SIGLEC12, a Human-specific Segregating (Pseudo)gene, Encodes a Signaling Molecule Expressed in Prostate Carcinomas*[s]

Nivedita Mitra[‡1], Kalyan Banda[‡], Tasha K. Altheide[‡2], Lana Schaffer[§], Teresa L. Johnson-Pais[¶], Joke Beuten[¶], Robin J. Leach[¶], Takashi Angata[‡3], Nissi Varki[‡], and Ajit Varki[‡4]

From the [‡]Glycobiology Research and Training Center, Departments of Medicine, Pathology and Cellular & Molecular Medicine, University of California at San Diego, La Jolla, California 92093, the [§]Scripps Research Institute, La Jolla, California 92037, and the [¶]Department of Cellular and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, Texas 78229

The primate *SIGLEC12* gene encodes one of the CD33-related Siglec family of signaling molecules in immune cells. We had previously reported that this gene harbors a human-specific missense mutation of the codon for an Arg residue required for sialic acid recognition. Here we show that this R122C mutation of the Siglec-XII protein is fixed in the human population, *i.e.* it occurred prior to the origin of modern humans. Additional mutations have since completely inactivated the *SIGLEC12* gene in some but not all humans. The most common inactivating mutation with a global allele frequency of 58% is a single nucleotide frameshift that markedly shortens the open reading frame. Unlike other CD33-related Siglecs that are primarily found on immune cells, we found that Siglec-XII protein is expressed not only on some macrophages but also on various epithelial cell surfaces in humans and chimpanzees. We also found expression on certain human prostate epithelial carcinomas and carcinoma cell lines. This expression correlates with the presence of the nonframeshifted, intact *SIGLEC12* allele. Although *SIGLEC12* allele status did not predict prostate carcinoma incidence, restoration of expression in a prostate carcinoma cell line homozygous for the frameshift mutation induced altered regulation of several genes associated with carcinoma progression. These stably transfected Siglec-XII-expressing prostate cancer cells also showed enhanced growth in nude mice. Finally, monoclonal antibodies against the protein were internalized by Siglec-XII-expressing prostate carcinoma cells, allowing targeting of a toxin to such cells. Polymorphic expression of Siglec-XII in humans thus has implications for prostate cancer biology and therapeutics.

Sialic acids (Sias)[5] are nine-carbon backbone sugar molecules typically found at terminal positions of glycan chains in *Deuterostomes* (vertebrates and some "higher" invertebrates),

making them potentially important in recognition events (1–3). One class of intrinsic Sia recognition proteins in vertebrates are Siglecs (sialic acid-binding immunoglobulin-like lectins). Siglecs are single-pass transmembrane proteins, with a Sia-binding site in the extracellular N-terminal Ig-like V-set domain (4–9). Such V-set domains are followed by one or more C2-set Ig-like domains. Siglecs can signal through one or more tyrosine-based signaling motif(s) in the cytoplasmic tail (5, 8, 9).

CD33-related Siglecs (CD33rSiglecs) are encoded by a subset of *SIGLEC* genes clustered on chromosome 19 in humans and chimpanzees. They are homologous in sequence and typically expressed on immune cells (10). Analyses of genomic *SIGLEC* sequences across humans, chimpanzees, baboons, rats, and mice showed that CD33rSiglecs are evolving rapidly (11). This is particularly pronounced in the Sia-recognizing V-set domain, suggesting that this domain is under the greatest selection pressure (11–14).

This study focuses on Siglec-12 (formerly Siglec-L1). We have shown previously that human Siglec-12 has an Arg → Cys (R122C) substitution mutation resulting in a protein unable to bind Sias (15). By convention, the protein is referred to as Siglec-XII[6] in humans, to differentiate it from Siglec-12 in other primates, where the Sia-binding arginine is present. The gene in both cases is referred to as *SIGLEC12*. Reversing this mutation *in vitro* restored Sia binding (15). Thus, except for the R122C mutation, the Sia-binding domain and reading frame were noted to remain intact.

The C-terminal signaling domain in CD33-related Siglecs has an immunoreceptor tyrosine-based inhibitory motif (ITIM) or an immunoreceptor tyrosine-based switch motif. ITIMs typically recruit the protein tyrosine phosphatases SHP-1 and SHP-2 or the lipid phosphatase SHIP-1, generally resulting in inhibitory downstream signaling (16). The function of the immunoreceptor tyrosine-based switch motif in CD33-related Siglecs is unclear. Siglec-XII has an ITIM motif in its cytoplasmic C terminus.

Analysis of chimpanzee, bonobo, gorilla, and orangutan sequences shows that they all have a functional *SIGLEC12* gene with the key Arg residue intact (15). Because humans and chimpanzees are typically ~99% identical in protein coding regions (17, 18) but have significant physiological, anatomical, and biomedical differences (such as a lower incidence of carcinomas in

the latter) (19), it is important to explore these genetic variations.

Here we ask whether the R122C mutation is universal to humans and explore the allele frequency of an additional polymorphic frameshift insertion mutation in the human *SIGLEC12* gene, which results in a premature stop codon in some individuals. Using newly generated monoclonal antibodies against Siglec-XII, we also confirm and extended earlier work where we noted unexpected expression in epithelial cells in addition to immune cells (15). Finally, we describe the mechanistic and potential therapeutic significance of Siglec-XII expression in human carcinomas, particularly in prostate carcinomas.

EXPERIMENTAL PROCEDURES

Cell Culture—All of the prostate cancer cell lines, PC-3, MDaPCa2b, and LnCAP, and breast cancer cell lines MDA-MB-231 and MCF-7 were obtained from ATCC and grown as directed.

Genomic Sequencing of the First V-set Domain of SIGLEC12—The first V-set domain of *SIGLEC12* was sequenced in 90 human individuals from diverse geographic origins. Characterization of the R122C mutation and the frequency of the frameshift mutation in globally distributed humans was performed on genomic DNA that was either kindly donated by Dr. Stephen Warren (Emory University) or Dr. Michael Hammer and the Y Chromosome Consortium (University of Arizona) (20) or obtained from the peripheral blood of healthy donors. The samples were amplified using primers 5′-UTR and 3′Chi3D as described (15), using the Roche long template PCR kit, with cycling conditions of (*a*) 94 °C denature for 10 s, (*b*) five cycles of annealing at 68 °C for 30 s, (*c*) five cycles of annealing at 65 °C for 30 s, and (*d*) 20 cycles of annealing at 62 °C for 30 s with a 3-min extension at 68 °C for each cycle, increasing by 20 s each cycle. The samples were sequenced on an ABI Prism 310 genetic analyzer using primer 5′-UTR and primer 5′-CTTTG-GCCTCTCTTGGAGCC-3′, and analyzed with Sequencher (Gene Codes Corporation, Ann Arbor, MI).

Analysis of Frameshift Mutation on Prostate Cancer and Control Samples—The frameshift mutation polymorphism was also analyzed in 242 lymphocyte DNA samples from patients with prostate cancer and on 244 control DNA samples, from men who had been screened for prostate cancer. These genomic samples, isolated from whole blood using the QIAamp DNA blood maxi kit (Qiagen), were a part of a cohort that has been previously used to assess risk for PCa using various genetic markers (21). To determine the genomic status, the primers 5′-ACCCCTGCTCTGTGGGAGAGT-3′ (forward) and 5′-AGGATCAGGAGGGGCATCCAAGGTGC-3′ (reverse) were used to amplify a 570-bp region. The PCR conditions were 95 °C for 3 min (94 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s) for 35 cycles. The amplified DNA was sequenced using the forward primer. The allele frequencies (presence/absence of frameshift) among the case control groups were compared using the chi-squared test. The odds ratio and its 95% confidence interval were estimated by unconditional logistic regression as a measure of the associations between genotypes and PCa risk, using R statistical software. Age was used as a covari-

ate. Statistical tests were two-sided, and significance was set at $p < 0.05$.

Analysis of Frameshift Mutation on Prostate Cancer Paraffin Sections—Genomic DNA extracted from a set of 50 paraffin-embedded prostate cancer samples was used to correlate Siglec-12 expression by immunohistochemistry with its genomic status (*i.e.* presence or absence of frameshift). Paraffin tissue blocks were acquired as explained in the immunohistochemistry section, and genomic DNA was isolated. Paraffin blocks were cut in 25-μm slices. Two such slices were used for each DNA prep. The slices were washed once with xylene and twice with ethanol (96–100%). The samples were air-dried, and genomic DNA was extracted using the Macherey-Nagel nucleospin genomic DNA extraction kit from Clontech. PCR was carried out using the forward primer, 5′-CAATGCA-GAAGTCCGTGACGGTGCAGG-3′, and the reverse primer, 5′-AGGATCAGGAGGGGCATCCAAGGTGC-3′, to obtain a 200-base pair product. PCR conditions were one cycle (94 °C for 2 min), 11 cycles (94 °C for 10 s, 62 °C for 30 s, 1 °C drop in annealing temperature per cycle, 68 °C for 45 s), and 24 cycles (94 °C for 15 s, 52 °C for 30 s, 68 °C for 45 s, 20-s increase in extension time per cycle) followed by 68 °C for 10 min. Sequencing was done using the reverse primer, 5′-GGAC-ATGTGTCCCGTCTCAGCCGTGC-3′.

Mouse Monoclonal Antibody against Human Siglec-XII—A fusion protein Siglec-XII-Fc including the first three Ig-like domains of human Siglec-XII and the human IgG Fc domain was prepared as described (15). The fusion protein was used to immunize mice to generate monoclonal antibodies (BD Pharmingen). Two final clones 1130 and 276 were obtained. Specificity was confirmed by lack of cross-reactivity with Siglec-7-Fc. Studies were done using a mixture of the two clones or clone 276 or 1130 alone.

Immunohistochemistry Studies—Anonymized archived paraffin blocks were obtained from the Cooperative Human Tissue Network or from the Veterans Affairs Medical Center (La Jolla, CA). Paraffin sections were deparaffinized, rehydrated, and blocked for endogenous peroxidases and endogenous biotin. Epitopes were revealed using heat-induced antigen retrieval at pH 6.0 in citrate buffer. The slides were then incubated with the mouse anti-Siglec-XII antibodies. Bound antibody was detected using the CSA kit from DAKO following the manufacturer's instructions. Digital photographs were taken using an Olympus BH2 microscope with an Olympus digital camera. The images were organized using Adobe Photoshop.

Flow Cytometry—The cells were stained with anti-Siglec-XII monoclonal antibody 1130 or 276 or a mixture of the two to probe for Siglec-XII expression. The cells were lifted using 10 mM EDTA and washed with 1% BSA-PBS. 500,000 cells were aliquoted and incubated with 1 μg of anti-Siglec-XII monoclonal antibody 1130 or 276 or a mixture of the two for 1 h on ice. The cells were washed with 1 ml of 1% BSA-PBS and incubated with 1:100 GAM-RPE (Caltag) for 30 min on ice in dark. The cells were washed and resuspended in 400 μl of 1% BSA-PBS and read on FACSCalibur flow cytometer using Cellquest. The data were analyzed using FlowJo.

Establishing Stable Transfectants of Prostate Carcinoma Cells Expressing Siglec-XII—PC-3 cells were transfected with PvuI linearized *hSIGLEC12*-pcDNA3.1(−) or empty pcDNA3.1(−) in six-well plates using Lipofectamine 2000 (Invitrogen). 48 h after transfection, the cells were trypsinized and grown with 800 μg/ml G418. After growing ~1 month, expression of Siglec-12 was determined by flow cytometry. Four independent cell lines from four independent transfections were obtained: two with *hSIGLEC12* and two with empty pcDNA 3.1(−).

Microarray Gene Expression Profile Comparison of Transfected and Sham-transfected Prostate Carcinoma Cells—RNA was isolated from the stably transfected PC-3 cell lines using the RNeasy mini kit (Qiagen). cDNA was synthesized and hybridized to Genechip Human Genome U133 Plus 2.0 Array (Affymetrix). Quantity and quality of final total RNA were examined using a nanodrop and with the RNA QC-Standard Bioanalyzer (Agilent). 5 μg of total RNA was used for cDNA synthesis, followed by *in vitro* transcription to incorporate biotin labels, and subsequent hybridization to Genechip human genome U133 Plus 2.0 array (Affymetrix) was performed by the GeneChip Microarray Core (University of California at San Diego) as described in the Affymetrix GeneChip protocol. The U133 Plus 2.0 interrogates ~54,000 probe identification codes. The raw expression values were normalized using DNA chip analyzer built May 8, 2008 (dChip), which is a Windows software package for probe level analysis of gene expression microarrays (22). Before further processing, the transcripts were filtered to 32,000 transcripts using the standard deviation for discrimination. The data were analyzed using rank products implemented within the Bioconductor project and the R program software (R is available as free software under the terms of the Free Software Foundation GNU General Public License). Heat maps were done using the dChip software. Functional analysis of genes was done using Ingenuity Pathways Analysis from Ingenuity Systems, Inc.

Growth of Stable Prostate Carcinoma Transfectants as Tumors in Nude Mice—1.5×10^6 stably transfected *hSIGLEC12*-PC3 and pcDNA3.1(−)-PC3 cells, suspended in 150 μl of PBS, were injected subcutaneously into the right and left flanks, respectively, of male athymic/nude mice aged 12–16 weeks. Tumor length, breath, and height were measured at regular intervals, and volume was calculated using the standardized formula $(\pi/6) \times$ length $\times$ width $\times$ height. The mice were sacrificed, and tumors were extracted. Histopathological analysis included staining for CD45 and CD31. All of the mice experiments were approved by the University of California at San Diego Institutional Animal Care and Use Committee.

Anti-Siglec-XII Antibody-mediated Endocytotic Toxin Delivery into Prostate Cancer Cells—2500 stably transfected *hSIGLEC12*-PC3 cells were plated in 96-well plates in growth medium. Next day, different amounts of MabZAP (Advanced Targeting Systems, San Diego, CA) (0–200 ng) and anti-Siglec-12 antibodies 1130 or 276 (0–130 ng) were incubated together in cell growth medium for 30 min on ice. After aspirating media from the cells, the MabZAP-antibody mixture was added. Each combination of MabZAP-anti-Siglec-XII antibody was repeated in triplicate. The cells were further incubated for 3 days. The number of viable cells in each well was determined colorimetrically using the CellTiter 96® AQ$_{ueous}$ One Solution cell proliferation assay (Promega).

RESULTS

Mutation of Essential Arginine of Siglec-XII Is Fixed in Humans—In our initial *SIGLEC12* paper (15), we reported that the few humans sequenced had a homozygous mutation that changed the Arg residue in the V-set Ig-like domain required for Sia recognition into a cysteine. We now asked whether this R122C mutation is fixed in human populations. We analyzed a set of human DNA samples representing globally diverse populations and found this mutation to be homozygous in all 90 humans tested. On the other hand, we previously reported that this R122C mutation was not present in multiple samples from "great apes," including bonobos, gorillas, chimpanzees, and orangutans (15). Thus, this functionally inactivating mutation apparently occurred prior to the common ancestor of all modern humans and sometime after the split of the hominin lineage from that of the common ancestor of chimpanzees and bonobos.

Polymorphic Pseudogenization of SIGLEC12 in the Human Lineage—The R122C missense mutation universal to humans still leaves an open reading frame encoding a full-length Siglec-XII protein. However, we have now found additional mutations in many humans that would result in complete pseudogenization of *SIGLEC12.* As shown in Fig. 1, the most common mutation was a single nucleotide insertion in the first V-set exon, which if translated would result in a truncated protein of only 115 amino acids. This insertion mutation, a guanidine, occurs within a string of three other guanidines between base pairs 194 and 197 of the nucleotide sequence relative to the standard reference human genome sequence (23), so we cannot know precisely which guanidine is the actual insertion. Evaluating the frequency of this mutation (*SIGLEC12P*, where *P* is for pseudogene) in human populations (Fig. 2), we found a global allele frequency of 0.58, with allele frequencies ranging from 0.38 in sub-Saharan Africa to 0.86 in Native American populations. Global genotype frequencies are 0.40 for homozygous frameshifted genotypes (designated G+/+), and 0.24 and 0.36 for homozygous wild type (G−/−) and heterozygotes (G+/−), respectively. Thus, following the fixation of the human universal R122C mutation prior to the origin of humans, the *SIGLEC12* gene has undergone an additional polymorphic pseudogenization in humans. In this regard, it is of interest that recent studies by others reported evidence for a selective sweep at the *SIGLEC12* locus in humans (24).

Macrophage and Epithelial Siglec-12 Expression in Some Humans and All Chimpanzees—A newly generated mouse monoclonal antibody was shown to recognize both human and chimpanzee Siglec-12-Fc protein in enzyme-linked immunosorbent assays and to have no cross-reactivity to Siglec-7-Fc. Flow cytometry showed no staining in human peripheral blood leukocytes and low levels of monocyte staining in some chimpanzee samples.[7] Next, we looked for expression of Siglec-XII on paraffin sections of human and chimpanzee tissues using

---

[7] N. Mitra, K. Banda, T. Altheide, T. Angata, and A. Varki, unpublished observations.

```
hSIGLEC12P  atg cta ctg ctg ctg cta ctg ct- --- --g cca ccc ctg ctc tgt ggg aga gtg ggg gct  54
hSIGLEC12   ... ... ... ... ... ... ... ..- --- --. ... ... ... ... ... ... ... ... ... ...  54
cSIGLEC12   ... ... ... ... ... ... ... ..a ctg ct. ... ... ... ... ... ... ... ... ... ...  60

hSIGLEC12P  aag gaa cag aag gat tac ctg ctg aca atg cag aag tcc gtg acg gtg cag gag ggc ctg 114
hSIGLEC12   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... 114
cSIGLEC12   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... 120

hSIGLEC12P  tgt gtc tct gtg ctt tgc tcc ttc tcc tac ccc caa aat ggc tgg act gcc tcc gat cca 174
hSIGLEC12   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... 174
cSIGLEC12   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... .a. ... ... ... ... 180

hSIGLEC12P  gtt cat ggc tac tgg ttc cgg ggc agg gga cca tgt aag ccg gaa cat tcc agt ggc cac 234
hSIGLEC12   ... ... ... ... ... ... ..- ... ... ... ... ... ... ... ... ... ... ... ... ... 233
cSIGLEC12   ... ... ... ... ... ... ..- ... ... ... ... ... ... ... .g. ... ... ... ... ... 239

hSIGLEC12P  aaa caa ccc agc tcg agc agt gca gga gga gac tcg gga ccg att cca cct cct tgg gga 294
hSIGLEC12   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... 293
cSIGLEC12   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... 299

hSIGLEC12P  ccc aca gaa caa gga ttg tac cct gag cat cag aga cac cag aga gag tga tgc agg gac 354
hSIGLEC12   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... 353
cSIGLEC12   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... 359

hSIGLEC12P  ata cgt ctt ttg tgt aga gag agg aaa tat gaa atg gaa tta taa ata tga cca gct ctc 414
hSIGLEC12   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... 413
cSIGLEC12   ... ... ... ... .c. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... 419

hSIGLEC12P  tgt gaa tgt gac agc gtc cca gga cct act gtc aag ata cag gct gga ggt gcc aga gtc 474
hSIGLEC12   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... 473
cSIGLEC12   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... 479

hSIGLEC12P  ggt gac tgt gca g 487
hSIGLEC12   ... ... ... ... . 486
cSIGLEC12   ... ... ... ... . 492
```

FIGURE 1. **Human-specific changes in the *SIGLEC12* gene.** The alignment of exon 1 of human and chimpanzee *SIGLEC12* is shown. *hSIGLEC12P* is the human allele with an extra G, resulting in a frameshift. *hSIGLEC12* is the wild type human allele. *cSIGLEC12* is the chimpanzee *SIGLEC12* allele. The signal peptide-coding region is *underlined* with a *dashed line*. The region with the extra G is *underlined*. The mutation from CGT (in chimp) to TGT (in humans) (*boxed*) results in an R122C mutation, which eliminates sialic acid binding. All of the human samples tested had the cysteine and thus do not bind sialic acid. *hSIGLEC12P* results in a premature termination codon, indicated by ***. Residues different from *hSIGLEC12* have a *gray background*.
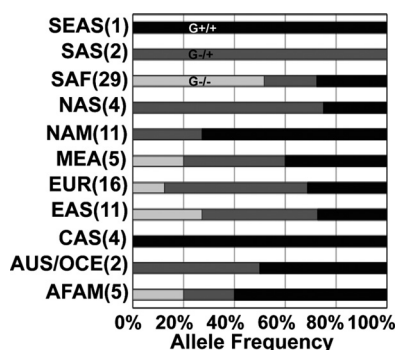
FIGURE 2. **Frequency of frameshift inactivation of *SIGLEC12* in human populations.** The number of individuals studied in each group is shown in parentheses. *SEAS*, Southeast Asian; *SAS*, South Asian; *SAF*, sub-Saharan African; *NAS*, North Asian; *NAM*, Native American; *MEA*, Middle Eastern; *EUR*, European; *EAS*, East Asian; *CAS*, Central Asian; *AUS/OCE*, Australia/Oceania; *AFAM*, African American. The *black bars* represent the percentage of homozygous frameshifted individuals, the *medium gray bars* represent the heterozygous, and the *light gray bars* represent the homozygous WT.

immunohistochemical methods. Cell pellets of 293T cells transfected with full-length Siglec-XII were used as positive controls and with empty vector as negative controls for all immunohistochemistry experiments (supplemental Fig. S1). Strong macrophage and epithelial expression of Siglec-12 was seen in the chimpanzee prostate, pancreas, kidney, and stomach (Fig. 3*A*). Fig. 3*B* shows expression of Siglec-12 on follicular dendritic cells in primate tonsil. As expected, not all humans showed expression because of the frameshift insertion mutation. Among humans who did have an intact ORF, expression was observed on macrophage-like cells in the tonsils (Fig. 3*C*) and also on epithelial cells of the kidney tubules and prostate epithelium. A possible explanation for the generally weaker

expression in humans over chimpanzees even with the homozygous WT allele is that the extra Cys residue arising from the R122C mutation in the human protein might have a destabilizing effect on the formation of disulfide bonds during protein folding, resulting in lesser expression of mature properly folded protein. Regardless of the reason, these data confirm and extend our earlier observation using a chicken polyclonal antibody that human Siglec-XII is expressed on epithelial cells (15).

Expression of Siglec-XII in Genotypically Positive Human Carcinomas and Carcinoma Cell Lines—Given the expression of Siglec-XII in certain epithelial cells, we next analyzed its expression on human carcinomas (cancers of epithelial origin). Siglec-XII expression was indeed seen in many human prostate carcinoma specimens and also occasionally in breast carcinoma and in melanoma. In keeping with easily detectable expression in normal prostate epithelium, we found clear expression in prostate carcinomas (PCa), and we studied this cancer further. In the initial 50 PCa samples studied, there was a genotype to phenotype correlation, with no expression in samples in which both alleles were *SIGLEC12P* or frameshifted (Fig. 4*A* and Table 1). Next, we looked for Siglec-XII expression in human breast and prostate carcinoma cell lines. Using flow cytometry, we found that although MDaPCa2b and LnCAP (PCa) and MCF-7 (breast cancer) lines were positive, MDaMb231 (breast cancer) and PC-3 (PCa) showed no expression (Fig. 4*B*). As expected, the lack of expression correlated with the presence of the homozygous genomic *SIGLEC12P* allele, *i.e.* the frameshift.

Siglec-XII Expression in a Genotypically Null Prostate Carcinoma Cell Line Alters Expression of Multiple Genes Associated with Carcinoma Progression—Given the prominent expression of Siglec-XII in some human PCa and its absence in others with
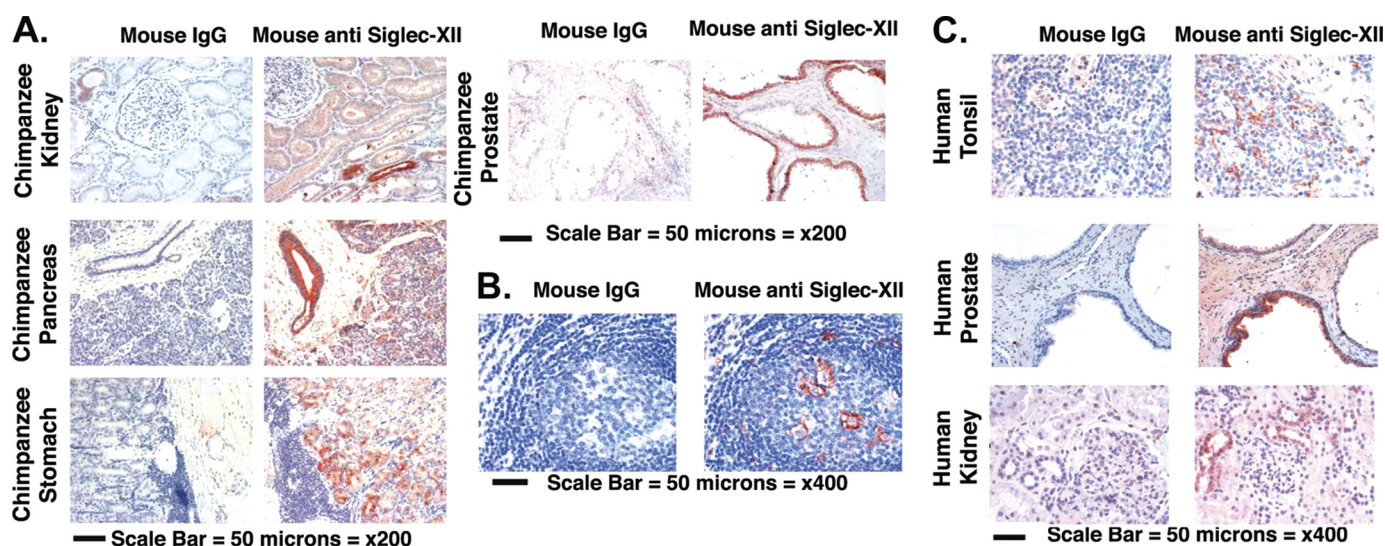
FIGURE 3. **Expression of Siglec-12/XII in human and chimpanzee tissues.** *A*, strong epithelial expression of Siglec-12 in chimpanzee kidney, pancreas, stomach, and prostate. In the *mouse IgG panel*, nonspecific mouse IgG is used instead of anti-Siglec-XII. *B*, Siglec-12 expression on primate follicular dendritic cells. *C*, epithelial expression of Siglec-XII in human tonsil, prostate, and kidney.
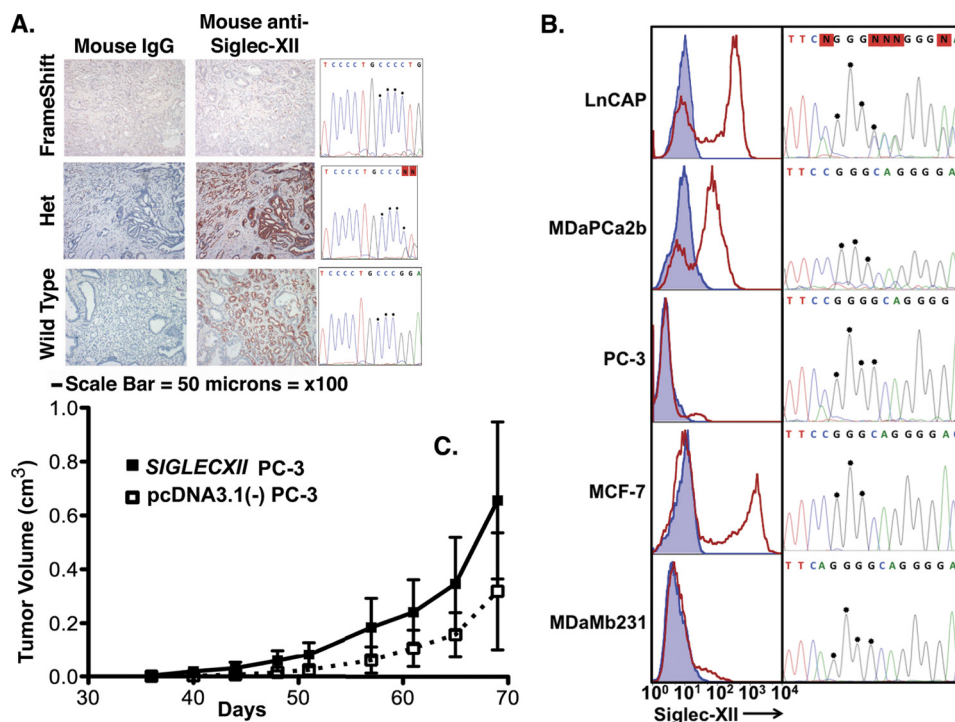

FIGURE 4. **Expression of Siglec-XII in human prostate carcinoma tissues.** *A*, the *left panels* are the mouse isotype controls. The *middle panels* show serial sections of the same tissue sample stained with anti-Siglec-XII mAb 276. The *right panels* show the sequencing chromatogram (using reverse sequencing primers) of the frameshift region. The heterozygous and WT samples show Siglec-XII staining, whereas the homozygous frameshifted sample does not. *B*, flow cytometry detection of Siglec-XII on human carcinoma cell lines and their sequencing chromatogram. PC-3 and MDaMb231 are homozygous for the frameshift and do not stain for Siglec-XII. MDaPCa2b and MCF-7 show expression and have WT alleles. LnCAP stains for Siglec-XII and upon sequencing is found to be heterozygous. *C*, effect of Siglec-XII expression on prostate cancer growth in nude mice. Comparison of tumor volumes produced by PC-3 cells in the absence and presence of Siglec-XII. By injecting Siglec-XII-positive and -negative cells on either flank of the same mice, we were able to obtain paired statistical comparisons (see text for details).

TABLE 1

Correspondence between genomic status of *SIGLEC12* and protein expression by immunohistochemistry (IHC) in prostate cancers

Genotype	Negative by IHC	Positive by IHC
G+/+	14	0
G+/−	1	32
G−/−	0	3

the homozygous *SIGLEC12P*, we wondered whether this had any functional consequences. As a first step toward addressing this question, we transfected the genotypically null PCa cell line, PC-3 with the intact human *SIGLEC12* cDNA in the pcDNA3.1(−) expression vector, and established two lines with stable expression of cell surface Siglec-XII, using G418 selection (supplemental Fig. S2*A*; two empty vector transfected cell

TABLE 2
SIGLEC12 pseudogene status does not affect prostate cancer risk

Genotype/model	Cases (n = 242)	Controls (n = 244)	Odds ratio* (95% confidence interval)	P
	%	%		
G+/+	123 (51)	120 (49)		
G+/−	89 (37)	97 (40)	0.90 (0.61–1.31)	0.6
G−/−	30 (12)	27 (110)	1.09 (0.61–1.94)	0.6
G+/− or G−/− *versus* G+/+ (G-Dom)	119	124	0.94 (0.66–1.34)	0.7
G−/− *versus* G+/+ or G+/− (G-Rec)	30	27	1.14 (0.66–1.99)	0.6

* adjusted for age.

lines did not give a positive signal). We next studied total mRNA from both cell lines by microarray, looking for gene expression differences. As shown in supplemental Fig. S2*B* we found limited but significant changes in gene expression between the *SIGLEC12* and empty vector transfected cell lines. In all, 67 transcripts were identified as being down-regulated over the control upon Siglec-XII expression, at a false positive rate of 15%. Only *MAP2K5* was up-regulated.

Interestingly, genes affected by Siglec-XII expression were involved in carcinoma progression such as matrix metalloproteinase1 (*MMT1*), growth differentiation factor 15 (*GDF-15/MIC-1*), and *RUNX2* (supplemental Fig. S2*B* and supplemental Table S1). A number of genes associated with cellular migration and adhesion such as *CDH1*, *FGA*, *GDF15*, *IGFBP5*, *ITGB4*, *ITGB8*, *MMP1*, *RUNX2*, *S100A9*, *SDC2*, *TFF3*, and *TGFA* were also down-regulated (supplemental Table S1).

Siglec-XII Expression Alters Human Prostate Carcinoma Growth in Nude Mice—PC-3 cells stably transfected with human *SIGLEC12* cDNA or with empty vector were injected into the two flanks of nude mice. Tumor growth was followed for 69 days. As shown in Fig. 4*C*, *SIGLEC12* expressing PC-3 cells showed a difference in tumor volumes when compared with empty vector transfected cells. A one-tailed *t* test comparing the tumor volumes of *SIGLEC12* PC-3 to pcDNA PC-3 tumors within each mouse showed a significant difference in mean size in four of five mice at $p < 0.05$ ($p = 0.0492$, 0.036, 0.0196, and 0.0134, respectively). Thus, the presence of the Siglec-XII resulted in significant increase in tumor volume. The same trend was seen in a second experiment in which the Siglec-XII-positive cells produced tumors and the control transfected cells did not. Histological analysis of these tumors (done by CD31 and CD45 staining) showed that Siglec-XII-positive tumors had decreased inflammatory cell infiltration.

SIGLEC12 Genomic Status Does Not Affect Prostate Cancer Incidence Risk—We next asked whether the presence or absence of an open reading frame in the *SIGLEC12* gene affected the risk of PCa in humans. The presence of the frameshift allele was determined in a European-American sample set of 242 PCa cases and 244 age-matched cancer-free controls. The PCa samples were obtained from men who were screened for PCa and consisted mainly of early stage cancer, although the samples were not necessarily selected on this basis. As shown in Table 2, we found no significant association between *SIGLEC12* genomic status and PCa incidence risk. No significant case/control difference in allele frequency distribution was observed ($p = 0.9$). Logistic regression analysis confirmed that the presence of the frameshift was not associated with any change in the risk of PCa ($p > 0.05$; Table 2).

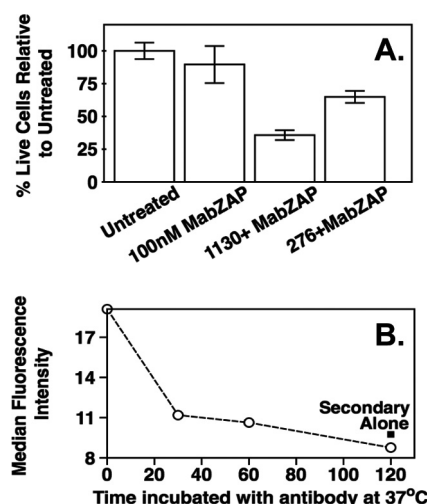FIGURE 5. **Targeting of carcinomas cells by a Siglec-XII antibody.** *A*, antibody-mediated endocytosis of Siglec-XII resulted in MabZAP (goat anti-mouse toxin conjugate) internalization as well and resulted in reduced cell viability in Siglec-XII expressing PC-3 cells. *B*, internalization of Siglec-XII in MDaPCa2b cells upon incubation with mAb276 for 2 h at 37 °C. The *closed square* is the level of nonspecific staining of the cells treated with secondary antibody alone.

Monoclonal Antibody Binding Induces Internalization of Cell Surface Siglec-XII and Targeting of a Toxin into Human Carcinoma Cells—Previous studies have shown that Siglec-3, -5, and -9 undergo rapid internalization upon cross-linking with antibodies (25–27). If a toxin is attached to such antibodies, toxin internalization also occurs, and this results in cell death (27). We wanted to see whether Siglec-XII could be utilized to deliver a toxin into cells and thus cause cell death. For this, we used MabZAP, a goat anti-mouse antibody that was conjugated to the toxin saporin (28). Saporin is a ribosome inactivating protein from the seeds of *Saponaria officinalis*. Internalization of the Siglec would also deliver the mAb-MabZAP complex into the cell. To induce cell death, we incubated cells in a 96-well plate with primary antibody and MabZAP for 72 h. Following this, cell viability was determined. Wells with both antibody and the MabZAP had only 35% of cells alive. Wells with MabZAP alone had nearly as many cells alive, as wells with no treatment and wells with antibody alone showed some cell killing (Fig. 5*A*). Although we only performed toxicity studies on the stably transfected *SIGLEC12* PC-3 cell line as a proof of principle experiment, we looked for Siglec-XII internalization in the prostate cell line MDaPCa2b (Fig. 5*B*), which natively expressed Siglec-XII, and saw that the Siglec-XII was completely internalized. Because plant and bacterial protein toxins are capable of killing cells at very low intracellular concentrations, we surmised that under the right conditions, cell death

would occur, even with small amounts of surface Siglec-XII on carcinoma cells.

DISCUSSION

The transmembrane position of Siglecs, along with their preference for binding sialic acids, the outermost sugars on cell surface glycans, make them an important part of the cell-cell communication system both within an organism and in between organisms such as host and pathogens. Previous work comparing *SIGLEC* loci in mice, rats, baboons, chimpanzees, and humans showed that CD33-related Siglecs are evolving rapidly, especially in the Sia-binding domain (11). This rapid evolution could be in response to slight changes in the Sia molecular structure in the host itself and/or on surface of pathogens. Apart from *SIGLEC12*, other Siglecs that have undergone human-specific changes in functional gene status, expression, or ligand binding include *SIGLEC1*, *SIGLEC5/14*, *SIGLEC6*, *SIGLEC7*, *SIGLEC9*, *SIGLEC11*, *SIGLEC13*, and *SIGLEC16* (29). Such large scale differences within a single class of gene indicate strong selection pressure on CD33-related *SIGLEC* loci in humans (29, 30).

Although Siglec-12 in chimpanzees and the reverse mutated C122R human protein are shown to bind Sias (15), its natural ligands are unknown. In humans, Siglec-XII with a Cys instead of the essential Arg is expressed as a non-Sia binding full-length protein with cytosolic signaling domains intact. It is unclear whether the protein can mediate normal downstream signaling in the absence the essential Arg, but our microarray experiment shows a large scale gene down-regulation upon stable expression. This suggests that Siglec-XII retains some downstream signaling activity. Despite the R122C mutation, it is possible that Siglec-XII weakly recognizes sialic acids that results in the observed signaling effect. As shown in a previous work by others (31), Cos-7 cells expressing Siglec-XII bind sialylated red blood cells marginally better (~4–5-fold) than nonexpressing Cos-7 cells, although this binding was undetectable when compared with a reverse mutated human C122R Siglec-XII (15). This phenomenon might arise from other interactions between sialic acids and Siglecs, such as that between the sialic acid glycerol group and a conserved hydrophobic amino acid as seen in crystal structures of Siglec-5 (Tyr-133), Siglec-7 (Trp-132), and Sialoadhesin (Trp-106) (32–34). This hydrophobic amino acid is conserved in all Siglecs (4). In addition there are other interactions between the sialic acid and the protein backbone (35). It has also been shown that certain Siglecs, such as Siglec-11, Siglec-6, and myelin-associated glycoprotein, still retain some Sia binding in the absence of the essential Arg (36–38). Siglec-XII might also have some residual sialic acid binding that might be enough to trigger downstream signaling.

In addition to the R122C mutation, if the frameshift mutation is also present, there is no Siglec-XII expression, as confirmed by comparing genotype with immunohistochemistry. Thus, individuals homozygous for *SIGLEC12P* will lack any Siglec-XII-mediated signaling. This suggests a potential biochemical difference between genotypically different individuals and thus a potential target for natural selection.

Unlike other Siglecs, Siglec-XII is expressed on human and chimpanzee epithelia, including human breast and prostate carcinomas. Epithelial carcinomas are rare in chimpanzees but occur at high frequencies in humans (19, 29, 39, 40). One might hypothesize that this difference results in part from the inability of human Siglec-XII to bind sialylated ligands, giving less efficient ITIM signaling. In contrast, chimpanzee Siglec-12 maintains sialic acid binding, thus giving full ITIM signaling and better inhibition. In keeping with this, we found it difficult to stably express chimpanzee Siglec-12 and human C122R Siglec-XII, *i.e.* the human Siglec reverse mutated to restore Sia binding.[8] In contrast, expressing the non-Sia binding full-length human Siglec-XII allows stable expression.

Regardless of these speculations, microarray data show that in presence of Siglec-XII, almost all of the observed changes in gene expression (with one exception) are down-regulations. Functional analyses of the affected genes show that many of them are associated with cellular movement and cancer (supplemental Table S1). Some genes known to be up-regulated in prostate cancers such as *GDF15*, *TTF3*, *RUNX2*, *ITGB4*, *MMP1*, and *S100A9* are down-regulated in Siglec-XII-expressing PC-3 cells. *GDF15* (alternatively MIC-1) expressed by PCa cells can disrupt cell adhesion and is associated with bone homeostasis (41, 42). *TFF3* is elevated in PCa, although its function is not clear (43). *RUNX2*, a transcription factor, has an anti-apoptotic effect in early cancer but contributes to subsequent bone metastasis (44). *ITGB4* overexpression is also associated with epithelial cancers and their proliferation (45). *IGFBP5* accelerates androgen independence of PCa cells *in vitro* and in mouse models after androgen withdrawal (46). Reducing *IGFBP5* levels reduces cell proliferation after androgen withdrawal.

Another down-regulated gene is *CDH1*, which codes for E-cadherin, important for adhesion and normal epithelium maintenance. Loss of E-cadherin expression is known to increase the likelihood of metastasis in epithelial carcinomas (47, 48). Other studies show strong expression of E-cadherin in metastatis of PCa (49, 50). The authors hypothesize that after a temporary loss of expression required for metastasis, the cancer cells regain E-cadherin expression for tumor establishment. Thus, the overall effect of CDH1 down-regulation is unclear.

The presence of the inhibitory ITIM motif in Siglec-XII could be the likely cause of the observed down-regulation. A previous study (31) showed that Siglec-XII (called S2V in that paper) binds to the tyrosine phosphatases SHP-1 and SHP-2 in a phosphorylation-dependent manner (co-immunoprecipitation worked only in the presence of pervanadate). Despite the R122C mutation, Siglec-XII was phosphorylated by the kinase c-Src. The authors also showed that SHP-1 interaction with Siglec-XII was mediated by the ITIM domain in Siglec-XII. Since Siglec-XII has been shown to associate with SHP-1 and SHP-2 and because the interaction of SHP-1 is ITIM-dependent, it is reasonable to conclude that the observed gene down-regulation is due to the inhibitory ITIM domain. It is theoretically possible that these observed effects could vary, based on the amount of Siglec-XII expression. However, our two independently stable lines with different levels of Siglec-

---

[8] N. Mitra, T. Angata, and A. Varki, unpublished observations.

XII expression (differing by almost one order of magnitude, see supplemental Fig. S1) still show a similar pattern of gene down-regulation.

U.S. population data show that PCa is the most prevalent form of cancer diagnosed among men and the second leading cause of cancer-related deaths (51). Lack of symptoms during the early stages makes periodic screening essential for timely treatment of this disease. The most common screening method is the prostate-specific antigen test (52), which has been shown to have a small effect on the mortality rate while picking up false positives (53, 54). Thus, there is a search for better markers to differentiate aggressive cancers from nonmetastatic ones. For example, transcriptome sequencing to detect gene fusions (55) has resulted in discovery of gene fusions in prostate cancer (56), especially the *TMPRSS2-ERG* gene fusions that occur in 50% of prostate cancers. The resulting overexpression of a chimeric fusion transcript encodes a truncated *ERG* product that can be detected by a rabbit anti-*ERG* monoclonal antibody (57). In addition deletions of the *PTEN* (phosphatase and tensin homolog) gene have been shown to correlate with aggressive and hormone resistant prostate cancer (58, 59). A review of recent advances in prostate cancer genetics can be found in Shen and Abate-Shen (60).

Our own analysis of *SIGLEC12* genomic status in PCa (mostly early stage PCa obtained during screening) and normal samples showed a similar allelic distribution in both groups, suggesting that the presence or absence of the frameshift does not predict the incidence of PCa. Indeed, this region of the genome was not picked in a recent large genome-wide association study that reported seven new prostate cancer susceptibility loci (61). However, based on our gene expression and tumor growth data, *SIGLEC12* genomic and expression status could still have an effect on the long term progression of the cancer and the eventual severity of the disease. Indeed, although the majority of patients diagnosed with early stage prostate cancer will develop microscopic disease with advancing age, only a minority will have metastatic disease (62).

For an initial indication of the role of Siglec-XII in prostate cancer development, we monitored the growth of human prostate cancer cells stably transfected with human *SIGLEC12* or empty vector in nude mice. Cells expressing Siglec-XII showed a significant growth advantage over nonexpressing cells. This small growth difference over 70 days (time period of the mouse experiment) could become pertinent over many years, the usual time that it takes for a clinically significant prostate cancer to develop in humans. It is currently unknown whether this result extends to humans, but this is testable by association studies on large cohorts with known outcomes.

A recent study on polymorphic nonsense single-nucleotide polymorphisms in the human genome found *SIGLEC12* to be one of the 12 outliers, among the 167 genes studied (24). By calculating a statistical parameter called F_{ST}, which is a measure of population differentiation, it was found that a majority of nonsense single-nucleotide polymorphisms have a low F_{ST} value (< 0.09), suggesting that these single-nucleotide polymorphisms were at least mildly deleterious. *SIGLEC12*, on the other hand, had a high F_{ST} of 0.221 and a relatively high heterozygosity of 0.317, suggesting a balancing selection or a

selective sweep. It is unlikely that PCa, which usually afflicts individuals in the postreproductive period, would affect this selective sweep. Future studies will work to uncover the selective target in or around the *SIGLEC12* region on chromosome 19.

All of the experiments in this study to understand the possible functions of Siglec-XII including microarray, tumor growth in nude mice, and cell killing experiments, were done in PC-3 cells stably expressing Siglec-XII under antibiotic selection. Although the effects observed in cells overexpressing Siglec-XII may not be relevant under normal physiological conditions in cancer cells, where Siglec-XII expression levels might be low, we would like to point out that the expression levels of the protein in Siglec-XII PC-3 cells and the three cancer cell lines natively expressing Siglec-XII are within range of each other (supplemental Fig. S3). Thus, the studies presented here are still relevant under conditions of native Siglec-XII expression. In the future, we aim to study the effect of Siglec-XII more directly by knocking down its expression in endogenously expressing cell lines.

Human Siglec-XII is expressed at low levels in the epithelium. Because 40% of human individuals are homozygous for *SIGLEC12P* and do not show any obvious phenotype, it can be concluded that this gene is not essential for survival, although its presence might become pertinent under special conditions. This apparently more specialized function makes it a target for *in vivo* experimentation. In addition, Siglec-XII has a low and very narrow expression profile (it is expressed in only a very few cell types), thus making it a suitable candidate for antibody-mediated targeting for drug delivery to cancer cells expressing it. Because Siglec-XII can be internalized upon antibody binding, a simultaneous delivery of a toxin attached to the antibody could be a viable approach to future cancer therapy.

REFERENCES

1. Angata, T., and Varki, A. (2002) *Chem. Rev.* **102,** 439–469
2. Schauer, R. (2009) *Curr. Opin. Struct. Biol.* **19,** 507–514
3. Chen, X., and Varki, A. (2010) *ACS Chem. Biol.* **5,** 163–176
4. Varki, A., and Angata, T. (2006) *Glycobiology* **16,** 1R–27R
5. Crocker, P. R., Paulson, J. C., and Varki, A. (2007) *Nat. Rev. Immunol.* **7,** 255–266
6. von Gunten, S., and Bochner, B. S. (2008) *Ann. N.Y. Acad. Sci.* **1143,** 61–82
7. Lopez, P. H., and Schnaar, R. L. (2009) *Curr. Opin. Struct. Biol.* **19,** 549–557
8. Varki, A., and Crocker, P. R. (2009) in *Essentials of Glycobiology* (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., and Etzler, M. E., eds) pp. 459–474, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
9. Cao, H., and Crocker, P. R. (2011) *Immunology.* **132,** 18–26
10. Cornish, A. L., Freeman, S., Forbes, G., Ni, J., Zhang, M., Cepeda, M., Gentz, R., Augustus, M., Carter, K. C., and Crocker, P. R. (1998) *Blood* **92,** 2123–2132
11. Angata, T., Margulies, E. H., Green, E. D., and Varki, A. (2004) *Proc. Natl.*

Acad. Sci. U.S.A. **101,** 13251–13256

12. Altheide, T. K., Hayakawa, T., Mikkelsen, T. S., Diaz, S., Varki, N., and Varki, A. (2006) *J. Biol. Chem.* **281,** 25689–25702

13. Cao, H., Lakner, U., de Bono, B., Traherne, J. A., Trowsdale, J., and Barrow, A. D. (2008) *Eur. J. Immunol.* **38,** 2303–2315

14. Cao, H., de Bono, B., Belov, K., Wong, E. S., Trowsdale, J., and Barrow, A. D. (2009) *Immunogenetics* **61,** 401–417

15. Angata, T., Varki, N. M., and Varki, A. (2001) *J. Biol. Chem.* **276,** 40282–40287

16. Avril, T., Attrill, H., Zhang, J., Raper, A., and Crocker, P. R. (2006) *Biochem. Soc Trans.* **34,** 1024–1027

17. Varki, A., and Altheide, T. K. (2005) *Genome Res.* **15,** 1746–1758

18. Go, Y., and Niimura, Y. (2008) *Mol. Biol. Evol.* **25,** 1897–1907

19. Varki, N. M., Strobert, E., Dick, E. J., Jr., Benirschke, K., and Varki, A. (2011) *Annu. Rev. Pathol.* **6,** 365–393

20. Consortium, Y. C. (2002) *Genome Res.* **12,** 339–348

21. Beuten, J., Gelfond, J. A., Franke, J. L., Shook, S., Johnson-Pais, T. L., Thompson, I. M., and Leach, R. J. (2010) *Cancer Epidemiol. Biomarkers Prev.* **19,** 588–599

22. Li, C., and Wong, W. H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98,** 31–36

23. International Human Genome Sequencing Consortium (2004) *Nature* **431,** 931–945

24. Yngvadottir, B., Xue, Y., Searle, S., Hunt, S., Delgado, M., Morrison, J., Whittaker, P., Deloukas, P., and Tyler-Smith, C. (2009) *Am. J. Hum. Genet.* **84,** 224–234

25. Caron, P. C., Co, M. S., Bull, M. K., Avdalovic, N. M., Queen, C., and Scheinberg, D. A. (1992) *Cancer Res.* **52,** 6761–6767

26. van Der Velden, V. H., te Marvelde, J. G., Hoogeveen, P. G., Bernstein, I. D., Houtsmuller, A. B., Berger, M. S., and van Dongen, J. J. (2001) *Blood* **97,** 3197–3204

27. Nguyen, D. H., Ball, E. D., and Varki, A. (2006) *Exp. Hematol.* **34,** 728–735

28. Kohls, M. D., and Lappi, D. A. (2000) *BioTechniques* **28,** 162–165

29. Varki, A. (2010) *Proc. Natl. Acad. Sci. U.S.A.* **107,** (Suppl. 2) 8939–8946

30. McMillan, S. J., and Crocker, P. R. (2008) *Carbohydr Res.* **343,** 2050–2056

31. Yu, Z., Lai, C. M., Maoui, M., Banville, D., and Shen, S. H. (2001) *J. Biol. Chem.* **276,** 23816–23824

32. Zhuravleva, M. A., Trandem, K., and Sun, P. D. (2008) *J. Mol. Biol.* **375,** 437–447

33. Attrill, H., Takazawa, H., Witt, S., Kelm, S., Isecke, R., Brossmer, R., Ando, T., Ishida, H., Kiso, M., Crocker, P. R., and van Aalten, D. M. (2006) *Biochem. J.* **397,** 271–278

34. May, A. P., Robinson, R. C., Vinson, M., Crocker, P. R., and Jones, E. Y. (1998) *Mol. Cell.* **1,** 719–728

35. Alphey, M. S., Attrill, H., Crocker, P. R., and van Aalten, D. M. (2003) *J. Biol. Chem.* **278,** 3372–3377

36. Angata, T., Kerr, S. C., Greaves, D. R., Varki, N. M., Crocker, P. R., and Varki, A. (2002) *J. Biol. Chem.* **277,** 24466–24474

37. Brinkman-Van der Linden, E. C., Hurtado-Ziola, N., Hayakawa, T., Wiggleton, L., Benirschke, K., Varki, A., and Varki, N. (2007) *Glycobiology* **17,** 922–931

38. Vinson, M., Strijbos, P. J., Rowles, A., Facci, L., Moore, S. E., Simmons, D. L., and Walsh, F. S. (2001) *J. Biol. Chem.* **276,** 20280–20285

39. McClure, H. M. (1973) *Am. J. Phys. Anthropol.* **38,** 425–429

40. Seibold, H. R., and Wolf, R. H. (1973) *Lab. Anim. Sci.* **23,** 533–539

41. Vanhara, P., Lincová, E., Kozubík, A., Jurdic, P., Soucek, K., and Smarda, J. (2009) *Differentiation* **78,** 213–222

42. Bauskin, A. R., Brown, D. A., Kuffner, T., Johnen, H., Luo, X. W., Hunter, M., and Breit, S. N. (2006) *Cancer Res.* **66,** 4983–4986

43. Vestergaard, E. M., Nexø, E., Tørring, N., Borre, M., Ørntoft, T. F., and Sørensen, K. D. (2010) *Int. J. Cancer* **127,** 1857–1865

44. Pratap, J., Lian, J. B., and Stein, G. S. (2011) *Bone* **48,** 30–36

45. Bon, G., Di Carlo, S. E., Folgiero, V., Avetrani, P., Lazzari, C., D'Orazi, G., Brizzi, M. F., Sacchi, A., Soddu, S., Blandino, G., Mottolese, M., and Falcioni, R. (2009) *Cancer Res.* **69,** 5978–5986

46. Miyake, H., Pollak, M., and Gleave, M. E. (2000) *Cancer Res.* **60,** 3058–3064

47. Umbas, R., Isaacs, W. B., Bringuier, P. P., Schaafsma, H. E., Karthaus, H. F., Oosterhof, G. O., Debruyne, F. M., and Schalken, J. A. (1994) *Cancer Res.* **54,** 3929–3933

48. Cheng, L., Nagabhushan, M., Pretlow, T. P., Amini, S. B., and Pretlow, T. G. (1996) *Am. J. Pathol.* **148,** 1375–1380

49. De Marzo, A. M., Knudsen, B., Chan-Tack, K., and Epstein, J. I. (1999) *Urology* **53,** 707–713

50. Rubin, M. A., Mucci, N. R., Figurski, J., Fecko, A., Pienta, K. J., and Day, M. L. (2001) *Hum. Pathol.* **32,** 690–697

51. Jemal, A., Siegel, R., Xu, J., and Ward, E. (2010) *CA Cancer J. Clin.* **60,** 277–300

52. Sarrats, A., Comet, J., Tabarés, G., Ramírez, M., Aleixandre, R. N., de Llorens, R., and Peracaula, R. (2010) *Prostate* **70,** 1–9

53. Schröder, F. H., Hugosson, J., Roobol, M. J., Tammela, T. L., Ciatto, S., Nelen, V., Kwiatkowski, M., Lujan, M., Lilja, H., Zappa, M., Denis, L. J., Recker, F., Berenguer, A., Määttänen, L., Bangma, C. H., Aus, G., Villers, A., Rebillard, X., van der Kwast, T., Blijenberg, B. G., Moss, S. M., de Koning, H. J., and Auvinen, A. (2009) *N. Engl. J. Med.* **360,** 1320–1328

54. Andriole, G. L., Crawford, E. D., Grubb, R. L., 3rd, Buys, S. S., Chia, D., Church, T. R., Fouad, M. N., Gelmann, E. P., Kvale, P. A., Reding, D. J., Weissfeld, J. L., Yokochi, L. A., O'Brien, B., Clapp, J. D., Rathmell, J. M., Riley, T. L., Hayes, R. B., Kramer, B. S., Izmirlian, G., Miller, A. B., Pinsky, P. F., Prorok, P. C., Gohagan, J. K., and Berg, C. D. (2009) *N. Engl. J. Med.* **360,** 1310–1319

55. Maher, C. A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., Sam, L., Barrette, T., Palanisamy, N., and Chinnaiyan, A. M. (2009) *Nature* **458,** 97–101

56. Mani, R. S., Tomlins, S. A., Callahan, K., Ghosh, A., Nyati, M. K., Varambally, S., Palanisamy, N., and Chinnaiyan, A. M. (2009) *Science* **326,** 1230

57. Park, K., Tomlins, S. A., Mudaliar, K. M., Chiu, Y. L., Esgueva, R., Mehra, R., Suleman, K., Varambally, S., Brenner, J. C., MacDonald, T., Srivastava, A., Tewari, A. K., Sathyanarayana, U., Nagy, D., Pestano, G., Kunju, L. P., Demichelis, F., Chinnaiyan, A. M., and Rubin, M. A. (2010) *Neoplasia* **12,** 590–598

58. Verhagen, P. C., van Duijn, P. W., Hermans, K. G., Looijenga, L. H., van Gurp, R. J., Stoop, H., van der Kwast, T. H., and Trapman, J. (2006) *J. Pathol.* **208,** 699–707

59. Schmitz, M., Grignard, G., Margue, C., Dippel, W., Capesius, C., Mossong, J., Nathan, M., Giacchi, S., Scheiden, R., and Kieffer, N. (2007) *Int. J. Cancer* **120,** 1284–1292

60. Shen, M. M., and Abate-Shen, C. (2010) *Genes Dev.* **24,** 1967–2000

61. Eeles, R. A., Kote-Jarai, Z., Al Olama, A. A., Giles, G. G., Guy, M., Severi, G., Muir, K., Hopper, J. L., Henderson, B. E., Haiman, C. A., Schleutker, J., Hamdy, F. C., Neal, D. E., Donovan, J. L., Stanford, J. L., Ostrander, E. A., Ingles, S. A., John, E. M., Thibodeau, S. N., Schaid, D., Park, J. Y., Spurdle, A., Clements, J., Dickinson, J. L., Maier, C., Vogel, W., Dörk, T., Rebbeck, T. R., Cooney, K. A., Cannon-Albright, L., Chappuis, P. O., Hutter, P., Zeegers, M., Kaneva, R., Zhang, H. W., Lu, Y. J., Foulkes, W. D., English, D. R., Leongamornlert, D. A., Tymrakiewicz, M., Morrison, J., Ardern-Jones, A. T., Hall, A. L., O'Brien, L. T., Wilkinson, R. A., Saunders, E. J., Page, E. C., Sawyer, E. J., Edwards, S. M., Dearnaley, D. P., Horwich, A., Huddart, R. A., Khoo, V. S., Parker, C. C., Van As, N., Woodhouse, C. J., Thompson, A., Christmas, T., Ogden, C., Cooper, C. S., Southey, M. C., Lophatananon, A., Liu, J. F., Kolonel, L. N., Le Marchand, L., Wahlfors, T., Tammela, T. L., Auvinen, A., Lewis, S. J., Cox, A., FitzGerald, L. M., Koopmeiners, J. S., Karyadi, D. M., Kwon, E. M., Stern, M. C., Corral, R., Joshi, A. D., Shahabi, A., McDonnell, S. K., Sellers, T. A., Pow-Sang, J., Chambers, S., Aitken, J., Gardiner, R. A., Batra, J., Kedda, M. A., Lose, F., Polanowski, A., Patterson, B., Serth, J., Meyer, A., Luedeke, M., Stefflova, K., Ray, A. M., Lange, E. M., Farnham, J., Khan, H., Slavov, C., Mitkova, A., Cao, G., and Easton, D. F. (2009) *Nat. Genet.* **41,** 1116–1121

62. Esserman, L., Shieh, Y., and Thompson, I. (2009) *J. Am. Med. Assoc.* **302,** 1685–1692