh	MetsMl R_297	Rhoh	GAGCTCAATCAAG TGCGTGC	Mouse Valid	297	Rhoh_297
MouseValid_MetsMIR_298_Rhoh	MetsMl R_298	Rhoh	GTCTACCAGCAG GCAGACG	Mouse Valid	298	Rhoh_298
MouseValid_MetsMIR_299_Rhoh	MetsMl R_299	Rhoh	TCCTGCATCAATG CCATAGA	Mouse Valid	299	Rhoh_299
MouseValid_MetsMIR_300_Rhoh	MetsMl R_300	Rhoh	CAAATACCTGCTG TACTCCC	Mouse Valid	300	Rhoh_300
MouseValid_MetsMIR_301_Rhox13	MetsMl R_301	Rhox13	CGAGTCGTATCA CTCTCGT	Mouse Valid	301	Rhox13_301
MouseValid_MetsMIR_302_Rhox13	MetsMl R_302	Rhox13	CGACTCGAGCAC CAGCGACG	Mouse Valid	302	Rhox13_302
MouseValid_MetsMIR_303_Rhox13	MetsMl R_303	Rhox13	ATGCTGAGAAATA TTCCGCC	Mouse Valid	303	Rhox13_303
MouseValid_MetsMIR_304_Rhox13	MetsMl R_304	Rhox13	CCTCTGCTTCGCT CTCCGAT	Mouse Valid	304	Rhox13_304
MouseValid_MetsMIR_305_Rhox13	MetsMl R_305	Rhox13	AAGAACTTTGAAT GTGCCTG	Mouse Valid	305	Rhox13_305
MouseValid_MetsMIR_306_Rhox13	MetsMl R_306	Rhox13	CTGACCTTCACTT TGACCTC	Mouse Valid	306	Rhox13_306
MouseValid_MetsMIR_307_Rhox13	MetsMl R_307	Rhox13	ATGGCGGTCCAC GGCGACGA	Mouse Valid	307	Rhox13_307
MouseValid_MetsMIR_308_Rhox13	MetsMl R_308	Rhox13	CGACCTGCCACTG TGCGAAG	Mouse Valid	308	Rhox13_308
MouseValid_MetsMIR_309_Rhox13	MetsMl R_309	Rhox13	ATCTGAAGTATCC TCGTGCG	Mouse Valid	309	Rhox13_309
MouseValid_MetsMIR_310_Rhox13	MetsMl R_310	Rhox13	CATGTTGCGCGTC GTGCGCG	Mouse Valid	310	Rhox13_310
MouseValid_MetsMIR_311_Sbp	MetsMl R_311	Sbp	AGCCAATCCCTCT GACGCAG	Mouse Valid	311	Sbp_311
MouseValid_MetsMIR_312_Sbp	MetsMl R_312	Sbp	TGGGGAAATATCA ATGCTAA	Mouse Valid	312	Sbp_312
MouseValid_MetsMIR_313_Sbp	MetsMl R_313	Sbp	CTCACTCTGGGCT CTGCATG	Mouse Valid	313	Sbp_313
MouseValid_MetsMIR_314_Sbp	MetsMl R_314	Sbp	CATGCTGCTGTTG CTGACTT	Mouse Valid	314	Sbp_314

Figure 18 (supplemental tables continued)

FIG. 18A0

MouseValid_MetsMIR_315_Sbp	MetsMI R_315	Sbp	ACTTGCACGACG ATTCCCC	Mouse Valid	315	Sbp_315
MouseValid_MetsMIR_316_Sbp	MetsMI R_316	Sbp	GATGTCCTGGGG AATGCTGC	Mouse Valid	316	Sbp_316
MouseValid_MetsMIR_317_Sbp	MetsMI R_317	Sbp	GCACCGGGCCTC TGCCTCAG	Mouse Valid	317	Sbp_317
MouseValid_MetsMIR_318_Sbp	MetsMI R_318	Sbp	GGTCGATACTTCA GTGACAC	Mouse Valid	318	Sbp_318
MouseValid_MetsMIR_319_Sbp	MetsMI R_319	Sbp	TCGACCCCTTCTA ACACCAA	Mouse Valid	319	Sbp_319
MouseValid_MetsMIR_320_Sbp	MetsMI R_320	Sbp	TATCTTCCTTGTA TTATAA	Mouse Valid	320	Sbp_320
MouseValid_MetsMIR_321_Scgb2b3	MetsMI R_321	Scgb2b3	GTCTATTTGAGCA CTGTCTC	Mouse Valid	321	Scgb2b3_321
MouseValid_MetsMIR_322_Scgb2b3	MetsMI R_322	Scgb2b3	TCTTATTTCAACC CTACTGA	Mouse Valid	322	Scgb2b3_322
MouseValid_MetsMIR_323_Scgb2b3	MetsMI R_323	Scgb2b3	AGCCAAAGCTCTCC AATCACC	Mouse Valid	323	Scgb2b3_323
MouseValid_MetsMIR_324_Scgb2b3	MetsMI R_324	Scgb2b3	TAAGTGTGATTAC CAAGAAA	Mouse Valid	324	Scgb2b3_324
MouseValid_MetsMIR_325_Scgb2b3	MetsMI R_325	Scgb2b3	TAACCTCTACATC CATTTCT	Mouse Valid	325	Scgb2b3_325
MouseValid_MetsMIR_326_Scgb2b3	MetsMI R_326	Scgb2b3	TTCCTGGTGATTG GAGAGCT	Mouse Valid	326	Scgb2b3_326
MouseValid_MetsMIR_327_Scgb2b3	MetsMI R_327	Scgb2b3	AGGACTAAAAGCT AAAAGTC	Mouse Valid	327	Scgb2b3_327
MouseValid_MetsMIR_328_Scgb2b3	MetsMI R_328	Scgb2b3	CTCTTTGTTTCAC CATCAGT	Mouse Valid	328	Scgb2b3_328
MouseValid_MetsMIR_329_Scgb2b3	MetsMI R_329	Scgb2b3	GAAGCATGTGCTT CTTTCTT	Mouse Valid	329	Scgb2b3_329
MouseValid_MetsMIR_330_Scgb2b3	MetsMI R_330	Scgb2b3	AAAGAGCTTCAAA AAAATCC	Mouse Valid	330	Scgb2b3_330
MouseValid_MetsMIR_331_Sk2	MetsMI R_331	Sk2	CTTTAAAACTGGT GAAGTGC	Mouse Valid	331	Sk2_331
MouseValid_MetsMIR_332_Sk2	MetsMI R_332	Sk2	ATCCTAATAGATTT CGGCTT	Mouse Valid	332	Sk2_332
MouseValid_MetsMIR_333_Sk2	MetsMI R_333	Sk2	CACAAGGTACAAC ATACTTT	Mouse Valid	333	Sk2_333
MouseValid_MetsMIR_334_Sk2	MetsMI R_334	Sk2	GGTCTCTGTGCAC TACCTTC	Mouse Valid	334	Sk2_334
MouseValid_MetsMIR_335_Sk2	MetsMI R_335	Sk2	TCACATCATTAAA CTGTATC	Mouse Valid	335	Sk2_335
MouseValid_MetsMIR_336_Sk2	MetsMI R_336	Sk2	TATTCATGTTGTTA TCCAGC	Mouse Valid	336	Sk2_336

Figure 18 (supplemental tables continued)

FIG. 18AP

MouseValid_MetsMIR_337_Slk2	MetsMi R_337	Slk2	TGATACAGTTTAA TGATGTG	Mouse Valid	337	Slk2_337
MouseValid_MetsMIR_338_Slk2	MetsMi R_338	Slk2	GTGACAGAATATG CCAAAAA	Mouse Valid	338	Slk2_338
MouseValid_MetsMIR_339_Slk2	MetsMi R_339	Slk2	GCAGTATGAAGGA CCACAGC	Mouse Valid	339	Slk2_339
MouseValid_MetsMIR_340_Slk2	MetsMi R_340	Slk2	AATACCCATATAT CCAGCTG	Mouse Valid	340	Slk2_340
MouseValid_MetsMIR_341_Slk38a3	MetsMi R_341	Slk38a3	ACGQTAGCCCAAC TGCTCGT	Mouse Valid	341	Slk38a3_341
MouseValid_MetsMIR_342_Slk38a3	MetsMi R_342	Slk38a3	CATCCGTGCCTAC GAGCAGT	Mouse Valid	342	Slk38a3_342
MouseValid_MetsMIR_343_Slk38a3	MetsMi R_343	Slk38a3	TCTGGGGCTCGC CTACGCCA	Mouse Valid	343	Slk38a3_343
MouseValid_MetsMIR_344_Slk38a3	MetsMi R_344	Slk38a3	CCACAATCCCCGA AGACTTG	Mouse Valid	344	Slk38a3_344
MouseValid_MetsMIR_345_Slk38a3	MetsMi R_345	Slk38a3	GAAAAGGATGATG CCCGTAT	Mouse Valid	345	Slk38a3_345
MouseValid_MetsMIR_346_Slk38a3	MetsMi R_346	Slk38a3	CAGCTCACATCGG TGAAGTG	Mouse Valid	346	Slk38a3_346
MouseValid_MetsMIR_347_Slk38a3	MetsMi R_347	Slk38a3	GGATGGAATAGCT AGACAAC	Mouse Valid	347	Slk38a3_347
MouseValid_MetsMIR_348_Slk38a3	MetsMi R_348	Slk38a3	ACTGAAGACACCC AACACTG	Mouse Valid	348	Slk38a3_348
MouseValid_MetsMIR_349_Slk38a3	MetsMi R_349	Slk38a3	TCCAATGTTCTGA AGCGTAA	Mouse Valid	349	Slk38a3_349
MouseValid_MetsMIR_350_Slk38a3	MetsMi R_350	Slk38a3	GCCATTACGCTTC AGAACAT	Mouse Valid	350	Slk38a3_350
MouseValid_MetsMIR_351_Snap47	MetsMi R_351	Snap47	ACATACCCTCCCA CGTGACC	Mouse Valid	351	Snap47_351
MouseValid_MetsMIR_352_Snap47	MetsMi R_352	Snap47	AGCATACGCCAGC GCTTCAT	Mouse Valid	352	Snap47_352
MouseValid_MetsMIR_353_Snap47	MetsMi R_353	Snap47	CCGACCTGGACG TGGCTGAC	Mouse Valid	353	Snap47_353
MouseValid_MetsMIR_354_Snap47	MetsMi R_354	Snap47	ACATCAGCTTTCT CATACGC	Mouse Valid	354	Snap47_354
MouseValid_MetsMIR_355_Snap47	MetsMi R_355	Snap47	TGTTCCAGCATCC AGACCGA	Mouse Valid	355	Snap47_355
MouseValid_MetsMIR_356_Snap47	MetsMi R_356	Snap47	CTGGCTTCCCAAT GAAGCGC	Mouse Valid	356	Snap47_356
MouseValid_MetsMIR_357_Snap47	MetsMi R_357	Snap47	TGCCATCCAGAGC CGCATCC	Mouse Valid	357	Snap47_357
MouseValid_MetsMIR_358_Snap47	MetsMi R_358	Snap47	CGGTGTGGAATG CAGCCCTT	Mouse Valid	358	Snap47_358

Figure 18 (supplemental tables continued)

FIG. 18AQ

MouseValid_MetsMIR_359_Snap47	MetsMi R_359	Snap47	ACTGCAGCCTAGT CGAAATG	Mouse Valid	359	Snap47_359
MouseValid_MetsMIR_360_Snap47	MetsMi R_360	Snap47	CTGATATGCGTGT CCACTCC	Mouse Valid	360	Snap47_360
MouseValid_MetsMIR_361_Tmem86a	MetsMi R_361	Tmem86a	ATCCTCTACGCCG CGGCCTT	Mouse Valid	361	Tmem86a_361
MouseValid_MetsMIR_362_Tmem86a	MetsMi R_362	Tmem86a	CTGACAGCACTCC GATCACC	Mouse Valid	362	Tmem86a_362
MouseValid_MetsMIR_363_Tmem86a	MetsMi R_363	Tmem86a	CTTAAAGAAGGGC ACCAATT	Mouse Valid	363	Tmem86a_363
MouseValid_MetsMIR_364_Tmem86a	MetsMi R_364	Tmem86a	AGACAAGTCCCAC GAAGATG	Mouse Valid	364	Tmem86a_364
MouseValid_MetsMIR_365_Tmem86a	MetsMi R_365	Tmem86a	CGCGCACTCACCA CAGTGAC	Mouse Valid	365	Tmem86a_365
MouseValid_MetsMIR_366_Tmem86a	MetsMi R_366	Tmem86a	CGCCGCCGCCAT GGTGCTC	Mouse Valid	366	Tmem86a_366
MouseValid_MetsMIR_367_Tmem86a	MetsMi R_367	Tmem86a	CTCGGCTCTCGAC AGCTGAC	Mouse Valid	367	Tmem86a_367
MouseValid_MetsMIR_368_Tmem86a	MetsMi R_368	Tmem86a	GCAGAACTTGTG AGAGCGA	Mouse Valid	368	Tmem86a_368
MouseValid_MetsMIR_369_Tmem86a	MetsMi R_369	Tmem86a	AGCCAGTGCCCG CATGCCAA	Mouse Valid	369	Tmem86a_369
MouseValid_MetsMIR_370_Tmem86a	MetsMi R_370	Tmem86a	GAAGAGCAGAGC ACCACCGC	Mouse Valid	370	Tmem86a_370
MouseValid_MetsMIR_371_Trim17	MetsMi R_371	Trim17	ACCTACCTGCCAC TCAGTTA	Mouse Valid	371	Trim17_371
MouseValid_MetsMIR_372_Trim17	MetsMi R_372	Trim17	GGACATAAAATAT CTACGGG	Mouse Valid	372	Trim17_372
MouseValid_MetsMIR_373_Trim17	MetsMi R_373	Trim17	GGACCAACAAAG CCGCTCCC	Mouse Valid	373	Trim17_373
MouseValid_MetsMIR_374_Trim17	MetsMi R_374	Trim17	CTTCAGCGTGCAG TACCCAG	Mouse Valid	374	Trim17_374
MouseValid_MetsMIR_375_Trim17	MetsMi R_375	Trim17	GCACAAGGTCTTG ATAGCCG	Mouse Valid	375	Trim17_375
MouseValid_MetsMIR_376_Trim17	MetsMi R_376	Trim17	TCAGGTCCAGGG AGCGGCTT	Mouse Valid	376	Trim17_376
MouseValid_MetsMIR_377_Trim17	MetsMi R_377	Trim17	ATGTGAAGGACAC CTTGAAC	Mouse Valid	377	Trim17_377
MouseValid_MetsMIR_378_Trim17	MetsMi R_378	Trim17	TTGAGCACTCTA TCTGTCC	Mouse Valid	378	Trim17_378
MouseValid_MetsMIR_379_Trim17	MetsMi R_379	Trim17	GAGGTGCTCAAAA GCTTCCA	Mouse Valid	379	Trim17_379
MouseValid_MetsMIR_380_Trim17	MetsMi R_380	Trim17	GACGAGACACT GCAAGCCA	Mouse Valid	380	Trim17_380

Figure 18 (supplemental tables continued)

FIG. 18AR

MouseValid_MetsMIR_381_Trim72	MetsMi R_381	Trim72	GCTTGACAGCTCC CGACGCA	Mouse Valid	381	Trim72_381
MouseValid_MetsMIR_382_Trim72	MetsMi R_382	Trim72	ATGCTGGGGTTGC CTTGCGT	Mouse Valid	382	Trim72_382
MouseValid_MetsMIR_383_Trim72	MetsMi R_383	Trim72	ATGTTCCGGGCTC TGATGCC	Mouse Valid	383	Trim72_383
MouseValid_MetsMIR_384_Trim72	MetsMi R_384	Trim72	GCGCAAGGAGAA GACTGTAG	Mouse Valid	384	Trim72_384
MouseValid_MetsMIR_385_Trim72	MetsMi R_385	Trim72	CCTGTCAGAGTCA CCACCAC	Mouse Valid	385	Trim72_385
MouseValid_MetsMIR_386_Trim72	MetsMi R_386	Trim72	TTGCTCACCTGCT GGTTACC	Mouse Valid	386	Trim72_386
MouseValid_MetsMIR_387_Trim72	MetsMi R_387	Trim72	CTTTCTCCAGAAA TTCTGCC	Mouse Valid	387	Trim72_387
MouseValid_MetsMIR_388_Trim72	MetsMi R_388	Trim72	ACAGCAGAAGATG CAGCTGC	Mouse Valid	388	Trim72_388
MouseValid_MetsMIR_389_Trim72	MetsMi R_389	Trim72	AGGAGACAGTGC GCCAGTTC	Mouse Valid	389	Trim72_389
MouseValid_MetsMIR_390_Trim72	MetsMi R_390	Trim72	GCTGCTCCCCGA CAGCTCCC	Mouse Valid	390	Trim72_390
MouseValid_MetsMIR_391_Tif2	MetsMi R_391	Tif2	CGTCCGCGATGG CCCGAATA	Mouse Valid	391	Tif2_391
MouseValid_MetsMIR_392_Tif2	MetsMi R_392	Tif2	CGAAGCCGCATGT GTTCTGC	Mouse Valid	392	Tif2_392
MouseValid_MetsMIR_393_Tif2	MetsMi R_393	Tif2	CAGACTGTCTAGT TGCCACC	Mouse Valid	393	Tif2_393
MouseValid_MetsMIR_394_Tif2	MetsMi R_394	Tif2	AGTGGTGTGGAAA CGTCCCA	Mouse Valid	394	Tif2_394
MouseValid_MetsMIR_395_Tif2	MetsMi R_395	Tif2	GCTCGCTCAGGG GCCTCTCC	Mouse Valid	395	Tif2_395
MouseValid_MetsMIR_396_Tif2	MetsMi R_396	Tif2	GCAATGCCTACCT GCCGCCA	Mouse Valid	396	Tif2_396
MouseValid_MetsMIR_397_Tif2	MetsMi R_397	Tif2	TTCTCACTGTTAT TGCATG	Mouse Valid	397	Tif2_397
MouseValid_MetsMIR_398_Tif2	MetsMi R_398	Tif2	TCAGGGTTTGTT CTACCAC	Mouse Valid	398	Tif2_398
MouseValid_MetsMIR_399_Tif2	MetsMi R_399	Tif2	GACGCCGGTCTTT AAGAAGC	Mouse Valid	399	Tif2_399
MouseValid_MetsMIR_400_Tif2	MetsMi R_400	Tif2	GAACACATGCCG CTTCGTGC	Mouse Valid	400	Tif2_400
MouseValid_MetsMIR_401_Tubb4a	MetsMi R_401	Tubb4a	GGAGTTAGTGGAT GCCGTCC	Mouse Valid	401	Tubb4a_401
MouseValid_MetsMIR_402_Tubb4a	MetsMi R_402	Tubb4a	TCTGTCCGCTCCG GCCCTTT	Mouse Valid	402	Tubb4a_402

Figure 18 (supplemental tables continued)

FIG. 18AS

MouseValid_MetsMIR_403_Tubb4a	MetsMI R_403	Tubb4a	TCAATGCGGTAAC CAGATCG	Mouse Valid	403	Tubb4a_403
MouseValid_MetsMIR_404_Tubb4a	MetsMI R_404	Tubb4a	GACGCACCTTGG CCCCGATC	Mouse Valid	404	Tubb4a_404
MouseValid_MetsMIR_405_Tubb4a	MetsMI R_405	Tubb4a	CAGGTCGACCAG CACGGCTC	Mouse Valid	405	Tubb4a_405
MouseValid_MetsMIR_406_Tubb4a	MetsMI R_406	Tubb4a	TCCCCAGGTCAAT CGGAGAC	Mouse Valid	406	Tubb4a_406
MouseValid_MetsMIR_407_Tubb4a	MetsMI R_407	Tubb4a	GTCCCCATGATAG GTCCGAG	Mouse Valid	407	Tubb4a_407
MouseValid_MetsMIR_408_Tubb4a	MetsMI R_408	Tubb4a	GACCCCACTGGG ACCTATCA	Mouse Valid	408	Tubb4a_408
MouseValid_MetsMIR_409_Tubb4a	MetsMI R_409	Tubb4a	TACCGCATTGACC GGCTTGC	Mouse Valid	409	Tubb4a_409
MouseValid_MetsMIR_410_Tubb4a	MetsMI R_410	Tubb4a	GAATCGCCCTCT TCAGCAG	Mouse Valid	410	Tubb4a_410
MouseValid_MetsMIR_411_Ube2g2	MetsMI R_411	Ube2g2	GCTTTCTGTTGTG AGCATGC	Mouse Valid	411	Ube2g2_411
MouseValid_MetsMIR_412_Ube2g2	MetsMI R_412	Ube2g2	GAGATGCAGACTC TGCCATC	Mouse Valid	412	Ube2g2_412
MouseValid_MetsMIR_413_Ube2g2	MetsMI R_413	Ube2g2	TCTCATCTTCGGA GGGCTCA	Mouse Valid	413	Ube2g2_413
MouseValid_MetsMIR_414_Ube2g2	MetsMI R_414	Ube2g2	TTACCTGCCACGA TTCCTTC	Mouse Valid	414	Ube2g2_414
MouseValid_MetsMIR_415_Ube2g2	MetsMI R_415	Ube2g2	GTGGGATGGAAC ATCTCAC	Mouse Valid	415	Ube2g2_415
MouseValid_MetsMIR_416_Ube2g2	MetsMI R_416	Ube2g2	TTAACCTGAATC CTCCAGA	Mouse Valid	416	Ube2g2_416
MouseValid_MetsMIR_417_Ube2g2	MetsMI R_417	Ube2g2	AAAAAAATTCTCTT CATTCA	Mouse Valid	417	Ube2g2_417
MouseValid_MetsMIR_418_Ube2g2	MetsMI R_418	Ube2g2	ATCTTCTCCACAC TCTGCAC	Mouse Valid	418	Ube2g2_418
MouseValid_MetsMIR_419_Ube2g2	MetsMI R_419	Ube2g2	GATCATCACCTGG AGCATGC	Mouse Valid	419	Ube2g2_419
MouseValid_MetsMIR_420_Ube2g2	MetsMI R_420	Ube2g2	GTTATGAGAGCAG TGCCGAG	Mouse Valid	420	Ube2g2_420
MouseValid_MetsMIR_421_Vav3	MetsMI R_421	Vav3	GTTCACAGTGTC GCCTGCC	Mouse Valid	421	Vav3_421
MouseValid_MetsMIR_422_Vav3	MetsMI R_422	Vav3	TCCAGGTTTGCCG AAGATGC	Mouse Valid	422	Vav3_422
MouseValid_MetsMIR_423_Vav3	MetsMI R_423	Vav3	TAGGGACGAAACT TAGCTTC	Mouse Valid	423	Vav3_423
MouseValid_MetsMIR_424_Vav3	MetsMI R_424	Vav3	ACTTCGGAGAGCT TCCAGAC	Mouse Valid	424	Vav3_424

Figure 18 (supplemental tables continued)

FIG. 18AT

MouseValid_MetsMIR_425_Vav3	MetsMIR_425	Vav3	TTCTTACAGAAAC GGACCAA	MouseValid	425	Vav3_425
MouseValid_MetsMIR_426_Vav3	MetsMIR_426	Vav3	AGGGCCATTCAAA CCACCAG	MouseValid	426	Vav3_426
MouseValid_MetsMIR_427_Vav3	MetsMIR_427	Vav3	TGAGAACTTACAC ATGGAGA	MouseValid	427	Vav3_427
MouseValid_MetsMIR_428_Vav3	MetsMIR_428	Vav3	CTTACCTCTGGTG GTTTGAA	MouseValid	428	Vav3_428
MouseValid_MetsMIR_429_Vav3	MetsMIR_429	Vav3	AGGACCACCATTG CATATCC	MouseValid	429	Vav3_429
MouseValid_MetsMIR_430_Vav3	MetsMIR_430	Vav3	CAGACCTGGGTC CACCTGTC	MouseValid	430	Vav3_430
MouseValid_MetsMIR_431_Vmn1r50	MetsMIR_431	Vmn1r50	AGCCTTTCTCATC GGACTCA	MouseValid	431	Vmn1r50_431
MouseValid_MetsMIR_432_Vmn1r50	MetsMIR_432	Vmn1r50	CAGTACCATCATG CTACTCA	MouseValid	432	Vmn1r50_432
MouseValid_MetsMIR_433_Vmn1r50	MetsMIR_433	Vmn1r50	GGTACGTGGTAGT TCTCCTA	MouseValid	433	Vmn1r50_433
MouseValid_MetsMIR_434_Vmn1r50	MetsMIR_434	Vmn1r50	GTTTGGGCTTGTT CTCATGA	MouseValid	434	Vmn1r50_434
MouseValid_MetsMIR_435_Vmn1r50	MetsMIR_435	Vmn1r50	AAGTTTGTACAT CAATGTG	MouseValid	435	Vmn1r50_435
MouseValid_MetsMIR_436_Vmn1r50	MetsMIR_436	Vmn1r50	GAAGCCCATGACT AGCATGA	MouseValid	436	Vmn1r50_436
MouseValid_MetsMIR_437_Vmn1r50	MetsMIR_437	Vmn1r50	TGCAGAGGTGGA CAACAAAA	MouseValid	437	Vmn1r50_437
MouseValid_MetsMIR_438_Vmn1r50	MetsMIR_438	Vmn1r50	TAGCAATAATTGA TACTAAT	MouseValid	438	Vmn1r50_438
MouseValid_MetsMIR_439_Vmn1r50	MetsMIR_439	Vmn1r50	CTTGTCTAGAGT AACTCAC	MouseValid	439	Vmn1r50_439
MouseValid_MetsMIR_440_Vmn1r50	MetsMIR_440	Vmn1r50	TTTCACAATGAT GACCATG	MouseValid	440	Vmn1r50_440
MouseValid_MetsMIR_441_Vwa5b2	MetsMIR_441	Vwa5b2	ATCGGCGCCGGC TATGCAAC	MouseValid	441	Vwa5b2_441
MouseValid_MetsMIR_442_Vwa5b2	MetsMIR_442	Vwa5b2	AGGCGTGTTCGTG TACCCTC	MouseValid	442	Vwa5b2_442
MouseValid_MetsMIR_443_Vwa5b2	MetsMIR_443	Vwa5b2	GCCTGAGAACAC GTCCGGAG	MouseValid	443	Vwa5b2_443
MouseValid_MetsMIR_444_Vwa5b2	MetsMIR_444	Vwa5b2	GAGCGGCGCGG CGTCCAGC	MouseValid	444	Vwa5b2_444
MouseValid_MetsMIR_445_Vwa5b2	MetsMIR_445	Vwa5b2	CGCTCACCGACTC CTGCCGTC	MouseValid	445	Vwa5b2_445
MouseValid_MetsMIR_446_Vwa5b2	MetsMIR_446	Vwa5b2	CTCCGGGGCTCT GTGACGAC	MouseValid	446	Vwa5b2_446

Figure 18 (supplemental tables continued)

FIG. 18AU

MouseValid_MetsMIR_447_Vwa5b2	MetsMIR_447	Vwa5b2	YCTTTCGAGATGC TAGTGAC	MouseValid	447	Vwa5b2_447
MouseValid_MetsMIR_448_Vwa5b2	MetsMIR_448	Vwa5b2	GCCACATGCATCA CCCCATC	MouseValid	448	Vwa5b2_448
MouseValid_MetsMIR_449_Vwa5b2	MetsMIR_449	Vwa5b2	TACACCCCGCCG CTGCGCAC	MouseValid	449	Vwa5b2_449
MouseValid_MetsMIR_450_Vwa5b2	MetsMIR_450	Vwa5b2	GGCCGCCTGCTG TCGCGCTC	MouseValid	450	Vwa5b2_450
MouseValid_MetsMIR_451_Wdly1	MetsMIR_451	Wdly1	CGGCGTGATCAC CGCTAGCG	MouseValid	451	Wdly1_451
MouseValid_MetsMIR_452_Wdly1	MetsMIR_452	Wdly1	CAAATATCCGCCT GCTGTCTG	MouseValid	452	Wdly1_452
MouseValid_MetsMIR_453_Wdly1	MetsMIR_453	Wdly1	TGGCGGCGGAGA TCCACTCC	MouseValid	453	Wdly1_453
MouseValid_MetsMIR_454_Wdly1	MetsMIR_454	Wdly1	CATACCACACGA CAGCAGG	MouseValid	454	Wdly1_454
MouseValid_MetsMIR_455_Wdly1	MetsMIR_455	Wdly1	TCCACAGAACCAT CCGAGTA	MouseValid	455	Wdly1_455
MouseValid_MetsMIR_456_Wdly1	MetsMIR_456	Wdly1	AAAGAGACAGCG GCCAGTAC	MouseValid	456	Wdly1_456
MouseValid_MetsMIR_457_Wdly1	MetsMIR_457	Wdly1	ACTCTACCTGTCC TCGCTAG	MouseValid	457	Wdly1_457
MouseValid_MetsMIR_458_Wdly1	MetsMIR_458	Wdly1	GCGCGGCGGTGA CGCGCTCC	MouseValid	458	Wdly1_458
MouseValid_MetsMIR_459_Wdly1	MetsMIR_459	Wdly1	CGCCGCGCTGCT CATCCCCA	MouseValid	459	Wdly1_459
MouseValid_MetsMIR_460_Wdly1	MetsMIR_460	Wdly1	TGATCACGCCGTC CTCCTTG	MouseValid	460	Wdly1_460
MouseValid_MetsMIR_461_Zfp800	MetsMIR_461	Zfp800	TTTCACTTCAATAC CGACGT	MouseValid	461	Zfp800_461
MouseValid_MetsMIR_462_Zfp800	MetsMIR_462	Zfp800	TCTGCACTTACGA AAGCCGT	MouseValid	462	Zfp800_462
MouseValid_MetsMIR_463_Zfp800	MetsMIR_463	Zfp800	ATAATTGAGTGCT TTCGATC	MouseValid	463	Zfp800_463
MouseValid_MetsMIR_464_Zfp800	MetsMIR_464	Zfp800	CTTCCATGTAGTT TATATCC	MouseValid	464	Zfp800_464
MouseValid_MetsMIR_465_Zfp800	MetsMIR_465	Zfp800	TTATGCCGTAGTC TCTTTAG	MouseValid	465	Zfp800_465
MouseValid_MetsMIR_466_Zfp800	MetsMIR_466	Zfp800	GTAATCCTCTAAA GAGACTA	MouseValid	466	Zfp800_466
MouseValid_MetsMIR_467_Zfp800	MetsMIR_467	Zfp800	AAGTATGAAGTAG CCGACGT	MouseValid	467	Zfp800_467
MouseValid_MetsMIR_468_Zfp800	MetsMIR_468	Zfp800	CTTGACTCGGTCA CTTCAGC	MouseValid	468	Zfp800_468

Figure 18 (supplemental tables continued)

FIG. 18AV

MouseValid	MetsMIR 469	Zfp800	MetsMi R 469	Zfp800	TTATCACATCACG TTTCCGA	Mouse Valid	469	Zfp800 469
MouseValid	MetsMIR 470	Zfp800	MetsMi R 470	Zfp800	AGGCCGAAGTACA AGATCTA	Mouse Valid	470	Zfp800 470
MouseValid	MetsMIR 471	Zfyve28	MetsMi R 471	Zfyve28	ATCCGGCATGACA ACTTGGC	Mouse Valid	471	Zfyve28 471
MouseValid	MetsMIR 472	Zfyve28	MetsMi R 472	Zfyve28	GCGGGGCACTCT CCGTGACC	Mouse Valid	472	Zfyve28 472
MouseValid	MetsMIR 473	Zfyve28	MetsMi R 473	Zfyve28	GAGCATGGCCAT GCGGCCAC	Mouse Valid	473	Zfyve28 473
MouseValid	MetsMIR 474	Zfyve28	MetsMi R 474	Zfyve28	CAGCCAGTTCCGC TCCTGTC	Mouse Valid	474	Zfyve28 474
MouseValid	MetsMIR 475	Zfyve28	MetsMi R 475	Zfyve28	AAC TTGACGCAGA AATCCCG	Mouse Valid	475	Zfyve28 475
MouseValid	MetsMIR 476	Zfyve28	MetsMi R 476	Zfyve28	CAGGGAGTACTAT GTGCAGC	Mouse Valid	476	Zfyve28 476
MouseValid	MetsMIR 477	Zfyve28	MetsMi R 477	Zfyve28	GAGGTCAGACCC ACAAC TGC	Mouse Valid	477	Zfyve28 477
MouseValid	MetsMIR 478	Zfyve28	MetsMi R 478	Zfyve28	CTGCTGCACATAG TAGTCCC	Mouse Valid	478	Zfyve28 478
MouseValid	MetsMIR 479	Zfyve28	MetsMi R 479	Zfyve28	TCCTCAGGACCG TGCCCT	Mouse Valid	479	Zfyve28 479
MouseValid	MetsMIR 480	Zfyve28	MetsMi R 480	Zfyve28	TTGTCATGCCGGA TCTCCTC	Mouse Valid	480	Zfyve28 480
MouseValid	MetsMIR 481	mmu-mir-345	MetsMi R 481	mmu-mir-345	TGAAGCCGACACC CAAGTCC	Mouse Valid	481	mmu-mir-345 481
MouseValid	MetsMIR 482	mmu-mir-345	MetsMi R 482	mmu-mir-345	GACCCCTAGTCCA GTGCTTG	Mouse Valid	482	mmu-mir-345 482
MouseValid	MetsMIR 483	mmu-mir-345	MetsMi R 483	mmu-mir-345	CCAGTGCTTG TGG TGGCTAC	Mouse Valid	483	mmu-mir-345 483
MouseValid	MetsMIR 484	mmu-mir-345	MetsMi R 484	mmu-mir-345	GCTACTGGGCCCT GAAC TAG	Mouse Valid	484	mmu-mir-345 484
MouseValid	MetsMIR 485	mmu-mir-345	MetsMi R 485	mmu-mir-345	GAACTAGGGGTCT GGAGACC	Mouse Valid	485	mmu-mir-345 485
MouseValid	MetsMIR 486	mmu-mir-345	MetsMi R 486	mmu-mir-345	GACCTGGGTTTGA TCTCCAC	Mouse Valid	486	mmu-mir-345 486
MouseValid	MetsMIR 487	mmu-mir-345	MetsMi R 487	mmu-mir-345	CCAGTAGCCACCA CAAGCAC	Mouse Valid	487	mmu-mir-345 487
MouseValid	MetsMIR 488	mmu-mir-345	MetsMi R 488	mmu-mir-345	CACCACAAGCACT GGACTAG	Mouse Valid	488	mmu-mir-345 488
MouseValid	MetsMIR 489	mmu-mir-345	MetsMi R 489	mmu-mir-345	GGACTAGGGGTC AGCAGGCC	Mouse Valid	489	mmu-mir-345 489
MouseValid	MetsMIR 490	mmu-mir-345	MetsMi R 490	mmu-mir-345	TGGGCCCTGAAC T AGGGGTC	Mouse Valid	490	mmu-mir-345 490

Figure 18 (supplemental tables continued)

FIG. 18AW

MouseValid_MetsMIR_491_mmu-mir-152	MetsMI R_491	mmu-mir-152	CGCCGCTGTTCCC CGGGCCCT	Mouse Valid	491	mmu-mir-152_491
MouseValid_MetsMIR_492_mmu-mir-152	MetsMI R_492	mmu-mir-152	TTCTGTGATACAC TCCGACT	Mouse Valid	492	mmu-mir-152_492
MouseValid_MetsMIR_493_mmu-mir-152	MetsMI R_493	mmu-mir-152	ATACACTCCGACT CGGGCTC	Mouse Valid	493	mmu-mir-152_493
MouseValid_MetsMIR_494_mmu-mir-152	MetsMI R_494	mmu-mir-152	AGTCAGTGCATGA CAGAACT	Mouse Valid	494	mmu-mir-152_494
MouseValid_MetsMIR_495_mmu-mir-152	MetsMI R_495	mmu-mir-152	TGCATGACAGAAC TTGGGCC	Mouse Valid	495	mmu-mir-152_495
MouseValid_MetsMIR_496_mmu-mir-152	MetsMI R_496	mmu-mir-152	GACTGCTCCAGAG CCCAGT	Mouse Valid	496	mmu-mir-152_496
MouseValid_MetsMIR_497_mmu-mir-152	MetsMI R_497	mmu-mir-152	CGGAGTGTATCAC AGAACTT	Mouse Valid	497	mmu-mir-152_497
MouseValid_MetsMIR_498_mmu-mir-152	MetsMI R_498	mmu-mir-152	TGTATCACAGAAC CTAGGCC	Mouse Valid	498	mmu-mir-152_498
MouseValid_MetsMIR_499_mmu-mir-152	MetsMI R_499	mmu-mir-152	TCTGTGATACACT CCGACTC	Mouse Valid	499	mmu-mir-152_499
MouseValid_MetsMIR_500_mmu-mir-152	MetsMI R_500	mmu-mir-152	GTCAGTGCATGAC AGAACTT	Mouse Valid	500	mmu-mir-152_500
MouseValid_MetsMIR_501_mmu-mir-367	MetsMI R_501	mmu-mir-367	TTGGAATTGCACT TTAGCAA	Mouse Valid	501	mmu-mir-367_501
MouseValid_MetsMIR_502_mmu-mir-367	MetsMI R_502	mmu-mir-367	TTGCACTTTAGCA ATGGTGA	Mouse Valid	502	mmu-mir-367_502
MouseValid_MetsMIR_503_mmu-mir-367	MetsMI R_503	mmu-mir-367	TTGCATGTTAGCA ACAGTGA	Mouse Valid	503	mmu-mir-367_503
MouseValid_MetsMIR_504_mmu-mir-367	MetsMI R_504	mmu-mir-367	CAACTCTGTTGAA TAGAAAT	Mouse Valid	504	mmu-mir-367_504
MouseValid_MetsMIR_505_mmu-mir-6537	MetsMI R_505	mmu-mir-6537	TTCTGTTCTACA GGCAGAG	Mouse Valid	505	mmu-mir-6537_505
MouseValid_MetsMIR_506_mmu-mir-6537	MetsMI R_506	mmu-mir-6537	GTTCTACAGGCAG AGAGGGT	Mouse Valid	506	mmu-mir-6537_506
MouseValid_MetsMIR_507_mmu-mir-6537	MetsMI R_507	mmu-mir-6537	AGTGGGAGAACT CACCCTT	Mouse Valid	507	mmu-mir-6537_507
MouseValid_MetsMIR_508_mmu-mir-6537	MetsMI R_508	mmu-mir-6537	CGCTAGGAACTC TTCTCTA	Mouse Valid	508	mmu-mir-6537_508
MouseValid_MetsMIR_509_mmu-mir-6537	MetsMI R_509	mmu-mir-6537	TCTCCCACTGGTG ACCAGGA	Mouse Valid	509	mmu-mir-6537_509
MouseValid_MetsMIR_510_mmu-mir-6537	MetsMI R_510	mmu-mir-6537	TGACCAGGAAGGT TATCAGG	Mouse Valid	510	mmu-mir-6537_510
MouseValid_MetsMIR_511_mmu-mir-6537	MetsMI R_511	mmu-mir-6537	AGTTTCTCCCACT GGTGACC	Mouse Valid	511	mmu-mir-6537_511
MouseValid_MetsMIR_512_mmu-mir-6537	MetsMI R_512	mmu-mir-6537	TAGCGGTGAGTTT CTCCAC	Mouse Valid	512	mmu-mir-6537_512

Figure 18 (supplemental tables continued)

FIG. 18AX

MouseValid_MetsMIR_513_mmu-mir-6537	MetsMIR_513	mmu-mir-6537	CTCCACCTGATA ACCTTCC	Mouse Valid	513	mmu-mir-6537_513
MouseValid_MetsMIR_514_mmu-mir-6537	MetsMIR_514	mmu-mir-6537	ATAACCTTCCTGG TCACCAG	Mouse Valid	514	mmu-mir-6537_514
MouseValid_MetsMIR_515_mmu-mir-6412	MetsMIR_515	mmu-mir-6412	ACTGCCTAGTTAG CTTGATG	Mouse Valid	515	mmu-mir-6412_515
MouseValid_MetsMIR_516_mmu-mir-6412	MetsMIR_516	mmu-mir-6412	TACTGCACCTCTA TAATCCC	Mouse Valid	516	mmu-mir-6412_516
MouseValid_MetsMIR_517_mmu-mir-6412	MetsMIR_517	mmu-mir-6412	CTCCTATAATCCC AGGGTCT	Mouse Valid	517	mmu-mir-6412_517
MouseValid_MetsMIR_518_mmu-mir-6412	MetsMIR_518	mmu-mir-6412	CTATAATCCCAGG GTCTTGG	Mouse Valid	518	mmu-mir-6412_518
MouseValid_MetsMIR_519_mmu-mir-6412	MetsMIR_519	mmu-mir-6412	TCCCAGGGTCTTG GAGGTCT	Mouse Valid	519	mmu-mir-6412_519
MouseValid_MetsMIR_520_mmu-mir-6412	MetsMIR_520	mmu-mir-6412	AGGGTCTTGGAG GTCTAGGC	Mouse Valid	520	mmu-mir-6412_520
MouseValid_MetsMIR_521_mmu-mir-6412	MetsMIR_521	mmu-mir-6412	GAGGTCTAGGCA GGATGATT	Mouse Valid	521	mmu-mir-6412_521
MouseValid_MetsMIR_522_mmu-mir-6412	MetsMIR_522	mmu-mir-6412	ACTGCACCTCTAT AATCCCA	Mouse Valid	522	mmu-mir-6412_522
MouseValid_MetsMIR_523_mmu-mir-6412	MetsMIR_523	mmu-mir-6412	CTCCAAGACCCTG GGATTAT	Mouse Valid	523	mmu-mir-6412_523
MouseValid_MetsMIR_524_mmu-mir-6412	MetsMIR_524	mmu-mir-6412	TGCCTAGACCTCC AAGACCC	Mouse Valid	524	mmu-mir-6412_524
MouseValid_MetsMIR_525_NonTargetingControlGuideForMouse_0001	MetsMIR_525	NonTargetingControlGuideForMouse_0001	GCGAGGTATTCG GCTCCGCG	Mouse Valid	525	NonTargetingControlGuideForMouse_0001_525
MouseValid_MetsMIR_526_NonTargetingControlGuideForMouse_0002	MetsMIR_526	NonTargetingControlGuideForMouse_0002	GCTTTACGGAGG TTCCAGC	Mouse Valid	526	NonTargetingControlGuideForMouse_0002_526
MouseValid_MetsMIR_527_NonTargetingControlGuideForMouse_0003	MetsMIR_527	NonTargetingControlGuideForMouse_0003	ATGTTGCAGTTCG GCTCGAT	Mouse Valid	527	NonTargetingControlGuideForMouse_0003_527
MouseValid_MetsMIR_528_NonTargetingControlGuideForMouse_0004	MetsMIR_528	NonTargetingControlGuideForMouse_0004	ACGTGTAAGGCGA ACGCCCT	Mouse Valid	528	NonTargetingControlGuideForMouse_0004_528
MouseValid_MetsMIR_529_NonTargetingControlGuideForMouse_0005	MetsMIR_529	NonTargetingControlGuideForMouse_0005	GACTCCGGGTACT AAATGTC	Mouse Valid	529	NonTargetingControlGuideForMouse_0005_529
MouseValid_MetsMIR_530_NonTargetingControlGuideForMouse_0006	MetsMIR_530	NonTargetingControlGuideForMouse_0006	CCGCGCCGTTAG GGAACGAG	Mouse Valid	530	NonTargetingControlGuideForMouse_0006_530
MouseValid_MetsMIR_531_NonTargetingControlGuideForMouse_0007	MetsMIR_531	NonTargetingControlGuideForMouse_0007	ATTGTTGACCGGT CTACGGG	Mouse Valid	531	NonTargetingControlGuideForMouse_0007_531
MouseValid_MetsMIR_532_NonTargetingControlGuideForMouse_0008	MetsMIR_532	NonTargetingControlGuideForMouse_0008	ACCCATCGGGTG CGATATGG	Mouse Valid	532	NonTargetingControlGuideForMouse_0008_532
MouseValid_MetsMIR_533_NonTargetingControlGuideForMouse_0009	MetsMIR_533	NonTargetingControlGuideForMouse_0009	CGGGCGTCACCT GCTAGTAA	Mouse Valid	533	NonTargetingControlGuideForMouse_0009_533
MouseValid_MetsMIR_534_NonTargetingControlGuideForMouse_0010	MetsMIR_534	NonTargetingControlGuideForMouse_0010	GCTTCTACTCGCA ACGTATT	Mouse Valid	534	NonTargetingControlGuideForMouse_0010_534

Figure 18 (supplemental tables continued)

FIG. 18AY

MouseValid_MetsMIR_535_NonTargetingControlGuideForMouse_0011	MetsMIR_535	NonTargetingControlGuideForMouse_0011	TACAGTTATACGT CGCGGTG	MouseValid	535	NonTargetingControlGuideForMouse_0011_535
MouseValid_MetsMIR_536_NonTargetingControlGuideForMouse_0012	MetsMIR_536	NonTargetingControlGuideForMouse_0012	AAGCACAAAGACG GTCCGCC	MouseValid	536	NonTargetingControlGuideForMouse_0012_536
MouseValid_MetsMIR_537_NonTargetingControlGuideForMouse_0013	MetsMIR_537	NonTargetingControlGuideForMouse_0013	CAGCCACCGCAC CGCGGTAA	MouseValid	537	NonTargetingControlGuideForMouse_0013_537
MouseValid_MetsMIR_538_NonTargetingControlGuideForMouse_0014	MetsMIR_538	NonTargetingControlGuideForMouse_0014	GTCAAGCCGAAC GCTGCCGG	MouseValid	538	NonTargetingControlGuideForMouse_0014_538
MouseValid_MetsMIR_539_NonTargetingControlGuideForMouse_0015	MetsMIR_539	NonTargetingControlGuideForMouse_0015	CCTTAGACCGGGT GTACCTC	MouseValid	539	NonTargetingControlGuideForMouse_0015_539
MouseValid_MetsMIR_540_NonTargetingControlGuideForMouse_0016	MetsMIR_540	NonTargetingControlGuideForMouse_0016	AAGTCTATGCGGG GCTCGTA	MouseValid	540	NonTargetingControlGuideForMouse_0016_540
MouseValid_MetsMIR_541_NonTargetingControlGuideForMouse_0017	MetsMIR_541	NonTargetingControlGuideForMouse_0017	TTGTCAACTTCGG CCAACGC	MouseValid	541	NonTargetingControlGuideForMouse_0017_541
MouseValid_MetsMIR_542_NonTargetingControlGuideForMouse_0018	MetsMIR_542	NonTargetingControlGuideForMouse_0018	ATAGATGTCTACG GCCCGTT	MouseValid	542	NonTargetingControlGuideForMouse_0018_542
MouseValid_MetsMIR_543_NonTargetingControlGuideForMouse_0019	MetsMIR_543	NonTargetingControlGuideForMouse_0019	CTCGGGCTATTCA GCGATAG	MouseValid	543	NonTargetingControlGuideForMouse_0019_543
MouseValid_MetsMIR_544_NonTargetingControlGuideForMouse_0020	MetsMIR_544	NonTargetingControlGuideForMouse_0020	GCGGTTACCGCG AAAACCAT	MouseValid	544	NonTargetingControlGuideForMouse_0020_544
MouseValid_MetsMIR_545_NonTargetingControlGuideForMouse_0021	MetsMIR_545	NonTargetingControlGuideForMouse_0021	ACCAACGCTACGA TCCCGGA	MouseValid	545	NonTargetingControlGuideForMouse_0021_545
MouseValid_MetsMIR_546_NonTargetingControlGuideForMouse_0022	MetsMIR_546	NonTargetingControlGuideForMouse_0022	CCCTATATGCGAG ATCCATA	MouseValid	546	NonTargetingControlGuideForMouse_0022_546
MouseValid_MetsMIR_547_NonTargetingControlGuideForMouse_0023	MetsMIR_547	NonTargetingControlGuideForMouse_0023	AGAAAGGCACGT GCGACGTC	MouseValid	547	NonTargetingControlGuideForMouse_0023_547
MouseValid_MetsMIR_548_NonTargetingControlGuideForMouse_0024	MetsMIR_548	NonTargetingControlGuideForMouse_0024	GACCAACCTTACG GTAACCTC	MouseValid	548	NonTargetingControlGuideForMouse_0024_548
MouseValid_MetsMIR_549_NonTargetingControlGuideForMouse_0025	MetsMIR_549	NonTargetingControlGuideForMouse_0025	ATTATTCCTCCGG ATGACGA	MouseValid	549	NonTargetingControlGuideForMouse_0025_549
MouseValid_MetsMIR_550_NonTargetingControlGuideForMouse_0026	MetsMIR_550	NonTargetingControlGuideForMouse_0026	TCTCAGTTCGTAG CGAACGA	MouseValid	550	NonTargetingControlGuideForMouse_0026_550
MouseValid_MetsMIR_551_NonTargetingControlGuideForMouse_0027	MetsMIR_551	NonTargetingControlGuideForMouse_0027	TCCCGGGAGGTA CGGTGTAC	MouseValid	551	NonTargetingControlGuideForMouse_0027_551
MouseValid_MetsMIR_552_NonTargetingControlGuideForMouse_0028	MetsMIR_552	NonTargetingControlGuideForMouse_0028	GGGGATGGCCTT ACGTCCGG	MouseValid	552	NonTargetingControlGuideForMouse_0028_552
MouseValid_MetsMIR_553_NonTargetingControlGuideForMouse_0029	MetsMIR_553	NonTargetingControlGuideForMouse_0029	GTATCCTCCTTAC GGCCCCGT	MouseValid	553	NonTargetingControlGuideForMouse_0029_553
MouseValid_MetsMIR_554_NonTargetingControlGuideForMouse_0030	MetsMIR_554	NonTargetingControlGuideForMouse_0030	GCTCGGACCTTTT AGACGTC	MouseValid	554	NonTargetingControlGuideForMouse_0030_554
MouseValid_MetsMIR_555_NonTargetingControlGuideForMouse_0031	MetsMIR_555	NonTargetingControlGuideForMouse_0031	GTGGTTACGTTAA CGACTAC	MouseValid	555	NonTargetingControlGuideForMouse_0031_555
MouseValid_MetsMIR_556_NonTargetingControlGuideForMouse_0032	MetsMIR_556	NonTargetingControlGuideForMouse_0032	TGCAACGATGTTT ACGGTAC	MouseValid	556	NonTargetingControlGuideForMouse_0032_556

Figure 18 (supplemental tables continued)

FIG. 18AZ

MouseValid_MetsMIR_557_NonTargetingControlGuideForMouse_0033	MetsMIR_557	NonTargetingControlGuideForMouse_0033	AACGGGCGCAATACCTTTT	MouseValid	557	NonTargetingControlGuideForMouse_0033_557
MouseValid_MetsMIR_558_NonTargetingControlGuideForMouse_0034	MetsMIR_558	NonTargetingControlGuideForMouse_0034	TCGTGTCTAGCTATCGAGTG	MouseValid	558	NonTargetingControlGuideForMouse_0034_558
MouseValid_MetsMIR_559_NonTargetingControlGuideForMouse_0035	MetsMIR_559	NonTargetingControlGuideForMouse_0035	GACGTCTAATTC TGGCCGT	MouseValid	559	NonTargetingControlGuideForMouse_0035_559
MouseValid_MetsMIR_560_NonTargetingControlGuideForMouse_0036	MetsMIR_560	NonTargetingControlGuideForMouse_0036	TAGTCCTAGTTAG ATTCGCG	MouseValid	560	NonTargetingControlGuideForMouse_0036_560
MouseValid_MetsMIR_561_NonTargetingControlGuideForMouse_0037	MetsMIR_561	NonTargetingControlGuideForMouse_0037	AAGGCCCTAACAC GTCGACC	MouseValid	561	NonTargetingControlGuideForMouse_0037_561
MouseValid_MetsMIR_562_NonTargetingControlGuideForMouse_0038	MetsMIR_562	NonTargetingControlGuideForMouse_0038	ATTCGTGCATCGC GGGGTTT	MouseValid	562	NonTargetingControlGuideForMouse_0038_562
MouseValid_MetsMIR_563_NonTargetingControlGuideForMouse_0039	MetsMIR_563	NonTargetingControlGuideForMouse_0039	CACCTCGCGTCAT ATCACTA	MouseValid	563	NonTargetingControlGuideForMouse_0039_563
MouseValid_MetsMIR_564_NonTargetingControlGuideForMouse_0040	MetsMIR_564	NonTargetingControlGuideForMouse_0040	GACCTCGCAATTG AGCGCTC	MouseValid	564	NonTargetingControlGuideForMouse_0040_564
MouseValid_MetsMIR_565_NonTargetingControlGuideForMouse_0041	MetsMIR_565	NonTargetingControlGuideForMouse_0041	TGTATCCACCGTG ACCCGGT	MouseValid	565	NonTargetingControlGuideForMouse_0041_565
MouseValid_MetsMIR_566_NonTargetingControlGuideForMouse_0042	MetsMIR_566	NonTargetingControlGuideForMouse_0042	GATCTTACCACCTC GTCGTAG	MouseValid	566	NonTargetingControlGuideForMouse_0042_566
MouseValid_MetsMIR_567_NonTargetingControlGuideForMouse_0043	MetsMIR_567	NonTargetingControlGuideForMouse_0043	TTCCGCGGCGAC GAAATGCA	MouseValid	567	NonTargetingControlGuideForMouse_0043_567
MouseValid_MetsMIR_568_NonTargetingControlGuideForMouse_0044	MetsMIR_568	NonTargetingControlGuideForMouse_0044	AGCACTAGGATCG CGGCCTT	MouseValid	568	NonTargetingControlGuideForMouse_0044_568
MouseValid_MetsMIR_569_NonTargetingControlGuideForMouse_0045	MetsMIR_569	NonTargetingControlGuideForMouse_0045	GCGTATCTACCCT ACCGCCG	MouseValid	569	NonTargetingControlGuideForMouse_0045_569
MouseValid_MetsMIR_570_NonTargetingControlGuideForMouse_0046	MetsMIR_570	NonTargetingControlGuideForMouse_0046	GCGCGAGGGCAC CGACAAGT	MouseValid	570	NonTargetingControlGuideForMouse_0046_570
MouseValid_MetsMIR_571_NonTargetingControlGuideForMouse_0047	MetsMIR_571	NonTargetingControlGuideForMouse_0047	CGCGTATATGTCA CACGGCA	MouseValid	571	NonTargetingControlGuideForMouse_0047_571
MouseValid_MetsMIR_572_NonTargetingControlGuideForMouse_0048	MetsMIR_572	NonTargetingControlGuideForMouse_0048	TAAACCGATCAC GATACGA	MouseValid	572	NonTargetingControlGuideForMouse_0048_572
MouseValid_MetsMIR_573_NonTargetingControlGuideForMouse_0049	MetsMIR_573	NonTargetingControlGuideForMouse_0049	TCCTGCGCGATGA CCGTCGG	MouseValid	573	NonTargetingControlGuideForMouse_0049_573
MouseValid_MetsMIR_574_NonTargetingControlGuideForMouse_0050	MetsMIR_574	NonTargetingControlGuideForMouse_0050	TGGCCACAAAGT GCGATAT	MouseValid	574	NonTargetingControlGuideForMouse_0050_574
MouseValid_MetsMIR_575_NonTargetingControlGuideForMouse_0051	MetsMIR_575	NonTargetingControlGuideForMouse_0051	GGGGTAGGCCTA ATTACGGA	MouseValid	575	NonTargetingControlGuideForMouse_0051_575
MouseValid_MetsMIR_576_NonTargetingControlGuideForMouse_0052	MetsMIR_576	NonTargetingControlGuideForMouse_0052	CGCTTCGTCTCTC GCAAACA	MouseValid	576	NonTargetingControlGuideForMouse_0052_576
MouseValid_MetsMIR_577_NonTargetingControlGuideForMouse_0053	MetsMIR_577	NonTargetingControlGuideForMouse_0053	TTATGTGCCTCTC GCGGCAT	MouseValid	577	NonTargetingControlGuideForMouse_0053_577
MouseValid_MetsMIR_578_NonTargetingControlGuideForMouse_0054	MetsMIR_578	NonTargetingControlGuideForMouse_0054	CCGTTCTGACGAC GCTAAAG	MouseValid	578	NonTargetingControlGuideForMouse_0054_578

Figure 18 (supplemental tables continued)

FIG. 18BA

MouseValid_MetsMIR_579_NonTargetingControlGuideForMouse_0055	MetsMIR_579	NonTargetingControlGuideForMouse_0055	AAACTCCCGTGTCAACCGAT	MouseValid	579	NonTargetingControlGuideForMouse_0055_579
MouseValid_MetsMIR_580_NonTargetingControlGuideForMouse_0056	MetsMIR_580	NonTargetingControlGuideForMouse_0056	GTTTGCGAGTCAAAGTACGC	MouseValid	580	NonTargetingControlGuideForMouse_0056_580
MouseValid_MetsMIR_581_NonTargetingControlGuideForMouse_0057	MetsMIR_581	NonTargetingControlGuideForMouse_0057	GGGTGCACACGCGCGCCTAT	MouseValid	581	NonTargetingControlGuideForMouse_0057_581
MouseValid_MetsMIR_582_NonTargetingControlGuideForMouse_0058	MetsMIR_582	NonTargetingControlGuideForMouse_0058	CTTTATACCGCGCGTCGCA	MouseValid	582	NonTargetingControlGuideForMouse_0058_582
MouseValid_MetsMIR_583_NonTargetingControlGuideForMouse_0059	MetsMIR_583	NonTargetingControlGuideForMouse_0059	TCAAACGCCCGGGCGCCCCA	MouseValid	583	NonTargetingControlGuideForMouse_0059_583
MouseValid_MetsMIR_584_NonTargetingControlGuideForMouse_0060	MetsMIR_584	NonTargetingControlGuideForMouse_0060	GACCGGTGTGTTTACGCGTG	MouseValid	584	NonTargetingControlGuideForMouse_0060_584
MouseValid_MetsMIR_585_NonTargetingControlGuideForMouse_0061	MetsMIR_585	NonTargetingControlGuideForMouse_0061	CGGTGAGGATCCCTCTTGC	MouseValid	585	NonTargetingControlGuideForMouse_0061_585
MouseValid_MetsMIR_586_NonTargetingControlGuideForMouse_0062	MetsMIR_586	NonTargetingControlGuideForMouse_0062	ACGGTCCCAACGACGCGCGG	MouseValid	586	NonTargetingControlGuideForMouse_0062_586
MouseValid_MetsMIR_587_NonTargetingControlGuideForMouse_0063	MetsMIR_587	NonTargetingControlGuideForMouse_0063	CGCTTCATTGCCGAACGCT	MouseValid	587	NonTargetingControlGuideForMouse_0063_587
MouseValid_MetsMIR_588_NonTargetingControlGuideForMouse_0064	MetsMIR_588	NonTargetingControlGuideForMouse_0064	TTATACCGTCACGAGTGCCC	MouseValid	588	NonTargetingControlGuideForMouse_0064_588
MouseValid_MetsMIR_589_NonTargetingControlGuideForMouse_0065	MetsMIR_589	NonTargetingControlGuideForMouse_0065	TGCACCGGTACGGGCGACTC	MouseValid	589	NonTargetingControlGuideForMouse_0065_589
MouseValid_MetsMIR_590_NonTargetingControlGuideForMouse_0066	MetsMIR_590	NonTargetingControlGuideForMouse_0066	GAGCGTACTTCGGCTCTAA	MouseValid	590	NonTargetingControlGuideForMouse_0066_590
MouseValid_MetsMIR_591_NonTargetingControlGuideForMouse_0067	MetsMIR_591	NonTargetingControlGuideForMouse_0067	TTCGAGACGGATACGTCTGC	MouseValid	591	NonTargetingControlGuideForMouse_0067_591
MouseValid_MetsMIR_592_NonTargetingControlGuideForMouse_0068	MetsMIR_592	NonTargetingControlGuideForMouse_0068	TAGCCCGTCGAGTACTCCCC	MouseValid	592	NonTargetingControlGuideForMouse_0068_592
MouseValid_MetsMIR_593_NonTargetingControlGuideForMouse_0069	MetsMIR_593	NonTargetingControlGuideForMouse_0069	AGCTTAGATCGTGCSTCGTA	MouseValid	593	NonTargetingControlGuideForMouse_0069_593
MouseValid_MetsMIR_594_NonTargetingControlGuideForMouse_0070	MetsMIR_594	NonTargetingControlGuideForMouse_0070	CGCCGGATGACTGGTGGATA	MouseValid	594	NonTargetingControlGuideForMouse_0070_594
MouseValid_MetsMIR_595_NonTargetingControlGuideForMouse_0071	MetsMIR_595	NonTargetingControlGuideForMouse_0071	CCGTTTCGCCTATTGGTGC	MouseValid	595	NonTargetingControlGuideForMouse_0071_595
MouseValid_MetsMIR_596_NonTargetingControlGuideForMouse_0072	MetsMIR_596	NonTargetingControlGuideForMouse_0072	TTCCACGTGCGCGTATTGG	MouseValid	596	NonTargetingControlGuideForMouse_0072_596
MouseValid_MetsMIR_597_NonTargetingControlGuideForMouse_0073	MetsMIR_597	NonTargetingControlGuideForMouse_0073	TTCATTGTTCGCGCGAACA	MouseValid	597	NonTargetingControlGuideForMouse_0073_597
MouseValid_MetsMIR_598_NonTargetingControlGuideForMouse_0074	MetsMIR_598	NonTargetingControlGuideForMouse_0074	GCGGATCTGACGGTTACTTA	MouseValid	598	NonTargetingControlGuideForMouse_0074_598
MouseValid_MetsMIR_599_NonTargetingControlGuideForMouse_0075	MetsMIR_599	NonTargetingControlGuideForMouse_0075	AATGAAGCACCGATTGCGGA	MouseValid	599	NonTargetingControlGuideForMouse_0075_599
MouseValid_MetsMIR_600_NonTargetingControlGuideForMouse_0076	MetsMIR_600	NonTargetingControlGuideForMouse_0076	TCCTATAATTGAGCGAACGG	MouseValid	600	NonTargetingControlGuideForMouse_0076_600

Figure 18 (supplemental tables continued)

FIG. 18BB

MouseValid_MetsMIR_601_NonTargetingControlGuideForMouse_0077	MetsMIR_601	NonTargetingControlGuideForMouse_0077	GCTACAGATTTCGAG	MouseValid_601	NonTargetingControlGuideForMouse_0077_601
MouseValid_MetsMIR_602_NonTargetingControlGuideForMouse_0078	MetsMIR_602	NonTargetingControlGuideForMouse_0078	GGACTGAAACCGA	MouseValid_602	NonTargetingControlGuideForMouse_0078_602
MouseValid_MetsMIR_603_NonTargetingControlGuideForMouse_0079	MetsMIR_603	NonTargetingControlGuideForMouse_0079	GCAATTCTGCAAC	MouseValid_603	NonTargetingControlGuideForMouse_0079_603
MouseValid_MetsMIR_604_NonTargetingControlGuideForMouse_0080	MetsMIR_604	NonTargetingControlGuideForMouse_0080	CACCCGACTCGG	MouseValid_604	NonTargetingControlGuideForMouse_0080_604
MouseValid_MetsMIR_605_NonTargetingControlGuideForMouse_0081	MetsMIR_605	NonTargetingControlGuideForMouse_0081	CAACCGTGCGATG	MouseValid_605	NonTargetingControlGuideForMouse_0081_605
MouseValid_MetsMIR_606_NonTargetingControlGuideForMouse_0082	MetsMIR_606	NonTargetingControlGuideForMouse_0082	TGATGGACGCGAT	MouseValid_606	NonTargetingControlGuideForMouse_0082_606
MouseValid_MetsMIR_607_NonTargetingControlGuideForMouse_0083	MetsMIR_607	NonTargetingControlGuideForMouse_0083	TACCCCTTGAGGG	MouseValid_607	NonTargetingControlGuideForMouse_0083_607
MouseValid_MetsMIR_608_NonTargetingControlGuideForMouse_0084	MetsMIR_608	NonTargetingControlGuideForMouse_0084	CCGTATACGTATC	MouseValid_608	NonTargetingControlGuideForMouse_0084_608
MouseValid_MetsMIR_609_NonTargetingControlGuideForMouse_0085	MetsMIR_609	NonTargetingControlGuideForMouse_0085	ACAGCGGCGCTAAT	MouseValid_609	NonTargetingControlGuideForMouse_0085_609
MouseValid_MetsMIR_610_NonTargetingControlGuideForMouse_0086	MetsMIR_610	NonTargetingControlGuideForMouse_0086	TATCGAACTGCAC	MouseValid_610	NonTargetingControlGuideForMouse_0086_610
MouseValid_MetsMIR_611_NonTargetingControlGuideForMouse_0087	MetsMIR_611	NonTargetingControlGuideForMouse_0087	TGTCGGACCCATT	MouseValid_611	NonTargetingControlGuideForMouse_0087_611
MouseValid_MetsMIR_612_NonTargetingControlGuideForMouse_0088	MetsMIR_612	NonTargetingControlGuideForMouse_0088	AAACCTAGCGTAG	MouseValid_612	NonTargetingControlGuideForMouse_0088_612
MouseValid_MetsMIR_613_NonTargetingControlGuideForMouse_0089	MetsMIR_613	NonTargetingControlGuideForMouse_0089	TAGATGGTGCTTC	MouseValid_613	NonTargetingControlGuideForMouse_0089_613
MouseValid_MetsMIR_614_NonTargetingControlGuideForMouse_0090	MetsMIR_614	NonTargetingControlGuideForMouse_0090	CATCGTAACACAC	MouseValid_614	NonTargetingControlGuideForMouse_0090_614
MouseValid_MetsMIR_615_NonTargetingControlGuideForMouse_0091	MetsMIR_615	NonTargetingControlGuideForMouse_0091	CAACGATCAGGC	MouseValid_615	NonTargetingControlGuideForMouse_0091_615
MouseValid_MetsMIR_616_NonTargetingControlGuideForMouse_0092	MetsMIR_616	NonTargetingControlGuideForMouse_0092	ACAGCAGGGGCC	MouseValid_616	NonTargetingControlGuideForMouse_0092_616
MouseValid_MetsMIR_617_NonTargetingControlGuideForMouse_0093	MetsMIR_617	NonTargetingControlGuideForMouse_0093	AGCCGGGCTTTCC	MouseValid_617	NonTargetingControlGuideForMouse_0093_617
MouseValid_MetsMIR_618_NonTargetingControlGuideForMouse_0094	MetsMIR_618	NonTargetingControlGuideForMouse_0094	ACGGTCCCTCTCG	MouseValid_618	NonTargetingControlGuideForMouse_0094_618
MouseValid_MetsMIR_619_NonTargetingControlGuideForMouse_0095	MetsMIR_619	NonTargetingControlGuideForMouse_0095	TCAATAGTTCTGC	MouseValid_619	NonTargetingControlGuideForMouse_0095_619
MouseValid_MetsMIR_620_NonTargetingControlGuideForMouse_0096	MetsMIR_620	NonTargetingControlGuideForMouse_0096	TCGTATACATCCC	MouseValid_620	NonTargetingControlGuideForMouse_0096_620
MouseValid_MetsMIR_621_NonTargetingControlGuideForMouse_0097	MetsMIR_621	NonTargetingControlGuideForMouse_0097	CATTCGGTCCGTT	MouseValid_621	NonTargetingControlGuideForMouse_0097_621
MouseValid_MetsMIR_622_NonTargetingControlGuideForMouse_0098	MetsMIR_622	NonTargetingControlGuideForMouse_0098	TTCGGCTCAATGG	MouseValid_622	NonTargetingControlGuideForMouse_0098_622

Figure 18 (supplemental tables continued)

FIG. 18BC

MouseValid_MetsMIR_623_NonTargetingControlGuideForMouse_0099	MetsMIR_623	NonTargetingControlGuideForMouse_0099	CATACCCGCGCCGTGACTCC	MouseValid	623	NonTargetingControlGuideForMouse_0099_623
MouseValid_MetsMIR_624_NonTargetingControlGuideForMouse_0100	MetsMIR_624	NonTargetingControlGuideForMouse_0100	GGACGGATGGGACGACTAGT	MouseValid	624	NonTargetingControlGuideForMouse_0100_624

Mouse Non-targeting control guides

long_name	uid	gene	seq	Pool name
MouseDummy_MetsMIR_625_NonTargetingControlGuideForMouse_0001	MetsMIR_625	NonTargetingControlGuideForMouse_0001	GCGAGGTATTCGGCTCCGCG	MouseDummy
MouseDummy_MetsMIR_626_NonTargetingControlGuideForMouse_0002	MetsMIR_626	NonTargetingControlGuideForMouse_0002	GCTTTCACGGAGGTTTCGACG	MouseDummy
MouseDummy_MetsMIR_627_NonTargetingControlGuideForMouse_0003	MetsMIR_627	NonTargetingControlGuideForMouse_0003	ATGTTGCAGTTCGGCTCGAT	MouseDummy
MouseDummy_MetsMIR_628_NonTargetingControlGuideForMouse_0004	MetsMIR_628	NonTargetingControlGuideForMouse_0004	ACGTGTAAGCGCAACGCCTT	MouseDummy
MouseDummy_MetsMIR_629_NonTargetingControlGuideForMouse_0005	MetsMIR_629	NonTargetingControlGuideForMouse_0005	GACTCCGGGTACTAAATGTC	MouseDummy
MouseDummy_MetsMIR_630_NonTargetingControlGuideForMouse_0006	MetsMIR_630	NonTargetingControlGuideForMouse_0006	CCGCGCCGTTAGGGAACGAG	MouseDummy
MouseDummy_MetsMIR_631_NonTargetingControlGuideForMouse_0007	MetsMIR_631	NonTargetingControlGuideForMouse_0007	ATTGTTGACCGTCTACGGG	MouseDummy
MouseDummy_MetsMIR_632_NonTargetingControlGuideForMouse_0008	MetsMIR_632	NonTargetingControlGuideForMouse_0008	ACCCATCGGGTGCGATATGG	MouseDummy
MouseDummy_MetsMIR_633_NonTargetingControlGuideForMouse_0009	MetsMIR_633	NonTargetingControlGuideForMouse_0009	CGGGCGTCACCTGCTAGTAA	MouseDummy
MouseDummy_MetsMIR_634_NonTargetingControlGuideForMouse_0010	MetsMIR_634	NonTargetingControlGuideForMouse_0010	GCTTCTACTCGCAACGTATT	MouseDummy
MouseDummy_MetsMIR_635_NonTargetingControlGuideForMouse_0011	MetsMIR_635	NonTargetingControlGuideForMouse_0011	TACAGTTATACGTCGCGGTG	MouseDummy
MouseDummy_MetsMIR_636_NonTargetingControlGuideForMouse_0012	MetsMIR_636	NonTargetingControlGuideForMouse_0012	AAGCACAGAAGCGGTCCGCC	MouseDummy
MouseDummy_MetsMIR_637_NonTargetingControlGuideForMouse_0013	MetsMIR_637	NonTargetingControlGuideForMouse_0013	CAGCCACCGCACCGGCGTAA	MouseDummy
MouseDummy_MetsMIR_638_NonTargetingControlGuideForMouse_0014	MetsMIR_638	NonTargetingControlGuideForMouse_0014	GTCAAGCCGAACGCTGCGG	MouseDummy
MouseDummy_MetsMIR_639_NonTargetingControlGuideForMouse_0015	MetsMIR_639	NonTargetingControlGuideForMouse_0015	CCTTAGACCGGGTGATCTC	MouseDummy
MouseDummy_MetsMIR_640_NonTargetingControlGuideForMouse_0016	MetsMIR_640	NonTargetingControlGuideForMouse_0016	AAGTCTATGCGGGGCTCGTA	MouseDummy
MouseDummy_MetsMIR_641_NonTargetingControlGuideForMouse_0017	MetsMIR_641	NonTargetingControlGuideForMouse_0017	TTGTCAACTTCGGCCAAAGC	MouseDummy

Figure 18 (supplemental tables continued)

FIG. 18BD

MouseDummy_MetsMIR_642_NonTargetingControlGuideFor Mouse_0018	MetsMIR_642	NonTargetingControlGuideForMo use_0018	ATAGATGTCTACGCGCC GTT	MouseDum my
MouseDummy_MetsMIR_643_NonTargetingControlGuideFor Mouse_0019	MetsMIR_643	NonTargetingControlGuideForMo use_0019	CTCGGGCTATTACGCGA TAG	MouseDum my
MouseDummy_MetsMIR_644_NonTargetingControlGuideFor Mouse_0020	MetsMIR_644	NonTargetingControlGuideForMo use_0020	GCGGTTACCGCGAAAAC CAT	MouseDum my
MouseDummy_MetsMIR_645_NonTargetingControlGuideFor Mouse_0021	MetsMIR_645	NonTargetingControlGuideForMo use_0021	ACCAACGCTACGATCCC GGA	MouseDum my
MouseDummy_MetsMIR_646_NonTargetingControlGuideFor Mouse_0022	MetsMIR_646	NonTargetingControlGuideForMo use_0022	CCCTATATGCGAGATCC ATA	MouseDum my
MouseDummy_MetsMIR_647_NonTargetingControlGuideFor Mouse_0023	MetsMIR_647	NonTargetingControlGuideForMo use_0023	AGAAAGGCACGTGCGAC GTC	MouseDum my
MouseDummy_MetsMIR_648_NonTargetingControlGuideFor Mouse_0024	MetsMIR_648	NonTargetingControlGuideForMo use_0024	GACCAACCTTACGGTAA CTC	MouseDum my
MouseDummy_MetsMIR_649_NonTargetingControlGuideFor Mouse_0025	MetsMIR_649	NonTargetingControlGuideForMo use_0025	ATTATTCTCTCCGGATGAC GA	MouseDum my
MouseDummy_MetsMIR_650_NonTargetingControlGuideFor Mouse_0026	MetsMIR_650	NonTargetingControlGuideForMo use_0026	TCTCAGTTCGTAGCGAA CGA	MouseDum my
MouseDummy_MetsMIR_651_NonTargetingControlGuideFor Mouse_0027	MetsMIR_651	NonTargetingControlGuideForMo use_0027	TCCCGGGAGGTACGGT GTAC	MouseDum my
MouseDummy_MetsMIR_652_NonTargetingControlGuideFor Mouse_0028	MetsMIR_652	NonTargetingControlGuideForMo use_0028	GGGGATGGCCTTACGTC GCG	MouseDum my
MouseDummy_MetsMIR_653_NonTargetingControlGuideFor Mouse_0029	MetsMIR_653	NonTargetingControlGuideForMo use_0029	GTATCCTCCTTACGGCC CGT	MouseDum my
MouseDummy_MetsMIR_654_NonTargetingControlGuideFor Mouse_0030	MetsMIR_654	NonTargetingControlGuideForMo use_0030	GCTCGGACCTTTTAGAC GTC	MouseDum my
MouseDummy_MetsMIR_655_NonTargetingControlGuideFor Mouse_0031	MetsMIR_655	NonTargetingControlGuideForMo use_0031	GTGGTTACGTTAACGAC TAC	MouseDum my
MouseDummy_MetsMIR_656_NonTargetingControlGuideFor Mouse_0032	MetsMIR_656	NonTargetingControlGuideForMo use_0032	TGCAACGATGGTTACGG TAC	MouseDum my
MouseDummy_MetsMIR_657_NonTargetingControlGuideFor Mouse_0033	MetsMIR_657	NonTargetingControlGuideForMo use_0033	AACGGGCGCAATACCCCT TTT	MouseDum my
MouseDummy_MetsMIR_658_NonTargetingControlGuideFor Mouse_0034	MetsMIR_658	NonTargetingControlGuideForMo use_0034	TCGTGTCTAGCTATCGA GTG	MouseDum my
MouseDummy_MetsMIR_659_NonTargetingControlGuideFor Mouse_0035	MetsMIR_659	NonTargetingControlGuideForMo use_0035	GACGTCTAATTTCTGGC CGT	MouseDum my
MouseDummy_MetsMIR_660_NonTargetingControlGuideFor Mouse_0036	MetsMIR_660	NonTargetingControlGuideForMo use_0036	TAGTCCTAGTTAGATTCC CG	MouseDum my
MouseDummy_MetsMIR_661_NonTargetingControlGuideFor Mouse_0037	MetsMIR_661	NonTargetingControlGuideForMo use_0037	AAGGCCCTAACACGTCG ACC	MouseDum my
MouseDummy_MetsMIR_662_NonTargetingControlGuideFor Mouse_0038	MetsMIR_662	NonTargetingControlGuideForMo use_0038	ATTCTGTCATCGCGGGG TTT	MouseDum my
MouseDummy_MetsMIR_663_NonTargetingControlGuideFor Mouse_0039	MetsMIR_663	NonTargetingControlGuideForMo use_0039	CACCTCGCGTCATATCA CTA	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BE

MouseDummy_MetsMIR_664_NonTargetingControlGuideFor Mouse_0040	MetsMIR_664	NonTargetingControlGuideForMo use_0040	GACCTCGCAATTGAGCG CTC	MouseDum my
MouseDummy_MetsMIR_665_NonTargetingControlGuideFor Mouse_0041	MetsMIR_665	NonTargetingControlGuideForMo use_0041	TGTATCCACCCTGACCC GGT	MouseDum my
MouseDummy_MetsMIR_666_NonTargetingControlGuideFor Mouse_0042	MetsMIR_666	NonTargetingControlGuideForMo use_0042	GATCTTACCACCTCGTCG TAG	MouseDum my
MouseDummy_MetsMIR_667_NonTargetingControlGuideFor Mouse_0043	MetsMIR_667	NonTargetingControlGuideForMo use_0043	TTCGCCGGCGACGAAGT GCA	MouseDum my
MouseDummy_MetsMIR_668_NonTargetingControlGuideFor Mouse_0044	MetsMIR_668	NonTargetingControlGuideForMo use_0044	AGCACTAGGATCGCGGC CTT	MouseDum my
MouseDummy_MetsMIR_669_NonTargetingControlGuideFor Mouse_0045	MetsMIR_669	NonTargetingControlGuideForMo use_0045	GCGTATCTACCCTACCG CCG	MouseDum my
MouseDummy_MetsMIR_670_NonTargetingControlGuideFor Mouse_0046	MetsMIR_670	NonTargetingControlGuideForMo use_0046	GCGCGAGGGCACCAGC AAGT	MouseDum my
MouseDummy_MetsMIR_671_NonTargetingControlGuideFor Mouse_0047	MetsMIR_671	NonTargetingControlGuideForMo use_0047	CGCGTATATGTCACACG GCA	MouseDum my
MouseDummy_MetsMIR_672_NonTargetingControlGuideFor Mouse_0048	MetsMIR_672	NonTargetingControlGuideForMo use_0048	TAAAACCGATCACGATA CGA	MouseDum my
MouseDummy_MetsMIR_673_NonTargetingControlGuideFor Mouse_0049	MetsMIR_673	NonTargetingControlGuideForMo use_0049	TCCTGCGCGATGACCGT CGG	MouseDum my
MouseDummy_MetsMIR_674_NonTargetingControlGuideFor Mouse_0050	MetsMIR_674	NonTargetingControlGuideForMo use_0050	TGGCCCAAGGTGCGCA TAT	MouseDum my
MouseDummy_MetsMIR_675_NonTargetingControlGuideFor Mouse_0051	MetsMIR_675	NonTargetingControlGuideForMo use_0051	GGGGTAGGCCTAATTAC GGA	MouseDum my
MouseDummy_MetsMIR_676_NonTargetingControlGuideFor Mouse_0052	MetsMIR_676	NonTargetingControlGuideForMo use_0052	CGCTTCGTCTCTCGCAA ACA	MouseDum my
MouseDummy_MetsMIR_677_NonTargetingControlGuideFor Mouse_0053	MetsMIR_677	NonTargetingControlGuideForMo use_0053	TTATGTGCCTCTCGGCG CAT	MouseDum my
MouseDummy_MetsMIR_678_NonTargetingControlGuideFor Mouse_0054	MetsMIR_678	NonTargetingControlGuideForMo use_0054	CCGTTCTGACGACGCTA AAG	MouseDum my
MouseDummy_MetsMIR_679_NonTargetingControlGuideFor Mouse_0055	MetsMIR_679	NonTargetingControlGuideForMo use_0055	AAACTCCCGTGTCAACC GAT	MouseDum my
MouseDummy_MetsMIR_680_NonTargetingControlGuideFor Mouse_0056	MetsMIR_680	NonTargetingControlGuideForMo use_0056	GTTTGCGAGTCAAAGTA CGC	MouseDum my
MouseDummy_MetsMIR_681_NonTargetingControlGuideFor Mouse_0057	MetsMIR_681	NonTargetingControlGuideForMo use_0057	GGGTGCACACGCCGGC CTAT	MouseDum my
MouseDummy_MetsMIR_682_NonTargetingControlGuideFor Mouse_0058	MetsMIR_682	NonTargetingControlGuideForMo use_0058	CTTTATACCGCGCTCG GCA	MouseDum my
MouseDummy_MetsMIR_683_NonTargetingControlGuideFor Mouse_0059	MetsMIR_683	NonTargetingControlGuideForMo use_0059	TCAAACGCCCGGGCGC CCCA	MouseDum my
MouseDummy_MetsMIR_684_NonTargetingControlGuideFor Mouse_0060	MetsMIR_684	NonTargetingControlGuideForMo use_0060	GACCGGTGTGTTTACGC GTG	MouseDum my
MouseDummy_MetsMIR_685_NonTargetingControlGuideFor Mouse_0061	MetsMIR_685	NonTargetingControlGuideForMo use_0061	CGGTGAGGATCCCTCTT GCG	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BF

MouseDummy_MetsMIR_686_NonTargetingControlGuideFor Mouse 0062	MetsMIR_686	NonTargetingControlGuideForMo use 0062	ACGGTCCCAACGAGCGC CGG	MouseDum my
MouseDummy_MetsMIR_687_NonTargetingControlGuideFor Mouse 0063	MetsMIR_687	NonTargetingControlGuideForMo use 0063	CGCTTCATTGCCCGAAC GCT	MouseDum my
MouseDummy_MetsMIR_688_NonTargetingControlGuideFor Mouse 0064	MetsMIR_688	NonTargetingControlGuideForMo use 0064	TTATACCGTCACGAGTG CCC	MouseDum my
MouseDummy_MetsMIR_689_NonTargetingControlGuideFor Mouse 0065	MetsMIR_689	NonTargetingControlGuideForMo use 0065	TGCACCGGTACGGGCG ACTC	MouseDum my
MouseDummy_MetsMIR_690_NonTargetingControlGuideFor Mouse 0066	MetsMIR_690	NonTargetingControlGuideForMo use 0066	GAGGCGTACTTCGGCTC TAA	MouseDum my
MouseDummy_MetsMIR_691_NonTargetingControlGuideFor Mouse 0067	MetsMIR_691	NonTargetingControlGuideForMo use 0067	TTCGAGACGGATACGTC TGC	MouseDum my
MouseDummy_MetsMIR_692_NonTargetingControlGuideFor Mouse 0068	MetsMIR_692	NonTargetingControlGuideForMo use 0068	TAGCCCGTCGAGTACTC CCC	MouseDum my
MouseDummy_MetsMIR_693_NonTargetingControlGuideFor Mouse 0069	MetsMIR_693	NonTargetingControlGuideForMo use 0069	AGCTTAGATCGTGCCTC GTA	MouseDum my
MouseDummy_MetsMIR_694_NonTargetingControlGuideFor Mouse 0070	MetsMIR_694	NonTargetingControlGuideForMo use 0070	CGCCGGATGACTGGTG GATA	MouseDum my
MouseDummy_MetsMIR_695_NonTargetingControlGuideFor Mouse 0071	MetsMIR_695	NonTargetingControlGuideForMo use 0071	CCGTTTCGCTATTGGT GCG	MouseDum my
MouseDummy_MetsMIR_696_NonTargetingControlGuideFor Mouse 0072	MetsMIR_696	NonTargetingControlGuideForMo use 0072	TTCGACGTGCGGCGTAT TGG	MouseDum my
MouseDummy_MetsMIR_697_NonTargetingControlGuideFor Mouse 0073	MetsMIR_697	NonTargetingControlGuideForMo use 0073	TTCAATTGTTCCGCCGAA CA	MouseDum my
MouseDummy_MetsMIR_698_NonTargetingControlGuideFor Mouse 0074	MetsMIR_698	NonTargetingControlGuideForMo use 0074	GCGGATCTGACGGTTAC TTA	MouseDum my
MouseDummy_MetsMIR_699_NonTargetingControlGuideFor Mouse 0075	MetsMIR_699	NonTargetingControlGuideForMo use 0075	AATGAAGCACCGATTGC GGA	MouseDum my
MouseDummy_MetsMIR_700_NonTargetingControlGuideFor Mouse 0076	MetsMIR_700	NonTargetingControlGuideForMo use 0076	TCCTATAATTGACGCGAA CGG	MouseDum my
MouseDummy_MetsMIR_701_NonTargetingControlGuideFor Mouse 0077	MetsMIR_701	NonTargetingControlGuideForMo use 0077	GCTACAGATTGCGTTC GAG	MouseDum my
MouseDummy_MetsMIR_702_NonTargetingControlGuideFor Mouse 0078	MetsMIR_702	NonTargetingControlGuideForMo use 0078	GGACTGAAACCGATAGT ATC	MouseDum my
MouseDummy_MetsMIR_703_NonTargetingControlGuideFor Mouse 0079	MetsMIR_703	NonTargetingControlGuideForMo use 0079	GCAATTCGCAACGCAC GTC	MouseDum my
MouseDummy_MetsMIR_704_NonTargetingControlGuideFor Mouse 0080	MetsMIR_704	NonTargetingControlGuideForMo use 0080	CACCCGACTCGGCCGTA AAG	MouseDum my
MouseDummy_MetsMIR_705_NonTargetingControlGuideFor Mouse 0081	MetsMIR_705	NonTargetingControlGuideForMo use 0081	CAACCGTGCGATGCGCG CTA	MouseDum my
MouseDummy_MetsMIR_706_NonTargetingControlGuideFor Mouse 0082	MetsMIR_706	NonTargetingControlGuideForMo use 0082	TGATGGACGCGATACGT TTA	MouseDum my
MouseDummy_MetsMIR_707_NonTargetingControlGuideFor Mouse 0083	MetsMIR_707	NonTargetingControlGuideForMo use 0083	TACCCCTTGAGGGGCGC ATA	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BG

MouseDummy_MetsMIR_708_NonTargetingControlGuideFor Mouse_0084	MetsMIR_708	NonTargetingControlGuideForMo use_0084	CCGTATACGTATCTATGC CG	MouseDum my
MouseDummy_MetsMIR_709_NonTargetingControlGuideFor Mouse_0085	MetsMIR_709	NonTargetingControlGuideForMo use_0085	ACAGCGGCCCTAATACTT CGC	MouseDum my
MouseDummy_MetsMIR_710_NonTargetingControlGuideFor Mouse_0086	MetsMIR_710	NonTargetingControlGuideForMo use_0086	TATCGAACTGCACACGC AAC	MouseDum my
MouseDummy_MetsMIR_711_NonTargetingControlGuideFor Mouse_0087	MetsMIR_711	NonTargetingControlGuideForMo use_0087	TGTCGGACCCATTGTCTG AAC	MouseDum my
MouseDummy_MetsMIR_712_NonTargetingControlGuideFor Mouse_0088	MetsMIR_712	NonTargetingControlGuideForMo use_0088	AAACCTAGCGTAGATTG GGC	MouseDum my
MouseDummy_MetsMIR_713_NonTargetingControlGuideFor Mouse_0089	MetsMIR_713	NonTargetingControlGuideForMo use_0089	TAGATGGTGCTTCTGTC CGC	MouseDum my
MouseDummy_MetsMIR_714_NonTargetingControlGuideFor Mouse_0090	MetsMIR_714	NonTargetingControlGuideForMo use_0090	CATCGTAACACACGTAC GAG	MouseDum my
MouseDummy_MetsMIR_715_NonTargetingControlGuideFor Mouse_0091	MetsMIR_715	NonTargetingControlGuideForMo use_0091	CAACGATCAGGCGTGTT ATC	MouseDum my
MouseDummy_MetsMIR_716_NonTargetingControlGuideFor Mouse_0092	MetsMIR_716	NonTargetingControlGuideForMo use_0092	ACAGCAGGGGCCCGCGA ATAA	MouseDum my
MouseDummy_MetsMIR_717_NonTargetingControlGuideFor Mouse_0093	MetsMIR_717	NonTargetingControlGuideForMo use_0093	AGCCGGGCTTTCCGTCA AGC	MouseDum my
MouseDummy_MetsMIR_718_NonTargetingControlGuideFor Mouse_0094	MetsMIR_718	NonTargetingControlGuideForMo use_0094	ACGGTCCCTCTCGGGTC AAT	MouseDum my
MouseDummy_MetsMIR_719_NonTargetingControlGuideFor Mouse_0095	MetsMIR_719	NonTargetingControlGuideForMo use_0095	TCAATAGTTCTGCGCGA ATT	MouseDum my
MouseDummy_MetsMIR_720_NonTargetingControlGuideFor Mouse_0096	MetsMIR_720	NonTargetingControlGuideForMo use_0096	TCGTTACATACCCCGCG GAA	MouseDum my
MouseDummy_MetsMIR_721_NonTargetingControlGuideFor Mouse_0097	MetsMIR_721	NonTargetingControlGuideForMo use_0097	CATTGGTGCCGTTTCATCT CG	MouseDum my
MouseDummy_MetsMIR_722_NonTargetingControlGuideFor Mouse_0098	MetsMIR_722	NonTargetingControlGuideForMo use_0098	TTCCGGCTCAATGGCGCG AGC	MouseDum my
MouseDummy_MetsMIR_723_NonTargetingControlGuideFor Mouse_0099	MetsMIR_723	NonTargetingControlGuideForMo use_0099	CATACCGCGCCGTGAC TCC	MouseDum my
MouseDummy_MetsMIR_724_NonTargetingControlGuideFor Mouse_0100	MetsMIR_724	NonTargetingControlGuideForMo use_0100	GGACGGATGGGACGAC TAGT	MouseDum my
MouseDummy_MetsMIR_725_NonTargetingControlGuideFor Mouse_0101	MetsMIR_725	NonTargetingControlGuideForMo use_0101	GTGCAACGACTGACCCG CGG	MouseDum my
MouseDummy_MetsMIR_726_NonTargetingControlGuideFor Mouse_0102	MetsMIR_726	NonTargetingControlGuideForMo use_0102	CGACACTTGGGCTGACG CGC	MouseDum my
MouseDummy_MetsMIR_727_NonTargetingControlGuideFor Mouse_0103	MetsMIR_727	NonTargetingControlGuideForMo use_0103	TGTGTAATTACGTTTCGC GG	MouseDum my
MouseDummy_MetsMIR_728_NonTargetingControlGuideFor Mouse_0104	MetsMIR_728	NonTargetingControlGuideForMo use_0104	AACCGTACTGCGAGGAG CAT	MouseDum my
MouseDummy_MetsMIR_729_NonTargetingControlGuideFor Mouse_0105	MetsMIR_729	NonTargetingControlGuideForMo use_0105	GATATTACCGGCGATA AGA	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BH

MouseDummy_MetsMIR_730_NonTargetingControlGuideFor Mouse_0106	MetsMIR_730	NonTargetingControlGuideForMo use_0106	GCGACGTGTATCGATCA CTC	Mou seDum my
MouseDummy_MetsMIR_731_NonTargetingControlGuideFor Mouse_0107	MetsMIR_731	NonTargetingControlGuideForMo use_0107	TAATGGAAGATGCGCGA TAC	Mou seDum my
MouseDummy_MetsMIR_732_NonTargetingControlGuideFor Mouse_0108	MetsMIR_732	NonTargetingControlGuideForMo use_0108	TGTCCTACCGAATGAAC CGT	Mou seDum my
MouseDummy_MetsMIR_733_NonTargetingControlGuideFor Mouse_0109	MetsMIR_733	NonTargetingControlGuideForMo use_0109	ACGCGAAGTGTGCGAGA GTG	Mou seDum my
MouseDummy_MetsMIR_734_NonTargetingControlGuideFor Mouse_0110	MetsMIR_734	NonTargetingControlGuideForMo use_0110	GCAGAGCGTAATCGGCA TCG	Mou seDum my
MouseDummy_MetsMIR_735_NonTargetingControlGuideFor Mouse_0111	MetsMIR_735	NonTargetingControlGuideForMo use_0111	GCCGCGCCATAATATGC CAT	Mou seDum my
MouseDummy_MetsMIR_736_NonTargetingControlGuideFor Mouse_0112	MetsMIR_736	NonTargetingControlGuideForMo use_0112	CGAATGCGCCGGAGAAT ATT	Mou seDum my
MouseDummy_MetsMIR_737_NonTargetingControlGuideFor Mouse_0113	MetsMIR_737	NonTargetingControlGuideForMo use_0113	CCCGCCTGGCGATTAC GGG	Mou seDum my
MouseDummy_MetsMIR_738_NonTargetingControlGuideFor Mouse_0114	MetsMIR_738	NonTargetingControlGuideForMo use_0114	AGCCTAGTCGCGCTAAT ATT	Mou seDum my
MouseDummy_MetsMIR_739_NonTargetingControlGuideFor Mouse_0115	MetsMIR_739	NonTargetingControlGuideForMo use_0115	GTGTGCCTCGACAGCGT AAG	Mou seDum my
MouseDummy_MetsMIR_740_NonTargetingControlGuideFor Mouse_0116	MetsMIR_740	NonTargetingControlGuideForMo use_0116	TCATGCGCCTCGTAATA CCT	Mou seDum my
MouseDummy_MetsMIR_741_NonTargetingControlGuideFor Mouse_0117	MetsMIR_741	NonTargetingControlGuideForMo use_0117	TGTGCCACGCCGCTGCA ACG	Mou seDum my
MouseDummy_MetsMIR_742_NonTargetingControlGuideFor Mouse_0118	MetsMIR_742	NonTargetingControlGuideForMo use_0118	GCAAATAGTTCGGAGCG TGT	Mou seDum my
MouseDummy_MetsMIR_743_NonTargetingControlGuideFor Mouse_0119	MetsMIR_743	NonTargetingControlGuideForMo use_0119	GGATAGCCCGGTTGGT CGT	Mou seDum my
MouseDummy_MetsMIR_744_NonTargetingControlGuideFor Mouse_0120	MetsMIR_744	NonTargetingControlGuideForMo use_0120	TCGCAAAATCCGATTTC CGT	Mou seDum my
MouseDummy_MetsMIR_745_NonTargetingControlGuideFor Mouse_0121	MetsMIR_745	NonTargetingControlGuideForMo use_0121	GGTGCAGTCCGTTTAGT CGG	Mou seDum my
MouseDummy_MetsMIR_746_NonTargetingControlGuideFor Mouse_0122	MetsMIR_746	NonTargetingControlGuideForMo use_0122	TGGGTTCCGGCCCCATG TAC	Mou seDum my
MouseDummy_MetsMIR_747_NonTargetingControlGuideFor Mouse_0123	MetsMIR_747	NonTargetingControlGuideForMo use_0123	CCTTCGGATTCTAGGC TGG	Mou seDum my
MouseDummy_MetsMIR_748_NonTargetingControlGuideFor Mouse_0124	MetsMIR_748	NonTargetingControlGuideForMo use_0124	TCCCCCGTACGTGTATG TCG	Mou seDum my
MouseDummy_MetsMIR_749_NonTargetingControlGuideFor Mouse_0125	MetsMIR_749	NonTargetingControlGuideForMo use_0125	AGGTGCGTGTCAACCGG TAG	Mou seDum my
MouseDummy_MetsMIR_750_NonTargetingControlGuideFor Mouse_0126	MetsMIR_750	NonTargetingControlGuideForMo use_0126	AAAAAGTCCGCGATTAC GTC	Mou seDum my
MouseDummy_MetsMIR_751_NonTargetingControlGuideFor Mouse_0127	MetsMIR_751	NonTargetingControlGuideForMo use_0127	TTATCGCACAAACCCGA AAG	Mou seDum my

Figure 18 (supplemental tables continued)

FIG. 18BI

MouseDummy_MetsMIR_752_NonTargetingControlGuideFor Mouse 0128	MetsMIR_752	NonTargetingControlGuideForMo use 0128	CACCCCGTAGCAACGAT AAA	MouseDum my
MouseDummy_MetsMIR_753_NonTargetingControlGuideFor Mouse 0129	MetsMIR_753	NonTargetingControlGuideForMo use 0129	TAACAGCTTCCGCGTAA TAT	MouseDum my
MouseDummy_MetsMIR_754_NonTargetingControlGuideFor Mouse 0130	MetsMIR_754	NonTargetingControlGuideForMo use 0130	GCCATAGCCAATCGCTA GTT	MouseDum my
MouseDummy_MetsMIR_755_NonTargetingControlGuideFor Mouse 0131	MetsMIR_755	NonTargetingControlGuideForMo use 0131	CATCTATGAGACGTGCG TAC	MouseDum my
MouseDummy_MetsMIR_756_NonTargetingControlGuideFor Mouse 0132	MetsMIR_756	NonTargetingControlGuideForMo use 0132	TTGATAAACCGCGGCCG AAA	MouseDum my
MouseDummy_MetsMIR_757_NonTargetingControlGuideFor Mouse 0133	MetsMIR_757	NonTargetingControlGuideForMo use 0133	GTATCCTCGCAATCGTTA GG	MouseDum my
MouseDummy_MetsMIR_758_NonTargetingControlGuideFor Mouse 0134	MetsMIR_758	NonTargetingControlGuideForMo use 0134	TGGCGCTCCTGCGCACA CGA	MouseDum my
MouseDummy_MetsMIR_759_NonTargetingControlGuideFor Mouse 0135	MetsMIR_759	NonTargetingControlGuideForMo use 0135	TTGGAAGTCTAACCCGC GGG	MouseDum my
MouseDummy_MetsMIR_760_NonTargetingControlGuideFor Mouse 0136	MetsMIR_760	NonTargetingControlGuideForMo use 0136	GTTTCCCGGGACTGTGCG CGT	MouseDum my
MouseDummy_MetsMIR_761_NonTargetingControlGuideFor Mouse 0137	MetsMIR_761	NonTargetingControlGuideForMo use 0137	AACGCCCGGATTTCGT TGA	MouseDum my
MouseDummy_MetsMIR_762_NonTargetingControlGuideFor Mouse 0138	MetsMIR_762	NonTargetingControlGuideForMo use 0138	ACAACACGCCGACACGT CTA	MouseDum my
MouseDummy_MetsMIR_763_NonTargetingControlGuideFor Mouse 0139	MetsMIR_763	NonTargetingControlGuideForMo use 0139	TAAGCGGCGTCATGTGCG CCC	MouseDum my
MouseDummy_MetsMIR_764_NonTargetingControlGuideFor Mouse 0140	MetsMIR_764	NonTargetingControlGuideForMo use 0140	CCGCCCTGCGAACTGCG GTT	MouseDum my
MouseDummy_MetsMIR_765_NonTargetingControlGuideFor Mouse 0141	MetsMIR_765	NonTargetingControlGuideForMo use 0141	TATCTAATCGCGGAGTC GTA	MouseDum my
MouseDummy_MetsMIR_766_NonTargetingControlGuideFor Mouse 0142	MetsMIR_766	NonTargetingControlGuideForMo use 0142	ACGCTATAGTGACGTC TAA	MouseDum my
MouseDummy_MetsMIR_767_NonTargetingControlGuideFor Mouse 0143	MetsMIR_767	NonTargetingControlGuideForMo use 0143	GTAGACATATTGCGTAAT CG	MouseDum my
MouseDummy_MetsMIR_768_NonTargetingControlGuideFor Mouse 0144	MetsMIR_768	NonTargetingControlGuideForMo use 0144	CGTAAACCATAACGTTG GTC	MouseDum my
MouseDummy_MetsMIR_769_NonTargetingControlGuideFor Mouse 0145	MetsMIR_769	NonTargetingControlGuideForMo use 0145	GGACGTAGATTAGGGCG TAA	MouseDum my
MouseDummy_MetsMIR_770_NonTargetingControlGuideFor Mouse 0146	MetsMIR_770	NonTargetingControlGuideForMo use 0146	CGGGCAATATGATCGTA GGC	MouseDum my
MouseDummy_MetsMIR_771_NonTargetingControlGuideFor Mouse 0147	MetsMIR_771	NonTargetingControlGuideForMo use 0147	GGACCAATGTTACCGTA GGT	MouseDum my
MouseDummy_MetsMIR_772_NonTargetingControlGuideFor Mouse 0148	MetsMIR_772	NonTargetingControlGuideForMo use 0148	CAGGGACGTAGCCGCC GTTA	MouseDum my
MouseDummy_MetsMIR_773_NonTargetingControlGuideFor Mouse 0149	MetsMIR_773	NonTargetingControlGuideForMo use 0149	AACCCCGGCTGTATCG CCG	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BJ

MouseDummy_MetsMIR_774_NonTargetingControlGuideFor Mouse_0150	MetsMIR_774	NonTargetingControlGuideForMo use_0150	GATTCATTATCGGCATAC GG	MouseDum my
MouseDummy_MetsMIR_775_NonTargetingControlGuideFor Mouse_0151	MetsMIR_775	NonTargetingControlGuideForMo use_0151	AGTCATCCTCTATGCGC GTA	MouseDum my
MouseDummy_MetsMIR_776_NonTargetingControlGuideFor Mouse_0152	MetsMIR_776	NonTargetingControlGuideForMo use_0152	CGTACGGGCGCGCCGT CCAT	MouseDum my
MouseDummy_MetsMIR_777_NonTargetingControlGuideFor Mouse_0153	MetsMIR_777	NonTargetingControlGuideForMo use_0153	TTGCGTGTGCGTTGTTA ACG	MouseDum my
MouseDummy_MetsMIR_778_NonTargetingControlGuideFor Mouse_0154	MetsMIR_778	NonTargetingControlGuideForMo use_0154	GACCATACGCCTCGTAT GCC	MouseDum my
MouseDummy_MetsMIR_779_NonTargetingControlGuideFor Mouse_0155	MetsMIR_779	NonTargetingControlGuideForMo use_0155	CGGCAATGTGTGCGCGAC CGC	MouseDum my
MouseDummy_MetsMIR_780_NonTargetingControlGuideFor Mouse_0156	MetsMIR_780	NonTargetingControlGuideForMo use_0156	TCGCTAAGCCGGTAAT ACT	MouseDum my
MouseDummy_MetsMIR_781_NonTargetingControlGuideFor Mouse_0157	MetsMIR_781	NonTargetingControlGuideForMo use_0157	TTTCTAATTACCCGATAC GT	MouseDum my
MouseDummy_MetsMIR_782_NonTargetingControlGuideFor Mouse_0158	MetsMIR_782	NonTargetingControlGuideForMo use_0158	ACCAGGACTGCCGCGTG AGG	MouseDum my
MouseDummy_MetsMIR_783_NonTargetingControlGuideFor Mouse_0159	MetsMIR_783	NonTargetingControlGuideForMo use_0159	GCGCGTACATATAAATA GGT	MouseDum my
MouseDummy_MetsMIR_784_NonTargetingControlGuideFor Mouse_0160	MetsMIR_784	NonTargetingControlGuideForMo use_0160	CTTATCCATAATAGGCG GCG	MouseDum my
MouseDummy_MetsMIR_785_NonTargetingControlGuideFor Mouse_0161	MetsMIR_785	NonTargetingControlGuideForMo use_0161	GTTATTGCGCCTTGGCC GTA	MouseDum my
MouseDummy_MetsMIR_786_NonTargetingControlGuideFor Mouse_0162	MetsMIR_786	NonTargetingControlGuideForMo use_0162	TAAAGTGGTGGCTCGT CGT	MouseDum my
MouseDummy_MetsMIR_787_NonTargetingControlGuideFor Mouse_0163	MetsMIR_787	NonTargetingControlGuideForMo use_0163	CTCGTGAACAAGATCC GAC	MouseDum my
MouseDummy_MetsMIR_788_NonTargetingControlGuideFor Mouse_0164	MetsMIR_788	NonTargetingControlGuideForMo use_0164	TAAACGATTACCGATAA CA	MouseDum my
MouseDummy_MetsMIR_789_NonTargetingControlGuideFor Mouse_0165	MetsMIR_789	NonTargetingControlGuideForMo use_0165	CGGCCGTAGTGACGAAT GGA	MouseDum my
MouseDummy_MetsMIR_790_NonTargetingControlGuideFor Mouse_0166	MetsMIR_790	NonTargetingControlGuideForMo use_0166	ACCGGTCGAAGTCTGGG ATT	MouseDum my
MouseDummy_MetsMIR_791_NonTargetingControlGuideFor Mouse_0167	MetsMIR_791	NonTargetingControlGuideForMo use_0167	TCCGTACATCGACCTATT AC	MouseDum my
MouseDummy_MetsMIR_792_NonTargetingControlGuideFor Mouse_0168	MetsMIR_792	NonTargetingControlGuideForMo use_0168	CGACACGATGGTCATAC TAC	MouseDum my
MouseDummy_MetsMIR_793_NonTargetingControlGuideFor Mouse_0169	MetsMIR_793	NonTargetingControlGuideForMo use_0169	ACTCGACCTAACGTCGA TGT	MouseDum my
MouseDummy_MetsMIR_794_NonTargetingControlGuideFor Mouse_0170	MetsMIR_794	NonTargetingControlGuideForMo use_0170	TGCAGATCGCGAAGCGA CTA	MouseDum my
MouseDummy_MetsMIR_795_NonTargetingControlGuideFor Mouse_0171	MetsMIR_795	NonTargetingControlGuideForMo use_0171	TCTAGGATACTCTTAACG GG	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BK

MouseDummy_MetsMIR_796_NonTargetingControlGuideFor Mouse_0172	MetsMIR_796	NonTargetingControlGuideForMo use_0172	GGTCTGTCCGTTGCGAC CAC	MouseDum my
MouseDummy_MetsMIR_797_NonTargetingControlGuideFor Mouse_0173	MetsMIR_797	NonTargetingControlGuideForMo use_0173	ACCGGACTGCGATATGC CGT	MouseDum my
MouseDummy_MetsMIR_798_NonTargetingControlGuideFor Mouse_0174	MetsMIR_798	NonTargetingControlGuideForMo use_0174	CTCGTATAGTATTGCGT GGT	MouseDum my
MouseDummy_MetsMIR_799_NonTargetingControlGuideFor Mouse_0175	MetsMIR_799	NonTargetingControlGuideForMo use_0175	CTTTAGTCACGATATACG TC	MouseDum my
MouseDummy_MetsMIR_800_NonTargetingControlGuideFor Mouse_0176	MetsMIR_800	NonTargetingControlGuideForMo use_0176	GCGGCCCGTGACCGTTC AAT	MouseDum my
MouseDummy_MetsMIR_801_NonTargetingControlGuideFor Mouse_0177	MetsMIR_801	NonTargetingControlGuideForMo use_0177	CCGGTTTATAAATTACGT GG	MouseDum my
MouseDummy_MetsMIR_802_NonTargetingControlGuideFor Mouse_0178	MetsMIR_802	NonTargetingControlGuideForMo use_0178	CAGCTCACCTGCGTAC GGT	MouseDum my
MouseDummy_MetsMIR_803_NonTargetingControlGuideFor Mouse_0179	MetsMIR_803	NonTargetingControlGuideForMo use_0179	CTCGGCTTTACGATCGA TCA	MouseDum my
MouseDummy_MetsMIR_804_NonTargetingControlGuideFor Mouse_0180	MetsMIR_804	NonTargetingControlGuideForMo use_0180	TCTGCACCCGACTTAG CTT	MouseDum my
MouseDummy_MetsMIR_805_NonTargetingControlGuideFor Mouse_0181	MetsMIR_805	NonTargetingControlGuideForMo use_0181	GGCTTATAGACGAGACT CGA	MouseDum my
MouseDummy_MetsMIR_806_NonTargetingControlGuideFor Mouse_0182	MetsMIR_806	NonTargetingControlGuideForMo use_0182	ACCGGATGTGGCGCCT CTC	MouseDum my
MouseDummy_MetsMIR_807_NonTargetingControlGuideFor Mouse_0183	MetsMIR_807	NonTargetingControlGuideForMo use_0183	CATTACGTGTCGAGCTC CGG	MouseDum my
MouseDummy_MetsMIR_808_NonTargetingControlGuideFor Mouse_0184	MetsMIR_808	NonTargetingControlGuideForMo use_0184	GATGCGGGTGGAAAACG TTA	MouseDum my
MouseDummy_MetsMIR_809_NonTargetingControlGuideFor Mouse_0185	MetsMIR_809	NonTargetingControlGuideForMo use_0185	GCAAAAAGCGACACGC GAC	MouseDum my
MouseDummy_MetsMIR_810_NonTargetingControlGuideFor Mouse_0186	MetsMIR_810	NonTargetingControlGuideForMo use_0186	CGAGGATGTACATACGT AAA	MouseDum my
MouseDummy_MetsMIR_811_NonTargetingControlGuideFor Mouse_0187	MetsMIR_811	NonTargetingControlGuideForMo use_0187	ACACGTCCTCGGCTATA CGC	MouseDum my
MouseDummy_MetsMIR_812_NonTargetingControlGuideFor Mouse_0188	MetsMIR_812	NonTargetingControlGuideForMo use_0188	TCGTGCCCTAGCTCGTT GAG	MouseDum my
MouseDummy_MetsMIR_813_NonTargetingControlGuideFor Mouse_0189	MetsMIR_813	NonTargetingControlGuideForMo use_0189	TCACACGGGATCTCGCC GGT	MouseDum my
MouseDummy_MetsMIR_814_NonTargetingControlGuideFor Mouse_0190	MetsMIR_814	NonTargetingControlGuideForMo use_0190	CGTGCGTTCGTAATAAA AGG	MouseDum my
MouseDummy_MetsMIR_815_NonTargetingControlGuideFor Mouse_0191	MetsMIR_815	NonTargetingControlGuideForMo use_0191	CCCCTCGATATCCGAT GAC	MouseDum my
MouseDummy_MetsMIR_816_NonTargetingControlGuideFor Mouse_0192	MetsMIR_816	NonTargetingControlGuideForMo use_0192	CGTGATTCCTAAGCCCC CGC	MouseDum my
MouseDummy_MetsMIR_817_NonTargetingControlGuideFor Mouse_0193	MetsMIR_817	NonTargetingControlGuideForMo use_0193	GCGTGCCATGGGAGGC CGTT	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BL

MouseDummy_MetsMIR_818_NonTargetingControlGuideForMouse_0184	MetsMIR_818	NonTargetingControlGuideForMouse_0184	GGCTTTACGTAAGGAGC GTA	MouseDummy
MouseDummy_MetsMIR_819_NonTargetingControlGuideForMouse_0185	MetsMIR_819	NonTargetingControlGuideForMouse_0185	ACCCTAGCCTCATCGCG ACC	MouseDummy
MouseDummy_MetsMIR_820_NonTargetingControlGuideForMouse_0186	MetsMIR_820	NonTargetingControlGuideForMouse_0186	GCTGATTAGACCCGGCG TAA	MouseDummy
MouseDummy_MetsMIR_821_NonTargetingControlGuideForMouse_0187	MetsMIR_821	NonTargetingControlGuideForMouse_0187	GTTACCCCTTTGGCCGG AAG	MouseDummy
MouseDummy_MetsMIR_822_NonTargetingControlGuideForMouse_0188	MetsMIR_822	NonTargetingControlGuideForMouse_0188	TAACGTTATGTCAAACG CTC	MouseDummy
MouseDummy_MetsMIR_823_NonTargetingControlGuideForMouse_0189	MetsMIR_823	NonTargetingControlGuideForMouse_0189	GGATCGATCCGAGGAAC GTG	MouseDummy
MouseDummy_MetsMIR_824_NonTargetingControlGuideForMouse_0200	MetsMIR_824	NonTargetingControlGuideForMouse_0200	CTGCACGGCACGACATC CAA	MouseDummy
MouseDummy_MetsMIR_825_NonTargetingControlGuideForMouse_0201	MetsMIR_825	NonTargetingControlGuideForMouse_0201	GGTCAAAGCGATGTTAG CCG	MouseDummy
MouseDummy_MetsMIR_826_NonTargetingControlGuideForMouse_0202	MetsMIR_826	NonTargetingControlGuideForMouse_0202	GCTTGTAATCTAAAGAC GCG	MouseDummy
MouseDummy_MetsMIR_827_NonTargetingControlGuideForMouse_0203	MetsMIR_827	NonTargetingControlGuideForMouse_0203	GAGATTAAATTAACGCC GGC	MouseDummy
MouseDummy_MetsMIR_828_NonTargetingControlGuideForMouse_0204	MetsMIR_828	NonTargetingControlGuideForMouse_0204	CGAATCGGGAAGGCGC GTGT	MouseDummy
MouseDummy_MetsMIR_829_NonTargetingControlGuideForMouse_0205	MetsMIR_829	NonTargetingControlGuideForMouse_0205	GGGACGCTCATCGAGTG ACG	MouseDummy
MouseDummy_MetsMIR_830_NonTargetingControlGuideForMouse_0206	MetsMIR_830	NonTargetingControlGuideForMouse_0206	AGCGCGTAAGGCGTAGT TAC	MouseDummy
MouseDummy_MetsMIR_831_NonTargetingControlGuideForMouse_0207	MetsMIR_831	NonTargetingControlGuideForMouse_0207	GTTTCAGCGCGAGTTGC GCG	MouseDummy
MouseDummy_MetsMIR_832_NonTargetingControlGuideForMouse_0208	MetsMIR_832	NonTargetingControlGuideForMouse_0208	TCCYTCGCTTCCACTCGT ACT	MouseDummy
MouseDummy_MetsMIR_833_NonTargetingControlGuideForMouse_0209	MetsMIR_833	NonTargetingControlGuideForMouse_0209	GTTATTGTCTGTGCAAA CG	MouseDummy
MouseDummy_MetsMIR_834_NonTargetingControlGuideForMouse_0210	MetsMIR_834	NonTargetingControlGuideForMouse_0210	AATAAGCCTACCCGGCG AGA	MouseDummy
MouseDummy_MetsMIR_835_NonTargetingControlGuideForMouse_0211	MetsMIR_835	NonTargetingControlGuideForMouse_0211	GTACATGCGGCAAGTCG ACT	MouseDummy
MouseDummy_MetsMIR_836_NonTargetingControlGuideForMouse_0212	MetsMIR_836	NonTargetingControlGuideForMouse_0212	AGCAGTTCGGGTAAACGC CCA	MouseDummy
MouseDummy_MetsMIR_837_NonTargetingControlGuideForMouse_0213	MetsMIR_837	NonTargetingControlGuideForMouse_0213	TGTAATCGCATAGCGGG GGC	MouseDummy
MouseDummy_MetsMIR_838_NonTargetingControlGuideForMouse_0214	MetsMIR_838	NonTargetingControlGuideForMouse_0214	AACTCGCGTGGGAAGTC CGG	MouseDummy
MouseDummy_MetsMIR_839_NonTargetingControlGuideForMouse_0215	MetsMIR_839	NonTargetingControlGuideForMouse_0215	CCTTGTGGCGTGCCAAA CGA	MouseDummy

Figure 18 (supplemental tables continued)

FIG. 18BM

MouseDummy_MetsMIR_840_NonTargetingControlGuideFor Mouse 0216	MetsMIR_840	NonTargetingControlGuideForMo use 0216	TGCCGACGCGACGACG CGTA	MouseDum my
MouseDummy_MetsMIR_841_NonTargetingControlGuideFor Mouse 0217	MetsMIR_841	NonTargetingControlGuideForMo use 0217	TGCGGCAATGTTAACCC TTA	MouseDum my
MouseDummy_MetsMIR_842_NonTargetingControlGuideFor Mouse 0218	MetsMIR_842	NonTargetingControlGuideForMo use 0218	TGTATAGTCATCGCCGT AAT	MouseDum my
MouseDummy_MetsMIR_843_NonTargetingControlGuideFor Mouse 0219	MetsMIR_843	NonTargetingControlGuideForMo use 0219	TAGGCGGACGGTTACAT ATA	MouseDum my
MouseDummy_MetsMIR_844_NonTargetingControlGuideFor Mouse 0220	MetsMIR_844	NonTargetingControlGuideForMo use 0220	GATAAGACTCCGCGAGC TTC	MouseDum my
MouseDummy_MetsMIR_845_NonTargetingControlGuideFor Mouse 0221	MetsMIR_845	NonTargetingControlGuideForMo use 0221	TCGTGCCGGAATAACCA CTA	MouseDum my
MouseDummy_MetsMIR_846_NonTargetingControlGuideFor Mouse 0222	MetsMIR_846	NonTargetingControlGuideForMo use 0222	CGAGTAATTATTGCGGT CG	MouseDum my
MouseDummy_MetsMIR_847_NonTargetingControlGuideFor Mouse 0223	MetsMIR_847	NonTargetingControlGuideForMo use 0223	TACTTACCGACCGACAA ACG	MouseDum my
MouseDummy_MetsMIR_848_NonTargetingControlGuideFor Mouse 0224	MetsMIR_848	NonTargetingControlGuideForMo use 0224	ATGGTGGCTGTACTCGT AAC	MouseDum my
MouseDummy_MetsMIR_849_NonTargetingControlGuideFor Mouse 0225	MetsMIR_849	NonTargetingControlGuideForMo use 0225	TCGTAAGTTCGCTATATG CC	MouseDum my
MouseDummy_MetsMIR_850_NonTargetingControlGuideFor Mouse 0226	MetsMIR_850	NonTargetingControlGuideForMo use 0226	CCGGCTTGAATACCGTG CGG	MouseDum my
MouseDummy_MetsMIR_851_NonTargetingControlGuideFor Mouse 0227	MetsMIR_851	NonTargetingControlGuideForMo use 0227	GGCCCAACGAACTAGC GTG	MouseDum my
MouseDummy_MetsMIR_852_NonTargetingControlGuideFor Mouse 0228	MetsMIR_852	NonTargetingControlGuideForMo use 0228	AGCCGCGCCGCAAGCT TTT	MouseDum my
MouseDummy_MetsMIR_853_NonTargetingControlGuideFor Mouse 0229	MetsMIR_853	NonTargetingControlGuideForMo use 0229	ACACGACCGACCGGTG GAAT	MouseDum my
MouseDummy_MetsMIR_854_NonTargetingControlGuideFor Mouse 0230	MetsMIR_854	NonTargetingControlGuideForMo use 0230	AAACTCATACGTAGCGA ATC	MouseDum my
MouseDummy_MetsMIR_855_NonTargetingControlGuideFor Mouse 0231	MetsMIR_855	NonTargetingControlGuideForMo use 0231	CGATTGACGTTGGGCTC TCA	MouseDum my
MouseDummy_MetsMIR_856_NonTargetingControlGuideFor Mouse 0232	MetsMIR_856	NonTargetingControlGuideForMo use 0232	AACCAGCATTTGACCGC GCT	MouseDum my
MouseDummy_MetsMIR_857_NonTargetingControlGuideFor Mouse 0233	MetsMIR_857	NonTargetingControlGuideForMo use 0233	ACGTCGGTCTAGAGTTA AGT	MouseDum my
MouseDummy_MetsMIR_858_NonTargetingControlGuideFor Mouse 0234	MetsMIR_858	NonTargetingControlGuideForMo use 0234	CTCCGGACGTGCATCCG AGA	MouseDum my
MouseDummy_MetsMIR_859_NonTargetingControlGuideFor Mouse 0235	MetsMIR_859	NonTargetingControlGuideForMo use 0235	ATAAGTCCCAGGTGCG CGC	MouseDum my
MouseDummy_MetsMIR_860_NonTargetingControlGuideFor Mouse 0236	MetsMIR_860	NonTargetingControlGuideForMo use 0236	GCCGCTCTTGATAACGA CGC	MouseDum my
MouseDummy_MetsMIR_861_NonTargetingControlGuideFor Mouse 0237	MetsMIR_861	NonTargetingControlGuideForMo use 0237	CGAAACCATACCTCCTT CGA	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BN

MouseDummy_MetsMIR_862_NonTargetingControlGuideFor Mouse 0238	MetsMIR_862	NonTargetingControlGuideForMo use 0238	CGCGCCTACCCCTTTAC CGC	MouseDum my
MouseDummy_MetsMIR_863_NonTargetingControlGuideFor Mouse 0239	MetsMIR_863	NonTargetingControlGuideForMo use 0239	TACCCCTCTGCGATGCC GGT	MouseDum my
MouseDummy_MetsMIR_864_NonTargetingControlGuideFor Mouse 0240	MetsMIR_864	NonTargetingControlGuideForMo use 0240	ATATTGTCCCATACGATC GG	MouseDum my
MouseDummy_MetsMIR_865_NonTargetingControlGuideFor Mouse 0241	MetsMIR_865	NonTargetingControlGuideForMo use 0241	GGATTGACGAGACGATA TCG	MouseDum my
MouseDummy_MetsMIR_866_NonTargetingControlGuideFor Mouse 0242	MetsMIR_866	NonTargetingControlGuideForMo use 0242	CACGTAAGACGCTCCAC TTA	MouseDum my
MouseDummy_MetsMIR_867_NonTargetingControlGuideFor Mouse 0243	MetsMIR_867	NonTargetingControlGuideForMo use 0243	GGAGTCTCACGCAATTA GCG	MouseDum my
MouseDummy_MetsMIR_868_NonTargetingControlGuideFor Mouse 0244	MetsMIR_868	NonTargetingControlGuideForMo use 0244	ATCGCAGGTCTACGCAG AGT	MouseDum my
MouseDummy_MetsMIR_869_NonTargetingControlGuideFor Mouse 0245	MetsMIR_869	NonTargetingControlGuideForMo use 0245	CGGTCGTGACAGACCTG GTG	MouseDum my
MouseDummy_MetsMIR_870_NonTargetingControlGuideFor Mouse 0246	MetsMIR_870	NonTargetingControlGuideForMo use 0246	GCCACTCCGCTCGTTCT AGA	MouseDum my
MouseDummy_MetsMIR_871_NonTargetingControlGuideFor Mouse 0247	MetsMIR_871	NonTargetingControlGuideForMo use 0247	CCGTCGAGCAATCCCGC CAA	MouseDum my
MouseDummy_MetsMIR_872_NonTargetingControlGuideFor Mouse 0248	MetsMIR_872	NonTargetingControlGuideForMo use 0248	GAGTCGAGTTAATAACG CTC	MouseDum my
MouseDummy_MetsMIR_873_NonTargetingControlGuideFor Mouse 0249	MetsMIR_873	NonTargetingControlGuideForMo use 0249	CAATATCTAAGCGCTAA CGA	MouseDum my
MouseDummy_MetsMIR_874_NonTargetingControlGuideFor Mouse 0250	MetsMIR_874	NonTargetingControlGuideForMo use 0250	ACCCATCCCCGCGTCCG AGA	MouseDum my
MouseDummy_MetsMIR_875_NonTargetingControlGuideFor Mouse 0251	MetsMIR_875	NonTargetingControlGuideForMo use 0251	TAATGAGTAACGCTCAT CGG	MouseDum my
MouseDummy_MetsMIR_876_NonTargetingControlGuideFor Mouse 0252	MetsMIR_876	NonTargetingControlGuideForMo use 0252	AACGGTAGCGTACCCGT GAA	MouseDum my
MouseDummy_MetsMIR_877_NonTargetingControlGuideFor Mouse 0253	MetsMIR_877	NonTargetingControlGuideForMo use 0253	CATCGAGGGTAAACGCC ATT	MouseDum my
MouseDummy_MetsMIR_878_NonTargetingControlGuideFor Mouse 0254	MetsMIR_878	NonTargetingControlGuideForMo use 0254	CGTCACCGCTAGTAATG ATG	MouseDum my
MouseDummy_MetsMIR_879_NonTargetingControlGuideFor Mouse 0255	MetsMIR_879	NonTargetingControlGuideForMo use 0255	AATCACCGACAACGTAA GAC	MouseDum my
MouseDummy_MetsMIR_880_NonTargetingControlGuideFor Mouse 0256	MetsMIR_880	NonTargetingControlGuideForMo use 0256	CACTCAGCGGTGGACG CCC	MouseDum my
MouseDummy_MetsMIR_881_NonTargetingControlGuideFor Mouse 0257	MetsMIR_881	NonTargetingControlGuideForMo use 0257	TCCGTAGGACGTATATA TTC	MouseDum my
MouseDummy_MetsMIR_882_NonTargetingControlGuideFor Mouse 0258	MetsMIR_882	NonTargetingControlGuideForMo use 0258	AAAACGTAATTATACCGA GC	MouseDum my
MouseDummy_MetsMIR_883_NonTargetingControlGuideFor Mouse 0259	MetsMIR_883	NonTargetingControlGuideForMo use 0259	TAGGCCGCTCAGGCCG CACT	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BO

MouseDummy_MetsMIR_884_NonTargetingControlGuideFor Mouse 0260	MetsMIR_884	NonTargetingControlGuideForMo use 0260	CCCACTTAATAACGCCG CTT	MouseDum my
MouseDummy_MetsMIR_885_NonTargetingControlGuideFor Mouse 0261	MetsMIR_885	NonTargetingControlGuideForMo use 0261	ACCATACGGGGTCTTGT CGA	MouseDum my
MouseDummy_MetsMIR_886_NonTargetingControlGuideFor Mouse 0262	MetsMIR_886	NonTargetingControlGuideForMo use 0262	ACGGAAACCCGATCGGAA CGG	MouseDum my
MouseDummy_MetsMIR_887_NonTargetingControlGuideFor Mouse 0263	MetsMIR_887	NonTargetingControlGuideForMo use 0263	GTTGCGTCCATTCCGTC GCC	MouseDum my
MouseDummy_MetsMIR_888_NonTargetingControlGuideFor Mouse 0264	MetsMIR_888	NonTargetingControlGuideForMo use 0264	ATTCCGTTTGCAGCGAG ACC	MouseDum my
MouseDummy_MetsMIR_889_NonTargetingControlGuideFor Mouse 0265	MetsMIR_889	NonTargetingControlGuideForMo use 0265	CCATCGGTTTCGACTTAC CGC	MouseDum my
MouseDummy_MetsMIR_890_NonTargetingControlGuideFor Mouse 0266	MetsMIR_890	NonTargetingControlGuideForMo use 0266	CATCGGCTATGTCGGGG ACA	MouseDum my
MouseDummy_MetsMIR_891_NonTargetingControlGuideFor Mouse 0267	MetsMIR_891	NonTargetingControlGuideForMo use 0267	AATGAGCGTCTCTCGAT CGC	MouseDum my
MouseDummy_MetsMIR_892_NonTargetingControlGuideFor Mouse 0268	MetsMIR_892	NonTargetingControlGuideForMo use 0268	TTACTGATCAGTCGGAC GCA	MouseDum my
MouseDummy_MetsMIR_893_NonTargetingControlGuideFor Mouse 0269	MetsMIR_893	NonTargetingControlGuideForMo use 0269	GATTGAGAAGCCGCGGT ATC	MouseDum my
MouseDummy_MetsMIR_894_NonTargetingControlGuideFor Mouse 0270	MetsMIR_894	NonTargetingControlGuideForMo use 0270	GTGATGGCCACGTCCGA ACC	MouseDum my
MouseDummy_MetsMIR_895_NonTargetingControlGuideFor Mouse 0271	MetsMIR_895	NonTargetingControlGuideForMo use 0271	TTGGACGTACACTTTCGT TC	MouseDum my
MouseDummy_MetsMIR_896_NonTargetingControlGuideFor Mouse 0272	MetsMIR_896	NonTargetingControlGuideForMo use 0272	ACTGCGCGTATAGGACG CAA	MouseDum my
MouseDummy_MetsMIR_897_NonTargetingControlGuideFor Mouse 0273	MetsMIR_897	NonTargetingControlGuideForMo use 0273	TAGCTGTTTACGCCGAC CTG	MouseDum my
MouseDummy_MetsMIR_898_NonTargetingControlGuideFor Mouse 0274	MetsMIR_898	NonTargetingControlGuideForMo use 0274	CATGCCCGTCTGCGCCG CAT	MouseDum my
MouseDummy_MetsMIR_899_NonTargetingControlGuideFor Mouse 0275	MetsMIR_899	NonTargetingControlGuideForMo use 0275	TCAACTATGAACGCCCG TGC	MouseDum my
MouseDummy_MetsMIR_900_NonTargetingControlGuideFor Mouse 0276	MetsMIR_900	NonTargetingControlGuideForMo use 0276	CGTTACGTTTCTTGCCA GGA	MouseDum my
MouseDummy_MetsMIR_901_NonTargetingControlGuideFor Mouse 0277	MetsMIR_901	NonTargetingControlGuideForMo use 0277	TCTATATCTAGTCTCGGC GC	MouseDum my
MouseDummy_MetsMIR_902_NonTargetingControlGuideFor Mouse 0278	MetsMIR_902	NonTargetingControlGuideForMo use 0278	CTCCTCGAGGCTGGCTA CGT	MouseDum my
MouseDummy_MetsMIR_903_NonTargetingControlGuideFor Mouse 0279	MetsMIR_903	NonTargetingControlGuideForMo use 0279	TCGTACCGCACGTGTAA GCC	MouseDum my
MouseDummy_MetsMIR_904_NonTargetingControlGuideFor Mouse 0280	MetsMIR_904	NonTargetingControlGuideForMo use 0280	CGGGCACTAACCCGATA CAC	MouseDum my
MouseDummy_MetsMIR_905_NonTargetingControlGuideFor Mouse 0281	MetsMIR_905	NonTargetingControlGuideForMo use 0281	GAACGTAACGGCATGCA TCA	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BP

MouseDummy_MetsMIR_906_NonTargetingControlGuideFor Mouse 0282	MetsMIR_906	NonTargetingControlGuideForMo use 0282	GCCCGATAGAATTACCC ATT	MouseDum my
MouseDummy_MetsMIR_907_NonTargetingControlGuideFor Mouse 0283	MetsMIR_907	NonTargetingControlGuideForMo use 0283	GCCCTCGAGCTCACGAT GAG	MouseDum my
MouseDummy_MetsMIR_908_NonTargetingControlGuideFor Mouse 0284	MetsMIR_908	NonTargetingControlGuideForMo use 0284	TTAGGACGTGCTTCCGA GGG	MouseDum my
MouseDummy_MetsMIR_909_NonTargetingControlGuideFor Mouse 0285	MetsMIR_909	NonTargetingControlGuideForMo use 0285	TAGTTAACCGTAAAGTG GGC	MouseDum my
MouseDummy_MetsMIR_910_NonTargetingControlGuideFor Mouse 0286	MetsMIR_910	NonTargetingControlGuideForMo use 0286	GTGGGCTGACCGTTCTC GAC	MouseDum my
MouseDummy_MetsMIR_911_NonTargetingControlGuideFor Mouse 0287	MetsMIR_911	NonTargetingControlGuideForMo use 0287	CAAGAGTTAACCTCGAC CGG	MouseDum my
MouseDummy_MetsMIR_912_NonTargetingControlGuideFor Mouse 0288	MetsMIR_912	NonTargetingControlGuideForMo use 0288	TATGACTGCACGACTCG CTA	MouseDum my
MouseDummy_MetsMIR_913_NonTargetingControlGuideFor Mouse 0289	MetsMIR_913	NonTargetingControlGuideForMo use 0289	GAACCGGCGTGCCTTAG CGG	MouseDum my
MouseDummy_MetsMIR_914_NonTargetingControlGuideFor Mouse 0290	MetsMIR_914	NonTargetingControlGuideForMo use 0290	TTCGATATAGGGGACGG CGG	MouseDum my
MouseDummy_MetsMIR_915_NonTargetingControlGuideFor Mouse 0291	MetsMIR_915	NonTargetingControlGuideForMo use 0291	GCCGTAAGCGGGCCGG TTGA	MouseDum my
MouseDummy_MetsMIR_916_NonTargetingControlGuideFor Mouse 0292	MetsMIR_916	NonTargetingControlGuideForMo use 0292	CGTAACCGGAGATAATA TTA	MouseDum my
MouseDummy_MetsMIR_917_NonTargetingControlGuideFor Mouse 0293	MetsMIR_917	NonTargetingControlGuideForMo use 0293	CGGCGATAACAGCGACA TCG	MouseDum my
MouseDummy_MetsMIR_918_NonTargetingControlGuideFor Mouse 0294	MetsMIR_918	NonTargetingControlGuideForMo use 0294	TCACCTCGGGCATTACG AGC	MouseDum my
MouseDummy_MetsMIR_919_NonTargetingControlGuideFor Mouse 0295	MetsMIR_919	NonTargetingControlGuideForMo use 0295	TAATAAACTATGTCCCGC CG	MouseDum my
MouseDummy_MetsMIR_920_NonTargetingControlGuideFor Mouse 0296	MetsMIR_920	NonTargetingControlGuideForMo use 0296	TGGCACGGAGTTGCATA CGC	MouseDum my
MouseDummy_MetsMIR_921_NonTargetingControlGuideFor Mouse 0297	MetsMIR_921	NonTargetingControlGuideForMo use 0297	GCTAGTTCTCCCGGGCG AAA	MouseDum my
MouseDummy_MetsMIR_922_NonTargetingControlGuideFor Mouse 0298	MetsMIR_922	NonTargetingControlGuideForMo use 0298	GTTCGGTTGCAGCTTAC ACG	MouseDum my
MouseDummy_MetsMIR_923_NonTargetingControlGuideFor Mouse 0299	MetsMIR_923	NonTargetingControlGuideForMo use 0299	GCACCTCTAGCGCGCTC GGC	MouseDum my
MouseDummy_MetsMIR_924_NonTargetingControlGuideFor Mouse 0300	MetsMIR_924	NonTargetingControlGuideForMo use 0300	CATTGTAGACTCGTACG GAT	MouseDum my
MouseDummy_MetsMIR_925_NonTargetingControlGuideFor Mouse 0301	MetsMIR_925	NonTargetingControlGuideForMo use 0301	TCGGAATTCGGCAGGCC GAG	MouseDum my
MouseDummy_MetsMIR_926_NonTargetingControlGuideFor Mouse 0302	MetsMIR_926	NonTargetingControlGuideForMo use 0302	GATTACGGGGTTCCGTA ACT	MouseDum my
MouseDummy_MetsMIR_927_NonTargetingControlGuideFor Mouse 0303	MetsMIR_927	NonTargetingControlGuideForMo use 0303	GCCTATGTGAATCGCGA ATT	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BQ

MouseDummy_MetsMIR_928_NonTargetingControlGuideFor Mouse_0304	MetsMIR_928	NonTargetingControlGuideForMo use_0304	CTCGGACGGCATAACGAC AAT	MouseDum my
MouseDummy_MetsMIR_929_NonTargetingControlGuideFor Mouse_0305	MetsMIR_929	NonTargetingControlGuideForMo use_0305	CATATGCTCGACGGTAT AAA	MouseDum my
MouseDummy_MetsMIR_930_NonTargetingControlGuideFor Mouse_0306	MetsMIR_930	NonTargetingControlGuideForMo use_0306	CTCATGTACGCCCTTCGC TAC	MouseDum my
MouseDummy_MetsMIR_931_NonTargetingControlGuideFor Mouse_0307	MetsMIR_931	NonTargetingControlGuideForMo use_0307	TTACAGTTCATACCGTCG CC	MouseDum my
MouseDummy_MetsMIR_932_NonTargetingControlGuideFor Mouse_0308	MetsMIR_932	NonTargetingControlGuideForMo use_0308	GATGGCCAGTAACGGCG TCA	MouseDum my
MouseDummy_MetsMIR_933_NonTargetingControlGuideFor Mouse_0309	MetsMIR_933	NonTargetingControlGuideForMo use_0309	ATTGAATGCTCCCGTC GAC	MouseDum my
MouseDummy_MetsMIR_934_NonTargetingControlGuideFor Mouse_0310	MetsMIR_934	NonTargetingControlGuideForMo use_0310	CAACGGTCACGCTAGAA TAA	MouseDum my
MouseDummy_MetsMIR_935_NonTargetingControlGuideFor Mouse_0311	MetsMIR_935	NonTargetingControlGuideForMo use_0311	CTTAAGTTCGCGACGGA ATG	MouseDum my
MouseDummy_MetsMIR_936_NonTargetingControlGuideFor Mouse_0312	MetsMIR_936	NonTargetingControlGuideForMo use_0312	TACTTAGGTCCGCGTAA AGC	MouseDum my
MouseDummy_MetsMIR_937_NonTargetingControlGuideFor Mouse_0313	MetsMIR_937	NonTargetingControlGuideForMo use_0313	GGTTAAAAATTAAAGCGG TCC	MouseDum my
MouseDummy_MetsMIR_938_NonTargetingControlGuideFor Mouse_0314	MetsMIR_938	NonTargetingControlGuideForMo use_0314	AAAACGGCTCGATCGGT GAT	MouseDum my
MouseDummy_MetsMIR_939_NonTargetingControlGuideFor Mouse_0315	MetsMIR_939	NonTargetingControlGuideForMo use_0315	ATGTTACGTACGTGATCT CC	MouseDum my
MouseDummy_MetsMIR_940_NonTargetingControlGuideFor Mouse_0316	MetsMIR_940	NonTargetingControlGuideForMo use_0316	AGTAGACGCTATGTTCC CGC	MouseDum my
MouseDummy_MetsMIR_941_NonTargetingControlGuideFor Mouse_0317	MetsMIR_941	NonTargetingControlGuideForMo use_0317	TACTGCGTACCGCAGTA AGC	MouseDum my
MouseDummy_MetsMIR_942_NonTargetingControlGuideFor Mouse_0318	MetsMIR_942	NonTargetingControlGuideForMo use_0318	CCCATCATTCGCGCTGA CGT	MouseDum my
MouseDummy_MetsMIR_943_NonTargetingControlGuideFor Mouse_0319	MetsMIR_943	NonTargetingControlGuideForMo use_0319	CGACTATTGCCGTCCAT CTC	MouseDum my
MouseDummy_MetsMIR_944_NonTargetingControlGuideFor Mouse_0320	MetsMIR_944	NonTargetingControlGuideForMo use_0320	ATCCGTACCAAACACGC TAC	MouseDum my
MouseDummy_MetsMIR_945_NonTargetingControlGuideFor Mouse_0321	MetsMIR_945	NonTargetingControlGuideForMo use_0321	TGCGTAAACTTGCGCT CGA	MouseDum my
MouseDummy_MetsMIR_946_NonTargetingControlGuideFor Mouse_0322	MetsMIR_946	NonTargetingControlGuideForMo use_0322	CATATGCCTATAACCGG CGG	MouseDum my
MouseDummy_MetsMIR_947_NonTargetingControlGuideFor Mouse_0323	MetsMIR_947	NonTargetingControlGuideForMo use_0323	GTTCCGGATATATACGG TTA	MouseDum my
MouseDummy_MetsMIR_948_NonTargetingControlGuideFor Mouse_0324	MetsMIR_948	NonTargetingControlGuideForMo use_0324	CTAATACACCCGGACGG TAC	MouseDum my
MouseDummy_MetsMIR_949_NonTargetingControlGuideFor Mouse_0325	MetsMIR_949	NonTargetingControlGuideForMo use_0325	GGGCGGGACGTAATATT ATG	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BR

MouseDummy_MetsMIR_950_NonTargetingControlGuideFor Mouse_0326	MetsMIR_950	NonTargetingControlGuideForMo use_0326	ACTTTACATCATGTCGTC GT	MouseDum my
MouseDummy_MetsMIR_951_NonTargetingControlGuideFor Mouse_0327	MetsMIR_951	NonTargetingControlGuideForMo use_0327	CTGATCGGTGCATATCT CGG	MouseDum my
MouseDummy_MetsMIR_952_NonTargetingControlGuideFor Mouse_0328	MetsMIR_952	NonTargetingControlGuideForMo use_0328	GAGCCCCAACGCGCGA AGCA	MouseDum my
MouseDummy_MetsMIR_953_NonTargetingControlGuideFor Mouse_0329	MetsMIR_953	NonTargetingControlGuideForMo use_0329	GCTAAGATTATCCGAA CAC	MouseDum my
MouseDummy_MetsMIR_954_NonTargetingControlGuideFor Mouse_0330	MetsMIR_954	NonTargetingControlGuideForMo use_0330	GTTACCGTGACGATAAG AAT	MouseDum my
MouseDummy_MetsMIR_955_NonTargetingControlGuideFor Mouse_0331	MetsMIR_955	NonTargetingControlGuideForMo use_0331	GCTAACGATTTTACG GGC	MouseDum my
MouseDummy_MetsMIR_956_NonTargetingControlGuideFor Mouse_0332	MetsMIR_956	NonTargetingControlGuideForMo use_0332	CCGAGCCGAATTGGGC GTGT	MouseDum my
MouseDummy_MetsMIR_957_NonTargetingControlGuideFor Mouse_0333	MetsMIR_957	NonTargetingControlGuideForMo use_0333	GTTGCGTGTGTCCGTAC AAA	MouseDum my
MouseDummy_MetsMIR_958_NonTargetingControlGuideFor Mouse_0334	MetsMIR_958	NonTargetingControlGuideForMo use_0334	CTTTATTCGGTTCATGT CG	MouseDum my
MouseDummy_MetsMIR_959_NonTargetingControlGuideFor Mouse_0335	MetsMIR_959	NonTargetingControlGuideForMo use_0335	AAGCGTACCCACTCGT TAA	MouseDum my
MouseDummy_MetsMIR_960_NonTargetingControlGuideFor Mouse_0336	MetsMIR_960	NonTargetingControlGuideForMo use_0336	ATCCGAGATCTGCGAAT TAT	MouseDum my
MouseDummy_MetsMIR_961_NonTargetingControlGuideFor Mouse_0337	MetsMIR_961	NonTargetingControlGuideForMo use_0337	TCTTGACTCCGACTTCG GGC	MouseDum my
MouseDummy_MetsMIR_962_NonTargetingControlGuideFor Mouse_0338	MetsMIR_962	NonTargetingControlGuideForMo use_0338	ACCATGATGTCACCGCC GCA	MouseDum my
MouseDummy_MetsMIR_963_NonTargetingControlGuideFor Mouse_0339	MetsMIR_963	NonTargetingControlGuideForMo use_0339	TCAACTTAACCTCGAGT CCG	MouseDum my
MouseDummy_MetsMIR_964_NonTargetingControlGuideFor Mouse_0340	MetsMIR_964	NonTargetingControlGuideForMo use_0340	CTACGAGGGCCGCGAG CGGT	MouseDum my
MouseDummy_MetsMIR_965_NonTargetingControlGuideFor Mouse_0341	MetsMIR_965	NonTargetingControlGuideForMo use_0341	ATGTGTATAGCGGCGT AGG	MouseDum my
MouseDummy_MetsMIR_966_NonTargetingControlGuideFor Mouse_0342	MetsMIR_966	NonTargetingControlGuideForMo use_0342	TGTAACGATCTGGGCGG TCT	MouseDum my
MouseDummy_MetsMIR_967_NonTargetingControlGuideFor Mouse_0343	MetsMIR_967	NonTargetingControlGuideForMo use_0343	CCGGATCGGCTACGCTA CGG	MouseDum my
MouseDummy_MetsMIR_968_NonTargetingControlGuideFor Mouse_0344	MetsMIR_968	NonTargetingControlGuideForMo use_0344	CAATATTCGACCTACGC TCC	MouseDum my
MouseDummy_MetsMIR_969_NonTargetingControlGuideFor Mouse_0345	MetsMIR_969	NonTargetingControlGuideForMo use_0345	TGATAACGATGCGAAC TCG	MouseDum my
MouseDummy_MetsMIR_970_NonTargetingControlGuideFor Mouse_0346	MetsMIR_970	NonTargetingControlGuideForMo use_0346	TGATTGGGGTTCGTTCG CCA	MouseDum my
MouseDummy_MetsMIR_971_NonTargetingControlGuideFor Mouse_0347	MetsMIR_971	NonTargetingControlGuideForMo use_0347	TAGTCAGTCGGCCTCCG TGC	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BS

MouseDummy_MetsMIR_972_NonTargetingControlGuideFor Mouse_0348	MetsMIR_972	NonTargetingControlGuideForMo use_0348	AACCGGCTGCGGTTTG CAA	MouseDum my
MouseDummy_MetsMIR_973_NonTargetingControlGuideFor Mouse_0349	MetsMIR_973	NonTargetingControlGuideForMo use_0349	GTGCCTGATAGTGTGAA GCG	MouseDum my
MouseDummy_MetsMIR_974_NonTargetingControlGuideFor Mouse_0350	MetsMIR_974	NonTargetingControlGuideForMo use_0350	CGACGACCCATTTCGGT TAT	MouseDum my
MouseDummy_MetsMIR_975_NonTargetingControlGuideFor Mouse_0351	MetsMIR_975	NonTargetingControlGuideForMo use_0351	GGGGGGTCTAATACCGA TTG	MouseDum my
MouseDummy_MetsMIR_976_NonTargetingControlGuideFor Mouse_0352	MetsMIR_976	NonTargetingControlGuideForMo use_0352	TTTTAGACCTAATTCGC GC	MouseDum my
MouseDummy_MetsMIR_977_NonTargetingControlGuideFor Mouse_0353	MetsMIR_977	NonTargetingControlGuideForMo use_0353	TCTACACGCGCGTTCAA CCG	MouseDum my
MouseDummy_MetsMIR_978_NonTargetingControlGuideFor Mouse_0354	MetsMIR_978	NonTargetingControlGuideForMo use_0354	CGGGGCGTACATGTGTG GCC	MouseDum my
MouseDummy_MetsMIR_979_NonTargetingControlGuideFor Mouse_0355	MetsMIR_979	NonTargetingControlGuideForMo use_0355	CGACTCGCTTAACCGTG CAG	MouseDum my
MouseDummy_MetsMIR_980_NonTargetingControlGuideFor Mouse_0356	MetsMIR_980	NonTargetingControlGuideForMo use_0356	CGCTACCTATCGATAAA CCA	MouseDum my
MouseDummy_MetsMIR_981_NonTargetingControlGuideFor Mouse_0357	MetsMIR_981	NonTargetingControlGuideForMo use_0357	ATATGAGCCCGACCTCT CGA	MouseDum my
MouseDummy_MetsMIR_982_NonTargetingControlGuideFor Mouse_0358	MetsMIR_982	NonTargetingControlGuideForMo use_0358	CGCGCCGAGGGCCTCG TTAC	MouseDum my
MouseDummy_MetsMIR_983_NonTargetingControlGuideFor Mouse_0359	MetsMIR_983	NonTargetingControlGuideForMo use_0359	CTCCGACGACTACGCAA GGA	MouseDum my
MouseDummy_MetsMIR_984_NonTargetingControlGuideFor Mouse_0360	MetsMIR_984	NonTargetingControlGuideForMo use_0360	TATCTCTGGATGCCGTC GGT	MouseDum my
MouseDummy_MetsMIR_985_NonTargetingControlGuideFor Mouse_0361	MetsMIR_985	NonTargetingControlGuideForMo use_0361	TTTCAGTTGTGCGCGCA CTC	MouseDum my
MouseDummy_MetsMIR_986_NonTargetingControlGuideFor Mouse_0362	MetsMIR_986	NonTargetingControlGuideForMo use_0362	AGTTTACGTGGTCCGAT GTC	MouseDum my
MouseDummy_MetsMIR_987_NonTargetingControlGuideFor Mouse_0363	MetsMIR_987	NonTargetingControlGuideForMo use_0363	GCGTGACGCGATCAAAC GGT	MouseDum my
MouseDummy_MetsMIR_988_NonTargetingControlGuideFor Mouse_0364	MetsMIR_988	NonTargetingControlGuideForMo use_0364	ACTAGCTCCATAACGTG TAC	MouseDum my
MouseDummy_MetsMIR_989_NonTargetingControlGuideFor Mouse_0365	MetsMIR_989	NonTargetingControlGuideForMo use_0365	CGCATAGTGATACCGTTG CGC	MouseDum my
MouseDummy_MetsMIR_990_NonTargetingControlGuideFor Mouse_0366	MetsMIR_990	NonTargetingControlGuideForMo use_0366	CCGCTTGGGCGCTATTA ATT	MouseDum my
MouseDummy_MetsMIR_991_NonTargetingControlGuideFor Mouse_0367	MetsMIR_991	NonTargetingControlGuideForMo use_0367	GCTTCATTATTAACCGGC GT	MouseDum my
MouseDummy_MetsMIR_992_NonTargetingControlGuideFor Mouse_0368	MetsMIR_992	NonTargetingControlGuideForMo use_0368	TATACTCGCCTGTCACA GCG	MouseDum my
MouseDummy_MetsMIR_993_NonTargetingControlGuideFor Mouse_0369	MetsMIR_993	NonTargetingControlGuideForMo use_0369	ACGCTGTTCTGTAACCGC GGG	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BT

MouseDummy_MetsMIR_994_NonTargetingControlGuideFor Mouse_0370	MetsMIR_994	NonTargetingControlGuideForMo use_0370	AGACGTTGATTACCGG CCA	MouseDum my
MouseDummy_MetsMIR_995_NonTargetingControlGuideFor Mouse_0371	MetsMIR_995	NonTargetingControlGuideForMo use_0371	AACGGTCAAATCCGTGA GGG	MouseDum my
MouseDummy_MetsMIR_996_NonTargetingControlGuideFor Mouse_0372	MetsMIR_996	NonTargetingControlGuideForMo use_0372	CATAGCAACGCCAAAC TCG	MouseDum my
MouseDummy_MetsMIR_997_NonTargetingControlGuideFor Mouse_0373	MetsMIR_997	NonTargetingControlGuideForMo use_0373	CGTCAATGTGTCCGGAC GGT	MouseDum my
MouseDummy_MetsMIR_998_NonTargetingControlGuideFor Mouse_0374	MetsMIR_998	NonTargetingControlGuideForMo use_0374	ACTGTCCATTGTACGA CGG	MouseDum my
MouseDummy_MetsMIR_999_NonTargetingControlGuideFor Mouse_0375	MetsMIR_999	NonTargetingControlGuideForMo use_0375	CTATCGGCCCGCAGTGA TGG	MouseDum my
MouseDummy_MetsMIR_1000_NonTargetingControlGuideFor Mouse_0376	MetsMIR_1000	NonTargetingControlGuideForMo use_0376	GATCGAGTGACACCCAC CCG	MouseDum my
MouseDummy_MetsMIR_1001_NonTargetingControlGuideFor Mouse_0377	MetsMIR_1001	NonTargetingControlGuideForMo use_0377	TCCGCGGCAACGTTTGG GTC	MouseDum my
MouseDummy_MetsMIR_1002_NonTargetingControlGuideFor Mouse_0378	MetsMIR_1002	NonTargetingControlGuideForMo use_0378	ACTTGTATACGACGGCT AGA	MouseDum my
MouseDummy_MetsMIR_1003_NonTargetingControlGuideFor Mouse_0379	MetsMIR_1003	NonTargetingControlGuideForMo use_0379	TTCGACTGCGCACGCCA TGA	MouseDum my
MouseDummy_MetsMIR_1004_NonTargetingControlGuideFor Mouse_0380	MetsMIR_1004	NonTargetingControlGuideForMo use_0380	AGCATGGAGTCAACGTC CGC	MouseDum my
MouseDummy_MetsMIR_1005_NonTargetingControlGuideFor Mouse_0381	MetsMIR_1005	NonTargetingControlGuideForMo use_0381	TGACTTGACACGTTCTGA TAT	MouseDum my
MouseDummy_MetsMIR_1006_NonTargetingControlGuideFor Mouse_0382	MetsMIR_1006	NonTargetingControlGuideForMo use_0382	CCGGACTTGTATACCTG AT	MouseDum my
MouseDummy_MetsMIR_1007_NonTargetingControlGuideFor Mouse_0383	MetsMIR_1007	NonTargetingControlGuideForMo use_0383	GTACCATGATAACCGTA CTA	MouseDum my
MouseDummy_MetsMIR_1008_NonTargetingControlGuideFor Mouse_0384	MetsMIR_1008	NonTargetingControlGuideForMo use_0384	CACCGCTGCCCTAGTAC CGG	MouseDum my
MouseDummy_MetsMIR_1009_NonTargetingControlGuideFor Mouse_0385	MetsMIR_1009	NonTargetingControlGuideForMo use_0385	GGATAATCGCGCGGTCT TCA	MouseDum my
MouseDummy_MetsMIR_1010_NonTargetingControlGuideFor Mouse_0386	MetsMIR_1010	NonTargetingControlGuideForMo use_0386	AGGCAGCCCGCGTTAGA GAT	MouseDum my
MouseDummy_MetsMIR_1011_NonTargetingControlGuideFor Mouse_0387	MetsMIR_1011	NonTargetingControlGuideForMo use_0387	TTCTCCATACCGTAACCTC CG	MouseDum my
MouseDummy_MetsMIR_1012_NonTargetingControlGuideFor Mouse_0388	MetsMIR_1012	NonTargetingControlGuideForMo use_0388	ACCCGCATATGCCGCCT AAG	MouseDum my
MouseDummy_MetsMIR_1013_NonTargetingControlGuideFor Mouse_0389	MetsMIR_1013	NonTargetingControlGuideForMo use_0389	CGAAGTACACGGTTCTC TCG	MouseDum my
MouseDummy_MetsMIR_1014_NonTargetingControlGuideFor Mouse_0390	MetsMIR_1014	NonTargetingControlGuideForMo use_0390	CACAATACGCGGCAGAG GTC	MouseDum my
MouseDummy_MetsMIR_1015_NonTargetingControlGuideFor Mouse_0391	MetsMIR_1015	NonTargetingControlGuideForMo use_0391	TGAGACCAATGCCGCGG AAT	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BU

MouseDummy_MetsMIR_1016_NonTargetingControlGuideFor Mouse_0392	MetsMIR_1016	NonTargetingControlGuideForMo use_0392	CTCGACCCAGTAGTAGA CGG	MouseDum my
MouseDummy_MetsMIR_1017_NonTargetingControlGuideFor Mouse_0393	MetsMIR_1017	NonTargetingControlGuideForMo use_0393	CGTACCGACAGTAGTCC AAG	MouseDum my
MouseDummy_MetsMIR_1018_NonTargetingControlGuideFor Mouse_0394	MetsMIR_1018	NonTargetingControlGuideForMo use_0394	AATTGTCTGATCGCGCC ATA	MouseDum my
MouseDummy_MetsMIR_1019_NonTargetingControlGuideFor Mouse_0395	MetsMIR_1019	NonTargetingControlGuideForMo use_0395	TTTAGTCAGACGAGCCG GAT	MouseDum my
MouseDummy_MetsMIR_1020_NonTargetingControlGuideFor Mouse_0396	MetsMIR_1020	NonTargetingControlGuideForMo use_0396	TGCGCGTCCTGTTATGA CCG	MouseDum my
MouseDummy_MetsMIR_1021_NonTargetingControlGuideFor Mouse_0397	MetsMIR_1021	NonTargetingControlGuideForMo use_0397	CTTATCGATTGGGTCA AC	MouseDum my
MouseDummy_MetsMIR_1022_NonTargetingControlGuideFor Mouse_0398	MetsMIR_1022	NonTargetingControlGuideForMo use_0398	AACATGTTAAGTCGCGT TAT	MouseDum my
MouseDummy_MetsMIR_1023_NonTargetingControlGuideFor Mouse_0399	MetsMIR_1023	NonTargetingControlGuideForMo use_0399	TACGACATCGCATGGTA ACG	MouseDum my
MouseDummy_MetsMIR_1024_NonTargetingControlGuideFor Mouse_0400	MetsMIR_1024	NonTargetingControlGuideForMo use_0400	CAGCTGACCGTTAATCG ATA	MouseDum my
MouseDummy_MetsMIR_1025_NonTargetingControlGuideFor Mouse_0401	MetsMIR_1025	NonTargetingControlGuideForMo use_0401	TACACGCCCTAAGCTAG TAA	MouseDum my
MouseDummy_MetsMIR_1026_NonTargetingControlGuideFor Mouse_0402	MetsMIR_1026	NonTargetingControlGuideForMo use_0402	CCGACTGCCGAGCTAGG CGT	MouseDum my
MouseDummy_MetsMIR_1027_NonTargetingControlGuideFor Mouse_0403	MetsMIR_1027	NonTargetingControlGuideForMo use_0403	GTTGCGGCGACCTAGAT ATC	MouseDum my
MouseDummy_MetsMIR_1028_NonTargetingControlGuideFor Mouse_0404	MetsMIR_1028	NonTargetingControlGuideForMo use_0404	TAGACGTCCACCGACTC TGA	MouseDum my
MouseDummy_MetsMIR_1029_NonTargetingControlGuideFor Mouse_0405	MetsMIR_1029	NonTargetingControlGuideForMo use_0405	TCACGCACGACTAAGGG AAT	MouseDum my
MouseDummy_MetsMIR_1030_NonTargetingControlGuideFor Mouse_0406	MetsMIR_1030	NonTargetingControlGuideForMo use_0406	GCTTATCGTTCCGCTAC GAT	MouseDum my
MouseDummy_MetsMIR_1031_NonTargetingControlGuideFor Mouse_0407	MetsMIR_1031	NonTargetingControlGuideForMo use_0407	ACCGGTTACGCCCGG GAAC	MouseDum my
MouseDummy_MetsMIR_1032_NonTargetingControlGuideFor Mouse_0408	MetsMIR_1032	NonTargetingControlGuideForMo use_0408	CTACCCGTCTATTACGAT CT	MouseDum my
MouseDummy_MetsMIR_1033_NonTargetingControlGuideFor Mouse_0409	MetsMIR_1033	NonTargetingControlGuideForMo use_0409	TGTCGCCGATGGTCAGT CGC	MouseDum my
MouseDummy_MetsMIR_1034_NonTargetingControlGuideFor Mouse_0410	MetsMIR_1034	NonTargetingControlGuideForMo use_0410	CGGGCGGTAAGAGCTCT ACG	MouseDum my
MouseDummy_MetsMIR_1035_NonTargetingControlGuideFor Mouse_0411	MetsMIR_1035	NonTargetingControlGuideForMo use_0411	GGCACCGTTCCGAAACC GAC	MouseDum my
MouseDummy_MetsMIR_1036_NonTargetingControlGuideFor Mouse_0412	MetsMIR_1036	NonTargetingControlGuideForMo use_0412	TCTCGAATAAATTTCTC GC	MouseDum my
MouseDummy_MetsMIR_1037_NonTargetingControlGuideFor Mouse_0413	MetsMIR_1037	NonTargetingControlGuideForMo use_0413	GCATTCTTGAGCTCCGC GCC	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BV

MouseDummy_MetsMIR_1038_NonTargetingControlGuideForMouse_0414	MetsMIR_1038	NonTargetingControlGuideForMouse_0414	ATTCGGACCGTTATCTCAC	MouseDummy
MouseDummy_MetsMIR_1039_NonTargetingControlGuideForMouse_0415	MetsMIR_1039	NonTargetingControlGuideForMouse_0415	GTTCTTCAAAGACGGGC	MouseDummy
MouseDummy_MetsMIR_1040_NonTargetingControlGuideForMouse_0418	MetsMIR_1040	NonTargetingControlGuideForMouse_0418	TTCCCCGCCCGTGGGT	MouseDummy
MouseDummy_MetsMIR_1041_NonTargetingControlGuideForMouse_0417	MetsMIR_1041	NonTargetingControlGuideForMouse_0417	TAGTAGGTTGATCGCGT	MouseDummy
MouseDummy_MetsMIR_1042_NonTargetingControlGuideForMouse_0418	MetsMIR_1042	NonTargetingControlGuideForMouse_0418	CGCTCGGTGACGTATAC	MouseDummy
MouseDummy_MetsMIR_1043_NonTargetingControlGuideForMouse_0419	MetsMIR_1043	NonTargetingControlGuideForMouse_0419	ATTATAGCAGCCCCCG	MouseDummy
MouseDummy_MetsMIR_1044_NonTargetingControlGuideForMouse_0420	MetsMIR_1044	NonTargetingControlGuideForMouse_0420	CGGTCACTCGTCTTCCC	MouseDummy
MouseDummy_MetsMIR_1045_NonTargetingControlGuideForMouse_0421	MetsMIR_1045	NonTargetingControlGuideForMouse_0421	TGGATCCGCTTCGCACG	MouseDummy
MouseDummy_MetsMIR_1046_NonTargetingControlGuideForMouse_0422	MetsMIR_1046	NonTargetingControlGuideForMouse_0422	AAACCCCCGCGCGGAG	MouseDummy
MouseDummy_MetsMIR_1047_NonTargetingControlGuideForMouse_0423	MetsMIR_1047	NonTargetingControlGuideForMouse_0423	CATAGATCCGCGATTGT	MouseDummy
MouseDummy_MetsMIR_1048_NonTargetingControlGuideForMouse_0424	MetsMIR_1048	NonTargetingControlGuideForMouse_0424	GTTTCAGGAACGACGGC	MouseDummy
MouseDummy_MetsMIR_1049_NonTargetingControlGuideForMouse_0425	MetsMIR_1049	NonTargetingControlGuideForMouse_0425	ACCGTCTCTATTATACGG	MouseDummy
MouseDummy_MetsMIR_1050_NonTargetingControlGuideForMouse_0426	MetsMIR_1050	NonTargetingControlGuideForMouse_0426	TGTTCCGCCCTCTACGA	MouseDummy
MouseDummy_MetsMIR_1051_NonTargetingControlGuideForMouse_0427	MetsMIR_1051	NonTargetingControlGuideForMouse_0427	AACTAACTCACTACGCA	MouseDummy
MouseDummy_MetsMIR_1052_NonTargetingControlGuideForMouse_0428	MetsMIR_1052	NonTargetingControlGuideForMouse_0428	TCGCTCCGCTAGTAGTG	MouseDummy
MouseDummy_MetsMIR_1053_NonTargetingControlGuideForMouse_0429	MetsMIR_1053	NonTargetingControlGuideForMouse_0429	TCGGCGCAGCCTAATGT	MouseDummy
MouseDummy_MetsMIR_1054_NonTargetingControlGuideForMouse_0430	MetsMIR_1054	NonTargetingControlGuideForMouse_0430	CAACCTGCCTAGCGACC	MouseDummy
MouseDummy_MetsMIR_1055_NonTargetingControlGuideForMouse_0431	MetsMIR_1055	NonTargetingControlGuideForMouse_0431	GTCGGCTCATCGGAAAA	MouseDummy
MouseDummy_MetsMIR_1056_NonTargetingControlGuideForMouse_0432	MetsMIR_1056	NonTargetingControlGuideForMouse_0432	GACAGTTGACGCGACGG	MouseDummy
MouseDummy_MetsMIR_1057_NonTargetingControlGuideForMouse_0433	MetsMIR_1057	NonTargetingControlGuideForMouse_0433	CCCTGATGTGCTATACG	MouseDummy
MouseDummy_MetsMIR_1058_NonTargetingControlGuideForMouse_0434	MetsMIR_1058	NonTargetingControlGuideForMouse_0434	TAACCTTGATCAACCG	MouseDummy
MouseDummy_MetsMIR_1059_NonTargetingControlGuideForMouse_0435	MetsMIR_1059	NonTargetingControlGuideForMouse_0435	CTATTGGGAGCGGCCTC	MouseDummy

Figure 18 (supplemental tables continued)

FIG. 18BW

MouseDummy_MetsMIR_1060_NonTargetingControlGuideFor Mouse_0436	MetsMIR_1060	NonTargetingControlGuideForMo use_0436	GCTTATCGTCATGCGGG TGA	MouseDum my
MouseDummy_MetsMIR_1061_NonTargetingControlGuideFor Mouse_0437	MetsMIR_1061	NonTargetingControlGuideForMo use_0437	TGATGCCGGTACCCGTA ACT	MouseDum my
MouseDummy_MetsMIR_1062_NonTargetingControlGuideFor Mouse_0438	MetsMIR_1062	NonTargetingControlGuideForMo use_0438	GGCGCTACTGTAATGAC GGT	MouseDum my
MouseDummy_MetsMIR_1063_NonTargetingControlGuideFor Mouse_0439	MetsMIR_1063	NonTargetingControlGuideForMo use_0439	ACGATTGCTGACGCTA TCA	MouseDum my
MouseDummy_MetsMIR_1064_NonTargetingControlGuideFor Mouse_0440	MetsMIR_1064	NonTargetingControlGuideForMo use_0440	TAATGCGAATGCGACCT CTC	MouseDum my
MouseDummy_MetsMIR_1065_NonTargetingControlGuideFor Mouse_0441	MetsMIR_1065	NonTargetingControlGuideForMo use_0441	TTAACTGGGGGACCGG ACG	MouseDum my
MouseDummy_MetsMIR_1066_NonTargetingControlGuideFor Mouse_0442	MetsMIR_1066	NonTargetingControlGuideForMo use_0442	CGCAGAGTTCAAATCCG CGC	MouseDum my
MouseDummy_MetsMIR_1067_NonTargetingControlGuideFor Mouse_0443	MetsMIR_1067	NonTargetingControlGuideForMo use_0443	TGTTTATGCGGAACCTCC GC	MouseDum my
MouseDummy_MetsMIR_1068_NonTargetingControlGuideFor Mouse_0444	MetsMIR_1068	NonTargetingControlGuideForMo use_0444	ACAGACCGAATAACCGA CGA	MouseDum my
MouseDummy_MetsMIR_1069_NonTargetingControlGuideFor Mouse_0445	MetsMIR_1069	NonTargetingControlGuideForMo use_0445	TGACGCTAAGGCCCGCT AAC	MouseDum my
MouseDummy_MetsMIR_1070_NonTargetingControlGuideFor Mouse_0446	MetsMIR_1070	NonTargetingControlGuideForMo use_0446	ACTCGCGAAACCGTACA TGA	MouseDum my
MouseDummy_MetsMIR_1071_NonTargetingControlGuideFor Mouse_0447	MetsMIR_1071	NonTargetingControlGuideForMo use_0447	ATAGGCGCACGGGCTAC TCC	MouseDum my
MouseDummy_MetsMIR_1072_NonTargetingControlGuideFor Mouse_0448	MetsMIR_1072	NonTargetingControlGuideForMo use_0448	AGATTCTCGCGTAACCA GAG	MouseDum my
MouseDummy_MetsMIR_1073_NonTargetingControlGuideFor Mouse_0449	MetsMIR_1073	NonTargetingControlGuideForMo use_0449	ACAATTCGGTTTATGCG CGT	MouseDum my
MouseDummy_MetsMIR_1074_NonTargetingControlGuideFor Mouse_0450	MetsMIR_1074	NonTargetingControlGuideForMo use_0450	CGGGTTATAGATAGTCG TCG	MouseDum my
MouseDummy_MetsMIR_1075_NonTargetingControlGuideFor Mouse_0451	MetsMIR_1075	NonTargetingControlGuideForMo use_0451	AGGGTAGGCTGACCCG CGA	MouseDum my
MouseDummy_MetsMIR_1076_NonTargetingControlGuideFor Mouse_0452	MetsMIR_1076	NonTargetingControlGuideForMo use_0452	AGCTATCCCCACGTTTCG CGG	MouseDum my
MouseDummy_MetsMIR_1077_NonTargetingControlGuideFor Mouse_0453	MetsMIR_1077	NonTargetingControlGuideForMo use_0453	ATCATGCCTTCGCATTAA CC	MouseDum my
MouseDummy_MetsMIR_1078_NonTargetingControlGuideFor Mouse_0454	MetsMIR_1078	NonTargetingControlGuideForMo use_0454	TGGGTCTTAAACGCGGT TCA	MouseDum my
MouseDummy_MetsMIR_1079_NonTargetingControlGuideFor Mouse_0455	MetsMIR_1079	NonTargetingControlGuideForMo use_0455	CATACTGTTATCGACCC GCA	MouseDum my
MouseDummy_MetsMIR_1080_NonTargetingControlGuideFor Mouse_0456	MetsMIR_1080	NonTargetingControlGuideForMo use_0456	TGAGCGCTTTCCCGATC CGG	MouseDum my
MouseDummy_MetsMIR_1081_NonTargetingControlGuideFor Mouse_0457	MetsMIR_1081	NonTargetingControlGuideForMo use_0457	GCCTTCCTCGCAGACCC GAC	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BX

MouseDummy_MetsMIR_1082_NonTargetingControlGuideFor Mouse_0458	MetsMIR_1082	NonTargetingControlGuideForMo use_0458	CCCTTAAAGTGACGGAC GAA	MouseDum my
MouseDummy_MetsMIR_1083_NonTargetingControlGuideFor Mouse_0459	MetsMIR_1083	NonTargetingControlGuideForMo use_0459	TGGTTCGACCAACATGG TTC	MouseDum my
MouseDummy_MetsMIR_1084_NonTargetingControlGuideFor Mouse_0460	MetsMIR_1084	NonTargetingControlGuideForMo use_0460	TGGGCCATAGTGGCGC GTGA	MouseDum my
MouseDummy_MetsMIR_1085_NonTargetingControlGuideFor Mouse_0461	MetsMIR_1085	NonTargetingControlGuideForMo use_0461	TCATATTAGACAATCTCC GC	MouseDum my
MouseDummy_MetsMIR_1086_NonTargetingControlGuideFor Mouse_0462	MetsMIR_1086	NonTargetingControlGuideForMo use_0462	TAGCCGAGTTCACGCCA GTA	MouseDum my
MouseDummy_MetsMIR_1087_NonTargetingControlGuideFor Mouse_0463	MetsMIR_1087	NonTargetingControlGuideForMo use_0463	ACCAGGCGCGGACCGC ACAT	MouseDum my
MouseDummy_MetsMIR_1088_NonTargetingControlGuideFor Mouse_0464	MetsMIR_1088	NonTargetingControlGuideForMo use_0464	GCGCGAGTGCCAAACGA GTG	MouseDum my
MouseDummy_MetsMIR_1089_NonTargetingControlGuideFor Mouse_0465	MetsMIR_1089	NonTargetingControlGuideForMo use_0465	AACGGCTGCGCCCGCG GCAA	MouseDum my
MouseDummy_MetsMIR_1090_NonTargetingControlGuideFor Mouse_0466	MetsMIR_1090	NonTargetingControlGuideForMo use_0466	ACGGTCGAGCAGGTTA TGA	MouseDum my
MouseDummy_MetsMIR_1091_NonTargetingControlGuideFor Mouse_0467	MetsMIR_1091	NonTargetingControlGuideForMo use_0467	TCATGTTGCGTCGTCCG TTA	MouseDum my
MouseDummy_MetsMIR_1092_NonTargetingControlGuideFor Mouse_0468	MetsMIR_1092	NonTargetingControlGuideForMo use_0468	CGCTCCGTATCCTCTCA TAA	MouseDum my
MouseDummy_MetsMIR_1093_NonTargetingControlGuideFor Mouse_0469	MetsMIR_1093	NonTargetingControlGuideForMo use_0469	AAACGAGGCTGTTCTGA CAC	MouseDum my
MouseDummy_MetsMIR_1094_NonTargetingControlGuideFor Mouse_0470	MetsMIR_1094	NonTargetingControlGuideForMo use_0470	CACGTATACTTAGGTCA GCG	MouseDum my
MouseDummy_MetsMIR_1095_NonTargetingControlGuideFor Mouse_0471	MetsMIR_1095	NonTargetingControlGuideForMo use_0471	ATCGAGCACCGAGTTGT GAT	MouseDum my
MouseDummy_MetsMIR_1096_NonTargetingControlGuideFor Mouse_0472	MetsMIR_1096	NonTargetingControlGuideForMo use_0472	TCACTTTAAGCACCCCG CGC	MouseDum my
MouseDummy_MetsMIR_1097_NonTargetingControlGuideFor Mouse_0473	MetsMIR_1097	NonTargetingControlGuideForMo use_0473	CTACTTGTGACGACCTC GCG	MouseDum my
MouseDummy_MetsMIR_1098_NonTargetingControlGuideFor Mouse_0474	MetsMIR_1098	NonTargetingControlGuideForMo use_0474	CTACGGATAGGCGCGG GTGA	MouseDum my
MouseDummy_MetsMIR_1099_NonTargetingControlGuideFor Mouse_0475	MetsMIR_1099	NonTargetingControlGuideForMo use_0475	GCTTACCTACTCCGCC CGC	MouseDum my
MouseDummy_MetsMIR_1100_NonTargetingControlGuideFor Mouse_0476	MetsMIR_1100	NonTargetingControlGuideForMo use_0476	TAAACGCACTCACGTG CGG	MouseDum my
MouseDummy_MetsMIR_1101_NonTargetingControlGuideFor Mouse_0477	MetsMIR_1101	NonTargetingControlGuideForMo use_0477	ACTCACCTCGCACGATC GTA	MouseDum my
MouseDummy_MetsMIR_1102_NonTargetingControlGuideFor Mouse_0478	MetsMIR_1102	NonTargetingControlGuideForMo use_0478	ACACCCGTGTATGCACC GGG	MouseDum my
MouseDummy_MetsMIR_1103_NonTargetingControlGuideFor Mouse_0479	MetsMIR_1103	NonTargetingControlGuideForMo use_0479	GCATAACGGCCGAGCAC CAC	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BY

MouseDummy_MetsMIR_1104_NonTargetingControlGuideFor Mouse 0480	MetsMIR_1104	NonTargetingControlGuideForMo use 0480	CCAGCGGACGGCGAAC TCCA	MouseDum my
MouseDummy_MetsMIR_1105_NonTargetingControlGuideFor Mouse 0481	MetsMIR_1105	NonTargetingControlGuideForMo use 0481	TGACCCCTCTTAATCTCCG GT	MouseDum my
MouseDummy_MetsMIR_1108_NonTargetingControlGuideFor Mouse 0482	MetsMIR_1108	NonTargetingControlGuideForMo use 0482	CCCTAAACCATAGTTTC GCC	MouseDum my
MouseDummy_MetsMIR_1107_NonTargetingControlGuideFor Mouse 0483	MetsMIR_1107	NonTargetingControlGuideForMo use 0483	CCCTAGCTCGGTTAGAG AAT	MouseDum my
MouseDummy_MetsMIR_1108_NonTargetingControlGuideFor Mouse 0484	MetsMIR_1108	NonTargetingControlGuideForMo use 0484	GAATTGAGCCGCAACTC GGC	MouseDum my
MouseDummy_MetsMIR_1109_NonTargetingControlGuideFor Mouse 0485	MetsMIR_1109	NonTargetingControlGuideForMo use 0485	TGCTTAGGGCCCTTCGG CGG	MouseDum my
MouseDummy_MetsMIR_1110_NonTargetingControlGuideFor Mouse 0488	MetsMIR_1110	NonTargetingControlGuideForMo use 0488	TCTCCGGGCCGTGTAG ACC	MouseDum my
MouseDummy_MetsMIR_1111_NonTargetingControlGuideFor Mouse 0487	MetsMIR_1111	NonTargetingControlGuideForMo use 0487	ATCACACGTTAAACGGG GCG	MouseDum my
MouseDummy_MetsMIR_1112_NonTargetingControlGuideFor Mouse 0488	MetsMIR_1112	NonTargetingControlGuideForMo use 0488	GATCCCTACGCTTCAA CCT	MouseDum my
MouseDummy_MetsMIR_1113_NonTargetingControlGuideFor Mouse 0489	MetsMIR_1113	NonTargetingControlGuideForMo use 0489	TTATTACGCGCTGAACT TG	MouseDum my
MouseDummy_MetsMIR_1114_NonTargetingControlGuideFor Mouse 0490	MetsMIR_1114	NonTargetingControlGuideForMo use 0490	ACCGTTGTTCCGTCGA AAC	MouseDum my
MouseDummy_MetsMIR_1115_NonTargetingControlGuideFor Mouse 0491	MetsMIR_1115	NonTargetingControlGuideForMo use 0491	CGAGACGAATCCATCAT GCG	MouseDum my
MouseDummy_MetsMIR_1116_NonTargetingControlGuideFor Mouse 0492	MetsMIR_1116	NonTargetingControlGuideForMo use 0492	ATATGCTCGACCCATCG TCC	MouseDum my
MouseDummy_MetsMIR_1117_NonTargetingControlGuideFor Mouse 0493	MetsMIR_1117	NonTargetingControlGuideForMo use 0493	AACCTCGTCTCATGTAC GAA	MouseDum my
MouseDummy_MetsMIR_1118_NonTargetingControlGuideFor Mouse 0494	MetsMIR_1118	NonTargetingControlGuideForMo use 0494	AATTTTTTCGGAATCTAG CG	MouseDum my
MouseDummy_MetsMIR_1119_NonTargetingControlGuideFor Mouse 0495	MetsMIR_1119	NonTargetingControlGuideForMo use 0495	AGCGAGTGTCCGTGACG TTC	MouseDum my
MouseDummy_MetsMIR_1120_NonTargetingControlGuideFor Mouse 0496	MetsMIR_1120	NonTargetingControlGuideForMo use 0496	AACGTCACCAACCTCGA TCC	MouseDum my
MouseDummy_MetsMIR_1121_NonTargetingControlGuideFor Mouse 0497	MetsMIR_1121	NonTargetingControlGuideForMo use 0497	GTCTAGATCCTATCGG GAG	MouseDum my
MouseDummy_MetsMIR_1122_NonTargetingControlGuideFor Mouse 0498	MetsMIR_1122	NonTargetingControlGuideForMo use 0498	CATATACTCTTGCCTAG AC	MouseDum my
MouseDummy_MetsMIR_1123_NonTargetingControlGuideFor Mouse 0499	MetsMIR_1123	NonTargetingControlGuideForMo use 0499	ACGTCGATGCTTATCCG TCT	MouseDum my
MouseDummy_MetsMIR_1124_NonTargetingControlGuideFor Mouse 0500	MetsMIR_1124	NonTargetingControlGuideForMo use 0500	CCGCGCATTACGATCA GTA	MouseDum my
MouseDummy_MetsMIR_1125_NonTargetingControlGuideFor Mouse 0501	MetsMIR_1125	NonTargetingControlGuideForMo use 0501	CCTATTAGTCCGGTTTAG TC	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18BZ

MouseDummy_MetsMIR_1126_NonTargetingControlGuideFor Mouse_0502	MetsMIR_1126	NonTargetingControlGuideForMo use_0502	ACCTCGTGCAAATCGGT GGC	MouseDum my
MouseDummy_MetsMIR_1127_NonTargetingControlGuideFor Mouse_0503	MetsMIR_1127	NonTargetingControlGuideForMo use_0503	CACGCAGGAGCGGCGA CACT	MouseDum my
MouseDummy_MetsMIR_1128_NonTargetingControlGuideFor Mouse_0504	MetsMIR_1128	NonTargetingControlGuideForMo use_0504	GTCGGCCGCTTAACCTT TTC	MouseDum my
MouseDummy_MetsMIR_1129_NonTargetingControlGuideFor Mouse_0505	MetsMIR_1129	NonTargetingControlGuideForMo use_0505	TGATCAACGTCGGTGGA CGG	MouseDum my
MouseDummy_MetsMIR_1130_NonTargetingControlGuideFor Mouse_0506	MetsMIR_1130	NonTargetingControlGuideForMo use_0506	GAACCACGAGCGAGCGT ATA	MouseDum my
MouseDummy_MetsMIR_1131_NonTargetingControlGuideFor Mouse_0507	MetsMIR_1131	NonTargetingControlGuideForMo use_0507	GAACCCGGGAAACACGT CCG	MouseDum my
MouseDummy_MetsMIR_1132_NonTargetingControlGuideFor Mouse_0508	MetsMIR_1132	NonTargetingControlGuideForMo use_0508	GGAAACGTTACATTGGA CGC	MouseDum my
MouseDummy_MetsMIR_1133_NonTargetingControlGuideFor Mouse_0509	MetsMIR_1133	NonTargetingControlGuideForMo use_0509	AATAGGAACCTCCGACC CGA	MouseDum my
MouseDummy_MetsMIR_1134_NonTargetingControlGuideFor Mouse_0510	MetsMIR_1134	NonTargetingControlGuideForMo use_0510	GTAGTGGGTGTGATGTC GGG	MouseDum my
MouseDummy_MetsMIR_1135_NonTargetingControlGuideFor Mouse_0511	MetsMIR_1135	NonTargetingControlGuideForMo use_0511	TTTGACATTTCTGCTCG CG	MouseDum my
MouseDummy_MetsMIR_1136_NonTargetingControlGuideFor Mouse_0512	MetsMIR_1136	NonTargetingControlGuideForMo use_0512	AGGTCCGCTGGCGCAAT GGG	MouseDum my
MouseDummy_MetsMIR_1137_NonTargetingControlGuideFor Mouse_0513	MetsMIR_1137	NonTargetingControlGuideForMo use_0513	CCGCCCGGGTGTGAGTT GAG	MouseDum my
MouseDummy_MetsMIR_1138_NonTargetingControlGuideFor Mouse_0514	MetsMIR_1138	NonTargetingControlGuideForMo use_0514	GCCGAGAGGCGTAAGC GCGA	MouseDum my
MouseDummy_MetsMIR_1139_NonTargetingControlGuideFor Mouse_0515	MetsMIR_1139	NonTargetingControlGuideForMo use_0515	TAATTCGCAACTCGGAT CAT	MouseDum my
MouseDummy_MetsMIR_1140_NonTargetingControlGuideFor Mouse_0516	MetsMIR_1140	NonTargetingControlGuideForMo use_0516	TAAGACTGGGTGTCCCG CGT	MouseDum my
MouseDummy_MetsMIR_1141_NonTargetingControlGuideFor Mouse_0517	MetsMIR_1141	NonTargetingControlGuideForMo use_0517	TTACGCAAAGCTGCGAT GAG	MouseDum my
MouseDummy_MetsMIR_1142_NonTargetingControlGuideFor Mouse_0518	MetsMIR_1142	NonTargetingControlGuideForMo use_0518	GGACACTCGCCGACCCC ACT	MouseDum my
MouseDummy_MetsMIR_1143_NonTargetingControlGuideFor Mouse_0519	MetsMIR_1143	NonTargetingControlGuideForMo use_0519	TACGTCTGCGCAAATAG AAT	MouseDum my
MouseDummy_MetsMIR_1144_NonTargetingControlGuideFor Mouse_0520	MetsMIR_1144	NonTargetingControlGuideForMo use_0520	ATACGCGTGATGACCTT ATG	MouseDum my
MouseDummy_MetsMIR_1145_NonTargetingControlGuideFor Mouse_0521	MetsMIR_1145	NonTargetingControlGuideForMo use_0521	CGCGCGACTGCAGCTAA TCT	MouseDum my
MouseDummy_MetsMIR_1146_NonTargetingControlGuideFor Mouse_0522	MetsMIR_1146	NonTargetingControlGuideForMo use_0522	ACTTCCCGCGGTTCCGT TGA	MouseDum my
MouseDummy_MetsMIR_1147_NonTargetingControlGuideFor Mouse_0523	MetsMIR_1147	NonTargetingControlGuideForMo use_0523	TTATGACCTCGATGCGA CAT	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18CA

MouseDummy_MetsMIR_1148_NonTargetingControlGuideFor Mouse 0524	MetsMIR_1148	NonTargetingControlGuideForMo use 0524	TGACGCGCCTCGGACCA ACC	MouseDum my
MouseDummy_MetsMIR_1149_NonTargetingControlGuideFor Mouse 0525	MetsMIR_1149	NonTargetingControlGuideForMo use 0525	TATCCCAGATCCGGAAC TAG	MouseDum my
MouseDummy_MetsMIR_1150_NonTargetingControlGuideFor Mouse 0526	MetsMIR_1150	NonTargetingControlGuideForMo use 0526	AGTACACTACATCGACTT CG	MouseDum my
MouseDummy_MetsMIR_1151_NonTargetingControlGuideFor Mouse 0527	MetsMIR_1151	NonTargetingControlGuideForMo use 0527	TTCCAGTATACCGAATTC GC	MouseDum my
MouseDummy_MetsMIR_1152_NonTargetingControlGuideFor Mouse 0528	MetsMIR_1152	NonTargetingControlGuideForMo use 0528	CCCGGCAGACTAACTAG CGC	MouseDum my
MouseDummy_MetsMIR_1153_NonTargetingControlGuideFor Mouse 0529	MetsMIR_1153	NonTargetingControlGuideForMo use 0529	GCAACGGACAGTCATCG AAC	MouseDum my
MouseDummy_MetsMIR_1154_NonTargetingControlGuideFor Mouse 0530	MetsMIR_1154	NonTargetingControlGuideForMo use 0530	GCTAGCGTTCAGCCCGA TGT	MouseDum my
MouseDummy_MetsMIR_1155_NonTargetingControlGuideFor Mouse 0531	MetsMIR_1155	NonTargetingControlGuideForMo use 0531	TATGAACCTAGGCGTAAA CGG	MouseDum my
MouseDummy_MetsMIR_1156_NonTargetingControlGuideFor Mouse 0532	MetsMIR_1156	NonTargetingControlGuideForMo use 0532	CGCGCGGATATTTAAAC TTC	MouseDum my
MouseDummy_MetsMIR_1157_NonTargetingControlGuideFor Mouse 0533	MetsMIR_1157	NonTargetingControlGuideForMo use 0533	AAGCAGCGACTACTCGA CGC	MouseDum my
MouseDummy_MetsMIR_1158_NonTargetingControlGuideFor Mouse 0534	MetsMIR_1158	NonTargetingControlGuideForMo use 0534	CCCCTATAGCGCGCTA AGG	MouseDum my
MouseDummy_MetsMIR_1159_NonTargetingControlGuideFor Mouse 0535	MetsMIR_1159	NonTargetingControlGuideForMo use 0535	CCGAGATGGCTCGGATA GAC	MouseDum my
MouseDummy_MetsMIR_1160_NonTargetingControlGuideFor Mouse 0536	MetsMIR_1160	NonTargetingControlGuideForMo use 0536	AGCACGCGCACAAGAGC CGC	MouseDum my
MouseDummy_MetsMIR_1161_NonTargetingControlGuideFor Mouse 0537	MetsMIR_1161	NonTargetingControlGuideForMo use 0537	TCCGTAAAGTCGAGCGA CGT	MouseDum my
MouseDummy_MetsMIR_1162_NonTargetingControlGuideFor Mouse 0538	MetsMIR_1162	NonTargetingControlGuideForMo use 0538	GTCTAACATCGGCGCAC GTG	MouseDum my
MouseDummy_MetsMIR_1163_NonTargetingControlGuideFor Mouse 0539	MetsMIR_1163	NonTargetingControlGuideForMo use 0539	TTCAACGACGGAAGACG CGC	MouseDum my
MouseDummy_MetsMIR_1164_NonTargetingControlGuideFor Mouse 0540	MetsMIR_1164	NonTargetingControlGuideForMo use 0540	CGCTCGGTAAGTCTGAT TGG	MouseDum my
MouseDummy_MetsMIR_1165_NonTargetingControlGuideFor Mouse 0541	MetsMIR_1165	NonTargetingControlGuideForMo use 0541	TAGGGGTGCGCATTAGA CTA	MouseDum my
MouseDummy_MetsMIR_1166_NonTargetingControlGuideFor Mouse 0542	MetsMIR_1166	NonTargetingControlGuideForMo use 0542	GCCCCGACCTCCACGTAA ATC	MouseDum my
MouseDummy_MetsMIR_1167_NonTargetingControlGuideFor Mouse 0543	MetsMIR_1167	NonTargetingControlGuideForMo use 0543	TATGAACATCTCGGCG ACG	MouseDum my
MouseDummy_MetsMIR_1168_NonTargetingControlGuideFor Mouse 0544	MetsMIR_1168	NonTargetingControlGuideForMo use 0544	CTGCTTGAACGCCTAGA CGG	MouseDum my
MouseDummy_MetsMIR_1169_NonTargetingControlGuideFor Mouse 0545	MetsMIR_1169	NonTargetingControlGuideForMo use 0545	GGGCCGGCACTCTGTCTG GAC	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18CB

MouseDummy_MetsMIR_1170_NonTargetingControlGuideFor Mouse_0546	MetsMIR_1170	NonTargetingControlGuideForMo use_0546	GCCGACCAACGATGACC ACG	MouseDum my
MouseDummy_MetsMIR_1171_NonTargetingControlGuideFor Mouse_0547	MetsMIR_1171	NonTargetingControlGuideForMo use_0547	ATCGCGGCGCGTTAAGG CAG	MouseDum my
MouseDummy_MetsMIR_1172_NonTargetingControlGuideFor Mouse_0548	MetsMIR_1172	NonTargetingControlGuideForMo use_0548	CGACGGTTGTGCTGAGG CTT	MouseDum my
MouseDummy_MetsMIR_1173_NonTargetingControlGuideFor Mouse_0549	MetsMIR_1173	NonTargetingControlGuideForMo use_0549	CGGATCCTCCCGTACTA TCC	MouseDum my
MouseDummy_MetsMIR_1174_NonTargetingControlGuideFor Mouse_0550	MetsMIR_1174	NonTargetingControlGuideForMo use_0550	CCATTAGTACGAACATT GCG	MouseDum my
MouseDummy_MetsMIR_1175_NonTargetingControlGuideFor Mouse_0551	MetsMIR_1175	NonTargetingControlGuideForMo use_0551	GCTCATACTGATGAACG TCC	MouseDum my
MouseDummy_MetsMIR_1176_NonTargetingControlGuideFor Mouse_0552	MetsMIR_1176	NonTargetingControlGuideForMo use_0552	GCCGAATCGCGTTATTC CAA	MouseDum my
MouseDummy_MetsMIR_1177_NonTargetingControlGuideFor Mouse_0553	MetsMIR_1177	NonTargetingControlGuideForMo use_0553	CAAGGATCGTGCCGTGA TTT	MouseDum my
MouseDummy_MetsMIR_1178_NonTargetingControlGuideFor Mouse_0554	MetsMIR_1178	NonTargetingControlGuideForMo use_0554	GTTAATATTGTGCCCG CAC	MouseDum my
MouseDummy_MetsMIR_1179_NonTargetingControlGuideFor Mouse_0555	MetsMIR_1179	NonTargetingControlGuideForMo use_0555	CCAACCCGGCATCGTCC GCT	MouseDum my
MouseDummy_MetsMIR_1180_NonTargetingControlGuideFor Mouse_0556	MetsMIR_1180	NonTargetingControlGuideForMo use_0556	GCGAGCGCTATCCCGT GGA	MouseDum my
MouseDummy_MetsMIR_1181_NonTargetingControlGuideFor Mouse_0557	MetsMIR_1181	NonTargetingControlGuideForMo use_0557	GTGTATGACGAGAAGCG AAG	MouseDum my
MouseDummy_MetsMIR_1182_NonTargetingControlGuideFor Mouse_0558	MetsMIR_1182	NonTargetingControlGuideForMo use_0558	GGTGTGACCGCTTTTA CGC	MouseDum my
MouseDummy_MetsMIR_1183_NonTargetingControlGuideFor Mouse_0559	MetsMIR_1183	NonTargetingControlGuideForMo use_0559	GTTATACCACCTACTATG AC	MouseDum my
MouseDummy_MetsMIR_1184_NonTargetingControlGuideFor Mouse_0560	MetsMIR_1184	NonTargetingControlGuideForMo use_0560	GATACGTGAGGTTGCCG GTG	MouseDum my
MouseDummy_MetsMIR_1185_NonTargetingControlGuideFor Mouse_0561	MetsMIR_1185	NonTargetingControlGuideForMo use_0561	TATGCCATATGCCGTTT TT	MouseDum my
MouseDummy_MetsMIR_1186_NonTargetingControlGuideFor Mouse_0562	MetsMIR_1186	NonTargetingControlGuideForMo use_0562	TCATCTACGGTATCGAA AGG	MouseDum my
MouseDummy_MetsMIR_1187_NonTargetingControlGuideFor Mouse_0563	MetsMIR_1187	NonTargetingControlGuideForMo use_0563	AGTGYCACGGCAAAGTC GAG	MouseDum my
MouseDummy_MetsMIR_1188_NonTargetingControlGuideFor Mouse_0564	MetsMIR_1188	NonTargetingControlGuideForMo use_0564	CAATGGCGCTGCCGTT CAG	MouseDum my
MouseDummy_MetsMIR_1189_NonTargetingControlGuideFor Mouse_0565	MetsMIR_1189	NonTargetingControlGuideForMo use_0565	GCTCGCTGATGTGAAT CCG	MouseDum my
MouseDummy_MetsMIR_1190_NonTargetingControlGuideFor Mouse_0566	MetsMIR_1190	NonTargetingControlGuideForMo use_0566	GGATACATCGGCGCGCT AGT	MouseDum my
MouseDummy_MetsMIR_1191_NonTargetingControlGuideFor Mouse_0567	MetsMIR_1191	NonTargetingControlGuideForMo use_0567	CAGCACCACGCGTCGTG CGG	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18CC

MouseDummy_MetsMIR_1182_NonTargetingControlGuideFor Mouse_0568	MetsMIR_1192	NonTargetingControlGuideForMo use_0568	CGATGTGCAAAGTCGGT CAA	MouseDum my
MouseDummy_MetsMIR_1183_NonTargetingControlGuideFor Mouse_0569	MetsMIR_1193	NonTargetingControlGuideForMo use_0569	CGATACGCTATAGAATA GTC	MouseDum my
MouseDummy_MetsMIR_1194_NonTargetingControlGuideFor Mouse_0570	MetsMIR_1194	NonTargetingControlGuideForMo use_0570	GCGTGGAGTTCGTCGCT CTT	MouseDum my
MouseDummy_MetsMIR_1195_NonTargetingControlGuideFor Mouse_0571	MetsMIR_1195	NonTargetingControlGuideForMo use_0571	CACAACGCCCTACCAGCG GAC	MouseDum my
MouseDummy_MetsMIR_1196_NonTargetingControlGuideFor Mouse_0572	MetsMIR_1196	NonTargetingControlGuideForMo use_0572	ATTTGCCCGTCCATACG CGG	MouseDum my
MouseDummy_MetsMIR_1197_NonTargetingControlGuideFor Mouse_0573	MetsMIR_1197	NonTargetingControlGuideForMo use_0573	CGGTAAAGATGGTTATAC CGT	MouseDum my
MouseDummy_MetsMIR_1198_NonTargetingControlGuideFor Mouse_0574	MetsMIR_1198	NonTargetingControlGuideForMo use_0574	GTTTCGCGGGGGCTTCTA TCA	MouseDum my
MouseDummy_MetsMIR_1199_NonTargetingControlGuideFor Mouse_0575	MetsMIR_1199	NonTargetingControlGuideForMo use_0575	GGCACCGCGTTTATTGC ACT	MouseDum my
MouseDummy_MetsMIR_1200_NonTargetingControlGuideFor Mouse_0576	MetsMIR_1200	NonTargetingControlGuideForMo use_0576	TGACAACCGCGGTACCT CTA	MouseDum my
MouseDummy_MetsMIR_1201_NonTargetingControlGuideFor Mouse_0577	MetsMIR_1201	NonTargetingControlGuideForMo use_0577	GGTCGCTCGATTTCATTT AA	MouseDum my
MouseDummy_MetsMIR_1202_NonTargetingControlGuideFor Mouse_0578	MetsMIR_1202	NonTargetingControlGuideForMo use_0578	GACGAAAGGTCTACGA AGT	MouseDum my
MouseDummy_MetsMIR_1203_NonTargetingControlGuideFor Mouse_0579	MetsMIR_1203	NonTargetingControlGuideForMo use_0579	GTACCACTTATCGACCTT GC	MouseDum my
MouseDummy_MetsMIR_1204_NonTargetingControlGuideFor Mouse_0580	MetsMIR_1204	NonTargetingControlGuideForMo use_0580	TTCAGTGTTATGGACG CGC	MouseDum my
MouseDummy_MetsMIR_1205_NonTargetingControlGuideFor Mouse_0581	MetsMIR_1205	NonTargetingControlGuideForMo use_0581	TGAGAGCAAGGCGCATA CGC	MouseDum my
MouseDummy_MetsMIR_1206_NonTargetingControlGuideFor Mouse_0582	MetsMIR_1206	NonTargetingControlGuideForMo use_0582	CATTCTCTGACGAATGC GCC	MouseDum my
MouseDummy_MetsMIR_1207_NonTargetingControlGuideFor Mouse_0583	MetsMIR_1207	NonTargetingControlGuideForMo use_0583	ATCCCCCGACTTAGGG ATT	MouseDum my
MouseDummy_MetsMIR_1208_NonTargetingControlGuideFor Mouse_0584	MetsMIR_1208	NonTargetingControlGuideForMo use_0584	CAGTGCCGGTAGCGGC ACGG	MouseDum my
MouseDummy_MetsMIR_1209_NonTargetingControlGuideFor Mouse_0585	MetsMIR_1209	NonTargetingControlGuideForMo use_0585	ACGACGTCCGCTGTGTG TAT	MouseDum my
MouseDummy_MetsMIR_1210_NonTargetingControlGuideFor Mouse_0586	MetsMIR_1210	NonTargetingControlGuideForMo use_0586	GCGCTCGAACAACATG GTC	MouseDum my
MouseDummy_MetsMIR_1211_NonTargetingControlGuideFor Mouse_0587	MetsMIR_1211	NonTargetingControlGuideForMo use_0587	GGCAACGCACGCTGGG TTGT	MouseDum my
MouseDummy_MetsMIR_1212_NonTargetingControlGuideFor Mouse_0588	MetsMIR_1212	NonTargetingControlGuideForMo use_0588	GCTATTCCGCTCGTCAA TTT	MouseDum my
MouseDummy_MetsMIR_1213_NonTargetingControlGuideFor Mouse_0589	MetsMIR_1213	NonTargetingControlGuideForMo use_0589	TTCATCGCAGATCGATTT CG	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18CD

MouseDummy_MetsMIR_1214_NonTargetingControlGuideFor Mouse_0590	MetsMIR_1214	NonTargetingControlGuideForMo use_0590	TTCATCCTTCGTAGCGC AGC	MouseDum my
MouseDummy_MetsMIR_1215_NonTargetingControlGuideFor Mouse_0591	MetsMIR_1215	NonTargetingControlGuideForMo use_0591	TGGTTCACCACTCGAGA TCG	MouseDum my
MouseDummy_MetsMIR_1218_NonTargetingControlGuideFor Mouse_0592	MetsMIR_1216	NonTargetingControlGuideForMo use_0592	TTGAGCGGACCCCCCTA CAA	MouseDum my
MouseDummy_MetsMIR_1217_NonTargetingControlGuideFor Mouse_0593	MetsMIR_1217	NonTargetingControlGuideForMo use_0593	ATGTGTATGAAGCCCGT CAT	MouseDum my
MouseDummy_MetsMIR_1218_NonTargetingControlGuideFor Mouse_0594	MetsMIR_1218	NonTargetingControlGuideForMo use_0594	TAGGTGGCGCCCAATCG GAC	MouseDum my
MouseDummy_MetsMIR_1219_NonTargetingControlGuideFor Mouse_0595	MetsMIR_1219	NonTargetingControlGuideForMo use_0595	CTCGTGCTACTCCTCGG TTC	MouseDum my
MouseDummy_MetsMIR_1220_NonTargetingControlGuideFor Mouse_0596	MetsMIR_1220	NonTargetingControlGuideForMo use_0596	ATGCCAATGCCGTTGTT AGC	MouseDum my
MouseDummy_MetsMIR_1221_NonTargetingControlGuideFor Mouse_0597	MetsMIR_1221	NonTargetingControlGuideForMo use_0597	CCTACGGCGGAGAACG AGT	MouseDum my
MouseDummy_MetsMIR_1222_NonTargetingControlGuideFor Mouse_0598	MetsMIR_1222	NonTargetingControlGuideForMo use_0598	CGCGCGGATCTTCCGTA CAA	MouseDum my
MouseDummy_MetsMIR_1223_NonTargetingControlGuideFor Mouse_0599	MetsMIR_1223	NonTargetingControlGuideForMo use_0599	CGGCAAGCGCATTCTTA TGG	MouseDum my
MouseDummy_MetsMIR_1224_NonTargetingControlGuideFor Mouse_0600	MetsMIR_1224	NonTargetingControlGuideForMo use_0600	TATGATCGTATGGCCCTT CC	MouseDum my
MouseDummy_MetsMIR_1225_NonTargetingControlGuideFor Mouse_0601	MetsMIR_1225	NonTargetingControlGuideForMo use_0601	ATCGCCTAGCCCCAAGCG ACG	MouseDum my
MouseDummy_MetsMIR_1226_NonTargetingControlGuideFor Mouse_0602	MetsMIR_1226	NonTargetingControlGuideForMo use_0602	GCGATCGCCGGTATAGC TTT	MouseDum my
MouseDummy_MetsMIR_1227_NonTargetingControlGuideFor Mouse_0603	MetsMIR_1227	NonTargetingControlGuideForMo use_0603	CCGTTCAATTATGCTGG CGT	MouseDum my
MouseDummy_MetsMIR_1228_NonTargetingControlGuideFor Mouse_0604	MetsMIR_1228	NonTargetingControlGuideForMo use_0604	CCAGCTAACGTTTATGTA CG	MouseDum my
MouseDummy_MetsMIR_1229_NonTargetingControlGuideFor Mouse_0605	MetsMIR_1229	NonTargetingControlGuideForMo use_0605	TCCCCCTTCGTCGGCGCC AGG	MouseDum my
MouseDummy_MetsMIR_1230_NonTargetingControlGuideFor Mouse_0606	MetsMIR_1230	NonTargetingControlGuideForMo use_0606	TTCTGATTAGATACGTAC GA	MouseDum my
MouseDummy_MetsMIR_1231_NonTargetingControlGuideFor Mouse_0607	MetsMIR_1231	NonTargetingControlGuideForMo use_0607	TTTTCGTCTGACTAAGTCA AG	MouseDum my
MouseDummy_MetsMIR_1232_NonTargetingControlGuideFor Mouse_0608	MetsMIR_1232	NonTargetingControlGuideForMo use_0608	TCAATTAGTGGCCGCCA AGC	MouseDum my
MouseDummy_MetsMIR_1233_NonTargetingControlGuideFor Mouse_0609	MetsMIR_1233	NonTargetingControlGuideForMo use_0609	TCAGCTACGATCGGACC CAA	MouseDum my
MouseDummy_MetsMIR_1234_NonTargetingControlGuideFor Mouse_0610	MetsMIR_1234	NonTargetingControlGuideForMo use_0610	GCACGAACCCGTTCTGTA TGG	MouseDum my
MouseDummy_MetsMIR_1235_NonTargetingControlGuideFor Mouse_0611	MetsMIR_1235	NonTargetingControlGuideForMo use_0611	CAATATGCACGTAGCCT CGT	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18CE

MouseDummy_MetsMIR_1236_NonTargetingControlGuideFor Mouse_0612	MetsMIR_1236	NonTargetingControlGuideForMo use_0612	GACCGACGGTATACCT ACT	MouseDum my
MouseDummy_MetsMIR_1237_NonTargetingControlGuideFor Mouse_0613	MetsMIR_1237	NonTargetingControlGuideForMo use_0613	TTGAAGTAGGGTCGGAT TGA	MouseDum my
MouseDummy_MetsMIR_1238_NonTargetingControlGuideFor Mouse_0614	MetsMIR_1238	NonTargetingControlGuideForMo use_0614	AAGGCGTAAACGAGTAC ACG	MouseDum my
MouseDummy_MetsMIR_1239_NonTargetingControlGuideFor Mouse_0615	MetsMIR_1239	NonTargetingControlGuideForMo use_0615	AATGAGCTTCGAGTTCG TCT	MouseDum my
MouseDummy_MetsMIR_1240_NonTargetingControlGuideFor Mouse_0616	MetsMIR_1240	NonTargetingControlGuideForMo use_0616	ACGTGTTCTCGTACTTAG CT	MouseDum my
MouseDummy_MetsMIR_1241_NonTargetingControlGuideFor Mouse_0617	MetsMIR_1241	NonTargetingControlGuideForMo use_0617	TTAGCCTTGCCCCGTCA TAC	MouseDum my
MouseDummy_MetsMIR_1242_NonTargetingControlGuideFor Mouse_0618	MetsMIR_1242	NonTargetingControlGuideForMo use_0618	GCCTTTTCCGCCCGTTC AAG	MouseDum my
MouseDummy_MetsMIR_1243_NonTargetingControlGuideFor Mouse_0619	MetsMIR_1243	NonTargetingControlGuideForMo use_0619	CAGAACTCTTACCGAG CGC	MouseDum my
MouseDummy_MetsMIR_1244_NonTargetingControlGuideFor Mouse_0620	MetsMIR_1244	NonTargetingControlGuideForMo use_0620	CTTGAAAAAGGGGCGAC TAT	MouseDum my
MouseDummy_MetsMIR_1245_NonTargetingControlGuideFor Mouse_0621	MetsMIR_1245	NonTargetingControlGuideForMo use_0621	CCTACTAACGACGAGTC AAA	MouseDum my
MouseDummy_MetsMIR_1246_NonTargetingControlGuideFor Mouse_0622	MetsMIR_1246	NonTargetingControlGuideForMo use_0622	CATTTCGGGGTCCGAT GCA	MouseDum my
MouseDummy_MetsMIR_1247_NonTargetingControlGuideFor Mouse_0623	MetsMIR_1247	NonTargetingControlGuideForMo use_0623	CCGGGCGCCGTATCCCC TAC	MouseDum my
MouseDummy_MetsMIR_1248_NonTargetingControlGuideFor Mouse_0624	MetsMIR_1248	NonTargetingControlGuideForMo use_0624	GAGACCCATTATGATCC TAG	MouseDum my

Figure 18 (supplemental tables continued)

FIG. 18CF

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2015/000194

A. CLASSIFICATION OF SUBJECT MATTER INV. C12Q1/68 C12N15/63 C12N15/10 ADD.		
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) C12Q C12N		
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) EPO-Internal, BIOSIS, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2014/204723 A1 (BROAD INST INC [US]; MASSACHUSETTS INST TECHNOLOGY [US]; BRIGHAM WOMEN) 24 December 2014 (2014-12-24) pages 3-5; claim 2	1-73
X	WO 2014/204727 A1 (BROAD INST INC [US]; MASSACHUSETTS INST TECHNOLOGY [US]; ZHANG FENG [U] 24 December 2014 (2014-12-24)	50-73
Y	example 9 ----- -/-	1-49
☒ 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 18 April 2016		Date of mailing of the international search report 11/05/2016
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Aguilera, Miguel

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2015/000194

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	T. WANG ET AL: "Genetic Screens in Human Cells Using the CRISPR-Cas9 System", SCIENCE, vol. 343, no. 6166, 12 December 2013 (2013-12-12), pages 80-84, XP055115509, ISSN: 0036-8075, DOI: 10.1126/science.1246981	50-73
Y	the whole document	1-49
X	O. SHALEM ET AL: "Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells", SCIENCE, vol. 343, no. 6166, 12 December 2013 (2013-12-12), pages 84-87, XP055115506, ISSN: 0036-8075, DOI: 10.1126/science.1247005	50-73
Y	the whole document	1-49
X	PLATT RANDALL J ET AL: "CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling", CELL, CELL PRESS, US, vol. 159, no. 2, 25 September 2014 (2014-09-25), pages 440-455, XP029073412, ISSN: 0092-8674, DOI: 10.1016/J.CELL.2014.09.014	50-73
Y	the whole document	1-49
X	WEN XUE ET AL: "CRISPR-mediated direct mutation of cancer genes in the mouse liver", NATURE, vol. 514, no. 7522, 6 August 2014 (2014-08-06), pages 380-384, XP055176815, ISSN: 0028-0836, DOI: 10.1038/nature13589	50-73
Y	the whole document	1-49
X,P	CHEN SIDI ET AL: "Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis", CELL, vol. 160, no. 6, 12 March 2015 (2015-03-12), pages 1246-1260, XP029203797, ISSN: 0092-8674, DOI: 10.1016/J.CELL.2015.02.038	1-73
	the whole document	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2015/000194

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2014204723 A1	24-12-2014	EP 3011035 A1	27-04-2016
		WO 2014204723 A1	24-12-2014

WO 2014204727 A1	24-12-2014	EP 3011033 A1	27-04-2016
		WO 2014204727 A1	24-12-2014
