



(51) International Patent Classification:

A61P 35/00 (2006.01) G01N 33/50 (2006.01)
C12N 15/11 (2006.01) G01N 33/574 (2006.01)
C12Q 1/68 (2018.01)

(21) International Application Number:

PCT/US2018/042243

(22) International Filing Date:

16 July 2018 (16.07.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/532,594 14 July 2017 (14.07.2017) US

(71) Applicant: DANA-FARBER CANCER INSTITUTE, INC. [US/US]; 450 Brookline Avenue, Boston, MA 02215-5450 (US).

(72) Inventors; and

(71) Applicants: HAINING, William, N. [GB/US]; 100 Homer Street, Newton, MA 02459 (US). MANGUSO, Robert [US/US]; 210 Calumet Street, Apt. 2, Boston, MA 02120 (US). LONG, Adrienne [US/US]; 99 Pond Avenue, Apt. 812, Brookline, MA 02445 (US).

(74) Agent: SMITH, DeAnn F. et al.; Foley Hoag LLP, 155 Seaport Boulevard, Boston, MA 02210-2600 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,

MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available):

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

Declarations under Rule 4.17:

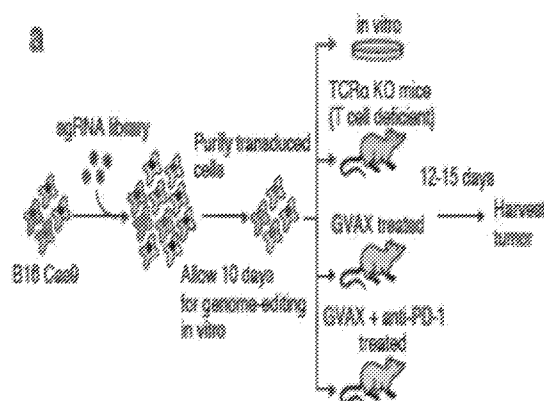
— of inventorship (Rule 4.17(iv))

Published:

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

(54) Title: MODULATING BIOMARKERS TO INCREASE TUMOR IMMUNITY AND IMPROVE THE EFFICACY OF CANCER IMMUNOTHERAPY

Figure 1



(57) Abstract: The present invention relates, in part, to methods of treating a cancer in a subject comprising administering to the subject a therapeutically effective amount of an agent that inhibits one or more biomarkers listed in Table 1 in combination with an immunotherapy.



MODULATING BIOMARKERS TO INCREASE TUMOR IMMUNITY AND IMPROVE THE EFFICACY OF CANCER IMMUNOTHERAPY

Cross-Reference to Related Applications

5 This application claims the benefit of U.S. Provisional Application No. 62/532,594, filed on 14 July 2017; the entire contents of said application are incorporated herein in their entirety by this reference.

Background of the Invention

10 The striking clinical success of cancer immunotherapy with checkpoint blockade suggests it is likely to form the foundation of curative therapy for many malignancies (Reck *et al.* (2016) *N. Engl. J. Med.* 375:1823-1833; Hodi *et al.* (2010) *N. Engl. J. Med.* 363:711-723; Postow *et al.* (2015) *N. Engl. J. Med.* 372:2006-2017; Wolchok *et al.* (2013) *N. Engl. J. Med.* 369:122-133; Ferris *et al.* (2016) *N. Engl. J. Med.* 375:1856-1867; Brahmer *et al.* (2012) *N. Engl. J. Med.* 366:2455-2465; Nghiem *et al.* (2016) *N. Engl. J. Med.* 374:2542-2552; Topalian *et al.* (2012) *N. Engl. J. Med.* 366:2443-2454); Motzer *et al.* (2015) *N. Engl. J. Med.* 373:1803-1813). However, despite these successes, checkpoint blockade does not achieve sustained clinical response in most patients (Tumeh *et al.* (2014) *Nature* 515:568-571; Kelderman *et al.* (2014) *Mol. Oncol.* 8:1132-1139; Zaretsky *et al.* (2016) *N. Engl. J. Med.* 375:819-829). Additional therapeutic strategies are therefore needed to increase the clinical efficacy of immunotherapy. Moreover, the optimal strategy for combining emerging cancer immunotherapies with checkpoint blockade remains uncertain.

25 A relatively small number of genes, such as PD-L1, that enable tumors to evade the immune system have been discovered and most of these are already the focus of intense efforts to develop new immunotherapies (Freeman *et al.* (2000) *J. Exp. Med.* 192:1027-1034; Hirano *et al.* (2005) *Cancer Res.* 65:1089-1096; Dong *et al.* (2002) *Nat. Med.* 8:793-800; Balachandran *et al.* (2011) *Nat. Med.* 17:1094-1100; Spranger *et al.* (2013) *Sci Transl Med.* 5:200ra116; Holmgaard *et al.* (2013) *J. Exp. Med.* 210:1389-1402; Sockolosky *et al.* (2016) *Proc. Natl. Acad. Sci. U.S.A.* 113:E2646-654; Liu *et al.* (2015) *Nat. Med.* 21:1209-1215; Weiskopf *et al.* (2016) *J. Clin. Invest.* 126:2610-2620; Tseng *et al.* (2013) *Proc. Natl. Acad. Sci. U.S.A.* 110: 11103-11108; Sica *et al.* (2003) *Immunity* 18:849-861; Zang *et al.* (2007) *Proc. Natl. Acad. Sci. U.S.A.* 104, 19458-19463). Although cancer cells could, in

theory, express many more genes that regulate their response or resistance to tumor immunity, strategies to systematically discover such genes are lacking.

Loss-of-function genetic screens have been increasingly used to study the functional consequences of gene deletion on tumor cells (Howard *et al.* (2016) Functional Genomic Characterization of Cancer Genomes. *Cold Spring Harb. Symp. Quant. Biol.* (2016); Ebert *et al.* (2008) *Nature* 451:335-339; Cowley *et al.* (2014) *Scientific Data* 1:article number 140035). These approaches include pooled genetic screens using CRISPR-Cas9-mediated genome editing that simultaneously test the role of a large number of genes on tumor cell growth, viability or drug resistance (Wang *et al.* (2014) *Science* 343:80-84; Shalem *et al.* (2014) *Science* 343:84-87). However, these screens have generally been conducted *in vitro*, where the contribution of the immune system is absent, or have studied phenotypes such as metastasis that do not directly evaluate the role of tumor immunity (Hart *et al.* (2015) *Cell* 163:1515-1526; Yu *et al.* (2016) *Nat. Biotechnol.* 34:419-423; Chen *et al.* (2015) *Cell* 160:1246-1260).

Despite the dramatic clinical success of cancer immunotherapy with PD-1 checkpoint blockade, most patients do not experience sustained clinical benefit from treatment. Accordingly a great need in the art exists for additional therapeutic strategies.

Summary of the Invention

The present invention is based, at least in part, on the discovery that inhibiting or blocking one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, and/or Erp1), in combination with an immunotherapy, results in a synergistic therapeutic benefit for treating cancers that is unexpected given the lack of such benefit observed for the immunotherapy alone.

In one aspect, a method of treating a subject afflicted with a cancer comprising administering to the subject a therapeutically effective amount of an agent that inhibits the copy number, the expression level, and/or the activity of one or more biomarkers listed in Table 1 or a fragment thereof, in combination with an immunotherapy, is provided.

Numerous embodiments are further provided that can be applied to any aspect of the present invention and/or combined with any other embodiment described herein. For example, in one embodiment, the agent described herein decreases the copy number, the expression level, and/or the activity of one or more biomarkers in Table 1 (*e.g.*, one or more

regulators of antigen processing and presentation (*e.g.*, HLA-E, Pdia3, Calr, Tap1, Tap2, Tapbp, and/or Erp1). In another embodiment, the agent selectively decreases the activity of one or more biomarkers in Table 1, such as decreasing the antigen processing and presentation activity and/or the substrate binding activity of one or more regulators of antigen processing and presentation (*e.g.*, HLA-E, Pdia3, Calr, Tap1, Tap2, Tapbp, and/or Erp1). In still another embodiment, the agent described herein antagonizes NKG2A-HLA-E interaction. In yet another embodiment, the agent described herein decreases the copy number, the expression level, and/or the activity of *HLA-E*, *Pdia3*, *Calr*, *Tap1*, *Tap2*, *Tapbp*, and/or *Erp1*. In one embodiment, the agent described herein selectively decreases the catalytic activity and/or the substrate binding activity of HLA-E. In still another embodiment, the agent described herein is a small molecule inhibitor, CRISPR single-guide RNA (sgRNA), RNA interfering agent, antisense oligonucleotide, peptide or peptidomimetic inhibitor, aptamer, or intrabody. In another embodiment, the RNA interfering agent is a small interfering RNA (siRNA), CRISPR RNA (crRNA), a small hairpin RNA (shRNA), a microRNA (miRNA), or a piwi-interacting RNA (piRNA). In still another embodiment, the RNA interfering agent is a CRISPR single-guide RNA (sgRNA). In yet another embodiment, the sgRNA comprises a nucleic acid sequence selected from the group consisting of nucleic acid sequence listed in Table 2. In one embodiment, the agent described herein comprises an intrabody, or an antigen binding fragment thereof, which specifically binds to the one or more biomarkers in Table 1 and/or a substrate of the one or more biomarkers in Table 1. In another embodiment, the agent described herein comprises an intrabody, or an antigen binding fragment thereof, which specifically binds to HLA-E and/or a substrate of HLA-E. In still another embodiment, the intrabody, or antigen binding fragment thereof, is murine, chimeric, humanized, composite, or human. In yet another embodiment, the intrabody, or antigen binding fragment thereof, is detectably labeled, comprises an effector domain, comprises an Fc domain, and/or is selected from the group consisting of Fv, Fav, F(ab')₂, Fab', dsFv, scFv, sc(Fv)₂, and diabodies fragments. In one another embodiment, the intrabody, or antigen binding fragment thereof, is conjugated to a cytotoxic agent. In another embodiment, the cytotoxic agent is selected from the group consisting of a chemotherapeutic agent, a biologic agent, a toxin, and a radioactive isotope. In still another embodiment, the agent described herein increases the sensitivity of the cancer cells to an immunotherapy. In another embodiment, the immunotherapy and/or a cancer therapy is administered before, after, or concurrently

with the agent. In still another embodiment, the immunotherapy comprises an anti-cancer vaccine and/or virus. In yet embodiment, the immunotherapy is cell-based. In one embodiment, immunotherapy inhibits an immune checkpoint. In another embodiment, the immune checkpoint is selected from the group consisting of CTLA-4, PD-1, VISTA, B7-H2, B7-H3, PD-L1, B7-H4, B7-H6, ICOS, HVEM, PD-L2, CD160, gp49B, PIR-B, KIR family receptors, TIM-1, TIM-3, TIM-4, LAG-3, GITR, 4-1BB, OX-40, BTLA, SIRP, CD47, CD48, 2B4 (CD244), B7.1, B7.2, ILT-2, ILT-4, TIGIT, HHLA2, butyrophilins, IDO, CD39, CD73 and A2aR. In still another embodiment, the immune checkpoint is selected from the group consisting of PD-1, PD-L1, and PD-L2. In yet another

10 embodiment, the immune checkpoint is PD-1. In one embodiment, the cancer therapy described herein is selected from the group consisting of radiation, a radiosensitizer, an immunogenic chemotherapy (*e.g.*, 5'-fluorouracil, anthracyclines, such as doxorubicin, and the platinum drug, oxaliplatin) that induce interferon production by the cancer cells or at the site of a tumor, interferon, an interferon-inducing agent, a topical inflammatory agent, and a

15 topical TLR agonist. In still another embodiment, the one or more biomarkers described herein comprise a nucleic acid sequence having at least 95% identity to a nucleic acid sequence listed in Table 1 and/or encode an amino acid sequence having at least 95% identity to an amino acid sequence listed in Table 1. In another embodiment, the the HLA-E is a human HLA-E, a mouse H2-T23, a chimeric HLA-E, or a fusion HLA-E. In still

20 another embodiment, the HLA-E comprises an amino acid sequence listed in Table 1, such as an amino acid sequence is selected from the group consisting of SEQ ID Nos: 2, 4, 6, 8, 10, 12, and 14. In yet another embodiment, the HLA-E is encoded by a nucleic acid sequence listed in Table 1, such as a nucleic acid sequence is selected from the group consisting of SEQ ID Nos: 1, 3, 5, 7, 9, 11, and 13. In another embodiment, the agent

25 described herein reduces the number of proliferating cells in the cancer and/or reduces the volume or size of a tumor comprising the cancer cells. In another embodiment, the agent increases the sensitivity of the cancer to the immunotherapy. In still another embodiment, the immunotherapy is T-cell-mediated. In yet another embodiment, the agent increases the amount of CD8⁺ T cells in a tumor comprising the cancer cells. In another embodiment,

30 the cancer described herein is selected from the group consisting of melanoma and colorectal cancer. In another embodiment, the cancer is melanoma. In still another embodiment, the subject described herein is an animal model of the cancer. In another embodiment, the animal model is a mouse model. In still another embodiment, the subject

is a mammal. In yet another embodiment, the mammal is a mouse or a human. In one embodiment, the mammal is a human. In another embodiment, the agent described herein is administered in a pharmaceutically acceptable formulation.

5 In another aspect, a method of killing cancer cells comprising contacting the cancer cells with an agent that inhibits the copy number, the expression level, and/or the activity of one or more biomarkers listed in Table 1 or a fragment thereof, in combination with an immunotherapy, is provided.

As described above, numerous embodiments are further provided that can be applied to any aspect of the present invention and/or combined with any other embodiment described herein. For example, in one embodiment, the agent described herein decreases the copy number, the expression level, and/or the activity of *HLA-E*, *Pdia3*, *Calr*, *Tap1*, *Tap2*, *Tapbp*, and/or *Erap1*. In another embodiment, the agent described herein selectively decreases the catalytic activity and/or the substrate binding activity of HLA-E. In still another embodiment, the agent is a small molecule inhibitor, CRISPR single-guide RNA (sgRNA), RNA interfering agent, antisense oligonucleotide, peptide or peptidomimetic inhibitor, aptamer, or intrabody. In yet another embodiment, the RNA interfering agent is a small interfering RNA (siRNA), CRISPR RNA (crRNA), a small hairpin RNA (shRNA), a microRNA (miRNA), or a piwi-interacting RNA (piRNA). In one embodiment, the RNA interfering agent is a CRISPR single-guide RNA (sgRNA). In another embodiment, the sgRNA comprises a nucleic acid sequence selected from the group consisting of nucleic acid sequence listed in Table 2. In still another embodiment, the agent comprises an intrabody, or an antigen binding fragment thereof, which specifically binds to HLA-E and/or a substrate of HLA-E. In yet another embodiment, the intrabody, or antigen binding fragment thereof, is murine, chimeric, humanized, composite, or human. In one embodiment, the intrabody, or antigen binding fragment thereof, is detectably labeled, comprises an effector domain, comprises an Fc domain, and/or is selected from the group consisting of Fv, Fav, F(ab')₂, Fab', dsFv, scFv, sc(Fv)₂, and diabodies fragments. In another embodiment, the intrabody, or antigen binding fragment thereof, is conjugated to a cytotoxic agent. In still another embodiment, the cytotoxic agent is selected from the group consisting of a chemotherapeutic agent, a biologic agent, a toxin, and a radioactive isotope. In yet another embodiment, the agent described herein increases the sensitivity of the cancer cells to an immunotherapy. In one embodiment, the cancer cells are contacted with

10
15
20
25
30

an immunotherapy and/or a cancer therapy before, after, or concurrently with the agent. In another embodiment,

the immunotherapy comprises an anti-cancer vaccine and/or virus. In still another embodiment, the immunotherapy is cell-based. In yet another embodiment, the

- 5 immunotherapy inhibits an immune checkpoint. In one embodiment, the immune checkpoint is selected from the group consisting of CTLA-4, PD-1, VISTA, B7-H2, B7-H3, PD-L1, B7-H4, B7-H6, ICOS, HVEM, PD-L2, CD160, gp49B, PIR-B, KIR family receptors, TIM-1, TIM-3, TIM-4, LAG-3, GITR, 4-1BB, OX-40, BTLA, SIRP, CD47, CD48, 2B4 (CD244), B7.1, B7.2, ILT-2, ILT-4, TIGIT, HHLA2, butyrophilins, IDO,
- 10 CD39, CD73 and A2aR. In one embodiment, the cancer therapy is selected from the group consisting of radiation, a radiosensitizer, an immunogenic chemotherapy (*e.g.*, 5'-fluorouracil, anthracyclines, such as doxorubicin, and the platinum drug, oxaliplatin) that induce interferon production by the cancer cells or at the site of a tumor, interferon, an interferon-inducing agent, a topical inflammatory agent, and a topical TLR agonist. In
- 15 another embodiment, the biomarker comprises a nucleic acid sequence having at least 95% identity to a nucleic acid sequence listed in Table 1 and/or encodes an amino acid sequence having at least 95% identity to an amino acid sequence listed in Table 1. In still another embodiment, the HLA-E is In still another embodiment, a human HLA-E, a mouse H2-T23, a chimeric HLA-E, or a fusion HLA-E. In one embodiment, the HLA-E described herein
- 20 comprises an amino acid sequence listed in Table 1. In another embodiment, the amino acid sequence described herein. In still another embodiment, is selected from the group consisting of SEQ ID Nos: 2, 4, 6, 8, 10, 12, and 14. In still another embodiment, the HLA-E described herein is encoded by a nucleic acid sequence listed in Table 1. In yet another embodiment, the nucleic acid sequence is selected from the group consisting of SEQ ID
- 25 Nos: 1, 3, 5, 7, 9, 11, and 13. In another embodiment, the agent reduces the number of proliferating cells in the cancer and/or reduces the volume or size of a tumor comprising the cancer cells. In another embodiment, the agent increases the sensitivity of the cancer to the immunotherapy. In still another embodiment, the immunotherapy is T-cell-mediated. In yet another embodiment, the agent increases the amount of CD8⁺ T cells in a tumor
- 30 comprising the cancer cells. In another embodiment, the cancer described herein is selected from the group consisting of melanoma and colorectal cancer. In another embodiment, the cancer is melanoma. In still another embodiment, the agent described herein is administered in a pharmaceutically acceptable formulation.

In still another aspect, a method of determining whether a subject afflicted with a cancer or at risk for developing a cancer would benefit from inhibiting the copy number, amount, and/or activity of at least one biomarker listed in Table 1 is provided, the method comprising a) obtaining a biological sample from the subject; b) determining the copy number, amount, and/or activity of at least one biomarker listed in Table 1; c) determining the copy number, amount, and/or activity of the at least one biomarker in a control; and d) comparing the copy number, amount, and/or activity of the at least one biomarker detected in steps b) and c); wherein the presence of, or a significant increase in, the copy number, amount, and/or activity of, the at least one biomarker listed in Table 1 in the subject sample relative to the control copy number, amount, and/or activity of the at least one biomarker indicates that the subject afflicted with the cancer or at risk for developing the cancer would benefit from inhibiting the copy number, amount, and/or activity of the at least one biomarker listed in Table 1.

As described above, numerous embodiments are further provided that can be applied to any aspect of the present invention and/or combined with any other embodiment described herein. In one embodiment, the method described herein further comprises recommending, prescribing, or administering an agent that inhibits the at least one biomarker listed in Table 1 if the cancer is determined to benefit from the agent. In another embodiment, the method described herein further comprises administering at least one additional cancer therapy. In still another embodiment, the method described herein further comprises recommending, prescribing, or administering cancer therapy other than an agent that inhibits the at least one biomarker listed in Table 1 if the cancer is determined to not benefit from the agent. In yet another embodiment, the cancer therapy is selected from the group consisting of immunotherapy, targeted therapy, chemotherapy, radiation therapy, hormonal therapy, an anti-cancer vaccine, an anti-cancer virus, a checkpoint inhibitor, a radiosensitizer, an immunogenic chemotherapy (*e.g.*, 5'-fluorouracil, anthracyclines, such as doxorubicin, and the platinum drug, oxaliplatin) that induce interferon production by the cancer cells or at the site of a tumor, interferon, an interferon-inducing agent, a topical inflammatory agent, and a topical TLR agonist. In another embodiment, the control sample described herein is determined from a cancerous or non-cancerous sample from either the patient or a member of the same species to which the patient belongs. In still another embodiment, the control sample comprises cells.

In yet another aspect, a method for predicting the clinical outcome of a subject afflicted with a cancer expressing one or more biomarkers listed in Table 1 or a fragment thereof is provided, the method comprising a) determining the copy number, amount, and/or activity of at least one biomarker listed in Table 1 in a subject sample; b) determining the copy number, amount, and/or activity of the at least one biomarker in a control having a good clinical outcome; and c) comparing the copy number, amount, and/or activity of the at least one biomarker in the subject sample and in the control, wherein the presence of, or a significant increase in, the copy number, amount, and/or activity of, the at least one biomarker listed in Table 1 in the subject sample as compared to the copy number, amount and/or activity in the control, is an indication that the subject has a poor clinical outcome.

In another aspect, a method for monitoring the progression of a cancer in a subject, wherein the subject is administered a therapeutically effective amount of an agent that inhibits the copy number, amount, and/or activity of at least one biomarker listed in Table 1, is provided, the method comprising a) detecting in a subject sample at a first point in time the copy number, amount, and/or activity of at least one biomarker listed in Table 1; b) repeating step a) at a subsequent point in time; and c) comparing the amount or activity of at least one biomarker listed in Table 1 detected in steps a) and b) to monitor the progression of the cancer in the subject.

In still another aspect, a method of assessing the efficacy of an agent that inhibits the copy number, amount, and/or activity of at least one biomarker listed in Table 1 for treating a cancer in a subject is provided, comprising a) detecting in a subject sample at a first point in time the copy number, amount, and/or or activity of at least one biomarker listed in Table 1; b) repeating step a) during at least one subsequent point in time after administration of the agent; and c) comparing the copy number, amount, and/or activity detected in steps a) and b), wherein the absence of, or a significant decrease in, the copy number, amount, and/or activity of, the at least one biomarker listed in Table 1, in the subsequent sample as compared to the copy number, amount, and/or activity in the sample at the first point in time, indicates that the agent treats the cancer in the subject.

As described above, numerous embodiments are further provided that can be applied to any aspect of the present invention and/or combined with any other embodiment described herein. For example, in one embodiment, between the first point in time and the subsequent point in time, as described herein, the subject has undergone treatment,

completed treatment, and/or is in remission for the cancer. In another embodiment, the cancer treatment described herein is selected from the group consisting of immunotherapy, targeted therapy, chemotherapy, radiation therapy, hormonal therapy, an anti-cancer vaccine, an anti-cancer virus, a checkpoint inhibitor, a radiosensitizer, an immunogenic chemotherapy (e.g., 5'-fluorouracil, anthracyclines, such as doxorubicin, and the platinum drug, oxaliplatin) that induce interferon production by the cancer cells or at the site of a tumor, interferon, an interferon-inducing agent, a topical inflammatory agent, and a topical TLR agonist. In still another embodiment, the first and/or at least one subsequent sample is selected from the group consisting of *ex vivo* and *in vivo* samples. In yet another embodiment, the first and/or at least one subsequent sample is a portion of a single sample or pooled samples obtained from the subject. In one embodiment, the sample described herein comprises cells, serum, peritumoral tissue, and/or intratumoral tissue obtained from the subject. In another embodiment, the one or more biomarkers listed in Table 1 comprise *HLA-E*, *Pdia3*, *Calr*, *Tap1*, *Tap2*, *Tapbp*, and/or *Erap1*. In still another embodiment, the one or more biomarkers described herein comprise a nucleic acid sequence having at least 95% identity to a nucleic acid sequence listed in Table 1 and/or encodes an amino acid sequence having at least 95% identity to an amino acid sequence listed in Table 1. In yet another embodiment, the HLA-E is a human HLA-E, a mouse H2-T23, a chimeric HLA-E, or a fusion HLA-E. In one embodiment, the HLA-E comprises an amino acid sequence listed in Table 1, such as an amino acid sequence selected from the group consisting of SEQ ID Nos: 2, 4, 6, 8, 10, 12, and 14. In another embodiment, the HLA-E is encoded by a nucleic acid sequence listed in Table 1, such as a nucleic acid sequence is selected from the group consisting of SEQ ID Nos: 1, 3, 5, 7, 9, 11, and 13. In still another embodiment, the agent described herein reduces the number of proliferating cells in the cancer and/or reduces the volume or size of a tumor comprising the cancer cells. In another embodiment, the agent described herein increases the sensitivity of the cancer to the immunotherapy. In still another embodiment, the immunotherapy is T-cell-mediated. In yet another embodiment, the agent described herein increases the amount of CD8+ T cells in a tumor comprising the cancer cells. In another embodiment, the cancer described herein is selected from the group consisting of melanoma and colorectal cancer. In another embodiment, the cancer is melanoma. In still another embodiment, the agent described herein is administered in a pharmaceutically acceptable formulation. In yet another embodiment, the subject described herein is an animal model of the cancer. In another embodiment, the

animal model is a mouse model. In still another embodiment, the subject is a mammal. In yet another embodiment, the mammal is a mouse or a human. In one embodiment, the mammal is a human.

5 **Brief Description of the Drawings**

Figure 1 includes 6 panels, identified as panels A, B, C, D, E, and F, which show that *in vivo* pooled loss-of-function screening using CRISPR/Cas9 in tumor cells recovers known mediators of immune evasion. Panel A shows a schematic diagram of the *in vivo* screening system using the B16 transplantable tumor model. Tumor volumes (in mm³) were compared under each conditions, averaged for each group at each time point (Panel B, 10 left) or for individual animals on the day of sacrifice (Panel B, right). Bars represent means, while whiskers represent standard deviation. Enrichment analysis was carried out using a hypergeometric test to show functional classes of genes, (from the Gene Ontology Consortium database (GO)) targeted by sgRNAs, that were enriched or depleted in tumors 15 in animals, including animals treated with irradiated tumor cell vaccine (GVAX) and anti-PD-1 antibody and the TCR $\alpha^{-/-}$ animals (Panel C). Frequency histogram (Panel D, top) and collapsed histograms (Panel D, middle) of enrichment or depletion (normalized as Z scores) are shown for all 9,992 sgRNAs screened. Enrichment/depletion scores are averaged from 10 mice per condition. sgRNAs targeting PD-L1 are indicated by the red lines (Panel D, 20 middle). PD-L1 expression is compared among Cas9-expressing B16 tumor cells transfected with one of the four sgRNAs targeting PD-L1 (red) or a control sgRNA (grey) (Panel D, bottom). Similar to Panel D, Panel E shows the depletion of CD47 by its specific sgRNAs (indicated in red (top and middle) and CD47 expression after CRISPR editing with sgRNAs targeting CD47 (bottom). Panel F compares tumor volumes over time between 25 CD47 null (red) and control (grey) tumors growing in mice treated with GVAX and PD-1 blockade (average and standard error of the mean; n = 10 animals per group). ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

Figure 2 includes 6 panels, identified as panels A, B, C, D, E, and F, which show the performance analysis of the screening in Figure 1. Panel A shows Western blot of B16 30 cell lysate for Cas9 and β -ACTIN with or without transduction with a lentiviral vector encoding Cas9. A pie chart shows the fraction of genes targeted in the screening in each of the GO term categories indicated (Panel B). Two-dimensional histograms show the pairwise distribution of sgRNAs abundance (averaged for each condition) (Panel C).

Saturation analysis of animal replicates from the three *in vivo* screening conditions is shown in Panel D. Pearson correlations are calculated for the library distribution in one animal vs. any other animal, then for two animals averaged versus any other two averaged, and so on. Saturation approaches $r = 0.95$. A matrix of the Pearson correlations of the library distribution from one animal compared to any other animal for B16 Pool 1 is shown (Panel E). Expression of CD47 by B16 cells transfected with either CD47-targeting (red) or control (grey) sgRNA is compared (Panel F).

Figure 3 includes 6 panels, identified as panels A, B, C, D, E, and F, which show that deletion of *H2-T23* or its related genes sensitizes tumor cells to immunotherapy. Panel A shows frequency histogram (top) and collapsed histogram (below) of enrichment or depletion (normalized as Z scores) for sgRNAs in GVAX + PD-1 blockade-treated mice relative to $\text{TCR}\alpha^{-/-}$ mice. sgRNAs targeting *H2-T23*, *Pdia3*, *Calr*, *Tap2*, *Tap1*, *Erap1*, and *Tapbp* are indicated in red. Panel B compares the expression of *H2-T23* by B16 cells transfected with the indicated *H2-T23*-targeting (red) or control (grey) sgRNA and Cas9. Representative flow plots show frequencies of *H2-T23* null and control B16 tumor cells in mixture *in vitro* or *in vivo* (Panel C). The cell number of either *H2-T23* null or control B16 cells grown in GVAX + PD-1 blockade-treated wild-type mice, relative to the cell number of those cells grown in $\text{TCR}\alpha^{-/-}$ mice, is shown in change in the ratios ($\log_2$ normalized fold change) (mean and standard deviation; $n = 5$ animals per group) (Panel D). Panel E compares tumor volume averaged for each group at each time point for *H2-T23* null (red) or control (grey) B16 cells growing in mice treated with GVAX and PD-1 blockade (mean and standard error of the mean; $n = 10$ animals per group). Panel F shows the survival analysis for the groups shown in panel D. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

Figure 4 shows the staining of HLA-E in different cancer and healthy tissues. HLA-E is generally overexpressed in cancer and the tissues having strong HLA-E expression include colorectal, lung, gastric, melanoma, vulva, HNSCC, breast, ovarium, etc. The figure and information is adapted from Gooden *et al.* (2011) *Proc. Natl. Acad. Sci. USA* 108:10656-10661; van Esch *et al.* (2014) *Int. J. Cancer*. 135:830-842; Yazdi *et al.* (2016) *Oncotarget* 7:3477-3488.

Figure 5 shows the interaction between tumor cells, NK cells, and CD8⁺ T cells through the ligand-receptor recognition between HLA-E and NKG2A.

Figure 6 shows the NKG2A expression on human tumor-infiltrating lymphocytes (TILs).

Figure 7 includes 2 panels, identified as panels A and B, which show the co-expression of PD-1 and NKG2A on human TILs. Panel A shows the flow cytometry analysis of TILs (in head and neck squamous cell carcinoma (HNSCC)). Panel B shows the Mass cytometry analysis of TILs.

Figure 8 shows the co-expression of multiple inhibitory receptors on mouse CD8 T cells. The mice model used here is TC-1 in C57BL/6 mice.

Figure 9 includes 3 panels, identified as panels A, B, and C, which show that therapeutic vaccination increases NKG2A and H2-T23 (Qa-1) expression in TC-1 C57BL/6 mice. H2-T23 (Qa-1) expression was also seen in B16 mice model.

Figure 10 includes 2 panels, identified as panels A and B, which show enhanced efficacy of vaccination in absence of H2-T23 (Qa-1).

Figure 11 includes 5 panels, identified as panels A, B, C, D, and E, which show that NKG2A-blocking monoclonal antibody recapitulates genetic model. Panel A shows the timeline of vaccination and administration of NKG2A mAb. Panel B compares mice survival rates after different treatments. Panel C summarizes the tumor size through time for each mouse without treatment. Panel D summarizes the tumor size through time for each mouse treated with vaccination and control mAb. Panel E summarizes the tumor size through time for each mouse treated with vaccination and anti-NKG2A mAb.

Figure 12 includes 3 panels, identified as panels A, B, and C, which summarize NKG2A expression on TILs (panel A) or CD8 T cells after PD-1 treatment (panel B) and compare the mice survival rates with different treatments.

The data in Figures 4-12 are adapted from a presentation titled “NKG2A checkpoint receptor expression on tumor-infiltrating CD8⁺ T cells restrains efficacy of immunotherapy” (Thorald van Hall *et al.*, AACR Annual Meeting 2017 dated April 3, 2017).

Detailed Description of the Invention

It has been determined herein that negative regulators of one or more biomarkers listed in Table 1, such as one or more proteins modulating antigen processing and presentation (*e.g.*, *HLA-E*, *Pdia3*, *Calr*, *Tap1*, *Tap2*, *Tapbp*, *Erap1*, etc.) can be used to increase interferon sensing by tumor cells and augment tumor immunity and

immunotherapies. Thus, the instant disclosure provides at least a method of treating cancers, *e.g.*, those cancer types otherwise not responsive or weakly responsive to immunotherapies, with a combination of a negative regulator of one or more biomarkers listed in Table 1, such as one or more proteins modulating antigen processing and presentation (*e.g.*, *HLA-E*, *Pdia3*, *Calr*, *Tap1*, *Tap2*, *Tapbp*, *Erap1*, etc.) and another immunotherapy. The results described herein are unexpected given that analyses of such regulator function has heretofore been largely confined to hematopoietic cells and not examined in cancer cells, as well as the fact that modulating sensitivity to interferon signaling is critical for immunotherapy effects rather than simply modulating interferon availability since interferon therapy is known to not significantly augment immunotherapy effects. Accordingly, the present invention provides exemplary RNA interfering agents inhibiting such regulators, which may be used in the combination therapy and other methods described herein, such as agents that inhibit the function and/or its ability of one or more biomarkers listed in Table 1 to interact/bind to its substrates described herein, or by increasing its degradation and/or stability and/or interaction/binding to its inhibitors. Similarly, methods of screening for inhibitors of such regulators and methods of diagnosing, prognosing, and monitoring cancer involving such inhibitors/immunotherapy combination therapies are provided.

I. Definitions

The articles “a” and “an” are used herein to refer to one or to more than one (*i.e.* to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

The term “altered amount” or “altered level” refers to increased or decreased copy number (*e.g.*, germline and/or somatic) of a biomarker nucleic acid, *e.g.*, increased or decreased expression level in a cancer sample, as compared to the expression level or copy number of the biomarker nucleic acid in a control sample. The term “altered amount” of a biomarker also includes an increased or decreased protein level of a biomarker protein in a sample, *e.g.*, a cancer sample, as compared to the corresponding protein level in a normal, control sample. Furthermore, an altered amount of a biomarker protein may be determined by detecting posttranslational modification such as methylation status of the marker, which may affect the expression or activity of the biomarker protein.

The amount of a biomarker in a subject is “significantly” higher or lower than the normal amount of the biomarker, if the amount of the biomarker is greater or less, respectively, than the normal level by an amount greater than the standard error of the assay employed to assess amount, and preferably at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 150%, 200%, 300%, 350%, 400%, 500%, 600%, 700%, 800%, 900%, 1000% or than that amount. Alternately, the amount of the biomarker in the subject can be considered “significantly” higher or lower than the normal amount if the amount is at least about two, and preferably at least about three, four, or five times, higher or lower, respectively, than the normal amount of the biomarker. Such “significance” can also be applied to any other measured parameter described herein, such as for expression, inhibition, cytotoxicity, cell growth, and the like.

The term “altered level of expression” of a biomarker refers to an expression level or copy number of the biomarker in a test sample, *e.g.*, a sample derived from a patient suffering from cancer, that is greater or less than the standard error of the assay employed to assess expression or copy number, and is preferably at least twice, and more preferably three, four, five or ten or more times the expression level or copy number of the biomarker in a control sample (*e.g.*, sample from a healthy subjects not having the associated disease) and preferably, the average expression level or copy number of the biomarker in several control samples. The altered level of expression is greater or less than the standard error of the assay employed to assess expression or copy number, and is preferably at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 150%, 200%, 300%, 350%, 400%, 500%, 600%, 700%, 800%, 900%, 1000% or more times the expression level or copy number of the biomarker in a control sample (*e.g.*, sample from a healthy subjects not having the associated disease) and preferably, the average expression level or copy number of the biomarker in several control samples. In some embodiments, the level of the biomarker refers to the level of the biomarker itself, the level of a modified biomarker (*e.g.*, phosphorylated biomarker), or to the level of a biomarker relative to another measured variable, such as a control (*e.g.*, phosphorylated biomarker relative to an unphosphorylated biomarker).

The term “altered activity” of a biomarker refers to an activity of the biomarker which is increased or decreased in a disease state, *e.g.*, in a cancer sample, as compared to the activity of the biomarker in a normal, control sample. Altered activity of the biomarker may be the result of, for example, altered expression of the biomarker, altered protein level

of the biomarker, altered structure of the biomarker, or, *e.g.*, an altered interaction with other proteins involved in the same or different pathway as the biomarker or altered interaction with transcriptional activators or inhibitors.

5 The term “altered structure” of a biomarker refers to the presence of mutations or allelic variants within a biomarker nucleic acid or protein, *e.g.*, mutations which affect expression or activity of the biomarker nucleic acid or protein, as compared to the normal or wild-type gene or protein. For example, mutations include, but are not limited to substitutions, deletions, or addition mutations. Mutations may be present in the coding or non-coding region of the biomarker nucleic acid.

10 Unless otherwise specified here within, the terms “antibody” and “antibodies” refers to antigen-binding portions adaptable to be expressed within cells as “intracellular antibodies.” (Chen *et al.* (1994) *Human Gene Ther.* 5:595-601). Methods are well-known in the art for adapting antibodies to target (*e.g.*, inhibit) intracellular moieties, such as the use of single-chain antibodies (scFvs), modification of immunoglobulin VL domains for
15 hyperstability, modification of antibodies to resist the reducing intracellular environment, generating fusion proteins that increase intracellular stability and/or modulate intracellular localization, and the like. Intracellular antibodies can also be introduced and expressed in one or more cells, tissues or organs of a multicellular organism, for example for prophylactic and/or therapeutic purposes (*e.g.*, as a gene therapy) (see, at least PCT Publs.
20 WO 08/020079, WO 94/02610, WO 95/22618, and WO 03/014960; U.S. Pat. No. 7,004,940; Cattaneo and Biocca (1997) *Intracellular Antibodies: Development and Applications* (Landes and Springer-Verlag publs.); Kontermann (2004) *Methods* 34:163-170; Cohen *et al.* (1998) *Oncogene* 17:2445-2456; Auf der Maur *et al.* (2001) *FEBS Lett.* 508:407-412; Shaki-Loewenstein *et al.* (2005) *J. Immunol. Meth.* 303:19-39).

25 Antibodies may be polyclonal or monoclonal; xenogeneic, allogeneic, or syngeneic; or modified forms thereof (*e.g.* humanized, chimeric, etc.). Antibodies may also be fully human. Preferably, antibodies of the present invention bind specifically or substantially specifically to a biomarker polypeptide or fragment thereof. The terms “monoclonal antibodies” and “monoclonal antibody composition”, as used herein, refer to a population
30 of antibody polypeptides that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of an antigen, whereas the term “polyclonal antibodies” and “polyclonal antibody composition” refer to a population of antibody polypeptides that contain multiple species of antigen binding sites capable of interacting

with a particular antigen. A monoclonal antibody composition typically displays a single binding affinity for a particular antigen with which it immunoreacts.

Antibodies may also be “humanized”, which is intended to include antibodies made by a non-human cell having variable and constant regions which have been altered to more closely resemble antibodies that would be made by a human cell. For example, by altering the non-human antibody amino acid sequence to incorporate amino acids found in human germline immunoglobulin sequences. The humanized antibodies of the present invention may include amino acid residues not encoded by human germline immunoglobulin sequences (*e.g.*, mutations introduced by random or site-specific mutagenesis *in vitro* or by somatic mutation *in vivo*), for example in the CDRs. The term “humanized antibody”, as used herein, also includes antibodies in which CDR sequences derived from the germline of another mammalian species, such as a mouse, have been grafted onto human framework sequences.

The term “assigned score” refers to the numerical value designated for each of the biomarkers after being measured in a patient sample. The assigned score correlates to the absence, presence or inferred amount of the biomarker in the sample. The assigned score can be generated manually (*e.g.*, by visual inspection) or with the aid of instrumentation for image acquisition and analysis. In certain embodiments, the assigned score is determined by a qualitative assessment, for example, detection of a fluorescent readout on a graded scale, or quantitative assessment. In one embodiment, an “aggregate score,” which refers to the combination of assigned scores from a plurality of measured biomarkers, is determined. In one embodiment the aggregate score is a summation of assigned scores. In another embodiment, combination of assigned scores involves performing mathematical operations on the assigned scores before combining them into an aggregate score. In certain, embodiments, the aggregate score is also referred to herein as the “predictive score.”

The term “biomarker” refers to a measurable entity of the present invention that has been determined to be predictive of effects of combinatorial therapies comprising one or more inhibitors of one or more biomarkers listed in Table 1, such as one or more proteins modulating antigen processing and presentation regulators (*e.g.*, *HLA-E*, *Pdia3*, *Calr*, *Tap1*, *Tap2*, *Tapbp*, *Erap1*, etc., or the corresponding proteins) and immunotherapy (*e.g.*, immune checkpoint inhibitors) on a cancer. Biomarkers can include, without limitation, nucleic acids and proteins, including those shown in the Tables, the Examples, the Figures, and otherwise described herein. As described herein, any relevant characteristic of a biomarker

can be used, such as the copy number, amount, activity, location, modification (*e.g.*, phosphorylation), and the like.

A “blocking” antibody or an antibody “antagonist” is one which inhibits or reduces at least one biological activity of the antigen(s) it binds. In certain embodiments, the
 5 blocking antibodies or antagonist antibodies or fragments thereof described herein substantially or completely inhibit a given biological activity of the antigen(s).

The term “body fluid” refers to fluids that are excreted or secreted from the body as well as fluids that are normally not (*e.g.* amniotic fluid, aqueous humor, bile, blood and blood plasma, cerebrospinal fluid, cerumen and earwax, cowper’s fluid or pre-ejaculatory
 10 fluid, chyle, chyme, stool, female ejaculate, interstitial fluid, intracellular fluid, lymph, menses, breast milk, mucus, pleural fluid, pus, saliva, sebum, semen, serum, sweat, synovial fluid, tears, urine, vaginal lubrication, vitreous humor, vomit).

The terms “cancer” or “tumor” or “hyperproliferative” refer to the presence of cells possessing characteristics typical of cancer-causing cells, such as uncontrolled proliferation,
 15 immortality, metastatic potential, rapid growth and proliferation rate, and certain characteristic morphological features. Unless otherwise stated, the terms include metaplasias. In some embodiments, such cells exhibit such characteristics in part or in full due to the expression and activity of signaling pathways regulated by one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and
 20 presentation described herein (*e.g.*, *HLA-E*, *Pdia3*, *Calr*, *Tap1*, *Tap2*, *Tapbp*, *Erap1*, etc.) (*e.g.*, NF-kappaB Signaling pathway, MAP kinase pathway, JAK-STAT signaling pathway, or other signaling pathways involving receptor tyrosine kinases, non-receptor tyrosine kinases, Src family kinases, and/or signal transducer and activator of transcription (STAT) proteins). In certain embodiments, the cancer cells are capable of responding to interferon
 25 because they express functional proteins of the type I interferon signaling pathway and/or type II interferon signaling pathway, such as those shown in panel E of Figure 11. In some embodiments, the cancer cells described herein are not sensitive to at least one of immunotherapies. Such insensitivity, without limitation, may be related to the inactivation or decreased activation, compared to control cells (*e.g.*, normal and/or wild-type non-cancer
 30 cells, and/or cancer cells without this insensitivity to immunotherapies), of interferon signaling (*e.g.*, IFN γ signaling) in such cancer cells and/or other surrounding cells and/or cells localized near to such cancer cells. Such inactivation or decreased activation of interferon signaling, without limitation, may be related to the inhibition of interferon

signaling by regulators of antigen processing and presentation described herein (*e.g.*, by overexpression and/or gain-of-function of the gene, RNA transcript, and/or protein, and/or by reduced expression and/or loss-of-function of endogenous negative regulator(s) for such regulators). In some embodiments, the cancer cells are treatable with an agent capable of antagonizing regulators of antigen processing and presentation described herein, such as inhibiting expression and/or function, as described herein. An exemplary agent, without limitation, may relieve the inhibition of signaling by such regulators described herein in such cancer cells and/or other cells surrounding or localized near such cancer cells, thus restoring the IFN γ signaling and the sensitivity of such cancer cells to immunotherapies, especially those immunotherapies related to interferon signaling pathways. In some embodiments, the treatment with the agent antagonizing such regulators of antigen processing and presentation as described herein would increase IFN γ signaling in such cancer cells, compared to pre-treatment situations, or would restore IFN γ signaling in such cancer cells to at least comparable to the levels in control cells, so that such cancer cells would regain sensitivity to immunotherapies. The term “interferon signaling” or “IFN γ signaling” used herein refers to any cell signaling downstream and/or related to the interaction of interferon (*e.g.*, IFN γ) and their receptor(s).

Cancer cells are often in the form of a tumor, but such cells may exist alone within an animal, or may be a non-tumorigenic cancer cell, such as a leukemia cell. As used herein, the term “cancer” includes premalignant as well as malignant cancers. Cancers include, but are not limited to, B cell cancer, *e.g.*, multiple myeloma, Waldenström's macroglobulinemia, the heavy chain diseases, such as, for example, alpha chain disease, gamma chain disease, and mu chain disease, benign monoclonal gammopathy, and immunocytic amyloidosis, melanomas, breast cancer, lung cancer, bronchus cancer, colorectal cancer, prostate cancer, pancreatic cancer, stomach cancer, ovarian cancer, urinary bladder cancer, brain or central nervous system cancer, peripheral nervous system cancer, esophageal cancer, cervical cancer, uterine or endometrial cancer, cancer of the oral cavity or pharynx, liver cancer, kidney cancer, testicular cancer, biliary tract cancer, small bowel or appendix cancer, salivary gland cancer, thyroid gland cancer, adrenal gland cancer, osteosarcoma, chondrosarcoma, cancer of hematologic tissues, and the like. Other non-limiting examples of types of cancers applicable to the methods encompassed by the present invention include human sarcomas and carcinomas, *e.g.*, fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma,

angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, colon carcinoma, colorectal cancer, pancreatic cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, liver cancer, choriocarcinoma, seminoma, embryonal carcinoma, Wilms' tumor, cervical cancer, bone cancer, brain tumor, testicular cancer, lung carcinoma, small cell lung carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, meningioma, melanoma, neuroblastoma, retinoblastoma; leukemias, *e.g.*, acute lymphocytic leukemia and acute myelocytic leukemia (myeloblastic, promyelocytic, myelomonocytic, monocytic and erythroleukemia); chronic leukemia (chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia); and polycythemia vera, lymphoma (Hodgkin's disease and non-Hodgkin's disease), multiple myeloma, Waldenstrom's macroglobulinemia, and heavy chain disease. In some embodiments, cancers are epithelial in nature and include but are not limited to, bladder cancer, breast cancer, cervical cancer, colon cancer, gynecologic cancers, renal cancer, laryngeal cancer, lung cancer, oral cancer, head and neck cancer, ovarian cancer, pancreatic cancer, prostate cancer, or skin cancer. In other embodiments, the cancer is breast cancer, prostate cancer, lung cancer, or colon cancer. In still other embodiments, the epithelial cancer is non-small-cell lung cancer, nonpapillary renal cell carcinoma, cervical carcinoma, ovarian carcinoma (*e.g.*, serous ovarian carcinoma), or breast carcinoma. The epithelial cancers may be characterized in various other ways including, but not limited to, serous, endometrioid, mucinous, clear cell, Brenner, or undifferentiated.

In certain embodiments, the cancer encompasses colorectal cancer (*e.g.*, colorectal carcinoma).

The term "colorectal cancer" as used herein, is meant to include cancer of cells of the intestinal tract below the small intestine (*e.g.*, the large intestine (colon), including the cecum, ascending colon, transverse colon, descending colon, and sigmoid colon, and rectum). Additionally, as used herein, the term "colorectal cancer" is meant to further include cancer of cells of the duodenum and small intestine (jejunum and ileum).

Colorectal cancer also includes neoplastic diseases involving proliferation of a single clone of cells of the colon and includes adenocarcinoma and carcinoma of the colon whether in a primary site or metastasized.

Colorectal cancer (CRC) is the third most commonly diagnosed cancer and ranks
5 second in cancer mortality. Extensive genetic and genomic analysis of human CRC has uncovered germline and somatic mutations relevant to CRC biology and malignant transformation (Fearon *et al.* (1990) *Cell* 61, 759-767). These mutations have been linked to well-defined disease stages from aberrant crypt proliferation or hyperplastic lesions to benign adenomas, to carcinoma *in situ*, and finally to invasive and metastatic disease,
10 thereby establishing a genetic paradigm for cancer initiation and progression. Genetic and genomic instability are catalysts for colon carcinogenesis (Lengauer *et al.* (1998) *Nature* 396:643-649). CRC can present with two distinct genomic profiles that have been termed (i) chromosomal instability neoplasia (CIN), characterized by rampant structural and numerical chromosomal aberrations driven in part by telomere dysfunction (Artandi *et al.* (2000) *Nature* 406:641-645; Fodde *et al.* (2001) *Nat. Rev. Cancer* 1:55-67; Maser and DePinho (2002) *Science* 297:565-569; Rudolph *et al.* (2001) *Nat. Genet.* 28:155-159) and mitotic aberrations (Lengauer *et al.* (1998) *Nature* 396:643-649) and (ii) microsatellite instability neoplasia (MIN), characterized by near diploid karyotypes with alterations at the nucleotide level due to mutations in mismatch repair (MMR) genes (Fishel *et al.* (1993)
20 *Cell* 75:1027-1038; Ilyas *et al.* (1999) *Eur. J. Cancer* 35:335-351; Modrich (1991) *Annu. Rev. Genet.* 25:229-253; Parsons *et al.* (1995) *Science* 268:738-740; Parsons *et al.* (1993) *Cell* 75:1227-1236). Germline MMR mutations are highly penetrant lesions which drive the MIN phenotype in hereditary nonpolyposis colorectal cancers, accounting for 1-5% of CRC cases (de la Chapelle (2004) *Nat. Rev. Cancer* 4:769-780; Lynch and de la Chapelle (1999) *J. Med. Genet.* 36:801-818; Umar *et al.* (2004) *Nat. Rev. Cancer* 4:153-158). While
25 CIN and MIN are mechanistically distinct, their genomic and genetic consequences emphasize the requirement of dominant mutator mechanisms to drive intestinal epithelial cells towards a threshold of oncogenic changes needed for malignant transformation.

A growing number of genetic mutations have been identified and functionally
30 validated in CRC pathogenesis. Activation of the WNT signaling pathway is an early requisite event for adenoma formation. Somatic alterations are present in *APC* in greater than 70% of nonfamilial sporadic cases and appear to contribute to genomic instability and induce the expression of *c-myc* and *Cyclin D1* (Fodde *et al.* (2001) *Nat. Rev. Cancer* 1:55-

67), while activating β -catenin mutations represent an alternative means of WNT pathway deregulation in CRC (Morin (1997) *Science* 275:1787-1790). *K-Ras* mutations occur early in neoplastic progression and are present in approximately 50% of large adenomas (Fearon and Gruber (2001) *Molecular abnormalities in colon and rectal cancer*, ed. J. Mendelsohn, P.H., M. Israel, and L. Liotta, W.B. Saunders, Philadelphia). The *BRAF* serine/threonine kinase and *PIK3CA* lipid kinase are mutated in 5-18% and 28% of sporadic CRCs, respectively (Samuels *et al.* (2004) *Science* 304:554; Davies *et al.* (2002) *Nature* 417:949-954; Rajagopalan *et al.* (2002) *Nature* 418:934; Yuen *et al.* (2002) *Cancer Res.* 62:6451-6455). *BRAF* and *K-ras* mutations are mutually exclusive in CRC, suggesting over-lapping oncogenic activities (Davies *et al.* (2002) *Nature* 417:949-954; Rajagopalan *et al.* (2002) *Nature* 418:934). Mutations associated with CRC progression, specifically the adenoma-to-carcinoma transition, target the *TP53* and the TGF- β pathways (Markowitz *et al.* (2002) *Cancer Cell* 1:233-236). Greater than 50% of CRCs harbor *TP53* inactivating mutations (Fearon and Gruber (2001) *Molecular abnormalities in colon and rectal cancer*, ed. J. Mendelsohn, P.H., M. Israel, and L. Liotta, W.B. Saunders, Philadelphia) and 30% of cases possess *TGF β -RII* mutations (Markowitz (2000) *Biochim. Biophys. Acta* 1470:M13-M20; Markowitz *et al.* (1995) *Science* 268:1336-1338). MIN cancers consistently inactivate *TGF β -RII* by frameshift mutations, whereas CIN cancers target the pathway via inactivating somatic mutations in the *TGF β -RII* kinase domain (15%) or in the downstream signaling components of the pathway, including *SMAD4* (15%) or *SMAD2* (5%) transcription factors (Markowitz (2000) *Biochim. Biophys. Acta* 1470:M13-M20). In some embodiments, the colorectal cancer is microsatellite instable (MSI) colorectal cancer (Llosa *et al.* (2014) *Cancer Disc.* CD-14-0863; published online Oct. 30, 2014). MSI represents about 15% of sporadic CRC and about 5-6% of stage IV CRCs. MSI is caused by epigenetic silencing or mutation of DNA mismatch repair genes and typically presents with lower stage disease than microsatellite stable subset (MSS) CRC. MSI highly express immune checkpoints, such as PD-1, PD-L1, CTLA-4, LAG-3, and IDO. In other embodiments, the colorectal cancer is MSS CRC.

In certain embodiments, the cancer encompasses melanoma. The term “melanoma” as used herein, is generally meant to include cancers that develop from the pigment-containing cells, known as melanocytes, in the basal layer of the epidermis. Melanomas typically occur in the skin but may rarely occur in the mouth, intestines, or eye. In women they most commonly occur on the legs, while in men they are most common on the back.

Sometimes they develop from a mole with concerning changes including an increase in size, irregular edges, change in color, itchiness, or skin breakdown. Thus, the term “melanoma” also includes cancers developing from these cells, tissues, and organs.

Melanomas are among the most dangerous forms of skin cancer and develop when
5 unrepaired DNA damage to skin cells (most often caused by ultraviolet radiation from
sunshine or tanning beds) triggers gene mutations that lead the skin cells to multiply rapidly
and form malignant tumors. The primary cause of melanoma is ultraviolet light (UV)
exposure in those with low levels of skin pigment. Melanomas often resemble moles; some
develop from moles. Those with many moles, a history of affected family members, and
10 who have poor immune function are at greater risk. A number of rare genetic defects such
as xeroderma pigmentosum also increase risk (Azoury and Lange, 2014 *Surg Clin North
Am.* 2014 94:945-962).

Melanoma can be divided into different types, including, at least, lentigo maligna,
lentigo maligna melanoma, superficial spreading melanoma, acral lentiginous melanoma,
15 mucosal melanoma, nodular melanoma, polypoid melanoma, desmoplastic melanoma,
amelanotic melanoma, soft-tissue melanoma, melanoma with small nevus-like cells,
melanoma with features of a Spitz nevus, uveal melanoma, etc. (see James, et al., 2006
Andrews' Diseases of the Skin: clinical Dermatology. Saunders Elsevier. pp. 694–9)

Diagnosis is by biopsy of any concerning skin lesion, including, at least, shave
20 (tangential) biopsy, punch biopsy, incisional and excisional biopsies, “optical” biopsies
(e.g., by reflectance confocal microscopy (RCM)), fine needle aspiration (FNA) biopsy,
surgical lymph node biopsy, sentinel lymph node biopsy, etc. In addition, visual inspection
may also be used for diagnosis, such as a popular method for the signs and symptoms of
melanoma as mnemonic “ABCDE”: Asymmetrical skin lesion, Border of the lesion is
25 irregular, Color: melanomas usually have multiple colors, Diameter: moles greater than 6
mm are more likely to be melanomas than smaller moles, and Enlarging: Enlarging or
evolving. Another method as the “ugly duckling sign” is also known in the art (Mascaro
and Mascaro, 1998 *Arch Dermatol.* 134: 1484-1485).

Treatment of melanoma includes surgery, chemotherapy (such as temozolomide,
30 dacarbazine (also termed DTIC), etc.), radiation therapy, oncolytic virotherapy (e.g., see
Forbes *et al.*, 2013 *Front. Genet.* 4:184), and immunotherapy (e.g., interleukin-2 (IL-2),
interferon, etc.). Targeted therapies (e.g., as in Maverakis *et al.*, 2015 *Acta Derm Venereol.*
95: 516-524) may include: 1) adoptive cell therapy (ACT) using TILs immune cells (tumor

infiltrating lymphocytes) isolated from a person's own melanoma tumor). Cells are grown in large numbers in a laboratory and returned to the patient after a treatment that temporarily reduces normal T cells in the patient's body. TIL therapy following lymphodepletion can result in durable complete response in a variety of setups (Besser *et al.*, 2010 *Clin. Cancer Res.* 16:2646-2655); and 2) adoptive transfer of genetically altered (expressing T cell receptors (TCRs)) autologous lymphocytes into patient's lymphocytes, where the altered lymphocytes recognize and bind to the surface of melanoma cells and kill them. Other therapies include, at least, B-Raf inhibitors (such as vemurafenib, see Chapman *et al.*, 2011 *N. Engl. J. Med.* 364:2507-2516) and ipilimumab (alone or in combination with dacarbazine, see, e.g., Robert *et al.* (2011) *N. Engl. J. Med.* 364:2517-2526).

The term “coding region” refers to regions of a nucleotide sequence comprising codons which are translated into amino acid residues, whereas the term “noncoding region” refers to regions of a nucleotide sequence that are not translated into amino acids (e.g., 5' and 3' untranslated regions).

The term “complementary” refers to the broad concept of sequence complementarity between regions of two nucleic acid strands or between two regions of the same nucleic acid strand. It is known that an adenine residue of a first nucleic acid region is capable of forming specific hydrogen bonds (“base pairing”) with a residue of a second nucleic acid region which is antiparallel to the first region if the residue is thymine or uracil. Similarly, it is known that a cytosine residue of a first nucleic acid strand is capable of base pairing with a residue of a second nucleic acid strand which is antiparallel to the first strand if the residue is guanine. A first region of a nucleic acid is complementary to a second region of the same or a different nucleic acid if, when the two regions are arranged in an antiparallel fashion, at least one nucleotide residue of the first region is capable of base pairing with a residue of the second region. Preferably, the first region comprises a first portion and the second region comprises a second portion, whereby, when the first and second portions are arranged in an antiparallel fashion, at least about 50%, and preferably at least about 75%, at least about 90%, or at least about 95% of the nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion. More preferably, all nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion.

The term “control” refers to any reference standard suitable to provide a comparison to the expression products in the test sample. In one embodiment, the control comprises obtaining a “control sample” from which expression product levels are detected and compared to the expression product levels from the test sample. Such a control sample may
5 comprise any suitable sample, including but not limited to a sample from a control cancer patient (can be stored sample or previous sample measurement) with a known outcome; normal tissue or cells isolated from a subject, such as a normal patient or the cancer patient, cultured primary cells/tissues isolated from a subject such as a normal subject or the cancer patient, adjacent normal cells/tissues obtained from the same organ or body location of the
10 cancer patient, a tissue or cell sample isolated from a normal subject, or a primary cells/tissues obtained from a depository. In another preferred embodiment, the control may comprise a reference standard expression product level from any suitable source, including but not limited to housekeeping genes, an expression product level range from normal tissue (or other previously analyzed control sample), a previously determined expression
15 product level range within a test sample from a group of patients, or a set of patients with a certain outcome (for example, survival for one, two, three, four years, etc.) or receiving a certain treatment (for example, standard of care cancer therapy). It will be understood by those of skill in the art that such control samples and reference standard expression product levels can be used in combination as controls in the methods of the present invention. In
20 one embodiment, the control may comprise normal or non-cancerous cell/tissue sample. In another preferred embodiment, the control may comprise an expression level for a set of patients, such as a set of cancer patients, or for a set of cancer patients receiving a certain treatment, or for a set of patients with one outcome versus another outcome. In the former case, the specific expression product level of each patient can be assigned to a percentile
25 level of expression, or expressed as either higher or lower than the mean or average of the reference standard expression level. In another preferred embodiment, the control may comprise normal cells, cells from patients treated with combination chemotherapy, and cells from patients having benign cancer. In another embodiment, the control may also comprise a measured value for example, average level of expression of a particular gene in
30 a population compared to the level of expression of a housekeeping gene in the same population. Such a population may comprise normal subjects, cancer patients who have not undergone any treatment (*i.e.*, treatment naive), cancer patients undergoing standard of care therapy, or patients having benign cancer. In another preferred embodiment, the control

comprises a ratio transformation of expression product levels, including but not limited to determining a ratio of expression product levels of two genes in the test sample and comparing it to any suitable ratio of the same two genes in a reference standard; determining expression product levels of the two or more genes in the test sample and
5 determining a difference in expression product levels in any suitable control; and determining expression product levels of the two or more genes in the test sample, normalizing their expression to expression of housekeeping genes in the test sample, and comparing to any suitable control. In particularly preferred embodiments, the control comprises a control sample which is of the same lineage and/or type as the test sample. In
10 another embodiment, the control may comprise expression product levels grouped as percentiles within or based on a set of patient samples, such as all patients with cancer. In one embodiment a control expression product level is established wherein higher or lower levels of expression product relative to, for instance, a particular percentile, are used as the basis for predicting outcome. In another preferred embodiment, a control expression
15 product level is established using expression product levels from cancer control patients with a known outcome, and the expression product levels from the test sample are compared to the control expression product level as the basis for predicting outcome. As demonstrated by the data below, the methods of the present invention are not limited to use of a specific cut-point in comparing the level of expression product in the test sample to the
20 control.

The “copy number” of a biomarker nucleic acid refers to the number of DNA sequences in a cell (*e.g.*, germline and/or somatic) encoding a particular gene product. Generally, for a given gene, a mammal has two copies of each gene. The copy number can be increased, however, by gene amplification or duplication, or reduced by deletion. For
25 example, germline copy number changes include changes at one or more genomic loci, wherein said one or more genomic loci are not accounted for by the number of copies in the normal complement of germline copies in a control (*e.g.*, the normal copy number in germline DNA for the same species as that from which the specific germline DNA and corresponding copy number were determined). Somatic copy number changes include
30 changes at one or more genomic loci, wherein said one or more genomic loci are not accounted for by the number of copies in germline DNA of a control (*e.g.*, copy number in germline DNA for the same subject as that from which the somatic DNA and corresponding copy number were determined).

The “normal” copy number (*e.g.*, germline and/or somatic) of a biomarker nucleic acid or “normal” level of expression of a biomarker nucleic acid or protein is the activity/level of expression or copy number in a biological sample, *e.g.*, a sample containing tissue, whole blood, serum, plasma, buccal scrape, saliva, cerebrospinal fluid, urine, stool, and bone marrow, from a subject, *e.g.*, a human, not afflicted with cancer, or from a corresponding non-cancerous tissue in the same subject who has cancer.

As used herein, the term “costimulate” with reference to activated immune cells includes the ability of a costimulatory molecule to provide a second, non-activating receptor mediated signal (a “costimulatory signal”) that induces proliferation or effector function. For example, a costimulatory signal can result in cytokine secretion, *e.g.*, in a T cell that has received a T cell-receptor-mediated signal. Immune cells that have received a cell-receptor mediated signal, *e.g.*, via an activating receptor are referred to herein as “activated immune cells.”

The term “determining a suitable treatment regimen for the subject” is taken to mean the determination of a treatment regimen (*i.e.*, a single therapy or a combination of different therapies that are used for the prevention and/or treatment of the cancer in the subject) for a subject that is started, modified and/or ended based or essentially based or at least partially based on the results of the analysis according to the present invention. One example is starting an adjuvant therapy after surgery whose purpose is to decrease the risk of recurrence, another would be to modify the dosage of a particular chemotherapy. The determination can, in addition to the results of the analysis according to the present invention, be based on personal characteristics of the subject to be treated. In most cases, the actual determination of the suitable treatment regimen for the subject will be performed by the attending physician or doctor.

The term “diagnosing cancer” includes the use of the methods, systems, and code of the present invention to determine the presence or absence of a cancer or subtype thereof in an individual. The term also includes methods, systems, and code for assessing the level of disease activity in an individual.

A molecule is “fixed” or “affixed” to a substrate if it is covalently or non-covalently associated with the substrate such that the substrate can be rinsed with a fluid (*e.g.* standard saline citrate, pH 7.4) without a substantial fraction of the molecule dissociating from the substrate.

The term “expression signature” or “signature” refers to a group of one or more coordinately expressed biomarkers related to a measured phenotype. For example, the genes, proteins, metabolites, and the like making up this signature may be expressed in a specific cell lineage, stage of differentiation, or during a particular biological response. The biomarkers can reflect biological aspects of the tumors in which they are expressed, such as the cell of origin of the cancer, the nature of the non-malignant cells in the biopsy, and the oncogenic mechanisms responsible for the cancer. Expression data and gene expression levels can be stored on computer readable media, *e.g.*, the computer readable medium used in conjunction with a microarray or chip reading device. Such expression data can be manipulated to generate expression signatures.

“Homologous” as used herein, refers to nucleotide sequence similarity between two regions of the same nucleic acid strand or between regions of two different nucleic acid strands. When a nucleotide residue position in both regions is occupied by the same nucleotide residue, then the regions are homologous at that position. A first region is homologous to a second region if at least one nucleotide residue position of each region is occupied by the same residue. Homology between two regions is expressed in terms of the proportion of nucleotide residue positions of the two regions that are occupied by the same nucleotide residue. By way of example, a region having the nucleotide sequence 5'-ATTGCC-3' and a region having the nucleotide sequence 5'-TATGGC-3' share 50% homology. Preferably, the first region comprises a first portion and the second region comprises a second portion, whereby, at least about 50%, and preferably at least about 75%, at least about 90%, or at least about 95% of the nucleotide residue positions of each of the portions are occupied by the same nucleotide residue. More preferably, all nucleotide residue positions of each of the portions are occupied by the same nucleotide residue.

The term “immune cell” refers to cells that play a role in the immune response. Immune cells are of hematopoietic origin, and include lymphocytes, such as B cells and T cells; natural killer cells; myeloid cells, such as monocytes, macrophages, eosinophils, mast cells, basophils, and granulocytes.

The term “immunotherapy” or “immunotherapies” refer to any treatment that uses certain parts of a subject’s immune system to fight diseases such as cancer. The subject’s own immune system is stimulated (or suppressed), with or without administration of one or more agent for that purpose. Immunotherapies that are designed to elicit or amplify an immune response are referred to as “activation immunotherapies.” Immunotherapies that

are designed to reduce or suppress an immune response are referred to as “suppression immunotherapies.” Any agent believed to have an immune system effect on the genetically modified transplanted cancer cells can be assayed to determine whether the agent is an immunotherapy and the effect that a given genetic modification has on the modulation of immune response. In some embodiments, the immunotherapy is cancer cell-specific. In some embodiments, immunotherapy can be “untargeted,” which refers to administration of agents that do not selectively interact with immune system cells, yet modulates immune system function. Representative examples of untargeted therapies include, without limitation, chemotherapy, gene therapy, and radiation therapy.

Immunotherapy is one form of targeted therapy that may comprise, for example, the use of cancer vaccines and/or sensitized antigen presenting cells. For example, an oncolytic virus is a virus that is able to infect and lyse cancer cells, while leaving normal cells unharmed, making them potentially useful in cancer therapy. Replication of oncolytic viruses both facilitates tumor cell destruction and also produces dose amplification at the tumor site. They may also act as vectors for anticancer genes, allowing them to be specifically delivered to the tumor site. The immunotherapy can involve passive immunity for short-term protection of a host, achieved by the administration of pre-formed antibody directed against a cancer antigen or disease antigen (*e.g.*, administration of a monoclonal antibody, optionally linked to a chemotherapeutic agent or toxin, to a tumor antigen). For example, anti-VEGF and mTOR inhibitors are known to be effective in treating renal cell carcinoma. Immunotherapy can also focus on using the cytotoxic lymphocyte-recognized epitopes of cancer cell lines. Alternatively, antisense polynucleotides, ribozymes, RNA interference molecules, triple helix polynucleotides and the like, can be used to selectively modulate biomolecules that are linked to the initiation, progression, and/or pathology of a tumor or cancer.

Immunotherapy can involve passive immunity for short-term protection of a host, achieved by the administration of pre-formed antibody directed against a cancer antigen or disease antigen (*e.g.*, administration of a monoclonal antibody, optionally linked to a chemotherapeutic agent or toxin, to a tumor antigen). Immunotherapy can also focus on using the cytotoxic lymphocyte-recognized epitopes of cancer cell lines. Alternatively, antisense polynucleotides, ribozymes, RNA interference molecules, triple helix polynucleotides and the like, can be used to selectively modulate biomolecules that are linked to the initiation, progression, and/or pathology of a tumor or cancer.

The term “immunogenic chemotherapy” refers to any chemotherapy that has been demonstrated to induce immunogenic cell death, a state that is detectable by the release of one or more damage-associated molecular pattern (DAMP) molecules, including, but not limited to, calreticulin, ATP and HMGB1 (Kroemer *et al.* (2013), *Annu. Rev. Immunol.*, 31:51-72). Specific representative examples of consensus immunogenic chemotherapies include 5'-fluorouracil, anthracyclines, such as doxorubicin, and the platinum drug, oxaliplatin, among others.

In some embodiments, immunotherapy comprises inhibitors of one or more immune checkpoints. The term “immune checkpoint” refers to a group of molecules on the cell surface of CD4+ and/or CD8+ T cells that fine-tune immune responses by down-modulating or inhibiting an anti-tumor immune response. Immune checkpoint proteins are well-known in the art and include, without limitation, CTLA-4, PD-1, VISTA, B7-H2, B7-H3, PD-L1, B7-H4, B7-H6, ICOS, HVEM, PD-L2, CD160, gp49B, PIR-B, KIR family receptors, TIM-1, TIM-3, TIM-4, LAG-3, GITR, 4-1BB, OX-40, BTLA, SIRP, CD47, CD48, 2B4 (CD244), B7.1, B7.2, ILT-2, ILT-4, TIGIT, HHLA2, butyrophilins, IDO, CD39, CD73 and A2aR (see, for example, WO 2012/177624). The term further encompasses biologically active protein fragment, as well as nucleic acids encoding full-length immune checkpoint proteins and biologically active protein fragments thereof. In some embodiment, the term further encompasses any fragment according to homology descriptions provided herein. In one embodiment, the immune checkpoint is PD-1.

Immune checkpoints and their sequences are well-known in the art and representative embodiments are described below. For example, the term “PD-1” refers to a member of the immunoglobulin gene superfamily that functions as a coinhibitory receptor having PD-L1 and PD-L2 as known ligands. PD-1 was previously identified using a subtraction cloning based approach to select for genes upregulated during TCR-induced activated T cell death. PD-1 is a member of the CD28/CTLA-4 family of molecules based on its ability to bind to PD-L1. Like CTLA-4, PD-1 is rapidly induced on the surface of T-cells in response to anti-CD3 (Agata *et al.* 25 (1996) *Int. Immunol.* 8:765). In contrast to CTLA-4, however, PD-1 is also induced on the surface of B-cells (in response to anti-IgM). PD-1 is also expressed on a subset of thymocytes and myeloid cells (Agata *et al.* (1996) *supra*; Nishimura *et al.* (1996) *Int. Immunol.* 8:773).

The nucleic acid and amino acid sequences of a representative human PD-1 biomarker is available to the public at the GenBank database under NM_005018.2 and

NP_005009.2 and is shown in Table 1 (see also Ishida *et al.* (1992) 20 *EMBO J* 11:3887; Shinohara *et al.* (1994) *Genomics* 23:704; U.S. Patent 5,698,520). PD-1 has an extracellular region containing immunoglobulin superfamily domain, a transmembrane domain, and an intracellular region including an immunoreceptor tyrosine-based inhibitory motif (ITIM) (Ishida *et al.* (1992) *EMBO J.* 11:3887; Shinohara *et al.* (1994) *Genomics* 23:704; and U.S. Patent 5,698,520) and an immunoreceptor tyrosine-based switch motif (ITSM). These features also define a larger family of polypeptides, called the immunoinhibitory receptors, which also includes gp49B, PIR-B, and the killer inhibitory receptors (KIRs) (Vivier and Daeron (1997) *Immunol. Today* 18:286). It is often assumed that the tyrosyl phosphorylated ITIM and ITSM motif of these receptors interacts with SH2-domain containing phosphatases, which leads to inhibitory signals. A subset of these immunoinhibitory receptors bind to MHC polypeptides, for example the KIRs, and CTLA4 binds to B7-1 and B7-2. It has been proposed that there is a phylogenetic relationship between the MHC and B7 genes (Henry *et al.* (1999) *Immunol. Today* 20(6):285-8).

Nucleic acid and polypeptide sequences of PD-1 orthologs in organisms other than humans are well-known and include, for example, mouse PD-1 (NM_008798.2 and NP_032824.1), rat PD-1 (NM_001106927.1 and NP_001100397.1), dog PD-1 (XM_543338.3 and XP_543338.3), cow PD-1 (NM_001083506.1 and NP_001076975.1), and chicken PD-1 (XM_422723.3 and XP_422723.2).

PD-1 polypeptides are inhibitory receptors capable of transmitting an inhibitory signal to an immune cell to thereby inhibit immune cell effector function, or are capable of promoting costimulation (*e.g.*, by competitive inhibition) of immune cells, *e.g.*, when present in soluble, monomeric form. Preferred PD-1 family members share sequence identity with PD-1 and bind to one or more B7 family members, *e.g.*, B7-1, B7-2, PD-1 ligand, and/or other polypeptides on antigen presenting cells.

The term “PD-1 activity,” includes the ability of a PD-1 polypeptide to modulate an inhibitory signal in an activated immune cell, *e.g.*, by engaging a natural PD-1 ligand on an antigen presenting cell. Modulation of an inhibitory signal in an immune cell results in modulation of proliferation of, and/or cytokine secretion by, an immune cell. Thus, the term “PD-1 activity” includes the ability of a PD-1 polypeptide to bind its natural ligand(s), the ability to modulate immune cell costimulatory or inhibitory signals, and the ability to modulate the immune response.

The term “PD-1 ligand” refers to binding partners of the PD-1 receptor and includes both PD-L1 (Freeman *et al.* (2000) *J. Exp. Med.* 192:1027-1034) and PD-L2 (Latchman *et al.* (2001) *Nat. Immunol.* 2:261). At least two types of human PD-1 ligand polypeptides exist. PD-1 ligand proteins comprise a signal sequence, and an IgV domain, an IgC domain, a transmembrane domain, and a short cytoplasmic tail. Both PD-L1 (See Freeman *et al.* (2000) for sequence data) and PD-L2 (See Latchman *et al.* (2001) *Nat. Immunol.* 2:261 for sequence data) are members of the B7 family of polypeptides. Both PD-L1 and PD-L2 are expressed in placenta, spleen, lymph nodes, thymus, and heart. Only PD-L2 is expressed in pancreas, lung and liver, while only PD-L1 is expressed in fetal liver. Both PD-1 ligands are upregulated on activated monocytes and dendritic cells, although PD-L1 expression is broader. For example, PD-L1 is known to be constitutively expressed and upregulated to higher levels on murine hematopoietic cells (*e.g.*, T cells, B cells, macrophages, dendritic cells (DCs), and bone marrow-derived mast cells) and non-hematopoietic cells (*e.g.*, endothelial, epithelial, and muscle cells), whereas PD-L2 is inducibly expressed on DCs, macrophages, and bone marrow-derived mast cells (see Butte *et al.* (2007) *Immunity* 27:111).

PD-1 ligands comprise a family of polypeptides having certain conserved structural and functional features. The term “family” when used to refer to proteins or nucleic acid molecules, is intended to mean two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology, as defined herein. Such family members can be naturally or non-naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin, as well as other, distinct proteins of human origin or alternatively, can contain homologues of non-human origin. Members of a family may also have common functional characteristics. PD-1 ligands are members of the B7 family of polypeptides. The term “B7 family” or “B7 polypeptides” as used herein includes costimulatory polypeptides that share sequence homology with B7 polypeptides, *e.g.*, with B7-1, B7-2, B7h (Swallow *et al.* (1999) *Immunity* 11:423), and/or PD-1 ligands (*e.g.*, PD-L1 or PD-L2). For example, human B7-1 and B7-2 share approximately 26% amino acid sequence identity when compared using the BLAST program at NCBI with the default parameters (Blosom62 matrix with gap penalties set at existence 11 and extension 1 (See the NCBI website). The term B7 family also includes variants of these polypeptides which are capable of modulating immune cell function. The B7 family of molecules share

a number of conserved regions, including signal domains, IgV domains and the IgC domains. IgV domains and the IgC domains are art-recognized Ig superfamily member domains. These domains correspond to structural units that have distinct folding patterns called Ig folds. Ig folds are comprised of a sandwich of two β sheets, each consisting of anti-parallel β strands of 5-10 amino acids with a conserved disulfide bond between the two sheets in most, but not all, IgC domains of Ig, TCR, and MHC molecules share the same types of sequence patterns and are called the C1-set within the Ig superfamily. Other IgC domains fall within other sets. IgV domains also share sequence patterns and are called V set domains. IgV domains are longer than IgC domains and contain an additional pair of β strands.

Preferred B7 polypeptides are capable of providing costimulatory or inhibitory signals to immune cells to thereby promote or inhibit immune cell responses. For example, B7 family members that bind to costimulatory receptors increase T cell activation and proliferation, while B7 family members that bind to inhibitory receptors reduce costimulation. Moreover, the same B7 family member may increase or decrease T cell costimulation. For example, when bound to a costimulatory receptor, PD-1 ligand can induce costimulation of immune cells or can inhibit immune cell costimulation, *e.g.*, when present in soluble form. When bound to an inhibitory receptor, PD-1 ligand polypeptides can transmit an inhibitory signal to an immune cell. Preferred B7 family members include B7-1, B7-2, B7h, PD-L1 or PD-L2 and soluble fragments or derivatives thereof. In one embodiment, B7 family members bind to one or more receptors on an immune cell, *e.g.*, CTLA4, CD28, ICOS, PD-1 and/or other receptors, and, depending on the receptor, have the ability to transmit an inhibitory signal or a costimulatory signal to an immune cell, preferably a T cell.

Modulation of a costimulatory signal results in modulation of effector function of an immune cell. Thus, the term “PD-1 ligand activity” includes the ability of a PD-1 ligand polypeptide to bind its natural receptor(s) (*e.g.* PD-1 or B7-1), the ability to modulate immune cell costimulatory or inhibitory signals, and the ability to modulate the immune response.

The term “PD-L1” refers to a specific PD-1 ligand. Two forms of human PD-L1 molecules have been identified. One form is a naturally occurring PD-L1 soluble polypeptide, *i.e.*, having a short hydrophilic domain and no transmembrane domain, and is referred to herein as PD-L1S. The second form is a cell-associated polypeptide, *i.e.*, having

a transmembrane and cytoplasmic domain, referred to herein as PD-L1M. The nucleic acid and amino acid sequences of representative human PD-L1 biomarkers regarding PD-L1M are also available to the public at the GenBank database under NM_014143.3 and NP_054862.1. PD-L1 proteins comprise a signal sequence, and an IgV domain and an IgC domain. The signal sequence of PD-L1S is shown from about amino acid 1 to about amino acid 18. The signal sequence of PD-L1M is shown from about amino acid 1 to about amino acid 18. The IgV domain of PD-L1S is shown from about amino acid 19 to about amino acid 134 and the IgV domain of PD-L1M is shown from about amino acid 19 to about amino acid 134. The IgC domain of PD-L1S is shown from about amino acid 135 to about amino acid 227 and the IgC domain of PD-L1M is shown from about amino acid 135 to about amino acid 227. The hydrophilic tail of the PD-L1 exemplified in PD-L1S comprises a hydrophilic tail shown from about amino acid 228 to about amino acid 245. The PD-L1 polypeptide exemplified in PD-L1M comprises a transmembrane domain shown from about amino acids 239 to about amino acid 259 and a cytoplasmic domain shown from about amino acid 260 to about amino acid 290. In addition, nucleic acid and polypeptide sequences of PD-L1 orthologs in organisms other than humans are well-known and include, for example, mouse PD-L1 (NM_021893.3 and NP_068693.1), rat PD-L1 (NM_001191954.1 and NP_001178883.1), dog PD-L1 (XM_541302.3 and XP_541302.3), cow PD-L1 (NM_001163412.1 and NP_001156884.1), and chicken PD-L1 (XM_424811.3 and XP_424811.3).

The term “PD-L2” refers to another specific PD-1 ligand. PD-L2 is a B7 family member expressed on various APCs, including dendritic cells, macrophages and bone-marrow derived mast cells (Zhong *et al.* (2007) Eur. J. Immunol. 37:2405). APC-expressed PD-L2 is able to both inhibit T cell activation through ligation of PD-1 and costimulate T cell activation, through a PD-1 independent mechanism (Shin *et al.* (2005) J. Exp. Med. 201:1531). In addition, ligation of dendritic cell-expressed PD-L2 results in enhanced dendritic cell cytokine expression and survival (Radhakrishnan *et al.* (2003) J. Immunol. 37:1827; Nguyen *et al.* (2002) J. Exp. Med. 196:1393). The nucleic acid and amino acid sequences of representative human PD-L2 biomarkers are well-known in the art and are also available to the public at the GenBank database under NM_025239.3 and NP_079515.2. PD-L2 proteins are characterized by common structural elements. In some embodiments, PD-L2 proteins include at least one or more of the following domains: a signal peptide domain, a transmembrane domain, an IgV domain, an IgC domain, an

extracellular domain, a transmembrane domain, and a cytoplasmic domain. For example, amino acids 1-19 of PD-L2 comprises a signal sequence. As used herein, a “signal sequence” or “signal peptide” serves to direct a polypeptide containing such a sequence to a lipid bilayer, and is cleaved in secreted and membrane bound polypeptides and includes a peptide containing about 15 or more amino acids which occurs at the N-terminus of secretory and membrane bound polypeptides and which contains a large number of hydrophobic amino acid residues. For example, a signal sequence contains at least about 10-30 amino acid residues, preferably about 15- 25 amino acid residues, more preferably about 18-20 amino acid residues, and even more preferably about 19 amino acid residues, and has at least about 35-65%, preferably about 38-50%, and more preferably about 40-45% hydrophobic amino acid residues (*e.g.*, valine, leucine, isoleucine or phenylalanine). In another embodiment, amino acid residues 220-243 of the native human PD-L2 polypeptide and amino acid residues 201-243 of the mature polypeptide comprise a transmembrane domain. As used herein, the term “transmembrane domain” includes an amino acid sequence of about 15 amino acid residues in length which spans the plasma membrane. More preferably, a transmembrane domain includes about at least 20, 25, 30, 35, 40, or 45 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, *e.g.*, leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta, W. N. *et al.* (1996) *Annu. Rev. Neurosci.* 19: 235-263. In still another embodiment, amino acid residues 20-120 of the native human PD-L2 polypeptide and amino acid residues 1-101 of the mature polypeptide comprise an IgV domain. Amino acid residues 121- 219 of the native human PD-L2 polypeptide and amino acid residues 102-200 of the mature polypeptide comprise an IgC domain. As used herein, IgV and IgC domains are recognized in the art as Ig superfamily member domains. These domains correspond to structural units that have distinct folding patterns called Ig folds. Ig folds are comprised of a sandwich of two β sheets, each consisting of antiparallel (3 strands of 5-10 amino acids with a conserved disulfide bond between the two sheets in most, but not all, domains. IgC domains of Ig, TCR, and MHC molecules share the same types of sequence patterns and are called the C1 set within the Ig superfamily. Other IgC domains fall within other sets. IgV domains also share sequence patterns and are called V set domains. IgV domains are longer than C-

domains and form an additional pair of strands. In yet another embodiment, amino acid residues 1-219 of the native human PD-L2 polypeptide and amino acid residues 1-200 of the mature polypeptide comprise an extracellular domain. As used herein, the term “extracellular domain” represents the N-terminal amino acids which extend as a tail from the surface of a cell. An extracellular domain of the present invention includes an IgV domain and an IgC domain, and may include a signal peptide domain. In still another embodiment, amino acid residues 244-273 of the native human PD-L2 polypeptide and amino acid residues 225-273 of the mature polypeptide comprise a cytoplasmic domain. As used herein, the term “cytoplasmic domain” represents the C-terminal amino acids which extend as a tail into the cytoplasm of a cell. In addition, nucleic acid and polypeptide sequences of PD-L2 orthologs in organisms other than humans are well-known and include, for example, mouse PD-L2 (NM_021396.2 and NP_067371.1), rat PD-L2 (NM_001107582.2 and NP_001101052.2), dog PD-L2 (XM_847012.2 and XP_852105.2), cow PD-L2 (XM_586846.5 and XP_586846.3), and chimpanzee PD-L2 (XM_001140776.2 and XP_001140776.1).

The term “PD-L2 activity,” “biological activity of PD-L2,” or “functional activity of PD-L2,” refers to an activity exerted by a PD-L2 protein, polypeptide or nucleic acid molecule on a PD-L2-responsive cell or tissue, or on a PD-L2 polypeptide binding partner, as determined *in vivo*, or *in vitro*, according to standard techniques. In one embodiment, a PD-L2 activity is a direct activity, such as an association with a PD-L2 binding partner. As used herein, a “target molecule” or “binding partner” is a molecule with which a PD-L2 polypeptide binds or interacts in nature, such that PD-L2-mediated function is achieved. In an exemplary embodiment, a PD-L2 target molecule is the receptor RGMb. Alternatively, a PD-L2 activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the PD-L2 polypeptide with its natural binding partner (*i.e.*, physiologically relevant interacting macromolecule involved in an immune function or other biologically relevant function), *e.g.*, RGMb. The biological activities of PD-L2 are described herein. For example, the PD-L2 polypeptides of the present invention can have one or more of the following activities: 1) bind to and/or modulate the activity of the receptor RGMb, PD-1, or other PD-L2 natural binding partners, 2) modulate intra-or intercellular signaling, 3) modulate activation of immune cells, *e.g.*, T lymphocytes, and 4) modulate the immune response of an organism, *e.g.*, a mouse or human organism.

“Anti-immune checkpoint therapy” refers to the use of agents that inhibit immune checkpoint nucleic acids and/or proteins. Inhibition of one or more immune checkpoints can block or otherwise neutralize inhibitory signaling to thereby upregulate an immune response in order to more efficaciously treat cancer. Exemplary agents useful for inhibiting immune checkpoints include antibodies, small molecules, peptides, peptidomimetics, natural ligands, and derivatives of natural ligands, that can either bind and/or inactivate or inhibit immune checkpoint proteins, or fragments thereof; as well as RNA interference, antisense, nucleic acid aptamers, etc. that can downregulate the expression and/or activity of immune checkpoint nucleic acids, or fragments thereof. Exemplary agents for upregulating an immune response include antibodies against one or more immune checkpoint proteins block the interaction between the proteins and its natural receptor(s); a non-activating form of one or more immune checkpoint proteins (*e.g.*, a dominant negative polypeptide); small molecules or peptides that block the interaction between one or more immune checkpoint proteins and its natural receptor(s); fusion proteins (*e.g.* the extracellular portion of an immune checkpoint inhibition protein fused to the Fc portion of an antibody or immunoglobulin) that bind to its natural receptor(s); nucleic acid molecules that block immune checkpoint nucleic acid transcription or translation; and the like. Such agents can directly block the interaction between the one or more immune checkpoints and its natural receptor(s) (*e.g.*, antibodies) to prevent inhibitory signaling and upregulate an immune response. Alternatively, agents can indirectly block the interaction between one or more immune checkpoint proteins and its natural receptor(s) to prevent inhibitory signaling and upregulate an immune response. For example, a soluble version of an immune checkpoint protein ligand such as a stabilized extracellular domain can binding to its receptor to indirectly reduce the effective concentration of the receptor to bind to an appropriate ligand. In one embodiment, anti-PD-1 antibodies, anti-PD-L1 antibodies, and/or anti-PD-L2 antibodies, either alone or in combination, are used to inhibit immune checkpoints. These embodiments are also applicable to specific therapy against particular immune checkpoints, such as the PD-1 pathway (*e.g.*, anti-PD-1 pathway therapy, otherwise known as PD-1 pathway inhibitor therapy).

The term “HLA-E,” also known as HLA class I histocompatibility antigen (alpha chain E) and Major Histocompatibility Complex (MHC) class I antigen E, refers to a member of a group of proteins in the non-classical MHC class I that are characterized by a limited polymorphism and a lower cell surface expression than the classical paralogues.

The functional homolog in mice is called Qa-1b, officially known as H2-T23. HLA-E has a very specialized role in cell recognition by natural killer cells (NK cells) (Braud *et al.* (1998) *Nature* 391:795-799). HLA-E binds a restricted subset of peptides derived from signal peptides of classical MHC class I molecules, namely HLA-A, B, C, G (Braud *et al.* (1997) *Eur. J. Immunol.* 27:1164-1169). These peptides are released from the membrane of the endoplasmic reticulum (ER) by the signal peptide peptidase and trimmed by the cytosolic proteasome (Lemberg *et al.* (2001) *J. Immunol.* 167:6441-6446; Bland *et al.* (2003) *J. Biol. Chem.* 278:33747-33752). Upon transport into the ER lumen by the transporter associated with antigen processing (TAP), these peptides bind to a peptide binding groove on the HLA-E molecule (Braud *et al.* (1998) *Curr. Biol.* 8:1-10). This allows HLA-E to assemble correctly and to be expressed on the cell surface. NK cells recognize the HLA-E+peptide complex using the heterodimeric inhibitory receptor CD94/NKG2A/B/C (Braud *et al.* (1998) *Nature* 391:795-799). When CD94/NKG2A or CD94/NKG2B is engaged, it produces an inhibitory effect on the cytotoxic activity of the NK cell to prevent cell lysis. However, binding of HLA-E to CD94/NKG2C results in NK cell activation. This interaction has been shown to trigger expansion of NK cell subsets in antiviral responses (Rölle *et al.* (2014) *J. Clin. Invest.* 124:5305-5316). HLA-E is involved in pathways such as allograft rejection (*e.g.*, Type I diabetes mellitus, graft-versus-host disease, autoimmune thyroid disease, viral myocarditis, viral carcinogenesis, viral infection (such as Epstein-Barr virus infection, herpes simplex infection, HTLV-I infection, etc.), interferon gamma signaling, CTL-mediated apoptosis, antigen processing and presentation, etc.), immune response role of DAP12 receptors in NK cells (*e.g.*, CD16 signaling in NK cells, NK cell-mediated cytotoxicity, etc.), antigen processing-cross presentation (*e.g.*, ER-phagosome pathway, endosomal/vacuolar pathway, etc.), class I MHC mediated antigen processing and presentation, and RET signaling (*e.g.*, DAP12 signaling, etc.). HLA-E preferably binds to a peptide derived from the signal sequence of most HLA-A, -B, -C and -G molecules, as well as beta-2-microglobulin (B2M), MHC class I protein, Major Histocompatibility Complex, Class II, DR Beta 1 (HLA-DRB1), NK cell lectin-like receptor, and other other peptide antigens.

The nucleic acid and amino acid sequences of a representative human HLA-E is available to the public at the GenBank database (Gene ID 3133) and is shown in Table 1 (*e.g.*, GenBank database numbers NM_005516.5 and NP_005507.3). HLA-E belongs to the HLA class I heavy chain paralogues. This class I molecule is a heterodimer consisting

of a heavy chain and a light chain (beta-2 microglobulin). The heavy chain is anchored in the membrane. HLA-E binds a restricted subset of peptides derived from the leader peptides of other class I molecules. The heavy chain is approximately 45 kDa and its gene contains 8 exons. Exon one encodes the leader peptide, exons 2 and 3 encode the alpha1 and alpha2 domains, which both bind the peptide, exon 4 encodes the alpha3 domain, exon 5 encodes the transmembrane region, and exons 6 and 7 encode the cytoplasmic tail. The domain structure of HLA-E polypeptide is well known and accessible in UniProtKB database under the accession number P13747, including, in the order from the 5' terminus to the 3' terminus, a signal peptide comprising, *e.g.*, amino acid positions 1-21 of NP_005507.3, an extracellular domain comprising, *e.g.*, amino acid positions 22-305 of NP_005507.3, a transmembrane domain comprising, *e.g.*, amino acid positions 306-329 of NP_005507.3, and a cytoplasmic domain comprising, *e.g.*, amino acid positions 330-358 of NP_005507.3. The extracellular domain further comprises an Alpha-1 region (*e.g.*, amino acid positions 22-111 of NP_005507.3), an Alpha-2 region (*e.g.*, amino acid positions 112-203 of NP_005507.3), an Alpha-3 region (*e.g.*, amino acid positions 204-295 of NP_005507.3, including an Ig-like C1-type domain comprising, *e.g.*, amino acid positions 206-294 of NP_005507.3), and a connecting peptide (*e.g.*, amino acid positions 296-305 of NP_005507.3).

Nucleic acid and polypeptide sequences of HLA-E orthologs in organisms other than humans are well-known and include, for example, chimpanzee (*Pan troglodytes*) PATR-E (NM_001045498.1 and NP_001038963.1), Rhesus monkey MAMU-E (NM_001114966.1 and NP_001108438.1), mouse H2-T23 (NM_010398.3 and NP_034528.1, and NM_001271005.1 and NP_001257934.1), and Norway rat (*Rattus norvegicus*) RT1-S3 (NM_001008886.2 and NP_001008886.2).

The term "HLA-E activity" includes the ability of an HLA-E polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein) to bind its substrate and/or catalyze serine/threonine-protein kinase activity. HLA-E activity may also include one or more of functions, such as those disclosed herein in the NF- κ B pathway, autoimmune response, antigen presentation, and viral infection. For example, HLA-E may interact with various proteins disclosed herein, such as HLA-A, -B, -C and -G molecules and other binding partners including multiple antigens and receptors, for its functions in signalings. HLA-E may also be proteolytically modified, such as being glycosylated, or otherwise disclosed herein, for its functions.

The term “HLA-E substrate(s)” refers to binding partners of a HLA-E polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein), *e.g.*, the proteins listed herein. Furthermore, HLA-E substrates may refer to downstream members in the signaling pathways where HLA-E has a functional role.

5 The term “HLA-E-regulated signaling pathway(s)” includes signaling pathways in which HLA-E (and its fragments, domains, and/or motifs thereof, discussed herein) binds to at least one of its substrate, through which at least one cellular function and/or activity and/or cellular protein profiles is changed (*e.g.*, antigen processing and presentation). In some embodiments, HLA-E promotes antigen presentation (including autoimmune
10 responses) for at least one of its substrates (*e.g.*, antigens) which bind to it. HLA-E-regulated signaling pathways include at least those described herein, such as allograft rejection, antigen processing-cross presentation, class I MHC mediated antigen processing and presentation, and RET signaling, etc.

 The term “HLA-E inhibitor(s)” includes any natural or non-natural agent prepared,
15 synthesized, manufactured, and/or purified by human that is capable of reducing, inhibiting, blocking, preventing, and/or that inhibits the ability of an HLA-E polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein). In one embodiment, such inhibitors may reduce or inhibit the binding/interaction between HLA-E and its substrates or other binding partners. In another embodiment, such inhibitors may reduce or inhibit
20 HLA-E as an MHC molecule for antigen presentation and/or other immune responses. In still another embodiment, such inhibitors may increase or promote the turnover rate, reduce or inhibit the expression and/or the stability (*e.g.*, the half-life), and/or change the cellular localization of HLA-E, resulting in at least a decrease in HLA-E levels and/or activity. Such inhibitors may be any molecule, including but not limited to small molecule
25 compounds, antibodies or intrabodies, RNA interfering (RNAi) agents (including at least siRNAs, shRNAs, microRNAs (miRNAs), piwi, and other well-known agents). Such inhibitors may be specific to HLA-E or also inhibit at least one of other serine/threonine-protein kinases. RNA interference for HLA-E polypeptides are well known and commercially available (*e.g.*, human or mouse shRNA (Cat. # TF312402, TG312402, TL312402, TL312402V, and TR312402) products from Origene (Rockville, MD),
30 siRNA/shRNA products (Cat. # sc-62470, sc-62471, and sc-42922) and human or mouse gene knockout kit via CRISPR (Cat. # sc-403124 and sc-437133) from Santa Cruz Biotechnology (Dallas, Texas), and Ready-to-package AAV shRNA clones from Vigene

Biosciences (Rockville, MD, Cat. # SH888322), etc.). Methods for detection, purification, and/or inhibition of HLA-E (*e.g.*, by anti-HLA-E antibodies) are also well known and commercially available (*e.g.*, multiple anti-HLA-E antibodies from Origene (Cat. # TA320346, AM03010BT-N, SM3053P, SM3053A, and others), Novus Biologicals (Littleton, CO, Cat. # NBP2-47435, NBP1-43124, NB500-310, etc.), abcam (Cambridge, MA, Cat. # ab2216, ab11820, ab11821, ab88090, ab203082, etc.), Santa Cruz Biotechnology (Cat. # sc-71262, sc-51621, sc-51622, and sc-51624), etc.). HLA-E knockout human cell lines are also well known and available at the Horizon (Cambridge, UK, Cat. # HZGHC3133).

The term “PDIA3,” also known as protein disulfide-isomerase A3 and glucose-regulated protein, 58-kD (GRP58), refers to a member of a group of isomerase enzymes which localize to the endoplasmic reticulum (ER) and interacts with lectin chaperones calreticulin and calnexin (CNX) to modulate folding of newly synthesized glycoproteins. Complexes of lectins and this protein mediate protein folding by promoting formation of disulfide bonds in their glycoprotein substrates. The PDIA3 protein is a thiol oxidoreductase that has protein disulfide isomerase activity (Koivunen *et al.* (1996) *Biochem. J.* 316:599-605; Dong *et al.* (2009) *Immunity* 30:21-32). PDIA3 is also part of the major histocompatibility complex (MHC) class I peptide loading complex, which is essential for formation of the final antigen conformation and export from the endoplasmic reticulum to the cell surface (Dong *et al.* (2009), *supra*; Garbi *et al.* (2006) *Nat.Immunol.* 7:93-102). This protein of the endoplasmic reticulum interacts with lectin chaperones such as calreticulin and CNX in order to modulate the folding of proteins that are newly synthesized. PDIA3 plays a role in protein folding by promoting the formation of disulfide bonds, and that CNX facilitates the positioning substrates next to the catalytic cysteines. This function allows it to serve as a redox sensor by activating mTORC1, which then mediates mTOR complex assembly to adapt cells to oxidative damage. Thus, PDIA3 regulates cell growth and death according to oxygen concentrations, such as in the hypoxic microenvironment of bones. Additionally, PDIA3 activates cell anchorage in bones by associating with cell division and cytoskeleton proteins, such as beta-actin and vimentin, to form a complex which controls TUBB3 folding and proper attachment of the microtubules to the kinetochore. PDIA3 also plays a role in cytokine-dependent signal transduction, including STAT3 signaling (Santana-Codina *et al.* (2013) *Mol. Cell. Proteomics* 12:2111-2125). PDIA3 is involved in pathways such as antigen processing-cross presentation (*e.g.*,

ER-phagosome pathway), Class I MHC mediated antigen processing and presentation, calnexin/calreticulin cycle, protein metabolisms (*e.g.*, post-translational protein modification), and allograft rejection. It has been demonstrated that the downregulation of PDIA3 expression is correlated with poor prognosis in early-stage cervical cancer (Chung *et al.* (2013) *Biomarkers* 18:573-579). It has also been demonstrated that ERp57/PDIA3 binds specific DNA fragments in a melanoma cell line (Aureli *et al.* (2013) *Gene* 524:390-395). PDIA3 is also involved in bone metastasis, which is the most common source of distant relapse in breast cancer (Santana-Codina *et al.* (2013), *supra*). In addition to cancer, overexpression of PDIA3 is linked to renal fibrosis, which is characterized by excess synthesis and secretion of ECM leading to ER stress (Dihazi *et al.* (2013) *J. Cell Sci.* 126:3649-3663). PDIA3 has been related to maxillary sinus squamous cell carcinoma, maxillary sinus cancer, and paranasal sinus cancer (Ogino *et al.* (2003) *Clin. Cancer Res.* 9:4043-4051).

The nucleic acid and amino acid sequences of a representative human PDIA3 is available to the public at the GenBank database (Gene ID 2923) and is shown in Table 1 (*e.g.*, GenBank database numbers NM_005313.4 and NP_005304.3). The PDIA3 protein consists of four thioredoxin-like domains: a, b, b', and a'. The a and a' domains have Cys-Gly-His-Cys active site motifs (C57-G58-H59-C60 and C406-G407-H408-C409) and are catalytically active (Dong *et al.* (2009), *supra*; Kozlov *et al.* (2006) *Structure* 14:1331-1339). The bb' domains contain a CNX binding site, which is composed of positively charged, highly conserved residues (K214, K274, and R282) that interact with the negatively charged residues of the CNX P domain. The b' domain comprises the majority of the binding site, but the β 4- β 5 loop of the b domain provides additional contact (K214) to strengthen the interaction. A transient disulfide bond forms between the N-terminal cysteine in the catalytic motif and a substrate, but in a step called "escape pathway", the bond is disrupted as the C-terminal cysteine attacks the N-terminal cysteine to release the substrate. The domain structure of PDIA3 polypeptide is well known and accessible in UniProtKB database under the accession number P30101, including, in the order from the 5' terminus to the 3' terminus, a Thioredoxin domain 1 comprising, *e.g.*, amino acid positions 25-133 of NP_005304.3, a Thioredoxin domain 2 comprising, *e.g.*, amino acid positions 343-485 of NP_005304.3, and a C-terminal motif to prevent secretion from ER (*e.g.*, amino acid positions 502-505 of NP_005304.3). Known binding partners of PDIA3 include, *e.g.*, BACE1 (Leach *et al.* (2002) *J. Biol. Chem.* 277:29686-29697), ERP27

(Alanen *et al.* (2006) *J. Biol. Chem.* 281:33727-33738), tapasin (Leach *et al.* (2002), *supra*), CRT, and CNX.

Nucleic acid and polypeptide sequences of PDIA3 orthologs in organisms other than humans are well-known and include, for example, chimpanzee (*Pan troglodytes*) PDIA3 (NM_001246452.1 and NP_001233381.1), Rhesus monkey PDIA3 (NM_001195716.1 and NP_001182645.1), dog PDIA3 (XM_535453.5 and XP_535453.3), cattle PDIA3 (NM_174333.3 and NP_776758.2), mouse PDIA3 (NM_007952.2 and NP_031978.2), chicken PDIA3 (NM_204110.3 and NP_989441.1), tropical clawed frog PDIA3 (NM_203998.1 and NP_989329.1), fruit fly ERp60 (NM_165849.3 and NP_725084.2), Norway rat (*Rattus norvegicus*) PDIA3 (NM_017319.1 and NP_059015.1) and zebrafish PDIA3 (NM_001199737.1 and NP_001186666.1).

The term “PDIA3 activity” includes the ability of a PDIA3 polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein) to bind its substrate and/or catalyze thiol oxidoreductase activity. PDIA3 activity may also include one or more of functions, such as those disclosed herein in antigen processing-cross presentation (*e.g.*, ER-phagosome pathway), Class I MHC mediated antigen processing and presentation, calnexin/calreticulin cycle, protein metabolisms (*e.g.*, post-translational protein modification), and allograft rejection. For example, PDIA3 may interact with various proteins disclosed herein, such as BACE1, ERP27, tapasin, CRT, and CNX and other binding partners including multiple antigens and receptors, for its functions in signalings. PDIA3 may also be proteolytically modified, such as being glycosylated, or otherwise disclosed herein, for its functions.

The term “PDIA3 substrate(s)” refers to binding partners of a PDIA3 polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein), *e.g.*, the proteins listed herein. Furthermore, PDIA3 substrates may refer to downstream members in the signaling pathways where PDIA3 has a functional role.

The term “PDIA3-regulated signaling pathway(s)” includes signaling pathways in which PDIA3 (and its fragments, domains, and/or motifs thereof, discussed herein) binds to at least one of its substrate, through which at least one cellular function and/or activity and/or cellular protein profiles is changed (*e.g.*, antigen processing and presentation). In some embodiments, PDIA3 promotes antigen presentation (including autoimmune responses) for at least one of its substrates (*e.g.*, antigens) which bind to it. PDIA3-regulated signaling pathways include at least those described herein, such as antigen

processing-cross presentation (e.g., ER-phagosome pathway), Class I MHC mediated antigen processing and presentation, calnexin/calreticulin cycle, protein metabolisms (e.g., post-translational protein modification), and allograft rejection, etc.

The term “PDIA3 inhibitor(s)” includes any natural or non-natural agent prepared, synthesized, manufactured, and/or purified by human that is capable of reducing, inhibiting, blocking, preventing, and/or that inhibits the ability of a PDIA3 polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein). In one embodiment, such inhibitors may reduce or inhibit the binding/interaction between PDIA3 and its substrates or other binding partners. In another embodiment, such inhibitors may reduce or inhibit PDIA3 as a protein disulfide isomerase and as an MHC molecule for antigen presentation and/or other immune responses. In still another embodiment, such inhibitors may increase or promote the turnover rate, reduce or inhibit the expression and/or the stability (*e.g.*, the half-life), and/or change the cellular localization of PDIA3, resulting in at least a decrease in PDIA3 levels and/or activity. Such inhibitors may be any molecule, including but not limited to small molecule compounds, antibodies or intrabodies, RNA interfering (RNAi) agents (including at least siRNAs, shRNAs, microRNAs (miRNAs), piwi, and other well-known agents). Such inhibitors may be specific to PDIA3 or also inhibit at least one of other serine/threonine-protein kinases. RNA interference for PDIA3 polypeptides are well known and commercially available (*e.g.*, human or mouse shRNA products (Cat. # TL310527) and siRNA products (Cat. # SR301974) and CRISPR knockout products (Cat. # KN205940 and KN313044) from Origene (Rockville, MD), siRNA/shRNA products (Cat. # sc-35341, sc-44299, sc-42876) and human or mouse gene knockout kit via CRISPR (Cat. # sc-420698) from Santa Cruz Biotechnonology (Dallas, Texas), and Ready-to-package AAV shRNA clones from Vigene Biosciences (Rockville, MD, Cat. # SH814529), etc.). Methods for detection, purification, and/or inhibition of PDIA3 (*e.g.*, by anti-PDIA3 antibodies) are also well known and commercially available (*e.g.*, multiple anti-PDIA3 antibodies from Origene (Cat. # AP16091PU-T, AP17626PU-N, TA504995BM, and others), Novus Biologicals (Littleton, CO, Cat. # NBP1-84796, NBP2-36765, NBP2-36766, etc.), abcam (Cambridge, MA, Cat. # ab13506, ab154191, ab183398, ab183397, etc.), Santa Cruz Biotechnology (Cat. # sc-136522, sc-71075, sc-166680, sc-80648, etc.), etc.). PDIA3 knockout human cell lines are also well known and available at the Horizon (Cambridge, UK, Cat. # HZGHC2923).

The term "CALR," also known as calreticulin, calregulin, CRP55, CaBP3, calsequestrin-like protein, and endoplasmic reticulum resident protein 60 (ERp60) glucose-regulated protein, 58-kD (GRP58), refers to a member of a group of multifunctional proteins that bind and inactivate (as a chaperone to store calcium in ER) Ca^{2+} ions (a second messenger in signal transduction). Calreticulin locates in the lumen of the endoplasmic reticulum and sarcoplasmic reticulum. It is also found in the nucleus, suggesting that it may have a role in transcription regulation. Calreticulin binds to the synthetic peptide KLGFfKR, which is almost identical to an amino acid sequence in the DNA-binding domain of the superfamily of nuclear receptors. CALR promotes folding, oligomeric assembly and quality control in the endoplasmic reticulum (ER) via the calreticulin/calnexin cycle. Calreticulin binds to misfolded proteins and prevents them from being exported from the endoplasmic reticulum to the Golgi apparatus. Calreticulin binds to antibodies in certain sera of systemic lupus and Sjogren patients which contain anti-Ro/SSA antibodies. Calreticulin is highly conserved among species. The amino terminus of calreticulin interacts with the DNA-binding domain of the glucocorticoid receptor and prevents the receptor from binding to its specific glucocorticoid response element. Calreticulin can inhibit the binding of androgen receptor to its hormone-responsive DNA element and can inhibit androgen receptor and retinoic acid receptor transcriptional activities *in vivo*, as well as retinoic acid-induced neuronal differentiation. Thus, calreticulin can act as an important modulator of the regulation of gene transcription by nuclear hormone receptors. Systemic lupus erythematosus is associated with increased autoantibody titers against calreticulin but calreticulin is not currently considered as a Ro/SS-A antigen. Increased autoantibody titer against human calreticulin is found in infants with complete congenital heart block of both the IgG and IgM classes. Diseases associated with CALR include myelofibrosis with myeloid metaplasia, somatic and thrombocythemia 1. CALR mutations are the second most common in myeloproliferative neoplasms. All mutations (insertions or deletions) affected the last exon, generating a reading frame shift of the resulting protein, that creates a novel terminal peptide and causes a loss of endoplasmic reticulum KDEL retention signal (Nangalia *et al.* (2013) *N. Engl. J. Med.* 369:2391-2405; Klampfl *et al.* (2013) *N. Engl. J. Med.* 369:2379-2390). CALR-related pathways include HTLV-I infection and transport to the Golgi and subsequent modification. GO annotations related to this gene include poly(A) RNA binding and ubiquitin protein ligase binding. CALR interacts transiently with almost all of the

monoglucosylated glycoproteins that are synthesized in the ER. It interacts with the DNA-binding domain of NR3C1 and mediates its nuclear export. CALR is involved in maternal gene expression regulation and may participate in oocyte maturation via the regulation of calcium homeostasis. Calreticulin has been shown to interact with perforin (Andrin *et al.* (1998) *Biochemistry* 37:10386-10394) and NK2 homeobox 1 (Perrone *et al.* (1999) *J. Biol. Chem.* 274:4640-4645).

The nucleic acid and amino acid sequences of a representative human CALR is available to the public at the GenBank database (Gene ID 811) and is shown in Table 1 (*e.g.*, GenBank database numbers NM_004343.3 and NP_004334.1). The domain structure of CALR polypeptide is well known and accessible in UniProtKB database under the accession number P27797, including, in the order from the 5' terminus to the 3' terminus, a signal peptide comprising, *e.g.*, amino acid positions 1-17 of NP_004334.1, a N-terminal globular domain (containing zinc binding sites) comprising, *e.g.*, amino acid positions 18-197 of NP_004334.1, a proline-rich P-domain forming an elongated arm-like structure comprising, *e.g.*, amino acid positions 198-308 of NP_004334.1, and a C-terminal acidic domain comprising, *e.g.*, amino acid positions 309-417 of NP_004334.1. In addition, other regions are indicated on CALR, including 4 X approximate repeats comprising, *e.g.*, amino acid positions 191-255 of NP_004334.1, a domain interacting with PPIB comprising, *e.g.*, amino acid positions 237-270 of NP_004334.1, and 3 X approximate repeats comprising, *e.g.*, amino acid positions 259-297 of NP_004334.1. The P-domain binds one molecule of calcium with high affinity, whereas the acidic C-domain binds multiple calcium ions with low affinity. The tip of the extended arm formed by the P-domain also associates with PDIA3.

Nucleic acid and polypeptide sequences of CALR orthologs in organisms other than humans are well-known and include, for example, chimpanzee CALR (XM_016935207.1 and XP_016790696.1), Rhesus monkey CALR (NM_001261131.1 and NP_001248060.1), dog CALR (XM_862217.4 and XP_867310.2), cattle CALR (NM_174000.2 and NP_776425.1), mouse CALR (NM_007591.3 and NP_031617.1), chicken CALR (NM_001348545.1 and NP_001335474.1), tropical clawed frog CALR (NM_001001253.1 and NP_001001253.1), fruit fly CALR (NM_001275501.1 and NP_001262430.1, and NM_079569.5 and NP_524293.2), rat CALR (NM_022399.2 and NP_071794.1), and zebrafish CALR (NM_199713.2 and NP_956007.2).

The term “CALR activity” includes the ability of a CALR polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein) to bind its substrate and/or catalyze thiol oxidoreductase activity. CALR activity may also include one or more of functions, such as those disclosed herein in calcium-binding, folding/assembly/quality control in the ER, interaction with glycoproteins in the ER and/or with NR3C1 for its nuclear export, maternal gene expression regulation, oocyte maturation via the regulation of calcium homeostasis.

The term “CALR substrate(s)” refers to binding partners of a CALR polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein), *e.g.*, the glycoproteins, NR3C1, perforin, NK2 homeobox 1, and PDIA3, as disclosed herein. Furthermore, CALR substrates may refer to downstream members in the signaling pathways where CALR has a functional role.

The term “CALR-regulated signaling pathway(s)” includes signaling pathways in which CALR (and its fragments, domains, and/or motifs thereof, discussed herein) binds to at least one of its substrate, through which at least one cellular function and/or activity and/or cellular protein profiles is changed (*e.g.*, antigen processing and presentation). In some embodiments, CALR promotes antigen presentation (including ER-Phagosome pathway), unfolded protein response (UPR), calnexin/calreticulin cycle, binding and uptake of ligands by scavenger receptors, Class I MHC mediated antigen processing and presentation, and/or virus assembly and release.

The term “CALR inhibitor(s)” includes any natural or non-natural agent prepared, synthesized, manufactured, and/or purified by human that is capable of reducing, inhibiting, blocking, preventing, and/or that inhibits the ability of a CALR polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein). In one embodiment, such inhibitors may reduce or inhibit the binding/interaction between CALR and its substrates or other binding partners. In another embodiment, such inhibitors may reduce or inhibit CALR functions. In one embodiment, such inhibitors may increase or promote the turnover rate, reduce or inhibit the expression and/or the stability (*e.g.*, the half-life), and/or change the cellular localization of CALR, resulting in at least a decrease in CALR levels and/or activity. Such inhibitors may be any molecule, including but not limited to small molecule compounds, antibodies or intrabodies, RNA interfering (RNAi) agents (including at least siRNAs, shRNAs, microRNAs (miRNAs), piwi, and other well-known agents). Such inhibitors may be specific to CALR or also inhibit at least one of other related proteins

(*e.g.*, calcium-binding chaperons). RNA interference for CALR polypeptides are well known and commercially available (*e.g.*, human or mouse shRNA products (Cat. # TL314216 and TG500254) and siRNA products (Cat # SR413808, SR300567, SR509128, etc.) and CRISPR knockout products (Cat. # KN203222 and KN302469) from Origene (Rockville, MD), siRNA/shRNA products (Cat. # sc-29234, sc-29895, and sc-63293) and human or mouse gene knockout kit via CRISPR (Cat. # sc-419428) from Santa Cruz Biotechnology (Dallas, Texas), and Ready-to-package AAV shRNA clones from Vigene Biosciences (Rockville, MD, Cat. # SH855911), etc.). Methods for detection, purification, and/or inhibition of CALR (*e.g.*, by anti-CALR antibodies) are also well known and commercially available (*e.g.*, multiple anti-CALR antibodies from Origene (Cat. # TA336776, AM06211SU-N, TA326909, etc.), Novus Biologicals (Littleton, CO, Cat. # NBP1-47518, NB600-101, NB600-103, etc.), abcam (Cambridge, MA, Cat. # ab2907, ab22683, ab204990, etc.), Santa Cruz Biotechnology (Cat. # sc-101436, sc-166837, sc-373863, etc.), etc.). CALR knockout human cell lines are also well known and available at the Horizon (Cambridge, UK, Cat. # HZGHC018222c012, HZGHC018222c019, etc.).

The term “TAP1,” also known as transporter 1, ATP Binding Cassette Subfamily B (MDR/TAP) member and transporter associated with antigen processing 1, refers to a member of a group of membrane-bound proteins of the superfamily of ATP-binding cassette (ABC) transporters. ABC proteins transport various molecules across extra- and intra-cellular membranes to the endoplasmic reticulum for association with MHC class I molecules. ABC genes are divided into seven distinct subfamilies (ABC1, MDR/TAP, MRP, ALD, OABP, GCN20, White). Members of the MDR/TAP subfamily are involved in multidrug resistance. TAP1 is involved in the pumping of degraded cytosolic peptides across the endoplasmic reticulum into the membrane-bound compartment where class I molecules assemble. TAP1 also acts as a molecular scaffold for the final stage of MHC class I folding. Nascent MHC class I molecules associate with TAP via tapasin. TAP1 is inhibited by the covalent attachment of herpes simplex virus ICP47 protein, which blocks the peptide-binding site of TAP. TAP1 is also inhibited by human cytomegalovirus US6 glycoprotein, which binds to the luminal side of the TAP complex and inhibits peptide translocation by specifically blocking ATP-binding to TAP1 and prevents the conformational rearrangement of TAP induced by peptide binding. TAP1 is further inhibited by human adenovirus E3-19K glycoprotein, which binds the TAP complex and acts as a tapasin inhibitor, preventing MHC class I/TAP association. Expression of TAP1 is

down-regulated by human Epstein-Barr virus vIL-10 protein, thereby affecting the transport of peptides into the endoplasmic reticulum and subsequent peptide loading by MHC class I molecules. Mutations in *Tap1* may be associated with ankylosing spondylitis, insulin-dependent diabetes mellitus, celiac disease, Bare Lymphocyte Syndrome (BLS), Type I, herpes simplex infection, autosomal recessive inheritance, bronchiectasis, bronchiolitis, chronic otitis media, chronic sinusitis, etc. TAP1-related pathways include, *e.g.*, antigen processing-cross presentation (such as ER-Phagosome pathway), Class I MHC mediated antigen processing and presentation, regulation of activated PAK-2p34 by proteasome mediated degradation, allograft rejection, Influenza A infection, immune System and Type II interferon signaling (IFNG), and NF-kappaB signaling. TAP1 has been shown to interact with HLA-A (Paulsson *et al.* (2002) *J. Biol. Chem.* 277:18266-18271), and Tapasin (Raghuraman *et al.* (2002) *J. Biol. Chem.* 277:41786-41794). TAP1 also forms heterodimers with TAP2 and interacts with Epstein-Barr virus BNLF2a, herpes simplex virus ICP47, PSMB5, and PSMB8 (Fruh *et al.* (1995) *Nature* 375:415-418; Begley *et al.* (2005) *Mol. Immunol.* 42:137-141; Horst *et al.* (2009) *J. Immunol.* 182:2313-2324).

The nucleic acid and amino acid sequences of a representative human TAP1 is available to the public at the GenBank database (Gene ID 6890) and is shown in Table 1 (*e.g.*, GenBank database numbers NM_000593.5 and NP_000584.2, which represent the longer transcript variant 1 and the encoded longer isoform 1; and NM_001292022.1 and NP_001278951.1, which represent the shorter transcript variant 2, which differs in the 5' UTR and coding sequence compared to variant 1, and the encode isoform 2 shorter at the N-terminus compared to isoform 1.). The domain structure of TAP1 polypeptide is well known and accessible in UniProtKB database under the accession number Q03518, including, in the order from the 5' terminus to the 3' terminus, a type I ABC transmembrane domain, *e.g.*, amino acid positions 247-530 of NP_000584.2, and an ABC transporter domain comprising, *e.g.*, amino acid positions 563-802 of NP_000584.2.

Nucleic acid and polypeptide sequences of TAP1 orthologs in organisms other than humans are well-known and include, for example, chimpanzee TAP1 (XM_001166911.5 and XP_001166911.2), Rhesus monkey TAP1 (XM_001115506.3 and XP_001115506.1), dog TAP1 (NM_001284496.2 and NP_001271425.1), cattle TAP1 (NM_001098058.1 and NP_001091527.1), mouse TAP1 (NM_013683.2 and NP_038711.2, representing the longer transcript variant 1 and the encoded longer isoform 1; and NM_001161730.1 and NP_001155202.1, representing the shorter transcript variant 2 lacking an alternate in-frame

segment compared to variant 1 and the encoded isoform 2 with the same N- and C-termini but being shorter compared to isoform 1), chicken TAP1 (NM_001135968.1 and NP_001129440.1), rat TAP1 (NM_032055.4 and NP_114444.2), and zebrafish TAP1 (XM_002665007.2 and XP_002665053.1).

5 The term “TAP1 activity” includes the ability of a TAP1 polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein) to bind its substrate (*e.g.*, ATP, HLA-A, Tapasin, TAP2, etc.) and/or catalyze antigen presentation activity. TAP1 activity may also include one or more of functions, such as those disclosed herein in MHC class I folding, dimerization with TAP2, and other signaling pathways.

10 The term “TAP1 substrate(s)” refers to binding partners of a TAP1 polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein), *e.g.*, ATP, HLA-A, Tapasin, TAP2, and others as disclosed herein. Furthermore, TAP1 substates may refer to downstream members in the signaling pathways where TAP1 has a functional role.

 The term “TAP1-regulated signaling pathway(s)” includes signaling pathways in
15 which TAP1 (and its fragments, domains, and/or motifs thereof, discussed herein) binds to at least one of its substrate, through which at least one cellular function and/or activity and/or cellular protein profiles is changed (*e.g.*, antigen processing and presentation). In some embodiments, TAP1 promotes antigen presentation (including ER-Phagosome pathway), regulation of activated PAK-2p34 by proteasome mediated degradation (*e.g.*,
20 ABC transporters), Class I MHC mediated antigen processing and presentation, allograft rejection, and/or virus assembly and release.

 The term “TAP1 inhibitor(s)” includes any natural or non-natural agent prepared, synthesized, manufactured, and/or purified by human that is capable of reducing, inhibiting, blocking, preventing, and/or that inhibits the ability of a TAP1 polypeptide (and its
25 fragments, domains, and/or motifs thereof, discussed herein). In one embodiment, such inhibitors may reduce or inhibit the binding/interaction between TAP1 and its substrates or other binding partners. In another embodiment, such inhibitors may reduce or inhibit TAP1 functions. In one embodiment, such inhibitors may increase or promote the turnover rate, reduce or inhibit the expression and/or the stability (*e.g.*, the half-life), and/or change the
30 cellular localization of TAP1, resulting in at least a decrease in TAP1 levels and/or activity. Such inhibitors may be any molecule, including but not limited to small molecule compounds, antibodies or intrabodies, RNA interfering (RNAi) agents (including at least siRNAs, shRNAs, microRNAs (miRNAs), piwi, and other well-known agents). Such

inhibitors may be specific to TAP1 or also inhibit at least one of other related proteins (*e.g.*, ATP-binding cassette (ABC) transporters). For example, Gromme *et al.* ((1997) *Eur. J. Immunol.* 27:898-904) teach rational design of TAP inhibitors by peptide substrate modification and peptidomimetics. RNA interference for TAP1 polypeptides are well known and commercially available (*e.g.*, human or mouse shRNA products (Cat. # TG308949, TR502210, TF711649, etc.) and siRNA products (Cat # SR321961, SR417657, SR504188, etc.) and CRISPR knockout products (Cat. # KN204242 and KN317175) from Origene (Rockville, MD), siRNA/shRNA products (Cat. # sc-42981 and sc-42982) and human or mouse gene knockout kit via CRISPR (Cat. # sc-423265 and sc-405210) from Santa Cruz Biotechnology (Dallas, Texas), and Ready-to-package AAV shRNA clones from Vigene Biosciences (Rockville, MD, Cat. # SH807891), etc.). Methods for detection, purification, and/or inhibition of TAP1 (*e.g.*, by anti-TAP1 antibodies) are also well known and commercially available (*e.g.*, multiple anti-TAP1 antibodies from Origene (Cat. # AP13226PU-N, R1028, TA311131, TA332358, etc.), Novus Biologicals (Littleton, CO, Cat. # NBP1-54435, H00006890-M01, H00006890-M04, NBP2-55934, etc.), abcam (Cambridge, MA, Cat. # ab13516, ab83817, ab10356, ab66074, etc.), Santa Cruz Biotechnology (Cat. # sc-376796, sc-20930, sc-11465, etc.), etc.). TAP1 knockout human cell lines are also well known and available at the Horizon (Cambridge, UK, Cat. # HZGHC6890, etc.).

The term "TAP2," also known as transporter 2, ATP Binding Cassette Subfamily B (MDR/TAP) Member, and transporter associated with antigen processing 2, refers to a member of a group of membrane-bound proteins of the superfamily of ATP-binding cassette (ABC) transporters. ABC proteins transport various molecules across extra- and intra-cellular membranes to the endoplasmic reticulum for association with MHC class I molecules. ABC genes are divided into seven distinct subfamilies (ABC1, MDR/TAP, MRP, ALD, OABP, GCN20, White). TAP2, same as TAP1, is a member of the MDR/TAP subfamily, involved in multidrug resistance and antigen presentation. *Tap2* is located 7 kb telomeric to gene family member ABCB2 (*Tap1*). TAP2 forms a heterodimer with TAP1 in order to transport peptides from the cytoplasm to the endoplasmic reticulum. Mutations in TAP2 may be associated with ankylosing spondylitis, insulin-dependent diabetes mellitus, schizophrenia, and celiac disease (Yu *et al.* (2014) *J. Psychiatr. Res.* 50:73-78). Alternative splicing of *Tap2* produces two products which differ in peptide selectivity and level of restoration of surface expression of MHC class I molecules. TAP2 is involved in

the transport of antigens from the cytoplasm to the endoplasmic reticulum for association with MHC class I molecules. TAP2 also acts as a molecular scaffold for the final stage of MHC class I folding. Nascent MHC class I molecules associate with TAP via tapasin.

TAP2 is inhibited by the covalent attachment of herpes simplex virus ICP47 protein, which
 5 blocks the peptide-binding site of TAP. TAP2 is also inhibited by human cytomegalovirus US6 glycoprotein, which binds to the luminal side of the TAP complex and inhibits peptide translocation by specifically blocking ATP-binding to TAP1 and prevents the conformational rearrangement of TAP induced by peptide binding. TAP2 is also inhibited by human adenovirus E3-19K glycoprotein, which binds the TAP complex and acts as a
 10 tapasin inhibitor, preventing MHC class I/TAP association. Mutations in *Tap2* may be associated with ankylosing spondylitis, insulin-dependent diabetes mellitus, celiac disease, Bare Lymphocyte Syndrome (BLS), Type I, herpes simplex infection, spondylitis, cardiac sarcoidosis, and posterior uveal melanoma. TAP2-related pathways include, *e.g.*, antigen processing-cross presentation (such as ER-Phagosome pathway), Class I MHC mediated
 15 antigen processing and presentation, regulation of activated PAK-2p34 by proteasome mediated degradation (*e.g.*, ABC transporters), allograft rejection, Influenza A infection, and NF-kappaB signaling. TAP2 interacts with, at least, TAP1, BLNF2a, ICP47, HLA-C, TAPBP, ABCA1, and ABCB7.

The nucleic acid and amino acid sequences of a representative human TAP2 is
 20 available to the public at the GenBank database (Gene ID 6891) and is shown in Table 1 (*e.g.*, GenBank database numbers NM_000544.3 and NP_000535.3, which represent the longer transcript variant 1 (B allele) and the encoded longest isoform 1; NM_001290043.1 and NP_001276972.1, which represent transcript variant 1 (A allele), which differs at 3 nt positions compared to variant 1, B allele, and the encoded isoform 2, which is shorter at the
 25 C-terminus compared to isoform 1; and NM_018833.2 and NP_061313.2, which represent transcript variant 2, which differs in the 5' UTR and coding region compared to variant 1, and the encoded isoform 3, which is shorter and has a distinct C-terminus compared to isoform 1). The domain structure of TAP2 polypeptide is well known and accessible in UniProtKB database under the accession number Q03519, including, in the order from the
 30 5' terminus to the 3' terminus, a type I ABC transmembrane domain comprising, *e.g.*, amino acid positions 152-435 of NP_000535.3, and an ABC transporter domain comprising, *e.g.*, amino acid positions 468-686 of NP_000535.3.

Nucleic acid and polypeptide sequences of TAP2 orthologs in organisms other than humans are well-known and include, for example, chimpanzee TAP2 (XM_009450990.2 and XP_009449265.2), tropical clawed frog TAP2 (XM_002942603.4 and XP_002942649.3), dog TAP2 (NM_001284495.1 and NP_001271424.1), cattle TAP2
 5 (NM_174222.2 and NP_776647.1), mouse TAP2 (NM_011530.3 and NP_035660.3), chicken TAP2 (NM_001099357.1 and NP_001092827.1), rat TAP2 (NM_032056.3 and NP_114445.2), and zebrafish TAP2 (NM_001006594.1 and NP_001006594.1).

The term “TAP2 activity” includes the ability of a TAP2 polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein) to bind its substrate (*e.g.*,
 10 ATP, BLNF2a, ICP47, HLA-C, TAPBP, ABCA1, ABCB7, etc.) and/or catalyze antigen presentation activity. TAP2 activity may also include one or more of functions, such as those disclosed herein in MHC class I folding, dimerization with TAP1, and other signaling pathways.

The term “TAP2 substrate(s)” refers to binding partners of a TAP2 polypeptide (and
 15 its fragments, domains, and/or motifs thereof, discussed herein), *e.g.*, ATP, BLNF2a, ICP47, HLA-C, TAPBP, ABCA1, ABCB7, and others as disclosed herein. Furthermore, TAP2 substrates may refer to downstream members in the signaling pathways where TAP2 has a functional role.

The term “TAP2-regulated signaling pathway(s)” includes signaling pathways in
 20 which TAP2 (and its fragments, domains, and/or motifs thereof, discussed herein) binds to at least one of its substrate, through which at least one cellular function and/or activity and/or cellular protein profiles is changed (*e.g.*, antigen processing and presentation). In some embodiments, TAP2 promotes antigen presentation (including ER-Phagosome pathway), regulation of activated PAK-2p34 by proteasome mediated degradation (*e.g.*,
 25 ABC transporters), Class I MHC mediated antigen processing and presentation, allograft rejection, virus assembly and release, and NF-kappaB signaling.

The term “TAP2 inhibitor(s)” includes any natural or non-natural agent prepared, synthesized, manufactured, and/or purified by human that is capable of reducing, inhibiting,
 30 blocking, preventing, and/or that inhibits the ability of a TAP2 polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein). In one embodiment, such inhibitors may reduce or inhibit the binding/interaction between TAP2 and its substrates or other binding partners. In another embodiment, such inhibitors may reduce or inhibit TAP2 functions. In one embodiment, such inhibitors may increase or promote the turnover rate,

reduce or inhibit the expression and/or the stability (*e.g.*, the half-life), and/or change the cellular localization of TAP2, resulting in at least a decrease in TAP2 levels and/or activity. Such inhibitors may be any molecule, including but not limited to small molecule compounds, antibodies or intrabodies, RNA interfering (RNAi) agents (including at least

5 siRNAs, shRNAs, microRNAs (miRNAs), piwi, and other well-known agents). Such inhibitors may be specific to TAP2 or also inhibit at least one of other related proteins (*e.g.*, ATP-binding cassette (ABC) transporters). For example, Gromme *et al.* ((1997) *Eur. J. Immunol.* 27:898-904) teach rational design of TAP inhibitors by peptide substrate modification and peptidomimetics. RNA interference for TAP2 polypeptides are well

10 known and commercially available (*e.g.*, human or mouse shRNA products (Cat. # TF301249, TR516069, etc.) and siRNA products (Cat. # SR503075, SR304711, SR316944, SR427405, etc.) and CRISPR knockout products (Cat. # KN222611 and KN317176) from Origene (Rockville, MD), siRNA/shRNA products (Cat. # sc-42983 and sc-42984) and human or mouse gene knockout kit via CRISPR (Cat. # sc-423266 and sc-404218-KO-2)

15 from Santa Cruz Biotechnology (Dallas, Texas), and Ready-to-package AAV shRNA clones from Vigene Biosciences (Rockville, MD, Cat. # SH836791 and SH812040), etc.). Methods for detection, purification, and/or inhibition of TAP2 (*e.g.*, by anti-TAP2 antibodies) are also well known and commercially available (*e.g.*, multiple anti-TAP2 antibodies from Origene (Cat. # TA317404, TA327099, TA351786, TA351787, etc.),

20 Novus Biologicals (Littleton, CO, Cat. # NBP1-54436, NBP1-89312, NBP1-40335, etc.), abcam (Cambridge, MA, Cat. # ab180611, ab130414, ab191067, etc.), Santa Cruz Biotechnology (Cat. # sc-25612, sc-515576, etc.), etc.). TAP2 knockout human cell lines are also well known and available at the Horizon (Cambridge, UK, Cat. # HZGHC6891, etc.).

25 The term "TAPBP," also known as TAP binding protein and Tapasin, refers to a member of a group of transmembrane glycoproteins which are present in the lumen of the endoplasmic reticulum and mediate interaction between newly assembled major histocompatibility complex (MHC) class I molecules and the transporter associated with antigen processing (TAP), which is required for the transport of antigenic peptides across

30 the endoplasmic reticulum membrane. This interaction facilitates optimal peptide loading on the MHC class I molecule. Up to four complexes of MHC class I and tapasin may be bound to a single TAP molecule. The peptide-loading complex consists of TAP, tapasin, MHC class I, calreticulin, and ERp57. Tapasin recruits MHC class I molecules to the TAP

peptide transporter, and also enhances loading of MHC class I with high-affinity peptides. Following loading of MHC class I with a high-affinity ligand, the interaction between tapasin and MHC class I disappears (Zhang and Williams (2006) *Immunol. Res.* 35:151-162). Tapasin contains a C-terminal double-lysine motif (KKKAE) known to maintain

5 membrane proteins in the endoplasmic reticulum. In humans, the tapasin gene lies within the major histocompatibility complex on chromosome 6. Alternative splicing results in three transcript variants encoding different isoforms. Tapasin has been shown to interact with HLA-A (Paulsson *et al.* (2002) *J. Biol. Chem.* 277:18266-18271), and TAP1 (Raghuraman *et al.* (2002) *J. Biol. Chem.* 277:41786-41794). Other predicted interacting

10 partners for TAPBP include HLA-B, HLA-C, and PDIA3. Diseases associated with Tapasin/TAPBP include Bare Lymphocyte Syndrome, Type I, maxillary sinus squamous cell carcinoma, paranasal sinus cancer (*e.g.*, adenoid cystic carcinoma of accessory sinus), and orbital melanoma. Among its related pathways are antigen processing-cross presentation (such as ER-Phagosome pathway), Class I MHC mediated antigen processing

15 and presentation, and immune system and allograft rejection. An important paralog of this gene is TAPBPL.

The nucleic acid and amino acid sequences of a representative human TAPBP is available to the public at the GenBank database (Gene ID 6892) and is shown in Table 1 (*e.g.*, GenBank database numbers NM_003190.4 and NP_003181.3, which represent the

20 transcript variant 1 and the encoded isoform 1; NM_172208.2 and NP_757345.2, which represent the transcript variant 2, which differs in the 3' coding region and 3' UTR, compared to variant 1, and the encoded isoform 2 having a distinct C-terminus than isoform 1; and NM_172209.2 and NP_757346.2, which represent the transcript variant 3, which lacks an alternate in-frame exon in the central coding region, compared to variant 1, and the

25 encoded isoform 3 that is shorter than isoform 1). The domain structure of TAPBP polypeptide is well known and accessible in UniProtKB database under the accession number O15533, including, in the order from the 5' terminus to the 3' terminus, a signal peptide comprising, *e.g.*, amino acid positions 1-20 of NP_003181.3, and a Tapasin domain comprising, *e.g.*, amino acid positions 21-448 of NP_003181.3.

30 Nucleic acid and polypeptide sequences of TAPBP orthologs in organisms other than humans are well-known and include, for example, chimpanzee TAPBP (XM_009451012.2 and XP_009449287.1, XM_009451014.2 and XP_009449289.1, XM_009451016.2 and XP_009449291.1, XM_009451017.2 and XP_009449292.1, and

XM_001170532.4 and XP_001170532.1), Rhesus monkey TAPBP (NM_001261727.2 and NP_001248656.1), dog TAPBP (NM_001048101.2 and NP_001041566.1), cattle TAPBP (NM_001045885.1 and NP_001039350.1), mouse TAPBP (NM_001025313.1 and NP_001020484.1, representing the longer transcript variant 1 and the encoded longer isoform 1; and NM_009318.2 and NP_033344.1, representing the shorter transcript variant 2 with an alternate splice site in the coding region and the encoded isoform 2 lacks an internal aa, compared to isoform 1), chicken TAPBP (NM_001034816.3 and NP_001029988.2, and NM_001206611.2 and NP_001193540.1), tropical clawed frog TAPBP (XM_002938682.4 and XP_002938728.2), rat TAPBP (NM_033098.2 and NP_149089.1), and zebrafish TAPBP (NM_130974.1 and NP_571049.1).

The term “TAPBP activity” includes the ability of a TAPBP polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein) to bind its substrate (*e.g.*, TAP1, HLA-A, HLA-B, HLA-C, PDIA3, etc.) and/or catalyze antigen presentation activity. TAPBP activity may also include one or more of functions, such as those disclosed herein in antigen processing-cross presentation (such as ER-Phagosome pathway), Class I MHC mediated antigen processing and presentation, and immune system and allograft rejection and other signaling pathways.

The term “TAPBP substrate(s)” refers to binding partners of a TAPBP polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein), *e.g.*, TAP1, HLA-A, HLA-B, HLA-C, PDIA3, and others as disclosed herein. Furthermore, TAPBP substrates may refer to downstream members in the signaling pathways where TAPBP has a functional role.

The term “TAPBP-regulated signaling pathway(s)” includes signaling pathways in which TAPBP (and its fragments, domains, and/or motifs thereof, discussed herein) binds to at least one of its substrate, through which at least one cellular function and/or activity and/or cellular protein profiles is changed (*e.g.*, antigen processing and presentation). In some embodiments, TAPBP promotes antigen processing-cross presentation (such as ER-Phagosome pathway), Class I MHC mediated antigen processing and presentation, immune system and allograft rejection, and other signaling pathways.

The term “TAPBP inhibitor(s)” includes any natural or non-natural agent prepared, synthesized, manufactured, and/or purified by human that is capable of reducing, inhibiting, blocking, preventing, and/or that inhibits the ability of a TAPBP polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein). In one embodiment, such

inhibitors may reduce or inhibit the binding/interaction between TAPBP and its substrates or other binding partners. In another embodiment, such inhibitors may reduce or inhibit TAPBP functions. In one embodiment, such inhibitors may increase or promote the turnover rate, reduce or inhibit the expression and/or the stability (*e.g.*, the half-life), and/or change the cellular localization of TAPBP, resulting in at least a decrease in TAPBP levels and/or activity. Such inhibitors may be any molecule, including but not limited to small molecule compounds, antibodies or intrabodies, RNA interfering (RNAi) agents (including at least siRNAs, shRNAs, microRNAs (miRNAs), piwi, and other well-known agents). Such inhibitors may be specific to TAPBP or also inhibit at least one of other related proteins (*e.g.*, TAP-associated proteins). RNA interference for TAPBP polypeptides are well known and commercially available (*e.g.*, human or mouse shRNA products (Cat. # TF301248) and siRNA products (Cat. # SR304712, SR324363, SR426898, etc.) and CRISPR knockout products (Cat. # KN317177 and KN217968) from Origene (Rockville, MD), siRNA/shRNA products (Cat. # sc-42986 and sc-42987) and human or mouse gene knockout kit via CRISPR (Cat. # sc-423267) from Santa Cruz Biotechnology (Dallas, Texas), and Ready-to-package AAV shRNA clones from Vigene Biosciences (Rockville, MD, Cat. # SH846826 and SH829641), etc.). Methods for detection, purification, and/or inhibition of TAPBP (*e.g.*, by anti-TAPBP antibodies) are also well known and commercially available (*e.g.*, multiple anti-TAPBP antibodies from Origene (Cat. # TA332478 and AP46045PU-N), Novus Biologicals (Littleton, CO, Cat. # NBP1-59071, NBP1-59070, NBP1-59938, etc.), abcam (Cambridge, MA, Cat. # ab13518, ab196764, and ab104806), Santa Cruz Biotechnology (Cat. # sc-136435), etc.). TAPBP knockout human cell lines are also well known and available at the Horizon (Cambridge, UK, Cat. # HZGHC6892, etc.).

The term “ERAP1,” also known as ER aminopeptidase 1 and Type 1 tumor necrosis factor receptor shedding aminopeptidase regulator or endoplasmic reticulum aminopeptidase 1 (ARTS-1), refers to a member of group of aminopeptidase involved in trimming HLA class I-binding precursors so that they can be presented on MHC class I molecules. ERAP1 acts as a monomer or as a heterodimer with ERAP2. It may also be involved in blood pressure regulation by inactivation of angiotensin II. Three transcript variants encoding two different isoforms have been found for *Erap1*. ERAP1 is an aminopeptidase that plays a central role in peptide trimming, a step required for the generation of most HLA class I-binding peptides. Peptide trimming is essential to

customize longer precursor peptides to fit them to the correct length required for presentation on MHC class I molecules. ERAP1 strongly prefers substrates 9-16 residues long and rapidly degrades 13-mer to a 9-mer and then stops. ERAP1 preferentially hydrolyzes the residue Leu and peptides with a hydrophobic C-terminus, while it has weak activity toward peptides with charged C-terminus, including Met, Cys, and Phe (Nguyen *et al.* (2011) *Nat. Struct. Mol. Biol.* 18:604-613). In addition, ERAP1 cleaves several cytokine receptors on the surface of cells and reduces their ability to transmit chemical signals into the cell, which affects the process of inflammation. ERAP1/ARTS-1 is a member of the M1 family of zinc metallopeptidases which acts as an aminopeptidase that degrades oligopeptides by cleavage starting at the amino terminus. One of the functions of aminopeptidases is to degrade potentially toxic peptides in the cytosol. ERAP1/ARTS-1 is a transmembrane protein that is localized to the endoplasmic reticulum. ERAP1 may play a role in the inactivation of peptide hormones and may be involved in the regulation of blood pressure through the inactivation of angiotensin II and/or the generation of bradykinin in the kidney. Mutations in the ERAP1/ARTS-1 have been linked to an increased risk of ankylosing spondylitis but only in HLA-B27 positive patients (Brionez and Reveille (2008) *Curr. Opin. Rheumatol.* 20:384-391). Other related diseases and disorders include, *e.g.*, birdshot chorioretinopathy, behcet syndrome, spondylitis, and essential hypertension. Tapasin has been suggested to interact with HIST3H3, HLA-B, IL6R, TNFRSF1A, and ACO2. Among its related pathways are antigen processing-cross presentation (folding, assembly, and peptide loading of class I MHC), Class I MHC mediated antigen processing and presentation, TNFR1 pathway (*e.g.*, apoptosis and survival TNFR1 signaling), and immune system rejection.

The nucleic acid and amino acid sequences of a representative human ERAP1 is available to the public at the GenBank database (Gene ID 51752) and is shown in Table 1 (*e.g.*, GenBank database numbers NM_016442.4 and NP_057526.3, which represent the longest transcript variant 1 and the encoded longer isoform a; NM_001349244.1 and NP_001336173.1, which represent the transcript variant 4, which differs in the 5' UTR, compared to variant 1, and the encoded same isoform a; NM_001040458.2 and NP_001035548.1, which represent the transcript variant 2, which differs in the 3' UTR and coding sequence compared to variant 1, and the encoded isoform b with a shorter and distinct C-terminus compared to isoform a; and NM_001198541.2 and NP_001185470.1, representing the transcript variant 3, which differs in the 3' UTR and coding sequence and

in the 5' UTR compared to variant 1, and the same encoded isoform b). The domain structure of ERAP1 polypeptide is well known and accessible in UniProtKB database under the accession number Q9NZ08, including, in the order from the 5' terminus to the 3' terminus, a transmembrane region comprising, *e.g.*, amino acid positions 2-21 of NP_057526.3, a peptidase M1 region comprising, *e.g.*, amino acid positions 53-441 of NP_057526.3, and an ERAP1-like C-terminal domain comprising, *e.g.*, amino acid positions 598-914 of NP_057526.3.

Nucleic acid and polypeptide sequences of ERAP1 orthologs in organisms other than humans are well-known and include, for example, chimpanzee ERAP1 (XM_009449387.2 and XP_009447662.2, XM_003310766.4 and XP_003310814.2, XM_016953521.1 and XP_016809010.1, XM_016953519.1 and XP_016809008.1, XM_016953520.1 and XP_016809009.1, XM_009449383.2 and XP_009447658.2, XM_009449384.2 and XP_009447659.2, XM_016953523.1 and XP_016809012.1, and XM_016953522.1 and XP_016809011.1), Rhesus monkey ERAP1 (XM_001094790.3 and XP_001094790.2), dog ERAP1 (XM_546015.5 and XP_546015.3, XM_005618086.2 and XP_005618143.1, XM_005618087.2 and XP_005618144.1, and XM_005618085.2 and XP_005618142.1), cattle ERAP1 (NM_001102003.1 and NP_001095473.1), mouse ERAP1 (NM_030711.4 and NP_109636.1), chicken ERAP1 (XM_001232417.4 and XP_001232418.2), tropical clawed frog ERAP1 (XM_002933429.2 and XP_002933475.2, and XM_012968052.2 and XP_012823506.1), and rat ERAP1 (NM_030836.1 and NP_110463.1).

The term "ERAP1 activity" includes the ability of an ERAP1 polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein) to bind its substrate (*e.g.*, ERAP2, HIST3H3, HLA-B, IL6R, TNFRSF1A, ACO2, etc.) and/or catalyze antigen presentation activity. ERAP1 activity may also include one or more of functions, such as those disclosed herein in antigen processing-cross presentation (folding, assembly, and peptide loading of class I MHC), Class I MHC mediated antigen processing and presentation, TNFR1 pathway (*e.g.*, apoptosis and survival TNFR1 signaling), immune system rejection, and other signaling pathways.

The term "ERAP1 substrate(s)" refers to binding partners of an ERAP1 polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein), *e.g.*, ERAP2, HIST3H3, HLA-B, IL6R, TNFRSF1A, ACO2, and others as disclosed herein.

Furthermore, ERAP1 substates may refer to downstream members in the signaling pathways where ERAP1 has a functional role.

The term “ERAP1-regulated signaling pathway(s)” includes signaling pathways in which ERAP1 (and its fragments, domains, and/or motifs thereof, discussed herein) binds to at least one of its substrate, through which at least one cellular function and/or activity and/or cellular protein profiles is changed (*e.g.*, antigen processing and presentation). In some embodiments, ERAP1 promotes antigen processing-cross presentation (folding, assembly, and peptide loading of class I MHC), Class I MHC mediated antigen processing and presentation, TNFR1 pathway (*e.g.*, apoptosis and survival TNFR1 signaling), immune system rejection, and other signaling pathways.

The term “ERAP1 inhibitor(s)” includes any natural or non-natural agent prepared, synthesized, manufactured, and/or purified by human that is capable of reducing, inhibiting, blocking, preventing, and/or that inhibits the ability of an ERAP1 polypeptide (and its fragments, domains, and/or motifs thereof, discussed herein). In one embodiment, such inhibitors may reduce or inhibit the binding/interaction between ERAP1 and its substrates or other binding partners. In another embodiment, such inhibitors may reduce or inhibit ERAP1 functions. In one embodiment, such inhibitors may increase or promote the turnover rate, reduce or inhibit the expression and/or the stability (*e.g.*, the half-life), and/or change the cellular localization of ERAP1, resulting in at least a decrease in ERAP1 levels and/or activity. Such inhibitors may be any molecule, including but not limited to small molecule compounds, antibodies or intrabodies, RNA interfering (RNAi) agents (including at least siRNAs, shRNAs, microRNAs (miRNAs), piwi, and other well-known agents). Such inhibitors may be specific to ERAP1 or also inhibit at least one of other related proteins (*e.g.*, aminopeptidases). For example, Węglarz-Tomczaka *et al.* teach discovery of multiple compound inhibitors for ERAP1 and ERAP2 by screening libraries of phosphorus-containing amino acid and dipeptide analogues (Węglarz-Tomczaka *et al.* (2016) *Bioorg. Med. Chem. Lett.* 26:4122-4126). RNA interference for ERAP1 polypeptides are well known and commercially available (*e.g.*, human or mouse shRNA products (Cat. # TG307017, TF505369, TF711082, etc.) and siRNA products (Cat. # SR309948, SR421275, SR505552, etc.) and CRISPR knockout products (Cat. # KN205249) from Origene (Rockville, MD), siRNA/shRNA products (Cat. # sc-43577 and sc-44435) and human or mouse gene knockout kit via CRISPR (Cat. # sc-429840) from Santa Cruz Biotechnology (Dallas, Texas), and Ready-to-package AAV shRNA clones

from Vigene Biosciences (Rockville, MD, Cat. # SH838580 and SH865692, etc.). Methods for detection, purification, and/or inhibition of ERAP1 (*e.g.*, by anti-ERAP1 antibodies) are also well known and commercially available (*e.g.*, multiple anti-ERAP1 antibodies from Origene (Cat. # TA350992, TA315501, TA341243, etc.), Novus Biologicals (Littleton, CO, Cat. # MAB2334, AF2334, NBP2-41232, H00051752-M01, etc.), abcam (Cambridge, MA, Cat. # ab124669, ab181390, ab150627, etc.), Santa Cruz Biotechnology (Cat. # sc-100727 and sc-271823), etc.). ERAP1 knockout human cell lines are also well known and available at the Horizon (Cambridge, UK, Cat. # HZGHC51752, etc.).

The term “immune response” includes T cell mediated and/or B cell mediated immune responses. Exemplary immune responses include T cell responses, *e.g.*, cytokine production and cellular cytotoxicity. In addition, the term immune response includes immune responses that are indirectly effected by T cell activation, *e.g.*, antibody production (humoral responses) and activation of cytokine responsive cells, *e.g.*, macrophages.

The term “immunotherapeutic agent” can include any molecule, peptide, antibody or other agent which can stimulate a host immune system to generate an immune response to a tumor or cancer in the subject. Various immunotherapeutic agents are useful in the compositions and methods described herein.

The term “inhibit” includes the decrease, limitation, or blockage, of, for example a particular action, function, or interaction. In some embodiments, cancer is “inhibited” if at least one symptom of the cancer is alleviated, terminated, slowed, or prevented. As used herein, cancer is also “inhibited” if recurrence or metastasis of the cancer is reduced, slowed, delayed, or prevented.

The term “interaction”, when referring to an interaction between two molecules, refers to the physical contact (*e.g.*, binding) of the molecules with one another. Generally, such an interaction results in an activity (which produces a biological effect) of one or both of said molecules.

An “isolated protein” refers to a protein that is substantially free of other proteins, cellular material, separation medium, and culture medium when isolated from cells or produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. An “isolated” or “purified” protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the antibody, polypeptide, peptide or fusion protein is derived, or substantially free from chemical precursors or other chemicals when chemically

synthesized. The language “substantially free of cellular material” includes preparations of a biomarker polypeptide or fragment thereof, in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language “substantially free of cellular material” includes preparations of a biomarker protein or fragment thereof, having less than about 30% (by dry weight) of non-biomarker protein (also referred to herein as a “contaminating protein”), more preferably less than about 20% of non-biomarker protein, still more preferably less than about 10% of non-biomarker protein, and most preferably less than about 5% non-biomarker protein. When antibody, polypeptide, peptide or fusion protein or fragment thereof, *e.g.*, a biologically active fragment thereof, is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

As used herein, the term “isotype” refers to the antibody class (*e.g.*, IgM, IgG1, IgG2C, and the like) that is encoded by heavy chain constant region genes.

As used herein, the term “K_D” is intended to refer to the dissociation equilibrium constant of a particular antibody-antigen interaction. The binding affinity of antibodies of the disclosed invention may be measured or determined by standard antibody-antigen assays, for example, competitive assays, saturation assays, or standard immunoassays such as ELISA or RIA.

A “kit” is any manufacture (*e.g.* a package or container) comprising at least one reagent, *e.g.* a probe or small molecule, for specifically detecting and/or affecting the expression of a marker of the present invention. The kit may be promoted, distributed, or sold as a unit for performing the methods of the present invention. The kit may comprise one or more reagents necessary to express a composition useful in the methods of the present invention. In certain embodiments, the kit may further comprise a reference standard, *e.g.*, a nucleic acid encoding a protein that does not affect or regulate signaling pathways controlling cell growth, division, migration, survival or apoptosis. One skilled in the art can envision many such control proteins, including, but not limited to, common molecular tags (*e.g.*, green fluorescent protein and beta-galactosidase), proteins not classified in any of pathway encompassing cell growth, division, migration, survival or apoptosis by GeneOntology reference, or ubiquitous housekeeping proteins. Reagents in the kit may be provided in individual containers or as mixtures of two or more reagents in a

single container. In addition, instructional materials which describe the use of the compositions within the kit can be included.

The term “neoadjuvant therapy” refers to a treatment given before the primary treatment. Examples of neoadjuvant therapy can include chemotherapy, radiation therapy, and hormone therapy. For example, in treating breast cancer, neoadjuvant therapy can allows patients with large breast cancer to undergo breast-conserving surgery.

The “normal” level of expression of a biomarker is the level of expression of the biomarker in cells of a subject, *e.g.*, a human patient, not afflicted with a cancer. An “over-expression” or “significantly higher level of expression” of a biomarker refers to an expression level in a test sample that is greater than the standard error of the assay employed to assess expression, and is preferably at least 10%, and more preferably 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10, 10.5, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 times or more higher than the expression activity or level of the biomarker in a control sample (*e.g.*, sample from a healthy subject not having the biomarker associated disease) and preferably, the average expression level of the biomarker in several control samples. A “significantly lower level of expression” of a biomarker refers to an expression level in a test sample that is at least 10%, and more preferably 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10, 10.5, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 times or more lower than the expression level of the biomarker in a control sample (*e.g.*, sample from a healthy subject not having the biomarker associated disease) and preferably, the average expression level of the biomarker in several control samples.

An “over-expression” or “significantly higher level of expression” of a biomarker refers to an expression level in a test sample that is greater than the standard error of the assay employed to assess expression, and is preferably at least 10%, and more preferably 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10, 10.5, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 times or more higher than the expression activity or level of the biomarker in a control sample (*e.g.*, sample from a healthy subject not having the biomarker associated disease) and preferably, the average expression level of the biomarker in several control samples. A “significantly lower level of expression” of a biomarker refers to an expression level in a test sample that is at least 10%, and more preferably 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0,

2.1, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10, 10.5, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 times or more lower than the expression level of the biomarker in a control sample (*e.g.*, sample from a healthy subject not having the biomarker associated disease) and preferably, the average expression level of the biomarker in several control samples.

The term “pre-determined” biomarker amount and/or activity measurement(s) may be a biomarker amount and/or activity measurement(s) used to, by way of example only, evaluate a subject that may be selected for a particular treatment, evaluate a response to a treatment such as inhibitor(s) of the regulators of one or more biomarkers listed in Table 1, such as one or more antigen processing and presentation as described herein/immunotherapy combination therapy, and/or evaluate the disease state. A pre-determined biomarker amount and/or activity measurement(s) may be determined in populations of patients with or without cancer. The pre-determined biomarker amount and/or activity measurement(s) can be a single number, equally applicable to every patient, or the pre-determined biomarker amount and/or activity measurement(s) can vary according to specific subpopulations of patients. Age, weight, height, and other factors of a subject may affect the pre-determined biomarker amount and/or activity measurement(s) of the individual. Furthermore, the pre-determined biomarker amount and/or activity can be determined for each subject individually. In one embodiment, the amounts determined and/or compared in a method described herein are based on absolute measurements. In another embodiment, the amounts determined and/or compared in a method described herein are based on relative measurements, such as ratios (*e.g.*, serum biomarker normalized to the expression of housekeeping or otherwise generally constant biomarker). The pre-determined biomarker amount and/or activity measurement(s) can be any suitable standard. For example, the pre-determined biomarker amount and/or activity measurement(s) can be obtained from the same or a different human for whom a patient selection is being assessed. In one embodiment, the pre-determined biomarker amount and/or activity measurement(s) can be obtained from a previous assessment of the same patient. In such a manner, the progress of the selection of the patient can be monitored over time. In addition, the control can be obtained from an assessment of another human or multiple humans, *e.g.*, selected groups of humans, if the subject is a human. In such a manner, the extent of the selection of the human for whom selection is being assessed can be compared to suitable other humans, *e.g.*, other humans who are in a similar situation to

the human of interest, such as those suffering from similar or the same condition(s) and/or of the same ethnic group.

The term “predictive” includes the use of a biomarker nucleic acid and/or protein status, *e.g.*, over- or under- activity, emergence, expression, growth, remission, recurrence or resistance of tumors before, during or after therapy, for determining the likelihood of response of a cancer to inhibitor(s) of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein/immunotherapy combination therapy (*e.g.*, treatment with a combination of such inhibitor and an immunotherapy, such as an immune checkpoint inhibitor). Such predictive use of the biomarker may be confirmed by, *e.g.*, (1) increased or decreased copy number (*e.g.*, by FISH, FISH plus SKY, single-molecule sequencing, *e.g.*, as described in the art at least at J. Biotechnol., 86:289-301, or qPCR), overexpression or underexpression of a biomarker nucleic acid (*e.g.*, by ISH, Northern Blot, or qPCR), increased or decreased biomarker protein (*e.g.*, by IHC), or increased or decreased activity, *e.g.*, in more than about 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 20%, 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 100%, or more of assayed human cancers types or cancer samples; (2) its absolute or relatively modulated presence or absence in a biological sample, *e.g.*, a sample containing tissue, whole blood, serum, plasma, buccal scrape, saliva, cerebrospinal fluid, urine, stool, or bone marrow, from a subject, *e.g.* a human, afflicted with cancer; (3) its absolute or relatively modulated presence or absence in clinical subset of patients with cancer (*e.g.*, those responding to a particular inhibitor/immunotherapy combination therapy or those developing resistance thereto).

The term “pre-malignant lesions” as described herein refers to a lesion that, while not cancerous, has potential for becoming cancerous. It also includes the term “pre-malignant disorders” or “potentially malignant disorders.” In particular this refers to a benign, morphologically and/or histologically altered tissue that has a greater than normal risk of malignant transformation, and a disease or a patient's habit that does not necessarily alter the clinical appearance of local tissue but is associated with a greater than normal risk of precancerous lesion or cancer development in that tissue (leukoplakia, erythroplakia, erytroleukoplakia lichen planus (lichenoid reaction) and any lesion or an area which histological examination showed atypia of cells or dysplasia. In one embodiment, a metaplasia is a pre-malignant lesion.

The terms “prevent,” “preventing,” “prevention,” “prophylactic treatment,” and the like refer to reducing the probability of developing a disease, disorder, or condition in a subject, who does not have, but is at risk of or susceptible to developing a disease, disorder, or condition.

5 The term “probe” refers to any molecule which is capable of selectively binding to a specifically intended target molecule, for example, a nucleotide transcript or protein encoded by or corresponding to a biomarker nucleic acid. Probes can be either synthesized by one skilled in the art, or derived from appropriate biological preparations. For purposes of detection of the target molecule, probes may be specifically designed to be labeled, as
10 described herein. Examples of molecules that can be utilized as probes include, but are not limited to, RNA, DNA, proteins, antibodies, and organic molecules.

 The term “prognosis” includes a prediction of the probable course and outcome of cancer or the likelihood of recovery from the disease. In some embodiments, the use of statistical algorithms provides a prognosis of cancer in an individual. For example, the
15 prognosis can be surgery, development of a clinical subtype of cancer (*e.g.*, solid tumors, such as esophageal cancer and gastric cancer), development of one or more clinical factors, or recovery from the disease.

 The term “response to immunotherapy” or “response to inhibitor(s) of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and
20 presentation as described herein /immunotherapy combination therapy” relates to any response of the hyperproliferative disorder (*e.g.*, cancer) to an anti-cancer agent, such as an inhibitor of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), and an immunotherapy, preferably to a change in tumor mass
25 and/or volume after initiation of neoadjuvant or adjuvant therapy. Hyperproliferative disorder response may be assessed, for example for efficacy or in a neoadjuvant or adjuvant situation, where the size of a tumor after systemic intervention can be compared to the initial size and dimensions as measured by CT, PET, mammogram, ultrasound or palpation. Responses may also be assessed by caliper measurement or pathological examination of the
30 tumor after biopsy or surgical resection. Response may be recorded in a quantitative fashion like percentage change in tumor volume or in a qualitative fashion like “pathological complete response” (pCR), “clinical complete remission” (cCR), “clinical partial remission” (cPR), “clinical stable disease” (cSD), “clinical progressive disease”

(cPD) or other qualitative criteria. Assessment of hyperproliferative disorder response may be done early after the onset of neoadjuvant or adjuvant therapy, *e.g.*, after a few hours, days, weeks or preferably after a few months. A typical endpoint for response assessment is upon termination of neoadjuvant chemotherapy or upon surgical removal of residual tumor cells and/or the tumor bed. This is typically three months after initiation of neoadjuvant therapy. In some embodiments, clinical efficacy of the therapeutic treatments described herein may be determined by measuring the clinical benefit rate (CBR). The clinical benefit rate is measured by determining the sum of the percentage of patients who are in complete remission (CR), the number of patients who are in partial remission (PR) and the number of patients having stable disease (SD) at a time point at least 6 months out from the end of therapy. The shorthand for this formula is $CBR = CR + PR + SD$ over 6 months. In some embodiments, the CBR for a particular cancer therapeutic regimen is at least 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, or more. Additional criteria for evaluating the response to cancer therapies are related to “survival,” which includes all of the following: survival until mortality, also known as overall survival (wherein said mortality may be either irrespective of cause or tumor related); “recurrence-free survival” (wherein the term recurrence shall include both localized and distant recurrence); metastasis free survival; disease free survival (wherein the term disease shall include cancer and diseases associated therewith). The length of said survival may be calculated by reference to a defined start point (*e.g.*, time of diagnosis or start of treatment) and end point (*e.g.*, death, recurrence or metastasis). In addition, criteria for efficacy of treatment can be expanded to include response to chemotherapy, probability of survival, probability of metastasis within a given time period, and probability of tumor recurrence. For example, in order to determine appropriate threshold values, a particular cancer therapeutic regimen can be administered to a population of subjects and the outcome can be correlated to biomarker measurements that were determined prior to administration of any cancer therapy. The outcome measurement may be pathologic response to therapy given in the neoadjuvant setting. Alternatively, outcome measures, such as overall survival and disease-free survival can be monitored over a period of time for subjects following cancer therapy for which biomarker measurement values are known. In certain embodiments, the doses administered are standard doses known in the art for cancer therapeutic agents. The period of time for which subjects are monitored can vary. For example, subjects may be monitored for at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 45, 50, 55, or 60

months. Biomarker measurement threshold values that correlate to outcome of a cancer therapy can be determined using well-known methods in the art, such as those described in the Examples section.

5 The term “resistance” refers to an acquired or natural resistance of a cancer sample or a mammal to a cancer therapy (*i.e.*, being nonresponsive to or having reduced or limited response to the therapeutic treatment), such as having a reduced response to a therapeutic treatment by 25% or more, for example, 30%, 40%, 50%, 60%, 70%, 80%, or more, to 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 15-fold, 20-fold or more. The reduction in response can be measured by comparing with the same cancer sample or mammal before the
10 resistance is acquired, or by comparing with a different cancer sample or a mammal that is known to have no resistance to the therapeutic treatment. A typical acquired resistance to chemotherapy is called “multidrug resistance.” The multidrug resistance can be mediated by P-glycoprotein or can be mediated by other mechanisms, or it can occur when a mammal is infected with a multi-drug-resistant microorganism or a combination of microorganisms.
15 The determination of resistance to a therapeutic treatment is routine in the art and within the skill of an ordinarily skilled clinician, for example, can be measured by cell proliferative assays and cell death assays as described herein as “sensitizing.” In some embodiments, the term “reverses resistance” means that the use of a second agent in combination with a primary cancer therapy (*e.g.*, chemotherapeutic or radiation therapy) is able to produce a
20 significant decrease in tumor volume at a level of statistical significance (*e.g.*, $p < 0.05$) when compared to tumor volume of untreated tumor in the circumstance where the primary cancer therapy (*e.g.*, chemotherapeutic or radiation therapy) alone is unable to produce a statistically significant decrease in tumor volume compared to tumor volume of untreated tumor. This generally applies to tumor volume measurements made at a time when the
25 untreated tumor is growing log rhythmically.

The terms “response” or “responsiveness” refers to an anti-cancer response, *e.g.* in the sense of reduction of tumor size or inhibiting tumor growth. The terms can also refer to an improved prognosis, for example, as reflected by an increased time to recurrence, which is the period to first recurrence censoring for second primary cancer as a first event or death
30 without evidence of recurrence, or an increased overall survival, which is the period from treatment to death from any cause. To respond or to have a response means there is a beneficial endpoint attained when exposed to a stimulus. Alternatively, a negative or detrimental symptom is minimized, mitigated or attenuated on exposure to a stimulus. It

will be appreciated that evaluating the likelihood that a tumor or subject will exhibit a favorable response is equivalent to evaluating the likelihood that the tumor or subject will not exhibit favorable response (*i.e.*, will exhibit a lack of response or be non-responsive).

An “RNA interfering agent” as used herein, is defined as any agent which interferes
5 with or inhibits expression of a target biomarker gene by RNA interference (RNAi). Such RNA interfering agents include, but are not limited to, nucleic acid molecules including RNA molecules which are homologous to the target biomarker gene of the present invention, or a fragment thereof, short interfering RNA (siRNA), and small molecules which interfere with or inhibit expression of a target biomarker nucleic acid by RNA
10 interference (RNAi).

“RNA interference (RNAi)” is an evolutionally conserved process whereby the expression or introduction of RNA of a sequence that is identical or highly similar to a target biomarker nucleic acid results in the sequence specific degradation or specific post-transcriptional gene silencing (PTGS) of messenger RNA (mRNA) transcribed from that
15 targeted gene (*see* Coburn and Cullen (2002) *J. Virol.* 76:9225), thereby inhibiting expression of the target biomarker nucleic acid. In one embodiment, the RNA is double stranded RNA (dsRNA). This process has been described in plants, invertebrates, and mammalian cells. In nature, RNAi is initiated by the dsRNA-specific endonuclease Dicer, which promotes processive cleavage of long dsRNA into double-stranded fragments termed
20 siRNAs. siRNAs are incorporated into a protein complex that recognizes and cleaves target mRNAs. RNAi can also be initiated by introducing nucleic acid molecules, *e.g.*, synthetic siRNAs or RNA interfering agents, to inhibit or silence the expression of target biomarker nucleic acids. As used herein, “inhibition of target biomarker nucleic acid expression” or “inhibition of marker gene expression” includes any decrease in expression or protein
25 activity or level of the target biomarker nucleic acid or protein encoded by the target biomarker nucleic acid. The decrease may be of at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or 99% or more as compared to the expression of a target biomarker nucleic acid or the activity or level of the protein encoded by a target biomarker nucleic acid which has not been targeted by an RNA interfering agent.

30 The term “sample” used for detecting or determining the presence or level of at least one biomarker is typically brain tissue, cerebrospinal fluid, whole blood, plasma, serum, saliva, urine, stool (*e.g.*, feces), tears, and any other bodily fluid (*e.g.*, as described above under the definition of “body fluids”), or a tissue sample (*e.g.*, biopsy) such as a small

intestine, colon sample, or surgical resection tissue. In certain instances, the method of the present invention further comprises obtaining the sample from the individual prior to detecting or determining the presence or level of at least one marker in the sample.

The term “sensitize” means to alter cancer cells or tumor cells in a way that allows
5 for more effective treatment of the associated cancer with a cancer therapy (*e.g.*, anti-immune checkpoint, chemotherapeutic, and/or radiation therapy). In some embodiments, normal cells are not affected to an extent that causes the normal cells to be unduly injured by the therapies. An increased sensitivity or a reduced sensitivity to a therapeutic treatment is measured according to a known method in the art for the particular treatment and
10 methods described herein below, including, but not limited to, cell proliferative assays (Tanigawa N, Kern D H, Kikasa Y, Morton D L, Cancer Res 1982; 42: 2159-2164), cell death assays (Weisenthal L M, Shoemaker R H, Marsden J A, Dill P L, Baker J A, Moran E M, Cancer Res 1984; 94: 161-173; Weisenthal L M, Lippman M E, Cancer Treat Rep 1985; 69: 615-632; Weisenthal L M, In: Kaspers G J L, Pieters R, Twentyman P R, Weisenthal L
15 M, Veerman A J P, eds. Drug Resistance in Leukemia and Lymphoma. Langhorne, P A: Harwood Academic Publishers, 1993: 415-432; Weisenthal L M, Contrib Gynecol Obstet 1994; 19: 82-90). The sensitivity or resistance may also be measured in animal by measuring the tumor size reduction over a period of time, for example, 6 month for human and 4-6 weeks for mouse. A composition or a method sensitizes response to a therapeutic
20 treatment if the increase in treatment sensitivity or the reduction in resistance is 25% or more, for example, 30%, 40%, 50%, 60%, 70%, 80%, or more, to 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 15-fold, 20-fold or more, compared to treatment sensitivity or resistance in the absence of such composition or method. The determination of sensitivity or resistance to a therapeutic treatment is routine in the art and within the skill of an ordinarily skilled
25 clinician. It is to be understood that any method described herein for enhancing the efficacy of a cancer therapy can be equally applied to methods for sensitizing hyperproliferative or otherwise cancerous cells (*e.g.*, resistant cells) to the cancer therapy.

“Short interfering RNA” (siRNA), also referred to herein as “small interfering
RNA” is defined as an agent which functions to inhibit expression of a target biomarker
30 nucleic acid, *e.g.*, by RNAi. An siRNA may be chemically synthesized, may be produced by *in vitro* transcription, or may be produced within a host cell. In one embodiment, siRNA is a double stranded RNA (dsRNA) molecule of about 15 to about 40 nucleotides in length, preferably about 15 to about 28 nucleotides, more preferably about 19 to about 25

nucleotides in length, and more preferably about 19, 20, 21, or 22 nucleotides in length, and may contain a 3' and/or 5' overhang on each strand having a length of about 0, 1, 2, 3, 4, or 5 nucleotides. The length of the overhang is independent between the two strands, *i.e.*, the length of the overhang on one strand is not dependent on the length of the overhang on the second strand. Preferably the siRNA is capable of promoting RNA interference through degradation or specific post-transcriptional gene silencing (PTGS) of the target messenger RNA (mRNA).

In another embodiment, an siRNA is a small hairpin (also called stem loop) RNA (shRNA). In one embodiment, these shRNAs are composed of a short (*e.g.*, 19-25 nucleotide) antisense strand, followed by a 5-9 nucleotide loop, and the analogous sense strand. Alternatively, the sense strand may precede the nucleotide loop structure and the antisense strand may follow. These shRNAs may be contained in plasmids, retroviruses, and lentiviruses and expressed from, for example, the pol III U6 promoter, or another promoter (*see, e.g.*, Stewart, *et al.* (2003) *RNA* Apr;9(4):493-501 incorporated by reference herein).

RNA interfering agents, *e.g.*, siRNA molecules, may be administered to a patient having or at risk for having cancer, to inhibit expression of a biomarker gene which is overexpressed in cancer and thereby treat, prevent, or inhibit cancer in the subject.

The term "small molecule" is a term of the art and includes molecules that are less than about 1000 molecular weight or less than about 500 molecular weight. In one embodiment, small molecules do not exclusively comprise peptide bonds. In another embodiment, small molecules are not oligomeric. Exemplary small molecule compounds which can be screened for activity include, but are not limited to, peptides, peptidomimetics, nucleic acids, carbohydrates, small organic molecules (*e.g.*, polyketides) (Cane *et al.* (1998) *Science* 282:63), and natural product extract libraries. In another embodiment, the compounds are small, organic non-peptidic compounds. In a further embodiment, a small molecule is not biosynthetic.

The term "specific binding" refers to antibody binding to a predetermined antigen. Typically, the antibody binds with an affinity (K_D) of approximately less than 10^{-7} M, such as approximately less than 10^{-8} M, 10^{-9} M or 10^{-10} M or even lower when determined by surface plasmon resonance (SPR) technology in a BIAcore® assay instrument using an antigen of interest as the analyte and the antibody as the ligand, and binds to the predetermined antigen with an affinity that is at least 1.1-, 1.2-, 1.3-, 1.4-, 1.5-, 1.6-, 1.7-,

1.8-, 1.9-, 2.0-, 2.5-, 3.0-, 3.5-, 4.0-, 4.5-, 5.0-, 6.0-, 7.0-, 8.0-, 9.0-, or 10.0-fold or greater than its affinity for binding to a non-specific antigen (*e.g.*, BSA, casein) other than the predetermined antigen or a closely-related antigen. The phrases “an antibody recognizing an antigen” and “an antibody specific for an antigen” are used interchangeably herein with
5 the term “an antibody which binds specifically to an antigen.” Selective binding is a relative term referring to the ability of an antibody to discriminate the binding of one antigen over another.

The term “subject” refers to any healthy animal, mammal or human, or any animal, mammal or human afflicted with a cancer, *e.g.*, brain, lung, ovarian, pancreatic, liver,
10 breast, prostate, and/or colorectal cancers, melanoma, multiple myeloma, and the like. The term “subject” is interchangeable with “patient.”

The term “survival” includes all of the following: survival until mortality, also known as overall survival (wherein said mortality may be either irrespective of cause or tumor related); “recurrence-free survival” (wherein the term recurrence shall include both
15 localized and distant recurrence); metastasis free survival; disease free survival (wherein the term disease shall include cancer and diseases associated therewith). The length of said survival may be calculated by reference to a defined start point (*e.g.* time of diagnosis or start of treatment) and end point (*e.g.* death, recurrence or metastasis). In addition, criteria for efficacy of treatment can be expanded to include response to chemotherapy, probability
20 of survival, probability of metastasis within a given time period, and probability of tumor recurrence.

The term “synergistic effect” refers to the combined effect of two or more anti-cancer agents (*e.g.*, inhibitor(s) of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein /immunotherapy
25 combination therapy) can be greater than the sum of the separate effects of the anti-cancer agents/therapies alone.

The term “T cell” includes CD4⁺ T cells and CD8⁺ T cells. The term T cell also includes both T helper 1 type T cells and T helper 2 type T cells. The term “antigen presenting cell” includes professional antigen presenting cells (*e.g.*, B lymphocytes,
30 monocytes, dendritic cells, Langerhans cells), as well as other antigen presenting cells (*e.g.*, keratinocytes, endothelial cells, astrocytes, fibroblasts, and oligodendrocytes).

The term “therapeutic effect” refers to a local or systemic effect in animals, particularly mammals, and more particularly humans, caused by a pharmacologically active

substance. The term thus means any substance intended for use in the diagnosis, cure, mitigation, treatment or prevention of disease or in the enhancement of desirable physical or mental development and conditions in an animal or human. The phrase “therapeutically-effective amount” means that amount of such a substance that produces some desired local or systemic effect at a reasonable benefit/risk ratio applicable to any treatment. In certain embodiments, a therapeutically effective amount of a compound will depend on its therapeutic index, solubility, and the like. For example, certain compounds discovered by the methods of the present invention may be administered in a sufficient amount to produce a reasonable benefit/risk ratio applicable to such treatment.

The terms “therapeutically-effective amount” and “effective amount” as used herein means that amount of a compound, material, or composition comprising a compound of the present invention which is effective for producing some desired therapeutic effect in at least a sub-population of cells in an animal at a reasonable benefit/risk ratio applicable to any medical treatment. Toxicity and therapeutic efficacy of subject compounds may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD₅₀ and the ED₅₀. Compositions that exhibit large therapeutic indices are preferred. In some embodiments, the LD₅₀ (lethal dosage) can be measured and can be, for example, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1000% or more reduced for the agent relative to no administration of the agent. Similarly, the ED₅₀ (*i.e.*, the concentration which achieves a half-maximal inhibition of symptoms) can be measured and can be, for example, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1000% or more increased for the agent relative to no administration of the agent. Also, Similarly, the IC₅₀ (*i.e.*, the concentration which achieves half-maximal cytotoxic or cytostatic effect on cancer cells) can be measured and can be, for example, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1000% or more increased for the agent relative to no administration of the agent. In some embodiments, cancer cell growth in an assay can be inhibited by at least about 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or even 100%. In another embodiment, at least about a 10% , 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or even 100% decrease in a solid malignancy can be achieved.

A “transcribed polynucleotide” or “nucleotide transcript” is a polynucleotide (*e.g.* an mRNA, hnRNA, a cDNA, or an analog of such RNA or cDNA) which is complementary to or homologous with all or a portion of a mature mRNA made by transcription of a biomarker nucleic acid and normal post-transcriptional processing (*e.g.* splicing), if any, of the RNA transcript, and reverse transcription of the RNA transcript.

As used herein, the term “unresponsiveness” includes refractivity of cancer cells to therapy or refractivity of therapeutic cells, such as immune cells, to stimulation, *e.g.*, stimulation via an activating receptor or a cytokine. Unresponsiveness can occur, *e.g.*, because of exposure to immunosuppressants or exposure to high doses of antigen. As used herein, the term “anergy” or “tolerance” includes refractivity to activating receptor-mediated stimulation. Such refractivity is generally antigen-specific and persists after exposure to the tolerizing antigen has ceased. For example, anergy in T cells (as opposed to unresponsiveness) is characterized by lack of cytokine production, *e.g.*, IL-2. T cell anergy occurs when T cells are exposed to antigen and receive a first signal (a T cell receptor or CD-3 mediated signal) in the absence of a second signal (a costimulatory signal). Under these conditions, reexposure of the cells to the same antigen (even if reexposure occurs in the presence of a costimulatory polypeptide) results in failure to produce cytokines and, thus, failure to proliferate. Anergic T cells can, however, proliferate if cultured with cytokines (*e.g.*, IL-2). For example, T cell anergy can also be observed by the lack of IL-2 production by T lymphocytes as measured by ELISA or by a proliferation assay using an indicator cell line. Alternatively, a reporter gene construct can be used. For example, anergic T cells fail to initiate IL-2 gene transcription induced by a heterologous promoter under the control of the 5' IL-2 gene enhancer or by a multimer of the AP1 sequence that can be found within the enhancer (Kang *et al.* (1992) *Science* 257:1134).

There is a known and definite correspondence between the amino acid sequence of a particular protein and the nucleotide sequences that can code for the protein, as defined by the genetic code (shown below). Likewise, there is a known and definite correspondence between the nucleotide sequence of a particular nucleic acid and the amino acid sequence encoded by that nucleic acid, as defined by the genetic code.

GENETIC CODE

Alanine (Ala, A)	GCA, GCC, GCG, GCT
Arginine (Arg, R)	AGA, ACG, CGA, CGC, CGG, CGT
Asparagine (Asn, N)	AAC, AAT

	Aspartic acid (Asp, D)	GAC, GAT
	Cysteine (Cys, C)	TGC, TGT
	Glutamic acid (Glu, E)	GAA, GAG
	Glutamine (Gln, Q)	CAA, CAG
5	Glycine (Gly, G)	GGA, GGC, GGG, GGT
	Histidine (His, H)	CAC, CAT
	Isoleucine (Ile, I)	ATA, ATC, ATT
	Leucine (Leu, L)	CTA, CTC, CTG, CTT, TTA, TTG
	Lysine (Lys, K)	AAA, AAG
10	Methionine (Met, M)	ATG
	Phenylalanine (Phe, F)	TTC, TTT
	Proline (Pro, P)	CCA, CCC, CCG, CCT
	Serine (Ser, S)	AGC, AGT, TCA, TCC, TCG, TCT
	Threonine (Thr, T)	ACA, ACC, ACG, ACT
15	Tryptophan (Trp, W)	TGG
	Tyrosine (Tyr, Y)	TAC, TAT
	Valine (Val, V)	GTA, GTC, GTG, GTT
	Termination signal (end)	TAA, TAG, TGA

An important and well-known feature of the genetic code is its redundancy,
 whereby, for most of the amino acids used to make proteins, more than one coding
 nucleotide triplet may be employed (illustrated above). Therefore, a number of different
 nucleotide sequences may code for a given amino acid sequence. Such nucleotide
 sequences are considered functionally equivalent since they result in the production of the
 same amino acid sequence in all organisms (although certain organisms may translate some
 sequences more efficiently than they do others). Moreover, occasionally, a methylated
 variant of a purine or pyrimidine may be found in a given nucleotide sequence. Such
 methylations do not affect the coding relationship between the trinucleotide codon and the
 corresponding amino acid.

In view of the foregoing, the nucleotide sequence of a DNA or RNA encoding a
 biomarker nucleic acid (or any portion thereof) can be used to derive the polypeptide amino
 acid sequence, using the genetic code to translate the DNA or RNA into an amino acid
 sequence. Likewise, for polypeptide amino acid sequence, corresponding nucleotide
 sequences that can encode the polypeptide can be deduced from the genetic code (which,

because of its redundancy, will produce multiple nucleic acid sequences for any given amino acid sequence). Thus, description and/or disclosure herein of a nucleotide sequence which encodes a polypeptide should be considered to also include description and/or disclosure of the amino acid sequence encoded by the nucleotide sequence. Similarly, description and/or disclosure of a polypeptide amino acid sequence herein should be considered to also include description and/or disclosure of all possible nucleotide sequences that can encode the amino acid sequence.

Finally, nucleic acid and amino acid sequence information for the loci and biomarkers of the present invention (*e.g.*, biomarkers listed in Tables 1 and 2) are well-known in the art and readily available on publicly available databases, such as the National Center for Biotechnology Information (NCBI). For example, exemplary nucleic acid and amino acid sequences derived from publicly available sequence databases are provided below and include, for example, PCT Publ. WO 2014/022759, which is incorporated herein in its entirety by this reference.

Table 1

SEQ ID NO: 1 Human HLA-E cDNA Sequence (NM_005516.5, CDS region from position 127-1203)

20	1	gcagactcag	ttctcattcc	caatgggtgt	cgggttttcta	gagaagccaa	tcagcgctgc
	61	cacgactccc	gactataaag	tccccatccg	gactcaagaa	gttctcagga	ctcagaggct
	121	gggatcatgg	tagatggaac	cctcctttta	ctcctctcgg	aggccctggc	ccttaccag
	181	acctgggcgg	gctcccaactc	cttgaagtat	ttccacactt	ccgtgtcccg	gccccggcgc
	241	ggggagcccc	gcttcatctc	tgtgggctac	gtggacgaca	cccagttcgt	gcgcttcgac
25	301	aacgacgcgg	cgagtcggag	gatgggtgccg	cgggcgcggc	ggatggagca	ggaggggtca
	361	gagtattggg	accgggagac	acggagcgcc	agggacaccg	cacagatttt	ccgagtgaat
	421	ctgcggacgc	tgcgcggcta	ctacaatcag	agcgaggccg	ggtctcacac	cctgcagtgg
	481	atgcatggct	gcgagctggg	gcccgcaggg	cgcttcctcc	gcgggtatga	acagttcgcc
	541	tacgacggca	aggattatct	caccctgaat	gaggacctgc	gctcctggac	cgcggtggac
30	601	acggcggttc	agatctccga	gcaaaagtca	aatgatgcct	ctgaggcgga	gcaccagaga
	661	gcctacctgg	aagacacatg	cgtggagtg	ctccacaaat	acctggagaa	ggggaaggag
	721	acgtgcttc	acctggagcc	cccaaagaca	cacgtgactc	accaccccat	ctctgacct
	781	gaggccaccc	tgaggtgctg	ggccctgggc	ttctaccctg	cggagatcac	actgacctgg
	841	cagcaggatg	gggagggcca	taccaggagc	acggagctcg	tggagaccag	gcctgcaggg
35	901	gatggaacct	tccagaagtg	ggcagctgtg	gtggtgcctt	ctggagagga	gcagagatac
	961	acgtgccatg	tgcagcatga	ggggctaccc	gagcccgtca	ccctgagatg	gaagccggct
	1021	tcccagccca	ccatccccat	cgtgggcatc	attgctggcc	tggttctcct	tggatctgtg
	1081	gtctctggag	ctgtggttgc	tgctgtgata	tggaggaaga	agagctcagg	tggaaaagga
	1141	gggagctact	ctaaggctga	gtggagcgac	agtgccccag	ggtctgagtc	tcacagcttg
40	1201	taaagcctga	gacagctgcc	ttgtgtgcga	ctgagatgca	cagctgcctt	gtgtgcgact
	1261	gagatgcagg	atttcctcac	gcctccccta	tgtgtcttag	gggactctgg	cttctctttt
	1321	tgcaagggcc	tctgaatctg	tctgtgtccc	tgtagcaca	atgtgaggag	gtagagaaac
	1381	agtccacctc	tgtgtctacc	atgaccccct	tcctcacact	gacctgtgtt	ccttccctgt
	1441	tctcttttct	attaaaaata	agaacctggg	cagagtgcgg	cagctcatgc	ctgtaatccc
45	1501	agcacttagg	gaggccgagg	agggcagatc	acgaggtcag	gagatcgaaa	ccatcctggc
	1561	taacacggtg	aaaccccgtc	tctactaaaa	aatacaaaaa	attagctggg	cgagagggca
	1621	cgggcctgta	gtcccagcta	ctcaggaggc	ggaggcagga	gaatggcgct	aacccgggag
	1681	gcggagggtt	cagtgcagca	ggattgtgct	actgcactcc	agcctgggtg	acagggtgaa

1741 acgccatctc aaaaaataaa aattgaaaaa taaaaaaaaga acctggatct caattttaatt
 1801 tttcatattc ttgcaatgaa atggacttga ggaagctaag atcatagcta gaaatacaga
 1861 taattccaca gcacatctct agcaaatTTta gcctattcct attctctagc ctatttcctta
 1921 ccacctgtaa tcttgaccat ataccttgga gttgaatatt gttttcatac tgctgtgggt
 5 1981 tgaatgttcc ctccaacact catgttgaga cttaatccct aatgtggcaa tactgaaagg
 2041 tggggccttt gagatgtgat tggatcgtaa ggctgtgcct tcattcatgg gttaatggat
 2101 taatgggtta tcacaggaat gggactggtg gctttataag aagaggaaaa gagaactgag
 2161 ctagcatgcc cagcccacag agagcctcca ctagagtgat gctaagtgga aatgtgaggt
 2221 gcagctgcca cagagggccc ccaccaggga aatgtctagt gtctagtgga tccaggccac
 10 2281 aggagagagt gccttgtgga gcgctgggag caggacctga ccaccaccag gaccccagaa
 2341 ctgtggagtc agtggcagca tgcagcggcc ccttgggaaa gctttaggca ccagcctgca
 2401 acccattcga gcagccacgt aggctgcacc cagcaaagcc acaggcacgg ggctacctga
 2461 ggccttgggg gcccaatccc tgctccagtg tgtccgtgag gcagcacacg aagtcaaaag
 2521 agattattct cttcccacag ataccttttc tctcccatga ccctttaaca gcactgtctt
 15 2581 cattcccctc accttcccag gctgatctga ggtaaaacttt gaagtaaaat aaaagctgtg
 2641 tttgagcatc atttgtatTTT caaaaaaaaa aaaaaaaaaa

SEQ ID NO: 2 Human HLA-E Amino Acid Sequence (NP_005507.3)

1 mvdgtlllll sealaltqtw agshslkyfh tsvsrpgrge prfisvgyvd dtqfvrfdnd
 20 61 aasprmvpra pwmeqegsey wdretsrard taqifrvnlr tlrqyynqse agshltqwmh
 121 gcelgpdgrf lrgyeqfayd gkdyltlnd lrswtavdta aqiseqksnd aseahqray
 181 ledtcvewlh kylekgketl lhleppkthv thhpisdhea tlrwalgfy paeitltwqq
 241 dgegthqgte lvetrpagdg tfqkwaavvv psgeeqrytc hvqheglpep vtlrwkpasq
 301 ptipivgiia glvllgsvvs gavvaaviwr kkssggkgs yskaewdsda qgseshsl
 25

SEQ ID NO: 3 Mouse H2-T23 cDNA Sequence (NM_010398.3, CDS region from position 33-1106)

1 attcaggttc ctcacagacc caggggggtga ggatgttgct ttttgccac ttgcttcagc
 61 tgctgggtcag cgccacagtc cgcagccaga gtagccaca ctgcgtgcgg tatttcacca
 30 121 ccgcgtgtc ccggcccggc ctccgggagc cccggttcat cattgtcggc tacgtggacg
 181 acacgcagtt cgtgcgcttc gacagcgacg cggaaaaatcc gaggatggag cctcggggcg
 241 ggtggattga gcaggagggg ccggagtatt gggagcggga gacttggaaa gccagggaca
 301 tggggaggaa cttcagagta aacctgagga ccctgctcgg ctactacaat cagagtaacg
 361 acgaatctca cacgctgcag tggatgtacg gctgcgacgt ggggcccgat gggcgctgc
 35 421 tccgcgggta ttgtcaggag gcctacgatg gccaggatta catctccctg aacgaggacc
 481 tgcgttcctg gaccgcgaat gacatagcct cacagatctc taagcacaag tcagaggcag
 541 tcgatgaggc ccaccaacag agggcatacc tgcaagggtcc ttgcgtggag tggctccata
 601 gatacctacg gctgggaaat gagacactgc agcgctcaga ccctccaaag gcacatgtga
 661 cccatcaccc tagatctgaa gatgaagtca ccctgaggtg ctgggccctg ggcttctacc
 40 721 ctgctgacat caccctgacc tggcagttga atggggagga gctgaccag gacatggagc
 781 ttgtggagac caggcctgca ggggatggaa ccttccagaa gtgggcagct gtcgtggtgc
 841 ctcttgggaa ggagcagtat tacacatgcc atgtgtacca tgaggggctg cctgagcccc
 901 tcaccctgag atgggagcct cctccatcca ctgtctccaa catggtaatc atagctgttc
 961 tggttgtcct tggagctgtg atcatccttg gagctgtggt ggcttttgtg atgaagagga
 45 1021 ggagacacat aggtgtaaaa ggatgctatg ctcatgttct aggcagcaag agcttcaga
 1081 cctctgactg gcctcagaag gcatgaaaat ccctgggggg gctggtgaga tggctcagt
 1141 ggtaagagca ctgactgctc ttctgaaggt ccagagttca aatcccagca accacatggt
 1201 ggctcacaa catccgtaac gagatctgac tccctcttct ggagtgtctg aagacagcta
 1261 caatgtactt acatataata aataaataaa taaataaata aataaataaa taaataa
 50

SEQ ID NO: 4 Mouse H2-T23 Amino Acid Sequence (NP_034528.1)

1 mllfahllql lvsatvptqs sphslryftt avsrpplgep rfiiivgyvdd tqfvrfdsd
 61 enprmeprar wieqegpeyw eretwkardm grnfrvnlr llgyynqsnd eshtlqwmgy
 121 cdvpgdgrll rgycqeydg qdyislndl rswtandias qiskhkseav deahqqrail
 55 181 ggpcvewlhr ylrlnetlq rsdppkahvt hhprsedevt lrcwalgfyp aditltwqln
 241 geeltqdmel vetrpagdgf fqkwaavvvp lgkeqyytch vyheglpepl tlrwepppst
 301 vsnmviiavl vvlgaviilg avvafvmkrr rhigvkgcya hvlgsksfqt sdwpqka

SEQ ID NO: 5 Mouse HLA-E/H2-T23 Pseudogene cDNA Sequence
(NM 001271005.1, CDS region from position 20-1024)

1 acagacccag ggagtgagga tgttgctttt tgcccacttg ctccagctgc tggtcagcgc
 5 61 cacagtcccg acccagagta gctcacactc gctgcggtat ttccacaccg tcgtatcccg
 121 gcccggcctc ggagagcccc gggttcatcat tgctcggtac gtggacgaca cgcagttcgt
 181 gcgcttcgac agcgactcgg agaatccgag gatggagcct cgggcgcggt ggattgagca
 241 ggagggggccg gagtattggg agcgggagac tcggaaagcc agggacatgg ggaggaactt
 301 cagagtaaac ctgaggaccc tgctcggcta ctacaatcag agtaaggacg aatctcacac
 361 gctgcagtggt atgtacggct gcgacgtggg gcccgatggg cgcctgctcc gcgggtattg
 10 421 tcaggaggcc tatgatggcc aggattacat ctccctgaac gaagacctgc gctcctggac
 481 cgcgaccaac ttagcctcgc atatctctaa gtgcaagtca gaggcggtcg atgaggccca
 541 ccaacagagg gcatacctgc aaggctcctt cgtggagtgg ctccatacat acctacagct
 601 gggaagtgag aactgctgc gctcagaccc tccaaaggca catgtgacct gtcaccccag
 661 acctgaaggt gatgtcaccg tgaggtgctg gccctggggc ttctatcctg ctgacatcac
 15 721 cctgacctgg cagttgaatg gggaggagct gaccaggac atggagtttg tggagaccag
 781 gcctgcaggg gatggaacct tccagaagtg ggcattctgt gtggtgcctc ttgggaagga
 841 gcagaattac acatgccatg tgtaccatga ggggctgcct gagcccctca ccctgagatg
 901 ggagcctcct ccatccactg tctccaacat ggtaatcatt gctgttcttg ttgtccttgg
 961 agctgtgata gtcattgttg ctgtgtgtgg ttttgtaatg aagaggagga aaacacaggt
 20 1021 gtaaaaggat gctatgctca tgttctaggg agcaagagct tccagacctc tgactggcct
 1081 cagaaggcat gacaatccct gggggccttg gactggacct gaggcacaag ggagaccggg
 1141 gttcaattcc tagcactgac atggcaggtg tctgttacta gtaacaaaca gttcatgtct
 1201 catgaatgta gacatgaatg cagctaaagc accaacatac ataattattaa accttataaa
 1261 aaataaaagg tgtatgagct ggacaaaca

SEQ ID NO: 6 Mouse HLA-E/H2-T23 Pseudogene Amino Acid Sequence
(NP 001257934.1)

1 mllfahllql lvsatvptqs sshslryfht vvsrpglgep rfiiivgyvdd tqfvrfdstds
 30 61 enprmeprar wieqegpeyw eretrkardm grnfrvnrlt llgyynqskd eshtlqwmvg
 121 cdvdpdgrll rgycqeydg qdyislndel rswtatnlas hiskckseav deahqgrayl
 181 ggpvcvwlht ylqlgsetll rsdppkahvt rhprpegdvt lrcwalgfyp aditltwqln
 241 geeltqdmef vetrpagdgt fqkwasvvvp lgkeqnytch vyheglpepl tlrwepppst
 301 vsnmviiavl vvlgaviviv avvaivmkrr ktqv

SEQ ID NO: 7 Human PDIA3 cDNA Sequence (NM 005313.4, CDS region from position 149-1666)

1 gccggggtttg ggggttggga cctccggctg cagggtccgcc tgggccagac gcgcgagcgc
 40 61 aagcagcggg ttagtggtcg cgcgcccagc ctccgcagtc ccagccgagc cgcgaccctt
 121 ccggccgctc ccaccccacc tcgcgcgcat gcgcctcgc cgcctagcgc tgttcccggg
 181 tgtggcgctg cttcttgccg cggcccgctt gcgcgctgcc tccgagctgc tagaactcac
 241 ggacgacaac ttcgagagtc gcatctccga cacgggctct gcgggcctca tgctcgtcga
 301 gttcttcgcc ccctgggtgtg gacactgcaa gagacttgca cctgagtatg aagctgcagc
 361 taccagatta aaaggaatag tcccattagc aaagggtgat tgcaactgca aactaacac
 45 421 ctgtaataaa tatggagtca gtggatatcc aaccctgaag atatttagag atggtgaaga
 481 agcaggtgct tatgatggac ctaggactgc tgatggaatt gtcagccact tgaagaagca
 541 ggcaggacca gcttcagtgc ctctcaggac tgaggaaaga tttaagaaat tcattagtga
 601 taaagatgcc tctatagtag gttttttcga tgattcattc agtgaggctc actccgagtt
 661 cctaaaagca gccagcaact tgagggataa ctaccgattt gcacatacga atgttgagtc
 50 721 tctggtgaac gagtatgatg ataattggaga gggatatcat ttatttcgtc cttcacatct
 781 cactaacaag tttgaggaca agactgtggc atatacagag caaaaaatga ccagtggcaa
 841 aattaaaaag tttatccagg aaaacatatt tggtatctgc ctcacatga cagaagacaa
 901 taaagatttg atacagggca aggacttact tattgcttac tatgatgtgg actatgaaaa
 961 gaacgctaaa ggttccaact actggagaaa cagggtaatg atggtggcaa agaaattcct
 55 1021 ggatgctggg cacaaactca actttgctgt agctagccgc aaaaccttta gccatgaact
 1081 ttctgatttt ggcttgagga gcaactgctg agagattcct gttgttgcta tcagaactgc
 1141 taaaggagag aagtttgtca tgcaggagga gttctcgcgt gatgggaagg ctctggagag
 1201 gttcctgcaat gattactttg atggcaatct gaagagatac ctgaagtctg aacctatccc
 1261 agagagcaat gatgggcctg tgaaggtagt ggtagcagag aattttgatg aaatagtga
 60 1321 taatgaaaat aaagatgtgc tgattgaatt ttatgccctt tgggtgtggtc actgtaagaa

1381 cctggagccc aagtataaag aacttggcga gaagctcagc aaagacccaa atatcgatcat
 1441 agccaagatg gatgccacag ccaatgatgt gccttctcca tatgaagtca gaggttttcc
 1501 taccatatac ttctctccag ccaacaagaa gctaaatcca aagaaatatg aagggtggccg
 1561 tgaattaagt gattttatta gctatctaca aagagaagct acaaaccccc ctgtaattca
 5 1621 agaagaaaaa cccaagaaga agaagaaggc acaggaggat ctctaaagca gtagccaaac
 1681 accactttgt aaaaggactc ttccatcaga gatgggaaaa ccattgggga ggactaggac
 1741 ccataatgga attattacct ctcagggccg agaggacaga atggatataa tctgaatcct
 1801 gttaaatttt ctctaaactg tttcttagct gcactgttta tggaaatacc aggaccagtt
 1861 tatgtttgtg gttttgggaa aaattatttg tgttggggga aatgttggg ggggtggggt
 10 1921 gagttggggg tattttctaa tttttttgt acatttggaa cagtgcacat aaatgagacc
 1981 cctttaaact gtcttatttt ccaccagatt gagaaccaga tgttctctac acactatcac
 2041 tgttcaatag agctttcttc agtgatggaa atgctctgta atctacactg ttcagtacag
 2101 gtagctacgg agcatctgaa atatggctag aaactacatt tttgttttga ttaatttaaa
 2161 tagccataca gttgctacca tactggtcac ggcagctgta gactgactgg gtccatagtt
 15 2221 catcacctca aaattctttc aaaatttatt atctctttct ctcttacct gtttatttcc
 2281 caggcctacc ctggtgatta gaacagctga agggcctttc ttgttaggct gtccatgccc
 2341 taaggatggg ttctgttta tccttgccac gcagctgagc ttactgcatg tttatatctc
 2401 ccaaggactg ttctctgctc agaaatgccc tgtcaagggt gtggcatcac gcagtttcat
 2461 ccaagtttgt tcaggaattg ctgacactgc tgggtgcagt tctatccctc aaagcctagg
 20 2521 gtgtggccct ttaacttccc ctctagtaga actgggaaaag gcagtacatc tcctagttat
 2581 gagtgaagca tccactttct tttgtaacaa tgaggaaacag aaaggataag actgaagagt
 2641 gatcttttgt ccaactaaac catttatctc cttttgtagg taaattcaaa tcctgccag
 2701 ttatagtttt cagtcaactg agaattccag gtaggagccc tacttttaggt gatectagga
 2761 actcctatgt tcagaaaaga attttcttcc tactatataa ttacagtatt tagctgtcaa
 25 2821 ttttaagatg aatttggtag agccttatag taaagtatgt atcttggtca cacacaaagc
 2881 ttggaaaaag taaaagatgt ctaaaccata atcttgtaac tcataacatc tgggctgggg
 2941 ccggggatc taattgttta gagagcccca aaagtggctc agatgtaaag ccagggttaa
 3001 gaaccatctg ccttggaaga ttaagggaa acatataaac ttgtacaaag gacaaaaaaa

30 **SEQ ID NO: 8 Human PDIA3 Amino Acid Sequence (NP_005304.3)**

1 mrlrrllalfp gvalllaaar laaasdvlel tddnfesris dtgsaglmvlv effapwgcghc
 61 krlapeyaaa atrlkgivpl akvdctantn tcnkygvsgy ptlkifrdge eagaydgprrt
 121 adgivshlkk qagpasvplr teeefkkfis dkdasivgff ddsfseahse flkaasnldr
 181 nyrfahtnve slvneyddng egiilfrpsh ltnkfedktv ayteqkmtsg kikkfiqeni
 35 241 fgicphmted nkdliqgkdl liayydvdye knakgsnywr nrvmnvakkf ldaghklnfa
 301 vasrktfshe lsdfglesta geipvvairt akgekfvme efsrdgkale rflqdyfdgn
 361 lkrylksepi pesndgpvk vvaenfdeiv nnenkdvlie fyapwgcghc nlepykelg
 421 eklskdpniv iakmdatand vpspyevrgf ptiyfspank klnpkkyegg relsdfisyl
 481 greatnppvi qeekpkkkkk aqedl

40

SEQ ID NO: 9 Mouse PDIA3 cDNA Sequence (NM_007952.2, CDS region from position 138-1655)

1 cgagcaggcc taggggggttg ggacctcggc agcgggtctg cccggggccag acgcgcgagc
 61 gcaggcaagc ggctgcagat tgccgggtct cccatctcat ttctccggtc ccggccctcc
 45 121 gaccgcgacc cgccgccatg cgcttcagct gcctagctct gctcccgggc gtggcgctgc
 181 tgctgcctc ggcccgctct gcccgccgct ccgatgtgtt ggaactgacg gacgaaaact
 241 tcgagagtcg cgtctccgac acgggctcgg cggggtctcat gctagtcgag ttcttcgccc
 301 cctgggtgtg acattgcaag aggtctgccc ctgagtatga agctgcagca accagattaa
 361 aaggaatagt cccattagca aagggtggatt gcaactgcaa cacaaacacc tgtaataagt
 50 421 atgggggtcag tggctaccca actcttaaga tcttttagaga tgggtgaagaa gcgggtgctt
 481 atgatgggccc taggactgct gatggaattg tcagccactt gaagaaacaa gcaggaccag
 541 cttcagttcc tctcaggact gaggaagaat ttaagaagtt cattagtgat aaagatgcct
 601 cagtgggtggg ttttttcagg gatttattca gtgatgggca ctctgaattc ctaaaagcag
 661 ccagcaactt gagagataac taccgatttg cacacaccaa cattgagtct ctggtgaagg
 55 721 agtacgatga taatggagag gggatcacta tattccgtcc attacatctt gctaacaagt
 781 ttgaagacaa aactgtggca tatactgaaa agaaaatgac cagtggcaag atcaagaaat
 841 ttattcagga tagcattttt ggtctctgtc ctcatatgac ggaagataat aaagatttga
 901 tacaaggcaa ggacttactc accgcttact atgatgtgga ctatgaaaag aatgctaaag
 961 gttctaaacta ctggagaaac agggatcatg tggtggcaaa gaaattcctt gatgctggac
 60 1021 acaaactcaa ctttgctgta gctagccgta aaacctttag ccatgaactg tcagattttg

1081 gcttagaaaag cactactgga gaggttcctg ttgtggctat cagaactgct aaaggagaga
 1141 agtttgtcat gcaggaggag ttctcgcgag atggcaaggc tcttgaacag ttcctgcaag
 1201 aatactttga tggcaacttg aagagatacc tgaagtctga acccatccca gagtccaacg
 1261 aagggcctgt caagggtgtg gtagcagaga attttgatga catagtgaat gaagaagata
 5 1321 aggacgtgct gattgaattt tacgccccctt ggtgtggcca ctgtaagaat ctggaaccca
 1381 agtataaaga gctgggagaa aaactcagca aagatccaaa tattgtcata gccaaagtgg
 1441 atgccacagc caatgatgtg ctttctccat atgaagtcaa ggggttttctt accatctact
 1501 tctcaccagc caacaagaag ctaactccaa agaagtatga aggtggccgt gaattaaatg
 1561 attttatttag ctatctacaa cgagaagcta caaaccccc tataattcaa gaagaaaaac
 10 1621 ctaagaagaa gaagaaggca caagaggacc tctaaagcaa cagccaaatg caccacttta
 1681 taaaaggact cttacaccag agaaggcaaa accagagagg acagaatgga tatactctga
 1741 atcctgttaa attttctcta aactgtttct tagctgcact gtttgaaaat acaaggacca
 1801 gtttatgttt gtggttttgg gagaaaatta tttgtgttgg gaagaatgtt gtgggggtgg
 1861 ggggaattga gttgggggtt tattttctaa tttttttgt acatttgga cagtacaat
 15 1921 aaatgtgccc cttttatact gtcgtttttt tcccctcatg ggatggagaa tcagggtctc
 1981 taaactgtaa tacagtaata gatagaagtg tttgagctat gctgttcaat gctgtagcta
 2041 aaccagatat gttatctcct ataatcgcaa cacttgggag gtgcaggcaa gagatctcaa
 2101 gttcaagatc ttccatggct acatagcagg ttttgaggct agcatcttaa acaggtttga
 2161 tctttcaaaa acaagaaaa ccagctatgc tgattcctgt aacttgagac tgaataggtc
 20 2221 cagttaccat ctccaaattg tataatttat tagtttcttt ctgtagctca agctgtcctt
 2281 gaacttgtga tgggtctccct ccctggccct gagatactga gatgacagtt gtttctagcc
 2341 tcattcatgt cctaaagggtg aatgcttact cttgccacac agccaagctt tctgcatgtg
 2401 catgcctcct gaggatggtt ctctgccag aaatgccctg tgtgagtgc agcagtttgg
 2461 ggttccatcc aatttttagga attattgaaa ctgctgggtg cagctctccc caaatcattg
 25 2521 ggtgtagcca tttacttccc tccagtgaac agggaaacag tatcaatcgc tggcaagaac
 2581 aagaaccac agagatgttt tcttttgaat ggattaaagt attgcctcc ttttatagct
 2641 aaaaaaaaa aaaaaaa

SEQ ID NO: 10 Mouse PDIA3 Amino Acid Sequence (NP_031978.2)

30 1 mrfscslallp gvallllasar laaasdvlel tdenfesrvs dtgsaglmvlv effapwgcghc
 61 krlapeyaaa atrlkgivpl akvdctantn tcnkygvsgy ptlkifrdge eagaydgprt
 121 adgivshlkk qagpasvplr teeefkkfis dkdasvvgff rdlsfdghse flkaasnldr
 181 nyrfahtnie slvkeyddng egitifrplh lankfedktv aytekkmtsg kikkfiqdsi
 241 fglcphmted nkdliqgkdl ltayydvdye knakgsnywr nrvmmvakkf ldaghklnfa
 35 301 vasrktfshe lsdfglestt gevppvairt akgekfvme efsrdgkale qflqeyfdgn
 361 lkrylksepi pesnegpvkv vvaenfddiv needkdvlie fyapwgcghck nlepkykelg
 421 eklskdpniv iakmdatand vpspyevkgf ptiyfspank kltpkkyegg relndfisyl
 481 qreatnppii qeekpkkkkk aqedl

SEQ ID NO: 11 Human CALR cDNA Sequence (NM_004343.3, CDS region from position 81-1334)

40 1 gggcggtccg tccgtactgc agagccgctg ccggagggtc gttttaaaagg gcccgcgctg
 61 tgccgcccc tggcccgcc atgctgctat ccgtgccgct gctgctcggc ctcctcggcc
 121 tggccgtcgc cgagcctgcc gtctacttca aggagcagtt tctggacgga gacgggtgga
 45 181 cttcccgtcg gatcgaatcc aaacacaagt cagatttttg caaatcgtt ctcagttccg
 241 gcaagtctta cggtagcag gagaaagata aaggtttgca gacaagccag gatgcacgct
 301 tttatgctct gtcggccagt ttcgagcctt tcagcaacaa aggccagacg ctggtggtgc
 361 agttcacggt gaaacatgag cagaacatcg actgtggggg cggctatgtg aagctgttct
 421 ctaatagttt ggaccagaca gacatgcacg gagactcaga atacaactac atgtttggtc
 50 481 ccgacatctg tggccctggc accaagaagg ttcattgtcat cttcaactac aagggaaga
 541 acgtgctgat caacaaggac atccgttgca aggatgatga gtttacacac ctgtacacac
 601 tgattgtgcg gccagacaac acctatgagg tgaagattga caacagccag gtggagtccg
 661 gctccttggg agacgattgg gacttctgac caccacaaga gataaaggat cctgatgctt
 721 caaaaccgga agactgggat gagcgggcca agatcgatga tcccacagac tccaagcctg
 55 781 aggactggga caagcccag catatccctg accctgatgc taagaagccc gaggactggg
 841 atgaagagat ggacggagag tgggaacccc cagtgattca gaaccctgag tacaagggtg
 901 agtggaaagg ccggcagatc gacaaccag attacaaggg cacttgatc caccagaaa
 961 ttgacaaccc cgagtattct cccgatccca gtatctatgc ctatgataac tttggcgtgc
 1021 tgggcctgga cctctggcag gtcaagtctg gcaccatctt tgacaacttc ctcacacca
 60 1081 acgatgaggc atacgctgag gagtttggca acgagacgtg gggcgtaaca aaggcagcag

1141 agaaacaaat gaaggacaaa caggacgagg agcagaggct taaggaggag gaagaagaca
 1201 agaaacgcaa agaggaggag gaggcagagg acaaggagga tgatgaggac aaagatgagg
 1261 atgaggagga tgaggaggac aaggagggaag atgaggagga agatgtcccc ggccaggcca
 1321 aggacgagct gtagagaggc ctgcctccag ggctggactg aggcctgagc gctcctgccg
 5 1381 cagagctggc cgcgccaaat aatgtctctg tgagactcga gaactttcat ttttttccag
 1441 gctggttcgg atttggggtg gattttgggt ttgttcccct cctccactct cccccacccc
 1501 ctccccgccc tttttttttt ttttttttaa actggtatatt tatctttgat tctccttcag
 1561 cctccacccc tggttctcat ctttcttgat caacatcttt tcttgccctct gtcccccttct
 1621 ctcatctctt agctcccctc caacctgggg ggagtggtg tggagaagcc acaggcctga
 10 1681 gatttcatct gctctccttc ctggagccca gaggaggga gcagaagggg gtggtgtctc
 1741 caacccccca gcactgagga agaacggggc tcttctcatt tcacccctcc ctttctcccc
 1801 tgccccaggg actgggcccac ttctgggtgg ggagtggtt cccagattgg ctcacactga
 1861 gaatgtaaga actacaaaca aaatttctat taaattaaat tttgtgtctc caaaaaaaaa
 1921 aaaaaaaaaa

SEQ ID NO: 12 Human CALR Amino Acid Sequence (NP_004334.1)

1 mllsvplllg llglavaepa vyfkeqfldg dgwtsrwies khksdfgkfv lssgkfygde
 61 ekdkglqtsq darfyalsas fepfsnkgqt lvvqftvkhe qnidcgggyv klfpnslldqt
 121 dmhgdseyini mfgpdicgpg tkkvhvifny kgknvlinkd irckddefth lytlivrpdn
 20 181 tyevkidnsq vesgsleddw dflppkkikd pdaskpedwd erakiddptd skpedwdkpe
 241 hipdpdakkp edwdeemdge weppviqnpe ykgewkprqi dnpdykgtwi hpeidnpeys
 301 pdpsiyaydn fgvlglldlwq vksgtifdnf litndeayae efgnetwgtv kaaekqmkdk
 361 qdeeqrlkee eedkkrkeee eaedkedded kdedeedeed keedeedevp gqakdel

SEQ ID NO: 13 Mouse CALR cDNA Sequence (NM_007591.3, CDS region from position 134-1384)

1 ggctgtgtca ggttcgggtg agaggtaggt gaatataaat tgaagcggcg gtggccgcgt
 61 ccgtcaatac cgcagagccg ctgcctgaag atcgtcttaa aaggcctgtg tgccgccgcc
 121 cctcgggccc gccatgctcc tttcggtgcc gctcctgctt ggctcctcgc gcttgccgcg
 30 181 ccgagaccct gccatctatt tcaaagagca gttcttggaac ggagatgcct ggaccaaccg
 241 ctgggtcgaa tccaaacata agtccgattt tggcaaat tgcctcagtt ctggcaaat
 301 ttacggggac ctggagaagg ataaagggtc gcagacaagc caagatgccc gattttacgc
 361 actgtccgcc aaattcgaac ccttcagcaa taagggccag aactgggtgg tacagttcac
 421 ggtgaagcat gagcagaata tcgactgtgg gggcggctac gtgaagctgt ttccgagtgg
 35 481 tttggaccag aaggacatgc atggagactc agaataaac atcatgtttg gtccggacat
 541 ctgcggtcct ggcaccaaga aggttcatgt catctttaac tacaaggga agaattgtgt
 601 gatcaacaag gatatccggt gtaaggatga tgaattcaca cacctataca cactgattgt
 661 gcgccagac aacacctatg aggtgaaaat tgacaacagc cagggtggagt caggctcctt
 721 ggaggatgat tgggactttc tgccacccaa gaagataaag gaccctgatg ctgccaagcc
 40 781 ggaagactgg gatgaacgag ccaagatcga tgacccca gattccaagc ctgaggactg
 841 ggacaagcca gagcacatcc ctgaccctga tgctaagaag cctgaggact gggatgaaga
 901 gatggatgga gagtgggaac caccagtgat tcaaaatcct gaatacaagg gcgagtggaa
 961 accacgtcaa attgacaacc cagattacaa ggttacctgg atacaccag aaattgacaa
 1021 ccctgaatac tccccgatg caaatatcta tgcctatgat agttttgctg tactgggect
 45 1081 agatctctgg caggtaagt ccgggacaat ctttgacaat ttcctcatca ccaatgatga
 1141 ggcctatgca gaggagtgtg gcaatgagac gtgggtgtt accaaggctg cagaagaagca
 1201 gatgaaggac aagcaggatg aggagcagag gcttaaggaa gaagaagagg acaagaagcg
 1261 taaagaggaa gaagaagctg aggataaaga ggatgatgat gacagagatg aagatgagga
 1321 cgaagaagat gagaaggagg aagatgagga agaattccct ggccaagcca aggatgagct
 50 1381 gtagaggcca caccacctgc cttcagggtc ggactgaggc ctgaacaccc tgccgcagag
 1441 ctggctgtct ccaataatgt ctctatgaga ctcaagaact tttcattttt tccaggcagg
 1501 ttcagatctg gggtagattc tgattttgtt cccctgcctc cccattacc cccccctt
 1561 ttttttttta ctggtgtttg tctttaattc tccttcagcc ctcatctggt ttctcatttt
 1621 tgaatcaaca tcttttccct ctgtccctcc ctttctccat cttttggtca ctaccctcca
 55 1681 actctaggaa caggggtgta gaggagaagc cctaggcttg agatttcatt tgctctcctt
 1741 cctgcatctc agaggagggc aggagaaggg ggtggtgtt tccctcccc cgcactgagg
 1801 aagaatgggg ctcttctcat cccctttctc ccttgcccc aggactgggc cacttggtgg
 1861 gcagccagtt ctagcacagc tcacactgag agtgtaagaa ctacaaacaa aatttctatt
 1921 aaattaagtt ttgtgtcttc cct

SEQ ID NO: 14 Mouse CALR Amino Acid Sequence (NP_031617.1)

1 mllsvpllllg llglaaadpa iyfkeqfldg dawtnrwves khksdfgkfv lssgkfygdl
 61 ekdkglqtsq darfyalsak fepfsnkgqt lvvqftvkhe qnidcgggyv klfpsgldqk
 121 dmhgdseydni mfgpdicgpg tkkvhvifny kgknvlinkd irckddefth lytlivrpdn
 5 181 tyevkidnsq vesgsleddw dflppkkikd pdaakpedwd erakiddptd skpedwdkpe
 241 hipdpdakp edwdeemdge weppviqnpe ykgewkprqi dnpdykgtwi hpeidnpeys
 301 pdaniyayds favlgldlwq vksgtifdnf litndeayae efgnetwgv tkaakqmkdk
 361 qdeeqrlkee eedkkrkeee eaedkedddd rdededeede keedeespq qakdel

10 SEQ ID NO: 15 Human Tap1 cDNA Variant 1 Sequence (NM_000593.5, CDS region from position 156-2582)

1 gtgtgctgta tggagaaaat tgggcaccag ggctgctccc gagattctca gatctgattt
 61 ccacgcttgc taccaaaata gtctgggcag gccacttttg gaagtaggcg tttatctagt
 121 agcaggcggc cgctttcgat ttctgctttcc cctaaatggc tgagcttctc gccagcgcag
 15 181 gatcagcctg ttcttgggac ttcccgagag ccccgccctc gttccctccc cagagcgcca
 241 gtaggggag actcggcggt acccgagact tcaggcccca ccggggcgcg gagactccca
 301 ggcccgggcg ggaccgggac ggcgctccgag tgccaatggc tagctctagg tgtcccgcctc
 361 ccgcggggtg ccgctgcctc cccggagact ctctcgcatg gctggggaca gtactgctac
 421 ttctcgccga ctgggtgctg ctccggaccg cgctgccccg catattctcc ctgctgggtg
 20 481 ccaccgctg gccactgctc cggtgtctgg cggtgggcct gagccgctgg gccgtgctct
 541 ggctgggggc ctgcggggtc ctgagggcaa cggttggtc caagagcgaa aacgcagggtg
 601 ccaggggctg gctggctgct ttgaagccat tagctgcggc actgggcttg gccctgccgg
 661 gacttgccct gttccgagag ctgatctcat ggggagcccc cgggtccgcg gatagacca
 721 ggctactgca ctggggaagt caccctaccg ccttcgttgt cagttatgca gcggcactgc
 25 781 ccgcagcagc cctgtggcac aaactcggga gcctctgggt gcccgcggt caggcggtc
 841 ctggaaaccc tgtgctgctg ctcttaggct gcctgggctc ggagacgcgc cgcctctcgc
 901 tgttcctggg cctggtgggc ctctcctctc ttggggagat ggccattcca ttctttacgg
 961 gccgcctcac tgactggatt ctacaagatg gctcagccga taccttcaact cgaaacttaa
 1021 ctctcatgtc cattctcacc atagccagtg cagtgtgga gttcgtgggt gacgggatct
 30 1081 ataacaacac catgggccac gtgcacagcc acttgaggag agaggtgttt ggggctgtcc
 1141 tgcgccagga gacggagttt ttccaacaga accagacagg taacatcatg tctcgggtaa
 1201 cacaggacac gtccaccctg agtgattctc tgagtgaaga tctgagctta tttctgtggt
 1261 acctggtgcg aggcctatgt ccttggggga tcatgctctg gggatcagtg ttcctcacca
 1321 tggtcaccct gatcaccctg cctctgcttt tccttctgcc caagaagtg ggaaaatggt
 35 1381 accagttgct ggaagtgcag gtgcgggaat ctctggcaaa gtccagccag gtggccattg
 1441 aggtctgtgc ggccatgcct acagttcgaa gctttgccaa cgaggagggc gaagcccaga
 1501 agtttaggga aaagctgcaa gaaataaaga cactcaacca gaaggaggct gtggcctatg
 1561 cagtcaactc ctggaccact agtatctcag gtatgctgct gaaagtggga atcctctaca
 1621 ttggtgggca gctggtgacc agtggggctg taagcagtg gaacctgtc acatttggtc
 40 1681 tctaccagat gcagttcacc caggctgtgg aggtactgct ctccatctac ccagagtac
 1741 agaaggctgt gggctcctca gagaaaatat ttgagtacct ggaccgcacc cctcgctgcc
 1801 caccagtggt tctgttgact cccttacact tggagggcct tgtccagttc caagatgtct
 1861 cctttgccta cccaaaccgc ccagatgtct tagtgctaca ggggctgaca ttcaccctac
 1921 gccctggcga ggtgacggcg ctggtgggac ccaatgggtc tgggaagagc acagtggctg
 45 1981 ccctgctgca gaatctgtac cagcccaccg ggggacagct gctgttggtt ggggaagccc
 2041 ttccccaata tgagcaccgc tacctgcaca ggcaggtggc tgagtgagg caagagccac
 2101 aggtatttgg aagaagtctt caagaaaata ttgcctatgg cctgaccagc aagccaacta
 2161 tggaggaaat cacagctgct gcagtaaagt ctggggccca tagtttcac tctggactcc
 2221 ccaggggcta tgacacagag gtagacgag ctgggagcca gctgtcagg ggtcagcgac
 50 2281 agcagtggtc gttggcccg gcaatgattc ggaaaccgtg tgtacttatc ctgatgatg
 2341 ccaccagtg cctggatgca aacagccagt tacagtgga gcagctcctg tacgaaagcc
 2401 ctgagcggta ctcccgtcga gtgcttctca tcacccagca cctcagcctg gtggagcagg
 2461 ctgaccacat cctctttctg gaaggaggcg ctatccggga ggggggaacc caccagcagc
 2521 tcatggagaa aaagggtgct tactgggcca tgggtgcaggc tctgcagat gctccagaat
 55 2581 gaaagccttc tcagacctgc gcaactccatc tccctccctt ttcttctctc tgtggtggag
 2641 aaccacagct gcagagtagg cagctgcctc caggatgagt tacttgaaat ttgcttgag
 2701 tgtgttacct cctttccaag ctctcgtga taatgcagac ttcctggagt acaaacacag
 2761 gatttgtaat tccttactgt aacggagttt agagccagg ctgatgcttt ggtgtggcca
 2821 gcactctgaa actgagaaat gttcagaatg tacggaaaga tgatcagcta ttttcaacat
 60 2881 aactgaaggc atatgctggc ccataaacac cctgtaggtt cttgatattt ataataaaat

2941 tgggtgttttg taaaaaaaaa aaaaaaaaaa aaaa

SEQ ID NO: 16 Human TAP1 Isoform 1 Amino Acid Sequence (NP_000584.2)

	1	maellasags	acswwdfprap	psfpppaasr	gglggtrsfr	phrgaesprp	grdrdgvrp
5	61	massrcpap	gcrclpgasl	awlgvtvlll	adwvllrtal	prifsllypt	alpllrwvav
	121	glswavlw	gacgvratv	gsksenagag	gwlaalkpla	aalglalpgl	alfreliswg
	181	apgsadstrl	lhwgshptaf	vvysaaalpa	aalwhklgsl	wvpggggsg	npvrrllgcl
	241	gsetrrlsf	lvlvvlssl	emaipfftgr	ltdwilqdgs	adtfrnltl	msiltiasav
	301	lefvgdgiyn	ntmghvhshl	qgevfgavlr	qeteffqqnq	tgnimsrvte	dtstlsdsls
10	361	enlsflwyl	vrglcllgim	lwgsvsltmv	tlitlplll	lpkkvgkwyq	llevqvresl
	421	akssqvaiea	lsamptvrsf	aneegaeqkf	reklqeiktl	nqkeavayav	nswttsisgm
	481	llkvgilyl	gqlvtsgavs	sgnlvtfvly	qmqtqavev	llsiyprvqk	avgssekife
	541	yldrtprc	sglltphle	glvqfvdvsf	aypnrdvlv	lqglftlrlp	gevtalvgpn
	601	gsgkstvaal	lqnlyqptgg	qllldgkplp	qyehrylhrq	vaavgqepqv	fgrslqenia
15	661	yglrtqkptme	eitaaavks	ahsfisglp	gydtevd	sqlsggqrq	valaralirk
	721	pcvlilddat	saldansql	veqllyespe	rysrsvllit	qhslsveqad	hilfleggai
	781	reggthqqlm	ekkgcywamv	qapadape			

SEQ ID NO: 17 Human Tap1 cDNA Variant 2 Sequence (NM_001292022.1, CDS region from position 199-1842)

20	1	gctctggggc	cgaagggcgg	ggcatcctca	tctctaacag	gaggcttttc	tacttcatga
	61	tctccagcct	tcctaataaa	atcctgaaag	ttctggtaga	gcaaccacag	ggtagttagt
	121	tccagggcag	cctatttagg	ttcgggattg	agacgtcagt	gtttcctttc	tgctgatgcc
	181	ctccaggata	atggggagat	ggccattcca	ttctttacgg	gccgcctcac	tgactggatt
25	241	ctacaagatg	gctcagccga	taccttcaact	cgaaacttaa	ctctcatgtc	cattctcacc
	301	atagccagtg	cagtgtctgga	gttcgtgggt	gacgggatct	ataacaacac	catgggccac
	361	gtgcacagcc	acttgacagg	agaggtgttt	ggggctgtcc	tgcgccagga	gacggagttt
	421	ttccaacaga	accagacagg	taacatcatg	tctcgggtaa	cagaggacac	gtccaccctg
	481	agtgtattctc	tgagttagaa	tctgagctta	tttctgtggt	acctggtgcg	aggcctatgt
30	541	ctcttgggga	tcatgtctctg	gggatcagtg	tccttcacca	tggtcaccct	gatcaccctg
	601	cctctgcttt	tccttctgcc	caagaagggtg	ggaaaatggt	accagttgct	ggaagtgcag
	661	gtgcgggaat	ctctggcaaa	gtccagccag	gtggccattg	aggctctgtc	ggcatgcct
	721	acagttcgaa	gctttgcca	cgaggagggtc	gaagccaga	agtttaggga	aaagctgcaa
	781	gaaataaaga	cactcaacca	gaagagggtc	gtggcctatg	cagtcaactc	ctggaccact
35	841	agtattttcag	gtatgtctgct	gaagagtggga	atcctctaca	ttggtgggca	gctggtgacc
	901	agtggggctg	taagcagtg	gaacctgtgc	acatttgttc	tctaccagat	gcagttcacc
	961	caggctgtgg	aggtactgct	ctccatctac	cccagagtac	agaaggctgt	gggctcctca
	1021	gagaaaatat	ttgagtacct	ggaccgcacc	cctcgctgcc	caccagtggt	tctgttgact
	1081	cccttacact	tgaggggcct	tgtccagttc	caagatgtct	cctttgccta	cccaaaccgc
40	1141	ccagatgtct	tagtgctaca	ggggctgaca	ttcaccctac	gccctggcga	ggtgacggcg
	1201	ctggtgggac	ccaatgggtc	tgggaagagc	acagtggctg	ccctgctgca	gaatctgtac
	1261	cagcccaccg	ggggacagct	gctgttggtg	gggaagcccc	ttcccccaata	tgagcaccgc
	1321	tacctgcaca	ggcaggtggc	tgcagtggga	caagagccac	aggtatttgg	aagaagtctt
	1381	caagaaaata	ttgcctatgg	cctgacccag	aagccaacta	tggaggaaat	cacagctgct
45	1441	gcagtaaagt	ctggggccca	tagtttcatc	tctggactcc	ctcagggcta	tgacacagag
	1501	gtagacgagg	ctgggagcca	gctgtcaggg	ggtcagcgac	aggcagtggt	gttgggccga
	1561	gcattgatcc	ggaaaccgtg	tgtacttatc	ctggatgatg	ccaccagtgc	cctggatgca
	1621	aacagccagt	tacaggtgga	gcagctcctg	tacgaaagcc	ctgagcggtg	ctcccgtctc
	1681	gtgcttctct	tcacccagga	cctcagcctg	gtggagcagg	ctgaccacat	ctcttttctg
50	1741	gaaggaggcg	ctatccggga	ggggggaaac	caccagcagc	tcattggaga	aaaggggtgc
	1801	tactggggca	tggtgcaggc	tccctgcagat	gctccagaat	gaaagccttc	tcagacctgc
	1861	gcactccatc	tccttccctt	ttcttctctc	tgtggtggag	aaccacagct	gcagagtagg
	1921	cagctgcctc	caggatgagt	tacttgaaat	ttgccttgag	tgtgttacct	cctttccaag
	1981	ctcctcgtga	taatgcagac	ttcctggagt	acaaacacag	gatttgaat	tccttactgt
55	2041	aacggagttt	agagccagg	ctgatgcttt	ggtgtggcca	gcactctgaa	actgagaaat
	2101	gttcagaatg	tacggaaaga	tgatcagcta	ttttcaacat	aactgaaggc	atatgctggc
	2161	ccataaacac	cctgtaggtt	cttgataatt	ataataaaat	tggtgttttg	taaaaaaaaa
	2221	aaaaaaaaaa	aaaa				

SEQ ID NO: 18 Human TAP1 Isoform 2 Amino Acid Sequence (NP_001278951)

	1	maipffftgrl	tdwilqdgsga	dtftrnltlm	siltiasavl	efvvgdgiynn	tmghvshshlq
	61	gevfagavlrlq	eteffqqnqt	gnimsrvted	tstlsdslse	nlsflflwylv	rglcllgiml
5	121	wgsvsltmvt	litlplllfll	pkkvqkwyql	levqvresla	kssqvaieal	samptvrsfa
	181	neegeaqkfr	eklqeiktl	qkeavayavn	swttsisgml	lkvgilyigg	qlvtsgavss
	241	gnlvtfvlyq	mqftqavevl	lsiyprvqka	vgssekifey	ldrtprcpps	glltphleg
	301	lvqfqdvsvfa	ypnrpdvlvl	qglftflrpg	evtalvgpng	sgkstvaall	qnlyqptggq
	361	llldgkplpq	yehrylhrqv	aavgqepqv	grslqeniay	gltkqptmee	itaaavksa
10	421	hsfisglpqq	ydtevddeags	qlsggqrqav	alaralirkp	cvlilddats	aldansqlqv
	481	eqllyesper	ysrsvllitq	hlslveqadh	ilfleggair	eggthqqlme	kkgcywamvq
	541	apadape					

SEQ ID NO: 19 Mouse Tap1 cDNA Variant 1 Sequence (NM_013683.2, CDS region from position 325-2499)

15	1	gaggagaatg	agattcatgg	agaagaacac	gacaggccag	ggctgctagg	cagaactcca
	61	actacagctt	cagcggcagc	ttccagaaca	gcctgagcaa	gccagtctca	gaaggaggcg
	121	tgtctagtca	ttcgaggtcg	gctttcgggt	tcttcttctc	ctaaacgcca	gcacttctag
	181	tcagctccac	cagctcgagc	gggttcccg	gactttacgc	gcacgcctc	ggacccgccc
20	241	ttcttccttc	cccacggaga	ctcctgtgca	gcgcggacgt	cgagagtccc	aggctaggac
	301	cagactctgg	acagctcacg	ctcgatggct	gcgcacgtct	ggctggcgcc	cgccctgctc
	361	cttctgggtg	actggctgct	gctgcggccc	atgctcccgg	gaatcttctc	cctgttggtt
	421	cccagaggtg	cgtgctccg	ggtctgggtg	gtgggcctga	gtcgtggggc	catectagga
	481	ctaggggtcc	gcggggtcct	cgggggtcacc	gcaggagccc	atggctggct	ggctgctttg
25	541	cagccgctgg	tggccgcact	gagtttggtc	ctgcctggac	ttgccttggt	ccgagagctg
	601	gccgcctggg	gaacactccg	ggagggtgac	agcgtggat	tactgtactg	gaacagtcgt
	661	ccagatgcct	tcgctatcag	ttatgtggca	gcattgccc	cagccgcctc	gtggcacaag
	721	ttggggagcc	tctgggcgcc	cagcggcaac	agggacgctg	gagacatgct	gtgtcggatg
	781	ctgggcttcc	tgggcccata	gaagagacgt	ctctacctgg	ttctggttct	cttgattctc
30	841	tcttgccctt	gggaaatggc	cattcccttc	ttcacggggc	gcatcactga	ctggattctt
	901	caggataaga	cagttccctag	cttcacccgc	aacatatggc	tcatgtccat	tctcaccata
	961	gccagcacag	cgttgaggtt	tgcaagtgat	ggaatctaca	acatcaccat	gggacacatg
	1021	cacggccgtg	tgcacagaga	ggtgtttcgg	gccgtccttc	gccaggagac	agggtttttc
	1081	ctgaagaacc	cagcaggttc	catcacatct	cgggtgactg	aggacacagc	caacgtgtgc
35	1141	gagtcacatta	gtggcacgct	gagcctgctg	ctgtggtacc	tggggcgagc	cctgtgtctc
	1201	ttggtgttca	tgttttgggg	gtcaccgtac	ctcactctgg	tcaccctgat	caacctgccc
	1261	ctgctttttc	ttttgcctaa	gaagctggga	aaagtgcac	agtcactggc	agtgaaggtg
	1321	caggagtctc	tagcaaagtc	cagcgagggtg	gcccttgagg	ccttatcggc	gtgacctact
	1381	gtgctggagct	ttgccaaacga	ggagggtgag	gcccaagaat	tcaggcagaa	gttggaagaa
40	1441	atgaagactc	taaacaagaa	ggaggccttg	gcttacgtgg	ctgaagtctg	gaccacgagt
	1501	gtctcgggaa	tgtctgtgaa	ggtgggaatt	ctgtacctgg	gcgggagcgt	ggtgatcaga
	1561	ggggtgttca	gcagcggcaa	ccttgtctca	ttcgttctct	accagcttca	gttcaccag
	1621	gctgttcagg	tctgtctctc	cctctacccc	tccatgcaga	aggctgtggg	ctcctcagag
	1681	aaaatatctg	aatacttgga	cgggactcct	tgctctccac	tcagtggctc	gttggcaccc
45	1741	tcaaacaatga	aaggccttgt	ggagttccaa	gatgtctctt	ttgcctaccc	aaaccagccc
	1801	aaagtccagg	tgcttcaggg	gctgacgttc	accctgcatc	ctggaacggt	gacagcgttg
	1861	gtgggaccca	atggatcagg	gaagagcacc	gtggctgccc	tgctgcagaa	cctgtaccag
	1921	cccaccgggg	gccagctgct	gctggatggc	cagtgcctgg	tccagtatga	tcaccattac
	1981	ctgcacactc	aggtggccgc	agtgggacaa	gagccgctgc	tatttggaa	aagctttcga
50	2041	gaaaatattg	cgatatggcct	gaaccggact	ccaacatgg	aggaaatcac	agctgtggcc
	2101	gtggagtctg	gagcccacga	tttcatctct	gggttccctc	agggtatga	cacagaggtg
	2161	ggtgagactg	ggaaccagct	gtcaggaggt	cagcgacagg	cagtggcctt	ggcccagacc
	2221	ttgatccgga	agccactcct	gcttatcttg	gatgatgcca	ccagtgcctc	ggatgctggc
	2281	aaccagctac	gggtccagcg	gctctgtgat	gagagcccca	agcgggcttc	tcggacgggt
55	2341	cttcttatca	cccagcagct	cagcctggca	gagcaggccc	accacatcct	ctttctcaga
	2401	gaaggctctg	tcggcgagca	gggcacccac	ctgcagctca	tgaagagagg	agggtgctac
	2461	cgggccatgg	tagaggctct	tgccgctcct	gcagactgac	aaggcctctg	gactgcacac
	2521	tgcgtgctgt	ccccctcctg	tcctctactc	tgtgtgacag	agaacctggg	agcaaagata
	2581	ttaccacacc	cacggagaca	gttgaggagc	ggagggtgctt	gttacatgag	aaaatgtaaa
60	2641	ctctaggaga	taccgcgaat	ttaccacgaa	tgtgtttccc	cgaccgcctc	cctattagac

2701 ggggttttcag gtacctcaca ccaacactga gctgctaaat ctccggtctc caccctcccc
 2761 gtgctgagct tgtatcacag catgcacaca caaacctggt ttatgtggcg ttgtgataga
 2821 atgagaagaa acgccccaaag tgtacagaga gaaaggacaa gcagcttgca attaaccaaa
 2881 ggcataggct ggcatTTTggg tgttctatgg gtgtttgata tttgtaataa aactggtggt
 5 2941 ttgtactgtg

SEQ ID NO: 20 Mouse TAP1 Isoform 1 Amino Acid Sequence (NP_038711.2)

1 maahvwlaaa llllvdwlll rpmlpgifsl lvpevpllrv wvvglsrwai lglgvrvgvl
 61 vtagahgwla alqplvaals lalpglalfr elaawgtlre gdsagllwyn srpdafaisy
 10 121 vaalpaaalw hklgslwaps gnrdagdmle rmlgflgpkk rrlylvlvll ilscldgemai
 181 pfftgritdw ilqdktvpsf trniwlmsil tiastalefa sdgiynitmg hmhgrvhrev
 241 fravlrqetg fflknpagsi tsrvtedtan vcesisgtls lllwylgral clldvfmfwgs
 301 pyltlvtlin lpillfllpkk lgkvhqslav kvqeslakst qvalealsam ptvrsfanee
 361 geaqkfrqkl eemktnkake alayvaevwt tsvsgmllkv gilylggqlv irgavssgnl
 15 421 vsfvlyqlqf tqavqvllsl ypsmqkavgs sekifeylr tpcsplsgsl apsnmkglve
 481 fqdvsfaypn qpkvqvlqgl tftlhpgtvt alvgpngsgk stvaallqnl yqptggqlll
 541 dgqclvqydh hylhtqvaav gqepllfgrs freniaygln rtptmeeita vavesgahdf
 601 isgfpggydt evgetgnqls ggqrqavala ralirkpill ilddatsald agnqlrvqrl
 661 lyespkrasr tvllitqqls laeqahhilf lregsvgegg thlqlmkrng cyramveala
 20 721 apad

SEQ ID NO: 21 Mouse Tap1 cDNA Variant 2 Sequence (NM_001161730.1, CDS region from position 325-2415)

1 gaggagaatg agattcatgg agaagaacac gacaggccag ggctgctagg cagaactcca
 25 61 actacagctt cagcggcagc ttccagaaca gcctgagcaa gccagtctca gaaggaggcg
 121 tgtctagtga ttcgaggtcg gctttcgggt tcttcttcct cttaaagcca gcacttctag
 181 tcagctccac cagctcgagc ggggttcccg gactttacgc gcacgcctc ggaccgccc
 241 ttcttccttc ccacgagaga ctctgtgca gcgcgagct cgagagctcc aggtaggac
 30 301 cagactctgg acagctcacg ctcgatggct gcgcacgtc ggctggcggc cgccctgctc
 361 cttctgggtg actggctgct gctgcggccc atgctcccg gaatcttct cctgttggtt
 421 cccgaggtgc cgctgctccg ggtctgggtg gtgggcctga gtcgctgggc catcctagga
 481 ctaggggtcc gcggggtcct cggggtcacc gcaggagccc atggctggct ggctgctttg
 541 cagccgctgg tggccgcaat gagtttgccc ctgcctggac ttgcctgtt ccgagagctg
 60 601 gccgctggg gaacactccg ggagggtgac agcgctggat tactgtactg gaacagctcg
 661 ccagatgcct tcgctatcag ttatgtggca gcattgccc cagccgcct gtggcacaag
 35 721 ttggggagcc tctgggcgcc cagcggcaac agggagcgtg gagacatgt gtgtcggtg
 781 ctgggcttcc tgggcccata gaagagacgt ctctacctgg ttctggttct cttgattctc
 841 tcttgccctt gggaaatggc cattcccttc ttcacgggcc gcatcactga ctggattctt
 901 caggataaga cagttcctag cttcacccgc aacatatggc tcatgtccat tctcaccata
 40 961 gccagcacag cgctggagtt tgcaagtgat ggaatctaca acatcaccat gggacacatg
 1021 cagggccgtg tgcacagaga ggtgtttcgg gccgtccttc gccaggagac aggggtttttc
 1081 ctgaagaacc cagcaggttc catcacatct cgggtgactg aggacacagc caacgtgtgc
 1141 gagtccatta gtggcacgct gagcctgctg ctgtggtacc tggggcgagc cctgtgtctc
 1201 ttggtgttca tgttttgggg gtcaccgtac ctactctgg tcaccctgat caacctgccc
 45 1261 ctgctttttc ttttgccata gaagctggga aaagtgcac agtactggc agtgaagggtg
 1321 caggagtctc tagcaaagtc cacgcaggtg gcccttgagg ccttatcggc gatgcctact
 1381 gtgctggagct ttgccaaaga ggagggtgag gccagaaagt tcaggcagaa gttggaagaa
 1441 atgaagactc taacaagaa ggaggccttg gcttacgtgg ctgaagtctg gaccacagat
 1501 gtctcgggaa tgetgctgaa gcttcagttc acccaggctg ttcaggctct gctctccctc
 50 1561 taccctcca tgcagaaggc tgtgggctcc tcagagaaaa tattcgaata cttggaccgg
 1621 actccttgct ctccactcag tggctcggtg gcaccctcaa acatgaaagg ccttggtggag
 1681 ttccaagatg tctcttttgc ctacccaaac cagcccaaag tccaggtgct tcaggggctg
 1741 acgttcaccc tgcattcctg aacgggtgac gcgttggtgg gacccaatgg atcaggaag
 1801 agcaccgtgg ctgcccctg gcagaacctg taccagccca ccgggggcca gctgctgctg
 55 1861 gatggccagt gcctgggtcca tggatgacac cattacctgc acactcaggt ggccgctg
 1921 ggacaagagc cgctgctatt ttgaagaagc tttcgagaaa atattgctga tggcctgaac
 1981 cggactccaa ccatggagga aatcacagct gtggccgtgg agtctggagc ccacgatttc
 2041 atctctgggt tccctcaggg ctatgacaca gaggtagggt agactgggaa ccagctgtca
 2101 ggaggctcagc gacaggcagt ggccctggcc cgagccttga tccggaagcc actcctgctt
 60 2161 atcttggtatg atgccaccag tgccctggat gctggcaacc agctacgggt ccagcggctc

2221 ctgtatgaga gccccaaagcg ggcttctcgg acggttcttc ttatcaccca gcagctcagc
 2281 ctggcagagc agggccacca catcctcttt ctacagagaag gctctgtcgg cgagcagggc
 2341 acccacctgc agctcatgaa gagaggaggg tgctaccggg ccatggtaga ggctcttgcg
 2401 gctcctgcag actgacaagg cctctggact gcacactgcg tgctgtcccc ctctgtcct
 5 2461 ctactctgtg tgacagagaa cctgggagca aagatattac cacaccacg gagacagttg
 2521 aggagcggag gtgcttgta catgagaaaa tgtaaactct aggagatacc cggaatttac
 2581 cacgaatgtg tttccccgac ccgctcccta ttagacgggg tttcaggtac ctcacaccaa
 2641 cactgagctg ctaaactctcc ggtctccacc ctcccgtgc tgagcttgta tcacagcatg
 2701 cacacacaaa cctggtttat gtggcgttgt gatagaatga gaagaaacgc ccaaagtgtg
 10 2761 cagagagaaa ggacaagcag cttgcaatta accaaaggca taggctggca tttgggtgtt
 2821 ctatgggtgt ttgatatttg taataaaact ggtgttttgt actgtg

SEQ ID NO: 22 Mouse TAP1 Isoform 2 Amino Acid Sequence (NP_001155202.1)

1 maahvwlaaa llllvdwlll rpmlpgifsl lvpevpllrv wvvglsrwai lglgvrvgvl
 15 61 vtagahgwla alqplvaals lalpglalfr elaawgtlre gdsagllwyn srpdaaisy
 121 vaalpaaalw hklglwaps gnrdagdmlc rmlgflgpkk rrlylvllvll ilscldgemai
 181 pfftgritdw ilqdktvpsf trniwlmsil tiastalefa sdgiynitmg hmhgrvhrev
 241 fravlrqetg fflknpagsi tsrvtedtan vcesisgtls lllwylgral cllvmfwgfs
 301 pyltlvtlin lp11fl1pkk lgkvhqslav kvqeslakst qvalealsam ptvrsfanee
 20 361 geaqkfrqkl eemktlnkke alayvaevwt tsvsgmllkl qftqavqvl slypsmqkav
 421 gssekifeyl drtpcsplsg slapsnmkg1 vefqdvsfay pnqpkvqvlq gltftlhpqt
 481 vtalvgpngs gkstvaallq nlyqptggql lldgqclvqy dhhyhltqva avggepllfq
 541 rsfreniayg lnrtptmeei tavavesgah dfisgfpqgy dtevgetgnq lsggqrqava
 601 lalarlirkpl llilddatsa ldagnqlrvq rllyespkra srtvllitqq ls1aeqahhi
 25 661 lflregsvge qgthlqlmkr ggcyravvea laapad

SEQ ID NO: 23 Human TAP2 cDNA Variant 1, B allele Sequence (NM_000544.3, CDS region from position 123-2234)

1 gcccgcctcg gccgagcgta gctggcggac cagagccggt agcgaggttg ggagagacgg
 30 61 agcggacctc agcgtgaag cagaagtccc cggagctgcg gtctccccgc cgcggttag
 121 ccatgaggct cctgacctg agaccctgga cctccctgct gctggtagac gcggctttac
 181 tgtggctgct tcagggccct ctggggactt tgcttctca agggctgcca ggactatggc
 241 tggaggggac cctgcggctg ggaggctgt gggggctgct aaagctaaga gggctcgtg
 301 gatttgtggg gacactgctg ctcccgtct gtctggccac cccctgact gtctccctga
 35 361 gagccctggg cgcgggggcc tcacgtgctc cccagccag agtcgcttca gccccttga
 421 gctggctgct ggtgggttac ggggctgcgg ggctcagctg gtactgtgg gctgttctga
 481 gccctcctgg agcccaggag aaggagcagg accaggtgaa caacaaagtc ttgatgtgga
 541 ggctgctgaa gctctccagg ccggacctgc ctctcctcgt tgccgccttc ttcttcttg
 601 tccttctgtg tttgggtgag acattaatcc ctactattc tggctgctg attgacatcc
 40 661 tgggaggtga ttttgacccc catgcctttg ccagtgccat cttcttcatg tgcctcttct
 721 cctttggcag ctactgtct gcaggctgcc gaggaggctg cttcacctac accatgtctc
 781 gaatcaactt gcggatccgg gagcagcttt tctcctccct gctgcgccag gacctcgggt
 841 tcttccagga gactaagaca ggggagctga actcacggct gagctcggat accaccctga
 901 tgagtaactg gcttccttta aatgccaatg tgctcttgcg aagcctgggt aaagtgggtg
 45 961 ggctgtatgg cttcatgctc agcatatcgc ctgcactcac cctcctttct ctgctgcaca
 1021 tgcccttcac aatagcagcg gagaagggtg acaacacccg ccatcaggaa gtgcttcggg
 1081 agatccagga tgacgtggcc agggcggggc aggtgggtgc ggaagccgtt ggagggtgc
 1141 agaccgttcg cagttttggg gccgaggagc atgaagtctg tcgctataaa gaggcccttg
 1201 aacaatgtcg gcagctgtat tggcggagag acctggaacg cgcttgtag ctgctcgtaa
 50 1261 ggaggggtgct gcacttgggg gtgcagatgc tgatgctgag ctgtgggctg cagcagatgc
 1321 aggatgggga gctcaccag ggacgctgc tttcctttat gatctaccag gagagcgtgg
 1381 ggagctatgt gcagaccctg gtatacatat atggggatat gctcagcaac gtgggagctg
 1441 cagagaagggt tttctcctac atggaccgac agccaaatct gccttcacct ggcacgcttg
 1501 ccccaaccac tctgcagggg gttgtgaaat tccaagacgt ctcttttgca tatcccaatc
 55 1561 gccctgacag gcctgtgctc aaggggctga cgtttaccct acgtcctggt gaggtgacgg
 1621 cgctgggtgg acccaatggg tctgggaaga gcacagtggc tgccctgctg cagaatctgt
 1681 accagcccac agggggacag gtgctgctgg atgaaaagcc catctcacag tatgaacact
 1741 gctacctgca cagccagggt gtttcagttg ggcaggagcc tgtgctgttc tccggttctg
 1801 tgaggaacaa cattgcttat gggctgcaga gctgcgaaga tgataagggt atggcggtg
 60 1861 ccagggctgc ccacgcagat gacttcatcc aggaaatgga gcatggaata tacacagatg

1921 tagggggagaa ggggaagccag ctggctgcgg gacagaaaca acgtctggcc attgcccggg
 1981 cccttgtagc agaccgcgg gtcctcatcc tggatgagcc tactagtgcc ctgatgtgc
 2041 agtgcgagca ggccctgcag gactggaatt cccgtgggga tcgcacagtg ctggtgattg
 2101 ctcacaggct gcaggcagtt cagcgcgccc accagatcct ggtgctccag gagggaagc
 5 2161 tgcagaagct tgcccagctc caggagggac aggacctcta ttcccgccgt gttcagcagc
 2221 ggctgatgga ctgaggcccc agggatactg ggccctcttc tcaggggctg ctccaggacc
 2281 cagagctgtt cctgctttga gtttccctag agctgtgcgg ccagatagct gttcctgagt
 2341 tgcaggcacg atggagattt ggacactgtg tgcttttggg ggggtagaga ggtggggtgg
 2401 ggtggggtgg gggctgtctg tgtccaggaa acttaattcc ctggtgacta gagctttgcc
 10 2461 tggatgagag gagtattttg tggcataata catatatttt aaaatatttt ctttcttaca
 2521 tgaactgtat acattcatat agaaaattta gacaatataa aaaagtacaa agaagaaaag
 2581 taaaagtacc cattgtttca cttcctggag ataaccatag ttgctatttt gctgcctgtc
 2641 ccatcagtcg tttatctgtt gtttgagata gaaattaacc aaaaatgaca taaatattca
 2701 tgagattgcc ttctatatc cttccttggt cctaccagtg tctgctattt tgaagaagct
 15 2761 agggctcggg gggacagaga acagttccct gattaacagt attaatagcg acattggtaa
 2821 cagctaccat ttatagagtt ttaatgggag taggagctat gctaagtgtt tttcatgtat
 2881 tatcggtttt aatcattatc cccaacccta tgagggttgg tattatcccc attttacaga
 2941 tagggaaact gaagctcaaa gaggtcacta gactttccca aggtggtcgt agtgggtggg
 3001 ttggagtttg aacacaggcc tgaccctaga gtccacacc tgaccacac aattatattg
 20 3061 catcttggtt ccataaacc taatccataa tcccatcaag aaaagctctg ctgctcttag
 3121 ctctaaataa ttcagaatct attctcttct ctccagtccc gttgttatag tcttactca
 3181 tagacttaag atgatcccat caccagagag gtttctctac cattagcttc cctcttccgg
 3241 ccattcttca caaagtcatt tttctaaatt ctgtgtcaca tacgatgatg gcatttctgg
 3301 aaattccttc aggtgctctc aagccctgct gcagagatcc ttttcagagc acacactgtt
 25 3361 ccagcccatc tgtctcacc tctcctgttg tatccagctc cacgacaaac ttctgccttc
 3421 cccaacacct ttgtgccttt gcataagggtg ttttcttgcc cattttctgc tcgactcgcc
 3481 cctgattttc aagttcaaga cttactcag ggttcaggctc ttccaggagg ccttacttat
 3541 gtcgtcagtc tggggaactc tccatgtgct tctatcactg tgcggttacc tctttcacag
 3601 ccctttttaa gttctatctt ccctttccca ctttttttga ccttccacta gaccatgagc
 30 3661 acctggggcg aaagccatat atcttattaa gctttatatc tgctacctgg ccgagggcct
 3721 aattcatagt ggagaataaa tagtcaattg aataaatgaa taaatatctc caccatcgta
 3781 ctaatcttaa tcctccctgc ccactcccac cactgaaaat gcaacattgt acacatcact
 3841 ggttggttggg agggacttac cttggaaagt tgctattcta ggaaagagaa accttcatat
 3901 tcctggaaac agcaggtagt ttccagtgtc ggcaatgaat tccccagaac tgctgttttg
 35 3961 gattttttct tgccctggcag ctggtgggag cagggtgcag tgaggatggg gtgagagtgg
 4021 gcagtttctt gtgcagattt gcctttcttt catcctgggg ctgacttgca gtccacacc
 4081 catccatctc tcaaatttca cagagggtaa aataggcatt tggagagaaa gaactctggc
 4141 ctgattcctt tctctccac aaatgtcctt tattcataaa acaggaataa taattcctgt
 4201 atctcccaac tacatggaag ctgcagccct cacagaagaa gatgatctga gaaattcttt
 40 4261 gatttcttca gtacagttat acccatgcat cataatactt taagcctgga aggcattcta
 4321 aaaataatgc aacagtcaaa cctaatttta cagagaaact gacatgaaat cacgcagcta
 4381 atcatgataa agctgggtgg aaaacttatc ttgatgggca gtacaggaag atgcagtaga
 4441 ccttaagatg tcctgaaagt ttcttatctc aggggaaact cccaggtagg ctttatgtca
 4501 gggacacaga aaaatgctcc ctgaaagtca aaatattcgg gctagacaga caaatcctg
 45 4561 taagtgtggt ttgtctggga accacagatg tcaactaatcc tggtttgctc cagagttctt
 4621 tttgttcaact cctaccccc atcaccattt gattgatctc cttaccctgt aatttcccct
 4681 tcttgctgct tacctgcagt atctttccca cccaggcatg cttattctt tctaaaggaa
 4741 agtatgaatg gagaggggaa agcttgggaa actgatagat ttcttggtg gccaaaacac
 4801 ctccatagcc tgtctgccc gcccattgtg gaaacagcat tgagtttcaa gtcccttatg
 50 4861 cctccacca ggatagcca cttgtaatcc acatggcaat tgtgaaacaa gcaggaaatg
 4921 cgtaattgtc agaattttgt ggggaaagga ctagggaata aggaaaacaa agatcttctc
 4981 tgtgttttag agctgtcagc tagaggagca cctgcttgag tctgatgcca tctaattgct
 5041 ccagaagaaa ctgggttttg aacctagagt tccatggact cttaggaatt agactactac
 5101 tactactaag cttactactg tgcctactat gtgctattgc tgtgccaagt atctgaaacc
 55 5161 tgtcttctta ccttattttt caagataatt ctatgtggca ggtattacta tctcaattct
 5221 aagagtgaga aaatggagtt ttagaaacat ttactaactt gcctgggtca catagctaag
 5281 gaagagggtg acttgcccag ctttgcataa aactcctcaa aagagttgcc tatactccct
 5341 gactccactt atcttctac tatcctcttt ttaaaatata ttattttatt atttaataa
 5401 gcaatatatg aatgtggttt gaaattcaaa agacacaaag aagtatacag aggaaagcct
 60 5461 cactctcaat ccttctcaag gtttgctaatt tcctcttgca taggcaatcc gttcttccag
 5521 ctttgtgttt atcttccag agaagtttac tgtgtattaa gcaaatatgt atatctttat

5581 tcttgctcag tatttttcgca aacagcagct gtctaagttc actgttctga actttatttt
 5641 ttaaattaaa aatataatggc tatgtagtat tctatttta

SEQ ID NO: 24 Human TAP2 Isoform 1 Amino Acid Sequence (NP_000535.3)

5 1 mrlpdlrpwt slllvdaall wllqgplgtl lpqglpglwl egtlrlgglw gllklrgllg
 61 fvgtllllplc latpltvslr alvagasrap parvasapws wllvgygaag lswslwavls
 121 ppgaqekeqd qvnnkvlmwr llklsrpdlp llvaaffflv lavlgetlip hysgrvidil
 181 ggdfdpahafa saiffmclfs fgsslsagcr ggcftytmr inlrireqlf ssllrqdlgf
 241 fgetktgeln srlssdttlm snwlpnanv llrslvkvvlg lygfmlsisp rltllsllhm
 10 301 pftiaaekvy ntrhgevlre iqdavaragq vvreaavgglq tvrsfgaeeh evcrykeale
 361 qcrqlywrrd leralyllvr rvlhlvgvml mlscglqgmq dgeltqgsll sfmiyqesvg
 421 syvqtlvyiy gdmalsnvga ekvfsymdrq pnlpspgtla pttlqgvvkf qdvsfaypnr
 481 pdrpvlkglt ftlrpgevta lvgpngsgks tvaallqnly qptggqvlll ekpisyehc
 541 ylhsqvsvsv qepvlfsqsv rnniayglqs ceddkvmaaa qaahaddfiq emehgiytdv
 15 601 gekgsqlaag qkqrlaiara lvrpdrvlll deatsaldvq ceqalqdwms rgdrtvlvia
 661 hrlqavqrah qilvlqegkl qklaqlqegq dlyslrvqqr lmd

**SEQ ID NO: 25 Human TAP2 cDNA Variant 1, A allele Sequence
 (NM_001290043.1, CDS region from position 176-2236)**

20 1 agaataagag ccttggtctg ggaagcgaaa gcgaaagctg cccgagccct gacgcccgc
 61 ctggccgagc gtagctggcg gaccagagcc ggtagcgagg ttgggagaga cggagcggac
 121 ctacagcgtg aagcagaagt ccccgagact gcggtctccc cgccgaggct gagccatgcg
 181 gctccctgac ctgagaccct ggacctccct gctgctggtg gacgaggctt tactgtggct
 241 gcttcagggc cctctgggga ctttgccttc tcaagggtg ccaggactat ggctggaggg
 25 301 gacctgctg ctgggagggc tgtgggggct gctaaagcta agagggtgc tgggatttgt
 361 ggggacactg ctgctcccg cctgtctggc caccctccct actgtctccc tgagagccct
 421 ggtcgcgggg gcctcacgtg ctccccagc cagagtcgct tcagccctt ggagctggct
 481 gctgggtggg tacggggctg cggggctcag ctggtcactg tgggctgttc tgagccctcc
 541 tggagcccag gagaaggagc aggaccaggt gaacaacaaa gtcttgatgt ggaggctgct
 30 601 gaagctctcc aggcgggacc tgcctctcct cgttgccgcc ttcttcttcc ttgtccttgc
 661 tgttttgggt gagacattaa tccctcacta ttctggtcgt gtgattgaca tctgggagg
 721 tgattttgac ccccatgcct ttggcagtc catcttcttc atgtgctct tctccttgg
 781 cagctcactg tctgcaggct gccgaggagg ctgcttcacc tacaccatgt ctccaatcaa
 841 cttgcggatc cgggagcagc ttttctcct cctgctgcgc caggacctg gtttcttcca
 35 901 ggagactaag acaggggagc tgaactcacg gctgagctcg gataccacc tgatgagtaa
 961 ctggcttctc ttaaatgcca atgtgctctt gcgaagcctg gtgaaagtgg tggggctgta
 1021 tggcttcatg ctacagcata cgctcgaact caccctcctt tctctgctgc acatgccctt
 1081 cacaatagca gcggagaagg tgtacaacac ccgccatcag gaagtgtctc gggagatcca
 1141 ggatgcagtg gccagggcgg ggcaggtggt gcgggaagcc gttggagggc tgcagaccgt
 40 1201 tcgcagtttt ggggcccagg agcatgaagt ctgtcgctat aaagaggccc ttgaacaatg
 1261 tcggcagctg tattggcgga gagacctgga acgcgccttg tacctgctcg taaggagggt
 1321 gctgcacttg ggggtgcaga tgctgatgct gagctgtggg ctgcagcaga tgcaggatgg
 1381 ggagctcacc cagggcagcc tgctttcctt tatgatctac caggagagcg tggggagcta
 1441 tgtgcagacc ctggtataca tatatgggga tatgctcagc aacgtgggag ctgcagagaa
 45 1501 ggttttctcc tacatggacc gacagccaaa tctgccttca cctggcacgc ttgccccac
 1561 cactctgcag ggggttgtga aattccaaga cgtctcctt gcataatcca atcgccctga
 1621 caggcctgtg ctcaaggggc tgacgtttac cctacgtcct ggtgaggtga cggcgtggt
 1681 gggaccatga gggctctggga agagcacagt ggctgccctg ctgcagaatc tgtaccagcc
 1741 cacaggggga cagggtgctgc tggatgaaa gcccatctca cagtatgaac actgtacct
 50 1801 gcacagccag gtggtttcag ttgggcagga gcctgtgctg ttctccggt ctgtgaggaa
 1861 caacattgct tatgggctgc agagctgcga agatgataag gtgatggcgg ctgccaggc
 1921 tgccacgca gatgacttca tccaggaaat ggagcatgga atatacacag atgtagggga
 1981 gaagggaagc cagctggctg cgggacagaa acaacgtctg gccattgccc gggcccttgt
 2041 acgagaccog cgggtcctca tcctggatga ggctactagt gccctagatg tgcagtgcga
 55 2101 gcaggccctg caggactgga attcccgtgg ggatcgca gtgctggtga ttgctcacag
 2161 gctgcagaca gttcagcgcg cccaccagat cctggtgctc caggagggca agctgcagaa
 2221 gcttgcccag ctctaggagg gacaggacct ctattcccgc ctggtgcagc agcggctgat
 2281 ggactgaggc cccagggata ctgggcccct ttctcagggg cgtctccagg acccagagct
 2341 gttcctgctt tgagtttccc tagagctgtg cggccagata gctgttctct agttgcaggc
 60 2401 acgatggaga tttggacact gtgtgctttt ggtggggtag agagggtggg tgggggtggg

2461 tgggggctgt ctgtgtccag gaaacttaat tccctggtga ctagagcttt gcctggtgat
 2521 gaggagtatt ttgtggcata atacatatat tttaaaaatat tttccttctt acatgaactg
 2581 tatacattca tatagaaaat ttagacaata taaaaaagta caaagaagaa aagtaaaagt
 2641 acccattgtt tcacttcctg gagataacca tagttgctat tttgctgcct gtcccatcag
 5 2701 tcgtttatct gttgtttgag atagaaaatta accaaaaatg acataaatat tcatgagatt
 2761 gccttcctat atccttcctt gttcctacca gtgtctgcta ttttgaagaa gctaggggtct
 2821 ggagggacag agaacagttc cctgattaac agtattaata gcgacattgg taacagctac
 2881 cttttataga gttttaatgg gagtaggagc tatgctaagt gtttttcatg tattatcggt
 2941 tttaatcatt atccccaacc ctatgagggt gggtattatc cccattttac agatgaggaa
 10 3001 actgaagctc aaagaggctc aatgactttc ccaagggtgg cgtagtgggtg gagttggagt
 3061 ttgaacacag gcctgaccct agagtccaca cctgaccca atcaattata ttgcatcttg
 3121 ggtccataaa ccctaatacca taatcccatc aagaaaaagct ctgctgctct tagctctaaa
 3181 taattcagaa tctattctct tctctccagt cccgttggtta tagtcttcac tcatagactt
 3241 aagatgatcc catcaccaga gaggtttctc taccattagc ttccctcttc cggccattct
 15 3301 tcacaaagtc atttttctaa attctgtgtc acatacgatg atggcatttc tggaaattcc
 3361 ttcaggtgct ctcaagccct gctgcagaga tccttttcag agcacacatc gttccagccc
 3421 atctgtctca cctctcctg ttgtatccag ctccacgaca aacttctgcc ttcccaaca
 3481 cctttgtgcc ttgtcatatg gtgttttctt gccattttc tgctcgactc gcccttgatt
 3541 ttcaagttca agacttaact cagggttcag gtcttccagg aggccttact tgcctcgta
 20 3601 gtctggggaa ctctccatgt gcttctatca ctgtgcggtt acctctttca cagccctttt
 3661 aaagttctat cttccctttc ccaccttttt tgaccttcca ctagaccatg agcacctggg
 3721 cggaaagcca tatatcttat taagctttat atctgctacc tggccgaggg cctaattcat
 3781 agtggagaat aaatagtcaa ttgaataaat gaataaatat ctccaccatc gtactaatct
 3841 taatcctccc tgccactcc caccactgaa aatgcaacat tgtacacatc actggttggt
 25 3901 gggagggact taccttgga agttgctatt ctaggaaaga gaaaccttca tattcctgga
 3961 aacagcaggt agtttccagt gctggcaatg aattcccag aactgctggt ttggattttt
 4021 tcttgccctg cagctgttg gagcaggggt cagtgaggat ggggtgagag tgggcagttt
 4081 cttgtgcaga tttgcctttc tttcatcctg gggctgactt gcagctccac acccatccat
 4141 ctctcaaatt tcacagaggg taaaataggg atttggagag aaagaactct ggctgattc
 30 4201 ctttctctcc cacaatgtc ctttattcat aaaacaggaa taataattcc tgtatctccc
 4261 aactacatgg aagctgcagc cctcacagaa gaagatgatc tgagaaattc tttgatttcc
 4321 tcagtacagt tatacccatg catcataata ctttaagcct ggaaggcatc ttaaaaaata
 4381 tgcaacagtc aaacctaatt ttacagagaa actgacatga aatcacgcag ctaatcatga
 4441 taaagctggg tggaaaactt atcttgatgg gcagtacagg aagatgcagt agacctaaag
 35 4501 atgtcctgaa agtttcttat ctcaggggaa actcccaggt aggccttatg tcagggacac
 4561 agaaaaatgc tccctgaaag tcaaaaatatt cgggctagac agacaaattc ctgtaagtgt
 4621 gggttgctc ggaaccacag atgtcactaa tctggtttg ctccagagtt cttttgttc
 4681 actcctaccc cccatcacca tttgattgat ctcttacc tgtaatctcc ctttctgtc
 4741 gcttacctgc agtatctttc ccaccagggc atgccttatt ctttctaaag gaaagtatga
 40 4801 atggagaggg gaaagcttg gaaactgata gatttccttg gatgcaaaa cacctccata
 4861 gcctgtctgc ccggccctat gtggaaacag cattgagttt caagtccttt atgcctccac
 4921 ccagggatag ccacttgtaa tccacatggc aattgtgaaa caagcaggaa atgcgtaatt
 4981 gtcagaattt tgtggggaaa ggactaggga ataaggaaaa caaagatctt ccttggtgtt
 5041 tagagctgtc agctagagga gcacctgctt gactctgatg ccatctaagt gtcccagaag
 45 5101 aaactgggtt ttgaacctag agttccatgg actcttagga attagactac tactactact
 5161 aagcattcac tgggtgcttac tatgtgctat tgctgtgcca agtatctgaa acctgtcttc
 5221 ttaccttatt tttcaagata attctatgtg gcaggatta ctatctcaat tctaagagtg
 5281 agaaaatgga gtttttagaaa catttactaa cttgcctggg tcacatagct aaggaagagg
 5341 tggacttgcc cagctttgca taaaactcct caaaagagtt gcctatactc cctgactcca
 50 5401 cttatcttcc tactatcctc tttttaaaat atattattta tttattttaa taagcaatat
 5461 atgaatgtgg tttgaaattc aaaagacaca aagaagtata cagaggaaag cctcactctc
 5521 aatccttctc aagggttgct aattcctctt gcataggcaa tccgttcttc cagctttgtg
 5581 tttatcttcc cagagaagtt tactgtgtat taagcaaata tgtatatctt tattcttgct
 5641 cagtattttc gcaaacagca gctgtctaag ttcactgttc tgaactttat tttttaaatt
 55 5701 aaaaatatat ggctatgtag tattctattt ta

SEQ ID NO: 26 Human TAP2 Isoform 2 Amino Acid Sequence**(NP_001276972.1)**

1 mrlpldlrptw slllvdaall wllqgplgtl lpqglpplwl egtlrlgglw gllklrgllg
 60 61 fvgtllllplc latpltvslr alvagasrap parvasapws wllvgygaag lswslwavls

121 ppgaqekekq qvnnkvlmwr llklsrpdlp llvaaffflv lavlgetlip hysgrvidil
 181 ggdfdpahafa saiffmclfs fgsslsagcr ggcftytmsr inlrireqlf ssllrqdlgf
 241 fgetktgeln srlssdttlm snwlpnanv llrslvkvg lygfmlsisp rltllsllhm
 301 pftiaaekvy ntrhgevlre iqdavaragq vvreaavgglq tvrsfgaeeh evcrykeale
 5 361 qcrqlywrrd leralyllvr rvlhlvgvml mlscglqqmq dgeltqgsll sfmiyqesvg
 421 syvqtlvyiy gdmalsnvгаа ekvfsymdrq pnlpspgtla pttlqgvvkf qdvsfaypnr
 481 pdrpvlkglt ftlrpgevta lvgpngsgks tvaallqnly qptggqvlll ekpisyehc
 541 ylhsqvsvvg qepvlfsgsv rnniayglqs ceddkvmaaa qaahaddfiq emehgiytdv
 601 gekgsqlaag qkqrlaiara lvrdrprvlil deatsaldvq ceqalqdwms rgdrtvlvia
 10 661 hrlqtvqrah qilvlqegkl qklaql

SEQ ID NO: 27 Human TAP2 cDNA Variant 2 Sequence (NM_018833.2, CDS region from position 123-2084)

1 gcccgccttg gccgagcgta gctggcggac cagagccggt agcgaggttg ggagagacgg
 15 61 agcggacctc agcgtgaag cagaagtccc cggagctgcg gtctccccgc cgcggctgag
 121 ccatgcccgt cctgacctg agaccctgga cctccctgct gctggtggac gcggctttac
 181 tgtggctgct tcaggccct ctggggactt tgcttcctca agggctgcca ggactatggc
 241 tggaggggac cctgcccgtg ggagggctgt gggggctgct aaagctaaga gggctgctgg
 301 gatttggtgg gacactgctg ctcccgtct gtctggccac ccccctgact gtctccctga
 20 361 gagccctggg cgcggggggc tcacgtgctc cccagccag agtcgcttca gccccttgga
 421 gctggctgct ggtggggtag ggggctgcgg ggctcagctg gtactgtgg gctgttctga
 481 gccctcctgg agcccaggag aaggagcagg accaggtgaa caacaaagtc ttgatgtgga
 541 ggctgctgaa gctctccagg ccggacctgc ctctcctcgt tgccgccttc ttcttccttg
 601 tccttgctgt tttgggtgag acattaatcc ctactattc tggctgtgtg attgacatcc
 25 661 tgggaggtga ttttgacccc catgcctttg ccagtgccat cttcttcagt tgcctcttct
 721 cctttggcag ctactgtct gcaggctgccc gaggaggctg cttcacctac accatgtctc
 781 gaatcaactt gcggatccgg gagcagcttt tctcctccct gctgcgccag gacctcggtt
 841 tcttcaggga gactaagaca ggggagctga actcacggct gagctcgat accaccctga
 901 tgagtaactg gcttccttta aatgccaatg tgctcttgcg aagcctgggt aaagtgggtg
 30 961 ggctgtatgg cttcatgctc agcatatcgc ctgcactcac cctcctttct ctgctgcaca
 1021 tgcccttcac aatagcagcg gagaagggtg acaacacccc ccatcaggaa gtgcttcggg
 1081 agatccaggga tgcagtggcc agggcggggc aggtgggtgc ggaagccgtt ggagggctgc
 1141 agaccgttcg cagttttggg gccgaggagc atgaagtctg tcgctataaa gaggcccttg
 1201 aacaatgtcg gcagctgtat tggcggagag acctggaacg cgccttgtag ctgctcgtaa
 35 1261 ggaggggtgct gcacttgggg gtgcagatgc tgatgctgag ctgtgggctg cagcagatgc
 1321 aggatgggga gctcaccagc ggcagcctgc tttcctttat gatctaccag gagagcgtgg
 1381 ggagctatgt gcagaccctg gtatacatat atggggatat gctcagcaac gtgggagctg
 1441 cagagaaggt tttctcctac atggaccgac agccaaatct gccttcacct ggcacgcttg
 1501 cccccaccac tctgcagggg gttgtgaaat tccaagacgt ctcttttgca tatcccaatc
 40 1561 gccctgacag gcctgtgctc aaggggctga cgtttaccct acgtcctggt gaggtgacgg
 1621 cgctggtggg acccaatggg tctgggaaga gcacagtggc tgccctgctg cagaatctgt
 1681 accagcccac agggggacag gtgctgctgg atgaaaagcc catctcacag tatgaacact
 1741 gctacctgca cagccaggtg gtttcagttg ggcaggagcc tgtgctgttc tccggttctg
 1801 tgaggaacaa cattgcttat gggctgcaga gctgcgaaga tgataagggt atggcggctg
 45 1861 ccagggctgc ccacgcagat gacttcatcc aggaaatgga gcatggaata tacacagatg
 1921 taggggagaa gggaagccag ctggctgcgg gacagaaaca acgtctggcc attgccggg
 1981 cccttgtagc agaccgcgg gtccctcatc tggatgaggc tactagtgc ctagatgtgc
 2041 agtgcgagca ggccaaaacc ctttggaagt tcatgatatt ttgaatttca atggatattt
 2101 cctgggaata atgagttcaa atgaacgaat atgtggaaca aagcatcacc aacatttatt
 50 2161 ttttcaggat gaggtgatgg acaaaaccat cacagggaaa ttgaggcaaa tagtacatgt
 2221 aaaacaatac ttcgggtgag tccacctatc ccaaagtcgt atcaaagaag tggctgcaga
 2281 ttggagccca aagcctttgg ttcctcagtt tccaaatgga ttctcactag gtgggatcat
 2341 gagtttgctt tggacacccc aaattctaac tatttctttt gtttcttaca tcctttccct
 2401 cttccccagc cccttcccct catgttacac ctcttgctgg ttgagacgt caatcaccac
 55 2461 tgagaaagaa ttaaaccagt attttgagct ggcaaaattc ttagcctagt acaattcctt
 2521 caattaaact gtagctcaac

SEQ ID NO: 28 Human TAP2 Isoform 3 Amino Acid Sequence (NP_061313.2)

1 mrlpdlrpwt slllvdaall wllqgplgtl lpqglpglwl egtlrlgglw gllklrgllg
 60 61 fvgtlllplc latpltvslr alvagasrap parvasapws wllvgygaag lswslwavls

121 ppgaqekeqd qvnnkvlmwr llklsrpdlp llvaaffflv lavlgetlip hysgrvidil
 181 ggdfdpahafa saiffmclfs fgsslsagcr ggcftytmr inlrireqlf ssllrqdlgf
 241 fgetktgeln srlssdttlm snwlpnanv llrslvkvg lygfmlsisp rltllsllhm
 301 pftiaaekvy ntrhqeavlre iqdavaragq vvreaavgglq tvrsfgaeeh evcrykeale
 5 361 qcrqlywrrd leralyllvr rvlhlvgvml mlscglqqmq dgeltqgsll sfmiyqesvg
 421 syvqtlvyiy gdmalsnvгаа ekvfsymdrq pnlpspgtla pttlqgvvkf qdvsfaypnr
 481 pdrpvlkglt ftlrpgevta lvgpngsgks tvaallqnly qptggqvlll ekpisyehc
 541 ylhsqvsvsv gepvlfsgsv rnniayglqs ceddkvmaaa qaahaddfiq emehgiytdv
 601 gekgsqlaag qkqlraiara lvr dprvlil deatsaldvq ceqaktlwkf mif

SEQ ID NO: 29 Mouse TAP2 cDNA Sequence (NM 011530.3, CDS region from position 154-2262)

1 accgcgacca cggcagtgaa gtgaaagcga aagccgcggg acccaggcgc gctccccgcg
 61 agggcgctgc tgcgcaccca ggagatccag tttgagaaga agcagattcc agaagctctc
 15 121 ctgagctgcc gctccgcagc cgcagaaccc accatggcgc tgcctacct gaggccctgg
 181 gtctctctgc tgctggcgga catggcttta cttgggttg tacaaggatc tctgggaaat
 241 ctgcttcccc aggggctgcc aggactctgg atagagggca ccctgcgact tggagtgtcg
 301 tggggactgc taaaagtggg agagctgctg ggacttgtgg ggacccttct gcccttgctc
 361 tgcccttgcca ctcccctgtt tttctcgcta agagctctgg tgggaggcac cgcgagcacc
 20 421 tcagtagtcc gagtggtctt tgccctcttg ggctggctgc tggctggcta tggggctgtt
 481 gcgctgagct gggccgtgtg ggctgtgctg agcccggctg gaggccagga gaaggaacca
 541 ggccaggaga acagaacact gatgaagcgg ttgctgaagc tgcacaggcc ggacctgect
 601 ttccctcatag ctgccttctt cttccttctg gtggctgtgt ggggggagac attaatccct
 661 cgctattcgg gtctgttaat tgacatcctg ggaggtgatt tcgaccccca cgcctttgca
 25 721 agcgccatct ttttcatgtg cctgttctct gttgggagct ctttctctgc aggctgtaga
 781 ggaggtcctc tccctctcac catgtccagg atcaacctgc ggatacaga gcagcttttc
 841 tcatctttgt tgcgccaaga ccttggtatt ttccaggaga ccaagacagg ggagctgaac
 901 tcgaggctga gctctgacac ctctctgatg agccgctggc tccctttcaa tgccaatatc
 961 ctgctgcgga gcctggtgaa ggtggtgggg ctctacttct tcatgctcca ggtatcgccc
 30 1021 cgactcacct tccctccct gctggacctg cccctcacga tagcagctga gaaggtgtac
 1081 aacccccgcc atcaggcggt gctaaaggag atccaggatg cagtggccaa ggcggggcag
 1141 gtggtgcgag aggcggtagg agggctgcag actgtgcgaa gctttggggc cgaggagcag
 1201 gaagtcagcc actacaagga ggccctggag cgatgtagac agctgtggtg gcgccgagac
 1261 ctggaaaaag acgtgtatct agtcatacgg aggggtgatg ccttgggcat gcaggtgtcg
 35 1321 attctgaact gcggcggtga gcagattctg gctggagagg tcacccgggg tggcctgctc
 1381 tccttcctgc tgtaccagga ggaagtggga caatatgtcc ggaacctggt ttacatgtac
 1441 ggggatatgc tgagcaacgt gggcgctgct gaaaagggtt ttcctacct ggaccgaaag
 1501 ccgaatctgc cccagcctgg gatcctggcc cctccctggc tggagggggc cgtggaattc
 1561 caagacgtct ccttttcgta tcccaggcgc cccgagaagc ctgtgctcca gggctgtgac
 40 1621 ttcacctgac atcctggaac ggtgacagcg ttggtgggac ccaatggatc agggagagc
 1681 accgtggccg ccctgctgca gaacctgtac cagccactg ggggccagct gctgtggtat
 1741 ggcgagcccg tgaccgagta tgatcaccac tacctgcacc gccaggtggt tctggtggg
 1801 caggagccct tgctgttctc ggggttctgt aaggacaata ttgcctatgg cctgagggac
 1861 tgtgaggacg ctcaagtgat ggcagctgcc caggcggcct gtgcagacga cttcataggg
 45 1921 gaaatgacta atggaataaa cacagaaatc ggggaaaaag ggggccagtt agctgtggga
 1981 cagaagcaac gtctggccat tgcccggggc cttgtgcgga acccacgggt cctcatcctg
 2041 gatgaggcta ccagcgccct ggacgcccag tgtgaacagg ccctacagaa ctggagatcg
 2101 cagggggaca ggacgatgct ggtgattgcc cacaggctgc acacggttca gaatgctgac
 2161 caagtctctg tgctcaagca gggacgtctg gtggagcatg accagctcag ggacggccag
 50 2221 gatgtctacg cccacctggt acagcagcgg ctggaggcat gaggcctcca gacctgagc
 2281 ccctctcagg actgtggcca ggatcagacc cacagggacc gtgcccagg aggctagggt
 2341 caggataaag attgggaccg ttttgactt tgtgtgtttt tgtggccttg atgggtgagg
 2401 gttgggggag tgggtggttt gtggttctag aaacattatt tttttgctga atataaaaat
 2461 aataaatata tataaaatgc ttttccttaa

SEQ ID NO: 30 Mouse TAP2 Amino Acid Sequence (NP 035660.3)

1 malsylrpwv sllladmall gllqgslgnl lpqglpplwi egtlrlglvw gllkvgellg
 61 lvgtlplllc latplffslr alvvggtasts vvrvasaswg wllagygava lswavwavls
 121 pagvqekepg qenrtlmkrl lklsrpdplf liaaffflvv avwgetlipr ysgrvidilg
 60 181 gdfdpdafas aiffmclfsv gssfsagcrg gsflftmsri nlrireqlfs slrrqdlgff

241 getktgelns rlssdtslms rwlpfnanil lrslvkvvgl yffmlqvspr ltflslldlp
 301 ltiaaekvyn prhqavlkei qdavaxagqv vreaavgglqt vrsfgaeeqe vshykealer
 361 crqlwrrrdl ekdvylvirr vmalgmqvli lncgvqqila gevtrgglls fllyqeevgq
 421 yvrnlvymyg dmlsnvgaae kvfsyldrkp nlpppgilap pwlegrvefq dvsfsyprp
 5 481 ekpvlqgltf tlhpgvtal vgpngsgkst vaallqnlyq ptggqlldg eplteydhhy
 541 lhrqvvlvgg epvlfsgsvk dniayglrdc edaqvmaaaq aacaddfige mtnginteig
 601 ekggqlavgg kqrlaiaral vrnprvlild eatsaldaqc equalqnwrsq gdrtmlviah
 661 rlhtvqnadq vlvlkqgrlv ehdqlrdgqd vyahlvqqrl ea

10 **SEQ ID NO: 31 Human TAPBP cDNA Variant 1 Sequence (NM_003190.4, CDS region from position 347-1693)**

1 ggggacgcgg cacagatagg gggaagccgg agtaatgggtt ttccgggcaag tggatgttgg
 61 agagcacaca caggagtggg ggggcggggg agggcctggg gttggggagg gctcgaactc
 121 ggggctgctg ggtagtccag gagggcgcgg taaggctggg gtgtcctggt gagaactgga
 15 181 gaggatctac ccgggtccct gcctggccag tggggaaaca ccgggtcccc aggcaccttc
 241 acctaaccag agcggggatt tccaccgccc ctcatgccgc cctttggagg aaagtgaag
 301 tgaaaggagg aagaggaggc ttcattggctg aggaggtcgc agcgccatga agtccctgtc
 361 tctgtctctc gctgtggctt tgggcctggc gaccgccgtc tcagcaggac ccgctgtgat
 421 cgagtgttgg ttctgtggagg atgcgagcgg aaagggcctg gccaaagagac ccggtgcact
 20 481 gctgttgccg cagggaccgg gggaaaccgg gccccggccg gacctcgacc ctgagctcta
 541 tctcagtgtg cacgaccccc cgggcgccct ccaggctgcc ttccaggcgg atccccgggg
 601 cgcccccgca ccacactgcg agatgagccg ctctgtgctt ctccccgctt ctgcgaaatg
 661 ggccagcggc ctgacccccg cgcagaactg cccgcggggc ctggatgggg ctgtgctgat
 721 ggtcagcata tccagcccag tcctcagcct ctccagcctc ttgcgaccac agccagagcc
 25 781 tcagcaggag cctgttctca tcaccatggc aacagtggta ctgactgtcc tcaccacac
 841 cctgcccctc cgagtgcgac tgggacaaga tgctctgctg gacttgagct ttgacctacat
 901 gccccccacc tccgagcccg cctcatctct ggctccgggt cccccctcct ttggcctaga
 961 gtggcgacgc cagcacctgg gtaagggaca tctgtctcct gctgcaactc ctgggctgaa
 1021 tggccagatg ccagcagccc aagaaggggc cgtggcattt gctgcttggg atgatgatga
 30 1081 gccatggggc ccatggaccg gaaatgggac cttctggctg cctacagtcc aaccttttca
 1141 ggagggcacc tatctggcca ccatacacct gccatacctg caaggacagg tcacctgga
 1201 gcttctgttg taaaaacccc ccaaagtgtc cctgatgcca gcaacccttg cacgggcccgc
 1261 cccaggggag gcacccccgg aattgctctg ccttgtgtcc cacttctacc cttctggggg
 1321 cctggagggtg gagggggaac tccgggggtg cccagggggc cgctctcaga aggccgaggg
 35 1381 gcagaggtgg ctctcgcccc tgcgccacca ttccgatggc tctgtcagcc tctctgggca
 1441 cttgcagccg cccccagtc aactgagcga gcatggggca cgctatgcct gtcgaattca
 1501 ccatcccagc ctgcctgcct cggggcgccg cgctgagggt accctggagg tagcaggctt
 1561 ttcaggggcc tcccttgagg acagcgtagg ccttttctct tctgcttttc ttctgcttgg
 1621 gctcttcaag gcactgggct gggctgctgt ctacctgtcc acctgcaagg attcaaagaa
 40 1681 gaaagcagag tgagggcact cactgccatc ctgtggaagc caccatcatc tctggcccaa
 1741 gcttctgtag tagctcccta aaataatacc ctatcatctg ctctaatacc ctccaatctc
 1801 tctccactga gtggctggaa tgcttttttt tttttctttt acttataaa gggataattt
 1861 tctttttttt tttttttttt agacggagtg tctactcttc gccaggctg gccgctcag
 1921 gcatgatctt ggcttactgc aacctccgcc tctgggttc aagcaattct gtggcttcag
 45 1981 cctccggagt agctgggatt acaggcacat gccaccacac ccagtgaatt tttgtatttt
 2041 tagtagagac ggggtttcac catgttggcc aggctggtct tgaattcctg acctcagggt
 2101 atctgcccac ctccagcctcc caaagtgtg ggattacagg cgtgagccac cacaccaggc
 2161 ccgagaaatg ctttttttaa aaacacacat cttatggcat tcaccttctt ggagctctag
 2221 gacagtgggt ctcaaaattt ttttctctca ggacctctta aaaatcatca aggaccccaa
 50 2281 aaagcttttg ggtatgtggg ttatagctat caatatttat ggtactagaa cttaaaagt
 2341 agaaaaattt aaaacacgag aatacatagg cacacattct attcatcgtg ggaaccatgg
 2401 tgtcaataca tatcatgtag cttctgaaaa actccactgt acacttatag aatgaagaag
 2461 gcaaaaaact tttttttttt tttttttgag acggagtctc gctctgtcgc ccaggctgga
 2521 gtgcagtggc gcgatctcgg ctcaactgaa gctccgcctc tcgggttcac gccattctcc
 55 2581 tgctcagcc tcccaagtag ctccggactac aggcgtcctc caccatgcct ggctaattat
 2641 ttgtattttt tagtagagac ggggtttcac cgtgttagcc aggatggtct cgatctccta
 2701 acctggtgat ccgcccgcct cggcctccca aagtattggg attaccgcg tgagccaccg
 2761 cgcccggctg caaataatct ttcttttttt ctgagacaga gtctcgctct gttgcccagg
 2821 ctggagtgca gtggcacgat ctccggctcac ggcacgctcc gcctccgggg ttacgacct
 60 2881 tctcctgcct cagcttcccg agtagctggg actacagggg cccgccacca cgcccggcta

2941 acttttttgtg ttttttagtag agacgggggtt tcaccgtgtt agccaggatg gtctcgatct
 3001 cctgaccttg tgatctgccc gcctcggcct cccaaagtgc tgggattaca ggctgagcc
 3061 accgcgcccc gcggcgaaac acgatattgt actaacatct taatttttgtt ataaaatctc
 3121 acaaaccccc tgacatagtc tcagagatct gtagggccga ggttacattt ggagaacccg
 5 3181 tactctaggg ccaaattccat tcttcttgcc ctggctcact tgtccccccc accgccccgc
 3241 gctggagcca ctgcctagtt cttcagccct agatgggtgct cgccagacct cctctcaatg
 3301 ctcatcacac acagggtctat tcctttcctc caatgaacca aacgcctccc gccacctcc
 3361 aggtcccagt cctctgttcc ctttgccctg tccacccttg ccctccctgg gtcgcagacg
 3421 aggtcggcct cgtcattccc cgcagaccgc cgcgcgtccc tcttggtcggg ttcaccacag
 10 3481 ttgtatttaa gtgatcgtgt gagtcgtcgt taaatgcctg tctccccgcg gatcatgggc
 3541 tcctcgagga cagggactgg cctgtctgtc cactgctgta accccgcgcc ggcataggga
 3601 cctaaggccc actggagggc gctcatcaag tagctgctgg atgttgacga aggaagcggc
 3661 ggcgcagctc agggatctcc gagtcaggac ggtcggccag acccacgggg taacgggtct
 3721 aatcgtgtag gaataaagct gtattccagt gcttccaaaa aaaaaaaaaa aaaaaaaaaa

SEQ ID NO: 32 Human TAPBP Isoform 1 Amino Acid Sequence (NP_003181.3)

1 mkslslllav alglatavsa gpaviecwfv edasgkglak rpgalllrqg pgeppprpdl
 61 dpelylsvhd pagalqaafr ryprgapaph cemsrfvplp asakwasglt paqncprald
 121 gawlmvsiss pvlsllsllr pqpepqqepv litmatvvlt vlthtpapr vrlgqdalldl
 20 181 sfaympptse aasslapgpp pfglewrrqh lgkghlllaa tpglngmpa aqegavafaa
 241 wdddepwgpw tngtffwlp vqpfqegtyl atihlpylqg qvtlelavyk ppkvsllmpat
 301 laraapgeap pellclvshf ypsggglewv elrggpggrs qkaegqrwls alrhhsdgsv
 361 slsghlqppp vtteqhgary acrihhpslp asgrsaevtl evaglsqpsl edsvglflsa
 421 fllllglfkal gwaavylstc kdskkkae

SEQ ID NO: 33 Human TAPBP cDNA Variant 2 Sequence (NM_172208.2, CDS region from position 347-1861)

1 ggggacgcgg cacagatagg ggggaagccgg agtaatggtt ttcgggcaag tggatgttgg
 61 agagcacaca caggagttgg ggggcggggg agggcctggg gttggggagg gctcgaactc
 121 ggggctgctg ggtagtccag gagggcgccg taaggctggg gtgtcctggt gagaactgga
 181 gaggatctac ccgggtccct gcctggccag tggggaaaca ccgggtcccc aggcaccttc
 241 acctaaccag agcggggatt tccaccgccc ctcatgccgc cctttggagg aaagtgaag
 301 tgaaaggagg aagaggaggc ttcattggctg aggaggtcgc agcgccatga agtcctgtc
 361 tctgtctctc gctgtggctt tgggcctggc gaccgccgtc tcagcaggac ccgcggtgat
 421 cgagtgttgg ttcgtggagg atgcgagcgg aaaggccctg gccaaagagac ccggtgact
 481 gctgttgccg cagggaaccg gggaaaccgc gccccggccg gacctcgacc ctgagctcta
 541 tctcagtgtg caccgccccg cgggcgcctt ccaggctgcc ttcaggcggg atccccgggg
 601 cggccccgca ccactctgc agatgagccg ctctgtgcct ctccccgcct ctgcgaaatg
 661 ggccagcggc ctgacccccg cgcagaactg cccgcgggcc ctggatgggg cttggctgat
 721 ggtcagcata tccagcccag tcctcagcct ctccagcctc ttgcgaccac agccagagcc
 781 tcagcaggag cctgttctca tcaccatggc aacagtggta ctgactgtcc tcaccacac
 841 ccctgccccct cgagtggagc tgggacaaga tgctctgctg gacttgagct ttgectacat
 901 gccccccacc tccgaggccg cctcatctct ggctccgggt cccctccctt ttgggctaga
 961 gtggcgacgc cagcacctgg gtaagggaca tctgtcctg gctgcaactc ctgggctgaa
 1021 tggccagatg ccagcagccc aagaaggggc cgtggcattt gctgcttggg atgatgatga
 1081 gccatggggc ccattggacc gaaatgggac cttctggctg cctacagttc aaccctttca
 1141 ggagggcacc tatctggcca ccatacact gccatacctg caaggacagg tcaccctgga
 1201 gcttgctgtg taaaaacccc ccaaagtgtc cctgatgcca gcaacccttg cacgggcccg
 1261 ccagggggag gcacccccgg aattgctctg ccttgtgtcc cacttctacc cttctggggg
 1321 cctggagggtg gactgggaac tccgggggtg ccaggggggc cgctctcaga aggccgaggg
 1381 gcagagggtg ctctcgccc tgccgccacca ttccgatggc tctgtcagcc tctctgggca
 1441 cttgcagccg cccccagtca ccactgagca gcatggggca cgctatgcct gtcgaattca
 1501 ccattccagc ctgcctgcct cggggcgccg cgctgaggtc accctggagg tagcaggtct
 1561 ttcaggggcc tcccttgagg acagcgtagg ctttttctct tctgccttct ttctgcttgg
 1621 gctcttcaag gactgggct gggctgctgt ctacctgtcc acctgcaagg attcaaagaa
 1681 ggtacagtgc tccacctctc tgtatctttc ccttgtcact ttatctcctc atcctatctc
 1741 aaaacccatg gagggaggct gctgggtgtg taggcagaac ctaggcttgg aattcacact
 1801 gatctgggtt aaaacctggc actatatcct aactgtagga ctctttgagc atgctactta
 1861 atctatatgt ttccttgggt gtagggattt taaaaagtta cttaggcagt gctgtccaat
 1921 agaaagataa tgcaagccac atatgtaaat taaatactc tagtactcac attaaaaaaa

1981 taagcagaaa caggtaaaat taa

SEQ ID NO: 34 Human TAPBP Isoform 2 Amino Acid Sequence (NP_757345.2)

	1	mkslsllllav	alglatavsa	gpaviecwfv	edasgkglak	rpgallllrqq	pgeppprpdl
5	61	dpelylsvhd	pagalqaafr	ryprgapaph	cemsrfvplp	asakwasgl	paqncprald
	121	gawlmvsiss	pvlslssllr	pqpepqqepv	litmatvvlt	vlthtpaprv	rlgqdalldl
	181	sfaympptse	aasslapgpp	pfglewrrqh	lgkghlllala	tpglngqmpa	aqegavafaa
	241	wdddepwgpw	tgngtfwlpt	vqpfqegtyl	atihlpylqq	qvtlelavyk	ppkvslnmpat
	301	laraapgeap	pellclvshf	ypsgglevew	elrggpggrs	qkaegqrwls	alrhhsdgs
10	361	slsghlqppp	vtteqhgary	acrihhpslp	asgrsaevtl	evaglsgpsl	edsvglflsa
	421	flllglfkal	gwaavylstc	kdkkvqcst	slyslvtls	phpiskpmeg	gcwcgrqnlq
	481	leftliwvkt	whyiltvqlf	ehat			

SEQ ID NO: 35 Human TAPBP cDNA Variant 3 Sequence (NM_172209.2, CDS region from position 347-1432)

15	1	ggggacgcgg	cacagatagg	gggaagccgg	agtaatggtt	ttcgggcaag	tggatgttgg
	61	agagcacaca	caggagttag	ggggcggggg	agggcctggg	gttggggagg	gctcgaactc
	121	ggggctgctg	ggtagtccag	gagggcgcg	taaggctggg	gtgtcctggt	gagaactgga
	181	gaggatctac	ccgggtccct	gcctggccag	tggggaaaca	ccgggtcccc	aggcaccttc
20	241	acctaaccag	agcggggatt	tccaccgccc	ctcatgccgc	cctttggagg	aaagtgaag
	301	tgaaaggagg	aagaggaggc	ttcatggctg	aggaggtcgc	agcgccatga	agtccctgtc
	361	tctgtctctc	gctgtggctt	tgggcctggc	gaccgccgtc	tcagcaggac	ccgcggtgat
	421	cgagtgttgg	ttcgtggagg	atgcgagcgg	aaagggcctg	gccaagagac	ccggtgcact
	481	gctgttgccg	cagggaccgg	gggaaccgcc	gccccggccg	gacctcgacc	ctgagctcta
25	541	tctcagtgtg	cacgtggtac	tgactgtcct	caccacacac	cctgcccctc	gagttagact
	601	gggacaagat	gctctgctgg	acttgagctt	tgccctacatg	ccccccacct	ccgaggccgc
	661	ctcatctctg	gctccgggtc	cccctccctt	tgggctagag	tggcgacgcc	agcacctggg
	721	taagggacat	ctgctcctgg	ctgcaactcc	tgggctgaat	ggccagatgc	cagcagccca
	781	agaagggggc	gtggcatttg	ctgcttgagg	tgatgatgag	ccatggggcc	catggaccgg
30	841	aaatgggacc	ttctggctgc	ctacagttca	accctttcag	gagggcacct	atctggccac
	901	catacacctg	ccatacctgc	aaggacaggt	caccctggag	cttgcgtgtg	acaaaccccc
	961	caaagtgtcc	ctgatgccag	caacccttgc	acgggcccgc	ccaggggagg	cacccccgga
	1021	attgtctctg	cttgtgtccc	actctacacc	ttctgggggc	ctggaggtgg	agtgggaact
	1081	ccggggttgg	ccagggggcc	gctctcagaa	ggccgagggg	cagaggtggc	tctcggccct
35	1141	gcgccaccat	tccgatggct	ctgtcagcct	ctctggggac	ttgcagccgc	ccccagtcac
	1201	cactgagcag	catggggcac	gctatgcctg	tcgaattcac	catcccagcc	tgccctgcctc
	1261	ggggcgccagc	gctgaggtca	ccctggaggt	agcaggtctt	tcagggccct	cccttgagga
	1321	cagcgtaggc	cttttcctgt	ctgcctttct	tctgcttggg	ctcttcaagg	cactgggctg
	1381	ggctgctgtc	tacctgtcca	cctgcaagga	ttcaaagaag	aaagcagagt	gagggcactc
40	1441	actgccatcc	tgtggaagcc	accatcatct	ctggcccaag	cttctgtagt	agctccctaa
	1501	aataataccc	tatcatctgc	tcctaataccc	tccaatctct	ctccactgag	tggctggaat
	1561	gctttttttt	ttttctttca	cttatataag	ggataatttt	tctttttttt	ttttttttga
	1621	gacggagtct	cactcttccg	cccaggctgc	agtgcagtgg	catgatcttg	gcttactgca
	1681	acctccgcct	cctgggttca	agcaattctg	tggcttcagc	ctccggagta	gctgggatta
45	1741	caggcacatg	ccaccacacc	cagtgaattt	ttgtattttt	agtagagacg	gggtttcacc
	1801	atgttggcca	ggctggtctt	gaattcctga	cctcaggtga	tctgcccacc	tcagcctccc
	1861	aaagtgtctg	gattacaggc	gtgagccacc	acaccaggcc	cgagaaatgc	tttttttaaaa
	1921	aacacacatc	ttatggcatt	caccttcttg	gagctctagg	acagtgggtc	tcaaaatttt
	1981	tttctctctc	gacctcttaa	aaatcatcaa	ggacccccaa	aagcttttgg	gtatgtgggt
50	2041	tatatgtatc	aataatttatg	gtactagaac	ttaaaagtga	gaaaaattta	aaacacgaga
	2101	atacataggc	acacattcta	ttcatcgtgg	gaaccatggt	gtcaatacat	atcatgtagc
	2161	ttctgaaaaa	ctccactgta	cacttataga	atgaagaagg	caaaaaactt	tttttttttt
	2221	tttttttgaga	cggagtctcg	ctctgtcgcc	caggctggag	tgacgtggcg	cgatctcggc
	2281	tactgcaag	ctccgcctct	cgggttcacg	ccattctcct	gcctcagcct	cccaagtagc
55	2341	tcggactaca	ggcgtcctcc	accatgcctg	gctaataattt	tgtatttttt	agtagagacg
	2401	gggtttcacc	gtgttagcca	ggatggtctc	gatctcctaa	cctgggtgatc	cgcccgctc
	2461	ggcctcccaa	agtattggga	ttacccgcgt	gagccaccgc	gcccggctgc	aaataatctt
	2521	tctttttttt	tgagacagag	tctcgtctctg	ttgcccaggc	tggagtgcag	tggcacgatc
	2581	tcggctcacg	gcacgctccg	cctcccgggt	tcacgccatt	ctcctgcctc	agcttcccga
60	2641	gtagctggga	ctacaggggc	ccgccaccac	gcccggctaa	ctttttgtgt	tttttagtaga

2701 gacgggggttt caccgtgtta gccaggatgg tctcgatctc ctgaccttgt gatctgcccc
 2761 cctcggcctc ccaaagtgc gggattacag gcgtgagcca ccgcgcccgg cggcgaaaca
 2821 cgatattgta ctaacatctt aattttgtta taaaatctca caaacccccct gacatagtct
 2881 cagagatctg tagggccgag gttacatttg gagaacccgt actctagggc caaatccatt
 5 2941 cttcttgccc tggctcactt gtccccccca ccgccccggc ctggagccac tgcctagttc
 3001 ttcagcccta gatggtgctc gccagacctc ctctcaatgc tcatcacaca cagggtattt
 3061 cctttcctcc aatgaaccaa acgcctcccg cccacctcca ggtcccagtc ctctgttccc
 3121 tttgcttggc ccacccttgc cctccctggg tcgcagacga ggtcggcctc gtcattcccc
 3181 gcagaccgcc gcgcgtccct cttgtgcggt tcaccacagt tgtatttaag tgatcgtgtg
 10 3241 agtcgtcgtt aaatgcctgt ccccccgcg atcatgggct cctcgaggac agggactggc
 3301 ctgtctgtcc actgctgttaa ccccgcgccg gcatagggac ctaaggccca ctggagggcg
 3361 ctcatacaagt agctgctgga tgttgacgaa ggaagcggcg gcgcagctca gggatctccg
 3421 agtcaggacg gtcggccaga cccacggggg aacgggtcta atcgtgtagg aataaagctg
 3481 tattccagtg cttccaaaaa aaaaaaaaaa aaaaaaaal

SEQ ID NO: 36 Human TAPBP Isoform 3 Amino Acid Sequence (NP_757346.2)

1 mkslslllav alglatavsa gpaviecwfv edasgkglak rpgalllrqg pgeppprpdl
 61 dpelylsvhv vltvlthtpa prvrlgqdal ldlsfaympp tseaasslap gpppfglewr
 121 rqlhlgkghll laaapglngq mpaaqegava faawdddepw gpwtgngtfw lptvqpfqeg
 20 181 tylatihlpy lggqvtlela vykppkvsml patlaraapg eappellclv shfypsggle
 241 vewelrggpg grsqkaegqr wlsalrhhsd gsvslsghlq pppvtteqhg aryacrihhp
 301 slpasgrsae vtlevaglsq psledsvglf lsaflllglf kalgwaaavyl stckdskkka
 361 e

SEQ ID NO: 37 Mouse TAPBP cDNA Variant 1 Sequence (NM_001025313.1, CDS region from position 211-1611)

1 caaggggggtg ggtgcagggc gaggagtggc tggagagcta cagacagggc ttcggcaact
 61 ctgttttgcc cctgtgggag gcgttcctcc aacgcgcagt cacgccctca cctgaccaga
 121 tcggggaatt tccagttoct caccacagcc ttagaaggaa aatgaaagtg aaaggggaag
 30 181 aaaagacttg gtaggagaga tcgtagcacc atgaagcctc tgctcctgct cgtcgtgtgtg
 241 gcactggggc tagcgaccgt cgtctccgtc gtctcggctg gaccagaggc gatcgagtgc
 301 tgggttcgtg aggatgcagg tgggggtggc ctgtctaaga aacctgccac actgctactg
 361 cgccatggac ccaggggacc gccgccccgg ccagatcttg acccaaagct atacttcaag
 421 gtggatgacc cggcggaat gctcctggcc gccttcaggc ggtaccccg aggcgcctcc
 35 481 gccccacact gcgagatgag ccgcttcac cgttccttg cctcggcgaa gtgggctaga
 541 agtctgagtc cggagcagaa ctgcccgcgg gccctggacg gggattggct gctggctcagc
 601 gtatccagca ctctcttcag cctctccagc ctgctgcgac cacagccgga gcctctgcgg
 661 gagcctgtcg tcatcaccat ggcaacagtg gtgctgaccg tcctcaccca taaccctgcc
 721 cctcgagtcc agctgggaaa ggatgcagtg ctggacctgc gcttcgccta cgcaccctcc
 40 781 gccctggaag gttctccctc tctggacgca ggccctcctc cctttgggct ggagtggcga
 841 cgccagcaca ggggaaaggg tcacctgctg ttggctgcca ccccggggct ggccgggaga
 901 atgccaccag cccaggaaaa ggctacggca tttgcagctt gggatgacga tgagccctgg
 961 ggcccggtga ctgggaatgg gaccttctgg cttccagccg tgaagccttc tcaggagggg
 1021 gtctacctgg ctacggtaca cctgccctac ctgcaaggac aggtctccct ggagctgact
 45 1081 gtgcacaagg cccacagagt gtctctaaca ccagcaccgg ttgtgtgggc tgccccagga
 1141 gagcaccgcc cagaactgct ctgtcttgca tccacttct tccctgcgga gggctctggag
 1201 gtcaagtggg agctcagagg cggcccagga ggaagtctca gaaaggtga ggggaagacg
 1261 tggtcttcca ccatccgcca ccattccgat ggctctgtca gccagtctgg gcacctgcag
 1321 ctacctccag tcaactgcaa gcagcatgga gttcactatg tctgtcgggt gtaccactct
 50 1381 agcctgccag catcggggcg cagtgtgac gtcaccctgg aggtggcagg cttctcaggg
 1441 cctccatcg aggacggcat cggcctgttc ctgtctgctt ttctctctct cggactctct
 1501 aagggtgctag gctggctggg agctgcctac tggaccattc ctgaagtctc aaaggagaag
 1561 gccacagctg ccagcctgac cattcccagg aactcaaaga agtcacagta aagaagttct
 1621 cgtctgtgga agccaccctc atctctggcc cagatgactc cagtagcccc tgccccagaa
 55 1681 caacagcctt cttctcttct tctccctcca ttgaatagct cggatttttt tttttaaggt
 1741 ttttttgttt gtttcaatct ttcttgtttt gttcttgttt ttagacaagg tctctctatg
 1801 tagccttggc tggctgtcct gaacctcac caagccaacc aggtggcctc cgaagtccca
 1861 gagatccacc tgtccctctg cctcggagtg acaaacggca ctggacacca tgcctagagg
 1921 tctttttttt ttttggtttt tggagacagg atttctctgt atagtcccg ctgtcctgga
 60 1981 actcactctg tagaccaggc tggcctcgaa ctcagaaatc cgctgcctc tgcctcccag

2041 gtgctgggat taaaggcgtg cgccaccacc gctggactcg aactaaactt ttaagatgag
 2101 aaaaaagcga gacttagcgc cacactcatt tgatgcccac agtcgggagg cagaggcagt
 2161 ggatctctga catttccatg ccagcctgat ctacaaatag ggagtttcag gtcagtcggg
 2221 gttacttggc ggggctttgt tgagagaact aaatataaga acacacatgc gtagagtccc
 5 2281 aggggttgaga agaaccagtg tatgcgccag agaggcgggt tcataacctg tggctcacia
 2341 ccccttttaa gtccatcgga aataccagac attaggattc acaacagtag taaaggcagt
 2401 tacgaagtag caaagaaaat aattttacaa cttgaactgc gttaatgtta agaagggtga
 2461 gaaccactgg ccaggggatt aggacgaaag gtcttcatag catcttaaga ttctagagac
 2521 ctcagggaca gccacagca gtagtctaga gacggaaagg tccatgtgtc tccctcgcgg
 10 2581 actaaagcct acctctccag ctctccattc ctaccacagc ttttcttttg cctgggtccac
 2641 ctttaccatc tccgggcctg gtgggtctcg gccctttacg aatgtccccc tgtgggtcgt
 2701 cctctcctgt tcgcccgtcc tcgtgagcta ataaagtctt tccgagtttc

SEQ ID NO: 38 Mouse TAPBP Isoform 1 Amino Acid Sequence

15 **(NP_001020484.1)**

1 mkpllllvav alglatvvsv vsagpeaiec wfvedagggg lskkpatlll rhgprgpppr
 61 pdldpklyfk vddpagmlla afrrypagas aphcemsrfi pfpasakwar slspeqncpr
 121 aldgdwllvs vsstlflslss llrpqpeplr epvvitmatv vltvlthnpa prvqlgkdav
 181 ldlrfayaps alegspsllda gpppfglewr rghrgkghll laatpglagr mppaqekata
 20 241 faawdddepw gpwtgngtfw lpavkpsgeg vylatvhlpy lggqvslelt vkhaprvslt
 301 papvvwaapg eappelclla shffpaegle vkwelrggpg gssrkvegkt wlstirhhsd
 361 gsvsqsglhl lppvtakqhg vhyvcrvyhs slpasgrsad vtlevagfsg psiedgiglf
 421 lsafllllgll kvlgwlvaay wtipevskek ataasltipr nskksq

25 **SEQ ID NO: 39 Mouse TAPBP cDNA Variant 2 Sequence (NM_009318.2, CDS region from position 211-1608)**

1 caaggggggtg ggtgcagggc gaggagtggc tggagagcta cagacagggt ttcggcaact
 61 ctgtttggcc cctgtgggag gcgttcctcc aacgcgcagt cacgccctca cctgaccaga
 121 tcggggaatt tccagttcct caccacagcc ttagaaggaa aatgaaagtg aaaggggaag
 30 181 aaaagacttg gtaggagaga tcgtagcacc atgaagcctc tgctcctgct cgtecgctgtg
 241 gcaactgggccc tagcgaccgt cgtctccgtc gtctcggctg gaccagaggc gatcgagtgc
 301 tgggttcgtgg aggatgcagg tgggggtggc ctgtctaaga aacctgccac actgctactg
 361 cgccatggac ccaggggacc gccgccccgg ccagatcttg acccaaagct atacttcaag
 421 gtggatgacc cggcgggaat gctcctggcc gccttcaggc ggtaccccg caggcctcc
 35 481 gccccacact gcgagatgag ccgcttcctc ccgttccttg cctcggcgaa gtgggctaga
 541 agtctgagtc cggagcagaa ctgcccgcgg gccctggacg gggattggct gctgtgcagc
 601 gtatccagca ctctcttcag cctctccagc ctgctgcgac cacagccgga gcctctgcgg
 661 gagcctgtcg tcatcaccat ggcaacagtg gtgctgaccg tcctcaccca taaccctgcc
 721 cctcgagtcc agctgggaaa ggatgcagtg ctggacctgc gcttcgccta cgcaccctcc
 40 781 gccctggaag gttctccctc tctggacgca ggccctcctc cctttgggct ggagtggcga
 841 cgccagcaca ggggaaaggg tcacctgctg ttggctgcca cccccgggct ggccgggaga
 901 atgccaccag cccaggaaaa ggctacggca tttgcagctt gggatgacga tgagccctgg
 961 ggcccggtga ctgggaatgg gaccttctgg cttccagccg tgaagccttc tcaggagggt
 1021 gtctacctgg ctacggtaca cctgccctac ctgcaaggac aggtctccct ggagctgact
 45 1081 gtgcacaagg cccccagagt gtctctaaca ccagacccg ttgtgtgggc tgccccagga
 1141 gaggcacccc cagaactgct ctgtcttgca tcccacttct tccctgcgga gggctctggag
 1201 gtcaagtggg agctcagagg cggcccagga ggaagtctta gaaaggttga ggggaagacg
 1261 tggtcttcca ccatccgcca ccattccgat ggctctgtca gccagtctgg gcacctgcag
 1321 ctacctccag tcaactgcaa gcagcatgga gttcactatg tctgtcgggt gtaccactct
 50 1381 agcctgccag catcggggag cagtgtgac gtcaccctgg aggtggcagg cttctcaggg
 1441 cctccatcg aggacggcat cggcctgttc ctgtctgctt ttctctcct cggactcctc
 1501 aagggtgctg gctggctggc tgcctactgg accattcctg aagtctcaaa ggagaaggcc
 1561 acagtggcca gcctgaccat tcccaggaac tcaaaagaag cacagtaaa aagtctcgt
 1621 ctgtggaagc caccctcatc tctggccag atgactccag tagccctgc cccagaacaa
 55 1681 cagccttctt ctcttcttct cctccattg aatagctcgg attttttttt ttaaggtttt
 1741 tttgtttgtt tcaatctttc ttgttttgtt cttgttttta gacaaggctc ctctatgtag
 1801 ccttggtctg ctgtcctgaa cctcaccaa gccaaaccag ctggcctcga agtcccagag
 1861 atccacctgt cctctgcct cggagtgaca aacggcactg gacaccatgc cttagaggtct
 1921 tttttttttt tggtttttgg agacaggatt tctctgtata gtcccggtct tcttggaaact
 60 1981 cactctgtag accaggctgg cctcgaactc agaaatccgc ctgcctctgc ctcccagggt

2041 ctgggattaa aggcgtgcgc caccaccgct ggactcgaac taaactttta agatgagaaa
 2101 aaagcgagac ttagcgccac actcatttga tgcccacagt cgggaggcag aggcagtgga
 2161 tctctgacat ttccatgcca gcctgatcta caaataggga gtttcaggctc agtcgggggtt
 2221 acttggcggg gctttgttga gagaactaaa tataagaaca cacatgcgta ggtcccagg
 5 2281 gttgagaaga accagtgtat gcgccagaga ggccgggttca taacctgtgg ctcacaaccc
 2341 cctttaagtc catcggaat accagacatt aggattcaca acagtagtaa aggcagttac
 2401 gaagtagcaa agaaaataat tttaacaactt gaactgcgtt aatgttaaga aggttgagaa
 2461 ccactggcca ggggattagg acgaaaggtc ttcatacat cttaagattc tagagacctc
 2521 agggacagcc cacagcagta gtctagagac ggaaagggtcc atgtgtctcc ctgcgggact
 10 2581 aaagcctacc tctccagctc tccattccta ccacagcttt tcttttgcct ggtccacctt
 2641 taccatctcc gggcctgggtg ggtctcggcc ctttacgaat gtccccgtgt ggttcgtcct
 2701 ctctgttctg ccggtcctcg tgagctaata aagtctttcc gaggtttc

SEQ ID NO: 40 Mouse TAPBP Isoform 2 Amino Acid Sequence (NP_033344.1)

15 1 mkpllllvav alglatvvsv vsagpeaiec wfvedagggg lskkpatlll rhgprgpppr
 61 pdldpklyfk vddpagmlla afrrypagas aphcemsrfi pfpasakwar slspeqncpr
 121 aldgdwllvs vsstlflslss llrpqpeplr epvvitmatv vltvlthnpa prvqlgkdav
 181 ldlrfayaps alegspsllda gpppfglewr rqrhgkghll laatpglagr mppaqekata
 241 faawdddepw gpwtgngtfw lpavkpsqeg vylatvhlpy lggqvslelt vkhaprsvlt
 20 301 papvvwaapg eappelclla shffpaegle vkwelrggpg gssrkvegkt wlstirhhsd
 361 gsvsqsghlq lppvtakqhg vhyvcrvyhs slpasgrsad vtlevagfsg psiedgiglf
 421 lsafllllgll kvlgwlaayw tipevskeka taasltiprn skksq

SEQ ID NO: 41 Human ERAP1 cDNA Variant 1 Sequence (NM_016442.4, CDS region from position 348-3194)

25 1 acggatccgc gttcagaaaag gcgtgcactt cctacgcctg atcccccgca tcgcaacctc
 61 gcagcttccc cggcgtgcag cgctcattta ccaattccct tcttgaggat tgccggttcc
 121 ctgcctcggc cccactcccg tttacccttt cccagctcc cgcttagacc aggggcttcc
 181 ccgcctgccc ctagggctcg ggccgaagcg ccgctcagcg ccagctgccc gctccccggg
 30 241 ctccactttc actttcgggtc ctgggggagc taggccggcg gcagtggtgg tggcgcgcg
 301 gcaagggtga gggcgggccc agaaccacag gtaggtagag caagaagatg gtgtttctgc
 361 cccctcaaatg gtcccttgca accatgtcat ttctactttc ctactgttg gctctcttaa
 421 ctgtgtccac tccctcatgg tgtcagagca ctgaagcatc tccaaaacgt agtgatggga
 481 caccatttcc ttggaataaa atacgacttc ctgagtacgt catcccagtt cattatgatc
 35 541 tcttgatcca tgcaaacctt accacgctga ctttctgggg aaccacgaaa gtagaaatca
 601 cagccagtca gccaccagc accatcatcc tgcatagtca ccacctgcag atatctaggg
 661 ccacctcag gaaggagct ggagagaggc tatcggaaga acccctgcag gtcctggaac
 721 acccccgta ggagcaaatt gcaactgctg cttccgagcc cctccttgct gggctcccgt
 781 acacagttgt cattcactat gctggcaatc tttcggagac tttccacgga ttttcaaaaa
 40 841 gcacctacag aaccaaggaa ggggaactga ggatactagc atcaacacaa tttgaaccca
 901 ctgcagctag aatggccttt ccctgctttg atgaacctgc cttcaaagca agtttctcaa
 961 tcaaaattag aagagagcca aggcacctag ccatctccaa tatgccattg gtgaaatctg
 1021 tgactgttgc tgaaggactc atagaagacc attttgatgt cactgtgaag atgagcacct
 1081 atctggtggc cttcatcatt tcagattttg agtctgtcag caagataacc aagagtggag
 45 1141 tcaaggtttc tgtttatgct gtgccagaca agataaatca agcagattat gcactggatg
 1201 ctgcggtgac tcttctagaa ttttatgagg attatttcag cataccgtat cccctaccga
 1261 aacaagatct tgcgtctatt cccgactttc agtctggtgc tatggaaaac tggggactga
 1321 caacatatag agaactctgct ctggtgtttg atgcagaaaa gtcttctgca tcaagtaagc
 1381 ttggcatcac aatgactgtg gcccatgaac tggctcacca gtggtttggg aacctggtca
 50 1441 ctatggaatg gtggaatgat ctttggtctaa atgaaggatt tgccaaattt atggagtgtg
 1501 tgtctgtcag tgtgacccat cctgaactga aagttggaga ttatttcttt ggcaaagtgt
 1561 ttgacgcaat ggaggtagat gcttttaaatt cctcacaccc tgtgtctaca cctgtggaaa
 1621 atcctgctca gatccgggag atgtttgatg atgtttctta tgataaggga gcttgtattc
 1681 tgaatatgct aaggaggtat cttagtgtcg acgcatttaa aagtggattt gtacagtatc
 55 1741 tccagaagca tagctataaa aatacaaaaa acgaggacct gtgggatagt atggcaagta
 1801 tttgccctac agatggtgta aaagggtatg atggcttttg ctctagaagt caacattcat
 1861 cttcatcctc acattggcat cagggaagggg tggatgtgaa aacctgatg aacacttgga
 1921 cactgcagaa gggttttccc ctaataacca tcacagttag ggggaggaa gtacacatga
 1981 agcaagagca ctacatgaag ggctctgacg gcgccccgga cactgggtac ctgtggcatg
 60 2041 ttccattgac attcatcacc agcaaatccg acatggtcca tcgatttttg ctaaaaaaca

2101 aaacagatgt gctcatcctc ccagaagagg tggaatggat caaatttaat gtgggcatga
 2161 atggctatta cattgtgcat tacgaggatg atggatggga ctctttgact ggccttttaa
 2221 aaggaacaca cacagcagtc agcagtaatg atcgggcgag tctcattaac aatgcatttc
 2281 agctcgtcag cattgggaag ctgtccattg aaaaggcctt ggatttatcc ctgtacttga
 5 2341 aacatgaaac tgaaattatg cccgtgtttc aagggttgaa tgagctgatt cctatgtata
 2401 agttaatgga gaaaagagat atgaatgaag tggaaactca attcaaggcc ttcctcatca
 2461 ggctgctaag ggacctcatt gataagcaga catggacaga cgagggctca gtctcagagc
 2521 gaatgctgag gagtcaacta ctactcctcg cctgtgtgca caactatcag ccgtgcgtac
 2581 agagggcaga aggctatttc agaaagtggg aggaatccaa tggaaacttg agcctgcctg
 10 2641 tgcagctgac cttggcagtg tttgctgtgg gggcccagag cacagaaggc tgggattttc
 2701 tttatagtaa atatcagttt tctttgtcca gtactgagaa aagccaaatt gaatttgccc
 2761 tctgcagaac ccaaaataag gaaaagcttc aatggctact agatgaaagc ttaaggagg
 2821 ataaaaataa aactcaggag tttccacaaa ttcttacact cattggcagg aaccagtag
 2881 gatacccaact ggcttgcaa tttctgagga aaaactggaa caaacttgta caaaagtttg
 15 2941 aacttggtc atcttcata gccacatgg taatgggtac aacaaatcaa ttctccacaa
 3001 gaacacggct tgaagaggta aaaggattct tcagctcttt gaaagaaaat ggttctcagc
 3061 tccgttgtgt ccaacagaca attgaaacca ttgaagaaaa catcggttgg atggataaga
 3121 attttgataa aatcagagtg tggctgcaaa gtgaaaagct tgaacatgat cctgaagctg
 3181 acgcaacagg atgaaaatcc atcagaatct cagactacag cactaaatat gctttgatgc
 20 3241 tacatcaaac ggaatggaag catagctgac ttcgctaaag ttacttcacg tccatctagc
 3301 aatgaggca ctgttctcaa ccaaaggaga tggggatctg gtttagggca atccctttat
 3361 aatttgatgt gctgtggtct ccttggtaat gtataatttg gtattgcaca ggtgattagt
 3421 caaggaagtc tggaaaagct ttggtcccac agccttgctt cacagcatgt aaataattaa
 3481 aacaatattg atgctgaggt tcttctactg ctagtatgaa agtgacaaat ttttactggt
 25 3541 gtgaattggg aagaaaacaa tgctattcca tgacgtttgt aaaatgtttg taaaagctca
 3601 aacatgacga ttccataaaa taaacttgag gttaaataat gggtagtaaa ttatagaatg
 3661 tataagaaaa aatataaagg agaaaatcaa ttatcaggaa agctaaagaa cttttcaaat
 3721 ctagtaattt gaatatagac acaatgcact ttattgcact ttcaattctt ataaagcaac
 3781 aataatatta aggtccttga ctatgtgtac aatgttttca catatatagt ttcattttaat
 30 3841 catttcaaag ttaatctctg ccatctcgct aaatcatcag tctcggtctt tctgaaatag
 3901 aagggtgctg atcttctctaa taattctgcc tattttcatt tgctttaaac aggcgcctta
 3961 ttttctttct agttgtggct gcgcaaaaac atttatctcc caaataagat gtgctgctta
 4021 ccgaggtatc acgggggtggg gctccagctt gggtcgttga agctggggtt tgggaaacca
 4081 cttcagagat ggcagcagca agtttagcat cttcaaattt cttttattga aaaaaatttt
 35 4141 attagtaaca tgttgtatat aaaattatga gcacaatgcc atcacttaac tataactctt
 4201 aaagatagct taatgactgt ttattctctt gaccaaatag actcataata acataaatt
 4261 ttaaaagaaa tttaaattct ttcttctcta ttgtattatt ttatacaatt tgctatttct
 4321 atttcttct catattgatt attctaaata ctatgcaata atataactta gagttccacg
 4381 gtttgttttac acatttctctg ttgtacattt aggttattca aagttttcag ctcttttaaa
 40 4441 attgctctga ataagttcta gtgagtgaat tatggtgctg gctatatatt gctaaactgc
 4501 cctctcaaat gttgctagga attcatactg cgaaaagcaa tgaataagca tgctgtttt
 4561 cccatggcct tgcttgccag aatttgactt ttattatgat aatcagtgtg aaatgatata
 4621 ctactattgc ttgtatattg tgggtatacgg tgtcagggtt cagggttttt tttcaacggt
 4681 aaatattcta gaaactttct gaaataattt ctgttttaaaa atattgaata tttgcttcat
 45 4741 ttcaataact cccttttgac aaaaaactt aggtataact gttgatgaaa aaccagaaaa
 4801 aagtcacagaa ctcttttggtg actccaacta tggatagctt attttgaaaa aggagaattg
 4861 caaattttac caaaagatgg agaaaagcac attaaaaaga taccaacatt cagaaattca
 4921 tttcagcatg ttattatttg aaattattta aactaattta gataactata agatacttat
 4981 tgtccattta taccctgtaa agccgtttta gaatgtaata ttttaggtaa tccaaaatgt
 50 5041 actaaattaa attcattttt agttatgaga aatctttgct tatatgacaa atgaaaagaa
 5101 taacaagttg tcaaatgaaa agaataacat tgaaacattt gtattgtctc ttcttaaaact
 5161 atcttattga cttattattt aagcctttta atactaagta tgaaacaacc tatggctcgg
 5221 aaatttgatg cgcaaagcta tatgtgcata tgttatttaa ttcacttaac gtcacacaaa
 5281 agcataaaaat aatgattttt cactctcttt aaaaatacta aatcatttat gtccatttct
 55 5341 caattttttt attgatctat gctttgagtt tgctttctca acattattgt attttccact
 5401 tattattact gtataacata tgctagtgtt tagttggatt aatcttacct aaaagtactg
 5461 aaaaatgctt tttagtactt tttcataatt tatacattta ttttccgaat gtatcattga
 5521 ataattttat tgagttataa aagtatctta ttgctattta ataaaaaatt aacacataaa
 5581 atgaaaaaaa aaa

SEQ ID NO: 42 Human ERAP1 Isoform a Amino Acid Sequence (NP_057526.3)

	1	mvflplkws	l	atmsfllssl	l	alltvstps	w	cqsteaspk	r	sdgtpfpwn	k	irlpeyv
	61	vhydlilian	l	ttltfwggt	k	veitasqpt	s	tiilhshhl	q	isratlrkg	a	gerlseep
5	121	qvlehprqeq	i	allapepll	v	glpytvvih	y	agnlsetfh	g	fykstyrk	e	gelrilast
	181	qfeptaarma	f	pcfdepafk	a	sfsikirre	p	rhlaismnp	l	vkstvae	l	iedhfdvtv
	241	kmstylvafi	i	sdfesvski	t	ksgvkvsy	a	vpdkinqad	y	aldaavtll	e	fyedyfsip
	301	yplpkqdlaa	i	pdfqsgame	n	wgltttyres	a	llfdaekss	a	sklgitmt	v	ahelahqwf
10	361	gnlvtmewwn	d	lwnegfak	f	mefvsvsvt	h	pelkvgydyf	f	gkcfamev	d	alnsshpvs
	421	tpvenpaqir	e	mfdvsvydk	g	acilnmlre	y	lsadafksg	i	vqylqkhsy	k	ntknedlwd
	481	smasicptdg	v	kgmdgfc	s	rshssssshw	h	qegvdkvm	m	ntwtlqkgf	p	lititvrgr
	541	nvhmkgqehym	k	sgdgapdtg	y	lwhvpltfi	t	sksdmvrhf	l	lktktvli	l	peevewikf
	601	nvgmngyyiv	h	yeddgwds	t	gllkgthta	v	ssndrasli	n	afqlvsig	k	lsiekaldl
	661	slylkhetei	m	pvfqqnl	e	ipmyklmekr	d	mnevetqfk	a	flirllrdl	i	dktwtdeg
15	721	svsermlrsq	l	lllacvhn	y	qpcvqraegy	f	rkkesngn	l	slpvdvta	v	favgagste
	781	gwdflyskyq	f	slssteksq	i	efalcrtqn	k	eklqwlde	s	fgdkiktq	e	fpqiltlig
	841	rnpvgypaw	q	flrkwnkl	v	qkfelgsss	i	ahmvmgttn	q	fstrtrlee	v	kgfssllke
	901	ngsqlrcvqq	t	ieteenig	w	mdknfdkir	v	wlqsekleh	d	peadatg		

20 **SEQ ID NO: 43 Human ERAP1 cDNA Variant 4 Sequence (NM_001349244.1, CDS region from position 344-3190)**

	1	acggatccgc	g	ttcagaaag	g	cgtgcactt	c	ctacgcctg	a	tcccccgca	t	cgaacctc
	61	gcagcttccc	c	ggcgtgcag	c	gctcattta	c	caattccct	t	cctgggagt	t	gcggcttcc
25	121	ctcgctcggc	c	ccactcccg	t	tacccttt	c	cccagctcc	c	gccttagcc	a	ggggcttcc
	181	cgcctgccc	c	tagggctcg	g	gccgaagcg	c	cgctcagcg	c	cagcctgcc	g	ctccccggg
	241	ctccactttc	a	ctttcggtc	c	tgggggagc	t	aggccggcg	g	cagtgggtg	t	ggcgggcggc
	301	gcaagggtga	g	ggcgggccc	a	gaaccccag	g	tagagcaag	a	agatgggtg	t	tctgcccct
	361	caaattgggtc	c	ttgcaacca	t	gtcattttc	a	cttttctca	c	tgttgggtc	t	cttaactgt
30	421	gtccactcct	t	catggtgtc	a	gagcactga	a	gcatctcca	a	aacgtagt	a	tgggacacc
	481	atttccttgg	a	ataaaatac	g	acttcctga	g	tacgtcatc	c	cagttcatt	a	tgatctctt
	541	gatccatgca	a	accttacca	c	gctgacctt	c	tgggggaacc	a	cgaagtag	a	aatcacagc
	601	cagtcagccc	a	ccagcacca	t	catcctgca	t	agtcaccac	c	tgcatat	c	tagggccac
	661	cctcaggaag	g	gagctggag	a	gaggtatc	g	gaagaaccc	c	tgcatgtcc	t	ggaaacacc
35	721	cgtcaggag	c	aaattgcac	t	gctggctcc	c	gagcccctc	c	ttgtcgggc	t	cccgtacac
	781	agttgtcatt	c	actatgctg	g	caatctttc	g	gagactttc	c	caggatttt	a	caaaagcac
	841	ctacagaacc	a	aggaagggg	a	actgaggat	a	tagcatca	a	acaaatttg	a	acccactgc
	901	agctagaatg	g	cccttccct	g	ctttgatga	a	ccctgcctt	a	agcaagtt	t	tcaatcaa
	961	aattagaaga	g	agccaaggc	a	ccatagccat	c	tccaatatg	c	cattgggtg	a	aatctgtgac
40	1021	tgttgctgaa	g	gactcatag	a	agaccat	t	tgtgtcact	g	tgaagatga	g	cacctatct
	1081	ggtggccttc	a	tcatttcag	a	ttttgagtc	t	gtcagcaag	a	ataaccaaga	g	tgaggatcaa
	1141	ggtttctgtt	t	atgctgtgc	c	agacaagat	a	aatcaagca	g	attatgcac	t	ggatgctgc
	1201	ggtgactctt	c	tagaatttt	a	tgaggatta	t	ttcagcata	c	cgatcccc	t	acccaaaca
	1261	agatcttgct	g	ctattccc	a	ctttcagtc	t	ggtgctatg	g	aaaactggg	g	actgacaac
45	1321	atatagagaa	t	ctgctctgt	t	gtttgatgc	a	gaaaagtct	t	ctgcatcaa	g	tgaagcttg
	1381	catcacaatg	a	ctgtggccc	a	tgaactggc	t	caccagtgg	t	ttgggaacc	t	gggtcactat
	1441	ggaatgggtg	a	atgatcttt	g	gctaaatga	a	ggatttgcc	a	aatttatgg	a	gtttgtgtc
	1501	tgtcagtgtg	a	cccatcctg	a	actgaaagt	t	ggagattat	t	ctttggca	a	atgttttga
	1561	cgcaatggag	g	tagatgctt	t	aaattcctc	a	caccctgtg	t	tacacctg	t	tgaaaatcc
	1621	tgtcagatc	c	gggagatgt	t	tgatgatgt	t	cttatgat	a	agggagctt	g	gtattctgaa
50	1681	tatgctaagg	g	agtatctta	g	tgtctgacg	c	atttaaaagt	g	gtattgtac	a	gtatctcca
	1741	gaagcatagc	t	ataaaaata	a	caaaaacga	g	gacctgtgg	g	atagtatgg	a	caagtatttg
	1801	ccctacagat	g	gtgtgaaaag	g	gatggatgg	c	ctttgtctc	t	agaagtcaac	a	attcatcttc
	1861	atcctcacat	t	ggcatcagg	a	aggggtgga	t	gtgaaaacc	a	tgatgaaca	c	ctggacact
	1921	gcagaagggt	t	ttcccctaa	t	aaccatcac	a	gtgaggggg	a	ggaatgtac	a	catgaagca
55	1981	agagcactac	a	tgaagggtc	c	tgacggcgc	c	ccggacact	g	gggtacctg	g	gcatgttcc
	2041	attgacattc	a	tcaccagca	a	atccgacat	g	gtccatcga	t	ttttgtctaa	a	acaaaaaac
	2101	agatgtgctc	a	tcctcccag	a	agaggtgga	a	tggatcaaa	t	ttaatgtgg	g	catgaatgg
	2161	ctattacatt	g	tcattacg	a	ggatgatgg	a	tgggactct	t	tgactggcc	t	tttaaaagg
	2221	aacacacaca	g	cagtcagca	g	taatgatcg	g	gcgagtctc	a	ttaacaatg	c	atttcagct
60	2281	cgtcagcatt	g	ggaagctgt	c	cattgaaaa	g	gccttggat	t	tatccctgt	a	cttgaaaca

2341 tgaactgaa attatgcccg tgtttcaagg tttgaatgag ctgattccta tgtataagtt
 2401 aatggagaaa agagatatga atgaagtggg aactcaattc aaggccttcc tcatcaggct
 2461 gctaagggac ctcatgtata agcagacatg gacagacgag ggctcagctc cagagcgaat
 2521 gctgcggagt caactactac tcctcgctcg tgtgcacaac tatcagccgt gcgtacagag
 5 2581 ggcagaaggc tatttcagaa agtggaaagg atccaatgga aacttgagcc tgccgtgtcg
 2641 cgtgaccttg gcagtgtttg ctgtgggggc ccagagcaca gaaggctggg attttcttta
 2701 tagtaaatat cagttttctt tgtccagtac tgagaaaagc caaattgaat ttgccctctg
 2761 cagaacccaa aataaggaaa agcttcaatg gctactagat gaaagcttta agggagataa
 2821 aataaaaact caggagtttc cacaaattct tacactcatt ggcaggaacc cagtaggata
 10 2881 cccactggcc tggcaatttc tgaggaaaaa ctggaacaaa cttgtacaaa agtttgaact
 2941 tggctcatct tccatagccc acatggtaat gggtaacaac aatcaattct ccacaagaac
 3001 acggcttgaa gaggtaaaag gattcttcag ctctttgaaa gaaaatgggt ctcagctccg
 3061 ttgtgtccaa cagacaattg aaaccattga agaaaacatc ggttgatggg ataagaattt
 3121 tgataaaatc agagtgtggc tgcaaaagtga aaagcttgaa catgatcctg aagctgacgc
 15 3181 aacaggatga aaatccatca gaatctcaga ctacagcact aaatatgctt tgatgtaca
 3241 tcaaacggaa tggaagcata gctgacttcg ctaaaagtac ttcactctcc tctagcaaat
 3301 gaggcactgt tctcaaccaa aggagatggg gatctgggtt agggcaatcc ctttataatt
 3361 tgatgtgctg tggctctcct ggtaattgat aatttggtat tgcacagggt attagtcaag
 3421 gaagtctgga aaagctttgg tcccacagcc ttgcctcaca gcatgtaaat aattaaaaca
 20 3481 atattgtatg tgaggttctt ctactgctag tatgaaagtg acaaattttt actggtgtga
 3541 attgggaaga aaacaatgct attccatgac gtttgtaaaa tgtttgtaaa agctcaaca
 3601 tgacgattcc ataaaaataa cttgagggtt aataatgggt agtaaattat agaattgata
 3661 agaaaaataa taaaggagaa aatcaattat caggaaaagc aaagaacttt tcaaacttag
 3721 taatttgaat atagacacaa tgcactttat tgcactttca attcttataa agcaacaata
 25 3781 atattaaggt ccttgactat gtgtacaatg ttttcacata tatagtttca tttaatcatt
 3841 tcaaagttaa tctctgccat ctgcctaaat catcagctct ggctcttctg aaatagaagg
 3901 tgccctgatct tccataataa tctgcctatt ttcatttgct taaacaggc gccctatttt
 3961 ctttctagtt gtggctgctc aaaaacattt atctcccaa taagatgtgc tgcttaccga
 4021 ggtatcacgg ggtggggctc cagcttgggt cggtgaagct ggggtttggg aaaccacttc
 30 4081 agagatggca gcagcaagtt tagcatcttc aaatttcttt tattgaaaaa aattttatta
 4141 gtaacatggt gtatataaaa ttatgagcac aatgccatca cttaactata actcttaaag
 4201 atagcttaat gactgtttat tctcttgacc aaatagactc ataataacat ataattttta
 4261 aagaaattta aattctttct tctctattgt attattttat acaatttgct atttctattt
 4321 ccttctcata ttgattattc taaatactat gcaataatat aacttagagt tccacgggtt
 35 4381 gtttacacat ttctgttgt acatttaggt tattcaaagt tttcagctct tttaaaattg
 4441 ctctgaataa gttctagtga gtgagttatg gtgctggcta tattttgcta aactgcctc
 4501 tcaaagtgtg ctaggaattc atactgcgaa aagcaatgaa taagcatgcc tgttttccca
 4561 tggccttgct tgccagaatt tgacttttat tatgataatc agtgtaaaat gatatactac
 4621 tattgcttgt atattgttgt atacggtgtc aggtttcagg gttttttttc aacgttaaat
 40 4681 attctagaaa ctttctgaaa taatttctgt ttaaaaaat tatgaatttg cttcatttca
 4741 aatactccct tttgacaaaa aaacttaggt ataactgttg atgaaaaacc agaaaaaggt
 4801 ccagaactct ttggtgactc caactatgga tagcttattt tgaaaaagga gaattgcaaa
 4861 ttttaccaaa agatggagaa aagcacatta aaaagatacc aacattcaga aattcatttc
 4921 agcatgttat tattggaaat tatttaaaact aatttagata actataagat acttattgtc
 45 4981 catttatacc ctgtaaagcc gttttagaat gtaaatattt aggtaatcca aaatgtacta
 5041 aattaaattc attttttagt atgagaaatc tttgcttata tgacaaatga aaagaataac
 5101 aagttgtcaa atgaaaagaa tgacattgaa acatttgtat tgtctcttct taaactatct
 5161 tattgactta ttatttaagc cttttaatac taagtatgaa acaacctatg gtctggaat
 5221 ttgtatcgca aagctatatg tgcataatgt atttaattca tctaagtcta caaaaagca
 50 5281 taaaataatg attttttact ctctttaaaa atactaaatc atttatgtcc atttctcaat
 5341 tttttcattg atctatgctt tgagtttgct ttctcaacat tattgtattt tccacttatt
 5401 attactgtat aacatatgct agtgtttagt tggattaatc ttacctaaaa gtactgaaaa
 5461 atgcttttta gtactttttc atattttata catttatttt ccgaatgtat cattgaataa
 5521 ttttattgag ttataaaagt atcttattgc tatttaataa aaaattaaca cataaaatga
 55 5581 aaaaaaaaa

SEQ ID NO: 44 Human ERAP1 cDNA Variant 2 Sequence (NM_001040458.2, CDS region from position 348-3173)

1 acggatccgc gttcagaaaag gctgtgactt cctacgcctg atcccccgca tgcgaacctc
 60 61 gcagcttccc cggcgtgcag cgctcattta ccaattccct tcctgggagt tgcggcttcc

121 ctcgctcggc cccactcccc tttacccttt cccagctcc cgccttagcc aggggcttcc
 181 cgcctgccc ctagggctcg ggccgaagcg ccgctcagcg ccagcctgcc gctccccggg
 241 ctccactttc actttcggtc ctgggggagc taggccggcg gcagtgggtg tggcgggcgg
 301 gcaagggtga gggcgccccc agaaccocag gtaggtagag caagaagatg gtgtttctgc
 5 361 ccctcaaagt gtcccttgca accatgtcat ttctactttc ctactgttg gctctcttaa
 421 ctgtgtccac tccttcatgg tgtcagagca ctgaagcatc tccaaaacgt agtgatggga
 481 caccatttcc ttggaataaa atacgacttc ctgagtacgt catcccagtt cattatgatc
 541 tcttgatcca tgcaaaccctt accacgctga ctttctgggg aaccacgaaa gtagaaatca
 601 cagccagtca gccaccagc accatcatcc tgcatagtca ccacctgcag atatctaggg
 10 661 ccaccctcag gaaggagct ggagagagcg tatcggaaga acccctgcag gtcctggaac
 721 acccccgta ggagcaaatt gactgctgg ctcggagcc ctccttgctc gggctcccgt
 781 acacagttgt cattcactat gctggcaatc tttcggagac tttccacgga ttttacaaaa
 841 gcacctacag aaccaaggaa ggggaactga ggatactagc atcaacacaa tttgaacca
 901 ctgcagctag aatggccttt ccctgctttg atgaacctgc cttcaaagca agtttctcaa
 15 961 tcaaaattag aagagagcca aggcacctag ccatctccaa tatgccattg gtgaaatctg
 1021 tgactgttgc tgaaggactc atagaagacc attttgatgt cactgtgaag atgagcacct
 1081 atctggtggc cttcatcatt tcagattttg agtctgtcag caagataacc aagagtggag
 1141 tcaagggttc tgtttatgct gtgccagaca agataaatca agcagattat gcactgtagt
 1201 ctgcggtgac tcttctagaa ttttatgagg attatttcag cataccgtat cccctaccca
 20 1261 aacagatct tgcgtctatt cccgactttc agtctggtgc tatggaaaac tggggactga
 1321 caacatatag agaactctgct ctgttgtttg atgcagaaaa gtcttctgca tcaagtaagc
 1381 ttggcatcac aatgactgtg gcccatgaac tggctcacca gtggtttggg aacctggtca
 1441 ctatggaatg gtggaatgat ctttggctaa atgaaggatt tgccaaattt atggagtgtg
 1501 tgtctgtcag tgtgacccat cctgaactga aagttggaga ttatttcttt ggcaaagtgt
 25 1561 ttgacgcaat ggaggtagat gctttaaatt cctcacaccc tgtgtctaca cctgtggaaa
 1621 atcctgctca gatccgggag atgtttgatg atgtttctta tgataaggga gcttgtattc
 1681 tgaatatgct aaggagatc cttagtctg acgcatttaa aagtgggtatt gtacagtatc
 1741 tccagaagca tagctataaa aatacaaaaa acgaggacct gtgggatagt atggcaagta
 1801 tttgccctac agatgggtgta aaagggatgg atggcttttg ctctagaagt caacattcat
 30 1861 cttcatcctc acattggcat caggaagggg tggatgtgaa aacctgatg aaccttggga
 1921 cactgcagaa gggttttccc ctaataacca tcacagttag ggggaggaat gtacacatga
 1981 agcaagagca ctacatgaag ggctctgacg gcgccccgga cactgggtac ctgtggcatg
 2041 ttccattgac attcatcacc agcaaatccg acatgggtcca tcgatttttg ctaaaaacaa
 2101 aaacagatgt gctcatcctc ccagaagagg tggaaatgat caaatttaat gtgggcatga
 35 2161 atggctatta cattgtgcat tacgaggatg atggatggga ctctttgact ggccttttaa
 2221 aaggaacaca cacagcagtc agcagtaatg atcgggcgag tctcattaac aatgcatttc
 2281 agctcgtag cattgggaag ctgtccattg aaaaggcctt ggatttatcc ctgtacttga
 2341 aacatgaaac tgaaattatg cccgtgtttc aaggtttgaa tgagctgatt cctatgtata
 2401 agttaatgga gaaaagagat atgaatgaag tggaaactca attcaaggcc ttcctcatca
 40 2461 ggctgctaag ggacctcatt gataagcaga catggacaga cgagggtcga gtctcagagc
 2521 gaatgctgag gagtcaacta ctactcctcg cctgtgtgca caactatcag ccgtgctgac
 2581 agagggcaga aggtatattc agaaagtgga aggaatccaa tggaaacttg agcctgctg
 2641 tcgacgtgac cttggcagtg tttgctgtgg gggcccagag cacagaaggc tgggattttc
 2701 tttatagtaa atatcagttt tctttgtcca gtactgagaa aagccaaatt gaatttgccc
 45 2761 tctgcagAAC ccaaaataag gaaaagcttc aatggctact agatgaaagc ttaaaggag
 2821 ataaaataaa aactcaggag tttccacaaa ttcttactac cattggcagg aaccagtag
 2881 gataccact ggcttgcaa tttctgagga aaaactggaa caaacttgta caaaagtgtg
 2941 aacttggtc atcttccata gccacatgg taatgggtac aacaaatcaa ttctccacaa
 3001 gaacacggct tgaagaggta aaaggattct tcagctcttt gaaagaaaat ggttctcagc
 50 3061 tccgttgtgt ccaacagaca attgaaacca ttgaagaaaa catcggttgg atggataaga
 3121 attttgataa aatcagagtgt tggctgcaaa gtgaaaagct tgaacgtatg taaaaattcc
 3181 tcccttgcca ggttcctgtt atctctaata accaacattt tgttgagtgt attttcaaac
 3241 tagataggc tgttttggtt ccaactggag atactttttt cccttcaact catttttga
 3301 ctatccctgt gaaaagaata gctgttagtt tttcatgaat gggctatcgc taccatgtgt
 55 3361 tttgttcata acaggtgttg ccctgcaacg taaacccaag tgttggttc cctgccacag
 3421 aagaataaag taccttattc ttctcatttt atagtttatg cttaagcacc cgtgtccaaa
 3481 accctgtacc ccatgtttat cattcataaa ctgtttcatc agtctcctcg aaagactctg
 3541 aatagtgcac tactgaacaa tgaacacctg gatctgagac taagccggac gatgactggg
 3601 ttaaagctct cccggtcac ccctccagac ccgctgcca tccctcttcc ttgctccatg
 60 3661 cccaggggct gacttgtaaa ggccaagtca tcaagctttc ttgcccttg gatgttggtc
 3721 agtggggagc cggagagctg gagctggggt cggaggaggt agtaggtgga ggtgttcttc

3781 cctgattccc ttgcgggatg cctcgggctg gcctcccctg agggctcttag ctccgagagg
 3841 ggacctcttt ttccacacag ccttctccac ctctggattt tggtaactgc tccctcctca
 3901 tcccttcagg attagtggcc tcagtgggag tctggctttt actagtcctg gcggacttgt
 3961 ggttttctaca taatgtgctc gcacttttgc aaaaaatctt tttatagaac cctcctcaga
 5 4021 taattctgag tgagtgtcat ctatttccct gactgggtaca gtatctcttc tgaaaaagca
 4081 gagtgcattc aagtctgtag gaaaaccctt ttcttaggga ggtgattttt tttctctctc
 4141 tgcttcttat ttggcctact ttacaatttc taactaacta gttattggca tttactgaca
 4201 gtaaattatt gcagtcacca ataaatgata gtacattgtg aaacaaaata tttgtcata
 4261 ttagcaataa ggacattctt tggctttgaa gtctttcttt cttttgtgaa gacttcacac
 10 4321 acggttgctt cagcacacag ttgctgctca ggttttatgt atagatgata ataatagaaa
 4381 gcacagttta ctaacatggt aaaccaacgg agttcaagtc aagtcagtta ataccctaag
 4441 aattagattt tatttcttat tctgaaaact tgctacacag ggacttatct aaccatagt
 4501 gtgctctgtt gctgacttga ttcaagttgc agcgtgtttt gcgctgactc taagggtcgg
 4561 aaatcctcac acctggcaaa ggagaattca aactgaactt tttgaatata aggcaaaaac
 15 4621 ttcaagataa gggaatatga ttgatgattg gtacgaaaaa tgtcaaatg tgttccccta
 4681 atacacgaca aaatagagtg acttctggac ataaatctgc catttattaa accattact
 4741 acaacaaata aataggtata aaagtggaat tggaattttt atacttattt gttgtatga
 4801 atggtttaat aaaaatagaa atcactggtg atttccaccc caaactaaac tatttccctt
 4861 ctttttaaaa aatacacaac caagatttta atgtaaaata ttttgcttta attgtatttt
 20 4921 atgccttgat taatgaaaca tggaaatatt gatttttcagt tttggtcacc tgaggaacct
 4981 atctttgttt gcttttgtaa aagcccatth tctaaacaga tacaatattg ccacaacaat
 5041 gtgcagaaac ctttttgata ataaaaaatt gttctttgcc tctaaaaaaa aaaaaaaaaa
 5101 aaaaaaaaaa

25 **SEQ ID NO: 45 Human ERAP1 Isoform b Amino Acid Sequence**
(NP_001035548.1)

1 mvflplkws1 atmsfllssl lalltvstps wcqsteaspk rsdgtfpfwn kirlpeyvip
 61 vhydllihan ltlftfwggt kveitasqpt stiihshhl qisratlrkg agerlseep1
 121 qvlehpqrqeq iallapepll vglpytvvfh yagnlsetfh gfykstyrk egerlilast
 30 181 qfeptaarma fpcfdepafk asfsikirre prhlaisnmp lvksvtvaeg liedhfdvtv
 241 kmstylvafi isdfesvski tkskvkvsy avpdkinqad yaldaavtll efyedyfsip
 301 yplpkqdlaa ipdfqsgame nwglttyres allfdaekss assklgitmt vahelahqwf
 361 gnlvtmewwn dlwlnegfak fmefvsvsvt hpelkvgdyf fgkcf damev dalnsshpvs
 421 tpvenpaqir emfddvsydk gacilnmlre ylsadafksg ivqylqkhsy kntknedlwd
 35 481 smasicptdg vkgmdgfcsr sqhssssshw hqegvdvktm mntwtlqkgf plititvrgr
 541 nvhmkqehym kgsdgapdtg ylwhvpltfi tsksdmvhrf llktktdvli lpeeewikf
 601 nvgmngyyiv hyeddgwds1 tgllkgthta vssndrasli nnaflvlsig klsiekaldl
 661 slylkhetei mpvfqglnel ipmyklmekr dmnevetqfk aflirllrld idkqtwtddeg
 721 svsermlrsq llllacvhny qpcvqraegy frkwkesngn ls1pvdvtla vfavgaqste
 40 781 gwdflyskyq fslssteksq iefalcrtqn keklqwlld sfkgdkiktq efpqiltlig
 841 rnpvgyp1aw qflrknwnkl vqkfelgsss iahvmvgttn qfstrtrlee vkgffss1ke
 901 ngsq1rcvqq tietieenig wmdknfdkir vwlqsekler m

45 **SEQ ID NO: 46 Human ERAP1 cDNA Variant 3 Sequence (NM_001198541.2,**
CDS region from position 106-2931)

1 ccttcccttc tgctgggctt caggtgagac tgcagctcca gtcatcatct tgattgcagc
 61 ccacagaaac taccaaccca cagaaactgt aggtagagca agaagatggt gtttctgccc
 121 ctcaaatggt cccttgcaac catgtcattt ctactttcct cactgttggc tctcttaact
 181 gtgtccactc cttcatggtg tcagagcact gaagcatctc caaaacgtag tgatgggaca
 50 241 ccatttccctt ggaataaaat acgacttcct gagtacgtca tcccagttca ttatgatctc
 301 ttgatccatg caaaccttac cacgctgacc ttctggggaa ccacgaaagt agaaatcaca
 361 gccagtcagc ccaccagcac catcatcctg catagtcacc acctgcagat atctagggcc
 421 accctcagga agggagctgg agagaggcta tcggaagaac ccctgcaggt cctggaacac
 481 ccccgctcagg agcaaatgac actgctggct cccgagcccc tccttgctgg gctcccgtac
 55 541 acagttgtca ttactatgc tggcaatctt tcggagactt tccacggatt ttacaaaagc
 601 acctacagaa ccaagggaagg ggaactgagg atactagcat caacacaatt tgaaccact
 661 gcagctagaa tggcctttcc ctgctttgat gaacctgcct tcaaagcaag tttctcaatc
 721 aaaattagaa gagagccaag gcacctagcc atctccaata tgccattggt gaaatctgtg
 781 actgttgctg aaggactcat agaagaccat tttgatgtca ctgtgaagat gagcacctat
 60 841 ctggtggcct tcatcatttc agattttgag tctgtcagca agataaccaa gaggaggagc

901 aagggtttctg tttatgctgt gccagacaag ataaatcaag cagattatgc actggatgct
 961 gcggtgactc ttctagaatt ttatgaggat tatttcagca taccgtatcc cctacccaaa
 1021 caagatcctt ctgctattcc cgactttcag tctggtgcta tggaaaactg gggactgaca
 1081 acatatagag aatctgctct gttgtttgat gcagaaaagt cttctgcatc aagtaagctt
 5 1141 ggcatcacia tgactgtggc ccatgaactg gctcaccagt ggtttgggaa cctggtcact
 1201 atggaatggg ggaatgatct ttggctaaat gaaggatttg ccaaatttat ggagtttgtg
 1261 tctgtcagtg tgacccatcc tgaactgaaa gttggagatt atttcttttg caaatgtttt
 1321 gacgcaatgg aggtagatgc tttaaattcc tcacaccctg tgtctacacc tgtggaaaat
 1381 cctgctcaga tccgggagat gtttgatgat gtttcttatg ataagggagc ttgtattctg
 10 1441 aatatgctaa gggagtatct tagtgctgac gcatttaaaa gtggtattgt acagtatctc
 1501 cagaagcata gctataaaaa tacaaaaaac gaggacctgt gggatagtat ggcaagtatt
 1561 tgccctacag atgggtgtaaa agggatggat ggcttttgct ctagaagtca acattcatct
 1621 tcatcctcac attggcatca ggaaggggtg gatgtgaaaa ccatgatgaa cacttgaca
 1681 ctgcagaagg gttttccctt aataaccatc acagtgaggg ggaggaatgt acacatgaag
 15 1741 caagagcact acatgaaggg ctctgacggc gccccggaca ctgggtacct gtggcatgtt
 1801 ccattgacat tcatcaccag caaatccgac atggtccatc gatttttgct aaaaaacaaa
 1861 acagatgtgc tcatcctccc agaagagggtg gaatggatca aatttaatgt gggcatgaat
 1921 ggctattaca ttgtgcatta cgaggatgat ggatgggact ctttgactgg ctttttaaaa
 1981 ggaacacaca cagcagtcag cagtaatgat cgggcgagtc tcattaacaa tgcatttcag
 20 2041 ctcgtcagca ttgggaagct gtccattgaa aaggccttgg atttatccct gtacttgaat
 2101 catgaaactg aaattatgcc cgtgtttcaa ggtttgaatg agctgattcc tatgtataag
 2161 ttaatggaga aaagagatat gaatgaagtg gaaactcaat tcaaggcctt cctcatcagg
 2221 ctgctaaggg acctcattga taagcagaca tggacagacg agggctcagt ctcagagcga
 2281 atgctgcgga gtcaactact actcctcgcc tgtgtgcaca actatcagcc gtgctacag
 25 2341 agggcagaag gctatttcag aaagtggaa gaaacttgag cctgcctgtc
 2401 gacgtgacct tggcagtggt tgctgtggg gcccagagca cagaaggctg ggattttctt
 2461 tatagtaaat atcagttttc tttgtccagt actgagaaaa gccaaattga atttgcctc
 2521 tgcagaacct aaaataagga aaagcttcaa tggctactag atgaaagctt taagggagat
 2581 aaaataaaaa ctccaggagt tccacaaatt cttacactca ttggcaggaa cccagtagga
 30 2641 taccactggg cctggcaatt tctgaggaaa aactggaaca aacttgtaca aaagtttgaa
 2701 cttggctcat cttccatagc ccacatggta atgggtacaa caaatcaatt ctccacaaga
 2761 acacggcctt aagaggtaaa aggattcttc agctctttga aagaaaatgg ttctcagctc
 2821 cgttgtgtcc aacagacaat tgaaaccatt gaagaaaaca tcggttggtg ggataagaat
 2881 tttgataaaa tcagagtgtg gctgcaaagt gaaaagcttg aacgtatgta aaaattcctc
 35 2941 ccttgccagg ttctgttat ctctaatac caacattttg ttgagtgtat tttcaacta
 3001 gattgggtg tttggctcc aactggagat acttttttcc cttcaactca ttttttgact
 3061 atccctgtga aaagaatagc tgttagtttt tcatgaatgg gctatcgcta ccatgtgttt
 3121 tgttcatcac aggtgttgcc ctgcaacgta aaccacagtg ttgggttccc tgccacagaa
 3181 gaataaagta cttatttctt ctcattttat agtttatgct taagcaccgg tgtccaaaac
 40 3241 cctgtacccc atgtttatca ttcataaact gtttcatcag tctcctcgaa agactctgaa
 3301 tagtcgacta ctgaacaatg aacacctgga tctgagacta agccggacga tgactgggtt
 3361 aaagctctcc cggctcacc ctccagaccc gctgcccac cctcttctt gctccatgcc
 3421 caggggctga cttgtaaagg ccaagtcac aagctttctt gccctttgga tgttggtcag
 3481 tggggagccg gagagctgga gctggggtcg gaggaggtag taggtggagg tgttcttccc
 45 3541 tgattccctt gcgggatgcc tccggctggc ctcccctgag ggtcttagct ccgagagggg
 3601 accctctttt ccacacagcc ttctccacct ctggattttg gtaactgctc cctcctcatc
 3661 ccttcaggat tagtggcctc agtgggagtc tggcttttac tagtctggc ggacttgtgg
 3721 tttctacata atgtgctcgc acttttgcaa aaaatctttt tatagaacct tcctcagata
 3781 attctgagtg agtgtcatct atttccctga ctggtacagt atctcttctg aaaaagcaga
 50 3841 gtgcattcaa gtctgtagga aaaccctttt cttagggagg tgattttttt tctctctctg
 3901 cttcttattt ggcctacttt acaatttcta actaactagt tattggcatt tactgacagt
 3961 aaattattgc agtcaccaat aatgatagt acattgtgaa acaaaaatatt tgctcatatt
 4021 agcaaatagg acattctttg gctttgaagt ctttctttct tttgtgaaga cttcacacac
 4081 ggttgcttca gcacacagtt gctgctcag ttttatgtat agatgataat aatagaaagc
 55 4141 acagtttact aacatggtaa accaacggag ttcaagtcaa gtcagttaat accctaagaa
 4201 ttagatttta tttcttattc tgaaaacttg ctacacaggg acttatctaa cccatagtgt
 4261 gctctgttgc tgacttgatt caagttgcag cgtgttttgc gctgactcta aggtgaggaa
 4321 atcctcacac ctggcaaagg agaattcaaa ctgaaactttt tgaatataag gcaaaaactt
 4381 caagataagg gaatatgatt gatgattggg acgaaaaatg tcaaaatgtg tttcccataat
 60 4441 acacgacaaa atagagtgc ttctggacat aaatctgcca tttattaaac cattcactac
 4501 aacaaataaa taggtataaa agtggaaattg gaatttttat acttatttgt tgtagtgaat

4561 ggtttaataa aaatagaaat cactggtaat ttccacccca aactaaacta tttcccttct
 4621 tttaaaaaaa tacacaacca agatttttaat gtaaaaatatt ttgctttaat tgtattttat
 4681 gccttgatta atgaaacatg gaaatatattga ttttcagttt tggtcacctg aggaacctat
 4741 ctttgtttgc ttttgaaaaa gcccattttc taaacagata caatattgcc acaacaatgt
 5 4801 gcagaaacct ttttgataat aaaaaattgt tctttgcctc taaaaaaaaa aaaaaaaaaa
 4861 aaaaaaaaaa

SEQ ID NO: 47 Mouse ERAP1 cDNA Sequence (NM_030711.4, CDS region from position 76-2868)

10 1 cagactggcg gctgtgatga aggtaatttg agcaagaaga tagtgtctct gctcctcaaa
 61 tagccacttg caagcatgcc ctctcttctt cccctagtat tgacattttt atctgtgtca
 121 tctccttctt ggtgtcagaa cagtgatata gaatctctaa aagctagtaa tggagactca
 181 ttccttggga ataatatgag acttctctgag tatatgacct cgattcatta tgatctcatg
 241 atccatgcaa acctcagcac tctgactttc tggggaaaaa cagaagtaga aatcatagct
 15 301 agccggccca ccagcaccat tattatgcat agtcaccacc tgcaaatact taaggccacc
 361 ctccagaagag gagctggaga gatgctgtct gaagaaccac tgaaggctct ggaatatcct
 421 gctcatgagc aagttgcact gctggctgcc cagccgcttc ttgctgggtc cctgtacaca
 481 gttatcatcg actatgctgc caacctgtct gagagtgttc atggattcta taagagcacc
 541 tacagaacac aggaaggatga aatgagaata cttgcagcaa cacagtgtga acccacagct
 20 601 gctagaatgg cttttccctg ctttgatgaa cctgccctca aggcaagttt ttccatcaag
 661 ataaagaggg atccaaggca cctggccatc tccaacatgc ccctggtgaa atctgtgaat
 721 gttgctgaag gactcataga agaccatttt gacatcactg tgaagatgag tacctacct
 781 gtggccttca tcatctctga ttttaagtct gtgagcaaga tgactaagag tggagtcaag
 841 gtttctgtgt atgctgtgcc agacaagata aatcaagccg attatgctct ggatgctgca
 25 901 gtgactcttc tagagtttta tgaggattat ttcaacattc catatccttt acccaagcaa
 961 gatcttgctg caatcccaga ctttcagtct ggtgctatgg aaaactgggg actgaccaca
 1021 tacagagagt cctctctgtt atacgataaa gaaaagtctt ctgcatctag taagcttggc
 1081 atcacaatga ttgtgtccca tgaactggct caccagtggg tcgggaacct ggtcaccatg
 1141 gaatgggtga atgacctttg gctcaatgaa ggatttgcca aatttatgga gtttgtgtct
 30 1201 gtcactgtga cccatcctga actaaaagtt gaagactatt tctttggcaa gtgttttaaa
 1261 gcaatggaag ttgatgcatt aaactcctct caccctgtat ccacacctgt ggagaatcct
 1321 ggcagatttc gggagatggt tgatgatggt tcttatgaaa agggagcttg tattctgaat
 1381 atgctaagag attatctgag tgcagataca tttaaaagag gaattgtaca gtatctccaa
 1441 aagtatagtt ataaaaaacac aaaaaacgag gacctgtgga atagcatgat gcataatttgc
 35 1501 cctacagatg gcacacaaac aatggatggc ttctgctcta gaagccaaca ctcatcatca
 1561 acctcacact ggcgtcagga ggtttagatg gtaaagacca tgatgaacac atggacactg
 1621 cagaagggtc ttctctctgat aaccatcact gtgagtgggc ggaatgtgca catgaagcaa
 1681 gaacactaca tgaagggttc agaacgcttc ccagagactg ggtatttgtg gcatgttcca
 1741 ctgacattca ttaccagcaa atcagactca gtccagagat ttttgctgaa gacaaagaca
 40 1801 gatgtgctca tctctccaga agcagtgcag tggatcaaata ttaacgtggg aatgaatggc
 1861 tattacattg tgcattacgc ggatgatgga tgggcttctc tgagtggcct tttaaaagaa
 1921 ggcacacaaa caatcagcag taatgatcgg gcaagtctca tcaacaatgc atttcagctg
 1981 gtcagcattg aaaagttatc aatagaaaaa gccctggatt taacctgtga cttgaaaaat
 2041 gaaactgaaa ttatgcccac attccaagct ttgaatgaac tcatacctat gtataagtta
 45 2101 atggagaaaa gagatatgat tgagggtgaa actcaattta aggacttctc tctcaagttg
 2161 ctgaaggacc ttattgataa gcagactgtg acagatgaag gctctgtctc atgagagtg
 2221 ctcaggagcc agctcctcct cctcgcattg gtgcgcaatt atcagccctg tgtgcagagg
 2281 gcagagcgct atttcagaga gtggaagtca tctaattgaa acatgagcat ccctattgat
 2341 gtgaccttgg ctgtgtttgc tgtcggggcc caaaacacag aaggctggga ttttctttat
 50 2401 agtaaatatc agtcttcttt gtctagtact gagaaaagcc aaattgaatt ttcctctgt
 2461 acaagcaaa atccagaaaa acttcagtgg cttctagatc aaagttttaa gggagagata
 2521 ataaaaactc aggaatttcc acataattctc aactcattg gcagaaaccc agtgggttat
 2581 ccattggcct ggaagtttct gagggaaaaat tggaaataagc ttgtacagaa atttgaactt
 2641 ggctcatcgt ccatagctca catggtaatg ggaacaacag accagttttc caccagagca
 55 2701 cgtcttgaag aggtaaaagg attcttcagt tccttgaaag aaaacgggtc tcagctccgt
 2761 tgtgttcaac aaaccattga gaccattgaa gaaaacatac gatggatgga taagaatttt
 2821 gataaaataa gactgtggct gcaaaaagaa aagccagagt tgctgtgaca atatgtttct
 2881 tgccaagttc cagaatttat gatgagcaaa attttgttca atttgagtat ttttaacta
 2941 aagatagttg ttttggttac tactgaagat attttcttat ttcatttcac ttgattatgt
 60 3001 ctgagaaaag cttcaatgac tgtgattgac aggttttaaac ccaagtgcct agctccttgc

3061 cacagaaaaa cccagggggc ttcttctcag tttatacttc agcttctgtg tcggccattc
 3121 atcagtat tttt ttgtaagact gaatagtgag gcataactgt gcagtgaatg cttggctcct
 3181 tagccatgct gggaagtggc tgtgcagcag ccctcctgtt gacctctctg gctctacagt
 3241 cctgcagtct ctcttcctgg gttcacttat ggggtgctgcc ttgcaggatc cacttaatca
 5 3301 gacacttttg ccctctggat tttgttaata gtgttttgta gttgttatcc atgagggatt
 3361 gcatagaggt gtttattttc ttgtacttgc ctctggctgg ctttcaacca ccgtaagtct
 3421 tactcctaag aggtggcctt tccccccccc cccccccgc actctgataa ctacttcctt
 3481 ttcattcctt ggggcatggg gcaacagtgg aaacctgctg ataccagtca tggtagtact
 3541 tgtcattttc actgaacctt cgtgcacatt tgtttttaaa aatattgtat atctataacc
 10 3601 ctcgttggat aagtcaaata tgagagccac ttactttctt tggaagggtta tttttaaaaa
 3661 tgcagagtgc attcagatct gtatcagaaa tatctcgttt taagatatac ttgtgagtca
 3721 ctgatccctt aagtagggag taggggttga aattcctaaa taactcataa tagcacttac
 3781 tgaaaataaa tcattgagga ctactaaca aaggacaata ttaataatat gtttatatta
 3841 gtataaggaa agattcttta aatcttaaaa agattttttt gatgtatggc ataaatgata
 15 3901 ggaaaaatat caacatatgg tgctgtaata tactatgaaa taaagcgatt ttaggggttc
 3961 aattatat ttcactggcct accaaaatac ttgagattaa atgaggaaag attt

SEQ ID NO: 48 Mouse ERAP1 Amino Acid Sequence (NP_109636.1)

1 mpsllplvlt flsvsspswc qnsdieslka sngdsfpwnn mrlpeymtpi hydmlmihanl
 20 61 stltfwgkte veiiasrpts tiimhshhlq iskatlrrga gemlseeplk vleyapaheqv
 121 allaaqplla gslytviidy aanlsesfhg fykstyrtqe gemrilaatq feptaarmaf
 181 pcfdepalka sfsikikrdp rhlaisnmp1 vksvnvaegl iedhfditvk mstylvafii
 241 sdfksvskmt ksgvkvsvya vpdkingady aldaavtllle fyedyfnipy plpkqdlaii
 301 pdfqsgamen wglttyress llydkekssa ssklgitmiv shelahqwfq nlvtmewwnd
 25 361 lwlnegfakf mefvsvtvth pelkvedyff gkcfnamevd alnsshpvst pvenpaqire
 421 mfdvsvyegk acilnmlrly lsadtfrkrgi vqylqkysyk ntknedlwns mmhicptdgt
 481 qtmddgfcrrs qhssstshwr qevvdvktmm ntwtlqkgfp lititvsgn vhmkgqehymk
 541 gserfpetgy lwhvpltfitt sksdsvqrfl lktktvdlil peavqwikfn vgmngyyivh
 601 yaddgwasls gllkeahhti ssndraslin nafqlvsiek lsiekaldlt lylkneteim
 30 661 pifqalneli pmyklmekrd mievetqfkd flklklkdli dkqtwtdqgs vsermlrsq1
 721 lllacvrnyq pcvgraeryf rewkssngnm sipidvtlav favgaqnteg wdflyskyqs
 781 slssteksqi efslctskdp eklqwlldqs fkgeiiktqe fphiltligr npvgyp1lawk
 841 flrenwnklv qkfelgsssi ahmvngtttdq fstararleev kgffss1lken gsqlrcvqq1
 901 ietieenirw mdknfdkir1 wlqkekpell

* Included in Table 1 are RNA nucleic acid molecules (e.g., thymines replaced with
 uridines), nucleic acid molecules encoding orthologs of the encoded proteins, as well as
 DNA or RNA nucleic acid sequences comprising a nucleic acid sequence having at least
 40 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%,
 95%, 96%, 97%, 98%, 99%, 99.5%, or more identity across their full length with the
 nucleic acid sequence of any SEQ ID NO listed in Table 1, or a portion thereof. Such
 nucleic acid molecules can have a function of the full-length nucleic acid as described
 further herein.

* Included in Table 1 are orthologs of the proteins, as well as polypeptide molecules
 comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%,
 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or more
 identity across their full length with an amino acid sequence of any SEQ ID NO listed in

Table 1, or a portion thereof. Such polypeptides can have a function of the full-length polypeptide as described further herein.

II. Subjects

5 In one embodiment, the subject for whom predicted likelihood of efficacy of an inhibitor of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), and an immunotherapy combination treatment is determined, is a mammal (*e.g.*, mouse, rat, primate, non-human mammal, domestic animal, such as a dog, 10 cat, cow, horse, and the like), and is preferably a human. In another embodiment, the subject is an animal model of cancer. For example, the animal model can be an orthotopic xenograft animal model of a human-derived cancer.

In another embodiment of the methods of the present invention, the subject has not undergone treatment, such as chemotherapy, radiation therapy, targeted therapy, and/or 15 immunotherapies. In still another embodiment, the subject has undergone treatment, such as chemotherapy, radiation therapy, targeted therapy, and/or immunotherapies.

In certain embodiments, the subject has had surgery to remove cancerous or precancerous tissue. In other embodiments, the cancerous tissue has not been removed, *e.g.*, the cancerous tissue may be located in an inoperable region of the body, such as in a 20 tissue that is essential for life, or in a region where a surgical procedure would cause considerable risk of harm to the patient.

The methods of the present invention can be used to determine the responsiveness to inhibitor(s) of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein and immunotherapy combination 25 treatment of many different cancers in subjects such as those described herein.

III. Sample Collection, Preparation and Separation

In some embodiments, biomarker amount and/or activity measurement(s) in a sample from a subject is compared to a predetermined control (standard) sample. The 30 sample from the subject is typically from a diseased tissue, such as cancer cells or tissues. The control sample can be from the same subject or from a different subject. The control sample is typically a normal, non-diseased sample. However, in some embodiments, such as for staging of disease or for evaluating the efficacy of treatment, the control sample can be from a diseased tissue. The control sample can be a combination of samples from

several different subjects. In some embodiments, the biomarker amount and/or activity measurement(s) from a subject is compared to a pre-determined level. This pre-determined level is typically obtained from normal samples. As described herein, a “pre-determined” biomarker amount and/or activity measurement(s) may be a biomarker amount and/or activity measurement(s) used to, by way of example only, evaluate a subject that may be selected for treatment (*e.g.*, based on the number of genomic mutations and/or the number of genomic mutations causing non-functional proteins for DNA repair genes), evaluate a response to an inhibitor of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erap1, etc.), and an immunotherapy combination treatment, and/or evaluate a response to an inhibitor of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein, and an immunotherapy combination treatment with one or more additional anti-cancer therapies. A pre-determined biomarker amount and/or activity measurement(s) may be determined in populations of patients with or without cancer. The pre-determined biomarker amount and/or activity measurement(s) can be a single number, equally applicable to every patient, or the pre-determined biomarker amount and/or activity measurement(s) can vary according to specific subpopulations of patients. Age, weight, height, and other factors of a subject may affect the pre-determined biomarker amount and/or activity measurement(s) of the individual. Furthermore, the pre-determined biomarker amount and/or activity can be determined for each subject individually. In one embodiment, the amounts determined and/or compared in a method described herein are based on absolute measurements.

In another embodiment, the amounts determined and/or compared in a method described herein are based on relative measurements, such as ratios (*e.g.*, biomarker copy numbers, level, and/or activity before a treatment vs. after a treatment, such biomarker measurements relative to a spiked or man-made control, such biomarker measurements relative to the expression of a housekeeping gene, and the like). For example, the relative analysis can be based on the ratio of pre-treatment biomarker measurement as compared to post-treatment biomarker measurement. Pre-treatment biomarker measurement can be made at any time prior to initiation of anti-cancer therapy. Post-treatment biomarker measurement can be made at any time after initiation of anti-cancer therapy. In some embodiments, post-treatment biomarker measurements are made 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 weeks or more after initiation of anti-cancer therapy,

and even longer toward indefinitely for continued monitoring. Treatment can comprise anti-cancer therapy, such as a therapeutic regimen comprising one or more inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein, and immunotherapy combination treatment alone or in
5 combination with other anti-cancer agents, such as with immune checkpoint inhibitors.

The pre-determined biomarker amount and/or activity measurement(s) can be any suitable standard. For example, the pre-determined biomarker amount and/or activity measurement(s) can be obtained from the same or a different human for whom a patient selection is being assessed. In one embodiment, the pre-determined biomarker amount
10 and/or activity measurement(s) can be obtained from a previous assessment of the same patient. In such a manner, the progress of the selection of the patient can be monitored over time. In addition, the control can be obtained from an assessment of another human or multiple humans, *e.g.*, selected groups of humans, if the subject is a human. In such a manner, the extent of the selection of the human for whom selection is being assessed can
15 be compared to suitable other humans, *e.g.*, other humans who are in a similar situation to the human of interest, such as those suffering from similar or the same condition(s) and/or of the same ethnic group.

In some embodiments of the present invention the change of biomarker amount and/or activity measurement(s) from the pre-determined level is about 0.1, 0.2, 0.3, 0.4, 0.5,
20 0.6, 0.7, 0.8, 0.9, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, or 5.0 fold or greater, or any range in between, inclusive. Such cutoff values apply equally when the measurement is based on relative changes, such as based on the ratio of pre-treatment biomarker measurement as compared to post-treatment biomarker measurement.

Biological samples can be collected from a variety of sources from a patient
25 including a body fluid sample, cell sample, or a tissue sample comprising nucleic acids and/or proteins. "Body fluids" refer to fluids that are excreted or secreted from the body as well as fluids that are normally not (*e.g.*, amniotic fluid, aqueous humor, bile, blood and blood plasma, cerebrospinal fluid, cerumen and earwax, cowper's fluid or pre-ejaculatory fluid, chyle, chyme, stool, female ejaculate, interstitial fluid, intracellular fluid, lymph,
30 menses, breast milk, mucus, pleural fluid, pus, saliva, sebum, semen, serum, sweat, synovial fluid, tears, urine, vaginal lubrication, vitreous humor, vomit). In a preferred embodiment, the subject and/or control sample is selected from the group consisting of cells, cell lines, histological slides, paraffin embedded tissues, biopsies, whole blood, nipple

aspirate, serum, plasma, buccal scrape, saliva, cerebrospinal fluid, urine, stool, and bone marrow. In one embodiment, the sample is serum, plasma, or urine. In another embodiment, the sample is serum.

The samples can be collected from individuals repeatedly over a longitudinal period of time (*e.g.*, once or more on the order of days, weeks, months, annually, biannually, *etc.*). Obtaining numerous samples from an individual over a period of time can be used to verify results from earlier detections and/or to identify an alteration in biological pattern as a result of, for example, disease progression, drug treatment, *etc.* For example, subject samples can be taken and monitored every month, every two months, or combinations of one, two, or three month intervals according to the present invention. In addition, the biomarker amount and/or activity measurements of the subject obtained over time can be conveniently compared with each other, as well as with those of normal controls during the monitoring period, thereby providing the subject's own values, as an internal, or personal, control for long-term monitoring.

Sample preparation and separation can involve any of the procedures, depending on the type of sample collected and/or analysis of biomarker measurement(s). Such procedures include, by way of example only, concentration, dilution, adjustment of pH, removal of high abundance polypeptides (*e.g.*, albumin, gamma globulin, and transferrin, *etc.*), addition of preservatives and calibrants, addition of protease inhibitors, addition of denaturants, desalting of samples, concentration of sample proteins, extraction and purification of lipids.

The sample preparation can also isolate molecules that are bound in non-covalent complexes to other protein (*e.g.*, carrier proteins). This process may isolate those molecules bound to a specific carrier protein (*e.g.*, albumin), or use a more general process, such as the release of bound molecules from all carrier proteins via protein denaturation, for example using an acid, followed by removal of the carrier proteins.

Removal of undesired proteins (*e.g.*, high abundance, uninformative, or undetectable proteins) from a sample can be achieved using high affinity reagents, high molecular weight filters, ultracentrifugation and/or electrodialysis. High affinity reagents include antibodies or other reagents (*e.g.*, aptamers) that selectively bind to high abundance proteins. Sample preparation could also include ion exchange chromatography, metal ion affinity chromatography, gel filtration, hydrophobic chromatography, chromatofocusing, adsorption chromatography, isoelectric focusing and related techniques. Molecular weight

filters include membranes that separate molecules on the basis of size and molecular weight. Such filters may further employ reverse osmosis, nanofiltration, ultrafiltration and microfiltration.

Ultracentrifugation is a method for removing undesired polypeptides from a sample.

5 Ultracentrifugation is the centrifugation of a sample at about 15,000-60,000 rpm while monitoring with an optical system the sedimentation (or lack thereof) of particles.

Electrodialysis is a procedure which uses an electromembrane or semipermeable membrane in a process in which ions are transported through semi-permeable membranes from one solution to another under the influence of a potential gradient. Since the membranes used
10 in electrodialysis may have the ability to selectively transport ions having positive or negative charge, reject ions of the opposite charge, or to allow species to migrate through a semipermeable membrane based on size and charge, it renders electrodialysis useful for concentration, removal, or separation of electrolytes.

Separation and purification in the present invention may include any procedure
15 known in the art, such as capillary electrophoresis (*e.g.*, in capillary or on-chip) or chromatography (*e.g.*, in capillary, column or on a chip). Electrophoresis is a method which can be used to separate ionic molecules under the influence of an electric field. Electrophoresis can be conducted in a gel, capillary, or in a microchannel on a chip. Examples of gels used for electrophoresis include starch, acrylamide, polyethylene oxides,
20 agarose, or combinations thereof. A gel can be modified by its cross-linking, addition of detergents, or denaturants, immobilization of enzymes or antibodies (affinity electrophoresis) or substrates (zymography) and incorporation of a pH gradient. Examples of capillaries used for electrophoresis include capillaries that interface with an electrospray.

Capillary electrophoresis (CE) is preferred for separating complex hydrophilic
25 molecules and highly charged solutes. CE technology can also be implemented on microfluidic chips. Depending on the types of capillary and buffers used, CE can be further segmented into separation techniques such as capillary zone electrophoresis (CZE), capillary isoelectric focusing (CIEF), capillary isotachophoresis (cITP) and capillary electrochromatography (CEC). An embodiment to couple CE techniques to electrospray
30 ionization involves the use of volatile solutions, for example, aqueous mixtures containing a volatile acid and/or base and an organic such as an alcohol or acetonitrile.

Capillary isotachophoresis (cITP) is a technique in which the analytes move through the capillary at a constant speed but are nevertheless separated by their respective

mobilities. Capillary zone electrophoresis (CZE), also known as free-solution CE (FSCE), is based on differences in the electrophoretic mobility of the species, determined by the charge on the molecule, and the frictional resistance the molecule encounters during migration which is often directly proportional to the size of the molecule. Capillary
5 isoelectric focusing (CIEF) allows weakly-ionizable amphoteric molecules, to be separated by electrophoresis in a pH gradient. CEC is a hybrid technique between traditional high performance liquid chromatography (HPLC) and CE.

Separation and purification techniques used in the present invention include any chromatography procedures known in the art. Chromatography can be based on the
10 differential adsorption and elution of certain analytes or partitioning of analytes between mobile and stationary phases. Different examples of chromatography include, but not limited to, liquid chromatography (LC), gas chromatography (GC), high performance liquid chromatography (HPLC), etc.

15 IV. Biomarker Nucleic Acids and Polypeptides

One aspect of the present invention pertains to the use of isolated nucleic acid molecules that correspond to biomarker nucleic acids that encode a biomarker polypeptide or a portion of such a polypeptide. As used herein, the term “nucleic acid molecule” is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules
20 (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

An “isolated” nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid molecule.

25 Preferably, an “isolated” nucleic acid molecule is free of sequences (preferably protein-encoding sequences) which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kB, 4 kB, 3 kB, 2 kB, 1 kB, 0.5 kB or 0.1 kB of
30 nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an “isolated” nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or

culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A biomarker nucleic acid molecule of the present invention can be isolated using standard molecular biology techniques and the sequence information in the database records described herein. Using all or a portion of such nucleic acid sequences, nucleic acid molecules of the present invention can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, ed., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

A nucleic acid molecule of the present invention can be amplified using cDNA, mRNA, or genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid molecules so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to all or a portion of a nucleic acid molecule of the present invention can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

Moreover, a nucleic acid molecule of the present invention can comprise only a portion of a nucleic acid sequence, wherein the full length nucleic acid sequence comprises a marker of the present invention or which encodes a polypeptide corresponding to a marker of the present invention. Such nucleic acid molecules can be used, for example, as a probe or primer. The probe/primer typically is used as one or more substantially purified oligonucleotides. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 7, preferably about 15, more preferably about 25, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 or more consecutive nucleotides of a biomarker nucleic acid sequence. Probes based on the sequence of a biomarker nucleic acid molecule can be used to detect transcripts or genomic sequences corresponding to one or more markers of the present invention. The probe comprises a label group attached thereto, *e.g.*, a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor.

A biomarker nucleic acid molecules that differ, due to degeneracy of the genetic code, from the nucleotide sequence of nucleic acid molecules encoding a protein which corresponds to the biomarker, and thus encode the same protein, are also contemplated.

In addition, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequence can exist within a population (*e.g.*, the human population). Such genetic polymorphisms can exist among individuals within a population due to natural allelic variation. An allele is one of a group
5 of genes which occur alternatively at a given genetic locus. In addition, it will be appreciated that DNA polymorphisms that affect RNA expression levels can also exist that may affect the overall expression level of that gene (*e.g.*, by affecting regulation or degradation).

The term “allele,” which is used interchangeably herein with “allelic variant,” refers
10 to alternative forms of a gene or portions thereof. Alleles occupy the same locus or position on homologous chromosomes. When a subject has two identical alleles of a gene, the subject is said to be homozygous for the gene or allele. When a subject has two different alleles of a gene, the subject is said to be heterozygous for the gene or allele. For example, biomarker alleles can differ from each other in a single nucleotide, or several nucleotides,
15 and can include substitutions, deletions, and insertions of nucleotides. An allele of a gene can also be a form of a gene containing one or more mutations.

The term “allelic variant of a polymorphic region of gene” or “allelic variant,” used interchangeably herein, refers to an alternative form of a gene having one of several possible nucleotide sequences found in that region of the gene in the population. As used
20 herein, allelic variant is meant to encompass functional allelic variants, non-functional allelic variants, SNPs, mutations and polymorphisms.

The term “single nucleotide polymorphism” (SNP) refers to a polymorphic site occupied by a single nucleotide, which is the site of variation between allelic sequences. The site is usually preceded by and followed by highly conserved sequences of the allele
25 (*e.g.*, sequences that vary in less than 1/100 or 1/1000 members of a population). A SNP usually arises due to substitution of one nucleotide for another at the polymorphic site. SNPs can also arise from a deletion of a nucleotide or an insertion of a nucleotide relative to a reference allele. Typically the polymorphic site is occupied by a base other than the reference base. For example, where the reference allele contains the base “T” (thymidine)
30 at the polymorphic site, the altered allele can contain a “C” (cytidine), “G” (guanine), or “A” (adenine) at the polymorphic site. SNP’s may occur in protein-coding nucleic acid sequences, in which case they may give rise to a defective or otherwise variant protein, or genetic disease. Such a SNP may alter the coding sequence of the gene and therefore

specify another amino acid (a “missense” SNP) or a SNP may introduce a stop codon (a “nonsense” SNP). When a SNP does not alter the amino acid sequence of a protein, the SNP is called “silent.” SNP’s may also occur in noncoding regions of the nucleotide sequence. This may result in defective protein expression, *e.g.*, as a result of alternative splicing, or it may have no effect on the function of the protein.

As used herein, the terms “gene” and “recombinant gene” refer to nucleic acid molecules comprising an open reading frame encoding a polypeptide corresponding to a marker of the present invention. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of a given gene. Alternative alleles can be identified by sequencing the gene of interest in a number of different individuals. This can be readily carried out by using hybridization probes to identify the same genetic locus in a variety of individuals. Any and all such nucleotide variations and resulting amino acid polymorphisms or variations that are the result of natural allelic variation and that do not alter the functional activity are intended to be within the scope of the present invention.

In another embodiment, a biomarker nucleic acid molecule is at least 7, 15, 20, 25, 30, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 450, 550, 650, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2200, 2400, 2600, 2800, 3000, 3500, 4000, 4500, or more nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule corresponding to a marker of the present invention or to a nucleic acid molecule encoding a protein corresponding to a marker of the present invention. As used herein, the term “hybridizes under stringent conditions” is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% (65%, 70%, 75%, 80%, preferably 85%) identical to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in sections 6.3.1-6.3.6 of *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989). A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50-65°C.

In addition to naturally-occurring allelic variants of a nucleic acid molecule of the present invention that can exist in the population, the skilled artisan will further appreciate that sequence changes can be introduced by mutation thereby leading to changes in the amino acid sequence of the encoded protein, without altering the biological activity of the protein encoded thereby. For example, one can make nucleotide substitutions leading to

amino acid substitutions at “non-essential” amino acid residues. A “non-essential” amino acid residue is a residue that can be altered from the wild-type sequence without altering the biological activity, whereas an “essential” amino acid residue is required for biological activity. For example, amino acid residues that are not conserved or only semi-conserved
5 among homologs of various species may be non-essential for activity and thus would be likely targets for alteration. Alternatively, amino acid residues that are conserved among the homologs of various species (*e.g.*, murine and human) may be essential for activity and thus would not be likely targets for alteration.

Accordingly, another aspect of the present invention pertains to nucleic acid
10 molecules encoding a polypeptide of the present invention that contain changes in amino acid residues that are not essential for activity. Such polypeptides differ in amino acid sequence from the naturally-occurring proteins which correspond to the markers of the present invention, yet retain biological activity. In one embodiment, a biomarker protein has an amino acid sequence that is at least about 40% identical, 50%, 60%, 70%, 75%,
15 80%, 83%, 85%, 87.5%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or identical to the amino acid sequence of a biomarker protein described herein.

An isolated nucleic acid molecule encoding a variant protein can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of nucleic acids of the present invention, such that one or more amino acid residue
20 substitutions, additions, or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A “conservative amino acid substitution” is one in which the amino acid residue is replaced with an amino acid residue having a similar
25 side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), non-polar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine,
30 tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Alternatively, mutations can be introduced randomly along all or part of the coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for biological activity to identify

mutants that retain activity. Following mutagenesis, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In some embodiments, the present invention further contemplates the use of anti-biomarker antisense nucleic acid molecules, *i.e.*, molecules which are complementary to a sense nucleic acid of the present invention, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule corresponding to a marker of the present invention or complementary to an mRNA sequence corresponding to a marker of the present invention. Accordingly, an antisense nucleic acid molecule of the present invention can hydrogen bond to (*i.e.* anneal with) a sense nucleic acid of the present invention. The antisense nucleic acid can be complementary to an entire coding strand, or to only a portion thereof, *e.g.*, all or part of the protein coding region (or open reading frame). An antisense nucleic acid molecule can also be antisense to all or part of a non-coding region of the coding strand of a nucleotide sequence encoding a polypeptide of the present invention. The non-coding regions ("5' and 3' untranslated regions") are the 5' and 3' sequences which flank the coding region and are not translated into amino acids.

An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides in length. An antisense nucleic acid can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-

oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been sub-cloned in an antisense orientation (*i.e.*, RNA transcribed from the
5 inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the present invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a polypeptide corresponding to a selected
10 marker of the present invention to thereby inhibit expression of the marker, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. Examples of a route of administration of antisense
15 nucleic acid molecules of the present invention includes direct injection at a tissue site or infusion of the antisense nucleic acid into a blood- or bone marrow-associated body fluid. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens
20 expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or
25 pol III promoter are preferred.

An antisense nucleic acid molecule of the present invention can be an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual α -units, the strands run parallel to each other (Gaultier *et al.*, 1987, *Nucleic Acids Res.* 15:6625-6641).
30 The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.*, 1987, *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987, *FEBS Lett.* 215:327-330).

The present invention also encompasses ribozymes. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes as described in Haselhoff and Gerlach, 1988, *Nature* 334:585-591) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of the protein encoded by the mRNA. A ribozyme having specificity for a nucleic acid molecule encoding a polypeptide corresponding to a marker of the present invention can be designed based upon the nucleotide sequence of a cDNA corresponding to the marker. For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved (see Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742). Alternatively, an mRNA encoding a polypeptide of the present invention can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (see, *e.g.*, Bartel and Szostak, 1993, *Science* 261:1411-1418).

The present invention also encompasses nucleic acid molecules which form triple helical structures. For example, expression of a biomarker protein can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the gene encoding the polypeptide (*e.g.*, the promoter and/or enhancer) to form triple helical structures that prevent transcription of the gene in target cells. See generally Helene (1991) *Anticancer Drug Des.* 6(6):569-84; Helene (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14(12):807-15.

In various embodiments, the nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acid molecules (see Hyrup *et al.*, 1996, *Bioorganic & Medicinal Chemistry* 4(1): 5-23). As used herein, the terms “peptide nucleic acids” or “PNAs” refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996), *supra*; Perry-O'Keefe *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:14670-675.

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 or translation arrest or inhibiting replication. PNAs can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*,
5 PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup (1996), *supra*; or as probes or primers for DNA sequence and hybridization (Hyrup, 1996, *supra*; Perry-O'Keefe *et al.*, 1996, *Proc. Natl. Acad. Sci. USA* 93:14670-675).

In another embodiment, PNAs can be modified, *e.g.*, to enhance their stability or
10 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 can be generated which can combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNASE H and DNA polymerases, to interact with the DNA portion while
15 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, 1996, *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996), *supra*, and Finn *et al.* (1996) *Nucleic Acids Res.* 24(17):3357-63. For example, a DNA chain can
20 be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs. Compounds such as 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite can be used as a link between the PNA and the 5' end of DNA (Mag *et al.*, 1989, *Nucleic Acids Res.* 17:5973-88). PNA monomers are then coupled in a step-wise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA
25 segment (Finn *et al.*, 1996, *Nucleic Acids Res.* 24(17):3357-63). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser *et al.*, 1975, *Bioorganic Med. Chem. Lett.* 5:1119-11124).

In other embodiments, the oligonucleotide can include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport
30 across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. WO 88/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered

cleavage agents (see, *e.g.*, Krol *et al.*, 1988, *Bio/Techniques* 6:958-976) or intercalating agents (see, *e.g.*, Zon, 1988, *Pharm. Res.* 5:539-549). To this end, the oligonucleotide can be conjugated to another molecule, *e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, *etc.*

5 Another aspect of the present invention pertains to the use of biomarker proteins and biologically active portions thereof. In one embodiment, the native polypeptide corresponding to a marker can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, polypeptides corresponding to a marker of the present invention are produced by
10 recombinant DNA techniques. Alternative to recombinant expression, a polypeptide corresponding to a marker of the present invention can be synthesized chemically using standard peptide synthesis techniques.

 An “isolated” or “purified” protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue
15 source from which the protein is derived, or substantially free of chemical precursors or other chemicals when chemically synthesized. The language “substantially free of cellular material” includes preparations of protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, protein that is substantially free of cellular material includes preparations of protein having less
20 than about 30%, 20%, 10%, or 5% (by dry weight) of heterologous protein (also referred to herein as a “contaminating protein”). When the protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, 10%, or 5% of the volume of the protein preparation. When the protein is produced by chemical synthesis, it is
25 preferably substantially free of chemical precursors or other chemicals, *i.e.*, it is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. Accordingly such preparations of the protein have less than about 30%, 20%, 10%, 5% (by dry weight) of chemical precursors or compounds other than the polypeptide of interest.

30 Biologically active portions of a biomarker polypeptide include polypeptides comprising amino acid sequences sufficiently identical to or derived from a biomarker protein amino acid sequence described herein, but which includes fewer amino acids than the full length protein, and exhibit at least one activity of the corresponding full-length

protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the corresponding protein. A biologically active portion of a protein of the present invention can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of the native form of a polypeptide of the present invention.

Preferred polypeptides have an amino acid sequence of a biomarker protein encoded by a nucleic acid molecule described herein. Other useful proteins are substantially identical (*e.g.*, at least about 40%, preferably 50%, 60%, 70%, 75%, 80%, 83%, 85%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%) to one of these sequences and retain the functional activity of the protein of the corresponding naturally-occurring protein yet differ in amino acid sequence due to natural allelic variation or mutagenesis.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = # of identical positions/total # of positions (*e.g.*, overlapping positions) x 100). In one embodiment the two sequences are the same length.

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the present invention. BLAST protein searches can be performed

with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecules of the present invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *Comput Appl Biosci*, 4:11-7. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Yet another useful algorithm for identifying regions of local sequence similarity and alignment is the FASTA algorithm as described in Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448. When using the FASTA algorithm for comparing nucleotide or amino acid sequences, a PAM120 weight residue table can, for example, be used with a *k*-tuple value of 2.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, only exact matches are counted.

The present invention also provides chimeric or fusion proteins corresponding to a biomarker protein. As used herein, a “chimeric protein” or “fusion protein” comprises all or part (preferably a biologically active part) of a polypeptide corresponding to a marker of the present invention operably linked to a heterologous polypeptide (*i.e.*, a polypeptide other than the polypeptide corresponding to the marker). Within the fusion protein, the term “operably linked” is intended to indicate that the polypeptide of the present invention and the heterologous polypeptide are fused in-frame to each other. The heterologous polypeptide can be fused to the amino-terminus or the carboxyl-terminus of the polypeptide of the present invention.

One useful fusion protein is a GST fusion protein in which a polypeptide corresponding to a marker of the present invention is fused to the carboxyl terminus of GST sequences. Such fusion proteins can facilitate the purification of a recombinant polypeptide of the present invention.

In another embodiment, the fusion protein contains a heterologous signal sequence, immunoglobulin fusion protein, toxin, or other useful protein sequence. Chimeric and fusion proteins of the present invention can be produced by standard recombinant DNA techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and re-amplified to generate a chimeric gene sequence (see, *e.g.*, Ausubel *et al.*, *supra*). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the present invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the polypeptide of the present invention.

A signal sequence can be used to facilitate secretion and isolation of the secreted protein or other proteins of interest. Signal sequences are typically characterized by a core of hydrophobic amino acids which are generally cleaved from the mature protein during secretion in one or more cleavage events. Such signal peptides contain processing sites that allow cleavage of the signal sequence from the mature proteins as they pass through the secretory pathway. Thus, the present invention pertains to the described polypeptides having a signal sequence, as well as to polypeptides from which the signal sequence has been proteolytically cleaved (*i.e.*, the cleavage products). In one embodiment, a nucleic acid sequence encoding a signal sequence can be operably linked in an expression vector to a protein of interest, such as a protein which is ordinarily not secreted or is otherwise difficult to isolate. The signal sequence directs secretion of the protein, such as from a eukaryotic host into which the expression vector is transformed, and the signal sequence is subsequently or concurrently cleaved. The protein can then be readily purified from the extracellular medium by art recognized methods. Alternatively, the signal sequence can be linked to the protein of interest using a sequence which facilitates purification, such as with a GST domain.

The present invention also pertains to variants of the biomarker polypeptides described herein. Such variants have an altered amino acid sequence which can function as either agonists (mimetics) or as antagonists. Variants can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation. An agonist can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of the protein. An

antagonist of a protein can inhibit one or more of the activities of the naturally occurring form of the protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the protein of interest. Thus, specific biological effects can be elicited by treatment with a variant of limited function.

- 5 Treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein can have fewer side effects in a subject relative to treatment with the naturally occurring form of the protein.

Variants of a biomarker protein which function as either agonists (mimetics) or as antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*,
10 truncation mutants, of the protein of the present invention for agonist or antagonist activity. In one embodiment, a variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of
15 potential protein sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display). There are a variety of methods which can be used to produce libraries of potential variants of the polypeptides of the present invention from a degenerate oligonucleotide sequence. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, 1983, *Tetrahedron*
20 39:3; Itakura *et al.*, 1984, *Annu. Rev. Biochem.* 53:323; Itakura *et al.*, 1984, *Science* 198:1056; Ike *et al.*, 1983 *Nucleic Acid Res.* 11:477).

In addition, libraries of fragments of the coding sequence of a polypeptide corresponding to a marker of the present invention can be used to generate a variegated population of polypeptides for screening and subsequent selection of variants. For
25 example, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of the coding sequence of interest with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed
30 duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes amino terminal and internal fragments of various sizes of the protein of interest.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify variants of a protein of the present invention (Arkin and Yourvan, 1992, *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.*, 1993, *Protein Engineering* 6(3):327- 331).

An isolated polypeptide or a fragment thereof (or a nucleic acid encoding such a polypeptide) corresponding to one or more biomarkers of the invention, including the biomarkers listed in Table 1 or fragments thereof, can be used as an immunogen to generate antibodies that bind to said immunogen, using standard techniques for polyclonal and monoclonal antibody preparation according to well-known methods in the art. An antigenic peptide comprises at least 8 amino acid residues and encompasses an epitope present in the respective full length molecule such that an antibody raised against the peptide forms a specific immune complex with the respective full length molecule. Preferably, the antigenic peptide comprises at least 10 amino acid residues. In one embodiment such epitopes can be specific for a given polypeptide molecule from one species, such as mouse or human (*i.e.*, an antigenic peptide that spans a region of the polypeptide molecule that is not conserved across species is used as immunogen; such non conserved residues can be determined using an alignment such as that provided herein).

In some embodiments, the immunotherapy utilizes an inhibitor of at least one immune checkpoint, such as an antibody binds substantially specifically to an immune checkpoint, such as PD-1, and inhibits or blocks its immunoinhibitory function, such as by interrupting its interaction with a binding partner of the immune checkpoint, such as PD-L1 and/or PD-L2 binding partners of PD-1. In one embodiment, an antibody, especially an intrbody, binds substantially specifically to one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.) and inhibits or blocks its biological

function, such as by interrupting its interaction with a substrate like STAT or JAK proteins. In another embodiment, an antibody, especially an intrbody, binds substantially specifically to the binding partner(s) of one or more biomarkers listed in Table 1, such as one or more proteins modulating antigen processing and presentation (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), such as substrates of such one or more biomarkers described herein, and inhibits or blocks its biological function, such as by interrupting its interaction to one or more biomarkers listed in Table 1.

For example, a polypeptide immunogen typically is used to prepare antibodies by immunizing a suitable subject (*e.g.*, rabbit, goat, mouse or other mammal) with the immunogen. A preferred animal is a mouse deficient in the desired target antigen. For example, a PD-1 knockout mouse if the desired antibody is an anti-PD-1 antibody, may be used. This results in a wider spectrum of antibody recognition possibilities as antibodies reactive to common mouse and human epitopes are not removed by tolerance mechanisms. An appropriate immunogenic preparation can contain, for example, a recombinantly expressed or chemically synthesized molecule or fragment thereof to which the immune response is to be generated. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic preparation induces a polyclonal antibody response to the antigenic peptide contained therein.

Polyclonal antibodies can be prepared as described above by immunizing a suitable subject with a polypeptide immunogen. The polypeptide antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized polypeptide. If desired, the antibody directed against the antigen can be isolated from the mammal (*e.g.*, from the blood) and further purified by well-known techniques, such as protein A chromatography, to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique (originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem.* 255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci.* 76:2927-31; Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol. Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985) *Monoclonal Antibodies and Cancer Therapy*,

Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well-known (see generally Kenneth, R. H. in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); Lerner, E. A. (1981) *Yale J. Biol. Med.* 54:387-402; 5 Gefter, M. L. *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with an immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds to the polypeptide antigen, preferably specifically. In some 10 embodiments, the immunization is performed in a cell or animal host that has a knockout of a target antigen of interest (*e.g.*, does not produce the antigen prior to immunization).

Any of the many well-known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating a monoclonal antibody against one or more biomarkers of the invention, including the biomarkers listed in Table 1, 15 or a fragment thereof (see, *e.g.*, Galfre, G. *et al.* (1977) *Nature* 266:55052; Gefter *et al.* (1977) *supra*; Lerner (1981) *supra*; Kenneth (1980) *supra*). Moreover, the ordinary skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (*e.g.*, a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be 20 made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, *e.g.*, the P3-NS1/1-Ag4-1, P3- 25 x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from the American Type Culture Collection (ATCC), Rockville, MD. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several 30 days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind a given polypeptide, *e.g.*, using a standard ELISA assay.

As an alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal specific for one of the above described polypeptides can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with the appropriate polypeptide to thereby isolate immunoglobulin library members that bind the polypeptide. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening an antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* International Publication No. WO 92/18619; Dower *et al.* International Publication No. WO 91/17271; Winter *et al.* International Publication WO 92/20791; Markland *et al.* International Publication No. WO 92/15679; Breitling *et al.* International Publication WO 93/01288; McCafferty *et al.* International Publication No. WO 92/01047; Garrard *et al.* International Publication No. WO 92/09690; Ladner *et al.* International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Biotechnology (NY)* 9:1369-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J.* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrard *et al.* (1991) *Biotechnology (NY)* 9:1373-1377; Hoogenboom *et al.* (1991) *Nucleic Acids Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al.* (1990) *Nature* 348:552-554.

Since it is well-known in the art that antibody heavy and light chain CDR3 domains play a particularly important role in the binding specificity/affinity of an antibody for an antigen, the recombinant monoclonal antibodies of the present invention prepared as set forth above preferably comprise the heavy and light chain CDR3s of variable regions of the antibodies described herein and well-known in the art. Similarly, the antibodies can further comprise the CDR2s of variable regions of said antibodies. The antibodies can further comprise the CDR1s of variable regions of said antibodies. In other embodiments, the antibodies can comprise any combinations of the CDRs.

The CDR1, 2, and/or 3 regions of the engineered antibodies described above can comprise the exact amino acid sequence(s) as those of variable regions of the present invention described herein. However, the ordinarily skilled artisan will appreciate that

some deviation from the exact CDR sequences may be possible while still retaining the ability of the antibody, especially an introbody, to bind a desired target, such as one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.) and/or a binding partner thereof, either alone or in combination with an immunotherapy, such as the one or more biomarkers, the binding partners/substrates of such biomarkers, or an immunotherapy effectively (*e.g.*, conservative sequence modifications). Accordingly, in another embodiment, the engineered antibody may be composed of one or more CDRs that are, for example, 50%, 60%, 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to one or more CDRs of the present invention described herein or otherwise publicly available.

For example, the structural features of non-human or human antibodies (*e.g.*, a rat anti-mouse/anti-human antibody) can be used to create structurally related human antibodies, especially introbodies, that retain at least one functional property of the antibodies of the present invention, such as binding to one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein, the binding partners/substrates of such one or more biomarkers, and/or an immune checkpoint. Another functional property includes inhibiting binding of the original known, non-human or human antibodies in a competition ELISA assay.

Antibodies, immunoglobulins, and polypeptides of the invention can be used in an isolated (*e.g.*, purified) form or contained in a vector, such as a membrane or lipid vesicle (*e.g.* a liposome). Moreover, amino acid sequence modification(s) of the antibodies described herein are contemplated. For example, it may be desirable to improve the binding affinity and/or other biological properties of the antibody. It is known that when a humanized antibody is produced by simply grafting only CDRs in VH and VL of an antibody derived from a non-human animal in FRs of the VH and VL of a human antibody, the antigen binding activity is reduced in comparison with that of the original antibody derived from a non-human animal. It is considered that several amino acid residues of the VH and VL of the non-human antibody, not only in CDRs but also in FRs, are directly or indirectly associated with the antigen binding activity. Hence, substitution of these amino acid residues with different amino acid residues derived from FRs of the VH and VL of the human antibody would reduce binding activity and can be corrected by replacing the amino acids with amino acid residues of the original antibody derived from a non-human animal.

Similarly, modifications and changes may be made in the structure of the antibodies described herein, and in the DNA sequences encoding them, and still obtain a functional molecule that encodes an antibody and polypeptide with desirable characteristics. For example, antibody glycosylation patterns can be modulated to, for example, increase stability. By “altering” is meant deleting one or more carbohydrate moieties found in the antibody, and/or adding one or more glycosylation sites that are not present in the antibody. Glycosylation of antibodies is typically N-linked. “N-linked” refers to the attachment of the carbohydrate moiety to the side chain of an asparagine residue. The tripeptide sequences asparagine-X-serine and asparagines-X-threonine, where X is any amino acid except proline, are the recognition sequences for enzymatic attachment of the carbohydrate moiety to the asparagine side chain. Thus, the presence of either of these tripeptide sequences in a polypeptide creates a potential glycosylation site. Addition of glycosylation sites to the antibody is conveniently accomplished by altering the amino acid sequence such that it contains one or more of the above-described tripeptide sequences (for N-linked glycosylation sites). Another type of covalent modification involves chemically or enzymatically coupling glycosides to the antibody. These procedures are advantageous in that they do not require production of the antibody in a host cell that has glycosylation capabilities for N- or O-linked glycosylation. Depending on the coupling mode used, the sugar(s) may be attached to (a) arginine and histidine, (b) free carboxyl groups, (c) free sulfhydryl groups such as those of cysteine, (d) free hydroxyl groups such as those of serine, threonine, or hydroxyproline, (e) aromatic residues such as those of phenylalanine, tyrosine, or tryptophan, or (f) the amide group of glutamine. For example, such methods are described in WO87/05330.

Similarly, removal of any carbohydrate moieties present on the antibody may be accomplished chemically or enzymatically. Chemical deglycosylation requires exposure of the antibody to the compound trifluoromethanesulfonic acid, or an equivalent compound. This treatment results in the cleavage of most or all sugars except the linking sugar (N-acetylglucosamine or N-acetylgalactosamine), while leaving the antibody intact. Chemical deglycosylation is described by *Sojahn et al.* (1987) and by *Edge et al.* (1981). Enzymatic cleavage of carbohydrate moieties on antibodies can be achieved by the use of a variety of endo- and exo-glycosidases as described by *Thotakura et al.* (1987).

Other modifications can involve the formation of immunoconjugates. For example, in one type of covalent modification, antibodies or proteins are covalently linked to one of a

variety of non proteinaceous polymers, *e.g.*, polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Pat. No. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

Conjugation of antibodies or other proteins of the present invention with
5 heterologous agents can be made using a variety of bifunctional protein coupling agents including but not limited to N-succinimidyl (2-pyridyldithio) propionate (SPDP), succinimidyl (N-maleimidomethyl)cyclohexane-1-carboxylate, iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido
10 compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as toluene 2,6diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, carbon labeled 1-isothiocyanatobenzyl methyl-diethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of
15 radionucleotide to the antibody (WO 94/11026).

In another aspect, the present invention features antibodies conjugated to a therapeutic moiety, such as a cytotoxin, a drug, and/or a radioisotope. When conjugated to a cytotoxin, these antibody conjugates are referred to as “immunotoxins.” A cytotoxin or cytotoxic agent includes any agent that is detrimental to (*e.g.*, kills) cells. Examples include
20 taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*,
25 methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics
30 (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine). An antibody of the present invention can be conjugated to a radioisotope, *e.g.*, radioactive iodine, to generate cytotoxic radiopharmaceuticals for treating a related disorder, such as a cancer.

Conjugated antibodies, in addition to therapeutic utility, can be useful for diagnostically or prognostically to monitor polypeptide levels in tissue as part of a clinical testing procedure, *e.g.*, to determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate (FITC), rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin (PE); an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S , or ^3H . [0134] As used herein, the term “labeled”, with regard to the antibody, is intended to encompass direct labeling of the antibody by coupling (*i.e.*, physically linking) a detectable substance, such as a radioactive agent or a fluorophore (*e.g.* fluorescein isothiocyanate (FITC) or phycoerythrin (PE) or Indocyanine (Cy5)) to the antibody, as well as indirect labeling of the antibody by reactivity with a detectable substance.

The antibody conjugates of the present invention can be used to modify a given biological response. The therapeutic moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, an enzymatically active toxin, or active fragment thereof, such as abrin, ricin A, *Pseudomonas* exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor or interferon- γ ; or, biological response modifiers such as, for example, lymphokines, interleukin-1 (“IL-1”), interleukin-2 (“IL-2”), interleukin-6 (“IL-6”), granulocyte macrophage colony stimulating factor (“GM-CSF”), granulocyte colony stimulating factor (“G-CSF”), or other cytokines or growth factors.

In one embodiment, an antibody for use in the instant invention is a bispecific or multispecific antibody. A bispecific antibody has binding sites for two different antigens within a single antibody polypeptide. Antigen binding may be simultaneous or sequential. Triomas and hybrid hybridomas are two examples of cell lines that can secrete bispecific

antibodies. Examples of bispecific antibodies produced by a hybrid hybridoma or a trioma are disclosed in U.S. Patent 4,474,893. Bispecific antibodies have been constructed by chemical means (Staerz *et al.* (1985) *Nature* 314:628, and Perez *et al.* (1985) *Nature* 316:354) and hybridoma technology (Staerz and Bevan (1986) *Proc. Natl. Acad. Sci. USA*, 83:1453, and Staerz and Bevan (1986) *Immunol. Today* 7:241). Bispecific antibodies are also described in U.S. Patent 5,959,084. Fragments of bispecific antibodies are described in U.S. Patent 5,798,229.

Bispecific agents can also be generated by making heterohybridomas by fusing hybridomas or other cells making different antibodies, followed by identification of clones producing and co-assembling both antibodies. They can also be generated by chemical or genetic conjugation of complete immunoglobulin chains or portions thereof such as Fab and Fv sequences. The antibody component can bind to a polypeptide or a fragment thereof of one or more biomarkers of the invention, including one or more biomarkers listed in Table 1, or a fragment thereof. In one embodiment, the bispecific antibody could specifically bind to both a polypeptide or a fragment thereof and its natural binding partner(s) or a fragment(s) thereof.

Techniques for modulating antibodies, such as humanization, conjugation, recombinant techniques, and the like are well-known in the art.

In another aspect of this invention, peptides or peptide mimetics can be used to antagonize the activity of one or more biomarkers of the invention, including one or more biomarkers listed in Table 1, or a fragment(s) thereof. In one embodiment, variants of one or more biomarkers listed in Table 1 which function as a modulating agent for the respective full length protein, can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, for antagonist activity. In one embodiment, a variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of variants can be produced, for instance, by enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential polypeptide sequences is expressible as individual polypeptides containing the set of polypeptide sequences therein. There are a variety of methods which can be used to produce libraries of polypeptide variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision,

in one mixture, of all of the sequences encoding the desired set of potential polypeptide sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang, S. A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477.

5 In addition, libraries of fragments of a polypeptide coding sequence can be used to generate a variegated population of polypeptide fragments for screening and subsequent selection of variants of a given polypeptide. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a polypeptide coding sequence with a nuclease under conditions wherein nicking occurs only
10 about once per polypeptide, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and
15 internal fragments of various sizes of the polypeptide.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of
20 polypeptides. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was
25 detected. Recursive ensemble mutagenesis (REM), a technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify variants of interest (Arkin and Youvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delagrave *et al.* (1993) *Protein Eng.* 6(3):327-331). In one embodiment, cell based assays can be exploited to analyze a variegated polypeptide library.
30 For example, a library of expression vectors can be transfected into a cell line which ordinarily synthesizes one or more biomarkers of the invention, including one or more biomarkers listed in Table 1, or a fragment thereof. The transfected cells are then cultured such that the full length polypeptide and a particular mutant polypeptide are produced and

the effect of expression of the mutant on the full length polypeptide activity in cell supernatants can be detected, *e.g.*, by any of a number of functional assays. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of full length polypeptide activity, and the individual clones further
5 characterized.

Systematic substitution of one or more amino acids of a polypeptide amino acid sequence with a D-amino acid of the same type (*e.g.*, D-lysine in place of L-lysine) can be used to generate more stable peptides. In addition, constrained peptides comprising a polypeptide amino acid sequence of interest or a substantially identical sequence variation
10 can be generated by methods known in the art (Rizo and Gierasch (1992) *Annu. Rev. Biochem.* 61:387, incorporated herein by reference); for example, by adding internal cysteine residues capable of forming intramolecular disulfide bridges which cyclize the peptide.

The amino acid sequences described herein will enable those of skill in the art to
15 produce polypeptides corresponding peptide sequences and sequence variants thereof. Such polypeptides can be produced in prokaryotic or eukaryotic host cells by expression of polynucleotides encoding the peptide sequence, frequently as part of a larger polypeptide. Alternatively, such peptides can be synthesized by chemical methods. Methods for expression of heterologous proteins in recombinant hosts, chemical synthesis of
20 polypeptides, and *in vitro* translation are well-known in the art and are described further in Maniatis *et al. Molecular Cloning: A Laboratory Manual* (1989), 2nd Ed., Cold Spring Harbor, N.Y.; Berger and Kimmel, *Methods in Enzymology*, Volume 152, Guide to Molecular Cloning Techniques (1987), Academic Press, Inc., San Diego, Calif.; Merrifield, J. (1969) *J. Am. Chem. Soc.* 91:501; Chaiken I. M. (1981) *CRC Crit. Rev. Biochem.* 11:
25 255; Kaiser *et al.* (1989) *Science* 243:187; Merrifield, B. (1986) *Science* 232:342; Kent, S. B. H. (1988) *Annu. Rev. Biochem.* 57:957; and Offord, R. E. (1980) *Semisynthetic Proteins*, Wiley Publishing, which are incorporated herein by reference).

Peptides can be produced, typically by direct chemical synthesis. Peptides can be produced as modified peptides, with nonpeptide moieties attached by covalent linkage to
30 the N-terminus and/or C-terminus. In certain preferred embodiments, either the carboxy-terminus or the amino-terminus, or both, are chemically modified. The most common modifications of the terminal amino and carboxyl groups are acetylation and amidation, respectively. Amino-terminal modifications such as acylation (*e.g.*, acetylation) or

alkylation (*e.g.*, methylation) and carboxy-terminal-modifications such as amidation, as well as other terminal modifications, including cyclization, can be incorporated into various embodiments of the invention. Certain amino-terminal and/or carboxy-terminal modifications and/or peptide extensions to the core sequence can provide advantageous physical, chemical, biochemical, and pharmacological properties, such as: enhanced stability, increased potency and/or efficacy, resistance to serum proteases, desirable pharmacokinetic properties, and others. Peptides described herein can be used therapeutically to treat disease, *e.g.*, by altering costimulation in a patient.

Peptidomimetics (Fauchere (1986) *Adv. Drug Res.* 15:29; Veber and Freidinger (1985) *TINS* p.392; and Evans *et al.* (1987) *J. Med. Chem.* 30:1229, which are incorporated herein by reference) are usually developed with the aid of computerized molecular modeling. Peptide mimetics that are structurally similar to therapeutically useful peptides can be used to produce an equivalent therapeutic or prophylactic effect. Generally, peptidomimetics are structurally similar to a paradigm polypeptide (*i.e.*, a polypeptide that has a biological or pharmacological activity), but have one or more peptide linkages optionally replaced by a linkage selected from the group consisting of: -CH₂NH-, -CH₂S-, -CH₂-CH₂-, -CH=CH- (*cis* and *trans*), -COCH₂-, -CH(OH)CH₂-, and -CH₂SO-, by methods known in the art and further described in the following references: Spatola, A. F. in “*Chemistry and Biochemistry of Amino Acids, Peptides, and Proteins*” Weinstein, B., ed., Marcel Dekker, New York, p. 267 (1983); Spatola, A. F., Vega Data (March 1983), Vol. 1, Issue 3, “Peptide Backbone Modifications” (general review); Morley, J. S. (1980) *Trends Pharm. Sci.* pp. 463-468 (general review); Hudson, D. *et al.* (1979) *Int. J. Pept. Prot. Res.* 14:177-185 (-CH₂NH-, CH₂CH₂-); Spatola, A. F. *et al.* (1986) *Life Sci.* 38:1243-1249 (-CH₂-S); Hann, M. M. (1982) *J. Chem. Soc. Perkin Trans. I.* 307-314 (-CH-CH-, *cis* and *trans*); Almquist, R. G. *et al.* (1990) *J. Med. Chem.* 23:1392-1398 (-COCH₂-); Jennings-White, C. *et al.* (1982) *Tetrahedron Lett.* 23:2533 (-COCH₂-); Szelke, M. *et al.* European Appln. EP 45665 (1982) CA: 97:39405 (1982)(-CH(OH)CH₂-); Holladay, M. W. *et al.* (1983) *Tetrahedron Lett.* (1983) 24:4401-4404 (-C(OH)CH₂-); and Hruby, V. J. (1982) *Life Sci.* (1982) 31:189-199 (-CH₂-S-); each of which is incorporated herein by reference. A particularly preferred non-peptide linkage is -CH₂NH-. Such peptide mimetics may have significant advantages over polypeptide embodiments, including, for example: more economical production, greater chemical stability, enhanced pharmacological properties (half-life, absorption, potency, efficacy,

etc.), altered specificity (e.g., a broad-spectrum of biological activities), reduced antigenicity, and others. Labeling of peptidomimetics usually involves covalent attachment of one or more labels, directly or through a spacer (e.g., an amide group), to non-interfering position(s) on the peptidomimetic that are predicted by quantitative
5 structure-activity data and/or molecular modeling. Such non-interfering positions generally are positions that do not form direct contacts with the macropolypeptides(s) to which the peptidomimetic binds to produce the therapeutic effect. Derivatization (e.g., labeling) of peptidomimetics should not substantially interfere with the desired biological or pharmacological activity of the peptidomimetic.

10 Also encompassed by the present invention are small molecules which can modulate (either enhance or inhibit) interactions, e.g., between biomarkers described herein or listed in Table 1 and their natural binding partners. The small molecules of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: spatially addressable parallel solid phase or solution
15 phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. (Lam, K. S. (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:6909; Erb *et al.* (1994)
20 *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994) *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds can be presented in solution (e.g., Houghten (1992)
25 *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-
30 310); (Ladner *supra.*). Compounds can be screened in cell based or non-cell based assays. Compounds can be screened in pools (e.g. multiple compounds in each testing sample) or as individual compounds.

Chimeric or fusion proteins can be prepared for the inhibitor(s) of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.) and/or agents for the immunotherapies described herein, such as inhibitors to the

5 biomarkers of the invention, including the biomarkers listed in Table 1, or fragments thereof. As used herein, a “chimeric protein” or “fusion protein” comprises one or more biomarkers of the invention, including one or more biomarkers listed in Table 1, or a fragment thereof, operatively linked to another polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the respective

10 biomarker. In a preferred embodiment, the fusion protein comprises at least one biologically active portion of one or more biomarkers of the invention, including one or more biomarkers listed in Table 1, or fragments thereof. Within the fusion protein, the term “operatively linked” is intended to indicate that the biomarker sequences and the non-biomarker sequences are fused in-frame to each other in such a way as to preserve functions

15 exhibited when expressed independently of the fusion. The “another” sequences can be fused to the N-terminus or C-terminus of the biomarker sequences, respectively.

Such a fusion protein can be produced by recombinant expression of a nucleotide sequence encoding the first peptide and a nucleotide sequence encoding the second peptide. The second peptide may optionally correspond to a moiety that alters the solubility, affinity,

20 stability or valency of the first peptide, for example, an immunoglobulin constant region. In another preferred embodiment, the first peptide consists of a portion of a biologically active molecule (*e.g.* the extracellular portion of the polypeptide or the ligand binding portion). The second peptide can include an immunoglobulin constant region, for example, a human C γ 1 domain or C γ 4 domain (*e.g.*, the hinge, CH2 and CH3 regions of human IgC γ 1, or human IgC γ 4, see *e.g.*, Capon *et al.* U.S. Patents 5,116,964; 5,580,756; 5,844,095 and the like, incorporated herein by reference). Such constant regions may retain regions which mediate effector function (*e.g.* Fc receptor binding) or may be altered to reduce effector function. A resulting fusion protein may have altered solubility, binding affinity, stability and/or valency (*i.e.*, the number of binding sites available per polypeptide) as compared to

25 the independently expressed first peptide, and may increase the efficiency of protein purification. Fusion proteins and peptides produced by recombinant techniques can be secreted and isolated from a mixture of cells and medium containing the protein or peptide. Alternatively, the protein or peptide can be retained cytoplasmically and the cells harvested,

30

lysed and the protein isolated. A cell culture typically includes host cells, media and other byproducts. Suitable media for cell culture are well-known in the art. Protein and peptides can be isolated from cell culture media, host cells, or both using techniques known in the art for purifying proteins and peptides. Techniques for transfecting host cells and purifying proteins and peptides are known in the art.

Preferably, a fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992).

The fusion proteins of the invention can be used as immunogens to produce antibodies in a subject. Such antibodies may be used to purify the respective natural polypeptides from which the fusion proteins were generated, or in screening assays to identify polypeptides which inhibit the interactions between one or more biomarkers polypeptide or a fragment thereof and its natural binding partner(s) or a fragment(s) thereof.

Also provided herein are compositions comprising one or more nucleic acids comprising or capable of expressing at least 1, 2, 3, 4, 5, 10, 20 or more small nucleic acids or antisense oligonucleotides or derivatives thereof, wherein said small nucleic acids or antisense oligonucleotides or derivatives thereof in a cell specifically hybridize (*e.g.*, bind) under cellular conditions, with cellular nucleic acids (*e.g.*, small non-coding RNAs such as miRNAs, pre-miRNAs, pri-miRNAs, miRNA*, anti-miRNA, a miRNA binding site, a variant and/or functional variant thereof, cellular mRNAs or a fragments thereof). In one embodiment, expression of the small nucleic acids or antisense oligonucleotides or derivatives thereof in a cell can inhibit expression or biological activity of cellular nucleic acids and/or proteins, *e.g.*, by inhibiting transcription, translation and/or small nucleic acid processing of, for example, one or more biomarkers of the invention, including one or more

biomarkers listed in Table 1, or fragment(s) thereof. In one embodiment, the small nucleic acids or antisense oligonucleotides or derivatives thereof are small RNAs (*e.g.*, microRNAs) or complements of small RNAs. In another embodiment, the small nucleic acids or antisense oligonucleotides or derivatives thereof can be single or double stranded and are at least six nucleotides in length and are less than about 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100, 50, 40, 30, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, or 10 nucleotides in length. In another embodiment, a composition may comprise a library of nucleic acids comprising or capable of expressing small nucleic acids or antisense oligonucleotides or derivatives thereof, or pools of said small nucleic acids or antisense oligonucleotides or derivatives thereof. A pool of nucleic acids may comprise about 2-5, 5-10, 10-20, 10-30 or more nucleic acids comprising or capable of expressing small nucleic acids or antisense oligonucleotides or derivatives thereof.

In one embodiment, binding may be by conventional base pair complementarity, or, for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix. In general, “antisense” refers to the range of techniques generally employed in the art, and includes any process that relies on specific binding to oligonucleotide sequences.

It is well-known in the art that modifications can be made to the sequence of a miRNA or a pre-miRNA without disrupting miRNA activity. As used herein, the term “functional variant” of a miRNA sequence refers to an oligonucleotide sequence that varies from the natural miRNA sequence, but retains one or more functional characteristics of the miRNA (*e.g.* cancer cell proliferation inhibition, induction of cancer cell apoptosis, enhancement of cancer cell susceptibility to chemotherapeutic agents, specific miRNA target inhibition). In some embodiments, a functional variant of a miRNA sequence retains all of the functional characteristics of the miRNA. In certain embodiments, a functional variant of a miRNA has a nucleobase sequence that is at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to the miRNA or precursor thereof over a region of about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or more nucleobases, or that the functional variant hybridizes to the complement of the miRNA or precursor thereof under stringent hybridization conditions. Accordingly, in certain embodiments the nucleobase sequence of a functional variant is capable of hybridizing to one or more target sequences of the miRNA.

miRNAs and their corresponding stem-loop sequences described herein may be found in miRBase, an online searchable database of miRNA sequences and annotation, found on the world wide web at microrna.sanger.ac.uk. Entries in the miRBase Sequence database represent a predicted hairpin portion of a miRNA transcript (the stem-loop), with information on the location and sequence of the mature miRNA sequence. The miRNA stem-loop sequences in the database are not strictly precursor miRNAs (pre-miRNAs), and may in some instances include the pre-miRNA and some flanking sequence from the presumed primary transcript. The miRNA nucleobase sequences described herein encompass any version of the miRNA, including the sequences described in Release 10.0 of the miRBase sequence database and sequences described in any earlier Release of the miRBase sequence database. A sequence database release may result in the re-naming of certain miRNAs. A sequence database release may result in a variation of a mature miRNA sequence.

In some embodiments, miRNA sequences of the invention may be associated with a second RNA sequence that may be located on the same RNA molecule or on a separate RNA molecule as the miRNA sequence. In such cases, the miRNA sequence may be referred to as the active strand, while the second RNA sequence, which is at least partially complementary to the miRNA sequence, may be referred to as the complementary strand. The active and complementary strands are hybridized to create a double-stranded RNA that is similar to a naturally occurring miRNA precursor. The activity of a miRNA may be optimized by maximizing uptake of the active strand and minimizing uptake of the complementary strand by the miRNA protein complex that regulates gene translation. This can be done through modification and/or design of the complementary strand.

In some embodiments, the complementary strand is modified so that a chemical group other than a phosphate or hydroxyl at its 5' terminus. The presence of the 5' modification apparently eliminates uptake of the complementary strand and subsequently favors uptake of the active strand by the miRNA protein complex. The 5' modification can be any of a variety of molecules known in the art, including NH_2 , NHCOCH_3 , and biotin.

In another embodiment, the uptake of the complementary strand by the miRNA pathway is reduced by incorporating nucleotides with sugar modifications in the first 2-6 nucleotides of the complementary strand. It should be noted that such sugar modifications can be combined with the 5' terminal modifications described above to further enhance miRNA activities.

In some embodiments, the complementary strand is designed so that nucleotides in the 3' end of the complementary strand are not complementary to the active strand. This results in double-strand hybrid RNAs that are stable at the 3' end of the active strand but relatively unstable at the 5' end of the active strand. This difference in stability enhances the uptake of the active strand by the miRNA pathway, while reducing uptake of the complementary strand, thereby enhancing miRNA activity.

Small nucleic acid and/or antisense constructs of the methods and compositions presented herein can be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of cellular nucleic acids (*e.g.*, small RNAs, mRNA, and/or genomic DNA). Alternatively, the small nucleic acid molecules can produce RNA which encodes mRNA, miRNA, pre-miRNA, pri-miRNA, miRNA*, anti-miRNA, or a miRNA binding site, or a variant thereof. For example, selection of plasmids suitable for expressing the miRNAs, methods for inserting nucleic acid sequences into the plasmid, and methods of delivering the recombinant plasmid to the cells of interest are within the skill in the art. See, for example, Zeng *et al.* (2002) *Mol. Cell* 9:1327-1333; Tuschl (2002), *Nat. Biotechnol.* 20:446-448; Brummelkamp *et al.* (2002) *Science* 296:550-553; Miyagishi *et al.* (2002) *Nat. Biotechnol.* 20:497-500; Paddison *et al.* (2002) *Genes Dev.* 16:948-958; Lee *et al.* (2002) *Nat. Biotechnol.* 20:500-505; and Paul *et al.* (2002) *Nat. Biotechnol.* 20:505-508, the entire disclosures of which are herein incorporated by reference.

Alternatively, small nucleic acids and/or antisense constructs are oligonucleotide probes that are generated *ex vivo* and which, when introduced into the cell, results in hybridization with cellular nucleic acids. Such oligonucleotide probes are preferably modified oligonucleotides that are resistant to endogenous nucleases, *e.g.*, exonucleases and/or endonucleases, and are therefore stable *in vivo*. Exemplary nucleic acid molecules for use as small nucleic acids and/or antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patents 5,176,996; 5,264,564; and 5,256,775). Additionally, general approaches to constructing oligomers useful in antisense therapy have been reviewed, for example, by Van der Krol *et al.* (1988) *BioTechniques* 6:958-976; and Stein *et al.* (1988) *Cancer Res* 48:2659-2668.

Antisense approaches may involve the design of oligonucleotides (either DNA or RNA) that are complementary to cellular nucleic acids (*e.g.*, complementary to biomarkers listed in Table 1). Absolute complementarity is not required. In the case of double-

stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with a nucleic acid (*e.g.*, RNA) it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the mRNA, *e.g.*, the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have recently been shown to be effective at inhibiting translation of mRNAs as well (Wagner (1994) *Nature* 372:333). Therefore, oligonucleotides complementary to either the 5' or 3' untranslated, non-coding regions of genes could be used in an antisense approach to inhibit translation of endogenous mRNAs. Oligonucleotides complementary to the 5' untranslated region of the mRNA may include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could also be used in accordance with the methods and compositions presented herein. Whether designed to hybridize to the 5', 3' or coding region of cellular mRNAs, small nucleic acids and/or antisense nucleic acids should be at least six nucleotides in length, and can be less than about 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100, 50, 40, 30, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, or 10 nucleotides in length.

Regardless of the choice of target sequence, it is preferred that *in vitro* studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. In one embodiment these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. In another embodiment these studies compare levels of the target nucleic acid or protein with that of an internal control nucleic acid or protein. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

Small nucleic acids and/or antisense oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. Small nucleic acids and/or antisense oligonucleotides can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc., and may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134), hybridization-triggered cleavage agents. (See, *e.g.*, Krol *et al.* (1988) *BioTech.* 6:958-976) or intercalating agents. (See, *e.g.*, Zon (1988) *Pharm. Res.* 5:539-549). To this end, small nucleic acids and/or antisense oligonucleotides may be conjugated to another molecule, *e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

Small nucleic acids and/or antisense oligonucleotides may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxyethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Small nucleic acids and/or antisense oligonucleotides may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

In certain embodiments, a compound comprises an oligonucleotide (*e.g.*, a miRNA or miRNA encoding oligonucleotide) conjugated to one or more moieties which enhance the activity, cellular distribution or cellular uptake of the resulting oligonucleotide. In

certain such embodiments, the moiety is a cholesterol moiety (*e.g.*, antagomirs) or a lipid moiety or liposome conjugate. Additional moieties for conjugation include carbohydrates, phospholipids, biotin, phenazine, folate, phenanthridine, anthraquinone, acridine, fluoresceins, rhodamines, coumarins, and dyes. In certain embodiments, a conjugate group is attached directly to the oligonucleotide. In certain embodiments, a conjugate group is attached to the oligonucleotide by a linking moiety selected from amino, hydroxyl, carboxylic acid, thiol, unsaturations (*e.g.*, double or triple bonds), 8-amino-3,6-dioxaoctanoic acid (ADO), succinimidyl 4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC), 6-aminohexanoic acid (AHEx or AHA), substituted C1-C10 alkyl, substituted or unsubstituted C2-C10 alkenyl, and substituted or unsubstituted C2-C10 alkynyl. In certain such embodiments, a substituent group is selected from hydroxyl, amino, alkoxy, carboxy, benzyl, phenyl, nitro, thiol, thioalkoxy, halogen, alkyl, aryl, alkenyl and alkynyl.

In certain such embodiments, the compound comprises the oligonucleotide having one or more stabilizing groups that are attached to one or both termini of the oligonucleotide to enhance properties such as, for example, nuclease stability. Included in stabilizing groups are cap structures. These terminal modifications protect the oligonucleotide from exonuclease degradation, and can help in delivery and/or localization within a cell. The cap can be present at the 5'-terminus (5'-cap), or at the 3'-terminus (3'-cap), or can be present on both termini. Cap structures include, for example, inverted deoxy abasic caps.

Suitable cap structures include a 4',5'-methylene nucleotide, a 1-(beta-D-erythrofuransyl) nucleotide, a 4'-thio nucleotide, a carbocyclic nucleotide, a 1,5-anhydrohexitol nucleotide, an L-nucleotide, an alpha-nucleotide, a modified base nucleotide, a phosphorodithioate linkage, a threo-pentofuransyl nucleotide, an acyclic 3',4'-seco nucleotide, an acyclic 3,4-dihydroxybutyl nucleotide, an acyclic 3,5-dihydroxypentyl nucleotide, a 3'-3'-inverted nucleotide moiety, a 3'-3'-inverted abasic moiety, a 3'-2'-inverted nucleotide moiety, a 3'-2'-inverted abasic moiety, a 1,4-butanediol phosphate, a 3'-phosphoramidate, a hexylphosphate, an aminohexyl phosphate, a 3'-phosphate, a 3'-phosphorothioate, a phosphorodithioate, a bridging methylphosphonate moiety, and a non-bridging methylphosphonate moiety 5'-amino-alkyl phosphate, a 1,3-diamino-2-propyl phosphate, 3-aminopropyl phosphate, a 6-aminohexyl phosphate, a 1,2-aminododecyl phosphate, a hydroxypropyl phosphate, a 5'-5'-inverted nucleotide moiety, a

5'-5'-inverted abasic moiety, a 5'-phosphoramidate, a 5'-phosphorothioate, a 5'-amino, a bridging and/or non-bridging 5'-phosphoramidate, a phosphorothioate, and a 5'-mercapto moiety.

Small nucleic acids and/or antisense oligonucleotides can also contain a neutral peptide-like backbone. Such molecules are termed peptide nucleic acid (PNA)-oligomers and are described, *e.g.*, in Perry-O'Keefe et al. (1996) Proc. Natl. Acad. Sci. U.S.A. 93:14670 and in Eglom et al. (1993) Nature 365:566. One advantage of PNA oligomers is their capability to bind to complementary DNA essentially independently from the ionic strength of the medium due to the neutral backbone of the DNA. In yet another embodiment, small nucleic acids and/or antisense oligonucleotides comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In a further embodiment, small nucleic acids and/or antisense oligonucleotides are α -anomeric oligonucleotides. An α -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gautier et al. (1987) Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al. (1987) Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue et al. (1987) FEBS Lett. 215:327-330).

Small nucleic acids and/or antisense oligonucleotides of the methods and compositions presented herein may be synthesized by standard methods known in the art, *e.g.*, by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein et al. (1988) Nucl. Acids Res. 16:3209, methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al. (1988) Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), etc. For example, an isolated miRNA can be chemically synthesized or recombinantly produced using methods known in the art. In some instances, miRNA are chemically synthesized using appropriately protected ribonucleoside phosphoramidites and a conventional DNA/RNA synthesizer. Commercial suppliers of synthetic RNA molecules or synthesis reagents include, *e.g.*, Proligo (Hamburg, Germany), Dharmacon Research (Lafayette, Colo., USA), Pierce Chemical (part of Perbio Science, Rockford, Ill., USA), Glen Research

(Sterling, Va., USA), ChemGenes (Ashland, Mass., USA), Cruachem (Glasgow, UK), and Exiqon (Vedbaek, Denmark).

Small nucleic acids and/or antisense oligonucleotides can be delivered to cells *in vivo*. A number of methods have been developed for delivering small nucleic acids and/or antisense oligonucleotides DNA or RNA to cells; *e.g.*, antisense molecules can be injected
5 directly into the tissue site, or modified antisense molecules, designed to target the desired cells (*e.g.*, antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systematically.

In one embodiment, small nucleic acids and/or antisense oligonucleotides may
10 comprise or be generated from double stranded small interfering RNAs (siRNAs), in which sequences fully complementary to cellular nucleic acids (*e.g.* mRNAs) sequences mediate degradation or in which sequences incompletely complementary to cellular nucleic acids (*e.g.*, mRNAs) mediate translational repression when expressed within cells, or piwiRNAs. In another embodiment, double stranded siRNAs can be processed into single stranded
15 antisense RNAs that bind single stranded cellular RNAs (*e.g.*, microRNAs) and inhibit their expression. RNA interference (RNAi) is the process of sequence-specific, post-transcriptional gene silencing in animals and plants, initiated by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene. *in vivo*, long dsRNA is cleaved by ribonuclease III to generate 21- and 22-nucleotide siRNAs. It has been shown
20 that 21-nucleotide siRNA duplexes specifically suppress expression of endogenous and heterologous genes in different mammalian cell lines, including human embryonic kidney (293) and HeLa cells (Elbashir et al. (2001) *Nature* 411:494-498). Accordingly, translation of a gene in a cell can be inhibited by contacting the cell with short double stranded RNAs having a length of about 15 to 30 nucleotides or of about 18 to 21 nucleotides or of about
25 19 to 21 nucleotides. Alternatively, a vector encoding for such siRNAs or short hairpin RNAs (shRNAs) that are metabolized into siRNAs can be introduced into a target cell (see, *e.g.*, McManus *et al.* (2002) *RNA* 8:842; Xia *et al.* (2002) *Nat. Biotechnol.* 20:1006; and Brummelkamp *et al.* (2002) *Science* 296:550). Vectors that can be used are commercially available, *e.g.*, from OligoEngine under the name pSuper RNAi SystemTM.

30 Ribozyme molecules designed to catalytically cleave cellular mRNA transcripts can also be used to prevent translation of cellular mRNAs and expression of cellular polypeptides, or both (See, *e.g.*, PCT International Publication WO90/11364, published October 4, 1990; Sarver et al. (1990) *Science* 247:1222-1225 and U.S. Patent No.

5,093,246). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy cellular mRNAs, the use of hammerhead ribozymes is preferred.

Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well-known in the art and is described more fully in Haseloff and Gerlach (1988) *Nature* 334:585-591. The ribozyme may be engineered so that the cleavage recognition site is located near the 5' end of cellular mRNAs; *i.e.*, to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

The ribozymes of the methods presented herein also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in *Tetrahymena thermophila* (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug *et al.* (1984) *Science* 224:574-578; Zaug *et al.* (1986) *Science* 231:470-475; Zaug *et al.* (1986) *Nature* 324:429-433; WO 88/04300; and Been *et al.* (1986) *Cell* 47:207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The methods and compositions presented herein encompasses those Cech-type ribozymes which target eight base-pair active site sequences that are present in cellular genes.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (*e.g.*, for improved stability, targeting, etc.). A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous cellular messages and inhibit translation. Because ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Nucleic acid molecules to be used in triple helix formation for the inhibition of transcription of cellular genes are preferably single stranded and composed of deoxyribonucleotides. The base composition of these oligonucleotides should promote triple helix formation via Hoogsteen base pairing rules, which generally require sizable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which will result in TAT and CGC triplets

across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine-rich, for example, containing a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in CGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called “switchback” nucleic acid molecule. Switchback molecules are synthesized in an alternating 5’-3’, 3’-5’ manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizable stretch of either purines or pyrimidines to be present on one strand of a duplex.

Small nucleic acids (*e.g.*, miRNAs, pre-miRNAs, pri-miRNAs, miRNA*, anti-miRNA, or a miRNA binding site, or a variant thereof), antisense oligonucleotides, ribozymes, and triple helix molecules of the methods and compositions presented herein may be prepared by any method known in the art for the synthesis of DNA and RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well-known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

Moreover, various well-known modifications to nucleic acid molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to the addition of flanking sequences of ribonucleotides or deoxyribonucleotides to the 5’ and/or 3’ ends of the molecule or the use of phosphorothioate or 2’ O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone. One of skill in the art will readily understand that polypeptides, small nucleic acids, and antisense oligonucleotides can be further linked to another peptide or polypeptide (*e.g.*, a heterologous peptide), *e.g.*, that serves as a means of

protein detection. Non-limiting examples of label peptide or polypeptide moieties useful for detection in the invention include, without limitation, suitable enzymes such as horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; epitope tags, such as FLAG, MYC, HA, or HIS tags; fluorophores such as green fluorescent protein; dyes; radioisotopes; digoxigenin; biotin; antibodies; polymers; as well as others known in the art, for example, in Principles of Fluorescence Spectroscopy, Joseph R. Lakowicz (Editor), Plenum Pub Corp, 2nd edition (July 1999).

The modulatory agents described herein (*e.g.*, antibodies, small molecules, peptides, fusion proteins, or small nucleic acids) can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The compositions may contain a single such molecule or agent or any combination of agents described herein. “Single active agents” described herein can be combined with other pharmacologically active compounds (“second active agents”) known in the art according to the methods and compositions provided herein.

The production and use of biomarker nucleic acid and/or biomarker polypeptide molecules described herein can be facilitated by using standard recombinant techniques. In some embodiments, such techniques use vectors, preferably expression vectors, containing a nucleic acid encoding a biomarker polypeptide or a portion of such a polypeptide. As used herein, the term “vector” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a “plasmid”, which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors, namely expression vectors, are capable of directing the expression of genes to which they are operably linked. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids (vectors). However, the present invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the present invention comprise a nucleic acid of the present invention in a form suitable for expression of the nucleic acid in a host cell. This means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operably linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, “operably linked” is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term “regulatory sequence” is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, *Methods in Enzymology: Gene Expression Technology* vol.185, Academic Press, San Diego, CA (1991). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the present invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein.

The recombinant expression vectors for use in the present invention can be designed for expression of a polypeptide corresponding to a marker of the present invention in prokaryotic (*e.g.*, *E. coli*) or eukaryotic cells (*e.g.*, insect cells {using baculovirus expression vectors}, yeast cells or mammalian cells). Suitable host cells are discussed further in Goeddel, *supra*. Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2)

to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988, *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, 1988, *Gene* 69:301-315) and pET 11d (Studier *et al.*, p. 60-89, In *Gene Expression Technology: Methods in Enzymology* vol.185, Academic Press, San Diego, CA, 1991). Target biomarker nucleic acid expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target biomarker nucleic acid expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a co-expressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21 (DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacterium with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, p. 119-128, In *Gene Expression Technology: Methods in Enzymology* vol. 185, Academic Press, San Diego, CA, 1990. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, 1992, *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the present invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the expression vector is a yeast expression vector.

Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari *et al.*, 1987, *EMBO J.* 6:229-234), pMFa (Kurjan and Herskowitz, 1982, *Cell* 30:933-943), pJRY88 (Schultz *et al.*, 1987, *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and pPicZ (Invitrogen Corp, San Diego, CA).

Alternatively, the expression vector is a baculovirus expression vector. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith *et al.*, 1983, *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers, 1989, *Virology* 170:31-39).

5 In yet another embodiment, a nucleic acid of the present invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, 1987, *Nature* 329:840) and pMT2PC (Kaufman *et al.*, 1987, *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly
10 used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook *et al.*, *supra*.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-
15 specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.*, 1987, *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton, 1988, *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989, *EMBO J.*
20 8:729-733) and immunoglobulins (Banerji *et al.*, 1983, *Cell* 33:729-740; Queen and Baltimore, 1983, *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle, 1989, *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.*, 1985, *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European
25 Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss, 1990, *Science* 249:374-379) and the α -fetoprotein promoter (Camper and Tilghman, 1989, *Genes Dev.* 3:537-546).

The present invention further provides a recombinant expression vector comprising
30 a DNA molecule cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operably linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to the mRNA encoding a polypeptide of the present invention. Regulatory sequences

operably linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue-specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid, or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes (see Weintraub *et al.*, 1986, *Trends in Genetics*, Vol. 1(1)).

Another aspect of the present invention pertains to host cells into which a recombinant expression vector of the present invention has been introduced. The terms “host cell” and “recombinant host cell” are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic (*e.g.*, *E. coli*) or eukaryotic cell (*e.g.*, insect cells, yeast or mammalian cells).

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms “transformation” and “transfection” are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*supra*), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, for resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Cells stably transfected with the introduced nucleic acid can be identified by

drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

V. Analyzing Biomarker Nucleic Acids and Polypeptides

5 Biomarker nucleic acids and/or biomarker polypeptides can be analyzed according to the methods described herein and techniques known to the skilled artisan to identify such genetic or expression alterations useful for the present invention including, but not limited to, 1) an alteration in the level of a biomarker transcript or polypeptide, 2) a deletion or addition of one or more nucleotides from a biomarker gene, 4) a substitution of one or more
10 nucleotides of a biomarker gene, 5) aberrant modification of a biomarker gene, such as an expression regulatory region, and the like.

a. Methods for Detection of Copy Number

Methods of evaluating the copy number of a biomarker nucleic acid are well-known to those of skill in the art. The presence or absence of chromosomal gain or loss can be
15 evaluated simply by a determination of copy number of the regions or markers identified herein.

In one embodiment, a biological sample is tested for the presence of copy number changes in genomic loci containing the genomic marker. A copy number of at least 3, 4, 5, 6, 7, 8, 9, or 10 is predictive of poorer outcome of inhibitor(s) of one or more biomarkers
20 listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), and immunotherapy combination treatments.

Methods of evaluating the copy number of a biomarker locus include, but are not limited to, hybridization-based assays. Hybridization-based assays include, but are not
25 limited to, traditional “direct probe” methods, such as Southern blots, *in situ* hybridization (*e.g.*, FISH and FISH plus SKY) methods, and “comparative probe” methods, such as comparative genomic hybridization (CGH), *e.g.*, cDNA-based or oligonucleotide-based CGH. The methods can be used in a wide variety of formats including, but not limited to, substrate (*e.g.* membrane or glass) bound methods or array-based approaches.

30 In one embodiment, evaluating the biomarker gene copy number in a sample involves a Southern Blot. In a Southern Blot, the genomic DNA (typically fragmented and separated on an electrophoretic gel) is hybridized to a probe specific for the target region. Comparison of the intensity of the hybridization signal from the probe for the target region

with control probe signal from analysis of normal genomic DNA (*e.g.*, a non-amplified portion of the same or related cell, tissue, organ, *etc.*) provides an estimate of the relative copy number of the target nucleic acid. Alternatively, a Northern blot may be utilized for evaluating the copy number of encoding nucleic acid in a sample. In a Northern blot, mRNA is hybridized to a probe specific for the target region. Comparison of the intensity of the hybridization signal from the probe for the target region with control probe signal from analysis of normal RNA (*e.g.*, a non-amplified portion of the same or related cell, tissue, organ, *etc.*) provides an estimate of the relative copy number of the target nucleic acid. Alternatively, other methods well-known in the art to detect RNA can be used, such that higher or lower expression relative to an appropriate control (*e.g.*, a non-amplified portion of the same or related cell tissue, organ, *etc.*) provides an estimate of the relative copy number of the target nucleic acid.

An alternative means for determining genomic copy number is *in situ* hybridization (*e.g.*, Angerer (1987) *Meth. Enzymol* 152: 649). Generally, *in situ* hybridization comprises the following steps: (1) fixation of tissue or biological structure to be analyzed; (2) prehybridization treatment of the biological structure to increase accessibility of target DNA, and to reduce nonspecific binding; (3) hybridization of the mixture of nucleic acids to the nucleic acid in the biological structure or tissue; (4) post-hybridization washes to remove nucleic acid fragments not bound in the hybridization and (5) detection of the hybridized nucleic acid fragments. The reagent used in each of these steps and the conditions for use vary depending on the particular application. In a typical *in situ* hybridization assay, cells are fixed to a solid support, typically a glass slide. If a nucleic acid is to be probed, the cells are typically denatured with heat or alkali. The cells are then contacted with a hybridization solution at a moderate temperature to permit annealing of labeled probes specific to the nucleic acid sequence encoding the protein. The targets (*e.g.*, cells) are then typically washed at a predetermined stringency or at an increasing stringency until an appropriate signal to noise ratio is obtained. The probes are typically labeled, *e.g.*, with radioisotopes or fluorescent reporters. In one embodiment, probes are sufficiently long so as to specifically hybridize with the target nucleic acid(s) under stringent conditions. Probes generally range in length from about 200 bases to about 1000 bases. In some applications it is necessary to block the hybridization capacity of repetitive sequences. Thus, in some embodiments, tRNA, human genomic DNA, or Cot-I DNA is used to block non-specific hybridization.

An alternative means for determining genomic copy number is comparative genomic hybridization. In general, genomic DNA is isolated from normal reference cells, as well as from test cells (*e.g.*, tumor cells) and amplified, if necessary. The two nucleic acids are differentially labeled and then hybridized *in situ* to metaphase chromosomes of a reference cell. The repetitive sequences in both the reference and test DNAs are either removed or their hybridization capacity is reduced by some means, for example by prehybridization with appropriate blocking nucleic acids and/or including such blocking nucleic acid sequences for said repetitive sequences during said hybridization. The bound, labeled DNA sequences are then rendered in a visualizable form, if necessary.

Chromosomal regions in the test cells which are at increased or decreased copy number can be identified by detecting regions where the ratio of signal from the two DNAs is altered. For example, those regions that have decreased in copy number in the test cells will show relatively lower signal from the test DNA than the reference compared to other regions of the genome. Regions that have been increased in copy number in the test cells will show relatively higher signal from the test DNA. Where there are chromosomal deletions or multiplications, differences in the ratio of the signals from the two labels will be detected and the ratio will provide a measure of the copy number. In another embodiment of CGH, array CGH (aCGH), the immobilized chromosome element is replaced with a collection of solid support bound target nucleic acids on an array, allowing for a large or complete percentage of the genome to be represented in the collection of solid support bound targets. Target nucleic acids may comprise cDNAs, genomic DNAs, oligonucleotides (*e.g.*, to detect single nucleotide polymorphisms) and the like. Array-based CGH may also be performed with single-color labeling (as opposed to labeling the control and the possible tumor sample with two different dyes and mixing them prior to hybridization, which will yield a ratio due to competitive hybridization of probes on the arrays). In single color CGH, the control is labeled and hybridized to one array and absolute signals are read, and the possible tumor sample is labeled and hybridized to a second array (with identical content) and absolute signals are read. Copy number difference is calculated based on absolute signals from the two arrays. Methods of preparing immobilized chromosomes or arrays and performing comparative genomic hybridization are well-known in the art (see, *e.g.*, U.S. Pat. Nos. 6,335,167; 6,197,501; 5,830,645; and 5,665,549 and Albertson (1984) *EMBO J.* 3: 1227-1234; Pinkel (1988) *Proc. Natl. Acad. Sci. USA* 85: 9138-9142; EPO Pub. No. 430,402; *Methods in Molecular Biology*, Vol. 33: *In situ* Hybridization Protocols,

Choo, ed., Humana Press, Totowa, N.J. (1994), *etc.*) In another embodiment, the hybridization protocol of Pinkel, *et al.* (1998) *Nature Genetics* 20: 207-211, or of Kallioniemi (1992) *Proc. Natl Acad Sci USA* 89:5321-5325 (1992) is used.

In still another embodiment, amplification-based assays can be used to measure copy number. In such amplification-based assays, the nucleic acid sequences act as a template in an amplification reaction (*e.g.*, Polymerase Chain Reaction (PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls, *e.g.* healthy tissue, provides a measure of the copy number.

Methods of “quantitative” amplification are well-known to those of skill in the art. For example, quantitative PCR involves simultaneously co-amplifying a known quantity of a control sequence using the same primers. This provides an internal standard that may be used to calibrate the PCR reaction. Detailed protocols for quantitative PCR are provided in Innis, *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.). Measurement of DNA copy number at microsatellite loci using quantitative PCR analysis is described in Ginzinger, *et al.* (2000) *Cancer Research* 60:5405-5409. The known nucleic acid sequence for the genes is sufficient to enable one of skill in the art to routinely select primers to amplify any portion of the gene. Fluorogenic quantitative PCR may also be used in the methods of the present invention. In fluorogenic quantitative PCR, quantitation is based on amount of fluorescence signals, *e.g.*, TaqMan and SYBR green.

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (see Wu and Wallace (1989) *Genomics* 4: 560, Landegren, *et al.* (1988) *Science* 241:1077, and Barringer *et al.* (1990) *Gene* 89: 117), transcription amplification (Kwoh, *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86: 1173), self-sustained sequence replication (Guatelli, *et al.* (1990) *Proc. Nat. Acad. Sci. USA* 87: 1874), dot PCR, and linker adapter PCR, *etc.*

Loss of heterozygosity (LOH) and major copy proportion (MCP) mapping (Wang, Z.C., *et al.* (2004) *Cancer Res* 64(1):64-71; Seymour, A. B., *et al.* (1994) *Cancer Res* 54, 2761-4; Hahn, S. A., *et al.* (1995) *Cancer Res* 55, 4670-5; Kimura, M., *et al.* (1996) *Genes Chromosomes Cancer* 17, 88-93; Li *et al.*, (2008) *MBC Bioinform.* 9, 204-219) may also be used to identify regions of amplification or deletion.

b. Methods for Detection of Biomarker Nucleic Acid Expression

Biomarker expression may be assessed by any of a wide variety of well-known

methods for detecting expression of a transcribed molecule or protein. Non-limiting examples of such methods include immunological methods for detection of secreted, cell-surface, cytoplasmic, or nuclear proteins, protein purification methods, protein function or activity assays, nucleic acid hybridization methods, nucleic acid reverse transcription
5 methods, and nucleic acid amplification methods.

In preferred embodiments, activity of a particular gene is characterized by a measure of gene transcript (*e.g.* mRNA), by a measure of the quantity of translated protein, or by a measure of gene product activity. Marker expression can be monitored in a variety of ways, including by detecting mRNA levels, protein levels, or protein activity, any of
10 which can be measured using standard techniques. Detection can involve quantification of the level of gene expression (*e.g.*, genomic DNA, cDNA, mRNA, protein, or enzyme activity), or, alternatively, can be a qualitative assessment of the level of gene expression, in particular in comparison with a control level. The type of level being detected will be clear from the context.

In another embodiment, detecting or determining expression levels of a biomarker and functionally similar homologs thereof, including a fragment or genetic alteration thereof (*e.g.*, in regulatory or promoter regions thereof) comprises detecting or determining
15 RNA levels for the marker of interest. In one embodiment, one or more cells from the subject to be tested are obtained and RNA is isolated from the cells. In a preferred
20 embodiment, a sample of breast tissue cells is obtained from the subject.

In one embodiment, RNA is obtained from a single cell. For example, a cell can be isolated from a tissue sample by laser capture microdissection (LCM). Using this technique, a cell can be isolated from a tissue section, including a stained tissue section, thereby assuring that the desired cell is isolated (*see, e.g.*, Bonner *et al.* (1997) Science 278:
25 1481; Emmert-Buck *et al.* (1996) Science 274:998; Fend *et al.* (1999) Am. J. Path. 154: 61 and Murakami *et al.* (2000) Kidney Int. 58:1346). For example, Murakami *et al.*, *supra*, describe isolation of a cell from a previously immunostained tissue section.

It is also possible to obtain cells from a subject and culture the cells *in vitro*, such as to obtain a larger population of cells from which RNA can be extracted. Methods for
30 establishing cultures of non-transformed cells, *i.e.*, primary cell cultures, are known in the art.

When isolating RNA from tissue samples or cells from individuals, it may be important to prevent any further changes in gene expression after the tissue or cells has

been removed from the subject. Changes in expression levels are known to change rapidly following perturbations, *e.g.*, heat shock or activation with lipopolysaccharide (LPS) or other reagents. In addition, the RNA in the tissue and cells may quickly become degraded. Accordingly, in a preferred embodiment, the tissue or cells obtained from a subject is snap
5 frozen as soon as possible.

RNA can be extracted from the tissue sample by a variety of methods, *e.g.*, the guanidium thiocyanate lysis followed by CsCl centrifugation (Chirgwin *et al.*, 1979, *Biochemistry* 18:5294-5299). RNA from single cells can be obtained as described in methods for preparing cDNA libraries from single cells, such as those described in Dulac,
10 C. (1998) *Curr. Top. Dev. Biol.* 36, 245 and Jena *et al.* (1996) *J. Immunol. Methods* 190:199. Care to avoid RNA degradation must be taken, *e.g.*, by inclusion of RNAsin.

The RNA sample can then be enriched in particular species. In one embodiment, poly(A)+ RNA is isolated from the RNA sample. In general, such purification takes advantage of the poly-A tails on mRNA. In particular and as noted above, poly-T
15 oligonucleotides may be immobilized within on a solid support to serve as affinity ligands for mRNA. Kits for this purpose are commercially available, *e.g.*, the MessageMaker kit (Life Technologies, Grand Island, NY).

In a preferred embodiment, the RNA population is enriched in marker sequences. Enrichment can be undertaken, *e.g.*, by primer-specific cDNA synthesis, or multiple rounds
20 of linear amplification based on cDNA synthesis and template-directed *in vitro* transcription (*see, e.g.*, Wang *et al.* (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86: 9717; Dulac *et al.*, *supra*, and Jena *et al.*, *supra*).

The population of RNA, enriched or not in particular species or sequences, can further be amplified. As defined herein, an “amplification process” is designed to
25 strengthen, increase, or augment a molecule within the RNA. For example, where RNA is mRNA, an amplification process such as RT-PCR can be utilized to amplify the mRNA, such that a signal is detectable or detection is enhanced. Such an amplification process is beneficial particularly when the biological, tissue, or tumor sample is of a small size or volume.

30 Various amplification and detection methods can be used. For example, it is within the scope of the present invention to reverse transcribe mRNA into cDNA followed by polymerase chain reaction (RT-PCR); or, to use a single enzyme for both steps as described in U.S. Pat. No. 5,322,770, or reverse transcribe mRNA into cDNA followed by symmetric

gap ligase chain reaction (RT-AGLCR) as described by R. L. Marshall, *et al.*, PCR Methods and Applications 4: 80-84 (1994). Real time PCR may also be used.

Other known amplification methods which can be utilized herein include but are not limited to the so-called "NASBA" or "3SR" technique described in PNAS USA 87: 1874-1878 (1990) and also described in Nature 350 (No. 6313): 91-92 (1991); Q-beta amplification as described in published European Patent Application (EPA) No. 4544610; strand displacement amplification (as described in G. T. Walker *et al.*, Clin. Chem. 42: 9-13 (1996) and European Patent Application No. 684315; target mediated amplification, as described by PCT Publication WO9322461; PCR; ligase chain reaction (LCR) (*see, e.g.*, Wu and Wallace, Genomics 4, 560 (1989), Landegren *et al.*, Science 241, 1077 (1988)); self-sustained sequence replication (SSR) (*see, e.g.*, Guatelli *et al.*, Proc. Nat. Acad. Sci. USA, 87, 1874 (1990)); and transcription amplification (*see, e.g.*, Kwoh *et al.*, Proc. Natl. Acad. Sci. USA 86, 1173 (1989)).

Many techniques are known in the state of the art for determining absolute and relative levels of gene expression, commonly used techniques suitable for use in the present invention include Northern analysis, RNase protection assays (RPA), microarrays and PCR-based techniques, such as quantitative PCR and differential display PCR. For example, Northern blotting involves running a preparation of RNA on a denaturing agarose gel, and transferring it to a suitable support, such as activated cellulose, nitrocellulose or glass or nylon membranes. Radiolabeled cDNA or RNA is then hybridized to the preparation, washed and analyzed by autoradiography.

In situ hybridization visualization may also be employed, wherein a radioactively labeled antisense RNA probe is hybridized with a thin section of a biopsy sample, washed, cleaved with RNase and exposed to a sensitive emulsion for autoradiography. The samples may be stained with hematoxylin to demonstrate the histological composition of the sample, and dark field imaging with a suitable light filter shows the developed emulsion. Non-radioactive labels such as digoxigenin may also be used.

Alternatively, mRNA expression can be detected on a DNA array, chip or a microarray. Labeled nucleic acids of a test sample obtained from a subject may be hybridized to a solid surface comprising biomarker DNA. Positive hybridization signal is obtained with the sample containing biomarker transcripts. Methods of preparing DNA arrays and their use are well-known in the art (*see, e.g.*, U.S. Pat. Nos: 6,618,679; 6,379,897; 6,664,377; 6,451,536; 5,482,257; U.S. 20030157485 and Schena *et al.* (1995)

Science 20, 467-470; Gerhold *et al.* (1999) *Trends In Biochem. Sci.* 24, 168-173; and Lennon *et al.* (2000) *Drug Discovery Today* 5, 59-65, which are herein incorporated by reference in their entirety). Serial Analysis of Gene Expression (SAGE) can also be performed (See for example U.S. Patent Application 20030215858).

5 To monitor mRNA levels, for example, mRNA is extracted from the biological sample to be tested, reverse transcribed, and fluorescently-labeled cDNA probes are generated. The microarrays capable of hybridizing to marker cDNA are then probed with the labeled cDNA probes, the slides scanned and fluorescence intensity measured. This intensity correlates with the hybridization intensity and expression levels.

10 Types of probes that can be used in the methods described herein include cDNA, riboprobes, synthetic oligonucleotides and genomic probes. The type of probe used will generally be dictated by the particular situation, such as riboprobes for *in situ* hybridization, and cDNA for Northern blotting, for example. In one embodiment, the probe is directed to nucleotide regions unique to the RNA. The probes may be as short as is required to
15 differentially recognize marker mRNA transcripts, and may be as short as, for example, 15 bases; however, probes of at least 17, 18, 19 or 20 or more bases can be used. In one embodiment, the primers and probes hybridize specifically under stringent conditions to a DNA fragment having the nucleotide sequence corresponding to the marker. As herein used, the term “stringent conditions” means hybridization will occur only if there is at least
20 95% identity in nucleotide sequences. In another embodiment, hybridization under “stringent conditions” occurs when there is at least 97% identity between the sequences.

 The form of labeling of the probes may be any that is appropriate, such as the use of radioisotopes, for example, ³²P and ³⁵S. Labeling with radioisotopes may be achieved, whether the probe is synthesized chemically or biologically, by the use of suitably labeled
25 bases.

 In one embodiment, the biological sample contains polypeptide molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject.

 In another embodiment, the methods further involve obtaining a control biological
30 sample from a control subject, contacting the control sample with a compound or agent capable of detecting marker polypeptide, mRNA, genomic DNA, or fragments thereof, such that the presence of the marker polypeptide, mRNA, genomic DNA, or fragments thereof, is detected in the biological sample, and comparing the presence of the marker polypeptide,

mRNA, genomic DNA, or fragments thereof, in the control sample with the presence of the marker polypeptide, mRNA, genomic DNA, or fragments thereof in the test sample.

c. Methods for Detection of Biomarker Protein Expression

The activity or level of a biomarker protein can be detected and/or quantified by
5 detecting or quantifying the expressed polypeptide. The polypeptide can be detected and quantified by any of a number of means well-known to those of skill in the art. Aberrant levels of polypeptide expression of the polypeptides encoded by a biomarker nucleic acid and functionally similar homologs thereof, including a fragment or genetic alteration thereof (*e.g.*, in regulatory or promoter regions thereof) are associated with the likelihood of
10 response of a cancer to inhibitor(s) of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), and immunotherapy combination treatments. Any method known in the art for detecting polypeptides can be used. Such methods include, but are not limited to, immunodiffusion, immunoelectrophoresis,
15 radioimmunoassay (RIA), enzyme-linked immunosorbent assays (ELISAs), immunofluorescent assays, Western blotting, binder-ligand assays, immunohistochemical techniques, agglutination, complement assays, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography, and the like (*e.g.*, Basic and Clinical Immunology, Sites and Terr, eds., Appleton and Lange, Norwalk,
20 Conn. pp 217-262, 1991 which is incorporated by reference). Preferred are binder-ligand immunoassay methods including reacting antibodies with an epitope or epitopes and competitively displacing a labeled polypeptide or derivative thereof.

For example, ELISA and RIA procedures may be conducted such that a desired biomarker protein standard is labeled (with a radioisotope such as ¹²⁵I or ³⁵S, or an
25 assayable enzyme, such as horseradish peroxidase or alkaline phosphatase), and, together with the unlabeled sample, brought into contact with the corresponding antibody, whereon a second antibody is used to bind the first, and radioactivity or the immobilized enzyme assayed (competitive assay). Alternatively, the biomarker protein in the sample is allowed to react with the corresponding immobilized antibody, radioisotope- or enzyme-labeled
30 anti-biomarker protein antibody is allowed to react with the system, and radioactivity or the enzyme assayed (ELISA-sandwich assay). Other conventional methods may also be employed as suitable.

The above techniques may be conducted essentially as a “one-step” or “two-step”

assay. A “one-step” assay involves contacting antigen with immobilized antibody and, without washing, contacting the mixture with labeled antibody. A “two-step” assay involves washing before contacting, the mixture with labeled antibody. Other conventional methods may also be employed as suitable.

5 In one embodiment, a method for measuring biomarker protein levels comprises the steps of: contacting a biological specimen with an antibody or variant (*e.g.*, fragment) thereof which selectively binds the biomarker protein, and detecting whether said antibody or variant thereof is bound to said sample and thereby measuring the levels of the biomarker protein.

10 Enzymatic and radiolabeling of biomarker protein and/or the antibodies may be effected by conventional means. Such means will generally include covalent linking of the enzyme to the antigen or the antibody in question, such as by glutaraldehyde, specifically so as not to adversely affect the activity of the enzyme, by which is meant that the enzyme must still be capable of interacting with its substrate, although it is not necessary for all of
15 the enzyme to be active, provided that enough remains active to permit the assay to be effected. Indeed, some techniques for binding enzyme are non-specific (such as using formaldehyde), and will only yield a proportion of active enzyme.

 It is usually desirable to immobilize one component of the assay system on a support, thereby allowing other components of the system to be brought into contact with
20 the component and readily removed without laborious and time-consuming labor. It is possible for a second phase to be immobilized away from the first, but one phase is usually sufficient.

 It is possible to immobilize the enzyme itself on a support, but if solid-phase enzyme is required, then this is generally best achieved by binding to antibody and affixing
25 the antibody to a support, models and systems for which are well-known in the art. Simple polyethylene may provide a suitable support.

 Enzymes employable for labeling are not particularly limited, but may be selected from the members of the oxidase group, for example. These catalyze production of hydrogen peroxide by reaction with their substrates, and glucose oxidase is often used for
30 its good stability, ease of availability and cheapness, as well as the ready availability of its substrate (glucose). Activity of the oxidase may be assayed by measuring the concentration of hydrogen peroxide formed after reaction of the enzyme-labeled antibody with the substrate under controlled conditions well-known in the art.

Other techniques may be used to detect biomarker protein according to a practitioner's preference based upon the present disclosure. One such technique is Western blotting (Towbin et al., Proc. Nat. Acad. Sci. 76:4350 (1979)), wherein a suitably treated sample is run on an SDS-PAGE gel before being transferred to a solid support, such as a nitrocellulose filter. Anti-biomarker protein antibodies (unlabeled) are then brought into contact with the support and assayed by a secondary immunological reagent, such as labeled protein A or anti-immunoglobulin (suitable labels including ^{125}I , horseradish peroxidase and alkaline phosphatase). Chromatographic detection may also be used.

Immunohistochemistry may be used to detect expression of biomarker protein, e.g., in a biopsy sample. A suitable antibody is brought into contact with, for example, a thin layer of cells, washed, and then contacted with a second, labeled antibody. Labeling may be by fluorescent markers, enzymes, such as peroxidase, avidin, or radiolabeling. The assay is scored visually, using microscopy.

Anti-biomarker protein antibodies, such as intrabodies, may also be used for imaging purposes, for example, to detect the presence of biomarker protein in cells and tissues of a subject. Suitable labels include radioisotopes, iodine (^{125}I , ^{121}I), carbon (^{14}C), sulphur (^{35}S), tritium (^3H), indium (^{112}In), and technetium ($^{99\text{m}}\text{Tc}$), fluorescent labels, such as fluorescein and rhodamine, and biotin.

For *in vivo* imaging purposes, antibodies are not detectable, as such, from outside the body, and so must be labeled, or otherwise modified, to permit detection. Markers for this purpose may be any that do not substantially interfere with the antibody binding, but which allow external detection. Suitable markers may include those that may be detected by X-radiography, NMR or MRI. For X-radiographic techniques, suitable markers include any radioisotope that emits detectable radiation but that is not overtly harmful to the subject, such as barium or cesium, for example. Suitable markers for NMR and MRI generally include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by suitable labeling of nutrients for the relevant hybridoma, for example.

The size of the subject, and the imaging system used, will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of technetium-99. The labeled antibody or antibody fragment will then

preferentially accumulate at the location of cells which contain biomarker protein. The labeled antibody or antibody fragment can then be detected using known techniques.

Antibodies that may be used to detect biomarker protein include any antibody, whether natural or synthetic, full length or a fragment thereof, monoclonal or polyclonal, that binds sufficiently strongly and specifically to the biomarker protein to be detected. An antibody may have a K_d of at most about $10^{-6}M$, $10^{-7}M$, $10^{-8}M$, $10^{-9}M$, $10^{-10}M$, $10^{-11}M$, $10^{-12}M$. The phrase “specifically binds” refers to binding of, for example, an antibody to an epitope or antigen or antigenic determinant in such a manner that binding can be displaced or competed with a second preparation of identical or similar epitope, antigen or antigenic determinant. An antibody may bind preferentially to the biomarker protein relative to other proteins, such as related proteins.

Antibodies are commercially available or may be prepared according to methods known in the art.

Antibodies and derivatives thereof that may be used encompass polyclonal or monoclonal antibodies, chimeric, human, humanized, primatized (CDR-grafted), veneered or single-chain antibodies as well as functional fragments, *i.e.*, biomarker protein binding fragments, of antibodies. For example, antibody fragments capable of binding to a biomarker protein or portions thereof, including, but not limited to, Fv, Fab, Fab' and F(ab')₂ fragments can be used. Such fragments can be produced by enzymatic cleavage or by recombinant techniques. For example, papain or pepsin cleavage can generate Fab or F(ab')₂ fragments, respectively. Other proteases with the requisite substrate specificity can also be used to generate Fab or F(ab')₂ fragments. Antibodies can also be produced in a variety of truncated forms using antibody genes in which one or more stop codons have been introduced upstream of the natural stop site. For example, a chimeric gene encoding a F(ab')₂ heavy chain portion can be designed to include DNA sequences encoding the CH₁ domain and hinge region of the heavy chain.

Synthetic and engineered antibodies are described in, *e.g.*, Cabilly *et al.*, U.S. Pat. No. 4,816,567 Cabilly *et al.*, European Patent No. 0,125,023 B1; Boss *et al.*, U.S. Pat. No. 4,816,397; Boss *et al.*, European Patent No. 0,120,694 B1; Neuberger, M. S. *et al.*, WO 86/01533; Neuberger, M. S. *et al.*, European Patent No. 0,194,276 B1; Winter, U.S. Pat. No. 5,225,539; Winter, European Patent No. 0,239,400 B1; Queen *et al.*, European Patent No. 0451216 B1; and Padlan, E. A. *et al.*, EP 0519596 A1. See also, Newman, R. *et al.*, BioTechnology, 10: 1455-1460 (1992), regarding primatized antibody, and Ladner *et al.*,

U.S. Pat. No. 4,946,778 and Bird, R. E. *et al.*, Science, 242: 423-426 (1988)) regarding single-chain antibodies. Antibodies produced from a library, *e.g.*, phage display library, may also be used.

5 In some embodiments, agents that specifically bind to a biomarker protein other than antibodies are used, such as peptides. Peptides that specifically bind to a biomarker protein can be identified by any means known in the art. For example, specific peptide binders of a biomarker protein can be screened for using peptide phage display libraries.

d. Methods for Detection of Biomarker Structural Alterations

10 The following illustrative methods can be used to identify the presence of a structural alteration in a biomarker nucleic acid and/or biomarker polypeptide molecule in order to, for example, identify the one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), or other biomarkers used in the immunotherapies described herein that are overexpressed, overfunctional, and the like.

15 In certain embodiments, detection of the alteration involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, *e.g.*, U.S. Pat. Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) Science 241:1077-1080; and Nakazawa *et al.* (1994) Proc. Natl. Acad. Sci. USA 91:360-364), the latter of which can be particularly
20 useful for detecting point mutations in a biomarker nucleic acid such as a biomarker gene (see Abravaya *et al.* (1995) Nucleic Acids Res. 23:675-682). This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a biomarker gene under conditions such that
25 hybridization and amplification of the biomarker gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

30 Alternative amplification methods include: self-sustained sequence replication (Guatelli, J. C. *et al.* (1990) Proc. Natl. Acad. Sci. USA 87:1874-1878), transcriptional amplification system (Kwoh, D. Y. *et al.* (1989) Proc. Natl. Acad. Sci. USA 86:1173-1177), Q-Beta Replicase (Lizardi, P. M. *et al.* (1988) Bio-Technology 6:1197), or any other

nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well-known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

5 In an alternative embodiment, mutations in a biomarker nucleic acid from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA
10 indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Pat. No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

 In other embodiments, genetic mutations in biomarker nucleic acid can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density
15 arrays containing hundreds or thousands of oligonucleotide probes (Cronin, M. T. *et al.* (1996) *Hum. Mutat.* 7:244-255; Kozal, M. J. *et al.* (1996) *Nat. Med.* 2:753-759). For example, biomarker genetic mutations can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin *et al.* (1996) *supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a
20 sample and control to identify base changes between the sequences by making linear arrays of sequential, overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one
25 complementary to the wild-type gene and the other complementary to the mutant gene. Such biomarker genetic mutations can be identified in a variety of contexts, including, for example, germline and somatic mutations.

 In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence a biomarker gene and detect mutations by comparing
30 the sequence of the sample biomarker with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert (1977) *Proc. Natl. Acad. Sci. USA* 74:560 or Sanger (1977) *Proc. Natl. Acad. Sci. USA* 74:5463. It is also contemplated that any of a variety of automated sequencing

procedures can be utilized when performing the diagnostic assays (Naeve (1995) *Biotechniques* 19:448-53), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

5 Other methods for detecting mutations in a biomarker gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of “mismatch cleavage” starts by providing heteroduplexes formed by hybridizing (labeled) RNA or DNA containing the wild-type biomarker sequence with
10 potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to base pair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with SI nuclease to enzymatically digest the mismatched regions. In other
15 embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:4397 and Saleeba *et al.* (1992) *Methods*
20 Enzymol. 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

 In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called “DNA mismatch repair” enzymes) in defined systems for detecting and mapping point mutations
25 in biomarker cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a biomarker sequence, *e.g.*, a wild-type biomarker treated with a DNA mismatch repair enzyme, and the cleavage products, if any,
30 can be detected from electrophoresis protocols or the like (*e.g.*, U.S. Pat. No. 5,459,039.)

 In other embodiments, alterations in electrophoretic mobility can be used to identify mutations in biomarker genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and

wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA* 86:2766; see also Cotton (1993) *Mutat. Res.* 285:125-144 and Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control biomarker nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet.* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to ensure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys. Chem.* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163; Saiki *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where,

under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

3. Anti-Cancer Therapies

The efficacy of inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erap1, etc.), and immunotherapy combination treatment is predicted according to biomarker amount and/or activity associated with a cancer in a subject according to the methods described herein. In one embodiment, such inhibitor and immunotherapy combination treatments (*e.g.*, one or more inhibitor and immunotherapy combination treatment in combination with one or more additional anti-cancer therapies, such as another immune checkpoint inhibitor) can be administered, particularly if a subject has first been indicated as being a likely responder to inhibitor and immunotherapy combination treatment. In another embodiment, such inhibitor and immunotherapy combination treatment can be avoided once a subject is indicated as not being a likely responder to inhibitor and immunotherapy combination treatment and an alternative treatment regimen, such as targeted and/or untargeted anti-cancer therapies can be administered. Combination therapies are also contemplated and can comprise, for example, one or more chemotherapeutic agents and radiation, one or more chemotherapeutic agents and immunotherapy, or one or more chemotherapeutic agents, radiation and chemotherapy, each combination of which can be with anti-immune checkpoint therapy. In addition, any representative embodiment of an agent to modulate a particular target can be adapted to any other target described herein by the ordinarily skilled artisan (*e.g.*, direct and indirect PD-1 inhibitors described herein can be applied to other immune checkpoint inhibitors and/or one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein, such as monospecific antibodies, bispecific

antibodies, non-activating forms, small molecules, peptides, interfering nucleic acids, and the like).

The term “targeted therapy” refers to administration of agents that selectively interact with a chosen biomolecule to thereby treat cancer. One example includes immunotherapies such as immune checkpoint inhibitors, which are well-known in the art. For example, anti-PD-1 pathway agents, such as therapeutic monoclonal blocking antibodies, which are well-known in the art and described above, can be used to target tumor microenvironments and cells expressing unwanted components of the PD-1 pathway, such as PD-1, PD-L1, and/or PD-L2.

For example, the term “PD-1 pathway” refers to the PD-1 receptor and its ligands, PD-L1 and PD-L2. “PD-1 pathway inhibitors” block or otherwise reduce the interaction between PD-1 and one or both of its ligands such that the immunoinhibitory signaling otherwise generated by the interaction is blocked or otherwise reduced. Anti-immune checkpoint inhibitors can be direct or indirect. Direct anti-immune checkpoint inhibitors block or otherwise reduce the interaction between an immune checkpoint and at least one of its ligands. For example, PD-1 inhibitors can block PD-1 binding with one or both of its ligands. Direct PD-1 combination inhibitors are well-known in the art, especially since the natural binding partners of PD-1 (*e.g.*, PD-L1 and PD-L2), PD-L1 (*e.g.*, PD-1 and B7-1), and PD-L2 (*e.g.*, PD-1 and RGMb) are known.

For example, agents which directly block the interaction between PD-1 and PD-L1, PD-1 and PD-L2, PD-1 and both PD-L1 and PD-L2, such as a bispecific antibody, can prevent inhibitory signaling and upregulate an immune response (*i.e.*, as a PD-1 pathway inhibitor). Alternatively, agents that indirectly block the interaction between PD-1 and one or both of its ligands can prevent inhibitory signaling and upregulate an immune response. For example, B7-1 or a soluble form thereof, by binding to a PD-L1 polypeptide indirectly reduces the effective concentration of PD-L1 polypeptide available to bind to PD-1. Exemplary agents include monospecific or bispecific blocking antibodies against PD-1, PD-L1, and/or PD-L2 that block the interaction between the receptor and ligand(s); a non-activating form of PD-1, PD-L1, and/or PD-L2 (*e.g.*, a dominant negative or soluble polypeptide), small molecules or peptides that block the interaction between PD-1, PD-L1, and/or PD-L2; fusion proteins (*e.g.* the extracellular portion of PD-1, PD-L1, and/or PD-L2, fused to the Fc portion of an antibody or immunoglobulin) that bind to PD-1, PD-L1, and/or PD-L2 and inhibit the interaction between the receptor and ligand(s); a non-activating form

of a natural PD-1, PD-L2, and/or PD-L2 ligand, and a soluble form of a natural PD-1, PD-L2, and/or PD-L2 ligand.

Indirect anti-immune checkpoint inhibitors block or otherwise reduce the immunoinhibitory signaling generated by the interaction between the immune checkpoint and at least one of its ligands. For example, an inhibitor can block the interaction between PD-1 and one or both of its ligands without necessarily directly blocking the interaction between PD-1 and one or both of its ligands. For example, indirect inhibitors include intrabodies that bind the intracellular portion of PD-1 and/or PD-L1 required to signal to block or otherwise reduce the immunoinhibitory signaling. Similarly, nucleic acids that reduce the expression of PD-1, PD-L1, and/or PD-L2 can indirectly inhibit the interaction between PD-1 and one or both of its ligands by removing the availability of components for interaction. Such nucleic acid molecules can block PD-L1, PD-L2, and/or PD-L2 transcription or translation.

Similarly, agents which directly block the interaction between one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), and the binding partners and/or substrates of such one or more biomarkers, and the like, can remove the inhibition to such one or more biomarkers-regulated signaling and its downstream immune responses, such as increasing sensitivity to interferon signaling.

Alternatively, agents that indirectly block the interaction between such one or more biomarkers and its binding partners/substrates can remove the inhibition to such one or more biomarkers-regulated signaling and its downstream immune responses. For example, a truncated or dominant negative form of such one or more biomarkers, such as biomarker fragments without functional activity, by binding to a substrate of such one or more biomarkers and indirectly reducing the effective concentration of such substrate available to bind to the one or more biomarkers in cell. Exemplary agents include monospecific or bispecific blocking antibodies, especially intrabodies, against the one or more biomarkers and/or their substrate(s) that block the interaction between the one or more biomarkers and their substrate(s); a non-active form of such one or more biomarkers and/or their substrate(s) (*e.g.*, a dominant negative polypeptide), small molecules or peptides that block the interaction between such one or more biomarkers and their substrate(s) or the activity of such one or more biomarkers; and a non-activating form of a natural biomarker and/or its substrate(s).

Immunotherapies that are designed to elicit or amplify an immune response are referred to as “activation immunotherapies.” Immunotherapies that are designed to reduce or suppress an immune response are referred to as “suppression immunotherapies.” Any agent believed to have an immune system effect on the genetically modified transplanted

5 cancer cells can be assayed to determine whether the agent is an immunotherapy and the effect that a given genetic modification has on the modulation of immune response. In some embodiments, the immunotherapy is cancer cell-specific. In some embodiments, immunotherapy can be “untargeted,” which refers to administration of agents that do not selectively interact with immune system cells, yet modulates immune system function.

10 Representative examples of untargeted therapies include, without limitation, chemotherapy, gene therapy, and radiation therapy.

Immunotherapy can involve passive immunity for short-term protection of a host, achieved by the administration of pre-formed antibody directed against a cancer antigen or disease antigen (*e.g.*, administration of a monoclonal antibody, optionally linked to a

15 chemotherapeutic agent or toxin, to a tumor antigen). Immunotherapy can also focus on using the cytotoxic lymphocyte-recognized epitopes of cancer cell lines. Alternatively, antisense polynucleotides, ribozymes, RNA interference molecules, triple helix polynucleotides and the like, can be used to selectively modulate biomolecules that are linked to the initiation, progression, and/or pathology of a tumor or cancer.

20 In one embodiment, immunotherapy comprises adoptive cell-based immunotherapies. Well-known adoptive cell-based immunotherapeutic modalities, including, without limitation, Irradiated autologous or allogeneic tumor cells, tumor lysates or apoptotic tumor cells, antigen-presenting cell-based immunotherapy, dendritic cell-based immunotherapy, adoptive T cell transfer, adoptive CAR T cell therapy, autologous immune

25 enhancement therapy (AIET), cancer vaccines, and/or antigen presenting cells. Such cell-based immunotherapies can be further modified to express one or more gene products to further modulate immune responses, such as expressing cytokines like GM-CSF, and/or to express tumor-associated antigen (TAA) antigens, such as Mage-1, gp-100, patient-specific neoantigen vaccines, and the like.

30 In another embodiment, immunotherapy comprises non-cell-based immunotherapies. In one embodiment, compositions comprising antigens with or without vaccine-enhancing adjuvants are used. Such compositions exist in many well-known forms, such as peptide compositions, oncolytic viruses, recombinant antigen comprising

fusion proteins, and the like. In still another embodiment, immunomodulatory interleukins, such as IL-2, IL-6, IL-7, IL-12, IL-17, IL-23, and the like, as well as modulators thereof (*e.g.*, blocking antibodies or more potent or longer lasting forms) are used. In yet another embodiment, immunomodulatory cytokines, such as interferons, G-CSF, imiquimod, TNFalpha, and the like, as well as modulators thereof (*e.g.*, blocking antibodies or more potent or longer lasting forms) are used. In another embodiment, immunomodulatory chemokines, such as CCL3, CCL26, and CXCL7, and the like, as well as modulators thereof (*e.g.*, blocking antibodies or more potent or longer lasting forms) are used. In another embodiment, immunomodulatory molecules targeting immunosuppression, such as STAT3 signaling modulators, NFkappaB signaling modulators, and immune checkpoint modulators, are used. The terms “immune checkpoint” and “anti-immune checkpoint therapy” are described above.

In still another embodiment, immunomodulatory drugs, such as immunocytostatic drugs, glucocorticoids, cytostatics, immunophilins and modulators thereof (*e.g.*, rapamycin, a calcineurin inhibitor, tacrolimus, ciclosporin (cyclosporin), pimecrolimus, abetimus, gusperimus, ridaforolimus, everolimus, temsirolimus, zotarolimus, etc.), hydrocortisone (cortisol), cortisone acetate, prednisone, prednisolone, methylprednisolone, dexamethasone, betamethasone, triamcinolone, beclometasone, fludrocortisone acetate, deoxycorticosterone acetate (doca) aldosterone, a non-glucocorticoid steroid, a pyrimidine synthesis inhibitor, leflunomide, teriflunomide, a folic acid analog, methotrexate, anti-thymocyte globulin, anti-lymphocyte globulin, thalidomide, lenalidomide, pentoxifylline, bupropion, curcumin, catechin, an opioid, an IMPDH inhibitor, mycophenolic acid, myriocin, fingolimod, an NF-xB inhibitor, raloxifene, drotrecogin alfa, denosumab, an NF-xB signaling cascade inhibitor, disulfiram, olmesartan, dithiocarbamate, a proteasome inhibitor, bortezomib, MG132, Prol, NPI-0052, curcumin, genistein, resveratrol, parthenolide, thalidomide, lenalidomide, flavopiridol, non-steroidal anti-inflammatory drugs (NSAIDs), arsenic trioxide, dehydroxymethylepoxyquinomycin (DHMEQ), I3C(indole-3-carbinol)/DIM(di-indolmethane) (I3C/DIM), Bay 11-7082, luteolin, cell permeable peptide SN-50, IKBa.-super repressor overexpression, NFkB decoy oligodeoxynucleotide (ODN), or a derivative or analog of any thereof, are used. In yet another embodiment, immunomodulatory antibodies or protein are used. For example, antibodies that bind to CD40, Toll-like receptor (TLR), OX40, GITR, CD27, or to 4-1BB, T-cell bispecific antibodies, an anti-IL-2 receptor antibody, an anti-CD3 antibody, OKT3 (muromonab), oteelixizumab, teplizumab,

visilizumab, an anti-CD4 antibody, clenoliximab, keliximab, zanolimumab, an anti-CD11 a antibody, efalizumab, an anti-CD18 antibody, erlizumab, rovelizumab, an anti-CD20 antibody, afutuzumab, ocrelizumab, ofatumumab, pascolizumab, rituximab, an anti-CD23 antibody, lumiliximab, an anti-CD40 antibody, teneliximab, toralizumab, an anti-CD40L antibody, ruplizumab, an anti-CD62L antibody, aselizumab, an anti-CD80 antibody, galiximab, an anti-CD147 antibody, gavilimomab, a B-Lymphocyte stimulator (BLyS) inhibiting antibody, belimumab, an CTLA4-Ig fusion protein, abatacept, belatacept, an anti-CTLA4 antibody, ipilimumab, tremelimumab, an anti-eotaxin 1 antibody, bertilimumab, an anti-a4-integrin antibody, natalizumab, an anti-IL-6R antibody, tocilizumab, an anti-LFA-1 antibody, odulimomab, an anti-CD25 antibody, basiliximab, daclizumab, inolimomab, an anti-CD5 antibody, zolimomab, an anti-CD2 antibody, siplizumab, nerelimomab, faralimomab, atlizumab, atorolimumab, cedelizumab, dorlimomab aritox, dorlixizumab, fontolizumab, gantenerumab, gomiliximab, lebrilizumab, maslimomab, morolimumab, pexelizumab, reslizumab, rovelizumab, talizumab, telimomab aritox, vapaliximab, vepalimomab, aflibercept, alefacept, rilonacept, an IL-1 receptor antagonist, anakinra, an anti-IL-5 antibody, mepolizumab, an IgE inhibitor, omalizumab, talizumab, an IL12 inhibitor, an IL23 inhibitor, ustekinumab, and the like.

Nutritional supplements that enhance immune responses, such as vitamin A, vitamin E, vitamin C, and the like, are well-known in the art (see, for example, U.S. Pat. Nos. 4,981,844 and 5,230,902 and PCT Publ. No. WO 2004/004483) can be used in the methods described herein.

Similarly, agents and therapies other than immunotherapy or in combination thereof can be used with in combination with inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), with or without immunotherapies to stimulate an immune response to thereby treat a condition that would benefit therefrom. For example, chemotherapy, radiation, epigenetic modifiers (*e.g.*, histone deacetylase (HDAC) modifiers, methylation modifiers, phosphorylation modifiers, and the like), targeted therapy, and the like are well-known in the art.

The term “untargeted therapy” refers to administration of agents that do not selectively interact with a chosen biomolecule yet treat cancer. Representative examples of untargeted therapies include, without limitation, chemotherapy, gene therapy, and radiation therapy.

In one embodiment, chemotherapy is used. Chemotherapy includes the administration of a chemotherapeutic agent. Such a chemotherapeutic agent may be, but is not limited to, those selected from among the following groups of compounds: platinum compounds, cytotoxic antibiotics, antimetabolites, anti-mitotic agents, alkylating agents, arsenic compounds, DNA topoisomerase inhibitors, taxanes, nucleoside analogues, plant alkaloids, and toxins; and synthetic derivatives thereof. Exemplary compounds include, but are not limited to, alkylating agents: cisplatin, treosulfan, and trofosfamide; plant alkaloids: vinblastine, paclitaxel, docetaxol; DNA topoisomerase inhibitors: teniposide, crisnatol, and mitomycin; anti-folates: methotrexate, mycophenolic acid, and hydroxyurea; pyrimidine analogs: 5-fluorouracil, doxifluridine, and cytosine arabinoside; purine analogs: mercaptopurine and thioguanine; DNA antimetabolites: 2'-deoxy-5-fluorouridine, aphidicolin glycinate, and pyrazoloimidazole; and antimitotic agents: halichondrin, colchicine, and rhizoxin. Compositions comprising one or more chemotherapeutic agents (*e.g.*, FLAG, CHOP) may also be used. FLAG comprises fludarabine, cytosine arabinoside (Ara-C) and G-CSF. CHOP comprises cyclophosphamide, vincristine, doxorubicin, and prednisone. In another embodiment, PARP (*e.g.*, PARP-1 and/or PARP-2) inhibitors are used and such inhibitors are well-known in the art (*e.g.*, Olaparib, ABT-888, BSI-201, BGP-15 (N-Gene Research Laboratories, Inc.); INO-1001 (Inotek Pharmaceuticals Inc.); PJ34 (Soriano *et al.*, 2001; Pacher *et al.*, 2002b); 3-aminobenzamide (Trevigen); 4-amino-1,8-naphthalimide; (Trevigen); 6(5H)-phenanthridinone (Trevigen); benzamide (U.S. Pat. Re. 36,397); and NU1025 (Bowman *et al.*). The mechanism of action is generally related to the ability of PARP inhibitors to bind PARP and decrease its activity. PARP catalyzes the conversion of .beta.-nicotinamide adenine dinucleotide (NAD⁺) into nicotinamide and poly-ADP-ribose (PAR). Both poly (ADP-ribose) and PARP have been linked to regulation of transcription, cell proliferation, genomic stability, and carcinogenesis (Bouchard V. J. *et al.* Experimental Hematology, Volume 31, Number 6, June 2003, pp. 446-454(9); Herceg Z.; Wang Z.-Q. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, Volume 477, Number 1, 2 Jun. 2001, pp. 97-110(14)). Poly(ADP-ribose) polymerase 1 (PARP1) is a key molecule in the repair of DNA single-strand breaks (SSBs) (de Murcia J. *et al.* 1997. Proc Natl Acad Sci USA 94:7303-7307; Schreiber V, Dantzer F, Ame J C, de Murcia G (2006) Nat Rev Mol Cell Biol 7:517-528; Wang Z Q, *et al.* (1997) Genes Dev 11:2347-2358). Knockout of SSB repair by inhibition of PARP1 function induces DNA double-strand breaks (DSBs) that can trigger synthetic

lethality in cancer cells with defective homology-directed DSB repair (Bryant H E, *et al.* (2005) Nature 434:913-917; Farmer H, *et al.* (2005) Nature 434:917-921). The foregoing examples of chemotherapeutic agents are illustrative, and are not intended to be limiting.

In another embodiment, radiation therapy is used. The radiation used in radiation
5 therapy can be ionizing radiation. Radiation therapy can also be gamma rays, X-rays, or
proton beams. Examples of radiation therapy include, but are not limited to, external-beam
radiation therapy, interstitial implantation of radioisotopes (I-125, palladium, iridium),
radioisotopes such as strontium-89, thoracic radiation therapy, intraperitoneal P-32
radiation therapy, and/or total abdominal and pelvic radiation therapy. For a general
10 overview of radiation therapy, see Hellman, Chapter 16: Principles of Cancer Management:
Radiation Therapy, 6th edition, 2001, DeVita *et al.*, eds., J. B. Lippencott Company,
Philadelphia. The radiation therapy can be administered as external beam radiation or
teletherapy wherein the radiation is directed from a remote source. The radiation treatment
can also be administered as internal therapy or brachytherapy wherein a radioactive source
15 is placed inside the body close to cancer cells or a tumor mass. Also encompassed is the use
of photodynamic therapy comprising the administration of photosensitizers, such as
hematoporphyrin and its derivatives, Vertoporphin (BPD-MA), phthalocyanine,
photosensitizer Pc4, demethoxy-hypocrellin A; and 2BA-2-DMHA.

In another embodiment, surgical intervention can occur to physically remove
20 cancerous cells and/or tissues.

In still another embodiment, hormone therapy is used. Hormonal therapeutic
treatments can comprise, for example, hormonal agonists, hormonal antagonists (*e.g.*,
flutamide, bicalutamide, tamoxifen, raloxifene, leuprolide acetate (LUPRON), LH-RH
antagonists), inhibitors of hormone biosynthesis and processing, and steroids (*e.g.*,
25 dexamethasone, retinoids, deltoids, betamethasone, cortisol, cortisone, prednisone,
dehydrotestosterone, glucocorticoids, mineralocorticoids, estrogen, testosterone,
progestins), vitamin A derivatives (*e.g.*, all-trans retinoic acid (ATRA)); vitamin D3
analogs; antigestagens (*e.g.*, mifepristone, onapristone), or antiandrogens (*e.g.*, cyproterone
acetate).

30 In yet another embodiment, hyperthermia, a procedure in which body tissue is
exposed to high temperatures (up to 106°F.) is used. Heat may help shrink tumors by
damaging cells or depriving them of substances they need to live. Hyperthermia therapy
can be local, regional, and whole-body hyperthermia, using external and internal heating

devices. Hyperthermia is almost always used with other forms of therapy (*e.g.*, radiation therapy, chemotherapy, and biological therapy) to try to increase their effectiveness. Local hyperthermia refers to heat that is applied to a very small area, such as a tumor. The area may be heated externally with high-frequency waves aimed at a tumor from a device

5 outside the body. To achieve internal heating, one of several types of sterile probes may be used, including thin, heated wires or hollow tubes filled with warm water; implanted microwave antennae; and radiofrequency electrodes. In regional hyperthermia, an organ or a limb is heated. Magnets and devices that produce high energy are placed over the region to be heated. In another approach, called perfusion, some of the patient's blood is removed,

10 heated, and then pumped (perfused) into the region that is to be heated internally. Whole-body heating is used to treat metastatic cancer that has spread throughout the body. It can be accomplished using warm-water blankets, hot wax, inductive coils (like those in electric blankets), or thermal chambers (similar to large incubators). Hyperthermia does not cause any marked increase in radiation side effects or complications. Heat applied directly to the

15 skin, however, can cause discomfort or even significant local pain in about half the patients treated. It can also cause blisters, which generally heal rapidly.

In still another embodiment, photodynamic therapy (also called PDT, photoradiation therapy, phototherapy, or photochemotherapy) is used for the treatment of some types of cancer. It is based on the discovery that certain chemicals known as photosensitizing agents

20 can kill one-celled organisms when the organisms are exposed to a particular type of light. PDT destroys cancer cells through the use of a fixed-frequency laser light in combination with a photosensitizing agent. In PDT, the photosensitizing agent is injected into the bloodstream and absorbed by cells all over the body. The agent remains in cancer cells for a longer time than it does in normal cells. When the treated cancer cells are exposed to

25 laser light, the photosensitizing agent absorbs the light and produces an active form of oxygen that destroys the treated cancer cells. Light exposure must be timed carefully so that it occurs when most of the photosensitizing agent has left healthy cells but is still present in the cancer cells. The laser light used in PDT can be directed through a fiber-optic (a very thin glass strand). The fiber-optic is placed close to the cancer to deliver the

30 proper amount of light. The fiber-optic can be directed through a bronchoscope into the lungs for the treatment of lung cancer or through an endoscope into the esophagus for the treatment of esophageal cancer. An advantage of PDT is that it causes minimal damage to healthy tissue. However, because the laser light currently in use cannot pass through more

than about 3 centimeters of tissue (a little more than one and an eighth inch), PDT is mainly used to treat tumors on or just under the skin or on the lining of internal organs.

Photodynamic therapy makes the skin and eyes sensitive to light for 6 weeks or more after treatment. Patients are advised to avoid direct sunlight and bright indoor light for at least 6 weeks. If patients must go outdoors, they need to wear protective clothing, including sunglasses. Other temporary side effects of PDT are related to the treatment of specific areas and can include coughing, trouble swallowing, abdominal pain, and painful breathing or shortness of breath. In December 1995, the U.S. Food and Drug Administration (FDA) approved a photosensitizing agent called porfimer sodium, or Photofrin®, to relieve symptoms of esophageal cancer that is causing an obstruction and for esophageal cancer that cannot be satisfactorily treated with lasers alone. In January 1998, the FDA approved porfimer sodium for the treatment of early non-small cell lung cancer in patients for whom the usual treatments for lung cancer are not appropriate. The National Cancer Institute and other institutions are supporting clinical trials (research studies) to evaluate the use of photodynamic therapy for several types of cancer, including cancers of the bladder, brain, larynx, and oral cavity.

In yet another embodiment, laser therapy is used to harness high-intensity light to destroy cancer cells. This technique is often used to relieve symptoms of cancer such as bleeding or obstruction, especially when the cancer cannot be cured by other treatments. It may also be used to treat cancer by shrinking or destroying tumors. The term “laser” stands for light amplification by stimulated emission of radiation. Ordinary light, such as that from a light bulb, has many wavelengths and spreads in all directions. Laser light, on the other hand, has a specific wavelength and is focused in a narrow beam. This type of high-intensity light contains a lot of energy. Lasers are very powerful and may be used to cut through steel or to shape diamonds. Lasers also can be used for very precise surgical work, such as repairing a damaged retina in the eye or cutting through tissue (in place of a scalpel). Although there are several different kinds of lasers, only three kinds have gained wide use in medicine: Carbon dioxide (CO₂) laser--This type of laser can remove thin layers from the skin's surface without penetrating the deeper layers. This technique is particularly useful in treating tumors that have not spread deep into the skin and certain precancerous conditions. As an alternative to traditional scalpel surgery, the CO₂ laser is also able to cut the skin. The laser is used in this way to remove skin cancers.

Neodymium:yttrium-aluminum-garnet (Nd:YAG) laser-- Light from this laser can penetrate

deeper into tissue than light from the other types of lasers, and it can cause blood to clot quickly. It can be carried through optical fibers to less accessible parts of the body. This type of laser is sometimes used to treat throat cancers. Argon laser--This laser can pass through only superficial layers of tissue and is therefore useful in dermatology and in eye surgery. It also is used with light-sensitive dyes to treat tumors in a procedure known as photodynamic therapy (PDT). Lasers have several advantages over standard surgical tools, including: Lasers are more precise than scalpels. Tissue near an incision is protected, since there is little contact with surrounding skin or other tissue. The heat produced by lasers sterilizes the surgery site, thus reducing the risk of infection. Less operating time may be needed because the precision of the laser allows for a smaller incision. Healing time is often shortened; since laser heat seals blood vessels, there is less bleeding, swelling, or scarring. Laser surgery may be less complicated. For example, with fiber optics, laser light can be directed to parts of the body without making a large incision. More procedures may be done on an outpatient basis. Lasers can be used in two ways to treat cancer: by shrinking or destroying a tumor with heat, or by activating a chemical--known as a photosensitizing agent--that destroys cancer cells. In PDT, a photosensitizing agent is retained in cancer cells and can be stimulated by light to cause a reaction that kills cancer cells. CO₂ and Nd:YAG lasers are used to shrink or destroy tumors. They may be used with endoscopes, tubes that allow physicians to see into certain areas of the body, such as the bladder. The light from some lasers can be transmitted through a flexible endoscope fitted with fiber optics. This allows physicians to see and work in parts of the body that could not otherwise be reached except by surgery and therefore allows very precise aiming of the laser beam. Lasers also may be used with low-power microscopes, giving the doctor a clear view of the site being treated. Used with other instruments, laser systems can produce a cutting area as small as 200 microns in diameter--less than the width of a very fine thread. Lasers are used to treat many types of cancer. Laser surgery is a standard treatment for certain stages of glottis (vocal cord), cervical, skin, lung, vaginal, vulvar, and penile cancers. In addition to its use to destroy the cancer, laser surgery is also used to help relieve symptoms caused by cancer (palliative care). For example, lasers may be used to shrink or destroy a tumor that is blocking a patient's trachea (windpipe), making it easier to breathe. It is also sometimes used for palliation in colorectal and anal cancer. Laser-induced interstitial thermotherapy (LITT) is one of the most recent developments in laser therapy. LITT uses the same idea as a cancer treatment called hyperthermia; that heat may

help shrink tumors by damaging cells or depriving them of substances they need to live. In this treatment, lasers are directed to interstitial areas (areas between organs) in the body. The laser light then raises the temperature of the tumor, which damages or destroys cancer cells.

5 The duration and/or dose of treatment with therapies may vary according to the particular therapeutic agent or combination thereof. An appropriate treatment time for a particular cancer therapeutic agent will be appreciated by the skilled artisan. The present invention contemplates the continued assessment of optimal treatment schedules for each cancer therapeutic agent, where the phenotype of the cancer of the subject as determined by
10 the methods of the present invention is a factor in determining optimal treatment doses and schedules.

 Any means for the introduction of a polynucleotide into mammals, human or non-human, or cells thereof may be adapted to the practice of this invention for the delivery of the various constructs of the present invention into the intended recipient. In one
15 embodiment of the present invention, the DNA constructs are delivered to cells by transfection, *i.e.*, by delivery of “naked” DNA or in a complex with a colloidal dispersion system. A colloidal system includes macromolecule complexes, nanocapsules, microspheres, beads, and lipid-based systems including oil-in-water emulsions, micelles, mixed micelles, and liposomes. The preferred colloidal system of this invention is a lipid-
20 complexed or liposome-formulated DNA. In the former approach, prior to formulation of DNA, *e.g.*, with lipid, a plasmid containing a transgene bearing the desired DNA constructs may first be experimentally optimized for expression (*e.g.*, inclusion of an intron in the 5' untranslated region and elimination of unnecessary sequences (Felgner, *et al.*, Ann NY Acad Sci 126-139, 1995). Formulation of DNA, *e.g.* with various lipid or liposome
25 materials, may then be effected using known methods and materials and delivered to the recipient mammal. See, *e.g.*, Canonico *et al.*, Am J Respir Cell Mol Biol 10:24-29, 1994; Tsan *et al.*, Am J Physiol 268; Alton *et al.*, Nat Genet. 5:135-142, 1993 and U.S. patent No. 5,679,647 by Carson *et al.*

 The targeting of liposomes can be classified based on anatomical and mechanistic
30 factors. Anatomical classification is based on the level of selectivity, for example, organ-specific, cell-specific, and organelle-specific. Mechanistic targeting can be distinguished based upon whether it is passive or active. Passive targeting utilizes the natural tendency of liposomes to distribute to cells of the reticulo-endothelial system (RES) in organs, which

contain sinusoidal capillaries. Active targeting, on the other hand, involves alteration of the liposome by coupling the liposome to a specific ligand such as a monoclonal antibody, sugar, glycolipid, or protein, or by changing the composition or size of the liposome in order to achieve targeting to organs and cell types other than the naturally occurring sites of localization.

The surface of the targeted delivery system may be modified in a variety of ways. In the case of a liposomal targeted delivery system, lipid groups can be incorporated into the lipid bilayer of the liposome in order to maintain the targeting ligand in stable association with the liposomal bilayer. Various linking groups can be used for joining the lipid chains to the targeting ligand. Naked DNA or DNA associated with a delivery vehicle, *e.g.*, liposomes, can be administered to several sites in a subject (see below).

Nucleic acids can be delivered in any desired vector. These include viral or non-viral vectors, including adenovirus vectors, adeno-associated virus vectors, retrovirus vectors, lentivirus vectors, and plasmid vectors. Exemplary types of viruses include HSV (herpes simplex virus), AAV (adeno associated virus), HIV (human immunodeficiency virus), BIV (bovine immunodeficiency virus), and MLV (murine leukemia virus). Nucleic acids can be administered in any desired format that provides sufficiently efficient delivery levels, including in virus particles, in liposomes, in nanoparticles, and complexed to polymers.

The nucleic acids encoding a protein or nucleic acid of interest may be in a plasmid or viral vector, or other vector as is known in the art. Such vectors are well-known and any can be selected for a particular application. In one embodiment of the present invention, the gene delivery vehicle comprises a promoter and a demethylase coding sequence. Preferred promoters are tissue-specific promoters and promoters which are activated by cellular proliferation, such as the thymidine kinase and thymidylate synthase promoters. Other preferred promoters include promoters which are activatable by infection with a virus, such as the α - and β -interferon promoters, and promoters which are activatable by a hormone, such as estrogen. Other promoters which can be used include the Moloney virus LTR, the CMV promoter, and the mouse albumin promoter. A promoter may be constitutive or inducible.

In another embodiment, naked polynucleotide molecules are used as gene delivery vehicles, as described in WO 90/11092 and U.S. Patent 5,580,859. Such gene delivery vehicles can be either growth factor DNA or RNA and, in certain embodiments, are linked

to killed adenovirus. Curiel *et al.*, Hum. Gene. Ther. 3:147-154, 1992. Other vehicles which can optionally be used include DNA-ligand (Wu *et al.*, J. Biol. Chem. 264:16985-16987, 1989), lipid-DNA combinations (Felgner *et al.*, Proc. Natl. Acad. Sci. USA 84:7413-7417, 1989), liposomes (Wang *et al.*, Proc. Natl. Acad. Sci. 84:7851-7855, 1987) and microprojectiles (Williams *et al.*, Proc. Natl. Acad. Sci. 88:2726-2730, 1991).

A gene delivery vehicle can optionally comprise viral sequences such as a viral origin of replication or packaging signal. These viral sequences can be selected from viruses such as astrovirus, coronavirus, orthomyxovirus, papovavirus, paramyxovirus, parvovirus, picornavirus, poxvirus, retrovirus, togavirus or adenovirus. In a preferred embodiment, the growth factor gene delivery vehicle is a recombinant retroviral vector. Recombinant retroviruses and various uses thereof have been described in numerous references including, for example, Mann *et al.*, Cell 33:153, 1983, Cane and Mulligan, Proc. Nat'l. Acad. Sci. USA 81:6349, 1984, Miller *et al.*, Human Gene Therapy 1:5-14, 1990, U.S. Patent Nos. 4,405,712, 4,861,719, and 4,980,289, and PCT Application Nos. WO 89/02,468, WO 89/05,349, and WO 90/02,806. Numerous retroviral gene delivery vehicles can be utilized in the present invention, including for example those described in EP 0,415,731; WO 90/07936; WO 94/03622; WO 93/25698; WO 93/25234; U.S. Patent No. 5,219,740; WO 93/11230; WO 93/10218; Vile and Hart, Cancer Res. 53:3860-3864, 1993; Vile and Hart, Cancer Res. 53:962-967, 1993; Ram *et al.*, Cancer Res. 53:83-88, 1993; Takamiya *et al.*, J. Neurosci. Res. 33:493-503, 1992; Baba *et al.*, J. Neurosurg. 79:729-735, 1993 (U.S. Patent No. 4,777,127, GB 2,200,651, EP 0,345,242 and WO91/02805).

Other viral vector systems that can be used to deliver a polynucleotide of the present invention have been derived from herpes virus, *e.g.*, Herpes Simplex Virus (U.S. Patent No. 5,631,236 by Woo *et al.*, issued May 20, 1997 and WO 00/08191 by Neurovex), vaccinia virus (Ridgeway (1988) Ridgeway, "Mammalian expression vectors," In: Rodriguez R L, Denhardt D T, ed. Vectors: A survey of molecular cloning vectors and their uses. Stoneham: Butterworth,; Baichwal and Sugden (1986) "Vectors for gene transfer derived from animal DNA viruses: Transient and stable expression of transferred genes," In: Kucherlapati R, ed. Gene transfer. New York: Plenum Press; Coupar *et al.* (1988) Gene, 68:1-10), and several RNA viruses. Preferred viruses include an alphavirus, a poxvirus, an arena virus, a vaccinia virus, a polio virus, and the like. They offer several attractive features for various mammalian cells (Friedmann (1989) Science, 244:1275-1281;

Ridgeway, 1988, *supra*; Baichwal and Sugden, 1986, *supra*; Coupar *et al.*, 1988; Horwich *et al.* (1990) *J. Virol.*, 64:642-650).

In other embodiments, target DNA in the genome can be manipulated using well-known methods in the art. For example, the target DNA in the genome can be manipulated by deletion, insertion, and/or mutation are retroviral insertion, artificial chromosome techniques, gene insertion, random insertion with tissue specific promoters, gene targeting, transposable elements and/or any other method for introducing foreign DNA or producing modified DNA/modified nuclear DNA. Other modification techniques include deleting DNA sequences from a genome and/or altering nuclear DNA sequences. Nuclear DNA sequences, for example, may be altered by site-directed mutagenesis.

In other embodiments, recombinant biomarker polypeptides, and fragments thereof, can be administered to subjects. In some embodiments, fusion proteins can be constructed and administered which have enhanced biological properties. In addition, the biomarker polypeptides, and fragment thereof, can be modified according to well-known pharmacological methods in the art (*e.g.*, pegylation, glycosylation, oligomerization, *etc.*) in order to further enhance desirable biological activities, such as increased bioavailability and decreased proteolytic degradation.

4. Clinical Efficacy

Clinical efficacy can be measured by any method known in the art. For example, the response to a therapy, such as inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, *etc.*), and immunotherapy combination treatment, relates to any response of the cancer, *e.g.*, a tumor, to the therapy, preferably to a change in tumor mass and/or volume after initiation of neoadjuvant or adjuvant chemotherapy. Tumor response may be assessed in a neoadjuvant or adjuvant situation where the size of a tumor after systemic intervention can be compared to the initial size and dimensions as measured by CT, PET, mammogram, ultrasound or palpation and the cellularity of a tumor can be estimated histologically and compared to the cellularity of a tumor biopsy taken before initiation of treatment. Response may also be assessed by caliper measurement or pathological examination of the tumor after biopsy or surgical resection. Response may be recorded in a quantitative fashion like percentage change in tumor volume or cellularity or using a semi-quantitative scoring system such as residual

cancer burden (Symmans *et al.*, *J. Clin. Oncol.* (2007) 25:4414-4422) or Miller-Payne score (Ogston *et al.*, (2003) *Breast* (Edinburgh, Scotland) 12:320-327) in a qualitative fashion like “pathological complete response” (pCR), “clinical complete remission” (cCR), “clinical partial remission” (cPR), “clinical stable disease” (cSD), “clinical progressive disease” (cPD) or other qualitative criteria. Assessment of tumor response may be performed early after the onset of neoadjuvant or adjuvant therapy, *e.g.*, after a few hours, days, weeks or preferably after a few months. A typical endpoint for response assessment is upon termination of neoadjuvant chemotherapy or upon surgical removal of residual tumor cells and/or the tumor bed.

In some embodiments, clinical efficacy of the therapeutic treatments described herein may be determined by measuring the clinical benefit rate (CBR). The clinical benefit rate is measured by determining the sum of the percentage of patients who are in complete remission (CR), the number of patients who are in partial remission (PR) and the number of patients having stable disease (SD) at a time point at least 6 months out from the end of therapy. The shorthand for this formula is $CBR = CR + PR + SD$ over 6 months. In some embodiments, the CBR for a particular anti-immune checkpoint therapeutic regimen is at least 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, or more.

Additional criteria for evaluating the response to immunotherapies, such as anti-immune checkpoint therapies, are related to “survival,” which includes all of the following: survival until mortality, also known as overall survival (wherein said mortality may be either irrespective of cause or tumor related); “recurrence-free survival” (wherein the term recurrence shall include both localized and distant recurrence); metastasis free survival; disease free survival (wherein the term disease shall include cancer and diseases associated therewith). The length of said survival may be calculated by reference to a defined start point (*e.g.*, time of diagnosis or start of treatment) and end point (*e.g.*, death, recurrence or metastasis). In addition, criteria for efficacy of treatment can be expanded to include response to chemotherapy, probability of survival, probability of metastasis within a given time period, and probability of tumor recurrence.

For example, in order to determine appropriate threshold values, a particular anti-cancer therapeutic regimen can be administered to a population of subjects and the outcome can be correlated to biomarker measurements that were determined prior to administration of any immunotherapy, such as anti-immune checkpoint therapy. The outcome

measurement may be pathologic response to therapy given in the neoadjuvant setting. Alternatively, outcome measures, such as overall survival and disease-free survival can be monitored over a period of time for subjects following immunotherapies for whom biomarker measurement values are known. In certain embodiments, the same doses of immunotherapy agents, if any, are administered to each subject. In related embodiments, the doses administered are standard doses known in the art for those agents used in immunotherapies. The period of time for which subjects are monitored can vary. For example, subjects may be monitored for at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 45, 50, 55, or 60 months. Biomarker measurement threshold values that correlate to outcome of an immunotherapy can be determined using methods such as those described in the Examples section.

5. Further Uses and Methods of the Present Invention

The compositions described herein can be used in a variety of diagnostic, prognostic, and therapeutic applications. In any method described herein, such as a diagnostic method, prognostic method, therapeutic method, or combination thereof, all steps of the method can be performed by a single actor or, alternatively, by more than one actor. For example, diagnosis can be performed directly by the actor providing therapeutic treatment. Alternatively, a person providing a therapeutic agent can request that a diagnostic assay be performed. The diagnostician and/or the therapeutic interventionist can interpret the diagnostic assay results to determine a therapeutic strategy. Similarly, such alternative processes can apply to other assays, such as prognostic assays.

a. Screening Methods

One aspect of the present invention relates to screening assays, including non-cell based assays and xenograft animal model assays. In one embodiment, the assays provide a method for identifying whether a cancer is likely to respond to inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), and immunotherapy combination treatments, such as in a human by using a xenograft animal model assay, and/or whether an agent can inhibit the growth of or kill a cancer cell that is unlikely to respond to inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*,

HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), and immunotherapy combination treatments.

In one embodiment, the present invention relates to assays for screening test agents which bind to, or modulate the biological activity of, at least one biomarker described
5 herein (*e.g.*, in the tables, figures, examples, or otherwise in the specification). In one embodiment, a method for identifying such an agent entails determining the ability of the agent to modulate, *e.g.* inhibit, the at least one biomarker described herein.

In one embodiment, an assay is a cell-free or cell-based assay, comprising contacting at least one biomarker described herein, with a test agent, and determining the
10 ability of the test agent to modulate (*e.g.*, inhibit) the enzymatic activity of the biomarker, such as by measuring direct binding of substrates or by measuring indirect parameters as described below.

For example, in a direct binding assay, biomarker protein (or their respective target polypeptides or molecules) can be coupled with a radioisotope or enzymatic label such that
15 binding can be determined by detecting the labeled protein or molecule in a complex. For example, the targets can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, the targets can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label
20 detected by determination of conversion of an appropriate substrate to product. Determining the interaction between biomarker and substrate can also be accomplished using standard binding or enzymatic analysis assays. In one or more embodiments of the above described assay methods, it may be desirable to immobilize polypeptides or molecules to facilitate separation of complexed from uncomplexed forms of one or both of
25 the proteins or molecules, as well as to accommodate automation of the assay.

Binding of a test agent to a target can be accomplished in any vessel suitable for containing the reactants. Non-limiting examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. Immobilized forms of the antibodies described
30 herein can also include antibodies bound to a solid phase like a porous, microporous (with an average pore diameter less than about one micron) or macroporous (with an average pore diameter of more than about 10 microns) material, such as a membrane, cellulose, nitrocellulose, or glass fibers; a bead, such as that made of agarose or polyacrylamide or latex; or a surface of a dish, plate, or well, such as one made of polystyrene.

In an alternative embodiment, determining the ability of the agent to modulate the interaction between the biomarker and a substrate or a biomarker and its natural binding partner can be accomplished by determining the ability of the test agent to modulate the activity of a polypeptide or other product that functions downstream or upstream of its position within the signaling pathway (*e.g.*, feedback loops). Such feedback loops are well-known in the art (see, for example, Chen and Guillemin (2009) *Int. J. Tryptophan Res.* 2:1-19).

The present invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein, such as in an appropriate animal model. For example, an agent identified as described herein can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an antibody identified as described herein can be used in an animal model to determine the mechanism of action of such an agent.

b. Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining the amount and/or activity level of a biomarker described herein in the context of a biological sample (*e.g.*, blood, serum, cells, or tissue) to thereby determine whether an individual afflicted with a cancer is likely to respond to inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), and immunotherapy combination treatments, such as in a cancer. Such assays can be used for prognostic or predictive purpose alone, or can be coupled with a therapeutic intervention to thereby prophylactically treat an individual prior to the onset or after recurrence of a disorder characterized by or associated with biomarker polypeptide, nucleic acid expression or activity. The skilled artisan will appreciate that any method can use one or more (*e.g.*, combinations) of biomarkers described herein, such as those in the tables, figures, examples, and otherwise described in the specification.

Another aspect of the present invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds, and small nucleic acid-based molecules) on the expression

or activity of a biomarker described herein. These and other agents are described in further detail in the following sections.

The skilled artisan will also appreciate that, in certain embodiments, the methods of the present invention implement a computer program and computer system. For example, a computer program can be used to perform the algorithms described herein. A computer system can also store and manipulate data generated by the methods of the present invention which comprises a plurality of biomarker signal changes/profiles which can be used by a computer system in implementing the methods of this invention. In certain embodiments, a computer system receives biomarker expression data; (ii) stores the data; and (iii) compares the data in any number of ways described herein (*e.g.*, analysis relative to appropriate controls) to determine the state of informative biomarkers from cancerous or pre-cancerous tissue. In other embodiments, a computer system (i) compares the determined expression biomarker level to a threshold value; and (ii) outputs an indication of whether said biomarker level is significantly modulated (*e.g.*, above or below) the threshold value, or a phenotype based on said indication.

In certain embodiments, such computer systems are also considered part of the present invention. Numerous types of computer systems can be used to implement the analytic methods of this invention according to knowledge possessed by a skilled artisan in the bioinformatics and/or computer arts. Several software components can be loaded into memory during operation of such a computer system. The software components can comprise both software components that are standard in the art and components that are special to the present invention (*e.g.*, dCHIP software described in Lin *et al.* (2004) *Bioinformatics* 20, 1233-1240; radial basis machine learning algorithms (RBM) known in the art).

The methods of the present invention can also be programmed or modeled in mathematical software packages that allow symbolic entry of equations and high-level specification of processing, including specific algorithms to be used, thereby freeing a user of the need to procedurally program individual equations and algorithms. Such packages include, *e.g.*, Matlab from Mathworks (Natick, Mass.), Mathematica from Wolfram Research (Champaign, Ill.) or S-Plus from MathSoft (Seattle, Wash.).

In certain embodiments, the computer comprises a database for storage of biomarker data. Such stored profiles can be accessed and used to perform comparisons of interest at a later point in time. For example, biomarker expression profiles of a sample

derived from the non-cancerous tissue of a subject and/or profiles generated from population-based distributions of informative loci of interest in relevant populations of the same species can be stored and later compared to that of a sample derived from the cancerous tissue of the subject or tissue suspected of being cancerous of the subject.

5 In addition to the exemplary program structures and computer systems described herein, other, alternative program structures and computer systems will be readily apparent to the skilled artisan. Such alternative systems, which do not depart from the above described computer system and programs structures either in spirit or in scope, are therefore intended to be comprehended within the accompanying claims.

10 c. Diagnostic Assays

The present invention provides, in part, methods, systems, and code for accurately classifying whether a biological sample is associated with a cancer that is likely to respond to inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erap1, etc.), and immunotherapy combination treatments. In some
15 embodiments, the present invention is useful for classifying a sample (*e.g.*, from a subject) as associated with or at risk for responding to or not responding to inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erap1, etc.),
20 and immunotherapy combination treatments using a statistical algorithm and/or empirical data (*e.g.*, the amount or activity of a biomarker described herein, such as in the tables, figures, examples, and otherwise described in the specification).

An exemplary method for detecting the amount or activity of a biomarker described herein, and thus useful for classifying whether a sample is likely or unlikely to
25 respond to inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erap1, etc.), and immunotherapy combination treatments involves obtaining a biological sample from a test subject and contacting the biological sample with an agent, such as a protein-binding agent like an antibody or antigen-binding fragment
30 thereof, or a nucleic acid-binding agent like an oligonucleotide, capable of detecting the amount or activity of the biomarker in the biological sample. In some embodiments, at least one antibody or antigen-binding fragment thereof is used, wherein two, three, four, five, six, seven, eight, nine, ten, or more such antibodies or antibody fragments can be used

in combination (*e.g.*, in sandwich ELISAs) or in serial. In certain instances, the statistical algorithm is a single learning statistical classifier system. For example, a single learning statistical classifier system can be used to classify a sample as a based upon a prediction or probability value and the presence or level of the biomarker. The use of a single learning statistical classifier system typically classifies the sample as, for example, a likely immunotherapy responder or progressor sample with a sensitivity, specificity, positive predictive value, negative predictive value, and/or overall accuracy of at least about 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%.

Other suitable statistical algorithms are well-known to those of skill in the art. For example, learning statistical classifier systems include a machine learning algorithmic technique capable of adapting to complex data sets (*e.g.*, panel of markers of interest) and making decisions based upon such data sets. In some embodiments, a single learning statistical classifier system such as a classification tree (*e.g.*, random forest) is used. In other embodiments, a combination of 2, 3, 4, 5, 6, 7, 8, 9, 10, or more learning statistical classifier systems are used, preferably in tandem. Examples of learning statistical classifier systems include, but are not limited to, those using inductive learning (*e.g.*, decision/classification trees such as random forests, classification and regression trees (C&RT), boosted trees, etc.), Probably Approximately Correct (PAC) learning, connectionist learning (*e.g.*, neural networks (NN), artificial neural networks (ANN), neuro fuzzy networks (NFN), network structures, perceptrons such as multi-layer perceptrons, multi-layer feed-forward networks, applications of neural networks, Bayesian learning in belief networks, etc.), reinforcement learning (*e.g.*, passive learning in a known environment such as naive learning, adaptive dynamic learning, and temporal difference learning, passive learning in an unknown environment, active learning in an unknown environment, learning action-value functions, applications of reinforcement learning, etc.), and genetic algorithms and evolutionary programming. Other learning statistical classifier systems include support vector machines (*e.g.*, Kernel methods), multivariate adaptive regression splines (MARS), Levenberg-Marquardt algorithms, Gauss-Newton algorithms, mixtures of Gaussians, gradient descent algorithms, and learning vector quantization (LVQ). In certain embodiments, the method of the present invention further comprises sending the sample classification results to a clinician, *e.g.*, an oncologist.

In another embodiment, the diagnosis of a subject is followed by administering to the individual a therapeutically effective amount of a defined treatment based upon the diagnosis.

In one embodiment, the methods further involve obtaining a control biological sample (*e.g.*, biological sample from a subject who does not have a cancer or whose cancer is susceptible to inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erap1, etc.), and immunotherapy combination treatments), a biological sample from the subject during remission, or a biological sample from the subject during treatment for developing a cancer progressing despite inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erap1, etc.), and immunotherapy combination treatments.

d. Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a cancer that is likely or unlikely to be responsive to inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erap1, etc.), and immunotherapy combination treatments. The assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with a misregulation of the amount or activity of at least one biomarker described herein, such as in cancer. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disorder associated with a misregulation of the at least one biomarker described herein, such as in cancer. Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, polypeptide, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with the aberrant biomarker expression or activity.

e. Treatment Methods

The therapeutic compositions described herein, such as the combination of inhibitors of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1,

Tap2, Tapbp, Erap1, etc.), and immunotherapy, can be used in a variety of *in vitro* and *in vivo* therapeutic applications using the formulations and/or combinations described herein. In one embodiment, the therapeutic agents can be used to treat cancers determined to be responsive thereto. For example, single or multiple agents that inhibit or block both an inhibitor of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (e.g., HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erap1, etc.), and a immunotherapy can be used to treat cancers in subjects identified as likely responders thereto.

Modulatory methods of the present invention involve contacting a cell, such as an immune cell with an agent that inhibits or blocks the expression and/or activity of such one or more biomarkers and an immunotherapy, such as an immune checkpoint inhibitor (e.g., PD-1). Exemplary agents useful in such methods are described above. Such agents can be administered *in vitro* or *ex vivo* (e.g., by contacting the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the present invention provides methods useful for treating an individual afflicted with a condition that would benefit from an increased immune response, such as an infection or a cancer like colorectal cancer.

Agents that upregulate immune responses can be in the form of enhancing an existing immune response or eliciting an initial immune response. Thus, enhancing an immune response using the subject compositions and methods is useful for treating cancer, but can also be useful for treating an infectious disease (e.g., bacteria, viruses, or parasites), a parasitic infection, and an immunosuppressive disease.

Exemplary infectious disorders include viral skin diseases, such as Herpes or shingles, in which case such an agent can be delivered topically to the skin. In addition, systemic viral diseases, such as influenza, the common cold, and encephalitis might be alleviated by systemic administration of such agents. In one preferred embodiment, agents that upregulate the immune response described herein are useful for modulating the arginase/iNOS balance during *Trypanosoma cruzi* infection in order to facilitate a protective immune response against the parasite.

Immune responses can also be enhanced in an infected patient through an *ex vivo* approach, for instance, by removing immune cells from the patient, contacting immune cells *in vitro* with an agent described herein and reintroducing the *in vitro* stimulated immune cells into the patient.

In certain instances, it may be desirable to further administer other agents that upregulate immune responses, for example, forms of other B7 family members that transduce signals via costimulatory receptors, in order to further augment the immune response. Such additional agents and therapies are described further below.

5 Agents that upregulate an immune response can be used prophylactically in vaccines against various polypeptides (*e.g.*, polypeptides derived from pathogens). Immunity against a pathogen (*e.g.*, a virus) can be induced by vaccinating with a viral protein along with an agent that upregulates an immune response, in an appropriate adjuvant.

10 In another embodiment, upregulation or enhancement of an immune response function, as described herein, is useful in the induction of tumor immunity.

In another embodiment, the immune response can be stimulated by the methods described herein, such that preexisting tolerance, clonal deletion, and/or exhaustion (*e.g.*, T cell exhaustion) is overcome. For example, immune responses against antigens to which a subject cannot mount a significant immune response, *e.g.*, to an autologous antigen, such as a tumor specific antigens can be induced by administering appropriate agents described herein that upregulate the immune response. In one embodiment, an autologous antigen, such as a tumor-specific antigen, can be coadministered. In another embodiment, the subject agents can be used as adjuvants to boost responses to foreign antigens in the process of active immunization.

20 In one embodiment, immune cells are obtained from a subject and cultured *ex vivo* in the presence of an agent as described herein, to expand the population of immune cells and/or to enhance immune cell activation. In a further embodiment the immune cells are then administered to a subject. Immune cells can be stimulated *in vitro* by, for example, providing to the immune cells a primary activation signal and a costimulatory signal, as is known in the art. Various agents can also be used to costimulate proliferation of immune cells. In one embodiment immune cells are cultured *ex vivo* according to the method described in PCT Application No. WO 94/29436. The costimulatory polypeptide can be soluble, attached to a cell membrane, or attached to a solid surface, such as a bead.

30

6. Administration of Agents

The immune modulating agents of the invention are administered to subjects in a biologically compatible form suitable for pharmaceutical administration *in vivo*, to enhance

immune cell mediated immune responses. By “biologically compatible form suitable for administration *in vivo*” is meant a form to be administered in which any toxic effects are outweighed by the therapeutic effects. The term “subject” is intended to include living organisms in which an immune response can be elicited, *e.g.*, mammals. Examples of
5 subjects include humans, dogs, cats, mice, rats, and transgenic species thereof.

Administration of an agent as described herein can be in any pharmacological form including a therapeutically active amount of an agent alone or in combination with a pharmaceutically acceptable carrier.

Administration of a therapeutically active amount of the therapeutic composition of
10 the present invention is defined as an amount effective, at dosages and for periods of time necessary, to achieve the desired result. For example, a therapeutically active amount of an agent may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability of peptide to elicit a desired response in the individual. Dosage regimens can be adjusted to provide the optimum therapeutic response. For example,
15 several divided doses can be administered daily or the dose can be proportionally reduced as indicated by the exigencies of the therapeutic situation.

Inhibiting or blocking expression and/or activity of one or more biomarkers listed in Table 1, such as one or more regulators of antigen processing and presentation as described herein (*e.g.*, HLA-E, Pdla3, Calr, Tap1, Tap2, Tapbp, Erp1, etc.), alone or in combination
20 with an immunotherapy, can be accomplished by combination therapy with the modulatory agents described herein. Combination therapy describes a therapy in which such one or more biomarkers is inhibited or blocked with an immunotherapy simultaneously. This may be achieved by administration of the modulatory agent described herein with the immunotherapy simultaneously (*e.g.*, in a combination dosage form or by simultaneous
25 administration of single agents) or by administration of single inhibitory agent for such one or more biomarkers and the immunotherapy, according to a schedule that results in effective amounts of each modulatory agent present in the patient at the same time.

The therapeutic agents described herein can be administered in a convenient manner such as by injection (subcutaneous, intravenous, etc.), oral administration, inhalation,
30 transdermal application, or rectal administration. Depending on the route of administration, the active compound can be coated in a material to protect the compound from the action of enzymes, acids and other natural conditions which may inactivate the compound. For example, for administration of agents, by other than parenteral administration, it may be

desirable to coat the agent with, or co-administer the agent with, a material to prevent its inactivation.

An agent can be administered to an individual in an appropriate carrier, diluent or adjuvant, co-administered with enzyme inhibitors or in an appropriate carrier such as liposomes. Pharmaceutically acceptable diluents include saline and aqueous buffer solutions. Adjuvant is used in its broadest sense and includes any immune stimulating compound such as interferon. Adjuvants contemplated herein include resorcinols, non-ionic surfactants such as polyoxyethylene oleyl ether and n-hexadecyl polyethylene ether. Enzyme inhibitors include pancreatic trypsin inhibitor, diisopropylfluorophosphate (DEEP) and trasylol. Liposomes include water-in-oil-in-water emulsions as well as conventional liposomes (Sterna *et al.* (1984) *J. Neuroimmunol.* 7:27).

As described in detail below, the pharmaceutical compositions of the present invention may be specially formulated for administration in solid or liquid form, including those adapted for the following: (1) oral administration, for example, drenches (aqueous or non-aqueous solutions or suspensions), tablets, boluses, powders, granules, pastes; (2) parenteral administration, for example, by subcutaneous, intramuscular or intravenous injection as, for example, a sterile solution or suspension; (3) topical application, for example, as a cream, ointment or spray applied to the skin; (4) intravaginally or intrarectally, for example, as a pessary, cream or foam; or (5) aerosol, for example, as an aqueous aerosol, liposomal preparation or solid particles containing the compound.

The phrase “therapeutically-effective amount” as used herein means that amount of an agent that modulates (*e.g.*, inhibits) biomarker expression and/or activity, or expression and/or activity of the complex, or composition comprising an agent that modulates (*e.g.*, inhibits) biomarker expression and/or activity, or expression and/or activity of the complex, which is effective for producing some desired therapeutic effect, *e.g.*, cancer treatment, at a reasonable benefit/risk ratio.

The phrase “pharmaceutically acceptable” is employed herein to refer to those agents, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

The phrase “pharmaceutically-acceptable carrier” as used herein means a pharmaceutically-acceptable material, composition or vehicle, such as a liquid or solid

filler, diluent, excipient, solvent or encapsulating material, involved in carrying or transporting the subject chemical from one organ, or portion of the body, to another organ, or portion of the body. Each carrier must be “acceptable” in the sense of being compatible with the other ingredients of the formulation and not injurious to the subject. Some

5 examples of materials which can serve as pharmaceutically-acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil,

10 cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) phosphate buffer solutions; and (21) other non-toxic compatible substances employed in

15 pharmaceutical formulations.

The term “pharmaceutically-acceptable salts” refers to the relatively non-toxic, inorganic and organic acid addition salts of the agents that modulates (*e.g.*, inhibits) biomarker expression and/or activity, or expression and/or activity of the complex

20 encompassed by the present invention. These salts can be prepared *in situ* during the final isolation and purification of the therapeutic agents, or by separately reacting a purified therapeutic agent in its free base form with a suitable organic or inorganic acid, and isolating the salt thus formed. Representative salts include the hydrobromide, hydrochloride, sulfate, bisulfate, phosphate, nitrate, acetate, valerate, oleate, palmitate,

25 stearate, laurate, benzoate, lactate, phosphate, tosylate, citrate, maleate, fumarate, succinate, tartrate, naphthylate, mesylate, glucoheptonate, lactobionate, and laurylsulphonate salts and the like (See, for example, Berge *et al.* (1977) “Pharmaceutical Salts”, *J. Pharm. Sci.* 66:1-19).

In other cases, the agents useful in the methods of the present invention may contain

30 one or more acidic functional groups and, thus, are capable of forming pharmaceutically-acceptable salts with pharmaceutically-acceptable bases. The term “pharmaceutically-acceptable salts” in these instances refers to the relatively non-toxic, inorganic and organic base addition salts of agents that modulates (*e.g.*, inhibits) biomarker expression and/or

activity, or expression and/or activity of the complex. These salts can likewise be prepared in situ during the final isolation and purification of the therapeutic agents, or by separately reacting the purified therapeutic agent in its free acid form with a suitable base, such as the hydroxide, carbonate or bicarbonate of a pharmaceutically-acceptable metal cation, with ammonia, or with a pharmaceutically-acceptable organic primary, secondary or tertiary amine. Representative alkali or alkaline earth salts include the lithium, sodium, potassium, calcium, magnesium, and aluminum salts and the like. Representative organic amines useful for the formation of base addition salts include ethylamine, diethylamine, ethylenediamine, ethanolamine, diethanolamine, piperazine and the like (see, for example, Berge *et al.*, *supra*).

Wetting agents, emulsifiers and lubricants, such as sodium lauryl sulfate and magnesium stearate, as well as coloring agents, release agents, coating agents, sweetening, flavoring and perfuming agents, preservatives and antioxidants can also be present in the compositions.

Examples of pharmaceutically-acceptable antioxidants include: (1) water soluble antioxidants, such as ascorbic acid, cysteine hydrochloride, sodium bisulfate, sodium metabisulfite, sodium sulfite and the like; (2) oil-soluble antioxidants, such as ascorbyl palmitate, butylated hydroxyanisole (BHA), butylated hydroxytoluene (BHT), lecithin, propyl gallate, alpha-tocopherol, and the like; and (3) metal chelating agents, such as citric acid, ethylenediamine tetraacetic acid (EDTA), sorbitol, tartaric acid, phosphoric acid, and the like.

Formulations useful in the methods of the present invention include those suitable for oral, nasal, topical (including buccal and sublingual), rectal, vaginal, aerosol and/or parenteral administration. The formulations may conveniently be presented in unit dosage form and may be prepared by any methods well-known in the art of pharmacy. The amount of active ingredient which can be combined with a carrier material to produce a single dosage form will vary depending upon the host being treated, the particular mode of administration. The amount of active ingredient, which can be combined with a carrier material to produce a single dosage form will generally be that amount of the compound which produces a therapeutic effect. Generally, out of one hundred per cent, this amount will range from about 1 per cent to about ninety-nine percent of active ingredient, preferably from about 5 per cent to about 70 per cent, most preferably from about 10 per cent to about 30 per cent.

Methods of preparing these formulations or compositions include the step of bringing into association an agent that modulates (*e.g.*, inhibits) biomarker expression and/or activity, with the carrier and, optionally, one or more accessory ingredients. In general, the formulations are prepared by uniformly and intimately bringing into association
5 a therapeutic agent with liquid carriers, or finely divided solid carriers, or both, and then, if necessary, shaping the product.

Formulations suitable for oral administration may be in the form of capsules, cachets, pills, tablets, lozenges (using a flavored basis, usually sucrose and acacia or tragacanth), powders, granules, or as a solution or a suspension in an aqueous or non-
10 aqueous liquid, or as an oil-in-water or water-in-oil liquid emulsion, or as an elixir or syrup, or as pastilles (using an inert base, such as gelatin and glycerin, or sucrose and acacia) and/or as mouth washes and the like, each containing a predetermined amount of a therapeutic agent as an active ingredient. A compound may also be administered as a bolus, electuary or paste.

15 In solid dosage forms for oral administration (capsules, tablets, pills, dragees, powders, granules and the like), the active ingredient is mixed with one or more pharmaceutically-acceptable carriers, such as sodium citrate or dicalcium phosphate, and/or any of the following: (1) fillers or extenders, such as starches, lactose, sucrose, glucose, mannitol, and/or silicic acid; (2) binders, such as, for example, carboxymethylcellulose,
20 alginates, gelatin, polyvinyl pyrrolidone, sucrose and/or acacia; (3) humectants, such as glycerol; (4) disintegrating agents, such as agar-agar, calcium carbonate, potato or tapioca starch, alginic acid, certain silicates, and sodium carbonate; (5) solution retarding agents, such as paraffin; (6) absorption accelerators, such as quaternary ammonium compounds; (7) wetting agents, such as, for example, acetyl alcohol and glycerol monostearate; (8)
25 absorbents, such as kaolin and bentonite clay; (9) lubricants, such a talc, calcium stearate, magnesium stearate, solid polyethylene glycols, sodium lauryl sulfate, and mixtures thereof; and (10) coloring agents. In the case of capsules, tablets and pills, the pharmaceutical compositions may also comprise buffering agents. Solid compositions of a similar type may also be employed as fillers in soft and hard-filled gelatin capsules using
30 such excipients as lactose or milk sugars, as well as high molecular weight polyethylene glycols and the like.

A tablet may be made by compression or molding, optionally with one or more accessory ingredients. Compressed tablets may be prepared using binder (for example,

gelatin or hydroxypropylmethyl cellulose), lubricant, inert diluent, preservative, disintegrant (for example, sodium starch glycolate or cross-linked sodium carboxymethyl cellulose), surface-active or dispersing agent. Molded tablets may be made by molding in a suitable machine a mixture of the powdered peptide or peptidomimetic moistened with an inert liquid diluent.

Tablets, and other solid dosage forms, such as dragees, capsules, pills and granules, may optionally be scored or prepared with coatings and shells, such as enteric coatings and other coatings well-known in the pharmaceutical-formulating art. They may also be formulated so as to provide slow or controlled release of the active ingredient therein using, for example, hydroxypropylmethyl cellulose in varying proportions to provide the desired release profile, other polymer matrices, liposomes and/or microspheres. They may be sterilized by, for example, filtration through a bacteria-retaining filter, or by incorporating sterilizing agents in the form of sterile solid compositions, which can be dissolved in sterile water, or some other sterile injectable medium immediately before use. These compositions may also optionally contain opacifying agents and may be of a composition that they release the active ingredient(s) only, or preferentially, in a certain portion of the gastrointestinal tract, optionally, in a delayed manner. Examples of embedding compositions, which can be used include polymeric substances and waxes. The active ingredient can also be in micro-encapsulated form, if appropriate, with one or more of the above-described excipients.

Liquid dosage forms for oral administration include pharmaceutically acceptable emulsions, microemulsions, solutions, suspensions, syrups and elixirs. In addition to the active ingredient, the liquid dosage forms may contain inert diluents commonly used in the art, such as, for example, water or other solvents, solubilizing agents and emulsifiers, such as ethyl alcohol, isopropyl alcohol, ethyl carbonate, ethyl acetate, benzyl alcohol, benzyl benzoate, propylene glycol, 1,3-butylene glycol, oils (in particular, cottonseed, groundnut, corn, germ, olive, castor and sesame oils), glycerol, tetrahydrofuryl alcohol, polyethylene glycols and fatty acid esters of sorbitan, and mixtures thereof.

Besides inert diluents, the oral compositions can also include adjuvants such as wetting agents, emulsifying and suspending agents, sweetening, flavoring, coloring, perfuming and preservative agents.

Suspensions, in addition to the active agent may contain suspending agents as, for example, ethoxylated isostearyl alcohols, polyoxyethylene sorbitol and sorbitan esters,

microcrystalline cellulose, aluminum metahydroxide, bentonite, agar-agar and tragacanth, and mixtures thereof.

Formulations for rectal or vaginal administration may be presented as a suppository, which may be prepared by mixing one or more therapeutic agents with one or more suitable
5 nonirritating excipients or carriers comprising, for example, cocoa butter, polyethylene glycol, a suppository wax or a salicylate, and which is solid at room temperature, but liquid at body temperature and, therefore, will melt in the rectum or vaginal cavity and release the active agent.

Formulations which are suitable for vaginal administration also include pessaries,
10 tampons, creams, gels, pastes, foams or spray formulations containing such carriers as are known in the art to be appropriate.

Dosage forms for the topical or transdermal administration of an agent that modulates (*e.g.*, inhibits) biomarker expression and/or activity include powders, sprays, ointments, pastes, creams, lotions, gels, solutions, patches and inhalants. The active
15 component may be mixed under sterile conditions with a pharmaceutically-acceptable carrier, and with any preservatives, buffers, or propellants which may be required.

The ointments, pastes, creams and gels may contain, in addition to a therapeutic agent, excipients, such as animal and vegetable fats, oils, waxes, paraffins, starch, tragacanth, cellulose derivatives, polyethylene glycols, silicones, bentonites, silicic acid,
20 talc and zinc oxide, or mixtures thereof.

Powders and sprays can contain, in addition to an agent that modulates (*e.g.*, inhibits) biomarker expression and/or activity, excipients such as lactose, talc, silicic acid, aluminum hydroxide, calcium silicates and polyamide powder, or mixtures of these substances. Sprays can additionally contain customary propellants, such as
25 chlorofluorohydrocarbons and volatile unsubstituted hydrocarbons, such as butane and propane.

The agent that modulates (*e.g.*, inhibits) biomarker expression and/or activity, can be alternatively administered by aerosol. This is accomplished by preparing an aqueous aerosol, liposomal preparation or solid particles containing the compound. A nonaqueous
30 (*e.g.*, fluorocarbon propellant) suspension could be used. Sonic nebulizers are preferred because they minimize exposing the agent to shear, which can result in degradation of the compound.

Ordinarily, an aqueous aerosol is made by formulating an aqueous solution or suspension of the agent together with conventional pharmaceutically acceptable carriers and stabilizers. The carriers and stabilizers vary with the requirements of the particular compound, but typically include nonionic surfactants (Tweens, Pluronic, or polyethylene glycol), innocuous proteins like serum albumin, sorbitan esters, oleic acid, lecithin, amino acids such as glycine, buffers, salts, sugars or sugar alcohols. Aerosols generally are prepared from isotonic solutions.

Transdermal patches have the added advantage of providing controlled delivery of a therapeutic agent to the body. Such dosage forms can be made by dissolving or dispersing the agent in the proper medium. Absorption enhancers can also be used to increase the flux of the peptidomimetic across the skin. The rate of such flux can be controlled by either providing a rate controlling membrane or dispersing the peptidomimetic in a polymer matrix or gel.

Ophthalmic formulations, eye ointments, powders, solutions and the like, are also contemplated as being within the scope of this invention.

Pharmaceutical compositions of this invention suitable for parenteral administration comprise one or more therapeutic agents in combination with one or more pharmaceutically-acceptable sterile isotonic aqueous or nonaqueous solutions, dispersions, suspensions or emulsions, or sterile powders which may be reconstituted into sterile injectable solutions or dispersions just prior to use, which may contain antioxidants, buffers, bacteriostats, solutes which render the formulation isotonic with the blood of the intended recipient or suspending or thickening agents.

Examples of suitable aqueous and nonaqueous carriers which may be employed in the pharmaceutical compositions of the present invention include water, ethanol, polyols (such as glycerol, propylene glycol, polyethylene glycol, and the like), and suitable mixtures thereof, vegetable oils, such as olive oil, and injectable organic esters, such as ethyl oleate. Proper fluidity can be maintained, for example, by the use of coating materials, such as lecithin, by the maintenance of the required particle size in the case of dispersions, and by the use of surfactants.

These compositions may also contain adjuvants such as preservatives, wetting agents, emulsifying agents and dispersing agents. Prevention of the action of microorganisms may be ensured by the inclusion of various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol sorbic acid, and the like. It may also

be desirable to include isotonic agents, such as sugars, sodium chloride, and the like into the compositions. In addition, prolonged absorption of the injectable pharmaceutical form may be brought about by the inclusion of agents which delay absorption such as aluminum monostearate and gelatin.

5 In some cases, in order to prolong the effect of a drug, it is desirable to slow the absorption of the drug from subcutaneous or intramuscular injection. This may be accomplished by the use of a liquid suspension of crystalline or amorphous material having poor water solubility. The rate of absorption of the drug then depends upon its rate of dissolution, which, in turn, may depend upon crystal size and crystalline form.

10 Alternatively, delayed absorption of a parenterally-administered drug form is accomplished by dissolving or suspending the drug in an oil vehicle.

 Injectable depot forms are made by forming microencapsule matrices of an agent that modulates (*e.g.*, inhibits) biomarker expression and/or activity, in biodegradable polymers such as polylactide-polyglycolide. Depending on the ratio of drug to polymer, and the nature of the particular polymer employed, the rate of drug release can be
15 controlled. Examples of other biodegradable polymers include poly(orthoesters) and poly(anhydrides). Depot injectable formulations are also prepared by entrapping the drug in liposomes or microemulsions, which are compatible with body tissue.

 When the therapeutic agents of the present invention are administered as
20 pharmaceuticals, to humans and animals, they can be given per se or as a pharmaceutical composition containing, for example, 0.1 to 99.5% (more preferably, 0.5 to 90%) of active ingredient in combination with a pharmaceutically acceptable carrier.

 Actual dosage levels of the active ingredients in the pharmaceutical compositions of this invention may be determined by the methods of the present invention so as to obtain an
25 amount of the active ingredient, which is effective to achieve the desired therapeutic response for a particular subject, composition, and mode of administration, without being toxic to the subject.

 The nucleic acid molecules of the present invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for
30 example, intravenous injection, local administration (see U.S. Pat. No. 5,328,470) or by stereotactic injection (see *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the

gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

In one embodiment, an agent of the invention is an antibody. As defined herein, a therapeutically effective amount of antibody (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The skilled artisan will appreciate that certain factors may influence the dosage required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of an antibody can include a single treatment or, preferably, can include a series of treatments. In a preferred example, a subject is treated with antibody in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody used for treatment may increase or decrease over the course of a particular treatment. Changes in dosage may result from the results of diagnostic assays.

7. Kits

The present invention also encompasses kits for detecting and/or modulating biomarkers described herein. A kit of the present invention may also include instructional materials disclosing or describing the use of the kit or an antibody of the disclosed invention in a method of the disclosed invention as provided herein. A kit may also include additional components to facilitate the particular application for which the kit is designed. For example, a kit may additionally contain means of detecting the label (*e.g.*, enzyme substrates for enzymatic labels, filter sets to detect fluorescent labels, appropriate secondary labels such as a sheep anti-mouse-HRP, *etc.*) and reagents necessary for controls (*e.g.*, control biological samples or standards). A kit may additionally include buffers and other reagents recognized for use in a method of the disclosed invention. Non-limiting examples include agents to reduce non-specific binding, such as a carrier protein or a detergent.

EXAMPLES**Example 1: Materials and methods for Examples 2-5****a. *In vivo* CRISPR screening in B16 tumor cells**

A Cas9-expressing version of the B16 melanoma cell line was created and confirmed that it could edit DNA efficiently with CRISPR using sgRNAs targeting the PD-L1 gene. For screening the B16-Cas9 cell line, a library of 9,992 optimized sgRNAs was created to target 2,398 genes, selected from the GO term categories: kinase, phosphatase, cell surface, plasma membrane, antigen processing and presentation, immune system process, and chromatin remodeling. The transcript abundance of the genes in these categories were then filtered to include only those that were expressed $> \text{RPKM} (\log_2) = 0.9$. These genes were then ranked for expression in the B16 cell line using RNAseq to select for the top 2,398 expressed genes. The library was divided into 4 sub-pools, each containing one sgRNA per gene and 100 non-targeting control sgRNAs. The 4 sub-pools were screened individually and sgRNAs were delivered to B16-Cas9 cells via lentiviral infection at an infection rate of 30%. Transduced B16 cells were purified using a hCD19 reporter by positive magnetic selection (Miltenyi Biotech, Cambridge, MA) and then expanded *in vitro* before being implanted into animals. For each sub-pool, B16 cells were implanted into 10 TCR $\alpha^{-/-}$ mice, 10 WT mice treated with GVAX, and 10 WT mice treated with GVAX and PD-1 blockade (see below for treatment protocols). B16 cells transduced with libraries were also grown *in vitro* at approximately 2000 x library coverage for the same time period as the animal experiment. Mice were sacrificed 12-14 days after tumor implantation tumor genomic DNA was prepared from whole tumor tissue using the Qiagen DNA Blood Midi kit (Qiagen, Hilden, Germany). PCR was used to amplify the sgRNA region and sequencing to determine sgRNA abundance was performed on an Illumina HiSeq system (Illumina, San Diego, CA). Significantly enriched or depleted sgRNAs from any comparison of conditions were identified using the STARS algorithm (Doench *et al.* (2014) *Nat. Biotechnol.* 32:1262-1267).

b. Animal treatment and tumor challenges

The designs of these animal studies and procedures were approved by the Dana Farber Cancer Institute IACUC committee. Dana Farber's specific-pathogen free facility was used for the storage and care of all mice. Seven-week old wild-type female C57BL/6J mice were obtained from Jackson laboratories (Bar Harbor, ME). A colony of B6.129S2-Tcr $^{tm1Mom}/J$ (Tcr α) T cell-deficient mice were bred on site. Mice were aged matched to be

7-12 weeks old at the time of tumor inoculation. For screening, 2.0×10^6 library-transduced B16-Cas9 cells resuspended in Hanks Balanced Salt Solution (Gibco, Thermo Fisher Scientific, Waltham, MA) were mixed 1:1 by volume with Matrigel® (Corning, Corning, NY) and subcutaneously injected into the right flank on day 0. Mice were vaccinated with 1.0×10^6 GM-CSF-secreting B16 (GVAX) cells that had been irradiated with 3500 Gy on days 1 and 4 to elicit an anti-tumor immune response. Subsequently, mice were treated with 100 µg of rat monoclonal anti-PD1 antibody (clone: 29F.1A12) on days 9 and 12 via intraperitoneal injection. For validation assays, 1.0×10^6 tumor cells were subcutaneously injected into the right flank without matrigel. Tumors were measured every 3 days beginning on day 6 after challenge until time of death. Measurements were taken manually by collecting the longest dimension (length) and the longest perpendicular dimension (width). Tumor volume was estimated with the formula: $(L \times W^2)/2$. CO₂ inhalation was used to euthanize mice on the day of sacrifice.

c. Creation of CRISPR edited tumor cell lines

Transient transfection of Cas9-sgRNA plasmid (pX459, Addgene, Cambridge, MA) was used to edit B16 and *Braf/Pten* melanoma cell lines. pX459 was digested with the enzyme Bpil (Thermo Fisher Scientific) as per the manufacturer's instructions and sgRNA oligos were cloned in using standard molecular cloning. For B16 cells, 5×10^5 cells were plated in a well of a 6-well plate and were transfected the following day using 2 µg of pX459 plasmid DNA and Turbofect™ (3:1 ratio, Thermo Fisher Scientific). Twenty-four hours after transfection, transfectants were selected in puromycin (6 µg/mL, Thermo Fisher Scientific). For *Braf/Pten* melanoma cells, 5×10^5 cells were plated in a well of a 6-well plate and were transfected the following day using 4 µg of pX459 plasmid DNA and Turbofect™ (3:1 ratio). After selection, cells were grown for 14 days *in vitro* before being implanted into mice.

d. In vivo competition assays

B16 cells were engineered to express GFP or TdTomato by lentiviral transduction to differentiate populations. Cas9-target sgRNA-transfected cells and Cas9-control sgRNA-transfected cells were mixed and then grown for at least two passages *in vitro* before implantation into mice. Mixes were analyzed by flow cytometry on the day of tumor inoculation. Mice were euthanized 15-21 days after tumor inoculation for tumor harvest. Tumors were macerated on ice and incubated in collagenase P (2 mg/mL, Sigma, St. Louis, MO) and DNase I (50 µg/mL, Sigma) supplemented DMEM for 10 minutes at 37°C. After

incubation, tumor cells were passed through 70µm filters to remove undigested tumor. Tumor cells were washed with ice-cold MACS media and stained with Near-IR LIVE/DEAD® (1:1000, BD Biosciences, Franklin Lakes, NJ) for 10 minutes on ice. Tumor cells were then washed and resuspended in ice-cold PBS with 2% FBS. An Accuri™ C6
5 flow cytometry system (BD Biosciences) was used to analyze final GFP/TdTomato tumor cell ratios.

e. In vitro cytokine stimulations

Differentially labeled B16 cells were plated in 12-well plates in DMEM + 10% FBS containing the indicated combinations of cytokines: TNFα (10 ng/mL, PeproTech, Rocky
10 Hill, NJ), IFNγ (100 ng/mL, Cell Signaling Technology, Danvers, MA), and IFNβ (1000 U/mL, Pbl Assay Science, Piscataway Township, NJ). Cells were passaged every 3 or 4 days. On day 14, tumor cells were analyzed by flow cytometry for changes in the ratio of GFP⁺/TdTomato⁺ cells.

f. In vitro T cell killing assays

Differentially labeled, ovalbumin-expressing B16 cells were plated onto 12 well
15 plates in DMEM + 10% FBS containing IFNβ at 50 U/mL. The following day, media was removed and RPMI + 10% FBS was added containing preactivated OT-1 T cells at the indicated ratio. Tumor cells were passaged and fresh OT-1 T cells were added every 2 days until day 6. On day 6, tumor cells were analyzed by flow cytometry for changes in the ratio
20 of GFP⁺/TdTomato⁺ cells.

g. Analysis tumor-infiltrating lymphocytes by flow cytometry

Mice were injected subcutaneously (s.c.) with 1.0×10^6 CRISPR/Cas9 modified B16 cells and treated with GVAX + anti-PD-1 mAb as described above. Tumors were harvested on day 12-13, weighed, mechanically diced, incubated with collagenase P (2
25 mg/mL, Sigma Aldrich) and DNase I (50 µg/mL, Sigma Aldrich) for 10 minutes, and pipetted into a single-cell suspension. After filtering through a 70 µm filter, cells were blocked with anti-mouse CD16/32 antibody (BioLegend, San Diego, CA) and stained with indicated antibodies for 30 minutes on ice. Dead cells were excluded using Aqua Live/Dead® (1:1000, ThermoFisher Scientific) added concurrently with surface antibodies.
30 After washing, cells were fixed with Foxp3/transcription factor staining buffer set (eBiosciences, San Diego, CA) as per manufacturer's instructions, blocked with mouse and rat serum, then stained with intracellular antibodies. AccuCount Fluorescent particles (Spherotech, Lake Forest, IL) were added for cell quantification prior to analysis on an LSR

Fortessa™ cell analyzer (BD Biosciences) using single-color compensation controls and fluorescence-minus-one thresholds to set gate margins. Comparisons between groups performed using Student's t test.

h. Flow cytometry analysis of B16 tumor cells

5 B16 cells were trypsinized and washed in PBS + 2% FBS, stained with antibodies for cell surface proteins as per the manufacturer's instructions and then analyzed on an Accuri™ C6 flow cytometry system (BD Biosciences).

i. RNAseq analysis of tumor cells

10 Null or control sgRNA-transfected B16 cells were stimulated with IFN γ (100 ng/mL, Cell Signaling Technology), TNF α (10 ng/mL, Peprotech) or both for 48 hours. RNA was extracted from cell pellets using the Qiagen RNeasy Mini kit according to manufacturer's instructions. First-strand Illumina-barcoded libraries were generated using the NEB RNA Ultra™ Directional kit according to manufacturer's instructions, including a 12-cycle PCR enrichment. Libraries were sequenced on an Illumina NextSeq™ 500
15 instrument using paired-end 37bp reads. Data were trimmed for quality using the Trimmomatic pipeline with the following parameters: LEADING:15 TRAILING:15 SLIDINGWINDOW:4:15 MINLEN:16. Data were aligned to mouse reference genome mm10 using the Bowtie 2 aligning sequencing tool (available at the World Wide Web website of Johns Hopkins University). HTSeq was used to map aligned reads to genes and
20 to generate a gene count matrix and it is available at the World Wide Web address of www.huber.embl.de/users/anders/HTSeq/doc/overview.html. Normalized counts and differential expression analysis was performed using the DESeq2 R package. The gene set enrichment analysis was performed as described previously in Subramanian *et al.* (2005) *Proc. Natl Acad Sci USA* 102:15545-15550. Principle Components Analysis (PCA) was performed on
25 the normalized gene counts including all genes that passed a minimal expression filter. Signature scores for the individual samples were generated using FastProject (available at the World Wide Web address of bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-016-1176-5) and the Hallmark gene signature collection (Liberzon *et al.* (2015) *Cell Sys.* 1:417-425). Pearson
30 correlation coefficients were calculated between the Hallmark gene signatures and PC1 and PC2. Selected signatures were plotted on a normalized PCA projection of the dataset.

j. Western blotting

Whole cell lysates were prepared in lysis buffer (60 mM Tris HCl, 2% SDS, 10% glycerol, complete EDTA-free protease-inhibitor (Roche, Basel, Switzerland), and 500 U/mL benzonase nuclease (Novagen, Merck, Darmstadt, Germany)). Samples were boiled at 100°C and clarified by centrifugation. Protein concentration was measured with a BCA protein assay kit (Pierce, Dallas, Texas). Fifty to one hundred and fifty micrograms of protein was loaded on 4-12% Bolt® Bis-Tris Plus gels (Life Technologies, Carlsbad, CA) in MES buffer (Life Technologies). Protein was transferred to 0.45 µm nitrocellulose membranes (Bio-Rad, Hercules, CA). Membranes were blocked in Tris-buffered saline plus 0.1% Tween 20 (TBS-T) containing 5% non-fat dry milk for 1 hour at room temperature followed by overnight incubation with primary antibody at 4°C. Membranes were washed with TBS-T and incubated with HRP-conjugated secondary antibodies for 1 hour at room temperature. HRP was activated with Supersignal® West Dura Extended Duration Substrate (Pierce) and visualized with a chemiluminescent detection system using Fuji ImageQuant™ LAS4000 (GE Healthcare Life Sciences, Pittsburgh, PA). Blots were then analyzed using ImageJ and Adobe® Photoshop® software.

k. Immunohistochemistry

Whole B16 tumors were fixed for 24 hours in 10% neutral buffered formalin and then permeabilized in 70% ethanol overnight. Fixed tissue was then embedded into paraffin, sectioned and then mounted onto slides for staining with mouse CD8α (eBiosciences Cat. #14-0808-82). Slides were imaged on a Leica Scanscope XT and analyzed using Aperio software (Leica Biosystems, Buffalo Grove, IL).

l. Antibodies

For Western blotting, primary antibodies against β-ACTIN (Abcam, Cambridge, UK, Cat. #8227), and FLAG (clone M2, Sigma Aldrich) were used. Peroxidase-conjugated secondaries against Rabbit-IgG (Cat. #111-035-046) and Mouse-IgG (Cat. #115-035-174) were purchased from Jackson Laboratories (Bar Harbor, ME).

For flow cytometry, the following anti-mouse (m) fluorochrome-conjugated antibodies were used: H2K(b)/H2D(b) (clone 28-8-6, BioLegend), CD47 (clone miap301, BioLegend), SIINFEKL-H2K(b) (clone 25-D1.16, BioLegend), Granzyme B (clone GB11, BioLegend), TNF (clone MP6-XT22, BioLegend), IFNγ (clone XMGI.2, BioLegend), CD8α (clone 53-6.7, BioLegend), CD4 (clone RM4-5 or GK15, BioLegend), TCR-β (clone H57-597, BioLegend), PD-1 (clone RPMI-30, BioLegend), Tim-3 (clone TMR3-2.3, BioLegend), CD45 (clone 104 or 30-F11, BioLegend), Ly6C (clone HK1.4, BioLegend), I-

A/I-E (clone M5/114.15.2, BioLegend), F4/80 (clone BM8, BioLegend), CD11c (clone N418, BioLegend), CD24 (clone M1/69, BioLegend), CD11b (clone M1/70, BioLegend), CD103 (clone 2E7, BioLegend), CD3 ϵ (clone 145-2C11, BioLegend), TCR γ/δ (clone GL3, BioLegend), NK1.1 (clone PK136, BioLegend), CD44 (clone IM7, BioLegend), Ki-67
 5 (clone B56, BD Biosciences), CD274 (clone MIH5, BD Biosciences), and Foxp3 (clone JFK-16s, eBioscience).

n. CRISPR sgRNA sequences.

CRISPR sgRNA sequences uses are described in Table 2 below.

Table 2

10	Gene Name/sg#	sgRNA Sequence	
	<i>Cd274</i> sgRNA 1	GCCTGCTGTCACCTTGCTACG	(SEQ ID NO: 49)
	<i>Cd274</i> sgRNA 2	AATCAACCAGAGAATTTCCG	(SEQ ID NO: 50)
	<i>Cd274</i> sgRNA 3	GGTCCAGCTCCCGTTCTACA	(SEQ ID NO: 51)
	<i>Cd274</i> sgRNA 4	GTATGGCAGCAACGTCACGA	(SEQ ID NO: 52)
15	<i>Cd47</i> sgRNA 1	TATAGAGCTGAAAAACCGCA	(SEQ ID NO: 53)
	<i>Cd47</i> sgRNA 2	CCACATTACGGACGATGCAA	(SEQ ID NO: 54)
	<i>Cd47</i> sgRNA 3	TCTTACGAGGAGGAGAAAGG	(SEQ ID NO: 55)
	<i>Cd47</i> sgRNA 4	GCAAGTGTAAGTTTCCCACCA	(SEQ ID NO: 56)
	control sgRNA 1	GCGAGGTATTCGGCTCCGCG	(SEQ ID NO: 57)
20	control sgRNA 2	GCTTTCACGGAGGTTTCGACG	(SEQ ID NO: 58)
	control sgRNA 3	ATGTTGCAGTTCGGCTCGAT	(SEQ ID NO: 59)
	control sgRNA 4	ACGTGTAAGGCGAACGCCTT	(SEQ ID NO: 60)
	control sgRNA 5	ATTGTTGACCGTCTACGGG	(SEQ ID NO: 61)
	<i>H2-T23</i> sgRNA 1	TAGCCGACAATGATGAACCG	(SEQ ID NO: 62)
25	<i>H2-T23</i> sgRNA 2	AGGCCTCCTGACAATACCCG	(SEQ ID NO: 63)
	<i>H2-T23</i> sgRNA 3	GGAACCTCAGAGTAAACCTG	(SEQ ID NO: 64)

Examples 2-5 disclose the development of a pooled loss-of-function *in vivo* genetic
 screening approach that uses CRISPR-Cas9 genome editing to discover genes that increase
 30 sensitivity or cause resistance to immunotherapy in a mouse transplantable tumor model.
 About 2,400 genes expressed by tumor cells were screened in the B16 murine melanoma
 model to identify those that increase or decrease sensitivity to immunotherapy with tumor
 vaccination and PD-1 checkpoint blockade. The screen identified known immune evasion

molecules PD-L1 and CD47, as tumor cells bearing sgRNAs for these targets were significantly depleted in animals treated with immunotherapy. In contrast, loss of function of any of the genes that sense or signal in response to interferon- γ (IFN γ) rendered cells resistant to immunotherapy with PD-1 checkpoint blockade and vaccination, recapitulating resistance mutations identified in melanoma patients (Zaretsky *et al.* (2016) *N. Engl. J. Med.* 375:819-829; Gao *et al.* (2016) *Cell* 167:397-404.e9). It was discovered and further validated that deletion of multiple new genes, such as those involved in antigen presentation (*e.g.*, *HLA-E/H2-T23*, *Pdia3*, *Calr*, *Tap2*, *Tap1*, *Erap1*, *Tapbp*, etc.), sensitizes tumor cells to immunotherapy. These findings reveal that therapeutic strategies, such as *HLA-E* inhibition, that sensitize tumor cells to the effects of IFN γ are capable of increasing the efficacy of cancer immunotherapy. Moreover, this screening approach can discover new immunotherapy targets and prioritize their combination with existing immunotherapies.

Example 2: A pooled loss-of-function *in vivo* genetic screen recovers known immune evasion molecules expressed by tumors

In order to systematically identify new cancer immunotherapy targets and resistance mechanisms, a pooled genetic screening approach was developed to identify genes that increase or decrease the fitness of tumor cells growing *in vivo* in animals treated with immunotherapy (Figure 1A). First, a B16 melanoma cell line was engineered to express Cas9 (Figure 2A), confirmed of efficient DNA editing using sgRNAs targeting PD-L1 (Figure 1D, bottom). Next, a library of lentiviral vectors was created to encode 9,992 sgRNAs targeting 2,398 genes from relevant functional classes that were expressed at detectable levels in the tumor cell line (Figure 2B). After transduction and *in vitro* passage to allow gene editing to take place, the tumor cells were transplanted into animals that were then treated with either a GM-CSF-secreting, irradiated tumor cell vaccine (GVAX) or GVAX plus PD-1 blockade using a monoclonal antibody for PD-1, in order to apply immune selective pressure on the tumor cells (Figure 1B) (see Dranoff (2003) *Oncogene* 22:3188-3192; Dranoff *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:3539-3543; and Duraiswamy *et al.* (2013) *Cancer Res.* 73:3591-3603; Curran & Allison (2009) *Cancer Res* 69:7747-7755; Curran *et al.* (2010) *Proc. Natl. Acad. Sci. U.S.A.* 107:4275-4280). In parallel, the library-transduced tumor cells were transplanted into TCR $\alpha^{-/-}$ mice, which lack CD4 $^{+}$ and CD8 $^{+}$ T cells and were therefore unable to apply adaptive immune selective pressure on the tumors. This allowed to distinguish the effect of immune selective pressure

on library representation from nonspecific effects on tumor cell viability. After 12 - 14 days, tumors were harvested (Figure 1B), with all sgRNAs recovered from each animal with good inter-animal reproducibility (Figures 2C-2E).

The library representation in tumors from immunotherapy-treated wild-type (WT) animals were compared with that found in tumors growing in $\text{TCR}\alpha^{-/-}$ mice, in which deletion of genes that result in resistance to immunotherapy would be expected to increase tumor sgRNA representation in WT animals, while deletion of genes that result in increased sensitivity of tumors to immunotherapy would decrease sgRNA representation. Analysis of sgRNAs enriched by immune selective pressure revealed those targeting genes involved in cytokine-mediating signaling and immune-system processes (Figure 1C). sgRNAs depleted by immunotherapy included those targeting genes involved in antigen processing, necroptosis, and regulation of immune responses (Figure 1C). These results indicate that the genetic screening approach used here identified genes expressed by tumors cells that play a role in interaction with the immune system.

Inspection of the list of genes targeted by sgRNAs depleted from tumors treated with immunotherapy revealed the known immune evasion molecule PD-L1, indicating that loss of PD-L1 increased the sensitivity of tumor cells to immune attack. sgRNAs targeting PD-L1 were not depleted from tumors in $\text{TCR}\alpha^{-/-}$ mice relative to cells growing *in vitro*, presumably due to the absence of T cell-mediated selective pressure (Figure 1D), but were significantly depleted in WT mice treated with GVAX relative to $\text{TCR}\alpha^{-/-}$ mice (FDR = 0.004). However, the depletion of PD-L1-targeting sgRNAs seen in GVAX-treated tumors was not observed in tumors treated with GVAX and anti-PD-1, indicating that loss of PD-L1 does not confer a selective disadvantage to tumors when PD-L1:PD-1 interactions are blocked (Figure 1D).

It was also found that sgRNAs targeting CD47, which enables immune evasion by impairing engulfment of tumors cells by phagocytes (as in Liu *et al.* (2015) *Nat. Med.* 21:1209-1215; Weiskopf *et al.* (2016) *J. Clin. Invest.* 126:2610-2620; and Tseng *et al.* (2013) *Proc. Natl. Acad. Sci. U.S.A.* 110: 11103-11108), were markedly depleted in tumors treated with either GVAX or with GVAX plus PD-1 blockade (FDR = 0.005, 0.002 respectively) (Figure 1E). To confirm that CD47 null tumors were more susceptible to GVAX and PD-1 blockade, CD47 null B16 melanoma cells were generated using transient transfection of a Cas9-sgRNA plasmid (as in Ran *et al.* (2013) *Nat. Protoc.* 8:2281-2308)

(Figure 2F). It was found that loss of CD47 significantly improved control of tumor growth mediated by GVAX plus anti-PD-1 immunotherapy (Figure 1F, $p < 0.01$).

Using the pooled loss-of-function *in vivo* genetic screen for identifying immune evasion molecules expressed by tumors described above, genes that increase or decrease the fitness of MC38 colon cancer cells growing *in vivo* in animal treated with immunotherapy. Thus, *in vivo* genetic screening recovered genes known to confer tumor evasion properties on cancer cells.

Example 3: Discovery of novel gene targets to increase the efficacy of immunotherapy

Deletion of novel candidate immunotherapy targets was found to increase sensitivity of tumor cells to immunotherapy. sgRNAs targeting genes involved in antigen processing and presentation (*e.g.*, *HLA-E/H2-T23*, *Pdia3*, *Calr*, *Tap2*, *Tap1*, *Erap1*, *Tapbp*, etc.) were markedly depleted in mice treated with GVAX and PD-1 blockade (Figure 3A) relative to growth in TCR $\alpha^{-/-}$ mice. In many cases, multiple members of the same pathway (*e.g.*, *Erap1*, *Tap2*, *Calr*) or even the same multi-protein complex were depleted under immune selective pressure, underscoring the importance of these diverse biological pathways.

Example 4: *In vivo* validation of HLA-E as a target for combination with immunotherapy

Representative genes affecting antigen processing and presentation, *e.g.*, *HLA-E/H2-T23* (a non-classical MHC-I gene), were selected to validate based on their highest cumulative score as ranked by the STARS algorithm (Doench *et al.* (2014) *Nat. Biotechnol.* 32:1262-1267). *In vivo* competition assays showed that tumor cells deleted for *H2-T23* were strongly selected against in WT animals treated with immunotherapy but grew at equivalent rates to control tumor cells in vitro and in TCR $\alpha^{-/-}$ mice (Figure 3C and 3D; $p < 0.01$, Student's t test). This indicates that these genes are synthetically lethal with an anti-tumor immune response, rendering tumor cells more sensitive to immunotherapy but not altering their cell growth or survival in the absence of T cells.

H2-T23 encodes Qa-1b (HLA-E in humans), a non-classical MHC molecule that binds the inhibitory receptor NKG2A on T cells and NK cells (Wada *et al.* (2004) *Eur J Immunol.* 34:81-90). It was confirmed that loss of function of Qa-1b in tumor cells enhances tumor immunity, by comparing the growth of H2-T23 null B16 melanoma cells

(Figure 3B) to control B16 tumors in mice treated with GVAX and PD-1 blockade. Immunotherapy in control tumors eradicated only 1 out of 10 tumors. In contrast, immunotherapy of H2-T23 null tumors was curative in all 10 animals (Figures 3E-3F; $p < 0.05$, Student's t test). These results indicate that Qa-1b functions as an immune evasion molecule in tumors and that its loss of function improves tumor immunity.

Example 5: Mechanism studies of HLA-E-induced sensitivity to checkpoint blockade

HLA-E is a non-classical MHC-I gene expressed on the surface of both hematopoietic and non-hematopoietic cells. HLA-E can present antigen (Hansen *et al.* (2016) *Science* 351:714-720), but in addition is also thought to have tolerogenic properties (Morandi & Pistoia (2014) *Front. Immunol.* 5:394). *In vivo* screening using mouse transplantable tumor models treated with anti-PD-1 immunotherapy showed that H2-T23, the mouse ortholog of HLA-E, increased the sensitivity of tumor cells to immunotherapy and synergized with PD-1 pathway blockade.

HLA-E expression on tumors has been variably associated with good and bad prognosis (Benevolo *et al.* (2011) *J. Transl. Med.* 9:184; Seliger *et al.* (2016) *Oncotarget* 7:67360-67372; Yazdi *et al.* (2016) *Oncotarget* 7:3477-3488) and therefore its role as an immune evasion molecule has previously been open to question. Recent reports suggest that the molecule restrains the positive prognostic impact of having tumor infiltrating CD8⁺ T cells (Yazdi *et al.* (2016) *Oncotarget* 7:3477-3488). Like other HLA molecules, expression of HLA-E on tumor cells is induced by interferons. Additionally, polymorphisms that encode single amino acid changes in the invariant peptide that normally occupies the HLA-E binding groove (derived from a leader sequence in HLA-B) can substantially alter the stability of HLA-E complex and impact its expression level on the cell surface (Horowitz *et al.* (2016) *Sci. Immunol.* 1:eaag1672).

HLA-E has been shown to interact with both NKG2C and NKG2A. NKG2C is a cell-surface receptor expressed on NK cells and ligation of this receptor has been shown to have a stimulatory effect on NK cells (Wada *et al.* (2004) *Eur. J. Immunol.* 34:81-90). NKG2A is an inhibitory cell-surface receptor expressed on both NK cells and T cells. Due to fact that NKG2A has a greater affinity for HLA-E, this interaction is thought to be the dominant one and HLA-E is thus considered to have a tolerogenic role in immunity (Wada *et al.* (2004), *supra*). Blockade of the interaction between HLA-E and NKG2A has been

shown to increase NK-cell mediated lysis in vitro (Levy *et al.* (2009) *Innate Immun.* 15:91-100) and NKG2A expression on antigen-specific T cells during viral infection limits their cytolytic abilities (Moser *et al.* (2002) *Nat. Immunol.* 3:189-195).

It has been demonstrated herein that H2-T23 expression on murine tumor cells is an important mechanism for immune evasion. H2-T23 deficient B16 melanoma tumor cells were completely rejected when implanted into mice that are treated with GVAX and PD-1 blocking antibodies, whereas control B16 tumors respond to treatment transiently and then continue to grow. The data indicate that, like PD-L1, H2-T23 is a major inhibitory ligand that protects B16 tumor cells from destruction by the immune system. Overexpression of HLA-E/H2-T23 correlates with worse prognosis, at least partly due to its neutralization of beneficial effects of T cell infiltration into the tumor. The data are consistent with another discovery that H2-T23 deficient tumor cells had a selective disadvantage compared to control tumor cells when the two populations were injected into mice as a mixture. Thus, the data indicate that depletion of H2-T23 deficient tumor cells is carried out by cytotoxic T cells, as the effect of depletion was not observed in TCR α KO mice that lack conventional T cells. Finally, when a monoclonal antibody (anti-NK1.1 mAb) was used to deplete NK cells, H2-T23 deficient tumor cells were still depleted from mice treated with immunotherapy, further indicating that the effect is mediated by cytotoxic T cells.

The data provided herein indicate that blocking the interaction between NKG2A and HLA-E using a monoclonal antibody would increase the efficacy of existing immunotherapies such as blockade of the PD-1 pathway. The checkpoint receptor, NKG2A, is expressed on intratumoral NK cells (~60%) and CD8 T cells (~15%) and its expression is found on nearly all activated cytotoxic T cell populations, analogous to the expression of PD-1. The frequencies of NKG2A is also enhanced after PD-1 blockade therapy and after vaccination. Similarly, the ligand HLA-E (Qa-1 in mice) is generally overexpressed in cancer and is upregulated after vaccination. Thus, blocking the NKG2A-HLA-E interaction would be expected to synergize with a broad array of immunotherapy strategies for most human cancers. In addition to increasing the efficacy of cancer immunotherapy or vaccination, blocking the interaction between HLA-E and NKG2A would be expected to improve the outcome of any therapy in which antigen specific T cells are being inhibited from eliminating target cells, such as in chronic viral infections including HIV and HCV. Because of its potent inhibitory role, recombinant HLA-E that could be administered systemically could also be used to control the progression of

autoimmune diseases in which activated, self-reactive T cells are the cause of pathogenesis, such as multiple sclerosis.

Incorporation by Reference

5 All publications, patents, and patent applications mentioned herein are hereby incorporated by reference in their entirety as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control.

Also incorporated by reference in their entirety are any polynucleotide and
10 polypeptide sequences which reference an accession number correlating to an entry in a public database, such as those maintained by The Institute for Genomic Research (TIGR) on the world wide web at tigr.org and/or the National Center for Biotechnology Information (NCBI) on the World Wide Web at ncbi.nlm.nih.gov.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

20

What is claimed is:

1. A method of treating a subject afflicted with a cancer comprising administering to the subject a therapeutically effective amount of an agent that inhibits the copy number, the expression level, and/or the activity of one or more biomarkers listed in Table 1 or a
5 fragment thereof, in combination with an immunotherapy.
2. The method of claim 1, wherein the agent antagonizes NKG2A-HLA-E interaction.
3. The method of claim 1 or 2, wherein the agent decreases the copy number, the
10 expression level, and/or the activity of *HLA-E*, *Pdia3*, *Calr*, *Tap1*, *Tap2*, *Tapbp*, and/or *Erap1*.
4. The method of any one of claims 1-3, wherein the agent selectively decreases the catalytic activity and/or the substrate binding activity of HLA-E.
- 15 5. The method of any one of claims 1-4, wherein the agent is a small molecule inhibitor, CRISPR single-guide RNA (sgRNA), RNA interfering agent, antisense oligonucleotide, peptide or peptidomimetic inhibitor, aptamer, or intrabody.
- 20 6. The method of claim 5, wherein the RNA interfering agent is a small interfering RNA (siRNA), CRISPR RNA (crRNA), a small hairpin RNA (shRNA), a microRNA (miRNA), or a piwi-interacting RNA (piRNA).
7. The method of claim 6, wherein the RNA interfering agent is a CRISPR single-
25 guide RNA (sgRNA).
8. The method of claim 7, wherein the sgRNA comprises a nucleic acid sequence selected from the group consisting of nucleic acid sequence listed in Table 2.
- 30 9. The method of claim 5, wherein the agent comprises an intrabody, or an antigen binding fragment thereof, which specifically binds to HLA-E and/or a substrate of HLA-E.
10. The method of claim 9, wherein the intrabody, or antigen binding fragment thereof, is murine, chimeric, humanized, composite, or human.

35

11. The method of claim 9 or 10, wherein the intrabody, or antigen binding fragment thereof, is detectably labeled, comprises an effector domain, comprises an Fc domain, and/or is selected from the group consisting of Fv, Fav, F(ab')₂, Fab', dsFv, scFv, sc(Fv)₂, and diabodies fragments.

5

12. The method of any one of claims 9-11, wherein the intrabody, or antigen binding fragment thereof, is conjugated to a cytotoxic agent.

10

13. The method of claim 12, wherein the cytotoxic agent is selected from the group consisting of a chemotherapeutic agent, a biologic agent, a toxin, and a radioactive isotope.

14. The method of any one of claims 1-13, wherein the agent increases the sensitivity of the cancer cells to an immunotherapy.

15

15. The method of any one of claims 1-14, wherein the immunotherapy and/or a cancer therapy is administered before, after, or concurrently with the agent.

16. The method of claim 14 or 15, wherein the immunotherapy comprises an anti-cancer vaccine and/or virus.

20

17. The method of any one of claims 14-16, wherein the immunotherapy is cell-based.

18. The method of any one of claims 14-17, wherein the immunotherapy inhibits an immune checkpoint.

25

19. The method of claim 18, wherein the immune checkpoint is selected from the group consisting of CTLA-4, PD-1, VISTA, B7-H2, B7-H3, PD-L1, B7-H4, B7-H6, ICOS, HVEM, PD-L2, CD160, gp49B, PIR-B, KIR family receptors, TIM-1, TIM-3, TIM-4, LAG-3, GITR, 4-1BB, OX-40, BTLA, SIRP, CD47, CD48, 2B4 (CD244), B7.1, B7.2, ILT-2, ILT-4, TIGIT, HHLA2, butyrophilins, IDO, CD39, CD73 and A2aR.

30

20. The method of claim 15, wherein the cancer therapy is selected from the group consisting of radiation, a radiosensitizer, an immunogenic chemotherapy that induce interferon production by the cancer cells or at the site of a tumor, interferon, an interferon-inducing agent, a topical inflammatory agent, and a topical TLR agonist.

35

21. The method of any one of claims 1-20, wherein the biomarker comprises a nucleic acid sequence having at least 95% identity to a nucleic acid sequence listed in Table 1 and/or encodes an amino acid sequence having at least 95% identity to an amino acid sequence listed in Table 1.

5

22. The method of any one of claims 1-21, wherein the HLA-E is a human HLA-E, a mouse H2-T23, a chimeric HLA-E, or a fusion HLA-E.

10

23. The method of any one of claims 1-22, wherein the HLA-E comprises an amino acid sequence listed in Table 1, optionally wherein the amino acid sequence is selected from the group consisting of SEQ ID Nos: 2, 4, 6, 8, 10, 12, and 14.

15

24. The method of any one of claims 1-23, wherein the HLA-E is encoded by a nucleic acid sequence listed in Table 1, optionally wherein the nucleic acid sequence is selected from the group consisting of SEQ ID Nos: 1, 3, 5, 7, 9, 11, and 13.

20

25. The method of any one of claims 1-24, wherein the agent reduces the number of proliferating cells in the cancer and/or reduces the volume or size of a tumor comprising the cancer cells.

26. The method of any one of claims 1-25, wherein the agent increases the sensitivity of the cancer to the immunotherapy, optionally wherein the immunotherapy is T-cell-mediated.

25

27. The method of any one of claims 1-26, wherein the agent increases the amount of CD8⁺ T cells in a tumor comprising the cancer cells.

28. The method of any one of claims 1-27, wherein the cancer is selected from the group consisting of melanoma and colorectal cancer.

30

29. The method of claim 28, wherein the cancer is melanoma.

30. The method of any one of claims 1-29, wherein the subject is an animal model of the cancer.

35

31. The method of claim 30, wherein the animal model is a mouse model.

32. The method of any one of claims 1-29, wherein the subject is a mammal.
33. The method of claim 32, wherein the mammal is a mouse or a human.
- 5 34. The method of claim 33, wherein the mammal is a human.
35. The method of any one of claims 1-34, wherein the agent is administered in a pharmaceutically acceptable formulation.
- 10 36. A method of killing cancer cells comprising contacting the cancer cells with an agent that inhibits the copy number, the expression level, and/or the activity of one or more biomarkers listed in Table 1 or a fragment thereof, in combination with an immunotherapy.
37. The method of claim 36, wherein the agent decreases the copy number, the
15 expression level, and/or the activity of *HLA-E*, *Pdia3*, *Calr*, *Tap1*, *Tap2*, *Tapbp*, and/or *Erap1*.
38. The method of claim 36 or 37, wherein the agent selectively decreases the catalytic activity and/or the substrate binding activity of HLA-E.
- 20 39. The method of any one of claims 36-38, wherein the agent is a small molecule inhibitor, CRISPR single-guide RNA (sgRNA), RNA interfering agent, antisense oligonucleotide, peptide or peptidomimetic inhibitor, aptamer, or intrabody.
- 25 40. The method of claim 39, wherein the RNA interfering agent is a small interfering RNA (siRNA), CRISPR RNA (crRNA), a small hairpin RNA (shRNA), a microRNA (miRNA), or a piwi-interacting RNA (piRNA).
41. The method of claim 40, wherein the RNA interfering agent is a CRISPR single-
30 guide RNA (sgRNA).
42. The method of claim 39, wherein the sgRNA comprises a nucleic acid sequence selected from the group consisting of nucleic acid sequence listed in Table 2.
- 35 43. The method of claim 39, wherein the agent comprises an intrabody, or an antigen binding fragment thereof, which specifically binds to HLA-E and/or a substrate of HLA-E.

44. The method of claim 43, wherein the intrabody, or antigen binding fragment thereof,
is murine, chimeric, humanized, composite, or human.

5 45. The method of claim 43 or 44, wherein the intrabody, or antigen binding fragment thereof, is detectably labeled, comprises an effector domain, comprises an Fc domain, and/or is selected from the group consisting of Fv, Fav, F(ab')₂, Fab', dsFv, scFv, sc(Fv)₂, and diabodies fragments.

10 46. The method of any one of claims 43-45, wherein the intrabody, or antigen binding fragment thereof, is conjugated to a cytotoxic agent.

47. The method of claim 46, wherein the cytotoxic agent is selected from the group consisting of a chemotherapeutic agent, a biologic agent, a toxin, and a radioactive isotope.

15 48. The method of any one of claims 36-47, wherein the agent increases the sensitivity of the cancer cells to an immunotherapy.

49. The method of any one of claims 36-48, , wherein the cancer cells are contacted
20 with an immunotherapy and/or a cancer therapy before, after, or concurrently with the agent.

50. The method of claim 48 or 49, wherein the immunotherapy comprises an anti-cancer vaccine and/or virus.

25 51. The method of any one of claims 48-50, wherein the immunotherapy is cell-based.

52. The method of any one of claims 48-51, wherein the immunotherapy inhibits an immune checkpoint.

30 53. The method of claim 52, wherein the immune checkpoint is selected from the group consisting of CTLA-4, PD-1, VISTA, B7-H2, B7-H3, PD-L1, B7-H4, B7-H6, ICOS, HVEM, PD-L2, CD160, gp49B, PIR-B, KIR family receptors, TIM-1, TIM-3, TIM-4, LAG-3, GITR, 4-1BB, OX-40, BTLA, SIRP, CD47, CD48, 2B4 (CD244), B7.1, B7.2, ILT-
35 2, ILT-4, TIGIT, HHLA2, butyrophilins, IDO, CD39, CD73 and A2aR.

54. The method of claim 53, wherein the cancer therapy is selected from the group consisting of radiation, a radiosensitizer, an immunogenic chemotherapy that induce interferon production by the cancer cells or at the site of a tumor, interferon, an interferon-inducing agent, a topical inflammatory agent, and a topical TLR agonist.

5

55. The method of any one of claims 36-54, wherein the biomarker comprises a nucleic acid sequence having at least 95% identity to a nucleic acid sequence listed in Table 1 and/or encodes an amino acid sequence having at least 95% identity to an amino acid sequence listed in Table 1.

10

56. The method of any one of claims 36-55, wherein the HLA-E is a human HLA-E, a mouse H2-T23, a chimeric HLA-E, or a fusion HLA-E.

57. The method of any one of claims 36-56, wherein the HLA-E comprises an amino acid sequence listed in Table 1, optionally wherein the amino acid sequence is selected from the group consisting of SEQ ID Nos: 2, 4, 6, 8, 10, 12, and 14.

15

58. The method of any one of claims 36-57, wherein the HLA-E is encoded by a nucleic acid sequence listed in Table 1, optionally wherein the nucleic acid sequence is selected from the group consisting of SEQ ID Nos: 1, 3, 5, 7, 9, 11, and 13.

20

59. The method of any one of claims 36-58, wherein the agent reduces the number of proliferating cells in the cancer and/or reduces the volume or size of a tumor comprising the cancer cells.

25

60. The method of any one of claims 36-59, wherein the agent increases the sensitivity of the cancer to the immunotherapy, optionally wherein the immunotherapy is T-cell-mediated.

61. The method of any one of claims 36-60, wherein the agent increases the amount of CD8+ T cells in a tumor comprising the cancer cells.

30

62. The method of any one of claims 36-61, wherein the cancer is selected from the group consisting of melanoma and colorectal cancer.

35

63. The method of claim 62, wherein the cancer is melanoma.

64. The method of any one of claims 36-63, wherein the agent is administered in a pharmaceutically acceptable formulation.

5 65. A method of determining whether a subject afflicted with a cancer or at risk for developing a cancer would benefit from inhibiting the copy number, amount, and/or activity of at least one biomarker listed in Table 1, the method comprising:

a) obtaining a biological sample from the subject;

10 b) determining the copy number, amount, and/or activity of at least one biomarker listed in Table 1;

c) determining the copy number, amount, and/or activity of the at least one biomarker in a control; and

d) comparing the copy number, amount, and/or activity of the at least one biomarker detected in steps b) and c);

15 wherein the presence of, or a significant increase in, the copy number, amount, and/or activity of, the at least one biomarker listed in Table 1 in the subject sample relative to the control copy number, amount, and/or activity of the at least one biomarker indicates that the subject afflicted with the cancer or at risk for developing the cancer would benefit from inhibiting the copy number, amount, and/or activity of the at least one biomarker
20 listed in Table 1.

66. The method of claim 65, further comprising recommending, prescribing, or administering an agent that inhibits the at least one biomarker listed in Table 1 if the cancer is determined to benefit from the agent, optionally further administering at least one
25 additional cancer therapy.

67. The method of claim 65, further comprising recommending, prescribing, or administering cancer therapy other than an agent that inhibits the at least one biomarker listed in Table 1 if the cancer is determined to not benefit from the agent.

30 68. The method of claim 66 or 67, wherein the cancer therapy is selected from the group consisting of immunotherapy, targeted therapy, chemotherapy, radiation therapy, hormonal therapy, an anti-cancer vaccine, an anti-cancer virus, a checkpoint inhibitor, a radiosensitizer, an immunogenic chemotherapy that induce interferon production by the

cancer cells or at the site of a tumor, interferon, an interferon-inducing agent, a topical inflammatory agent, and a topical TLR agonist.

69. The method of any one of claims 65-68, wherein the control sample is determined
5 from a cancerous or non-cancerous sample from either the patient or a member of the same species to which the patient belongs.

70. The method of any one of claims 65-69, wherein the control sample comprises cells.

10 71. A method for predicting the clinical outcome of a subject afflicted with a cancer expressing one or more biomarkers listed in Table 1 or a fragment thereof, the method comprising:

a) determining the copy number, amount, and/or activity of at least one biomarker listed in Table 1 in a subject sample;

15 b) determining the copy number, amount, and/or activity of the at least one biomarker in a control having a good clinical outcome; and

c) comparing the copy number, amount, and/or activity of the at least one biomarker in the subject sample and in the control;

20 wherein the presence of, or a significant increase in, the copy number, amount, and/or activity of, the at least one biomarker listed in Table 1 in the subject sample as compared to the copy number, amount and/or activity in the control, is an indication that the subject has a poor clinical outcome.

72. A method for monitoring the progression of a cancer in a subject, wherein the
25 subject is administered a therapeutically effective amount of an agent that inhibits the copy number, amount, and/or activity of at least one biomarker listed in Table 1, the method comprising:

a) detecting in a subject sample at a first point in time the copy number, amount, and/or activity of at least one biomarker listed in Table 1;

30 b) repeating step a) at a subsequent point in time; and

c) comparing the amount or activity of at least one biomarker listed in Table 1 detected in steps a) and b) to monitor the progression of the cancer in the subject.

73. A method of assessing the efficacy of an agent that inhibits the copy number, amount, and/or activity of at least one biomarker listed in Table 1 for treating a cancer in a subject, comprising:

- a) detecting in a subject sample at a first point in time the copy number, amount, and/or or activity of at least one biomarker listed in Table 1;
- b) repeating step a) during at least one subsequent point in time after administration of the agent; and
- c) comparing the copy number, amount, and/or activity detected in steps a) and b), wherein the absence of, or a significant decrease in, the copy number, amount, and/or activity of, the at least one biomarker listed in Table 1, in the subsequent sample as compared to the copy number, amount, and/or activity in the sample at the first point in time, indicates that the agent treats the cancer in the subject.

74. The method of claim 72 or 73, wherein between the first point in time and the subsequent point in time, the subject has undergone treatment, completed treatment, and/or is in remission for the cancer.

75. The method of claim 74, wherein the cancer treatment is selected from the group consisting of immunotherapy, targeted therapy, chemotherapy, radiation therapy, hormonal therapy, an anti-cancer vaccine, an anti-cancer virus, a checkpoint inhibitor, a radiosensitizer, an immunogenic chemotherapy that induce interferon production by the cancer cells or at the site of a tumor, interferon, an interferon-inducing agent, a topical inflammatory agent, and a topical TLR agonist.

76. The method of any one of claims 72-75, wherein the first and/or at least one subsequent sample is selected from the group consisting of *ex vivo* and *in vivo* samples.

77. The method of any one of claims 72-76, wherein the first and/or at least one subsequent sample is a portion of a single sample or pooled samples obtained from the subject.

78. The method of any one of claims 65-77, wherein the sample comprises cells, serum, peritumoral tissue, and/or intratumoral tissue obtained from the subject.

79. The method of any one of claims 65-78, wherein the one or more biomarkers listed in Table 1 comprise *HLA-E*, *Pdia3*, *Calr*, *Tap1*, *Tap2*, *Tapbp*, and/or *Erap1*.

80. The method of any one of claims 65-79, wherein the one or more biomarkers
5 comprise a nucleic acid sequence having at least 95% identity to a nucleic acid sequence listed in Table 1 and/or encode an amino acid sequence having at least 95% identity to an amino acid sequence listed in Table 1.

81. The method of claim 79, wherein the HLA-E is a human HLA-E, a mouse H2-T23,
10 a chimeric HLA-E, or a fusion HLA-E.

82. The method of claim 81, wherein the HLA-E comprises an amino acid sequence listed in Table 1, optionally wherein the amino acid sequence is selected from the group consisting of SEQ ID Nos: 2, 4, 6, 8, 10, 12, and 14.

15 83. The method of claim 81 or 82, wherein the HLA-E is encoded by a nucleic acid sequence listed in Table 1, optionally wherein the nucleic acid sequence is selected from the group consisting of SEQ ID Nos: 1, 3, 5, 7, 9, 11, and 13.

20 84. The method of any one of claims 65-83, wherein the agent reduces the number of proliferating cells in the cancer and/or reduces the volume or size of a tumor comprising the cancer cells.

25 85. The method of any one of claims 65-84, wherein the agent increases the sensitivity of the cancer to the immunotherapy, optionally wherein the immunotherapy is T-cell-mediated.

86. The method of any one of claims 65-85, wherein the agent increases the amount of CD8+ T cells in a tumor comprising the cancer cells.

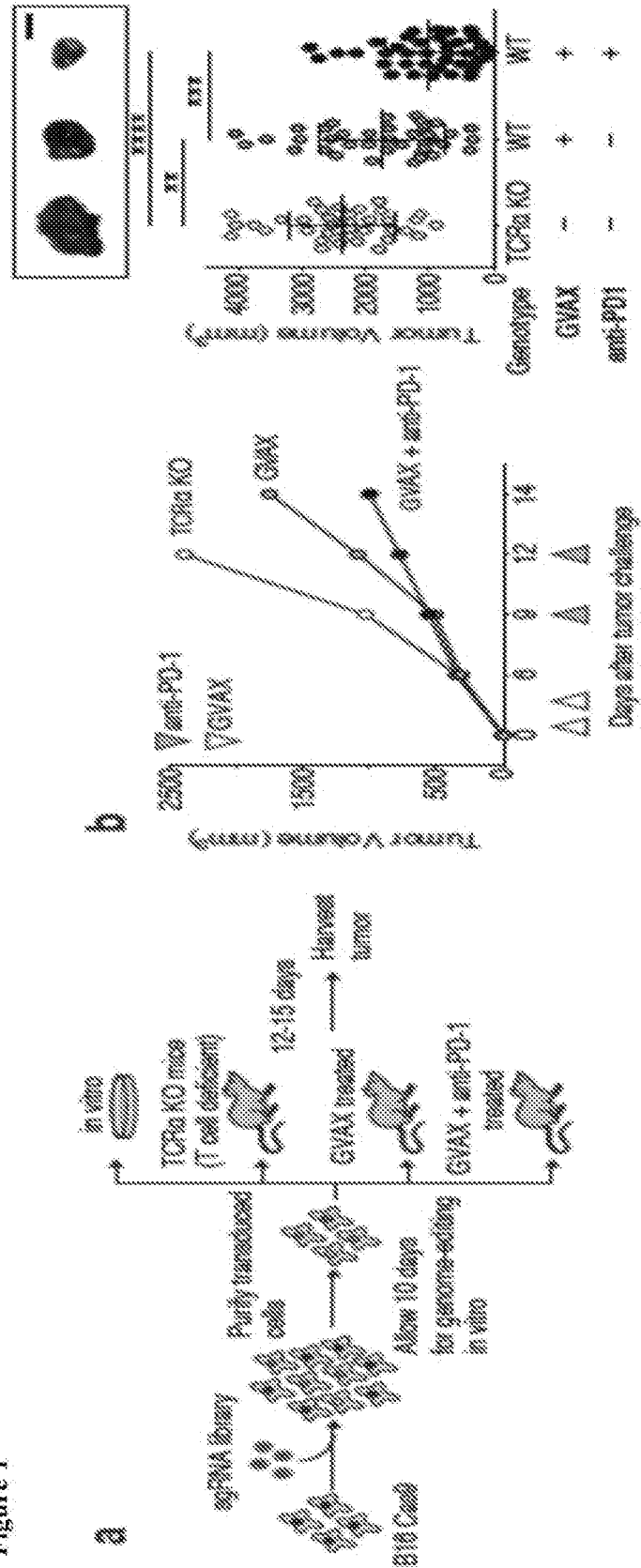
30 87. The method of any one of claims 65-86, wherein the cancer is selected from the group consisting of melanoma and colorectal cancer.

88. The method of claim 87, wherein the cancer is melanoma.

35 89. The method of any one of claims 65-88, wherein the agent is administered in a pharmaceutically acceptable formulation.

90. The method of any one of claims 65-89, wherein the subject is an animal model of the cancer.
91. The method of claim 90, wherein the animal model is a mouse model.
- 5 92. The method of any one of claims 65-91, wherein the subject is a mammal.
93. The method of claim 3292 wherein the mammal is a mouse or a human.
- 10 94. The method of claim 93, wherein the mammal is a human.

Figure 1



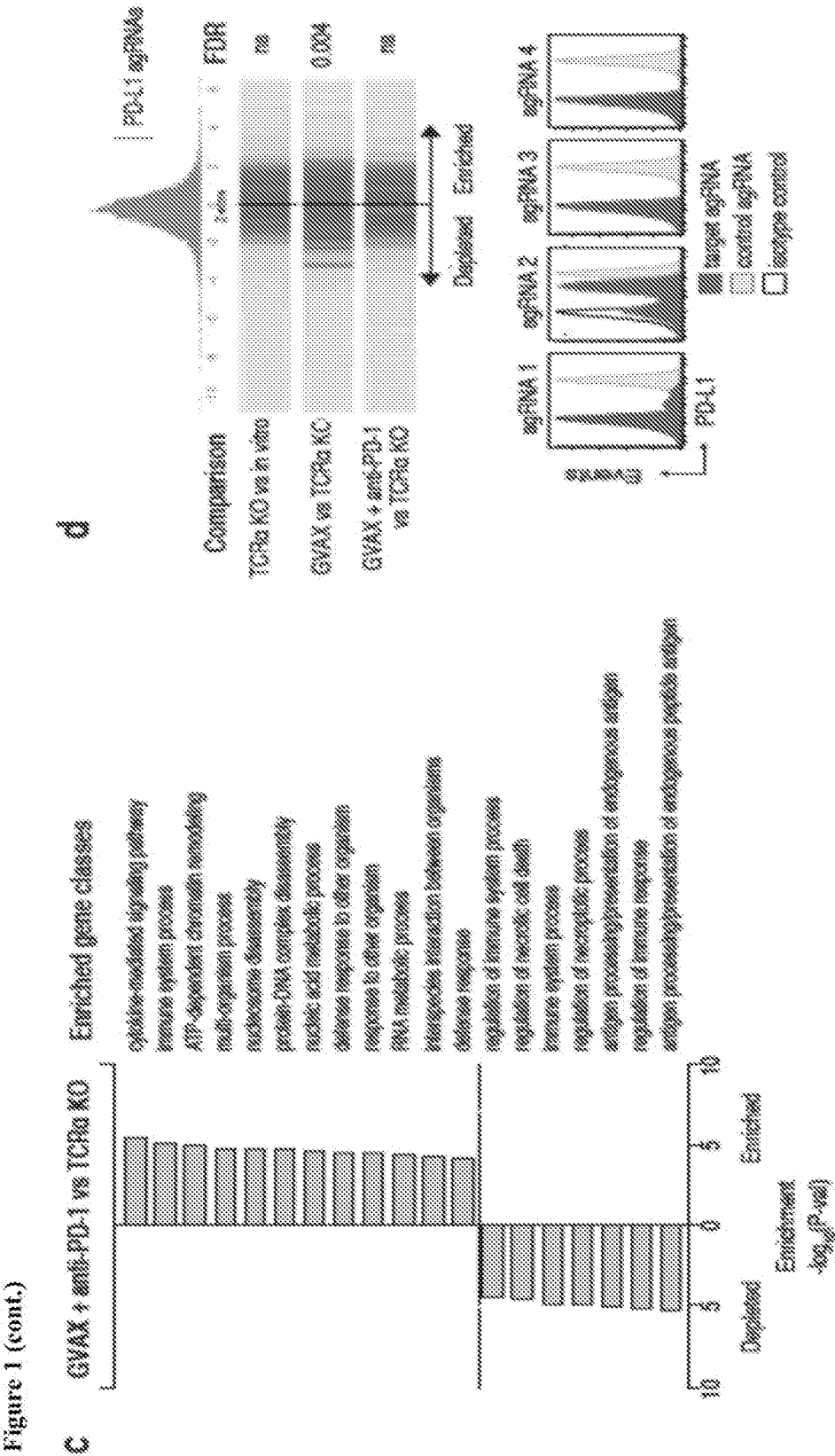
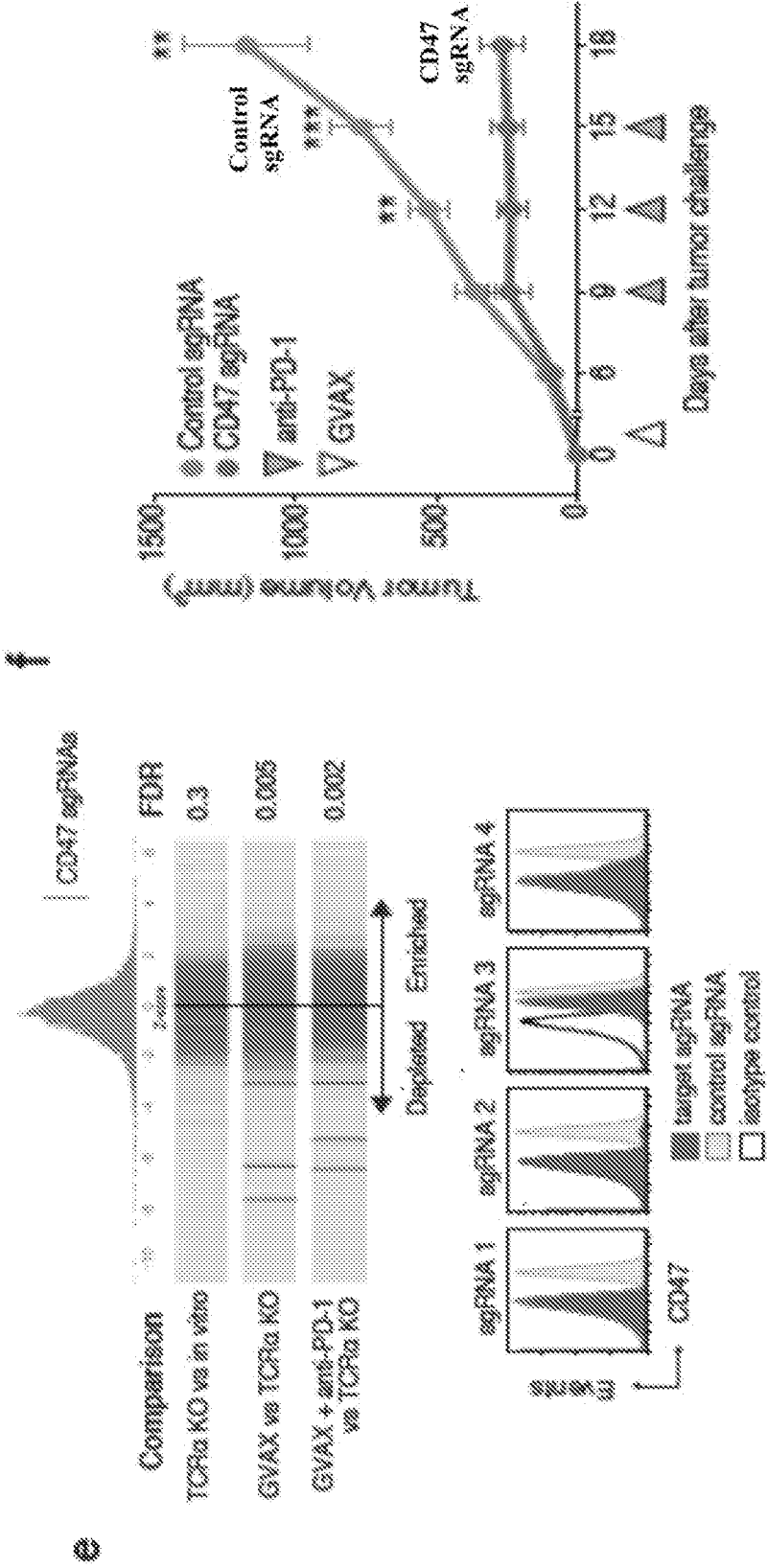


Figure 1 (cont.)



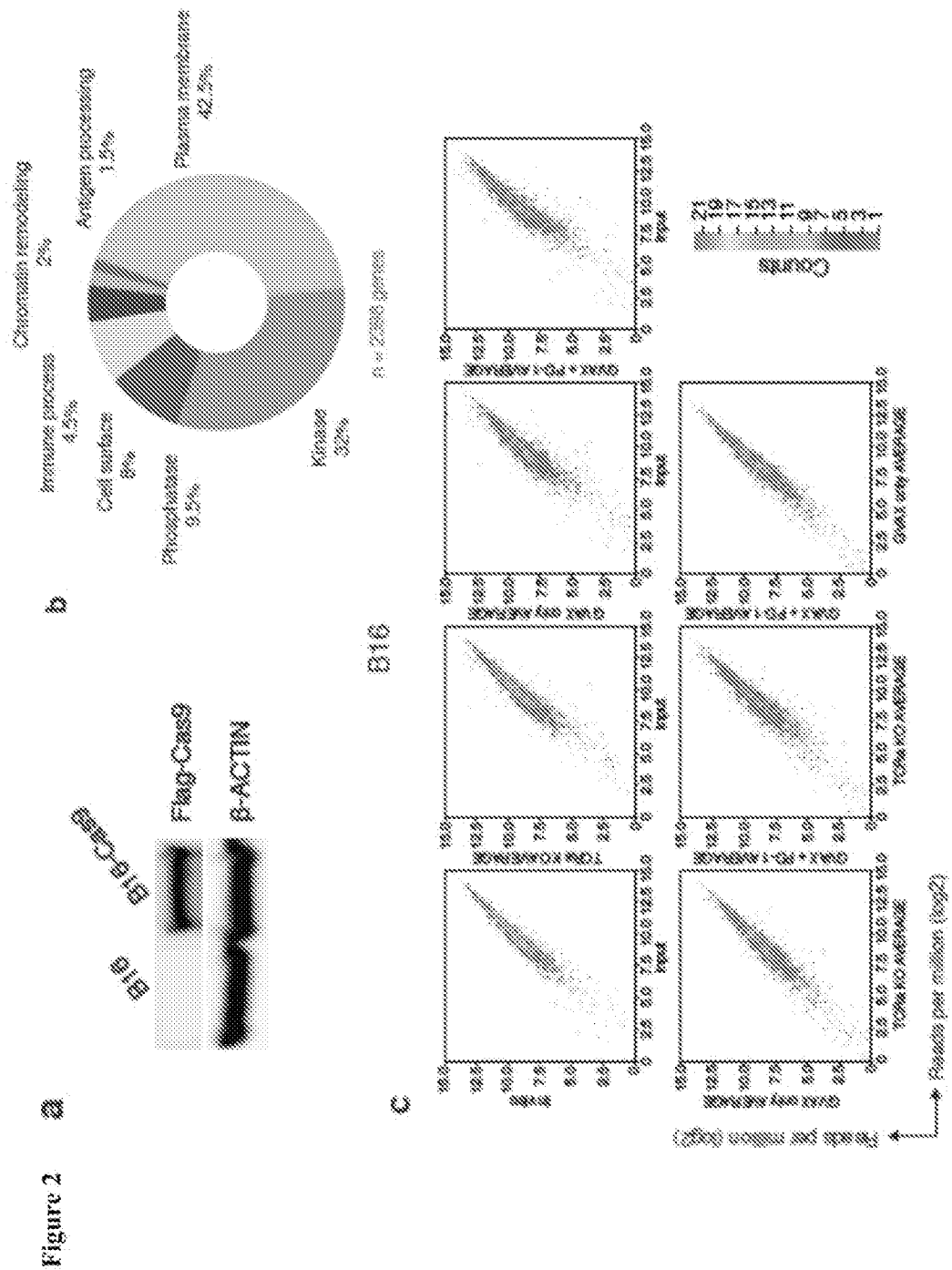


Figure 2 (cont.)

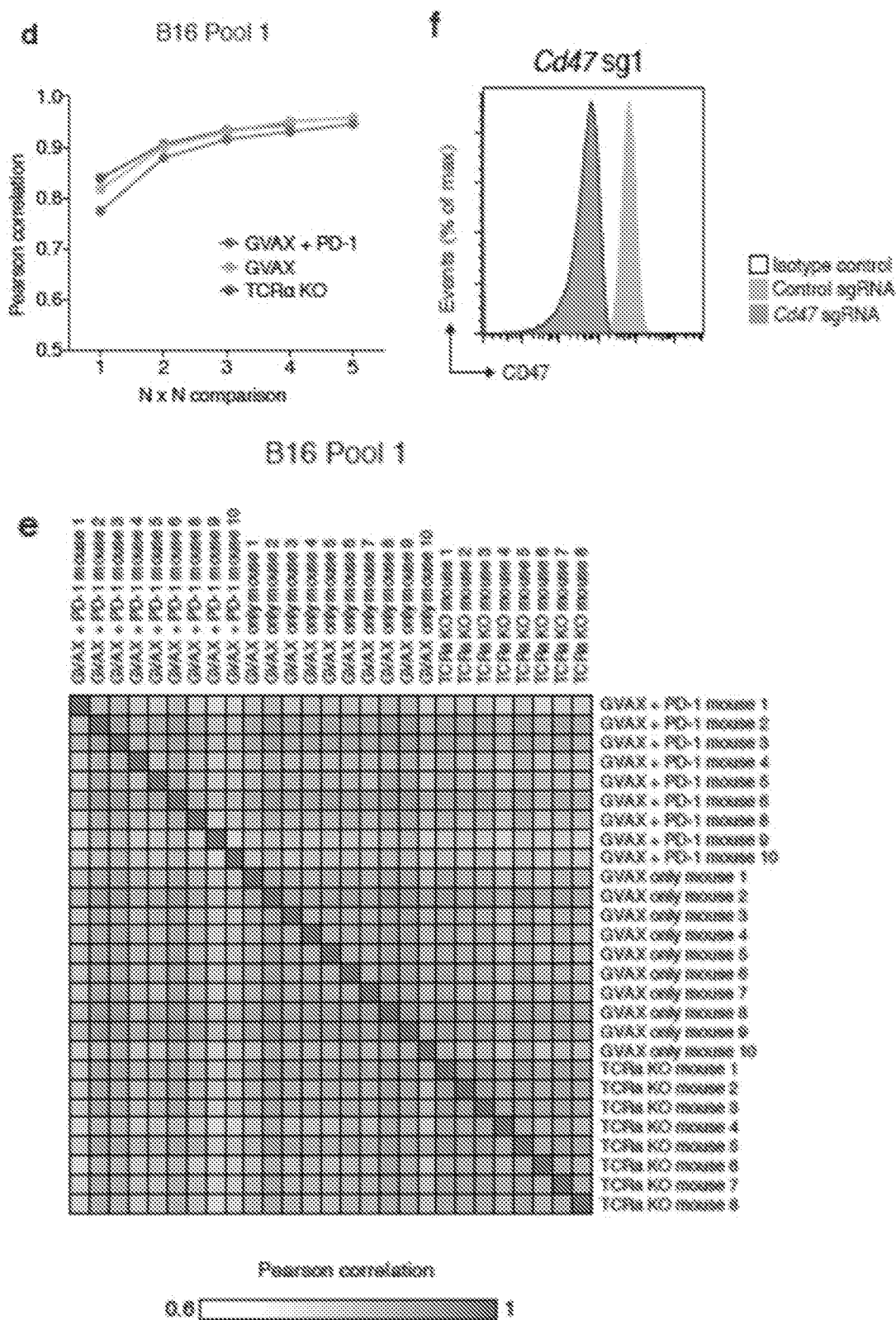


Figure 3

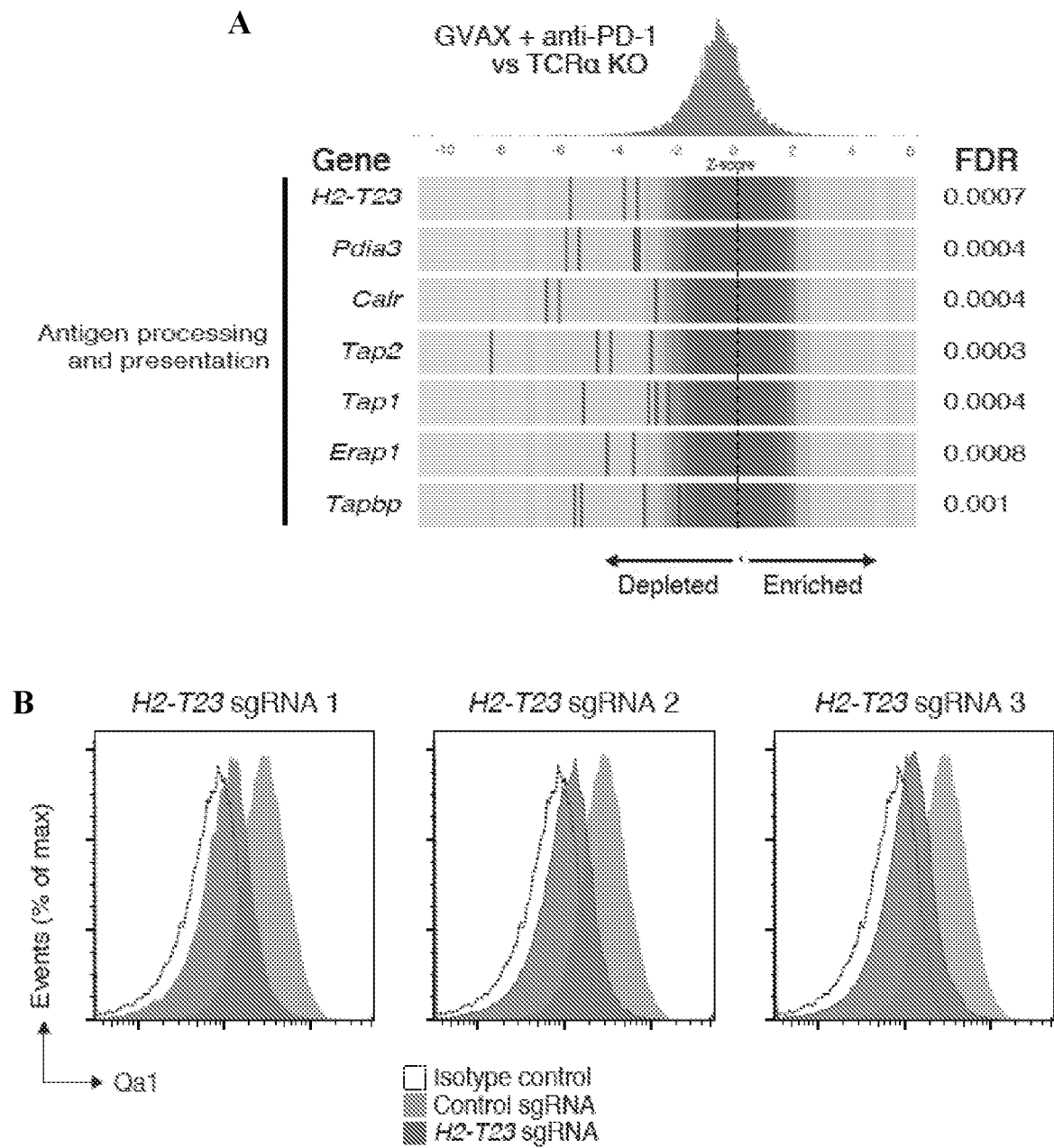


Figure 3 (cont.)

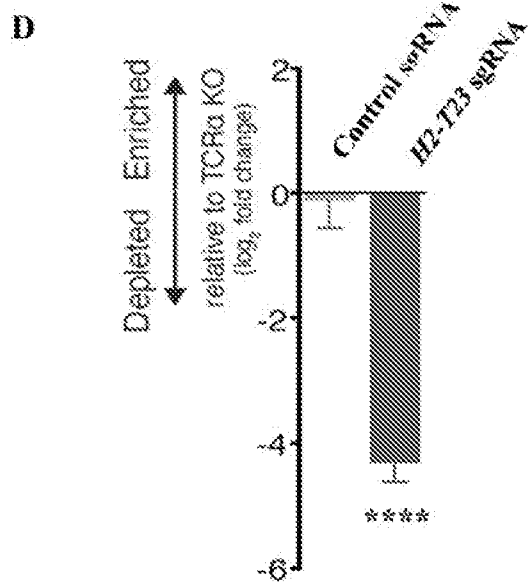
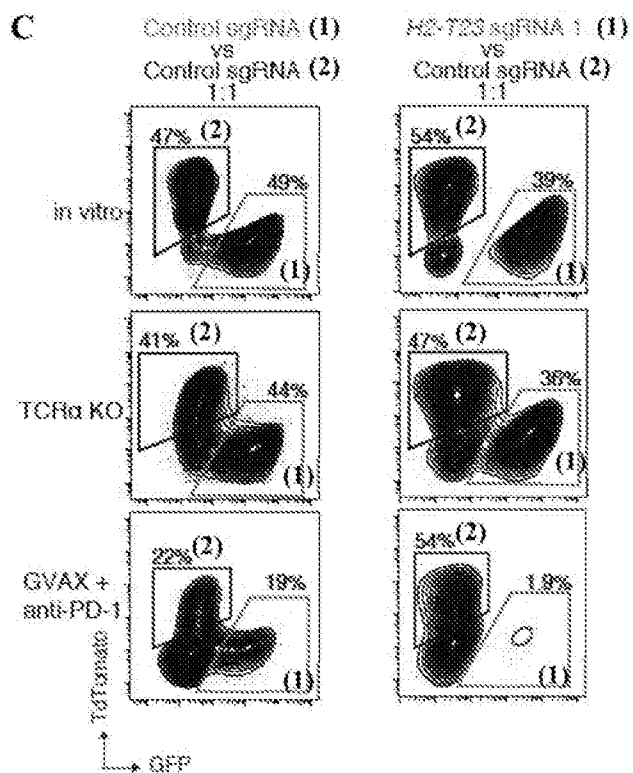
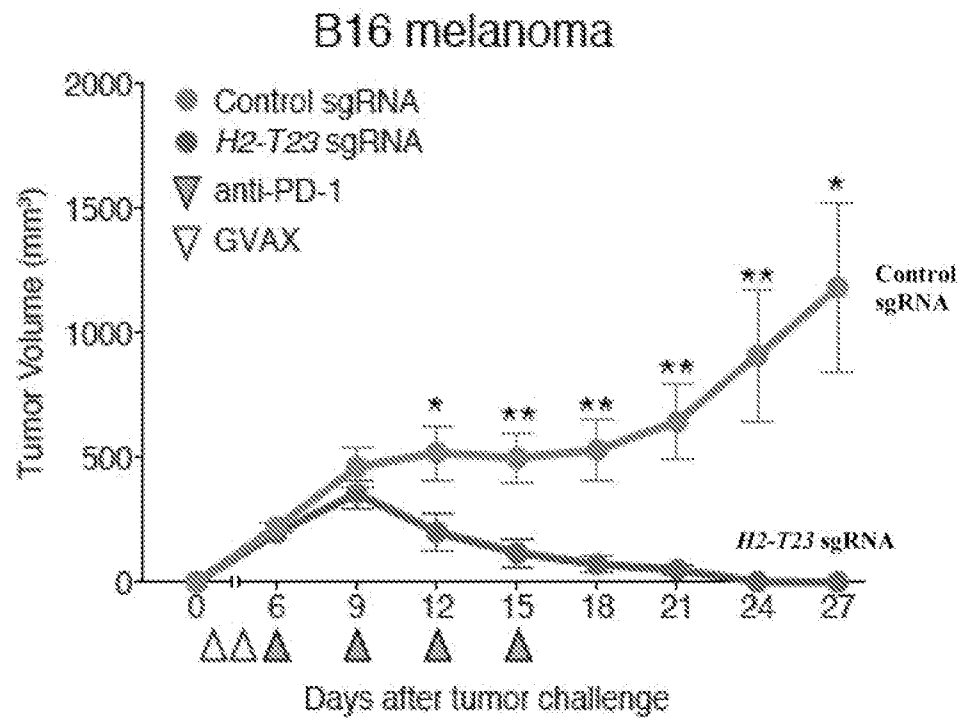


Figure 3 (cont.)

E



F

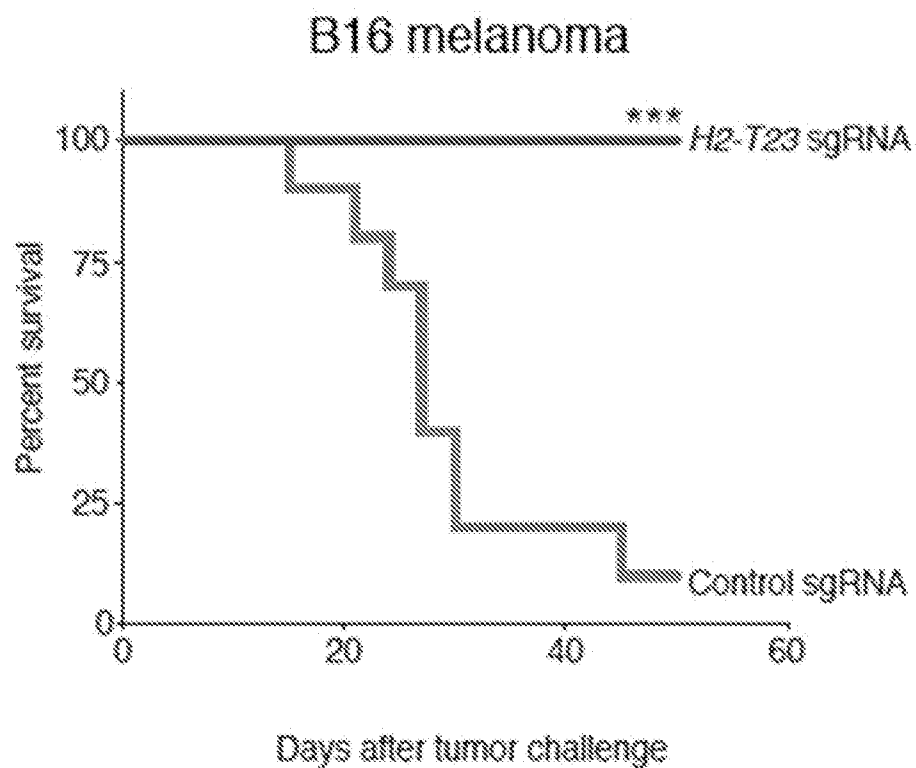


Figure 4

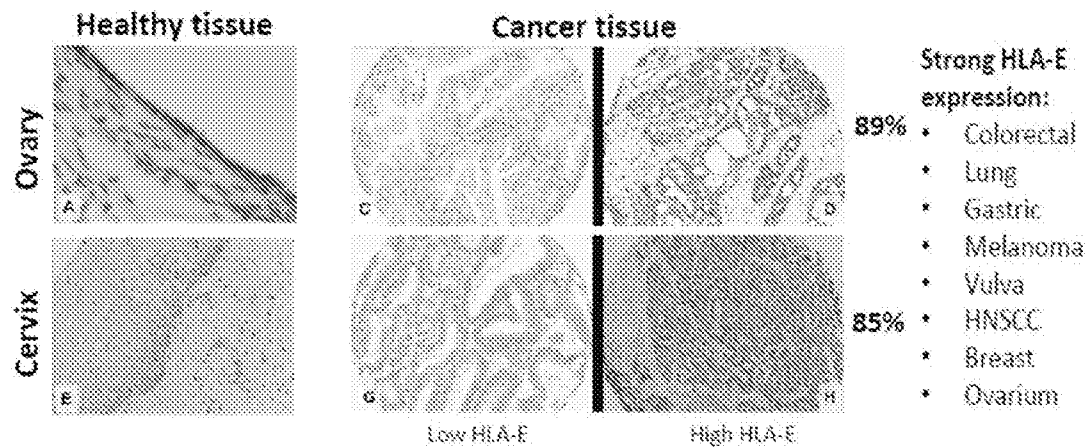


Figure 5

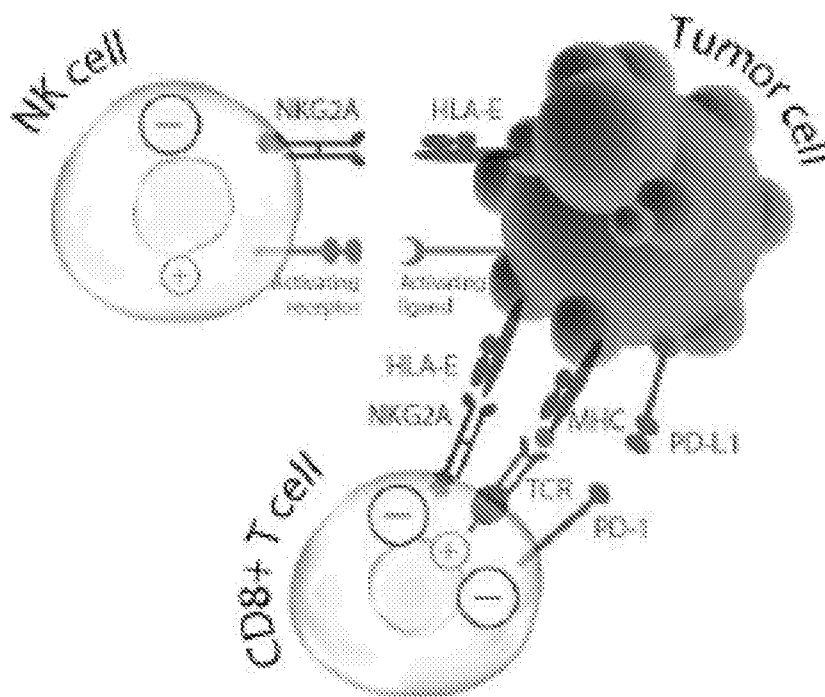
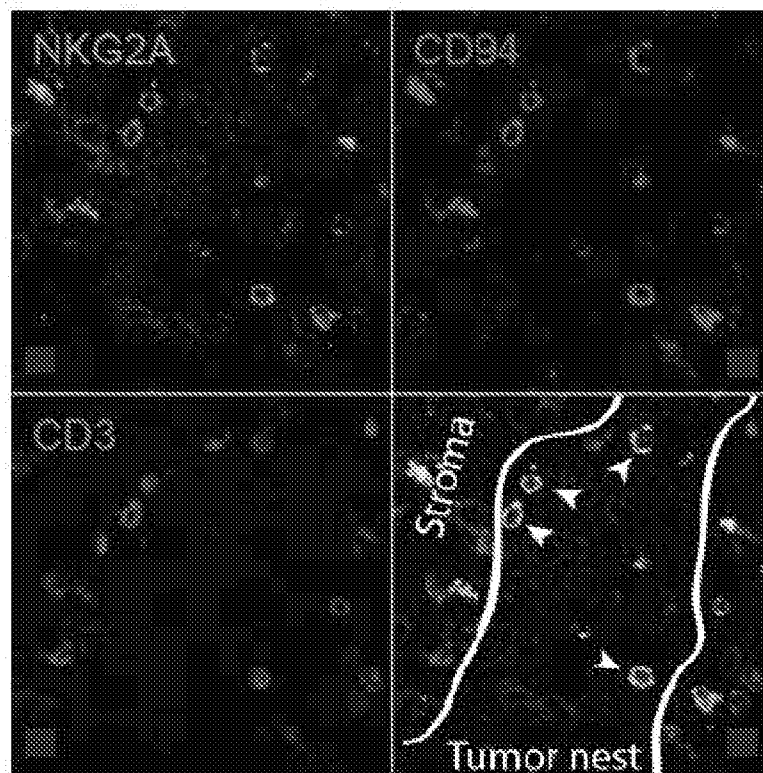
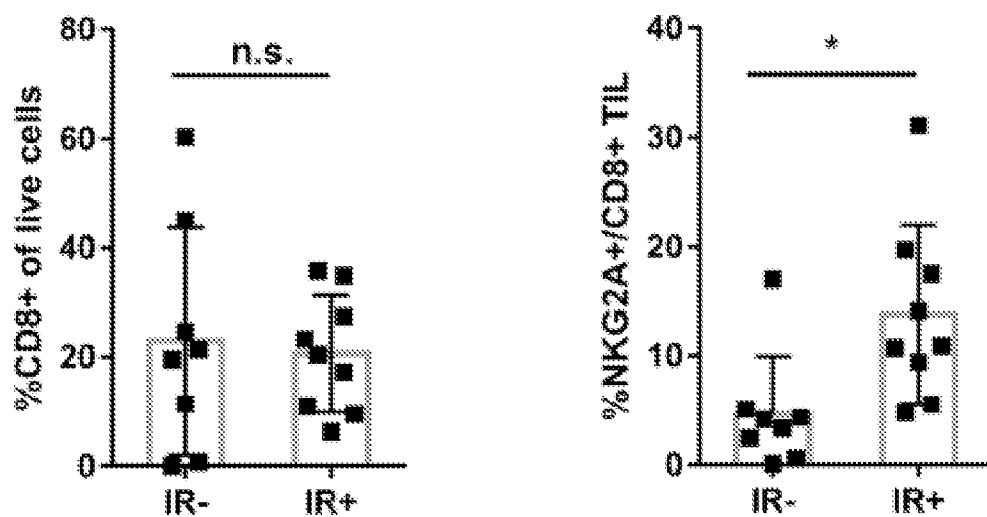


Figure 6



Gooden et al, PNAS 2011

% CD8 T cells of TIL

% NKG2A⁺ of CD8⁺HNSCC: IR= *ex vivo* immune response

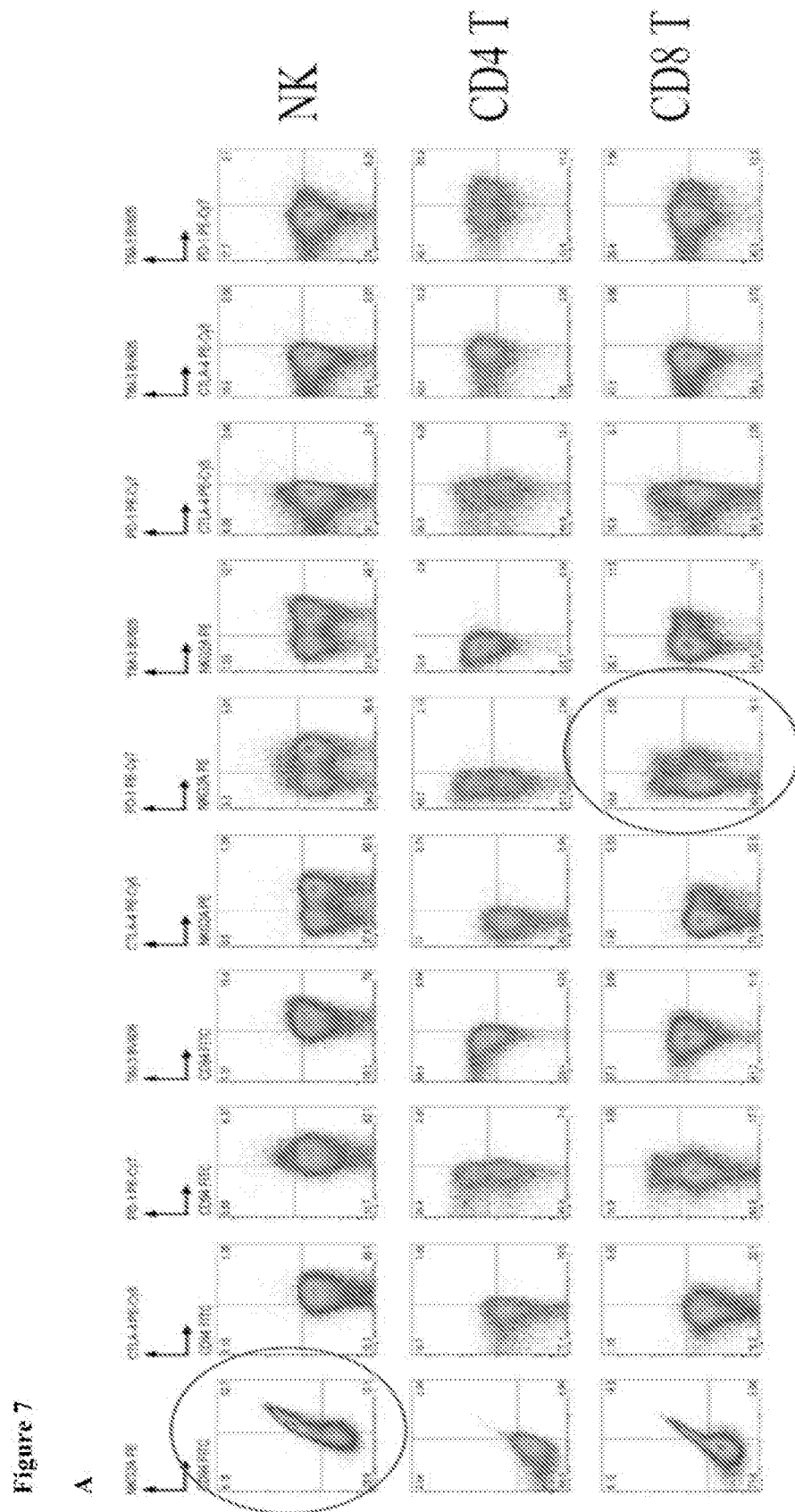
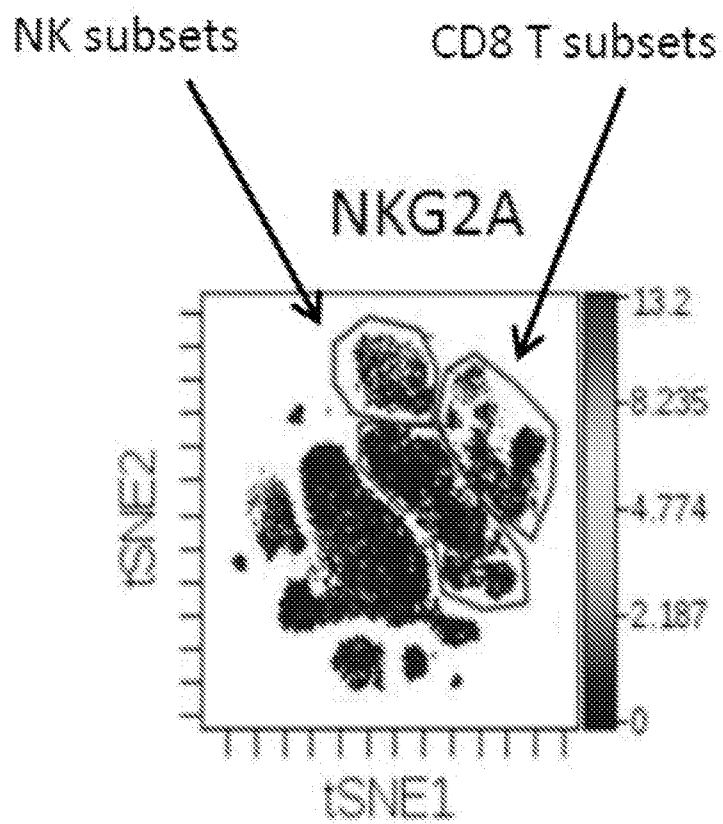


Figure 7 (cont.)

B



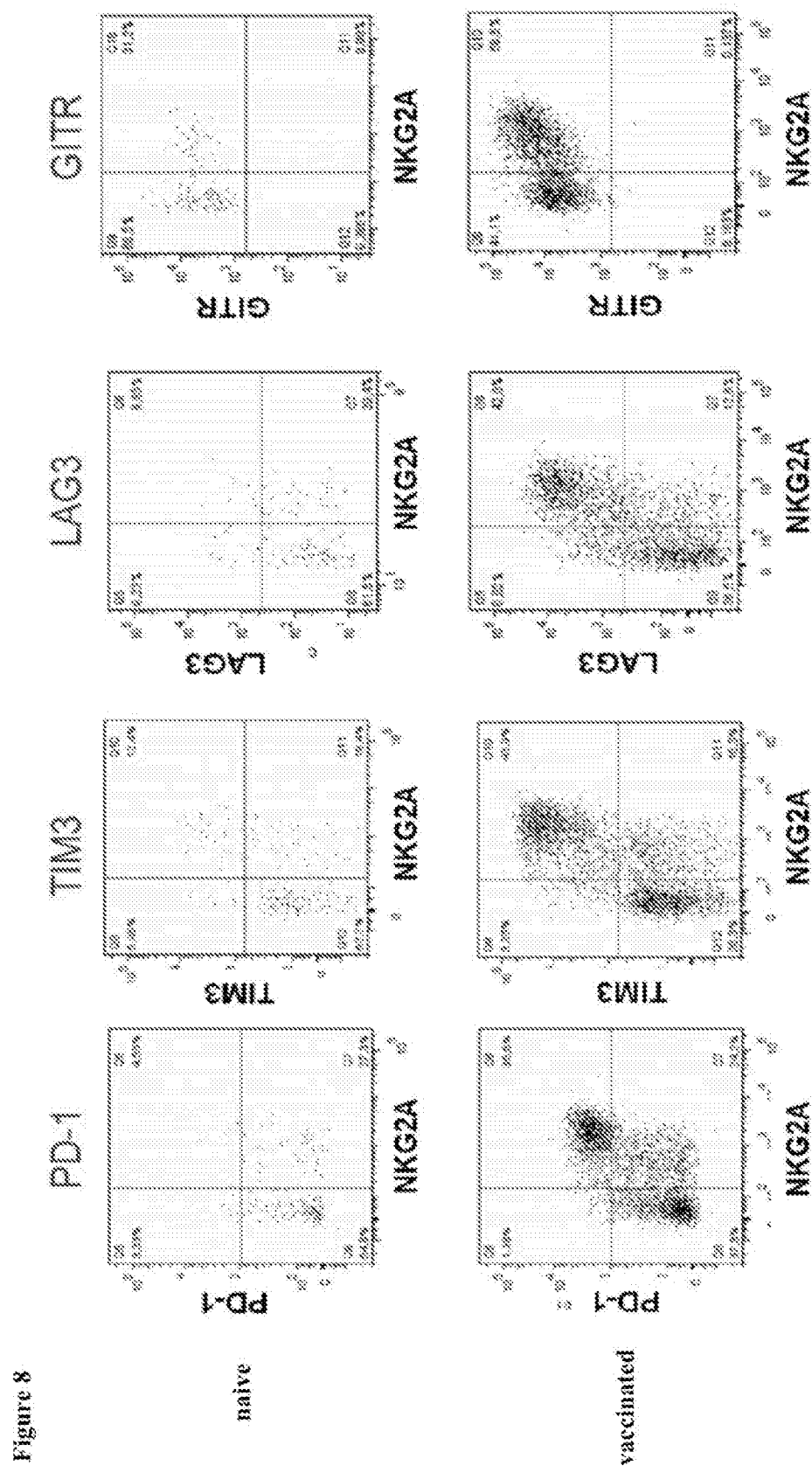
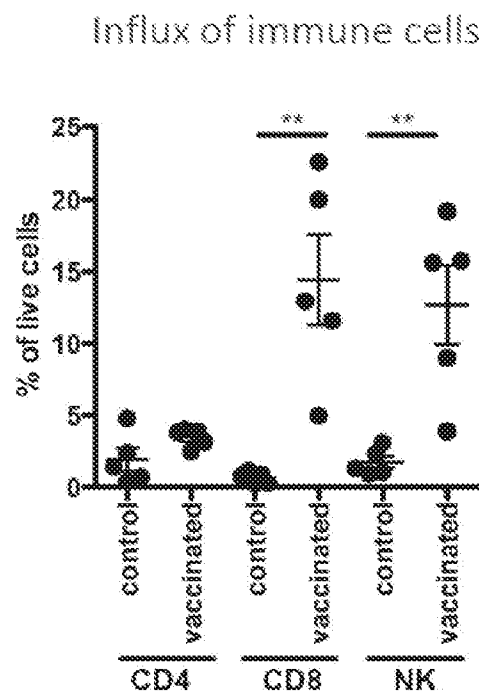
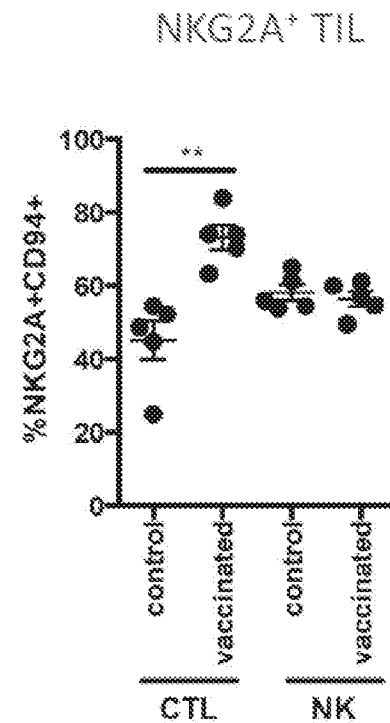
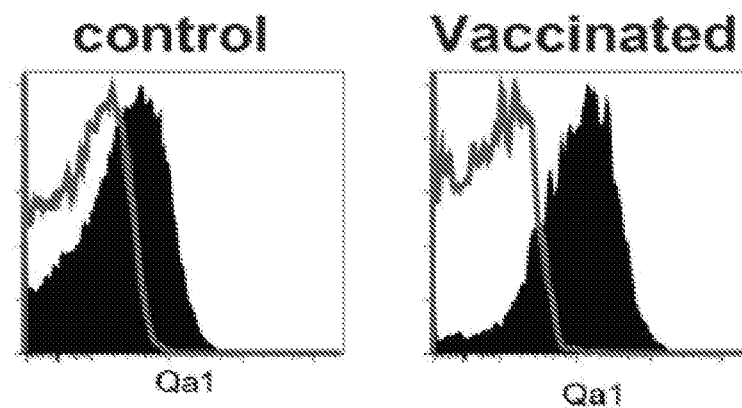


Figure 9**A****B****C**

Enhanced Qa-1 expression
on tumor

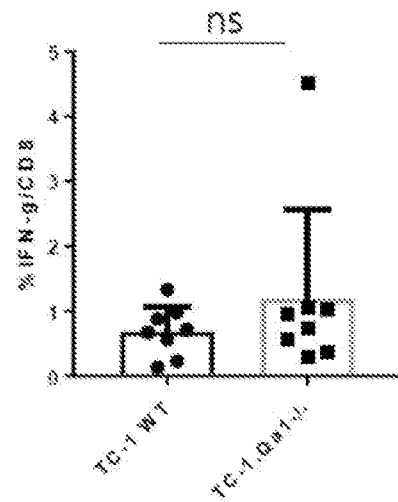


Also found in B16 model

Figure 10

A

Similar T cell induction



B

Stronger vaccination efficacy

WT TC-1 tumor

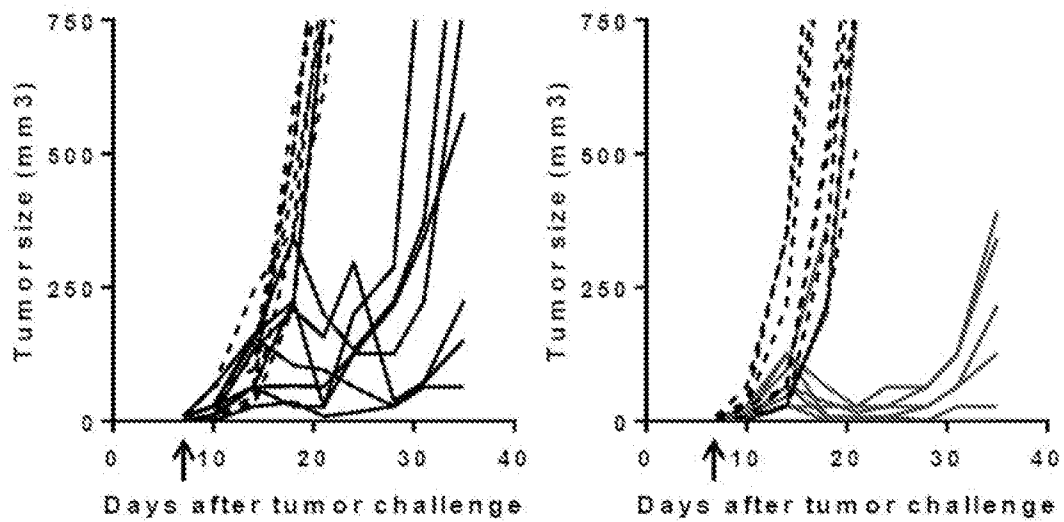
Qa-1^{ko} TC-1 tumor

Figure 11

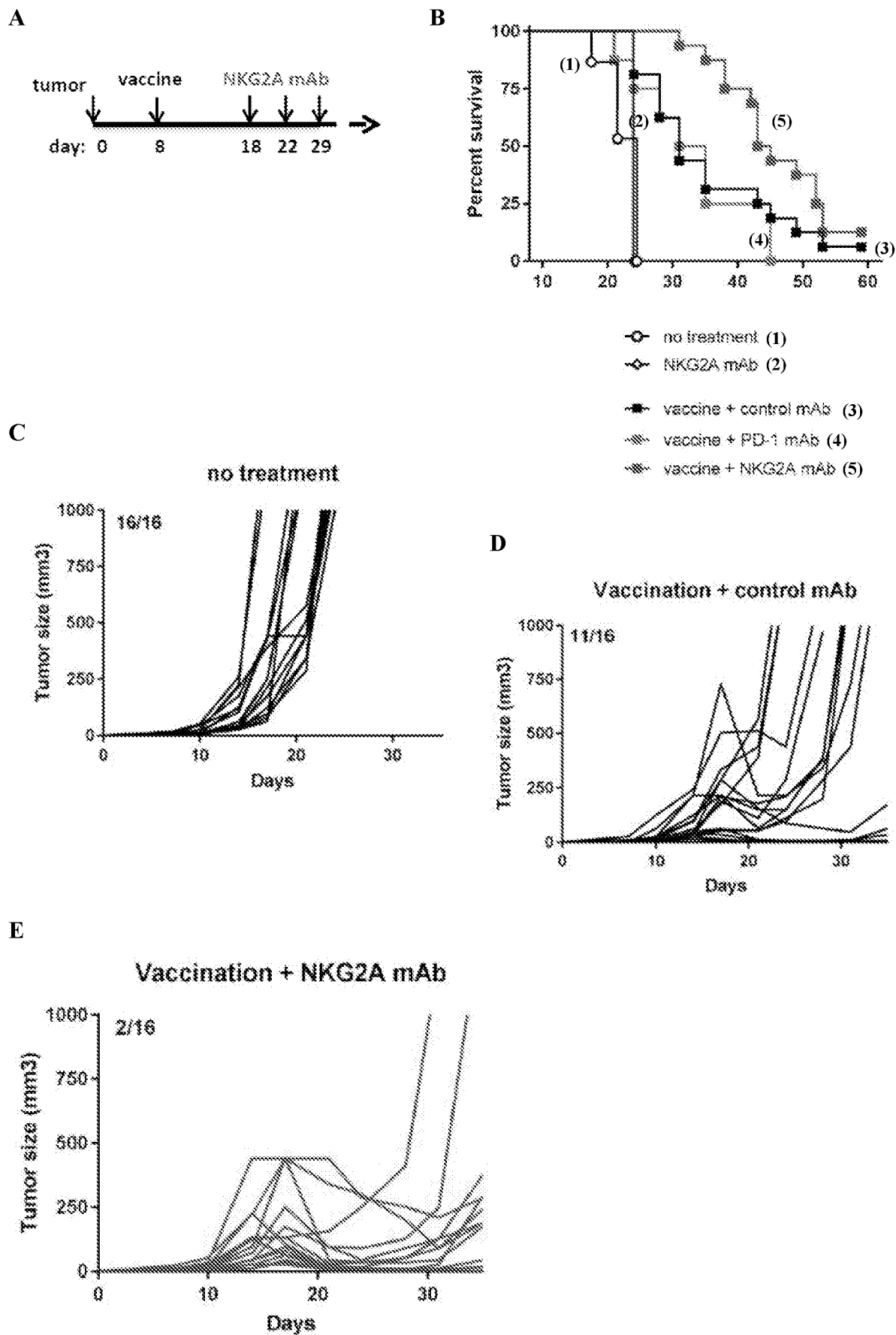
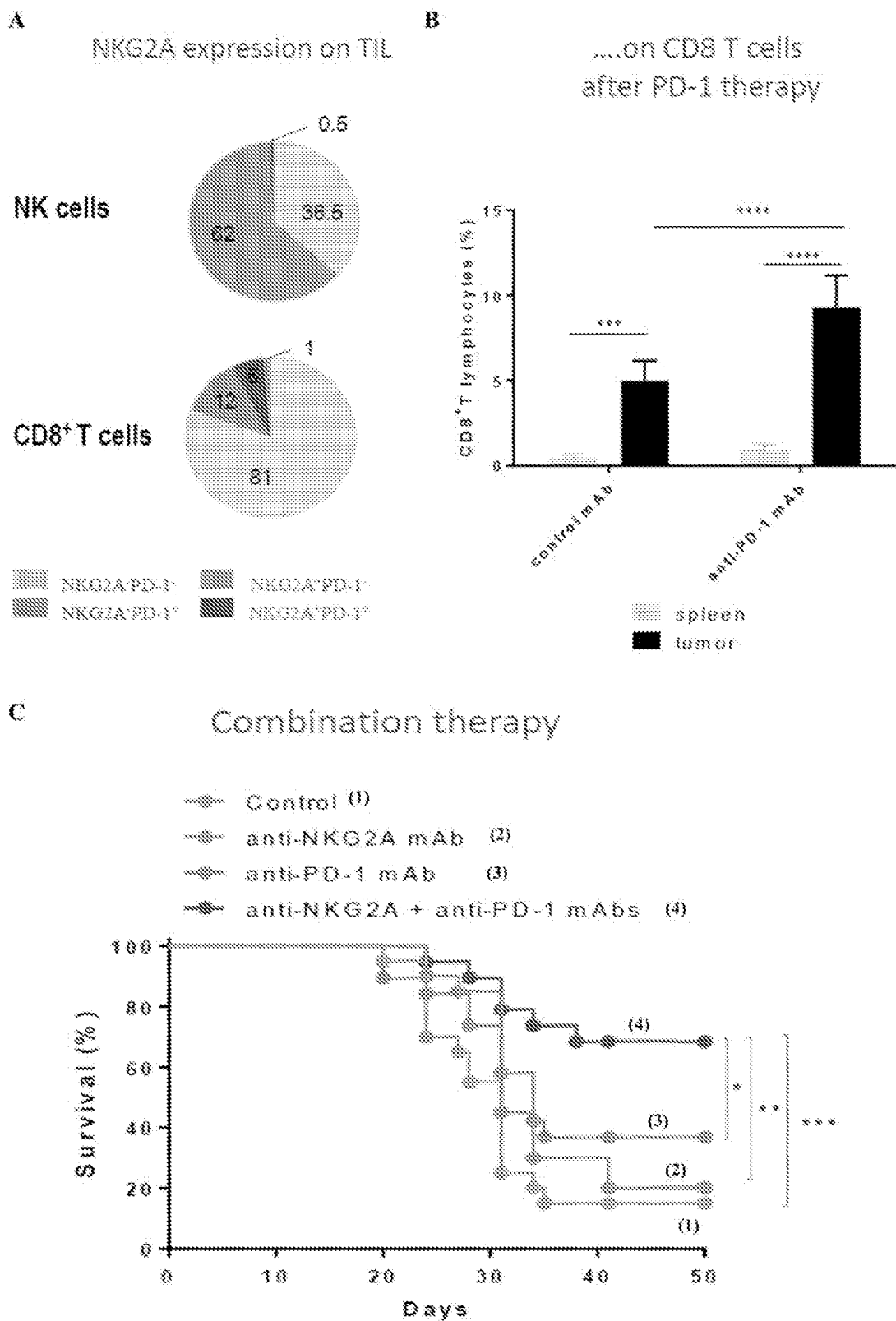


Figure 12



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2018/042243

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - A61P 35/00; C12N 15/11; C12Q 1/68; G01N 33/50; G01N 33/574 (2018.01)

CPC - A61P 35/00; C12N 15/11; C12Q 1/6881; C12Q 1/6883; C12Q 2600/106; C12Q 2600/158; G01N 33/574; G01N 2800/52 (2018.08)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

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

USPC - 435/6; 435/375; 435/7.1; 514/19.3; 536/23.1 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2014/0010825 A1 (RAVINDRANATH et al) 09 January 2014 (09.01.2014) entire document	1-3, 36-38, 65-68, 71-75
Y	US 2013/0165504 A1 (MODERNA THERAPEUTICS et al) 27 June 2013 (27.06.2013) entire document	1-3, 36-38, 65-68, 71-75
Y	US 2016/0122829 A1 (DANA-FARBER CANCER INSTITUTE, INC.) 05 May 2016 (05.05.2016) entire document	65-68, 71-75
A	WO 2012/094252 A1 (THE TERASAKI FAMILY FOUNDATION et al) 12 July 2012 (12.07.2012) entire document	1-3, 36-38, 65-68, 71-75
A	ZHEN et al. "HLA-E inhibitor enhances the killing of neuroblastoma stem cells by co-cultured dendritic cells and cytokine-induced killer cells loaded with membrane-based microparticles," Am J Cancer Res, 01 February 2017 (01.02.2017), Vol. 7, Pgs. 334-345. entire document	1-3, 36-38, 65-68, 71-75
P, A	MANGUSO et al. "In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target," Nature, 19 July 2017 (19.07.2017), Vol. 547, Pgs. 413-418. entire document	1-3, 36-38, 65-68, 71-75

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

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

26 November 2018

Date of mailing of the international search report

11 DEC 2018

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2018/042243

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

a. ☐ forming part of the international application as filed:

☐ in the form of an Annex C/ST.25 text file.

☐ on paper or in the form of an image file.

b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

c. ☒ furnished subsequent to the international filing date for the purposes of international search only:

☒ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☒ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

SEQ ID NOs: 1-8, 10, 12, and 14 were searched.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2018/042243

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

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

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 4-35, 39-64, 69, 70, 76-94
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

This International Searching Authority found multiple inventions in this international application, as follows:
See extra sheet(s).

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-3, 36-38, 65-68, and 71-75 to the extent that they read on a biomarker of SEQ ID NO:1 and/or SEQ ID NO:2.

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

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

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

Group I+: claims 1-3, 36-38, 65-68, and 71-75 are drawn to biomarkers, and methods comprising the same.

The first invention of Group I+ is restricted to a biomarker, and methods comprising the same, wherein the biomarker is selected to be SEQ ID NO:2 (encoded by SEQ ID NO:1). It is believed that claims 1-3, 36-38, 65-68, and 71-75 read on this first named invention and thus these claims will be searched without fee to the extent that they read on SEQ ID NO:1 and/or SEQ ID NO:2.

Applicant is invited to elect additional biomarkers, each with specified SEQ ID NO, to be searched in a specific combination by paying additional fee for each set of election. An exemplary election would a biomarker, and methods comprising the same, wherein the biomarker is selected to be SEQ ID NO:4 (encoded by SEQ ID NO:3). Additional biomarkers will be searched upon the payment of additional fees. Applicants must specify the claims that read on any additional elected inventions. Applicants must further indicate, if applicable, the claims which read on the first named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined.

The inventions listed in Groups I+ do not relate to a single general inventive concept under PCT Rule 13.1, because under PCT Rule 13.2 they lack the same or corresponding special technical features for the following reasons:

The Groups I+ formulas do not share a significant structural element responsible for treating a subject afflicted with a cancer, requiring the selection of alternatives for the biomarker, where "one or more biomarkers listed in Table 1 or a fragment thereof".

Additionally, even if Groups I+ were considered to share the technical features of a method of treating a subject afflicted with a cancer comprising administering to the subject a therapeutically effective amount of an agent that inhibits the copy number, the expression level, and/or the activity of one or more biomarkers or a fragment thereof, in combination with an immunotherapy; method of killing cancer cells comprising contacting the cancer cells with an agent that inhibits the copy number, the expression level, and/or the activity of one or more biomarkers or a fragment thereof, in combination with an immunotherapy; a method of determining whether a subject afflicted with a cancer or at risk for developing a cancer would benefit from inhibiting the copy number, amount, and/or activity of at least one biomarker, the method comprising: a) obtaining a biological sample from the subject; b) determining the copy number, amount, and/or activity of at least one biomarker; c) determining the copy number, amount, and/or activity of the at least one biomarker in a control; and d) comparing the copy number, amount, and/or activity of the at least one biomarker detected in steps b) and c); wherein the presence of, or a significant increase in, the copy number, amount, and/or activity of, the at least one biomarker in the subject sample relative to the control copy number, amount, and/or activity of the at least one biomarker indicates that the subject afflicted with the cancer or at risk for developing the cancer would benefit from inhibiting the copy number, amount, and/or activity of the at least one biomarker; a method for predicting the clinical outcome of a subject afflicted with a cancer expressing one or more biomarkers or a fragment thereof to treatment with an immunotherapy, the method comprising: a) determining the copy number, amount, and/or activity of at least one biomarker in a subject sample; b) determining the copy number, amount, and/or activity of the at least one biomarker in a control having a good clinical outcome; and c) comparing the copy number, amount, and/or activity of the at least one biomarker in the subject sample and in the control; wherein the presence of, or a significant increase in, the copy number, amount, and/or activity of, the at least one biomarker in the subject sample as compared to the copy number, amount and/or activity in the control, is an indication that the subject has a poor clinical outcome; a method for monitoring the progression of a cancer in a subject, wherein the subject is administered a therapeutically effective amount of an agent that inhibits the copy number, amount, and/or activity of at least one biomarker and an immunotherapy, the method comprising: a) detecting in a subject sample at a first point in time the copy number, amount, and/or activity of at least one biomarker; b) repeating step a) at a subsequent point in time; and c) comparing the amount or activity of at least one biomarker detected in steps a) and b) to monitor the progression of the cancer in the subject; a method of assessing the efficacy of an agent that inhibits the copy number, amount, and/or activity of at least one biomarker and an immunotherapy for treating a cancer in a subject, comprising: a) detecting in a subject sample at a first point in time the copy number, amount, and/or activity of at least one biomarker; b) repeating step a) during at least one subsequent point in time after administration of the agent and the immunotherapy; and c) comparing the copy number, amount, and/or activity detected in steps a) and b), wherein the absence of, or a significant decrease in, the copy number, amount, and/or activity of, the at least one biomarker, in the subsequent sample as compared to the copy number, amount, and/or activity in the sample at the first point in time, indicates that the agent and immunotherapy treats the cancer in the subject; these shared technical features do not represent a contribution over the prior art.

Specifically, US 2016/0122829 A1 to Dana Farber Cancer Institute, Inc. discloses a method of treating a subject afflicted with a cancer (The present invention relates to methods for identifying, assessing, preventing, and treating cancer, Abstract) comprising administering to the subject a therapeutically effective amount of an agent that inhibits the copy number, the expression level, and/or the activity of one or more biomarkers or a fragment thereof (the method comprises administering an agent that downregulates the copy number, level of expression, or level of activity of one or more biomarkers, Para. [0023]), in combination with an immunotherapy (treatment with one or more modulators of PD-1 and/or PD-L1 levels is provided, Para. [0018]); method of killing cancer cells comprising contacting the cancer cells with an agent that inhibits the copy number, the expression level, and/or the activity of one or more biomarkers or a fragment thereof (the method comprises administering an agent that downregulates the copy number, level of expression, or level of activity of one or more biomarkers, Para. [0023]; cancer cell proliferation inhibition, induction of cancer cell apoptosis, enhancement of cancer cell susceptibility to chemotherapeutic agents, Para. [0174]); a method of determining whether a subject afflicted with a cancer or at risk for developing a cancer would benefit from inhibiting the copy number, amount, and/or activity of at least one biomarker (a method of determining whether a subject afflicted with a head, neck, or lung cancer or at risk for developing a head, neck, or lung cancer would

benefit from modulating PD-1 and/or PD-L1 levels is provided, Para. [0015]), the method comprising: a) obtaining a biological sample from the subject (obtaining a biological sample from the subject; Para. [0015]); b) determining the copy number, amount, and/or activity of at least one biomarker (determining the copy number, level of expression, or level of activity of one or more biomarkers, Para. [0015]); c) determining the copy number, amount, and/or activity of the at least one biomarker in a control; and d) comparing the copy number, amount, and/or activity of the at least one biomarker detected in steps b) and c) (determining the copy number, level of expression, or level of activity of the one or more biomarkers in a control; and d) comparing the copy number, level of expression, or level of activity of said one or more biomarkers detected in steps b) and c), Para. [0015]); wherein the presence of, or a significant increase in, the copy number, amount, and/or activity of, the at least one biomarker in the subject sample relative to the control copy number, amount, and/or activity of the at least one biomarker indicates that the subject afflicted with the cancer or at risk for developing the cancer would benefit from inhibiting the copy number, amount, and/or activity of the at least one biomarker (wherein a significant increase in the copy number, level of expression, or level of activity of the one or more biomarkers in the subject sample relative to the control copy number, level of expression, or level of activity of the one or more biomarkers indicates that the subject afflicted with the head, neck, or lung cancer or at risk for developing the head, neck, or lung cancer would benefit from modulating PD-L1 and/or PD-L1 levels, Para. [0015]); a method for predicting the clinical outcome of a subject afflicted with a cancer expressing one or more biomarkers or a fragment thereof to treatment with an immunotherapy (aspect, a method for stratifying subjects afflicted with a head, neck, or lung cancer according to predicted clinical outcome of treatment with one or more modulators of PD-1 and/or PD-L1 levels is provided, Para. [0018]), the method comprising: a) determining the copy number, amount, and/or activity of at least one biomarker in a subject sample (determining the copy number, level of expression, or level of activity of one or more biomarkers listed in Table 1 or a fragment thereof in a subject sample, Para. [0018]); b) determining the copy number, amount, and/or activity of the at least one biomarker in a control having a good clinical outcome (determining the copy number, level of expression, or level of activity of the one or more biomarkers in a control sample, Para. [0018]); and c) comparing the copy number, amount, and/or activity of the at least one biomarker in the subject sample and in the control; wherein the presence of, or a significant increase in, the copy number, amount, and/or activity of, the at least one biomarker in the subject sample as compared to the copy number, amount and/or activity in the control, is an indication that the subject has a poor clinical outcome (comparing the copy number, level of expression, or level of activity of said one or more biomarkers detected in steps a) and b); wherein a significant modulation in the copy number, level of expression, or level of activity of the one or more biomarkers in the subject sample relative to the normal copy number, level of expression, or level of activity of the one or more biomarkers in the control sample predicts the clinical outcome of the patient to treatment with one or more modulators of PD-1 and/or PD-L levels, Para. [0018]); a method for monitoring the progression of a cancer in a subject (a method for monitoring the progression of a head, neck, or lung cancer in a subject is provided, Para. [0016]), wherein the subject is administered a therapeutically effective amount of an agent that inhibits the copy number, amount, and/or activity of at least one biomarker and an immunotherapy (detecting in a subject sample at a first point in time the copy number, level of expression, or level of activity of one or more biomarkers, Para. [0016]), the method comprising: a) detecting in a subject sample at a first point in time the copy number, amount, and/or activity of at least one biomarker; b) repeating step a) at a subsequent point in time; and c) comparing the amount or activity of at least one biomarker detected in steps a) and b) to monitor the progression of the cancer in the subject (repeating step a) at a subsequent point in time; and c) comparing the copy number, level of expression, or level of activity of said one or more biomarkers detected in steps a) and b) to monitor the progression of the head, neck, or lung cancer, wherein the one or more biomarkers comprises soluble PD-L1, Para. [0016]; to determine appropriate threshold values, a particular cancer therapeutic regimen can be administered to a population of subjects and the outcome can be correlated to copy number, level of expression, level of activity, etc. of one or more biomarkers, Para. [0093]); a method of assessing the efficacy of an agent (a method of determining the efficacy of a therapy for inhibiting a head, neck, or lung cancer in a subject is provided, Para. [0020]) that inhibits the copy number, amount, and/or activity of at least one biomarker and an immunotherapy for treating a cancer in a subject (wherein the method comprises: a) determining the copy number, level of expression, or level of activity of one or more biomarkers listed in Table 1 or a fragment thereof in a first sample obtained from the subject prior to providing at least a portion of the therapy to the subject, Para. [0020]), comprising: a) detecting in a subject sample at a first point in time the copy number, amount, and/or activity of at least one biomarker (wherein the method comprises: a) determining the copy number, level of expression, or level of activity of one or more biomarkers listed in Table 1 or a fragment thereof in a first sample obtained from the subject prior to providing at least a portion of the therapy to the subject, Para. [0020]); b) repeating step a) during at least one subsequent point in time after administration of the agent and the immunotherapy; and c) comparing the copy number, amount, and/or activity detected in steps a) and b), wherein the absence of, or a significant decrease in, the copy number, amount, and/or activity of, the at least one biomarker, in the subsequent sample as compared to the copy number, amount, and/or activity in the sample at the first point in time (b) repeating step a) at a subsequent point in time; and c) comparing the copy number, level of expression, or level of activity of said one or more biomarkers detected in steps a) and b) to monitor the progression of the head, neck, or lung cancer, wherein the one or more biomarkers comprises soluble PD-L1, Para. [0016]; comparing the copy number, level of expression, or level of activity of the one or more biomarkers in the first and second samples, wherein a significantly decreased copy number, level of expression, or level of activity of the one or more biomarkers in the second sample, relative to the first sample, is an indication that the therapy as efficacious for inhibiting the head, neck, or lung cancer in the subject. In one embodiment, the one or more biomarkers comprises soluble PD-L1, Para. [0020]), indicates that the agent and immunotherapy treats the cancer in the subject (is an indication that the therapy as efficacious for inhibiting the head, neck, or lung cancer in the subject, Para. [0020]).

The inventions listed in Groups I+ therefore lack unity under Rule 13 because they do not share a same or corresponding special technical features.