Supplementary Table 1. Haplotype frequencies and r^2 values of rs922483 with *BLK* disease associated SNPs in major racial groups*

Haplotype frequency		rs2736340	rs13277113	rs922483	rs2736345	rs2618476
European	r^2	0.76	0.76	-	0.94	0.74
	73%	C	C	G	A	T
	20%	**T**	**T**	**A**	**G**	**C**
African American	r^2	0.46	<0.4	-	0.66	0.46
	61%	C	C	G	A	T
	15%	**T**	**T**	**A**	**G**	**C**
	14%	C	C	**A**	**G**	T
Asian	r^2	0.74	0.86	-	0.91	0.74
	21%	C	C	G	A	T
	66%	**T**	**T**	**A**	**G**	**C**

*****Healthy subjects of the GAP cohort genotyped using the Immunochip were divided into ancestral groups using principal component analysis. Of the subjects analyzed (n=1395): 77% grouped with subjects of European ancestry, 18% were of African American ancestry and 5% were of Asian ancestry. The linkage disequilibrium of rs922483 with autoimmune disease-associated *BLK* SNPs (rs2736340, rs13277113, rs2736345 and rs2618476) is shown for each group as r^2 values. Haplotypes with a frequency of >5% for European, African American and Asian's within the GAP cohort are shown and risk alleles are bold.

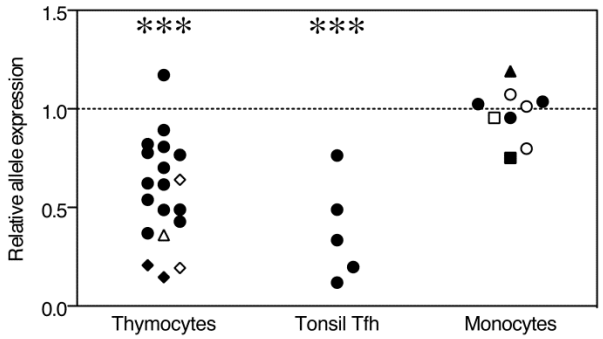
Supplementary Methods

GAP subjects genotyped on the Illumina Immunochip (1) were analyzed by principal component analysis using markers of continental origin and ethnicity (2) to stratify ancestral populations. Linkage disequilibrium, haplotype and population substructure analysis was performed using SNP and Variation Suite (Golden Helix).

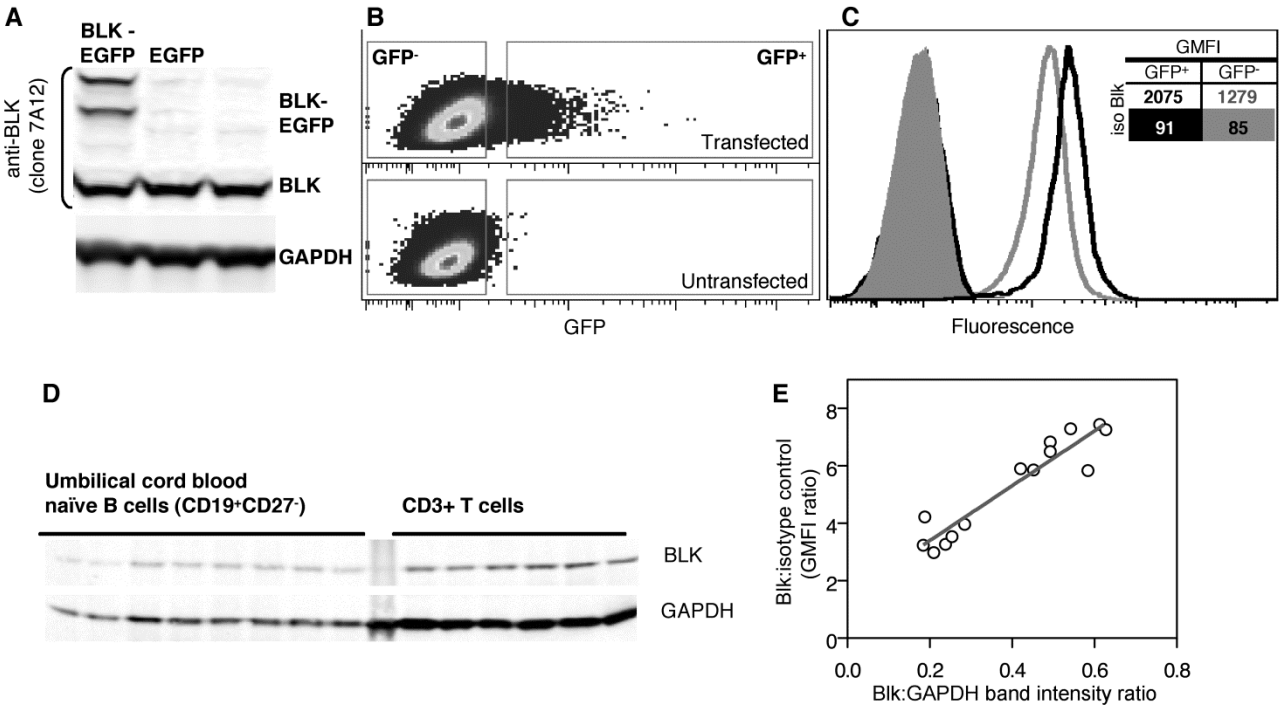
References

1. Cortes, A. and Brown, M.A. (2011) Promise and pitfalls of the Immunochip. *Arthritis Res. Ther.*, **13**, 101.

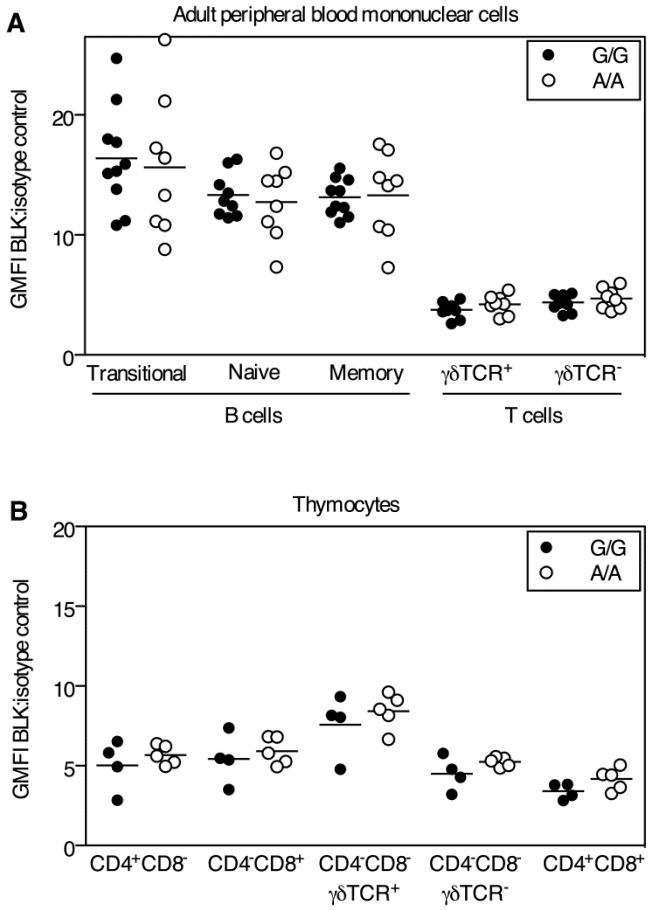
2. Tian, C., Kosoy, R., Nassir, R., Lee, A., Villoslada, P., Klareskog, L., Hammarstrom, L., Garchon, H.J., Pulver, A.E., Ransom, M. *et al.* (2009) European population genetic substructure: further definition of ancestry informative markers for distinguishing among diverse European ethnic groups. *Mol. Med.*, **15**, 371-383.



Supplementary Figure 1. Allelic expression bias of *BLK* transcript is restricted to T cells (including thymus and tonsil) and absent in monocytes and dendritic cells. Relative expression of SNP alleles was assayed using pyrosequencing and relative expression of SNP alleles normalized to DNA ratios are shown for: total thymocytes (filled circles), $CD4^+CD3^+CD19^-CD25^+$ thymocytes (unfilled diamonds), $CD8^+CD3^+CD19^-$ (unfilled triangles), $CD4^+CD8^+CD3^+CD19^-$ thymocytes (filled diamonds), tonsil T follicular helper cells (Tfh: $CD3^+CD4^+ICOS^+CD19^-$), untreated monocytes (filled circles) and activated monocytes (open circles), immature (filled squares) and mature (open squares) dendritic cells and macrophages (filled triangles). A value of one (dotted line) represents equal allele expression and statistically significant differences as determined by the Mann Whitney non-parametric *U*-test to compare DNA values to other groups are denoted as
$***P < 0.001$.



Supplementary Figure 2. Validation of intracellular Blk flow assay. (**A**) Western blot analysis of Jurkat cells transduced with a *BLK-GFP* lentivirus construct (lane 1), *GFP* lentivirus construct (lane 2), and untransduced Jurkat cells (lane 3). Protein bands were detected using anti-Blk clone 7A12 and anti-GAPDH (**B**) Freshly transduced and untransduced Jurkat cells were analyzed using intracellular staining and flow cytometry. Cells were gated on GFP^+ or GFP^- (**C**) The histogram shows cells intracellularly labeled with anti-Blk clone 7A12 that are GFP^+ (over-expressing Blk, black unfilled) or GFP^- cells (untransduced cells, grey unfilled). Cells labeled with an isotype control are shown as GFP^+ cells (black filled) or GFP^- cells (grey filled), the geometric mean of fluorescence intensity (GMFI) values are shown (**D**) Western blot analysis of cell lysates from umbilical cord blood naïve and transitional B cells ($CD19^+CD3^-CD27^-$) (lanes 1-8) and $CD3^+$ T cells (lanes 10-15). Blk was detected using anti-Blk clone 7A12 and the protein loading control was determined by GAPDH band intensity (**E**) Normalized Western blot values are plotted against GMFI values of intracellular Blk staining of the same samples as determined by flow and normalized to isotype controls. The line fit through the data points represents a linear regression (R-squared = 0.89, $P < 0.0001$).



Supplementary Figure 3. Blk protein levels are not reduced in thymocytes and adult peripheral blood B or T cells from subjects homozygous for the risk allele. Intracellular Blk was measured by intracellular staining and flow cytometry. The GMFI of cells labeled with anti-Blk was normalized to the GMFI of isotype control antibody labeled cells. Shown are Blk protein levels in (**A**) adult peripheral blood* cell populations pooled data from two experiments: transitional B cells ($CD19^+IgD^+CD27^-CD38^{hi}CD10^+$), naïve/mature B cells ($CD19^+IgD^+CD27^-CD38^{int/lo}CD10^-$), unswitched memory B cells ($CD19^+IgD^+CD27^+$), γδ TCR positive ($CD3^+CD19^-$γδ TCR^+) and negative ($CD3^+CD19^-$ γδ TCR^-) T cells (**B**) $CD3^+CD19^-$ thymus cell populations were gated as $CD4^+CD8^-$, $CD4^-CD8^+$, $CD4^-CD8^-$γδTCR^+, $CD4^-CD8^-$γδTCR^- and $CD4^+CD8^+$. The mean of each group is represented by a horizontal line.