SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure S1: GAG-HERV-K ch22q11.23 features. Endogenous retrovirus K from chromosome 22 is located on the q11.23 region, nearby the BCR, ZDHHC8P, IGLL1 and RGL4 genes. HERV-K ch22q11.23 has GAG-PRO, POL and ENV genes flanked by LTR sequences. The GAG gene has an ORF encoding a 715aa protein that contains P10, P24 and Zinc finger domains. ∆: deletions; I: early stop codons; → primer binding regions; ABS: antibody binding site for monoclonal antibody TI-35 (see Supplementary Fig. S3C).

Supplementary Figure S2: qPCR gene expression analysis of chromosome 22q11.23 region in cancer tissues. Total RNA from cancer tissues was submitted to DNAse treatment and reverse transcriptase reaction for cDNA synthesis. The resulting cDNA was used on qPCR to detect the expression of GAG-HERV-K and the neighbor genes BCR, IGGL1, RGL4 and ZDHHC8P. The results are expressed as relative expression to GAG-HERV-K levels in normal prostate. GAG-HERV-K expression is relative to GAPDH. BCR, IGGL1, RGL4 and ZDHHC8P expression is relative to TFRC.

Supplementary Figure S3: Alignment of GAG proteins derived from human endogenous retroviruses and respective epitopes recognized by prostate cancer patient sera. A, Detail alignment from several GAG-HERV-K family members and the TI-35 binding site (red) highlighting the amino-acid differences present in GAG-HERV-K ch22q11.23 (blue). B, Western-blot analysis of immunoprecipitation of GAG-HERV-K from VCaP and LNCaP prostate cancer cell lines. WFP prostate cancer patient serum (WFP) was incubated with VCaP or LNCaP protein extract for 2 hours at 4ºC. Protein-antibody complexes were pulled down using anti-human IgG sepharose beads. The pull-

down pellet was submitted to SDS-PAGE and Western-blot analysis. Reactivity to GAG-HERV-K protein was developed by incubating the membrane with TI-35 monoclonal antibody. A pool of healthy donor sera was used as control (Con). C, Summary of ELISA results from epitope mapping using 20mers peptides overlapping by 10aa (71 peptides tested in total from JPT Innovative Peptide Solution, Berlin, Germany) spanning GAG-HERV-K and tested individually for recognition by monoclonal antibody TI-35 and different prostate cancer patient sera seropositive for GAG-HERV-K protein. The ■ symbol represents a positive reaction to the corresponding peptide (labeled by aminoacid start and end position). A pool of healthy donor sera was used as control (Con pool). Plates were coated with 5 μM peptide and the reaction was performed as previously described in methods.

Supplementary Figure S4: Characterization of TI-35 mAb by subcellular colocalization of staining with GAG-HERV-K expression and by transfection. A-B, Confocal images of HeLa cells transfected with GAG-HERV-K_GFP (A) or GFP only (B) and stained with TI-35. Blue represents DAPI nuclear staining, green represents GFP, and purple represents TI-35 staining. Colocalization is observed in the merged data as white, in areas indicated by arrows. Axes in the corner of each quadrant show size/magnification information. C-D, TI-35 immunostaining with formalin-fixed Sf9 insect cells infected with GAG-HERV-K-recombinant baculovirus (C) or control baculovirus (D). Original magnification (objective) was 4x, and insets show zoomed area at 20x.

Supplementary Figure S5: Correlation between antibodies against GAG-HERV-K and group risk or Gleason score in cancer patients. Prostate cancer patients were split (A) in 3 groups according to Gleason score and classified from group risk 1 to 3 and

compared to healthy donors (HD), or (B) in 3 groups according to Gleason score as indicated. Antibodies anti-GAG-HERV-K were detected and analyzed as in Fig. 4. Asterisk denotes significantly higher frequency of seropositive patients (titer > 100) in risk group 3 and Gleason ≥ 8 (Chi-square test, $p<0.003$).