**Supplemental Digital content 1: Supplemental Methods**

***Cell separation***

MACS magnetic cell separation technology was used (Miltenyi Biotec, Bergisch Gladbach, Germany) following the manufacturer's protocol. Briefly, cells were mechanically separated from fresh biopsies by rotation-agitation in 10 ml of RPMI medium supplemented 2% FBS, 1% DTT and 5nM EDTA for 1 hour. The *lamina propia* and debris were separated by filtration through 30 μm preseparation filters and dead cell removal kit was used to prepare a viable single-cell homogenous suspension. Live cells were labeled with anti-CD45 magnetic microbeads (Miltenyi Biotec) and FcR blocking reagent (Miltenyi Biotec) was used to increase the specificity of antibody labeling. The antibody-labeled cell-suspension was applied to a MS separation column and unlabelled CD45- cells were collected with the pass through, while CD45+ cells were recovered after removing the column from the magnetic separator. The ratio of *EPCAM* (CD326) *vs.* *PTPRC* (CD45) gene expression in the epithelial fraction was >99% and the inverse ratio in the immune fraction was 90%.

***MHC genotyping***

To analyze 2,360 SPNs across a region of 4.91bp in the MHC, a commercially available MHC Panel Set (Illumina Inc., San Diego, CA) was genotyped in the discovery samples using a Golden Gate protocol (Illumina Inc.). The set includes two oligonucleotide pools, one containing 1,293 SNPs evenly spaced across the MHC region, and the other 1,228 SNPs in or near coding regions. Genotyping results were analyzed using Beadstudio 2.0 (Illumina Inc.) and threshold for genotype confidence (GC) score was set at 0.25. Genotyping data quality was assessed by controlling for discrepancies between overlapping SNPs in the two panels (161 SNPs). Analysis of the genotyping results was performed using PLINK v2.044 association analysis toolset (http://pngu.mgh.hravard.edu/~purcell/plink/). Before performing disease association analysis, SNPs with minor allele frequency lower than 0.01 and those failing Hardy-Weinberg equilibrium test (p<0.001) in either cases or controls were eliminated. Fisher’s exact test was used for single marker allelic association analyses, and p-values below 10-4 were regarded significant.

To confirm the results of the most significant SNPs, samples were also genotyped using Taqman SNP Genotyping Assays (Applied Biosystems, Foster city, CA). Since a pre-designed genotyping assay was not available for rs3130838, it was replaced by a tag SNP (rs3135316; r2=1). The same assays were used for the replication study in the larger sample set. In this case, allelic frequencies were compared in 2 x 2 contingency tables using Fisher’s exact test and statistical analyses were performed using EPI-INFO v.6.0 (Center for Disease Control, Atlanta, GA). The strength of association was estimated by the odds ratio with 95% confidence interval and p-values below 0.05 were regarded significant.