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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 July 2015 (09.07.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/103287 A2

(51) International Patent Classification:
*C12Q 1/68* (2006.01)

(21) International Application Number:
PCT/US2014/072793

(22) International Filing Date:
30 December 2014 (30.12.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/921,780 30 December 2013 (30.12.2013) US

(71) Applicants: **THE HENRY M. JACKSON FOUNDATION FOR THE ADVANCEMENT OF MILITARY MEDICINE, INC.** [US/US]; 6720a Rockledge Drive, Suite 100, Bethesda, MD 20817 (US). **GENOMATIX** [DE/DE]; Bayerstrasse 85a, 80335 Munich (DE).

(72) Inventors: **SRIVASTAVA, Shiv, K.**; 10520 Boswell Lane, Potomac, MD 20854 (US). **DOBI, Albert**; 14071 Pellita Terr, Rockville, MD 20850 (US). **PETROVICS, Gyorgy**; 9103 Friars Road, Bethesda, MD 20817 (US). **WERNER, Thomas**; Manresastrasse 48, 80999 Munich (DE). **SEIFERT, Martin**; Am Romerhugel 9, 82396 Fischen/pahl (DE). **SCHERF, Matthias**; Kochelseestrasse 29, 80335 Grobenzell (DE).

(74) Agent: **DONALDSON, Timothy, B.**; Mh2 Technology Law Group, 1951 Kidwell Drive, Suite 550, Tysons Corner, VA 22182 (US).

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

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

**Published**:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) Title: GENOMIC REARRANGEMENTS ASSOCIATED WITH PROSTATE CANCER AND METHODS OF USING THE SAME

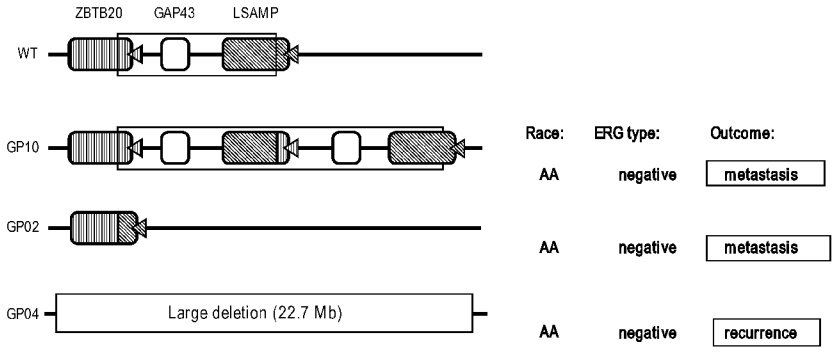


FIG. 1

(57) Abstract: The present disclosure provides genomic arrangements of the chromosome 3q13 region that are associated with prostate cancer, such as rearrangements between the ZBTB20 and LSAMP genes, including gene fusions between the ZBTB20 gene and the LSAMP gene and deletions spanning both genes. The ZBTB20/LSAMP genomic rearrangement serves as a biomarker for prostate cancer and can be used to stratify prostate cancer based on ethnicity or the severity or aggressiveness of prostate cancer and/or identify a patient for prostate cancer treatment. Another aspect involves discovering that deletions of the PTEN gene are observed predominately in prostate cancer from subjects of Caucasian descent. Also provided are kits for diagnosing and prognosing prostate cancer.

WO 2015/103287 A2

Customer No. 39,878
Attorney Docket No. HMJ-149-PCT

# GENOMIC REARRANGEMENTS ASSOCIATED WITH PROSTATE CANCER AND METHODS OF USING THE SAME

## CROSS REFERENCE TO RELATED APPLICATIONS

[001] This application claims the benefit of, and relies on the filing date of, U.S. provisional patent application number 61/921,780, filed 30 December 2013, the entire disclosure of which is incorporated herein by reference.

## GOVERNMENT INTEREST

[002] This invention was made in part with Government support. The Government has certain rights in the invention.

## SEQUENCE LISTING

[003] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on December 30, 2014, is named HMJ-149-PCT_SL.txt and is 205,868 bytes in size.

## BACKGROUND

[004] In 2013 an estimated 238,590 men will be diagnosed with carcinoma of the prostate (CaP) and an estimated 29,720 men will die from the disease [1]. This malignancy is the second leading cause of cancer-related death in men in the United States. In addition, African American (AA) men have the highest incidence and mortality from CaP compared with other races [1]. The racial disparity exists from presentation and diagnosis through treatment, survival, and quality of life [2]. Researchers have suggested that socio-economic status (SES) contributes significantly to these disparities including CaP-specific mortality [3]. As well, there is evidence that reduced access to care is associated with poor CaP outcomes, which is more prevalent among AA men than Caucasian American (CA) men [4].

[005] However, there are populations in which AA men have similar outcomes to CA men. Sridhar and colleagues [5] published a meta-analysis in which they concluded that when SES is accounted for, there are no differences in the overall and CaP-specific survival between AA and CA men. Similarly, the military and veteran populations (systems of equal access and screening) do not observe differences in survival across race [6], and differences

in pathologic stage at diagnosis narrowed by the early 2000s in a veterans' cohort [7]. Of note, both of these studies showed that AA men were more likely to have higher Gleason scores and PSA levels than CA men [6, 7].

[006] While socio-economic factors may contribute to CaP outcomes, they do not seem to account for all variables associated with the diagnosis and disease risk. Several studies support that AA men have a higher incidence of CaP compared to CA men [1, 8, 9]. Studies also show that AA men have a significantly higher PSA at diagnosis, higher grade disease on biopsy, greater tumor volume for each stage, and a shorter PSA doubling time before radical prostatectomy [10–12]. Biological differences between prostate cancers from CA and AA men have been noted in the tumor microenvironment with regard to stress and inflammatory responses [13]. Although controversy remains over the role of biological differences, observed differences in incidence and disease aggressiveness at presentation indicate a potential role for different pathways of prostate carcinogenesis between AA and CA men.

[007] Over the past decade, much research has focused on alterations of cancer genes and their effects in CaP [14–16]. Variations in prevalence across ethnicity and race have been noted in the TMPRSS2/ERG gene fusion that is overexpressed in CaP and is the most common known oncogene in CaP [17, 18]. Accumulating data suggest that there are differences of *ERG* oncogenic alterations across ethnicities [17, 19–21]. Significantly greater *ERG* expression in CA men compared to AA men was noted in initial papers describing *ERG* overexpression and *ERG* splice variants [17, 21]. The difference is even more pronounced between CA and AA (50% versus 16%) in patients with high Gleason grade (8-10) tumors. [Ferrell et al., manuscript]. Thus, ERG is a major somatic gene alteration between these ethnic groups. Yet beyond TMPRSS2/ERG, little is known regarding the genetic basis for the CaP disparity between AA and CA men remains unknown [24].

[008] Therefore, new biomarkers and therapeutic markers that are specific for distinct ethnic populations and provide more accurate diagnostic and/or prognostic potential are needed.

## SUMMARY

[009] The present disclosure provides a genomic arrangement that occurs in chromosome region 3q13 and involves a ZBTB20 gene and an LSAMP gene and methods of diagnosing and prognosing prostate cancer based on the detection of the ZBTB20/LSAMP

genomic arrangement in a biological sample comprising prostate cells. The ZBTB20/LSAMP genomic arrangement can be a gene fusion between the ZBTB20 gene and the LSAMP gene, a gene inversion, a gene deletion, or a gene duplication.

[0010] Detecting the ZBTB20/LSAMP genomic rearrangement in prostate cells from a subject indicates that the subject has prostate cancer or an increased likelihood to develop prostate cancer or characterizes the prostate cancer in the subject as being an aggressive form of prostate cancer or as having an increased risk of developing into an aggressive form of prostate cancer. The ZBTB20/LSAMP gene rearrangements can be measured at either the nucleic acid or protein level.

[0011] In one embodiment, prostate cancer from the subject does not express a gene fusion between ERG and an androgen regulated gene, such as TMPRSS2. In another embodiment, the subject is of African descent. In this way, the ZBTB20/LSAMP genomic rearrangement can be used to prognose the severity of prostate cancer within a particular ethnic group, as the examples show that subjects of African descent who possess the ZBTB20/LSAMP genomic rearrangement in prostate cells, but not the TMPRSS2/ERG fusion, consistently develop an aggressive form of prostate cancer.

[0012] Given the prognostic value of the ZBTB20/LSAMP genomic rearrangement, the methods may further comprise a step of selecting a treatment regimen for the subject based on the detection of the ZBTB20/LSAMP genomic rearrangement or of treating the subject if the genomic rearrangement is detected in the biological sample obtained from the subject. Alternatively, the methods may further comprise a step of increasing the frequency of monitoring the subject for the development of prostate cancer or a more aggressive form of prostate cancer.

[0013] Another aspect is directed to compositions for diagnosing or prognosing prostate cancer. In one embodiment, the composition comprises a polynucleotide probe, wherein the polynucleotide probe hybridizes under high stringency conditions to a junction of a chimeric nucleic acid, wherein the chimeric nucleic acid comprises a first portion from a ZBTB20 gene and a second portion from a LSAMP gene ("the ZBTB20/LSAMP polynucleotide probe"). In one embodiment, the first portion comprises exon 1 of the ZBTB20 gene and the second portion comprises exon 3* or exon 4 of LSAMP. In other embodiments, the polynucleotide probe hybridizes under high stringency conditions to exon 3* of a LSAMP gene or a junction of a chimeric nucleic acid, wherein the chimeric nucleic acid comprises a first portion from exon 3* of a LSAMP gene and a second portion from

exon 4 of a LSAMP gene ("the exon 3*/exon 4 polynucleotide probe"). In other embodiments, the polynucleotide probe hybridizes under high stringency conditions to exon 0* of a LSAMP gene or a junction of a chimeric nucleic acid, wherein the chimeric nucleic acid comprises a first portion from exon 0* of a LSAMP gene and a second portion from exon 1 of a LSAMP gene "the exon 0*/exon 1 polynucleotide probe"). The polynucleotide probe is optionally labeled.

[0014] Another aspect is directed to a kit comprising the composition with the ZBTB20/LSAMP polynucleotide probe, the exon 3*/exon 4 polynucleotide probe, or the exon 0*/exon 1 polynucleotide probe and a second composition comprising a polynucleotide probe that hybridizes under high stringency conditions to a gene selected from COL10A1, HOXC4, ESPL1, MMP9, ABCA13, PCDHGA1, and AGSK1. The polynucleotide probe is optionally labeled. Alternatively, the kit comprises the composition with the ZBTB20/LSAMP polynucleotide probe, the exon 3*/exon 4 polynucleotide probe, or the exon 0*/exon 1 polynucleotide probe and a second composition comprising a polynucleotide probe that hybridizes under high stringency conditions to a gene selected from ERG, AMACR, PCA3, and PSA. The polynucleotide probe is optionally labeled.

[0015] Yet another aspect is directed to a composition comprising a double stranded oligonucleotide duplex, wherein the oligonucleotide duplex comprises a first nucleic acid hybridized to a second nucleic acid, wherein the first nucleic acid comprises a first portion from a ZBTB20 gene fused to a second portion from a LSAMP gene and wherein the second nucleic acid is a polynucleotide probe that is hybridized to a junction between the first portion from the ZBTB20 gene and the second portion from the LSAMP gene. In one embodiment, the first portion comprises exon 1 of the ZBTB20 gene and the second portion comprises exon 3* or exon 4 of LSAMP. In another embodiment, the first nucleic acid comprises a first portion from exon 3* of a LSAMP gene fused to a second portion from exon 4 of a LSAMP gene and wherein the second nucleic acid is a polynucleotide probe that is hybridized to a junction between the first portion from exon 3* of the LSAMP gene and the second portion from exon 4 of the LSAMP gene. In another embodiment, the first nucleic acid comprises a first portion from exon 0* of a LSAMP gene fused to a second portion from exon 1 of a LSAMP gene and wherein the second nucleic acid is a polynucleotide probe that is hybridized to a junction between the first portion from exon 0* of the LSAMP gene and the second portion from exon 4 of the LSAMP gene. The polynucleotide probe is optionally labeled.

[0016] Another aspect is directed to an isolated antibody that binds to a polypeptide encoded by a gene fusion, wherein the gene fusion has a first portion from a ZBTB20 gene and a second portion from a LSAMP gene. In one embodiment the polypeptide is a truncated LSAMP polypeptide. In another embodiment, the antibody binds to an epitope present in the polypeptide encoded by the gene fusion that is not present in either the wild type ZBTB20 protein or the wild type LSAMP protein. The antibody is optionally labeled.

[0017] Another aspect is directed to a composition for amplifying a gene fusion, wherein the gene fusion has a first portion from a ZBTB20 gene and a second portion from a LSAMP gene. In one embodiment, the first portion comprises exon 1 of the ZBTB20 gene and the second portion comprises exon 3* or exon 4 of LSAMP. In one embodiment, the composition comprises a first and a second primer, wherein the first and the second primer are capable of amplifying a nucleotide sequence from the gene fusion that spans the junction between the first portion of the gene fusion from the ZBTB20 gene and the second portion of the gene fusion from the LSAMP gene. In one embodiment, the first primer hybridizes to the first portion of the gene fusion from the ZBTB20 gene and the second primer hybridizes to the second portion of the gene fusion from the LSAMP gene.

[0018] Exon 3* of the LSAMP locus is a newly recognized LSAMP exon that arises as a result of a genomic rearrangement of the ZBTB20 and LSAMP genes. Thus, another aspect is directed to a composition for amplifying exon 3* of a LSAMP gene or a gene fusion, wherein the gene fusion has a first portion from exon 3* of a LSAMP gene and a second portion from exon 4 of a LSAMP gene. In one embodiment, the composition comprises a first and a second primer, wherein the first and the second primer are capable of amplifying a nucleotide sequence within exon 3* of LSAMP that is unique to exon 3* and is not found in other gene sequences or a nucleotide sequence from the gene fusion that spans the junction between a first portion from exon 3* of a LSAMP gene and a second portion from exon 4 of a LSAMP gene.

[0019] Exon 0* of the LSAMP locus is a newly recognized LSAMP exon that arises as a result of a genomic rearrangement of the ZBTB20 and LSAMP genes. Thus, another aspect is directed to a composition for amplifying exon 0* of a LSAMP gene or a gene fusion, wherein the gene fusion has a first portion from exon 0* of a LSAMP gene and a second portion from exon 1 of a LSAMP gene. In one embodiment, the composition comprises a first and a second primer, wherein the first and the second primer are capable of amplifying a nucleotide sequence within exon 0* of LSAMP that is unique to the 0* exon and

is not found in other gene sequences or a nucleotide sequence from the gene fusion that spans the junction between a first portion from exon 0* of a LSAMP gene and a second portion from exon 1 of a LSAMP gene.

[0020] Another genomic rearrangement of interest that is associated with prostate cancer is the PTEN deletion. While the PTEN gene is a common tumor suppressor and its deletion is known to be associated with cancer, it has been surprisingly discovered that the PTEN deletion occurs with significantly different frequencies in different ethnic groups and is markedly absent in subjects of African descent. Understanding the stratification of cancer-related genomic rearrangements, such as the PTEN deletion, between different patient populations provides important information to instruct treatment options for prostate cancer patients.

[0021] Accordingly, one aspect is directed to a method of selecting a targeted prostate cancer treatment for a patient of African descent, wherein the method comprises (a) excluding prostate cancer therapy that targets the PI3K/PTEN/Akt/mTOR pathway as a treatment option; and selecting an appropriate prostate cancer treatment. In one embodiment, the method further comprises a step of testing a biological sample from the patient, wherein the biological sample comprises prostate cells to confirm that the prostate cells to do not contain a PTEN gene deletion.

## BRIEF DESCRIPTION OF THE DRAWINGS

[0022] The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate certain embodiments, and together with the written description, serve to explain certain principles of the antibodies and methods disclosed herein.

[0023] **Figure 1** is a map of the wild type chromosome region 3q13, showing the ZBTB20, GAP43, and LSAMP genes, as well genomic rearrangements of chromosome region 3q13 identified in three different AA patients, all of whom developed an aggressive form of prostate cancer.

[0024] **Figure 2** shows the genomic arrangement of chromosome region 3q13 in patient GP10 and provides the cDNA sequence of the junction between exon 1 of ZBTB20 and exon 4 of LSAMP that results from the genomic arrangement. Figure 2 discloses SEQ ID NOS 49-50, respectively, in order of appearance.

[0025] **Figure 3** shows a schematic diagram of a system according to some embodiments of the invention. In particular, this figure illustrates various hardware,

software, and other resources that may be used in implementations of computer system 106 according to disclosed systems and methods. In embodiments as shown, computer system 106 may include one or more processors 110 coupled to random access memory operating under control of or in conjunction with an operating system. The processor(s) 110 in embodiments may be included in one or more servers, clusters, or other computers or hardware resources, or may be implemented using cloud-based resources. The operating system may be, for example, a distribution of the Linux™ operating system, the Unix™ operating system, or other open-source or proprietary operating system or platform. Processor(s) 110 may communicate with data store 112, such as a database stored on a hard drive or drive array, to access or store program instructions other data.

[0026] Processor(s) 110 may further communicate via a network interface 108, which in turn may communicate via the one or more networks 104, such as the Internet or other public or private networks, such that a query or other request may be received from client 102, or other device or service. Additionally, processor(s) 110 may utilize network interface 108 to send information, instructions, workflows query partial workflows, or other data to a user via the one or more networks 104. Network interface 104 may include or be communicatively coupled to one or more servers. Client 102 may be, e.g., a personal computer coupled to the internet.

[0027] Processor(s) 110 may, in general, be programmed or configured to execute control logic and control operations to implement methods disclosed herein. Processors 110 may be further communicatively coupled (i.e., coupled by way of a communication channel) to co-processors 114. Co-processors 114 can be dedicated hardware and/or firmware components configured to execute the methods disclosed herein. Thus, the methods disclosed herein can be executed by processor 110 and/or co-processors 114.

[0028] Other configurations of computer system 106, associated network connections, and other hardware, software, and service resources are possible.

[0029] **Figure 4A** shows the hybridization of two PTEN FISH probes and two chromosome 10 centromeric probes, indicating the presence of two (diploid) wild type PTEN alleles on chromosome 10 in AD CaPs. Heterozygous PTEN deletion indicated by the loss of one copy of PTEN (absence of one PTEN-specific FISH signal) in the nuclei of a CD CaP. Deletion may occur on either or both the maternal or paternal chromosomes.

[0030] **Figure 4B** shows heterozygous deletion between the ZBTB20-LSAMP region indicated by the loss of one copy of the signal within the nuclei of an AD CaP. Centromeric probes detect two copies of chromosome 3.

[0031] **Figure 5** shows the exon (E) structure (in the 5' to 3' direction) of 10 different ZBTB20-LSAMP fusion transcripts and the exon structure of an alternatively spliced LSAMP (LPCS1) transcript. E1, E1A, E1B, and E1C represent four variants of exon 1 of ZBTB20. E0, E1, E2, E3, E3*, E4, E5, E6, and E7 represent the exons of LSAMP. The numbers in parentheses represent the number of nucleotides in each exon. The asterix indicates previously unannotated exon variants.

## DETAILED DESCRIPTION

[0032] Reference will now be made in detail to various exemplary embodiments, examples of which are illustrated in the accompanying drawings. It is to be understood that the following detailed description is provided to give the reader a fuller understanding of certain embodiments, features, and details of aspects of the invention, and should not be interpreted as a limitation of the scope of the invention.

### Definitions

[0033] In order that the present invention may be more readily understood, certain terms are first defined. Additional definitions are set forth throughout the detailed description.

[0034] The term "**of African descent**" refers to individuals who self-identify as being of African descent, including individuals who self-identify as being African-American, and individuals determined to have genetic markers correlated with African ancestry, also called Ancestry Informative Markers (AIM), such as the AIMs identified in Judith Kidd et al., Analyses of a set of 128 ancestry informative single-nucleotide polymorphisms in a global set of 119 population samples, Investigative Genetics, (2):1, 2011, which reference is incorporated by reference in its entirety.

[0035] The term "**of Caucasian descent**" refers to individuals who self-identify as being of Caucasian descent, including individuals who self-identify as being Caucasian-American, and individuals determined to have genetic markers correlated with Caucasian (e.g., European, North African, or Asian (Western, Central or Southern) ancestry, also called Ancestry Informative Markers (AIM), such as the AIMs identified in Judith Kidd et al., Analyses of a set of 128 ancestry informative single-nucleotide polymorphisms in a global set

of 119 population samples, Investigative Genetics, (2):1, 2011, which reference is incorporated by reference in its entirety.

[0036] The term "**antibody**" refers to an immunoglobulin or antigen-binding fragment thereof, and encompasses any polypeptide comprising an antigen-binding fragment or an antigen-binding domain. The term includes but is not limited to polyclonal, monoclonal, monospecific, polyspecific, humanized, human, single-chain, chimeric, synthetic, recombinant, hybrid, mutated, grafted, and *in vitro* generated antibodies. Unless preceded by the word "intact", the term "antibody" includes antibody fragments such as Fab, $F(ab')_2$, Fv, scFv, Fd, dAb, and other antibody fragments that retain antigen-binding function. Unless otherwise specified, an antibody is not necessarily from any particular source, nor is it produced by any particular method.

[0037] The terms "**antigen-binding domain**" and "**antigen-binding fragment**" refer to a part of an antibody molecule that comprises amino acids responsible for the specific binding between antibody and antigen. For certain antigens, the antigen-binding domain or antigen-binding fragment may only bind to a part of the antigen. The part of the antigen that is specifically recognized and bound by the antibody is referred to as the "epitope" or "antigenic determinant." Antigen-binding domains and antigen-binding fragments include Fab (Fragment antigen-binding); a $F(ab')_2$ fragment, a bivalent fragment having two Fab fragments linked by a disulfide bridge at the hinge region; Fv fragment; a single chain Fv fragment (scFv) *see e.g.*, Bird *et al.* (1988) *Science* 242:423-426; and Huston *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:5879-5883); a Fd fragment having the two $V_{H}$ and $C_{H}1$ domains; dAb (Ward *et al.*, (1989) *Nature* 341:544-546), and other antibody fragments that retain antigen-binding function. The Fab fragment has $V_H$-$C_H1$ and $V_L$-$C_L$ domains covalently linked by a disulfide bond between the constant regions. The $F_v$ fragment is smaller and has $V_{H}$ and $V_L$ domains non-covalently linked. To overcome the tendency of non-covalently linked domains to dissociate, a $scF_v$ can be constructed. The $scF_v$ contains a flexible polypeptide that links (1) the C-terminus of $V_H$ to the N-terminus of $V_L$, or (2) the C-terminus of $V_L$ to the N-terminus of $V_H$. A 15-mer $(Gly_4Ser)_3$ peptide (SEQ ID NO:48) may be used as a linker, but other linkers are known in the art. These antibody fragments are obtained using conventional techniques known to those with skill in the art, and the fragments are evaluated for function in the same manner as are intact antibodies.

[0038] The term "**detecting**" or "**detection**" means any of a variety of methods known in the art for determining the presence or amount of a nucleic acid or a protein. As

used throughout the specification, the term "detecting" or "detection" includes either qualitative or quantitative detection.

[0039] The term "**therapeutically effective amount**" refers to a dosage or amount that is sufficient for treating an indicated disease or condition.

[0040] The term "**gene expression profile**" refers to the expression levels of a plurality of genes in a sample. As is understood in the art, the expression level of a gene can be analyzed by measuring the expression of a nucleic acid (e.g., genomic DNA or mRNA) or a polypeptide that is encoded by the nucleic acid.

[0041] The term "**isolated**," when used in the context of a polypeptide or nucleic acid refers to a polypeptide or nucleic acid that is substantially free of its natural environment and is thus distinguishable from a polypeptide or nucleic acid that might happen to occur naturally. For instance, an isolated polypeptide or nucleic acid is substantially free of cellular material or other polypeptides or nucleic acids from the cell or tissue source from which it was derived.

[0042] The terms "**polypeptide**," "**peptide**," and "**protein**" are used interchangeably herein to refer to polymers of amino acids.

[0043] The term "**polypeptide probe**" as used herein refers to a labeled (e.g., isotopically labled) polypeptide that can be used in a protein detection assay (e.g., mass spectrometry) to quantify a polypeptide of interest in a biological sample.

[0044] The term "**primer**" means a polynucleotide capable of binding to a region of a target nucleic acid, or its complement, and promoting nucleic acid amplification of the target nucleic acid. Generally, a primer will have a free 3' end that can be extended by a nucleic acid polymerase. Primers also generally include a base sequence capable of hybridizing via complementary base interactions either directly with at least one strand of the target nucleic acid or with a strand that is complementary to the target sequence. A primer may comprise target-specific sequences and optionally other sequences that are non-complementary to the target sequence. These non-complementary sequences may comprise, for example, a promoter sequence or a restriction endonuclease recognition site.

[0045] A "**variation**" or "**variant**" refers to an allele sequence that is different from the reference at as little as a single base or for a longer interval.

[0046] The term "**genomic rearrangement of the ZBTB20 and LSAMP genes**" and the like refers to any rearrangement of the ZBTB20 and LSAMP genes that is associated with prostate cancer and can include a gene fusion between the ZBTB20 gene and the LSAMP

gene, a gene inversion involving the ZBTB20 gene and the LSAMP gene, a gene deletion involving the ZBTB20 and LSAMP genes (or a portion of one or both genes), or a gene duplication involving the ZBTB20 and LSAMP genes.

[0047] The term "**ERG**" or "**ERG gene**" refers to Ets-related gene (ERG), which has been assigned the unique Hugo Gene Nomenclature Committee (HGNC) identifier code: HGNC:3446, and includes ERG gene fusion products that are prevalent in prostate cancer, including TMPRSS2-ERG fusion products. Analyzing the expression of ERG or the ERG gene includes analyzing the expression of ERG gene fusion products that are associated with prostate cancer, such as TMPRSS2-ERG.

[0048] As used herein, the term "**aggressive form of prostate cancer**" refers to prostate cancer with a primary Gleason score of 4 or 5 (also known as "poorly differentiated" prostate cancer or prostate cancer that has metastasized or has recurred following prostatectomy).

[0049] As used herein, the term "**Gleason 6-7**" refers to Gleason grade 3 +3 and 3+4. It is also referred to in the art as primary pattern 3 or primary Gleason pattern 3.

**ZBTB20/LSAMP Genomic Rearrangement**

[0050] Next generation sequencing techniques were used to identify new biomarkers and therapeutic targets for CaP. High quality genome sequence data and coverage obtained from histologically defined and precisely dissected primary CaP specimens (80-95% tumor, primary Gleason pattern 3) was compared between cohorts of 7 patients of Caucasian descent and 7 patients of African descent (28 samples total including matched controls from each patient) to evaluate the observed disparities of CaP incidence and mortality between the two ethnic groups. These data and analyses provide the first evaluation of prostate cancer genomes from CaP patients of African descent ("AD") and Caucasian descent ("CD") that have been matched for clinic-pathologic features.

[0051] Whole genome sequence analysis of these prostate cancer samples identified a novel genomic rearrangement between the ZBTB20 (zinc finger and BTB containing 20) and LSAMP (limbic system associated membrane protein) genes. Four of the 7 samples from subjects of African descent were negative for the TMPRSS2/ERG fusion, the most prevalent gene fusion identified to date in prostate cancer. The ZBTB20-LSAMP region was rearranged or deleted in three of the four TMPRSS2/ERG negative samples from subjects of African descent. All three of the ZBTB20/LSAMP positive, TMPRSS2/ERG negative

subjects developed an aggressive form of prostate cancer, with two experiencing metastasis and one developing recurrence, suggesting that the ZBTB20/LSAMP genomic rearrangement is an indicator of a more aggressive prostate cancer phenotype, particularly in patients who do not express the TMPRSS2/ERG fusion. As patients of AD express the TMPRSS2/ERG fusion at lower frequencies than patients of CD, the genomic rearrangement of the ZBTB20 and LSAMP genes, may represent a particularly useful biomarker for detecting a more aggressive prostate cancer phenotype in patients of AD.

[0052] The unique identifier code assigned by HGNC for the LSAMP gene is HGNC:6705. The Entrez Gene code for LSAMP is 4045. The nucleotide and amino acid sequences of LSAMP are known and represented by the NCBI Reference Sequence NM_002338.3, GI:257467557 (SEQ ID NO:9 and SEQ ID NO:10), which sequences are incorporated by reference in their entirety. The chromosomal location of the LSAMP gene is 3q13.2-q21. The LSAMP gene encodes a neuronal surface glycoprotein found in cortical and subcortical regions of the limbic system. LSAMP has been reported as a tumor suppressor gene (Baroy et al., 2014, Mol Cancer 28;13:93). For example, Kuhn et al. reported a recurrent deletion in chromosome region 3q13.31, which contains the LSAMP gene, in a subset of core binding factor acute myeloid leukemia [29]. In osteosarcoma, chromosome region 3q13.31 was identified as the most altered genomic region, with most alterations taking the form of a deletion, including, in certain instances, deletion of a region that contains the LSAMP gene [30]. A chromosomal translocation (t1;3) with a breakpoint involving the NORE1 gene of chromosome region 1q32.1 and the LSAMP gene of chromosome region 3q13.3 was identified in clear cell renal carcinomas [31]. A chromosomal translocation in epithelial ovarian carcinoma has also been identified [32]. Although single nucleotide variations of LSAMP has been shown to be a significant predictor of prostate cancer-specific mortality [33], genomic rearrangement of LSAMP has never been reported in prostate cancer and has never been described as a fusion with ZBTB20 in any type of cancer.

[0053] The unique identifier code assigned by HGNC for the ZBTB20 gene is HGNC:13503. The Entrez Gene code for ZBTB20 is 26137. ZBTB20 is a DNA binding protein and is believed to be a transcription factor. There are at least 7 alternative transcript variants. There are at least four distinct promoters that can initiate transcription from at least four distinct sites within the ZBTB20 locus, producing four variants of exon 1 of ZBTB20: E1, E1A, E1B, and E1C. Representative nucleotide and amino acid sequences of ZBTB20 variant 1 are known and represented by the NCBI Reference Sequence NM_001164342.1

GI:257900532 (SEQ ID NO:11 and SEQ ID NO:12), which sequences are incorporated by reference in their entirety. Variant 2 differs from variant 1 in the 5' untranslated region, lacks a portion of the 5' coding region, and initiates translation at a downstream start codon, compared to variant 1. The encoded isoform (2) has a shorter N-terminus compared to isoform 1. Variants 2-7 encode the same isoform (2). Representative nucleotide and amino acid sequences of ZBTB20 variant 2 are known and represented by the NCBI Reference Sequence NM_015642.4, GI:257900536 (SEQ ID NO:13 and SEQ ID NO:14), which sequences are incorporated by reference in their entirety. The chromosomal location of the ZBTB20 gene is 3q13.2.

[0054] Certain embodiments are directed to a method of collecting data for use in diagnosing or prognosing CaP, the method comprising detecting in a biological sample comprising prostate cells (or nucleic acid or polypeptides isolated from prostate cells) a genomic rearrangement of the ZBTB20 and LSAMP genes. The method may optionally include an additional step of diagnosing or prognosing CaP using the collected gene expression data. In one embodiment, detecting a genomic rearrangement of the ZBTB20 and LSAMP genes indicates the presence of CaP in the biological sample or an increased likelihood of developing CaP. In another embodiment detecting a genomic rearrangement of the ZBTB20 and LSAMP genes indicates the presence of an aggressive form of CaP in the biological sample or an increased likelihood of developing an aggressive form of CaP.

[0055] In one embodiment, the genomic rearrangement comprises a gene fusion between the ZBTB20 gene and the LSAMP gene, such as a fusion between exon 1 (e.g., E1, E1A, E1B, or E1C) of the ZBTB20 gene and exon 4 of the LSAMP gene. In another embodiment, the genomic rearrangement comprises a gene inversion involving the ZBTB20 gene and the LSAMP gene. In another embodiment, the genomic rearrangement comprises a deletion in chromosome region 3q13, wherein the deletion spans both the ZBTB20 and LSAMP genes (or a portion of one or both genes). In yet another embodiment, the genomic rearrangement comprises a gene duplication involving the ZBTB20 and LSAMP genes.

[0056] The methods of collecting data or diagnosing and/or prognosing CaP may further comprise detecting expression of other genes associated with prostate cancer, including, but not limited to COL10A1, HOXC4, ESPL1, MMP9, ABCA13, PCDHGA1, and AGSK1. The unique identifier codes assigned by HGNC and Entrez Gene for these genes that are more frequently overexpressed in patients of African descent and the accession

number of representative sequences are provided in Table 1, which sequences are hereby incorporated by reference in their entirety.

[0057] **Table 1**

| Gene | HGNC ID | Entrez Gene ID | NCBI Reference | SEQ ID NOs. |
|---|---|---|---|---|
| COL10A1 | 2185 | 1300 | NM_000493.3 GI:98985802 | 17 and 18 |
| HOXC4 | 5126 | 3221 | NM_014620.5 GI:546232084 | 19 and 20 |
| ESPL1 | 16856 | 9700 | NM_012291.4 GI:134276942 | 21 and 22 |
| MMP9 | 7176 | 4318 | NM_004994.2 GI:74272286 | 23 and 24 |
| ABCA13 | 14638 | 154664 | AY204751.1 GI:30089663 | 25 and 26 |
| PCDHGA1 | 8696 | 56114 | NM_018912.2 GI:14196453 | 27 and 28 |
| AGSK1 | N/A | 80154 | NR_026811 GI:536293433<br>NR_033936.3 GI:536293365<br>NR_103496.2 GI:536293435 | 29-31 |

[0058] In another embodiment, the methods of collecting data or diagnosing and/or prognosing CaP may further comprise detecting expression of other genes associated with prostate cancer, including, but not limited to ERG, PSA, and PCA3.

**PTEN Deletion**

[0059] PTEN (phosphatase and tensin homolog) is a known tumor suppressor gene that is mutated in a large number of cancers at high frequency. The protein encoded by this gene is a phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase. It contains a tensin like domain as well as a catalytic domain similar to that of the dual specificity protein tyrosine phosphatases. Unlike most of the protein tyrosine phosphatases, PTEN preferentially dephosphorylates phosphoinositide substrates. It negatively regulates intracellular levels of phosphatidylinositol-3,4,5-trisphosphate in cells and functions as a tumor suppressor by negatively regulating AKT/PKB signaling pathway. Activation of growth factor receptors by binding of a growth factor to its receptor or by mutation of the growth factor receptor leads to activation of the PI3K/PTEN/Akt/mTOR cascade, which, among other things, leads to the activation of certain transcription factors [28], which reference is hereby incorporated by reference in its entirety. PTEN normally acts to down regulate this pathway. Thus, in

cancers that contain a PTEN gene deletion, the expression of the Akt gene and activation of mTOR is frequently increased.

[0060] The unique identifier codes assigned by HGNC and Entrez Gene for the PTEN gene are HGNC:9588 and Entrez Gene:5728, respectively. The accession number of representative PTEN nucleic acid and polypeptide sequences is NM_000314.4, GI:257467557 (SEQ ID NO:15 and SEQ ID NO:16), which sequences are incorporated by reference in their entirety. The chromosomal location of the PTEN gene is 10q23.

[0061] Whole genome sequence analysis of CD and AD prostate cancer samples disclosed a significant disparity between the genomic rearrangement of the PTEN locus in the different ethnic groups. More specifically, PTEN deletion was detected only in patients of Caucasian descent. Additional FISH analysis in a tissue microarray confirmed that PTEN deletion is an infrequent event in the development of prostate cancer in AD men as compared to CD men.

[0062] Accordingly, one aspect is directed to using this discovery about the disparity in the PTEN deletion across ethnic groups to make informed decisions about treatment options available to a subject who has prostate cancer. In particular, given the disclosed disparity in the PTEN deletion in prostate cancer from patients of Caucasian and African descent, as a general rule, prostate cancer therapies that target the PI3K/PTEN/Akt/mTOR pathway [28] should not be selected for patients of African descent. Or, at a minimum, a prostate cancer therapy that targets the PI3K/PTEN/Akt/mTOR pathway [28] should not be considered for a patient of African descent unless it is first confirmed by genetic testing that prostate cells from the patient contain the PTEN deletion. As such, one embodiment is directed to a method of selecting a targeted prostate cancer treatment for a patient of African descent, wherein the method comprises excluding a prostate cancer therapy that targets the PI3K/PTEN/Akt/mTOR pathway [28] as a treatment option; and selecting an appropriate prostate cancer treatment. In one embodiment, the method further comprises a step of testing a biological sample from the patient, wherein the biological sample comprises prostate cells to confirm that the prostate cells to do not contain a PTEN gene deletion.

[0063] There are various inhibitors that target the PI3K/PTEN/Akt/mTOR pathway, including PI3K inhibitors, Akt inhibitors, mTOR inhibitors, and dual PI3K/mTOR inhibitors. PI3K inhibitors include, but are not limited to LY-294002, wortmannin, PX-866, GDC-0941, CAL-10, XL-147, XL-756, IC87114, NVP-BKM120, and NVP-BYL719. Akt inhibitors include, but are not limited to, A-443654, GSK690693, VQD-002 (a.k.a. API-2, triciribine),

KP372-1, KRX-0401 (perifosine), MK-2206, GSK2141795, LY317615 (enzasturin), erucylphosphocholine (ErPC), erucylphosphohomocholine (ErPC3), PBI-05204, RX-0201, and XL-418. mTOR inhibitors include, but are not limited to, rapamycin, modified rapamycins (rapalogs, e.g., CCI-779, afinitor, torisel, temsirolimus), AP-23573 (ridaforolimus), and RAD001 (afinitor, everolimus), metformin, OSI-027, PP-242, AZD8055, AZD2014, palomid 529, WAY600, WYE353, WYE687, WYE132, Ku0063794, and OXA-01. Dual PI3K/mTOR inhibitors include, but are not limited to, PI-103, NVP-BEZ235, PKI-587, PKI-402, PF-04691502, XL765, GNE-477, GSK2126458, and WJD008.

**Detecting ZBTB20/LSAMP or PTEN Deletion**

[0064] Measuring or detecting the expression of a genomic rearrangement of the ZBTB20 and LSAMP genes in the methods described herein comprises measuring or detecting any nucleic acid transcript (e.g., mRNA, cDNA, or genomic DNA) thereof or any protein encoded by such a nucleic acid transcript. Thus, in one embodiment, detecting the presence of the ZBTB20/LSAMP genomic rearrangement in the biological sample comprises detecting a chromosomal rearrangement of genomic DNA having a first portion from the ZBTB20 gene and a second portion from the LSAMP gene. In one embodiment, the chromosomal rearrangement gives rise to a fusion between exon 1 (e.g., E1, E1A, E1B, or E1C) of the ZBTB20 gene and exon3* or exon 4 of the LSAMP gene. In one embodiment, the chromosomal rearrangement comprises the nucleotide sequence of SEQ ID NO:1.

[0065] In another embodiment, the chromosomal rearrangement results in the deletion of the ZBTB20 and LSAMP genes. Thus, in one embodiment, detecting the presence of the ZBTB20/LSAMP genomic rearrangement in the biological sample comprises detecting a deletion in chromosome region 3q13, wherein the deletion spans the ZBTB20 and LSAMP genes.

[0066] In another embodiment, detecting the presence of the ZBTB20/LSAMP genomic rearrangement in the biological sample comprises detecting a chimeric mRNA or cDNA transcript having a first nucleic acid portion from the ZBTB20 gene and a second nucleic acid portion from the LSAMP gene. In one embodiment, the chimeric mRNA or cDNA transcript comprises a fusion between exon 1 (e.g., E1, E1A, E1B, or E1C) of the ZBTB20 gene and exon 4 of the LSAMP gene. For example, the chimeric mRNA or cDNA transcript may comprise the nucleotide sequence of SEQ ID NO:4, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, or SEQ ID NO:40.

In another embodiment, the chimeric mRNA or cDNA transcript comprises a fusion between exon 1 (e.g., E1, E1A, E1B, or E1C) of the ZBTB20 gene and exon 3* of the LSAMP gene. For example, the chimeric mRNA or cDNA transcript may comprise the nucleotide sequence of SEQ ID NO:33, SEQ ID NO:39, or SEQ ID NO:41.

[0067] Exon 3* (SEQ ID NO:45) represents a novel exon sequence from the LSAMP locus that has not been previously annotated and has now been shown to arise following genomic rearrangement of the ZBTB20 and LSAMP genes. Therefore, detecting the presence of the ZBTB20/LSAMP genomic rearrangement in the biological sample may comprise detecting a mRNA or cDNA transcript corresponding to a region of exon 3* that is not present in other genes and, thus, can be used to positively identify exon 3* of the LSAMP gene. Alternatively, detecting the presence of the ZBTB20/LSAMP genomic rearrangement in the biological sample may comprise detecting a mRNA or cDNA transcript comprising exon 3* and exon 4 of the LSAMP gene or a portion thereof that spans the junction between exon 3* and exon 4 of LSAMP and is not present in other genes and, thus, can be used to positively identify the transcript as coming from a fusion of exon 3* and exon 4 of LSAMP.

[0068] Exon 0* (SEQ ID NO:47) represents a novel exon sequence from the LSAMP locus that has not been previously annotated and was identified in an alternatively spliced transcript in a patient having a genomic rearrangement of the ZBTB20 and LSAMP genes. Therefore, detecting the presence of the ZBTB20/LSAMP genomic rearrangement in the biological sample may comprise detecting a mRNA or cDNA transcript corresponding to a region of exon 0* that is not present in other genes and, thus, can be used to positively identify exon 0* of the LSAMP gene. Alternatively, detecting the presence of the ZBTB20/LSAMP genomic rearrangement in the biological sample may comprise detecting a mRNA or cDNA transcript comprising exon 0* and exon 1 of the LSAMP gene or a portion thereof that spans the junction between exon 0* and exon 1 of the LSAMP gene and is not present in other genes and, thus, can be used to positively identify the transcript as coming from a fusion of exon 0* and exon 1 of LSAMP.

[0069] The expression of the ZBTB20/LSAMP genomic rearrangement can be measured or detected by measuring or detecting one or more of the genomic sequences or mRNA/cDNA transcripts corresponding to the genomic rearrangement of the genes, or to all of the genomic sequences or mRNA/cDNA transcripts associated with the genomic rearrangement of the ZBTB20 and LSAMP genes.

[0070] Detecting a deletion in the PTEN gene comprises detecting a deletion in chromosome region 10q23, wherein the deletion spans the PTEN gene or a portion thereof. The PTEN deletion can be measured or detected by measuring or detecting one or more of the genomic sequences or mRNA/cDNA transcripts corresponding to the PTEN deletion, or to all of the genomic sequences or mRNA/cDNA transcripts associated with the PTEN gene.

[0071] Chromosomal rearrangements can be detected using known techniques. For example, fluorescent *in situ* hybridization (FISH) analysis can be used to detect chromosomal rearrangements. In these embodiments, nucleic acid probes that hybridize under conditions of high stringency to the chromosomal rearrangement, such as the ZBTB20/LSAMP chromosomal rearrangement or PTEN deletion, are incubated with a biological sample comprising prostate cells (or nucleic acid obtained therefrom). Other known *in situ* hybridization techniques can be used to detect chromosomal rearrangements, such as ZBTB20/LSAMP or the PTEN deletion. The nucleic acid probes (DNA or RNA) can hybridize to DNA or mRNA and can be designed to detect genomic rearrangements in the ZBTB20 and LSAMP genes or the PTEN gene, such as gene fusion events, amplifications, deletions, or mutations.

[0072] Typically, gene expression can be detected or measured on the basis of mRNA or cDNA levels, although protein levels also can be used when appropriate. Any quantitative or qualitative method for measuring mRNA levels, cDNA, or protein levels can be used. Suitable methods of detecting or measuring mRNA or cDNA levels include, for example, Northern Blotting, RNAse protection assays, microarray analysis, or a nucleic acid amplification procedure, such as reverse-transcription PCR (RT-PCR) or real-time RT-PCR, also known as quantitative RT-PCR (qRT-PCR). Such methods are well known in the art. *See e.g.*, Sambrook et al., Molecular Cloning: A Laboratory Manual, 4$^{th}$ Ed., Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 2012. Other techniques include digital, multiplexed analysis of gene expression, such as the nCounter® (NanoString Technologies, Seattle, WA) gene expression assays, which are further described in [22], [23], US20100112710 and US20100047924, all of which are hereby incorporated by reference in their entirety.

[0073] Detecting a nucleic acid of interest generally involves hybridization between a target (e.g. mRNA, cDNA, or genomic DNA) and a probe. One of skill in the art can readily design hybridization probes for detecting the genomic rearrangement of the ZBTB20 and LSAMP genes or deletion of the PTEN gene. *See e.g.*, Sambrook et al., Molecular Cloning: A Laboratory Manual, 4$^{th}$ Ed., Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 2012.

Each probe should be substantially specific for its target, to avoid any cross-hybridization and false positives. An alternative to using specific probes is to use specific reagents when deriving materials from transcripts (e.g., during cDNA production, or using target-specific primers during amplification). In both cases specificity can be achieved by hybridization to portions of the targets that are substantially unique within the group of genes being analyzed, e.g. hybridization to the polyA tail would not provide specificity. If a target has multiple splice variants, it is possible to design a hybridization reagent that recognizes a region common to each variant and/or to use more than one reagent, each of which may recognize one or more variants.

[0074] Stringency of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured nucleic acid sequences to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature that can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

[0075] "Stringent conditions" or "high stringency conditions," as defined herein, are identified by, but not limited to, those that: (1) use low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) use during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) use 50% formamide, 5XSSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5X Denhardt's solution, sonicated salmon sperm DNA (50μg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2XSSC (sodium chloride/sodium. citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1XSSC containing EDTA at 55°C. "Moderately stringent conditions" are described by, but not

limited to, those in Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and % SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5XSSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5X Denhardt's solution, 10% dextran sulfate, and 20 mg/mL denatured sheared salmon sperm DNA, followed by washing the filters in 1XSSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

[0076] In certain embodiments, microarray analysis or a PCR-based method is used. In this respect, measuring the expression of the genomic rearrangement of the ZBTB20 and LSAMP genes or PTEN deletion in prostate cancer cells can comprise, for instance, contacting a sample containing or suspected of containing prostate cancer cells with polynucleotide probes specific to the ZBTB20/LSAMP genomic rearrangement or PTEN deletion, or with primers designed to amplify a portion of the ZBTB20/LSAMP or PTEN genomic rearrangement, and detecting binding of the probes to the nucleic acid targets or amplification of the nucleic acids, respectively. Detailed protocols for designing PCR primers are known in the art. *See e.g.*, Sambrook et al., Molecular Cloning: A Laboratory Manual, 4th Ed., Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 2012. Similarly, detailed protocols for preparing and using microarrays to analyze gene expression are known in the art and described herein.

[0077] Thus, one aspect is directed to a method of determining if a biological sample comprising nucleic acid contains a gene fusion, where the gene fusion has a first portion from a ZBTB20 gene and a second portion from a LSAMP gene, the method comprising:

[0078] combining the biological sample with at least a first and a second polynucleotide primer under hybridizing conditions, wherein the first polynucleotide primer comprises a sequence that hybridizes to the first portion of the gene fusion from the ZBTB20 gene (e.g., exon 1), and the second polynucleotide primer comprises a sequence that hybridizes to the second portion of the gene fusion from the LSAMP gene (e.g., exon 3* or exon 4), wherein the first and second polynucleotide primers are capable of amplifying a target sequence from the gene fusion that spans the junction between the first portion of the gene fusion from the ZBTB20 gene and the second portion of the gene fusion from the LSAMP gene;

[0079] adding a polymerase activity under conditions that allow for the amplification of the target sequence and production of an amplification product that comprises the target sequence if the gene fusion is present in the biological sample; and

[0080] determining whether the biological sample contains the gene fusion based on the presence or absence of the amplification product.

[0081] Another aspect is directed to a method of determining if a biological sample comprising nucleic acid contains an mRNA or cDNA sequence corresponding to a region of exon 3* that is not present in other genes or a gene fusion where the gene fusion has a first portion from exon 3* of the LSAMP gene and a second portion from exon 4 of the LSAMP gene, the method comprising:

[0082] combining the biological sample with at least a first and a second polynucleotide primer under hybridizing conditions,

(a) wherein the first polynucleotide primer comprises a sequence that hybridizes to a first region of exon 3* and the second polynucleotide primer comprises a sequence that hybridizes to a second region of exon 3*, wherein the first and second polynucleotide primers are capable of amplifying a target sequence from exon 3* that is unique to exon 3* of the LSAMP gene; or

(b) wherein the first polynucleotide primer comprises a sequence that hybridizes to the first portion of the gene fusion from exon 3* and the second polynucleotide primer comprises a sequence that hybridizes to the second portion of the gene fusion from exon 4 of the LSAMP gene, wherein the first and second polynucleotide primers are capable of amplifying a target sequence from the gene fusion that spans the junction between the first portion of the gene fusion from exon 3* of the LSAMP gene and the second portion of the gene fusion from exon 4 of the LSAMP gene;

[0083] adding a polymerase activity under conditions that allow for the amplification of the target sequence and production of an amplification product that comprises the target sequence if the gene fusion is present in the biological sample; and

[0084] determining whether the biological sample contains the mRNA or cDNA sequence corresponding to a region of exon 3* that is not present in other genes or the gene fusion based on the presence or absence of the amplification product.

[0085] Another aspect is directed to a method of determining if a biological sample comprising nucleic acid contains an mRNA or cDNA sequence corresponding to a region of

exon 0* of the LSAMP gene that is not present in other genes or a gene fusion where the gene fusion has a first portion from exon 0* of the LSAMP gene and a second portion from exon 1 of the LSAMP gene, the method comprising:

[0086] combining the biological sample with at least a first and a second polynucleotide primer under hybridizing conditions,

(a) wherein the first polynucleotide primer comprises a sequence that hybridizes to a first region of exon 0* and the second polynucleotide primer comprises a sequence that hybridizes to a second region of exon 0*, wherein the first and second polynucleotide primers are capable of amplifying a target sequence from exon 0* that is unique to exon 0* of the LSAMP gene; or

(b) wherein the first polynucleotide primer comprises a sequence that hybridizes to the first portion of the gene fusion from exon 0* and the second polynucleotide primer comprises a sequence that hybridizes to the second portion of the gene fusion from exon 1 of the LSAMP gene, wherein the first and second polynucleotide primers are capable of amplifying a target sequence from the gene fusion that spans the junction between the first portion of the gene fusion from exon 0* of the LSAMP gene and the second portion of the gene fusion from exon 1 of the LSAMP gene;

[0087] adding a polymerase activity under conditions that allow for the amplification of the target sequence and production of an amplification product that comprises the target sequence if the gene fusion is present in the biological sample; and

[0088] determining whether the biological sample contains the mRNA or cDNA sequence corresponding to a region of exon 0* that is not present in other genes or the gene fusion based on the presence or absence of the amplification product.

[0089] Alternatively or additionally, expression levels of the ZBTB20/LSAMP genomic rearrangement can be determined at the protein level, meaning that when the ZBTB20/LSAMP genomic rearrangement results in a truncated LSAMP protein or a chimeric ZBTB20/LSAMP protein, the levels of such proteins encoded by the ZBTB20/LSAMP genomic rearrangement are measured. Several methods and devices are well known for determining levels of proteins including immunoassays such as described in e.g., U.S. Pat. Nos. 6,143,576; 6,113,855; 6,019,944; 5,985,579; 5,947,124; 5,939,272; 5,922,615; 5,885,527; 5,851,776; 5,824,799; 5,679,526; 5,525,524; 5,458,852; and 5,480,792, each of which is hereby incorporated by reference in its entirety. These assays

include various sandwich, competitive, or non-competitive assay formats, to generate a signal that is related to the presence or amount of a protein of interest. Any suitable immunoassay may be utilized, for example, lateral flow, enzyme-linked immunoassays (ELISA), radioimmunoassays (RIAs), competitive binding assays, and the like. Numerous formats for antibody arrays have been described. Such arrays typically include different antibodies having specificity for different proteins intended to be detected. For example, at least 100 different antibodies are used to detect 100 different protein targets, each antibody being specific for one target. Other ligands having specificity for a particular protein target can also be used, such as the synthetic antibodies disclosed in WO/2008/048970, which is hereby incorporated by reference in its entirety. Other compounds with a desired binding specificity can be selected from random libraries of peptides or small molecules. U.S. Pat. No. 5,922,615, which is hereby incorporated by reference in its entirety, describes a device that uses multiple discrete zones of immobilized antibodies on membranes to detect multiple target antigens in an array. Microtiter plates or automation can be used to facilitate detection of large numbers of different proteins.

[0090] One type of immunoassay, called nucleic acid detection immunoassay (NADIA), combines the specificity of protein antigen detection by immunoassay with the sensitivity and precision of the polymerase chain reaction (PCR). This amplified DNA-immunoassay approach is similar to that of an enzyme immunoassay, involving antibody binding reactions and intermediate washing steps, except the enzyme label is replaced by a strand of DNA and detected by an amplification reaction using an amplification technique, such as PCR. Exemplary NADIA techniques are described in U.S. Patent No. 5,665,539 and published U.S. Application 2008/0131883, both of which are hereby incorporated by reference in their entirety. Briefly, NADIA uses a first (reporter) antibody that is specific for the protein of interest and labelled with an assay-specific nucleic acid. The presence of the nucleic acid does not interfere with the binding of the antibody, nor does the antibody interfere with the nucleic acid amplification and detection. Typically, a second (capturing) antibody that is specific for a different epitope on the protein of interest is coated onto a solid phase (e.g., paramagnetic particles). The reporter antibody/nucleic acid conjugate is reacted with sample in a microtiter plate to form a first immune complex with the target antigen. The immune complex is then captured onto the solid phase particles coated with the capture antibody, forming an insoluble sandwich immune complex. The microparticles are washed to remove excess, unbound reporter antibody/nucleic acid conjugate. The bound nucleic acid

label is then detected by subjecting the suspended particles to an amplification reaction (e.g. PCR) and monitoring the amplified nucleic acid product.

[0091] Although immunoassays have typically been used for the identification and quantification of proteins, recent advances in mass spectrometry (MS) techniques have led to the development of sensitive, high throughput MS protein analyses. The MS methods can be used to detect low concentrations of proteins in complex biological samples. For example, it is possible to perform targeted MS by fractionating the biological sample prior to MS analysis. Common techniques for carrying out such fractionation prior to MS analysis include two-dimensional electrophoresis, liquid chromatography, and capillary electrophoresis [25], which reference is hereby incorporated by reference in its entirety. Selected reaction monitoring (SRM), also known as multiple reaction monitoring (MRM), has also emerged as a useful high throughput MS-based technique for quantifying targeted proteins in complex biological samples, including prostate cancer biomarkers that are encoded by gene fusions (e.g., TMPRSS2/ERG) [26, 27], which references are hereby incorporated by reference in their entirety.

**Samples**

[0092] The methods described in this application involve analysis of the genomic rearrangement of the ZBTB20 and LSAMP genes or PTEN gene in cells, including prostate cells. These prostate cells are found in a biological sample, such as prostate tissue, blood, serum, plasma, urine, saliva, or prostatic fluid. Nucleic acids or polypeptides may be isolated from the cells prior to detecting gene expression.

[0093] In one embodiment, the biological sample comprises prostate tissue and is obtained through a biopsy, such as a transrectal or transperineal biopsy. In another embodiment, the biological sample is urine. Urine samples may be collected following a digital rectal examination (DRE) or a prostate biopsy. In another embodiment, the sample is blood, serum, or plasma, and contains circulating tumor cells that have detached from a primary tumor. The sample may also contain tumor-derived exosomes. Exosomes are small (typically 30 to 100 nm) membrane-bound particles that are released from normal, diseased, and neoplastic cells and are present in blood and other bodily fluids. The methods disclosed in this application can be used with samples collected from a variety of mammals, but preferably with samples obtained from a human subject.

**Prostate Cancer**

[0094] This application discloses certain chromosomal rearrangements between the ZBTB20 and LSAMP genes that are associated with prostate cancer. Detecting a ZBTB20/LSAMP genomic rearrangement in a biological sample can be used to identify cancer cells, such as prostate cancer cells, in a sample or to measure the severity or aggressiveness of prostate cancer, for example, distinguishing between well differentiated prostate (WD) cancer and poorly differentiated (PD) prostate cancer and/or identifying prostate cancer that has metastasized or recurred following prostatectomy or is more likely to metastasize or recur following prostatectomy. This application also discloses that deletion of the tumor suppressor gene, PTEN, occurs predominately, if not exclusively in subjects of Caucasian descent. Conversely, the PTEN deletion is an infrequent event in prostate cancer from subjects of African descent (AD), particularly in Gleason 6-7 prostate cancer from AD subjects. Of note, Gleason 6-7 (also called primary pattern 3) CaP represents the most commonly diagnosed form of CaPs in the PSA screened patient population.

[0095] When prostate cancer is found in a biopsy, it is typically graded to estimate how quickly it is likely to grow and spread. The most commonly used prostate cancer grading system, called Gleason grading, evaluates prostate cancer cells on a scale of 1 to 5, based on their pattern when viewed under a microscope.

[0096] Cancer cells that still resemble healthy prostate cells have uniform patterns with well-defined boundaries and are considered well differentiated (Gleason grades 1 and 2). The more closely the cancer cells resemble prostate tissue, the more the cells will behave like normal prostate tissue and the less aggressive the cancer. Gleason grade 3, the most common grade, shows cells that are moderately differentiated, that is, still somewhat well-differentiated, but with boundaries that are not as well-defined. Poorly-differentiated cancer cells have random patterns with poorly defined boundaries and no longer resemble prostate tissue (Gleason grades 4 and 5), indicating a more aggressive cancer.

[0097] Prostate cancers often have areas with different grades. A combined Gleason score is determined by adding the grades from the two most common cancer cell patterns within the tumor. For example, if the most common pattern is grade 4 and the second most common pattern is grade 3, then the combined Gleason score is 4+3=7. If there is only one pattern within the tumor, the combined Gleason score can be as low as 1+1=2 or as high as 5+5=10. Combined scores of 2 to 4 are considered well-differentiated, scores of 5 to 6 are

considered moderately-differentiated and scores of 7 to 10 are considered poorly-differentiated. Cancers with a high Gleason score are more likely to have already spread beyond the prostate gland (metastasized) at the time they were found.

[0098] In general, the lower the Gleason score, the less aggressive the cancer and the better the prognosis (outlook for cure or long-term survival). The higher the Gleason score, the more aggressive the cancer and the poorer the prognosis for long-term, metastasis-free survival.

**Patient Treatment**

[0099] This application describes methods of diagnosing and prognosing prostate cancer in a sample obtained from a subject, in which gene expression in prostate cells and/or tissues are analyzed. If a sample shows expression of a genomic rearrangement of the ZBTB20 and LSAMP genes, then there is an increased likelihood that the subject has prostate cancer or a more advanced/aggressive form (e.g., PD prostate cancer) of prostate cancer. In the event of such a result, the methods of detecting or prognosing prostate cancer may include one or more of the following steps: informing the patient that they are likely to have prostate cancer or PD prostate cancer; performing confirmatory histological examination of prostate tissue; and/or treating the subject.

[00100] Thus, in certain aspects, if the detection step indicates that prostate cells from the subject have a genomic rearrangement of the ZBTB20 and LSAMP genes, the methods further comprise a step of taking a prostate biopsy from the subject and examining the prostate tissue in the biopsy (e.g., histological examination) to confirm whether the patient has prostate cancer or an aggressive form of prostate cancer. Alternatively, the methods of detecting or prognosing prostate cancer may be used to assess the need for therapy or to monitor a response to a therapy (e.g., disease-free recurrence following surgery or other therapy), and, thus may include an additional step of treating a subject having prostate cancer.

[00101] Prostate cancer treatment options include surgery, radiation therapy, hormone therapy, chemotherapy, biological therapy, or high intensity focused ultrasound. Drugs approved for prostate cancer include: Abiraterone Acetate, Cabazitaxel, Degarelix, Enzalutamide (XTANDI), Jevtana (Cabazitaxel), Prednisone, Provenge (Sipuleucel-T), Sipuleucel-T, or Docetaxel. Thus a method as described in this application may, after a positive result, include a further step of surgery, radiation therapy, hormone therapy, chemotherapy, biological therapy, or high intensity focused ultrasound.

**Computer-Implemented Models**

[00102] In accordance with all aspects and embodiments of the invention, the methods provided may be computer-implemented.

[00103] The status of the genomic rearrangement of the ZBTB20 and LSAMP genes or the PTEN gene can be analyzed and associated with status of a subject (e.g., presence of prostate cancer or severity of disease (e.g., WD or PD prostate cancer)) in a digital computer. Optionally, such a computer is directly linked to a scanner or the like receiving experimentally determined signals related to the expression of a genomic rearrangement of the ZBTB20 and LSAMP genes or the deletion of the PTEN gene. Alternatively, expression levels can be input by other means. The computer can be programmed to convert raw signals into expression levels (absolute or relative), compare measured expression levels with one or more reference expression levels, or a scale of such values. The computer can also be programmed to assign values or other designations to expression levels based on the comparison with one or more reference expression levels, and to aggregate such values or designations for multiple genes in an expression profile. The computer can also be programmed to output a value or other designation providing an indication of the presence or severity of prostate cancer as well as any of the raw or intermediate data used in determining such a value or designation.

[00104] A typical computer (*see* U.S. Pat. No. 6,785,613; FIGS. 4 and 5) includes a bus which interconnects major subsystems such as a central processor, a system memory, an input/output controller, an external device such as a printer via a parallel port, a display screen via a display adapter, a serial port, a keyboard, a fixed disk drive and a port (e.g., USB port) operative to receive an external memory storage device. Many other devices can be connected such as a scanner via I/O controller, a mouse connected to serial port or a network interface. The computer contains computer readable media holding codes to allow the computer to perform a variety of functions. These functions include controlling automated apparatus, receiving input and delivering output as described above. The automated apparatus can include a robotic arm for delivering reagents for determining expression levels, as well as small vessels, e.g., microtiter wells for performing the expression analysis.

[00105] A typical computer system 106 may also include one or more processors 110 coupled to random access memory operating under control of or in conjunction with an operating system as set forth in Figure 3 and discussed above.

[00106] In one embodiment, any of the computer-implemented methods of the invention may comprise a step of obtaining by at least one processor information reflecting whether a biological sample contains a genomic rearrangement of the ZBTB20 and LSAMP genes or a deletion of the PTEN gene. In one embodiment, the biological sample is obtained from a subject of African descent. In another embodiment, the biological sample is obtained from a subject, wherein prostate cancer from the subject does not express a TMPRSS2/ERG gene fusion.

[00107] In another embodiment, any of the computer-implemented methods of the invention may further comprise a step of obtaining by at least one processor information reflecting the expression level of at least 2, 3, 4, 5, 6, or 7 of the following human genes: COL10A1, HOXC4, ESPL1, MMP9, ABCA13, PCDHGA1, and AGSK1 in a biological sample obtained from a patient of African descent.

[00108] In another embodiment of the computer-implemented methods of the invention, the methods may additionally comprise the steps outputting in user readable format the information obtained in the obtaining step.

[00109] In another embodiment of the computer-implemented methods of the invention, the methods may further comprise outputting in user readable format a determination that the subject has prostate cancer or poorly differentiated prostate cancer based on the information conveyed in the outputting step.

**Compositions and Kits**

[00110] The polynucleotide probes and/or primers or antibodies or polypeptide probes that are used in the methods described in this application can be arranged in a composition or a kit. Thus, one embodiment is directed to a composition for diagnosing or prognosing prostate cancer comprising a polynucleotide probe for detecting a genomic rearrangement of the ZBTB20 and LSAMP genes. In certain embodiments, where the genomic rearrangement comprises a fusion between the ZBTB20 and LSAMP genes, such as a fusion between exon 1 (e.g., E1, E1A, E1B, or E1C) of the ZBTB20 gene and exon 3* or exon 4 of the LSAMP gene, the polynucleotide probe hybridizes under high stringency conditions to a junction of a chimeric nucleic acid, wherein the chimeric nucleic acid comprises a first portion from a ZBTB20 gene (e.g., all or part of exon 1, which includes E1, E1A, E1B, or E1C) and a second portion from a LSAMP gene (e.g., all or part of exon 3* or exon 4). All of the polynucleotide probes described herein may be optionally labeled.

[00111] In one embodiment, the composition for diagnosing or prognosing prostate cancer comprises a polynucleotide probe, wherein the polynucleotide probe is designed to detect a chromosomal rearrangement of genomic DNA having a first portion from the ZBTB20 gene (e.g., all or part of exon 1, which includes E1, E1A, E1B, or E1C) and a second portion from the LSAMP gene (e.g., all or part of exon 3* or exon 4). In one embodiment, the polynucleotide probe hybridizes under high stringency conditions to a chromosomal rearrangement comprising the nucleotide sequence of SEQ ID NO:1.

[00112] In another embodiment, the composition for diagnosing or prognosing prostate cancer comprises a polynucleotide probe, wherein the polynucleotide probe is designed to detect a chimeric mRNA or cDNA transcript having a first nucleic acid portion from the ZBTB20 gene and a second nucleic acid portion from the LSAMP gene. In one embodiment, the chimeric mRNA or cDNA transcript comprises a fusion between exon 1 (e.g., E1, E1A, E1B, or E1C) of the ZBTB20 gene and exon 4 of the LSAMP gene. For example, in certain embodiments, the polynucleotide probe hybridizes under high stringency conditions to a chimeric mRNA or cDNA transcript having the nucleotide sequence of SEQ ID NO:4, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, or SEQ ID NO:40. In another embodiment, the chimeric mRNA or cDNA transcript comprises a fusion between exon 1 (e.g., E1, E1A, E1B, or E1C) of the ZBTB20 gene and exon 3* of the LSAMP gene. For example, in certain embodiments, the polynucleotide probe hybridizes under high stringency conditions to a chimeric mRNA or cDNA transcript having the nucleotide sequence of SEQ ID NO:33, SEQ ID NO:39, or SEQ ID NO:41.

[00113] In yet another embodiment, the composition for diagnosing or prognosing prostate cancer comprises a polynucleotide probe, wherein the polynucleotide probe is designed to detect a deletion in chromosome region 3q13, wherein the deletion spans the ZBTB20 and LSAMP genes.

[00114] In other embodiments, the polynucleotide probe hybridizes under high stringency conditions to exon 3* of a LSAMP gene (SEQ ID NO:45) or a junction of a chimeric nucleic acid, wherein the chimeric nucleic acid comprises a first portion from exon 3* of a LSAMP gene and a second portion from exon 4 of a LSAMP gene.

[00115] In yet other embodiments, the polynucleotide probe hybridizes under high stringency conditions to exon 0* of a LSAMP gene (SEQ ID NO:47) or a junction of a chimeric nucleic acid, wherein the chimeric nucleic acid comprises a first portion from exon 0* of a LSAMP gene and a second portion from exon 1 of a LSAMP gene.

[00116] The compositions for diagnosing or prognosing prostate cancer may also comprise primers. In one embodiment, wherein the genomic rearrangement comprises a gene fusion between the ZBTB20 and LSAMP genes, such as a fusion between exon 1 (e.g., E1, E1A, E1B, or E1C) of the ZBTB20 gene and exon 3* or exon 4 of the LSAMP gene, the composition comprises primers for amplifying the ZBTB20/LSAMP gene fusion and, in particular, primers for amplifying a nucleotide sequence from the gene fusion that spans the junction between a first portion of the gene fusion from the ZBTB20 gene (e.g., all or part of exon 1, which includes E1, E1A, E1B, or E1C) and a second portion of the gene fusion from the LSAMP gene (e.g., all or part of exon 3* or exon 4). In this way, the primers can be used to specifically identify a gene fusion between the ZBTB20 and LSAMP genes. In one embodiment, the composition comprises a first polynucleotide primer comprising a sequence that hybridizes under high stringency conditions to a first portion of the gene fusion from the ZBTB20 gene (e.g., all or part of exon 1, which includes E1, E1A, E1B, or E1C); and a second polynucleotide primer comprising a sequence that hybridizes under high stringency conditions to the second portion of the gene fusion from the LSAMP gene (e.g., all or part of exon 3* or exon 4), wherein the first and second polynucleotide primers are capable of amplifying a nucleotide sequence from the gene fusion that spans the junction between the first portion of the gene fusion from the ZBTB20 gene and the second portion of the gene fusion from the LSAMP gene. In certain embodiments, the composition comprises a first and a second polynucleotide primer for amplifying a gene fusion between a ZBTB20 gene and an LSAMP gene, wherein the first and the second polynucleotide primer are capable of amplifying one or more of SEQ ID NOs 32-41 or a portion thereof that spans the junction between a first portion of the gene fusion from the ZBTB20 gene and a second portion of the gene fusion from the LSAMP gene.

[00117] In one embodiment, the first portion of the gene fusion from the LSAMP gene comprises the nucleotide sequence of SEQ ID NO:2 and the second portion of the gene fusion from the ZBTB20 gene comprises the nucleotide sequence of SEQ ID NO:3. In another embodiment, the first portion of the gene fusion from the ZBTB20 gene comprises the nucleotide sequence of SEQ ID NO:5 and the second portion of the gene fusion from the LSAMP gene comprises the nucleotide sequence of SEQ ID NO:6. In yet another embodiment, the first polynucleotide primer comprises the nucleotide sequence of SEQ ID NO:7 and the second polynucleotide primer comprises the nucleotide sequence of SEQ ID NO:8.

[00118] Exon 3* (SEQ ID NO:45) represents a novel exon sequence from the LSAMP locus that has not been previously annotated and has now been found in fusion transcripts resulting from the genomic rearrangement of the ZBTB20 and LSAMP genes. Therefore, amplifying an mRNA/cDNA sequence corresponding to a region of exon 3* that is not present in other genes provides yet another mechanism for detecting the genomic rearrangement of the ZBTB20 and LSAMP genes. In one embodiment,the the composition comprises a first polynucleotide primer comprising a sequence that hybridizes under high stringency conditions to a first portion of exon 3* of LSAMP, and a second polynucleotide primer comprising a sequence that hybridizes under high stringency conditions to a sequence complementary to a second portion of exon 3* of LSAMP, wherein the first and second portions of the exon 3* of LSAMP do not overlap and wherein the polynucleotide primers are capable of amplifying a nucleotide sequence within exon 3* of LSAMP that is unique to the 3* exon and is not found in other gene sequences and, thus, can be used to positively identify the amplified seqeuence as coming from exon 3* of LSAMP. In other embodiments, the second polynucleotide primer hybridizes under high stringency conditions to a sequence complementary to a region within exon 4 of LSAMP, such that the primer pair yields an amplification product that spans the junction between exon 3* and exon 4 of LSAMP. The second primer can also hybridize to a sequence complementary to a region with exon 5, 6, or 7 of LSAMP.

[00119] Exon 0* (SEQ ID NO:47) represents a novel exon sequence the LSAMP locus that has not been previously annotated and was identified in an alternatively spliced transcript in a patient having a genomic rearrangement of the ZBTB20 and LSAMP genes. Therefore, amplifying an mRNA/cDNA sequence corresponding to a region of exon 0* that is not present in other genes provides yet another mechanism for detecting the genomic rearrangement of the ZBTB20 and LSAMP genes. In one embodiment,the the composition comprises a first polynucleotide primer comprising a sequence that hybridizes under high stringency conditions to a first portion of exon 0* of LSAMP, and a second polynucleotide primer comprising a sequence that hybridizes under high stringency conditions to a sequence complementary to a second portion of exon 0* of LSAMP, wherein the first and second portions of exon 0* of LSAMP do not overlap and wherein the polynucleotide primers are capable of amplifying a nucleotide sequence within exon 0* of LSAMP that is unique to the 0* exon and is not found in other gene sequences and, thus, can be used to positively identify the amplified seqeuence as coming from exon 0* of LSAMP. In other embodiments, the

second polynucleotide primer to hybridizes under high stringency conditions to a sequence complementary to a region within exon 1 of LSAMP, such that the primer pair yields an amplification product that spans the junction between exon 0* and exon 1 of LSAMP. The second primer can also hybridize to a sequence complementary to a region with exon 2, 3, or 4 of LSAMP.

[00120] In another embodiment, wherein the genomic rearrangement comprises a deletion at chromosome region 3q13, wherein the deletion spans both the ZBTB20 and LSAMP genes, the composition for diagnosing or prognosing prostate cancer comprises primers for amplifying a chimeric junction created by the deletion. Thus, in one embodiment, the composition comprises a first polynucleotide primer comprising a sequence that hybridizes under high stringency conditions to a first nucleic acid that borders a 5' end of the deletion; and a second polynucleotide primer comprising a sequence that hybridizes under high stringency conditions to the second nucleic acid that borders a 3' end of the deletion, wherein the first and second polynucleotide primers are capable of amplifying a nucleotide sequence that spans the chimeric junction created by the deletion, wherein the deletion occurs at chromosome region 3q13 and spans the ZBTB20 and LSAMP genes.

[00121] Another aspect is directed to a double stranded oligonucleotide duplex, wherein the oligonucleotide duplex comprises a first nucleic acid hybridized to a second nucleic acid, wherein the first nucleic acid comprises a first portion from a ZBTB20 gene fused to a second portion from a LSAMP gene and wherein the second nucleic acid is a polynucleotide probe that is hybridized to a junction between the first portion from the ZBTB20 gene and the second portion from the LSAMP gene. The polynucleotide probe is optionally labeled.

[00122] In another embodiment, the oligonucleotide duplex comprises a first nucleic acid hybridized to a second nucleic acid, wherein the first nucleic acid comprises a first portion from exon 3* of a LSAMP gene fused to a second portion from exon 4 of a LSAMP gene and wherein the second nucleic acid is a polynucleotide probe that is hybridized to a junction between the first portion from exon 3* of the LSAMP gene and the second portion from exon 4 of the LSAMP gene. The polynucleotide probe is optionally labeled.

[00123] In yet another embodiment, the oligonucleotide duplex comprises a first nucleic acid hybridized to a second nucleic acid, wherein the first nucleic acid comprises a first portion from exon 0* of a LSAMP gene fused to a second portion from exon 1 of a LSAMP gene and wherein the second nucleic acid is a polynucleotide probe that is

hybridized to a junction between the first portion from exon 0* of the LSAMP gene and the second portion from exon 4 of the LSAMP gene. The polynucleotide probe is optionally labeled.

[00124] Another aspect is directed to kits for diagnosing or prognosing prostate cancer. In one embodiment, the kit for diagnosing or prognosing prostate cancer comprises a first composition comprising one or more polynucleotide probes and/or primers for detecting a ZBTB20/LSAMP genomic rearrangement, as discussed above, and a second composition comprising a polynucleotide probe that hybridizes under high stringency conditions to a gene selected from COL10A1, HOXC4, ESPL1, MMP9, ABCA13, PCDHGA1, and AGSK1.

[00125] In another embodiment, the kit for diagnosing or prognosing prostate cancer comprises a first composition comprising one or more polynucleotide probes and/or primers for detecting a ZBTB20/LSAMP genomic rearrangement, as discussed above, and a second composition comprising a polynucleotide probe that hybridizes under high stringency conditions to a gene selected from ERG, AMACR, PCA3, and PSA.

[00126] The kit for diagnosing or prognosing prostate cancer may also comprise antibodies. Thus, in one embodiment, the kit for diagnosing or prognosing prostate cancer comprises an antibody that binds to a polypeptide encoded by a ZBTB20/LSAMP gene fusion. The antibody may be optionally labeled. In another embodiment, the kit further comprises one or more antibodies for detecting at least 1, 2, 3, 4, 5, 6, or 7 of the polypeptides encoded by following human genes: COL10A1, HOXC4, ESPL1, MMP9, ABCA13, PCDHGA1, and AGSK1. In another embodiment, the kit further comprises one or more antibodies for detecting ERG, AMACR, or PSA.

[00127] In one embodiment, a kit for diagnosing or prognosing prostate cancer includes instructional materials disclosing methods of use of the kit contents in a disclosed method. The instructional materials may be provided in any number of forms, including, but not limited to, written form (e.g., hardcopy paper, etc.), in an electronic form (e.g., computer diskette or compact disk) or may be visual (e.g., video files). The kits may also include additional components to facilitate the particular application for which the kit is designed. Thus, for example, the kits may additionally include other reagents routinely used for the practice of a particular method, including, but not limited to buffers, enzymes (e.g., polymerase), labeling compounds, and the like. Such kits and appropriate contents are well known to those of skill in the art. The kit can also include a reference or control sample. The reference or control sample can be a biological sample or a data base.

[00128] The polynucleotide probes and antibodies described in this application are optionally labeled with a detectable label. Any detectable label used in conjunction with probe or antibody technology, as known by one of ordinary skill in the art, can be used. In a particular embodiment, the probe is labeled with a detectable label selected from the group consisting of: a fluorescent label, a chemiluminescent label, a quencher, a radioactive label, biotin, mass tags and/or gold.

**Antibodies that Bind to a Chimeric ZBTB20/LSAMP Fusion Protein**

[00129] This disclosure provides antibodies that bind to the protein encoded by the ZBTB20/LSAMP gene fusion. Antibodies, also known as immunoglobulins, are typically tetrameric glycosylated proteins composed of two light (L) chains of approximately 25 kDa each and two heavy (H) chains of approximately 50 kDa each. Two types of light chain, termed lambda and kappa, may be found in antibodies. Depending on the amino acid sequence of the constant domain of heavy chains, immunoglobulins can be assigned to five major classes: A, D, E, G, and M, and several of these may be further divided into subclasses (isotypes), e.g., $IgG_1$, $IgG_2$, $IgG_3$, $IgG_4$, $IgA_1$, and $IgA_2$. Each light chain includes an N-terminal variable (V) domain (VL) and a constant (C) domain (CL). Each heavy chain includes an N-terminal V domain (VH), three or four C domains (CHs), and a hinge region. The CH domain most proximal to VH is designated as CH1. The VH and VL domains consist of four regions of relatively conserved sequences called framework regions (FR1, FR2, FR3, and FR4), which form a scaffold for three regions of hypervariable sequences (complementarity determining regions, CDRs). The CDRs contain most of the residues responsible for specific interactions of the antibody with the antigen. CDRs are referred to as CDR1, CDR2, and CDR3. Accordingly, CDR constituents on the heavy chain are referred to as H1, H2, and H3, while CDR constituents on the light chain are referred to as L1, L2, and L3. Identification and numbering of framework and CDR residues is as described by Chothia *et al.*, Structural determinants in the sequences of immunoglobulin variable domain, J Mol Biol 1998, 278:457-79, which is hereby incorporated by reference in its entirety.

[00130] In one embodiment, the protein encoded by the ZBTB20/LSAMP gene fusion is a truncated LSAMP protein. In certain embodiments, the antibody that binds to the truncated LSAMP protein binds to an epitope present in the truncated LSAMP protein but not present in the wild type LSAMP protein. In one embodiment, the antibody binds to a chimeric polypeptide encoded by a gene fusion between exon 1 of the ZBTB20 gene and

exon 4 of the LSAMP gene. In certain embodiments, the antibody that binds to a chimeric polypeptide encoded by a gene fusion between exon 1 of the ZBTB20 gene and exon 4 of the LSAMP gene, binds to an epitope present in the polypeptide encoded by the gene fusion that is not present in either the wild type ZBTB20 or wild type LSAMP protein. In another embodiment, the gene fusion comprises the nucleotide sequence of SEQ ID NO:4.

[00131] Methods of making antibodies, or antigen-binding fragments thereof, and formulating the same for therapeutic administration are well known as discussed, for example, in PCT/US2010/032714, which is hereby incorporated by reference in its entirety.

[00132] The antibodies described herein that bind to a polypeptide encoded by a ZBTB20/LSAMP gene fusion can be used in a variety of research and medical applications. In one aspect, the disclosure provides a method of treating prostate cancer in a subject, comprising administering to said subject a therapeutically effective amount of an antibody that binds to a polypeptide encoded by a ZBTB20/LSAMP gene fusion formulated in a pharmaceutically acceptable vehicle.

[00133] The disclosure also provides compositions comprising an antibody that binds to a polypeptide encoded by a ZBTB20/LSAMP gene fusion. In certain embodiments, the compositions are suitable for pharmaceutical use and administration to patients. These compositions comprise an antibody that binds to a polypeptide encoded by a ZBTB20/LSAMP gene fusion and a pharmaceutically acceptable excipient. The compositions may also contain other active compounds providing supplemental, additional, or enhanced therapeutic functions. The pharmaceutical compositions may also be included in a container, pack, or dispenser together with instructions for administration. In one embodiment, the composition comprises a monoclonal antibody that binds to a polypeptide encoded by a ZBTB20/LSAMP gene fusion for use in treating prostate cancer.

[00134] A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Methods to accomplish the administration are known to those of ordinary skill in the art. This includes, for example, injections, by parenteral routes such as intravenous, intravascular, intraarterial, subcutaneous, intramuscular, intratumor, intraperitoneal, intraventricular, intraepidural, or others as well as oral, nasal, ophthalmic, rectal, or topical. Sustained release administration is also specifically contemplated, by such means as depot injections or erodible implants. Localized delivery is particularly contemplated, by such means as delivery via a catheter to one or more arteries, such as the renal artery or a vessel supplying a localized tumor.

[00135] Toxicity and therapeutic efficacy of the composition can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., determining the $LD_{50}$ (the dose lethal to 50% of the population) and the $ED_{50}$ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio $LD_{50}/ED_{50}$. Antibodies that exhibit large therapeutic indices may be less toxic and/or more therapeutically effective.

[00136] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

## EXAMPLES

[00137] **Example 1. Comparative Genomic DNA Analysis**

[00138] A comparative full genome analysis was conducted using primary prostate tumors and corresponding normal tissue (blood) in a cohort of seven AA and seven CA CaP patients (28 specimens). The cohort was selected based on the following criteria: primary treatment radical prostatectomy, no neo-adjuvant treatment, Gleason 6-7 (aka, primary pattern 3 or Gleason grade 3 +3 and 3+4, which represents the majority of PSA-screened CaP at diagnosis/primary treatment), frozen tumor tissue with 80% or more tumor cell content, dissected tumor tissue yielding over 2 μg high molecular weight genomic DNA, availability of corresponding blood genomic DNA and patient clinico-pathological data.

[00139] 28 samples were sent to Illumina Inc. (UK) for sequencing. Sequences from tumor samples were mapped to the reference genome using Illumina's ELAND alignment algorithm. Sequencing reported good coverage (average 37). Variant calling for single nucleotide polymorphisms (SNPs), small insertions and deletions (InDels), copy number variants (CNVs), and structural variants (SVs) was performed concurrently using the Strelka algorithm. All established CaP mutations (TMPRSS2/ERG, SPOP, CHD1, and PTEN) were identified at expected frequencies in this cohort. A genome sequence coverage summary of the 14 patients is presented in Table 2.

[00140] Table 2

| Sample ID | GP-10 | GP-04 | GP-02 | GP18 | GP-12 | GP-13 | GP-15 | GP-06 | GP-11 | GP-16 | GP-07 | GP-01 | GP-09 | GP-17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ethnicity | AA | AA | AA | AA | AA | AA | AA | CA | CA | CA | CA | CA | CA | CA |
| TMPRSS2-ERG Status | - | - | - | - | + | + | + | - | - | - | + | + | + | + |
| Tumor: estimated purity (%) | 90 | 80 | 80 | 80 | 90 | 85 | 80 | 95 | 90 | 90 | 90 | 80 | 80 | 80 |
| Tumor Differentiation (%) | WD 95 PD 5 | WD 95 PD 5 | WD 100 | WD 100 | WD 100 | WD 92 PD 8 | WD 100 | WD 95 PD 5 | WD 100 | WD 85 PD 15 | WD 100 | WD 95 PD 5 | WD 97 PD 3 | WD 95 PD 5 |
| Tumor Gleason | 7(3+4) | 7(3+4) | 6(3+3) | 6(3+3) | 6(3+3) | 7(3+4) | 6(3+3) | 7(3+4) | 6(3+3) | 7(3+4) | 6(3+3) | 7(3+4) | 7(3+4) | 7(3+4) |
| Tumor: bases sequenced and aligned (in Gb) | 118.5 | 116.7 | 114.7 | 116.2 | 112.6 | 107.6 | 114.2 | 123.5 | 107.5 | 108.6 | 106.5 | 117.5 | 111.9 | 111.3 |
| Tumor: haploid coverage | 39.2 | 38.3 | 37.8 | 38.6 | 37.2 | 34.9 | 38.1 | 41 | 35 | 36 | 34.9 | 39 | 37.2 | 36.6 |
| Coverage-% Positions>=1x | 98 | 98 | 98 | 98 | 98 | 98 | 98 | 98.1 | 98 | 98 | 98 | 98 | 98 | 98 |
| Normal: bases sequenced and aligned (in Gb) | 111.6 | 111.8 | 112.2 | 102.4 | 117.7 | 113.4 | 108.1 | 113.3 | 109.8 | 115.3 | 121.9 | 104 | 105.3 | 112.2 |
| Normal: haploid coverage | 37.1 | 37.5 | 37.3 | 34.1 | 39.2 | 37.6 | 36 | 37.5 | 36.4 | 38.5 | 40.3 | 34.4 | 35.2 | 37.4 |
| Coverage-% Positions>=1x | 98 | 98 | 98 | 98 | 98 | 98 | 98.1 | 98 | 98 | 98 | 98 | 98 | 98 | 98 |
| Mutation rate/Mb | 0.9 | 0.98 | 0.82 | 0.69 | 0.72 | 0.68 | 0.76 | 1.27 | 0.75 | 1.08 | 0.74 | 2.33 | 1.02 | 1.13 |
| Non-silent point mutations (>25% read coverage) | 23 | 25 | 25 | 20 | 8 | 13 | 12 | 38 | 19 | 26 | 16 | 38 | 15 | 35 |

[00141] Thirty one genes (including known mutations) with SNP, CNV or InDel somatic mutations in at least two of 14 patients were identified: AC091435.2; APC; ASMTL; ASMTL-AS1; CDC73; CHD1; CSF2RA; EYS; FRG1; FRG1B; HK2; IL3RA; KLLN; LIPF; LOC100293744; MT-ATP6; MT-BD4; MT-CO1; MT-CYB; MT-ND2; MT-ND3; MUC16; MUC6; NOX3; PDHA2; PTEN; SLC25A6; SLC9B1; SPOP; TRAV20; and USH2A. The top SVs and CNVs (highest confidence) present in at least 2 of 14 patients are set forth in Table 3:

[00142] Table 3

| **Gene** | **Structural and Copy Number Variants** | **Ethnicity** | **Score** |
|---|---|---|---|
| ZBTB20-LSAMP | Structural and Copy Number Variation | AD | 17 |
| TMPRSS2-ERG* | Structural Variation (Gene Fusion) | CD>AD | 27 |
| HLA-DRB5 | Structural Variation | AD>CD | 42 |
| MLL3-BAGE | Structural Variation | No | 30 |
| HLA-B | Structural Variation | AD>CD | 19 |
| FOXP1 | Structural and Copy Number Variation | No | 19 |

| Gene | Structural and Copy Number Variants | Ethnicity | Score |
|---|---|---|---|
| CHD1* | Copy Number Variation | No | 9 |
| TRAV20 | Structural and Copy Number Variation | AD>CD | 8 |
| PTEN* | Copy Number Variant | CD>AD | 5 |
| PRDM2-VS13D | Structural and Copy Number Variation | No | 5 |
| RPL11-SLC30A2 | Structural Variation | AD | 5 |
| SLC45A3 | Structural Variation | No | 5 |
| PCDH10 | Structural and Copy Number Variation | AD | 4 |
| * Known gene alteration in CaP<br>African descent = AD; Caucasian descent = CD | | | |

[00143] In Table 3, the "Score" is the sum of individual scores of maximum 5 for SV or 4 for CNV for each patient (theoretical maximum score for the 14 patients is 14 x 5 = 70 for SV, or 14 x 4 = 56 for CNV). A score of 1, 4 or 5 is given for SV (where 5 means SV is supported with RNA data, 4 means SV with splice reads, and 1 if no splice reads or RNA data is available), and a score of 2 or 4 is given for CNV (2 if it is predicted by one algorithm, and 4 if it is predicted by more than one algorithm).

[00144] Certain SNV, CNV or InDel somatic mutations exhibited a preferred association with a specific ethnic group (AD (African Descent), CD (Caucasian Descent), AD>CD, or CD>AD). In particular, the absence of PTEN deletions in AA patients was unexpected. PTEN deletion was detected in 4 of 7 CD, indicating the potential exclusivity of PTEN deletions in CD cases.

**[00145] Example 2. Stratification of ZBTB20/LSAMP and PTEN Genetic Alterations Based on Ethnicity**

[00146] Established prostate cancer genomic defects (TMPRSS2/ERG, PTEN) as well as a novel recurrent rearrangement affecting the ZBTB20/LSAMP loci on chromosome 3q13 were identified in this cohort of 14 patients. All three (ZBTB20/LSAMP (GP10), TMPRSS2/ERG (GP1-14), and PTEN (GP1-14)) were validated by RT-PCR or FISH analyses.

[00147] Table 4

| Sample ID | Ethnicity | TMPRSS2 ERG | ZBTB20 LSAMP | PTEN Deletion |
|---|---|---|---|---|
| GP-02 | AA | - | + | - |
| GP-04 | AA | - | + | - |
| GP-10 | AA | - | + | - |
| GP-18 | AA | - | - | - |
| GP-12 | AA | + | - | - |
| GP-13 | AA | + | - | - |
| GP-15 | AA | + | - | - |
| GP-06 | CA | - | - | - |
| GP-11 | CA | - | - | + |
| GP-16 | CA | - | - | + |
| GP-01 | CA | + | - | + |
| GP-07 | CA | + | - | - |
| GP-09 | CA | + | - | - |
| GP-17 | CA | + | - | + |

+ indicates presence; - indicates absence

[00148] The PTEN deletion was not observed in any of the AD patients (0 of 7). It was detected only in CD patients (4 of 7), suggesting that PTEN loss is an infrequent event in CaP development of men of AD. Activation of ERG (via gene fusion with androgen regulated genes, such as TMPRSS2), which has been associated with PTEN deletion, are also less frequent in CaP patients of AD [20]. The observed association of PTEN deletion with CD ethnicity was apparent and similar only to the ERG rearrangement when compared to other structural (SV) or copy number (CNV) variations (see Table 3 above).

[00149] PTEN deletions identified by whole genome sequencing were independently confirmed by FISH assay in identical tumor foci in consecutive sections of whole mount prostates from the 7 AD and 7 CD. Figure 4A.

[00150] To validate these findings PTEN deletions were assessed by FISH assay in a tissue microarray (TMA) of 41 AD and 58 CD cases representing Gleason 6, Gleason 7 or Gleason 8-10 tumors. Multiple samples including different tumor foci from each case were represented in the TMA. Examining all cores in the TMA, a significantly lower overall

frequency of PTEN deletion was observed in AD (19.5%) when compared to CD (62.1) cases (Table 5A). Consistent with the indicated low frequency of PTEN deletion from the whole-genome sequences of Gleason 6-7 tumors of AD CaPs, PTEN deletions were found in only 1 out of 15 (6.7%) Gleason sum 6 and only 4 out of 15 (27%) Gleason sum 7 AD cases (Table 5B). This is in sharp contrast to PTEN deletion frequencies found in 10 out of 19 (52.6%) Gleason sum 6 and 14 out of 21 (66.7%) Gleason sum 7 CD CaPs. Taken together, these results validate the finding that PTEN loss is an infrequent event in CaPs of AD. Ethnic disparities noted in CaP include higher genomic frequencies of PTEN deletion in CD when compared to Asian men [34]. Recent reports have shown lower frequency of PTEN deletions in AD CaP patients in comparison to CD CaP patients [35, 36].

**[00151] Table 5A**

| Race | PTEN deletion | | P value |
|---|---|---|---|
| | No | Yes | |
| AD, N=41 | 33 (80.5) | 8 (**19.5**) | |
| CD, N=58 | 22 (37.9) | 36 (**62.1**) | <.0001 |

**[00152] Table 5B**

| Worst Gleason Sum | Overall (N=91) | | P value | AD (N=39) | | P value | CD (N=45) | | P value |
|---|---|---|---|---|---|---|---|---|---|
| | PTEN no deletion | PTEN deletion | | PTEN no deletion | PTEN deletion | | PTEN no deletion | PTEN deletion | |
| 6 or less | 23 (67.6) | 11 (32.4) | | **14 (93.3)** | **1 (6.7%)** | | **9 (47.4)** | **10 (52.6)** | |
| 7 | 18 (50.0) | 18 (50.0) | | 11 (73.3) | 4 (26.7) | | 7 (33.3) | 14 (66.7) | |
| 8 to 10 | 11 (50.0) | 11 (50.0) | 0.2571 | 7 (70.0) | 3 (30.0) | 0.3185 | 4 (33.3) | 8 (66.7) | 0.6055 |

[00153] In patients of AD with ERG negative CaP (i.e., no ERG/TMPRSS2 fusion), genomic rearrangements of the ZBTB20 - LSAMP chromosomal loci (Chr. 3q13) was noted in 3 of 4 patients. Detailed analysis of the affected loci and RNA-Seq data revealed three

different genomic rearrangements of the 3q13 loci.

[00154] Detailed analysis of the affected loci and RNA-Seq data indicated that in one instance, the genomic rearrangement of the 3q13 loci involved a tandem duplication giving rise to a fusion between exon 1 of ZBTB20 and exon 4 of LSAMP. Figures 1 and 2. More specifically, the genomic rearrangement gave rise to a fusion between

[00155] (a) a portion of the LSAMP gene having the sequence:

AGAGTCTCTTCTTTGGGTCTCTTCCACATAGCTTGTTTGTAATCTCCAAGAAAGAC
TTCACATTACAGGCTGAAAAGAATCACCTACGGTTTCCATATTTTGAAAGAAATT
TTTAAAAACCATGAAAACAAACAAACAAAAATCCTAGTTTCCTTTATAAAATAGC
AAAGGAAAGTTCTCTCTCCTGTCACCAGGAATATGATTATGATCAGTTGGTTATT
TAGGTCACATGTGAAAGAAATGAAAGAGGAGGCATGGGAATGTAAGGGAGAAT
AGTAGTCTGCCCTCAAGTCTGCAAACG (SEQ ID NO:2); and

[00156] (b) a portion of the ZBTB20 gene having the sequence:

CTCGAGAACAGTGAGCAATAAATTTTTCTTTATACATTACCCAGTCTGTGGTATTC
TGTTATGGCAACACAAAATAAACTAAGACAGTATTATGTATTTTTTCTTTTGTTTT
ACATTTTACTAAGTGCCGACTTATTCGAAAAGGTAATTAGCTTTGGTTAATTATC
AAAGTTTTGTCTTTCCTTTCCTACTTTTGTCCCACTAAGCAAAAAACAAAACAATG
AGCATTGACCTTTACCTTTCTCTGGTAAGGGAGTATGGAAGGTTTTCTACTACTTT
GTAAAAATACTGCTACAGATGG (SEQ ID NO:3).

[00157] When combined, the genomic rearrangement of the ZBTB20 and LSAMP genes, in this one instance, had the following sequence:

CTCGAGAACAGTGAGCAATAAATTTTTCTTTATACATTACCCAGTCTGTGGTATTC
TGTTATGGCAACACAAAATAAACTAAGACAGTATTATGTATTTTTTCTTTTGTTTT
ACATTTTACTAAGTGCCGACTTATTCGAAAAGGTAATTAGCTTTGGTTAATTATC
AAAGTTTTGTCTTTCCTTTCCTACTTTTGTCCCACTAAGCAAAAAACAAAACAATG
AGCATTGACCTTTACCTTTCTCTGGTAAGGGAGTATGGAAGGTTTTCTACTACTTT
GTAAAAATACTGCTACAGATGGAGAGTCTCTTCTTTGGGTCTCTTCCACATAGCT
TGTTTGTAATCTCCAAGAAAGACTTCACATTACAGGCTGAAAAGAATCACCTACG
GTTTCCATATTTTGAAAGAAATTTTTAAAAACCATGAAAACAAACAAACAAAAA
TCCTAGTTTCCTTTATAAAATAGCAAAGGAAAGTTCTCTCTCCTGTCACCAGGAA
TATGATTATGATCAGTTGGTTATTTAGGTCACATGTGAAAGAAATGAAAGAGGA
GGCATGGGAATGTAAGGGAGAATAGTAGTCTGCCCTCAAGTCTGCAAACG (SEQ

ID NO:1)

[00158] A cDNA sequence of the ZBTB20/LSAMP gene fusion resulting from this genomic rearrangement was identified as:

[00159] CACAACATCAAGAGCAGGAAAATGGACTCATTAGGGAGGCAGGCA GTCATTACCACTCACACTGTACTTCCAGGGAGACACCGATTATAAGAAGAGAAA CTCAGCGCTGGGGAAGAAGGAAGGGAATTTGAAGGAGAAGAAGAATATCTGGA GATCCTTGGCATCACCAGGGAGCAGTCAGGCAAATATGAGTGCAAAGCTGCCAA CGAGGTCTCCTCGGCGGATGTCAAACAAGTCAAGGTCACTGTGAACTATCCTCCC ACTATCACAGAATCCAAGAGCAATGAAGCCACCACAGGACGACAAGCTTCACTC AAATGTGAGGCCTCGGCAGTGCCTGCACCTGACTTTGAGTGGTACCGGGATGAC ACTAGGATAAATAGTGCCAATGGCCTTGAGATTAAGAGCACGGAGGGCCAGTCT TCCCTGACGGTGACCAACGTCACTGAGGAGCACTACGGCAACTACACCTGTGTG GCTGCCAACAAGCTGGGGGTCACCAATGCCAGCCTAGTCCTTTTCAGACCTGGGT CGGTGAGAGGAATAAATGGATCCATCAGTCTGGCCGTACCACTGTGGCTGCTGG CAGCATCTCTGCTCTGCCTTCTCAGCAAATGTTAA (SEQ ID NO:4).

[00160] In this ZBTB20/LSAMP gene fusion, the following sequence was derived from exon 1 of ZBTB20:
CACAACATCAAGAGCAGGAAAATGGACTCATTAGGGAGGCAGGCAGTCATTACC ACTCACACTGTACTTCCAGGGAGACACCGATTATAAGAAGAGAAACTCAGCGCT GGGGAAGAAG (SEQ ID NO:5), with the remainder of the sequence (SEQ ID NO:6) derived from exons 4-7 of LSAMP.

[00161] One of ordinary skill in the art can design primers for amplifying this ZBTB20/LSAMP gene fusion. In one embodiment, the forward primer comprises the nucleotide sequence of SEQ ID NO:5 or a sequence complementary thereto and the reverse primer comprises the nucleotide sequence of SEQ ID NO:6 or a sequence complementary thereto. The following exemplary primers were designed to amplify this particular ZBTB20/LSAMP gene fusion:

[00162] ZBTB20/LSAMP Forward Primer:

[00163] GCAGGCAGTCATTACCACTC (SEQ ID NO:7)

[00164] ZBTB20/LSAMP Reverse Primer:

[00165] TGACTTGTTTGACATCCGCC (SEQ ID NO:8)

[00166] In another instance, the genomic rearrangement of the 3q13 loci involved a deletion, giving rise to a fusion between the ZBTB20 gene and the LSAMP gene. Figure 1.

In the third instance, the genomic rearrangement of the 3q13 loci involved a large deletion of at least 22.7 Mb, spanning both the ZBTB20 and LSAMP genes. Figure 1. Strikingly, two of the 3 patients with the chromosome 3q13 rearrangement developed metastasis (the only two metastasis in this cohort), and the third had biochemical recurrence of CaP (1 of 3 in the cohort), indicating that the ZBTB20/LSAMP genomic rearrangement is associated with an aggressive form of prostate cancer or an increased likelihood to develop an aggressive form of prostate cancer. Figure 1.

[00167] To validate the association of the ZBTB20/LSAMP fusion transcript with AA ethnicity, negative ERG status, and adverse disease outcome, 20 additional CaP tumors and 24 CaP tumor-derived cell lines of AA and CA descent were evaluated using PCR and the primers described above. Of the 20 AA tumor samples analyzed, one contained the same ZBTB20/LSAMP fusion transcript, which was present in GP10 specimen. This patient had an ERG negative, aggressive form of prostate cancer (poorly differentiated with progression to metastasis). The ZBTB20/LSAMP genomic rearrangement was not detected in mRNA or genomic DNA of 34 normal samples (20 normal prostate tissue specimens and constitutional DNA from 7 CA and 7 AA patients). A ZBTB20/LSAMP rearrangement was detected in one CaP cell line (CPDR RC92), which was derived from a patient of AD with poorly differentiated prostate cancer.

[00168] ZBTB20-LSAMP deletions, identified in the tumor genome of two AD cases, were validated by FISH assay by probing the genomic region of chromosome 3 (from the ZBTB20 promoter upstream sequences through the GAP43 gene to the LSAMP locus 3' adjacent region) Figure 4B. One of the AD cases harbored ZBTB20-LSAMP duplication rearrangement predicting an inactivating gene fusion between the promoter and first exon of ZBTB20 and LSAMP coding sequences. This gene fusion was confirmed by 5'-RACE method defining complete cDNA sequences of the 5' fusion partner, first exon of ZBTB20, the fusion junction, and the 3' fusion partner, LSAMP exon 4. The fusion transcript eliminates the natural translation initiation (ATG) of the LSAMP gene leading to the premature truncation of LSAMP protein.

[00169] To validate the findings a tissue microarray (TMA) was assessed by FISH assay for detecting the absence or presence of a DNA region between the ZBTB20 and LSAMP gene loci in 23 AD and 7 CD cases representing Gleason 6, Gleason 7 or Gleason 8-10 tumors. A deletion was detected in 26% (6 out of 23) of AD cases, whereas, the deletion was observed in only 1 out of 7 (14%) in CD cases (Table 6A). Among the seven cases with

deletion, six had biochemical recurrence (Table 6B). ERG expression rearrangement was found only in two of the seven patients with the deletion in the ZBTB20-LSAMP region. PTEN deletion was found in only one case out of seven (Table 6B). Six of the seven patients having the deletion in the ZBTB20-LSAMP region experienced biochemical recurrence (BCR) or metastisis (Met) (Table 6B). These data support our original discovery showing an association of a significantly higher proportion of ZBTB20-LSAMP region deletions in AD CaP patients with poor prognosis (biochemical recurrence after prostatectomy or metastis).

[00170] Table 6A

| | AD | CD |
|---|---|---|
| Deletion | 6 (26.09%) | 1 (14.29%) |
| No deletion | 17 (73.91%) | 6 (85.71%) |
| Total | 23 | 7 |

[00171] Table 6B

| Ethnicity | ERG | PTEN | BCR/Met |
|---|---|---|---|
| AD | Negative | wt | Yes |
| AD | Negative | wt | No |
| AD | Negative | wt | Yes |
| AD | Positive | wt | Yes/Yes |
| AD | Negative | wt | Yes |
| AD | Positive | wt | Yes |
| CD | Negative | deletion | Yes |

[00172] Evaluation of the affected 3q13 loci and RNA-Seq data indicated that the genomic rearrangement in one AD ERG negative case involved a tandem duplication giving rise to 5'-ZBTB20-LSAMP-3' fusion transcripts. The 5'-ZBTB20-LSAMP-3' fusion transcripts from this AD ERG negative case were also assessed by the "Rapid Amplification of cDNA Ends" (RACE) method. The sequence of RACE products were cloned in an M13 sequencing vector and confirmed by forward and reverse DNA sequencing for six clones for each RACE cDNA products. Ten prototypical CaP-associated fusion cDNA were identified,

ZBTB20-LSAMP Fusion Type 1-to 10, respectively and one alternatively spliced cDNA of the LSAMP locus (LSAMP prostate cancer alternatively spliced form Type 1: LPCS1). Figure 5. The sequences of these transcripts are set forth below with bold text indicating the portion from the ZBTB20 gene, and underlining indicating a sequence that was used as a reverse primer in the RACE method. A universal 5' RACE sequence was used as the forward primer in these amplification methods. The exons of the fusion transcripts and LPCS1 alternate between normal text and italicized text.

[00173] Type 1: **ZBTB20-E1**-LSAMP-*E4*-E5-*E6*-E7

**AGAGTACATGCGGCGGGGGGAAGTTTAGGAGTTGAGGAAAGAAGATTAAAGAGCGCGAGGAG**
*GAAGGGAATTTGAAGGAGAAGAAGAATATCTGGAGATCCTTGGCATCACCAGGGAGCAGTCA*
*GGCAAATATGAGTGCAAAGCTGCCAACGAGGTCTCCTCGGCGGATGTCAAACAAGTCAAGGT*
*CACTGTGAACT*ATCCTCCCACTATCACAGAATCCAAGAGCAATGAAGCCACCACAGGACGAC
AAGCTTCACTCAAATGTGAGGCCTCGGCAGTGCCTGCACCTGACTTTGAGTGGTACCGGGAT
GACACTAG*GATAAATAGTGCCAATGGCCTTGAGATTAAGAGCACGGAGGGCCAGTCTTCCCT*
*GACGGTGACCAACGTCACTGAGGAGCACTACGGCAACTACACCTGTGTGGCTGCCAACAAGC*
*TGGGGGTCACCAATGCCAGCCTAGTCCTTTTC*AGACCTGGGTCGGTGAGAGGAATAAATGGA
TCCATCAGTCTGGCCGTACCACTGTGGCTGCTGGCAGCATCTCTGCTCTGCCTTCTCAGCAA
ATGTTAATAGAATAAAAATTTAAAAATAATTTAAAAAACACACAAAAATGTGTCACACAGAA
TACAGAGAGAGAGAGACAGAGAGAGAGAGAGAGAGAGATGGGGGAGACCGTTTATTTCAC
AACTTTGTGTGTTTATACATGAAGGGGGAAATAAGAAAGTGAAGAAGAAAATNACAACATTT
AAAACAATTTTACAGTCCATCATTAAAAATTTATGTATCATTCAGGATGGAGAAGGTTCTAC
TGGGATATGTTTATATCTACTAAGCAAATGTATGCTGTGTAAAGACTACACCACACTAAGGA
CATCTGGATGCTGTAAAAATAAGAGAAGAACCAGATGGATATTAAGCCCCCCAACACACACT
TTATCCTTCCTTCCTTCATCTTTTTTCATCTGTGGGGAAGAAAATAAGGTCTTGCCTTTGGT
GTTTATATTTCCATAACCTTTTAATTCTATTTTTCATTTGAGCTGACTTGTAGCCACTTCAG
ACTATCAATGGAATCTTATGTTGAGCCTTTCTCTGGCTTTCCTTCCTCCACTATCTCTCCAA
CTTTAGAGATCATCCCCTCTCCCTCCAGTGCGTTCTATCTCCCCCACACCCACCCAA (SEQ
ID NO:32)

[00174] Type 2: **ZBTB20-E1C** LSAMP *E3A*-E4-*E5*-E6-*E7*

**ACATGGGGAGGTTGCAGTGTGTGTATATACACAACATCAAGAGCAGGAAAATGGACTCATTA**
**GGGAGGCAGGCAGTCATTACCACTCACACTGTACTTCCAGGGAGACACCGATTATAAGAAGA**
**GAAACTCAGCGCTGGGGAAGA***AGATTAACTTACTCTTAATGATCTTCCAACACTTGAGAAGG*
*TCAGTAGCCCTCCATCTGTCATTCTCCAAGTTCACCAACAGCTTATCCACCCATCAAAGGTG*

*CTTTTGTAACAAAATCCATGCATAATGAAACCAAGAAAG*GAAGGGAATTTGAAGGAGAAGAA
GAATATCTGGAGATCCTTGGCATCACCAGGGAGCAGTCAGGCAAATATGAGTGCAAAGCTGC
CAACGAGGTCTCCTCGGCGGATGTCAAACAAGTCAAGGTCACTGTGAAC*TATCCTCCCACTA*
*TCACAGAATCCAAGAGCAATGAAGCCACCACAGGACGACAAGCTTCACTCAAATGTGAGGCC*
*TCGGCAGTGCCTGCACCTGACTTTGAGTGGTACCGGGATGACACTAG*GATAAATAGTGCCAA
TGGCCTTGAGATTAAGAGCACGGAGGGCCAGTCTTCCCTGACGTGACCAACGTCACTGAGAG
GNGAGCACTACGGCAACTACACCTGTGTGGCTGCCAACAAGCTGGGGGTCACAATGCCAGCC
TAGTCCTTTTCAG*ACNTGGKYSGTGAGAGGAATAAATGGATCCATCAGTCTGGCCGTACCAC*
*TNGTGGCTGCTGGCAGCAATNNTCTCTGCTCTGCCGTCTCAGCAAATGTTAATAGAATAAAA*
*ATTTAAAAATAATTTAAAAAACACACAAAAATGCGTCACACAGAATACAGAGAGAGAGAGAC*
*AGAGAGAGAGAGAGAGAGAGAGATGGGGGAGACCGTTTATTTCACAACTTTGTGTGTTTATA*
*CATGAAGGGGGAAATAAGAAAGTGAAGAAGAAAATACAACATTTAAAACAATTTTACAGTCC*
*ATCATTAAAAATTTATGTATCATTCAGGATGGAGAAGGTTCTACTGGGATATGTTTATATCT*
*ACTAAGCAAATGTATGCTGTGTAAAGACTACACCACACTAAGGACATCTGGATGCTGTAAAA*
*ATAAGAGAAGAACCAGATGGATATTAAGCCCCCCAACACACACTTTATCCTTCCTTCCTTCA*
*TCTTTTTTCATCTGTGGGGAAGAAAATAAGGTCTTGCCTTTGGTGTTTATATTTCCATAACC*
*TTTTAATTCTATTTTTCATTTGAGCTGACTTGTAGCCACTTCAGACTATCAATGGAATCTTA*
*TGTTGAGCCTTTCTCTGGCTTTCCTTCCTCCACTATCTCTCCAACTTTAGAGATCATCCCCT*
*CTCCCTCCAGT*GCGTTCTATCTCCCCCACA*CCCACCCAAGCTTGGCGTAATC* (SEQ ID
NO:33)

[00175] Type 3: **ZBTB20-E1C** LSAMP-*E4*-E5-*E6*-E7

**CAACGCAGAGTACATGGGACACAACATCAAGAGCAGGAAAATGGACTCATTAGGGAGGCAGG**
**CAGTCATTACCACTCACACTGTACTTCCAGGGAGACACCGATTATAAGAAGAGAAACTCAGC**
**GCTGGGGAAG***AAGGAAGGGAATTTGAAGGAGAAGAAGAATATCTGGAGATCCTTGGCATCAC*
*CAGGGAGCAGTCAGGCAAATATGAGTGCAAAGCTGCCAACGAGGTCTCCTCGGCGGATGTCA*
*AACAAGTCAAGGTCACTGTGAACTA*TCCTCCCACTATCACAGAATCCAAGAGCAATGAAGCC
ACCACAGGACGACAAGCTTCACTCAAATGTGAGGCCTCGGCAGTGCCTGCACCTGACTTTGA
GTGGTACCGGGATGACACTAG*GATAAATAGTGCCAATGGCCTTGAGATTAAGAGCACGGAGG*
*GCCAGTCTTCCCTGACGGTGACCAACGTCACTGAGGAGCACTACGGCAACTACACCTGTGTG*
*GCTGCCAACAAGCTGGGGGTCACCAATGCCAGCCTAGTCCTTTTC*AGACCTGGGTCGGTGAG
AGGAATAAATGGATCCATCAGTCTGGCCGTACCACTGTGGCTGCTGGCAGCATCTCTGCTCT
GCCTTCTCAGCAAATGTTAATAGAATAAAAATTTAAAAATAATTTAAAAAACACACAAAAAT
GCGTCACACAGAATACAGAGAGAGAGACAGAGAGAGAGAGAGAGAGAGAGATGGGGGAGA

CCGTTTATTTCACAACTTTGTGTGTTTATACATGAAGGGGGAAATAAGAAAGTGAAGAAGAA
AATACAACATTTAAAACAATTTTACAGTCCATCATTAAAAATTTATGTATCATTCAGGATGG
AGAAGGTTCTACTGGGATATGTTTATATCTACTAAGCAAATGTATGCTGTGTAAAGACTACA
CCACACTAAGGACATCTGGATGCTGTAAAAATAAGAGAAGAACCAGATGGATATTAAGCCCC
CCAACACACACTTTATCCTTCCTTCCTTCATCTTTTTTCATCTGTGGGGAAGAAAATAAGGT
CTTGCCTTTGGTGTTTATATTTCCATAACCTTTTAATTCTATTTTTCATTTGAGCTGACTTG
TAGCCACTTCAGACTATCAATGGAATCTTATGTTGAGCCTTTCTCTGGCTTTCCTTCCTCCA
CTATCTCTCCAACTTTAGAGATCATCCCCTCTCCCTCCAGTGCGTTCTATCTCCCCCACACC
CACCCAAGCTTGGCGTAATC (SEQ ID NO:34)

[00176] Type 4: **ZBTB20-E1-*E1A*-E1B**-LSAMP-*E4*

**ACATGGGGAAGTTTAGGAGTTGAGGAAAGAAGATTAAAGAGCGCGAGGA*GATTTTATAGACC***
***AGTGGAATACAGCCCTTGTGCATATGAAGATCAGGTGACAAGTTTGSTGCCTACCAGCCTCC***
***ACAGCAATATGCCCTTTCACG*AGTCCCTATCGCCCAGGCTGGAGTGCAGTGGCGTGATCTCT**
**GCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCTTGCCTCAGCCTCCCGAGTAGC**
**TGGGATTACAGGA***AGGGAATTTGAAGGAGAAGAAGAATATCTGGAGATCCTTGGCATCACCA*
*GGGAGCAGTCAGGCAAAAGCTTGGCGTAATC* (SEQ ID NO:35)

[00177] Type 5: **ZBTB20-E1*A**-LSAMP-*E4*

**ACATGGGGGAGGAAAGAAGATTAAAGAGCGCGAGGAGATTTTATAGACCAGTGGAATACAGG**
**CCTTGTGCATATGAAGATCAGGTGACAAGTTTGCTGCCTACCAGCCTCCACAGCAATATGCC**
**CTTTCACGGA***AGGGAATTTGAAGGAGAAGAAGAATATCTGGAGATCCTTGGCATCACCAGGG*
*AGCAGTCAGGCAAAAGCTTGGCGTAATC* (SEQ ID NO:36)

[00178] Type 6: **ZBTB20-E1-*E1B*-**LSAMP-E4

**ACATGGGGGAGGAAAGAAGATTAAAGAGCGCGAGGAGACAG***AGTCCCTATCGCCCAG*
*GCTGGAGTGCAGTGGCGTGATCTCTGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATT*
*CTCTTGCCTCAGCCTCCCGAGTAGCTGGGATTACAG*GAAGGGAATTTGAAGGAGAAGAAGAA
TATCTGGAGATCCTTGGCATCACCAGGGAGCAGTCAGGCAAAAG (SEQ ID NO:37)

[00179] Type 7: **ZBTB20-E1-LSAMP-*E4***

**ACATGGGGAAGTTTAGGAGTTGAGGAAAGAAGATTAAAGAGCGCGAGGAG***GAAGGGAATTTG*
*AAGGAGAAGAAGAATATCTGGAGATCCTTGGCATCACCAGGGAGCAGTCAGGCAAAAGCTTG*
*GSGTAATC* (SEQ ID NO:38)

[00180] Type 8: **ZBTB20-E1**-LSAMP-*E3**- E4

**ACATGGGGGGGCGGGGGGAAGTTTAGGAGTTGAGGAAAGAAGATTAAAGAGCGCGAGG***AGAT*
*TAACTTACTCTTAATGATCTTCCAACACTTGAGAAGGTCAGTAGCCCTCCATCTGTCATTCT*

*CCAAGTTCACCAACAGCTTATCCACCCATCAAAGGTGCTTTTGTAACAAAATCCATGCATAA TGAAACCAAGAAAGGAAGGGAATTTGAAGGAGAAGAAGAATATCTGGAGATCCTTGGCATCA* CCAGGGAGCAGTCAGGCAAAAGCTTGGCGTAATC (SEQ ID NO:39)

[00181] Type 9: **ZBTB20-E1C** LSAMP-*E4*

**ACATGGGGAAGAGCAGGAAAATGGACTCATTAGGGAGGCAGGCAGTCATTACCACTCACACT GTACTTCCAGGGAGACACCGATTATAAGAAGAGAAACTCAGCGCTGGGGAAGAAG***GAAGGGA ATTTGAAGGAGAAGAAGAATATCTGGAGATCCTTGGCATCACCAGGGAGCAGTCAGGCAAAA GCTTGGCGTAATC* (SEQ ID NO:40)

[00182] Type 10: **ZBTB20-E1C**-LSAMP-*E3**-E4

**ACATGGGGAGTACATGGGGATATACACAACATCAAGAGCAGGAAAATGGACTCATTAGGGAG GCAGGCAGTCATTACCACTCACACTGTACTTCCAGGGAGACACCGATTATAAGAAGAGAAAC TCAGCGCTGGGGAAGA***AGATTAACTTACTCTTAATGATCTTCCAACACTTGAGAAGGTCAGT AGCCCTCCATCTGTCATTCTCCAAGTTCACCAACAGCTTATCCACCCATCAAAGGTGCTTTT GTAACAAAATCCATGCATAATGAAACCAAGAAAGGAAGGGAATTT*GAAGGAGAAGAAGAATA TCTGGAGATCCTTGGCATCACCAGGGAGCAGTCAGGCAAAAGCTTGGCGTAATC (SEQ ID NO:41)

[00183] LPCS1: LSAMP-E0*-*E1*-E2-*E3*-E4

GGAGGAGGATAGGAAGCAGGAAAGCGGGAGAGCTCGAGGGACAAGGGGGCTCGGTGTGTTTA CACCAGGCACGGGCTACGAGCGTCCATCCCGGCCCCTGGCTTGCGCTCCCGAAGAGGAGAGC AAGGCTGTTCTGGGATCCGGCCGTCGTGCGGCAAGAGGCTTGTCTGTCCGGGTTGCCGGAAC CAGGAGAACCCAGAGGGAAACCGAGGGAAAGGAGCGGCGCGTTTTACTAGAGAGAGCGCGAG CGGAAGAGGCGAGAGCAGGAGCGCGCGAGGGAGCATCGAGCGCAGCGGAGACATGAGGACCT ACTGGCTGCACAGCGTCTGGGTGCTGGGCTTTTTCCTGTCCCTCTTCTCATTGCA*AGGACTG CCTGTTCGCAGCGTGGATTTTAACCGAGGCACGGACAACATCACCGTGAGGCAGGGGGACAC AGCCATCCTCAG*GTGCGTTGTAGAAGACAAGAACTCAAAGGTGGCCTGGTTGAACCGTTCTG GCATCATTTTTGCTGGACATGACAAGTGGTCTCTGGACCCACGGGTTGAGCTGGAGAAACGC CATTCTCTGGAATACAGCCTCCGAATCCAGAAGGTGGATGTCTATGATGAGGGTTCCTACAC TTGCTCAGTTCAGACACAGCATGAGCCCAAGACCTCCCAAGTTTACTTGATCGTACAAG*TCC CACCAAAGATCTCCAATATCTCCTCGGATGTCACTGTGAATGAGGGCAGCAACGTGACTCTG GTCTGCATGGCCAATGGCCGTCCTGAACCTGTTATCACCTGGAGACACCTTACACCAACTGG* AAGGGAATTTGAAGGAGAAGAAGAATATCTGGAGATCCTTGGCATCACCAGGGAGCAGTCAG GCAAAAGCTTGGCGTAATCC (SEQ ID NO:46)

[00184] All cDNA sequences were distinct from the wild type LSAMP and ZBTB20 sequences. Seven of the ten fusion transcripts involved a fusion between exon 1 of ZBTB20 and exon 4 of LSAMP. Figure 5. The remaining three fusion transcripts involved a fusion between exon 1 of ZBTB20 and exon 3* of LSAMP. Figure 5. Exon 3* represents a novel exon sequence from the LSAMP locus that has not been previously annotated and that is associated with the genomic rearrangement of the ZBTB20 and LSAMP genes. The nucleotide sequence of exon 3* corresponds to SEQ ID NO:45. Open reading frame (ORF) searches predicted severe N-terminal truncation of the LSAMP protein or the absence of ORF in the LSAMP cDNA.

[00185] One of ordinary skill in the art can design primers for amplifying these ZBTB20/LSAMP gene fusion transcripts. In one embodiment, the forward primer is designed to hybridize to a region of exon 1 of ZBTB20 (e.g., E1, E1A, E1B, or E1C) and the reverse primer is designed to hybridize to exon 3* or exon 4 of LSAMP. In another embodiment, the primers are designed to amplify an amplification product that comprises a first region from exon 1 of ZBTB20 (e.g., E1, E1A, E1B, or E1C) and a second region from exon 3* or exon 4 of LSAMP, wherein detecting the amplification product indicates the presence of a ZBTB20/LSAMP gene fusion transcript. In certain embodiments, the forward primer hybridizes under conditions of high stringency to a region within E1 of the ZBTB20 gene and the reverse primer (e.g., SEQ ID NO:44) hybridizes under conditions of high stringency to a region within exon 4 of LSAMP. Exon 4 of LSAMP corresponds to nucleotides 1022-1156 of SEQ ID NO:10. These primers could be used to amplify, for example, the Type 1, 4, 5, 6, 7, and 8 fusion transcripts in Figure 5. In other embodiments, the forward primer hybridizes under stringent conditions to a region within E1C of the ZBTB20 gene and the reverse primer (e.g., SEQ ID NO:44) hybridizes under stringent conditions to a region within exon 4 of LSAMP. These primers could be used to amplify, for example, the Type 2, 3, 9, and 10 fusion transcripts in Figure 5. In certain embodiments, the reverse primer (e.g., SEQ ID NO:42 or SEQ ID NO:43) hybridizes to a region within exon 7 of LSAMP or a sequence complementary thereto. Because exon 3* is a unique LSAMP exon associated with the genomic arrangement of the ZBTB20 and LSAMP genes, it is also possible to design primers to amplify unique regions of exon 3* or a unique region that spans the junction between exon 3* and exon 4 of LSAMP and, thus, can be used to specifically identify the genomic arrangement of the ZBTB20 and LSAMP genes.

[00186] LPCS1 (SEQ ID NO:46) represents an alternatively spliced cDNA of the LSAMP locus and includes exon 0*, a novel exon sequence from the LSAMP locus that has not been previously annotated and that is associated with the genomic rearrangement of the ZBTB20 and LSAMP genes. The nucleotide sequence of exon 0* corresponds to SEQ ID NO:47. Because exon 0* is a unique LSAMP exon associated with the genomic arrangement of the ZBTB20 and LSAMP genes, one of skill in the art can design primers to amplify unique regions of exon 0* or a unique region that spans the junction between exon 0* and exon 1 of LSAMP and, thus, can be used to specifically identify the genomic arrangement of the ZBTB20 and LSAMP genes.

[00187] All patents, patent applications, and published references cited herein are hereby incorporated by reference in their entirety. While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

REFERENCES

[00188] The following references are cited in the application and provide general information on the field of the invention and provide assays and other details discussed in the application. The following references are incorporated herein by reference in their entirety.

[00189] 1. Siegel, R.; Naishadham, D.; Jemal, A. Cancer statistics. *CA Cancer J. Clin.* **2013**, *63*, 11–30.

[00190] 2. Chornokur, G.; Dalton, K.; Borysova, M.E.; Kumar, N.B. Disparities at presentation, diagnosis, treatment, and survival in African American men affected by prostate cancer. *Prostate* **2011**, *71*, 985–997.

[00191] 3. Schwartz, K.; Powell, I.J.; Underwood, W., 3rd; George, J.; Yee, C.; Banerjee, M. Interplay of race, socioeconomic status, and treatment on survival of patients with prostate cancer. *Urology* **2009**, *74*, 1296–1302.

[00192] 4. Major, J.M.; Oliver, M.N.; Doubeni, C.A.; Hollenbeck, A.R.; Graubard, B.I.; Sinha, R. Socioeconomic status, healthcare density, and risk of prostate cancer among African American and Caucasian men in a large prospective study. *Cancer Causes Control* **2012**, *23*, 1185–1191.

[00193] 5. Sridhar, G.; Masho, S.W.; Adera, T.; Ramakrishnan, V.; Roberts, J.D. Do African American men have lower survival from prostate cancer compared with White men? A meta-analysis. *Am. J Mens. Health* **2010**, *4*, 189–206.

[00194] 6. Cullen, J.; Brassell, S.; Chen, Y.; Porter, C.; L'Esperance, J.; Brand, T.; McLeod, D.G. Racial/ethnic patterns in prostate cancer outcomes in an active surveillance cohort. *Prostate Cancer* **2011**, *2011*, doi:10.1155/2011/234519.

[00195] 7. Berger, A.D.; Satagopan, J.; Lee, P.; Taneja, S.S.; Osman, I. Differences in clinicopathologic features of prostate cancer between black and white patients treated in the 1990s and 2000s. *Urology* **2006**, *67*, 120–124.

[00196] 8. Kheirandish, P.; Chinegwundoh, F. Ethnic differences in prostate cancer. *Br. J. Cancer* **2011**, *105*, 481–485.

[00197] 9. Odedina, F.T.; Akinremi, T.O.; Chinegwundoh, F.; Roberts, R.; Yu, D.; Reams, R.R.; Freedman, M.L.; Rivers, B.; Green, B.L.; Kumar, N. Prostate cancer disparities in black men of African descent: A comparative literature review of prostate cancer burden among black men in the United States, Caribbean, United Kingdom, and West Africa. *Infect. Agents Cancer* **2009**, *4*, doi:10.1186/1750-9378-4S1-S2.

[00198] 10. Heath, E.I.; Kattan, M.W.; Powell, I.J.; Sakr, W.; Brand, T.C.; Rybicki,

B.A.; Thompson, I.M.; Aronson, W.J.; Terris, M.K.; Kane, C.J.; *et al.* The effect of race/ethnicity on the accuracy of the 2001 Partin Tables for predicting pathologic stage of localized prostate cancer. *Urology* **2008**, *71*, 151–155.

[00199] 11. Moul, J.W.; Sesterhenn, I.A.; Connelly, R.R.; Douglas, T.; Srivastava, S.; Mostofi, F.K.; McLeod, D.G. Prostate-specific antigen values at the time of prostate cancer diagnosis in African-American men. *JAMA* **1995**, *274*, 1277–1281.

[00200] 12. Tewari, A.; Horninger, W.; Badani, K.K.; Hasan, M.; Coon, S.; Crawford, E.D.; Gamito, E.J.; Wei, J.; Taub, D.; Montie, J.; *et al.* Racial differences in serum prostate-specific (PSA) doubling time, histopathological variables and long-term PSA recurrence between African-American and white American men undergoing radical prostatectomy for clinically localized prostate cancer. *BJU Int.* **2005**, *96*, 29–33.

[00201] 13. Wallace, T.A.; Prueitt, R.L.; Yi, M.; Howe, T.M.; Gillespie, J.W.; Yfantis, H.G.; Stephens, R.M.; Caporaso, N.E.; Loffredo, C.A.; Ambs, S. Tumor immunobiological differences in prostate cancer between African-American and Caucasian-American men. *Cancer Res.* **2008**, *68*, 927–936.

[00202] 14. Prensner, J.R.; Rubin, M.A.; Wei, J.T.; Chinnaiyan, A.M. Beyond PSA: The next generation of prostate cancer biomarkers. *Sci. Transl. Med.* **2012**, *4*, doi:10.1126/scitranslmed.3003180.

[00203] 15. Rubin, M.A.; Maher, C.A.; Chinnaiyan, A.M. Common gene rearrangements in prostate cancer. *J. Clin. Oncol.* **2011**, *29*, 3659–3668.

[00204] 16. Sreenath, T.L.; Dobi, A.; Petrovics, G.; Srivastava, S. Oncogenic activation of ERG: A predominant mechanism in prostate cancer. *J. Carcinog.* **2011**, *11*, 10–21.

[00205] 17. Petrovics, G.; Liu, A.; Shaheduzzaman, S.; Furasato, B.; Sun, C.; Chen, Y.; Nau, M. Ravindranath, L.; Chen, Y.; Dobi, A.; *et al.* Frequent overexpression of ETS-related gene-1 (*ERG1*) in prostate cancer transcriptome. *Oncogene* **2005**, *24*, 3847–3852.

[00206] 18. Tomlins, S.A.; Rhodes, D.R.; Perner, S.; Dhanasekaran, S.M.; Mehra, R.; Sun, X.W.; Varambally, S.; Cao, X.; Tchinda, J.; Kuefer, R.; *et al.* Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **2005**, *310*, 644–648.

[00207] 19. Magi-Galluzzi, C.; Tsusuki, T.; Elson, P.; Simmerman, K.; LaFarque, C.; Esqueva, R.; Klein, E.; Rubin, M.A.; Zhou, M. TMPRSS2-ERG gene fusion prevalence and class are significantly different in prostate cancer of Caucasian, African-American and

Japanese patients. *Prostate* **2011**, *71*, 489–497.

[00208] 20. Rosen, P.; Pfister, D.; Young, D.; Petrovics, G.; Chen, Y.; Cullen, J.; Bohm, D.; Perner, S.; Dobi, A.; McLeod, D.G.; *et al.* Differences in frequency of ERG oncoprotein expression between index tumors of Caucasian and African American patients with prostate cancer. *Urology* **2012**, *80*, 749–753.

[00209] 21. Hu, Y.; Dobi, A.; Sreenath, T.; Cook, C.; Tadase, A.Y.; Ravindranath, L.; Cullen, J.; Furusato, B.; Chen, Y.; Thanqapazham, R.L.; *et al.* Delineation of TMPRSS2-ERG splice variants in prostate cancer. *Clin. Cancer Res.* **2008**, *14*, 4719–4725.

[00210] 22. Gary K Geiss, et al. (2008) Direct multiplexed measurement of gene expression with color-coded probe pairs, *Nature Biotechnology* **26**:317-25.

[00211] 23. Paolo Fortina and Saul Surrey, (2008) Digital mRNA Profiling, *Nature Biotechnology* **26**:317-25.

[00212] Farrell J, Petrovics G, McLeod DG, Srivastava S.: Genetic and molecular differences in prostate carcinogenesis between African American and Caucasian American men. *International Journal of Molecular Sciences*. 2013; 14(8):15510-31.

[00213] 25. Rodriquez-Suarez et al., Urine as a source for clinical proteome analysis: From discovery to clinical application, *Biochimica et Biophysica Acta* (2013).

[00214] 26. Shi et al., Antibody-free, targeted mass-spectrometric approach for quantification of proteins at low picogram per milliliter levels in human plasma/serum, *PNAS*, 109(38):15395-15400 (2012).

[00215] 27. Elentiboba-Johnson and Lim, Fusion peptides from oncogenic chimeric proteins as specific biomarkers of cancer, *Mol Cell Proteomics*, 12:2714 (2013).

[00216] 28. Ras/Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR Cascade Inhibitors: How Mutations Can Result in Therapy Resistance and How to Overcome Resistance, *Oncotarget*, 3(10):1068-1111 (2012).

[00217] 29. Kuhn et al., High-resolution genomic profiling of adult and pediatric core-binding factor acute myeloid leukemia reveals new recurrent genomic alterations, Blood, 119(10):e67 (2012).

[00218] 30. Pasic et al., Recurrent Focal Copy Number Changes and Loss of Heterozygosity Implicate Two Non-Coding RNAs and One Tumor Suppressor Gene at Chromosome 3q13.31 in Osteosarcoma, Cancer Research, 70(1):160-71 (2010).

[00219] 31. Chen et al., The t(1;3) breakpoint-spanning genes *LSAMP* and *NORE1* are involved in clear cell renal cell carcinomas, Cancer Cell, 4:405-413 (2003).

[00220] 32. Ntougkos et al., Clin Cancer Res, 11:5764-5768 (2005).

[00221] 33. Huang et al., Eur J Cancer 49:3729-37 (2013).

[00222] 34. Mao et al., Cancer Res, 70:5207-5212 (2010).

[00223] 35. Blattner et al., Neoplasia 16(1):14-20 (2014).

[00224] 36. Khani et al., Clin Cancer Res 20(18):4925-34 (2014).

What is claimed:

1. A method of identifying or characterizing prostate cancer in a subject, the method comprising detecting the presence or absence of a genomic rearrangement in a biological sample obtained from the subject, wherein the biological sample comprises prostate cells or nucleic acid or polypeptides isolated from the prostate cells and wherein the genomic rearrangement occurs in chromosome region 3q13 and involves a ZBTB20 gene and an LSAMP gene and wherein detecting the genomic rearrangement identifies prostate cancer in the subject or characterizes the prostate cancer in the subject as being an aggressive form of prostate cancer or as having an increased risk of developing into an aggressive form of prostate cancer.

2. The method of claim 1, wherein the biological sample comprises prostate cells or nucleic acids or polypeptides isolated therefrom.

3. The method of claim 1 or 2, wherein the genomic rearrangement is a gene fusion, a gene inversion, a gene deletion, or a gene duplication.

4. The method of any one of claims 1-3, wherein prostate cancer from the subject does not express a TMPRSS2/ERG gene fusion.

5. The method of any one of claims 1-4, wherein the subject is of African descent.

6. The method of claim 1 or 2, wherein detecting the presence of the genomic rearrangement in the biological sample comprises detecting a chromosomal rearrangement of genomic DNA having a first portion from the ZBTB20 gene and a second portion from the LSAMP gene.

7. The method of claim 1 or 2, wherein detecting the presence of the genomic rearrangement in the biological sample comprises detecting a chimeric mRNA or cDNA transcript having a first nucleic acid portion from the ZBTB20 gene and a second nucleic acid portion from the LSAMP gene.

8. The method of claim 7, wherein the first nucleic acid portion is exon 1of the ZBTB20 gene and the second nucleic acid portion is exon 3* or exon 4 of LSAMP.

9. The method of claim 1 or 2, wherein detecting the presence of the genomic rearrangement in the biological sample comprises detecting a deletion in chromosome region 3q13, wherein the deletion spans the ZBTB20 and LSAMP genes.

10. The method of claim 1 or 2, wherein detecting the presence of the genomic rearrangement in the biological sample comprises detecting a mRNA or cDNA transcript comprising exon 0* (SEQ ID NO: 47) or exon 3* (SEQ ID NO:45).

11. The method of any one of claims 1-10, further comprising measuring the expression of one or more of the following genes: COL10A1, HOXC4, ESPL1, MMP9, ABCA13, PCDHGA1, or AGSK1.

12. The method of any one of claims 1-11, further comprising measuring the expression of one or more of the following genes: ERG, AMACR, PCA3, or PSA.

13. The method of any one of claims 1-12, further comprising a step of selecting a treatment regimen for the subject based on the detection of the genomic rearrangement.

14. The method of any one of claims 1-13, further comprising a step of treating the subject if the genomic rearrangement is detected in the biological sample obtained from the subject.

15. A composition comprising a polynucleotide probe, wherein the polynucleotide probe hybridizes under high stringency conditions to a junction of a chimeric nucleic acid, wherein the chimeric nucleic acid comprises:

(a) a first portion from a ZBTB20 gene and a second portion from a LSAMP gene;

(b) a first portion from exon 3* of a LSAMP gene and a second portion from exon 4 of a LSAMP gene; or

(c) a first portion from exon 0* of a LSAMP gene and a second portion from exon 1 of a LSAMP gene.

16. The composition of claim 15, wherein the first portion is exon 1of the ZBTB20 gene and the second portion is exon 3* or exon 4 of LSAMP.

17. The composition of claim 15 or 16, wherein the polynucleotide probe is labeled.

18. A kit comprising the composition of claim 17 and a second composition comprising a polynucleotide probe that hybridizes under high stringency conditions to a gene selected from COL10A1, HOXC4, ESPL1, MMP9, ABCA13, PCDHGA1, and AGSK1.

19. A kit comprising the composition of claim 17 and a second composition comprising a polynucleotide probe that hybridizes under high stringency conditions to a gene selected from ERG, AMACR, PCA3, and PSA.

20. A composition comprising a double stranded oligonucleotide duplex, wherein the oligonucleotide duplex comprises:

(a) a first nucleic acid hybridized to a second nucleic acid, wherein the first nucleic acid comprises a first portion from a ZBTB20 gene fused to a second portion from a LSAMP gene and wherein the second nucleic acid is a polynucleotide probe that is hybridized to a junction between the first portion from the ZBTB20 gene and the second portion from the LSAMP gene;

(b) a first nucleic acid hybridized to a second nucleic acid, wherein the first nucleic acid comprises a first portion from exon 3* of a LSAMP gene fused to a second portion from exon 4 of a LSAMP gene and wherein the second nucleic acid is a polynucleotide probe that is hybridized to a junction between the first portion from exon 3* of the LSAMP gene and the second portion from exon 4 of the LSAMP gene; or

(c) a first nucleic acid hybridized to a second nucleic acid, wherein the first nucleic acid comprises a first portion from exon 0* of a LSAMP gene fused to a second portion from exon 1 of a LSAMP gene and wherein the second nucleic acid is a polynucleotide probe that is hybridized to a junction between the first portion from exon 0* of the LSAMP gene and the second portion from exon 1 of the LSAMP gene

21. The composition of claim 20, wherein the first portion is exon 1of the ZBTB20 gene and the second portion is exon 3* or exon 4 of LSAMP.

22. The composition of claim 20 or 21, wherein the polynucleotide probe is labeled.

23. An isolated antibody that binds to a truncated LSAMP polypeptide encoded by a gene fusion, wherein the gene fusion has a first portion from a ZBTB20 gene and a second portion from a LSAMP gene.

24. A composition for amplifying a gene fusion, wherein the gene fusion has a first portion from a ZBTB20 gene and a second portion from a LSAMP gene, the composition comprising:

a) a first polynucleotide primer comprising a sequence that hybridizes to the first portion of the gene fusion from the ZBTB20 gene; and

b) a second polynucleotide primer comprising a sequence that hybridizes to the second portion of the gene fusion from the LSAMP gene,

wherein the first and second polynucleotide primers are capable of amplifying a nucleotide sequence from the gene fusion that spans the junction between the first portion of the gene fusion from the ZBTB20 gene and the second portion of the gene fusion from the LSAMP gene.

25. The method of claim 24, wherein the first portion of the gene fusion is exon 1of the ZBTB20 gene and the second portion of the gene fusion is exon 3* or exon 4 of LSAMP.

26. A composition for amplifying a gene fusion, wherein the gene fusion has a first portion from exon 3* of a LSAMP gene and a second portion from exon 4 of a LSAMP gene, the composition comprising:

a) a first polynucleotide primer comprising a sequence that hybridizes to the first portion of the gene fusion from exon 3* of the LSAMP gene; and

b) a second polynucleotide primer comprising a sequence that hybridizes to the second portion of the gene fusion from exon 4 of the LSAMP gene,

wherein the first and second polynucleotide primers are capable of amplifying a nucleotide sequence from the gene fusion that spans the junction between the first portion of the gene fusion from exon 3* of the LSAMP gene and the second portion of the gene fusion from exon 4 of the LSAMP gene.

27. A composition for amplifying a gene fusion, wherein the gene fusion has a first portion from exon 0* of a LSAMP gene and a second portion from exon 1 of a LSAMP gene, the composition comprising:

a) a first polynucleotide primer comprising a sequence that hybridizes to the first portion of the gene fusion from exon 0* of the LSAMP gene; and

b) a second polynucleotide primer comprising a sequence that hybridizes to the second portion of the gene fusion from exon 1 of the LSAMP gene,

wherein the first and second polynucleotide primers are capable of amplifying a nucleotide sequence from the gene fusion that spans the junction between the first portion of the gene fusion from exon 0* of the LSAMP gene and the second portion of the gene fusion from exon 1 of the LSAMP gene.

28. A method of selecting a targeted prostate cancer treatment for a patient of African descent, wherein the method comprises:

(a) excluding a prostate cancer therapy that targets the PI3K/PTEN/Akt/mTOR pathway as a treatment option; and

(b) selecting an appropriate prostate cancer treatment.

29. The method of 28, further comprising a step of testing a biological sample from the patient, wherein the biological sample comprises prostate cells to confirm that the prostate cells to do not contain a PTEN gene deletion.

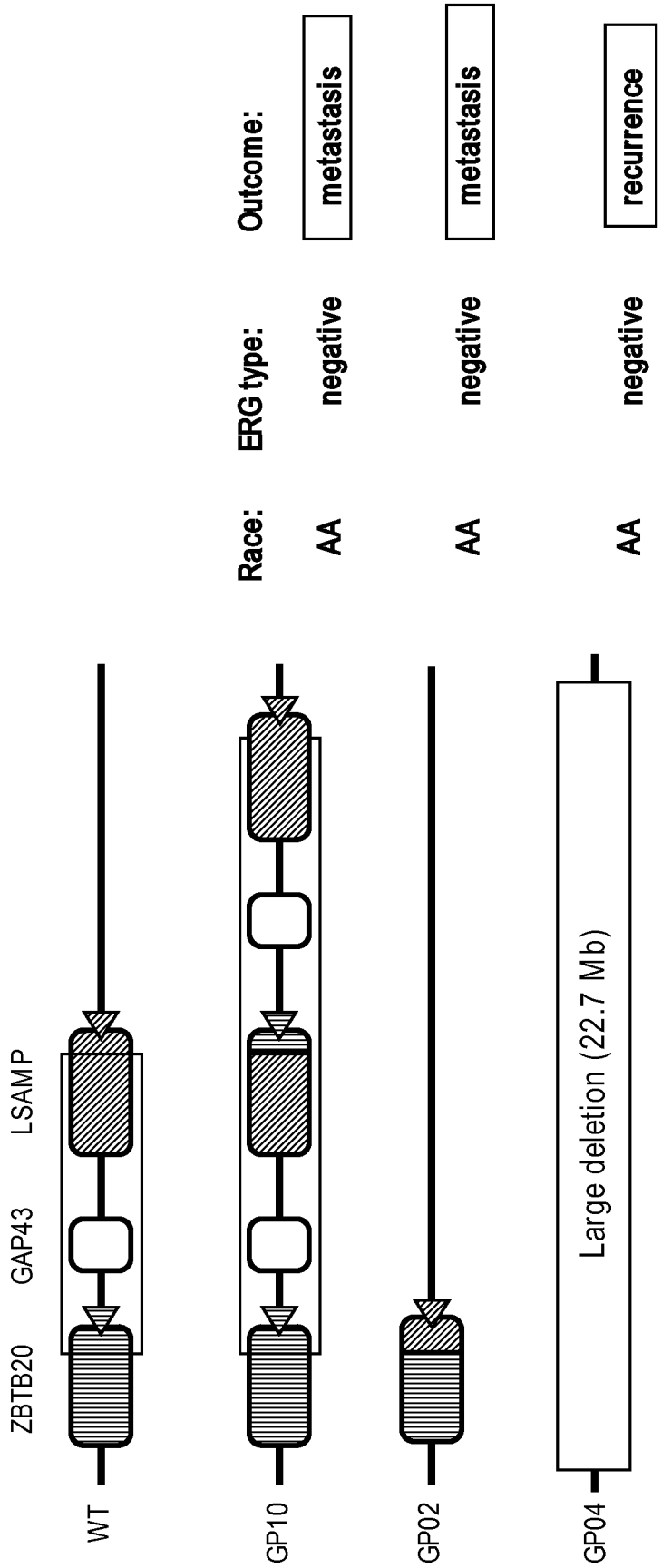
Outcome:
ERG type:
Race:
metastasis
negative
AA
metastasis
negative
AA
recurrence
negative
AA
ZBTB20
GAP43
LSAMP
WT
GP10
GP02
GP04
Large deletion (22.7 Mb)

FIG. 1

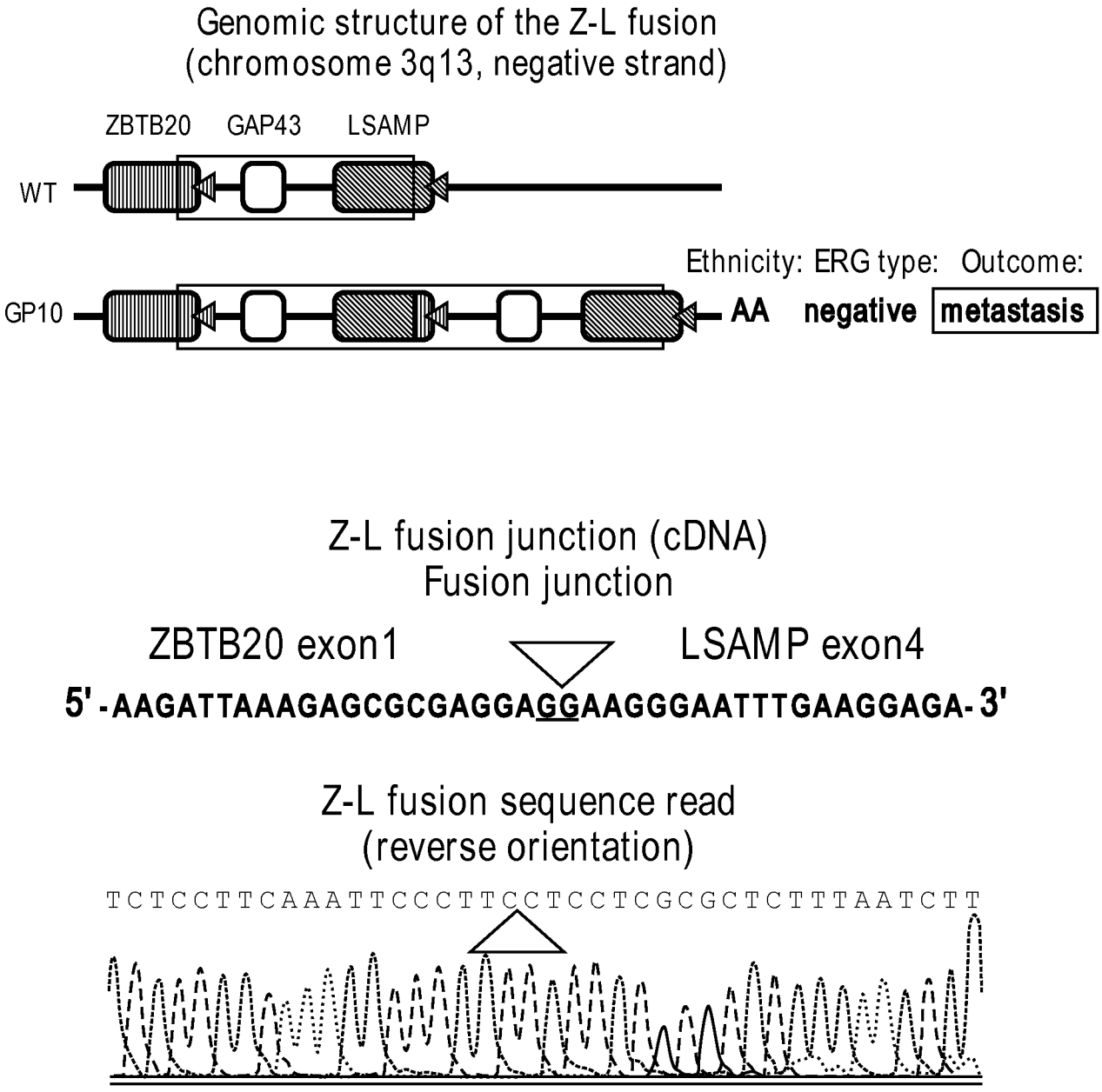
Genomic structure of the Z-L fusion
(chromosome 3q13, negative strand)
ZBTB20
GAP43
LSAMP
WT
GP10
Ethnicity:
ERG type:
Outcome:
AA
negative
metastasis
Z-L fusion junction (cDNA)
Fusion junction
ZBTB20 exon1
LSAMP exon4
5'-AAGATTAAAGAGCGCGAGGAGGAAGGGAATTTGAAGGAGA-3'
Z-L fusion sequence read
(reverse orientation)
TCTCCTTCAAATTCCCTTCCTCCTCGCGCTCTTTAATCTT

**FIG. 2**

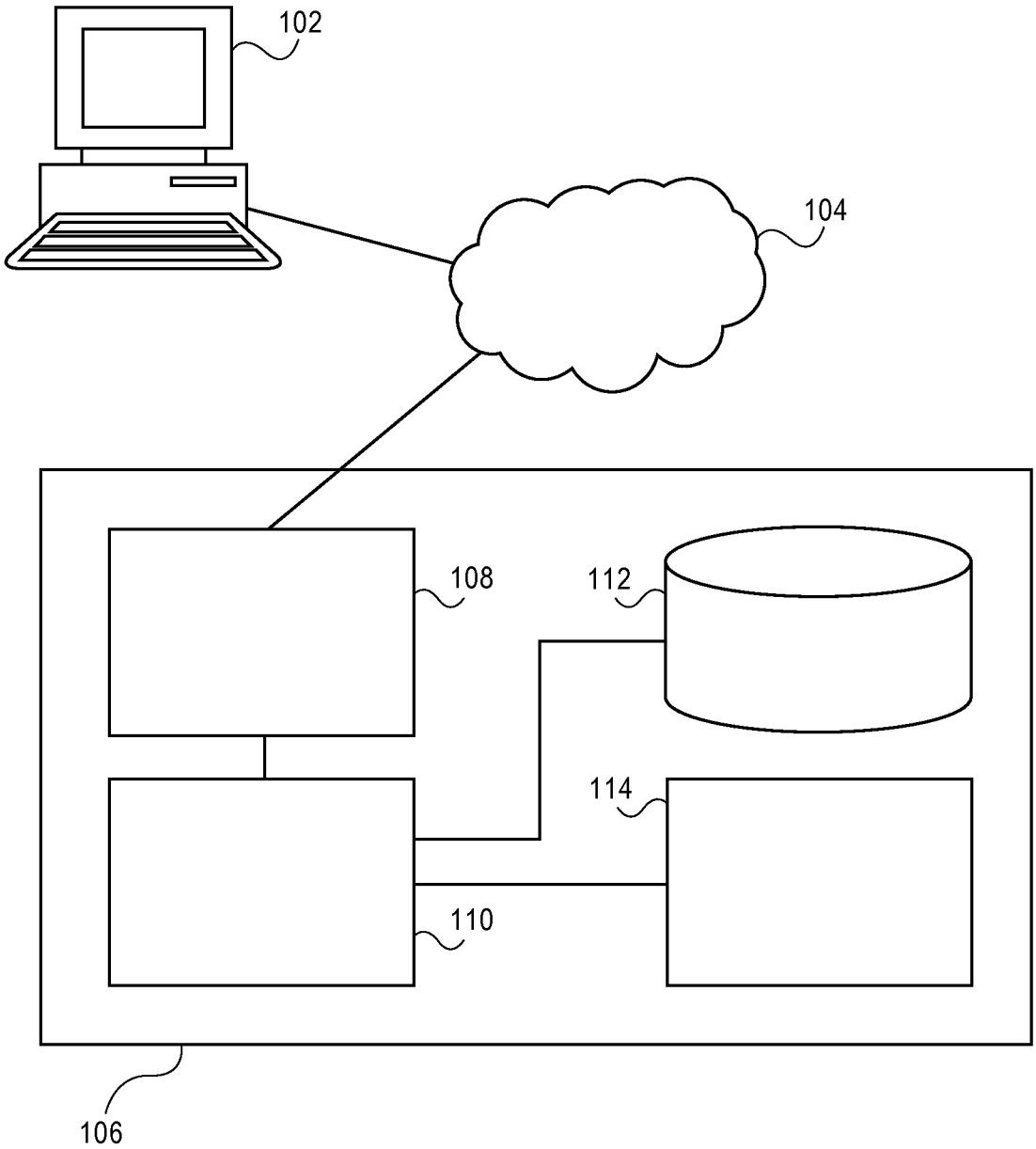


FIG. 3

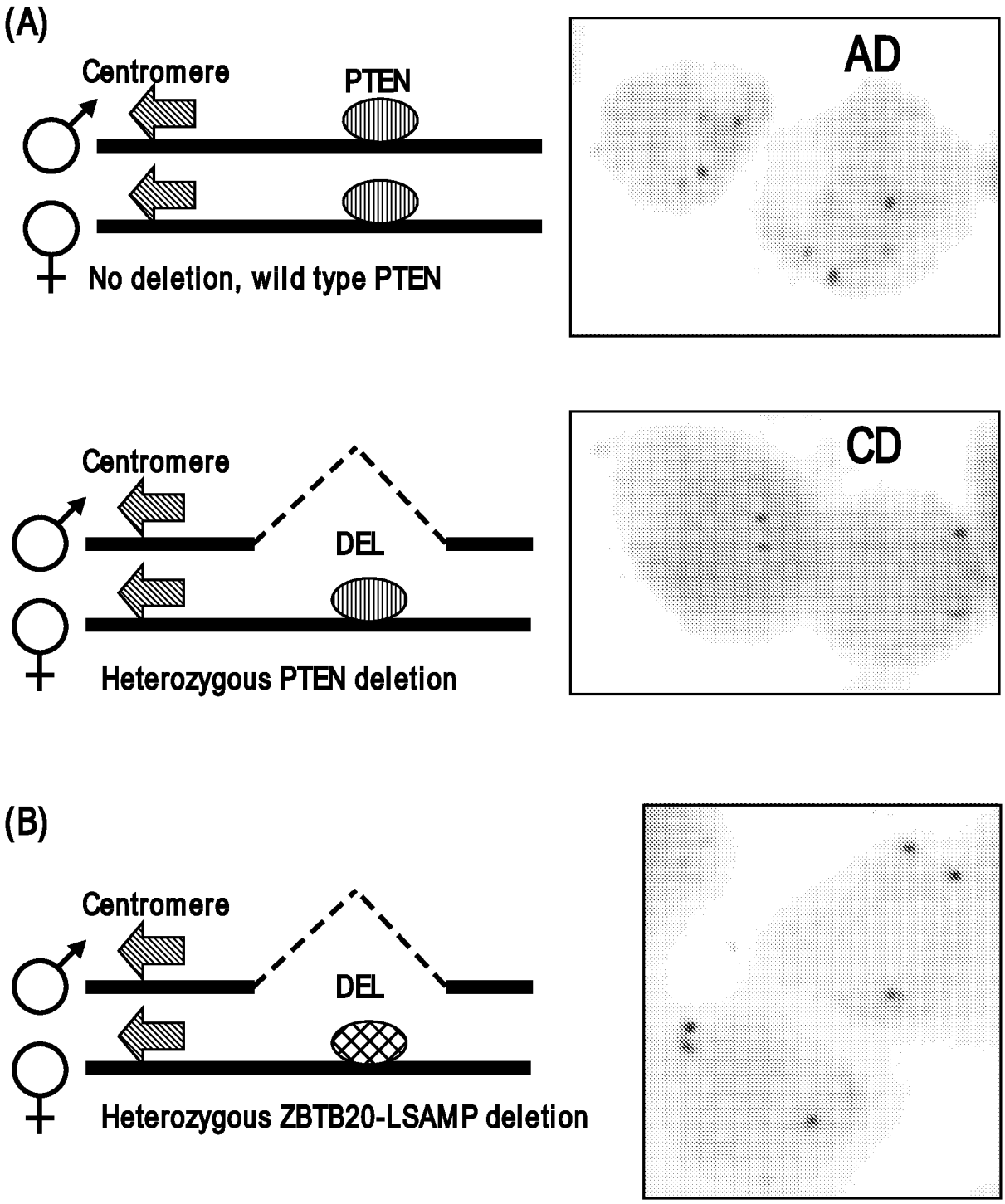
(A)
Centromere
PTEN
No deletion, wild type PTEN
AD
Centromere
DEL
Heterozygous PTEN deletion
CD
(B)
Centromere
DEL
Heterozygous ZBTB20-LSAMP deletion

FIG. 4

**5' ZBTB20 -> 3' LSAMP cDNA Fusion and Exon Map**

| | E1 (62) | E1A (95) | E1B (119) | E1C (121) | E0* (365) | E1 (662) | E2 (233) | E3 (126) | E3* (140) | E4 (135) | E5 (121) | E6 (149) | E7 (8052) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Type 1 | E1 | | | | | | | | | E4 | E5 | E6 | E7 |
| Type 2 | | | | E1C (121) | | | | | E3* (140) | E4 | E5 | E6 | E7 |
| Type 3 | | | | E1C (121) | | | | | | E4 | E5 | E6 | E7 |
| Type 4 | E1 | E1A(93) | E1B(116) | | | | | | | E4 | | | |
| Type 5 | E1 | E1* A(95) | | | | | | | | E4 | | | |
| Type 6 | E1 | | E1B(119) | | | | | | | E4 | | | |
| Type 7 | E1 | | | | | | | | | E4 | | | |
| Type 8 | E1 | | | | | | | | E3* (140) | E4 | | | |
| Type 9 | | | | E1C (121) | | | | | | E4 | | | |
| Type 10 | | | | E1C (121) | | | | | E3* (140) | E4 | | | |
| LPCS1 | | | | | E0* (365) | E1 (81) | E2 | E3 | | E4 | | | |

FIG. 5