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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 June 2016 (30.06.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/105517 A1

(51) International Patent Classification:
C07K 16/40 (2006.01) *G01N 33/574* (2006.01)

(21) International Application Number:
PCT/US2015/000270

(22) International Filing Date:
23 December 2015 (23.12.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/096,311 23 December 2014 (23.12.2014) US

(71) Applicant: **THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK** [US/US]; 412 Low Memorial Library, 535 West 116th Street, New York, NY 10027 (US).

(72) Inventors: **LASORELLA, Anna**; 245 East 72nd Street, Apt. 14A, New York, NY 10021 (US). **IAVARONE, Antonio**; 245 East 72nd Street, Apt. 14A, New York, NY 10021 (US).

(74) Agents: **SUPERKO, Colleen** et al.; Wilmer Cutler Pickering Hale and Dorr LLP, 60 State Street, Boston, MA 02109 (US).

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

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

Published:

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

— *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))*

[Continued on next page]

(54) Title: FUSION PROTEINS AND METHODS THEREOF

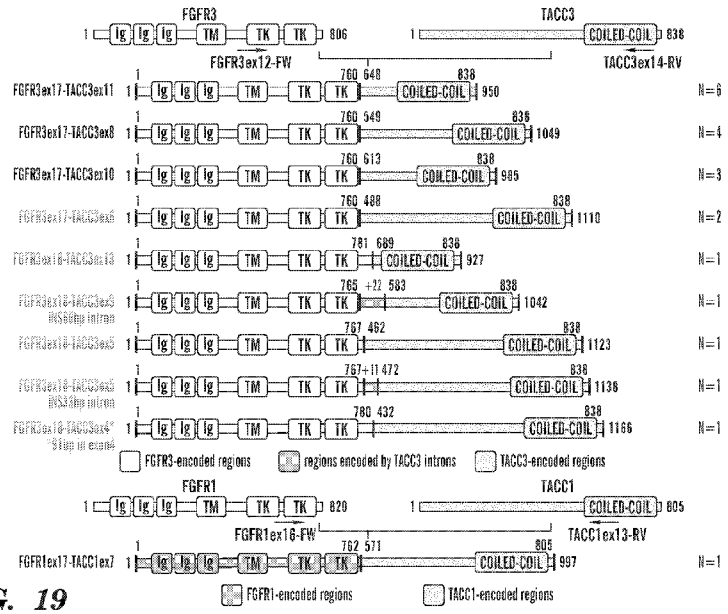


FIG. 19

(57) Abstract: The invention discloses oncogenic fusion proteins. The invention provides methods for treating and diagnosing gene-fusion based cancers.

WO 2016/105517 A1

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

FUSION PROTEINS AND METHODS THEREOF

This application claims the benefit of and priority to U.S. Provisional Patent Application No. 62/096311, filed on December 23, 2014, the content of which is hereby incorporated by reference in its entirety.

All patents, patent applications and publications cited herein are hereby incorporated by reference in their entirety. The disclosures of these publications in their entireties are hereby incorporated by reference into this application.

GOVERNMENT SUPPORT

This invention was made with government support under Grant Nos. R01CA101644, R01CA085628, R01CA131126, and R01CA178546 awarded by the National Cancer Institute and R01NS061776 awarded by the National Institute of Neurological Disorders and Stroke. The Government has certain rights in the invention.

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted electronically in ASCII format on compact disc and on paper as a separate part of the description and is hereby incorporated by reference in its entirety. Said ASCII copy, created on December 22, 2015, is named 19240.984WO2_SL.txt and is 470,083 bytes in size. The sequence listing forms part of the application. Sequences 1-494 are also described in WO2014/018673, the contents of which are hereby incorporated by reference in its entirety.

BACKGROUND OF THE INVENTION

Glioblastoma multiforme (GBM) is the most common form of brain cancer and among the most incurable and lethal of all human cancers. The current standard of care includes surgery, chemotherapy, and radiation therapy. However, the prognosis of GBM remains uniformly poor. There are few available targeted therapies and none that specifically target GBM.

The target population of GBM patients who may carry FGFR-TACC gene fusions and would benefit from targeted inhibition of FGFR kinase activity is estimated to correspond to 6,000 patients per year world-wide.

SUMMARY OF THE INVENTION

The invention is based, at least in part, on the discovery of a highly expressed class of gene fusions in GBM, which join the tyrosine kinase domain of FGFR genes to the TACC domain of TACC1 or TACC3. The invention is based, at least in part, on the finding that FGFR-TACC fusions identify a subset of GBM patients who will benefit from targeted inhibition of the tyrosine kinase activity of FGFR. Identification of fusions of FGFR and TACC genes in glioblastoma patients and other subjects afflicted with a gene-fusion associated cancer (such as an epithelial cancer) are useful therapeutic targets.

The invention is also based, at least in part, on the discovery of gene fusions joining the tyrosine kinase

domain of FGFR genes to the TACC domain of TACC1 or TACC3 in grade II and III glioma. The invention is based, at least in part, on the finding that FGFR-TACC fusions identify a subset of grade II and III glioma patients who will benefit from targeted inhibition of the tyrosine kinase activity of FGFR. Identification of fusions of FGFR and TACC genes in glioma patients are useful therapeutic targets.

An aspect of the invention provides for a purified fusion protein comprising a tyrosine kinase domain of an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein. In one embodiment, the FGFR protein is FGFR1, FGFR2, FGFR3, or FGR4. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein comprising a transforming acidic coiled-coil (TACC) domain fused to a polypeptide with a tyrosine kinase domain, wherein the TACC domain constitutively activates the tyrosine kinase domain. In one embodiment, the TACC protein is TACC1, TACC2, or TACC3. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein comprising the tyrosine kinase domain of an FGFR protein fused 5' to the TACC domain of a transforming acidic coiled-coil-containing (TACC) protein. In one embodiment, the FGFR protein is FGFR1, FGFR2, FGFR3, or FGR4. In another embodiment, the TACC protein is TACC1, TACC2, or TACC3. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein encoded by an FGFR1-TACC1 nucleic acid, wherein FGFR1-TACC1 comprises a combination of exons 1-17 of FGFR1 located on human chromosome 8p11 spliced 5' to a combination of exons 7-13 of TACC1 located on human chromosome 8p11, wherein a genomic breakpoint occurs in any one of exons 1-17 of FGFR1 and any one of exons 7-13 of TACC1. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein encoded by an FGFR2-TACC2 nucleic acid, wherein FGFR2-TACC2 comprises a combination of any exons 1-18 of FGFR2 located on human chromosome 10q26 spliced 5' to a combination of any exons 1-23 of TACC2 located on human chromosome 10q26. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein FGFR3-TACC3 comprises a combination of exons 1-16 of FGFR3 located on human chromosome 4p16 spliced 5' to a combination of exons 8-16 of TACC3 located on human chromosome 4p16, wherein a genomic breakpoint occurs in any one of exons 1-16 of FGFR3 and any one of exons 8-16 of TACC3. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein FGFR3-TACC3 comprises a combination of exons 1-18 of FGFR3 located on human chromosome 4p16 spliced 5' to a combination of exons 4-16 of TACC3 located on human chromosome 4p16, wherein a genomic breakpoint occurs in any one of exons 1-18 of FGFR3 and any one of exons 4-16

of TACC3. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein FGFR3-TACC3 comprises a combination of exons 1-16 of FGFR3 located on human chromosome 4p16 spliced 5' to a combination of exons 8-16 of TACC3 located on human chromosome 4p16, wherein a genomic breakpoint occurs in any one of introns 1-16 of FGFR3 and any one of exons 8-16 of TACC3. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein FGFR3-TACC3 comprises a combination of exons 1-18 of FGFR3 located on human chromosome 4p16 spliced 5' to a combination of exons 4-16 of TACC3 located on human chromosome 4p16, wherein a genomic breakpoint occurs in any one of introns 1-18 of FGFR3 and any one of exons 4-16 of TACC3. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein FGFR3-TACC3 comprises a combination of exons 1-16 of FGFR3 located on human chromosome 4p16 spliced 5' to a combination of exons 8-16 of TACC3 located on human chromosome 4p16, wherein a genomic breakpoint occurs in any one of exons 1-16 of FGFR3 and any one of introns 7-16 of TACC3. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein FGFR3-TACC3 comprises a combination of exons 1-18 of FGFR3 located on human chromosome 4p16 spliced 5' to a combination of exons 4-16 of TACC3 located on human chromosome 4p16, wherein a genomic breakpoint occurs in any one of exons 1-18 of FGFR3 and any one of introns 3-16 of TACC3. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein FGFR3-TACC3 comprises a combination of exons 1-16 of FGFR3 located on human chromosome 4p16 spliced 5' to a combination of exons 8-16 of TACC3 located on human chromosome 4p16, wherein a genomic breakpoint occurs in any one of introns 1-16 of FGFR3 and any one of introns 7-16 of TACC3. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein FGFR3-TACC3 comprises a combination of exons 1-18 of FGFR3 located on human chromosome 4p16 spliced 5' to a combination of exons 4-16 of TACC3 located on human chromosome 4p16, wherein a genomic breakpoint occurs in any one of introns 1-18 of FGFR3 and any one of introns 3-16 of TACC3. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a synthetic nucleic acid encoding the fusion proteins described above.

An aspect of the invention provides for a purified FGFR3-TACC3 fusion protein comprising SEQ ID NO: 79, 158, 159, 160, 161, 539, 540, 541, 542, 543, 544, 545, 546, 547. In another embodiment, the purified

fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified FGFR3-TACC3 fusion protein having a genomic breakpoint comprising at least 3 consecutive amino acids from amino acids 730-758 of SEQ ID NO: 90 and comprising at least 3 consecutive amino acids from amino acids 549-838 of SEQ ID NO: 92. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified FGFR3-TACC3 fusion protein having a genomic breakpoint comprising at least 3 consecutive amino acids from amino acids 730-781 of SEQ ID NO: 90 and comprising at least 3 consecutive amino acids from amino acids 432-838 of SEQ ID NO: 92. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified FGFR3-TACC3 fusion protein having a genomic breakpoint comprising SEQ ID NO: 78. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified FGFR3-TACC3 fusion protein having a genomic breakpoint comprising any one of SEQ ID NOS: 85, 86, 87, 89, 516, or 518. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified FGFR1-TACC1 fusion protein comprising SEQ ID NO: 150. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified FGFR1-TACC1 fusion protein having a genomic breakpoint comprising at least 3 consecutive amino acids from amino acids 746-762 of SEQ ID NO: 146 and comprising at least 3 consecutive amino acids from amino acids 572-590 of SEQ ID NO: 148. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified FGFR1-TACC1 fusion protein having a genomic breakpoint comprising at least 3 consecutive amino acids from amino acids 746-762 of SEQ ID NO: 146 and comprising at least 3 consecutive amino acids from amino acids 571-590 of SEQ ID NO: 148. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified FGFR1-TACC1 fusion protein having a genomic breakpoint comprising SEQ ID NO: 88. In another embodiment, the purified fusion protein is essentially free of other human proteins.

An aspect of the invention provides for a purified DNA encoding an FGFR3-TACC3 fusion protein comprising SEQ ID NO: 94, 530, 531, 532, 533, 534, 535, 536, 537, or 538. In another embodiment, the purified fusion protein is essentially free of other human proteins. An aspect of the invention provides for a purified cDNA encoding an FGFR3-TACC3 fusion protein comprising SEQ ID NO: 94, 530, 531, 532, 533, 534, 535, 536, 537, or 538.

An aspect of the invention provides for a synthetic nucleic acid encoding an FGFR3-TACC3 fusion protein having a genomic breakpoint comprising at least 9 consecutive in-frame nucleotides from nucleotides 2443-2530 of SEQ ID NO: 91 and comprising at least 9 consecutive in-frame nucleotides from nucleotides 1800-2847 of SEQ ID NO: 93.

An aspect of the invention provides for a synthetic nucleic acid encoding an FGFR3-TACC3 fusion protein having a genomic breakpoint comprising any one of SEQ ID NOS: 1-77, or 519-527.

An aspect of the invention provides for a synthetic nucleic acid encoding an FGFR1-TACC1 fusion protein comprising SEQ ID NO: 151.

An aspect of the invention provides for a synthetic nucleic acid encoding an FGFR1-TACC1 fusion protein having a genomic breakpoint comprising at least 9 consecutive in-frame nucleotides from nucleotides 3178-3228 of SEQ ID NO: 147 and comprising at least 9 consecutive in-frame nucleotides from nucleotides 2092-2794 of SEQ ID NO: 149.

An aspect of the invention provides for a synthetic nucleic acid encoding an FGFR1-TACC1 fusion protein having a genomic breakpoint comprising SEQ ID NO: 83.

An aspect of the invention provides for an antibody or antigen-binding fragment thereof, that specifically binds to a purified fusion protein comprising a tyrosine kinase domain of an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein. In one embodiment, the FGFR protein is FGFR1, FGFR2, FGFR3, or FGFR4. In another embodiment, the fusion protein is an FGFR-TACC fusion protein. In a further embodiment, the FGFR-TACC fusion protein is FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3. In some embodiments, the FGFR1-TACC1 fusion protein comprises the amino acid sequence of SEQ ID NO: 150. In other embodiments, the FGFR3-TACC3 fusion protein comprises the amino acid sequence of SEQ ID NO: 79, 158, 159, 160, 161, 539, 540, 541, 542, 543, 544, 545, 546, or 547.

An aspect of the invention provides for a composition for decreasing in a subject the expression level or activity of a fusion protein comprising the tyrosine kinase domain of an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein, the composition in an admixture of a pharmaceutically acceptable carrier comprising an inhibitor of the fusion protein. In one embodiment, the fusion protein is an FGFR-TACC fusion protein. In another embodiment, the inhibitor comprises an antibody that specifically binds to a FGFR-TACC fusion protein or a fragment thereof; a small molecule that specifically binds to a FGFR protein; a small molecule that specifically binds to a TACC protein; an antisense RNA or antisense DNA that decreases expression of a FGFR-TACC fusion polypeptide; a siRNA that specifically targets a FGFR-TACC fusion gene; or a combination of the listed inhibitors. In a further embodiment, the FGFR protein is FGFR1, FGFR2, FGFR3, or FGFR4. In some embodiments, the FGFR-TACC fusion protein is FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises AZD4547, NVP-BGJ398, PD173074, NF449, TK1258, BIBF-1120, BMS-582664, AZD-2171, TSU68, AB1010, AP24534, E-7080, LY2874455, or a combination of the listed small molecules. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises an oral pan-FGFR tyrosine kinase inhibitor. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises JNJ-42756493.

An aspect of the invention provides for a method for decreasing in a subject in need thereof the expression

level or activity of a fusion protein comprising the tyrosine kinase domain of an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein. In one embodiment, the method comprises administering to the subject a therapeutic amount of a composition for decreasing the expression level or activity in a subject of a fusion protein comprising the tyrosine kinase domain of an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein. In one embodiment, the method comprises obtaining a sample from the subject to determine the level of expression of an FGFR fusion molecule in the subject. In some embodiments, the sample is incubated with an agent that binds to an FGFR fusion molecule, such as an antibody, a probe, a nucleic acid primer, and the like. In one embodiment, the detection or determining comprises nucleic acid sequencing, selective hybridization, selective amplification, gene expression analysis, or a combination thereof. In another embodiment, the detection or determination comprises protein expression analysis, for example by western blot analysis, ELISA, immunostaining, or other antibody detection methods. In a further embodiment, the method comprises determining whether the fusion protein expression level or activity is decreased compared to fusion protein expression level or activity prior to administration of the composition, thereby decreasing the expression level or activity of the fusion protein. In one embodiment, the fusion protein is an FGFR-TACC fusion protein. In a further embodiment, the FGFR protein is FGFR1, FGFR2, FGFR3, or FGFR4. In some embodiments, the FGFR-TACC fusion protein is FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3. In one embodiment, the composition for decreasing the expression level or activity of a fusion protein comprises an antibody that specifically binds to a FGFR-TACC fusion protein or a fragment thereof; a small molecule that specifically binds to a FGFR protein; a small molecule that specifically binds to a TACC protein; an antisense RNA or antisense DNA that decreases expression of a FGFR-TACC fusion polypeptide; a siRNA that specifically targets a FGFR-TACC fusion gene; or a combination of the listed inhibitors. In a further embodiment, the FGFR protein is FGFR1, FGFR2, FGFR3, or FGFR4. In some embodiments, the FGFR-TACC fusion protein is FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises AZD4547, NVP-BGJ398, PD173074, NF449, TK1258, BIBF-1120, BMS-582664, AZD-2171, TSU68, AB1010, AP24534, E-7080, LY2874455, or a combination of the small molecules listed. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises an oral pan-FGFR tyrosine kinase inhibitor. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises JNJ-42756493.

An aspect of the invention provides for a method for treating a gene-fusion associated cancer in a subject in need thereof, the method comprising administering to the subject an effective amount of a FGFR fusion molecule inhibitor. In one embodiment, the gene-fusion associated cancer comprises an epithelial cancer. In one embodiment, the gene-fusion associated cancer comprises glioblastoma multiforme, breast cancer, lung cancer, prostate cancer, or colorectal carcinoma. In one embodiment, the gene-fusion associated cancer comprises bladder carcinoma, squamous lung carcinoma and head and neck carcinoma. In one embodiment, the gene-fusion associated cancer comprises glioma. In one embodiment, the gene-fusion associated cancer

comprises grade II or III glioma. In one embodiment, the gene-fusion associated cancer comprises IDH wild-type grade II or III glioma. In one embodiment, the method comprises obtaining a sample from the subject to determine the level of expression of an FGFR fusion molecule in the subject. In some embodiments the sample from the subject is a tissue sample. In some embodiments, the sample is a paraffin-embedded tissue section. In some embodiments, the tissue sample from the subject is a tumor sample. In some embodiments, the sample is incubated with an agent that binds to an FGFR fusion molecule, such as an antibody, a probe, a nucleic acid primer, and the like. In one embodiment, the detection or determining comprises nucleic acid sequencing, selective hybridization, selective amplification, gene expression analysis, or a combination thereof. In another embodiment, the detection or determination comprises protein expression analysis, for example by western blot analysis, ELISA, immunostaining, or other antibody detection methods. In another embodiment, the FGFR fusion protein comprises an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein. In one embodiment, the fusion protein is an FGFR-TACC fusion protein. In another embodiment, the inhibitor comprises an antibody that specifically binds to a FGFR-TACC fusion protein or a fragment thereof; a small molecule that specifically binds to a FGFR protein; a small molecule that specifically binds to a TACC protein; an antisense RNA or antisense DNA that decreases expression of a FGFR-TACC fusion polypeptide; a siRNA that specifically targets a FGFR-TACC fusion gene; or a combination of the listed inhibitors. In a further embodiment, the FGFR protein is FGFR1, FGFR2, FGFR3, or FGFR4. In some embodiments, the FGFR-TACC fusion protein is FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises AZD4547, NVP-BGJ398, PD173074, NF449, TK1258, BIBF-1120, BMS-582664, AZD-2171, TSU68, AB1010, AP24534, E-7080, LY2874455, or a combination of the small molecules listed. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises an oral pan-FGFR tyrosine kinase inhibitor. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises JNJ-42756493.

An aspect of the invention provides for a method of decreasing growth of a solid tumor in a subject in need thereof, the method comprising administering to the subject an effective amount of a FGFR fusion molecule inhibitor, wherein the inhibitor decreases the size of the solid tumor. In one embodiment, the solid tumor comprises glioblastoma multiforme, breast cancer, lung cancer, prostate cancer, or colorectal carcinoma. In one embodiment, the solid tumor comprises bladder carcinoma, squamous lung carcinoma and head and neck carcinoma. In one embodiment, the solid tumor comprises glioma. In one embodiment, the solid tumor comprises grade II or III glioma. In one embodiment, the solid tumor comprises IDH wild-type grade II or III glioma. In one embodiment, the method comprises obtaining a sample from the subject to determine the level of expression of an FGFR fusion molecule in the subject. In some embodiments the sample from the subject is a tissue sample. In some embodiments, the sample is a paraffin-embedded tissue section. In some embodiments, the tissue sample from the subject is a tumor sample. In some embodiments, the sample is incubated with an agent that binds to an FGFR fusion molecule, such as an antibody, a probe, a nucleic acid

primer, and the like. In one embodiment, the detection or determining comprises nucleic acid sequencing, selective hybridization, selective amplification, gene expression analysis, or a combination thereof. In another embodiment, the detection or determination comprises protein expression analysis, for example by western blot analysis, ELISA, immunostaining, or other antibody detection methods. In another embodiment, the FGFR fusion protein comprises an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein. In one embodiment, the fusion protein is an FGFR-TACC fusion protein. In another embodiment, the inhibitor comprises an antibody that specifically binds to a FGFR-TACC fusion protein or a fragment thereof; a small molecule that specifically binds to a FGFR protein; a small molecule that specifically binds to a TACC protein; an antisense RNA or antisense DNA that decreases expression of a FGFR-TACC fusion polypeptide; a siRNA that specifically targets a FGFR-TACC fusion gene; or a combination of the listed inhibitors. In a further embodiment, the FGFR protein is FGFR1, FGFR2, FGFR3, or FGFR4. In some embodiments, the FGFR-TACC fusion protein is FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises AZD4547, NVP-BGJ398, PD173074, NF449, TK1258, BIBF-1120, BMS-582664, AZD-2171, TSU68, AB1010, AP24534, E-7080, LY2874455, or a combination of the small molecules listed. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises an oral pan-FGFR tyrosine kinase inhibitor. In other embodiments, the small molecule that specifically binds to a FGFR protein comprises JNJ-42756493.

An aspect of the invention provides for a diagnostic kit for determining whether a sample from a subject exhibits a presence of a FGFR fusion, the kit comprising at least one oligonucleotide that specifically hybridizes to a FGFR fusion, or a portion thereof. In one embodiment, the oligonucleotides comprise a set of nucleic acid primers or *in situ* hybridization probes. In another embodiment, the oligonucleotide comprises SEQ ID NO: 162, 163, 164, 165, 166, 167, 168, 169, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514 or a combination of the listed oligonucleotides. In one embodiment, the primers prime a polymerase reaction only when a FGFR fusion is present. In another embodiment, the determining comprises gene sequencing, selective hybridization, selective amplification, gene expression analysis, or a combination thereof. In a further embodiment, the FGFR-fusion is an FGFR-TACC fusion. In some embodiments, the FGFR is FGFR1, FGFR2, FGFR3, or FGFR4. In other embodiments, the FGFR-TACC fusion is FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3.

An aspect of the invention provides for a diagnostic kit for determining whether a sample from a subject exhibits a presence of a FGFR fusion protein, the kit comprising an antibody that specifically binds to a FGFR fusion protein comprising SEQ ID NO: 79, 85, 86, 87, 88, 89, 150, 158, 159, 160, 161, 516, 518, 539, 540, 541, 542, 543, 544, 545, 546, or 547 wherein the antibody will recognize the protein only when a FGFR fusion protein is present. In one embodiment, the antibody directed to and FGFR fusion comprising SEQ ID NO: 79, 85, 86, 87, 88, 89, 150, 158, 159, 160, 161, 516, 518, 539, 540, 541, 542, 543, 544, 545, 546, or 547. In a further embodiment, the FGFR-fusion is an FGFR-TACC fusion. In some embodiments, the FGFR is FGFR1, FGFR2, FGFR3, or FGFR4. In other embodiments, the FGFR-TACC fusion is

FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3. In some embodiments the sample from the subject is a tissue sample. In some embodiments, the sample is a paraffin-embedded tissue section. In some embodiments, the tissue sample from the subject is a tumor sample.

An aspect of the invention provides for a method for detecting the presence of a FGFR fusion in a human subject. In one embodiment, the method comprises obtaining a biological sample from the human subject. In some embodiments the sample from the subject is a tissue sample. In some embodiments, the sample is a paraffin-embedded tissue section. In some embodiments, the tissue sample from the subject is a tumor sample. In some embodiments, the sample is incubated with an agent that binds to an FGFR fusion molecule, such as an antibody. In another embodiment, the detection or determination comprises protein expression analysis, for example by western blot analysis, ELISA, immunostaining or other antibody detection methods. In some embodiments, the method further comprises assessing whether to administer a FGFR fusion molecule inhibitor based on the expression pattern of the subject. In further embodiments, the method comprises administering a FGFR fusion molecule inhibitor to the subject. In other embodiments, the FGFR fusion molecule inhibitor comprises an oral pan-FGFR tyrosine kinase inhibitor. In other embodiments, the FGFR fusion molecule inhibitor comprises JNJ-42756493. In another embodiment, the method comprises detecting whether or not there is a FGFR fusion present in the subject. In one embodiment, the detecting comprises measuring FGFR fusion protein levels by ELISA using an antibody directed to SEQ ID NO: 79, 85, 86, 87, 88, 89, 150, 158, 159, 160, 161, 516, 518, 539, 540, 541, 542, 543, 544, 545, 546, or 547; western blot using an antibody directed to SEQ ID NO: 79, 85, 86, 87, 88, 89, 150, 158, 159, 160, 161, 516, 518, 539, 540, 541, 542, 543, 544, 545, 546, or 547; immunostaining using an antibody directed to SEQ ID NO: 79, 85, 86, 87, 88, 89, 150, 158, 159, 160, 161, 516, 518, 539, 540, 541, 542, 543, 544, 545, 546, or 547; mass spectroscopy, isoelectric focusing, or a combination of the listed methods. In some embodiments, the FGFR-fusion is an FGFR-TACC fusion. In other embodiments, the FGFR is FGFR1, FGFR2, FGFR3, or FGFR4. In other embodiments, the FGFR-TACC fusion is FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3.

An aspect of the invention provides for a method for detecting the presence of a FGFR fusion in a human subject. In one embodiment, the method comprises obtaining a biological sample from a human subject. In some embodiments, the sample is incubated with an agent that binds to an FGFR fusion molecule, such as a probe, a nucleic acid primer, and the like. In other embodiments, the detection or determination comprises nucleic acid sequencing, selective hybridization, selective amplification, gene expression analysis, or a combination thereof. In some embodiments, the method further comprises assessing whether to administer a FGFR fusion molecule inhibitor based on the expression pattern of the subject. In further embodiments, the method comprises administering a FGFR fusion molecule inhibitor to the subject. In another embodiment, the method comprises detecting whether or not there is a nucleic acid sequence encoding a FGFR fusion protein in the subject. In one embodiment, the nucleic acid sequence comprises any one of SEQ ID NOS: 1-77, 80-84, 95-145, 515, 517, 519-527, or 530-538. In another embodiment, the detecting comprises using hybridization, amplification, or sequencing techniques to detect a FGFR fusion. In a further embodiment, the

amplification uses primers comprising SEQ ID NO: 162, 163, 164,165, 166, 167, 168, 169, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, or 514. In some embodiments, the FGFR-fusion is an FGFR-TACC fusion. In other embodiments, the FGFR is FGFR1, FGFR2, FGFR3, or FGFR4. In other embodiments, the FGFR-TACC fusion is FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3.

An aspect of the invention provides for a method of initiating oncogenic transformation *in vitro*. The method comprises (a) transducing cells cultured *in vitro* with FGFR-TACC fusion DNA; and (b) determining whether the cells acquire the ability to grow in anchorage-independent conditions, form multi-layered foci, or a combination thereof.

An aspect of the invention provides for a method of initiating oncogenic transformation *in vivo*. The method comprises (a) transducing cells cultured *in vitro* with FGFR-TACC fusion DNA; (b) injecting a mouse with the transduced cells; and (c) determining whether a tumor grows in the mouse. In one embodiment, the injecting is a subcutaneous or intracranial injection.

An aspect of the invention provides a method of identifying a compound that decreases the oncogenic activity of a FGFR-TACC fusion. The method comprises (a) transducing a cell cultured *in vitro* with FGFR-TACC DNA; (b) contacting a cell with a ligand source for an effective period of time; and (c) determining whether the cells acquire the ability to grow in anchorage-independent conditions, form multi-layered foci, or a combination thereof, compared to cells cultured in the absence of the test compound.

In one embodiment, the method comprises contacting a sample from the subject with an antibody specific for a FGFR fusion molecule, and determining the presence of an immune complex. In another embodiment, the method comprises contacting a sample from the subject with an antibody specific for a FGFR molecule, or a TACC molecule, and determining the presence of an immune complex. In another embodiment, the antibody recognizes the FGFR3 C-terminal region, or the TACC3 N-terminal region, or a combination thereof. In another embodiment, the antibody recognizes the FGFR3 C-terminal region, or the TACC3 N-terminal region, or a combination thereof. In another embodiment, the method comprises contacting a sample from the subject with an antibody specific for a FGFR molecule, or a TACC molecule, or a FGFR fusion molecule, and determining the amount of an immune complex formed compared to the amount of immune complex formed in non-tumor cells or tissue, wherein an increased amount of an immune complex indicates the presence of an FGFR fusion.

In one embodiment, the method comprises contacting a sample from the subject with primers specific for a FGFR fusion molecule, and determining the presence of an PCR product. In another embodiment, the method comprises contacting a sample from the subject with primer specific for a FGFR molecule, or a TACC molecule, and determining the presence of a PCR product. In another embodiment, the primers recognizes the nucleic acids encoding a FGFR3 C-terminal region, or nucleic acids encoding a TACC3 N-terminal region, or a combination thereof. In another embodiment, the method comprises contacting a sample from the subject with primers specific for a FGFR molecule, or a TACC molecule, or a FGFR fusion molecule, and determining the amount of PCR product formed compared to the amount of PCR product

formed in non-tumor cells or tissue, wherein an increased amount of PCR product indicates the presence of an FGFR fusion.

An aspect of the invention provides for a purified fusion protein comprising the tyrosine kinase domain of an FGFR protein fused to the TACC domain of a transforming acidic coiled-coil-containing (TACC) protein. In one embodiment, the FGFR protein is FGFR1, FGFR2, FGFR3, or FGFR4. In one embodiment, the TACC protein is TACC1, TACC2, or TACC3. In one embodiment, the fusion protein is FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3. In one embodiment, the fusion protein comprises SEQ ID NO: 79, SEQ ID NO: 158, SEQ ID NO: 159, SEQ ID NO: 160, SEQ ID NO: 161, SEQ ID NO: 539, SEQ ID NO: 540, SEQ ID NO: 541, SEQ ID NO: 542, SEQ ID NO: 543, SEQ ID NO: 545, SEQ ID NO: 546, or SEQ ID NO: 547. In one embodiment, the fusion protein has a breakpoint comprising at least 3 consecutive amino acids from amino acids 730-758 of SEQ ID NO: 90 and comprising at least 3 consecutive amino acids from amino acids 549-838 of SEQ ID NO: 92. In one embodiment, the fusion protein has a breakpoint comprising SEQ ID NO: 78, SEQ ID NO: 85, SEQ ID NO: 86, SEQ ID NO: 87, SEQ ID NO:89, SEQ ID NO: 516, or SEQ ID NO:518. In one embodiment, the fusion protein comprises SEQ ID NO: 150. In one embodiment, the fusion protein has a breakpoint comprising at least 3 consecutive amino acids from amino acids 746-762 of SEQ ID NO: 146 and comprising at least 3 consecutive amino acids from amino acids 572-590 of SEQ ID NO: 148. In one embodiment, the fusion protein has a breakpoint comprising SEQ ID NO: 88.

An aspect of the invention provides for a cDNA encoding a fusion protein comprising the tyrosine kinase domain of FGFR fused to the TACC domain of TACC. In one embodiment the FGFR is FGFR1, FGFR2, FGFR3, or FGFR4. In one embodiment, the TACC is TACC1, TACC2, or TACC3. In one embodiment, the fusion protein is FGFR1-TACC1, FGFR2-TACC2, or FGFR3-TACC3. In one embodiment, the cDNA comprises SEQ ID NO: 94, SEQ ID NO: 530, SEQ ID NO: 531, SEQ ID NO: 532, SEQ ID NO: 533, SEQ ID NO: 534, SEQ ID NO: 535, SEQ ID NO: 536, SEQ ID NO: 537 or SEQ ID NO: 538. In one embodiment, the cDNA has a breakpoint comprising at least 9 consecutive in-frame nucleotides from nucleotides 2443-2530 of SEQ ID NO: 91 and comprising at least 9 consecutive in-frame nucleotides from nucleotides 1800-2847 of SEQ ID NO: 93. In one embodiment, the cDNA has a breakpoint comprising any one of SEQ ID NOs: 1-77, or SEQ ID NOs: 519-527. In one embodiment, the cDNA comprises SEQ ID NO: 151. In one embodiment, the cDNA has a breakpoint comprising at least 9 consecutive in-frame nucleotides from nucleotides 3178-3228 of SEQ ID NO: 147 and comprising at least 9 consecutive in-frame nucleotides from nucleotides 2092-2794 of SEQ ID NO: 149. In one embodiment, the cDNA has a breakpoint comprising SEQ ID NO: 83. In one embodiment, the cDNA comprises a combination of exons 1-16 of FGFR3 spliced 5' to a combination of exons 8-16 of TACC3, wherein a breakpoint occurs in: a) any one of exons 1-16 of FGFR3 and any one of exons 8-16 of TACC3; b) any one of introns 1-16 of FGFR3 and any one of exons 8-16 of TACC3; c) any one of exons 1-16 of FGFR3 and any one of introns 7-16 of TACC3; or d) any one of introns 1-16 of FGFR3 and any one of introns 7-16 of TACC3. In one embodiment, the cDNA comprises a combination of exons 1-17 of FGFR1 spliced 5' to a combination of

exons 7-13 of TACC1, wherein a breakpoint occurs in any one of exons 1-17 of FGFR3 and any one of exons 7-13 of TACC3. In one embodiment, the cDNA comprises a combination of exons 1-18 of FGFR2 spliced 5' to a combination of exons 1-23 of TACC2.

BRIEF DESCRIPTION OF THE FIGURES

To conform to the requirements for PCT patent applications, many of the figures presented herein are black and white representations of images originally created in color. In the below descriptions and the examples, the colored plots, and images are described in terms of its appearance in black and white. The original color versions can be viewed in Singh *et al., Science* (2012), 337(6099):1231-5 (including the accompanying Supplementary Information available in the on-line version of the manuscript available on the Science web site) and in Di Stefano *et al., Clin. Cancer Res.* (2015), 21(14):3307-17 (including the accompanying Supplementary Information available in the on-line version of the manuscript). For the purposes of the PCT, the contents of Singh *et al., Science* (2012), 337(6099):1231-5 and Di Stefano *et al., Clin. Cancer Res.* (2015), 21(14):3307-17including the accompanying "Supplementary Information," are herein incorporated by reference.

FIG. 1A is a graph that shows genes recurrently involved in gene fusions in TCGA. Only genes involved in at least three gene fusions across different samples are displayed. **FIG. 1B** shows an *FGFR3-TACC3* gene fusion identified by whole transcriptome sequencing of GSCs. On the left, *FGFR3-TACC3*-specific PCR from cDNA derived from GSCs and GBM is shown. On the right, Sanger sequencing chromatogram shows the reading frame at the breakpoint (SEQ ID NO: 80) and putative translation of the fusion protein (SEQ ID NO: 85) in the positive samples. **FIG. 1C** shows an *FGFR3-TACC3* gene fusion identified by whole transcriptome sequencing of GSCs. The fusion protein joins the tyrosine kinase domain of FGFR3 to the TACC domain of TACC3. **FIG. 1D** shows an *FGFR3-TACC3* gene fusion identified by whole transcriptome sequencing of GSCs. Genomic fusion of *FGFR3* exon 17 with intron 7 of *TACC3* is shown. In the fused mRNA, exon 16 of *FGFR3* is spliced 5' to exon 8 of *TACC3*. Filled arrows indicate the position of the fusion-genome primers, which generate fusion-specific PCR products in GSC-1123 and GBM-1123.

FIG. 2A and **FIG. 2B** show recurrent gene fusions between *FGFR* and *TACC* genes in GBM. On the left, a gel of *FGFR-TACC*-specific PCR is shown for *FGFR3-TACC3* from a GBM cDNA sample. On the right, Sanger sequencing chromatograms show the reading frame at the breakpoint (SEQ ID NO: 81 (FIG. 2B); SEQ ID NO: 82 (FIG. 2C) and putative translation of the fusion protein (SEQ ID NO: 86 (FIG. 2B); SEQ ID NO: 87 (FIG. 2D) in the positive samples. **FIG. 2C** shows recurrent gene fusions between *FGFR* and *TACC* genes in GBM. Co-outlier expression of *FGFR3* and *TACC3* in four GBM tumors from Atlas-TCGA is shown in the plot. **FIG. 2D** shows recurrent gene fusions between *FGFR* and *TACC* genes in GBM. CNV analysis shows micro-amplifications of the rearranged portions of the *FGFR3* and *TACC3* genes in the same four Atlas-TCGA GBM samples. **FIG. 2E-F** shows recurrent gene fusions between *FGFR* and *TACC* genes in GBM. On the left, a gel of *FGFR-TACC*-specific PCR is shown for *FGFR1-TACC1* (FIG. 2F) or *FGFR3-TACC3* (FIG. 2G) from a GBM cDNA sample. On the right, Sanger sequencing chromatograms

show the reading frame at the breakpoint (SEQ ID NO: 83 (FIG. 2F); SEQ ID NO: 84 (FIG. 2G) and putative translation of the fusion protein (SEQ ID NO: 88 (FIG. 2F); SEQ ID NO: 89 (FIG. 2G) in the positive samples.

FIG. 3A shows transforming activity of FGFR-TACC fusion proteins. FGFR1-TACC1 and FGFR3-TACC3 induce anchorage-independent growth in Rat1A fibroblasts. The number of soft agar colonies was scored from triplicate samples 14 days after plating. Representative microphotographs are shown.

FIG. 3B are photomicrographs showing of immunofluoresence staining of tumors from mice injected with *Ink4A;Arf-/-* astrocytes expressing FGFR3-TACC3 showing positivity for glioma-specific (Nestin, Oig2 and GFAP) and proliferation markers (Ki67 and pHH3). Sub-cutaneous tumors were generated by *Ink4A;Arf-/-* astrocytes expressing FGFR-TACC fusions. **FIG. 3C** shows Kaplan-Meier survival curves of mice injected intracranially with pTomo-shp53 (n = 8) or pTomo-EGFRvIII-shp53 (n = 7) (green line; "light grey" in black and white image) and pTomo-FGFR3-TACC3-shp53 (n = 8, red line; "dark grey" in black and white image). Points on the curves indicate deaths (log-rank test, p = 0.00001, pTomo-shp53 vs. pTomo-FGFR3-TACC3-shp53). **FIG. 3D** shows representative photomicrographs of Hematoxylin and Eosin staining of advanced FGFR3-TACC3-shp53 generated tumors showing histological features of high-grade glioma. Of note is the high degree of infiltration of the normal brain by the tumor cells. Immunofluorescence staining shows that glioma and stem cell markers (Nestin, Olig2 and GFAP), the proliferation markers (Ki67 and pHH3) and the FGFR3-TACC3 protein are widely expressed in the FGFR3-TACC3-shp53 brain tumors. F1-T1: FGFR1-TACC1; F3-T3: FGFR3-TACC3; F3-T3-K508M: FGFR3-TACC3-K508M.

FIG. 4A shows that FGFR3-TACC3 localizes to spindle poles, delays mitotic progression and induces chromosome segregation defects and aneuploidy Constitutive auto-phosphorylation of FGFR3-TACC3 fusion. *Ink4A;Arf-/-* astrocytes transduced with empty lentivirus or a lentivirus expressing FGFR3-TACC3 or FGFR3-TACC3-K508M were left untreated (0) or treated with 100 nM of the FGFR inhibitor PD173074 for the indicated times. Phospho-proteins and total proteins were analyzed by Western blot using the indicated antibodies. **FIG. 4B** shows quantitative analysis of segregation defects in Rat1A expressing FGFR1-TACC1 and FGFR3-TACC3. F3-T3: FGFR3-TACC3; F3-T3-K508M: FGFR3-TACC3-K508M (top panel). FGFR3-TACC3 protein induces chromosomal mis-segregation, chromatid cohesion defects and defective spindle checkpoint. Quantitative analysis of metaphase spreads for chromosome segregation defects in *Ink4A;ARF-/-* astrocytes expressing vector control or FGFR3-TACC3 is shown in the bottom panel.

FIG. 5A shows karyotype analysis of Rat1A cells transduced with control, FGFR3, TACC3 or FGFR3-TACC3 expressing lentivirus. Distribution of chromosome counts of cells arrested in mitosis and analyzed for karyotypes using DAPI. Chromosomes were counted in 100 metaphase cells for each condition to determine the ploidy and the diversity of chromosome counts within the cell population. FGFR3-TACC3 fusion induces aneuploidy. **FIG. 5B** shows representative karyotypes and **FIG. 5C** shows distribution of chromosome counts of human astrocytes transduced with control or FGFR3-TACC3 expressing lentivirus. Chromosomes were counted in 100 metaphase cells for each condition to determine the ploidy and the

diversity of chromosome counts within the cell population. **FIG. 5D** shows quantitative analysis of chromosome number in 100 metaphase cells for each condition to determine the ploidy and the diversity of chromosome counts within the cell population. (n=3 independent experiments).

FIG. 6A shows inhibition of FGFR-TK activity corrects the aneuploidy initiated by FGFR3-TACC3. The upper panel is a karyotype analysis of Rat1A cells transduced with control or FGFR3-TACC3 lentivirus and treated with vehicle (DMSO) or PD173470 (100 nM) for five days. The lower panel shows the ploidy and the diversity of chromosome counts within the cell population were determined by quantitative analysis of chromosome number in 100 metaphase cells for each condition. **FIG. 6B** shows inhibition of FGFR-TK activity corrects the aneuploidy initiated by FGFR3-TACC3. Correction of premature sister chromatid separation (PMSCS) by PD173470 in cells expressing FGFR3-TACC3. Panels show representative metaphase spreads. DNA was stained by DAPI. **FIG. 6C** shows quantitative analysis of metaphases with loss of sister chromatid cohesion (n=3; p = 0.001, FGFR3-TACC3 treated with DMSO vs. FGFR3-TACC3 treated with PD173470).

FIG. 7A shows inhibition of FGFR-TK activity suppresses tumor growth initiated by FGFR3-TACC3. Growth rate of Rat1A transduced with the indicated lentiviruses and treated for three days with increasing concentrations of PD173074. Cell growth was determined by the MTT assay. Data are presented as the means±standard error (n=4). **FIG. 7B** shows the growth rate of GSC-1123 treated with PD173470 at the indicated concentrations for the indicated times. Cell growth was determined by the MTT assay. Data are presented as the means±standard error (n=4). **FIG. 7C** shows the growth inhibitory effect of silencing FGFR3-TACC3 fusion. At the left, parallel cultures of GSC-1123 cells were transduced in triplicate. Rat1A cells expressing FGFR3-TACC3 fusion were transduced with lentivirus expressing a non-targeting shRNA (Ctr) or shRNA sequences targeting FGFR3 (sh2, sh3, sh4). Five days after infection cells were plated at density of $2X10^4$ cells/well in triplicate and the number of trypan blue excluding cells was scored at the indicated times. Infection with lentivirus expressing sh-3 and sh-4, the most efficient FGFR3 silencing sequences reverted the growth rate of FGFR3-TACC3 expressing cultures to levels comparable to those of Rat1A transduced with empty vector. Values are the means ± standard deviation (n = 3). At th right sided figure, GSC-1123 cells were transduced with lentivirus expressing a non-targeting shRNA (sh-Ctr) or lentivirus expressing sh-3 and sh-4 sequences targeting FGFR3. Western Blot analysis was performed on parallel cultures using the FGFR3 antibody to the detect FGFT3-TACC3 fusion protein. β-actin is shown as a control for loading. **FIG. 7D** shows that the FGFR inhibitor PD173074 suppresses tumor growth of glioma sub-cutaneous xenografts generated by *Ink4A;Arf-/-* astrocytes expressing FGFR3-TACC3. After tumor establishment (200-300 mm^3, arrow) mice were treated with vehicle or PD173074 (50 mg/kg) for 14 days. Values are mean tumor volumes ± standard error (n = 7 mice per group). **FIG. 7E** is a Kaplan-Meier analysis of glioma-bearing mice following orthotopic implantation of *Ink4A;Arf-/-* astrocytes transduced with FGFR3-TACC3. After tumor engraftment mice were treated with vehicle (n=9) or AZD4547 (50 mg/kg) (n=7) for 20 days (p = 0.001).

FIG. 8 shows a schematic of the TX-Fuse pipeline for the identification of fusion transcripts from RNA-Seq

data generated from nine GSC cultures. The continued figure shows a schematic of the Exome-Fuse pipeline for the identification of gene fusion rearrangements from DNA exome sequences of 84 GBM TCGA tumor samples.

FIGS. 9A-D shows the validation of fusion transcripts identified by RNA-seq of nine GSCs. Sanger sequencing chromatograms show the reading frames at the breakpoint and putative translation of the fusion proteins in the positive samples (right side). The left side shows gels of RT-PCR conducted. (**FIG. 9A**) POLR2A-WRAP53. DNA sequence disclosed as SEQ ID NO: 319 and protein sequence disclosed as SEQ ID NO: 320. (**FIG. 9B**) CAPZB-UBR4. DNA sequence disclosed as SEQ ID NO: 321 and protein sequence disclosed as SEQ ID NO: 322. (**FIG. 9C**) ST8SIA4-PAM. DNA sequence disclosed as SEQ ID NO: 323 and protein sequence disclosed as SEQ ID NO: 324. (**FIG. 9D**) PIGU-NCOA6. DNA sequence disclosed as SEQ ID NO: 325 and protein sequence disclosed as SEQ ID NO: 326. Also (below), 6 split-reads (SEQ ID NOS 485-490, respectively, in order of appearance) are shown aligning on the breakpoint of the IFNAR2-IL10RB fusion (SEQ ID NO: 483). The predicted reading frame at the breakpoint is shown at the top with IFNAR2 sequences in red (left) and IL10RB in blue (right). Protein sequence disclosed as SEQ ID NO: 484.

FIG. 10A shows the analysis and validation of the expression of fused transcripts in GSCs and GBM sample. Expression measured by read depth from RNA-seq data. Light grey arcs indicate predicted components of transcripts fused together. Overall read depth (blue; "grey" in black and white image) and split insert depth (red; "dark grey" in black and white image) are depicted in the graph, with a 50-read increment and a maximum range of 1800 reads. Note the very high level of expression in the regions of the genes implicated in the fusion events, particularly for FGFR3-TACC3. **FIG. 10B** shows the analysis and validation of the expression of fused transcripts in GSCs and GBM sample. Top panel, qRT-PCR showing the very high expression of FGFR3 and TACC3 mRNA sequences included in the FGFR3-TACC3 fusion transcript in GSC-1123. Bottom panel, for comparison the expression of sequences of WRAP53 mRNA included in the POLR2A-WRAP53 fusion in GSC-0114 is also shown. **FIG. 10C** shows the expression of the FGFR3-TACC3 protein in GSC-1123 and GBM-1123. Western blot analysis with a monoclonal antibody, which recognizes the N-terminal region of human FGFR3 shows expression of a ~150 kD protein in GSC-1123 but not in the GSC cultures GSC-0331 and GSC-0114, which lack the FGFR3-TACC3 rearrangement. **FIG. 10D** shows the analysis and validation of the expression of fused transcripts in GSCs and GBM sample. Immunostaining analysis with the FGFR3 antibody of the tumor GBM-1123 (top panel) and a GBM tumor lacking the FGFR3-TACC3 rearrangement. FGFR3 (red; "light grey" in black and white image), DNA (DAPI, blue; "grey" in black and white image). The pictures were taken at low (left) and high (right) magnification. **FIG. 10E1-E6** shows MS/MS analysis of the ~150 kD fusion protein immunoprecipitated by the monoclonal anti-FGFR3 antibody from GSC-1123, identifying three unique peptides mapping to the FGFR3 (FGFR3 Peptide 1 (SEQ ID NO: 492), 2 (SEQ ID NO: 493), and 3 (SEQ ID NO: 494)) and three peptides mapping to the C-terminal region of TACC3 (TACC Peptide 1 (SEQ ID NO: 156), 2 (SEQ ID NO: 157), and 3 (SEQ ID NO: 491)).

FIGS. 11A-C shows Rat1A cells transduced with control lentivirus or lentivurus expressing FGFR3, TACC3, FGFR3-TACC3 (**FIG. 11A**) that were analyzed by Western blot with an antibody recognizing the N-terminus of FGFR3 (included in the FGFR3-TACC3 fusion protein) or the N-terminus of TACC3 (not included in the FGFR3-TACC3 fusion protein). **FIG. 11B** shows quantitative Western blot analysis of endogenous FGFR3-TACC3 in GSC-1123 compared with lentivirally expressed FGFR3-TACC3 in Rat1A. **FIG. 11C** shows Western blot analysis of FGFR3-TACC3 and FGFR3-TACC3-K508M in Rat1A. α-tubulin is shown as a control for loading. **FIGS. 11D-F** shows expression analyses of FGFR3-TACC3 fusion construct (**FIG. 11D**) FGFR3 immunostaining of GBM-1123 (left, upper panel), BTSC1123 (right, upper panel), mouse GBM induced by FGFR3-TACC3 expressing lentivirus (left, lower panel), and sub-cutaneous xenograft of mouse astrocytes transformed by FGFR3-TACC3 fusion (right, lower panel); FGFR3-TACC3, red ("light grey" in black and white image); DNA (DAPI), blue ("grey" in black and white image). **FIG. 11E** shows quantification of FGFR3-TACC3 positive cells in the tumors and cultures of cells shown in **FIG. 11D**. **FIG. 11F** shows a quantitative Western blot analysis of ectopic FGFR3-TACC3 fusion protein in mouse astrocytes and FGFR3-TACC3 induced mouse GBM (mGBM-15 and mGBM-17) compared with the endogenous expression in GBM1123. β-actin is shown as a control for loading. F3-T3: FGFR3-TACC3. α-tubulin or β-actin is shown as a control for loading.

FIGS. 12A-B shows growth curves of human primary astrocytes transduced with lentivirus expressing FGFR3-TACC3 fusion or the empty vector. An analysis was conducted of FGFR3-TACC3 fusion mediated growth alteration and specific effect of RTK inhibitors on cells carrying FGFR-TACC fusions. **FIG. 12A** is a graph that shows cell proliferation of human primary astrocytes transduced with lentivirus expressing FGFR3-TACC3 fusion or the empty vector was determined by the MTT assay 7 days after infection (passage 1). Values are the means ± standard deviation (n = 4). *p*-value: 0.0033. **FIG. 12B** is a graph that shows cell proliferation of human primary astrocytes transduced with lentivirus expressing FGFR3-TACC3 fusion or the empty vector was determined by the MTT assay six weeks after the infection (passage 10). Values are the means±standard deviation (n=4). *p*-value: 0.0018.

FIGS. 12C-D shows specific growth inhibitory effect by FGFR inhibitors on FGFR-TACC fusion expressing cells. Cell growth was determined by MTT assay. Rat1A cells transduced with the indicated lentivirus were treated for three days with BGJ398 **(FIG. 12C)** or AZD4547 **(FIG. 12D)** at the indicated concentration. Values are the means ± standard error (n = 4). **FIG. 12E** shows the growth inhibitory effect of silencing FGFR3-TACC3 fusion. **(left)** GSC-1123 cells were transduced in triplicate with lentivirus expressing a non-targeting shRNA (Ctr) or lentivirus expressing sh-3 and sh-4 sequences targeting FGFR3. Five days after infection cells were plated at density of $2X10^4$ cells/well in triplicate and the number of trypan blue excluding cells was scored at the indicated times. Values are the means ± standard deviation (n = 3). **(right)** Western Blot analysis was performed on parallel cultures collected five days after infection using the FGFR3 antibody to the detect FGFT3-TACC3 fusion protein. β-actin is shown as a control for loading. (**: *p*-value = < 0.005; ***: *p*-value = < 0.0001).

FIG. 13 shows a survival plot of cells treated with PD173074 , NVP-BGJ398, or AZD4547.

FIG. 14 shows that inhibition of FGFR-TK activity corrects the aneuploidy and suppresses tumor growth initiated by FGFR3-TACC3. Short-term growth inhibition assays are shown of Rat1A transduced with the indicated lentivirus and treated with PD173470 at the indicated concentrations. Cells were treated for three days. Cell viability was determined by the MTT assay. Error bars show means ± standard error (n = 4).

FIG. 15 is a growth inhibition assay of human astrocytes transduced with the indicated lentivirus and treated for four days with PD173470 at the indicated concentration. Cell viability was determined by the MTT assay. Error bars show means ± standard error (n = 4).

FIG. 16 is a graph showing a growth inhibition assay of human astrocytes transduced with the indicated lentivirus and treated for four days with PD173470 at the indicated concentration. Cell viability was determined by the MTT assay. Error bars show means ± standard error (n = 4).

FIG. 17 shows Kaplan-Meier analysis of IDH mutant and *FGFR3-TACC3* positive human GBM. Log rank test *p*-value: 0.0169.

FIG. 18A-B is a picture that shows tumor xenografts that were induced following sub-cutaneous injection of *Ink4A;Arf-/-* mouse astrocytes transduced with lentivirus expressing FGFR3-TACC3 (upper panel A, right flank) or FGFR1-TACC1 (lower panel B, right flank) fusion, but not with the empty vector (upper panel, left flank) or FGFR3-TACC3 carrying a K508M mutation in the kinase domain (FGFR3-TACC3-K508M; lower panel, left flank).

FIG. 19 shows the structure of FGFR-TACC gene fusions identified by RT-PCR-Sanger sequencing (see also SEQ ID NOs: 530-547). Predicted FGFR-TACC fusion proteins encoded by the transcripts identified by RT-PCR. Regions corresponding to FGFR3 or TACC3 are shown in red or blue, respectively. FGFR1 and TACC1 corresponding regions are shown in yellow and green. On the left are indicated the FGFR and TACC exons joined in the fused mRNA; the presence of TACC3 introns is also reported when they are spliced in the fusion cDNA. On the right, the number of patients harboring the corresponding fusion variant is indicated. The novel transcripts discovered in this study are highlighted in red. Black arrows indicate the position of the primers used for the FGFR-TACC fusions screening.

FIGS. 20A-H show the identification and immunostaining of FGFR3-TACC3-positive tumors. Results from RT-PCR screening in representative samples from the Pitié-Salpêtrière Hospital (A, C) and the Besta (B, D) datasets. M, DNA ladder. Schematic representation of the FGFR3-TACC3 fusion transcripts identified in samples GBM-4620 (C) and GBM-021 (D). The junction sequences on the mRNA (GBM-4620 (C) SEQ ID NO: 515; GBM-021 (D) SEQ ID NO: 517) and the reading frame and translation (GBM-4620 (C) SEQ ID NO: 516; GBM-021 (D) SEQ ID NO: 518) at the breakpoint are reported. Representative microphotographs of H&E and FGFR3 immunostaining in the FGFR3-TACC3 positive samples GBM-4620 (E) and GBM-021 (F) and two FGFR3-TACC3 negative samples (panels G and H); a, H&E, 10X magnification; b, H&E, 40X magnification; c, FGFR3, 10X magnification; d, FGFR3, 40X magnification. Figures 20C and 20D disclose chromatogram readings as SEQ ID NOS 550 and 551, respectively.

FIGS. 21A-D show pre-clinical evaluation of FGFR3-TACC3 inhibition by JNJ-42756493. (A) Mouse astrocytes expressing FGFR3-TACC3 (F3T3), FGFR3-TACC3-KD (F3T3-KD) or the empty vector

(Vector) were treated with the indicated concentration of JNJ-42756493. Cell viability was determined by the MTT assay. Error bars show mean±SEM (n=6). (B) Survival analysis of GIC28 1123 treated with JNJ-42756493. (C) The FGFR-TK inhibitor JNJ-42756493 suppresses tumor growth of subcutaneous tumors generated by GIC-1123. After tumor establishment (arrow) mice were treated with vehicle or JNJ-42756493 (12 mg/kg) for 14 days. Values are mean tumor volumes±SD, (n=9 mice per group). P-value of the slope calculated from the treatment starting point (arrow) is 0.04. (D) Photograph showing the tumors dissected from vehicle or JNJ-42756493 treated mice after two weeks of treatment.

FIGS. 22A-G show baseline and post-treatment Magnetic Resonance Imaging (MRI) of patients treated with JNJ-42756493. Patient 1 (Panels A-D). (A) Post-gadolinium T1 weighted images show the target lesion on the right parietal lobe. The interval (days) from the beginning of follow-up is indicated above each MRI. (B) Analysis of sum of product diameters (SPD) before and during the anti-FGFR treatment (RANO criteria). (C) Analysis of tumor volume (cm3) before and during the anti-FGFR treatment. During anti-FGFR treatment a stabilization of the tumor was observed according to RANO criteria and volumetry. (D) Perfusion images at baseline and after 20 days of anti-FGFR treatment. rCBV (relative cerebral blood volume). Post-gadolinium T1 weighted images with color overlay of rCBV are shown. Patient 2 (Panels E-G). (E) Two different MRI slice levels of superior and middle part of the lesion are presented. (F) Analysis of sum of product diameters (SPD) before and during the anti-FGFR treatment. During the anti-FGFR treatment a reduction of 22% of tumor size was observed. (G) Volumetric evaluation showed a 28% tumor reduction. Vertical red arrow indicates the start of anti-FGFR treatment (baseline).

FIG. 23 shows the genomic PCR images and Sanger sequences of FGFR3-TACC3 genomic breakpoints. Fusion specific PCR products and Sanger sequencing chromatograms showing the FGFR3-TACC3 genomic breakpoints (Sample #4451 SEQ ID NO:519; Sample #OPK-14 SEQ ID NO: 520; Sample #MB-22 SEQ ID NO: 521; Sample #3048 SEQ ID NO: 522; Sample #4373 SEQ ID NO: 523; Sample #4867 SEQ ID NO: 524; Sample #3808 SEQ ID NO: 525; Sample #27-1835 SEQ ID NO: 526; Sample #06-6390 SEQ ID NO: 527). The genomic sequences corresponding to FGFR3 and TACC3 are indicated in red or blue, respectively. M, DNA ladder; C-, Negative Control. Figure 23 discloses chromatogram readings as SEQ ID NOS 552-560, respectively, in order of appearance.

FIG. 24 shows schematics of FGFR3-TACC3 genomic breakpoints. Schematic representation of the genomic fusions between FGFR3 and TACC3 compared to the corresponding mRNA. In red and blue are reported the regions belonging to FGFR3 and TACC3, respectively. The genomic breakpoint coordinates, according to the genome build GRCh37/hg19, are indicated above each fusion gene.

FIGS. 25A-B show evaluation of the expression of FGFR3-TACC3 fusion elements. (A) Microphotographs of immunofluorescence staining of a representative GBM harboring FGFR3- TACC3 fusion using antibodies that recognize the N- and C- termini of FGFR3 (FGFR3-N, FGFR3-C) and TACC3 (TACC3-N, TACC3-C), red. Nuclei are counterstained with DAPI, blue. (B) Quantitative RT-PCR of four representative GBM carrying FGFR3-TACC3 fusion and three negative controls using primer pairs that amplify FGFR3 and TACC3 regions included in or excluded from the fusion transcripts, as indicated in the diagram.

OAW28: ovarian cystoadenocarcinoma cell line harboring wild type FGFR3 and TACC3 genes; GBM55 and GBM0822: GBM harboring wild type FGFR3 and TACC3 genes; GBM3808; GBM1133; GBM0826; GBM3048: GBM harboring FGFR3-TACC3 (F3-T3) fusion. Error bars are SD of triplicate samples.

FIGS 26A-C show the FGFR3-TACC3 fusion gene and protein are retained in recurrent GBM. (FIG. 26A) FGFR3-TACC3 fusion specific RT-PCR product from untreated and recurrent GBM from patient #3124. (FIG. 26B) Sanger sequencing chromatogram showing the identical reading frame at the breakpoint (SEQ ID NO: 517) and the putative translation of the fusion protein (SEQ ID NO: 86) in the untreated and recurrent tumor from the same patient. The fused exons at mRNA level are shown. Regions corresponding to FGFR3 and TACC3 are indicated in red and blue, respectively. T, threonine; S, serine; D, aspartic acid; V, valine; K, lysine; A, alanine. (FIG. 26C) Representative microphotographs of FGFR3 immunofluorescence (IF) staining in both untreated and recurrent GBM. Blue staining, DAPI; Red staining, FGFR3. 10X Magnification. Figure 26B discloses chromatogram readings as SEQ ID NOS 551 and 551, respectively, in order of appearance

FIGS. 27A-B show PFS and OS of FGFR3-TACC3-positive glioma patients. (A) Kaplan-Meier curves in IDH wild-type glioma patients don't show significant differences in Progression Free Survival (PFS) between FGFR3-TACC3 positive (N=12, median PFS=11.20 months) and FGFR3-TACC3 negative (N=274, Median PFS= 12.27 months) (P=0.85). (B) Kaplan-Meier curves in IDH wild-type glioma patients don't show significant differences in Overall Survival (OS) between FGFR3-TACC3 positive (N=12, Median OS=32.80 months) and FGFR3-TACC3 negative (N=326, Median OS=18.60 months) (P=0.6). In red FGFR3-TACC3 positive patients, in green FGFR3-TACC3 negative patients. Open circles represent censored patients.

FIG. 28 shows analysis of SNP6.0 arrays of GBM harboring CNVs of FGFR3 and TACC3 genomic loci. CNVs of the FGFR3/TACC3 genomic loci in "gain labeled" (LRR > 0.2) TCGA samples. The CNA magnitudes (expressed as log2 ratio) were classified using simple thresholds: deletion ($x < -1$), loss ($-1 < x \leq -0.2$), gain ($0.2 \leq x < 1$) or amplification ($x > 1$). Gains are in gradient of red, loss in gradient of blue. Samples with uniform gains/amplification of FGFR3 and TACC3 lack FGFR3-TACC3 fusions. Samples harboring FGFR3-TACC3 fusions (F3-T3) show microamplifications involving the first FGFR3 exons, which are spliced in the fusion gene.

DETAILED DESCRIPTION OF THE INVENTION

Glioblastoma multiformes (GBMs) are the most common form of brain tumors in adults accounting for 12-15% of intracranial tumors and 50-60% of primary brain tumors. GBM is among the most lethal forms of human cancer. The history of successful targeted therapy of cancer largely coincides with the inactivation of recurrent and oncogenic gene fusions in hematological malignancies and recently in some types of epithelial cancer. GBM is among the most lethal and incurable forms of human cancer. Targeted therapies against common genetic alterations in GBM have not changed the dismal clinical outcome of the disease, most likely because they have systematically failed to eradicate the truly addicting oncoprotein activities of GBM. Recurrent chromosomal rearrangements resulting in the creation of oncogenic gene fusions have not been

found in GBM.

GBM is among the most difficult forms of cancer to treat in humans (*1*). So far, the therapeutic approaches that have been tested against potentially important oncogenic targets in GBM have met limited success (*2-4*). Recurrent chromosomal translocations leading to production of oncogenic fusion proteins are viewed as initiating and addicting events in the pathogenesis of human cancer, thus providing the most desirable molecular targets for cancer therapy (*5, 6*). Recurrent and oncogenic gene fusions have not been found in GBM. Chromosomal rearrangements are hallmarks of hematological malignancies but recently they have also been uncovered in subsets of solid tumors (breast, prostate, lung and colorectal carcinoma) (*7, 8*). Important and successful targeted therapeutic interventions for patients whose tumors carry these rearrangements have stemmed from the discovery of functional gene fusions, especially when the translocations involve kinase-coding genes (*BCR-ABL, EML4-ALK*) (*9, 10*).

A hallmark of GBM is rampant chromosomal instability (CIN), which leads to aneuploidy (*11*). CIN and aneuploidy are early events in the pathogenesis of cancer (*12*). It has been suggested that genetic alterations targeting mitotic fidelity might be responsible for missegregation of chromosomes during mitosis, resulting in aneuploidy (*13, 14*).

Fibroblast growth factor receptors (FGFR) are transmembrane receptors that bind to members of the fibroblast growth factor family of proteins. The structure of the FGFRs consist of an extracellular ligand binding domain comprised of three Ig-like domains, a single transmembrane helix domain, and an intracellular domain with tyrosine kinase activity (Johnson, D.E., Williams, E. T. Structural and functional diversity in the FGF receptor multigene family. (1993) *Adv. Cancer Res*, 60:1–41).

Transforming acidic coiled-coiled protein (TACC) stabilize microtubules during mitosis by recruiting minispindles (Msps)/XMAP215 proteins to centrosomes. TACCs have been implicated in cancer.

From a medical perspective, the FGFR-TACC fusions provide the first "bona-fide" oncogenically addictive gene fusions in GBM whose identification has long been overdue in this disease.

Beside GBM, which features the highest grade of malignancy among glioma (grade IV), lower grade glioma which include grade II and grade III are a heterogeneous group of tumors in which specific molecular features are associated with divergent clinical outcome. The majority of grade II-III glioma (but only a small subgroup of GBM) harbor mutations in IDH genes (IDH1 or IDH2), which confer a more favorable clinical outcome. Conversely, the absence of IDH mutations is associated with the worst prognosis (5).

Described herein is the identification of FGFR-TACC gene fusions (mostly FGFR3-TACC3, and rarely FGFR1-TACC1) as the first example of highly oncogenic and recurrent gene fusions in GBM. The FGFR-TACC fusions that have been identified so far include the Tyrosine Kinase (TK) domain of FGFR and the coiled-coil domain of TACC proteins, both necessary for the oncogenic function of FGFR-TACC fusions. FGFR3-TACC3 fusions have been identified in pediatric and adult glioma, bladder carcinoma, squamous lung carcinoma and head and neck carcinoma, thus establishing FGFR-TACC fusions as one of the chromosomal translocation most frequently found across multiple types of human cancers (6-15).

Here, a screening method for FGFR-TACC fusions is reported that includes a RT-PCR assay designed to

identify the known and novel FGFR3-TACC3 fusion transcripts, followed by confirmation of the inframe breakpoint by Sanger sequencing. Using this assay, a dataset of 584 GBM and 211 grade II and grade III gliomas has been analyzed. It was determined that brain tumors harboring FGFR-TACC fusions manifest strong and homogeneous intra-tumor expression of the FGFR3 and TACC3 component invariably included in the fusion protein, when analyzed by immunostaining. A significant clinical benefit following treatment with a specific inhibitor of FGFR-TK is reported in two GBM patients who harbored FGFR3-TACC3 rearrangement.

DNA and AminoAcid Manipulation Methods and Purification Thereof

The practice of aspects of the present invention can employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. *See*, e.g., Molecular Cloning A Laboratory Manual, 3rd Ed., ed. by Sambrook (2001), Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); DNA Cloning, Volumes I and II (D. N. Glover ed., 1985); Oligonucleotide Synthesis (M. J. Gait ed., 1984); Mullis et al. U.S. Pat. No: 4,683,195; Nucleic Acid Hybridization (B. D. Hames & S. J. Higgins eds. 1984); Transcription and Translation (B. D. Hames & S. J. Higgins eds. 1984); Culture Of Animal Cells (R. I. Freshney, Alan R. Liss, Inc., 1987); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide To Molecular Cloning (1984); the series, Methods In Enzymology (Academic Press, Inc., N.Y.), specifically, Methods In Enzymology, Vols. 154 and 155 (Wu et al. eds.); Gene Transfer Vectors For Mammalian Cells (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); Immunochemical Methods In Cell And Molecular Biology (Caner and Walker, eds., Academic Press, London, 1987); Handbook Of Experimental Immunology, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); Manipulating the Mouse Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). All patents, patent applications and references cited herein are incorporated by reference in their entireties.

One skilled in the art can obtain a protein in several ways, which include, but are not limited to, isolating the protein via biochemical means or expressing a nucleotide sequence encoding the protein of interest by genetic engineering methods.

A protein is encoded by a nucleic acid (including, for example, genomic DNA, complementary DNA (cDNA), synthetic DNA, as well as any form of corresponding RNA). For example, it can be encoded by a recombinant nucleic acid of a gene. The proteins of the invention can be obtained from various sources and can be produced according to various techniques known in the art. For example, a nucleic acid that encodes a protein can be obtained by screening DNA libraries, or by amplification from a natural source. A protein can be a fragment or portion thereof. The nucleic acids encoding a protein can be produced via recombinant DNA technology and such recombinant nucleic acids can be prepared by conventional techniques, including chemical synthesis, genetic engineering, enzymatic techniques, or a combination thereof. For example, a fusion protein of the invention comprises a tyrosine kinase domain of an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein. For example, a

fusion protein of the invention comprises a transforming acidic coiled-coil (TACC) domain fused to a polypeptide with a tyrosine kinase domain, wherein the TACC domain constitutively activates the tyrosine kinase domain. An example of a FGFR1-TACC1 polypeptide has the amino acid sequence shown in SEQ ID NO: 150. An example of a FGFR3-TACC3 protein is the polypeptide encoded by the nucleic acid having the nucleotide sequence shown in SEQ ID NOs: 94, 530, 531, 532, 533, 534, 535, 536, 537, or 538. Examples of a FGFR3-TACC3 polypeptide has the amino acid sequence shown in SEQ ID NO: 79, 158, 159, 160, 161, 539, 540, 541, 542, 543, 544, 545, 546, or 547.

The Genbank ID for the FGFR3 gene is 2261. Three isoforms are listed for FGFGR3, e.g., having Genebank Accession Nos. NP_000133 (corresponding nucleotide sequence NM_000142); NP_001156685 (corresponding nucleotide sequence NM_001163213); NP_075254 (corresponding nucleotide sequence NM_022965).

SEQ ID NO: 90 is the FGFR3 Amino Acid Sequence, Transcript Variant 1 (NP_000133; 806 aa). The location of exons are 1-37, 38-126, 127-149, 150-205, 206-247, 248-310, 311-359, 360-422, 423-471, 472-511, 512-549, 550-612, 613-653, 654-676, 677-723, 724-758, 759-7806. Amino acids encoded by nucleotides spanning exons are E37, D127, G149, E247, A359, R471, D512, G549, V677, and L723.

SEQ ID NO: 91 is the FGFR3 Nucleotide Sequence, Transcript Variant 1 (NM_000142; 4304 bp).

SEQ ID NO: 528 is the FGFR3 wt cDNA Nucleotide Sequence corresponding to the coding sequence of FGFR3 (2421 bp) (NM_000142.4 NP_000133.1). The location of exons are 110-379, 446-615, 740-930, 1076-1266, 1413-1534, 1646-1836, 1960-2030, and 2169-2274.

The Genbank ID for the TACC3 gene is 10460. SEQ ID NO: 92 is the TACC3 Amino Acid Sequence (NP_006333) (838 aa). The location of exons are 1-54, 55-101, 102-462, 463-487, 488-531, 532-548, 549-583, 584-612, 613-647, 648-672, 673-688, 689-741, 742-777, 778-817. Amino acids encoded by nucleotides spanning exons are N102, Q462, V531, S583, G673, E688, and L777. Double underlining indicates K112, the amino acid encoded by the nucleotides 334-336 of SEQ ID NO: 529.

SEQ ID NO: 93 is the TACC3 Nucleotide Sequence (NM_006342) (2847 bp).

SEQ ID NO: 529 is the TACC3 wt cDNA Nucleotide Sequence corresponding to the coding sequence of TACC3 (2517 bp) (NM_006342.2, NP_006333.1). The location of exons are 163- 305, 1386-1461, 1592-1644, 1749-1836, 1942-2018, 2063-2223, and 2331-2451.

SEQ ID NO: 94 is the nucleotide sequence of FGFR3-TACC3.

SEQ ID NO: 530 is the nucleotide sequence (cDNA) of FGFR3ex17-TACC3ex11. The sequence corresponding to FGFR3 is 1-2274. The sequence corresponding to TACC3 is 2275-2850.

SEQ ID NO: 531 is the nucleotide sequence (cDNA) of FGFR3ex17-TACC3ex8. The sequence corresponding to FGFR3 is 1-2274. The sequence corresponding to TACC3 is 2275-3147.

SEQ ID NO: 532 is the nucleotide sequence (cDNA) of FGFR3ex17-TACC3ex10. The sequence corresponding to FGFR3 is 1-2274. The sequence corresponding to TACC3 is 2275-2955.

SEQ ID NO: 533 is the nucleotide sequence (cDNA) of FGFR3ex17-TACC3ex6. The sequence corresponding to FGFR3 is 1-2274. The sequence corresponding to TACC3 is 2275-3330.

SEQ ID NO: 534 is the nucleotide sequence (cDNA) of FGFR3ex18-TACC3ex13. The sequence corresponding to FGFR3 is 1-2336. The sequence corresponding to TACC3 is 2337-2780.

SEQ ID NO: 535 is the nucleotide sequence (cDNA) of FGFR3ex18-TACC3ex9_INS66BP. The sequence corresponding to FGFR3 is 1-2296. The sequence corresponding to TACC3 is 2363-3135. The sequence corresponding the the 66bp intronic insert is 2297-2362.

SEQ ID NO: 536 is the nucleotide sequence (cDNA) of FGFR3ex18-TACC3ex5. The sequence corresponding to FGFR3 is 1-2295. The sequence corresponding to TACC3 is 2296-3396.

SEQ ID NO: 537 is the nucleotide sequence (cDNA) of FGFR3ex18-TACC3ex5_INS33bp. The sequence corresponding to FGFR3 is 1-2355. The sequence corresponding to TACC3 is 2389-3503. The sequence corresponding the the 33bp intronic insert is 2356-2388.

SEQ ID NO: 538 is the nucleotide sequence (cDNA) of FGFR3ex18-TACC3ex4. The sequence corresponding to FGFR3 is 1-2331. The sequence corresponding to TACC3 is 2332-3555.

The Genbank ID for the FGFR1 gene is 2260. Eight isoforms are listed for FGFR1, e.g., having Genebank Accession Nos. NP_001167534 (corresponding nucleotide sequence NM_001174063); NP_001167535 (corresponding nucleotide sequence NM_001174064); NP_001167536 (corresponding nucleotide sequence NM_001174065); NP_001167537 (corresponding nucleotide sequence NM_001174066); NP_001167538 (corresponding nucleotide sequence NM_001174067); NP_056934 (corresponding nucleotide sequence NM_015850); NP_075593 (corresponding nucleotide sequence NM_023105); NP_075594 (corresponding nucleotide sequence NM_023106); NP_075598 (corresponding nucleotide sequence NM_023110).

SEQ ID NO: 146 is the FGFR1 Amino Acid Sequence for isoform 10, having Genebank Accession No. NP_001167534 (820 aa).

SEQ ID NO: 147 is the FGFR1 Nucleotide Sequence for isoform 10, having Genebank Accession No. NM_001174063 (5895 bp). The start codon is 943-945.

SEQ ID NO: 185 is the FGFR1 Amino Acid Sequence for isoform 1, having Genebank Accession No. NP_075598 (822 aa).

SEQ ID NO: 186 is the FGFR1 Nucleotide Sequence for isoform 1, having Genebank Accession No. NM_023110 (5917 bp).

The Genbank ID for the TACC1 gene is 6867. Three isoforms are listed for TACC1, e.g., having Genebank Accession Nos. NP_006274 (corresponding nucleotide sequence NM_001174063); NP_001167535 (corresponding nucleotide sequence NM_001174064); NP_001167536 (corresponding nucleotide sequence NM_001174065).

SEQ ID NO: 148 is the TACC1 Amino Acid Sequence for isoform 1, having Genebank Accession No. NP_006274 (805 aa).

SEQ ID NO: 149 is the TACC1 Nucleotide Sequence for isoform 1, having Genebank Accession No. NM_006283 (7802 bp).

SEQ ID NO: 150 is the amino acid sequence of the FGFR1-TACC1 fusion protein.

SEQ ID NO: 151 is the nucleotide sequence that encodes the FGFR1-TACC1 fusion protein (FGFR1 – 1-

2286; TACC1 – 2287-2989.

The Genbank ID for the FGFR2 gene is 2263. Eight isoforms are listed for FGFR2, e.g., having Genebank Accession Nos. NP_000132 (corresponding nucleotide sequence NM_000141); NP_001138385 (corresponding nucleotide sequence NM_001144913); NP_001138386 (corresponding nucleotide sequence NM_001144914); NP_001138387 (corresponding nucleotide sequence NM_001144915); NP_001138388 (corresponding nucleotide sequence NM_001144916); NP_001138389 (corresponding nucleotide sequence NM_001144917); NP_001138390 (corresponding nucleotide sequence NM_001144918); NP_001138391 (corresponding nucleotide sequence NM_001144919); NP_075259 (corresponding nucleotide sequence NM_022970).

SEQ ID NO: 152 is the FGFR2 Amino Acid Sequence for isoform 1, having Genebank Accession No. NP_000132 (821 aa).

SEQ ID NO: 153 is the FGFR2 Nucleotide Sequence for isoform 1, having Genebank Accession No. NM_000141 (4654 bp).

The Genbank ID for the TACC2 gene is 10579. Four isoforms are listed for TACC2, e.g., having Genebank Accession Nos. NP_996744 (corresponding nucleotide sequence NM_206862); NP_996743 (corresponding nucleotide sequence NM_206861); NP_996742 (corresponding nucleotide sequence NM_206860); NP_008928 (corresponding nucleotide sequence NM_006997).

SEQ ID NO: 154 is the TACC2 Amino Acid Sequence for isoform a, having Genebank Accession No. NP_996744 (2948 aa).

SEQ ID NO: 155 is the TACC2 Nucleotide Sequence for isoform a, having Genebank Accession No. NM_206862 (9706 bp).

SEQ ID NO: 158 is the amino acid sequence of the FGFR3-TACC3-1 fusion protein. Position 1-758 corresponds to the FGFR3 protein. Position 759- 1048 corresponds to the TACC3 protein.

SEQ ID NO: 159 is the amino acid sequence of the FGFR3-TACC3-2 fusion protein. Position 1-758 corresponds to the FGFR3 protein. Position 759- 1027 corresponds to the TACC3 protein.

SEQ ID NO: 160 is the amino acid sequence of the FGFR3-TACC3-3 fusion protein. Position 1-758 corresponds to the FGFR3 protein. Position 759- 984 corresponds to the TACC3 protein.

SEQ ID NO: 161 is the amino acid sequence of the.FGFR3-TACC3-4 fusion protein. Position 1-758 corresponds to the FGFR3 protein. Position 759- 949 corresponds to the TACC3 protein.

SEQ ID NO: 539 is the amino acid sequence of FGFR3ex17-TACC3ex11. The sequence corresponding to FGFR3 is 1-758 (underlined). The sequence corresponding to TACC3 is 759-949 (shaded gray):

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCPPPGGGPMG
PTVWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCHFSVRVTDAPSSGDDEDGEDE
AEDTGVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAGNPTPSISWLKNGREFRGEHRIGGIKLRHQQ
WSLVMESVVPSDRGNYTCVVENKFGSIRQTYTLDVLERSPHRPILQAGLPANQTAVLGSDVEFHCKVYSD
AQPHIQWLKHVEVNGSKVGPDGTPYVTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCLAGNSIGFSHH
SAWLVVLPAEEELVEADEAGSVYAGILSYGVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLK
RQVSLESNASMSSNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGKPLGEGCFGQVVMAE

AIGIDKDRAAKPVTVAVKMLKDDATDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAK
GNLREFLRARRPPGLDYSFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVTEDNV
MKIADFGLARDVHNLDYYKKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPV
EELFKLLKEGHRMDKPANCTHDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDVKATQEENRELR
SRCEELHGKNLELGKIMDRFEEVVYQAMEEVQKQKELSKAEIQKVLKEKDQLTTDLNSMEKSFSDLFKRF
EKQKEVIEGYRKNEESLKKCVEDYLARITQEGQRYQALKAHAEEKLQLANEEIAQVRSKAQAEALALQAS
LRKEQMRIQSLEKTVEQKTKENEELTRICDDLISKMEKI

SEQ ID NO: 540 is the amino acid sequence of FGFR3ex17-TACC3ex8. The sequence corresponding to FGFR3 is 1-758 (underlined). The sequence corresponding to TACC3 is 759-1048 (shaded gray):

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCPPPGGGPMG
PTVWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCHFSVRVTDAPSSGDDEDGEDE
AEDTGVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAGNPTPSISWLKNGREFRGEHRIGGIKLRHQQ
WSLVMESVVPSDRGNYTCVVENKFGSIRQTYTLDVLERSPHRPILQAGLPANQTAVLGSDVEFHCKVYSD
AQPHIQWLKHVEVNGSKVGPDGTPYVTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCLAGNSIGFSHH
SAWLVVLPAEEELVEADEAGSVYAGILSYGVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLK
RQVSLESNASMSSNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGKPLGEGCFGQVVMAE
AIGIDKDRAAKPVTVAVKMLKDDATDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAK
GNLREFLRARRPPGLDYSFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVTEDNV
MKIADFGLARDVHNLDYYKKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPV
EELFKLLKEGHRMDKPANCTHDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDFKESALRKQSLY
LKFDPLLRDSPGRPVPVATETSSMHGANETPSGRPREAKLVEFDFLGALDIPVPGPPPGVPAPGGPPLST
GPIVDLLQYSQKDLDAVVKATQEENRELRSRCEELHGKNLELGKIMDRFEEVVYQAMEEVQKQKELSKAE
IQKVLKEKDQLTTDLNSMEKSFSDLFKRFEKQKEVIEGYRKNEESLKKCVEDYLARITQEGQRYQALKAH
AEEKLQLANEEIAQVRSKAQAEALALQASLRKEQMRIQSLEKTVEQKTKENEELTRICDDLISKMEKI

SEQ ID NO: 541 is the amino acid sequence of FGFR3ex17-TACC3ex10. The sequence corresponding to FGFR3 is 1-758 (underlined). The sequence corresponding to TACC3 is 759-984 (shaded gray):

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCPPPGGGPMGPT
VWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCHFSVRVTDAPSSGDDEDGEDEAEDT
GVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAGNPTPSISWLKNGREFRGEHRIGGIKLRHQQWSLVME
SVVPSDRGNYTCVVENKFGSIRQTYTLDVLERSPHRPILQAGLPANQTAVLGSDVEFHCKVYSDAQPHIQWL
KHVEVNGSKVGPDGTPYVTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCLAGNSIGFSHHSAWLVVLPAE
EELVEADEAGSVYAGILSYGVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLKRQVSLESNASMS
SNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGKPLGEGCFGQVVMAEAIGIDKDRAAKPVT
VAVKMLKDDATDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRPPGLD
YSFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVTEDNVMKIADFGLARDVHNLDYY
KKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMDKPANCT
HDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDVPGPPPGVPAPGGPPLSTGPIVDLLQYSQKDLDA
VVKATQEENRELRSRCEELHGKNLELGKIMDRFEEVVYQAMEEVQKQKELSKAEIQKVLKEKDQLTTDLNSM

EKSFSDLFKRFEKQKEVIEGYRKNEESLKKCVEDYLARITQEGQRYQALKAHAEEKLQLANEEIAQVRSKAQ
AEALALQASLRKEQMRIQSLEKTVEQKTKENEELTRICDDLISKMEKI

SEQ ID NO: 542 is the amino acid sequence of FGFR3ex17-TACC3ex6. The sequence corresponding to FGFR3 is 1-758 (underlined). The sequence corresponding to TACC3 is 759-1109 (shaded gray):

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCPPPGGGPMGPT
VWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCHFSVRVTDAPSSGDDEDGEDEAEDT
GVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAGNPTPSISWLKNGREFRGEHRIGGIKLRHQQWSLVME
SVVPSDRGNYTCVVENKFGSIRQTYTLDVLERSPHRPILQAGLPANQTAVLGSDVEFHCKVYSDAQPHIQWL
KHVEVNGSKVGPDGTPYVTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCLAGNSIGFSHHSAWLVVLPAE
EELVEADEAGSVYAGILSYGVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLKRQVSLESNASMS
SNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGKPLGEGCFGQVVMAEAIGIDKDRAAKPVT
VAVKMLKDDATDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRPPGLD
YSFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVTEDNVMKIADFGLARDVHNLDYY
KKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMDKPANCT
HDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDERALNSASTSLPTSCPGSEPVPTHQQGQPALELK
EESFRDPAEVLGTGAEVDYLEQFGTSSFKESALRKQSLYLKFDPLLRDSPGRPVPVATETSSMHGANETPSG
RPREAKLVEFDFLGALDIPVPGPPPGVPAPGGPPLSTGPIVDLLQYSQKDLDAVVKATQEENRELRSRCEEL
HGKNLELGKIMDRFEEVVYQAMEEVQKQKELSKAEIQKVLKEKDQLTTDLNSMEKSFSDLFKRFEKQKEVIE
GYRKNEESLKKCVEDYLARITQEGQRYQALKAHAEEKLQLANEEIAQVRSKAQAEALALQASLRKEQMRIQS
LEKTVEQKTKENEELTRICDDLISKMEKI

SEQ ID NO: 543 is the amino acid sequence of FGFR3ex18-TACC3ex13. The sequence corresponding to FGFR3 is 1-779 (underlined). The sequence corresponding to TACC3 is 780-926 (shaded gray):

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCPPPGGGPMGPT
VWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCHFSVRVTDAPSSGDDEDGEDEAEDT
GVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAGNPTPSISWLKNGREFRGEHRIGGIKLRHQQWSLVME
SVVPSDRGNYTCVVENKFGSIRQTYTLDVLERSPHRPILQAGLPANQTAVLGSDVEFHCKVYSDAQPHIQWL
KHVEVNGSKVGPDGTPYVTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCLAGNSIGFSHHSAWLVVLPAE
EELVEADEAGSVYAGILSYGVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLKRQVSLESNASMS
SNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGKPLGEGCFGQVVMAEAIGIDKDRAAKPVT
VAVKMLKDDATDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRPPGLD
YSFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVTEDNVMKIADFGLARDVHNLDYY
KKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMDKPANCT
HDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDEYLDLSAPFEQYSPGGQDTPSKQKELSKAEIQKV
LKEKDQLTTDLNSMEKSFSDLFKRFEKQKEVIEGYRKNEESLKKCVEDYLARITQEGQRYQALKAHAEEKLQ
LANEEIAQVRSKAQAEALALQASLRKEQMRIQSLEKTVEQKTKENEELTRICDDLISKMEKI

SEQ ID NO: 544 is the amino acid sequence of FGFR3ex18-TACC3ex9_INS66BP. The sequence corresponding to FGFR3 is 1-765 (underlined) . The sequence corresponding to TACC3 is 788-1043

(shaded gray). The sequence corresponding the the 66bp intronic insert is 766-787 (double-underlined):

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCPPPGGGPMGPT
VWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCHFSVRVTDAPSSGDDEDGEDEAEDT
GVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAGNPTPSISWLKNGREFRGEHRIGGIKLRHQQWSLVME
SVVPSDRGNYTCVVENKFGSIRQTYTLDVLERSPHRPILQAGLPANQTAVLGSDVEFHCKVYSDAQPHIQWL
KHVEVNGSKVGPDGTPYVTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCLAGNSIGFSHHSAWLVVLPAE
EELVEADEAGSVYAGILSYGVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLKRQVSLESNASMS
SNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGKPLGEGCFGQVVMAEAIGIDKDRAAKPVT
VAVKMLKDDATDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRPPGLD
YSFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVTEDNVMKIADFGLARDVHNLDYY
KKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMDKPANCT
HDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDEYLDLSAQEQRQPTLQPQGCCLAGYSHRSSMHGA
NETPSGRPREAKLVEFDFLGALDIPVPGPPPGVPAPGGPPLSTGPIVDLLQYSQKDLDAVVKATQEENRELR
SRCEELHGKNLELGKIMDRFEEVVYQAMEEVQKQKELSKAEIQKVLKEKDQLTTDLNSMEKSFSDLFKRFEK
QKEVIEGYRKNEESLKKCVEDYLARITQEGQRYQALKAHAEEKLQLANEEIAQVRSKAQAEALALQASLRKE
QMRIQSLEKTVEQKTKENEELTRICDDLISKMEKI

SEQ ID NO: 545 is the amino acid sequence of FGFR3ex18-TACC3ex5. The sequence corresponding to FGFR3 is 1-765 (underlined). The sequence corresponding to TACC3 is 766-1131 (shaded gray):

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCPPPGGGPMGPT
VWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCHFSVRVTDAPSSGDDEDGEDEAEDT
GVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAGNPTPSISWLKNGREFRGEHRIGGIKLRHQQWSLVME
SVVPSDRGNYTCVVENKFGSIRQTYTLDVLERSPHRPILQAGLPANQTAVLGSDVEFHCKVYSDAQPHIQWL
KHVEVNGSKVGPDGTPYVTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCLAGNSIGFSHHSAWLVVLPAE
EELVEADEAGSVYAGILSYGVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLKRQVSLESNASMS
SNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGKPLGEGCFGQVVMAEAIGIDKDRAAKPVT
VAVKMLKDDATDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRPPGLD
YSFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVTEDNVMKIADFGLARDVHNLDYY
KKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMDKPANCT
HDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDEYLDLSAPVVQLAAETPTAESKERALNSASTSLP
TSCPGSEPVPTHQQGQPALELKEESFRDPAEVLGTGAEVDYLEQFGTSSFKESALRKQSLYLKFDPLLRDSP
GRPVPVATETSSMHGANETPSGRPREAKLVEFDFLGALDIPVPGPPPGVPAPGGPPLSTGPIVDLLQYSQKD
LDAVVKATQEENRELRSRCEELHGKNLELGKIMDRFEEVVYQAMEEVQKQKELSKAEIQKVLKEKDQLTTDL
NSMEKSFSDLFKRFEKQKEVIEGYRKNEESLKKCVEDYLARITQEGQRYQALKAHAEEKLQLANEEIAQVRS
KAQAEALALQASLRKEQMRIQSLEKTVEQKTKENEELTRICDDLISKMEKI

SEQ ID NO: 546 is the amino acid sequence of FGFR3ex18-TACC3ex5_INS33bp. The sequence corresponding to FGFR3 is 1-784 (underlined). The sequence corresponding to TACC3 is 796-1166 (shaded gray). The sequence corresponding the the 33bp intronic insert is 785-795 (double underlined):

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCPPPGGGPMGPT
VWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCHFSVRVTDAPSSGDDEDGEDEAEDT
GVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAGNPTPSISWLKNGREFRGEHRIGGIKLRHQQWSLVME
SVVPSDRGNYTCVVENKFGSIRQTYTLDVLERSPHRPILQAGLPANQTAVLGSDVEFHCKVYSDAQPHIQWL
KHVEVNGSKVGPDGTPYVTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCLAGNSIGFSHHSAWLVVLPAE
EELVEADEAGSVYAGILSYGVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLKRQVSLESNASMS
SNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGKPLGEGCFGQVVMAEAIGIDKDRAAKPVT
VAVKMLKDDATDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRPPGLD
YSFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVTEDNVMKIADFGLARDVHNLDYY
KKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMDKPANCT
HDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDEYLDLSAPFEQYSPGGQDTPSSSSSGVREPPHPA
FPXSAEDTPVVQLAAETPTAESKERALNSASTSLPTSCPGSEPVPTHQQGQPALELKEESFRDPAEVLGTGA
EVDYLEQFGTSSFKESALRKQSLYLKFDPLLRDSPGRPVPVATETSSMHGANETPSGRPREAKLVEFDFLGA
LDIPVPGPPPGVPAPGGPPLSTGPIVDLLQYSQKDLDAVVKATQEENRELRSRCEELHGKNLELGKIMDRFE
EVVYQAMEEVQKQKELSKAEIQKVLKEKDQLTTDLNSMEKSFSDLFKRFEKQKEVIEGYRKNEESLKKCVED
YLARITQEGQRYQALKAHAEEKLQLANEEIAQVRSKAQAEALALQASLRKEQMRIQSLEKTVEQKTKENEEL
TRICDDLISKMEKI

SEQ ID NO: 547 is the amino acid sequence of FGFR3ex18-TACC3ex4. The sequence corresponding to FGFR3 is 1-777 (underlined). The sequence corresponding to TACC3 is 778-1184 (shaded gray).

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCPPPGGGPM
GPTVWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCHFSVRVTDAPSSGDDEDGE
DEAEDTGVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAGNPTPSISWLKNGREFRGEHRIGGIKLR
HQQWSLVMESVVPSDRGNYTCVVENKFGSIRQTYTLDVLERSPHRPILQAGLPANQTAVLGSDVEFHCK
VYSDAQPHIQWLKHVEVNGSKVGPDGTPYVTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCLAGNSI
GFSHHSAWLVVLPAEEELVEADEAGSVYAGILSYGVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKI
SRFPLKRQVSLESNASMSSNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGKPLGEGCF
GQVVMAEAIGIDKDRAAKPVTVAVKMLKDDATDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLY
VLVEYAAKGNLREFLRARRPPGLDYSFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAAR
NVLVTEDNVMKIADFGLARDVHNLDYYKKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIFTL
GGSPYPGIPVEELFKLLKEGHRMDKPANCTHDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDE
YLDLSAPFEQYSPGGQDTPESPETRLGQPAAEQLHAGPATEEPGPCLSQQLHSASAEDTPVVQLAAETP
TAESKERALNSASTSLPTSCPGSEPVPTHQQGQPALELKEESFRDPAEVLGTGAEVDYLEQFGTSSFKE
SALRKQSLYLKFDPLLRDSPGRPVPVATETSSMHGANETPSGRPREAKLVEFDFLGALDIPVPGPPPGV
PAPGGPPLSTGPIVDLLQYSQKDLDAVVKATQEENRELRSRCEELHGKNLELGKIMDRFEEVVYQAMEE
VQKQKELSKAEIQKVLKEKDQLTTDLNSMEKSFSDLFKRFEKQKEVIEGYRKNEESLKKCVEDYLARIT
QEGQRYQALKAHAEEKLQLANEEIAQVRSKAQAEALALQASLRKEQMRIQSLEKTVEQKTKENEELTRI
CDDLISKMEKI

The Genbank ID for the FGFR4 gene is 2264. Three isoforms are listed for FGFR4, e.g., having Genebank Accession Nos. NP_002002 (corresponding nucleotide sequence NM_002011); NP_075252 (corresponding nucleotide sequence NM_022963); NP_998812 (corresponding nucleotide sequence NM_213647).

As used herein, a "FGFR fusion molecule" can be a nucleic acid (e.g., synthetic, purified, and/or recombinant) which encodes a polypeptide corresponding to a fusion protein comprising a tyrosine kinase domain of an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein, or a nucleic acid encoding a fusion protein comprising a transforming acidic coiled-coil (TACC) domain fused to a polypeptide with a tyrosine kinase domain, wherein the TACC domain constitutively activates the tyrosine kinase domain. It can also be a fusion protein comprising a tyrosine kinase domain of an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein, or a fusion protein comprising a transforming acidic coiled-coil (TACC) domain fused to a polypeptide with a tyrosine kinase domain, wherein the TACC domain constitutively activates the tyrosine kinase domain. For example, a FGFR fusion molecule can include a FGFR1-TACC1 (e.g., comprising the amino acid sequence shown in SEQ ID NO: 150, or comprising the nucleic acid sequence shown in SEQ ID NO: 88), FGFR2-TACC2, FGFR3-TACC3 (e.g., comprising the amino acid sequence shown in SEQ ID NOS: 79, 158-161, or 539-547 or comprising the nucleic acid sequence shown in SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538), or other FGFR-TACC fusion proteins (e.g., an N-terminal fragment of FGFR4 containing its tyrosine kinase domain fused to a fragment of TACC1, TACC2, or TACC3). For example, a FGFR fusion molecule can include a FGFR1-containing fusion comprising the amino acid sequence corresponding to Genebank Accession no. NP_001167534, NP_001167535, NP_001167536, NP_001167537, NP_001167538, NP_056934, NP_075593, NP_075594, or NP_075598; or a FGFR1-containing fusion comprising the nucleotide sequence corresponding to Genebank Accession no. NM_001174063, NM_001174064, NM_001174065, NM_001174066, NM_001174067, NM_015850, NM_023105, NM_023106, or NM_023110. For example, a FGFR fusion molecule can include a FGFR2-containing fusion comprising the amino acid sequence corresponding to Genebank Accession no. NP_000132, NP_001138385, NP_001138386, NP_001138387, NP_001138388, NP_001138389, NP_001138390, NP_001138391, or NP_075259; or a FGFR2-containing fusion comprising the nucleotide sequence corresponding to Genebank Accession no. NM_000141, NM_001144913, NM_001144914, NM_001144915, NM_001144916, NM_001144917, NM_001144918, NM_001144919, or NM_022970. For example, a FGFR fusion molecule can include a FGFR3-containing fusion comprising the amino acid sequence corresponding to Genebank Accession no. NP_000133, NP_001156685, or NP_075254; or a FGFR3-containing fusion comprising the nucleotide sequence corresponding to Genebank Accession no. NM_000142, NM_001163213, or NM_022965. For example, a FGFR fusion molecule can include a FGFR4-containing fusion comprising the amino acid sequence corresponding to Genebank Accession no. NP_002002, NP_075252, or NP_998812; or a FGFR4-containing fusion comprising the nucleotide sequence corresponding to Genebank Accession no. NM_002011, NM_022963, or NM_213647. A FGFR fusion molecule can also include a tyrosine kinase domain of an FGFR protein fused to a protein

encoded by any one of the genes listed in Table 7. A FGFR fusion molecule can include a variant of the above described examples, such as a fragment thereof. **Table 7**. Fusion Partners

gene	gene	gene	gene
ABCA13	C21orf29	CAMKK1	DNAJC6
ABCC1	CACNA1C	CAMSAP1	DYRK3
ABCC12	CACNA1G	CAMTA1	EIF2C2
ABCC6	CNTNAP4	CAP2	FAM184B
ABL1	CUL3	CCDC147	FREM2
ADAM12	DMD	CCDC158	GDPD2
ADCY10	DUSP27	CELF2	GLI3
ADCY2	ECE1	CILP	IL1RN
ADCY8	EYS	CMYA5	ISX
AGBL4	FAM172A	COL14A1	KIDINS220
AHNAK	FAM184B	CORO7	LRBA
ANXA7	FGFR4	CSMD2	LY75
AP4S1	ITGAV	CUL3	MDH2
AQP2	LRP1	DDI2	MMP12
ARMC6	LY75	DEPDC5	N4BP2L2
ATP5B	MAPKAP1	DEPDC7	NCF2
ATP6AP1L	MYT1	DI10L	NCOR1
ATP6V0D2	NCF2	DMD	NCRNA00157
ATXN1	NCOR1	EDA	NRXN3
BAHD1	NHSL2	EFHC1	PARP16
BBX	NKAIN2	EFS	PLA2G2F
BCA10	NR3C1	EIF2C2	PLEK2
C15orf23	NUP188	ENTPD2	PRKCH
C15orf33	OSBPL10	EYS	PTPRS
C21orf29	PACSIN1	FAM160A1	ROBO1
C2CD3	PARP16	MUSK	SASH3
C6orf170	PDZRN4	NEUROG1	SH3BP5
C7orf44	POLM	NHSL2	SLC44A2
CACNA1C	PPP1R3A	NR3C1	SLC5A4
CACNA1G	PSEN1	ODZ1	SNX5
FAM168A	PTPRD	PCDH12	SORCS2
FAM172A	PTPRS	PLCL1	SRRM1
FAM192A	RALYL	PLEKHM3	SSX3
FAM19A2	RERE	PLOD3	STAG2
FBXL4	RIMBP2	PRKCH	STK24
FH	RNF216	PSEN1	SURF6
FREM2	SDAD1	SEPT5	SYNPO2
GAPVD1	SEC14L3	SLC44A2	TAF1
GLI3	SH3RF3	SNTA1	TMEM80
GPR182	SLC9A1	USP48	TNFRSF10B
GSTA3	SMOC2	VSNL1	TTYH1
IGFBP3	SNX5	WDFY1	UNC93B1
ITGA9	TACC2	WISP2	VSNL1
ITGB2	SRGAP1	XRRA1	XRCC4
JOSD2	SSX3	LRRC4B	ZNF410
KIDINS220	SUMF1	LRRK2	TRIOBP
LAMA2	SYNPO2	MAPKAP1	TTYH1
LCLAT1	TNFRSF10B	MST1R	LRBA
LIN9			

For example, a FGFR fusion molecule can include a FGFR3-containing fusion comprising the amino acid sequence corresponding to residues 1-760 of FGFR3 (e.g. SEQ ID NO: 90) fused to the amino acid sequence corresponding residues 648-838 of TACC3 (e.g. SEQ ID NO: 92). A FGFR fusion molecule can also include a FGFR3-containing fusion comprising the amino acid sequence corresponding to residues 1-760 of FGFR3 (e.g. SEQ ID NO: 90) fused to the amino acid sequence corresponding residues 549-838 of TACC3 (e.g. SEQ ID NO: 92). A FGFR fusion molecule can include a FGFR3-containing fusion comprising the amino acid sequence corresponding to residues 1-760 of FGFR3 (e.g. SEQ ID NO: 90) fused to the amino acid sequence corresponding residues 613-838 of TACC3 (e.g. SEQ ID NO: 92). A FGFR fusion molecule can include a FGFR3-containing fusion comprising the amino acid sequence corresponding to residues 1-760 of FGFR3 (e.g. SEQ ID NO: 90) fused to the amino acid sequence corresponding residues 488-838 of TACC3 (e.g. SEQ ID NO: 92). A FGFR fusion molecule can include a FGFR3-containing fusion comprising the amino acid sequence corresponding to residues 1-781 of FGFR3 (e.g. SEQ ID NO: 90) fused to the amino acid sequence corresponding residues 689-838 of TACC3 (e.g. SEQ ID NO: 92). A FGFR fusion molecule can include a FGFR3-containing fusion comprising the amino acid sequence corresponding to residues 1-765 of FGFR3 (e.g. SEQ ID NO: 90) fused to the amino acid sequence corresponding residues 583-838 of TACC3 (e.g. SEQ ID NO: 92). A FGFR fusion molecule can include a FGFR3-containing fusion comprising the amino acid sequence corresponding to residues 1-767 of FGFR3 (e.g. SEQ ID NO: 90) fused to the amino acid sequence corresponding residues 462-838 of TACC3 (e.g. SEQ ID NO: 92). A FGFR fusion molecule can include a FGFR3-containing fusion comprising the amino acid sequence corresponding to residues 1-767 of FGFR3 (e.g. SEQ ID NO: 90) fused to the amino acid sequence corresponding residues 472-838 of TACC3 (e.g. SEQ ID NO: 92). A FGFR fusion molecule can include a FGFR3-containing fusion comprising the amino acid sequence corresponding to residues 1-780 of FGFR3 (e.g. SEQ ID NO: 90) fused to the amino acid sequence corresponding residues 432-838 of TACC3 (e.g. SEQ ID NO: 92). A FGFR fusion molecule can include a FGFR1-containing fusion comprising the amino acid sequence corresponding to residues 1-762 of FGFR1 (e.g. SEQ ID NOS: 146, 185) fused to the amino acid sequence corresponding residues 571-805 of TACC1 (e.g. SEQ ID NO: 148). For example, a FGFR fusion molecule can include SEQ ID NOs: 539-543, 545, and 547.

The alteration in a chromosome region occupied by a FGFR fusion molecule, e.g., a FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid, can result in amino acid substitutions, RNA splicing or processing, product instability, the creation of stop codons, production of oncogenic fusion proteins, frame-shift mutations, and/or truncated polypeptide production. A FGFR fusion molecule can include FGFR and TACC exons joined in the fused mRNA or cDNA. A FGFR fusion molecule can also include FGFR and TACC exons joined in the fused mRNA or cDNA along with the presence of FGFR or TACC introns that are spliced in the fusion cDNA. FGFR or TACC introns can encode amino acids of the FGFR fusion molecule. For example, a FGFR fusion molecule can include a FGFR3-containing fusion comprising the amino acid sequence corresponding to residues 1-765 of FGFR3 (e.g. SEQ ID NO: 90) fused

to a 22 amino acid sequence encoded by a TACC3 intron fused to the amino acid sequence corresponding to residues 583-838 of TACC3 (e.g. SEQ ID NO: 92). For example, a FGFR fusion molecule can include a FGFR3-containing fusion comprising the amino acid sequence corresponding to residues 1-767 of FGFR3 (e.g. SEQ ID NO: 90) fused to a 11 amino acid sequence encoded by a TACC3 intron fused to the amino acid sequence corresponding to residues 472-838 of TACC3 (e.g. SEQ ID NO: 92). For example, a FGFR fusion molecule can include SEQ ID NOs: 544 and 546.

A FGFR fusion protein can also include a fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein FGFR3-TACC3 comprises a combination of introns and exons 1-18 of FGFR3 located on human chromosome 4p16 spliced 5' to a combination of introns and exons 4-16 of TACC3 located on human chromosome 4p16, wherein a genomic breakpoint occurs in any one of introns or exons 1-18 of FGFR3 and any one of introns or exons 3-16 of TACC3. For example, a FGFR fusion protein can also include a fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein the nucleic acid comprises exons 1-17 of FGFR3 located on human chromosome 4p16 spliced 5' to exons 11-16 of TACC3 located on human chromosome 4p16. In one embodiment, a genomic breakpoint occurs in intron 17 of FGFR3 and in intron 10 of TACC3. For example, a FGFR fusion protein can also include a fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein the nucleic acid comprises exons 1-17 of FGFR3 located on human chromosome 4p16 spliced 5' to exons 8-16 of TACC3 located on human chromosome 4p16. In one embodiment, a genomic breakpoint occurs in intron 17 of FGFR3 and in intron 7 of TACC3. For example, a FGFR fusion protein can also include a fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein the nucleic acid comprises exons 1-17 of FGFR3 located on human chromosome 4p16 spliced 5' to exons 10-16 of TACC3 located on human chromosome 4p16. In one embodiment, a genomic breakpoint occurs in exon 18 of FGFR3 and in intron 9 of TACC3. For example, a FGFR fusion protein can also include a fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein the nucleic acid comprises exons 1-17 of FGFR3 located on human chromosome 4p16 spliced 5' to exons 6-16 of TACC3 located on human chromosome 4p16. In one embodiment, a genomic breakpoint occurs in intron 17 of FGFR3 and in intron 5 of TACC3. In one embodiment, a genomic breakpoint occurs in intron 17 of FGFR3 and in exon 5 of TACC3. For example, a FGFR fusion protein can also include a fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein the nucleic acid comprises exons 1-18 of FGFR3 located on human chromosome 4p16 spliced 5' to exons 13-16 of TACC3 located on human chromosome 4p16. In one embodiment, a genomic breakpoint occurs in exon 18 of FGFR3 and in intron 12 or exon 13 of TACC3. For example, a FGFR fusion protein can also include a fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein the nucleic acid comprises exons 1-18 of FGFR3 located on human chromosome 4p16 spliced 5' to a portion of intron 8 of TACC3 and exons 9-16 of TACC3 located on human chromosome 4p16. In one embodiment, a genomic breakpoint occurs in exon 18 of FGFR3 and in intron 8 of TACC3. For example, a FGFR fusion protein can also include a fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein the nucleic acid comprises exons 1-18 of FGFR3 located on human chromosome 4p16 spliced 5' to exons 5-16 of TACC3 located on human chromosome 4p16. In one embodiment, a genomic

breakpoint occurs in exon 18 of FGFR3 and in intron 4 or exon 5 of TACC3. For example, a FGFR fusion protein can also include a fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein the nucleic acid comprises exons 1-18 of FGFR3 located on human chromosome 4p16 spliced 5' to a portion of intron 4 of TACC 3 and exons 5-16 of TACC3 located on human chromosome 4p16. In one embodiment, a genomic breakpoint occurs in exon 18 of FGFR3 and in intron 4 or exon 5 of TACC3. For example, a FGFR fusion protein can also include a fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein the nucleic acid comprises exons 1-18 of FGFR3 located on human chromosome 4p16 spliced 5' to exons 4-16 of TACC3 located on human chromosome 4p16. In one embodiment, a genomic breakpoint occurs in exon 18 of FGFR3 and in intron 3 or exon 4 of TACC3. For example, a FGFR fusion protein can also include a fusion protein encoded by an FGFR3-TACC3 nucleic acid, wherein the nucleic acid comprises exons 1-17 of FGFR3 and a portion of intron 17 of FGFR3 located on human chromosome 4p16 spliced 5' to exons 9-16 of TACC3 located on human chromosome 4p16. In one embodiment, a genomic breakpoint occurs in intron 17 of FGFR3 and in exon 9 of TACC3. For example, a FGFR fusion protein can also include a fusion protein encoded by an FGFR1-TACC1 nucleic acid, wherein the nucleic acid comprises exons 1-17 of FGFR1 located on human chromosome 8p11 spliced 5' to exons 7-13 of TACC1 located on human chromosome 8p11. In one embodiment, a genomic breakpoint occurs in intron 17 or exon 17 of FGFR1 and in intron 6 or exon 7 of TACC1.

In one embodiment, the coordinates comprising FGFR3 translocations comprise chr4:1,795,039-1,810,599. In a further embodiment, the genomic breakpoint coordinate according to the genome build GRCh37/hg19 for FGFR3 is chr4:1,808,808, chr4:1,808,843, chr4:1,809,083, chr4:1,808,785, chr4:1,808,700, chr4:1,808,864, chr4:1,808,678, chr4:1, 808,798, or chr4:1,808,723. In a further embodiment, the coordinates comprising TACC3 translocations comprise chr4:1,723,217-1,746,905. In a further embodiment, the genomic breakpoint coordinate according to the genome build GRCh37/hg19 for TACC3 is chr4:1,732,648, chr4:1,732,757, chr4:1,739,187, chr4:1,737,091, chr4:1,737,062, chr4:1,737741, chr4:1,739,662, chr4:1,739,600, or chr4:1,738,989.

The nucleic acid can be any type of nucleic acid, including genomic DNA, complementary DNA (cDNA), recombinant DNA, synthetic or semi-synthetic DNA, as well as any form of corresponding RNA. A cDNA is a form of DNA artificially synthesized from a messenger RNA template and is used to produce gene clones. A synthetic DNA is free of modifications that can be found in cellular nucleic acids, including, but not limited to, histones and methylation. For example, a nucleic acid encoding a FGFR fusion molecule can comprise a recombinant nucleic acid encoding such a protein. The nucleic acid can be a non-naturally occurring nucleic acid created artificially (such as by assembling, cutting, ligating or amplifying sequences). It can be double-stranded or single-stranded.

The invention further provides for nucleic acids that are complementary to a FGFR fusion molecule. Complementary nucleic acids can hybridize to the nucleic acid sequence described above under stringent hybridization conditions. Non-limiting examples of stringent hybridization conditions include temperatures above 30°C, above 35°C, in excess of 42°C, and/or salinity of less than about 500 mM, or less than 200 mM.

Hybridization conditions can be adjusted by the skilled artisan via modifying the temperature, salinity and/or the concentration of other reagents such as SDS or SSC.

According to the invention, protein variants can include amino acid sequence modifications. For example, amino acid sequence modifications fall into one or more of three classes: substitutional, insertional or deletional variants. Insertions can include amino and/or carboxyl terminal fusions as well as intrasequence insertions of single or multiple amino acid residues. Insertions ordinarily will be smaller insertions than those of amino or carboxyl terminal fusions, for example, on the order of one to four residues. Deletions are characterized by the removal of one or more amino acid residues from the protein sequence. These variants ordinarily are prepared by site-specific mutagenesis of nucleotides in the DNA encoding the protein, thereby producing DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture.

In one embodiment, a FGFR fusion molecule comprises a protein or polypeptide encoded by a nucleic acid sequence encoding a FGFR fusion molecule, such as the sequences shown in SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538. In some embodiments, the nucleic acid sequence encoding a FGFR fusion molecule is about 70%, about 75%, about 80%, about 85%, about 90%, about 93%, about 95%, about 97%, about 98%, or about 99% identical to SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538.

In another embodiment, the polypeptide can be modified, such as by glycosylations and/or acetylations and/or chemical reaction or coupling, and can contain one or several non-natural or synthetic amino acids. An example of a FGFR fusion molecule is the polypeptide having the amino acid sequence shown in SEQ ID NOS: 79, 88, 150, 158-161, or 539-547. In some embodiments, the FGFR fusion molecule that is a polypeptide is about 70%, about 75%, about 80%, about 85%, about 90%, about 93%, about 95%, about 97%, about 98%, or about 99% identical to SEQ ID NOS: 79, 88, 150, 158-161, or 539-547. In another embodiment, a FGFR fusion molecule can be a fragment of a FGFR fusion protein. For example, the FGFR fusion molecule can encompass any portion of at least about 8 consecutive amino acids of SEQ ID NOS: 79, 88, 150, 158-161, or 539-547. The fragment can comprise at least about 10 amino acids, a least about 20 amino acids, at least about 30 amino acids, at least about 40 amino acids, a least about 50 amino acids, at least about 60 amino acids, or at least about 75 amino acids of SEQ ID NOS: 79, 88, 150, 158-161, or 539-547. Fragments include all possible amino acid lengths between about 8 and 100 about amino acids, for example, lengths between about 10 and 100 amino acids, between about 15 and 100 amino acids, between about 20 and 100 amino acids, between about 35 and 100 amino acids, between about 40 and 100 amino acids, between about 50 and 100 amino acids, between about 70 and 100 amino acids, between about 75 and 100 amino acids, or between about 80 and 100 amino acids. Fragments include all possible amino acid lengths between about 100 and 800 amino acids, for example, lengths between about 125 and 800 amino acids, between about 150 and 800 amino acids, between about 175 and 800 amino acids, between about 200 and 800 amino acids, between about 225 and 800 amino acids, between about 250 and 800 amino acids, between about 275 and 800 amino acids, between about 300 and 800 amino acids, between about 325 and 800 amino acids, between about 350 and 800 amino acids, between about 375 and 800 amino acids, between about 400 and 800 amino acids, between about 425 and 800 amino acids, between about 450 and 800 amino

acids, between about 475 and 800 amino acids, between about 500 and 800 amino acids, between about 525 and 800 amino acids, between about 550 and 800 amino acids, between about 575 and 800 amino acids, between about 600 and 800 amino acids, between about 625 and 800 amino acids, between about 650 and 800 amino acids, between about 675 and 800 amino acids, between about 700 and 800 amino acids, between about 725 and 800 amino acids, between about 750 and 800 amino acids, or between about 775 and 800 amino acids.

Chemical Synthesis. Nucleic acid sequences encoding a FGFR fusion molecule can be synthesized, in whole or in part, using chemical methods known in the art. Alternatively, a polypeptide can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques. Protein synthesis can either be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer).

Optionally, polypeptides fragments can be separately synthesized and combined using chemical methods to produce a full-length molecule. For example, these methods can be utilized to synthesize a fusion protein of the invention. In one embodiment, the fusion protein comprises a tyrosine kinase domain of an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein. In another embodiment, a fusion protein comprises a transforming acidic coiled-coil (TACC) domain fused to a polypeptide with a tyrosine kinase domain, wherein the TACC domain constitutively activates the tyrosine kinase domain. An example of a fusion protein is the FGFR1-TACC1 polypeptide, which comprises the amino acid sequence shown in SEQ ID NO: 150. An example of a fusion protein is the FGFR3-TACC3 polypeptide, which has the amino acid sequence comprising SEQ ID NO: 79, 158, 159, 160, 161,539, 540, 541, 542, 543, 544, 545, 546, or 547.

Obtaining, Purifying and Detecting FGFR fusion molecules. A polypeptide encoded by a nucleic acid, such as a nucleic acid encoding a FGFR fusion molecule, or a variant thereof, can be obtained by purification from human cells expressing a protein or polypeptide encoded by such a nucleic acid. Non-limiting purification methods include size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis.

A synthetic polypeptide can be substantially purified via high performance liquid chromatography (HPLC), such as ion exchange chromatography (IEX-HPLC). The composition of a synthetic polypeptide, such as a FGFR fusion molecule, can be confirmed by amino acid analysis or sequencing.

Other constructions can also be used to join a nucleic acid sequence encoding a polypeptide/protein of the claimed invention to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). Including cleavable linker sequences (i.e., those specific for Factor Xa or enterokinase (Invitrogen, San Diego, Calif.)) between the

purification domain and a polypeptide encoded by a nucleic acid of the invention also can be used to facilitate purification. For example, the skilled artisan can use an expression vector encoding 6 histidine residues that precede a thioredoxin or an enterokinase cleavage site in conjunction with a nucleic acid of interest. The histidine residues facilitate purification by immobilized metal ion affinity chromatography, while the enterokinase cleavage site provides a means for purifying the polypeptide encoded by, for example, an FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3, other FGFR-TACC, FGFR-containing, or TACC containing nucleic acid.

Host cells which contain a nucleic acid encoding a FGFR fusion molecule, and which subsequently express the same, can be identified by various procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein. For example, the presence of a nucleic acid encoding a FGFR fusion molecule can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments of nucleic acids encoding the same. In one embodiment, a nucleic acid fragment of a FGFR fusion molecule can encompass any portion of at least about 8 consecutive nucleotides of SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538. In another embodiment, the fragment can comprise at least about 10 consecutive nucleotides, at least about 15 consecutive nucleotides, at least about 20 conseutive nucleotides, or at least about 30 consecutive nucleotides of SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, 530-538. Fragments can include all possible nucleotide lengths between about 8 and about 100 nucleotides, for example, lengths between about 15 and about 100 nucleotides, or between about 20 and about 100 nucleotides. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding a FGFR fusion molecule nucleic acid, or FGFR fusion molecule nucleic acid to detect transformants which contain a nucleic acid encoding a protein or polypeptide of the same.

Protocols are known in the art for detecting and measuring the expression of a polypeptide encoded by a nucleic acid, such as a nucleic acid encoding a FGFR fusion molecule, using either polyclonal or monoclonal antibodies specific for the polypeptide. Non-limiting examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), immunostaining, and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a polypeptide encoded by a nucleic acid, such as a nucleic acid encoding a FGFR fusion molecule, can be used, or a competitive binding assay can be employed.

Labeling and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Methods for producing labeled hybridization or PCR probes for detecting sequences related to nucleic acid sequences encoding a protein, such as FGFR fusion molecule, include, but are not limited to, oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, nucleic acid sequences, such as nucleic acids encoding a FGFR fusion molecule, can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled

nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, and/or magnetic particles.

A fragment can be a fragment of a protein, such as a FGFR fusion protein. For example, a fragment of a FGFR fusion molecule can encompass any portion of at least about 8 consecutive amino acids of SEQ ID NOS: 79, 88, 150, 158-161, or 539-547. The fragment can comprise at least about 10 consecutive amino acids, at least about 20 consecutive amino acids, at least about 30 consecutive amino acids, at least about 40 consecutive amino acids, a least about 50 consecutive amino acids, at least about 60 consecutive amino acids, at least about 70 consecutive amino acids, at least about 75 consecutive amino acids, at least about 80 consecutive amino acids, at least about 85 consecutive amino acids, at least about 90 consecutive amino acids, at least about 95 consecutive amino acids, at least about 100 consecutive amino acids, at least about 200 consecutive amino acids, at least about 300 consecutive amino acids, at least about 400 consecutive amino acids, at least about 500 consecutive amino acids, at least about 600 consecutive amino acids, at least about 700 consecutive amino acids, or at least about 800 consecutive amino acids of SEQ ID NOS: 79, 88, 150, 158-161, or 539-547. Fragments include all possible amino acid lengths between about 8 and 100 about amino acids, for example, lengths between about 10 and about 100 amino acids, between about 15 and about 100 amino acids, between about 20 and about 100 amino acids, between about 35 and about 100 amino acids, between about 40 and about 100 amino acids, between about 50 and about 100 amino acids, between about 70 and about 100 amino acids, between about 75 and about 100 amino acids, or between about 80 and about 100 amino acids.

Cell Transfection

Host cells transformed with a nucleic acid sequence of interest can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. Expression vectors containing a nucleic acid sequence, such as a nucleic acid encoding a FGFR fusion molecule, can be designed to contain signal sequences which direct secretion of soluble polypeptide molecules encoded by the nucleic acid. Cell transfection and culturing methods are described in more detail below.

A eukaryotic expression vector can be used to transfect cells in order to produce proteins encoded by nucleotide sequences of the vector, e.g. those encoding a FGFR fusion molecule. Mammalian cells can contain an expression vector (for example, one that contains a nucleic acid encoding a fusion protein comprising a tyrosine kinase domain of an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein, or a nucleic acid encoding a fusion protein comprises a transforming acidic coiled-coil (TACC) domain fused to a polypeptide with a tyrosine kinase domain,

wherein the TACC domain constitutively activates the tyrosine kinase domain) via introducing the expression vector into an appropriate host cell via methods known in the art.

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed polypeptide encoded by a nucleic acid, in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, Va. 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

An exogenous nucleic acid can be introduced into a cell via a variety of techniques known in the art, such as lipofection, microinjection, calcium phosphate or calcium chloride precipitation, DEAE-dextran-mediated transfection, or electroporation. Electroporation is carried out at approximate voltage and capacitance to result in entry of the DNA construct(s) into cells of interest (such as glioma cells (cell line SF188), neuroblastoma cells (cell lines IMR-32, SK-N-SH, SH-F and SH-N), astrocytes and the like). Other transfection methods also include modified calcium phosphate precipitation, polybrene precipitation, liposome fusion, and receptor-mediated gene delivery.

Cells that will be genetically engineered can be primary and secondary cells obtained from various tissues, and include cell types which can be maintained and propagated in culture. Non-limiting examples of primary and secondary cells include epithelial cells, neural cells, endothelial cells, glial cells, fibroblasts, muscle cells (such as myoblasts) keratinocytes, formed elements of the blood (e.g., lymphocytes, bone marrow cells), and precursors of these somatic cell types.

Vertebrate tissue can be obtained by methods known to one skilled in the art, such a punch biopsy or other surgical methods of obtaining a tissue source of the primary cell type of interest. In one embodiment, a punch biopsy or removal (e.g., by aspiration) can be used to obtain a source of cancer cells (for example, glioma cells, neuroblastoma cells, and the like). A mixture of primary cells can be obtained from the tissue, using methods readily practiced in the art, such as explanting or enzymatic digestion (for examples using enzymes such as pronase, trypsin, collagenase, elastase dispase, and chymotrypsin). Biopsy methods have also been described in United States Patent No. 7,419,661 and PCT application publication WO 2001/32840, and each are hereby incorporated by reference.

Primary cells can be acquired from the individual to whom the genetically engineered primary or secondary cells are administered. However, primary cells can also be obtained from a donor, other than the recipient, of the same species. The cells can also be obtained from another species (for example, rabbit, cat, mouse, rat, sheep, goat, dog, horse, cow, bird, or pig). Primary cells can also include cells from a purified vertebrate tissue source grown attached to a tissue culture substrate (for example, flask or dish) or grown in a suspension; cells present in an explant derived from tissue; both of the aforementioned cell types plated for

the first time; and cell culture suspensions derived from these plated cells. Secondary cells can be plated primary cells that are removed from the culture substrate and replated, or passaged, in addition to cells from the subsequent passages. Secondary cells can be passaged one or more times. These primary or secondary cells can contain expression vectors having a gene that encodes a FGFR fusion molecule.

Cell Culturing

Various culturing parameters can be used with respect to the host cell being cultured. Appropriate culture conditions for mammalian cells are well known in the art (Cleveland WL, et al., *J Immunol Methods*, 1983, 56(2): 221-234) or can be determined by the skilled artisan (see, for example, Animal Cell Culture: A Practical Approach 2nd Ed., Rickwood, D. and Hames, B. D., eds. (Oxford University Press: New York, 1992)). Cell culturing conditions can vary according to the type of host cell selected. Commercially available medium can be utilized. Non-limiting examples of medium include, for example, Minimal Essential Medium (MEM, Sigma, St. Louis, Mo.); Dulbecco's Modified Eagles Medium (DMEM, Sigma); Ham's F10 Medium (Sigma); HyClone cell culture medium (HyClone, Logan, Utah); RPMI-1640 Medium (Sigma); and chemically-defined (CD) media, which are formulated for various cell types, e.g., CD-CHO Medium (Invitrogen, Carlsbad, Calif.).

The cell culture media can be supplemented as necessary with supplementary components or ingredients, including optional components, in appropriate concentrations or amounts, as necessary or desired. Cell culture medium solutions provide at least one component from one or more of the following categories: (1) an energy source, usually in the form of a carbohydrate such as glucose; (2) all essential amino acids, and usually the basic set of twenty amino acids plus cysteine; (3) vitamins and/or other organic compounds required at low concentrations; (4) free fatty acids or lipids, for example linoleic acid; and (5) trace elements, where trace elements are defined as inorganic compounds or naturally occurring elements that can be required at very low concentrations, usually in the micromolar range.

The medium also can be supplemented electively with one or more components from any of the following categories: (1) salts, for example, magnesium, calcium, and phosphate; (2) hormones and other growth factors such as, serum, insulin, transferrin, and epidermal growth factor; (3) protein and tissue hydrolysates, for example peptone or peptone mixtures which can be obtained from purified gelatin, plant material, or animal byproducts; (4) nucleosides and bases such as, adenosine, thymidine, and hypoxanthine; (5) buffers, such as HEPES; (6) antibiotics, such as gentamycin or ampicillin; (7) cell protective agents, for example pluronic polyol; and (8) galactose. In one embodiment, soluble factors can be added to the culturing medium.

The mammalian cell culture that can be used with the present invention is prepared in a medium suitable for the type of cell being cultured. In one embodiment, the cell culture medium can be any one of those previously discussed (for example, MEM) that is supplemented with serum from a mammalian source (for example, fetal bovine serum (FBS)). In another embodiment, the medium can be a conditioned medium to sustain the growth of host cells.

Three-dimensional cultures can be formed from agar (such as Gey's Agar), hydrogels (such as matrigel, agarose, and the like; Lee et al., (2004) *Biomaterials* 25: 2461-2466) or polymers that are cross-linked. These polymers can comprise natural polymers and their derivatives, synthetic polymers and their derivatives, or a combination thereof. Natural polymers can be anionic polymers, cationic polymers, amphipathic polymers, or neutral polymers. Non-limiting examples of anionic polymers can include hyaluronic acid, alginic acid (alginate), carageenan, chondroitin sulfate, dextran sulfate, and pectin. Some examples of cationic polymers, include but are not limited to, chitosan or polylysine. (Peppas et al., (2006) *Adv Mater.* 18: 1345-60; Hoffman, A. S., (2002) *Adv Drug Deliv Rev.* 43: 3-12; Hoffman, A. S., (2001) *Ann NY Acad Sci* 944: 62-73). Examples of amphipathic polymers can include, but are not limited to collagen, gelatin, fibrin, and carboxymethyl chitin. Non-limiting examples of neutral polymers can include dextran, agarose, or pullulan. (Peppas et al., (2006) *Adv Mater.* 18: 1345-60; Hoffman, A. S., (2002) *Adv Drug Deliv Rev.* 43: 3-12; Hoffman, A. S., (2001) *Ann NY Acad Sci* 944: 62-73).

Cells to be cultured can harbor introduced expression vectors, such as plasmids. The expression vector constructs can be introduced via transformation, microinjection, transfection, lipofection, electroporation, or infection. The expression vectors can contain coding sequences, or portions thereof, encoding the proteins for expression and production. Expression vectors containing sequences encoding the produced proteins and polypeptides, as well as the appropriate transcriptional and translational control elements, can be generated using methods well known to and practiced by those skilled in the art. These methods include synthetic techniques, *in vitro* recombinant DNA techniques, and *in vivo* genetic recombination which are described in J. Sambrook et al., 2001, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y. and in F. M. Ausubel et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y.

FGFR Fusion Molecule Inhibitors

The invention provides methods for use of compounds that decrease the expression level or activity of a FGFR fusion molecule in a subject. In addition, the invention provides methods for using compounds for the treatment of a gene-fusion associated cancer. In one embodiment, the gene-fusion associated cancer is an epithelial cancer. In one embodiment, the gene-fusion associated cancer comprises glioblastoma multiforme, breast cancer, lung cancer, prostate cancer, or colorectal carcinoma.

As used herein, a "FGFR fusion molecule inhibitor" refers to a compound that interacts with a FGFR fusion molecule of the invention and modulates its activity and/or its expression. For example, the compound can decrease the activity or expression of a FGFR fusion molecule. The compound can be an antagonist of a FGFR fusion molecule (e.g., a FGFR fusion molecule inhibitor). Some non-limiting examples of FGFR fusion molecule inhibitors include peptides (such as peptide fragments comprising a FGFR fusion molecule, or antibodies or fragments thereof), small molecules, and nucleic acids (such as siRNA or antisense RNA specific for a nucleic acid comprising a FGFR fusion molecule). Antagonists of a FGFR fusion molecule decrease the amount or the duration of the activity of an FGFR fusion protein. In one embodiment, the fusion protein comprises a tyrosine kinase domain of an FGFR protein fused to a polypeptide that

constitutively activates the tyrosine kinase domain of the FGFR protein (e.g., FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC), or a fusion protein comprises a transforming acidic coiled-coil (TACC) domain fused to a polypeptide with a tyrosine kinase domain, wherein the TACC domain constitutively activates the tyrosine kinase domain. Antagonists include proteins, nucleic acids, antibodies, small molecules, or any other molecule which decrease the activity of a FGFR fusion molecule.

The term "modulate," as it appears herein, refers to a change in the activity or expression of a FGFR fusion molecule. For example, modulation can cause a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of a FGFR fusion molecule, such as an FGFR fusion protein.

In one embodiment, a FGFR fusion molecule inhibitor can be a peptide fragment of a FGFR fusion protein that binds to the protein itself.

For example, the FGFR fusion polypeptide can encompass any portion of at least about 8 consecutive amino acids of SEQ ID NOS: 79, 88, 150, 158-161, or 539-547. The fragment can comprise at least about 10 consecutive amino acids, at least about 20 consecutive amino acids, at least about 30 consecutive amino acids, at least about 40 consecutive amino acids, a least about 50 consecutive amino acids, at least about 60 consecutive amino acids, at least about 70 consecutive amino acids, at least about 75 consecutive amino acids, at least about 80 consecutive amino acids, at least about 85 consecutive amino acids, at least about 90 consecutive amino acids, at least about 95 consecutive amino acids, at least about 100 consecutive amino acids, at least about 200 consecutive amino acids, at least about 300 consecutive amino acids, at least about 400 consecutive amino acids, at least about 500 consecutive amino acids, at least about 600 consecutive amino acids, at least about 700 consecutive amino acids, or at least about 800 consecutive amino acids of SEQ ID NOS: 79, 88, 150, 158-161, or 539-547. Fragments include all possible amino acid lengths between about 8 and 100 about amino acids, for example, lengths between about 10 and about 100 amino acids, between about 15 and about 100 amino acids, between about 20 and about 100 amino acids, between about 35 and about 100 amino acids, between about 40 and about 100 amino acids, between about 50 and about 100 amino acids, between about 70 and about 100 amino acids, between about 75 and about 100 amino acids, or between about 80 and about 100 amino acids. These peptide fragments can be obtained commercially or synthesized via liquid phase or solid phase synthesis methods (Atherton et al., (1989) Solid Phase Peptide Synthesis: a Practical Approach. IRL Press, Oxford, England). The FGFR fusion peptide fragments can be isolated from a natural source, genetically engineered, or chemically prepared. These methods are well known in the art.

A FGFR fusion molecule inhibitor can be a protein, such as an antibody (monoclonal, polyclonal, humanized, chimeric, or fully human), or a binding fragment thereof, directed against a FGFR fusion moleculeof the invention. An antibody fragment can be a form of an antibody other than the full-length form and includes portions or components that exist within full-length antibodies, in addition to antibody fragments that have been engineered. Antibody fragments can include, but are not limited to, single chain Fv (scFv), diabodies, Fv, and $(Fab')_2$, triabodies, Fc, Fab, CDR1, CDR2, CDR3, combinations of CDR's,

variable regions, tetrabodies, bifunctional hybrid antibodies, framework regions, constant regions, and the like (*see,* Maynard et al., (2000) *Ann. Rev. Biomed. Eng.* 2:339-76; Hudson (1998) *Curr. Opin. Biotechnol.* 9:395-402). Antibodies can be obtained commercially, custom generated, or synthesized against an antigen of interest according to methods established in the art (*see* United States Patent Nos. 6,914,128, 5,780,597, 5,811,523; Roland E. Kontermann and Stefan Dübel (editors), Antibody Engineering, Vol. I & II, (2010) 2nd ed., Springer; Antony S. Dimitrov (editor), Therapeutic Antibodies: Methods and Protocols (Methods in Molecular Biology), (2009), Humana Press; Benny Lo (editor) Antibody Engineering: Methods and Protocols (Methods in Molecular Biology), (2004) Humana Press, each of which are hereby incorporated by reference in their entireties). For example, antibodies directed to a FGFR fusion molecule can be obtained commercially from Abcam, Santa Cruz Biotechnology, Abgent, R&D Systems, Novus Biologicals, etc. Human antibodies directed to a FGFR fusion molecule (such as monoclonal, humanized, fully human, or chimeric antibodies) can be useful antibody therapeutics for use in humans. In one embodiment, an antibody or binding fragment thereof is directed against SEQ ID NOS: 79, 88, 150, 158-161, or 539-547.

Inhibition of RNA encoding a FGFR fusion molecule can effectively modulate the expression of a FGFR fusion molecule. Inhibitors are selected from the group comprising: siRNA; interfering RNA or RNAi; dsRNA; RNA Polymerase III transcribed DNAs; ribozymes; and antisense nucleic acids, which can be RNA, DNA, or an artificial nucleic acid.

Antisense oligonucleotides, including antisense DNA, RNA, and DNA/RNA molecules, act to directly block the translation of mRNA by binding to targeted mRNA and preventing protein translation. For example, antisense oligonucleotides of at least about 15 bases and complementary to unique regions of the DNA sequence encoding a FGFR fusion molecule can be synthesized, e.g., by conventional phosphodiester techniques (Dallas et al., (2006) *Med. Sci. Monit.*12(4):RA67-74; Kalota et al., (2006) *Handb. Exp. Pharmacol.* 173:173-96; Lutzelburger et al., (2006) *Handb. Exp. Pharmacol.* 173:243-59). Antisense nucleotide sequences include, but are not limited to: morpholinos, 2'-O-methyl polynucleotides, DNA, RNA and the like.

siRNA comprises a double stranded structure containing from about 15 to about 50 base pairs, for example from about 21 to about 25 base pairs, and having a nucleotide sequence identical or nearly identical to an expressed target gene or RNA within the cell. The siRNA comprise a sense RNA strand and a complementary antisense RNA strand annealed together by standard Watson-Crick base-pairing interactions. The sense strand comprises a nucleic acid sequence which is substantially identical to a nucleic acid sequence contained within the target miRNA molecule. "Substantially identical" to a target sequence contained within the target mRNA refers to a nucleic acid sequence that differs from the target sequence by about 3% or less. The sense and antisense strands of the siRNA can comprise two complementary, single-stranded RNA molecules, or can comprise a single molecule in which two complementary portions are base-paired and are covalently linked by a single-stranded "hairpin" area. See also, McMnaus and Sharp (2002) *Nat Rev Genetics*, 3:737-47, and Sen and Blau (2006) *FASEB J.*, 20:1293-99, the entire disclosures of which are herein incorporated by reference.

The siRNA can be altered RNA that differs from naturally-occurring RNA by the addition, deletion, substitution and/or alteration of one or more nucleotides. Such alterations can include addition of non-nucleotide material, such as to the end(s) of the siRNA or to one or more internal nucleotides of the siRNA, or modifications that make the siRNA resistant to nuclease digestion, or the substitution of one or more nucleotides in the siRNA with deoxyribo-nucleotides. One or both strands of the siRNA can also comprise a 3' overhang. As used herein, a 3' overhang refers to at least one unpaired nucleotide extending from the 3'-end of a duplexed RNA strand. For example, the siRNA can comprise at least one 3' overhang of from 1 to about 6 nucleotides (which includes ribonucleotides or deoxyribonucleotides) in length, or from 1 to about 5 nucleotides in length, or from 1 to about 4 nucleotides in length, or from about 2 to about 4 nucleotides in length. For example, each strand of the siRNA can comprise 3' overhangs of dithymidylic acid ("TT") or diuridylic acid ("uu").

siRNA can be produced chemically or biologically, or can be expressed from a recombinant plasmid or viral vector (for example, see U.S. Patent No. 7,294,504 and U.S. Patent No. 7,422,896, the entire disclosures of which are herein incorporated by reference). Exemplary methods for producing and testing dsRNA or siRNA molecules are described in U.S. Patent Application Publication No. 2002/0173478 to Gewirtz, U.S. Patent No. 8,071,559 to Hannon et al., and in U.S. Patent No. 7,148,342 to Tolentino et al., the entire disclosures of which are herein incorporated by reference.

In one embodiment, an siRNA directed to a human nucleic acid sequence comprising a FGFR fusion molecule can be generated against any one of SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527 or 530-538. In another embodiment, an siRNA directed to a human nucleic acid sequence comprising a breakpoint of anFGFR fusion molecule can be generated against any one of SEQ ID NOS: 1-77, 80-82, 84-145, 515, 517, 519-527 or 530-538. In one embodiment, the hairpin sequences targeting the *FGFR3* gene comprise SEQ ID NOS: 182, 183, or 184.

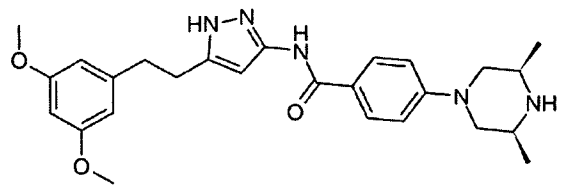
RNA polymerase III transcribed DNAs contain promoters, such as the U6 promoter. These DNAs can be transcribed to produce small hairpin RNAs in the cell that can function as siRNA or linear RNAs, which can function as antisense RNA. The FGFR fusion molecule inhibitor can comprise ribonucleotides, deoxyribonucleotides, synthetic nucleotides, or any suitable combination such that the target RNA and/or gene is inhibited. In addition, these forms of nucleic acid can be single, double, triple, or quadruple stranded. (see for example Bass (2001) *Nature*, 411:428-429; Elbashir et al., (2001) *Nature*, 411:494 498; U.S. Patent No. 6,509,154; U.S. Patent Application Publication No. 2003/0027783; and PCT Publication Nos. WO 00/044895, WO 99/032619, WO 00/01846, WO 01/029058, WO 00/044914).

FGFR fusion molecule inhibitor can be a small molecule that binds to a FGFR fusion protein described herein and disrupts its function. Small molecules are a diverse group of synthetic and natural substances generally having low molecular weights. They can be isolated from natural sources (for example, plants, fungi, microbes and the like), are obtained commercially and/or available as libraries or collections, or synthesized. Candidate small molecules that inhibit a FGFR fusion protein can be identified via *in silico* screening or high-through-put (HTP) screening of combinatorial libraries according to methods established

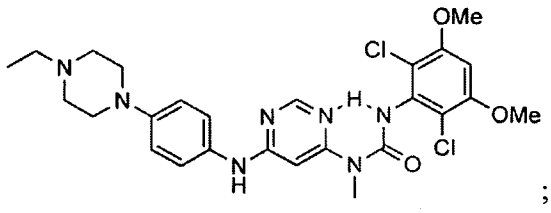
in the art (e.g., see Potyrailo et al., (2011) *ACS Comb Sci.* 13(6):579-633; Mensch et al., (2009) *J Pharm Sci.* 98(12):4429-68; Schnur (2008) *Curr Opin Drug Discov Devel.* 11(3):375-80; and Jhoti (2007) *Ernst Schering Found Symp Proc.* (3):169-85, each of which are hereby incorporated by reference in their entireties.) Most conventional pharmaceuticals, such as aspirin, penicillin, and many chemotherapeutics, are small molecules, can be obtained commercially, can be chemically synthesized, or can be obtained from random or combinatorial libraries as described below (*see*, e.g., Werner et al., (2006) *Brief Funct. Genomic Proteomic* 5(1):32-6).

Non-limiting examples of FGFR fusion molecule inhibitors include the FGFR inhibitors AZD4547 (*see* Gavine et al., (2012) *Cancer Res*, 72(8); 2045–56; *see also* PCT Application Publication No. WO 2008/075068, each of which are hereby incorporated by reference in their entireties); NVP-BGJ398 (*see* Guagnano et al., (2011) *J. Med. Chem.*, 54:7066–7083; *see also* U.S. Patent Application Publication No. 2008-0312248 A1, each of which are hereby incorporated by reference in their entireties); PD173074 (*see* Guagnano et al., (2011) *J. Med. Chem.*, 54:7066–7083; see also Mohammadi et al., (1998) *EMBO J.*, 17:5896–5904, each of which are hereby incorporated by reference in their entireties); NF449 (EMD Millipore (Billerica, MA) Cat. No. 480420; *see also* Krejci, (2010) *the Journal of Biological Chemistry*, 285(27):20644–20653, which is hereby incorporated by reference in its entirety); LY2874455 (Active Biochem; *see* Zhao et al. (2011) *Mol Cancer Ther.* (11):2200-10; *see also* PCT Application Publication No. WO 2010129509, each of which are hereby incorporated by reference in their entireties); TKI258 (Dovitinib); BIBF-1120 (Intedanib-Vargatef); BMS-582664 (Brivanib alaninate); AZD-2171 (Cediranib); TSU-68 (Orantinib); AB-1010 (Masitinib); AP-24534 (Ponatinib); and E-7080 (by Eisai). A non-limiting example of an FGFR fusion molecule inhibitor includes the TACC inhibitor KHS101 (Wurdak et al., (2010) *PNAS*, 107(38): 16542-47, which is hereby incorporated by reference in its entirety).

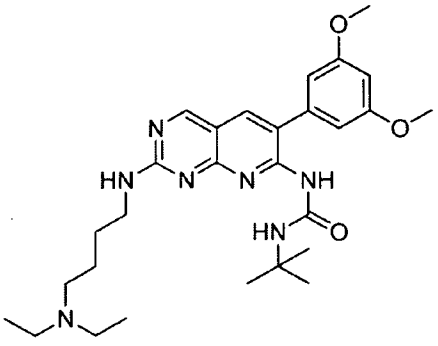
Structures of FGFR fusion molecule inhibitors useful for the invention include, but are not limited to: the FGFR inhibitor AZD4547,



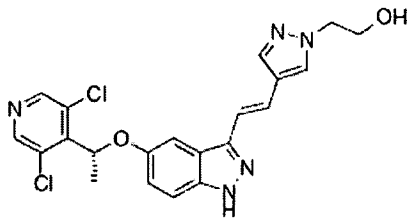
; the FGFR inhibitor NVP-BGJ398,



; the FGFR inhibitor PD173074,

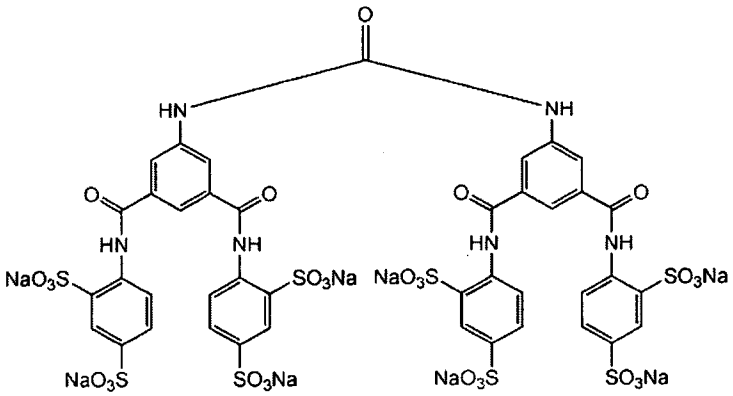


; the FGFR inhibitor LY2874455

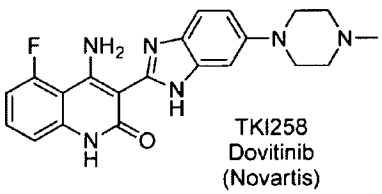


;

and the FGFR inhibitor NF449 (EMD Millipore (Billerica, MA) Cat. No. 480420),

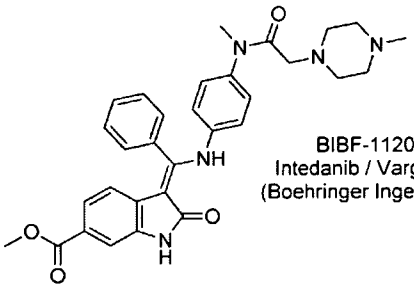


Other FGFR inhibitors include, but are not limited to:



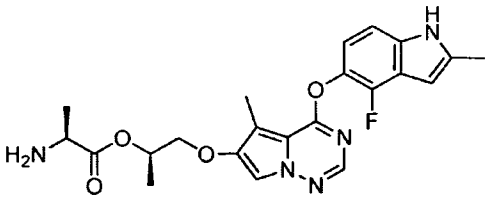
TKI258
Dovitinib
(Novartis)

,



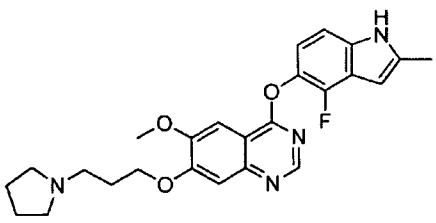
BIBF-1120
Intedanib / Vargatef
(Boehringer Ingelheim)

,



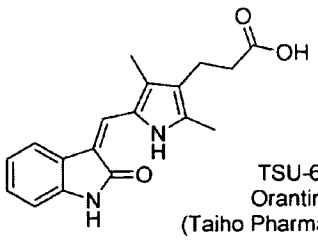
BMS-582664
Brivanib alaninate
(Bristol-Myers Squibb)

,

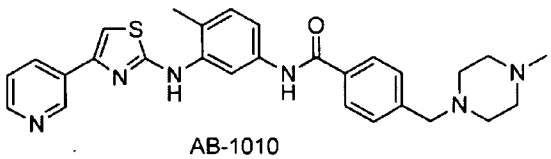


AZD-2171
Cediranib
(AstraZeneca)

,

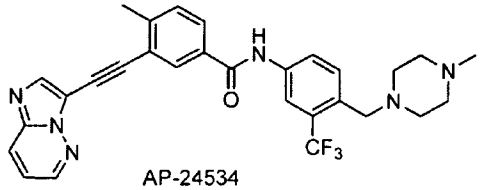


TSU-68
Orantinib
(Taiho Pharmaceutical)



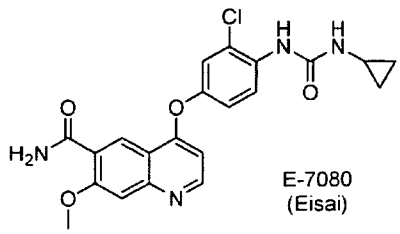
AB-1010
Masitinib
(AB Science)

,



AP-24534
Ponatinib
(Ariad Pharmaceuticals)

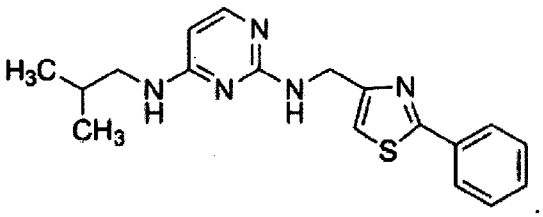
, and



E-7080
(Eisai)

In other embodiments, the FGFR fusion molecule inhibitor comprises an oral pan-FGFR tyrosine kinase inhibitor. In other embodiments, the FGFR fusion molecule inhibitor comprises JNJ-42756493. Structures of FGFR fusion molecule inhibitors useful for the invention include, but are not limited to, the FGFR inhibitor JNJ-42756493.

A structure of an FGFR fusion molecule inhibitor useful for the invention include, but is not limited to the TACC inhibitor KHS101,



Assessment and Therapuetic Treatment

The invention provides a method of decreasing the growth of a solid tumor in a subject. The tumor is associated with, but not limited to, glioblastoma multiforme, breast cancer, lung cancer, prostate cancer, or colorectal carcinoma. In another embodiment, the tumor is associated with, but not limited to, bladder carcinoma, squamous lung carcinoma and head and neck carcinoma. In one embodiment, the tumor is associated with, but not limited to, glioma. In one embodiment, the tumor is associated with, but not limited to, grade II or III glioma. In one embodiment, the tumor is associated with, but not limited to, IDH wild-type grade II or III glioma. In one embodiment, the method comprises detecting the presence of a FGFR fusion molecule in a sample obtained from a subject. In some embodiments, the sample is incubated with an agent that binds to an FGFR fusion molecule, such as an antibody, a probe, a nucleic acid primer, and the like. In further embodiments, the method comprises administering to the subject an effective amount of a FGFR fusion molecule inhibitor, wherein the inhibitor decreases the size of the solid tumor. In further embodiments, the method comprises further detecting the presence or absence of IDH1 mutations, IDH2 mutations, EGFR gene amplification, CDK4 gene amplification, or MDM2 gene amplification. In further embodiments, a FGFR fusion molecule inhibitor can be administered in combination with CDK4 inhibitors, MDM2 inhibitors, or a combination thereof.

The invention also provides a method for treating or preventing a gene-fusion associated cancer in a subject. In one embodiment, the gene-fusion associated cancer comprises an epithelial cancer. In one embodiment, the gene-fusion associated cancer comprises glioblastoma multiforme, breast cancer, lung cancer, prostate cancer, or colorectal carcinoma. In some embodiments, the epithelial cancer comprises bladder urothelial carcinoma, breast carcinoma, colorectal cancer, prostate carcinoma, lung squamous cell carcinoma, head and neck squamous cell carcinoma, or a combination of the epithelial cancers decribed. In one embodiment, the

gene-fusion associated cancer comprises glioma. In one embodiment, the gene-fusion associated cancer comprises grade II or III glioma. In one embodiment, the gene-fusion associated cancer comprises IDH wild-type grade II or III glioma. In one embodiment, the method comprises detecting the presence of a FGFR fusion molecule in a sample obtained from a subject, the presence of the fusion being indicative of a gene-fusion associated cancer, and, administering to the subject in need a therapeutic treatment against a gene-fusion associated cancer. In some embodiments, the sample is incubated with an agent that binds to an FGFR fusion molecule, such as an antibody, a probe, a nucleic acid primer, and the like. In further embodiments, the method comprises further detecting the presence or absence of IDH1 mutations, IDH2 mutations, EGFR gene amplification, CDK4 gene amplification, or MDM2 amplification. In further embodiments, an agent that binds to an FGFR fusion molecule can be administered in combination with CDK4 inhibitors, MDM2 inhibitors, or a combination thereof.

The invention also provides a method for decreasing in a subject in need thereof the expression level or activity of a fusion protein comprising the tyrosine kinase domain of an FGFR protein fused to a polypeptide that constitutively activates the tyrosine kinase domain of the FGFR protein. In some embodiments, the method comprises obtaining a biological sample from the subject. In some embodiments, the sample is incubated with an agent that binds to an FGFR fusion molecule, such as an antibody, a probe, a nucleic acid primer, and the like. In some embodiments, the method comprises administering to the subject a therapeutic amount of a composition comprising an admixture of a pharmaceutically acceptable carrier an inhibitor of the fusion protein of the invention. In one embodiment, the inhibitor is JNJ-42756493. In another embodiment, the method further comprises determining the fusion protein expression level or activity. In another embodiment, the method further comprises detecting whether the fusion protein expression level or activity is decreased as compared to the fusion protein expression level or activity prior to administration of the composition, thereby decreasing the expression level or activity of the fusion protein. In some embodiments, the fusion protein is an FGFR-TACC fusion protein. In further embodiments, the method comprises further detecting the presence or absence of IDH1 mutations, IDH2 mutations, EGFR amplification, CDK4 amplification, or MDM2 amplification.

The administering step in each of the claimed methods can comprise a drug administration, such as FGFR fusion molecule inhibitor (for example, a pharmaceutical composition comprising an antibody that specifically binds to a FGFR fusion molecule or a fragment thereof; a small molecule that specifically binds to a FGFR protein; a small molecule that specifically binds to a TACC protein; an antisense RNA or antisense DNA that decreases expression of a FGFR fusion molecule; a siRNA that specifically targets a gene encoding a FGFR fusion molecule; a small molecule such as JNJ-42756493; or a combination thereof). In one embodiment, the therapeutic molecule to be administered comprises a polypeptide of a FGFR fusion molecule, comprising at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 93%, at least about 95%, at least about 97%, at least about 98%, at least about 99%, or 100% of the amino acid sequence of SEQ ID NOS: 79, 88, 150, 158-161, or 539-547 and exhibits the function of decreasing expression of such a protein, thus treating a gene fusion-associated cancer. In another

embodiment, administration of the therapeutic molecule decreases the size of the solid tumor associated with glioblastoma multiforme, breast cancer, lung cancer, prostate cancer, colorectal carcinoma, bladder carcinoma, squamous lung carcinoma and head and neck carcinoma, glioma, grade II or III glioma, or IDH wild-type grade II or III glioma. In further embodiments, the therapeutic molecule can be administered in combination with CDK4 inhibitors, MDM2 inhibitors, or a combination thereof.

In another embodiment, the therapeutic molecule to be administered comprises an siRNA directed to a human nucleic acid sequence comprising a FGFR fusion molecule. In one embodiment, the siRNA is directed to any one of SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538. In another embodiment, the siRNA is directed to any one of SEQ ID NOS: 1-77, 80-82, 84-145, 515, 517, 519-527, or 530-538. In a further embodiment, the therapeutic molecule to be administered comprises an antibody or binding fragment thereof, which is directed against SEQ ID NOS: 79, 88, 150, 158-161, or 539-547. In some embodiments, the therapeutic molecule to be administered comprises a small molecule that specifically binds to a FGFR protein, such as AZD4547, NVP-BGJ398, PD173074, NF449, TK1258, BIBF-1120, BMS-582664, AZD-2171, TSU68, AB1010, AP24534, E-7080, or LY2874455. In some embodiments, the therapeutic molecule to be administered is JNJ-42756493. In other embodiments, the therapeutic molecule to be administered comprises a small molecule that specifically binds to a TACC protein, such as KHS101.

In one embodiment, the invention provides for the detection of a chromosomal rearrangement at given chromosomal coordinates. In another embodiment, the detection or determination comprises nucleic acid sequencing, selective hybridization, selective amplification, gene expression analysis, or a combination thereof. In another embodiment, the detection or determination comprises protein expression analysis, for example by western blot analysis, immunostaining, ELISA, or other antibody detection methods.

In one embodiment, the biological sample comprises neuronal cells, serum, bone marrow, blood, peripheral blood, lymph nodes, cerebro-spinal fluid, urine, a saliva sample, a buccal swab, a serum sample, a sputum sample, a lacrimal secretion sample, a semen sample, a vaginal secretion sample, a fetal tissue sample, or a combination thereof. In some embodiments the sample is a tissue sample. In some embodiments, the sample is a paraffi- embedded tissue section. In some embodiments, the tissue sample is a tumor sample.

A FGFR fusion molecule, for example, a fusion between FGFR1, FGFR2, FGFR3, or any other FGFR, and TACC1, TACC2, TACC3 or any other TACC, can be determined at the level of the DNA, RNA, or polypeptide. Optionally, detection can be determined by performing an oligonucleotide ligation assay, a confirmation based assay, a hybridization assay, a sequencing assay, an allele-specific amplification assay, a microsequencing assay, a melting curve analysis, a denaturing high performance liquid chromatography (DHPLC) assay (for example, *see* Jones et al, (2000) *Hum Genet.*, 106(6):663-8), or a combination thereof.

In one embodiment, the detection is performed by sequencing all or part of a FGFR fusion molecule (e.g., a FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid, or a FGFR1, TACC1, FGFR2, TACC2, FGFR3, TACC3 or other FGFR or TACC nucleic acid), or by selective hybridization or amplification of all or part of a FGFR fusion molecule (e.g., a FGFR1-TACC1, FGFR2-

TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid, or a FGFR1, TACC1, FGFR2, TACC2, FGFR3, TACC3 or other FGFR or TACC nucleic acid). A FGFR fusion molecule specific amplification (e.g., a FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid specific amplification) can be carried out before the fusion identification step.

The invention provides for a method of detecting a chromosomal alteration in a subject afflicted with a gene-fusion associated cancer. In one embodiment, the chromosomal alteration is an in-frame fused transcript described herein, for example an FGFR fusion molecule. In some embodiments, the chromosomal alteration is a chromosomal translocation, for example an FGFR fusion molecule. An alteration in a chromosome region occupied by a FGFR fusion molecule, such as a FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid, can be any form of mutation(s), deletion(s), rearrangement(s) and/or insertions in the coding and/or non-coding region of the locus, alone or in various combination(s). Mutations can include point mutations. Insertions can encompass the addition of one or several residues in a coding or non-coding portion of the gene locus. Insertions can comprise an addition of between 1 and 50 base pairs in the gene locus. Deletions can encompass any region of one, two or more residues in a coding or non-coding portion of the gene locus, such as from two residues up to the entire gene or locus. Deletions can affect smaller regions, such as domains (introns) or repeated sequences or fragments of less than about 50 consecutive base pairs, although larger deletions can occur as well. Rearrangement includes inversion of sequences. The alteration in a chromosome region occupied by a FGFR fusion molecule, e.g., a FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid, can result in amino acid substitutions, RNA splicing or processing, product instability, the creation of stop codons, production of oncogenic fusion proteins, frame-shift mutations, and/or truncated polypeptide production. The alteration can result in the production of a FGFR fusion molecule, for example, one encoded by a FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid, with altered function, stability, targeting or structure. The alteration can also cause a reduction, or even an increase in protein expression. In one embodiment, the alteration in the chromosome region occupied by a FGFR fusion molecule can comprise a chromosomal rearrangement resulting in the production of a FGFR fusion molecule, such as a FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC fusion. This alteration can be determined at the level of the DNA, RNA, or polypeptide. In another embodiment, the detection or determination comprises nucleic acid sequencing, selective hybridization, selective amplification, gene expression analysis, or a combination thereof. In another embodiment, the detection or determination comprises protein expression analysis, for example by western blot analysis, ELISA, immunostaining or other antibody detection methods. In one embodiment, the coordinates comprising FGFR1 translocations comprise chr8:38,268,656-38,325,363. In another embodiment, the coordinates comprising FGFR2 translocations comprise chr10:123,237,844-123,357,972. In a further embodiment, the coordinates comprising FGFR3 translocations comprise chr4:1,795,039-1,810,599. In yet another embodiment, the coordinates comprising FGFR4 translocations comprise chr5:176,513,921-176,525,126. In one embodiment, the coordinates comprising TACC1 translocations comprise chr8:38,644,722-38,710,546. In

another embodiment, the coordinates comprising TACC2 translocations comprise chr10:123,748,689-124,014,057. In a further embodiment, the coordinates comprising TACC3 translocations comprise chr4:1,723,217-1,746,905.

The present invention provides a method for treating a gene-fusion associated cancer in a subject in need thereof. In one embodiment, the method comprises obtaining a sample from the subject to determine the level of expression of an FGFR fusion molecule in the subject. In some embodiments, the sample is incubated with an agent that binds to an FGFR fusion molecule, such as an antibody, a probe, a nucleic acid primer, and the like. In another embodiment, the detection or determination comprises nucleic acid sequencing, selective hybridization, selective amplification, gene expression analysis, or a combination thereof. In another embodiment, the detection or determination comprises protein expression analysis, for example by western blot analysis, immunostaining, ELISA, or other antibody detection methods. In some embodiments, the method further comprises assessing whether to administer a FGFR fusion molecule inhibitor based on the expression pattern of the subject. In further embodiments, the method comprises administering a FGFR fusion molecule inhibitor to the subject. In one embodiment, the FGFR fusion molecule inhibitor is JNJ-42756493. In one embodiment, the gene-fusion associated cancer comprises an epithelial cancer. In one embodiment, the gene-fusion associated cancer comprises glioblastoma multiforme, breast cancer, lung cancer, prostate cancer, or colorectal carcinoma. In some embodiments, the epithelial cancer comprises bladder urothelial carcinoma, breast carcinoma, colorectal cancer, prostate carcinoma, lung squamous cell carcinoma, head and neck squamous cell carcinoma, or a combination of the epithelial cancers decribed. In one embodiment, the gene-fusion associated cancer comprises glioma, grade II or III glioma, or IDH wild-type grade II or III glioma. In further embodiments, the method comprises further detecting the presence or absence of IDH1 mutations, IDH2 mutations, EGFR gene amplification, CDK4 gene amplification, or MDM2 gene amplification. In further embodiments, a FGFR fusion molecule inhibitor can be administered in combination with CDK4 inhibitors, MDM2 inhibitors, or a combination thereof.

In one embodiment, the invention provides for a method of detecting the presence of altered RNA expression of an FGFR fusion molecule in a subject, for example one afflicted with a gene-fusion associated cancer. In another embodiment, the invention provides for a method of detecting the presence of an FGFR fusion molecule in a subject. In some embodiments, the method comprises obtaining a sample from the subject to determine whether the subject expresses an FGFR fusion molecule. In some embodiments, the sample is incubated with an agent that binds to an FGFR fusion molecule, such as an antibody, a probe, a nucleic acid primer, and the like. In other embodiments, the detection or determination comprises nucleic acid sequencing, selective hybridization, selective amplification, gene expression analysis, or a combination thereof. In another embodiment, the detection or determination comprises protein expression analysis, for example by western blot analysis, ELISA, or other antibody detection methods. In some embodiments, the method further comprises assessing whether to administer a FGFR fusion molecule inhibitor based on the expression pattern of the subject. In further embodiments, the method comprises administering a FGFR

fusion molecule inhibitor to the subject. In one embodiment, the FGFR fusion molecule inhibitor is JNJ-42756493. Altered RNA expression includes the presence of an altered RNA sequence, the presence of an altered RNA splicing or processing, or the presence of an altered quantity of RNA. These can be detected by various techniques known in the art, including sequencing all or part of the RNA or by selective hybridization or selective amplification of all or part of the RNA. In a further embodiment, the method can comprise detecting the presence or expression of a FGFR fusion molecule, such as one encoded by a FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid. Altered polypeptide expression includes the presence of an altered polypeptide sequence, the presence of an altered quantity of polypeptide, or the presence of an altered tissue distribution. These can be detected by various techniques known in the art, including by sequencing and/or binding to specific ligands (such as antibodies). In one embodiment, the detecting comprises using a northern blot; real time PCR and primers directed to SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538; a ribonuclease protection assay; a hybridization, amplification, or sequencing technique to detect an FGFR fusion molecule, such as one comprising SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538; or a combination thereof. In another embodiment, the PCR primers comprise SEQ ID NOS: 162, 163, 164, 165 166, 167, 168, 169, 495, 496, 497, 498, 507, 508, 509, 510, 511, 512, 513, or 514. In a further embodiment, primers used for the screening of FGFR fusion molecules, such as FGFR-TACC fusions, comprise SEQ ID NOS: 166, 167, 168, 169, 495, 496, 497, 498, 507, 508, 509, 510, 511, 512, 513, or 514. In some embodiments, primers used for genomic detection of an FGFR3-TACC3 fusion comprise SEQ ID NOS: 170 171, 499, 500, 501, 502, 503, 504, 505, or 506.

In some aspects of the invention, the method comprises further detecting the presence or absence of IDH1 mutations, IDH2 mutations, EGFR gene amplification, CDK4 gene amplification, or MDM2 gene amplification. MDM2 encodes a nuclear-localized E3 ubiquitin ligase. Alternative splicing results in a multitude of transcript variants, many of which may be expressed only in tumor cells. EGFR (epidermal growth factor receptor) is a transmembrane glycoprotein that is a member of the protein kinase superfamily. This protein is a receptor for members of the epidermal growth factor family. EGFR is a cell surface protein that binds to epidermal growth factor. Multiple alternatively spliced transcript variants that encode different protein isoforms have been found for this gene. CDK4 (cyclin dependent kinase 4) is a member of the Ser/Thr protein kinase family. It is a catalytic subunit of the protein kinase complex that is important for cell cycle G1 phase progression. The activity of this kinase is restricted to the G1-S phase, which is controlled by the regulatory subunits D-type cyclins and CDK inhibitor p16(INK4a). Multiple polyadenylation sites of this gene have been reported. IDH1 (isocitrate dehydrogenase 1 (NADP+), soluble) catalyzes the oxidative decarboxylation of isocitrate to 2-oxoglutarate. Alternatively spliced transcript variants encoding the same protein have been found for this gene. IDH1 mutations, EGFR amplification, CDK4 amplification, or MDM2 amplification can be detected using various techniques know in the art, including, but not limited to sequencing, qPCR, and microarrays (e.g. CGH arrays, SNP arrays).

Various techniques known in the art can be used to detect or quantify altered gene or RNA expression or nucleic acid sequences, which include, but are not limited to, hybridization, sequencing, amplification,

and/or binding to specific ligands (such as antibodies). Other suitable methods include allele-specific oligonucleotide (ASO), oligonucleotide ligation, allele-specific amplification, Southern blot (for DNAs), Northern blot (for RNAs), single-stranded conformation analysis (SSCA), PFGE, fluorescent in situ hybridization (FISH), gel migration, clamped denaturing gel electrophoresis, denaturing HLPC, melting curve analysis, heteroduplex analysis, RNase protection, chemical or enzymatic mismatch cleavage, ELISA, radio-immunoassays (RIA) and immuno-enzymatic assays (IEMA).

Some of these approaches (such as SSCA and constant gradient gel electrophoresis (CGGE)) are based on a change in electrophoretic mobility of the nucleic acids, as a result of the presence of an altered sequence. According to these techniques, the altered sequence is visualized by a shift in mobility on gels. The fragments can then be sequenced to confirm the alteration. Some other approaches are based on specific hybridization between nucleic acids from the subject and a probe specific for wild type or altered gene or RNA. The probe can be in suspension or immobilized on a substrate. The probe can be labeled to facilitate detection of hybrids. Some of these approaches are suited for assessing a polypeptide sequence or expression level, such as Northern blot, ELISA and RIA. These latter require the use of a ligand specific for the polypeptide, for example, the use of a specific antibody.

Hybridization. Hybridization detection methods are based on the formation of specific hybrids between complementary nucleic acid sequences that serve to detect nucleic acid sequence alteration(s). A detection technique involves the use of a nucleic acid probe specific for a wild type or altered gene or RNA, followed by the detection of the presence of a hybrid. The probe can be in suspension or immobilized on a substrate or support (for example, as in nucleic acid array or chips technologies). The probe can be labeled to facilitate detection of hybrids. In one embodiment, the probe according to the invention can comprise a nucleic acid directed to SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538. For example, a sample from the subject can be contacted with a nucleic acid probe specific for a gene encoding a FGFR fusion molecule, and the formation of a hybrid can be subsequently assessed. In one embodiment, the method comprises contacting simultaneously the sample with a set of probes that are specific for an FGFR fusion molecule. Also, various samples from various subjects can be investigated in parallel.

According to the invention, a probe can be a polynucleotide sequence which is complementary to and specifically hybridizes with a, or a target portion of a, gene or RNA corresponding to a FGFR fusion molecule. Useful probes are those that are complementary to the gene, RNA, or target portion thereof. Probes can comprise single-stranded nucleic acids of between 8 to 1000 nucleotides in length, for instance between 10 and 800, between 15 and 700, or between 20 and 500. Longer probes can be used as well. A useful probe of the invention is a single stranded nucleic acid molecule of between 8 to 500 nucleotides in length, which can specifically hybridize to a region of a gene or RNA that corresponds to a FGFR fusion molecule.

The sequence of the probes can be derived from the sequences of the FGFR fusion genes provided herein. Nucleotide substitutions can be performed, as well as chemical modifications of the probe. Such chemical modifications can be accomplished to increase the stability of hybrids (e.g., intercalating groups) or to label

the probe. Some examples of labels include, without limitation, radioactivity, fluorescence, luminescence, and enzymatic labeling.

A guide to the hybridization of nucleic acids is found in e.g., Sambrook, ed., Molecular Cloning: A Laboratory Manual (3rd Ed.), Vols. 1-3, Cold Spring Harbor Laboratory, 1989; Current Protocols In Molecular Biology, Ausubel, ed. John Wiley & Sons, Inc., New York, 2001; Laboratory Techniques In Biochemistry And Molecular Biology: Hybridization With Nucleic Acid Probes, Part I. Theory and Nucleic Acid Preparation, Tijssen, ed. Elsevier, N.Y., 1993.

Sequencing. Sequencing can be carried out using techniques well known in the art, using automatic sequencers. The sequencing can be performed on the complete FGFR fusion molecule or on specific domains thereof.

Amplification. Amplification is based on the formation of specific hybrids between complementary nucleic acid sequences that serve to initiate nucleic acid reproduction. Amplification can be performed according to various techniques known in the art, such as by polymerase chain reaction (PCR), ligase chain reaction (LCR), strand displacement amplification (SDA) and nucleic acid sequence based amplification (NASBA). These techniques can be performed using commercially available reagents and protocols. Useful techniques in the art encompass real-time PCR, allele-specific PCR, or PCR based single-strand conformational polymorphism (SSCP). Amplification usually requires the use of specific nucleic acid primers, to initiate the reaction. For example, nucleic acid primers useful for amplifying sequences corresponding to a FGFR fusion molecule are able to specifically hybridize with a portion of the gene locus that flanks a target region of the locus. In one embodiment, amplification comprises using forward and reverse PCR primers directed to SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538. Nucleic acid primers useful for amplifying sequences from a FGFR fusion molecule (e.g., a FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid); the primers specifically hybridize with a portion of an FGFR fusion molecule. In certain subjects, the presence of an FGFR fusion molecule corresponds to a subject with a gene fusion-associated cancer. In one embodiment, amplification can comprise using forward and reverse PCR primers comprising nucleotide sequences of SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538. In one embodiment, amplification can comprise using forward and reverse PCR primers comprising nucleotide sequences of SEQ ID NOS: 162-169, or 495-514.

Non-limiting amplification methods include, e.g., polymerase chain reaction, PCR (PCR Protocols, A Guide To Methods And Applications, ed. Innis, Academic Press, N.Y., 1990 and PCR Strategies, 1995, ed. Innis, Academic Press, Inc., N.Y.); ligase chain reaction (LCR) (Wu (1989) *Genomics* 4:560; Landegren (1988) *Science* 241:1077; Barringer (1990) *Gene* 89:117); transcription amplification (Kwoh (1989) *PNAS* 86:1173); and, self-sustained sequence replication (Guatelli (1990) *PNAS* 87:1874); Q Beta replicase amplification (Smith (1997) *J. Clin. Microbiol.* 35:1477-1491), automated Q-beta replicase amplification assay (Burg (1996) *Mol. Cell. Probes* 10:257-271) and other RNA polymerase mediated techniques (e.g., NASBA, Cangene, Mississauga, Ontario; *see also* Berger (1987) *Methods Enzymol.* 152:307-316; U.S. Pat.

Nos. 4,683,195 and 4,683,202; and Sooknanan (1995) *Biotechnology* 13:563-564). All the references stated above are incorporated by reference in their entireties.

The invention provides for a nucleic acid primer, wherein the primer can be complementary to and hybridize specifically to a portion of a FGFR fusion molecule, such as a FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid (e.g., DNA or RNA) in certain subjects having a gene fusion-associated cancer. In one embodiment, the gene-fusion associated cancer comprises glioblastoma multiforme, breast cancer, lung cancer, prostate cancer, or colorectal carcinoma. Primers of the invention can be specific for fusion sequences in a FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid (DNA or RNA). By using such primers, the detection of an amplification product indicates the presence of a fusion of a FGFR1 and TACC1, FGFR2 and TACC2, FGFR3 and TACC3 or other FGFR and TACC nucleic acid. Examples of primers of this invention can be single-stranded nucleic acid molecules of about 5 to 60 nucleotides in length, or about 8 to about 25 nucleotides in length. The sequence can be derived directly from the sequence of a FGFR fusion molecule, e.g. FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC nucleic acid. Perfect complementarity is useful to ensure high specificity; however, certain mismatch can be tolerated. For example, a nucleic acid primer or a pair of nucleic acid primers as described above can be used in a method for detecting the presence of a gene fusion-associated cancer in a subject. In one embodiment, primers can be used to detect an FGFR fusion molecule, such as a primer comprising SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538; or a combination thereof. In another embodiment, the PCR primers comprise SEQ ID NOS: 162, 163, 164, 165, 166, 167, 168, 169, 495, 496, 497, 498, 507, 508, 509, 510, 511, 512, 513, or 514. In a further embodiment, primers used for the screening of FGFR fusion molecules, such as FGFR-TACC fusions, comprise SEQ ID NOS: 166, 167, 168, 169, 495, 496, 497, 498, 507, 508, 509, or 510. In some embodiments, primers used for genomic detection of an FGFR3-TACC3 fusion comprise SEQ ID NOS: 170, 171, 499, 500, 501, 502, 503, 504, 505, or 506.In one embodiment, the method can comprise contacting a sample from the subject with primers specific for a FGFR fusion molecule, and determining the presence of an PCR product. In another embodiment, the method can comprise contacting a sample from the subject with primer specific for a FGFR molecule, or a TACC molecule, and determining the presence of a PCR product. In another embodiment, the primers can recognize the nucleic acids encoding a FGFR3 C-terminal region, or nucleic acids encoding a TACC3 N-terminal region, or a combination thereof. In another embodiment, the method can comprise contacting a sample from the subject with primers specific for a FGFR molecule, or a TACC molecule, or a FGFR fusion molecule, and determining the amount of PCR product formed compared to the amount of PCR product formed in non-tumor cells or tissue, wherein an increased amount of PCR product indicates the presence of an FGFR fusion. In one embodiment, primers and/or the PCR product are labeled to enable detection of the PCR product. For example, nucleic acid primers useful for amplifying sequences corresponding to a FGFR fusion molecules can be labeled with fluorescent molecules, radioactive molecules, chemiluminescent molecules, or affinity molecules (e.g. biotin) which can then be detected by methods known in the art (e.g. fluorescently labeled streptavidin). PCR products can also be detected by

using dyes that can be incorporated into newly formed PCR products, such as, but not limited to, SYBR Green.

Specific Ligand Binding. As discussed herein, a nucleic acid encoding a FGFR fusion molecule or expression of a FGFR fusion molecule, can also be detected by screening for alteration(s) in a sequence or expression level of a polypeptide encoded by the same. Different types of ligands can be used, such as specific antibodies. In one embodiment, the sample is contacted with an antibody specific for a polypeptide encoded by a FGFR fusion molecule and the formation of an immune complex is subsequently determined. Various methods for detecting an immune complex can be used, such as ELISA, immunostaining, radioimmunoassays (RIA) and immuno-enzymatic assays (IEMA).

For example, an antibody can be a polyclonal antibody, a monoclonal antibody, as well as fragments or derivatives thereof having substantially the same antigen specificity. Fragments include Fab, Fab'2, or CDR regions. Derivatives include single-chain antibodies, humanized antibodies, or poly-functional antibodies. An antibody specific for a polypeptide encoded by a FGFR fusion molecule can be an antibody that selectively binds such a polypeptide. In one embodiment, the antibody is raised against a polypeptide encoded by a FGFR fusion molecule (such as FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC fusion) or an epitope-containing fragment thereof. Although non-specific binding towards other antigens can occur, binding to the target polypeptide occurs with a higher affinity and can be reliably discriminated from non-specific binding. In one embodiment, the method can comprise contacting a sample from the subject with an antibody specific for a FGFR fusion molecule, and determining the presence of an immune complex. Optionally, the sample can be contacted to a support coated with antibody specific for a FGFR fusion molecule. In one embodiment, the sample can be contacted simultaneously, or in parallel, or sequentially, with various antibodies specific for different forms of a FGFR fusion molecule, e.g., FGFR1-TACC1, FGFR2-TACC2, FGFR3-TACC3 or other FGFR-TACC fusion.

In one embodiment, the method can comprise contacting a sample from the subject with an antibody specific for a FGFR fusion molecule, and determining the presence of an immune complex. In another embodiment, the method can comprise contacting a sample from the subject with an antibody specific for a FGFR molecule, or a TACC molecule, and determining the presence of an immune complex. In another embodiment, the antibody can recognize the FGFR3 C-terminal region, or the TACC3 N-terminal region, or a combination thereof. In another embodiment, the antibody can recognize the FGFR3 C-terminal region, or the TACC3 N-terminal region, or a combination thereof. In another embodiment, the method can comprise contacting a sample from the subject with an antibody specific for a FGFR molecule, or a TACC molecule, or a FGFR fusion molecule, and determining the amount of an immune complex formed compared to the amount of immune complex formed in non-tumor cells or tissue, wherein an increased amount of an immune complex indicates the presence of an FGFR fusion.

Detection the formation of a complex between an antibody and a protein can be performed by a variety of method known in the art. For example, an antibody-protein complex can be detected by using antibodies or secondary antibodies labeled with fluorescent molecules, chromogenic molecules, chemiluminescent

molecules, radioactive isotopes, or affinity molecules (e.g. biotin) which can then be detected by methods known in the art (e.g. fluorescently labeled streptavidin).

The invention also provides for a diagnostic kit comprising products and reagents for detecting in a sample from a subject the presence of a FGFR fusion molecule. The kit can be useful for determining whether a sample from a subject exhibits increased or reduced expression of a FGFR fusion molecule. For example, the diagnostic kit according to the present invention comprises any primer, any pair of primers, any nucleic acid probe and/or any ligand, or any antibody directed specifically to a FGFR fusion molecule. The diagnostic kit according to the present invention can further comprise reagents and/or protocols for performing a hybridization, amplification, or antigen-antibody immune reaction. In one embodiment, the kit can comprise nucleic acid primers that specifically hybridize to and can prime a polymerase reaction from a FGFR fusion molecule comprising SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538, or a combination thereof. In one embodiment, primers can be used to detect a FGFR fusion molecule, such as a primer directed to SEQ ID NOS: 80-82, 84, 94-145, 515, 517, 519-527, or 530-538; or a combination thereof. In another embodiment, the PCR primer comprises SEQ ID NOS: 162, 163, 164, 165, 166, 167, 168, 169, 495, 496, 497, 498, 507, 508, 509, 510, 511, 512, 513, or 514. In a further embodiment, primers used for the screening of FGFR fusion molecules, such as FGFR-TACC fusions, comprise SEQ ID NOS: 166, 167, 168, 169, 495, 496, 497, 498, 507, 508, 509, or 510. In some embodiments, primers used for genomic detection of an FGFR3-TACC3 fusion comprise SEQ ID NOS: 170, 171, 499, 500, 501, 502, 503, 504, 505, or 506. In some embodiments, the kit comprises an antibody that specifically binds to a FGFR fusion molecule comprising SEQ ID NOS: 79, 85-89, 150, 158-161, or 539-547, wherein the antibody will recognize the protein only when a FGFR fusion molecule is present.The diagnosis methods can be performed *in vitro*, *ex vivo*, or *in vivo*. These methods utilize a sample from the subject in order to assess the status of a FGFR fusion molecule. The sample can be any biological sample derived from a subject, which contains nucleic acids or polypeptides. Examples of such samples include, but are not limited to, fluids, tissues, cell samples, organs, and tissue biopsies. Non-limiting examples of samples include blood, liver, plasma, serum, saliva, urine, or seminal fluid. In some embodiments the sample is a tissue sample. In some embodiments, the sample is a paraffin embedded tissue section. In some embodiments, the tissue sample is a tumor sample. The sample can be collected according to conventional techniques and used directly for diagnosis or stored. The sample can be treated prior to performing the method, in order to render or improve availability of nucleic acids or polypeptides for testing. Treatments include, for instance, lysis (e.g., mechanical, physical, or chemical), centrifugation. The nucleic acids and/or polypeptides can be pre-purified or enriched by conventional techniques, and/or reduced in complexity. Nucleic acids and polypeptides can also be treated with enzymes or other chemical or physical treatments to produce fragments thereof. In one embodiment, the sample is contacted with reagents, such as probes, primers, or ligands, in order to assess the presence of a FGFR fusion molecule. Contacting can be performed in any suitable device, such as a plate, tube, well, or glass. In some embodiments, the contacting is performed on a substrate coated with the reagent, such as a nucleic acid array or a specific ligand array. The substrate can

be a solid or semi-solid substrate such as any support comprising glass, plastic, nylon, paper, metal, or polymers. The substrate can be of various forms and sizes, such as a slide, a membrane, a bead, a column, or a gel. The contacting can be made under any condition suitable for a complex to be formed between the reagent and the nucleic acids or polypeptides of the sample.

Nucleic Acid Delivery Methods

Delivery of nucleic acids into viable cells can be effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, such as viral vectors (e.g., lentivirus, adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). Non-limiting techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, and the calcium phosphate precipitation method (See, for example, Anderson, *Nature*, 1998) supplement to 392(6679):25(). Introduction of a nucleic acid or a gene encoding a polypeptide of the invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells can also be cultured ex vivo in the presence of therapeutic compositions of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

Nucleic acids can be inserted into vectors and used as gene therapy vectors. A number of viruses have been used as gene transfer vectors, including papovaviruses, e.g., SV40 (Madzak et al., (1992) *J Gen Virol.* 73(Pt 6):1533-6), adenovirus (Berkner (1992) *Curr Top Microbiol Immunol.*158:39-66; Berkner (1988) *Biotechniques*, 6(7):616-29; Gorziglia and Kapikian (1992) *J Virol.* 66(7):4407-12; Quantin et al., (1992) *Proc Natl Acad Sci U S A.* 89(7):2581-4; Rosenfeld et al., (1992) *Cell.* 68(1):143-55; Wilkinson et al., (1992) *Nucleic Acids Res.* 20(9):2233-9; Stratford-Perricaudet et al., (1990) *Hum Gene Ther.* 1(3):241-56), vaccinia virus (Moss (1992) *Curr Opin Biotechnol.* 3(5):518-22), adeno-associated virus (Muzyczka, (1992) *Curr Top Microbiol Immunol.* 158:97-129; Ohi et al., (1990) *Gene.* 89(2):279-82), herpesviruses including HSV and EBV (Margolskee (1992) *Curr Top Microbiol Immunol.* 158:67-95; Johnson et al., (1992) *Brain Res Mol Brain Res.*12(1-3):95-102; Fink et al., (1992) *Hum Gene Ther.* 3(1):11-9; Breakefield and Geller (1987) *Mol Neurobiol.* 1(4):339-71; Freese et al., (1990) *Biochem Pharmacol.* 40(10):2189-99), and retroviruses of avian (Bandyopadhyay and Temin (1984) *Mol Cell Biol.* 4(4):749-54; Petropoulos et al., (1992) *J Virol.* 66(6):3391-7), murine (Miller et al. (1992) *Mol Cell Biol.* 12(7):3262-72; Miller et al., (1985) *J Virol.* 55(3):521-6; Sorge et al., (1984) *Mol Cell Biol.* 4(9):1730-7; Mann and Baltimore (1985) *J Virol.* 54(2):401-7; Miller et al., (1988) *J Virol.* 62(11):4337-45), and human origin (Shimada et al., (1991) *J Clin Invest.* 88(3):1043-7; Helseth et al., (1990) *J Virol.* 64(12):6314-8; Page et al., (1990) *J Virol.* 64(11):5270-6; Buchschacher and Panganiban (1992) *J Virol.* 66(5):2731-9).

Non-limiting examples of *in vivo* gene transfer techniques include transfection with viral (e.g., retroviral) vectors (*see* U.S. Pat. No. 5,252,479, which is incorporated by reference in its entirety) and viral coat protein-liposome mediated transfection (Dzau et al., (1993) *Trends in Biotechnology* 11:205-210), incorporated entirely by reference). For example, naked DNA vaccines are generally known in the art; *see*

Brower, (1998) *Nature Biotechnology*, 16:1304-1305, which is incorporated by reference in its entirety. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (*see*, e.g., U.S. Pat. No. 5,328,470) or by stereotactic injection (*see*, e.g., Chen, et al., (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g., retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

For reviews of nucleic acid delivery protocols and methods *see* Anderson et al. (1992) *Science* 256:808-813; U.S. Pat. Nos. 5,252,479, 5,747,469, 6,017,524, 6,143,290, 6,410,010 6,511,847; and U.S. Application Publication No. 2002/0077313, which are all hereby incorporated by reference in their entireties. For additional reviews, *see* Friedmann (1989) *Science*, 244:1275-1281; Verma, *Scientific American*: 68-84 (1990); Miller (1992) *Nature*, 357: 455-460; Kikuchi et al. (2008) *J Dermatol Sci.* 50(2):87-98; Isaka et al. (2007) *Expert Opin Drug Deliv.* 4(5):561-71; Jager et al.(2007) *Curr Gene Ther.* 7(4):272-83; Waehler et al.(2007) *Nat Rev Genet.* 8(8):573-87; Jensen et al. (2007) *Ann Med.* 39(2):108-15; Herweijer et al. (2007) *Gene Ther.* 14(2):99-107; Eliyahu et al. (2005) *Molecules* 10(1):34-64; and Altaras et al. (2005) *Adv Biochem Eng Biotechnol.* 99:193-260, all of which are hereby incorporated by reference in their entireties.

A FGFR fusion nucleic acid can also be delivered in a controlled release system. For example, the FGFR fusion molecule can be administered using intravenous infusion, an implantable osmotic pump, a transdermal patch, liposomes, or other modes of administration. In one embodiment, a pump can be used (*see* Sefton (1987) *Biomed. Eng.* 14:201; Buchwald et al. (1980) *Surgery* 88:507; Saudek et al. (1989) *N. Engl. J. Med.* 321:574). In another embodiment, polymeric materials can be used (*see* Medical Applications of Controlled Release, Langer and Wise (eds.), CRC Pres., Boca Raton, Fla. (1974); Controlled Drug Bioavailability, Drug Product Design and Performance, Smolen and Ball (eds.), Wiley, New York (1984); Ranger and Peppas, (1983) *J. Macromol. Sci. Rev. Macromol. Chem.* 23:61; *see also* Levy et al. (1985) *Science* 228:190; During et al. (1989) *Ann. Neurol.* 25:351; Howard et al. (1989) *J. Neurosurg.* 71:105). In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target thus requiring only a fraction of the systemic dose (*see*, e.g., Goodson, in Medical Applications of Controlled Release, supra, vol. 2, pp. 115-138 (1984)). Other controlled release systems are discussed in the review by Langer (*Science* (1990) 249:1527-1533).

Pharmaceutical Compositions and Administration for Therapy

An inhibitor of the invention can be incorporated into pharmaceutical compositions suitable for administration, for example the inhibitor and a pharmaceutically acceptable carrier

A FGFR fusion molecule or inhibitor of the invention (e.g. JNJ-42756493) can be administered to the subject once (e.g., as a single injection or deposition). Alternatively, a FGFR fusion molecule or inhibitor can be administered once or twice daily to a subject in need thereof for a period of from about two to about twenty-eight days, or from about seven to about ten days. A FGFR fusion molecule or inhibitor can also be

administered once or twice daily to a subject for a period of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 times per year, or a combination thereof. A FGFR fusion molecule or inhibitor can also be administered in seven to ten day repeating cycles (i.e. administration of a FGFR fusion molecule or inhibitor for seven to ten days, followed by no administration of a FGFR fusion molecule or inhibitor for seven to ten days). Furthermore, a FGFR fusion molecule or inhibitor of the invention can be co-administrated with another therapeutic. Where a dosage regimen comprises multiple administrations, the effective amount of the FGFR fusion molecule or inhibitor administered to the subject can comprise the total amount of gene product administered over the entire dosage regimen.

A FGFR fusion molecule or inhibitor can be administered to a subject by any means suitable for delivering the FGFR fusion molecule or inhibitor to cells of the subject, such as cancer cells, e.g., glioblastoma multiforme, breast cancer, lung cancer, prostate cancer, colorectal carcinoma, bladder carcinoma, squamous lung carcinoma, head and neck carcinoma, glioma, grade II or III glioma, or IDH wild-type grade II or III glioma. For example, a FGFR fusion molecule or inhibitor can be administered by methods suitable to transfect cells. Transfection methods for eukaryotic cells are well known in the art, and include direct injection of the nucleic acid into the nucleus or pronucleus of a cell; electroporation; liposome transfer or transfer mediated by lipophilic materials; receptor mediated nucleic acid delivery, bioballistic or particle acceleration; calcium phosphate precipitation, and transfection mediated by viral vectors.

The compositions of this invention can be formulated and administered to reduce the symptoms associated with a gene fusion-associated cancer, e.g., glioblastoma multiforme, breast cancer, lung cancer, prostate cancer, colorectal carcinoma, bladder carcinoma, squamous lung carcinoma, head and neck carcinoma, glioma, grade II or III glioma, or IDH wild-type grade II or III glioma, by any means that produces contact of the active ingredient with the agent's site of action in the body of a subject, such as a human or animal (e.g., a dog, cat, or horse). They can be administered by any conventional means available for use in conjunction with pharmaceuticals, either as individual therapeutic active ingredients or in a combination of therapeutic active ingredients. They can be administered alone, but are generally administered with a pharmaceutical carrier selected on the basis of the chosen route of administration and standard pharmaceutical practice.

A therapeutically effective dose of FGFR fusion molecule or inhibitor (e.g. JNJ-42756493) can depend upon a number of factors known to those or ordinary skill in the art. The dose(s) of the FGFR fusion molecule inhibitor can vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the a FGFR fusion molecule inhibitor to have upon the nucleic acid or polypeptide of the invention. For example, 12 mg of JNJ-42756493 can be orally administered daily. JNJ-42756493 can be administered in seven to ten day repeating cycles (i.e. administration of JNJ-42756493 for seven to ten days, followed by no administration of JNJ-42756493 for seven to ten days). These amounts can be readily determined by a skilled artisan. Any of the therapeutic

applications described herein can be applied to any subject in need of such therapy, including, for example, a mammal such as a dog, a cat, a cow, a horse, a rabbit, a monkey, a pig, a sheep, a goat, or a human.

Pharmaceutical compositions for use in accordance with the invention can be formulated in conventional manner using one or more physiologically acceptable carriers or excipients. The therapeutic compositions of the invention can be formulated for a variety of routes of administration, including systemic and topical or localized administration. Techniques and formulations generally can be found in Remmington's Pharmaceutical Sciences, Meade Publishing Co., Easton, Pa (20^{th} Ed., 2000), the entire disclosure of which is herein incorporated by reference. For systemic administration, an injection is useful, including intramuscular, intravenous, intraperitoneal, and subcutaneous. For injection, the therapeutic compositions of the invention can be formulated in liquid solutions, for example in physiologically compatible buffers such as Hank's solution or Ringer's solution. In addition, the therapeutic compositions can be formulated in solid form and redissolved or suspended immediately prior to use. Lyophilized forms are also included. Pharmaceutical compositions of the present invention are characterized as being at least sterile and pyrogen-free. These pharmaceutical formulations include formulations for human and veterinary use.

According to the invention, a pharmaceutically acceptable carrier can comprise any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Any conventional media or agent that is compatible with the active compound can be used. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition containing FGFR fusion molecule inhibitor can be administered in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed herein. Such pharmaceutical compositions can comprise, for example antibodies directed to a FGFR fusion molecule, or a variant thereof, or antagonists of a FGFR fusion molecule, or JNJ-42756493. The compositions can be administered alone or in combination with at least one other agent, such as a stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

Sterile injectable solutions can be prepared by incorporating the FGFR fusion molecule inhibitor (e.g., a polypeptide or antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated herein, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated herein. In the case of sterile powders for the preparation of sterile injectable solutions, examples of useful preparation methods are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

In some embodiments, the FGFR fusion molecule inhibitor can be applied via transdermal delivery systems, which slowly releases the active compound for percutaneous absorption. Permeation enhancers can be used to facilitate transdermal penetration of the active factors in the conditioned media. Transdermal patches are described in for example, U.S. Pat. No. 5,407,713; U.S. Pat. No. 5,352,456; U.S. Pat. No. 5,332,213; U.S. Pat. No. 5,336,168; U.S. Pat. No. 5,290,561; U.S. Pat. No. 5,254,346; U.S. Pat. No. 5,164,189; U.S. Pat. No. 5,163,899; U.S. Pat. No. 5,088,977; U.S. Pat. No. 5,087,240; U.S. Pat. No. 5,008,110; and U.S. Pat. No. 4,921,475.

"Subcutaneous" administration can refer to administration just beneath the skin (i.e., beneath the dermis). Generally, the subcutaneous tissue is a layer of fat and connective tissue that houses larger blood vessels and nerves. The size of this layer varies throughout the body and from person to person. The interface between the subcutaneous and muscle layers can be encompassed by subcutaneous administration. This mode of administration can be feasible where the subcutaneous layer is sufficiently thin so that the factors present in the compositions can migrate or diffuse from the locus of administration. Thus, where intradermal administration is utilized, the bolus of composition administered is localized proximate to the subcutaneous layer.

Administration of the cell aggregates (such as DP or DS aggregates) is not restricted to a single route, but can encompass administration by multiple routes. For instance, exemplary administrations by multiple routes include, among others, a combination of intradermal and intramuscular administration, or intradermal and subcutaneous administration. Multiple administrations can be sequential or concurrent. Other modes of application by multiple routes will be apparent to the skilled artisan.

In other embodiments, this implantation method will be a one-time treatment for some subjects. In further embodiments of the invention, multiple cell therapy implantations will be required. In some embodiments, the cells used for implantation will generally be subject-specific genetically engineered cells. In another embodiment, cells obtained from a different species or another individual of the same species can be used. Thus, using such cells can require administering an immunosuppressant to prevent rejection of the implanted cells. Such methods have also been described in United States Patent No. 7,419,661and PCT application publication WO 2001/32840, and are hereby incorporated by reference.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation or ingestion), transdermal (topical), transmucosal, and rectal administration. For example, JNJ-42756493 can be orally administered. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral

preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic. Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, a pharmaceutically acceptable polyol like glycerol, propylene glycol, liquid polyetheylene glycol, and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it can be useful to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the inhibitor (e.g., a polypeptide or antibody or small molecule) of the invention in the required amount in an appropriate solvent with one or a combination of ingredients enumerated herein, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated herein. In the case of sterile powders for the preparation of sterile injectable solutions, examples of useful preparation methods are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions (e.g. of JNJ-42756493) generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier and subsequently swallowed.

Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal

administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

In some embodiments, the effective amount of the administered FGFR fusion molecule inhibitor (e.g. JNJ-42756493) is at least about 0.0001 μg/kg body weight, at least about 0.00025 μg/kg body weight, at least about 0.0005 μg/kg body weight, at least about 0.00075 μg/kg body weight, at least about 0.001 μg/kg body weight, at least about 0.0025 μg/kg body weight, at least about 0.005 μg/kg body weight, at least about 0.0075 μg/kg body weight, at least about 0.01 μg/kg body weight, at least about 0.025 μg/kg body weight, at least about 0.05 μg/kg body weight, at least about 0.075 μg/kg body weight, at least about 0.1 μg/kg body weight, at least about 0.25 μg/kg body weight, at least about 0.5 μg/kg body weight, at least about 0.75 μg/kg body weight, at least about 1 μg/kg body weight, at least about 5 μg/kg body weight, at least about 10 μg/kg body weight, at least about 25 μg/kg body weight, at least about 50 μg/kg body weight, at least about 75 μg/kg body weight, at least about 100 μg/kg body weight, at least about 150 μg/kg body weight, at least about 200 μg/kg body weight, at least about 250 μg/kg body weight, at least about 300 μg/kg body weight, at least about 350 μg/kg body weight, at least about 400 μg/kg body weight, at least about 450 μg/kg body weight, at least about 500 μg/kg body weight, at least about 550 μg/kg body weight, at least about 600 μg/kg body weight, at least about 650 μg/kg body weight, at least about 700 μg/kg body weight, at least about 750 μg/kg body weight, at least about 800 μg/kg body weight, at least about 850 μg/kg body weight, at least about 900 μg/kg body weight, at least about 950 μg/kg body weight, at least about 1,000 μg/kg body weight, at least about 2,000 μg/kg body weight, at least about 3,000 μg/kg body weight, at least about 4,000 μg/kg body weight, at least about 5,000 μg/kg body weight, at least about 6,000 μg/kg body weight, at least about 7,000 μg/kg body weight, at least about 8,000 μg/kg body weight, at least about 9,500 μg/kg body weight, or at least about 10,000 μg/kg body weight.

In some embodiments, the effective amount of the administered FGFR fusion molecule inhibitor (e.g. JNJ-42756493) is at least about 1mg, 2mg, 3mg, 4mg, 5mg, 6mg, 7mg, 8mg, 9mg, 10mg, 11mg, 12mg, 14mg, 15 mg, 16 mg, 17 mg, 18 mg, 19 mg, 20 mg, 21 mg, 22 mg, 23 mg, 24 mg, 25 mg, 26 mg, 27 mg, 28 mg, 29 mg, 30 mg.

***

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Exemplary methods and materials are described below, although methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present invention.

All publications and other references mentioned herein are incorporated by reference in their entirety, as if each individual publication or reference were specifically and individually indicated to be incorporated by reference. Publications and references cited herein are not admitted to be prior art.

EXAMPLES

Examples are provided below to facilitate a more complete understanding of the invention. The following examples illustrate the exemplary modes of making and practicing the invention. However, the scope of the invention is not limited to specific embodiments disclosed in these Examples, which are for purposes of illustration only, since alternative methods can be utilized to obtain similar results.

The invention is further illustrated in Singh *et al.*, *Science* (2012), 337(6099):1231-5 (including the accompanying Supplementary Information). The entire contents of Singh *et al.*, *Science* (2012), 337(6099):1231-5, including the accompanying "Supplementary Information," is expressly incorporated by reference. The invention is also further illustrated in Di Stefano et al., "Detection, characterization and inhibition of FGFR-TACC fusions in IDH wild type glioma" Clin. Cancer Res. (2015), the entire contents of which are expressly incorporated by reference, including the accompanying Supplementary Information.

Example 1: *Transforming and recurrent fusions of FGFR and TACC gene in glioblastoma*

The history of successful targeted therapy of cancer largely coincides with the inactivation of recurrent, oncogenic and addicting gene fusions in hematological malignancies and recently in some types of epithelial cancer. Glioblastoma multiforme (GBM) is among the most lethal forms of human cancer. Here, an integrated gene fusion discovery pipeline was developed for the detection of in-frame fused transcripts from RNA-seq and genomic fusions from whole exome sequences. The application of the pipeline to human GBM unraveled recurrent chromosomal translocations, which fuse in-frame the tyrosine kinase domain of FGFR genes (FGFR1 or FGFR3) to the TACC domain of TACC1 or TACC3, respectively. The frequency of FGFR-TACC fusions is 3 of 97 GBM (3.1%). The FGFR-TACC fusion protein displays strong oncogenic activity when introduced into astrocytes or transduced by lentivirus-mediated stereotactic delivery to the adult mouse brain. The FGFR-TACC fusion protein mis-localizes over the mitotic spindle pole, has constitutive tyrosine kinase activity and dysregulates the mitotic cycle with delayed mitotic progression. The impaired mitotic fidelity triggers chromatid cohesion defects, defective spindle checkpoint activation, chromosomal mis-segregation, and rampant aneuploidy. Inhibition of FGFR kinase corrects the aneuploidy and oral administration of a specific FGFR tyrosine kinase inhibitor under clinical investigation arrests tumor growth and prolongs survival of mice harboring intracranial FGFR3-TACC3-initiated glioma. FGFR-TACC fusions identify a subset of GBM patients who may benefit from targeted inhibition of the tyrosine kinase activity of FGFR.

Glioblastoma multiforme (GBM) is among the most difficult forms of cancer to treat in humans (Ohgaki and Kleihues, 2005). So far, the targeted therapeutic approaches that have been tested against potentially important oncogenic drivers in GBM have met limited success (Lo, 2010; Reardon et al., 2010; van den Bent et al., 2009). Recurrent chromosomal translocations leading to production of oncogenic fusion proteins are viewed as initiating and addicting events in the pathogenesis of human cancer, thus providing the most desirable molecular targets for cancer therapy (Ablain et al., 2011; Mitelman et al., 2007). Chromosomal rearrangements resulting in recurrent and oncogenic gene fusions are hallmarks of hematological malignancies and recently they have also been uncovered in subsets of solid tumors (breast, prostate, lung

and colorectal carcinoma), but they have not been found in GBM (Bass et al., 2011; Prensner and Chinnaiyan, 2009). Important and successful targeted therapeutic interventions for patients whose tumors carry these rearrangements have stemmed from the discovery of functional gene fusions, especially when the translocations involve kinase-coding genes (*BCR-ABL*, *EML4-ALK*) (Druker, 2009; Gerber and Minna, 2010).

A hallmark of GBM is rampant chromosomal instability (CIN), which leads to aneuploidy (Furnari et al., 2007). CIN and aneuploidy are early events in the pathogenesis of cancer (Cahill et al., 1999). It has been suggested that genetic alterations targeting mitotic fidelity might be responsible for mis-segregation of chromosomes during mitosis, resulting in aneuploidy (Gordon et al., 2012; Solomon et al., 2011). Here, the first cases of recurrent and oncogenic gene fusions in human GBM are described. The resulting fusion protein localizes to mitotic cells, disrupts the normal control of chromosome segregation and induces aneuploidy. A therapeutic strategy with FGFR tyrosine kinase inhibitors is also reported for the targeted therapy of GBM patients harboring these chromosomal rearrangements.

Identification of recurrent fusions of FGFR and TACC genes. To identify genomic rearrangements in GBM that generate functional fusion proteins and are recurrent, gene pairs discovered as in-frame fused transcripts from the analysis of massively parallel, paired-end sequencing of expressed transcripts (RNA-seq) would also emerge as fused gene pairs from the genomic analysis of human GBM. Towards this aim, two complementary gene fusion discovery methods were devised and were applied to two GBM cohorts. The first, TX-Fuse, is an algorithm for the discovery of candidate fusion transcripts from RNA-seq (**Figure 8**). The second, Exome-Fuse, detects fusion genes from whole exome DNA sequences (**Figure 8**). As first step for the detection of fused transcripts, RNA-seq data was generated from short-term cultures of glioma stem-like cells (GSCs) freshly isolated from nine patients carrying primary GBM. The culture of primary GBM tumors under serum-free conditions selects cells that retain phenotypes and genotypes closely mirroring primary tumor profiles as compared to serum-cultured glioma cell lines that have largely lost their developmental identities (Lee et al., 2006). Therefore, without being bound by theory, if glioma cells carry gene fusions causally responsible for the most aggressive hallmarks of GBM, they should be selected in GSCs. RNA-seq generated an average of 60.3 million paired reads for each GSC culture, of which over 80% were mapped to the reference transcriptome and genome. TX-Fuse detects two main sources of evidence: split reads and split inserts (see Experimental Procedures). The application of TX-Fuse to the RNA-seq dataset from nine GSCs led to the discovery of five candidate rearrangements (all of which were intrachromosomal) that give rise to in-frame fusion transcripts (Table 1).

# Split Inserts	# Split Reads	Sample	Gene1	Gene2	Ref Seq1	Ref Seq2	Tx Pos1	Tx Pos2
294	76	GSC-1123	FGFR3	TACC3	NM_000142	NM_006342	2530	1751
37	54	GSC-	POLR2A	WRAP53	NM_000937	NM_001143990	479	798

		0114						
7	48	GSC-0114	CAPZB	UBR4	NM_001206540	NM_020765	228	12111
8	29	GSC-0517	ST8SIA4	PAM	NM_005668	NM_000919	1125	730
6	17	GSC-0308	PIGU	NCOA6	NM_080476	NM_014071	729	6471
1	6	GSC-0127	IFNAR2	IL10RB	NM_000874	NM_000628	1083	149

Table 1: Predicted in-frame fusion proteins from RNA-Seq of nine GSCs

#Split Inserts	#Split Reads	Sample	Chr 1	Strand 1	hg19_ GenPos1	Chr 2	Strand 2	hg19_ GenPos2
294	76	GSC-1123	4	+	1808842	4	+	1737004
37	54	GSC-0114	17	+	7399259	17	+	7604059
7	48	GSC-0114	1	-	19712098	1	-	19433440
8	29	GSC-0517	5	-	100147809	5	+	102260661
6	17	GSC-0308	20	-	33203914	20	-	33303130
1	6	GSC-0127	21	+	34632901	21	+	34640699

Next, genomic rearrangements leading to gene fusions were identified in GBM by applying Exome-Fuse to a dataset of paired-end exome DNA sequences from 84 GBM samples from TCGA (Table 2).

This analysis detected 147 paired gene fusions, thus producing an average of 1.75 gene fusion events per tumor (Table 3).

The FGFR and TACC families of genes were markedly enriched among those recurrently involved in genomic fusions, with eight tumors harboring FGFR rearrangements and seven tumors harboring fusions that implicate TACC genes (**Figure 1A**). The comparative analysis of the TX-Fuse and Exon-Fuse outputs revealed that FGFR3-TACC3 was the only fusion pair identified as either an in-frame transcript by TX-Fuse and genomic fusions by Exome-Fuse (Tables 1, 2 and 3).

A list of recurrent gene fusions identified by Exome-fuse analysis of 84 GBM from TCGA can be found in Table 2 of WO2014/018673, the contents of which is hereby incorporated by reference in it entirety. As multiple junctions may exist in each fusion candidate, information for all breakpoints is displayed. Column definitions include: sample = TCGA sample ID, virtForSplitReads/virtRevSplitReads/virtTotSplitReads = # forward/reverse/total split reads, splitInserts = # split inserts, dirA/dirB = forward (1) or reverse (0) direction of split read portion mapping to gene A/B, dirAB_matepair = direction of mate pair of split read, cosmicA+B = # recorded mutations of gene A+B in COSMIC.

Table 3 Recurrent gene fusion pairs from Exome-fuse analysis of 84 GBM from TCGA.

Sample	gene A	gene B	Sample	gene A	gene B
TCGA-12-0820	ABCA13	NHSL2	TCGA-12-0820	CAMKK1	FAM184B
TCGA-12-1089	ABCC1	RNF216	TCGA-12-1089	CAMSAP1	NCF2
TCGA-12-1088	ABCC1	AGBL4	TCGA-12-1088	CAMTA1	TMPRSS3
TCGA-12-0827	ABCC12	FGFR4	TCGA-06-1801	CAMTA1	GDPD2
TCGA-12-0829	ABCC6	SUMF1	TCGA-06-1801	CAP2	DNAJC6
TCGA-06-1801	ABCC6	CMTM7	TCGA-19-0957	CCDC147	STK4
TCGA-12-1088	ABL1	TNFRSF10B	TCGA-12-0829	CCDC147	ISX
TCGA-06-1802	ADAM12	PTPRD	TCGA-06-0166	CCDC158	SNX5
TCGA-12-0829	ADAM12	DAPK1	TCGA-19-0957	CDH11	RERE
TCGA-12-1088	ADCY10	DUSP27	TCGA-06-1802	CELF2	PLA2G2F
TCGA-19-0957	ADCY10	AKT3	TCGA-12-0826	CELF2	NME4
TCGA-12-0828	ADCY2	SDAD1	TCGA-12-1600	CILP	PARP16
TCGA-12-0829	ADCY2	C14orf174	TCGA-12-1089	CLK3	LRP1
TCGA-12-0829	ADCY8	SSX3	TCGA-12-1088	CMYA5	STK24
TCGA-12-0829	AGBL4	NUP188	TCGA-27-1835	CMYA5	SRRM1
TCGA-06-1805	AGBL4	NOX4	TCGA-12-1092	CNTN2	DNAJC6
TCGA-12-1089	AHNAK	C21orf29	TCGA-06-1805	COL14A1	NCRNA00157
TCGA-12-0829	ANXA7	CACNA1C	TCGA-12-0829	COL14A1	MMP12
TCGA-06-1801	AP4S1	EYS	TCGA-12-1093	CORO7	PLEK2
TCGA-12-0828	AQP2	ECE1	TCGA-12-0829	CORO7	DYRK3
TCGA-19-0957	AQP2	CDH4	TCGA-06-1801	CROCC	CSMD2
TCGA-19-1790	ARMC6	FAM184B	TCGA-19-0957	CSMD2	MDH2
TCGA-19-1786	ATP5B	USP48	TCGA-06-1805	CUL3	SLC44A2
TCGA-12-0829	ATP5B	PRC1	TCGA-12-0827	CUL3	LY75
TCGA-12-1600	ATP6AP1L	FAM172A	TCGA-12-0829	DDI2	KIDINS220
TCGA-12-0829	ATP6V0D2	RERE	TCGA-19-5958	DEPDC5	SLC5A4
TCGA-12-0829	ATXN1	CACNA1G	TCGA-06-1801	DEPDC5	ROBO1
TCGA-06-1802	BAHD1	OSBPL10	TCGA-06-1801	DEPDC7	EIF2C2
TCGA-12-0820	BBX	CUL3	TCGA-12-0829	DIS3L	GLI3
TCGA-19-2621	BCAS3	TTYH1	TCGA-12-0829	DMD	N4BP2L2
TCGA-12-1088	BCAS3	CACNA1G	TCGA-06-1802	DNM1L	SYNPO2
TCGA-06-1801	C15orf23	DMD	TCGA-12-3644	EDA	SSX3
TCGA-12-3644	C15orf33	PARP16	TCGA-12-3644	EFHC1	LRBA
TCGA-12-0829	C21orf29	MYT1	TCGA-12-0829	EFS	NRXN3
TCGA-06-1805	C2CD3	XRRA1	TCGA-06-1802	EIF2C2	TNFRSF10B
TCGA-12-0829	C2CD3	CAPZB	TCGA-19-0957	EML1	NRXN3
TCGA-12-1089	C6orf170	NKAIN2	TCGA-12-0829	ENTPD2	FREM2
TCGA-12-0822	C7orf44	TACC2	TCGA-12-0829	EYS	IL1RN
TCGA-12-1089	CACNA1C	ITGAV	TCGA-14-0781	FAM160A1	UNC93B1
TCGA-12-0829	CACNA1G	CNTNAP4	TCGA-12-0829	FAM160A1	LY75

Table 3 above shows recurrent gene fusion pairs from Exome-fuse analysis of 84 GBM from TCGA. Fusion candidates have been nominated if they have at least two split inserts and at least two split reads. To further filter the list on recurrence, any fusion candidate was kept in which one of the genes is involved in at least two fusions across different samples.

To experimentally validate the computational predictions that emerged from TX-Fuse, the PCR products

spanning the fusion breakpoint were sequenced and validated each of the five in-frame fusion predictions (**Figures 1A-D** and **9A-D**). FGFR3-TACC3 gene fusions were identified by whole transcriptome sequencing of GSCs. SEQ ID NOs: 2-77 disclose 76 split-reads . SEQ ID NO: 1 discloses the predicted reading frame at the breakpoint (2530bp of NM_000142 fused to 1751bp of NM_006342), with the FGFR3 sequence (nucleotides 1-88) preceding the TACC3 sequence (nucleotides 89-178). SEQ ID NO: 78 discloses the putative amino acid sequence corresponding to SEQ ID NO: 1. Amino acids 1-29 of SEQ ID NO: 78 correspond to the FGFR3 sequence and amino acids 30-59 correspond to the TACC3 sequence. In **Figure 1B**, the cDNA sequence validation for the fusion with the highest read support involving FGFR3 fused in-frame with TACC3 in GSC-1123 is shown. The same FGFR3-TACC3 fusion transcript was also detected in the primary GBM-1123 tumor specimen from which the GSC-1123 culture was established (**Figure 1B**). The amplified cDNA contained an open reading frame for a protein of 1,048 amino acids resulting from the fusion of a FGFR3 amino-terminal portion of residues 1-758 with a TACC3 carboxy-terminal portion of residues 549-838 (**Figure 1C**). SEQ ID NO: 79 discloses the amino acid sequence of the FGFR3-TACC3 protein. Amino acids 1-758 of SEQ ID NO: 79 correspond to FGFR3 and amino acids 759-1048 of SEQ ID NO: 79 correspond to TACC3. FGFR3 is a member of the FGFR receptor tyrosine kinase (TK) family that transduces intracellular signals after binding to FGF ligands (Turner and Grose, 2010). TACC3 belongs to the evolutionarily conserved TACC gene family, which also includes TACC1 and TACC2. The distinctive feature of TACC proteins is the presence of a coiled-coil domain at the C-terminus, known as the TACC domain. Through the TACC domain, TACC proteins localize to the mitotic spindle during metaphase and stabilize the microtubule spindle network (Hood and Royle, 2011; Peset and Vernos, 2008). In the predicted fusion protein the intracellular TK domain of FGFR3 is fused upstream of the TACC domain of TACC3 (**Figure 1C**).

Exon-specific gene expression analysis from the RNA-seq coverage in GSC-1123 demonstrated that the FGFR3 and TACC3 exons implicated in the fusion are highly overexpressed compared with the mRNA sequences not included in the fusion event (**Figure 10A**). Quantitative RT-PCR showed that the expression of the fused FGFR3-TACC3 exons is significantly higher in GSC-1123 than other GSCs and the normal brain (80 to 130-fold, **Figure 10B**). Without being bound by theory, functionally significant genetic rearrangements may result in marked overexpression (outlier) of the genes implicated in the fusion events (Tomlins et al., 2007; Tomlins et al., 2005). The FGFR3-TACC3 fusion protein was also abundantly expressed in GSC-1123 and in the primary tumor GBM-1123, as shown by Western blot and immunohistochemistry (**Figures 10C and 10D**). On a Western Blot, the FGFR3-TACC3 fusion protein migrated at a size of ~150 kD and immunoprecipitation followed by mass spectrometry revealed the presence of FGFR3 and TACC3 peptides consistent with the cDNA translation prediction (**Figure 10E1-E6**). The position of the peptides from **FIG. 10E1-E6** in the amino acid sequence of the FGFR3-TACC3 fusion protein (SEQ ID NO: 79) are 77-93 and 298-310 (FGFR3) and 773-778, 985-996, and 999-1011 (TACC3).

Using PCR, the genomic breakpoint coordinates were mapped to chromosome 4 (#1,808,966 for FGFR3 and

#1,737,080 for TACC3, genome build GRCh37/hg19) falling within FGFR3 exon 17 and TACC3 intron 7, which gives rise to a transcript in which the 5' FGFR3 exon 16 is spliced to the 3' TACC3 exon 8. The DNA junctions of FGFR3 and TACC3 show microhomology within a 10-base region, an observation consistent with results previously reported for other chromosomal rearrangements in human cancer (Bass et al., 2011; Stephens et al., 2009) (**Figure 1D**).

The experimental validation of the inferred genomic fusions was focused on FGFR3-TACC3. Exome-Fuse identified FGFR3-TACC3 gene fusions in four GBM samples with breakpoints spanning invariably within intron 16 of FGFR3 (which is downstream to the coding region for the TK domain) and intron 7-10 of TACC3 (which is upstream to the TACC domain) (Tables 4 and 5).

For sample TCGA-27-1835, SEQ ID NOS: 96-107 disclose split-reads of the genomic breakpoints of *FGFR3* within intron 16 and *TACC3* genes within intron 10. SEQ ID NO: 95 discloses the reading frame at the breakpoint (1778595 of FGFR3 fused to 1709397 of TACC3), with the FGFR3 sequence (nucleotides 1-76) preceding the TACC3 sequence (nucleotides 77-152).

For sample TCGA-19-5958, SEQ ID NOS: 109-111 disclose split-reads of the genomic breakpoints of *FGFR3* within intron 16 and *TACC3* genes within intron 7. SEQ ID NO: 108 discloses the reading frame at the breakpoint (1778539 of FGFR3 fused to 1707202 of TACC3), with the FGFR3 sequence (nucleotides 1-76) preceding the TACC3 sequence (nucleotides 77-153).

For sample TCGA-06-6390, SEQ ID NOS: 113-131 disclose split-reads of the genomic breakpoints of *FGFR3* within intron 16 and *TACC3* genes within exon 9. SEQ ID NO: 112 discloses the reading frame at the breakpoint (1778521 of FGFR3 fused to 1708787 of TACC3), with the FGFR3 sequence (nucleotides 1-76) preceding the TACC3 sequence (nucleotides 77-152).

For sample TCGA-12-0826, SEQ ID NOS: 133-145 disclose split-reads of the genomic breakpoints of *FGFR3* within intron 16 and *TACC3* genes within intron 7. SEQ ID NO: 132 discloses the reading frame at the breakpoint (1778502 of FGFR3 fused to 1707185 of TACC3), with the FGFR3 sequence (nucleotides 1-75) preceding the TACC3 sequence (nucleotides 76-151).

Among the four positive TCGA GBM specimens, two were available from TCGA centers for molecular analysis (TCGA-27-1835 and TCGA-06-6390) and, by Sanger sequencing, each of them were confirmed to carry an in-frame fusion transcript that is consistent with the predicted genomic breakpoints (**Figures 2A** and **2B**). Thus, the frames of the FGFR3-TACC3 fusion proteins invariably result in juxtaposing the TK domain of FGFR3 upstream of the TACC domain of TACC3. Consistent with the abundant expression of FGFR3-TACC3 in GSC-1123 and GBM-1123, the mRNA expression analysis of the TCGA tumors revealed that the four FGFR3-TACC3-positive GBM display marked co-outlier expression of FGFR3 and TACC3 (**Figure 2C**). Recurrent gene fusions can be associated with local copy number variations (CNV) of the breakpoint regions (Wang et al., 2009). Accordingly, the analysis of SNP arrays in the TCGA dataset revealed the presence of microamplification events of the FGFR3 and TACC3 genes in all four FGFR3-TACC3-positive GBM (**Figure 2D**).

SEQ ID NOS 187-224 disclose the split inserts supporting the identification of FGFR3-TACC3 fusion

genes in four GBM samples from the ATLAS-TCGA exome collection (TCGA-12-0826; TCGA-06-6390; TCGA-27-1835; TCGA-19-5958).

SEQ ID NOS 225-318 disclose the split reads supporting the identification of FGFR3-TACC3 fusion genes in four GBM samples from the ATLAS-TCGA exome collection (TCGA-12-0826; TCGA-06-6390; TCGA-27-1835; TCGA-19-5958).

The FGFR3 and TACC3 genes are located 48-Kb apart on human chromosome 4p16. The other members of the FGFR and TACC families retain the close physical association of FGFR3 and TACC3, with FGFR1 and TACC1 paired on chromosome 8p11 and FGFR2 and TACC2 paired on chromosome 10q26. Without being bound by theory, the ancestral FGFR and TACC genes were physically linked and that this tandem gene cluster was duplicated at least twice to generate the FGFR1-TACC1, FGFR2-TACC2 and FGFR3-TACC3 pairs that mark mammalian evolution (Still et al., 1999). The highly conserved TK domains among FGFR genes and TACC domains among TACC genes together with their invariable fusion in the FGFR3-TACC3 rearrangements prompted to ask whether other intra-chromosomal FGFR-TACC fusion combinations exist in human GBM.

cDNA from a panel of 88 primary GBM were screened using pairs of upstream PCR primers that bind the amino-terminal coding region of the TK domains of FGFR1, FGFR2 and FGFR3 and downstream primers that bind to the carboxy-terminal coding region of the TACC domains of TACC1, TACC2 and TACC3 genes, respectively. The screening resulted in the identification of intrachromosomal FGFR-TACC fusions in two additional cases (one harboring FGFR1-TACC1 and one FGFR3-TACC3), corresponding to three of 97 total GBM (3.1%), including the GBM-1123 case. The FGFR1-TACC1 fusion breakpoint in GBM-51 joined in-frame exon 17 of FGFR1 to exon 7 of TACC1, resulting in a novel protein in which the TK domain of FGFR1 is fused upstream of the TACC domain of TACC1 (**Figure 2E**). The same structure was conserved again in GBM-22 in which exon 16 of FGFR3 is joined in-frame to exon 10 of TACC3 (**Figure 2F**). None of the tumors harboring FGFR-TACC fusions had mutations in IDH1 or IDH2 genes, thus indicating that FGFR-TACC-positive GBM mark an independent subgroup of patients from those carrying IDH mutations (Table 6) (Yan et al., 2009). The constant linkage of the FGFR-TK to the TACC domain created in each of the seven GBM harboring FGFR-TACC rearrangements suggests that FGFR-TACC fusion proteins may generate important functional consequences for oncogenesis in the brain.

Table 6.

Samples	Type	Time	Status	Age at initial pathologic diagnosis	IDH1-2 status (Sanger)	IDH1-2 status (exome)
TCGA-12-0826	FGFR3-TACC3	845	DECEASED	38	WT	WT
TCGA-27-1835	FGFR3-TACC3	648	DECEASED	53	NA	WT
TCGA-19-5958	FGFR3-TACC3	164	LIVING	56	NA	WT
TCGA-06-6390	FGFR3-TACC3	163	DECEASED	58	WT	WT
GBM-22	FGFR3-TACC3	390	DECEASED	60	WT	NA
GBM-1123	FGFR3-TACC3	NA	DECEASED	62	WT	NA
GBM-51	FGFR1-TACC1	NA	NA	NA	WT	NA

Time = Survival (days after diagnosis)
Sanger = analysis done by Sanger sequencing of genomic DNA
Exome = alnalysis done by the SAVI (Statistical Algorithm for Variant Identification), an algorithm developed to detect point mutation in cancer (BRAF Mutations in Hairy-Cell Leukemia, Tiacci E et al. The New England Journal of Medicine 2011 Jun 16;364(24):2305-15)
NA = Not Available
WT = Wild type sequence for R132 and R172 of IDH1 and IDH2, respectively

Transforming activity of FGFR-TACC fusions. To test the functional importance of the FGFR-TACC fusions in GBM, the FGFR3-TACC3 cDNA was cloned from GSC-1123 and recombinant lentiviruses were prepared expressing FGFR3-TACC3, FGFR1-TACC1, a kinase-dead FGFR3-TACC3 protein (FGFR3-TACC3-K508M), *wild type* FGFR3 and *wild type* TACC3. Transduction of Rat1A fibroblasts and *Ink4A;Arf-/-* astrocytes with the FGFR3-TACC3 lentivirus resulted in the expression of the fusion protein at levels comparable to those present in GSC-1123 (**Figure 11A-C**). Having reconstituted in non-transformed cells the endogenous level of the FGFR-TACC protein that accumulates in GBM cells, it was determined whether it was sufficient to initiate oncogenic transformation *in vitro* and *in vivo*. Rat1A cells expressing FGFR3-TACC3 and FGFR1-TACC1 but not those expressing FGFR3-TACC3-K508M, FGFR3, TACC3 or the empty lentivirus acquired the ability to grow in anchorage-independent conditions in soft agar (**Figure 3A**). Transduction of the same lentiviruses in primary *Ink4A;Arf-/-* astrocytes followed by subcutaneous injection into immunodeficient mice revealed that only astrocytes expressing FGFR3-TACC3 and FGFR1-TACC1 formed tumors. The tumors emerged in 100% of the mice injected with astrocytes expressing the fusion proteins and were glioma-like lesions with strong positivity for Ki67, phospho-histone H3, nestin, GFAP and Olig2 (**Figure 3B**).

Next, it was determined whether the FGFR3-TACC3 fusion protein is oncogenic when transduced to a small number of cells directly into the brain of immunocompetent animals. A recently described mouse glioma model was used in which brain tumors are initiated by lentiviral transduction of oncogenes and inactivation of p53 in the mouse brain (Marumoto et al., 2009). To target adult NSCs, the adult mouse hippocampus was stereotactically transduced with purified lentivirus expressing the FGFR3-TACC3 protein and shRNA against p53 (pTomo-FGFR3-TACC3-shp53). Seven of eight mice (87.5%) transduced with FGFR3-TACC3 succumbed from malignant brain tumors within 240 days (**Figure 3C**). None of the mice transduced with a lentivirus expressing the most frequent gain-of-function mutation in GBM (the constitutively active EGFRvIII, pTomo-EGFRvIII-shp53) or the pTomo-shp53 control lentivirus died or developed clinical signs of brain tumors (**Figure 3C**). The FGFR3-TACC3 tumors were high-grade glioma with strong propensity to invade the normal brain and stained positive for the glioma stem cell markers nestin and Olig2 and the glial marker GFAP. They were also highly positive for Ki67 and phospho-histone H3, thus displaying rapid

tumor growth (**Figure 3D**). The expression of FGFR3-TACC3 in the xenograft and intracranial tumor models was comparable to the expression of the endogenous protein in the human GSCs and tumor (**Figure 11D, 11E and 11F**).

These data show that FGFR-TACC fusion proteins possess transforming activity in two independent cellular models and this activity is not the result of the overexpression of individual FGFR and TACC genes. They also show that direct transduction of the FGFR3-TACC3 protein to the adult mouse brain leads to efficient development of malignant glioma.

The FGFR-TACC fusions interfere with mitotic progression and induce chromosome missegregation and aneuploidy. To elucidate the mechanism by which the FGFR-TACC fusion drives oncogenesis, it was explored whether it activates downstream FGFR signaling. FGFR3-TACC3 failed to hyperactivate the canonical signaling events downstream of FGFR (pERK and pAKT) in the presence or absence of the ligands FGF-1, FGF-2 or FGF-8 (Wesche et al., 2011). However, FGFR3-TACC3 displayed constitutive phosphorylation of its TK domain and the adaptor protein FRS2, both of which were abolished by the specific inhibitor of FGFR-associated TK activity PD173074 (Mohammadi et al., 1998) or the K508M mutation (**Figure 4A**). Thus, FGFR3-TACC3 gains constitutive kinase activity that is essential for oncogenic transformation but the downstream signaling of this aberrant activity is distinct from the canonical signaling events downstream to FGFR. By driving the localization of the fusion protein, the TACC domain can create entirely novel TK-dependent functions. The TACC domain is essential for the localization of TACC proteins to the mitotic spindle (Hood and Royle, 2011; Peset and Vernos, 2008). Confocal imaging showed that FGFR3-TACC3 designed an arc-shaped structure bending over and encasing the metaphase spindle poles, frequently displaying asymmetry towards one of the two poles and relocated to the midbody as cells progressed into the late stages of mitosis (telophase and cytokinesis). Conversely, the localization of TACC3 was restricted to spindle microtubules and did not relocalize to the midbody . Wild type FGFR3 lacked discrete localization patterns in mitosis.

The mitotic localization of FGFR3-TACC3 indicates that it may impact the fidelity of mitosis and perturb the accurate delivery of the diploid chromosomal content to daughter cells, thus generating aneuploidy. Mitotic progression of individual cells was examined in vector-transduced and FGFR3-TACC3 expressing cells co-expressing histone H2B-GFP by time-lapse microscopy. The average time from nuclear envelope breakdown to anaphase onset was increased in cells expressing FGFR3-TACC3 in comparison with control cells. The mitotic delay was further exacerbated by difficulties in completing cytokinesis .

Next, it was determined whether the expressions of the FGFR-TACC fusion proteins induce defects of chromosomal segregation. Quantitative analyses of mitoses revealed that cells expressing FGFR3-TACC3 or FGFR1-TACC1 exhibit a three to five fold increase of chromosomal segregation errors than control cells. The most frequent mitotic aberrations triggered by the fusion proteins were misaligned chromosomes during metaphase, lagging chromosomes at anaphase and chromosome bridges that impaired cytokinesis and generated micronuclei in the daughter cells (**Figure 4B**). Aberrations at the metaphase-anaphase transition frequently lead to the inability of mitotic cells to maintain a metaphase arrest after treatment with a spindle

poison. Over 18% of cells expressing FGFR3-TACC3 displayed prematurely separated sister chromatids in contrast with less than 3% in control, FGFR3 or TACC3-expressing cells. Accordingly, cells expressing the fusion protein were unable to efficiently arrest in metaphase after nocodazole treatment.

The above findings indicate that expression of the FGFR3-TACC3 fusion protein may spark aneuploidy. Karyotype analysis revealed that FGFR3-TACC3 increased over 2.5 fold the percent of aneuploidy and led to the accumulation of cells with broad distribution of chromosome counts in comparison with cells transduced with empty vector, FGFR3 or TACC3 (**Figure 5A**). Accordingly, GSC-1123 contained aneuploid modal number of chromosomes (49) and manifested a broad distribution of chromosome counts characterized by 60% of metaphase spreads that deviate from the mode.

Next, it was determined whether aneuploidy is a direct consequence of FGFR3-TACC3 expression and is induced in human diploid neural cells. Primary human astrocytes analyzed six days after transduction with the FGFR3-TACC3 lentivirus exhibited a 5-fold increase of the rate of aneuploidy and a significantly wider distribution of chromosome counts (**Figures 5B, 5C and 5D**). Consistent that aneuploidy is detrimental to cellular fitness, acute expression of FGFR3-TACC3 compromised the proliferation capacity of human astrocytes. However, continuous culture of FGFR3-TACC3-expressing human astrocytes led to progressive gain of proliferative capacity that overrode that of control cells (**Figure 12A, 12B**). Thus, the acute expression of FGFR3-TACC3 in primary normal human cells from the central nervous system causes CIN and aneuploidy with an acute fitness cost manifested by slower proliferation.

It was also determined whether the CIN and aneuploidy caused by FGFR3-TACC3 requires the TK activity of FGFR3 and can be corrected. Treatment with PD173074 rescued the aneuploidy caused by FGFR3-TACC3 by over 80%, restored the narrow distribution of chromosome counts typical of control cells and largely corrected the cohesion defect (**Figures 6A, 6B and 6C**). Together, these findings indicate that the CIN and aneuploidy caused by rearrangements of FGFR and TACC genes are reversible and suggest that specific FGFR kinase inhibition may be a valuable therapeutic strategy in tumor cells expressing FGFR-TACC fusion proteins.

FGFR-TACC fusion proteins are new therapeutic targets in GBM. Driver genetic alterations trigger a state of oncogene addiction in the cancer cells harboring them that can be exploited therapeutically. To ask whether FGFR-TACC fusions confer addiction to FGFR-TK activity, cell growth was analyzed in the presence of PD173074, AZD4547 or BGJ398, the latter being two highly specific inhibitors of FGFR-TK under clinical investigation (Gavine et al., 2012; Guagnano et al., 2011). Each of the three drugs inhibited growth of cells expressing FGFR3-TACC3 and FGFR1-TACC1 at concentrations <10 nM whereas they were ineffective at concentrations as high as 1 μM in cells transduced with vector, FGFR3, TACC3 and the FGFR3-TACC3-K508M mutant (**Figures 7A, 12C** and **12D**). These findings underscore the elevated degree of specificity for FGFR kinase inhibition towards cells carrying the fusion protein. The growth of GSC-1123 cells, which naturally harbor the FGFR3-TACC3 translocation, was also abolished by nanomolar concentrations of FGFR-TK inhibitors (**Figure 7B**). Targeting of the fusion gene by FGFR3 shRNA inhibited the growth of cells ectopically expressing FGFR3-TACC3 and GSC-1123 proportionally to the

silencing efficiency of FGFR3-TACC3 (**Figures 7C and 12E**).

Finally, it was determined whether treatment with PD173074 of mice bearing glioma xenografts of FGFR3-TACC3 transformed astrocytes inhibits tumor growth. Twelve days after injection of tumor cells, subcutaneous tumors were present in all animals. The mice were randomized in two cohorts and treated with PD173074 or vehicle. PD173074 elicited a potent growth inhibition of FGFR3-TACC3 glioma (**Figure 7D**). To confirm the efficacy of a clinically meaningful FGFR-TK inhibitor using a more anatomically relevant model, the AZD4547 FGFR inhibitor, a compound under clinical investigation (Gavine et al., 2012), was used against intracranial luciferase-expressing FGFR3-TACC3-driven glioma xenografts. After an engraftment period, tumor-bearing animals were treated with either AZD4547 or vehicle. Oral administration of AZD4547 markedly prolonged survival (**Figure 7E**). Taken together, the data provide a strong rationale for a clinical trial based on FGFR inhibitors in GBM harboring FGFR-TACC rearrangements.

DISCUSSION

This work has established that recurrent, oncogenic and addicting gene fusions identify a subset of GBM patients. The functional characterization of FGFR-TACC fusions indicates that the constitutively active FGFR-TK and the TACC domain of the fusion protein are both essential for oncogenesis. The TACC-dependent mis-localization to mitotic cells of the FGFR kinase results in aberrant compartmentalization of a constitutively active TK to the mitotic spindle pole, thus providing a mechanistic explanation for the impaired mitotic fidelity, chromosome mis-segregation and aneuploidy instigated by the fusion protein. Without being bound by theory, mutation of the genes that control chromosome segregation during mitosis can explain the high rate of CIN and aneuploidy, which is typical of most solid tumors including GBM (Gordon et al., 2012). A few examples of mutational inactivation of candidate genes have been reported in human cancer (Solomon et al., 2011; Thompson et al., 2010). However, gain-of-function mutations causally implicated in the control of mitotic fidelity have not been described. This clashes with the classical observation from cell fusion experiments that the underlying mechanisms that cause CIN behave as dominant traits, indicating that the CIN phenotype results from gain-of-function events rather than gene inactivation (Lengauer et al., 1997, 1998). The FGFR-TACC gene fusion is a novel mechanism for the initiation of CIN and provides a clue to the nature of dominant mutations responsible for aneuploidy in human cancer.

The rapid emergence of mitotic defects and aneuploid cell populations triggered by the fusion protein in normal human astrocytes, combined with the correction of aneuploidy after short inhibition of FGFR-TK activity indicate that aneuploidy is a key event in tumor induction by the FGFR-TACC gene fusions. Induction of aneuploidy *per se* is detrimental to cellular fitness (Sheltzer and Amon, 2011). Full-blown tumorigenesis requires cooperation between aneuploidy and genetic lesions that confer growth advantage and protect cells against the detrimental effects of aneuploidy (Coschi and Dick, 2012; Holland and Cleveland, 2009; Weaver and Cleveland, 2009). Therefore, the potent tumor-initiating activity of FGFR-TACC fusions shows that the novel oncoproteins have growth-promoting signaling functions that

complement the loss of mitotic fidelity with ensuing karyotypic alterations (Sheltzer and Amon, 2011). Targeted therapies against common genetic alterations in GBM have not changed the dismal clinical outcome of the disease, most likely because they have systematically failed to eradicate the truly addicting oncoprotein activities of GBM. The highly specific anti-tumor effects and the correction of aneuploidy precipitated by FGFR-TK inhibition of FGFR-TACC-driven GBM provide a strong rationale for clinical trials based on FGFR inhibitors in patients harboring FGFR-TACC rearrangements. The computational gene fusion discovery pipeline reported here detected other GBM cases in which FGFR family genes are implicated in additional gene fusions beyond the FGFR-TACC rearrangements. Therefore, the frequency of 3.1% is likely to be an underestimate of the target GBM patient population that may benefit from FGFR-TK inhibition.

EXPERIMENTAL PROCEDURES

Cell culture and isolation and maintenance of GSCs. Rat1A, mouse astrocytes *Ink4A;Arf-/-*, and human astrocytes were cultured in DMEM supplemented with 10% FBS. Isolation and culture of GSCs was performed as described (Carro et al., 2010). For treatment *in vitro* with PD173074, AZD4547 or BJG398, cells infected with vector control, FGFR3, TACC3, FGFR-TACC fusions or FGFR3-TACC3-K508M were seeded in 96-well plates and treated with increasing concentrations of FGFR inhibitors. After 72-120 h, growth rate was measured using the 3-(4,5-dimethythiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) assay. Data were expressed as mean±SD. Proliferation rate in GSC-1123 infected with FGFR3 shRNA lentivirus was determined by plating dissociated gliomaspheres at $2x10^4$ cells/well in twelve-well plates 5 days after infection. The number of viable cells was determined by trypan blue exclusion in triplicate cultures obtained from triplicate independent infections. Cell number was scored every other day.

DNA, RNA preparation, genomic and real-time quantitative PCR (qRT-PCR).

The validation of fusion transcripts was performed using both genomic and RT-PCR with forward and reverse primer combinations designed within the margins of the paired-end read sequences detected by RNA-seq. DNA, RNA preparation and qRT-PCR were performed as described (Carro et al., 2010; Zhao et al., 2008). To identify novel fusion transcripts within the GBM cohort, PCR primers pairs were designed to bind upstream to the TK domain of the FGFR genes and inside or downstream the Coiled Coil domain of the TACC genes. Expressed fusion transcript variants were subjected to direct sequencing to confirm sequence and frame. Primer sequences are included below.

Subcutaneous xenografts and drug treatment. Rat1A or *Ink4A;Arf-/-* astrocytes ($5x10^5$) transduced with different lentiviral constructs were suspended in 150 µl of PBS, together with 30 µl of Matrigel (BD Biosciences), and injected subcutaneously in the flank of athymic nude (Nu/Nu) mice (Charles River Laboratories, Wilmington, MA). For experiments with FGFR inhibitors, mice carrying ~200-300 mm^3 subcutaneous tumors derived from *Ink4A;Arf-/-* astrocytes were randomized to receive 50 mg/kg PD173074 in 0.05 M lactate buffer (pH 5) or an equal volume of lactate buffer by oral gavage. Treatment was administered for three cycles consisting of four consecutive days followed by two days of rest. Tumor diameters were measured with caliper, and tumor volumes estimated using the formula: 0.5 x length x

width^2. Data are expressed as mean±SE. Mice were sacrificed when tumors in the control group reached the maximal size allowed.

Orthotopic transplantation and drug treatment. *Ink4A;Arf-/-* astrocytes carrying a luciferase expressing vector were transduced with FGFR3-TACC3 lentivitus. 1x10^3 cells in 2 μl of saline were injected in the caudate-putamen of 4-6 week old male athymic nude (Nu/Nu) mice using a stereotaxic frame (coordinates relative to bregma: 0.5 mm anterior; 1.1 mm lateral; 3.0 mm ventral) and a 26 gauge Hamilton syringe. Six days after injection, mice underwent bioluminescence imaging using a Xenogen CCD apparatus and were randomized to receive 50 mg/kg AZD4547 in 1% Tween 80 (treatment group) or DMSO in an equal volume of vehicle by oral gavage (control group). AZD4547 was administered daily for two cycles of 10 days with a two day interval. Mice were monitored daily and sacrificed when neurological symptoms appeared. Kaplan-Meier survival curve was generated using the DNA Statview software package (AbacusConcepts, Berkeley, CA). Log-rank analysis was performed on the Kaplan-Meier survival curve to determine statistical significance.

Intracranial injections of lentiviruses. Intracranial injection of FGFR3-TACC3-shp53, EGFRvIII-shp53 or shp53 pTomo lentiviruses was performed in 4 week-old C57/BL/6J mice in accordance with guidelines of IACUC Committee. Briefly, 1.8 μl of purified lentiviral particles in PBS (1X10^9/ml) were injected into the dentate gyrus using a stereotaxic frame (coordinates relative to bregma: 1.45 mm posterior; 1.65 mm lateral; 2.4 mm ventral) and a 26 gauge Hamilton syringe. Mice were monitored daily and sacrificed when neurological symptoms appeared. Mouse brain was analyzed histopathologically and by immunofluorescence staining.

Histology and immunostaining. Tissue preparation and immunohistochemistry on brain tumors and immunofluorescence staining were performed as previously described (Carro et al., 2010; Zhao et al., 2009; Zhao et al., 2008). Antibodies used in immunostaining and immunoblotting are listed below.

Cloning and Lentiviral production. Lentivirus preparation and infections were performed as described (Carro et al., 2010) and are detailed in Extended Experimental Procedures.

Karyotype analysis. Cultured cells were colcemid (20 ng/ml) treated for 90 minutes before harvesting for karyotopic analysis as detailed in Extended Experimental procedures. At least one hundred cells in metaphase were examined for chromosome count. PMSCS was scored in cells where a majority of the sister chromosomes were no longer associated. Two-tailed unpaired t-tests with Welch's correction were performed for comparison of means analysis.

Immunofluorescence and live-cell microscopy. Immunofluorescence microscopy was performed on cells fixed with 4% PFA in PHEM (60 mM Pipes, 27 mM Hepes, 10 mM EGTA, 4 mM $MgSO_4$, pH 7.0). Cells were permeabilized using 1% Triton X-100. Mitotic spindles were visualized by anti-α-tubulin antibody (Sigma). Secondary antibodies conjugated to Alexa Fluor-488/-594 (Molecular Probes) were used. All staining with multiple antibodies were performed in a sequential manner. DNA was stained by DAPI (Sigma). Fluorescence microscopy was performed on a Nikon A1R MP microscope.

Identification of gene fusions from whole transcriptome (RNA-seq) and exome sequencing. RNA-

Sequencing was performed from total RNA extracted from GSC cultures isolated from nine GBM patients using Illumina HiSeq 2000, producing roughly 60.3 million paired reads per sample. Using the global alignment software Burrows-Wheeler Aligner (BWA) (Li and Durbin, 2009) with modified Mott's trimming, an initial seed length of 32, maximum edit distance of 2 and a maximum gap number of 1, on average 43.1 million reads were mapped properly to the RefSeq transcriptome and, of the remaining, 8.6 million were mapped to the hg19 genome per sample. The remaining 14.3% of paired reads—including those that failed to map to either transcriptome or genome with proper forward-reverse (F-R) orientation, within expected insert size, and with minimal soft clipping (unmapped portions at the ends of a read)— were considered to be appropriate for gene fusion analysis.

A novel computational pipeline was constructed called ***TX-Fuse*** that identifies two sources of evidence for the presence of a gene fusion: 1. *Split inserts*, in which each read of a mate pair maps entirely to one side of a breakpoint, and 2. Individual *split reads* that span a breakpoint. Split inserts are readily detected from BWA mapping. On the other hand, split reads demand precision alignment of smaller nucleotide stretches. To that end, the pipeline employs the local alignment package BLAST with word size of 20, identity cutoff of 95%, expectation cutoff of 10^{-4}, and soft filtering to map raw paired reads against the RefSeq transcriptome. From this procedure, a list of potential split reads were obtained that were filtered to ensure maintenance of coding frame in the predicted fusion transcript given the proper F-R orientation in the read pair. False positive candidates produced from paralogous gene pairs were also screened out using the Duplicated Genes Database and the EnsemblCompara GeneTrees (Vilella et al., 2009). Pseudogenes in the candidate list were annotated using the list from HUGO Gene Nomenclature Committee (HGNC) database (Seal et al., 2011) and given lower priority. For each remaining gene fusion candidate, a virtual reference was created based on the predicted fusion transcript and re-mapped all unmapped reads using BLAST with word size of 16, identity cutoff of 85%, query coverage greater than 85%, and expectation cutoff of 10^{-4} to obtain a final count of split reads and inserts. Moreover, sequencing depth per base of the virtual reference was calculated to corroborate that components of each gene participating in the gene fusion were highly expressed.

To establish the recurrence of the initial panel of gene fusion candidates, the gene fusion discovery pipeline was modified to produce ***EXome-Fuse***, which probes for fusions within the available dataset of paired-read exome DNA sequencing of 84 matched GBM samples from TCGA. To increase sensitivity for gene fusion identification, reads unmapped by BWA were aligned to the gene pair participating in each fusion candidate using a BLAST word size of 24 for split inserts and 16 for split read and split insert discovery. Given that the breakpoint detected in DNA cannot directly indicate the resulting breakpoint in the transcribed RNA, no restriction was made on split insert orientation. For split reads, it was only required that the component of the split read mapped to the same gene as its mate maintained F-R directionality.

Co-outlier expression and CNV analysis from TCGA GBM samples. Tomlins et al. (Tomlins et al., 2005) reported that outlier gene expression from microarray datasets identifies candidate oncogenic gene fusions. Wang et al. (Wang et al., 2009) suggested a "breakpoint principle" for intragenic copy number aberrations in

fusion partners. The two principles (outlier expression and intragenic CNV) were combined to identify candidate gene fusions in GBM samples from Atlas-TCGA. Genomic and expression data sets were downloaded from TCGA public data portal as available on December 1, 2011, where a description of TCGA data types, platforms, and analyses is also available (2008). Specific data sources were (according to Data Levels and Data Types) as follows: Expression data, "Level 2" normalized signals per probe set (Affymetrix HT_HG-U133A) of 84 samples; Copy number data, "Level 1" raw signals per probe (Affymetrix Genome-Wide Human SNP *Array* 6.0) of the 4 *FGFR3-TACC3* gene fusion positive samples (tumor and matched normal control).

The gene expression analysis was performed first using R^3. The median absolute deviation (MAD) was calculated and then a gene was labeled as an outlier according to the following formula: $Z_{i,j} = 0.6745(x_{i,j} - \text{mean}(x_i))/MAD_i > 3.5$ (Iglewicz and Hoaglin, 1993). Samples were identified as ECFS (expression candidate fusion sample) if both genes of interest (e. g. FGFR3 and TACC3) displayed outlier behavior (co-outliers). Next, ECFS were analyzed for CNV using pennCNV (Wang et al., 2007). Tumors samples were paired to their normal controls to obtain the log ratio values and the VEGA algorithm was used to obtain a more accurate segmentation (Morganella et al., 2010).

Karyotypic Analysis. The colcemid treated cells were trypsinized, centrifuged for 7 minutes at 200 x g, and the cell pellet re-suspended in warmed hypotonic solution and incubated at 37°C for 13 minutes. The swollen cells were then centrifuged and the pellet re-suspended in 8 ml of Carnoy's fixative (3:1 methanol:glacial acetic acid). The cell suspension was centrifuged and washed twice in Carnoy's fixative. After the last centrifugation, the cells were resuspended in 0.5 to 1 ml of freshly prepared fixative to produce an opalescent cell suspension. Drops of the final cell suspension were placed on clean slides and air-dried. Slides were stained with DAPI and metaphases were analyzed under a fluorescent microscope.

Cloning and Lentiviral production. Lentiviral expression vectors, pLOC-GFP (Open Biosystems) and pTomo-shp53, were used to clone FGFR3, TACC3, FGFR3-TACC3, FGFR3-TACC3-K508M, and FGFR1-TACC1. pTomo-shp53 was a gift of Inder Verma and Dinorah Friedman-Morvinski (Salk Institute, San Diego). The FGFR3-TACC3-K508M mutant was generated using the Phusion Site Direct Mutagenesis kit (NEB, USA). MISSION shRNAs clones (pLKO.1 lentiviral expression vectors) against FGFR3 were purchased from Sigma. The hairpin sequences targeting the *FGFR3* gene are SEQ ID NO: 182 (#TRCN0000000372; Sh#2); SEQ ID NO: 183 (#TRCN0000430673; Sh#3); and SEQ ID NO: 184 (#TRCN0000000374; Sh#4).

Genomic and mRNA RT-PCR. Total RNA was extracted from cells by using RNeasy Mini Kit (QIAGEN), following the manufacturer instructions. 500 ng of total RNA was retro-transcribed by using the Superscript III kit (Invitrogen), following the manufacturer instructions. The cDNAs obtained after the retro-transcription was used as templates for qPCR. The reaction was performed with a Roche480 thermal cycler, by using the Absolute Blue QPCR SYBR Green Mix from Thermo Scientific. The relative amount of specific mRNA was normalized to 18S. Results are presented as the mean ± SD of triplicate amplifications. Primers used are: hFGFR3-RT-FW1 (SEQ ID NO: 162); hFGFR3-RT-REV1 (SEQ ID NO: 163); hTACC3-

RT-FW2 (SEQ ID NO: 164); hTACC3-RT-REV2 (SEQ ID NO: 165); hWRAP53-RT-FW1 (SEQ ID NO: 180); hWRAP53-RT-REV1 (SEQ ID NO: 181).

Primers used for the screening of FGFR-TACC fusions are: FGFR3-FW1 (SEQ ID NO: 166); TACC3-REV1 (SEQ ID NO: 167); FGFR1-FW1 (SEQ ID NO: 168); TACC1-REV1 (SEQ ID NO: 169).

Primers used for genomic detection of FGFR3-TACC3 fusion in GBM-1123 and GSC-1123 are: Genomic FGFR3-FW1 (SEQ ID NO: 170); genomicTACC3-REV1 (SEQ ID NO: 171).

Primers used to validate fusions detected by RNA-seq are: POLR2A-FW1 (SEQ ID NO: 172); WRAP53-REV1 (SEQ ID NO: 173); PIGU-FW1 (SEQ ID NO: 174); NCOA6-REV1 (SEQ ID NO: 175); ST8SIA4-FW1 (SEQ ID NO: 176); PAM-REV1 (SEQ ID NO: 177); CAPZB-FW (SEQ ID NO: 178);UBR4-REV1 (SEQ ID NO: 179).

Confocal microscopy imaging. For immunofluorescence of fixed cells, images were recorded with a *Z*-optical spacing of 0.25 μm using a Nikon A1R MP and a 60X1.3 oil objective and analyzed using ImageJ software (National Institute of Health). For live-cell analyses, Rat1A cells infected with pLNCX-H2B retrovirus and transduced with lentiviral vector or FGFR3-TACC3 fusion were seeded in glass bottom dishes in phenol red free DMEM and followed by time-lapse microscopy using the Nikon A1R MP biostation at 37°C and 5% CO_2/95% air. Images with a *Z*-optical spacing of 1 μm were recorded every 4 min for 8 h. Images of unchallenged mitosis from early prophase until cytokinesis were processed using ImageJ software (National Institute of Health). The time-point of nuclear envelope breakdown (NEB) was defined as the first frame showing loss of smooth appearance of chromatin and anaphase was the first frame when chromosome movement towards the poles became apparent. Nuclear envelope reconstitution (NER) was defined as the first frame showing nuclei decondensation.

Box and whisker plots were calculated from image sequences from at least 50 recorded cells. Two-tailed unpaired t-tests with Welch's correction were performed for comparison of means analysis using StatView software (AbacusConcepts, Berkeley, CA).

Immunofluorescence. Antibodies and concentrations used in immunofluorescence staining are:

Anti-Ki67	Rabbit	1:1000	Vector Labs
Anti-pHH3	Rabbit	1:500	Millipore
Anti-FGFR3	Mouse	1:1000	Santa Cruz
Anti-Tacc3	Goat	1:1000	USBiological
Anti-a-tubulin	Mouse	1:1000	Sigma
Anti-Nestin	Mouse	1:1000	BD Pharmingen
Anti-Olig2	Rabbit	1:200	IBL
Anti-GFAP	Rabbit	1:200	Dako
Anti-ERK	Rabbit	1:1000	Cell Signaling
Anti-pERK	Rabbit	1:1000	Cell Signaling
AntiFRS	Rabbit	1:250	Santa Cruz
Anti-pFRS	Rabbit	1:1000	Cell Signaling
Anti-AKT	Rabbit	1:1000	Cell Signaling
Anti-pAKT473	Rabbit	1:1000	Cell Signaling

Example 2- *Fusions in GBM*

Table 8: Soft agar colony assay

Cell line	Vector	FGFR3	TACC3	F1-T1 Fusion	F3-T3 Fusion	F3-T3-K508M Fusion
Rat1	0	0	0	225.3±10.0	198.7±8.0	0
Balb 3T3	0	0	0	n.d.	45.5±8.9	n.d.

n.d.: not done

Table 9: Subcutaneous tumor xenografts

Cell line	Vector	FGFR3	TACC3	F1-T1 Fusion	F3-T3 Fusion	F3-T3-K508M Fusion
Rat1	0/5	0/5	0/5	n.d.	5/5	n.d.
Ink4A/Arf-/- Astrocytes	0/9	0/5	0/5	8/8	12/12	0/8

n.d.: not done

Table 10: Analysis of chromosomal number in Rat1 cells

Cell line	Number of cells counted	Percent aneuploidy	Range	Mean number	Average variation from mean number	*p*- value
Rat1A Vector	100	27	35-43	41.2	1.2	
Rat1A FGFR3	100	33	35-44	42.1	1.3	n.s.
Rat1A TACC3	100	41	34-46	40.7	1.1	n.s.
Rat1A FGFR3-TACC3	100	69	35-73	43.8	3.1	<0.0001

Table 11: Analysis of chromosomal number in human astrocytes

Cell line	Number of cells counted	Percent aneuploidy	Range	Mean number	Average variation from mean number	*p*- value
Human Astrocytes Vector	100	8	42-46	45.85	0.28	$p = <0.001$
Human Astrocytes FGFR3-TACC3	100	42	28-48	42.24	3.33	

Example 3 - *Fusions in Other Cancers*

The inventors reported in Example 1 that 3.1% of human glioblastoma harbor FGFR3-TACC3 and FGFR1-TACC1 gene fusions. Tumors harboring FGFR3-TACC3 gene fusions are identified by the presence of highly specific focal micro-amplification events of the rearranged portions of the FGFR3 and TACC3 genes (*See* **FIG. 2D**). Therefore, these micro-amplification events can be used as distinctive marks for the presence of FGFR3-TACC3 gene fusions. It was asked whether other types of human tumors also harbor FGFR3-

TACC3 gene fusions from the analysis of Copy Number Variations (CNVs) of SNP arrays generated from the Atlas-TCGA project. This analysis was performed using segmented CNVs data visualized using the Integrated Genomic Viewers software. The analysis revealed that the following tumors display focal micro-amplification events of FGFR3 and TACC3 that indicate the presence of FGFR3-TACC3 gene fusions: Bladder Urothelial Carcinoma; Breast Carcinoma; Colorectal Carcinoma; Lung Squamous Cell Carcinoma; Head and Neck Squamous Cell Carcinoma.

Taken together, these data indicate that the same FGFR3-TACC3 gene fusions reported for the first time in Glioblastoma also occur in several other types of human tumors. Therefore, as for Glioblastoma and other epithelial cancers (such as the human tumors discussed herein), the identification of FGFR-TACC gene fusions also provides a new diagnostic and therapeutic target for treatment with drugs that inhibit FGFR-TACC gene fusions.

Example 4 –*fusion transcripts identified by RNA-seq of nine GSCs*

POLR2A-WRAP53, CAPZB-UBR4, ST8SIA4-PAM, PIGU-NCOA6, IFNAR2-IL10RB fusions were identified by RNA-seq of nine GSCs (**FIGS. 9A-D).**

The POLR2A-WRAP53 gene fusion was identified by whole transcriptome sequencing of GSCs. SEQ ID NOs: 329-382 disclose 54 split-reads . SEQ ID NO: 327 discloses the predicted reading frame at the breakpoint (479bp of NM_000937 fused to 798bp of NM_001143990), with the POLR2A sequence (nucleotides 1-88) preceding the WRAP53 sequence (nucleotides 89-176). SEQ ID NO: 328 discloses the putative amino acid sequence corresponding to SEQ ID NO: 327. Amino acids 1-29 of SEQ ID NO: 328 correspond to the POLR2A sequence and amino acids 30-59 correspond to the WRAP53 sequence.

The CAPZB-UBR4 gene fusion was identified by whole transcriptome sequencing of GSCs. SEQ ID NOs: 385-432 disclose 48 split-reads . SEQ ID NO: 383 discloses the predicted reading frame at the breakpoint (227bp of NM_00120650 fused to 12110bp of NM_020/65), with the CAPZB sequence (nucleotides 1-82) preceding the UBR4 sequence (nucleotides 83-170). SEQ ID NO: 384discloses the putative amino acid sequence corresponding to SEQ ID NO: 383. Amino acids 1-27 of SEQ ID NO: 384 correspond to the CAPZB sequence and amino acids 27-56 correspond to the UBR4 sequence.

The ST8SIA4-PAM gene fusion was identified by whole transcriptome sequencing of GSCs. SEQ ID NOs: 435-463 disclose 29 split-reads . SEQ ID NO: 433 discloses the predicted reading frame at the breakpoint (1125bp of NM_005668 fused to 730bp of NM_000919), with the ST8SIA4 sequence (nucleotides 1-87) preceding the PAM sequence (nucleotides 88-176). SEQ ID NO: 434 discloses the putative amino acid sequence corresponding to SEQ ID NO: 433. Amino acids 1-29 of SEQ ID NO: 434 correspond to the ST8SIA4 sequence and amino acids 30-59 correspond to the PAM sequence.

The PIGU-NCOA6 gene fusion was identified by whole transcriptome sequencing of GSCs. SEQ ID NOs: 466-482 disclose 17 split-reads . SEQ ID NO: 464 discloses the predicted reading frame at the breakpoint (729bp of NM_080476 fused to 6471bp of NM_014071), with the PIGU sequence (nucleotides 1-84) preceding the NCOA6 sequence (nucleotides 85-167). SEQ ID NO: 465 discloses the putative amino acid

sequence corresponding to SEQ ID NO: 464. Amino acids 1-28 of SEQ ID NO: 465 correspond to the PIGU sequence and amino acids 39-56 correspond to the NCOA6 sequence.

The IFNAR2-IL10RB gene fusion was identified by whole transcriptome sequencing of GSCs. SEQ ID NOs: 495-490 disclose 6 split-reads . SEQ ID NO: 483 discloses the predicted reading frame at the breakpoint (1083bp of NM_000874 fused to 149bp of NM_000628), with the IFNAR2 sequence (nucleotides 1-78) preceding the IL10RB sequence (nucleotides 79-166). SEQ ID NO: 484 discloses the putative amino acid sequence corresponding to SEQ ID NO: 483. Amino acids 1-26 of SEQ ID NO: 484 correspond to the IFNAR2 sequence and amino acids 27-55 correspond to the IL10RB sequence.

Example 4 - ***Detection, characterization and inhibition of FGFR-TACC fusions in IDH wild type glioma***

TRANSLATIONAL RELEVANCE

Described herein is an unbiased screening assay for FGFR-TACC fusions in glioma that overcomes the great variability of variants that are generated by FGFR-TACC chromosomal translocation in human cancer. FGFR-TACC fusions occur in grade II and III glioma harboring wildtype IDH1 with frequency similar to glioblastoma (GBM), therefore providing a clue to the aggressive clinical behavior of this glioma subtype. The molecular characterization of fusion-positive glioma revealed that FGFR-TACC is mutually exclusive with EGFR amplification but co-occurs with CDK4 amplification. FGFR-TACC-positive glioma displays strikingly uniform and strong expression of the fusion protein at the single cell level. Preclinical experiments with FGFR3-TACC3-positive glioma cells treated with the FGFR inhibitor JNJ-42756493 showed strong antitumor effects and treatment of two patients with recurrent GBM harboring FGFR3-TACC3 resulted in clinical improvement and radiological tumor reduction. These findings validate the treatment with FGFR inhibitors of glioma patients harboring FGFR-TACC chromosomal translocations.

ABSTRACT

Purpose. Oncogenic fusions consisting of FGFR and TACC are present in a subgroup of glioblastoma (GBM) and other human cancers and have been proposed as new therapeutic targets. Frequency, molecular features of FGFR-TACC fusions, and the therapeutic efficacy of inhibiting FGFR kinase in GBM and grade-II-III glioma were analyzed.

Experimental Design. Overall, 795 gliomas (584 GBM, 85 grade-II-III with wild-type and 126 with IDH1/2 mutation) were screened for FGFR-TACC breakpoints and associated molecular profile. Expression of the FGFR3 and TACC3 components of the fusions were also analyzed. The effects of the specific FGFR inhibitor JNJ-42756493 for FGFR3-TACC3-positive glioma were determined in preclinical experiments. Two patients with advanced FGFR3-TACC3-positive GBM received JNJ-42756493 and were assessed for therapeutic response.

Results. Three of 85 IDH1/2 wild type (3.5%) but none of 126 IDH1/2 mutant grade-II-III glioma harbored FGFR3-TACC3 fusions. FGFR-TACC rearrangements were present in 17 of 584 GBM (2.9%). FGFR3-TACC3 fusions were associated with strong and homogeneous FGFR3 immunostaining. They are mutually exclusive with IDH1/2 mutations and EGFR amplification whereas co-occur with CDK4 amplification. JNJ-42756493 inhibited growth of glioma cells harboring FGFR3-TACC3 in vitro and in vivo. The two patients

with FGFR3-TACC3 rearrangements who received JNJ-42756493 manifested clinical improvement with stable disease and minor response, respectively.

Conclusions. RT-PCR-sequencing is a sensitive and specific method to identify FGFR-TACC-positive patients. FGFR3-TACC3 fusions are associated with uniform intra-tumor expression of the fusion protein. The clinical response observed in the FGFR3-TACC3-positive patients treated with a FGFR inhibitor supports clinical studies of FGFR inhibition in FGFR-TACC-positive patients.

INTRODUCTION

The history of successful targeted therapy of cancer largely coincides with the inactivation of recurrent, oncogenic and addicting gene fusions in hematological malignancies and recently in some types of epithelial cancer (1, 2). Glioblastoma multiforme (GBM) is among the most lethal forms of human cancer and targeted therapies against common genetic alterations in GBM have not changed the dismal outcome of the disease (3, 4). Underlying biological features including infiltrative growth behavior, intratumoral heterogeneity, and adaptive resistance mechanisms coupled with the unique challenges of intracranial location present significant problems in its effective management. Despite surgery and chemo-radiotherapy, most patients rapidly recur and no effective treatment options are available at that stage. Beside GBM, which features the highest grade of malignancy among glioma (grade IV), lower grade glioma which include grade II and grade III are a heterogeneous group of tumors in which specific molecular features are associated with divergent clinical outcome. The majority of grade II-III glioma (but only a small subgroup of GBM) harbor mutations in IDH genes (IDH1 or IDH2), which confer a more favorable clinical outcome. Conversely, the absence of IDH mutations is associated with the worst prognosis (5).

Described herein is the identification of FGFR-TACC gene fusions (mostly FGFR3-TACC3, and rarely FGFR1-TACC1) as the first example of highly oncogenic and recurrent gene fusions in GBM. The FGFR-TACC fusions that have been identified so far include the Tyrosine Kinase (TK) domain of FGFR and the coiled-coil domain of TACC proteins, both necessary for the oncogenic function of FGFR-TACC fusions. Tumor dependency on FGFR-TACC fusions was also tested in preclinical mouse models of FGFR-TACC glioma and observed marked anti-tumor effects by FGFR inhibition (6). FGFR3-TACC3 fusions have been identified in pediatric and adult glioma, bladder carcinoma, squamous lung carcinoma and head and neck carcinoma, thus establishing FGFR-TACC fusions as one of the chromosomal translocation most frequently found across multiple types of human cancers (6-15).

From a mechanistic standpoint, the unexpected capacity of FGFR-TACC fusions to trigger aberrant chromosome segregation during mitosis, thus initiating chromosome instability (CIN) and aneuploidy, two hallmarks of cancer, is described herein. However, the full repertoire of the structural variants of FGFR-TACC fusions occurring in GBM and lower grade glioma is not completely understood. Furthermore, it remains unknown whether FGFR-TACC fusions mark distinct grades of glioma and GBM subtypes.

To date eight variants of the FGFR3-TACC3 fusion have been reported that mostly differ for the breakpoint in the TACC3 gene (6-15). Because of the close proximity of FGFR3 and TACC3 (the two genes map at a distance of 70 Kb on chromosome 4p16.3), detection of FGFR3-TACC3 rearrangements by FISH is not a

feasible option with the currently available methods. Here a screening method for FGFR-TACC fusions is reported that includes a RT-PCR assay designed to identify the known and novel FGFR3-TACC3 fusion transcripts, followed by confirmation of the inframe breakpoint by Sanger sequencing. Using this assay, a dataset of 584 GBM and 211 grade II and grade III gliomas has been analyzed.

A crucial question with fundamental clinical relevance for any novel candidate target mutation is the frequency of the alteration in the cancer cell population, thus discriminating between a clonal or sub-clonal origin of the mutation. In fact, GBM is characterized by a formidable degree of subclonal heterogeneity, whereby neighboring cells display amplification and expression of different Receptor Tyrosine Kinase (RTK)-coding genes (16-19). This notion poses major therapeutic challenges for targeting any individual RTK will result, at best, in the eradication of a limited tumor sub-clone. Described herein, it was determined that brain tumors harboring FGFR-TACC fusions manifest strong and homogeneous intra-tumor expression of the FGFR3 and TACC3 component invariably included in the fusion protein, when analyzed by immunostaining. A significant clinical benefit following treatment with a specific inhibitor of FGFR-TK is reported in two GBM patients who harbored FGFR3-TACC3 rearrangement.

MATERIALS AND METHODS

Patients and tissue samples. This example includes a cohort of 746 untreated patients with histologic diagnosis of glioma from 5 institutions. Forty-nine recurrent gliomas from Pitié-Salpêtrière Hospital and one recurrent glioma from the University of Calgary were also included. A summary of the patient cohort is provided in Table 12.

Tumor sample source		No of case (GBM)	No of detected fusions	Immunostaining FGFR3 positive /Sample analyzed
Pitié-Salpêtrière Hospital		380	9	9/9
Besta Neurological Institute		85	5	2/2
University of Calgary		60 + 1R[§]	2 + 1R[§]	1/1 +1/1R[§]
Montreal Neurological Institute		51	1	-
University of British Columbia		8	0	-
Total		584 (100%)[£]	17 (2.9%)	
Tumor sample source	IDH Status	No of cases (Grade II-III)	No of detected fusions	Immunostaining FGFR3 positive/Sample analyzed
Pitié-Salpêtrière Hospital	IDH wt	85* (100%)	3 (3.5%)	3/3
	IDH1/IDH2 Mut	126 (100%)	0 (0%)	0

Table 12 - Frequency of FGFR3-TACC3 Fusions in GBM and Grade II-III glioma. Distribution of the FGFR3-TACC3 fusions in GBM (upper panel) and lower grade glioma (lower panel) samples stratified according to the Institution of origin. The table reports number of cases analyzed, number of tumors harboring FGFR3-TACC3 fusion transcripts, and results of FGFR3 immunostaining. Lower grade glioma samples are further classified according to *IDH* status (*IDH1* and *IDH2*). The respective frequency of

FGFR3-TACC3 in GBM, Glioma grade II-III *IDH* wild type (wt), and *IDH* mutant (Mut) glioma is reported in parentheses. R[§], Recurrent GBM. [£], Recurrent GBM from the University of Calgary Dataset is not included in the total count of GBM. *, 25 cases out of 85 are unknown for *IDH2* status.

Tumor specimens, blood samples and clinico-pathological information were collected with informed consent and relevant ethical board approval in accordance with the tenets of the Declaration of Helsinki. For the samples from the Pitié-Salpêtrière Hospital, clinical data and follow-up are available in the neuro-oncology database (Onconeurotek, GH Pitié-Salpêtrière, Paris).

Two recurrent GBM patients harboring FGFR3-TACC3 were enrolled in the dose escalation part of JNJ-42756493 trial at the Gustave Roussy Institute.

Identification of fusion transcripts and analysis of genomic breakpoints. Total RNA was extracted from frozen tissues using Trizol (Invitrogen) according to manufacturer instructions. Two to three hundred nanograms of total RNA were retro-transcribed with the Maxima First Strand cDNA Synthesis Kit (Thermo Scientific) or SuperScript II (Invitrogen). RT-PCR was performed using AccuPrime Taq DNA Polymerase (Invitrogen). Primer pairs used for the FGFR3-TACC3 fusions screening were: FGFR3ex12-FW: 5'-CGTGAAGATGCTGAAAGACGATG-3 (SEQ ID NO: 495) and TACC3ex14-RV: 5'-AAACGCTTGAAGAGGTCGGAG-3 (SEQ ID NO: 496); amplification conditions were 94°C-3min, (94°C-30sec/61°C-30sec/68°C-1min40sec) for 35 cycles, 68°C-7min. FGFR1-TACC1 fusions were amplified with FGFR1ex16-FW: 5'-TGCCTGTGGAGGAACTTTTCA-3' (SEQ ID NO: 497) and TACC1ex13-RV: 5'-CCCAAACTCAGCAGCCTAAG-3' (SEQ ID NO: 498) primers (94°C-30sec/60°C-30sec/68°C-1min40sec for 35 cycles). PCR products were subjected to Sanger sequencing.

FGFR3-TACC3 genomic breakpoints were analyzed in 6 FGFR3-TACC3 positive samples, 5 of which from the Pitié-Salpêtrière Hospital and 1 from Montreal Neurological Institute. Three additional samples (MB-22, TCGA 27-1835 and TCGA 06-6390) available from the previous study (6) were also included in the analysis. Fifty nanograms of genomic DNA were used in the PCR reaction, performed with Accuprime Taq Polymerase (Invitrogen) and PCR products were Sanger sequenced. Primers used in genomic PCR were designed according to the breakpoint sequence in the mRNA; the list of primers used are: FGFR3ex17-FW 5'- TGGACCGTGTCCTTACCGT-3' (SEQ ID NO: 499) (PCR Samples 3048, 4373, 4867, 4451, MB-22, OPK-14, 06-6390, 27-1835 and Sequencing samples 3048, 4373, 4867, 4451, MB-22, OPK14, 06-6390, 27-1835); FGFR3ex16-FW 5'- GGTCCTTTGGGGTCCTGCT-3' (SEQ ID NO: 500) (PCR and Sequencing Sample 3808); TACC3ex6-RV 5'- CCTCTTTCAGCTCCAAGGCA-3' (SEQ ID NO: 501) (PCR and Sequencing Samples PCR 4451 and OPK-14); TACC3ex8-RV 5'-TCTACCAGGACTGTCCCTCAG-3' (SEQ ID NO: 502) (Sequencing Samples 3048 and 4373); TACC3ex9-RV 5'-GGGAGTCTCATTTGCACCGT-3' (SEQ ID NO: 503) (PCR Samples 3048,4373,4867 and Sequencing Sample 4867); TACC3ex10-RV 5'- CTGCATCCAGGTCCTTCTGG-3' (SEQ ID NO: 504) (PCR and Sequencing Samples MB-22 and 06-6390); TACC3ex11-RV 5'- CCAGTTCCAGGTTCTTCCCG-3' (SEQ ID NO: 505) (Sequencing Samples 27-1837 and 3808); TACC3ex12-RV 5'-CAACCTCTTCGAACCTGTCCA-3' (SEQ ID NO: 506) (PCR and Sequencing Samples 27-1837 and

3808). PCR conditions were 94°C-30sec/60°C-30sec/68°C-2min30sec for 40 cycles. For amplifications performed with the primer TACC3ex9-RV, the program was 94°C-30sec/56°C-30sec/68°C-2min30sec) for 40 cycles.

Quantitation of FGFR3 and TACC3 transcripts in GBM. The relative expression of FGFR3 and TACC3 regions included in or excluded from the fusion transcript was assessed by qRT-PCR. Primer pairs with comparable efficiency of amplification were identified and efficiency was assessed using serial dilutions of cDNA (20) prepared from OAW28 ovarian carcinoma cells that contain wild type FGFR3 and TACC3 (21). Primers used are: N-terminal region of FGFR3, FGFR3-N: Forward 5'-AAGACGATGCCACTGACAAG-3' (SEQ ID NO: 507), Reverse 5'-CCCAGCAGGTTGATGATGTTTTTG-3' (SEQ ID NO: 508); C-terminal region of TACC3, TACC3-C: Forward 5'-TCCTTCTCCGACCTCTTCAAGC-3' (SEQ ID NO: 509), Reverse 5'-TAATCCTCCACGCACTTCTTCAG-3' (SEQ ID NO: 510). To amplify transcripts in regions excluded from FGFR3-TACC3 fusion, primers were designed in the C-terminal region of FGFR3, FGFR3-C: Forward 5'-TACCTGGACCTGTCGGCG-3' (SEQ ID NO: 511), Reverse 5'-TGGGCAAACACGGAGTCG-3' (SEQ ID NO: 512) and N-terminal domain of TACC3, TACC3-N: Forward 5'-CCACAGACGCACAGGATTCTAAGTC-3' (SEQ ID NO: 513), Reverse 5'-TGAGTTTTCCAGTCCAAGGGTG-3' (SEQ ID NO: 514). All reactions were performed in triplicate and the data are reported as Fold Change ± Standard Deviation.

Immunofluorescence and immunohistochemistry. For immunofluorescence (IF) staining of FGFR3, 5μm FFPE sections subjected to antigen retrieval with citrate buffer for 8 min. Primary antibodies were: FGFR3-N (1:400, sc-13121, Santa Cruz Biotechnology), FGFR3-C (1:2000, sc-123, Santa Cruz Biotechnology), TACC3-N (1:600, ab134153, Abcam), and TACC3-C (1:300, NBP1-01032, Novus Biological). Secondary biotinylated antibodies were used at 1:50,000 followed by streptavidin and TSA Cy3-conjugated. Nuclei were counterstained with DAPI. For immunohistochemical analysis (IHC) of FGFR3 expression, antigen retrieval was performed for 12 min and FGFR-3 antibody (sc-13121, Santa Cruz Biotechnology) was diluted 1:500. Biotinylated anti-mouse antibody (1:30,000) and streptavidin were added before incubation with the chromogen. Nuclei were counterstaining with hematoxylin.

Molecular characterization of tumor samples. Mutational status of IDH1, IDH2, TERT promoter, as well as the methylation status of the MGMT promoter was analyzed in the Pitié-Salpêtrière cohort. Expression of IDH1-R132H mutant was analyzed by IHC in 500 cases as previously described (22). IDH1 and IDH2 gene mutations were identified by Sanger sequencing in 464 and 388 gliomas, respectively (5). IDH wild-type tumors are defined according to the absence of IDH1-R132H immunopositivity and/or mutations in IDH1 and IDH2 genes. TERT promoter status was determined by the same technique in 277 samples (23). Hyper-methylation of the MGMT promoter was tested in 242 samples by bisulfite pyro-sequencing (24). The presence of EGFRvIII was evaluated by RT-PCR in 118 samples using EGFR-FW 5'-CTTCGGGGAGCAGCGATGCGAC-3' (SEQ ID NO: 548) and EGFR-RV 5'CTGTCCATCCAGAGG AGGAGTA-3' (SEQ ID NO: 549) primers (25).

Copy number variations analyses have been performed on 192 tissue samples using CGH arrays using BAC arrays (N=187), Agilent 4x180K (N=2), Nimblegen 3x720K (N=2), Agilent 8x60K (N=1). Results were normalized using control DNA from matched blood samples as previously described (26). Additional analyses of 193 tumor specimens were performed by SNP array, using Illumina Omni (N=110), Illumina HumCore (N=32), Illumina 370K (N=27), or Illumina 610K (N=24), as previously described (27). Array processing was outsourced to Integragen. Raw copy numbers were estimated at each of the SNP and copy-number markers. Biodiscovery property SNP-FASST2 algorithm was then used to segment copy number data. Segments were mapped to hg18 genome assembly (28). Copy number alterations (CAN) magnitudes called log-R ratio (LRR) were classified using simple thresholds: deletion ($x \leq -1$), loss ($-1 < x \leq -0.2$), gain ($0.2 \leq x < 1$) or amplification ($x \geq 1$) according to default Nexus 7.5 software. For additional 56 gliomas, 10q loss was assessed on tumor and blood DNA by microsatellite analysis, while amplification of EGFR, MDM2 and CDK4, and deletion of CDKN2A gene, were determined by qPCR, as previously reported (29, 30).

The molecular profiles obtained in Pitié-Salpêtrière dataset were combined with those available in the TCGA dataportal. TCGA GBM segmented copy number variation profile was downloaded from The UCSC Cancer Genomics Browser (31). Copy Number Variations (CNVs) were measured experimentally using the Affymetrix Genome-Wide Human SNP Array 6.0 platform at the Broad TCGA genome characterization center (32). Raw copy numbers were estimated at each of the SNP and copy-number markers. Circular binary segmentation was then used to segment the copy number data (28). Segments are mapped to hg18 genome assembly at Broad.

For CNV analysis of the regions across FGFR3 and TACC3 genes, samples for which RNAseq and CNV data were available or samples for which only CNV data were available and RT-PCR-sequencing of FGFR3-TACC3 fusion had been performed were considered. Overall, 158 GBM (all with a wild type IDH1 gene) satisfied these criteria. Among them, 5 harbored an FGFR3-TACC3 fusion whereas 153 were FGFR-TACC-negative. The CNV magnitudes, called log-R ratio (LRR), were classified using the following thresholds: deletion ($x < -1$), loss ($-1 \leq x \leq -0.2$), gain ($0.2 \leq x \leq 1$) or amplification ($x > 1$), according to the Atlas-TCGA (32). The analysis of the genomic regions encompassing EGFR, MDM2, CDK4, CDKN2A, 7p, 10q, according to hg18 genome assembly, was performed to evaluate their CNV. EGFRvIII mutation status was inferred according to Brennan et al. (32). The frequencies of the aberrations of these genes in FGFR3-TACC3 positive and negative samples were calculated and the obtained data were then combined with the Pitié-Salpêtrière Hospital dataset.

Statistical Analysis. Differences in the distribution on categorical variables were analyzed using Fisher Exact test. The *p*-values were adjusted for multiple testing according to Benjamini and Hochberg false discovery rate (FDR). A *q*-value ≤ 0.05 (two-sided) was considered to be statistically significant.

Overall survival (OS) was defined as the time between the diagnosis and death or last follow-up. Patients who were still alive at the last follow-up were considered as censored events in the analysis. Progression-free survival (PFS) was defined as the time between the diagnosis and recurrence or last follow-up. Patients

who were recurrence-free at the last follow-up were considered as censored events in the analysis. Survival curves were calculated by the Kaplan-Meier method and differences between curves assessed using the Log-Rank test. A Log-Rank test p-value ≤ 0.05 (two-sided) was considered to be statistically significant.

Cell culture and cell growth assay. GIC-1123 gliomaspheres were cultured in neurobasal medium (Invitrogen) supplemented with B27, N2 (Invitrogen), EGF and FGF2 (20 ng/ml, PeproTech). Mouse astrocytes Ink4A-Arf-/- were cultured in DMEM supplemented with 10% Fetal Bovine Serum. Cells were seeded at 1,000 cells/well in a 96-well plate and treated with JNJ-42756493. After 72 hours cell viability was assessed using the 3-(4,5- dimethythiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) assay. Data are mean ± SEM of six replicates. Experiments were performed three times.

Subcutaneous xenografts and drug treatment. GIC-1123 cells ($5x10^5$) were injected subcutaneously in the flank of athymic nude (Nu/Nu) mice (Charles River Laboratories). Mice carrying ~200 mm^3 subcutaneous tumors were randomized to receive 12 mg/kg JNJ-42756493 or DMSO in 1% Tween 80 by oral gavage. Tumor diameters were measured with caliper and tumor volumes estimated using the formula: 0.5 x length x $width^2$. Data are mean ± SD of nine mice in each group. Mice were sacrificed when tumors in the control group reached the maximal size allowed by the IACUC Committee at Columbia University.

MRI imaging and evaluation of clinical response to JNJ-42756493. Baseline and follow-up imaging assessments were performed on 1.5 Tesla MR imaging systems, including at least axial T1 weighted images before gadolinium injection, Axial or 3D FLAIR (Fluid-Attenuated Inversion-Recovery), dynamic susceptibility contrast MR perfusion (0.1 mmol/kg of gadobutrol), axial and 3D T1 weighted images after gadolinium injection. Tumor response was assessed according to the RANO criteria (33). Contrast-enhancing lesion volume was assessed with the help of a semi-automated volumetry tool (SegmentiX), based on shape-detection and thresholding, with control and manual correction of edges when necessary. Since exclusion of cystic or necrotic portions of the lesion may be affected by operator subjectivity, both were included for volumetric and axial measurements.

DSC (dynamic susceptibility contrast) perfusion datasets were processed with vendor's software suite (Neuroperfusion, Philips), including coregistration and rCBV (relative cerebral blood volume) parametric maps generation with 3 different algorithms (Gamma-variate fitting, Arterial Input Function based deconvolution and Model Free).

RESULTS

Detection of FGFR1-TACC1 and FGFR3-TACC3 fusions in GBM and grade II-III glioma. To determine the frequency and molecular features of FGFR-TACC fusions in human glioma patients, a cohort of 584 GBM and 211 grade II-III glioma treated at five Neurooncology centers (**Table 12**) were screened. 108 were grade III (49 IDH wild type, 52 IDH1 mutant and 7 IDH2 mutant) and 103 were grade II (36 IDH wild type, 63 IDH1 mutant and 4 IDH2 mutant). The IDH mutational status of 333 GBM was also established and it was determined that 303 harbored wild type IDH1/2 and 30 were mutated at codon 132 of IDH1. A RT-PCR assay was designed for the detection of all known and possibly new variants of FGFR1-TACC1 and FGFR3-TACC3 fusions that retain the mRNA sequences coding for the key FGFR-TK and

TACC domains required for the oncogenic activity of the fusion protein (Fig. 19 and Figs. 20A-D). Overall, 20 tumors with an FGFR3-TACC3 fusion were found, of which 17 were GBM (2.9% positives) and 3 lower grade glioma harboring wild type IDH1/2 genes (3.5% positives). The size of the FGFR3-TACC3 RT-PCR amplicons ranged from 928 bp (for FGFR3ex18-TACC3ex13) to 1706 bp (for FGFR3ex18-TACC3ex4). The FGFR1-TACC1 fusion was detected in one grade II IDH wild type glioma (Fig. 19). Conversely, an IDH1/2 mutant glioma harboring FGFR-TACC fusions ($p<0.02$) was not found. Sanger sequencing of the fusion amplicons revealed that each FGFR-TACC cDNA joined in-frame the sequence coding for the entire TK domain upstream of TACC-coding sequences that invariably include the coiled-coil TACC domain (Fig. 19). However, a notable variability among FGFR3-TACC3 fusion isoforms was detected, whereby 5 of the identified variants occurred only in individual cases (Fig. 19). Furthermore, 6 fusion transcripts emerged as new variants that have not been reported before in human cancer (marked in red in Fig. 19).

Next, suitable PCR primers were designed to map the genomic breakpoint coordinates for 9 FGFR3-TACC3-positive samples for which genomic DNA was available (Figs. 23 and 24). The genomic breakpoints were successfully reconstructed by Sanger sequencing and found that they differ for each of the 9 positive cases. Interestingly, even cases harboring the same FGFR3-TACC3 transcript splice variants (#4451 and #OPK-14 joining exon 17 of FGFR3 to exon 6 of TACC3; #3048 and #4373 joining exon 17 of FGFR3 to exon 8 of TACC3; #3808 and #27-1835 joining exon 17 of FGFR3 to exon 11 of TACC3) had different genomic breakpoints (Fig. 24). Taken together, the above findings indicate that the noticeable variability among FGFR3-TACC3 fusion transcripts and genomic breakpoints is efficiently resolved by the RT-PCR screening assay.

Immunostaining analysis of FGFR3-TACC3-positive tumors. The expression of the FGFR3 fusion protein was analyzed by IHC or IF using an antibody that recognizes the N-terminal region of FGFR3 (FGFR3-N) in 12 GBM and 3 lower grade glioma harboring FGFR3-TACC3 fusions for which sufficient tissue was available. Remarkably, each of the 15 positive tumors but none of those that had scored negative in the RT-PCR assay, displayed strong positivity for FGFR3 in the vast majority of tumor cells but not endothelial cells throughout the analyzed tumor section (Figs. 20A-H). Notably, IF using an antibody that recognizes an epitope at the C-terminus of TACC3, which is invariably retained within FGFR3-TACC3 variants (TACC3-C), reproduced the staining pattern of the FGFR3-N antibody in FGFR3-TACC3 positive tumors. Conversely, negative or very weak staining was obtained in FGFR3-TACC3-positive tumors with antibodies recognizing the regions of FGFR3 (FGFR3 C-terminal region, FGFR3-C) and TACC3 (TACC3 N-terminal region, TACC3-N) constantly excluded from FGFR3-TACC3 fusion proteins (Fig. 25A). Consistently, quantitative RT-PCR of GBM harboring FGFR3-TACC3 fusions showed that the expression of the N-terminal coding region of FGFR3 and the C-terminal coding region of TACC3 (which are included in the fusion genes) is markedly higher than the expression of the C-terminal coding region of FGFR3 and the N-terminal coding region of TACC3, which are excluded from the fusion transcripts (Fig. 25B). One recurrent GBM from a patient whose tumor had been found positive for FGFR3-TACC3 at the initial diagnosis and who had recurred after concurrent radiotherapy and temozolomide treatment was analyzed.

The recurrent tumor retained the same FGFR3-TACC3 fusion gene and protein that was present in the untreated GBM as determined by RT-PCR-sequencing and FGFR3 IF, respectively (Fig. 26A-C). Although this requires additional evaluation, the retained uniform positivity for FGFR3 in this recurrent GBM suggests that targeting the FGFR3-TACC3 fusion protein at relapse is a valid therapeutic strategy.

Clinical and molecular characteristics of glioma patients with FGFR3-TACC3 fusions. Clinical and molecular profiling data were available for 591 patients including 380 GBM (9 with FGFR3-TACC3 fusions) and all 211 lower grade glioma (3 with FGFR3-TACC3 fusions). Of these 12 patients 5 are males and 7 females, aged 48y to 82y (median= 61y). The molecular profile of FGFR3-TACC3-positive glioma was determined. To do so, the analysis of CNVs and somatic mutations of key GBM genes in the dataset was combined with the SNP6.0 high-density genomic array analysis of 158 TCGA-derived GBM samples fully annotated for FGFR3-TACC3 fusion genes (the RNA-seq and/or RT-PCR analysis of these samples had revealed that 5 of them harbor FGFR3-TACC3 fusions) (6). Patients with FGFR3-TACC3 fusions displayed unique characteristics (Table 13). FGFR3-TACC3 fusions were mutually exclusive with EGFR amplification (0/16 vs. 166/411; p=0.0004, FDR q-value corrected for multiple comparisons=0.0012) and showed a clear trend against the presence of the EGFRvIII transcript variant (0/16 vs. 37/219; p=0.083). Conversely, CDK4 amplification was significantly more frequent in FGFR3-TACC3-positive tumors (7/16 vs 41/408, p=0.0008; FDR q-value= 0.0024). A less significant association of FGFR3-TACC3 fusions was also seen with amplification of MDM2, which as CDK4, maps to chromosome 12q (4/16 vs 24/408, p=0.016; FDR q-value= 0.048). No statistical association between FGFR3-TACC3 fusions and other genetic and epigenetic alterations that commonly occur in gliomas harboring wild type IDH genes was found (CDKN2A deletion, TERT promoter mutations, gain of chromosome 7p, loss of chromosome 10q and methylation of the MGMTpromoter, Table 13). When compared with the IDH wild type patient population of grade II and grade III glioma and GBM, there was no significant difference in progression free survival (PFS) or overall survival (OS) between patients positive or negative for FGFR3-TACC3 (Figs. 27A-B). Finally, it was established whether the CNV analysis of the FGFR3 and TACC3 genomic loci could be used to predict positivity for FGFR3-TACC3 fusions. The analysis of high-density SNP6.0 arrays of the 158 GBM samples from the Atlas-TCGA revealed that 10 samples displayed different degrees of copy number gains encompassing the entire FGFR3 and TACC3 loci (Fig. 28). However, none of them harbored FGFR3-TACC3 fusions. Conversely, the 5 FGFR3-TACC3-positive samples in the dataset harbor micro-amplification events involving only the exons of the FGFR3 gene that are included in the fusion breakpoint. This finding suggests that any CNV survey that is less accurate than high-density SNP arrays, could fail to identify the genomic marks associated with true FGFR3-TACC3-positive cases.

Table 13 - Molecular alterations in IDH wild type glioma harboring FGFR3-TACC3 fusions. The table reports the absolute number and frequency (percentage) of individual glioma-specific molecular alterations in tumors scoring positive or negative for FGFR3-TACC3 fusions. The analysis is done on the Union dataset (TCGA and "Pitié-Salpêtrière Hospital" datasets, see methods for details). Statistically significant associations are indicated in bold (Fisher Exact test, q-values adjusted with FDR).

	N of FGFR3-TACC3 Positive	% of FGFR3-TACC3 Positive	N of FGFR3-TACC3 Negative	% of FGFR3-TACC Negative	*P*-value (Fisher test)	*q*-value (FDR)
EGFR amplification	0/16	0.0 %	166/411	40.4 %	4.E-04	0.0012
CDK4 amplification	7/16	43.7 %	41/408	10.0 %	8.E-04	0.0024
MDM2 amplification	4/16	25.0 %	24/408	5.9 %	0.016	0.048
EGFRvIII	0/16	0.0 %	37/219	16.9 %	0.083	0.25
CDKN2A deletion	4/16	25.0 %	188/411	45.7 %	0.13	0.39
Chr. 7p gain	12/15	80.0 %	242/374	64.7 %	0.28	0.84
Chr. 10q deletion	12/16	75.0 %	253/420	60.2 %	0.3	0.9
TERT promoter mutation	9/11	81.8 %	128/163	78.5 %	0.8	1
MGMT promoter hypermethylation	6/12	50.0%	73/160	45.6 %	0.7	1

Preclinical and clinical relevance of targeting FGFR3-TACC3 fusions. JNJ-42756493 is a potent, oral pan-FGFR tyrosine kinase inhibitor with IC50 values in the low nanomolar range for all members of the FGFR family. It has demonstrated potent antitumor activities in nonclinical models with FGFR aberrations including squamous non-small cell lung cancer, gastric, breast, hepatocellular cancer (HCC), endometrial, and bladder (34, 35). To ask whether JNJ-42756493 is effective in targeting specifically FGFR-TACC-positive cells, mouse astrocytes expressing FGFR3-TACC3, FGFR3-TACC3 containing a mutation that inactivates the kinase activity of FGFR3 (FGFR3-TACC3-KD), or the empty vector were treated with JNJ-42756493. The effect of JNJ-42756493 on human glioma stem cells GIC-1123 that harbor the FGFR3-TACC3 gene fusion (6) was also studied. These experiments revealed that both mouse astrocytes and GIC-1123 that express FGFR3-TACC3 but not cells expressing the KD mutant fusion or the empty vector are highly sensitive to FGFR inhibition by JNJ-42756493 with an IC50 of 3.03 nM and 1.55 nM, respectively (Figs. 21A-B). Next, the effect of oral treatment with JNJ-42756493 of mice bearing xenografts of human GIC-1123 affects tumor growth was tested. Mice were randomized to receive vehicle or JNJ-42756493 (12 mg/kg). Mirroring the in vitro results, JNJ-42756493 elicited a potent growth inhibition of GIC-1123 tumor xenografts (Figs. 21C-D) with a statistically significant tumor regression after two weeks (p-value of the slope calculated from the treatment starting point = 0.04). The above findings provide a strong foundation for the treatment of GBM patients harboring FGFR-TACC rearrangements with JNJ-42756493.

Two patients with recurrent GBM harboring FGFR3-TACC3 fusions were treated with JNJ-42756493 in a first-in-man phase I trial. Patient 1, male aged 52, underwent partial surgical resection of a right parietal GBM, followed by fractionated radiotherapy and concomitant temozolomide (TMZ) as first line treatment (36). The RT-PCR-sequencing analysis of the GBM specimen revealed positivity for the FGFR3-TACC3 fusion (FGFR3-exon17-TACC3-exon 6, sample 4451, Figs 40 and 41) and the immunostaining using FGFR3 antibody on paraffin embedded sections showed strong positivity in a large fraction of tumor cells.

After 5 cycles of TMZ, the patient presented with dizziness and headache and brain MRI revealed tumor progression (Figs. 22A). At this time the patient was enrolled in the JNJ-42756493 trial and received JNJ-42756493 (12 mg/day administered in cycles of 7 days followed by 7 days off treatment). After 3 weeks the patient reported a marked clinical improvement (complete regression of dizziness and headache). On MRI, the sum of product diameters (RANO criteria, Fig. 22B) and volumetry (Fig. 22C) measured without excluding cystic and necrotic components showed disease stabilization. However, the tumor mass underwent significant decrease of the enhancing parenchyma (-44%) with formation of a cystic portion in the central core (33). The objective response was further corroborated by the marked reduction of the extent of tumor vascularity estimated by quantitative analysis of rCBV (relative cerebral blood volume) from dynamic susceptibility MR perfusion maps (37) (Fig. 22D). Stabilization lasted for 115 days. During JNJ-42756493 treatment mild and manageable toxicity was observed (grade I hyperphosphatemia, asthenia, dysgueusia, dry mouth, keratitis, and grade II nail changes). After 4 months, tumor progressed on MRI locally both on T1 contrast-enhanced area and T2/FLAIR hypersignal. The patient was re-operated and subsequently treated with CCNU. He is still alive, but in progression after 21 months from diagnosis and 287 days from the start of the anti-FGFR therapy.

Patient 2 is a 64 years old woman, affected by left parietal GBM, diagnosed by stereotactic biopsy. The tumor was positive for FGFR3-TACC3 gene fusion by RT-PCR-sequencing and showed diffuse FGFR3 expression in most tumor cells (Figs. 20A, 20C, 20E, sample 4620). The patient received as first line treatment fractionated radiotherapy and TMZ according to the Stupp protocol (36), but after 2 cycles of monthly TMZ she presented with clinical deterioration including progressive headaches, right homonymous hemianopsia and memory impairment. Brain MRI performed 3 and 4 months after the completion of concomitant chemo-radiotherapy revealed tumor progression with increase of the left parietal mass and the appearance of a small contralateral lesion (Fig. 22E). The patient was thus enrolled in the JNJ-42756493 trial (12 mg/day administered in cycles of 7 days followed by 7 days off treatment) and showed clinical improvement after 4 weeks (regression of headaches, visual field defect and memory impairment). Best response was observed after 104 days of treatment with a 22% reduction of tumor size according to the RANO criteria (Fig. 22F) and 28% according to volumetry (Fig. 22G). Grade I hyperphosphatemia, nail changes, and mucositis were observed. Clinical status remained stable until disease progression occurring 134 days after the start of the anti-FGFR. The patient is still alive and is receiving a third-line chemotherapy with nitrosoureas and bevacizumab.

FGFR-TACC Fusion Variants		N Cases	Tumor Type
FGFR3-TACC3	FGFR3exon17-TACC3exon11	30	Brain Tumors, N=10 (N=2,[6]; N=2,[9,15]; N=6, Present Study).
			Bladder Cancer, N=6 (N=3,[12,13]; N=3,[11]).
			Lung Cancer, N=13 (N=4,[12,15]; N=9,[10]).
			Renal Carcinoma, N=1,[15].
	FGFR3exon17-TACC3exon10	18	Brain Tumors, N=5 (N=1,[8]; N=1,[9]; N=3, Present study).
			Oral Cancer, N=1,[14].
			Head and Neck Cancer, N=2,[12,15].
			Bladder Cancer, N=3,[7].
			Lung Cancer, N=7 (N=4,[8]; N=2,[10]; N=1,[14]).
	FGFR3exon17-TACC3exon8	8	Brain Tumors, N=6 (N=2,[8], N=4, Present study).
			Lung Cancer, N=2 (N=1,[10]; N=1,[14]).
	FGFR3exon17-TACC3exon4	4	Brain Tumors, N=2 (N=1,[9]; N=1,[13]).
			Bladder Cancer, N=1[9].
			Lung Cancer, N=1[14].
	FGFR3exon17-TACC3exon6	2	Brain Tumors, N=2, Present study.
	FGFR3exon18-TACC3exon4	1	Brain Tumors, N=1, Present study.
	FGFR3exon17-TACC3exon9 INS63bp	1	Brain Tumors, N=1,[8].
	FGFR3exon18-TACC3exon9 INS66bp	1	Brain Tumors, N=1, Present study.
	FGFR3exon18-TACC3exon5	1	Brain Tumors, N=1, Present study.
	FGFR3exon18-TACC3exon5 INS33bp	1	Brain Tumors, N=1, Present study.
	INS71bp	1	Lung Cancer, N=1,[10].
	FGFR3exon18-TACC3exon13	1	Brain Tumors, N=1, Present study.
	FGFR3exon18-TACC3exon11	1	Lung Cancer, N=1,[10].
FGFR1-TACC1	FGFR1exon17-TACC1exon7	5	Brain Tumors, N=5 (N=1,[8]; N=3,[13]; N=1, Present study).
FGFR2-TACC2		1	Stomach Adenocarcinoma, N=1[15].

Table 14 –Summary of FGFR-TACC fusion transcripts identified in all cancer types. FGFR3-TACC3 fusion variants are ranked according to their prevalence across any cancer type. The number of FGFR-TACC fusions identified in each tumor type, including those identified in the present study, is also indicated.

DISCUSSION

FGFR-TACC fusions are potent oncogenic events that when present in brain tumor cells confer sensitivity to FGFR inhibitors (6). Since the original identification of recurrent FGFR-TACC fusions in GBM, small subgroups of patients harboring FGFR-TACC translocations have been identified in several other tumor types (7-15). Here, an unbiased RT-PCR-sequencing analysis for the identification of all possible functional FGFR-TACC fusion transcripts is reported. The screening of a large glioma dataset from multiple Institutions not only confirmed that FGFR-TACC rearrangements occur in ~3% of human GBM but also revealed that FGFR-TACC fusions are present in the subgroup of IDH wild type lower grade glioma (grade II-III) with prevalence similar to that of GBM. IDH wild type grade II and III glioma have a significantly worse clinical outcome than the IDH mutant glioma and manifests molecular and clinical features that resemble GBM (5). The finding that FGFR-TACC fusions occur in IDH wild type but not IDH mutant glioma provides an important clue for the molecular characterization of this glioma subtype. Furthermore, the clustering of such potent oncogenic events in IDH wild type glioma underscores the particularly aggressive nature of this group of glioma. While it was shown that FGFR-TACC fusions cluster within the poor clinical outcome subgroup of IDH wild type glioma, these translocations do not seem to carry prognostic value within the IDH wild type subgroup of glioma patients. Without being bound by theory, the sample size of patients harboring FGFR-TACC fusions is too small to draw definitive conclusions with

respect to the impact on survival and larger studies may be necessary to clarify the prognostic role of FGFR-TACC fusions in IDH wild type glioma.

Beside mutual exclusivity between IDH1 mutations and FGFR-TACC fusions, the results showed that patients with FGFR3-TACC3 rearrangements lack EGFR amplification and EGFRvIII but are significantly enriched for amplification of CDK4 (and MDM2 to a lesser extent). Knowledge of these molecular characteristics will help select those patients who most likely harbor FGFR-TACC rearrangements and design combinatorial targeted therapies that might be more effective in the FGFR-TACC-positive glioma subgroup.

The molecular screen uncovered 6 new FGFR3-TACC3 fusion events. Together with the previously identified variants, 12 distinct isoforms of FGFR3-TACC3 have been reported, thus revealing a remarkable variability of FGFR3-TACC3 transcripts in human cancer (see Table 14 summarizing the structure of all the FGFR-TACC variants identified to date). The structural heterogeneity of FGFR3-TACC3 fusions is yet more pronounced at the genomic level, whereby each fusion event harbors distinct genomic breakpoints, even for identical fusion transcripts. This finding underscores the notion that targeted genomic analyses are unlikely to be suitable approaches for the molecular diagnosis of FGFR3-TACC3 positivity. Conversely, the unbiased identification of FGFR3-TACC3-positive tumors with the RT-PCR-sequencing assay reported here overcomes the limitations of screening only for previously identified FGFR3-TACC3 fusions and provides a simple molecular diagnostic assay.

Rather than displaying uniform amplifications of the FGFR3 and TACC3 genomic loci, FGFR3-TACC3-positive samples harbor small, intragenic micro-amplification events typically encompassing only the exons of the FGFR3 and TACC3 genes included in the breakpoint (6). This finding is consistent with the notion that a "fusion breakpoint principle" sustains the CNVs of driver gene fusions such as FGFR3-TACC3 in which local CNVs target exclusively the breakpoint region (38). It is noted that such small and irregular CNVs may easily go undetected from CNV analyses performed using platforms less sensitive than the high-density SNP6.0 genomic arrays. Furthermore, the notion that FGFR3-TACC3-negative GBM may harbor uniform amplifications across the FGFR3 and TACC3 loci argues against the standard analysis of FGFR3 and/or TACC3 CNVs as a method for the selection of FGFR3-TACC3-positive tumors.

There is a growing body of evidence supporting the notion that GBM is a markedly heterogeneous tumor. The formidable degree of intra-tumor heterogeneity of GBM is a potential cause of failure of targeted therapies in these tumors. In particular, the intra-tumor heterogeneity of GBM has previously been recognized in light of the mosaic expression of the RTK genes EGFR, PDGFRA and MET by neighboring cells (16-19). Thus, in the majority of GBM, amplification or overexpression of individual RTK genes are present in a sub-clonal fraction of tumor cells and co-exist with amplification/expression of other RTK-coding genes within the tumor mass. Therefore, it was essential to determine whether such heterogeneity was also present in gliomas harboring FGFR-TACC translocations. The immunostaining of FGFR3-TACC3-positive tumors revealed that positive specimens manifest strong and uniform expression of the fusion protein, which is also retained after recurrence. This behavior is reminiscent of other driver

chromosome translocations (BCR-ABL, EML4-ALK) and is compatible with the glioma-initiating functions of FGFR-TACC fusions (6). It is also the scenario expected for a driver oncogene whose activity remains essential for tumor maintenance regardless of secondary genetic alterations that occur during tumor progression. The strong antitumor effects obtained with JNJ-42756493 in glioma cells harboring FGFR3-TACC3 fusions have built a compelling rationale for the treatment of glioma patients positive for FGFR-TACC rearrangements. JNJ-42756493 is an oral ATP-competitive pan-FGFR selective inhibitor that inhibits tyrosine phosphorylation of activated FGFR at nanomolar concentrations (34, 35). The enrollment of two patients with recurrent FGFR3-TACC3-positive GBM in a phase I trial with JNJ-42756493 showed that this treatment has tolerable toxicity and clear anti-tumor activity, thus validating FGFR-TACC as a therapeutic target. Therefore, targeted inhibition of FGFR-TK in preselected IDH wild type FGFR-TACC-positive glioma may provide clinical benefits for patients with recurrent glioma who currently lack valuable therapeutic options. In conclusion, described herein is the importance and feasibility of prospective genotyping for FGFR-TACC fusions in glioma patients and provided a preliminary evidence of clinical response that warrants the investigation of the sensitivity of gliomas harboring FGFR-TACC rearrangements to FGFR kinase inhibition in clinical trials.

What is claimed is:

1. A method for detecting the presence of a FGFR fusion protein and the presence or absence of a gene amplification of EGFR, CDK4 or MDM2 in a human subject, wherein the FGFR fusion protein comprises the tyrosine kinase domain of a FGFR protein fused to the TACC domain of a TACC protein, the method comprising:

   (a) obtaining a biological sample from the human subject;

   (b) detecting whether or not there is a FGFR fusion present in the subject; and

   (c) detecting whether or not there is gene amplification of EGFR, CDK4, or MDM2.

2. The method of claim 1, wherein the detecting of step (b) comprises measuring FGFR fusion protein levels by ELISA using an antibody directed to SEQ ID NO: 79, 85, 86, 87, 88, 89, 150, 158, 159, 160, 161, 542, 543, 544, 545, 546, or 547; western blot using an antibody directed to SEQ ID NO: 79, 85, 86, 87, 88, 89, 150, 158, 159, 160, 161, 542, 543, 544, 545, 546, or 547; mass spectroscopy, isoelectric focusing; immunostaining using an antibody directed to the N-terminus of a FGFR protein or to the C-terminus of a TACC protein, or a combination thereof; or

   wherein the detecting of step (b) comprises detecting whether or not there is a nucleic acid sequence encoding a FGFR fusion protein in the subject; and optionally wherein the nucleic acid sequence comprises any one of SEQ ID NOS: 1-77, 80-84, 95-145, 515, 517, 519-527, or 533-538; or

   wherein the detecting of step (b) comprises using hybridization, amplification, or sequencing techniques to detect a FGFR fusion protein; and optionally wherein the amplification uses primers comprising SEQ ID NO: 162, 163, 164, 165, 166, 167, 168, 169, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, or 514.

3. The method of claim 1, wherein the detecting step of step (c) comprises measuring gene amplification by microarray or quantative PCR (qPCR); and optionally wherein the microarray is a CGH array or a SNP array.

4. The method of claim 1, further comprising (d) detecting whether or not there is a mutation in IDH1 and/or, IDH2.

5. The method of claim 1, further comprising (d) diagnosing the human subject with a FGFR fusion associated cancer if a FGFR fusion is detected in the biological sample and the EGFR gene is not amplified, and/or the CDK4 gene is amplified, and/or the MDM2 gene is amplified in the biological sample.

6. The method of claim 5, further comprising (e) administering a treatment for a FGFR fusion associated cancer to the human subject if the human subject is diagnosed with a FGFR fusion associated cancer.

7. A diagnostic kit for determining whether a sample from a subject exhibits a presence of a FGFR fusion protein and the presence or absence of a gene amplification of EGFR, CDK4 or MDM2, the kit comprising at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding FGFR fusion protein, or a portion thereof, wherein the FGFR fusion protein comprises the tyrosine kinase domain of a FGFR protein fused to the TACC domain of a TACC protein; the kit further comprising at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding EGFR, CDK4, or MDM2.

8. The kit of claim 7, wherein the at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding FGFR fusion protein comprises a set of nucleic acid primers or *in situ* hybridization probes; and optionally wherein the primers prime a polymerase reaction only when a nucleic acid encoding a FGFR fusion is present; or

wherein the at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding FGFR fusion protein comprises SEQ ID NO: 162, 163, 164, 165, 166, 167, 168, 169, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, or 514, or a combination thereof; or

wherein the at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding FGFR fusion protein are for use in gene sequencing, selective hybridization, selective amplification, gene expression analysis, or a combination thereof; or

wherein the at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding EGFR, CDK4, or MDM2 comprises a set of nucleic acid primers; and optionally wherein the primers are for use in quantitative PCR (qPCR).

9. A diagnostic kit for determining whether a sample from a subject exhibits a presence of a FGFR fusion protein and the presence or absence of a gene amplification of EGFR, CDK4 or MDM2, the kit comprising an antibody that specifically binds to a FGFR fusion protein, wherein the antibody will recognize the protein only when a FGFR fusion protein is present, and wherein the FGFR fusion protein comprises a tyrosine kinase domain of an FGFR protein fused to the TACC domain of a transforming acidic coiled-coil-containing (TACC) protein; the kit further comprising at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding EGFR, CDK4, or MDM2.

10. The kit of claim 9, wherein the antibody is directed to the FGFR N-terminus or the TACC C-terminus; or

wherein the antibody specifically binds to a FGFR fusion protein comprising SEQ ID NO: 79, 85, 86, 87, 88, 89, 150, 158, 159, 160, 161, 542, 543, 544, 545, 546, or 547; or

wherein the at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding EGFR, CDK4, or MDM2 comprises a set of nucleic acid primers; and optionally wherein the primers are for use quantitative PCR (qPCR).

11. An antibody or antigen-binding fragment thereof, that specifically binds to a purified FGFR3-TACC3 fusion protein comprising the amino acid sequence of SEQ ID NO: 542, 543, 544, 545, 546, or 547.

12. A composition for decreasing in a subject the expression level or activity of a FGFR3-TACC3 fusion protein comprising the amino acid sequence of SEQ ID NO: 542, 543, 544, 545, 546, or 547, the composition in an admixture of a pharmaceutically acceptable carrier comprising an inhibitor of the fusion protein.

13. A method of decreasing growth of a solid tumor in a subject in need thereof, the method comprising administering to the subject an effective amount of a FGFR3-TACC3 fusion protein molecule inhibitor, wherein the inhibitor decreases the size of the solid tumor, and wherein the FGFR3-TACC3 fusion protein comprises the amino acid sequence of SEQ ID NO: 542, 543, 544, 545, 546, or 547.

14. The method of claim 13, wherein the solid tumor comprises glioma, glioblastoma multiforme, breast cancer, lung cancer, prostate cancer, or colorectal carcinoma; and optionally wherein the glioma is grade II or grade III glioma; and further optionally wherein the grade II or grade III glioma is IDH wild-type glioma.

15. The method of claim 13, wherein the inhibitor comprises AZD4547, NVP-BGJ398, PD173074, NF449, TK1258, BIBF-1120, BMS-582664, AZD-2171, TSU68, AB1010, AP24534, E-7080, LY2874455, or a combination thereof; or

wherein the inhibitor comprises an oral pan-FGFR tyrosine kinase inhibitor; and optionally wherein the oral pan-FGFR tyrosine kinase inhibitor comprises JNJ-42756493.

16. A purified FGFR2-TACC3 fusion protein comprising the amino acid sequence of SEQ ID NO: 542, 543, 544, 545, 546, or 547.

17. A cDNA encoding a FGFR2-TACC3 fusion protein comprising the nucleotide sequence of SEQ ID NO: 533, 534, 535, 536, 537, or 538.

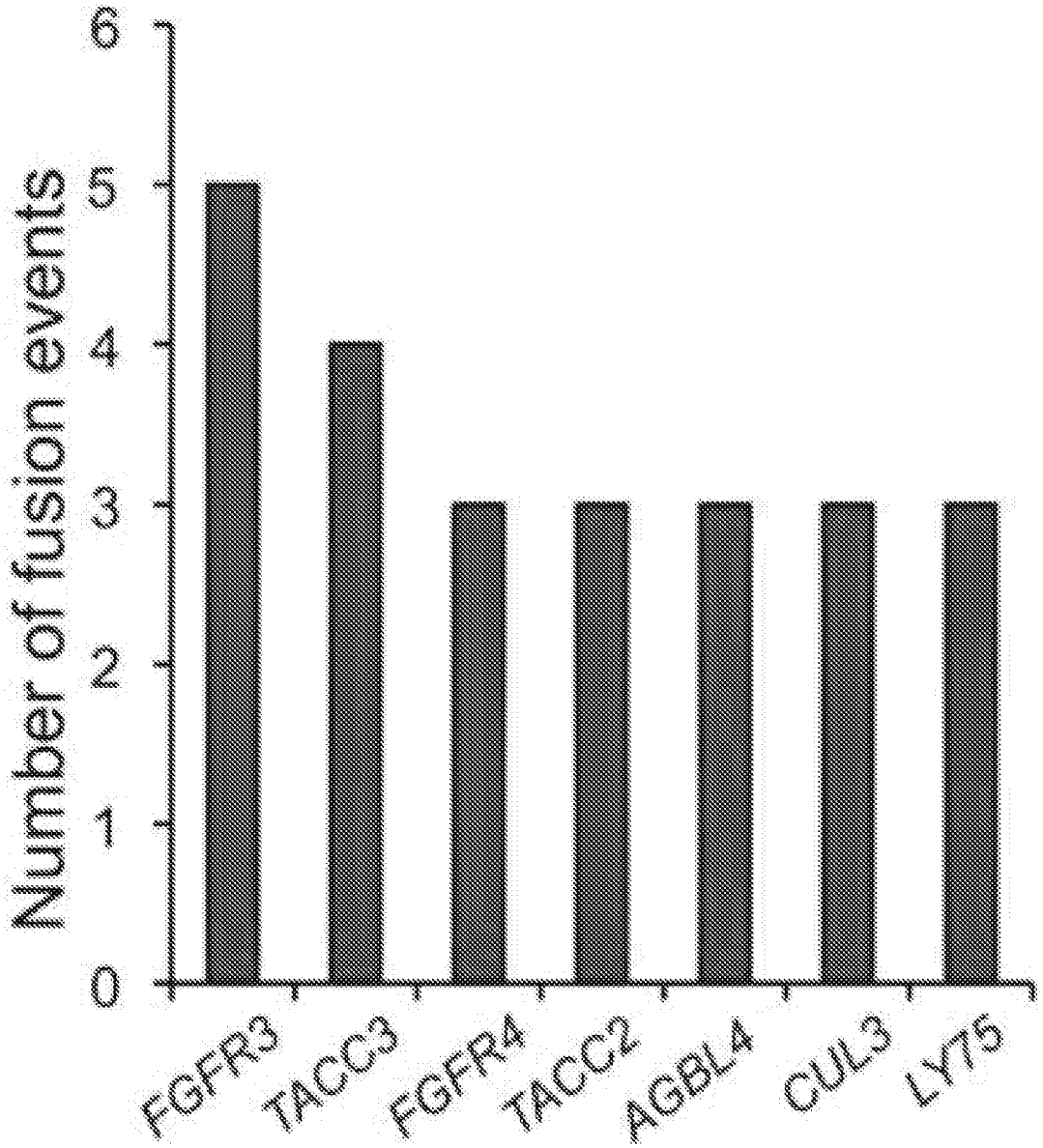

Number of fusion events
0
1
2
3
4
5
6
FGFR3
TACC3
FGFR4
TACC2
AGBL4
CUL3
LY75

FIG. 1A

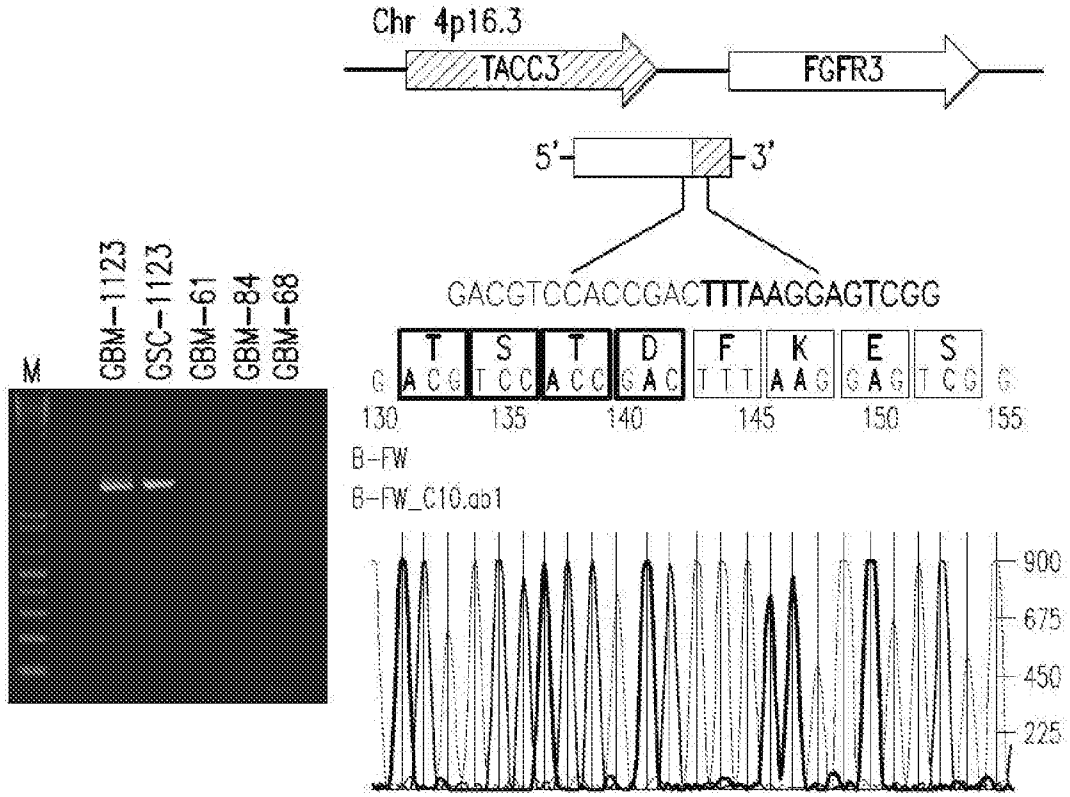

Chr 4p16.3
TACC3
FGFR3
5'
3'
GACGTCCACCGACTTTAAGGAGTCGG
M
GBM-1123
GSC-1123
GBM-61
GBM-84
GBM-68
T S T D F K E S
A C G T C C A C C G A C T T T A A G G A G T C G G
130 135 140 145 150 155
B-FW
B-FW_C10.ab1
900
675
450
225

FIG. 1B

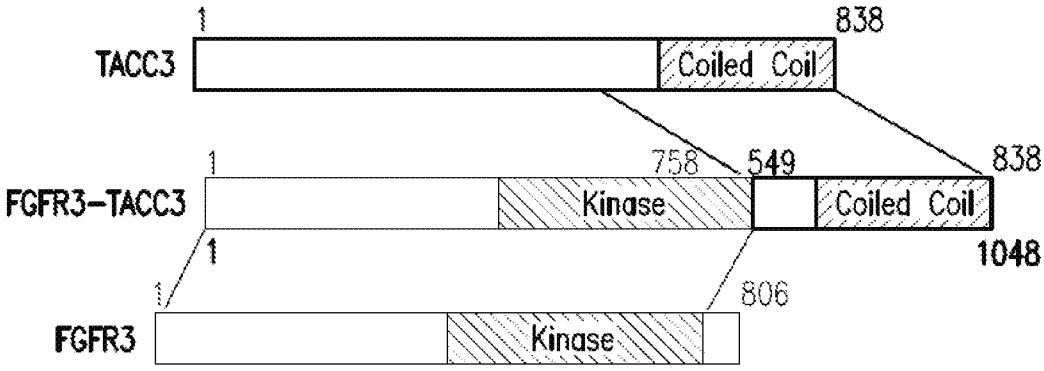

TACC3
1
838
Coiled Coil
FGFR3-TACC3
1
758
549
838
Kinase
Coiled Coil
1048
FGFR3
1
806
Kinase

FIG. 1C

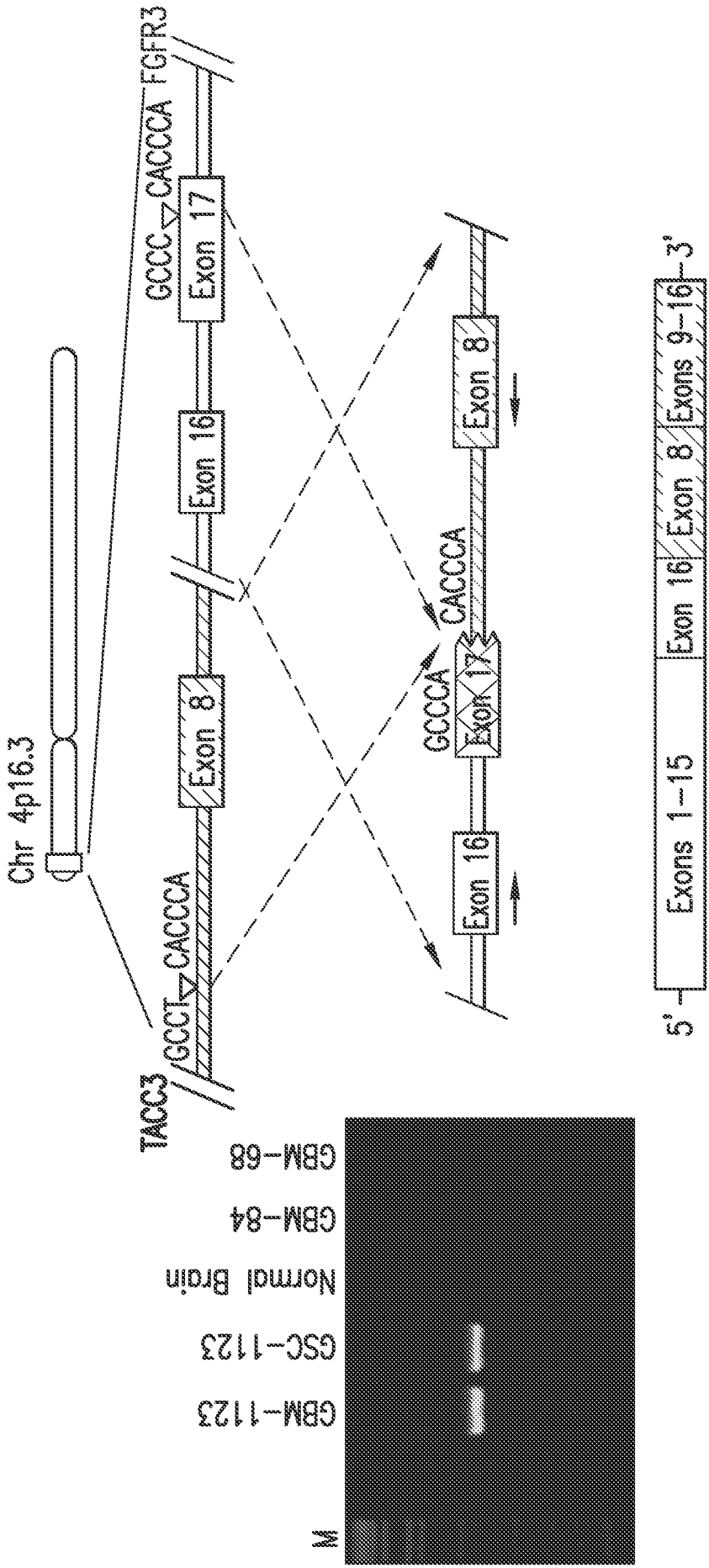

Chr 4p16.3
TACC3
GCCT
CACCCA
Exon 8
Exon 16
GCCC
CACCCA
Exon 17
FGFR3
GCCCA
CACCCA
Exon 17
Exon 8
Exon 16
5'
Exons 1–15
Exon 16
Exon 8
Exons 9–16
3'
M
GBM–113
GSC–113
Normal Brain
GBM–84
GBM–68

FIG. 1D

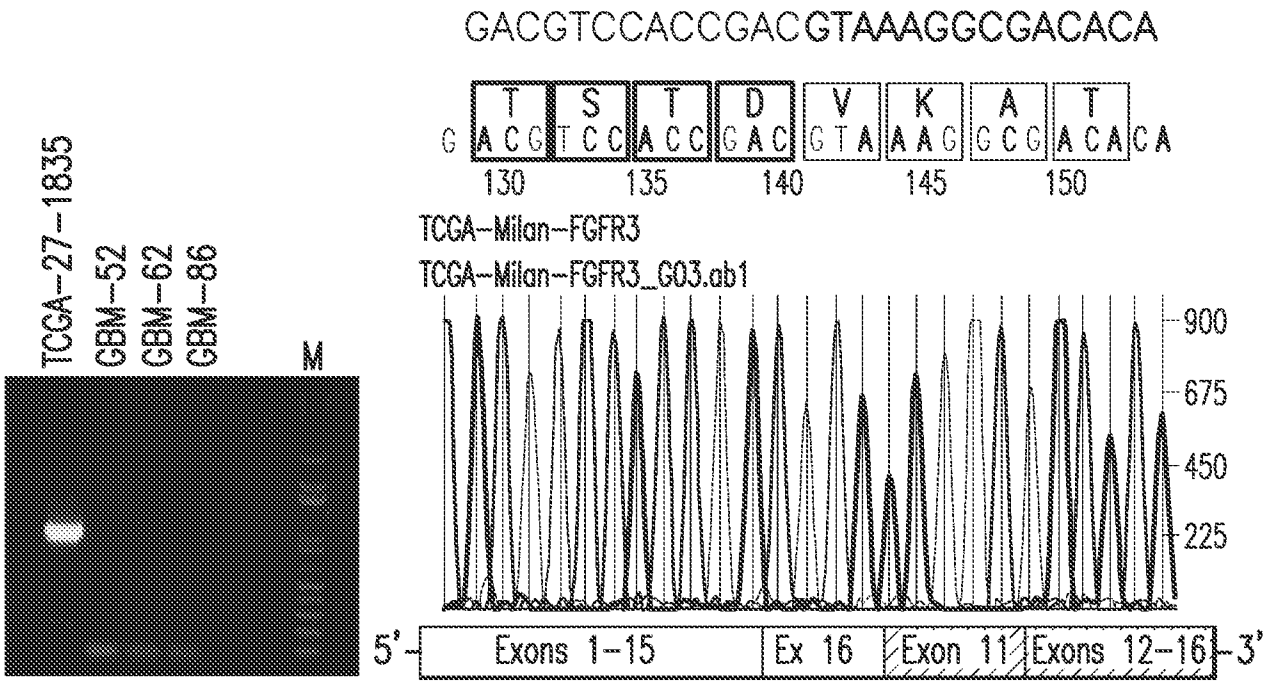

TCGA-27-1835
GBM-52
GBM-62
GBM-86
M
GACGTCCACCGACGTAAAGGCGACACA
T S T D V K A T
G ACG TCC ACC GAC GTA AAG GCG ACA CA
130 135 140 145 150
TCGA-Milan-FGFR3
TCGA-Milan-FGFR3_G03.ab1
900
675
450
225
5'
Exons 1-15
Ex 16
Exon 11
Exons 12-16
3'

FIG. 2A

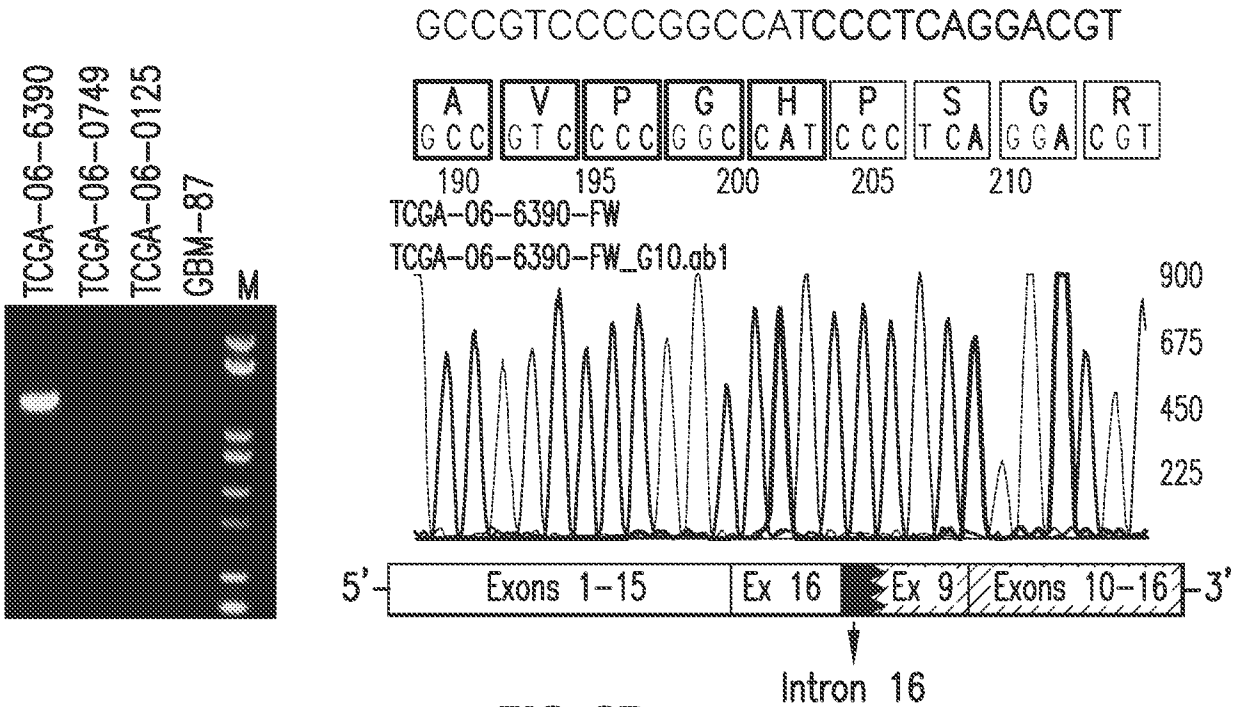

TCGA-06-6390
TCGA-06-0749
TCGA-06-0125
GBM-87
M
GCCGTCCCCGGCCATCCCTCAGGACGT
A V P G H P S G R
GCC GTC CCC GGC CAT CCC TCA GGA CGT
190 195 200 205 210
TCGA-06-6390-FW
TCGA-06-6390-FW_G10.ab1
900
675
450
225
5'
Exons 1-15
Ex 16
Ex 9
Exons 10-16
3'
Intron 16

FIG. 2B

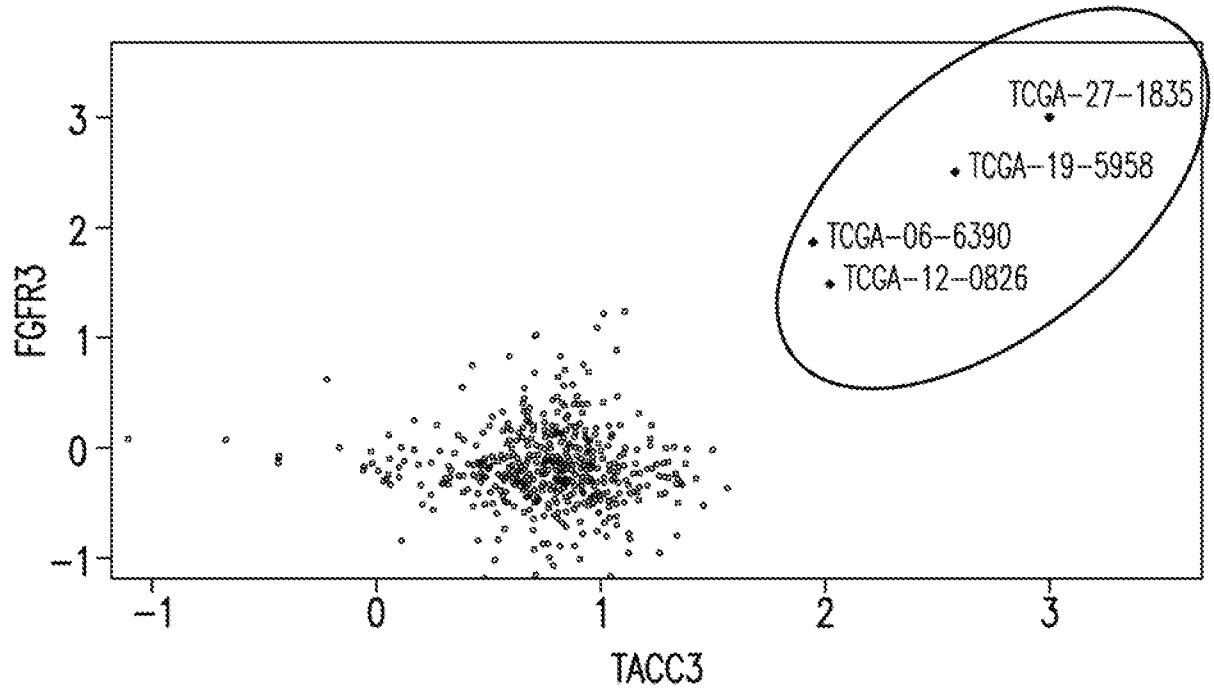

TCGA-27-1835
TCGA-19-5958
TCGA-06-6390
TCGA-12-0826
FGFR3
TACC3
3
2
1
0
−1
−1
0
1
2
3

FIG. 2C

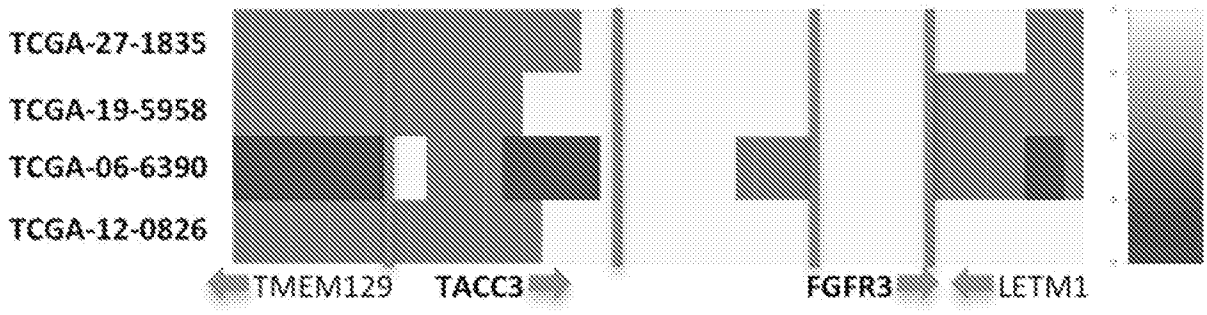

TCGA-27-1835
TCGA-19-5958
TCGA-06-6390
TCGA-12-0826
TMEM129
TACC3
FGFR3
LETM1

FIG. 2D

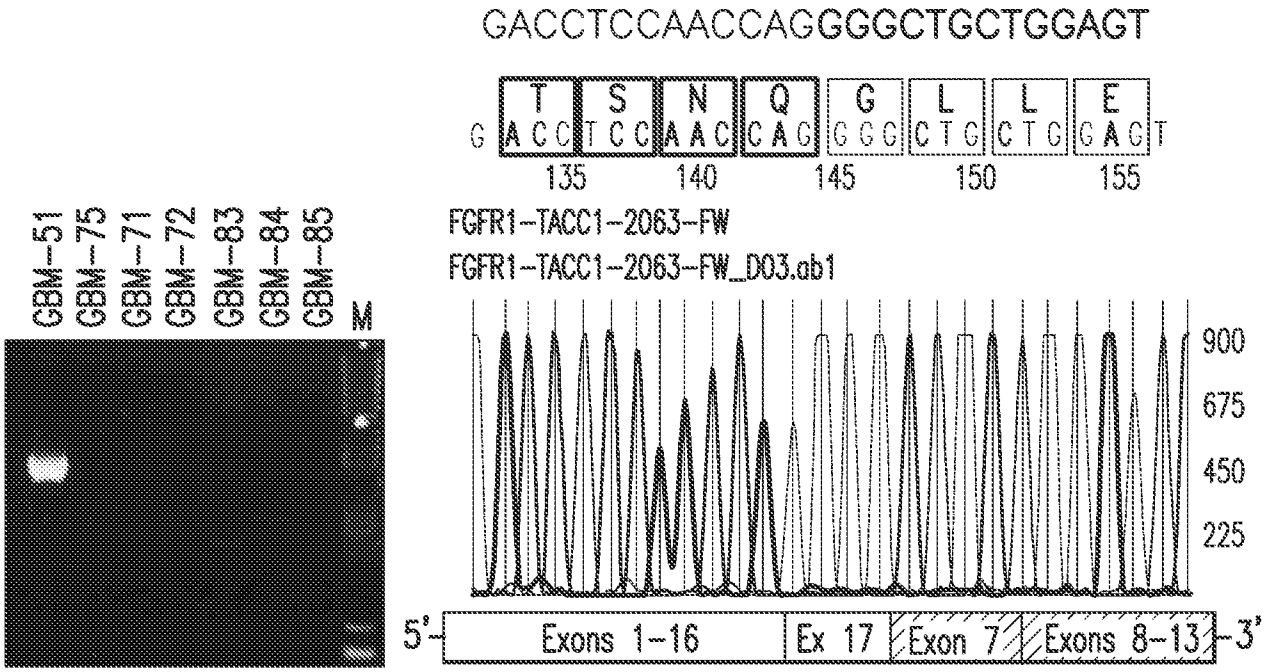

GACCTCCAACCAGGGGCTGCTGGAGT
G T S N Q G L L E
A C C T C C A A C C A G G G C T C T G C T C T G A C T
135 140 145 150 155
FGFR1-TACC1-2063-FW
FGFR1-TACC1-2063-FW_D03.ab1
GBM-51
GBM-75
GBM-71
GBM-72
GBM-83
GBM-84
GBM-85
M
900
675
450
225
5'
Exons 1-16
Ex 17
Exon 7
Exons 8-13
3'

FIG. 2E

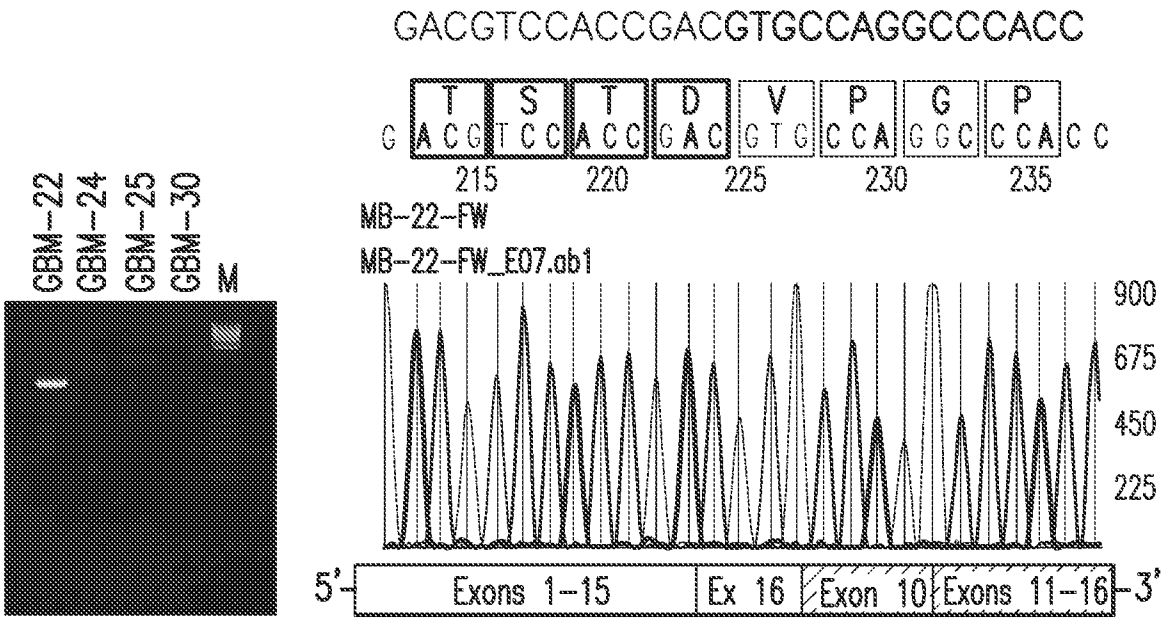

GACGTCCACCGACGTGCCAGGCCCACC
G T S T D V P G P
A C G T C C A C C G A C G T C C A G G C C C A C C
215 220 225 230 235
MB-22-FW
MB-22-FW_E07.ab1
GBM-22
GBM-24
GBM-25
GBM-30
M
900
675
450
225
5'
Exons 1-15
Ex 16
Exon 10
Exons 11-16
3'

FIG. 2F

	Vector	FGFR3	TACC3	F1-T1	F3-T3	F3-T3-K508M
Number of colonies	0	0	0	225.3±10.0	198.7±8.0	0

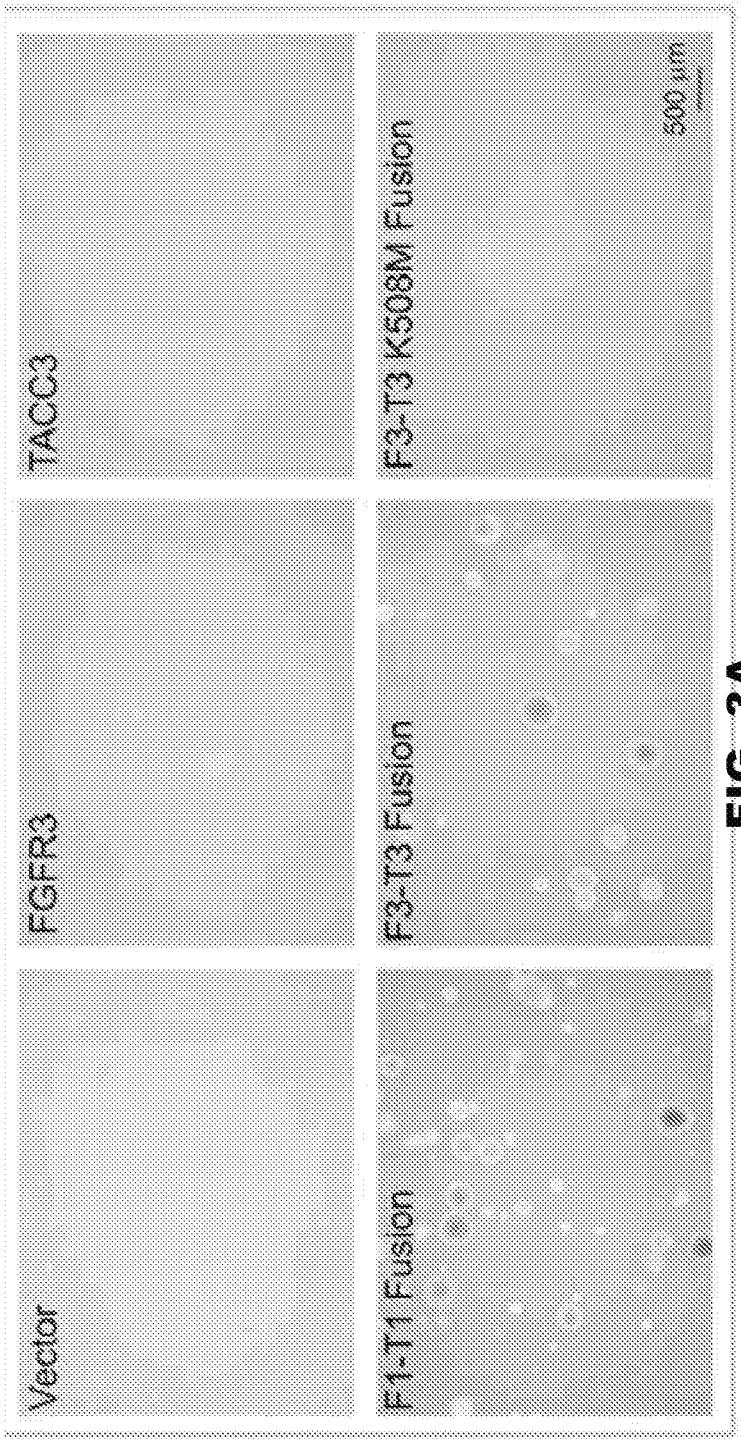


FIG. 3A

	Vector	FGFR3	TACC3	F1-T1	F3-T3	F3-T3-K508M
Number of mice with tumor	0/9	0/5	0/5	8/8	12/12	0/8

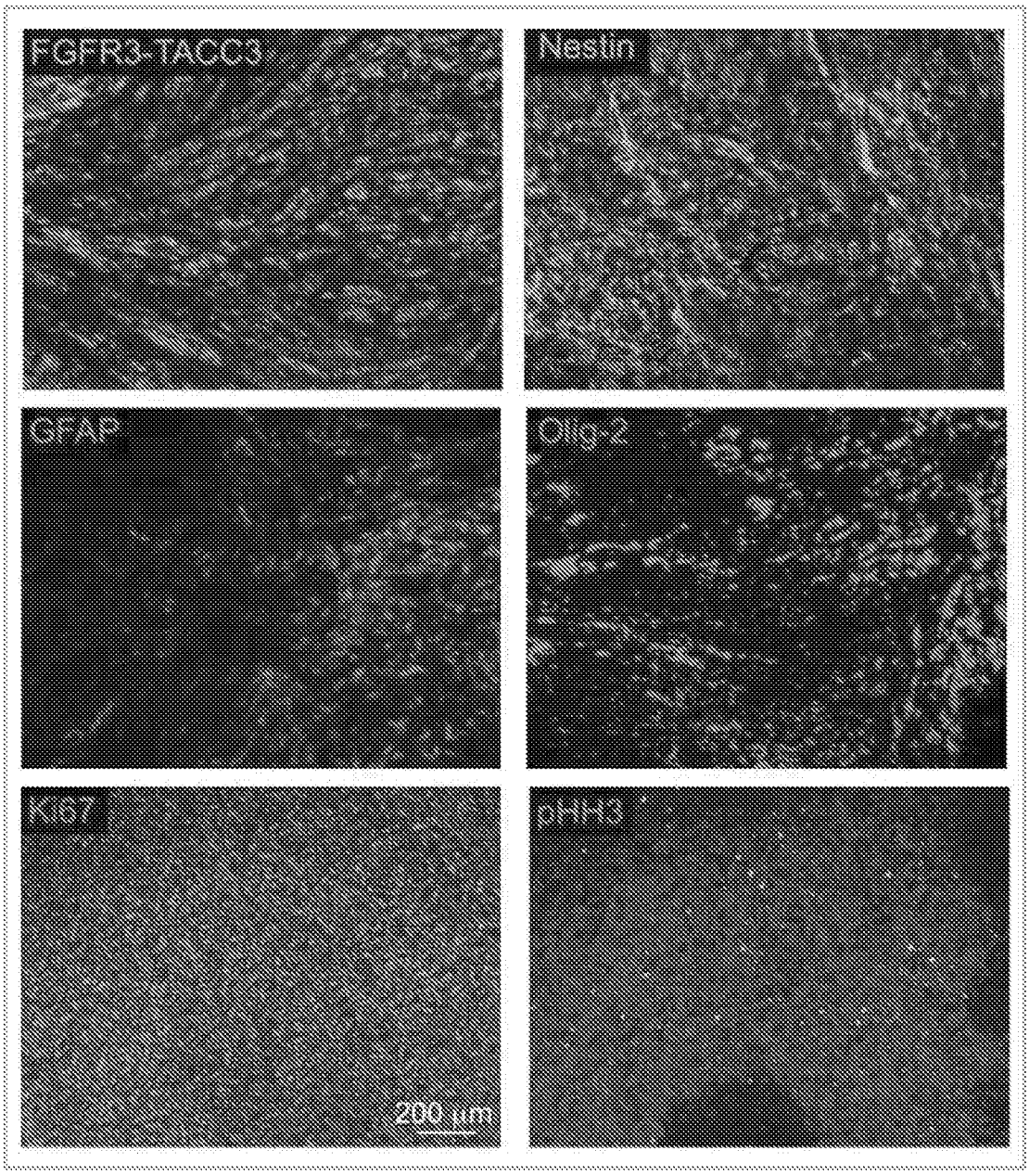


FIG. 3B

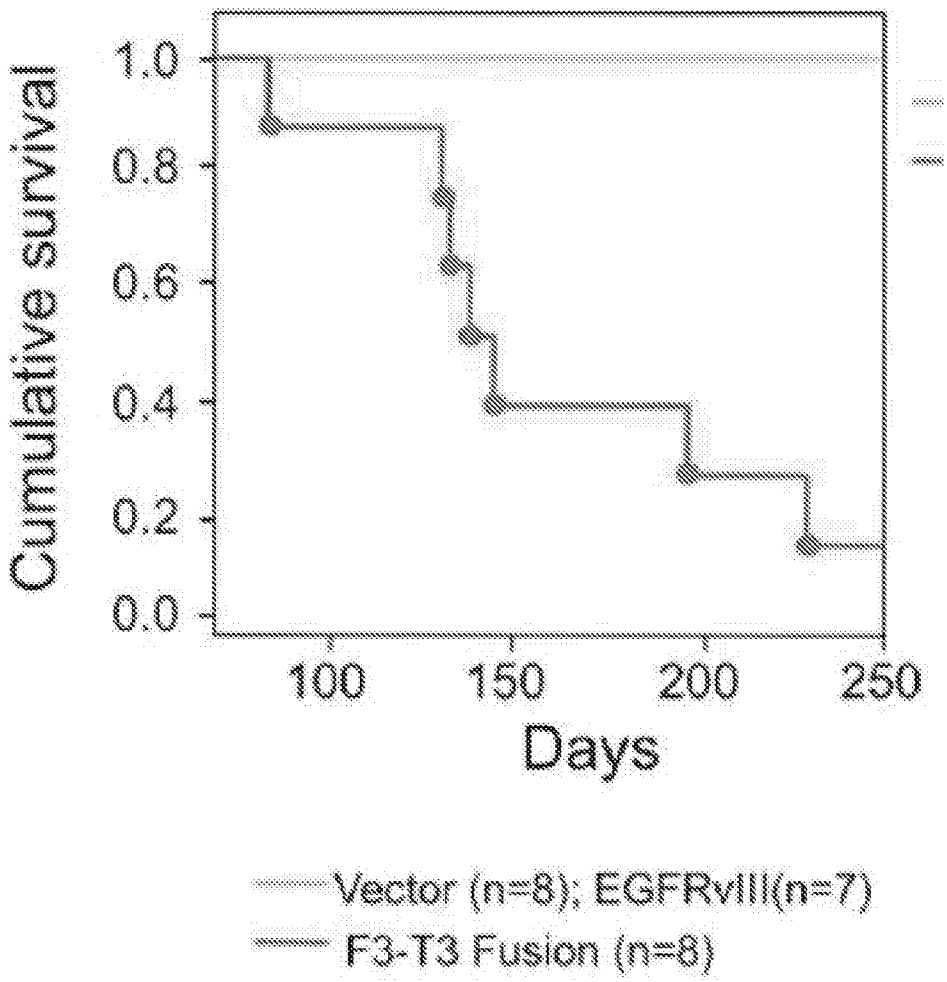

Cumulative survival
1.0
0.8
0.6
0.4
0.2
0.0
100
150
200
250
Days
Vector (n=8); EGFRvIII(n=7)
F3-T3 Fusion (n=8)

FIG. 3C

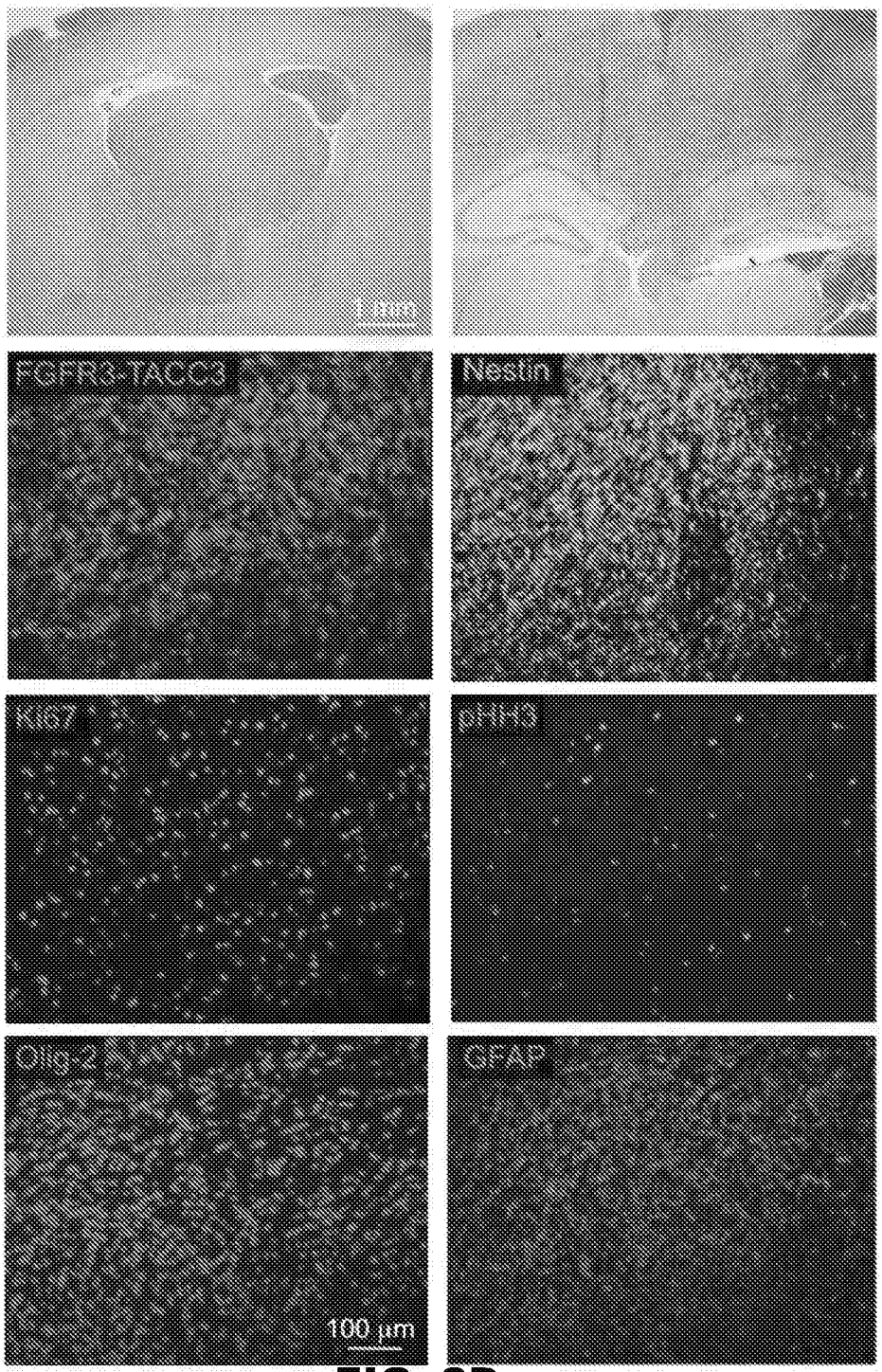


FIG. 3D

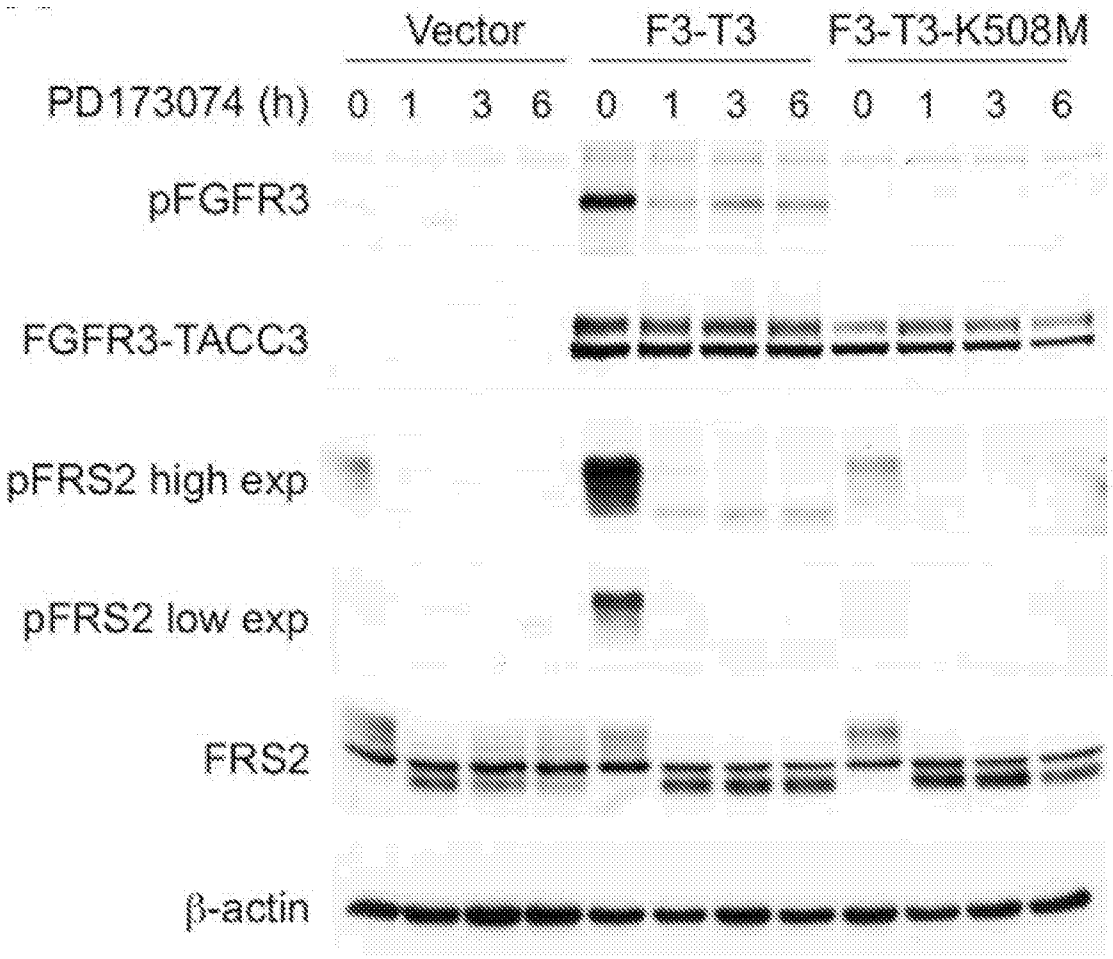

Vector
F3-T3
F3-T3-K508M
PD173074 (h)
0 1 3 6 0 1 3 6 0 1 3 6
pFGFR3
FGFR3-TACC3
pFRS2 high exp
pFRS2 low exp
FRS2
β-actin

FIG. 4A

Cell line	Metaphases inspected	Cells with segregation defects (%±SD)	Metaphases with misaligned chromosomes (%±SD)	Anaphases with lagging chromosomes (%±SD)	Anaphases/telophases with chromosome bridges (%±SD)
Rat1A Vector	150	9.5±3.8	2.3±1.7	2.4±0.9	5.4±2.4
Rat1A FGFR3-TACC3	150	27.4±3.9	8.0±2.6	8.0±1.0	11.2±0.7
Rat1A FGFR1-TACC1	100	45.5±3.1	5±2.8	10±1.4	31.5±4.9

Cell line	Metaphases inspected	Cells with segregation defects (%±SD)	Metaphases with misaligned chromosomes (%±SD)	Anaphases with lagging chromosomes (%±SD)	Anaphases/telophases with chromosome bridges (%±SD)
Ink4A/ARF-/- Vector	100	7.3±1.5	2.2±1.2	2.0±0.7	3.3±0.6
Ink4A/ARF-/- F3-T3 Fusion	100	18.3±1.5	6.0±1.1	6.5±1.3	11.2±0.7

FIG. 4B

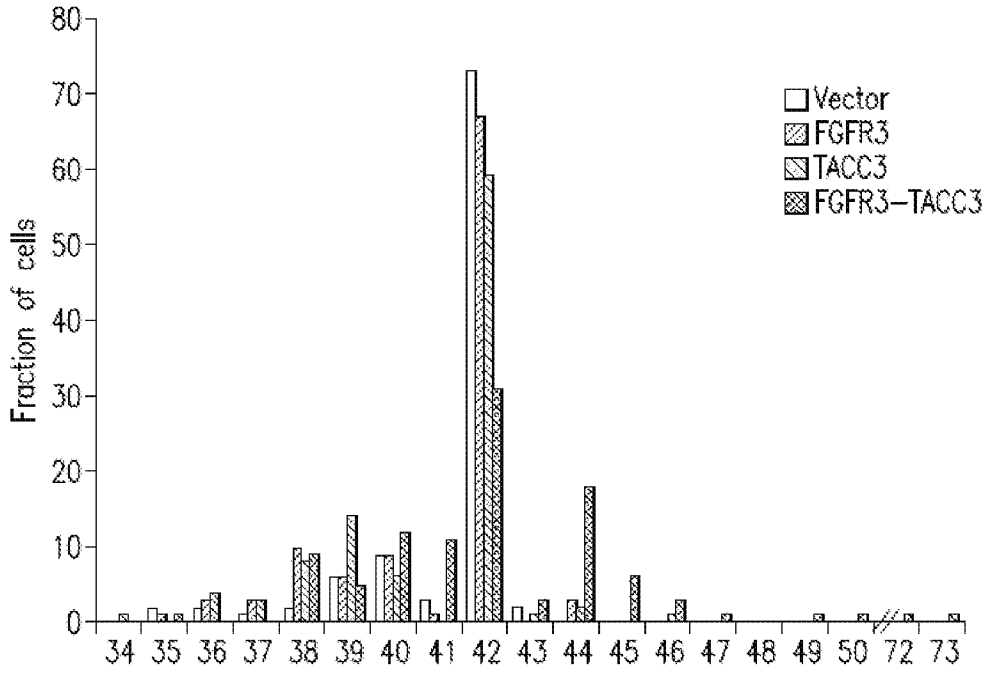


FIG.5A

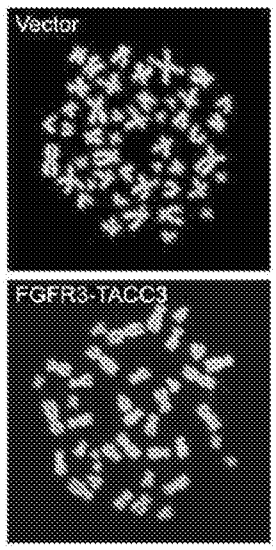


FIG. 5B

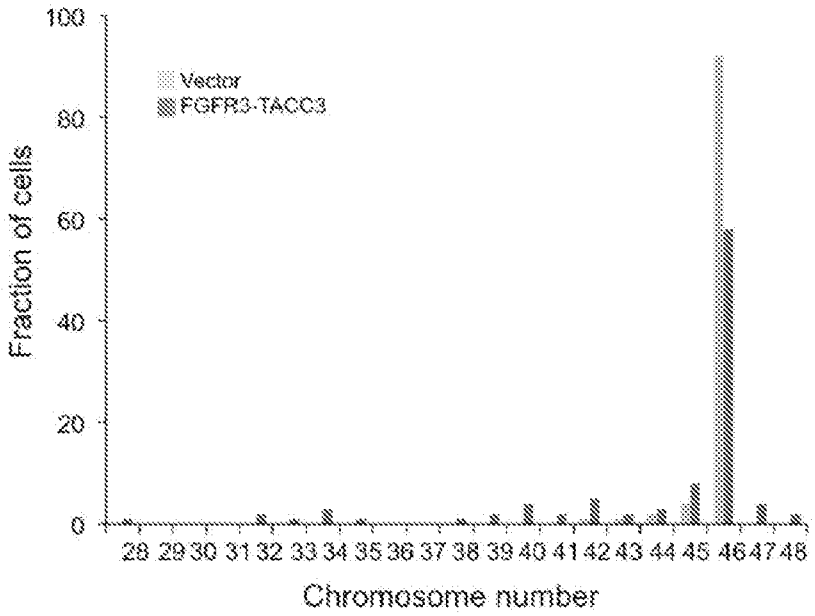
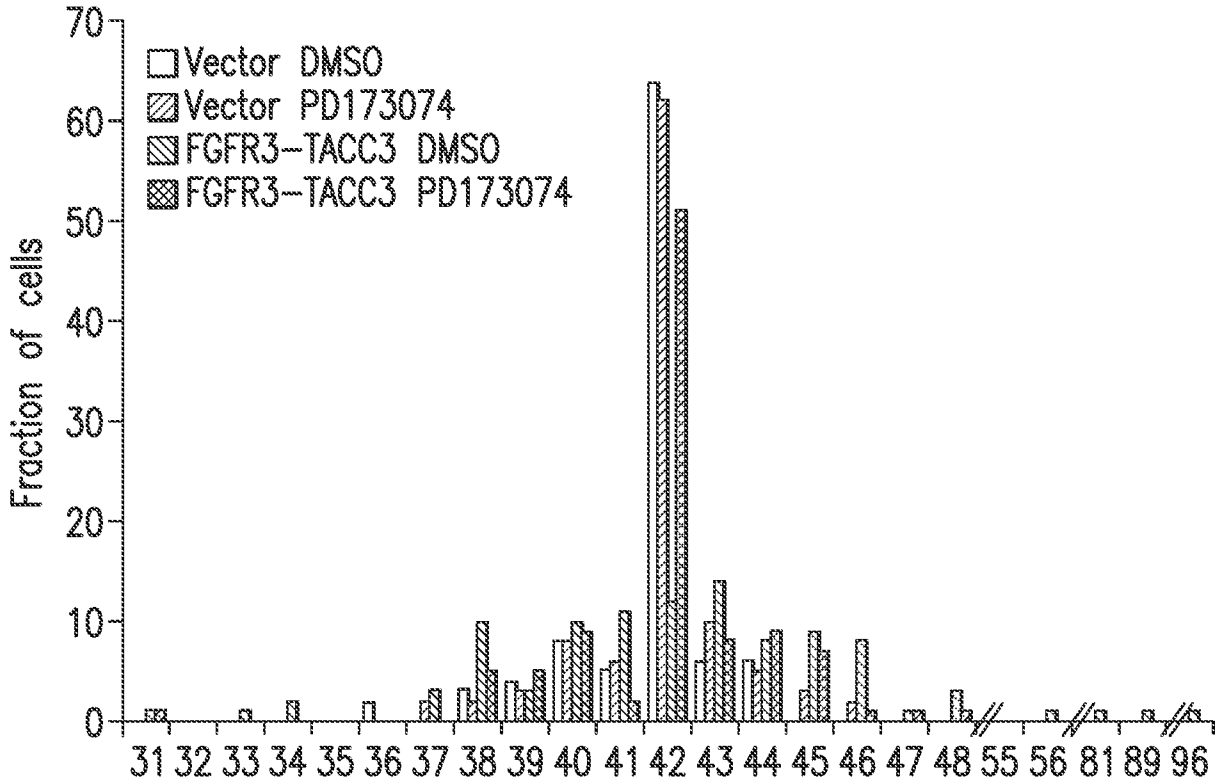


FIG. 5C

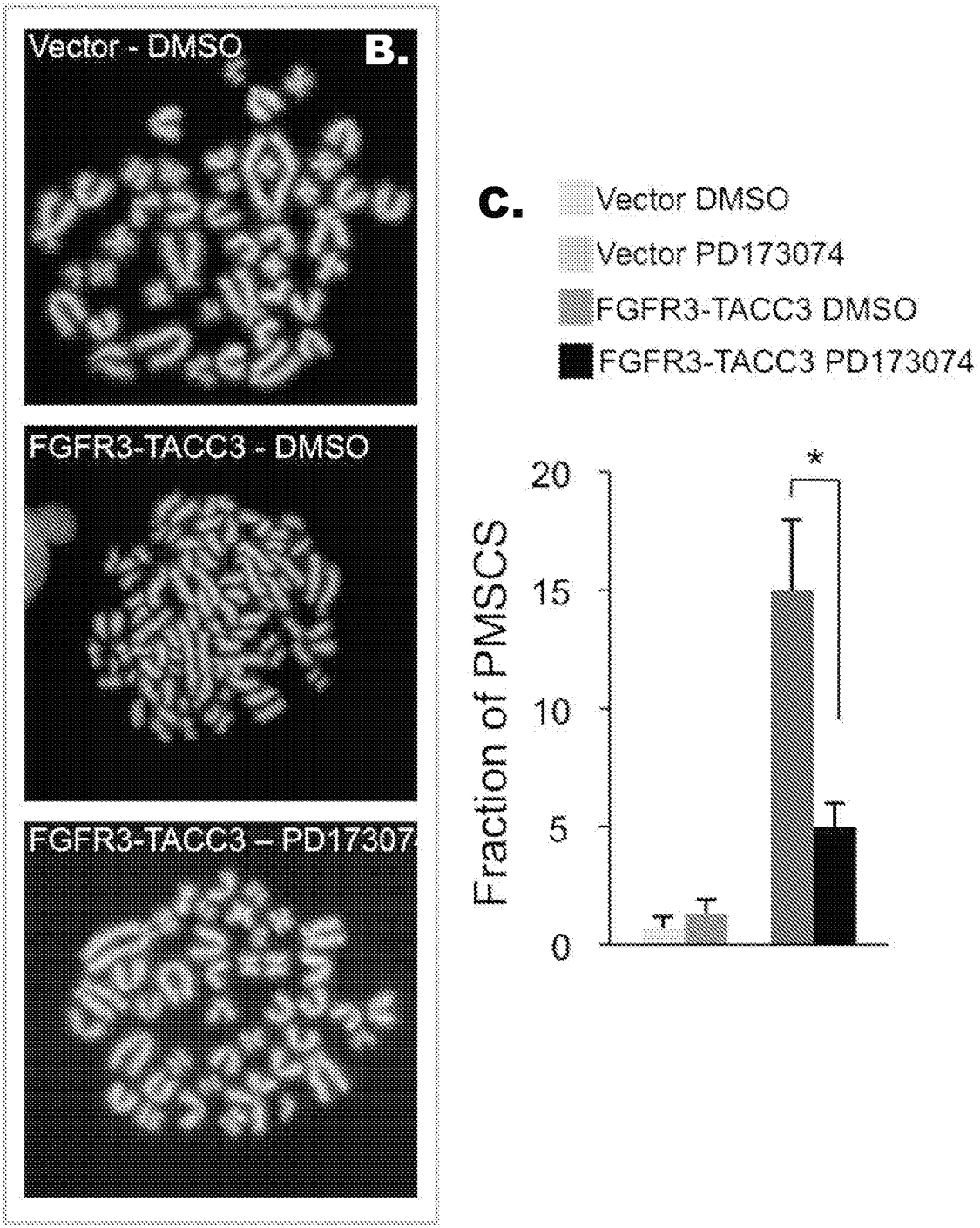
Cell line	Number of cells counted	Percent aneuploidy	Range	Mean number	Average variation from mean number	p- value
Vector	100	8	42-46	45.85	0.28	$p = <0.001$
FGFR3-TACC3	100	42	28-48	42.24	3.33	

FIG. 5D



Cell line	Number of cells counted	Percent aneuploidy	Range	Mean number	Average variation from mean number	*p*-value
Rat1A Vector DMSO	100	38	37–55	41.7	1.10	
Rat1A Vector PD173074	100	39	37–46	42.0	0.77	n.s.
Rat1A FGFR3-TACC3 DMSO	100	88	31–96	43.7	4.11	<0.0001
Rat1A FGFR3-TACC3 PD173074	100	49	31–48	42.0	1.28	<0.0001

FIG.6A


Vector - DMSO
B.
FGFR3-TACC3 - DMSO
FGFR3-TACC3 – PD17307
C.
Vector DMSO
Vector PD173074
FGFR3-TACC3 DMSO
FGFR3-TACC3 PD173074
Fraction of PMSCS
20
15
10
5
0
*

FIGS. 6B-C

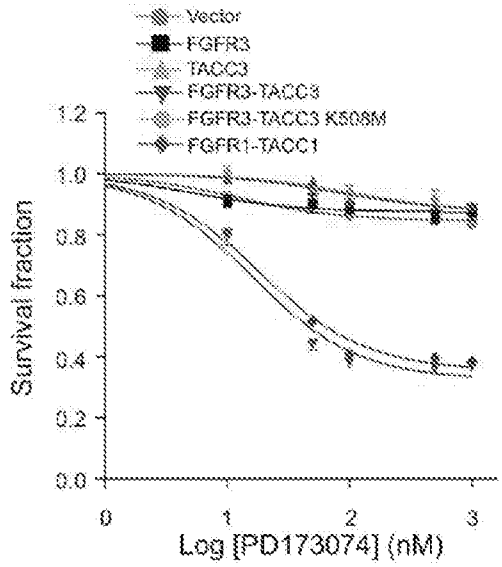


FIG. 7A

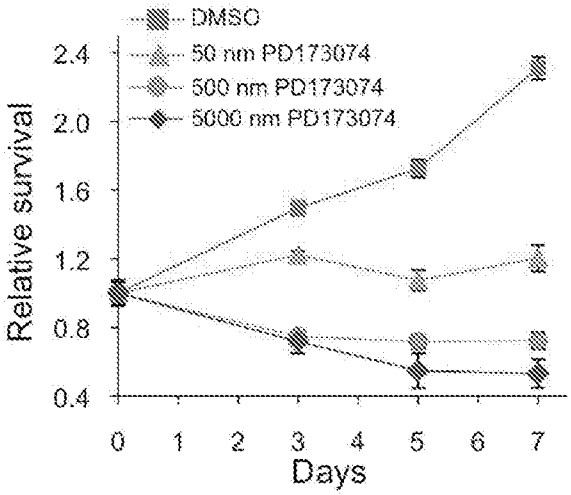
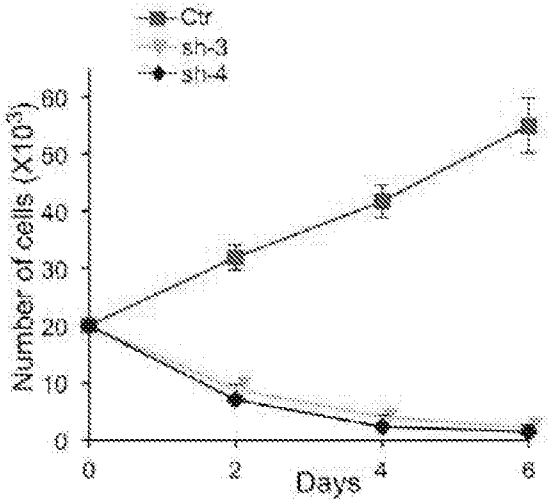


FIG. 7B



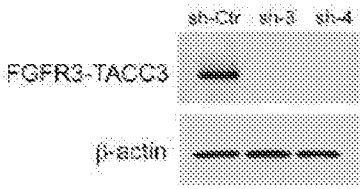


FIG. 7C

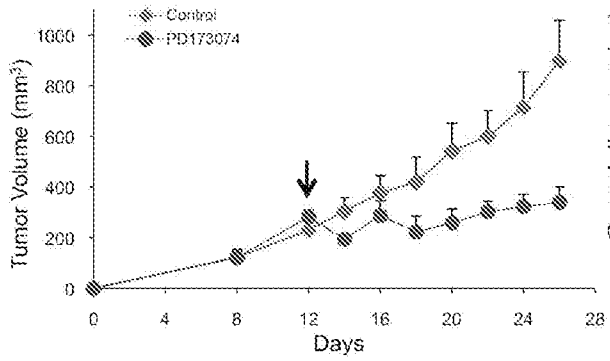


FIG. 7D

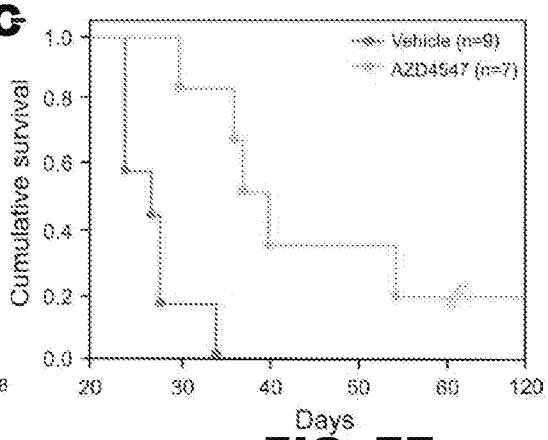


FIG. 7E

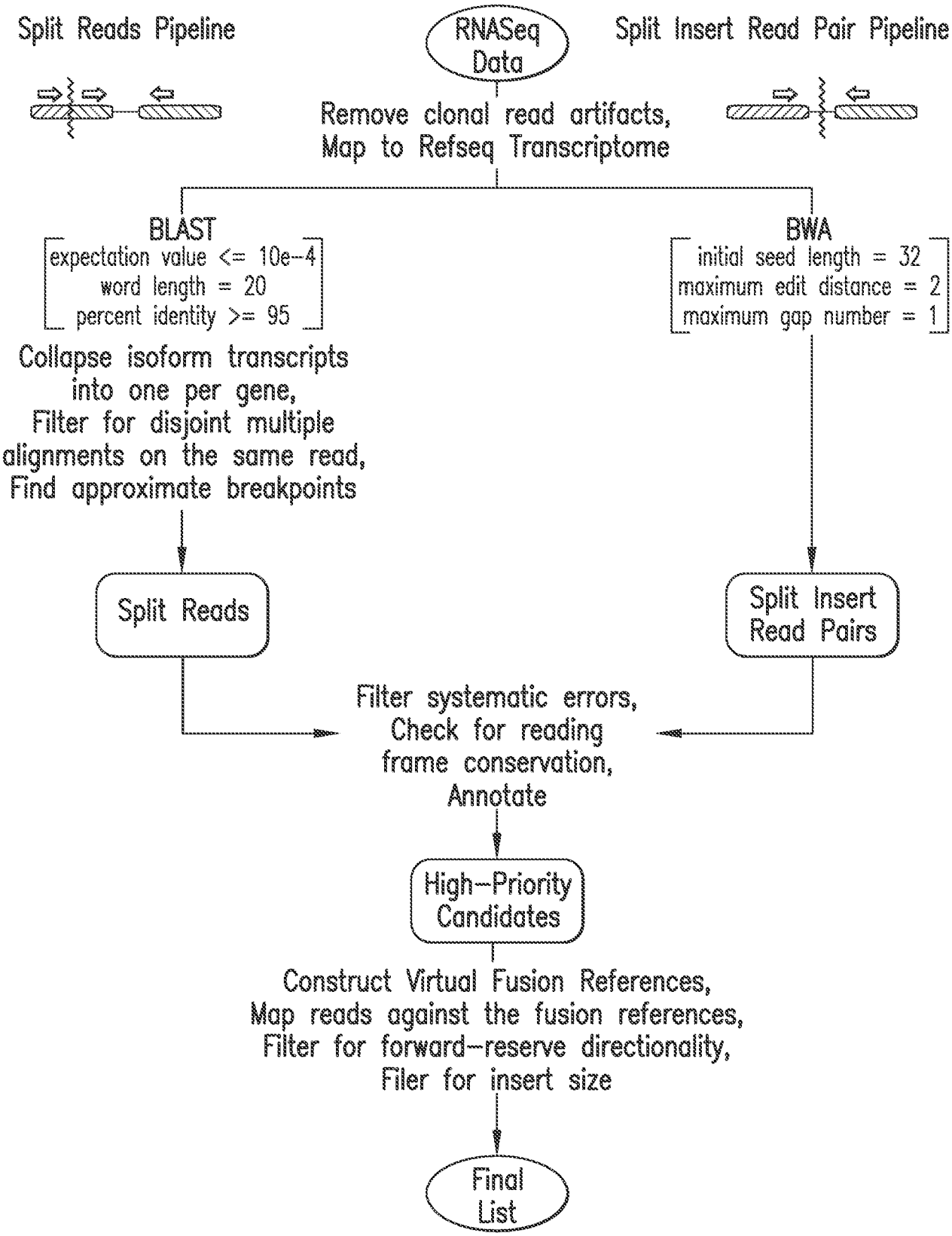

Split Reads Pipeline
RNASeq Data
Split Insert Read Pair Pipeline
Remove clonal read artifacts,
Map to Refseq Transcriptome
BLAST
expectation value <= 10e-4
word length = 20
percent identity >= 95
BWA
initial seed length = 32
maximum edit distance = 2
maximum gap number = 1
Collapse isoform transcripts into one per gene,
Filter for disjoint multiple alignments on the same read,
Find approximate breakpoints
Split Reads
Split Insert Read Pairs
Filter systematic errors,
Check for reading frame conservation,
Annotate
High-Priority Candidates
Construct Virtual Fusion References,
Map reads against the fusion references,
Filter for forward-reserve directionality,
Filer for insert size
Final List

FIG.8

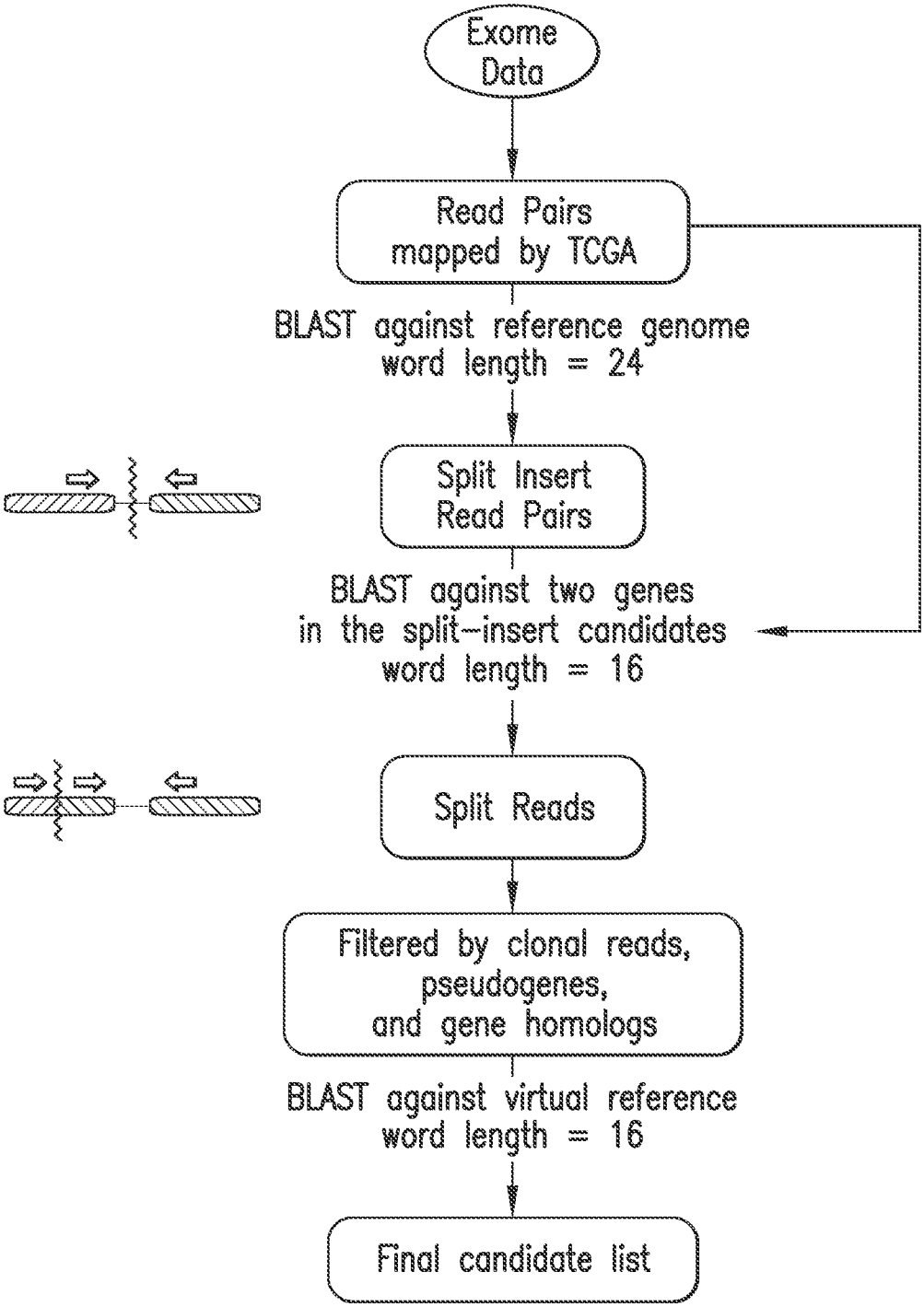

Exome Data
Read Pairs mapped by TCGA
BLAST against reference genome word length = 24
Split Insert Read Pairs
BLAST against two genes in the split-insert candidates word length = 16
Split Reads
Filtered by clonal reads, pseudogenes, and gene homologs
BLAST against virtual reference word length = 16
Final candidate list

FIG.8 - CONT.

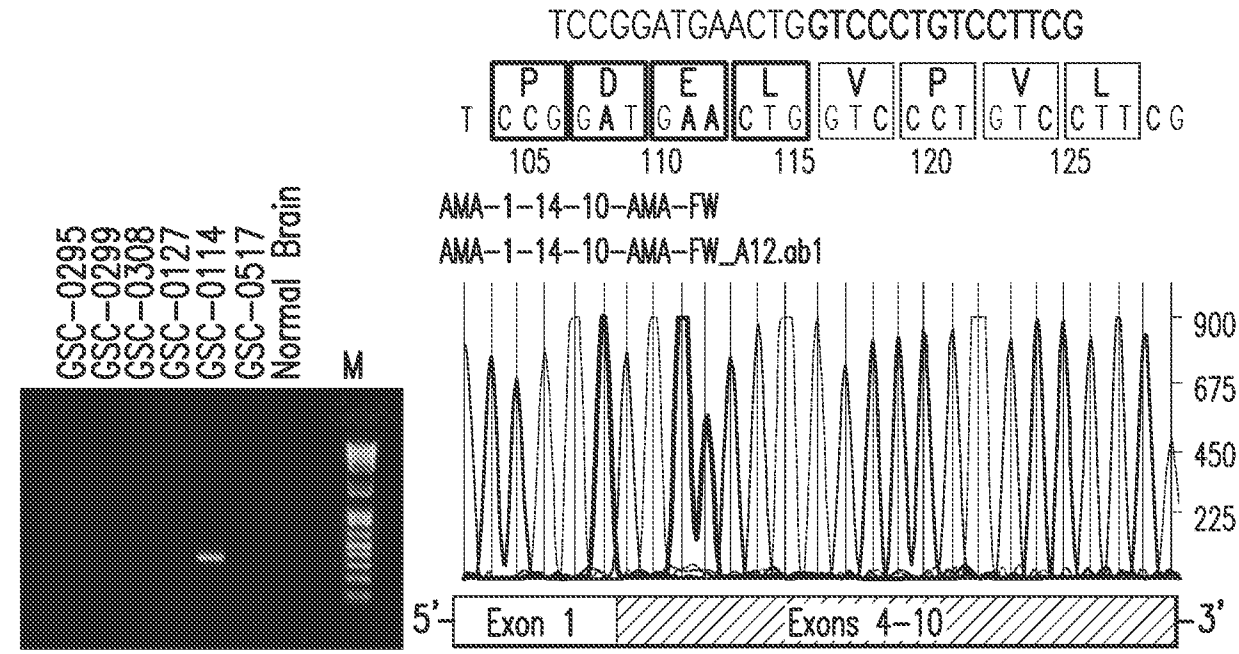

POLR2A-WRAP53 GSC-0114
TCCGGATGAACTGGTCCCTGTCCTTCG
P D E L V P V L
T C C G G A T G A A C T G G T C C C T G T C C T T C G
105 110 115 120 125
AMA-1-14-10-AMA-FW
AMA-1-14-10-AMA-FW_A12.ab1
GSC-0295
GSC-0299
GSC-0308
GSC-0127
GSC-0114
GSC-0517
Normal Brain
M
900
675
450
225
5'
Exon 1
Exons 4-10
3'

FIG.9A

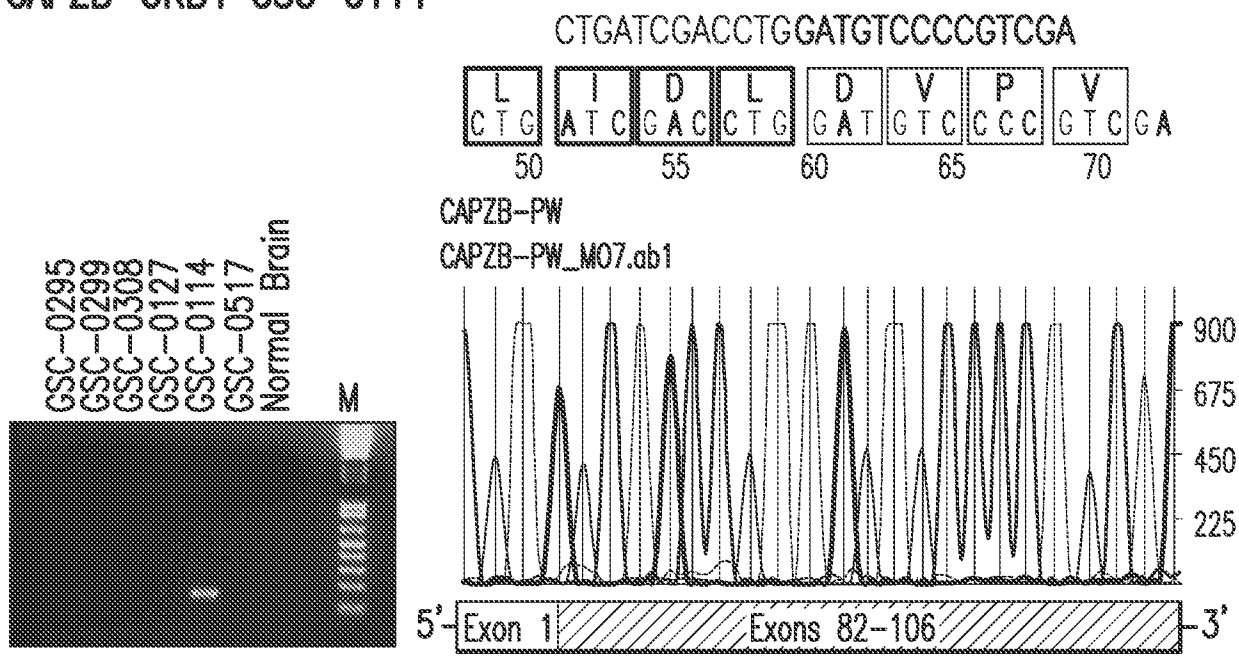

CAPZB-URB4 GSC-0114
CTGATCGACCTGGATGTCCCCGTCGA
L I D L D V P V
C T G A T C G A C C T G G A T G T C C C C G T C G A
50 55 60 65 70
CAPZB-PW
CAPZB-PW_M07.ab1
GSC-0295
GSC-0299
GSC-0308
GSC-0127
GSC-0114
GSC-0517
Normal Brain
M
900
675
450
225
5'
Exon 1
Exons 82-106
3'

FIG.9B

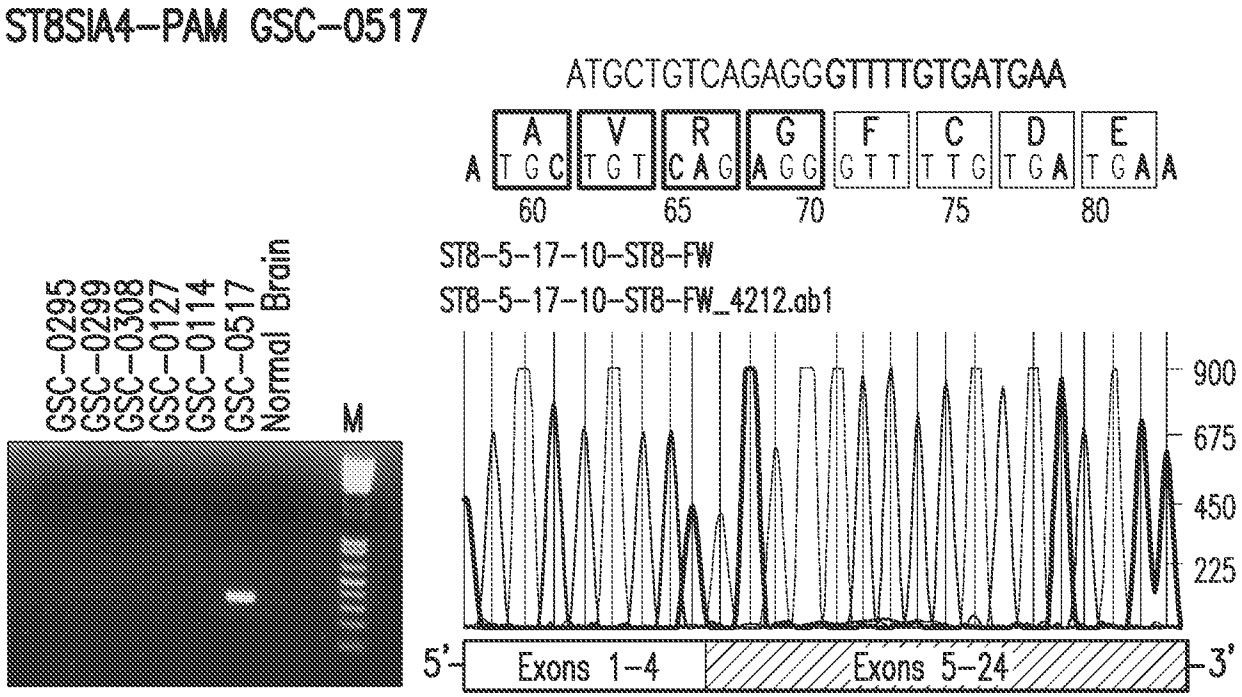

ST8SIA4-PAM GSC-0517
ATGCTGTCAGAGGGTTTTGTGATGAA
A V R G F C D E
A T G C T G T C A G A G G G T T T T G T G A T G A A
60 65 70 75 80
ST8-5-17-10-ST8-FW
ST8-5-17-10-ST8-FW_4212.ab1
GSC-0295
GSC-0299
GSC-0308
GSC-0127
GSC-0114
GSC-0517
Normal Brain
M
900
675
450
225
5'
Exons 1-4
Exons 5-24
3'

FIG.9C

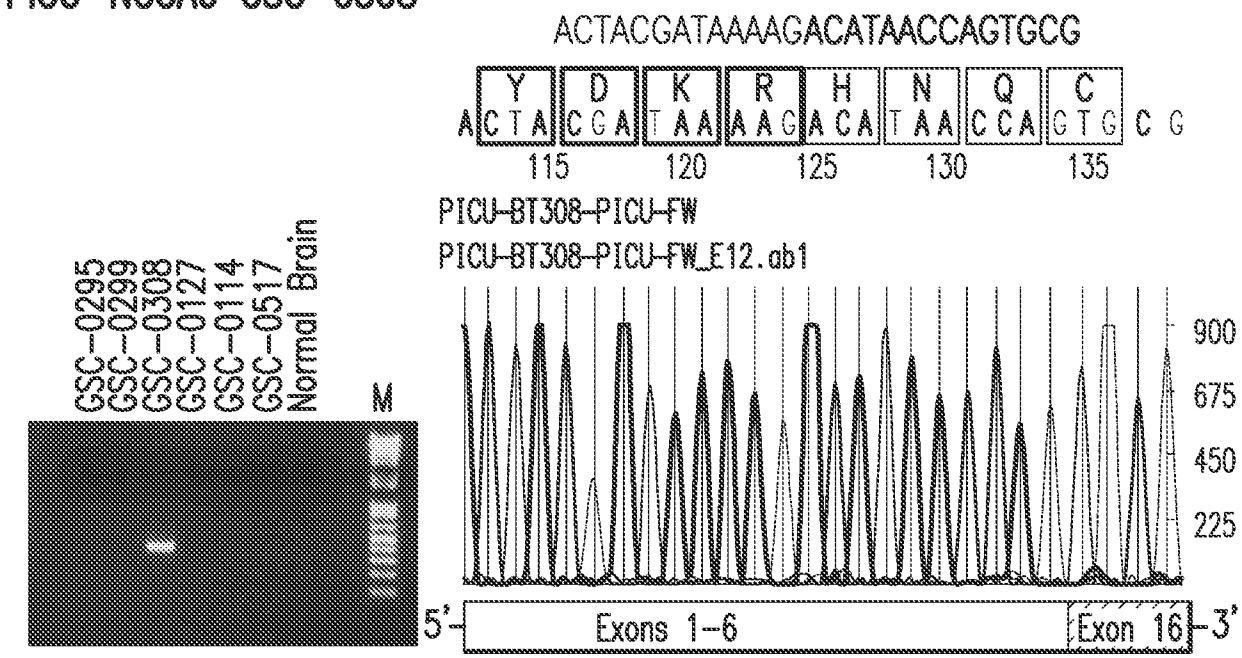

PIGU-NCOA6 GSC-0308
ACTACGATAAAAGACATAACCAGTGCG
Y D K R H N Q C
A C T A C G A T A A A A G A C A T A A C C A G T G C G
115 120 125 130 135
PICU-BT308-PICU-FW
PICU-BT308-PICU-FW_E12.ab1
GSC-0295
GSC-0299
GSC-0308
GSC-0127
GSC-0114
GSC-0517
Normal Brain
M
900
675
450
225
5'
Exons 1-6
Exon 16
3'

FIG.9D

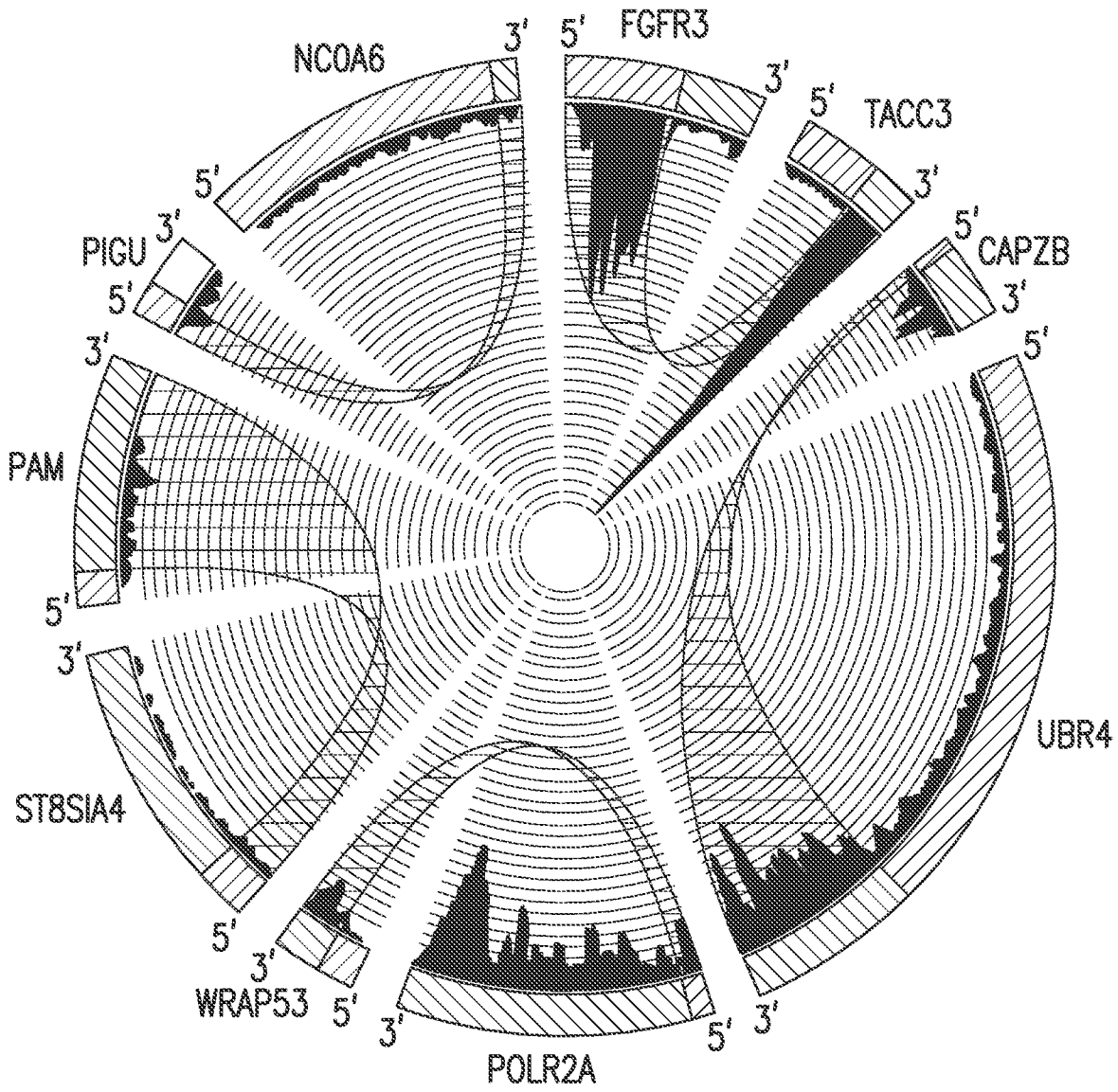

NCOA6
FGFR3
TACC3
CAPZB
PIGU
PAM
UBR4
ST8SIA4
WRAP53
POLR2A
5'
3'

FIG. 10A

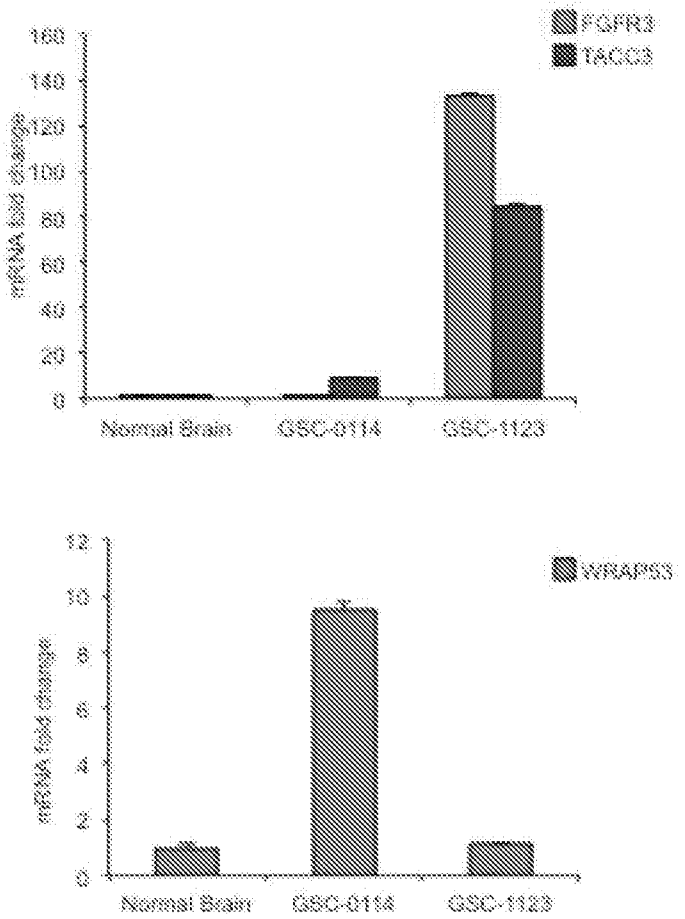


FIG. 10B

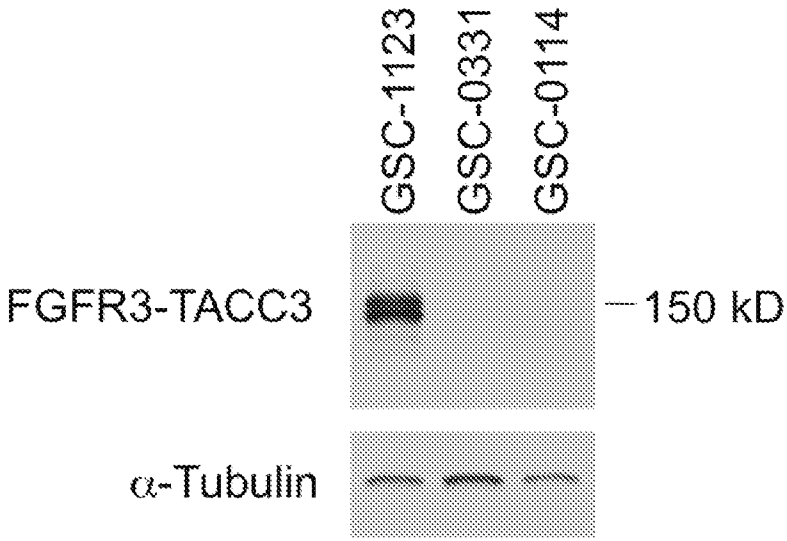


FIG. 10C

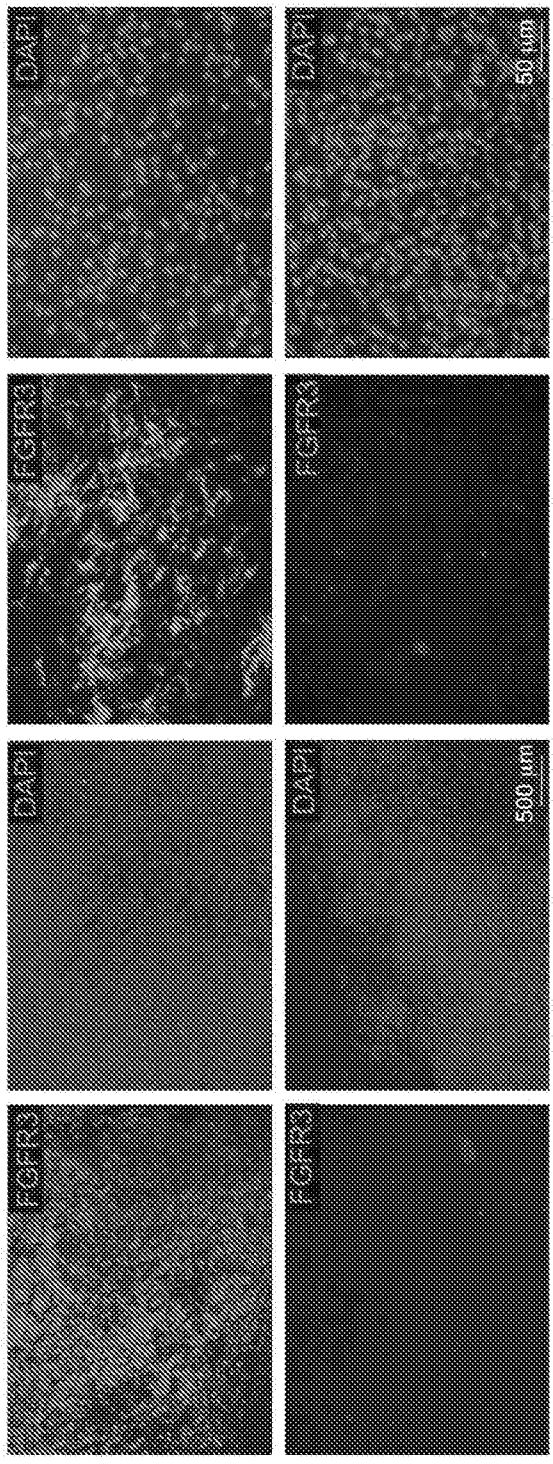


FIG. 10D

Seq	#	b	y	+1
A	1	72.0	–	13
Q	2	**200.1**	1270.7	12
A	3	271.1	1142.6	11
E	4	400.2	1071.6	10
A	5	**471.2**	942.6	9
L	6	**584.3**	871.5	8
A	7	**655.3**	758.4	7
L	8	**768.4**	687.4	6
Q	9	**896.5**	574.3	5
A	10	**967.5**	446.3	4
S	11	**1054.5**	375.2	3
L	12	**1167.6**	288.2	2
R	13	–	175.1	1

TACC3
Peptide 1

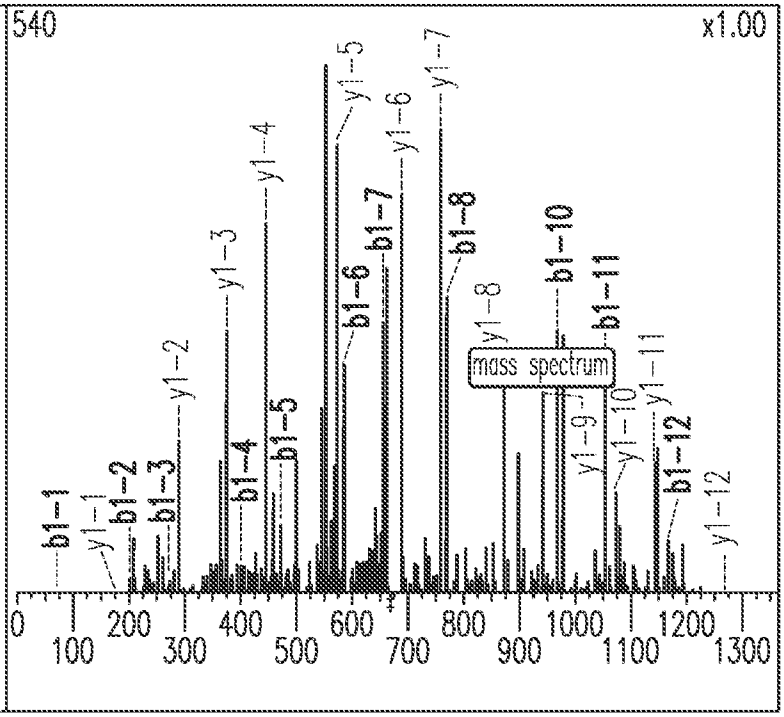


FIG.10E-1

Seq	#	b	y	+1
F	1	148.1	–	6
D	2	**263.1**	613.3	5
P	3	360.1	498.3	4
L	4	473.2	401.3	3
L	5	**586.3**	288.2	2
R	6	–	175.1	1

TACC3
Peptide 2

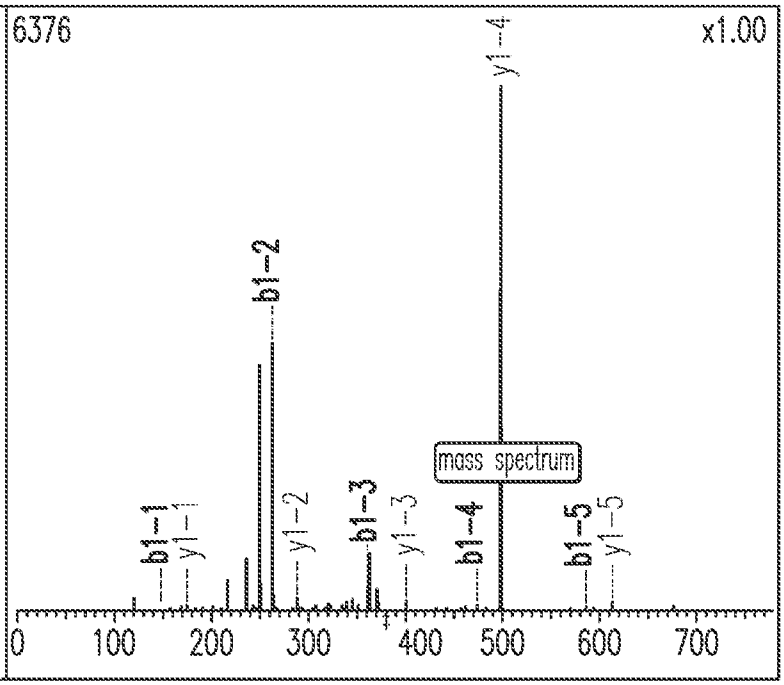


FIG.10E-2

Seq	#	b	y	+1
L	1	114.1	–	12
Q	2	242.1	1270.7	11
L	3	355.2	1142.6	10
A	4	426.3	1029.5	9
N	5	540.3	958.5	8
E	6	669.3	844.4	7
E	7	798.4	715.4	6
I	8	911.5	586.3	5
A	9	982.5	473.3	4
Q	10	1110.6	402.2	3
V	11	1209.6	274.2	2
R	12	–	175.1	1

TACC3
Peptide 3

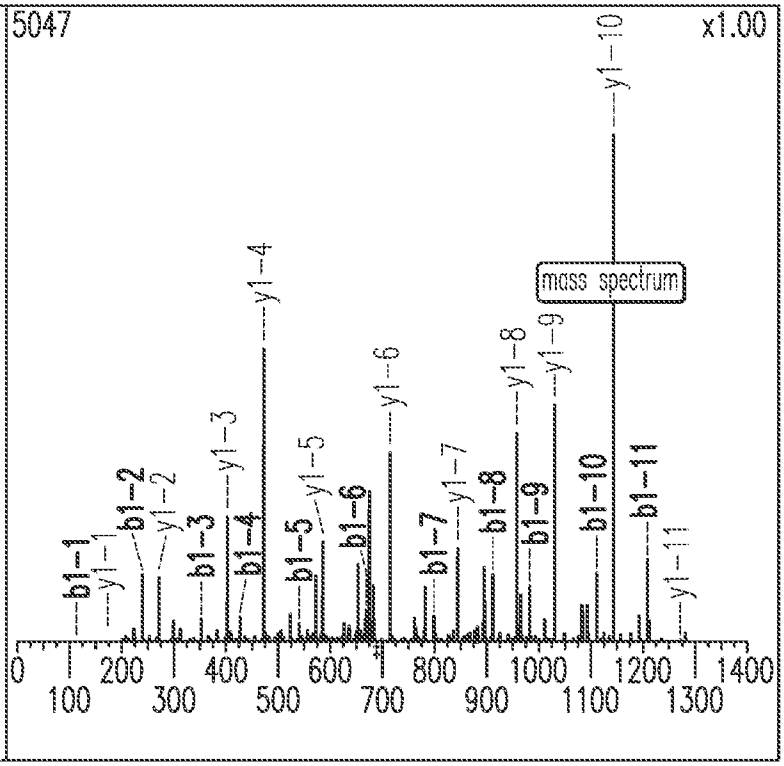


FIG.10E-3

Seq	#	b	y	+1
D	1	116.0	–	10
G	2	173.0	915.5	9
T	3	274.1	858.5	8
G	4	331.1	757.4	7
L	5	444.2	700.4	6
V	6	543.3	587.3	5
P	7	640.3	488.2	4
S	8	727.4	391.2	3
E	9	856.4	304.1	2
R	10	–	175.1	1

FGFR3
Peptide 1

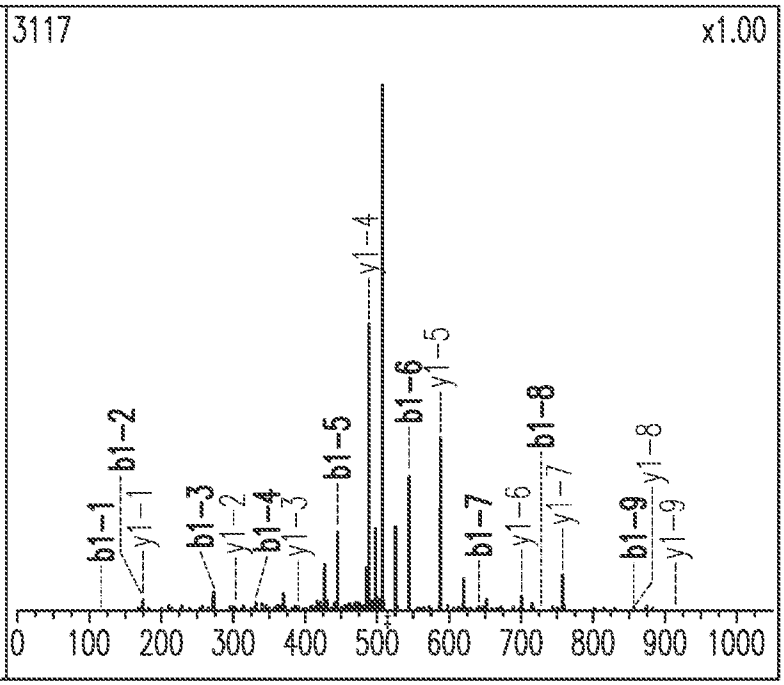


FIG.10E-4

Seq	#	b	y	+1
V	1	100.1	–	13
G	2	157.1	1246.6	12
P	3	**254.1**	1189.6	11
D	4	**369.2**	1092.6	10
G	5	**426.2**	977.5	9
T	6	**527.2**	920.5	8
P	7	**624.3**	819.5	7
Y	8	787.4	722.4	6
V	9	**886.4**	**559.4**	5
T	10	987.5	460.3	4
V	11	**1086.5**	**359.2**	3
L	12	**1199.6**	260.2	2
K	13	–	147.1	1

FGFR3
Peptide 2

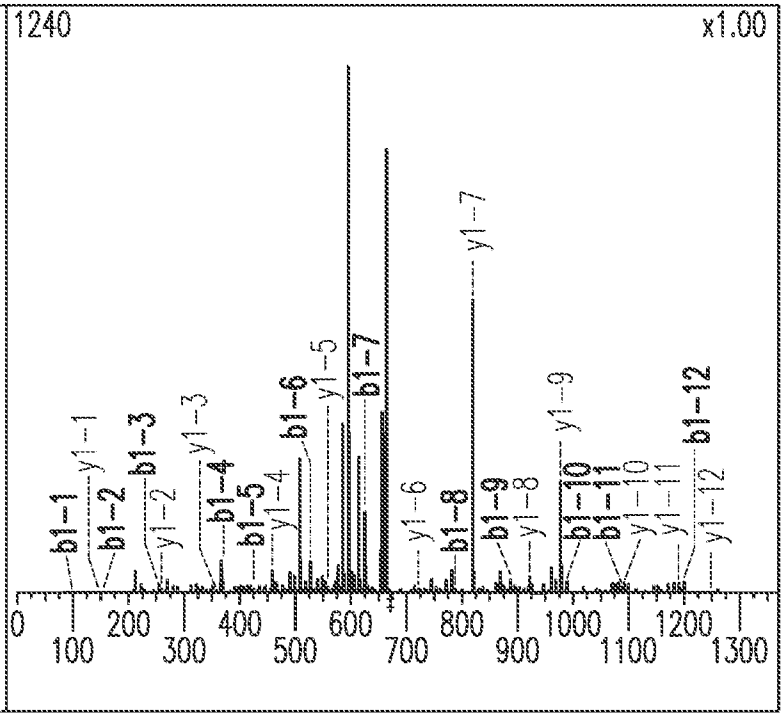


FIG.10E-5

Seq	#	b	y	+1
V	1	100.1	–	7
L	2	**213.2**	669.4	6
V	3	**312.2**	556.3	5
G	4	**369.2**	457.2	4
P	5	466.3	400.2	3
Q	6	**594.4**	303.2	2
R	7	–	175.1	1

FGFR3
Peptide 3

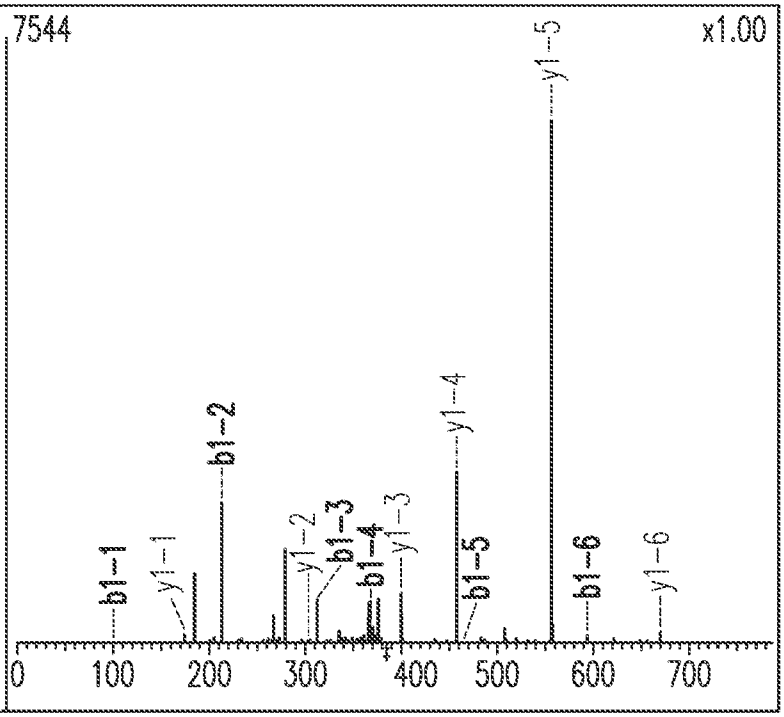
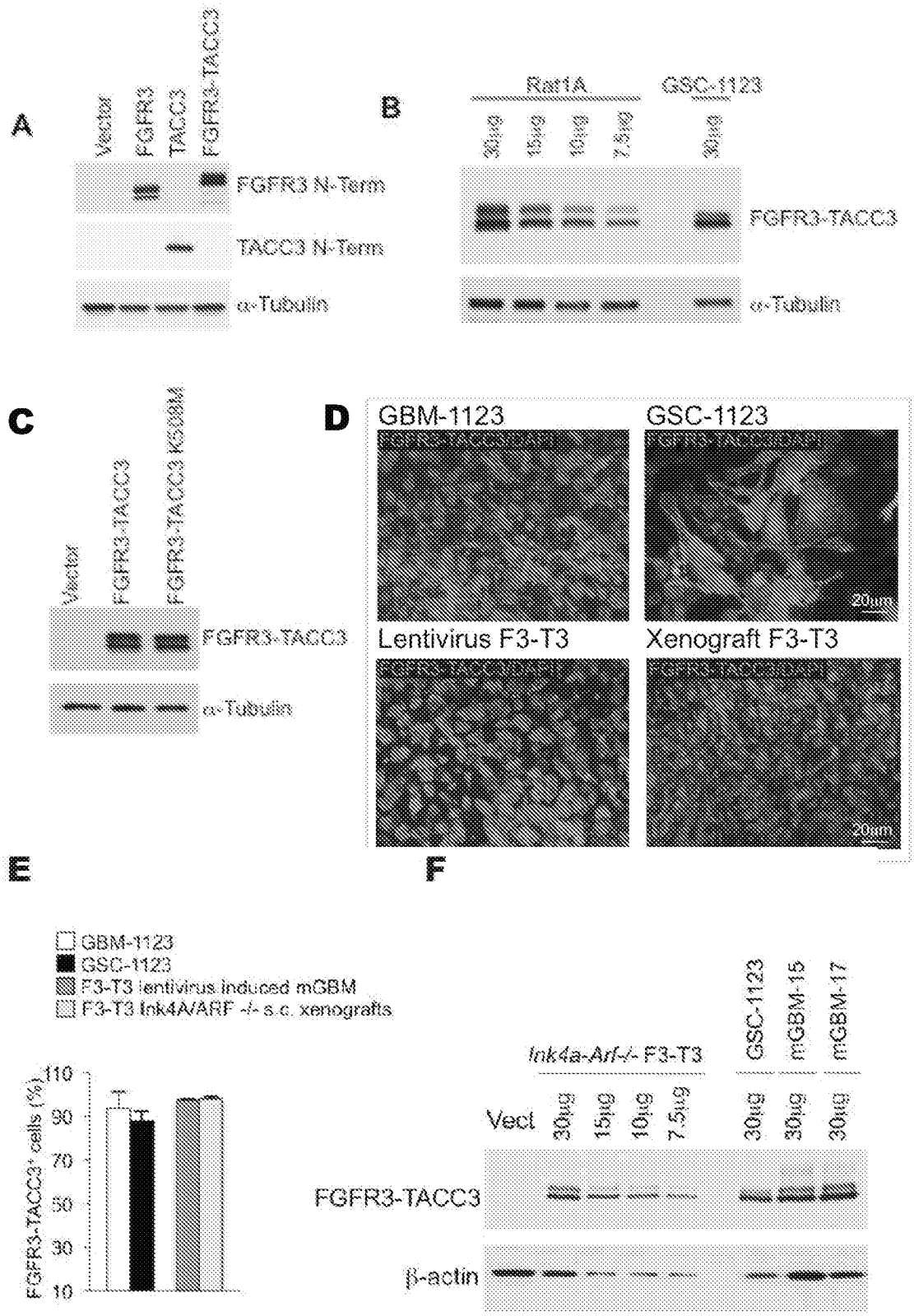
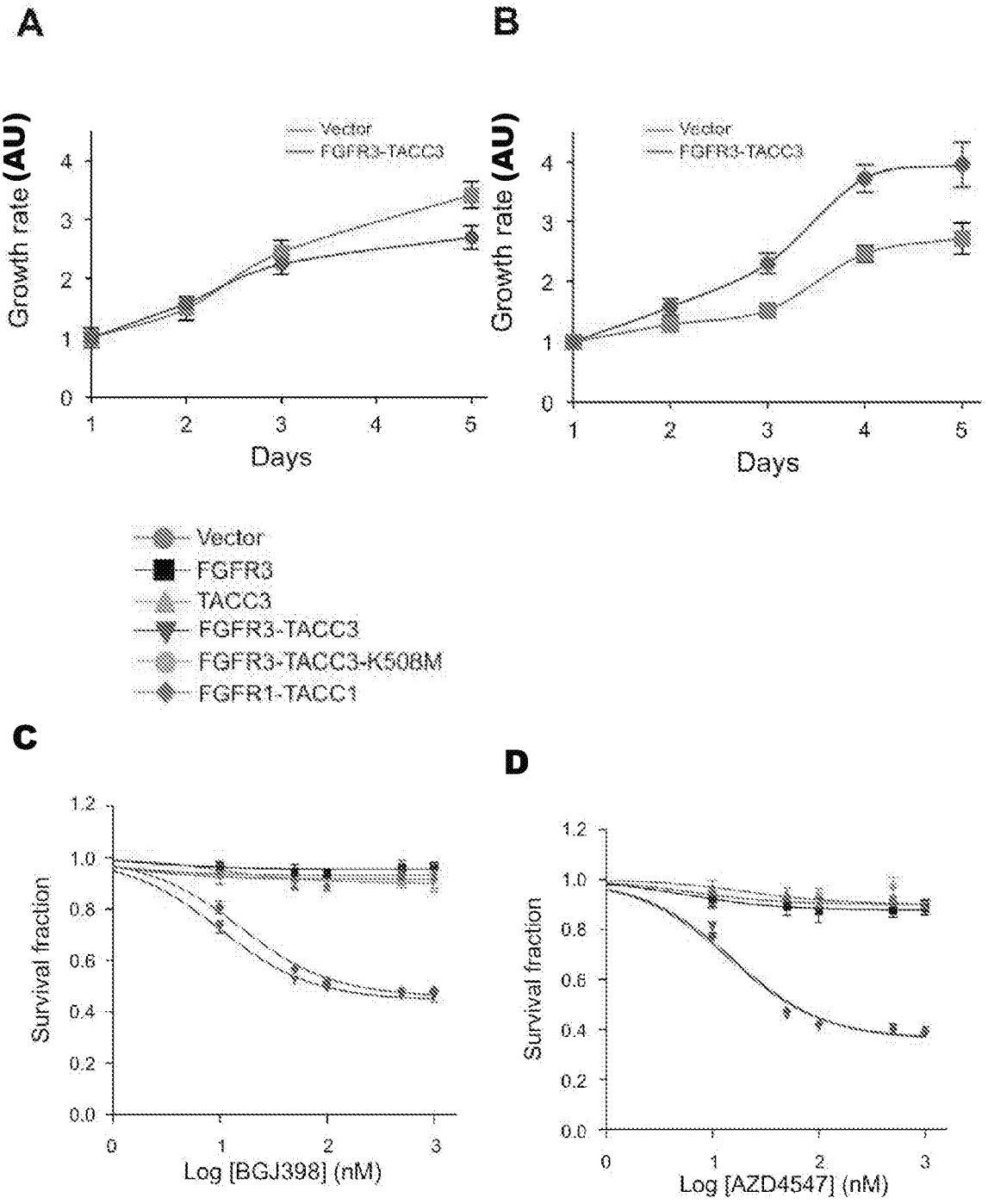


FIG.10E-6


A
Vector
FGFR3
TACC3
FGFR3-TACC3
FGFR3 N-Term
TACC3 N-Term
α-Tubulin
B
Rat1A
GSC-1123
30μg
15μg
10μg
7.5μg
30μg
FGFR3-TACC3
α-Tubulin
C
Vector
FGFR3-TACC3
FGFR3-TACC3 K508M
FGFR3-TACC3
α-Tubulin
D
GBM-1123
GSC-1123
20μm
Lentivirus F3-T3
Xenograft F3-T3
20μm
E
GBM-1123
GSC-1123
F3-T3 lentivirus induced mGBM
F3-T3 Ink4A/ARF -/- s.c. xenografts
FGFR3-TACC3+ cells (%)
110
90
70
50
30
10
F
Ink4a-Arf-/- F3-T3
GSC-1123
mGBM-15
mGBM-17
Vect
30μg
15μg
10μg
7.5μg
30μg
30μg
30μg
FGFR3-TACC3
β-actin

FIGS. 11A-F


A
B
Growth rate (AU)
Vector
FGFR3-TACC3
Days
Vector
FGFR3
TACC3
FGFR3-TACC3
FGFR3-TACC3-K508M
FGFR1-TACC1
C
D
Survival fraction
Log [BGJ398] (nM)
Log [AZD4547] (nM)

FIGS. 12A-D

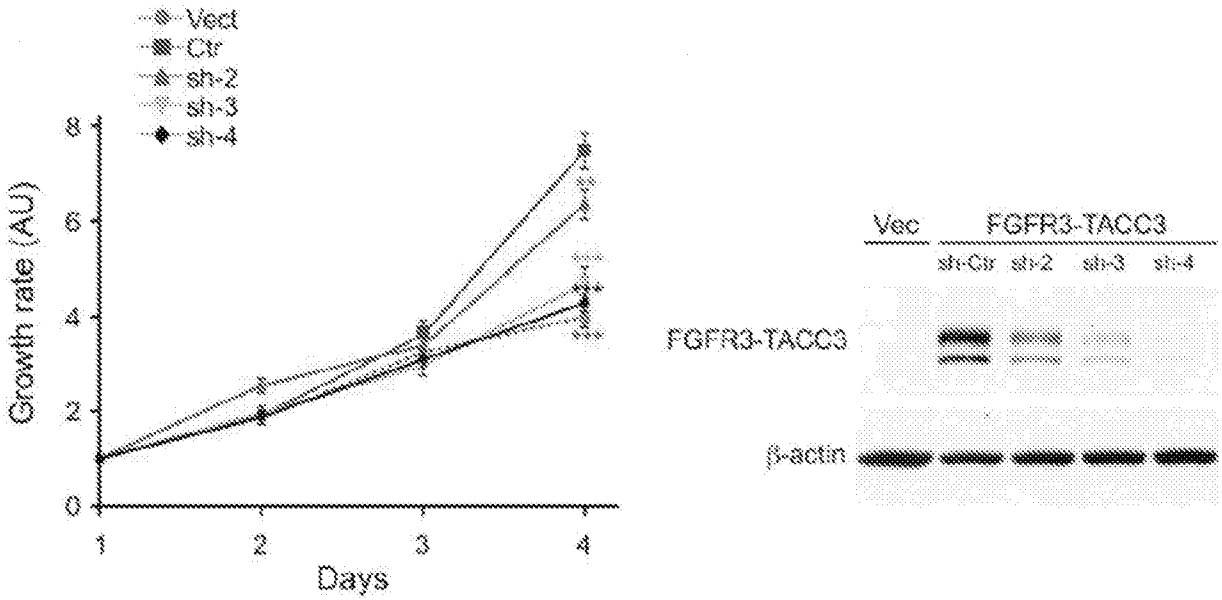


FIG. 12E

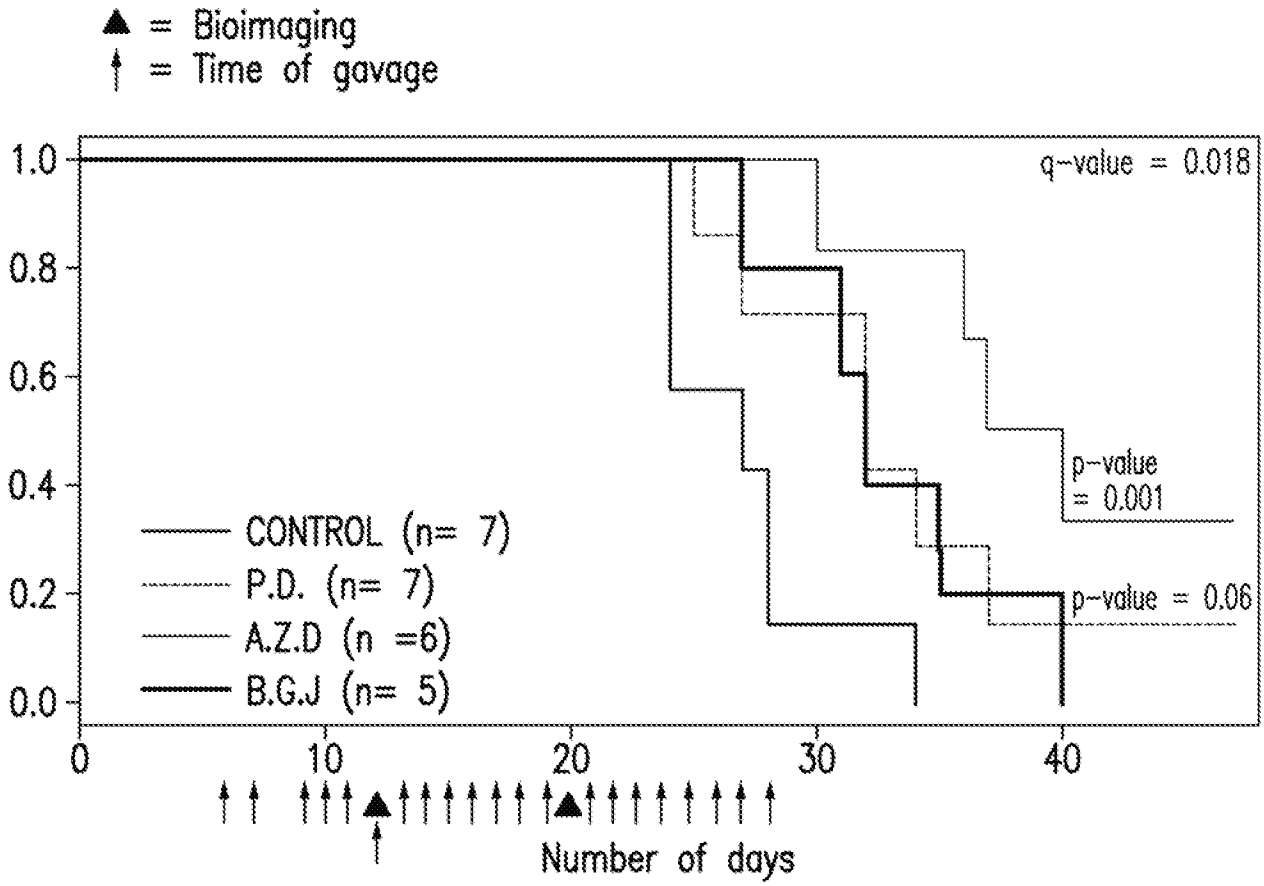


FIG. 13

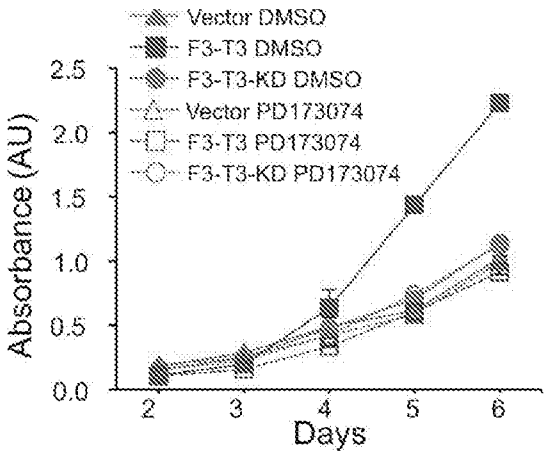


FIG. 14

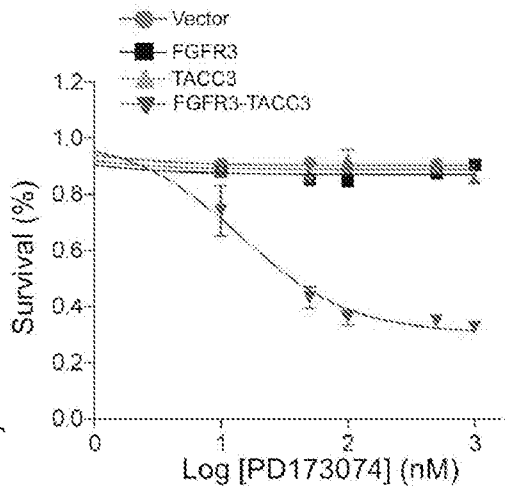


FIG. 15

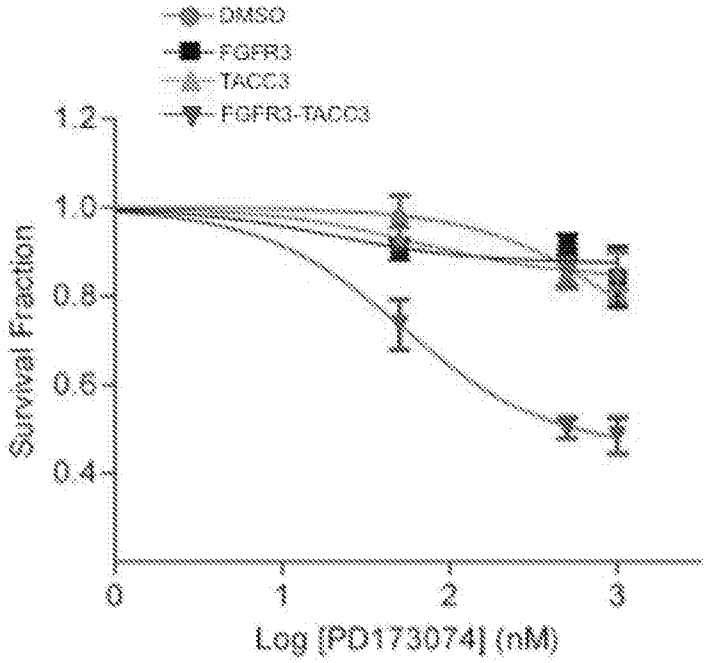


FIG. 16

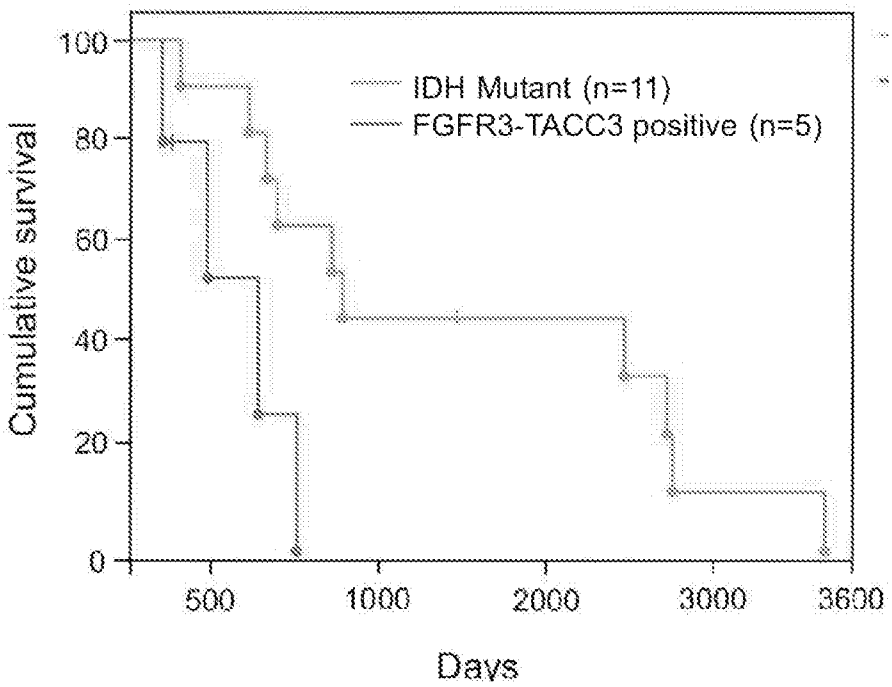

Cumulative survival
100
80
60
40
20
0
IDH Mutant (n=11)
FGFR3-TACC3 positive (n=5)
500
1000
2000
3000
3600
Days

FIG. 17

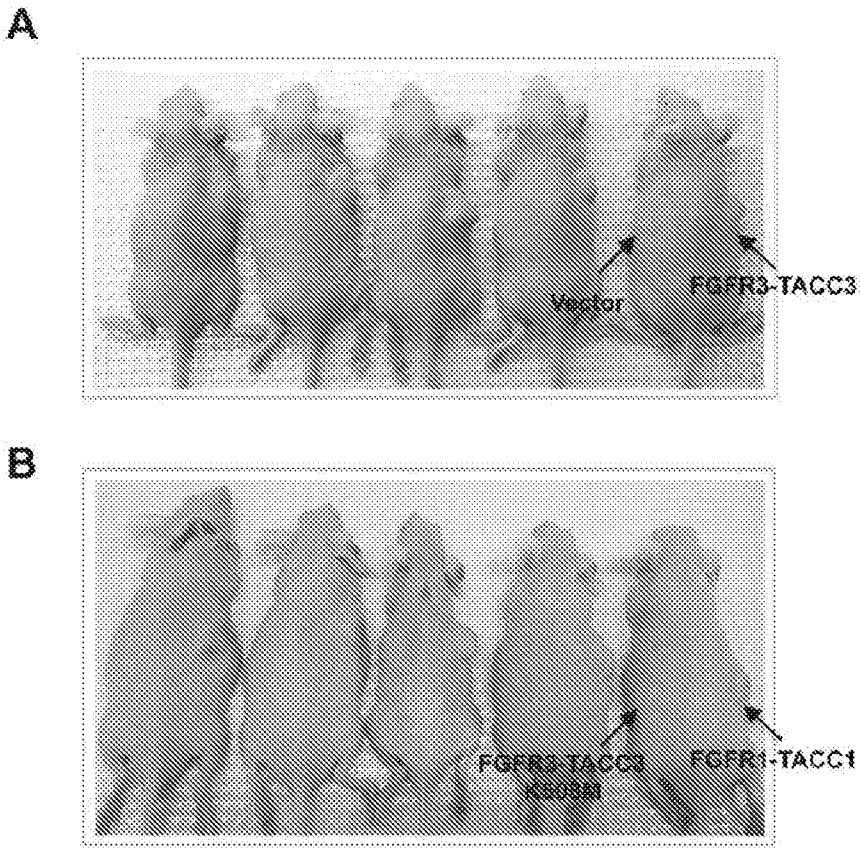

A
Vector
FGFR3-TACC3
B
FGFR1-TACC1

FIG. 18

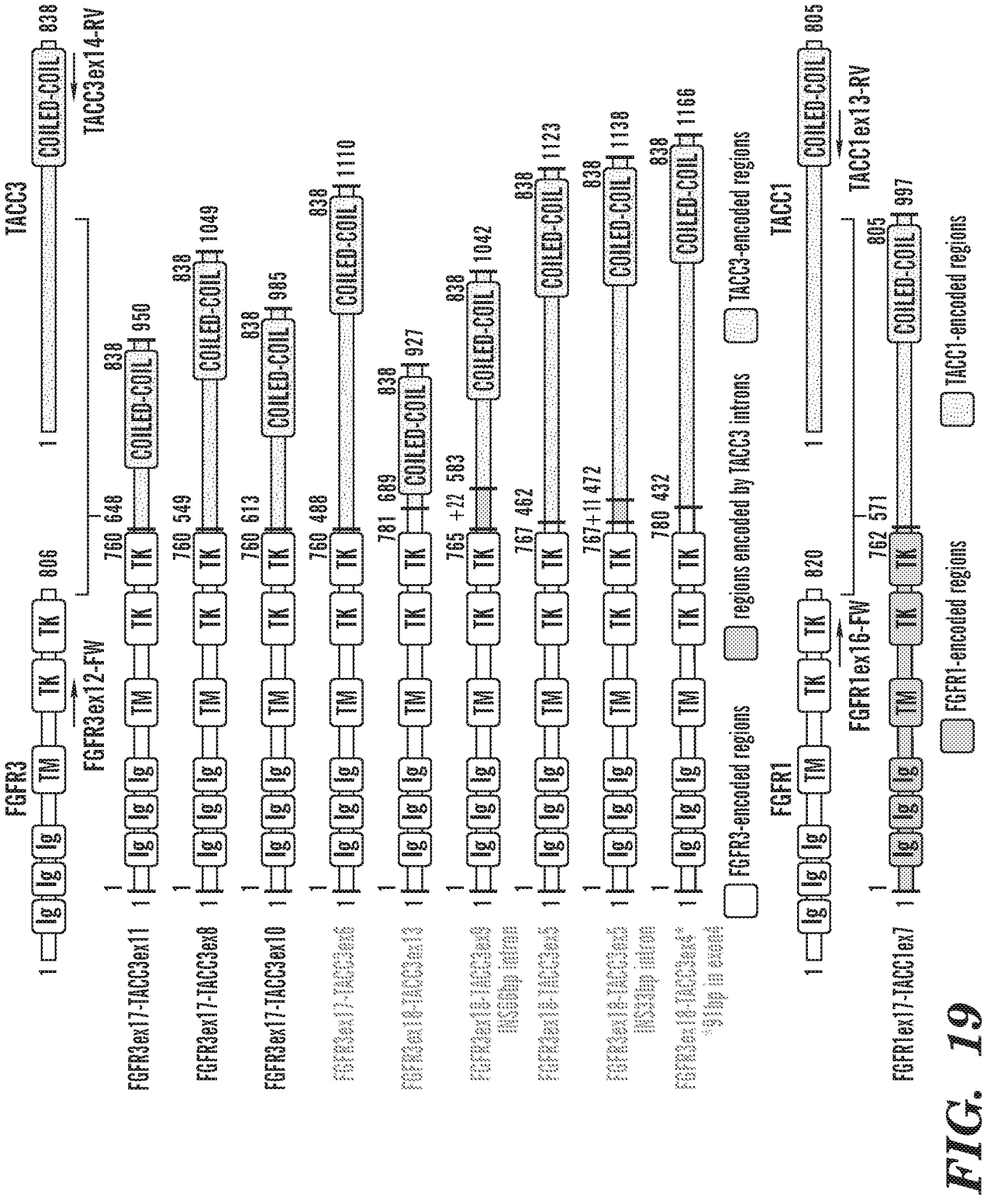
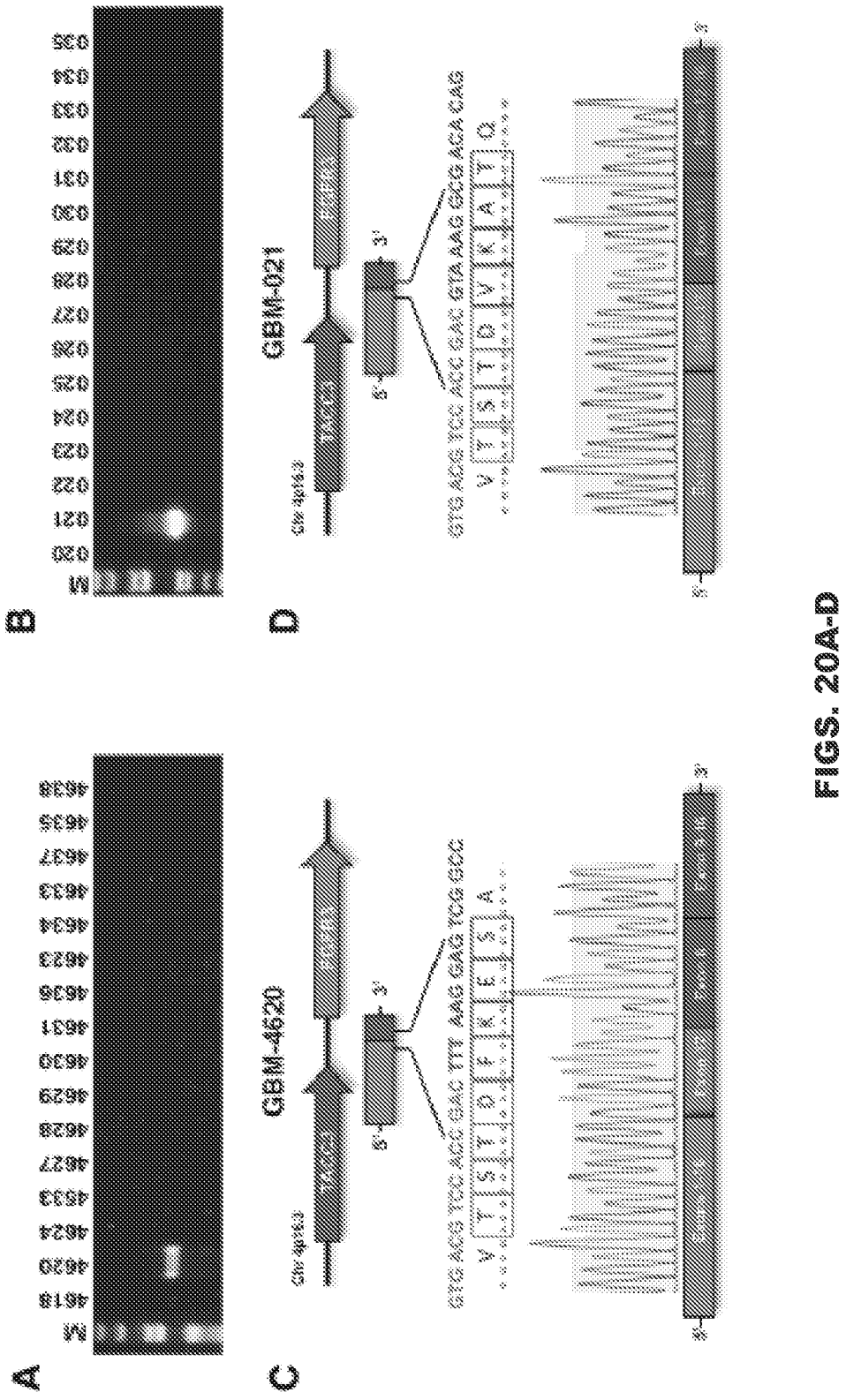
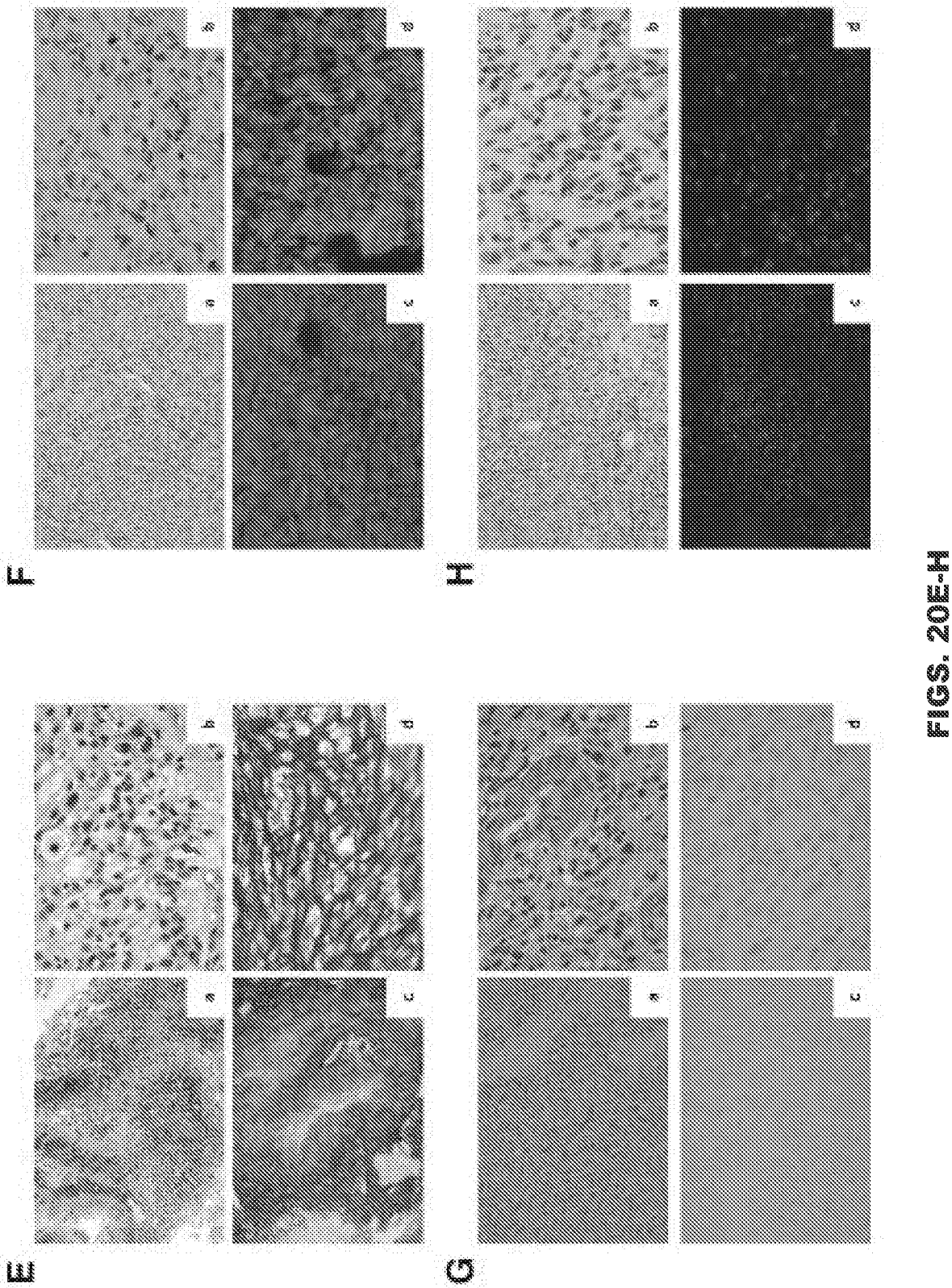

FGFR3
Ig
Ig
Ig
TM
TK
TK
806
1
FGFR3ex12-FW
TACC3
1
COILED-COIL
838
TACC3ex14-RV
FGFR3ex17-TACC3ex11
760 648
838
950
N=6
FGFR3ex17-TACC3ex8
760 549
838
1049
N=4
FGFR3ex17-TACC3ex10
760 613
838
985
N=3
760 488
838
1110
N=2
781 689
838
927
N=1
765 +22 583
838
1042
N=1
767 462
838
1123
N=1
767+11 472
838
1138
N=1
780 432
838
1166
N=1
FGFR3-encoded regions
regions encoded by TACC3 introns
TACC3-encoded regions
FGFR1
TACC1
FGFR1ex16-FW
762 571
805
997
TACC1ex13-RV
FGFR1ex17-TACC1ex7
FGFR1-encoded regions
TACC1-encoded regions

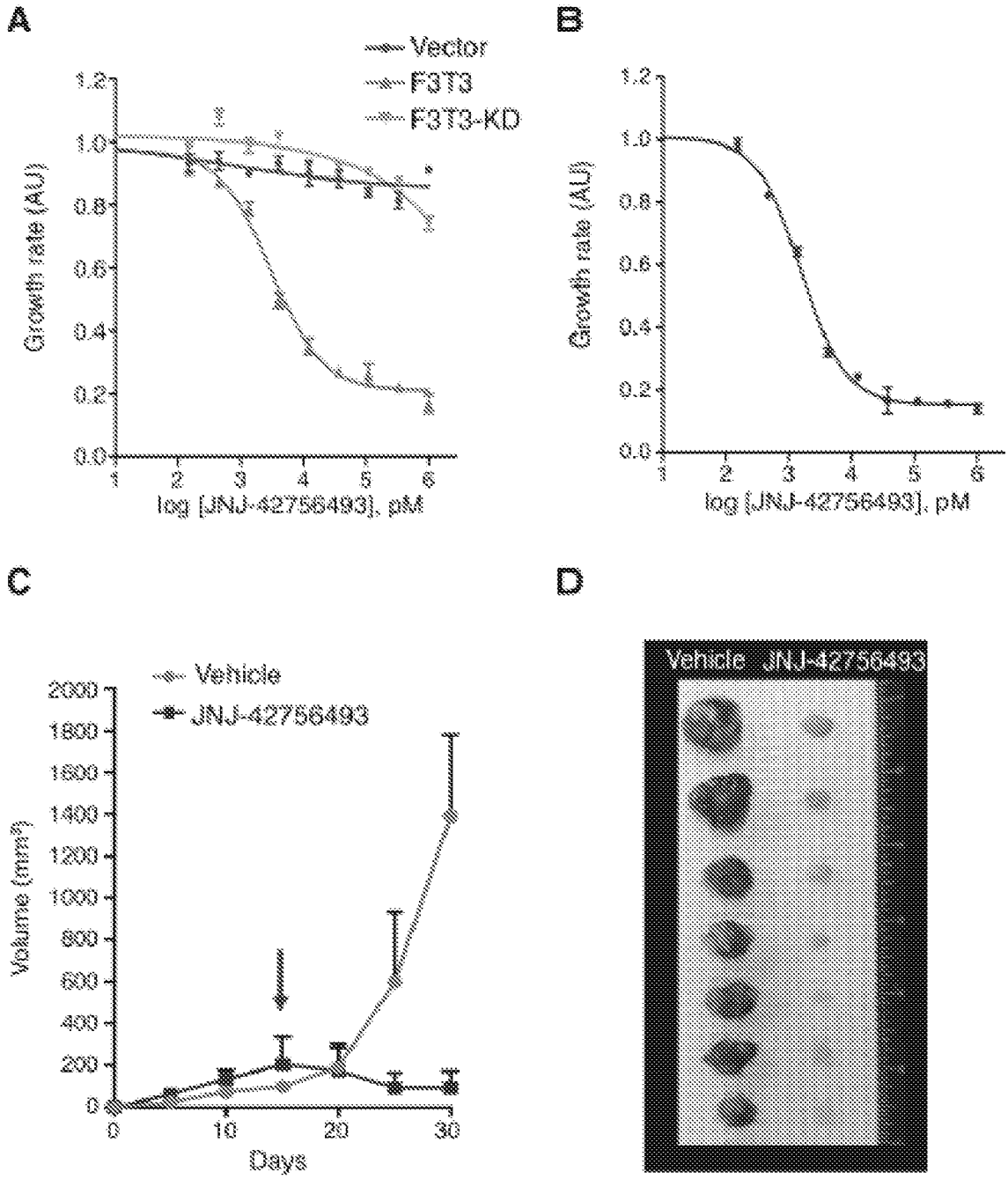
FIG. 19


A
B
C
D
GBM-4620
GBM-021
GTG ACG TCC ACC GAC TTT AAG GAG TCG GCC
V T S T D F K E S A
GTG ACG TCC ACC GAC GTA AAG GCG ACA CAG
V T S T D V K A T Q

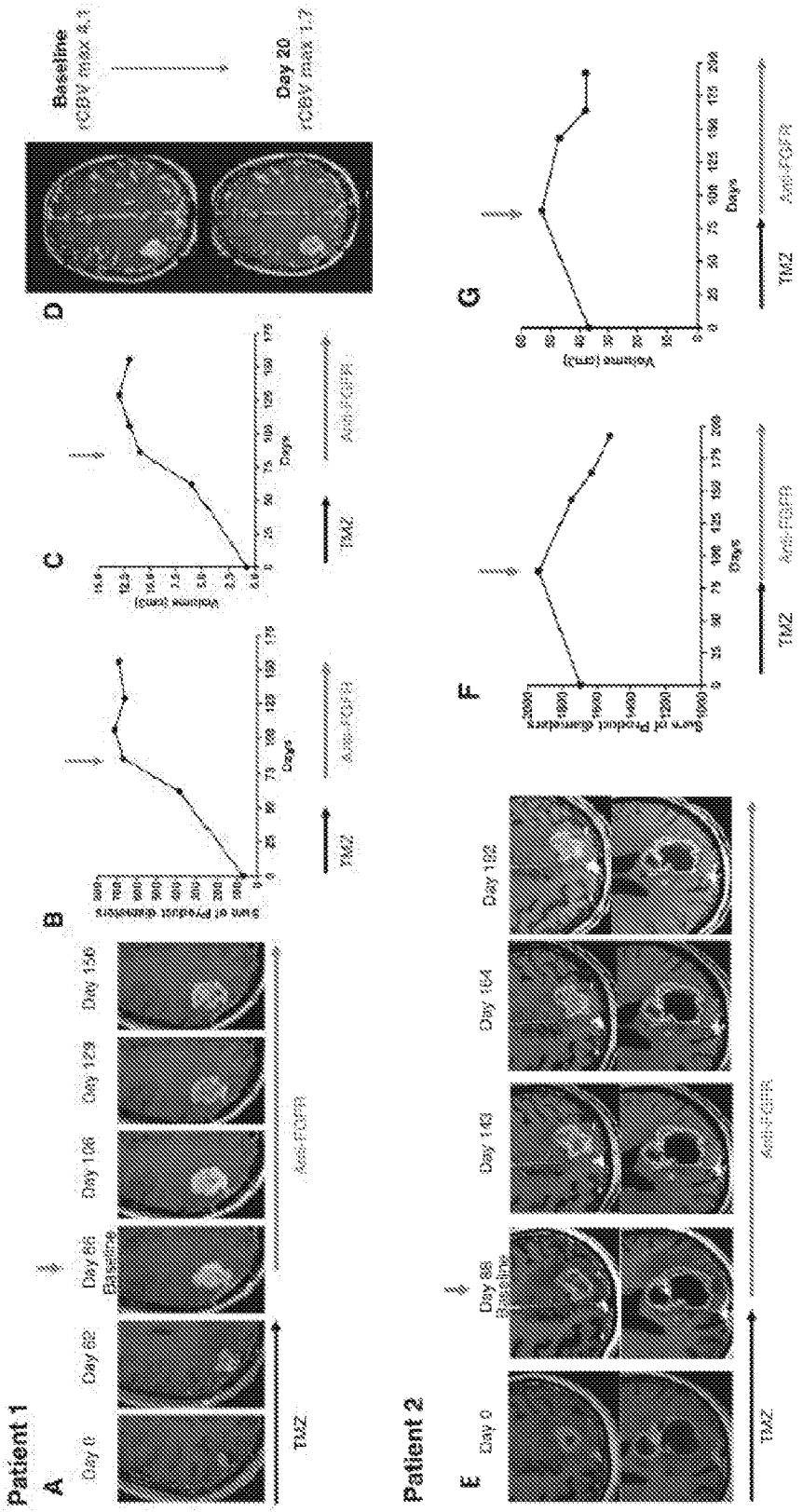
FIGS. 20A-D



FIGS. 20E-H

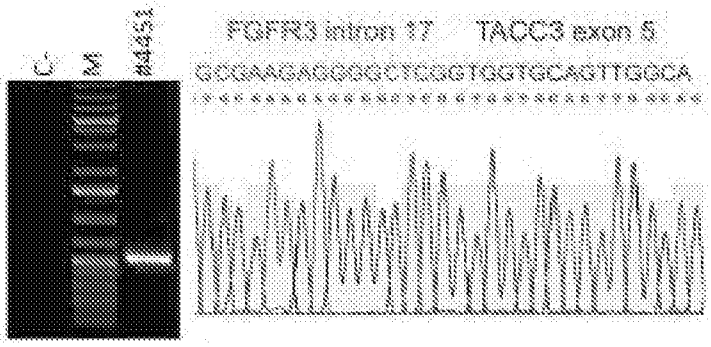

A
Vector
F3T3
F3T3-KD
Growth rate (AU)
log [JNJ-42756493], pM
B
Growth rate (AU)
log [JNJ-42756493], pM
C
Vehicle
JNJ-42756493
Volume (mm³)
Days
D
Vehicle
JNJ-42756493

FIGS. 21A-D

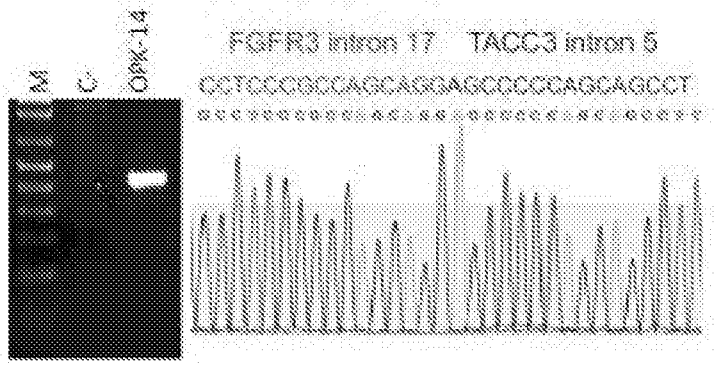

Patient 1
A
Day 0
Day 62
Day 86
Baseline
Day 106
Day 129
Day 156
TMZ
Anti-FGFR
B
Sum of Product diameters
Days
TMZ
Anti-FGFR
C
Volume (cm3)
Days
TMZ
Anti-FGFR
D
Baseline
rCBV max 4.3
Day 20
rCBV max 1.7
Patient 2
E
Day 0
Day 88
Baseline
Day 143
Day 164
Day 198
TMZ
Anti-FGFR
F
Sum of Product diameters
Days
TMZ
Anti-FGFR
G
Volume (cm3)
Days
TMZ
Anti-FGFR

FIGS. 22A-D

Sample #4451



Sample #OPK-14



Sample #MB-22

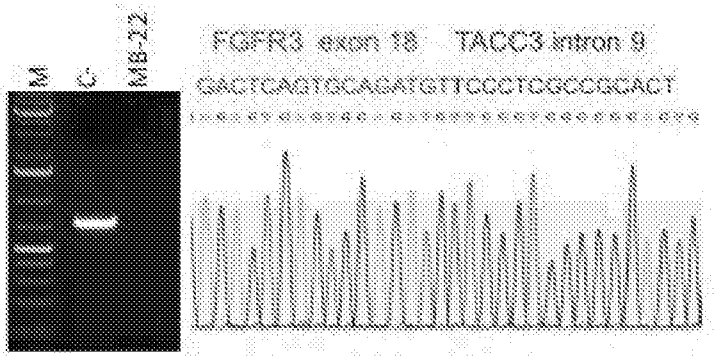
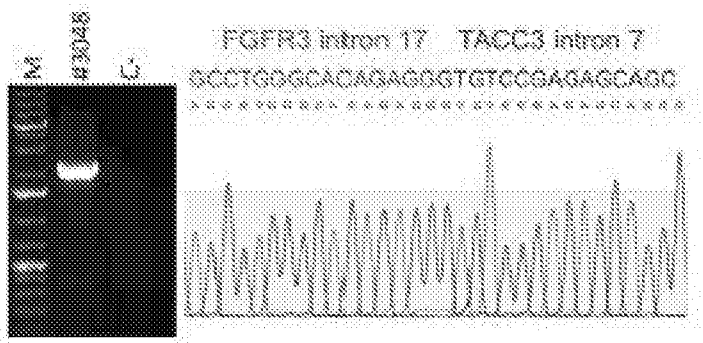
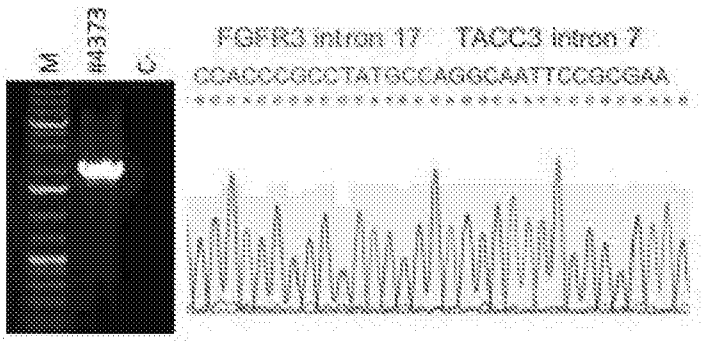


FIG. 23

Sample #3048



Sample #4373



Sample #4867

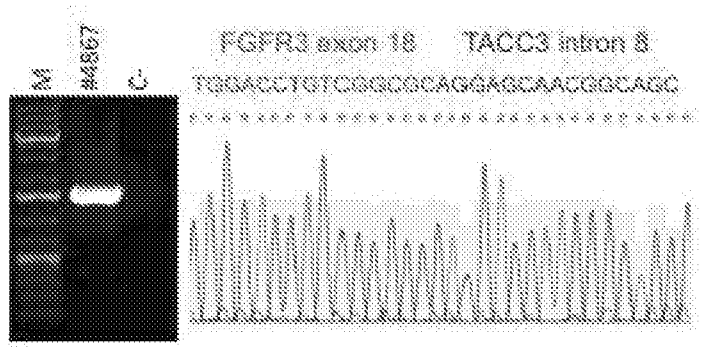
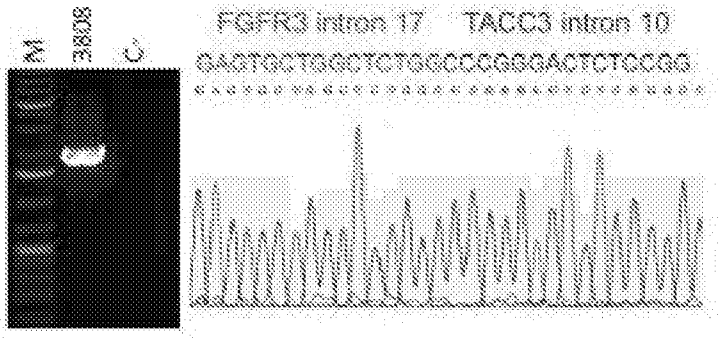
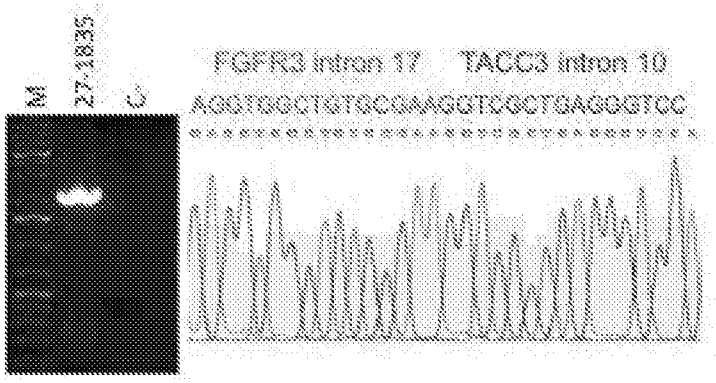


FIG. 23 cont.

Sample #3808



Sample #27-1835



Sample #06-6390

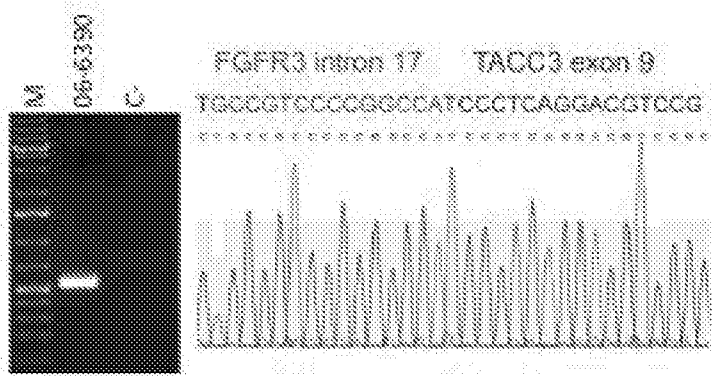


FIG. 23 cont.

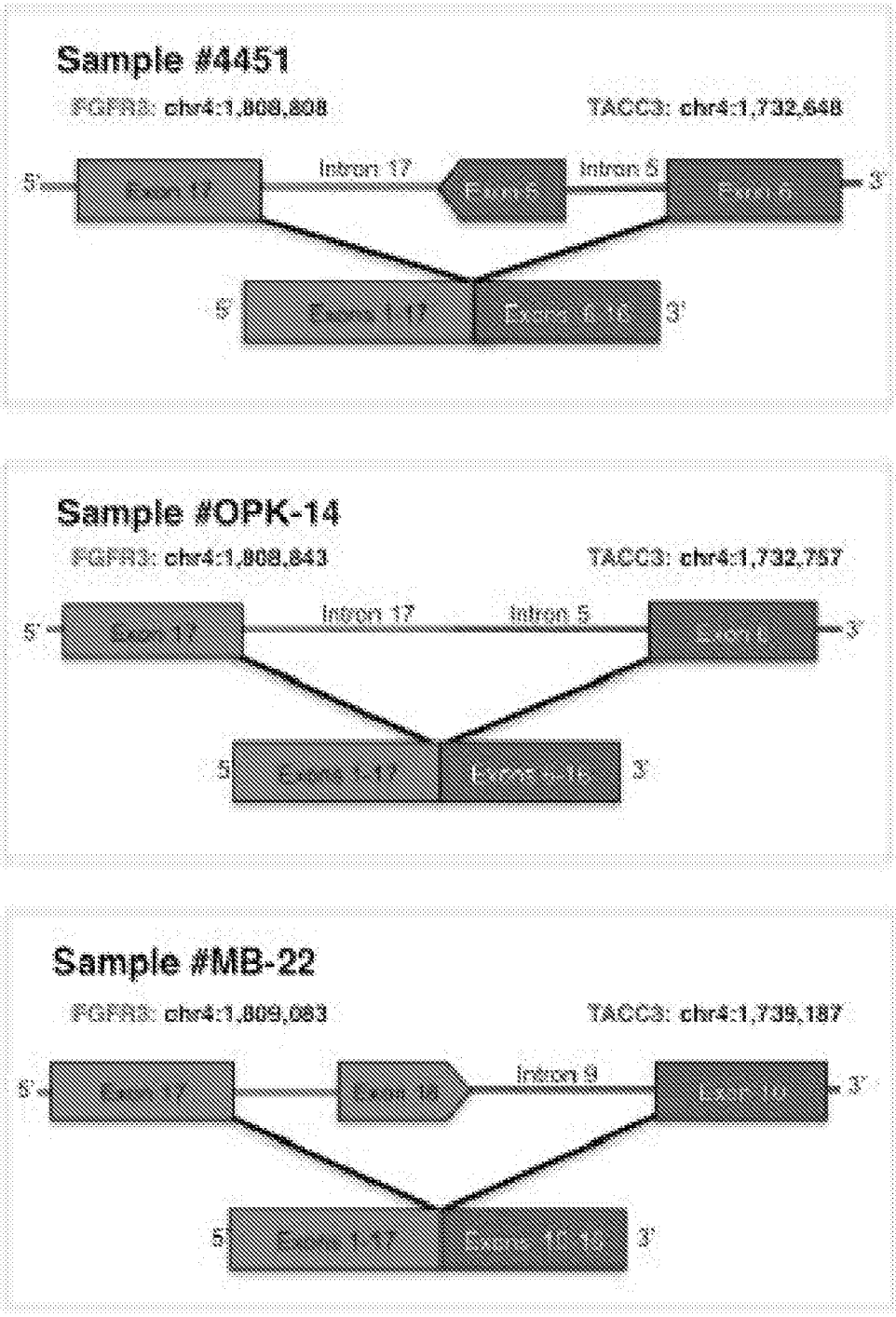

Sample #4451
FGFR3: chr4:1,808,808
TACC3: chr4:1,732,648
5'
Intron 17
Intron 8
3'
5'
3'
Sample #OPK-14
FGFR3: chr4:1,808,843
TACC3: chr4:1,732,757
5'
Intron 17
Intron 5
3'
5'
3'
Sample #MB-22
FGFR3: chr4:1,809,083
TACC3: chr4:1,739,187
5'
Intron 9
3'
5'
3'

FIG. 24

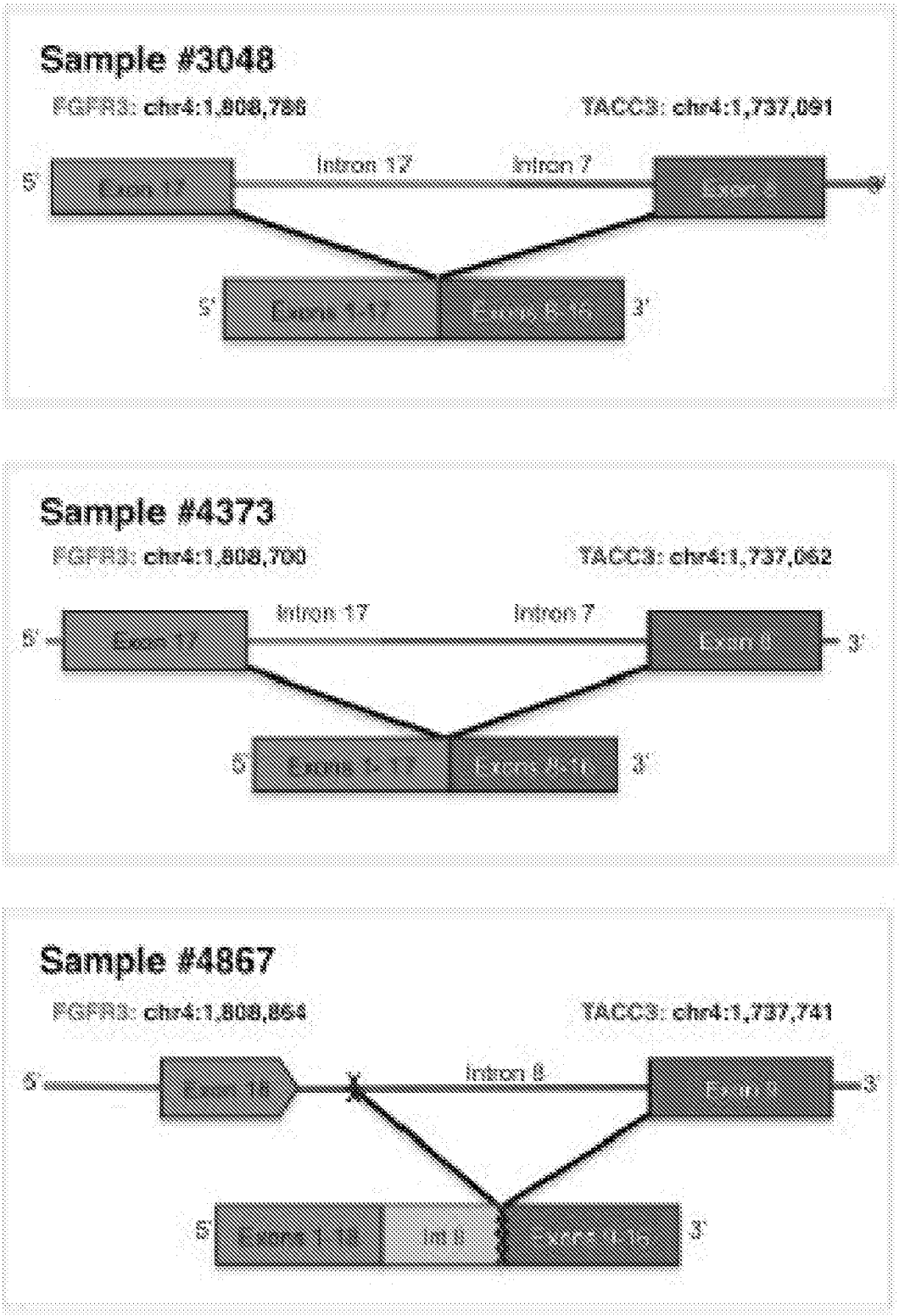

Sample #3048
FGFR3: chr4:1,808,786
TACC3: chr4:1,737,091
5'
Intron 17
Intron 7
3'
Sample #4373
FGFR3: chr4:1,808,700
TACC3: chr4:1,737,052
5'
Intron 17
Intron 7
3'
Sample #4867
FGFR3: chr4:1,808,864
TACC3: chr4:1,737,741
5'
Intron 8
3'

FIG. 24 cont.

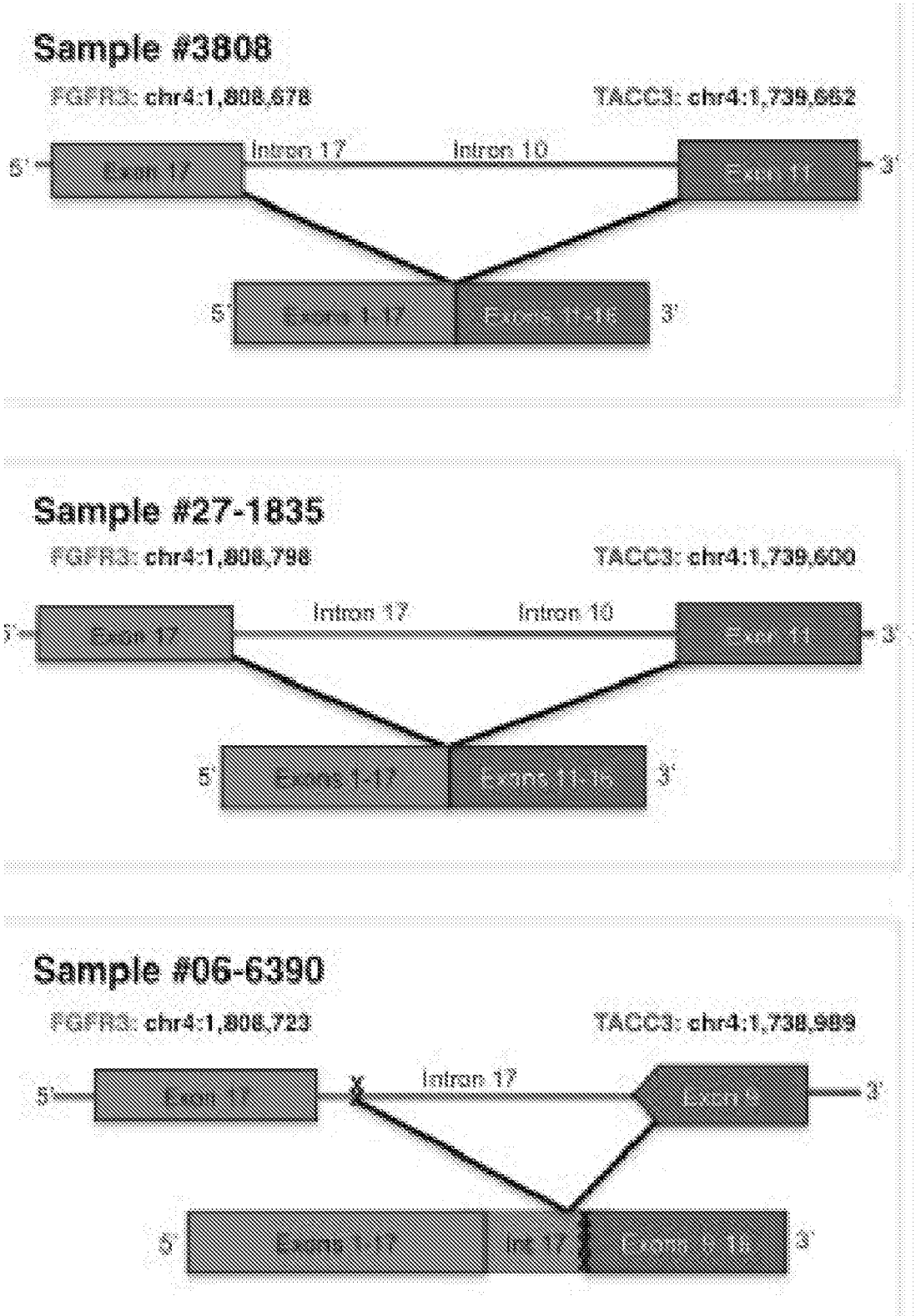

Sample #3808
FGFR3: chr4:1,808,678
TACC3: chr4:1,739,662
5'
Exon 17
Intron 17
Intron 10
3'
5'
3'
Sample #27-1835
FGFR3: chr4:1,808,798
TACC3: chr4:1,739,600
5'
Exon 17
Intron 17
Intron 10
3'
5'
3'
Sample #06-6390
FGFR3: chr4:1,808,723
TACC3: chr4:1,738,989
5'
Intron 17
3'
5'
3'

FIG. 24 cont.

A

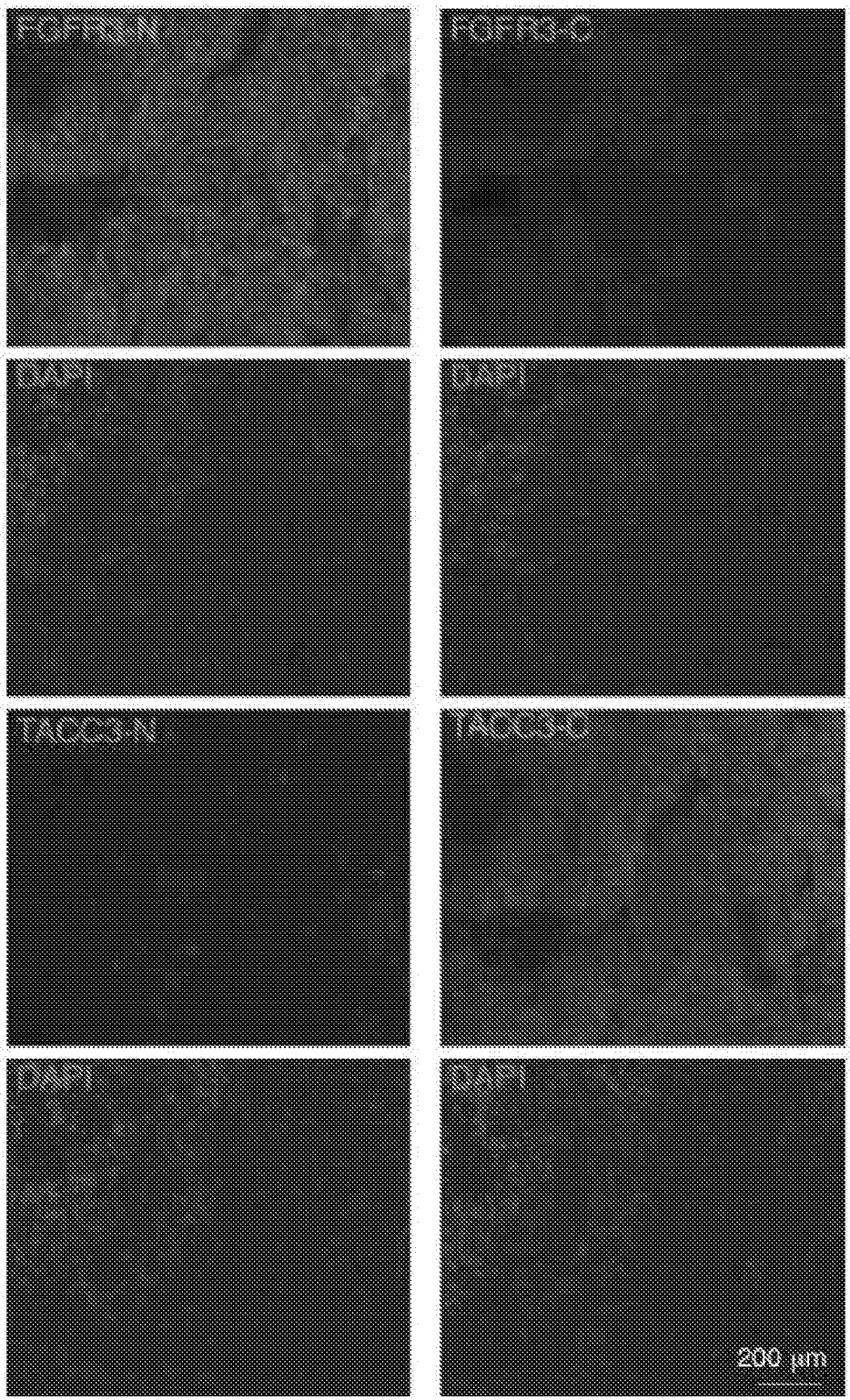


FIG. 25A

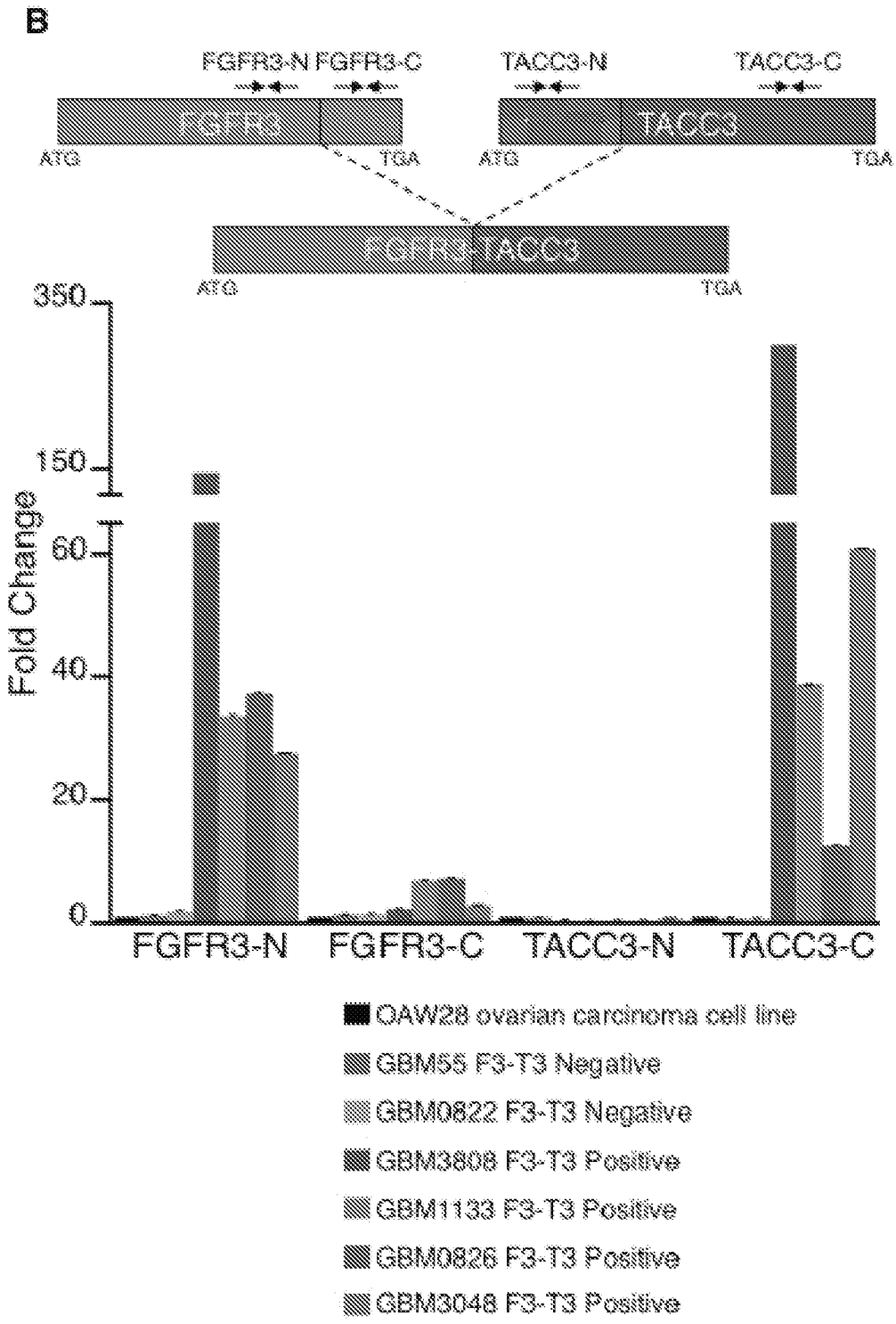
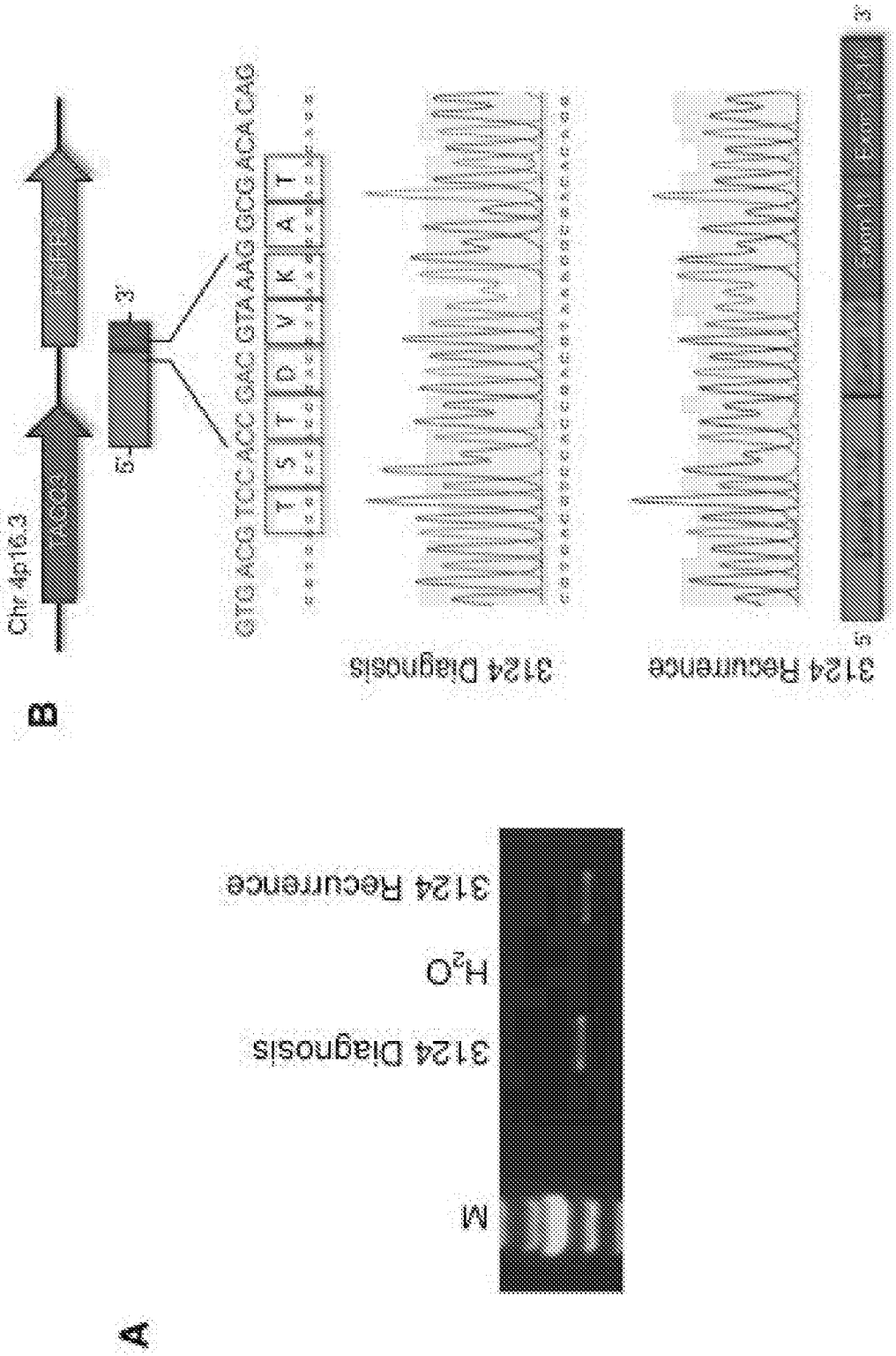

B
FGFR3-N
FGFR3-C
TACC3-N
TACC3-C
FGFR3
TACC3
ATG
TGA
ATG
TGA
FGFR3-TACC3
ATG
TGA
350
150
60
40
20
0
Fold Change
FGFR3-N
FGFR3-C
TACC3-N
TACC3-C
OAW28 ovarian carcinoma cell line
GBM55 F3-T3 Negative
GBM0822 F3-T3 Negative
GBM3808 F3-T3 Positive
GBM1133 F3-T3 Positive
GBM0826 F3-T3 Positive
GBM3048 F3-T3 Positive

FIG. 25B


A
M
3124 Diagnosis
H2O
3124 Recurrence
B
Chr 4p16.3
FGFR3
5'
3'
GTG ACG TCC ACC GAC GTA AAG GCG ACA CAG
V T S T D V K A T
3124 Diagnosis
3124 Recurrence
5'
3'

FIGS. 26A-B

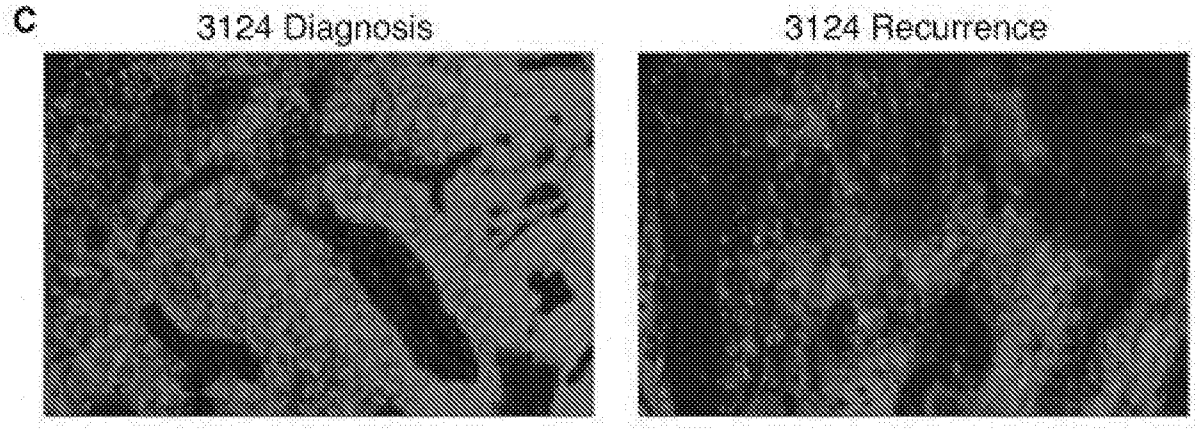
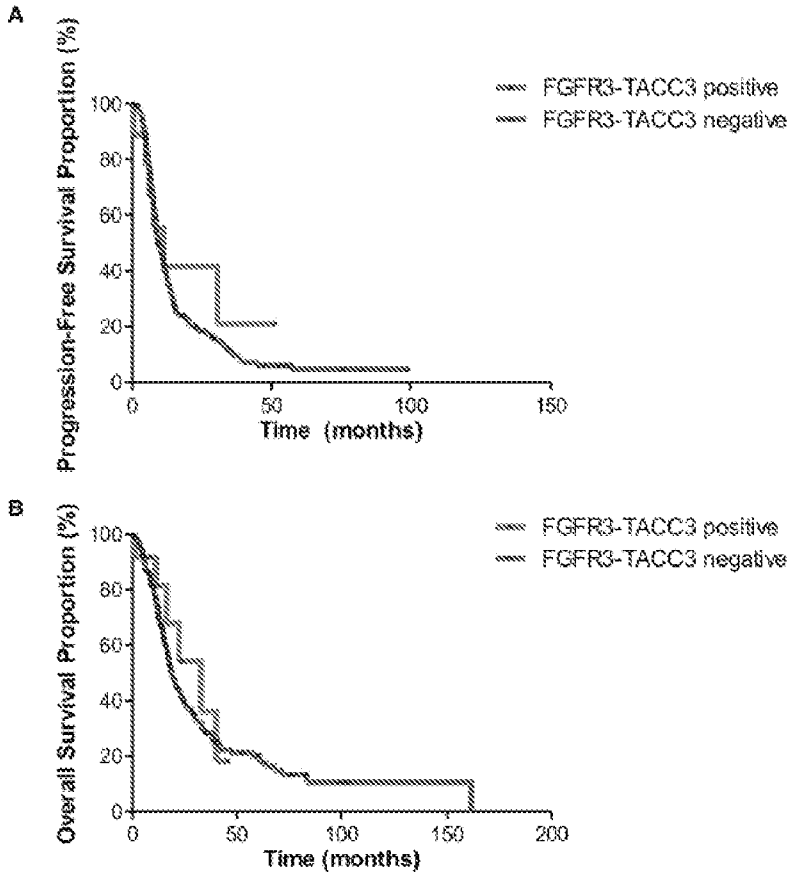


FIG. 26C



FIGS. 27A-B

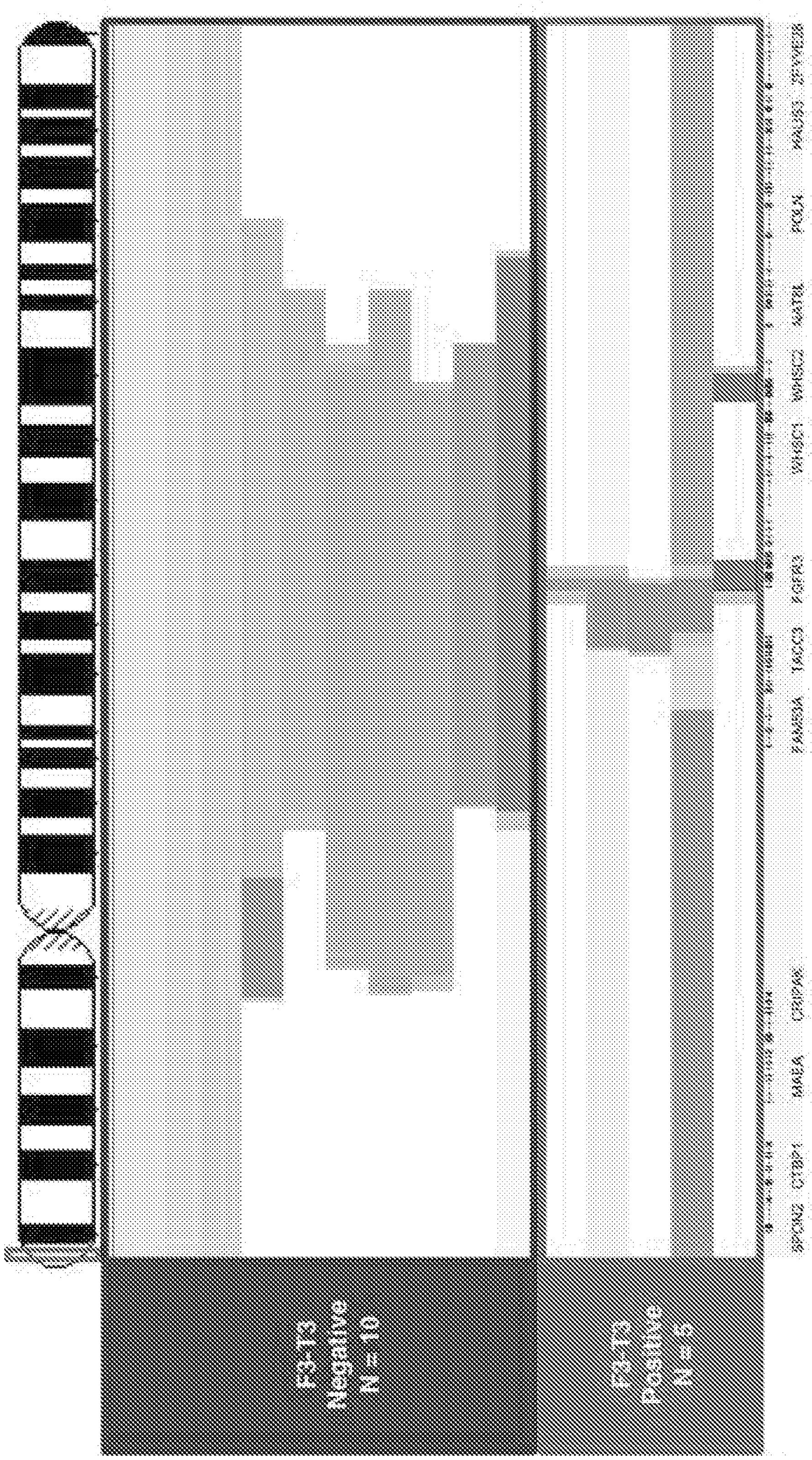

SPON2
CTBP1
MAEA
CRIPAK
FAM53A
TACC3
FGFR3
WHSC1
WHSC2
NAT8L
POLN
HAUS3
ZFYVE28
F3-T3 Negative N = 10
F3-T3 Positive N = 5

FIG. 28

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US15/00270

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C07K 16/40; G01N 33/574 (2016.01)
CPC - C07K 16/40, 14/82; C12N 9/12
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) Classification(s): C07K 16/40; G01N 33/574;
CPC Classification(s): C07K 16/40, 14/82; C12N 9/12

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)
USPTO Web Page; Google Patents; PatSeer (US, EP, WO); Google; Google Scholar; EBSCO; Entrez Pubmed; Science Direct; Search Terms -- FGFR, 'fusion protein, TCAA, amplification, EGFR, hybridization, antibody, IDH1, primers, probe, ELISA, 'Western blot'

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2014/018673 A2 (TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK) January 30, 2014; paragraphs [0007], [0010], [0033], [0036]-[0039], [00199], [00220], [00262], [00266], [00287], [00288]; SEQ ID NOs: 1, 79, 162, 163	1-10
Y	WO 2014/151734 A1 (TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK) September 25, 2014; paragraphs [0036], [0039], [00109], [00171], [00240], [00398], [00321]	1-10
A	(SHAH, N. et al.) Exploration of the Gene Fusion Landscape of Glioblastoma Using Transcriptome Sequencing and Copy Number Data. BMC Genomics. 22 November 2013, Vol 14; page 818; DOI: 10.1186/1471-2164-14-818	1-10
A	(BORAD, MJ et al.) Integrated Genomic Characterization Reveals Novel, Therapeutically Relevant Drug Targets in FGFR and EGFR Pathways in Sporadic Intrahepatic Cholangiocarcinoma. PLoS Genetics. 13 February 2014, Vol. 10, No. 2; e1004135; DOI: 10.1371/journal.pgen.1004135	1-10

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

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

Date of the actual completion of the international search	Date of mailing of the international search report
3 May 2016 (03.05.2016)	13 MAY 2016
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer Shane Thomas PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

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

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US15/00270

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
-***-Please see supplemental page-***-

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-10; SEQ ID NOs 1, 79, 162, 163

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

-***-Continued from Box No. III: Observations where unity of invention is lacking-***-

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Groups I+, Claims 1-17, SEQ ID NO: 79, SEQ ID NO: 1 and SEQ ID NOs: 162 and 163 are directed toward a method for detecting the presence of a FGFR fusion protein and the presence or absence of a gene amplification of EGFR, CDK4 or MDM2 in a human subject; kits for said method; antibodies for said method; purified proteins, antibodies and cDNA for said method, and a composition and method for decreasing in a subject the expression level or activity of a FGFR3-TACC3 fusion protein.

The methods, kits, composition antibodies, proteins and cDNA will be searched to the extent that they encompass an antibody directed to SEQ ID NO: 79 (first exemplary target protein sequence), a nucleic acid sequence encompassing SEQ ID NO: 1 (first exemplary target nucleic acid sequence), and primers encompassing SEQ ID NO: 162 (first exemplary primer sequence) and SEQ ID NO: 163 (second exemplary primer sequence). Applicant is invited to elect additional target protein(s) with specified SEQ ID NO: for each, target nucleic acid(s) with specified SEQ ID NO: for each, and/or specified primer set(s) with specified pair(s) of SEQ ID NOs for each pair of primers, to be searched. Additional target protein sequence(s), target nucleic acid sequence(s) and/or pair(s) of primer sequences will be searched upon the payment of additional fees. It is believed that claims 1, 2 (in-part), 3-7, 8 (in-part), 9, and 10 (in-part) encompass this first named invention and thus these claims will be searched without fee to the extent that they encompass SEQ ID NO: 79 (target protein sequence), SEQ ID NO: 1 (target nucleic acid sequence), SEQ ID NO: 162 (first primer sequence) and SEQ ID NO: 163 (second primer sequence). Applicants must specify the claims that encompass any additionally elected target protein, target nucleic acid and/or pair(s) of primer sequence(s). Applicants must further indicate, if applicable, the claims which encompass the first named invention, if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined. An exemplary first election would be a target protein encompassing SEQ ID NO: 85 (first exemplary elected target protein sequence). An exemplary second election would be a target nucleic acid encompassing SEQ ID NO: 2 (first exemplary elected target nucleic acid sequence). An exemplary third election would be a pair of primers encompassing SEQ ID NO: 164 (first exemplary elected primer sequence) and SEQ ID NO: 165 (second exemplary elected primer sequence).

No technical features are shared between the target protein sequences, target nucleic acid sequences and/or primer sequences of Groups I+ and, accordingly, these groups lack unity a priori.

Groups I+ share the technical features including: a method for detecting the presence of a FGFR fusion protein and the presence or absence of a gene amplification of EGFR, CDK4 or MDM2 in a human subject, wherein the FGFR fusion protein comprises the tyrosine kinase domain of a FGFR protein fused to the TACC domain of a TACC protein, the method comprising: (a) obtaining a biological sample from the human subject; (b) detecting whether or not there is a FGFR fusion present in the subject; and (c) detecting whether or not there is gene amplification of EGFR, CDK4, or MDM2; a diagnostic kit for determining whether a sample from a subject exhibits a presence of a FGFR fusion protein and the presence or absence of a gene amplification of EGFR, CDK4 or MDM2, the kit comprising at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding FGFR fusion protein, or a portion thereof, wherein the FGFR fusion protein comprises the tyrosine kinase domain of a FGFR protein fused to the TACC domain of a TACC protein; the kit further comprising at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding EGFR, CDK4, or MDM2; a diagnostic kit for determining whether a sample from a subject exhibits a presence of a FGFR fusion protein and the presence or absence of a gene amplification of EGFR, CDK4 or MDM2, the kit comprising an antibody that specifically binds to a FGFR fusion protein, wherein the antibody will recognize the protein only when a FGFR fusion protein is present, and wherein the FGFR fusion protein comprises a tyrosine kinase domain of an FGFR protein fused to the TACC domain of a transforming acidic coiled-coil-containing (TACC) protein; the kit further comprising at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding EGFR, CDK4, or MDM2; an antibody or antigen-binding fragment thereof, that specifically binds to a purified FGFR3-TACC3 fusion protein; a composition for decreasing in a subject the expression level or activity of a FGFR3-TACC3 fusion protein, the composition in an admixture of a pharmaceutically acceptable carrier comprising an inhibitor of the fusion protein; a method of decreasing growth of a solid tumor in a subject in need thereof, the method comprising administering to the subject an effective amount of a FGFR3-TACC3 fusion protein molecule inhibitor, wherein the inhibitor decreases the size of the solid tumor; a purified FGFR2-TACC3 fusion protein; and a cDNA encoding a FGFR2-TACC3 fusion protein.

However, these shared technical features are previously disclosed by WO 2014/018673 A2 to The Trustees of Columbia University in the City of New York (hereinafter 'Columbia') in view of the article 'Exploration of the gene fusion landscape of glioblastoma using transcriptome sequencing and copy number data' by Shah et al. (hereinafter 'Shah'), and the article 'Integrated Genomic Characterization Reveals Novel, Therapeutically Relevant Drug Targets in FGFR and EGFR Pathways in Sporadic Intrahepatic Cholangiocarcinoma' by Borad et al. (hereinafter 'Borad').

-***-Continued Within the Next Supplemental Box-***-

-***-Continued from Previous Supplemental Box-***-

Columbia discloses a method for detecting the presence of a FGFR fusion protein (a method for detecting the presence of a FGFR fusion protein; paragraphs [0033], [0038]) in a human subject (in a sample from a human subject; paragraph [0038]), wherein the FGFR fusion protein comprises the tyrosine kinase domain of a FGFR protein fused to the TACC domain of a TACC protein (wherein the FGFR fusion protein comprises the tyrosine kinase domain of a FGFR protein fused to the TACC domain of a TACC protein; paragraph [0038]), the method comprising: (a) obtaining a biological sample from the human subject (the method comprising: (a) obtaining a biological sample from the human subject; paragraph [0038]); (b) detecting whether or not there is a FGFR fusion present in the subject ((b) detecting whether or not there is a FGFR fusion present in the subject; paragraph [0038]); a diagnostic kit for determining whether a sample from a subject exhibits a presence of a FGFR fusion protein (a diagnostic kit for determining whether a sample from a subject exhibits a presence of a FGFR fusion protein; paragraph [0036]), the kit comprising at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding FGFR fusion protein, or a portion thereof (the kit comprising at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding FGFR fusion protein, or a portion thereof; paragraph [0036]), wherein the FGFR fusion protein comprises the tyrosine kinase domain of a FGFR protein fused to the TACC domain of a TACC protein (wherein the FGFR fusion protein comprises the tyrosine kinase domain of a FGFR protein fused to the TACC domain of a TACC protein; paragraph [0036]); a diagnostic kit for determining whether a sample from a subject exhibits a presence of a FGFR fusion protein (a diagnostic kit for determining whether a sample from a subject exhibits a presence of a FGFR fusion protein; paragraph [0037]), the kit comprising an antibody that specifically binds to a FGFR fusion protein (the kit comprising an antibody that specifically binds to a FGFR fusion protein; paragraph [0037]), wherein the antibody will recognize the protein only when a FGFR fusion protein is present (wherein the antibody will recognize the protein only when a FGFR fusion protein is present; paragraph [0037]), and wherein the FGFR fusion protein comprises a tyrosine kinase domain of an FGFR protein fused to the TACC domain of a transforming acidic coiled-coil-containing (TACC) protein (and wherein the FGFR fusion protein comprises a tyrosine kinase domain of an FGFR protein fused to the TACC domain of a transforming acidic coiled-coil-containing (TACC) protein; paragraph [0037]); an antibody or antigen-binding fragment thereof, that specifically binds to a purified FGFR3-TACC3 fusion protein (an antibody or antigen-binding fragment thereof, that specifically binds to a purified FGFR3-TACC3 fusion protein; paragraph [0031]); a composition for decreasing in a subject the expression level or activity of a FGFR3-TACC3 fusion protein (a composition for decreasing in a subject the expression level or activity of a FGFR3-TACC3 fusion protein; paragraphs [0031], [0032]), the composition in an admixture of a pharmaceutically acceptable carrier comprising an inhibitor of the fusion protein (the composition in an admixture of a pharmaceutically acceptable carrier comprising an inhibitor of the fusion protein; paragraph [0032]); a method of decreasing growth of a solid tumor in a subject in need thereof (a method of decreasing growth of a solid tumor in a subject in need thereof; paragraph [0035]), the method comprising administering to the subject an effective amount of a FGFR3-TACC3 fusion protein molecule inhibitor (the method comprising administering to the subject an effective amount of a FGFR3-TACC3 fusion protein molecule inhibitor; paragraph [0035]), wherein the inhibitor decreases the size of the solid tumor (wherein the inhibitor decreases the size of the solid tumor; paragraph [0035]). Columbia further discloses wherein the FGFR protein is an FGFR2 protein (wherein the FGFR protein is an FGFR2 protein; paragraph [0010]), and wherein the TACC protein is a TACC3 protein (wherein the TACC protein is a TACC3 protein; paragraph [0010]); cDNA molecules obtained after retro-transcription of RNA (cDNA molecules obtained after retro-transcription of RNA; paragraph [00306]); wherein measuring the FGFR fusion protein includes an antibody directed to SEQ ID NO: 79 (wherein measuring the FGFR fusion protein includes an antibody directed to SEQ ID NO: 79; paragraphs [0031], [0038]; SEQ ID NO: 79, wherein SEQ ID NO: 79 is 100% identical to Applicants' SEQ ID NO: 79); wherein the FGFR fusion protein comprises SEQ ID NO: 1 (wherein the FGFR fusion protein comprises SEQ ID NO: 1; paragraph [0045]; SEQ ID NO: 1, wherein SEQ ID NO: 1 is 100% identical to Applicants' SEQ ID NO: 1), and the use of primers for amplification including SEQ ID NOs: 162 and 163 (use of primers for amplification including SEQ ID NOs: 162 and 163; paragraph [0208]; SEQ ID NOs: 162 and 163, wherein SEQ ID NOs: 162 and 163 are each 100% identical to Applicants' SEQ ID NOs: 162 and 163, respectively).

Columbia does not disclose the presence or absence of a gene amplification of EGFR, CDK4 or MDM2; and (c) detecting whether or not there is gene amplification of EGFR, CDK4, or MDM2; the kit further comprising at least one oligonucleotide that specifically hybridizes to a nucleic acid encoding EGFR, CDK4, or MDM2; a purified FGFR2-TACC3 fusion protein; and a cDNA encoding a FGFR2-TACC3 fusion protein.

Shah discloses an analysis of the gene fusion landscape in glioblastoma patients (analysis of the gene fusion landscape in glioblastoma patients; abstract), including fusions of FGFR with TACC (including fusions of FGFR with TACC; page 3, paragraph 3 - page 4, paragraph 1); wherein amplifications of EGFR were more prevalent in samples with at least one fusion (wherein amplifications of EGFR were more prevalent in samples with at least one fusion; page 8, paragraph 3).

Borad discloses FGFR2-TACC3 fusions in cancer cells (FGFR2-TACC3 fusions in cancer cells; abstract).

It would have been obvious to a person of ordinary skill in the art at the time of the invention was made to have modified the disclosure of Columbia to have included the determination of the presence of additional characteristics of glioblastoma, including amplification of a gene including EGFR, as disclosed by Shah, in the methods and kits disclosed by Columbia by providing oligonucleotides that specifically hybridize to a target EGFR sequence, in a similar manner to the oligonucleotides disclosed by Columbia, in order to better determine an appropriate course of treatment for a subject having both an FGFR fusion, as disclosed by Columbia, as well as amplification of a normal or mutant EGFR, as disclosed by Shah, in order to enable more effective treatment of a subject with glioblastoma. It further would have been obvious to a person of ordinary skill in the art at the time of the invention was made to have modified the disclosure of Columbia to have included the detection of additional FGFR fusion proteins associated with cancers, such as an FGFR2-TACC3 fusion, as disclosed by Borad, using the reagents disclosed by Columbia for the detection of FGFR2 and TACC3 fusion proteins, in order to better determine the presence of a wider range of specific FGFR fusion proteins in a sample from a subject in order to determine an appropriate course of treatment for said subject.

Since none of the special technical features of the Groups I+ inventions is found in more than one of the inventions, and since all of the shared technical features are previously disclosed by a combination of the Columbia, Shah and Borad references, unity of invention is lacking.