

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
18 May 2006 (18.05.2006)

PCT

(10) International Publication Number  
**WO 2006/053155 A2**

(51) International Patent Classification: Not classified

**OKAMURA, Heidi** [US/US]; 372 Longwood Avenue,  
No. 11, Boston, Massachusetts 02215 (US).

(21) International Application Number:  
PCT/US2005/040814

(74) Agent: **HALEY, James, F., Jr.**; Fish & Neave IP Group,  
Ropes & Gray LLP, 1251 Avenue Of The Americas, New  
York, New York 10020 (US).

(22) International Filing Date:  
12 November 2005 (12.11.2005)

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

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/627,552 12 November 2004 (12.11.2004) US

(84) Designated States (*unless otherwise indicated, for every  
kind of regional protection available*): ARIPO (BW, GH,  
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,  
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,  
FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,  
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,  
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): **AVEO  
PHARMACEUTICALS, INC.** [US/US]; 75 Sidney  
Street, 4th Floor, Cambridge, Massachusetts 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **CHIU, Maria, I.**  
[US/US]; 71 Montvale Crescent, Newton Centre, Mass-  
achusetts 02459 (US). **ROBINSON, Murray** [US/US];  
1200 Washington Street, Boston, Massachusetts 02118  
(US). **O'HAGAN, Ronan** [CA/US]; 15 Garrison Road,  
Apt. 2, Brookline, Massachusetts 02445 (US). **KAN-  
NAN, Karuppiah** [IN/US]; 170 Gore Street, Cambridge,  
Massachusetts 02141 (US). **BAILEY, David** [GB/US]; 68  
Hobart Steet, #2, Brighton, Massachusetts 02135 (US).  
**LERNER, Lorena** [AR/US]; 37 Westgate Road, Newton  
Centre, Massachusetts 02459 (US). **CAI, Ti** [CN/US]; 4  
Wi Ld Rose Drive, Andover, Massachusetts 01810 (US).

**Published:**

— without international search report and to be republished  
upon receipt of that report

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

(54) Title: GP202: METHODS AND COMPOSITIONS FOR TREATING CANCER

(57) Abstract: The use of molecules relating to the GP202 (KIRREL) gene for treating cancer and other hyperproliferative condi-  
tions is disclosed. Non-human mammals harboring a genetic modification relating to the GP202 gene, and their use as experimental  
cancer models, are disclosed.



WO 2006/053155 A2

**GP202: METHODS AND COMPOSITIONS FOR TREATING CANCER****Cross-Reference To Related Application**

This application claims the benefit of U.S. provisional patent application no. 60/627,552, filed November 12, 2004, the disclosure of which is incorporated  
5 herein by reference in its entirety.

**Field of the Invention**

The field of the invention is molecular biology and oncology.

**Background of the Invention**

Surprisingly little is known about the genetic lesions responsible for the  
10 genesis, progression, and clinical behavior of cancer. For example, in the case of melanoma, although many genes have been implicated in the genesis of this disease, only the INK4a, RAS and BRAF genes have been shown to be true etiologic lesions in a formal genetic sense (Chin et al., Genes Devel. 11:2822-34 (1997); Davies et al., Nature 417:949-54 (2002)). In the case of breast cancer,  
15 much progress has been made toward improvement in early diagnosis of early-stage disease, while options for treating advanced breast carcinomas remain limited. The genetic identities of key players in breast cancer initiation, progression and maintenance will serve in the design of novel, targeted therapies for this disease (Minna et. al, Cancer Cell:17-32 (2004) and van der Vijver et al.,

NEJM 347:1999-2009 (2002)). Moreover, advanced malignancy represents the phenotypic endpoint of many successive genetic lesions that affect many oncogene and tumor suppressor gene pathways. Lesions that lead to such a condition therefore may differ from those required to maintain it. Both types of lesions  
5 represent rational therapeutic targets.

### **Summary of the Invention**

It has been discovered that a gene designated GP202 functionally complements the HER2 oncogene in an inducible, spontaneous, *in vivo* cancer model (mouse). It has also been discovered that interfering RNAs that target  
10 GP202 expression inhibit the growth of certain tumor cells *in vitro*.

Based on these discoveries, the invention provides GP202 antagonists that inhibit GP202 gene expression or GP202 protein activity. Antagonists that inhibit GP202 gene expression include an interfering RNA that inhibits the expression of GP202, a GP202 antisense nucleic acid, and an anti-GP202 ribozyme. The sense  
15 strand sequences of four exemplary interfering RNAs of the invention include SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 and SEQ ID NO: 6. Antagonists that inhibit GP202 protein activity include blocking antibodies that bind to that portion of the GP202 protein that is exposed on the surface of a GP202-expressing cell. GP202 antagonists of the invention inhibit tumorigenesis, tumor development,  
20 tumor maintenance, tumor recurrence, tumor growth, or the growth of tumor cells *in vitro*.

The invention also provides methods of inducing apoptosis in a cell. The methods include contacting the cell with an effective amount of a GP202 antagonist.

25 The invention also provides methods of treating a hyperproliferative condition in a mammal, e.g., a human patient. The method includes administering to the mammal an effective amount of a GP202 antagonist. Cancer is an example of such a hyperproliferative condition. Other examples of hyperproliferative conditions are uncontrolled angiogenesis, psoriasis, arteriosclerosis, arthritis and  
30 diabetic retinopathy.

In some embodiments, the method of treating a hyperproliferative condition includes administering a second therapeutic agent. The second therapeutic agent can be, for example, an anti-angiogenic agent, anti-metastatic agent, agent that induces hypoxia, agent that induces apoptosis, or an agent that inhibits cell survival signals. Examples of cancer therapeutics include farnesyl transferase inhibitors, tamoxifen, herceptin, taxol, STI571, cisplatin, fluorocil, cytoxan, and ionizing radiation.

The invention also provides a host cell containing a recombinant DNA construct that includes a GP202-encoding sequence operably linked to an expression control sequence, and a genetic mutation that causes the host cell to have a greater likelihood of becoming a cancer cell than a cell not comprising the genetic mutation. Such a mutation can be, e.g., a mutation that deletes or inactivates a tumor suppressor gene, or a mutation that activates an oncogene. Examples of tumor suppressor genes include INK4a, P53, Rb, PTEN, LATS, APAF1, Caspase 8, APC, DPC4, KLF6, GSTP1, ELAC2/HPC2 and NKX3.1. Examples of oncogenes include K-RAS, H-RAS, N-RAS, EGFR, MDM2, TGF- $\beta$ , RhoC, AKT family members, myc,  $\beta$ -catenin, PGDF, C-MET, PI3K-CA, CDK4, cyclin B1, cyclin D1, estrogen receptor gene, progesterone receptor gene, HER2 (also known as neu or ErbB2), ErbB1, ErbB3, ErbB4, TGF $\alpha$ , ras-GAP, Shc, Nck, Src, Yes, Fyn, Wnt, and Bcl2.

The invention also provides a genetically modified non-human mammal, e.g., a mouse, at least some of whose cells contain a genome that includes: (a) a recombinant GP202-encoding nucleic acid operably linked to an expression control sequence, and (b) a genetic mutation that causes the mammal to have a greater susceptibility to cancer than a mammal whose cells do not contain the genetic mutation. In preferred embodiments, the genetic mutation involves a tumor suppressor gene and renders the tumor suppressor gene non-functional. The genetically modified nonhuman mammal can be a conventional transgenic mammal, all of whose cells contain the recombinant GP202-encoding nucleic acid operably linked to an expression control sequence, and the genetic mutation that causes the mammal to have increased susceptibility to cancer. Alternatively, the mammal is a chimeric mammal at least some of whose, but not all of whose,

somatic cells contain the recombinant GP202-encoding nucleic acid operably linked to an expression control sequence, and the genetic mutation that causes the mammal to have a increased susceptibility to cancer. In such a chimeric mammal, the percentage of somatic cells containing the recombinant GP202-encoding  
5 nucleic acid operably linked to an expression control sequence, and a genetic mutation that causes the mammal to have a greater susceptibility to cancer is between 5% and 95%. Preferably it is between 15% and 85%. In some embodiments of the invention, the GP202-encoding nucleic acid is operably linked to a tissue-specific expression system.

10 The invention also provides a genetically modified nonhuman mammal, wherein the genetic modification reduces or eliminates expression of one or both of the mammal's endogenous GP202 alleles. Such reduction or elimination of GP202 expression can be achieved, for example, when the genetic modification is addition of an RNAi expression construct targeting GP202 gene expression, or  
15 when the genetic modification is a knockout of one or both of the GP202 alleles. Such a genetic modification can reduce or eliminate GP202 expression in a tissue-specific manner. In some embodiments of the invention, the genetically modified mammal is chimeric with respect to the genetic modification.

The invention also provides a screening method for identifying a compound  
20 useful for treating a hyperproliferative condition such as cancer. The method includes: (a) identifying a biomarker whose level correlates with inhibition of GP202 activity; and (b) detecting a change in the level of the biomarker in the presence of a test compound relative to the level of the biomarker detected in the absence of the test compound.

25 The invention also provides a screening method for identifying a compound useful in treatment of a hyperproliferative condition such as cancer. The method includes: (a) providing an inhibitor of GP202 expression or activity; (b) identifying a negative control biomarker pattern formed by a plurality of biomarkers in a cancer cell wherein the cell is not contacted with the inhibitor of GP202 expression  
30 or activity; (c) identifying a positive control biomarker pattern formed by a plurality of biomarkers in the cancer cell wherein the cancer cell is contacted with the inhibitor of GP202 expression or activity; (d) identifying a test biomarker

pattern formed by a plurality of biomarkers in the cancer cell wherein the cancer cell is contacted with a candidate compound but not contacted with the inhibitor of GP202 expression or activity; and (e) comparing the negative control biomarker pattern, positive control biomarker pattern and test biomarker pattern, and  
5 detecting a greater similarity between the positive control biomarker pattern and the test biomarker pattern than between the negative control biomarker pattern and the test biomarker pattern.

The invention also provides methods of diagnosing an abnormal hyperproliferative condition, e.g., cancer, in a subject. These methods involve  
10 detecting the expression level of a GP202 gene or the activity level of a GP202 protein. An abnormally high level relative to control, e.g., at least about 50%, 100%, 150%, 200%, 250%, or 300% higher, indicates an abnormal hyperproliferative condition.

Unless otherwise defined, all technical and scientific terms used herein  
15 have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains. In case of conflict, the present specification, including definitions, will control. All publications, patents and other references mentioned herein are incorporated by reference in their entirety.

Throughout this specification and claims, the word "comprise," or  
20 variations such as "comprises" or "comprising" is intended to include the stated integer or group of integers, but not to exclude any other integer or group of integers.

Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the invention, the preferred methods  
25 and materials are described below. The materials, methods and examples are illustrative only, and are not intended to be limiting. Other features and advantages of the invention will be apparent from the detailed description and from the claims.

### **Detailed Description of the Invention**

Up-regulation of GP202 contributes to tumorigenesis and tumor  
30 maintenance. GP202 was identified as a cancer therapeutic target by using the

Mammalian Second Site Suppression ("MaSS") screening system described below and in WO 02/079419.

The GP202 protein, i.e., KIRREL, also known as NEPH1, FLJ10845, LOC348416 or 1\_154843723 is a member of the nephrin-like protein family and contains an immunoglobulin domain. It maps on chromosome 1 at 1q21-q25. An exemplary human GP202 protein contains 757 amino acid residues and has the following amino acid sequence (SEQ ID NO:1):

```

10  MLSLLVWILT LSDTFSQGTQ TRFSQEPADQ TVVAGQRAVL PCVLLNYSGL
    VQWTKDGLAL GMGQGLKAWP RYRVVGSADA GQYNLEITDA ELSDDASYEC
    QATEAALRSR RAKLTVLIPP EDTRIDGGPV ILLQAGTPHN LTCRAFNAKP
    AATIIWFRDG TQQEGAVAST ELLKDGKRET TVSQLLINPT DLDIGRVFTC
    RSMNEAIPSG KETSIELDVH HPPTVTLISIE PQTVQEGERV VFTCQATANP
    EILGYRWAKG GFLIEDAHES RYETNVDSYF FTEPVSCVH NKVGSTNVST
    LVNVHFAPRI VVDPKPTTDD IGSDVTLTCV WVGNPPLTLT WTKKDSNMVL
15  SNSNQLLLKS VTQADAGTYT CRAIVPRIGV AEREVPLYVN GPPIISSEAV
    QYAVRGDGGK VECFIGSTPP PDRIAWAWKE NFLEVGTLEK YTVERTNSGS
    GVLSTLTINN VMEADFQTHY NCTAWNSFGP GTAI IQLEER EVLPVGI IAG
    ATIGASILLI FFFIALVFFL YRRRKGSRKD VTLRKLDIKV ETVNREPLTM
    HSDREDDTAS VSTATRVMKA IYSSFKDDVD LKQDLRCDTI DTREEYEMKD
20  PTNGYYNVRA HEDRPSSRAV LYADYRAPGP ARFDGRPSSR LSHSSGYAQL
    NTYSRGPASD YGPEPTPPGP AAPAGTDTTT QLSYENYEKF NSHPFPGAAG
    YPTYRLGYPQ APPSGLERTP YEAYDPIGKY ATATRFSTYS QHSDYGQRFQ
    QRMQTHV

```

25 This amino acid sequence has been assigned the following GenBank<sup>TM</sup> numbers:  
NP\_060710.1 and GI:54036152.

The nucleotide sequence of one exemplary cDNA encoding the above polypeptide is (SEQ ID NO:2):

```

30  ggcgccgtcg gagggagagt agacgtcagg cggagggagg gaggcaggca
    ggcagggagg cagggaggga gagagggaag gaagagaagg ggagagaaag
    ggggagcaaa gaccgaaaga gaccagaga gatacagctt gagaggagaa
    ataattaaag ggggagggcg accaggagga gaaaaagggtc tgggaggaaa
    gagagataaa gagaatcgga gggcagggaa atgagtttgt gcgcgtgagt
    gtgcgcgcgc gcgtgtgtgt gtgcaggggc gggcgcgggc gggcgggggg
35  cgggctcccg gnetcccggg ctcccctcct cccagccccg ctccctctggc
    tcgctcgcca actccggggc ccagccgcgg gcgggcgcac ggcgggaggga
    cagcatgctg agcctcctcg tctggatcct cactctctcc gatactttct
    cccaaggtaa ggaccagac ccgcttcagc caggagccag ctgaccagac
    ggtggtggct ggacagcggg ccgtgctccc ctgtgtgctg ctcaactact
40  ctggaattgt gcaatggacc aaggacgggc tggccctggg cATGGGCCAG
    GGCCCTCAAAG CCTGGCCACG GTACCGGGTT GTGGGCTCCG CAGACNCTGG
    GCAGTACAAC CTGAGATCA CAGATGCTGA GCTCTCTGAC GACGCCTCTT

```

|    |             |             |             |            |             |
|----|-------------|-------------|-------------|------------|-------------|
|    | ACGAGTGCCA  | GGCCACGGAG  | GCCGCCCTGC  | GCTCTCGGCG | GGCCAAACTC  |
|    | ACCGTGCTCA  | ACCCCTCCTAC | AGTGACCCTG  | TCCATTGAGC | CACAGACGGT  |
|    | GCAGGAGGGT  | GAGCGTGTTG  | TCTTTACCTG  | CCAGGCCACA | GCCAACCCCG  |
|    | AGATCTTGGG  | CTACAGGTGG  | GCCAAAGGGG  | GTTTCTTGAT | TGAAGACGCC  |
| 5  | CACGAGAGTC  | GCTATGAGAC  | AAATGTGGAT  | TATTCCTTTT | TCACGGAGCC  |
|    | TGTGTCTTGT  | GAGGTTTACA  | ACAAAGTGGG  | AAGCACCAAT | GTCAGCACTT  |
|    | TAGTAAATGT  | CCACTTTGCT  | CCCCGGATTG  | TAGTTGACCC | CAAACCCACA  |
|    | ACCACAGACA  | TTGGCTCTGA  | TGTGACCCTT  | ACCTGTGTCT | GGGTTGGGAA  |
|    | TCCCCCCTC   | ACTCTCACCT  | GGACCAAAAA  | GGACTCAAAT | ATGGTCCTGA  |
| 10 | GTAACAGCAA  | CCAGCTGCTG  | CTGAAGTCGG  | TGACTCAGGC | AGACGCTGGC  |
|    | ACCTACACCT  | GCCGGGCCAT  | CGTGCCCTCGA | ATCGGAGTGG | CTGAGCGGGA  |
|    | GGTGCCGCTC  | TATGTGAANG  | GGCCCCCAT   | CATCTCCAGT | GAGGCAGTGC  |
|    | AGTATGCTGT  | GAGGGGTGAC  | GGTGGCAAGG  | TGGAGTGTTT | CATTGGGAGC  |
|    | ACACCACCCC  | CAGACCGCAT  | AGCATGGGCC  | TGGAAGGAGA | ACTTCTTGGA  |
| 15 | GGTGGGGACC  | CTGGAACGCT  | ATACAGTGGA  | GAGGACCAAC | TCAGGCAGTG  |
|    | GGGTGCTATC  | CACGCTCACC  | ATCAACAATG  | TCATGGAGGC | CGACTTTCAG  |
|    | ACTCACTACA  | ACTGCACCGC  | CTGGAACAGC  | TTCGGGCCAG | GCACAGCCAT  |
|    | CATCCAGCTG  | GAAGAGCGAG  | AGGTGTTACC  | TGTGGGCATC | ATAGCTGGGG  |
|    | CCACCATCGG  | CGCGAGCATC  | CTGCTCATCT  | TCTTCTTCAT | CGCCTTG GTA |
| 20 | TTCTTCTCT   | ACCGGCGCCG  | CAAAGGCAGT  | CGCAAAGACG | TGACCCTGAG  |
|    | GAAGCTGGAT  | ATCAAGGTGG  | AGACAGTGAA  | CCGAGAGCCA | CTTACGATGC  |
|    | ATTCTGACCG  | GGAGGATGAC  | ACCGCCAGCG  | TCTCCACAGC | AACCCGGGTC  |
|    | ATGAAGGCCA  | TCTACTCGTC  | GTTTAAGGAT  | GATGTGGATC | TGAAGCAGGA  |
|    | CCTGCGCTGC  | GACACCATCG  | ACACCCGGGA  | GGAGTATGAG | ATGAAGGACC  |
| 25 | CCACCAATGG  | CTACTACAAC  | GTGCGTGCCC  | ATGAAGACCG | CCCGTCTTCC  |
|    | AGGGCAGTGC  | TCTATGCTGA  | CTACCGTGCC  | CCTGGCCCTG | CCCGCTTCGA  |
|    | CGGCCGCCCC  | TCATCCCGTC  | TCTCCCACTC  | CAGCGGCTAT | GCCCAGCTCA  |
|    | ACACCTATAG  | CCGGGGCCCT  | GCCTCTGACT  | ATGGCCCTGA | GCCCACACCC  |
|    | CCTGGCCCTG  | CTGCCCCAGC  | TGGCACTGAC  | ACAACCAGCC | AGCTGTCCTA  |
| 30 | CGAGAACTAT  | GAGAAGTTCA  | ACTCCCATCC  | CTTCCCTGGG | GCAGCTGGGT  |
|    | ACCCACCTA   | CCGACTGGGC  | TACCCCCAGG  | CCCCACCCTC | TGGCCTGGAG  |
|    | CGGACCCCAT  | ATGAGGCGTA  | TGACCCCAT   | GGCAAGTACG | CCACAGCCAC  |
|    | TCGATTCTCC  | TACACCTCCC  | AGCACTCGGA  | CTACGGCCAG | CGATTCCAGC  |
|    | AGCGCATGCA  | GACTCACGTG  | TAGggggccag | agcctggctg | gggcatctct  |
| 35 | gcgggggcaga | ggagaaggct  | ttcacagctg  | ttccctgata | ttcagggaca  |
|    | ttgctcattg  | ctcccttctc  | ggaccagcct  | tcttcctccc | accatggcag  |
|    | gtggggagca  | ggtctcccag  | aaacaccccg  | ttccgaggat | ggtgctctgt  |
|    | gcatgcccc   | gcctcctggg  | cctgcccctc  | cctcttcttc | gggaggatgt  |
|    | gtctcttctg  | acctgcactc  | ttgcctgacc  | ctagaatggg | gacagggaaa  |
| 40 | gtgaagggtta | gggaaagcag  | aggggggcac  | tttttagcat | tccctttcta  |
|    | tcccacccct  | ctgatctccc  | ataagtggaa  | atgggggtac | ccagggatgg  |
|    | gcaggccttg  | gcctagggac  | atgaagtatg  | ggagtgggtg | gctgtggcac  |
|    | agacagggtg  | aaaacgggat  | agcctggcca  | gtccctctgt | tgtctgcatt  |
|    | cgtgccctgg  | gtgcctctct  | ccttcctcag  | ggtactgcag | aagggagcga  |
| 45 | acagggtact  | gttcgctctt  | gtctacagaa  | cagccctggc | actgcattca  |
|    | aatccagtct  | tcattcagct  | gggatcaaaa  | tgccagtcac | cttggctacc  |
|    | cactgtggac  | agctgtctgt  | cagcatgcag  | agggatccag | gaatcccccc  |
|    | ggcagcacgg  | cccgccttcc  | ttctcctcca  | tgctgggcca | gccagataag  |
|    | tcagggtcct  | ggtggagaaa  | gaaaggctag  | gaccatgtcc | tcattgaccc  |

agatactgct gtgtgctgca cagcagtga ccaacactag agggagccac  
 acaagcctcc tctccccagt ctgccccact tcctggcttt aactcttgng  
 ctggtttggg gagggtgag gtaggggtgg ggggtgctgta ggctcttttt  
 caaaaaaaaa c

5

The above nucleotide sequence has been assigned GenBank<sup>™</sup> numbers  
 NM\_018240.3 and GI:34916054. The open reading frame of the above sequence  
 is nucleotides 592-2373. Human GP202 has at least four alternatively spliced  
 transcript variants. Other human GP202 sequences include GenBank<sup>™</sup> numbers  
 10 GI:1457252, GI:31872094, GI:13141188, GI: 19696844, GI: 12402663.

GP202 is expressed in normal podocytes in the kidney, in normal ovary and  
 in prostatic epithelium. GP202 expression is elevated in chondrosarcoma,  
 glioblastomas, including glioblastomas expressing mutant activated EGFR,  
 astrocytomas and medulloblastomas, pancreatic adenocarcinoma, breast  
 15 carcinomas and colon adenocarcinoma, and melanomas.

The GP202 gene is likely to be involved in development (including  
 maintenance, progression, angiogenesis, and/or metastasis) of cancers, e.g., cancers  
 found in skin (e.g., melanoma), lung, prostate, breast, colorectal, liver, pancreatic,  
 brain, testicular, ovarian, uterine, cervical, kidney, thyroid, bladder, esophageal,  
 20 and hematological tissues.

The cancer-related functions of this gene can be confirmed by, e.g., (1) its  
 overexpression in more than about 10% of human cancers; (2) the inhibition of  
 human cancer cell proliferation and growth in soft agar by RNA interference  
 ("RNAi") of this gene; (3) the ability of this gene to enhance the growth of mouse  
 25 tumor cell lines in soft agar; and/or (4) the prevention by inhibiting this gene of  
 maintenance or formation of tumors arising *de novo* in a mouse, or tumors derived  
 from human cancer cell lines.

## **I. GP202-RELATED NUCLEIC ACIDS**

The nucleic acid sequences specifically provided herein are sequences of  
 30 deoxyribonucleotides. However, the given sequences are to be interpreted as  
 would be appropriate to the polynucleotide composition. For example, if the  
 isolated nucleic acid is RNA, the given sequence intends ribonucleotides, with

uridine substituted for thymidine. In some embodiments, differences from naturally occurring nucleic acids, e.g., non-native bases, altered internucleoside linkages, and post-synthesis modification, can be present throughout the length of the GP202 nucleic acid or can be usefully localized to discrete portions thereof.

- 5 For example, a chimeric nucleic acid can be synthesized with discrete DNA and RNA domains and demonstrated utility for targeted gene repair. See, e.g., U.S. Pat. Nos. 5,760,012 and 5,731,181.

Polymorphisms such as single nucleotide polymorphisms (SNPs) occur frequently in eukaryotic genomes. Additionally, small deletions and insertions, rather than SNPs, are not uncommon in the general population, and often do not alter the function of the protein. Accordingly, this invention provides not only isolated nucleic acids identical in sequence to those described with particularity herein, but also isolated nucleic acids that are allelic variants of those particularly described nucleic acid sequences. In some embodiments, such sequence variations result from human intervention, e.g., by random or directed mutagenesis.

#### **A. Nucleic Acids Encoding GP202 Protein or Portions Thereof**

The invention provides isolated nucleic acid molecules that encode the entirety or part (e.g., at least five, seven, or nine contiguous amino acid residues) of the GP202 protein, including allelic variants of this protein. As is well known, the genetic code is degenerate and codon choice for optimal expression varies from species to species. Thus, the coding sequences of this invention include degenerate variants of the sequences described herein with particularity. Thus, the isolated polynucleotide comprises a nucleotide sequence encoding SEQ ID NO: 1.

These nucleic acids can be used, for example, to express the GP202 protein or specific portions of the protein, either alone or as elements of a fusion protein, e.g., to express epitopic or immunogenic fragments of the GP202 protein. For example, such nucleic acids are used to produce non-human mammals of the invention. These nucleic acids also can be used as probes to hybridize to GP202 nucleic acids and related nucleic acid sequences.

30 This invention also relates to nucleic acids comprising sequences coding for polypeptides containing conservative amino acid substitutions or moderately

conservative amino acid substitutions from those polypeptides described with particularity herein. These amino acid substitutions can be due to, e.g., allelic variations, naturally occurring mutations, or man-made mutations.

### **B. Cross-Hybridizing Nucleic Acids**

5 This invention also relates to isolated polynucleotides that hybridize to one or more of the above-described GP202 nucleic acids. These cross-hybridizing nucleic acids can be used, e.g., as hybridization probes, primers, and/or for expression of proteins that are related to GP202 as isoforms and homologs, e.g., paralogs, and orthologs. In some such embodiments, the invention relates to an  
10 isolated nucleic acid comprising a sequence that hybridizes under high stringency conditions to a probe comprising a fragment of SEQ ID NO: 2 having at least 15, 16, 18, 20, 24, or 25 nucleotides. As used herein, "high stringency conditions" are defined for solution phase hybridization as aqueous hybridization (i.e., free of formamide) in 6X SSC (where 20X SSC contains 3.0 M NaCl and 0.3 M sodium  
15 citrate), 1% SDS at 65°C for 8-12 hours, followed by two washes in 0.2X SSC, 0.1% SDS at 65°C for 20 minutes. It will be appreciated by the skilled worker that hybridization at 65°C will occur at different rates depending on a number of factors including the length and percent identity of the sequences which are hybridizing.

The hybridizing portion of a reference nucleic acid (i.e., target nucleic acid)  
20 is typically at least 15 nucleotides in length, and often at least 17, 20, 25, 30, 35, 40 or 50 nucleotides in length. Cross-hybridizing nucleic acids that hybridize to a larger portion of the reference nucleic acid – for example, to a portion of at least 50, 100, 150, 200, 250, 300, 350, 400, 450, 500 or more nucleotides, up to and including the entire length of the reference nucleic acid, are also useful.

### **C. Nucleic Acid Fragments**

25 Fragments of the above-described nucleic acids also relate to this invention. They can be used as region-specific probes, as amplification primers, regulatory sequences to direct expression of a gene, and/or to direct expression of a GP202 polypeptide fragment, e.g., immunogenic fragment.

The nucleic acid probes may comprise a detectable label, e.g., a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic kit for identifying cells or tissues that (i) incorrectly express a GP202 protein, e.g., aberrant splicing, or abnormal mRNA levels, or (ii) harbor a mutation in the GP202 gene, such as a deletion, an insertion, or a point mutation. Such diagnostic kits preferably include labeled reagents and instructional inserts for their use.

The nucleic acid primers can be used in PCR, primer extension and the like. They can be, e.g., at least 6 nucleotides (e.g., at least 7, 8, 9, or 10) in length. The primers can hybridize to an exon sequence of a GP202 gene, e.g., for amplification of a GP202 mRNA or cDNA. Alternatively, the primers can hybridize to an intron sequence or an upstream or downstream regulatory sequence of a GP202 gene, to utilize non-transcribed, e.g., regulatory portions of the genomic structure of a GP202 gene. The nucleic acid primers also can be used, e.g., to prime single base extension (SBE) for SNP detection (see, e.g., U.S. Pat. No. 6,004,744). Isothermal amplification approaches, such as rolling circle amplification, are also now well-described. See, e.g., Schweitzer et al., *Curr. Opin. Biotechnol.* 12(1):21-7 (2001); U.S. Patent Nos. 5,854,033 and 5,714,320 and PCT international patent publications WO 97/19193 and WO 00/15779. Rolling circle amplification can be combined with other techniques to facilitate SNP detection. See, e.g., Lizardi et al., *Nature Genet.* 19:225-32 (1998).

Nucleic acid fragments that encode 5 or more contiguous amino acids, e.g., fragments of 15, 18, 21, 24, or 27 nucleotides or more, are useful in directing the synthesis of peptides that have utility in mapping the epitopes of the GP202 protein. See, e.g., Geysen et al., *Proc. Natl. Acad. Sci. USA* 81:3998-4002 (1984); and U.S. Pat. Nos. 4,708,871 and 5,595,915. Such nucleic acid fragments are also useful in directing the synthesis of peptides that have utility as immunogens. See, e.g., Lerner, "Tapping the immunological repertoire to produce antibodies of predetermined specificity," *Nature* 299:592-596 (1982); Shinnick et al., *Annu. Rev. Microbiol.* 37:425-46 (1983); Sutcliffe et al., *Science* 219:660-6 (1983). Larger fragments containing at least 25, 30, 35, 40, 45, 50, 100, 150, 200, 250,

300, 350, 400, 450, 500 or more nucleotides are also useful and sometimes preferred.

#### **D. Single Exon Probes**

The invention further relates to single exon probes having portions of no  
5 more than one exon of the GP202 gene. Such single exon probes have particular utility in identifying and characterizing splice variants. In particular, these probes are useful for identifying and discriminating the expression of distinct isoforms of GP202.

#### **E. Antisense Reagents**

##### **1. Antisense Nucleic Acids**

Some embodiments of the invention relate to isolated nucleic acids that are  
antisense polynucleotides that specifically hybridize to GP202 sense  
polynucleotides. The antisense nucleic acid molecule can be complementary to the  
entire coding or non-coding region of GP202, but more often is antisense to only a  
15 portion of the coding or non-coding region of GP202 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of GP202 mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length.

The antisense nucleic acids of this invention, for example, may form a  
20 stable duplex with its target sequence, or, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. In other embodiments, the antisense nucleic acid molecule of the invention is an  $\alpha$ -anomeric nucleic acid molecule. An  $\alpha$ -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary  
25 RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gaultier et al. Nucl. Acids Res 15: 6625-6641 (1987)).

An antisense target sequence is a nucleotide sequence specific to GP202, and can be designed through use of a publicly available sequence database, and/or through use of commercially available sequence comparison programs. Antisense  
30 nucleic acids of the invention can then be constructed using procedures known in

the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be produced biologically using an expression vector into which a nucleic acid has been inserted in an antisense orientation, i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest.

Alternatively, the antisense nucleic acid can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecule or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids. For example, phosphorothioate derivatives and acridine substituted nucleotides can be used. Phosphorothioate and methylphosphonate antisense oligonucleotides are useful in practicing the invention. The antisense nucleic acid molecule can also comprise a 2'-O-methylribonucleotide (Inoue et al., Nucl. Acids Res. 15: 6131-6148 (1987)) or a chimeric RNA-DNA analogue (Inoue et al., FEBS Lett. 215: 327-330 (1987)). The antisense oligonucleotides may be further modified by adding poly-L-lysine, transferrin, polylysine, or cholesterol moieties at their 5' end.

The antisense molecules will be tested for undesired non-specific effects such as the induction of ds-RNA stress response genes in the interferon pathway. Only those molecules that do not induce a significant non-specific response will be subsequently used.

Antisense molecules can be administered to a mammal or generated *in situ* via an expression vector, such that they bind to cellular RNA and/or genomic DNA encoding a GP202 protein, thereby inhibiting GP202 expression. Suppression of GP202 expression at either the transcriptional or translational level is useful to treat certain cancer conditions in patients or to generate cellular or animal models for cancer characterized by aberrant GP202 expression. An antisense molecule can be administered by direct injection at a tissue site of a subject. Alternatively, an antisense molecule can be designed to target selected cells, e.g., cancer cells overexpressing GP202, and then administered systemically.

## 2. Ribozymes and Catalytic Nucleic Acids

In some embodiments, an antisense nucleic acid of the invention is part of a GP202-specific ribozyme (or, as modified, a "nucleozyme"). Ribozymes are

catalytic RNA molecules with ribonuclease activity capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes such as hammerhead, hairpin, and Group I intron ribozymes can cleave GP202 mRNA transcripts catalytically, thereby inhibiting translation of GP202 mRNA. A ribozyme having specificity for a GP202-  
5 encoding nucleic acid can be designed based upon the nucleotide sequence of a GP202 polynucleotide disclosed herein (SEQ ID NO: 2). See, e.g., U.S. Patent Nos. 5,116,742; 5,334,711; 5,652,094; and 6,204,027. For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide  
10 sequence of the active site is complementary to the nucleotide sequence to be cleaved in a GP202-encoding mRNA. See, e.g., Cech et al. U.S. Pat. Nos. 4,987,071 and 5,116,742. Alternatively, GP202 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., Science 261:1411-1418 (1993).

15 In some embodiments, the ribozymes and other antisense reagents of this invention include appended groups such as peptides. This is useful for targeting host cell receptors, facilitating transport across the cell membrane (Letsinger et al., Proc. Natl. Acad. Sci. USA 86:6553-6556 (1989); Lemaitre et al., Proc. Natl. Acad. Sci. USA 84:648-652 (1987); WO 88/09810), or facilitating transport across  
20 the blood-brain barrier (WO 89/10134).

Expression of the GP202 gene can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the GP202, e.g., the GP202 promoter and/or enhancers, to form triple helical structures that prevent transcription of the GP202 gene in target cells. See generally, Helene, Anticancer  
25 Drug Des. 6: 569-84 (1991); Helene et al., Ann. N.Y. Acad. Sci. 660:27-36 (1992); and Maher, Bioassays 14: 807-15 (1992).

### 3. Peptide Nucleic Acids (PNA)

Some preferred oligonucleotide mimetics, especially those useful for *in vivo* administration, are peptide nucleic acids (PNA). See, e.g., Hyrup et al.,  
30 Bioorg. Med. Chem. Lett. 4:5-23 (1996). In PNA compounds, the phosphodiester backbone of the nucleic acid is replaced with an amide-containing backbone, in

particular by repeating N-(2-aminoethyl) glycine units linked by amide bonds. Nucleobases are bound directly or indirectly to aza-nitrogen atoms of the amide portion of the backbone, typically by methylene carbonyl linkages. The synthesis of PNA oligomers can be performed using conventional solid phase peptide  
5 synthesis as described in Hyrup et al., *supra*; and Perry-O'Keefe et al., Proc. Natl. Acad. Sci. USA 93:14670-675 (1996).

GP202-based PNAs can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription arrest,  
10 inducing translation arrest, or inhibiting replication. GP202-based PNAs also can be used in the analysis of single base pair mutations in a gene. This can be done by PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases, or as probes or primers for DNA sequence and hybridization (Hyrup et al., *supra*; Perry-O'Keefe, *supra*).

15 In other embodiments, PNAs of GP202 are modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of GP202 can be generated that may combine the advantageous properties of PNA  
20 and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup et al., *supra*). The  
25 synthesis of PNA-DNA chimeras can be performed as described in Hyrup, *supra* and Finn et al., Nucl. Acids Res. 24:3357-63 (1996).

#### 4. RNA Interference

The invention provides RNA interference (RNAi) molecules for use in silencing the expression of the GP202 gene. RNAi is a sequence-specific  
30 posttranscriptional gene silencing mechanism triggered by double-stranded RNA (dsRNA). RNAi results in degradation of mRNAs homologous in sequence to the

dsRNA. The mediators of the degradation are 21- to 23-nucleotide small interfering RNAs (siRNAs) generated by ribonuclease III cleavage from the longer dsRNAs. Molecules of siRNA typically have 2- to 3-nucleotide 3' overhanging ends resembling the RNase III processing products of long dsRNAs that normally initiate RNAi. When introduced into a cell, they assemble with an endonuclease complex (RNA-induced silencing complex), which then guides cleavage of the targeted mRNA. This results in a phenotype with suppressed expression of the protein encoded by the targeted mRNA. The small size of siRNAs avoids activation of the dsRNA-inducible interferon system in mammalian cells. This helps avoid the nonspecific phenotypes normally produced by dsRNA larger than 30 base pairs in somatic cells. For review, see, e.g., Elbashir et al., *Methods* 26:199-213 (2002); McManus and Sharp, *Nature Reviews* 3:737-747 (2002); Hannon, *Nature* 418:244-251 (2002); Tuschl, *Nature Biotechnology* 20:446-448 (2002); and Tuschl, U.S. Application US2002/0086356 A1.

Small interfering RNA oligonucleotides can be designed using commercially available software programs, e.g., the OligoEngine<sup>™</sup> siRNA design tool available at <http://www.oligoengine.com>, or purchased from commercial vendors, e.g., Dharmacon RNA Technologies (LaFayette, CO). Preferred siRNAs of this invention range about 19-29 basepairs in length for the double-stranded portion. In some embodiments, the siRNAs are hairpin RNAs having an about 19-29 bp stem and an about 4-34 nucleotide loop. In some embodiments, siRNAs are highly specific for a GP202 target region and may comprise any 19-29 bp fragment of a GP202 mRNA that has at least 1 (e.g., at least 2 or 3) bp mismatch with a non-GP202-related sequence. In some embodiments, the GP202 siRNAs do not bind to RNAs having more than 3 mismatches. The sense strand sequences of four exemplary siRNAs for GP202 are:

5'- CAACGUGCGUGCCCAUGAA -3' (SEQ ID NO: 3);

5'- CAACGUGCGUGCCCAUGAA -3' (SEQ ID NO: 4);

5'- GCCAUCUACUCGUCGUUUA -3' (SEQ ID NO: 5); and

5'- GAGAGCCACUUACGAUGCA -3' (SEQ ID NO: 6).

These siRNAs can be used individually or pooled and used in combination. Small interfering RNAs that target an RNA region having 10 or more nucleotide overlap with an aforementioned exemplary target region are also useful.

Intracellular transcription of siRNAs can be achieved by cloning the siRNA  
5 templates into RNA polymerase III (Pol III) transcription units, which normally encode the small nuclear RNA U6 or the human RNase P RNA H1. Two approaches can be used for expressing siRNA: (1) sense and antisense strands constituting the siRNA duplex are transcribed by individual promoters; or (2) siRNAs are expressed as fold-back stem-loop structures that give rise to siRNAs  
10 after intracellular processing. Inducible promoters can also be used to drive the expression of the siRNA.

To target specific regions of a GP202 mRNA, the following short hairpin RNAs (shRNAs) can be expressed by inserting the following sequences (minus the two terminal UU) after the U6 small RNA promoter. The first base of the  
15 following sequences is the transcriptional start site of the U6 promoter. When additional sequence complementary to the 3' end of the U6 promoter is added to the reverse complement of the sequences below, the sequence can be used with a primer specific to the U6 small RNA promoter to form double-stranded DNA in a polymerase chain reaction, using a vector containing this U6 promoter as a  
20 template. The PCR product can then be ligated into a vector. Alternatively, the sequences below can be added downstream to the U6 promoter by annealing two oligonucleotides, one containing the sense sequence (below) and the other containing the antisense sequence and containing appropriate overhang nucleotides to allow ligation into U6 promoter vector digested by a restriction enzyme. The  
25 short hairpin sequence is followed by 4 or more Ts to allow termination of transcription. Expression of the insert leads to expression of a short hairpin RNA. The hairpin structure displays inhibitory effects on GP202 expression.

The sequence of the sense-loop-antisense sequence is shown for each shRNA construct: (1) The following shRNA targets nucleotides 1803-1827 of  
30 GP202:

GCCAUCUUCU CGUCGUUUA GGCUCCUGU CACCUAAAAC GACGAGUAGA  
UGGC ( SEQ ID NO: 7).

(2) The following shRNA targets nucleotides 979-1003 of GP202:

GCACCAAAGU CAGCACUUA GGCUCCUGU CACCUAAAGU GCUGACAUUG  
5 GUGC (SEQ ID NO: 8).

(3) The following shRNA targets nucleotides 991-1015 of GP202:

GCACUUUCGU AAAUGUCCUC UGCUCCUGU CACAGUGGAC AUUUACUAAA  
GUGC (SEQ ID NO: 9).

(4) The following shRNA targets nucleotides 948-972 of GP202:

10 GUGUCUUCUG AGGUUCACUA CGCUCCUGU CACGUUGUGA ACCUCACAAG  
ACAC (SEQ ID NO: 10).

An siRNA oligonucleotide or its coding sequence can be delivered into a  
target cell via a variety of methods, including but not limited to, liposome fusion  
(transposomes), routine nucleic acid transfection methods such as electroporation,  
15 calcium phosphate precipitation, or microinjection, and infection by viral vectors.

## **F. Exemplary Uses of GP202 Nucleic Acids**

### **1. Characterization of Genetic Mutations**

The above-described isolated nucleic acids can be used as hybridization  
probes to characterize GP202 nucleic acids in both genomic and transcript-derived  
20 nucleic acid samples. For example, the probes can be used to detect gross  
alterations in the GP202 genomic locus, such as deletions, insertions,  
translocations, and duplications of the GP202 genomic locus. Methods of  
detection include fluorescence *in situ* hybridization (FISH) to chromosome  
spreads, comparative genomic hybridization (CGH), array CGH (e.g., on  
25 microarrays containing GP202-coding sequences or BAC comprising GP202-  
coding sequences), and spectral karyotyping (SKY). See, e.g., Andreeff et al.  
(eds.), *Introduction to Fluorescence In Situ Hybridization: Principles and Clinical  
Applications*, John Wiley & Sons (1999) (ISBN: 0471013455). The probes also  
can be used to assess smaller genomic alterations using, e.g., Southern blot  
30 detection of restriction fragment length polymorphisms. The nucleic acid probes

can be also used to isolate genomic clones that include the nucleic acids of the present invention, which thereafter can be restriction mapped and sequenced to identify deletions, insertions, amplifications, translocations, and substitutions (e.g., SNPs) at the sequence level. The nucleic acid probes can also be used to isolate  
5 GP202 nucleic acids from cDNA libraries, permitting sequence level characterization of GP202 RNA messages, including identification of deletions, insertions, truncations (including deletions, insertions, and truncations of exons in alternatively spliced forms) and single nucleotide polymorphisms. Some of the nucleic acids also can be used as amplification primers for real time PCR to detect  
10 the above-described genomic alterations. Such genomic alterations of the GP202 gene often play a role in tumor genesis, maintenance and development, and thus can be used as markers for diagnosis and prognosis of GP202-mediated cancers.

## 2. Quantification of Expression Levels

The nucleic acid probes can be used to measure the representation of  
15 GP202 clones in a cDNA library, used as primers for quantitative real time PCR, or otherwise used to measure expression level of the GP202 gene. Measurement of GP202 expression has particular utility in diagnostic assays for cancer-related conditions associated with abnormal GP202 expression. Moreover, differences in the expression levels of the gene before and after a cancer event (e.g., cancer  
20 genesis, maintenance, regression, and metastasis) are useful in determining the effect of a candidate cancer drug, identifying cancer types, designing diagnostics and prognostics, and predicting likely outcome of a cancer therapy.

## 3. Genetic Alterations

The nucleic acids also can be used to introduce mutations (e.g., null  
25 mutations, dominant negative mutations, dominant acting mutations) into a GP202 locus of an animal via homologous recombination. Such animals (e.g., knock out mice) are useful in delineating the role of GP202 in tumor genesis and development and in facilitating cancer drug development.

Where the genomic region includes transcription regulatory elements,  
30 homologous recombination can be used to replace the endogenous regulatory elements with heterologous regulatory elements, i.e., elements not natively

associated with the gene in the same manner. This can alter the expression of GP202, both for production of GP202 protein, and for gene therapy. See, e.g., U.S. Pat. Nos. 5,981,214; 6,048,524 and 5,272,071.

5 Fragments of the above-described polynucleotides smaller than those typically used for homologous recombination can also be used for targeted gene correction or alteration, possibly by cellular mechanisms different from those engaged during homologous recombination. See, e.g., U.S. Pat. Nos. 5,945,339; 5,888,983; 5,871,984; 5,795,972; 5,780,296; 5,760,012; 5,756,325; 5,731,181; and  
10 Culver et al., Nature Biotechnol. 17:989-93 (1999); Gamper et al., Nucl. Acids Res. 28:4332-9 (2000).

## **II. VECTORS AND HOST CELLS**

### **A. General Consideration**

This invention relates to nucleic acid constructs containing one or more of the isolated nucleic acid molecules encoding all or part of GP202. The  
15 vectors can be used to propagate the new nucleic acid molecules in host cells, to shuttle the molecules between host cells derived from disparate organisms, to insert the molecules into host genomes, to express sense or antisense RNA transcripts or interfering RNAs, and/or to express GP202 polypeptides. Typically, the vectors are derived from virus, plasmid, prokaryotic or eukaryotic  
20 chromosomal elements, or some combination thereof, and may optionally include at least one origin of replication, at least one site for insertion of heterologous nucleic acid, and at least one selectable marker.

This invention relates to host cells, which can be either prokaryotic (bacteria) or eukaryotic (e.g., yeast, insect, plant and animal cells). A host cell  
25 strain may be chosen for its ability to process the expressed protein in the desired fashion. Such post-translational modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, hydroxylation, sulfation, lipidation, and acylation. Some embodiments of the invention may involve GP202 proteins with such post-translational modifications.

Exemplary prokaryotic host cells are *E. coli*, *Caulobacter crescentus*, *Streptomyces species*, and *Salmonella typhimurium* cells. Vectors useable in these cells include, without limitation, those related to pBR322 and the pUC plasmids.

Exemplary yeast host cells are *Saccharomyces cerevisiae*,  
5 *Schizosaccharomyces pombe*, *Pichia pastoris*, and *Pichia methanolica*. Any suitable vector can be used, e.g., YIp vectors, replicating episomal YEp vectors containing centromere sequences CEN and autonomously replicating sequences ARS.

Insect cells may be advantageous, e.g., for high efficiency protein  
10 expression. Exemplary insect host cells are those from *Spodoptera frugiperda*, e.g., Sf9 and Sf21 cell lines, and EXPRESSF<sup>TM</sup> cells (Protein Sciences Corp., Meriden, CT, USA), *Drosophila* S2 cells, and *Trichoplusia ni* HIGH FIVE<sup>®</sup> Cells (Invitrogen, Carlsbad, CA). Where the host cells are *Spodoptera frugiperda* cells, the vector replicative strategy is typically based upon the baculovirus life cycle.

Exemplary mammalian host cells are COS1 and COS7 cells, NS0 cells,  
15 Chinese hamster ovary (CHO) cells, NIH 3T3 cells, 293 cells, HEPG2 cells, HeLa cells, L cells, MDCK, HEK293, WI38, murine ES cell lines (e.g., from strains 129/SV, C57/BL6, DBA-1, 129/SVJ), K562, Jurkat cells, BW5147 and any other commercially available human cancer cell lines. Cells with elevated Her2  
20 expression, such as human cancer lines BT474 and SKOV3, can also be used. Other useful mammalian cell lines are well known and are available from the American Type Culture Collection (ATCC<sup>®</sup>) (Manassas, VA, USA) and the National Institute of General medical Sciences (NIGMS) Human Genetic Cell Repository at the Coriell Cell Repositories (Camden, NJ, USA).

Vectors intended for autonomous extrachromosomal replication in  
25 mammalian cells typically include a viral origin, such as the SV40 origin (for replication in cell lines expressing the large T-antigen, such as COS1 and COS7 cells), the papillomavirus origin, or the EBV origin for long term episomal replication (for use in, e.g., 293-EBNA cells, which constitutively express the EBV  
30 EBNA-1 gene product and adenovirus E1A). Vectors intended for integration, and thus replication as part of the mammalian chromosome optionally include an origin

of replication functional in mammalian cells, e.g., the SV40 origin. Useful vectors also include vectors based on viruses such as lentiviruses, adenovirus, adeno-associated virus, vaccinia virus, parvoviruses, herpesviruses, poxviruses, Semliki Forest viruses, and retroviruses.

- 5           Plant cells also can be used for expression, with the vector replicon typically derived from a plant virus (e.g., cauliflower mosaic virus; tobacco mosaic virus) and selectable markers chosen for suitability in plants.

          The invention relates to artificial chromosomes, e.g., bacterial artificial chromosomes (BACs), yeast artificial chromosomes (YACs), mammalian artificial  
10 chromosomes (MACs), and human artificial chromosomes (HACs), that contain the GP202 nucleic acid of interest.

- Vectors often include elements that permit *in vitro* transcription of RNA from the inserted heterologous nucleic acid. Such vectors typically include a phage promoter, such as that from T7, T3, or SP6, flanking the nucleic acid insert.  
15       Often two different such promoters flank the inserted nucleic acid, permitting separate *in vitro* production of both sense and antisense strands.

#### **B.       Transcription Regulators for Expression Vectors**

- Expression vectors often include a variety of other genetic elements operatively linked to the protein-encoding heterologous nucleic acid insert.  
20       Examples of other such genetic elements are promoters and enhancer elements. Other elements are those that facilitate RNA processing, e.g., transcription termination signals, splicing signals and polyadenylation signals. Other elements are those that facilitate translation, e.g., ribosomal consensus sequences. Examples of other transcription control sequences are operators and silencers. Use of such  
25       expression control elements, including those that confer constitutive or inducible expression, and developmental or tissue-regulated expression are known in the art.

          Constitutively active promoters include, without limitation, a CMV promoter, EF1 $\alpha$ , retroviral LTRs, and SV40 early region.

- Inducible promoters useful in this invention include, without limitation, a  
30       tetracycline-inducible promoter, a metallothionine promoter, the IPTG/lacI

promoter system, the ecdysone promoter system, and the “lox stop lox” system for irreversibly deleting inhibitory sequences for translation or transcription. In some embodiments, a GP202 gene is placed between lox sites,. Upon expression of the cre enzyme, the GP202 gene is deleted from the genome so that the GP202 activity  
5 is permanently eliminated.

Instead of inducible promoters, the activity of a GP202 protein also can be inducibly switched on or off by fusing the GP202 protein to, e.g., an estrogen receptor polypeptide sequence, where administration of estrogen or an estrogen analog (e.g., hydroxytamoxifen) will allow the correct folding of the GP202  
10 polypeptide into a functional protein.

Tissue-specific promoters that can be used in driving expression of GP202 in animal models include, without limitation: a tyrosinase promoter or a TRP2 promoter in the case of melanoma cells and melanocytes; an MMTV or WAP promoter in the case of breast cells and/or cancers; a Villin or FABP promoter in  
15 the case of intestinal cells and/or cancers; a RIP promoter in the case of pancreatic beta cells; a Keratin promoter in the case of keratinocytes; a Probasin promoter in the case of prostatic epithelium; a Nestin or GFAP promoter in the case of CNS cells and/or cancers; a Tyrosine Hydroxylase, S100 promoter or neurofilament promoter in the case of neurons; the pancreas-specific promoter described in  
20 Edlund et al., Science 230:912-916 (1985); a Clara cell secretory protein promoter in the case of lung cancer; and an Alpha myosin promoter in the case of cardiac cells.

Developmentally regulated promoters also can be used. They include, without limitation, the murine hox promoters (Kessel et al., Science 249:374-379  
25 (1990)) and the  $\alpha$ -fetoprotein promoter (Campes et al., Genes Dev. 3:537-546 (1989)).

### **C. Expression Vectors Encoding Peptide Tags**

Expression vectors can be designed to fuse the expressed polypeptide to small protein tags that facilitate purification and/or visualization. Many such tags  
30 are known and available. Expression vectors can also be designed to fuse proteins encoded by the heterologous nucleic acid insert to polypeptides larger than

purification and/or identification tags. Useful protein fusions include those that permit display of the encoded protein on the surface of a phage or cell, fusions to intrinsically fluorescent proteins, such as luciferase or those that have a green fluorescent protein (GFP)-like chromophore, and fusions for use in two hybrid  
5 selection systems.

For secretion of expressed proteins, a wide variety of vectors are available which include appropriate sequences that encode secretion signals, such as leader peptides. Vectors designed for phage display, yeast display, and mammalian display, for example, target recombinant proteins using an N-terminal cell surface  
10 targeting signal and a C-terminal transmembrane anchoring domain.

A wide variety of vectors now exist that fuse proteins encoded by heterologous nucleic acids to the chromophore of the substrate-independent, intrinsically fluorescent green fluorescent protein from *Aequorea victoria* ("GFP") and its many color-shifted and/or stabilized variants.

15 For long-term, high-yield recombinant production of the proteins, protein fusions, and protein fragments of the present invention, stable expression is preferred. Stable expression is readily achieved by integration into the host cell genome of vectors (preferably having selectable markers), followed by selection for integrants.

### 20 **III. GP202 PROTEINS, POLYPEPTIDES AND FRAGMENTS**

The present invention relates to GP202 proteins and various fragments suitable for use, e.g., as antigens, biomarkers for diseases, and in therapeutic compositions. The invention also relates to fusions of GP202 polypeptides to heterologous polypeptides or other moieties.

#### 25 **A. GP202 Polypeptides of Particular Sequences**

The invention relates to an isolated GP202 polypeptide (SEQ ID NO: 1), optionally containing one or more conservative amino acid substitutions. The invention also relates to fragments of the GP202 polypeptide, particularly fragments having at least 5, 6, 8, or 15 amino acids of SEQ ID NO: 1. Larger  
30 fragments of at least 20, 25, 30, 35, 50, 75, 100, 150 or more amino acids are also

useful, and at times preferred. The GP202 fragments of the invention may be continuous portions of the native GP202 protein. However, it will be appreciated that knowledge of the GP202 gene and protein sequences permits recombining of various domains that are not contiguous in the native GP202 protein.

5                    **B.      Fusion Proteins And Other Protein Conjugates**

This invention also relates to fusions of GP202 polypeptides to heterologous polypeptides. As used herein, "fusion" means that the GP202 polypeptide is linearly contiguous to the heterologous polypeptide in a peptide-bonded polymer of amino acids or amino acid analogues. As used herein,  
10      "heterologous polypeptide" means a polypeptide that does not naturally occur in contiguity with the GP202 fusion partner. The fusion polypeptide can consist entirely of a plurality of fragments of the GP202 protein in altered arrangement. In such a case, any of the GP202 fragments can be considered heterologous to the other GP202 fragments in the fusion protein.

15                    The heterologous polypeptide included within the fusion protein is at least 6 amino acids in length, often at least 8 amino acids in length, and preferably, at least 15, 20, and 25 amino acids in length. The heterologous sequences can target the GP202 polypeptide to a selected cell by binding to a cell surface receptor, prolong the serum life of the GP202 polypeptide (e.g., an IgG Fc region), make the  
20      GP202 polypeptide detectable (e.g., a luciferase or a green fluorescent protein), facilitate purification (e.g., His tag, FLAG, etc.), facilitate secretion of recombinantly expressed proteins (e.g., into the periplasmic space or extracellular milieu for prokaryotic hosts, into the culture medium for eukaryotic cells, through incorporation of secretion signals and/or leader sequences). Other useful protein  
25      fusions include fusions that permit use of the protein of the present invention as bait in a yeast two-hybrid system, fusions that display the encoded protein on the surface of a phage or cell, and fusions to intrinsically detectable proteins, such as fluorescent or light-emitting proteins.

30                    The proteins and protein fragments also can be fused to protein toxins, such as *Pseudomonas* exotoxin A, diphtheria toxin, shiga toxin A, anthrax toxin lethal

factor, ricin, or other toxic moieties in order to effect specific ablation of cells that bind or take up the proteins.

### C. Other Modifications of the Polypeptides

The polypeptides can be composed of natural amino acids linked by native  
5 peptide bonds, or can contain any or all of nonnatural amino acid analogues,  
nonnative bonds, and post-synthetic (post-translational) modifications, either  
throughout the length of the polypeptide or localized to one or more portions  
thereof. However, the range of such nonnatural analogues, nonnative inter-residue  
bonds, or post-synthesis modifications will be limited to those that do not interfere  
10 with the biological function of the polypeptide.

Techniques for incorporating non-natural amino acids during solid phase  
chemical synthesis or by recombinant methods are well established in the art. For  
instance, D-enantiomers of natural amino acids can readily be incorporated during  
chemical peptide synthesis: peptides assembled from D-amino acids are more  
15 resistant to proteolytic attack; incorporation of D-enantiomers can also be used to  
confer specific three dimensional conformations on the peptide. Other amino acid  
analogues commonly added during chemical synthesis include ornithine,  
norleucine, phosphorylated amino acids (typically phosphoserine,  
phosphothreonine, phosphotyrosine), L-malonyltyrosine, a non-hydrolyzable  
20 analog of phosphotyrosine (Kole et al., Biochem. Biophys. Res. Com. 209:817-821  
(1995)), and various halogenated phenylalanine derivatives.

The isolated GP202 polypeptides can also include non-native inter-residue  
bonds, including bonds that lead to circular and branched forms. The isolated  
GP202 polypeptides can also include post-translational and post-synthetic  
25 modifications, either throughout the length of the protein or localized to one or  
more portions thereof. For example, when produced by recombinant expression in  
eukaryotic cells, the isolated polypeptide can include N-linked and/or O-linked  
glycosylation, the pattern of which will reflect both the availability of  
glycosylation sites on the protein sequence and the identity of the host cell.  
30 Further modification of glycosylation pattern can be performed enzymatically. As  
another example, recombinant polypeptides of the invention may also include an

initial modified methionine residue, in some cases resulting from host-mediated processes.

Amino acid analogues having detectable labels are also usefully incorporated during synthesis to provide a labeled polypeptide (e.g., biotin, various  
5 chromophores, or fluorophores). The GP202 polypeptides of this invention can also usefully be conjugated to polyethylene glycol (PEG). PEGylation increases the serum half life of proteins administered intravenously for replacement therapy.

#### **D. Purification of the Polypeptides**

Production of the isolated polypeptides optionally can be followed by  
10 purification from the producing cells. Producing cells include, without limitation, recombinant cells overexpressing the polypeptides, naturally occurring cells (e.g., cancer cells) overexpressing the polypeptides, or established cancer cell lines overexpressing the polypeptides. If purification tags have been fused through use of an expression vector that appends such tags, purification can be effected, at least  
15 in part, by means appropriate to the tags, such as use of immobilized metal affinity chromatography for polyhistidine tags. Other techniques common in the art include ammonium sulfate fractionation, immuno-precipitation, fast protein liquid chromatography (FPLC), high performance liquid chromatography (HPLC), and preparative gel electrophoresis. Purification of chemically-synthesized peptides  
20 can readily be effected, e.g., by HPLC.

Accordingly, the isolated GP202 proteins may be in pure or substantially pure form. A purified protein is an isolated protein that is present at a concentration of at least 95%, as measured on a mass basis (w/w) with respect to total protein in a composition. Such purities can often be obtained during chemical  
25 synthesis without further purification, as, e.g., by HPLC. Purified proteins can be present at a concentration (measured on a mass basis with respect to total protein in a composition) of 96%, 97%, 98%, and even 99%. The proteins of the present invention can even be present at levels of 99.5%, 99.6%, and even 99.7%, 99.8%, or even 99.9% following purification.

30 Although high levels of purity are preferred when the isolated proteins are used as therapeutic agents, the isolated proteins are also useful at lower purity. For

example, partially purified proteins can be used as immunogens to raise antibodies in laboratory animals.

The isolated proteins are generally used in substantially purified form. As used herein, "substantially purified protein" means a protein present at a concentration of at least 70%, measured on a mass basis with respect to total protein in a composition. Usefully, the substantially purified protein is present at a concentration, measured on a mass basis with respect to total protein in a composition, of at least 75%, 80%, or even at least 85%, 90%, 91%, 92%, 93%, 94%, 94.5% or even at least 94.9%.

In preferred embodiments, the purified and substantially purified proteins are in compositions that lack detectable ampholytes, acrylamide monomers, bis-acrylamide monomers, and polyacrylamide.

## **E. Exemplary Uses of GP202 Polypeptides**

### **1. Therapeutic Use**

Certain fragments of at least 8, 9, 10 or 12 contiguous amino acids are also useful as competitive inhibitors of binding of the entire protein, or a portion thereof, to its ligand. Thus, such fragments can be used as anti-cancer agents to reduce the activity of GP202.

### **2. Epitope Mapping**

Fragments of at least six contiguous amino acids are useful in mapping B cell and T cell epitopes of the reference protein. See, e.g., Geysen et al., "Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid," Proc. Natl. Acad. Sci. USA 81:3998-4002 (1984) and U.S. Pat. Nos. 4,708,871 and 5,595,915. Because the fragment need not itself be immunogenic, part of an immunodominant epitope, nor even recognized by native antibody, to be useful in such epitope mapping, all fragments of at least six amino acids of the proteins of the present invention have utility in such a study.

### **3. Immunogens**

Fragments of at least eight contiguous amino acids, often fifteen to thirty contiguous amino acids, have utility as immunogens for raising antibodies that

recognize GP202 proteins or as vaccines for GP202-mediated diseases such as cancers.

The GP202 proteins, fragments, and fusions of the present invention can usefully be attached to a substrate. When bound to a substrate, the polypeptides  
5 can be used to detect and quantify antibodies, e.g., in serum, that bind specifically to the immobilized protein.

#### **IV. ANTIBODIES AND APTAMERS**

##### **A. General Consideration**

The invention relates to antibodies that bind specifically to the GP202  
10 polypeptides. The antibodies can be specific for linear epitopes, discontinuous epitopes, or conformational epitopes of such polypeptides, either as present on the polypeptide in its native conformation or as present on the polypeptides when denatured, e.g., by solubilization in SDS. In some embodiments, the antibodies, both polyclonal and monoclonal, bind specifically to a polypeptide having an  
15 amino acid sequence presented in SEQ ID NO: 1.

As used herein, an “antibody” means a full antibody, e.g., an antibody comprising two heavy chains and two light chains, or to an antigen-binding fragment of a full antibody. Such fragments include, but are not limited to, those produced by digestion with various proteases, those produced by chemical  
20 cleavage and/or chemical dissociation, and those produced recombinantly, so long as the fragment remains capable of specific binding to an antigen. Among these fragments are Fab, Fab', F(ab')<sub>2</sub> and single chain Fv (scFv) fragments.

An antibody can be a murine or hamster antibody or a homolog thereof, or a fully human antibody. An antibody also can be a humanized antibody, a  
25 chimeric antibody, an antibody fusion, a diabody, an intrabody, or a single-chained antibody. An antibody can be of any isotype and subtype, for example, IgA (e.g., IgA1 and IgA2), IgG (e.g., IgG1, IgG2, IgG3 and IgG4), IgE, IgD, IgM, wherein the light chains of the immunoglobulin may be of type kappa or lambda. While the useful antibodies are generally monoclonal, polyclonal antibodies from mice,  
30 rabbits, turkeys, or sheep may also be used.

Typically, the affinity or avidity of an antibody (or antibody multimer, as in the case of an IgM pentamer) will be at least  $1 \times 10^{-6}$  molar (M), preferably at least  $5 \times 10^{-7}$  M, more preferably at least  $1 \times 10^{-7}$  M, with affinities and avidities of at least about  $1 \times 10^{-8}$  M,  $5 \times 10^{-9}$  M, or  $1 \times 10^{-10}$  M being especially useful.

5           **B.       Moieties Conjugated to the Antibodies**

The antibodies are useful in a variety of *in vitro* immunoassays, such as western blotting and ELISA, in isolating and purifying GP202 proteins (e.g., by immunoprecipitation, immunoaffinity chromatography, or magnetic bead-mediated purification). The antibodies are also useful as modulators (i.e., antagonists or  
10 agonists) of a GP202 protein *in vivo* to modulate the protein's interaction with its natural ligand. The antibodies can also be used to conjugate to cytotoxic reagents for site-specific delivery.

The new antibodies can be variously associated with moieties appropriate for their uses. When the antibodies are used for immunohistochemical staining of  
15 tissue samples, the moieties can be an enzyme that catalyzes production and local deposition of a detectable product. Exemplary enzymes are alkaline phosphatase,  $\beta$ -galactosidase, glucose oxidase, horseradish peroxidase (HRP), and urease. The antibodies also can be labeled using colloidal gold. When the antibodies are used for flow cytometric detection and scanning laser cytometric detection, they can be  
20 labeled with fluorophores. For secondary detection using labeled avidin, streptavidin, captavidin or neutravidin, the antibodies of the present invention can usefully be labeled with biotin. When the antibodies of the present invention are used, e.g., for western blotting, they can usefully be labeled with radioisotopes. When the antibodies are to be used for *in vivo* diagnoses, they can be rendered  
25 detectable by conjugation to MRI contrast agents, such as radioisotopic labeling or gadolinium diethylenetriaminepentaacetic acid (DTPA).

The antibodies also can be conjugated to toxic agents so as to direct the agents to a tumor site. By way of example, the antibody is conjugated to  
30 *Pseudomonas* exotoxin A, diphtheria toxin, shiga toxin A, anthrax toxin lethal factor, or ricin. See, e.g., Hall (ed.), Immunotoxin Methods and Protocols (Methods in Molecular Biology, Vol 166), Humana Press (2000)

(ISBN:0896037754); and Frankel et al. (eds.), Clinical Applications of Immunotoxins, Springer-Verlag New York, Incorporated (1998) (ISBN:3540640975). Small molecule toxins such as calicheamycin or chemotherapeutic agents can also be delivered via chemical conjugation to the antibodies. The antibodies may also be used to deliver DNA to the tumor site as gene therapy to inhibit or otherwise modify the behavior of the tumor, e.g., to deliver an antisense reagent to the GP202 gene.

For some uses, the antibodies can be bound to a substrate through a linker moiety. For example, the antibodies can be conjugated to filtration media, such as NHS-activated Sepharose or CNBr-activated Sepharose for immunoaffinity chromatography. The antibodies also can be attached to paramagnetic microspheres, by, e.g., biotin-streptavidin interaction. The microsphere then can be used for isolation of cells that express or display the above-described proteins. The antibodies also can be attached to the surface of a microtiter plate for ELISA.

### C. Nucleic Acid Aptamers

Aptamers for GP202 are DNA-based or RNA-based molecules that bind to GP202 proteins with high affinity and specificity. Large libraries of nucleic acid compounds can be screened to identify specific molecules that e.g., bind GP202 polypeptides, or inhibit GP202 enzymatic activity. Like antibodies and small molecules, they can be used as research tools or therapeutic agents. For a review, see "Nucleic acid aptamers as tools and drugs: recent developments" by Rimmelle, *Chembiochem* 4:963-71 (2003)).

## V. PHARMACEUTICAL COMPOSITIONS

As a protein involved in tumor genesis, development and/or maintenance, GP202 is a suitable therapeutic target for treating neoplasia, hyperplasia, malignant cancers, or any other hyperproliferative conditions. For example, the GP202 gene can be a target in cancers of the breast, skin, lung, prostate, colorectal tissue, liver, pancreas, brain, testis, ovary, uterus, cervix, kidney, thyroid, bladder, esophagus, and blood.

The invention relates to pharmaceutical compositions comprising GP202 nucleic acids, proteins, and antibodies or aptamers that bind GP202, as well as

mimetics, agonists, antagonists, or modulators of GP202 activity, and methods of using them to prevent or treat (i.e., ameliorate, mitigate, alleviate, slow, or inhibit) tumor growth, angiogenesis, metastasis or any other inappropriate cell proliferation.

5           Inhibitors of GP202 also can be administered in combination with one or more other therapeutic agents, for improved cancer treatment. Other therapeutic agents suitable for co-administration with a GP202 inhibitor include, for example, an anti-angiogenic agent, an anti-metastatic agent, or an agent that creates a hypoxic environment. Chemotherapeutic agents that can be co-administered with  
10           inhibitors of GP202 include folate antagonists, pyrimidine and purine antimetabolites, alkylating agents, platinum antitumor compounds, DNA interchelators, other agents that induce DNA damage, microtubule targeting products, small molecule inhibitors of protein kinases and biological inhibitors of growth factor receptors. The GP202 inhibitor and additional therapeutic agent(s)  
15           may be used concurrently or sequentially. In some embodiments, the subject is pre-treated with one or more agents, followed by treatment with a GP202 inhibitor.

          The ability of tumor cells to detach from the primary site and produce metastases in a distant organ is due to the survival and growth of a unique subpopulation of cells with metastatic properties. Tumor growth and metastasis  
20           are angiogenesis-dependent. Inhibition of angiogenesis will generate a hypoxic environment in the tumor, forcing tumor cells to become more dependent upon the glycolytic pathway for energy generation. Accordingly, preventing angiogenesis in combination with inhibiting GP202 is a promising therapeutic strategy. Angiogenesis inhibitors, e.g. angiostatin, endostatin, Avastin<sup>™</sup> or Regeneron's  
25           VEGF trap technology, can be used in combination with GP202 inhibitors as an effective anti-cancer therapy. Such a combination can be expected to have a synergistic effect. This may also allow the use of a lower dose of GP202 inhibitor or anti-angiogenic agent or both in chemotherapy. This is desirable because it is likely to cause less toxicity in patients. In addition, the use of combinations of  
30           therapeutic agents may circumvent drug resistance problems.

GP202 inhibitors can also be used in combination with agents that create a hypoxic environment to enhance the effect of GP202 inhibitor. Hypoxia, i.e., lack of oxygen, plays a fundamental role in many pathologic processes. In response to hypoxia, mammalian cells activate and express multiple genes. Tumor cells may respond to hypoxia by diminishing their proliferative rates, thereby leaving the cells viable but nonproliferating. Some transformed cell lines can also undergo apoptosis in extreme hypoxia and an acidic environment. Similar to inhibitors of angiogenesis, other agents that induce a hypoxic environment may sensitize tumor cells to inhibition of GP202 and use of hypoxia inducing agents in combination with inhibiting GP202 is therefore another promising therapeutic strategy.

An increase in apoptosis (programmed cell death) has been associated with a decrease in tumor proliferation. Tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) induces apoptosis in several human tumors. A number of cell regulatory pathways such as Rb/E2F pathway, the c-Myc transcription factor, and the Ras signaling molecule have also been shown to control not only cell proliferation but also pathways leading to apoptosis. Further, a combination of GP202 inhibitor with reagents that activate additional apoptotic signals, or inhibit survival signals, can also be used for cancer therapy. Survival signals that have recently been shown to modulate apoptotic signaling include the focal adhesion kinase (FAK), the phosphoinositol 3' kinase (PI3'K), and protein kinase B (PKB, also known as Akt).

A composition of the invention typically contains from about 0.1 to 90% by weight (such as 1 to 20% or 1 to 10%) of a therapeutic agent of the invention in a pharmaceutically accepted carrier. Solid formulations of the compositions for oral administration can contain suitable carriers or excipients, such as corn starch, gelatin, lactose, acacia, sucrose, microcrystalline cellulose, kaolin, mannitol, dicalcium phosphate, calcium carbonate, sodium chloride, or alginic acid. Disintegrators that can be used include, without limitation, microcrystalline cellulose, corn starch, sodium starch glycolate, and alginic acid. Tablet binders that can be used include acacia, methylcellulose, sodium carboxymethylcellulose, polyvinylpyrrolidone (Povidone<sup>TM</sup>), hydroxypropyl methylcellulose, sucrose,

starch, and ethylcellulose. Lubricants that can be used include magnesium stearates, stearic acid, silicone fluid, talc, waxes, oils, and colloidal silica.

Liquid formulations of the compositions for oral administration prepared in water or other aqueous vehicles can contain various suspending agents such as methylcellulose, alginates, tragacanth, pectin, kelgin, carrageenan, acacia, polyvinylpyrrolidone, and polyvinyl alcohol. The liquid formulations can also include solutions, emulsions, syrups and elixirs containing, together with the active compound(s), wetting agents, sweeteners, and coloring and flavoring agents. Various liquid and powder formulations can be prepared by conventional methods for inhalation into the lungs of the mammal to be treated.

Injectable formulations of the compositions can contain various carriers such as vegetable oils, dimethylacetamide, dimethylformamide, ethyl lactate, ethyl carbonate, isopropyl myristate, ethanol, polyols (glycerol, propylene glycol, liquid polyethylene glycol, and the like). Physiologically acceptable excipients can include, for example, 5% dextrose, 0.9% saline, Ringer's solution or other suitable excipients. Intramuscular preparations, e.g., a sterile formulation of a suitable soluble salt form of the compounds, can be dissolved and administered in a pharmaceutical excipient such as Water-for-Injection, 0.9% saline, or 5% glucose solution. A suitable insoluble form of the compound can be prepared and administered as a suspension in an aqueous base or a pharmaceutically acceptable oil base, such as an ester of a long chain fatty acid (e.g., ethyl oleate).

A topical semi-solid ointment formulation typically contains a concentration of the active ingredient from about 1 to 20%, e.g., 5 to 10%, in a carrier such as a pharmaceutical cream base. Various formulations for topical use include drops, tinctures, lotions, creams, solutions, and ointments containing the active ingredient and various supports and vehicles. The optimal percentage of the therapeutic agent in each pharmaceutical formulation varies according to the formulation itself and the therapeutic effect desired in the specific pathologies and correlated therapeutic regimens.

Pharmaceutical formulation is a well-established art, and is further described in Gennaro (ed.), Remington: The Science and Practice of Pharmacy,

20th ed., Lippincott, Williams & Wilkins (2000) (ISBN: 0683306472); Ansel et al., Pharmaceutical Dosage Forms and Drug Delivery Systems, 7th ed., Lippincott Williams & Wilkins Publishers (1999) (ISBN: 0683305727); and Kibbe (ed.), Handbook of Pharmaceutical Excipients American Pharmaceutical Association, 3rd ed. (2000) (ISBN: 091733096X), the disclosures of which are incorporated  
5 herein by reference in their entireties. Conventional methods known to those of ordinary skill in the art can be used to administer the pharmaceutical formulation(s) to the patient.

Typically, the pharmaceutical formulation will be administered to the  
10 patient by applying to the skin of the patient a transdermal patch containing the pharmaceutical formulation, and leaving the patch in contact with the patient's skin (generally for 1 to 5 hours per patch). Other transdermal routes of administration (e.g., through use of a topically applied cream, ointment, or the like) can be used by applying conventional techniques. The pharmaceutical formulation(s) can also  
15 be administered via other conventional routes (e.g., parenteral, subcutaneous, intrapulmonary, transmucosal, intraperitoneal, intrauterine, sublingual, intrathecal, or intramuscular routes) by using standard methods. In addition, the pharmaceutical formulations can be administered to the patient via injectable depot routes of administration such as by using 1-, 3-, or 6-month depot injectable or  
20 biodegradable materials and methods.

Regardless of the route of administration, the therapeutic protein or antibody agent typically is administered at a daily dosage of 0.01 mg to 30 mg/kg of body weight of the patient (e.g., 1mg/kg to 5 mg/kg). The pharmaceutical formulation can be administered in multiple doses per day, if desired, to achieve  
25 the total desired daily dose. The effectiveness of the method of treatment can be assessed by monitoring the patient for known signs or symptoms of a disorder.

The pharmaceutical compositions of the invention may be included in a container, package or dispenser alone or as part of a kit with labels and instructions for administration. These compositions can also be used in combination with other  
30 cancer therapies involving, e.g., radiation, photosensitizing compounds, anti-neoplastic agents and immunotoxics.

## **VI. GP202-RELATED ANIMALS**

### **A. General Consideration**

This invention provides genetically modified non-human mammals at least some of whose somatic cells and germ cells contain one of the above-described  
5 GP202-coding nucleic acid of this invention (including both heterozygotes and homozygotes). Such mammals can be used to study the effect of the GP202 gene on tumorigenicity and tumor development, to study the role of GP202 on normal tissue development and differentiation, to identify via array CGH regions of the genome whose amplification or deletion is correlated with GP202 status, and to  
10 screen for and establish toxicity profiles of anti-cancer drugs. This invention also provides genetically modified non-human mammals with targeted disruption of one or both copies of the endogenous GP202 gene. Animal models according to the invention can be conventional germline transgenic animals or chimeric animals.

### **B. Inducible Cancer Model**

This invention provides an inducible cancer model to study tumor biology and to screen for anti-cancer drugs. In some embodiments, the inducible cancer model is a mouse whose genome has been modified to include: (a) an expression construct comprising a GP202 gene linked operably to an inducible promoter, and  
20 (b) a genetic mutation that causes the mouse to have greater susceptibility to cancer than a mouse not comprising said genetic mutation. Expression of the GP202 gene leads to formation of cancer in the mouse. The cancer regresses when expression of the GP202 gene is reduced. Mutations that render the mammal more susceptible to cancer include disabling mutations in a tumor suppressor gene (e.g., INK4a),  
25 disabling mutations in a DNA repair gene (e.g., MSH2), and activating mutations in an oncogene (e.g., Her2, *myc* and *ras*). Such testing also can be carried out in cells (e.g., human cells) that are engineered to contain an inducible oncogene and endowed with tumorigenic capacity by the presence of an appropriate combination of oncogenes, tumor suppressor genes, and/or telomerase.

30 In one particular embodiment, the mammal's genome comprises (i) a first expression construct containing a gene encoding a reverse tetracycline

transactivator operably linked to a promoter, such as any tissue or cell type-specific promoter or any general promoter, and (ii) a second expression construct containing the GP202 gene operably linked to a promoter that is regulated by the reverse tetracycline transactivator and tetracycline (or a tetracycline analogue, for example, doxycycline). The mammal is observed with and without administration of tetracycline (or analogue thereof) for the development, maintenance, or progression of a tumor that is tetracycline-dependent. Other inducible systems such as those described above also can be employed.

This animal model can be used to determine the efficacy of a candidate compound in preventing or treating cancer. This method involves administering to the mammal a candidate compound and observing the effect of the compound on tumor development, maintenance, angiogenesis and/or progression in the mammal. Regression and/or reduction of tumor size in the presence of the compound indicates effectiveness of the compound. Similarly, the effect of a candidate compound on the level of GP202 mRNA, protein, or activity in the mammal or cell lines derived from the mammal (or cell lines transfected with the gene) can be used to identify the candidate as an agonist or antagonist. The ability to compare the effect of a test compound to that of genetically switching off the inducible oncogene in this system allows the identification of surrogate markers that are predictive of the clinical response to the compound. The inducible model can be used to determine whether a compound can eradicate minimal residual tumor. Normally in the inducible model, a tumor regresses when the GP202 gene is switched from "on" to "off" via the inducible promoter. But if a compound can eradicate minimal residual tumor, switching the gene back on after administration of the compound will not bring back the tumor.

The animal model also can be used to identify other cancer-related elements. To do this, a detailed expression profile of gene expression in tumors undergoing regression or regrowth due to the inactivation or activation of the GP202 transgene is established. Techniques used to establish the profile include the use of suppression subtraction (in cell culture), differential display, proteomic analysis, serial analysis of gene expression (SAGE), and expression/transcription profiling using cDNA and/or oligonucleotide microarrays. Then, comparisons of

expression profiles at different stages of cancer development can be performed to identify genes whose expression patterns are altered.

This animal model can also be used to identify molecular surrogates of GP202 in another manner. To do this, the expression of GP202 gene is turned off  
5 (by removal of the inducer), and another round of MaSS screening is performed using retroviral integration, cDNA complementation, or the genetic suppressor elements (GSE) method. Genes whose activation results in transformation of the cells are likely to be in a tumorigenic pathway related to GP202.

The animal model also can be used to identify surrogate biomarkers for  
10 diagnosis or for following disease progression in patients. The biomarkers can be identified based on the differences between the expression profiles of the “on” and “off” states in the animal model. Blood or urine samples from the animal can be tested with ELISAs or other assays to determine which biomarkers are released from the tumor into circulation during tumor genesis, maintenance, or regression  
15 (when GP202 is turned off). These biomarkers are particularly useful clinically in following disease progression post anti-GP202 therapy.

## VII. CANCER DIAGNOSTICS

Since GP202 activity is up-regulated in tumor cells, one can use GP202 as a marker in diagnosing cancer or any other abnormal hyperplasia conditions. To  
20 do this, the above-described nucleic acid probes or antibodies are used to quantify the expression level of GP202 in a tissue sample. An increase in that level relative to control indicates cancerous, neoplastic, or hyperplastic pathology of the tissue sample. This type of test can be performed using conventional techniques such as RT-PCR, ribonuclease protection assays, *in situ* hybridization, Northern blot  
25 analysis, FISH, CGH, array CGH, SKY, and immunohistochemistry.

Because up-regulation of GP202 is generally associated with a malignant state, a GP202 protein or GP202 protein fragments may be found to be elevated in a tissue sample (e.g., blood or urine) of cancer patients relative to that of normal individuals. This elevation can be detected by, e.g., specific ELISAs,  
30 radioimmunoassays, or protein chip assays. Such tests may not only be useful for

diagnosis of GP202-related diseases such as cancers, but also for monitoring the progress of therapy using GP202 inhibitors.

## VIII. EXAMPLES

The invention is further illustrated by the following examples. The examples are provided for illustrative purposes only, and are not to be construed as limiting the scope or content of the invention in any way.

### Example 1: MaSS Screening Identification of the Gene

This example describes the procedures for identifying the GP202 gene by MaSS screening.

#### 10 a. Retroviral Infection of Tumor Cells

Viral production: Mo-MuLV producer cell line TMJ (NIH3T3 based cell line) was plated to the required number of plates (100 mm). These cells were cultured and maintained in RPMI media with 10% FBS. For viral production, TMJ cells were fed with 4-5 ml of fresh culture media, and culture supernatant was harvested 8-12 hours later. The supernatant was filtered through a 0.45  $\mu$ M filter.

Cells for MaSS screen: Material for MaSS screen was prepared from the spontaneous, oncogene-induced tumors derived from the animal model of choice. In this example, HER2-induced mammary tumors arising from the breast HER2 model (BH) were dissociated by protease digestion and cultured in vitro for limited passages to produce individual culture lines such as BH3, BH1-1, BH1-2, BH1-3, BH8 and BH15. These culture lines retained the *in vivo* tumorigenic potential of their parent tumor material, in a doxycycline-dependent manner.

Viral infection: Doxycycline-dependent, HER2-induced mammary tumor cells such as BH1-1, BH1-2, BH1-3 or BH3 cells were maintained in RPMI media with 10% fetal bovine serum and 10 nM  $\beta$ -estradiol in the presence of doxycycline (2  $\mu$ g/ml). One day prior to viral infection, BH cells were seeded at 30 percent confluence. Next day, filtered Mo-MuLV viral supernatant was added to the BH cells in the presence of polybrene (6-8  $\mu$ g/ml),  $\beta$ -estradiol (10 nM) and doxycycline (2  $\mu$ g/ml).

Forty eight hours after addition of the virus, infected BH cells were trypsinized, rinsed and resuspended in Hanks' Balanced Salt Solution. Cell suspensions were kept on ice and the handling time after trypsinization was kept to a minimum. About  $1 \times 10^6$  Mo-MLV infected cells were injected onto the flank of SCID mice. Subsequent to injection, the recipient SCID mice were fed with water without doxycycline. The animals were observed for tumor development. Positive control animals were injected with  $1 \times 10^6$  BH uninfected cells, but the recipient mice were fed with water containing doxycycline. Negative control animals were similarly injected with  $1 \times 10^6$  uninfected BH cells. Mo-MuLV-derived tumors typically developed after approximately 20-120 days, depending on the culture cells used. Tumors were harvested and tumor tissue processed for various uses, such as genomic DNA preparation, mRNA preparation, histological staining, and cryopreservation.

#### **b. Recovery of integration sites by ligation-mediated PCR**

This recovery protocol involves a restriction endonuclease digestion of the genomic DNA from each Mo-MuLV derived tumor, followed by linker ligation, two PCR-amplification steps, and final cloning of the products into destination plasmids. Several restriction enzymes can be used, including, but not limited to, BstYI and Tsp 509I. In the examples below, linkers and primers used for the BstYI and the Tsp 509I protocols are given.

DNA was isolated from tumor tissues using the PureGene<sup>®</sup> DNA isolation kit (Gentra Systems, Minneapolis, MN). Genomic DNA (1  $\mu$ g) was digested to completion with either BstYI or Tsp509I and the reaction was terminated by either incubation at 65°C for 20 minutes or phenol/chloroform extraction. The digested samples were ligated with an annealed linker (BstYI linkers or Tsp509I linkers) in the molar ratio 1:10 (genomic fragment : linker) using NEB Quick Ligation<sup>™</sup> kit (New England BioLabs<sup>®</sup>, Beverly, MA). Ligation mediated-PCR was performed with one primer specific to the LTR (5' or 3' LTR primer 1) and the other primer to the linker (Linker primer 1) with the following conditions: pre-incubation at 94°C for 10 min, then 28 cycles of (95°C for 15 sec, 55°C for 30 sec and 72°C for 2 min), and a final extension step at 72°C for 10 min.

The PCR products were diluted 1:100 and nested PCR was performed under the following conditions: pre-incubation at 94°C for 10 min, then 28 cycles of (95°C for 15 sec, 58°C for 30 sec and 72°C for 2 min), a final extension step at 72°C for 10 min. The second pair of primers have attB sites at the ends, and are designed to bind to the LTR and the linker region (5' or 3' LTR primer 2 and Linker primer 2). The resulting nested PCR products were cleaned with Millipore Montage<sup>®</sup> column (Millipore, Bedford, MA) and cloned into the pDONR221 vector using Invitrogen Gateway<sup>®</sup> BP Clonase<sup>™</sup>. (Invitrogen, Carlsbad, CA) and transformed into DH5 $\alpha$  to form plasmid libraries of integration junction fragments. The library agar lawns were sent to Agencourt (Beverly, MA) for colony picking and sequencing.

### c. Identification of Candidate Genes

Mapping integration sites: The precise site and orientation of retroviral integration into the mouse genome was mapped for the recovered sequences as follows. Retroviral leader sequences were trimmed from the recovered sequences, and homology searches for the trimmed sequences were performed in the NCBI MGSCV3 database by using the BLAST software program. Unique BLAST hits were analyzed and mapped to the mouse genome, and recurrent sites of integration in multiple mouse tumors were identified. NCBI MapView was used to identify the site of each recurrent retroviral integration onto the mouse genome. Recurrence was defined as two or more integrations within a specified region, for example, 20 kb. Genes either containing the integration sites or immediately neighboring the sites were identified by using the MGSCV3 Gene map. Published literature shows that Mo-MuLV favors integration events near transcribed genes, frequently resulting in alterations in expression of the latter. Integration events within a gene implicates that particular gene as the best candidate target for the biological effects of said integrations. Using these criteria, a set of candidate genes involved in cancer were identified from each MaSS screen. Genes identified in the BH screen, in the specific example, possess the ability to functionally complement the HER2 oncogene in this inducible, *in vivo* cancer model (mouse).

#### **d. Identification of GP202 as a cancer gene from a Mass Screen**

The set of candidate genes were analyzed according to several criteria, including tumor recurrence and transcript abundance in the target tumors.

Tumor recurrence. Tumor recurrence was defined as the number of tumors in which insertions implicating a particular gene have been identified. For example, in one screen, from the total set of 888 recurrent hits, 275 genes were each identified as recurrent hits in at least 3 tumors, whereas only 9 genes were each identified in at least 8 tumors. One of these 9 genes was GP202. GP202 was identified as a major target of viral insertion in the BH MaSS screens in eight independent tumors within 20 kb of the gene (tumor ID no. 153, 101, 140, 121, 195, 180, 91, 156). The probability that these insertions occurred at random is  $p = 5.5 \times 10^{-11}$ .

Expression in target tumors. We tested whether MoMuLV insertions in or near GP202 lead to increased expression of the target gene, by quantitative RT-PCR of mRNA extracted from target tumors. Elevated expression of GP202 was detected in target tumor 180 (about 6-12 fold over control tumors). Expression of GP202 could not be monitored by microarray analysis because it is not represented in the Agilent 60-mer array. Taken together, the recurrence of GP202 hits among eight independent tumors in the MaSS screen, along with confirmation of its overexpression in at least one target tumor identify GP202 as a significant gene emerging from the HER2 complementation screen, and a novel therapeutic candidate for cancer.

#### **Example 2: Expression in Human Tumors and Tumor Cell Lines**

This example describes the protocols for expressing the candidate gene identified above in human cancer cells.

Primer pairs for each human gene were designed as described for the mouse gene. Expression of each candidate gene was assessed in a panel of 31 human cancer cell lines and 47 human primary tumors by using real-time reverse transcription PCR. The forward and reverse primers were

5'-cacatggaggatgaggacag-3' (SEQ ID NO: 11) and

5'-gagctctggcttcggattctc-3' (SEQ ID NO: 12), respectively.

RNA was prepared from the cells using QIAGEN<sup>®</sup> RNeasy<sup>®</sup> mini-prep kits and QIAGEN<sup>®</sup> RNeasy<sup>®</sup> maxi-prep columns. RNA preparations were treated with DNase during column purification according to manufacturer's instructions.

5 Expression of each gene was determined in triplicate for all tumors and cell lines using SYBR<sup>®</sup> green-based real-time PCR. To do this, 2X SYBR<sup>®</sup> green PCR master mix (ABI) was mixed with the MultiScribe<sup>™</sup> reverse transcriptase (ABI) and RNase inhibitor (ABI). Forward and reverse primers were added at ratios previously optimized for each gene using control human reference RNA  
10 (Stratagene). 50 ng of RNA template was used per reaction, and the reactions were performed in a total volume of 20  $\mu$ l. Real-time quantification was performed using the ABI 7900HT and SDS2.0 software. RNA loading was normalized for  $\beta$ -actin and 18S rRNA. RNA quantity was determined relative to human universal reference RNA (Stratagene) to permit run-to-run comparisons.

### 15 **Example 3: RNAi Inhibition of Human Cancer Cell Lines**

This example describes the protocols used to inhibit the expression of the candidate gene in human cancer cells by RNA interference (RNAi).

#### **a. Transfection with siRNA**

A pool of double-stranded siRNA oligonucleotides was purchased from a  
20 commercial vendor (SMARTPOOL<sup>™</sup>, Dharmacon RNA Technologies, Lafayette, CO). Human cancer cell lines were transfected with siRNA using the Oligofectamine<sup>™</sup> reagent from Invitrogen. Transfected cells were allowed to grow for at least 24 hours. The cells were then trypsinized and reseeded for growth curve analysis or growth in soft agar.

#### 25 **b. Soft Agar Assay**

One day prior to seeding cells for assay, 6-well plates containing bottom agar were prepared. The bottom agar was made of 0.7% agarose (SeaKem<sup>®</sup> GTG<sup>®</sup> agarose) and 1X DME plus 10% fetal bovine serum. The next day, cells were seeded into each well by adding 5ml of a mixture containing  $1 \times 10^4$  cells in

0.32% agarose and 1 X DME plus 10% FBS. The cell mixture was allowed to solidify at room temperature for 30 min. The plates were then incubated at 37°C in a 5% CO<sub>2</sub> atmosphere for 1 to 14 days. Colony formation was analyzed and photographs taken at various time points.

#### 5                    **d. Sequence of siRNAs for Target Genes**

The sense sequences of four independent siRNAs that inhibit expression of GP202 are shown below.

5'- CAACGUGCGUGCCCAUGAA -3' (SEQ ID NO: 3);

5'- CAACGUGCGUGCCCAUGAA -3' (SEQ ID NO: 4);

10                    5'- GCCAUCUACUCGUCGUUUA -3' (SEQ ID NO: 5); and

5'- GAGAGCCACUUACGAUGCA -3' (SEQ ID NO: 6).

The effect of transfecting a pool of the above GP202-specific siRNAs in human cancer cell lines U-251 (glioblastoma), NCI-H460 (lung carcinoma), HCT-116 (colon carcinoma) and SKOV-3 (ovarian carcinoma) was compared to the effect of transfecting a negative control siRNA that targets firefly luciferase and a positive control siRNA that targets PLK1 (GenBank<sup>™</sup> NM\_005030), an essential gene for mitosis. The data are summarized in Table 1.

**Table 1**  
**siRNA Soft Agar Assay Results**

| <b>Cancer Cell Line</b>       | <b>GP202 siRNA</b> | <b>Positive Control</b> | <b>Negative Control</b> |
|-------------------------------|--------------------|-------------------------|-------------------------|
| U-251<br>(glioblastoma)       | 49 ± 13            | 10 ± 0.2                | 100 ± 15                |
| NCI-H460 (lung<br>carcinoma)  | 23 ± 0.7           | 11 ± 2                  | 100 ± 15                |
| HCT-116 (colon<br>carcinoma)  | 32 ± 4             | 37 ± 5                  | 100 ± 8                 |
| SKOV-3 (ovarian<br>carcinoma) | 26 ± 12            | 21 ± 3                  | 100 ± 10                |

Significant inhibition of colony formation was observed in all four cancer cell lines upon inhibition of GP202 expression. This indicates that these cancer cell lines are dependent upon GP202 for tumorigenic growth.

5     **Example 4: RNAi on Human or Mouse model Cancer Cell Lines in SCIDs**

Human cancer cell lines expressing high levels of the candidate, or cell lines established from the inducible mouse cancer model described above, or cells derived from MaSS screen tumors are transfected with a vector encoding a short hairpin RNA homologous to the candidate cancer-related gene. Expression of the  
10     RNA is placed under the control of an inducible U6 promoter. Stable cell lines are established.  $5 \times 10^5$  cells are injected subcutaneously into the flank of 6 week old female inbred SCID mice. Tumor formation is observed visually. For cell lines derived from the inducible mouse model, tumor formation is induced by doxycycline. Expression of the RNAi is induced once tumors are visually  
15     identified. In the case of MaSS screen-derived tumor cells, the mice are not fed with doxycycline. Tumor regression is followed using calipers to measure the shrinking tumor diameter.

**Example 5: Tumor Formation in SCIDs**

Tumor cells derived from the tumor suppressor null (INK4/arf -/-)  
20     doxycycline-inducible oncogene mouse model are infected with retrovirus encoding the candidate gene under the control of an IPTG-inducible promoter. Stable cell lines are established.  $10^6$  cells are injected subcutaneously into the flanks of 6-week-old female inbred SCID mice. Mice are fed with doxycycline for 7-12 days. 24 hours prior to doxycycline withdrawal, mice are fed IPTG. IPTG  
25     feeding is maintained after withdrawal of doxycycline and tumor regression is monitored using calipers.

**Example 6: Directed Complementation in Mouse Breast Tumor Cells**

This example describes the protocols used to demonstrate the ability of Kirrel1 (GP202) cDNA to complement a non-induced HER2 oncogene in mouse  
30     breast cells, thereby producing a tumor phenotype.

**a. Constructing the retroviral vector**

The vector pB-Kirrel were constructed as follows. A 1.8 kb mouse Kirrel1 cDNA fragment (ATCC, Manassas, VA; cat# MGC-38329) was cloned into vector pENTR11 (Invitrogen; cat#11819-018) to generate vector pENTR-Kirrel. A  
5 recombination reaction was performed between vector pENTR-Kirrel and vector pBGD-1 using the Gateway<sup>®</sup> LR method. The resulting vector, pB-Kirrel, was confirmed by restriction enzyme digestion and nucleic acid sequencing.

**b. Generating ecotropic retrovirus**

Ecotropic retrovirus packaging cell line EcoPack2-293 and Phoenix<sup>™</sup> were  
10 obtained from BD Biosciences Clontech (Palo Alto, CA; cat#631507) and Origene Inc. (San Diego, CA; cat# PVL-10014). EcoPack2-293 or Phoenix<sup>™</sup> cells were cultured in DMEM medium (containing 10% FBS, 50 U/ml penicillin, 50 mg/ml streptomycin) until 70% confluent. 20 mg DNA of pB-Kirrel retroviral vector was transfected into EcoPack2-293 or Phoenix<sup>™</sup> cells using Fugene6  
15 transfection reagent (Roche Applied Science, Indianapolis, IN; cat# 11814443001). Transfection efficiency was monitored by GFP visualization at 24 hours after transfection. Retrovirus was harvested at 48 hours after transfection by filtering the medium supernatant through a 0.45 micron filter. The filtered supernatant was stored at -80°C.

**c. Retroviral infection of tumor cells**

20 Mouse breast HER2 tumor cells were cultured in RPMI 1640 medium (containing 10% FBS, 50 U/ml penicillin, 50 mg/ml streptomycin and 2 mg/ml doxycycline). At approximately 18-24 hours after plating, or when the plated cells are 70-80% confluent, the breast tumor cells were infected with thawed retroviral  
25 supernatant in the presence of polybrene (8 mg/ml).

**d. Generation of complemented tumor**

Approximately 24 hours after retroviral infection, infected breast tumor cells were trypsinized, rinsed and resuspended in Hans's Balanced Salt Solution. About  $1 \times 10^6$  of these cells were injected into the flank of each SCID mouse that  
30 had been fed without doxycycline. The animals were observed for tumor

development. Control animals were each similarly injected with  $1 \times 10^6$  uninfected cells.

**e. Results**

5 Tumors complemented with pB-Kirrel retrovirus were observed after approximately 31 days in 3 out of 6 injection sites. As examined by real time PCR, mouse Kirrel1 expression levels in complemented tumors were about 10-25 fold higher than a normal mouse reference. The expression of retroviral constructs in tumor cells was also confirmed by GFP protein immunohistochemistry on formalin fixed tumor samples.

10

Other embodiments are within the following claims.

What is Claimed is:

1. An interfering RNA that inhibits the expression of GP202.
2. The interfering RNA of claim 1, wherein the interfering RNA targets the sequence of SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, or SEQ ID NO: 6.
3. The interfering RNA of claim 1, wherein the interfering RNA inhibits tumorigenesis, tumor development, tumor maintenance, tumor recurrence, tumor growth, or growth of tumor cells *in vitro*.
4. A method of inducing apoptosis in a cell, comprising contacting the cell with an effective amount of the interfering RNA of claim 1.
5. A method of treating a hyperproliferative condition in a mammal, comprising administering to the mammal an effective amount of the interfering RNA of claim 1.
6. The method of claim 5, wherein the hyperproliferative condition is a cancer.
7. The method of claim 5, further comprising the step of administering a second therapeutic agent to the mammal.
8. The method of claim 6, wherein said second therapeutic agent is selected from the group consisting of an anti-angiogenic agent, an anti-metastatic agent, an agent that induces hypoxia, an agent that induces apoptosis, and an agent that inhibits cell survival signals.
9. An antibody that specifically binds to GP202 polypeptide and inhibits GP202 activity.
10. The antibody of claim 1, wherein the antibody inhibits tumorigenesis, tumor development, tumor maintenance, tumor recurrence, tumor growth, or growth of tumor cells *in vitro*.
11. A method of inducing apoptosis in a cell, comprising contacting the cell with an effective amount of the antibody of claim 9.

12. A method of treating a hyperproliferative condition in a mammal, comprising administering to the mammal an effective amount of the antibody of claim 9.
13. The method of claim 12, wherein the hyperproliferative condition is a cancer.
14. The method of claim 12, further comprising the step of administering a second therapeutic agent to the mammal.
15. The method of claim 14, wherein the second therapeutic agent is selected from the group consisting of an anti-angiogenic agent, an anti-metastatic agent, an agent that induces hypoxia, an agent that induces apoptosis, and an agent that inhibits cell survival signals.
16. A host cell comprising a recombinant DNA comprising a GP202-encoding sequence operably linked to an expression control sequence, wherein the host cell further comprises a genetic mutation that causes the host cell to have a greater likelihood of becoming a cancer cell than a cell not comprising the genetic mutation.
17. The cell of claim 16, where the genetic mutation disables a tumor suppressor gene.
18. A genetically modified non-human mammal at least some of whose cells comprise a genome comprising: (a) a recombinant GP202-encoding nucleic acid operably linked to an expression control sequence, and (b) a genetic mutation that causes the mammal to have a greater susceptibility to cancer than a mammal not comprising the genetic mutation.
19. The genetically modified nonhuman mammal of claim 18, where the genetic mutation disables a tumor suppressor gene.
20. The genetically modified nonhuman mammal of claim 18, wherein the mammal is a transgenic mammal, all of whose cells comprise a recombinant GP202-encoding nucleic acid operably linked to an expression control sequence, and a genetic mutation that causes the mammal to have a greater susceptibility to cancer than a mammal not comprising the genetic mutation.

21. The genetically modified nonhuman mammal of claim 18, wherein the mammal is a chimeric mammal at least some of whose, but not all of whose, somatic cells comprise a recombinant GP202-encoding nucleic acid operably linked to an expression control sequence, and a genetic mutation that causes the mammal to have a greater susceptibility to cancer than a mammal not comprising the genetic mutation.
22. The chimeric mammal of claim 21, wherein the percentage of somatic cells comprising a recombinant GP202-encoding nucleic acid operably linked to an expression control sequence, and a genetic mutation that causes the mammal to have a greater susceptibility to cancer is between 5% and 95%.
23. The chimeric mammal of claim 22, wherein the percentage of somatic cells comprising the recombinant GP202-encoding nucleic acid operably linked to an expression control sequence, and the genetic mutation that causes the mammal to have a greater susceptibility to cancer is between 15% and 85%.
24. The genetically modified nonhuman mammal of claim 18, wherein the expression control sequence comprises a tissue-specific expression system.
25. A genetically modified nonhuman mammal, wherein the genetic modification reduces or eliminates expression of the mammal's endogenous GP202 genes.
26. The mammal of claim 25, wherein the genetic modification is a knockout of at least one of the mammal's endogenous GP202 alleles.
27. The mammal of claim 25, wherein the genetic modification is addition of an RNAi expression construct targeting GP202 gene expression.
28. The mammal of claim 25, wherein the genetic modification eliminates expression of the mammal's endogenous GP202 genes in a tissue-specific manner.
29. The mammal of claim 25, wherein the mammal is chimeric with respect to the genetic modification.
30. A screening method for identifying a compound useful for treating a hyperproliferative condition, comprising:

a) identifying a biomarker whose level correlates with inhibition of GP202 activity; and

b) detecting a change in the level of the biomarker in the presence of a test compound relative to the level of the biomarker detected in the absence of the test compound.

31. The method of claim 30, wherein said hyperproliferative condition is cancer.

32. A screening method for identifying a compound useful in treatment of a hyperproliferative condition comprising:

(a) providing an inhibitor of GP202 expression or activity;

(b) identifying a negative control biomarker pattern formed by a plurality of biomarkers in a cancer cell, wherein the negative control biomarker pattern is formed when the cell is not contacted with the inhibitor of GP202 expression or activity;

(c) identifying a positive control biomarker pattern formed by a plurality of biomarkers in the cancer cell, wherein the positive control biomarker pattern is formed when the cancer cell is contacted with the inhibitor of GP202 expression or activity;

(d) identifying a test biomarker pattern formed by a plurality of biomarkers in the cancer cell, wherein the test biomarker pattern is formed when the cancer cell is contacted with a candidate compound but not contacted with the inhibitor of GP202 expression or activity; and

(e) comparing the negative control biomarker pattern, the positive control biomarker pattern and the test biomarker pattern,

whereby detecting a greater similarity between the positive control biomarker pattern and the test biomarker pattern than between the negative control biomarker pattern and the test biomarker pattern is indicative that the candidate compound is useful for the treatment of the hyperproliferative condition.

33. The method of claim 30, wherein the hyperproliferative condition is a cancer.