

Identifying regulatory pathways of spleen tyrosine kinase expression in human basophils



Xia Peng, PhD,^{a,b,*} Mingming Zhao, PhD,^{c,*} Li Gao, PhD,^a Ranjan Sen, PhD,^b and Donald MacGlashan, Jr, MD, PhD^{a,*}
Baltimore, Md, and Shanghai, China

Background: Expression levels of spleen tyrosine kinase (SYK), a critical signaling tyrosine kinase in basophils, are uniquely low relative to all other circulating leukocytes, and levels are highly variable in the population.

Objective: We sought to determine whether transcriptional regulation of SYK through unique silencing of the *SYK* gene determines its basophil-specific expression patterns.

Methods: Culture-derived basophils (CD34B cells) were derived from cultures of CD34⁺ progenitor cells by using 2 methods (G1 or G3). Peripheral blood basophils (PBBs; relative SYK protein level = 1), B cells (SYK = 8), CD34B-G1 cells (SYK = 11), and CD34B-G3 cells (SYK = 5) were examined by using assay for transposase-accessible chromatin sequencing (ATAC-seq) methods. In addition, the transcriptomes of 6 cell types, PBBs, peripheral blood eosinophils (SYK = 11), plasmacytoid dendritic cells (SYK = 30), CD34⁺ progenitors (SYK = 11), CD34B-G1 cells, and CD34B-G3 cells, were analyzed for patterns that matched patterns of SYK expression in these cells, with a focus on transcription factors.

Results: ATAC-seq showed that PBBs have multiple open regions in the *SYK* gene, suggesting a nonsilenced state with 1 region unique to PBBs (low SYK expression), 1 region unique to both PBBs (low SYK expression) and both G1 and G3 CD34B cells (high and moderate SYK expression, respectively), and 5 regions unique to B cells (high SYK expression). SYK expression across the 6 cell types explored showed a unique pattern that was matched to expression patterns of 3 transcription factors: Kruppel-like factor 5 (KLF5), zinc-finger protein 608 (ZNF608), and musculoaponeurotic fibrosarcoma protein (c-MAF).

Conclusions: Two new potential regulatory pathways for SYK expression were identified. One appears independent of transcriptional regulation, and one appears to be dependent on transcriptional control in the *SYK* gene. (*J Allergy Clin Immunol* 2020;145:947-57.)

Key words: Human, basophil, allergy, development, signal transduction

From ^athe Asthma and Allergy Center, Johns Hopkins University, Baltimore; ^bthe Department of Laboratory Medicine, Shanghai General Hospital, Shanghai JiaoTong University School of Medicine, Shanghai; and ^cthe Laboratory of Molecular Biology and Immunology, Biomedical Research Center, National Institutes of Health/National Institute on Aging, Baltimore.

*These authors contributed equally to this work as first authors.

These studies were funded by National Institutes of Health grant AI100952.

Disclosure of potential conflict of interest: The authors declare that they have no relevant conflicts of interest.

Received for publication June 17, 2019; revised August 27, 2019; accepted for publication October 1, 2019.

Available online October 17, 2019.

Corresponding author: Donald W. MacGlashan, Jr, MD, PhD, Johns Hopkins University Asthma and Allergy Center, Department of Medicine, 5501 Hopkins Bayview Circle, Unit Office 3, Baltimore, MD 21224. E-mail: dmacglas@jhmi.edu.

The CrossMark symbol notifies online readers when updates have been made to the article such as errata or minor corrections

0091-6749/\$36.00

© 2019 American Academy of Allergy, Asthma & Immunology

<https://doi.org/10.1016/j.jaci.2019.10.005>

Abbreviations used

ABP:	Alcian blue-positive
ATAC-seq:	Assay for transposase-accessible chromatin sequencing
CD34B:	Culture-derived basophil
eQTL:	Expression quantitative trait locus
Flt3L:	Flt3 ligand
JUN:	c-Jun N-terminal kinase
PBB:	Peripheral blood basophil
PBE:	Peripheral blood eosinophil
PDC:	Plasmacytoid dendritic cell
SCF:	Stem cell factor
SNP:	Single nucleotide polymorphism
SYK:	Spleen tyrosine kinase
TSS:	Transcription start site
UTR:	Untranslated region

IgE-mediated secretion from human basophils (and mast cells) requires the activity of spleen tyrosine kinase (SYK), a tyrosine kinase in the SYK/ζ chain-associated protein of 70 kDa 2-member family.¹⁻³ SYK expression in human basophils is uniquely low, despite the absolute requirement for SYK expression in generating an IgE-mediated response.⁴ With one exception, circulating αβ T cells, which use ζ chain-associated protein of 70 kDa, basophils show the lowest expression of SYK of all the common circulating leukocytes. Indeed, the SYK level lies at the cusp of functionality for FcεRI-mediated activation of these cells.^{5,6} For example, there are subjects whose basophils do not express enough SYK to drive an IgE-mediated reaction despite there being an excess of FcεRI (so-called nonreleasers).^{5,7-10} Therefore SYK expression correlates well with the maximum secretion possible through FcεRI.

There are only 2 known mechanisms of regulation of SYK expression in human basophils. The first is a posttranslational process of downregulation after cross-linking of cell-surface FcεRI through a process of ubiquitination and proteasomal degradation.¹¹ Although this is a strong mechanism, as a mechanistic explanation for the highly variable expression of SYK in basophils among the general population, it is unsatisfactory. For example, recent studies have not been able to detect autoantibodies that could drive SYK downregulation in most subjects (while possibly explaining low SYK levels in basophils from patients with chronic spontaneous urticaria).¹²

A second mechanism, upregulation by means of exposure to high concentrations of IL-3, is weak.^{5,13} Changes in SYK protein expression induced by IL-3 are in the range of 30% and not the 30-fold range found in the general population or the 100-fold range when comparing peripheral blood basophils (PBBs) with other leukocytes.

All leukocytes are derived from a CD34⁺ progenitor cell that responds to factors that include the cytokine environment.

A simple *in vitro* culture of CD34⁺ cells with IL-3 alone will drive the production of a highly heterogeneous mixture of cells with the overall characteristics of mature basophils.⁴ However, the phenotype is not precisely that of mature circulating basophils. For example, SYK protein expression in culture-derived basophils (CD34B cells) is unchanged from that in CD34⁺ progenitor cells, with levels that are 11-fold higher than those in mature PBBs. With the exception of dendritic cells and monocytes, most other leukocytes express SYK at levels similar to those in CD34⁺ progenitors, suggesting that there is no program during differentiation that modifies the basic mechanics of SYK transcription or translation during maturation. Monocytes and dendritic cells express approximately 3 times the amount of SYK of the CD34⁺ progenitor, suggesting there is a program to upregulate expression of this kinase. Therefore, in contrast, basophils appear to undergo a program of downregulation during maturation.

The mechanism of SYK downregulation during maturation is not clear. As noted above, one strong mechanism of SYK regulation is related to FcεRI aggregation. However, previous studies have also determined that SYK protein expression was reasonably well correlated with steady-state SYK mRNA levels among subjects' basophils or among leukocyte cell types.⁴ Preliminary studies suggested that mRNA stability is not different between CD34B cells and PBBs (and unpublished studies)^{4,14} nor was there any evidence for a microRNA that might regulate SYK mRNA expression.^{4,14} Together, the various observations suggest that SYK protein expression might be under transcriptional control of the SYK gene, control that is unique to the human basophil.

The studies in this report provide a first look at how SYK can be regulated outside of the IgE-dependent processes. During preliminary studies of CD34⁺ cells maturing into basophils, we discovered that these cells can be induced to downregulate SYK during maturation through an IgE-independent mechanism. This report characterizes this discovery and explores some of its implications. In addition, this study posed the hypothesis that the SYK gene is silenced during maturation, either completely with no active open/accessible chromatin or through a unique pattern of "open" chromatin.

METHODS

Materials, buffers, and methods that have been used extensively in prior publications are presented in detail in the Methods section in this article's Online Repository at www.jacionline.org.

Generation of CD34-derived basophils from CD34⁺ progenitors

Two sources of CD34⁺ cells were used in these studies. For most of the studies, residual leukocytes from plateletpheresis were enriched by a combination of Percoll gradient separation (53%/62% Percoll) and positive selection. The upper layer of the gradient separation was found to be enriched in CD34⁺ cells (personal communication from Dr John Schroeder). Magnetic positive selection of CD34⁺ cells (Miltenyi Biotec, Bergisch Gladbach, Germany) followed the manufacturer's guidelines, with slight modifications. Generally, no more than 1×10^9 total cells were incubated for 30 minutes at 4°C with 200 μL of Kit reagent before passage over Miltenyi LS columns (Miltenyi Biotec, Bergisch Gladbach, Germany). In early pilot studies, the CD34⁺ purity after column separation was determined by using flow cytometry with a Miltenyi antibody designed to work with the purification reagents. Typically, postcolumn purities were $42\% \pm 13\%$, and yields ranged from 0.2 to 3.0 million CD34⁺ cells (average, $1.0 \pm 0.7 \times 10^6$). For some preliminary

studies, CD34⁺ cells were obtained from the laboratory of Dr Shelley Heimfeld (Fred Hutchinson Cancer Research Center, Seattle, Wash) and IL-3 cultures started without further handling.

The CD34⁺ progenitors were resuspended at a density of 0.1 to 0.2×10^6 cells/mL in StemPro serum-free medium (Thermo Fisher, Waltham, Mass), which contained the provided nutrient supplement, 2 mmol/L L-glutamine, 100 U/mL penicillin, 100 μg/mL streptomycin, 5 ng/mL rhIL-3, and different concentrations of Flt3 ligand (Flt3L) and stem cell factor (SCF). Cells were cultured at 37°C in a 5% CO₂ atmosphere. Cell counts and viability were checked on days 14 to 21 by using erythrosin B staining. The cells generated are heterogeneous in size, granularity, and nuclear morphology. Although not presented in the figures, a rough set of categories (eg, large vs small and strong vs weak Alcian blue staining) were enumerated to monitor qualitative differences. By and large, there were only minor differences in these characteristics for the culture conditions tested.

In many cases a low centrifugation rate (40g for 1 minute in a 70-μL volume) was used to remove a significant fraction of dead cells from the harvested cultures before use in subsequent assays. Previous studies of these cells established that dead cells do not contribute mRNA, and therefore cell lysis procedures for mRNA were based on live cell counts. However, protein remains present in dead cells, and therefore lysis for protein measurements was based on the total cell count. See [Tables E1 and E2](#) in this article's Online Repository at www.jacionline.org for details about culture conditions.

Assay for transposase-accessible chromatin sequencing analysis

Samples for assay for transposase-accessible chromatin sequencing (ATAC-seq) were prepared, as described in the literature (ATAC-seq: a method for assaying chromatin accessibility genome-wide). Briefly, 2×10^5 cells were washed with cold PBS and resuspended in 50 μL of cold lysis buffer. Cell pellets were harvested by means of centrifugation for 10 minutes at 500g (4°C) and placed on ice. The transposition reaction mix was prepared according to the manufacturer's guidelines with adjustments to optimize the fragmentation of the DNA for deep sequencing. Each cell type studied was examined with a range of cell number/reagent ratios to optimize the outcome. Generally, the reaction mixture was 25 μL of TD buffer (2 × reaction buffer from the Nextera Kit [Illumina, San Diego, Calif]), 1 μL of TDE1 buffer (Nextera Tn5 Transposase from the Nextera Kit), and 24 μL of nuclease-free H₂O. The nuclei pellet was resuspended in the transposition reaction mix and incubated at 37°C for 30 minutes. Immediately after transposition, DNA was purified with a Qiagen MinElute PCR Purification Kit (Qiagen, Hilden, Germany), and transposed DNA was eluted in 10 μL of elution buffer (Buffer EB from the MinElute Kit consisting of 10 mmol/L Tris · Cl [pH 8]). The transposed DNA fragments were amplified by using PCR, and the amplified library was purified by using the Qiagen MinElute PCR Purification Kit, eluted in 20 μL of elution buffer, and separated by means of gel electrophoresis. The gel was excised in the 200- to 1000-bp region, and DNA was eluted and suspended in 20 μL of TE buffer. Sequencing was performed by using Illumina NextSeq sequencing with the high-output 75-cycle reagent kit for single 75-bp reads. Sequencing included 150 to 400 million reads, depending on the number of samples (with their unique bar coding) combined for a particular run. Data analysis workflow included Bowtie2 (2.2.5), Samtools (0.1.19), and peak calling with the MACS2 algorithm by using the hg19 human genome build for reference and HOMER used for peak annotation. Images of the .bedgra files produced from the workflow were generated with the UCSC genome browser.

Data preparation

All microarray and ATAC-seq data relevant to the article are located in 3 repositories. The primary repository is the Gene Expression Omnibus (accession no. GSE138610). Additional data, including an in-house-designed application for exploring the ultimate raw data set, are available through <http://www.basophil.net> and <http://162.129.217.250/basophilMicroarrays>. The Web site includes a variety of files, one of which provides a detailed description of the analysis methodology for the microarray results.

RESULTS

IgE-independent downregulation of SYK in CD34⁺-derived basophils

Previous studies used only IL-3 during culture of CD34⁺ progenitors to generate basophils with characteristics that have been partially described.⁴ These cells were phenotypically similar to basophils, but SYK levels were similar to those in CD34⁺ progenitors (and other leukocytes) and not PBBs. Other studies have shown that during production of eosinophils from CD34⁺ progenitors, high concentrations of Flt3L and SCF during the first 3 days would enhance recovery of mature eosinophils.¹⁵ Evidence suggests that basophils and eosinophils have a similar lineage,¹⁶⁻¹⁸ and therefore we examined whether this same methodology could be used to enhance the recovery of mature basophils.

For the purposes of presenting these results, several labels were adopted to describe the different methods of culture that were explored. The traditional method of culture (5 ng/mL IL-3 only for 2-3 weeks) was labeled group 1 (G1). A second method used Flt3L, SCF, and IL-3 during the first 3 days of culture and 5 ng/mL IL-3 alone thereafter for 2 to 3 weeks and was labeled group 2 (G2). A third group (G3) incorporated Flt3L, SCF, and IL-3 throughout the culture period. Cultured cells were evaluated for overall cell viability, Alcian blue–positive (ABP) cell/live cell ratios, ABP cell numbers, surface marker levels, histamine content, and SYK levels. Fig E1 in this article's Online Repository at www.jacionline.org shows the various culture schemes tested.

These studies developed iteratively toward a method in which the resulting preparations contained a greater number of ABP cells appearing morphologically similar to PBBs, with similar cell-surface expression of FcεRI to G1 cells but expression levels of SYK protein that were reduced considerably compared with the G1 method. The Results section in this article's Online Repository at www.jacionline.org presents in greater detail the progression of these studies. After the 3 experimental series presented in the Results section in this article's Online Repository, the final approach incorporated a low concentration of Flt3L and SCF throughout culture.

Fig 1 summarizes results from the experimental series that were used to generate Fig E4 in this article's Online Repository at www.jacionline.org and a subsequent distinct experimental series only comparing G1 and G3 preparations. As shown in Fig 1 and Figs E2-E4 in this article's Online Repository at www.jacionline.org, these 2 culture models generate ABP cells that have basophil characteristics but are distinguished by downregulation of SYK expression in the G3 group (Fig 1, A and C). In addition, as noted above, cultures with Flt3L and SCF consistently generated more ABP cells (Fig 1, F) but with similar histamine content per cell (Fig 1, G) and FcεRI expression (Fig 1, B). Furthermore, based on morphology (size and kidney-shaped condensed nucleus), there were 10-fold greater numbers of PBB-like cells in G3 cultures versus G1 cultures (data not shown). Western blots (Fig 1, D) comparing SYK expression in the G1 and G3 groups supported the conclusions of flow cytometry in which SYK expression was lower in the G3 group. With respect to other cell markers of mature basophils (eg, CD203c, CD63, and CD32b), the Results section in this article's Online Repository includes other tests of similarity between the G1 and G3 groups that suggest that the 2 culture methods generate cells similar in phenotype.

The G1 and G3 cultures represented a contrasting model of basophil development with a particular focus on SYK expression. These 2 culture conditions were explored for characteristics that might provide insights into the regulation of SYK expression.

ATAC-seq differences

Differences in SYK expression when comparing CD34B cells and PBBs at both the protein and mRNA level⁴ and the similarity in the decay rate of mRNA between PBBs and CD34B cells (unpublished experiments and reference⁴) suggest that there is a transcriptional rate difference for SYK mRNA expression between these 2 cell types. Polymorphonuclear cells, such as PBBs, are characterized by their apparent inability to engage in further replication. In addition, differentiation can silence nonrelevant genes. These observations raise the possibility that low SYK expression results from rendering the SYK gene inaccessible to transcription, perhaps late in basophil differentiation (because there is some mRNA with an unaltered decay rate). As a step in characterizing the SYK gene, PBBs were examined with ATAC-seq methods. To provide context, 2 cells, B cells and CD34B-G1 cells, that express similar levels of SYK protein to each other but 10 times greater than in PBBs, were also examined. The B-cell was chosen because it represented a nonpolymorphonuclear leukocyte, unlike the basophil. Finally, CD34B-G3 cells (which were developed in parallel from the same progenitors used for the G1 results) were examined for differences that might explain the transition from G1 to G3 under the influence of low levels of Flt3L and SCF to downregulate SYK protein.

ATAC-seq generates a wealth of information about the cell's DNA state. These results represent the first look at the human basophil's "open/accessible" and "closed" DNA status. A first pass through the results showed that there were many accessible genes in PBBs that were consistent with the mRNA being present and many "closed" genes for mRNA (or protein) that were not present in basophils but present in B cells or CD34B cells. For example, for PBBs, IL-4, histidine decarboxylase (*HDC*), the β subunit of FcεRI (*MS4A2*), and *FCERIA* showed peak calls in PBBs (and CD34B cells) that were not present in B cells. Conversely, *CD79A*, *BLNK*, and *EBF1* were present in B-cell profiles but not in PBBs.

Fig 2, A, shows the SYK gene region. The same region is shown for B cells and CD34B-G1 and CD34B-G3 cells. There are shared active regions in these profiles and several that are not shared; regions of interest are labeled "1" to "7" and "A." The region labeled A includes the transcription start site (TSS) that was identified in the FANTOM5 project (<http://fantom.gsc.riken.jp>). Although there are no obvious start sites (TATA or CAAT boxes) in the region immediately upstream of exon 1, the FANTOM5 project included a TSS analysis of the human SYK gene in mast cells and basophils, and 99% of the TSSs for this gene were 27 bp inside of the identified exon 1. Fig E5 in this article's Online Repository at www.jacionline.org shows the peak "A" region expanded and sequence information as it relates to the 5' end of basophil SYK mRNA. Note that the identified exon 1 contains the TSS but encodes none of the sequences found in the mRNA (in unrepresented data, this mRNA 5' untranslated region [UTR] was roughly mapped by using sequential PCR primers). The 5' UTR begins downstream of the GenBank-identified exon 1, stops

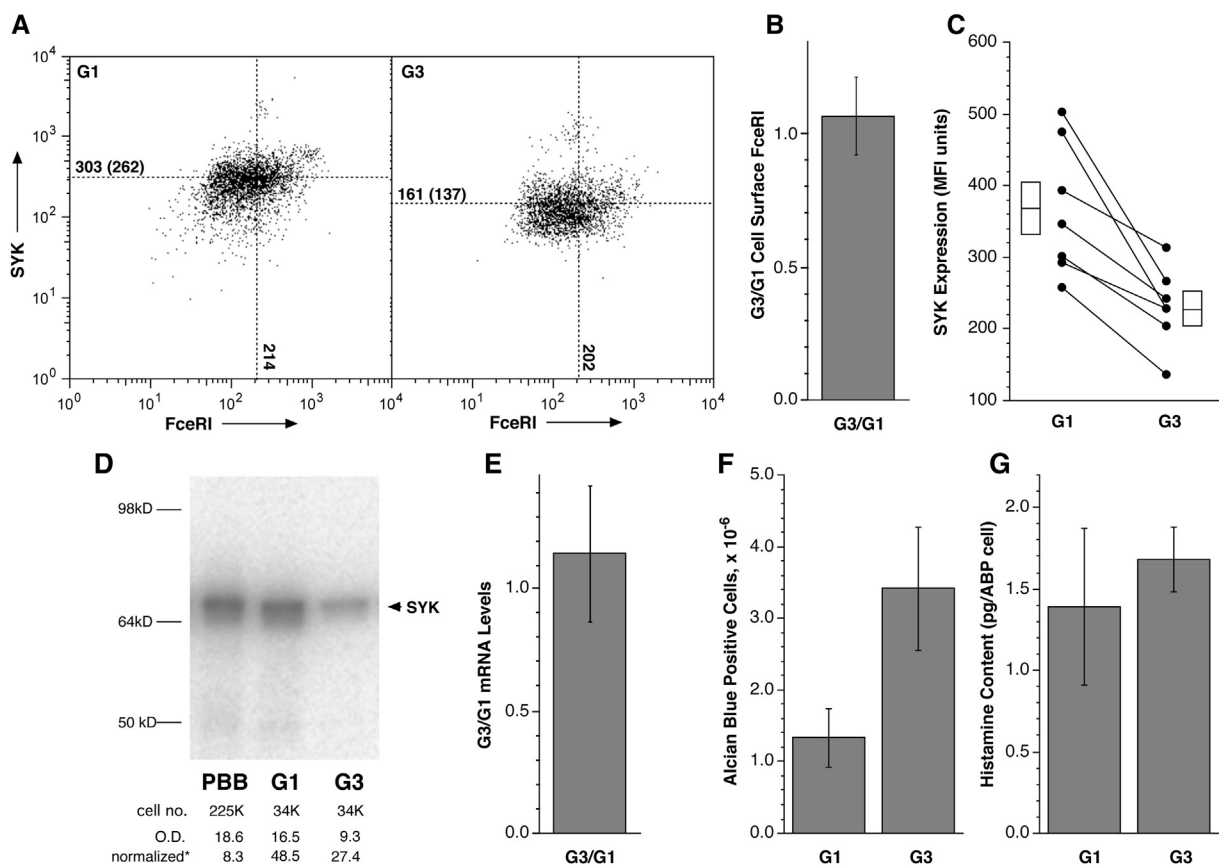


FIG 1. Primary characteristics of G1 and G3 cultures. Using the methodologies described in the text and similar to Fig E4, CD34B-G1 and CD34B-G3 cells were generated and analyzed for various metrics of the basophil phenotype. **A**, An example of SYK and FcεRI expression for the G1 and G3 cultures, as determined by using flow cytometry (gating for these plots was exclusion of “debris” or dead cells by using forward and side scatter followed by gating for CD123⁺ cells, which represented approximately 80% of live cells). **B**, Relative FcεRI expression for the G1 and G3 comparison. **C**, Compilation of SYK expression studies, as assessed by using flow cytometry. **D**, Example Western blots comparing CD34B-G1 and CD34B-G3 cell preparations compared with a PBB standard. The number of cells represented in the lysate run in the gel is noted as the cell number. Below this line is the OD for the bands and the normalized OD per 100,000 cells. As found previously, SYK expression in CD34B cells was significantly greater than that in PBBs (approximately 6-fold in this instance). **E**, Ratio of SYK mRNA expression for G1/G3 cells (see Table E4 for details). **F**, Total number of ABP cells. **G**, Histamine content per ABP cell.

at 141 bp 3' of the TSS, and continues in exon 2 (see Fig E6 in this article's Online Repository at www.jacionline.org).

The bottom of Fig 2, A, shows the GenBank assignments of SYK exons and introns; there are 14 expected exons covering approximately 100 kbp, and the known hypersensitivity regions are also shown in the tracing in Fig 3, A. The peak calls colored blue (numbered 4-7) are most evident in the B cells. The peak calls colored green (numbered 1-3) are most evident in basophils, and the orange region is at the common exon 1 region. Peak call 1 might be shared between B cells and basophils. It is notable that CD34B cells are similar to PBBs; in particular, the peak labeled “2” is found in both CD34B cells and PBBs but not B cells. Peak 3 is weak but present in PBBs but not apparent in CD34B cells. An expanded version of the peak 2 and peak 3 regions is shown in Fig E6. It is also notable that basophils and B cells show differences in activity in intron 1 or at exon 2 (peaks 2, 3 [basophil unique], 4, and 5 [B-cell unique]). B cells also show a peak at -27 kbp upstream of exon 1 (peak 7) that is not apparent in basophils, as well as a peak at intron 11 (peak 6). Activity in

basophils, B cells, and CD34B cells centers around the TSS and includes approximately 1200 bp but only 500 to 600 bp upstream of the TSS. The average results for these peaks are shown in Fig 2, B.

Regulatory regions for SYK were evaluated by using the ENCODE regulatory tracks.¹⁹ Interestingly, 3 single nucleotide variants (rs290988, rs290987, and rs2562397, which are in complete linkage disequilibrium and located at -1084, -771, and -46 bp of the TSS) within 1.1 kb upstream of SYK overlaps with many regulatory features, including enhancer histone H3K27Ac marks (which are often found near active regulatory elements) and DNase I hypersensitivity clusters (Fig 3, A). The ATAC-seq results for basophils showed a strong peak spanning the likely transcriptional start site within exon 1 at ±600 bp (labeled “A” in Fig 2).

By analyzing global RNA expression within individual tissues or cells and treating the expression levels of genes as quantitative traits, variations in gene expression that are highly correlated with genetic variation can be identified as expression quantitative trait

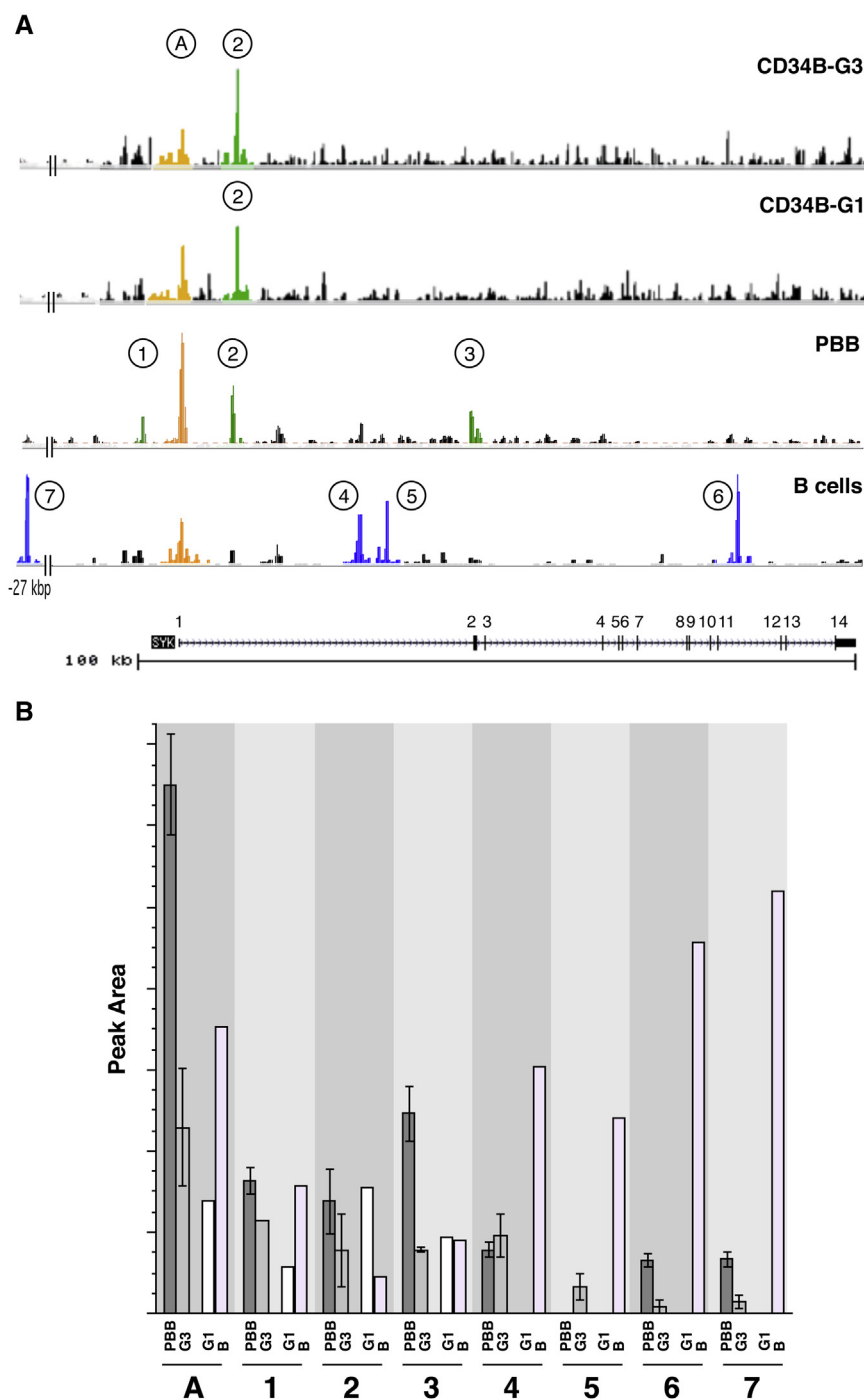


FIG 2. ATAC-seq information for the *SYK* gene. **A**, Four cell types, B cells, PBBs, and CD34B culture-derived basophils using the G1 and G3 protocols, were analyzed by using ATAC-seq and data for the *SYK* gene (shown at the bottom of the figure, both exons [black regions] and introns [line with arrows indicating the direction of transcription]), with numbered peaks collated by using the UCSC browser. Eight peaks (identified by using MACS2/HOMER annotation of the peak calls) are labeled as follows: A, peak associated with the generally identified TSS for *SYK*; 1-7, peaks 5' and throughout the gene. The bottom-most trace is the hypersensitivity profile identified in GenBank. **B**, Summary of ATAC results for the 8 identified peaks in the *SYK* gene. ATAC-seq results were averaged for preparations that were of high quality: PBBs ($n = 3$), CD34B-G3 cells ($n = 2$), CD34B-G1 cells ($n = 1$), and B cells ($n = 1$). See this article's Online Repository at www.jacionline.org for methodological details.

loci (eQTLs). The publicly available database (Genotype-Tissue Expression [GTEx]) with newly released RNA sequencing data from 1641 samples across 43 human tissues from 175 subjects

was searched for *SYK* eQTL-related single nucleotide polymorphisms (SNPs). Significant SNP-gene eQTL association was observed in samples from multiple tissue types, including the

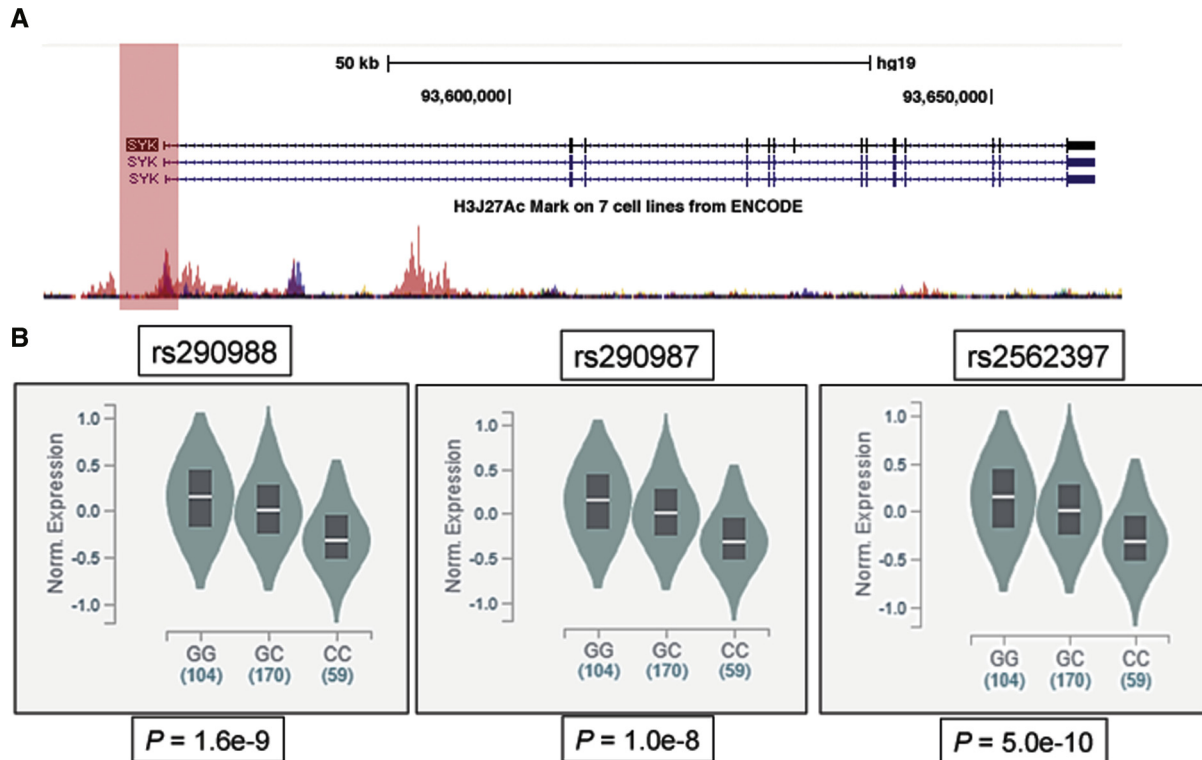


FIG 3. ENCODE regulation tracks on the *SYK* gene region (chromosome 9: 90,795,787-90,898,549) and eQTL analysis for *SYK* promoter single nucleotide variants. **A**, *SYK* gene structure, position of 3 variants (rs290988 [–1.1 kb], rs290987 [–771 bp], and rs2562397 [–76 bp]) in complete linkage disequilibrium (positions relative to TSS) within 1.1 kb upstream of *SYK* (depicted by the user track), and regulatory regions with ENCODE regulation tracks, including UCSC gene (*SYK*), layered H3K27Ac, and DNase clusters. **B**, Significant SNP gene eQTL association in tissues for these *SYK* promoter variants (with the CC homozygote showing lower *SYK* expression).

skin ($P < 10^{-8}$), and EBV-transformed lymphocytes ($P < 10^{-3}$) for 3 *SYK* promoter variants (rs290988, rs290987, and rs2562397) that are in complete linkage disequilibrium (Fig 3, B). Homozygous carriers had significantly lower gene expression compared with the wild-type. These findings provided additional evidence suggesting *SYK* promoter variants could be potential modifiers of *SYK* expression. Indeed, one of the 3 SNPs (rs2562397) is located within the first GC box motif (GGGCGG) that centered on the peak around exon 1 (see Fig E5 and the next section).

Transcriptomes of CD34 cells, CD34B-G1 and CD34B-G3 cells, PBBs, peripheral blood eosinophils, and plasmacytoid dendritic cells

As a next step, differences in the transcriptomes of various cell types that differentially express *SYK* protein were of interest. Specifically, this experimental series represented a search for transcription factors with expression that was consistent with the pattern of *SYK* expression in 6 cell types. The transcriptome for PBBs has been published.¹³ This report will add information about CD34B cells and, to enrich the comparison, add information about peripheral blood eosinophils (PBEs) and plasmacytoid dendritic cells (PDCs).

The utility of PBEs is that during development from CD34⁺ progenitors, basophils and eosinophils are thought to develop

along similar pathways,¹⁶⁻¹⁸ and in these artificial culture conditions, eosinophil-like cells are observed to be generated transiently (unpublished reports). In addition, PBEs express *SYK* at levels 10 times those in PBBs and equal to those in CD34⁺ progenitors.¹³ The implication is that the genetic program that determines *SYK* expression in CD34B cells might be more similar to that in PBEs than that in PBBs (see the Results section and Fig E7 in this article's Online Repository at www.jacionline.org). PDCs express 30 times more *SYK* than PBBs¹³ and provide a useful perspective on regulation of this gene when it is not simply maintained at the level of the CD34⁺ progenitor but upregulated during maturation to the cell obtained in circulation.

The final comparisons included changes that occur between CD34⁺ (day 0 vs day 21) and CD34B cells and differences between the G1 and G3 culture conditions that lead to lower *SYK* levels in G3. For some of the issues analyzed, the focus was on the presence of transcription factors and whether a pattern of mRNA expression for a particular transcription factor fit with the pattern of *SYK* expression across these various cell types.

There were several approaches to the general data analysis of the transcriptomes of these cells. Four issues related to general issues of transcriptomics are presented in Tables E5-E7 in this article's Online Repository at www.jacionline.org. The fifth comparison relates to using all 6 transcriptomes to explore patterns consistent with *SYK* expression.

The entire data set was used to discover a pattern in specific genes that could be used to guide future studies and to suggest possible mechanisms of regulation, and transcription factors that could be matched to the accessible regions of the *SYK* gene shown in Fig 2. Fig 4 summarizes the results of this analysis as a heat map, and Table E3 in this article's Online Repository at www.jacionline.org enumerates the specific data. The analysis began by defining a list of 2-fold changes when comparing mRNA expression in PBEs versus PBBs. A 2-fold change would be considered a very liberal threshold because the Bonferroni, Benjamini-Hochberg false discovery rate, and paired *t* test thresholds were 31-, 20- and 2.67-fold, respectively, to capture a broad selection of differences. This list was then used to generate ratio data for 4 additional comparisons. At the top of the heat map is the pattern for *SYK* mRNA levels. Although not part of this study (except for the G3/G1 comparison), comparative levels for SYK protein are also noted.^{4,14} The remainder of the heat map shows genes that follow this pattern, either directly or in an inverted manner.

There were 5 Boolean logic gates in the algorithm to generate the 4 transcript patterns noted in the figure. For example, one logic gate would be a test of whether the ratio of CD34B/CD34-D0 for a gene was unchanged (in these studies high-purity CD34 progenitors [day 0] were obtained from the Fred Hutchinson Cancer Research Center and used to generate the G1 culture transcriptomes). The logic gate for the G3/G1 difference was allowed to either reflect no change (defined as a fold change between 0.5 and 1.5) or a decrease/increase (>2-fold). The logic gates for CD34B cells/PBBs, PBEs/PBBs, and PDCs/PBBs were set for an increase or decrease. The heat map and Table E3 show the logic gates and then list the transcripts that followed the pattern. The reasoning for this specific analysis is that (1) there might be changes in other elements that either follow the pattern of SYK and therefore are part of a similar regulatory network, (2) they are controlling elements in SYK regulation, or (3) they might be inhibitory elements that directly or indirectly influence SYK expression. There were 61 transcripts that fit these patterns, of which 3 were known transcription factors: KLF5, ZNF608, and MAF (to be discussed below). Interestingly, c-Jun N-terminal kinase (*JUN*), FOS, and MAF are transcription factors that regulate the *IL4* gene, which is only weakly expressed in CD34B cells (see the Discussion section).

The microarray results also provided a first look at whether there were differences in SYK mRNA expression between CD34B-G1 and CD34B-G3 cells. Quantitative PCR was used on the 2 paired preparations used for the microarray and one additional G1/G3 sample pair to confirm the absence of a difference. Fig 1, E, summarizes the results and Table E4 in this article's Online Repository at www.jacionline.org provides greater detail on these results, showing that there was no difference in SYK mRNA expression between the 2 types of cells. See Table E8 in this article's Online Repository at www.jacionline.org for a broader context of the differences between these 2 cell types.

DISCUSSION

As noted in the introduction, SYK expression in PBBs is uniquely low relative to that in nearly all other leukocytes, despite being a required early tyrosine kinase linking FcεRI aggregation to secretion. A model of basophil development

(CD34B) that involved culture of CD34⁺ progenitors cells with IL-3 alone develops characteristics roughly consistent with PBBs. However, it is clear that these cultures only generate basophil-like cells that are morphologically identical to PBBs at a low frequency. In the context of SYK, these cells clearly are not phenotypically like PBBs. In our efforts to induce more productive CD34B cell cultures, we discovered that the addition of low concentrations of Flt3L and SCF would induce 2.5-fold increases in ABP cells and nearly 10-fold greater numbers of cells with condensed bilobed nuclei and cell size consistent with circulating basophils. In addition, this change was accompanied by decreases in SYK expression. This second model of basophil maturation allowed a side-by-side comparison to explore potential regulatory pathways. When coupled with comparisons to eosinophils, PDCs, PBBs, or B cells, it becomes apparent that there are multiple potential controlling elements in regulation of SYK expression.

Therefore these studies created a new model of basophil generation from CD34⁺ progenitors. Although our interest in the Flt3/SCF-induced changes in SYK expression motivates our interest in this model, it is apparent that these cytokines lead to cultures more enriched in cells resembling circulating basophils. Profiling the transcriptome shows that the G1 culture model produced a cell that resembles the basophil more than the eosinophil. Two thirds of the transcripts that were used to test the similarity between CD34B cells and PBBs move further toward the basophil profile in the G2/G3 culture model. However, the difference between G3 and G1 is not always in the direction of the basophil, with one third of the transcripts moving away from the phenotype of the PBB, although not necessarily toward the eosinophil.

Fig 5 summarizes the pathways that this study and prior studies now suggest result in regulation of SYK expression. At the far left of the cartoon is the well-established FcεRI-mediated ubiquitinylation of SYK as a posttranslational process of downregulation. This mechanism of SYK downregulation is well studied, and aggregation is found to have a significant effect; the median reduction after an optimal level of FcεRI aggregation will result in 70% loss of SYK,²⁰ but in some preparations reductions of 90% are observed *in vitro*. Studies of this specific process led to a recent study that explored the possibility that autoantibodies could be a driver of the high variability of SYK protein expression in the general population.¹² With some caveats, this recent study concluded that there was no evidence for such a process in most subjects. However, when it was possible to demonstrate the presence of autoantibodies, as found in 5% to 10% of patients with chronic spontaneous urticaria, it was evident that such a mechanism could be a possible cause of significant SYK reduction. These autoantibodies could also reduce SYK expression in basophils maturing in CD34B cell cultures. However, in general, this is not a satisfactory explanation for the low levels of SYK expression in PBBs compared with that in other circulating leukocytes or a reasonable explanation for the high variability in SYK expression among patients without chronic spontaneous urticaria.

The current studies frame the regulatory pathways in a general sense. Previous studies showed that mRNA expression of SYK showed similar rates of disappearance in PBBs and CD34B cells after addition of actinomycin D (and unpublished results),^{4,14} suggesting that the increased steady-state level of mRNA in CD34B cells relative to PBBs was not a result of differences in

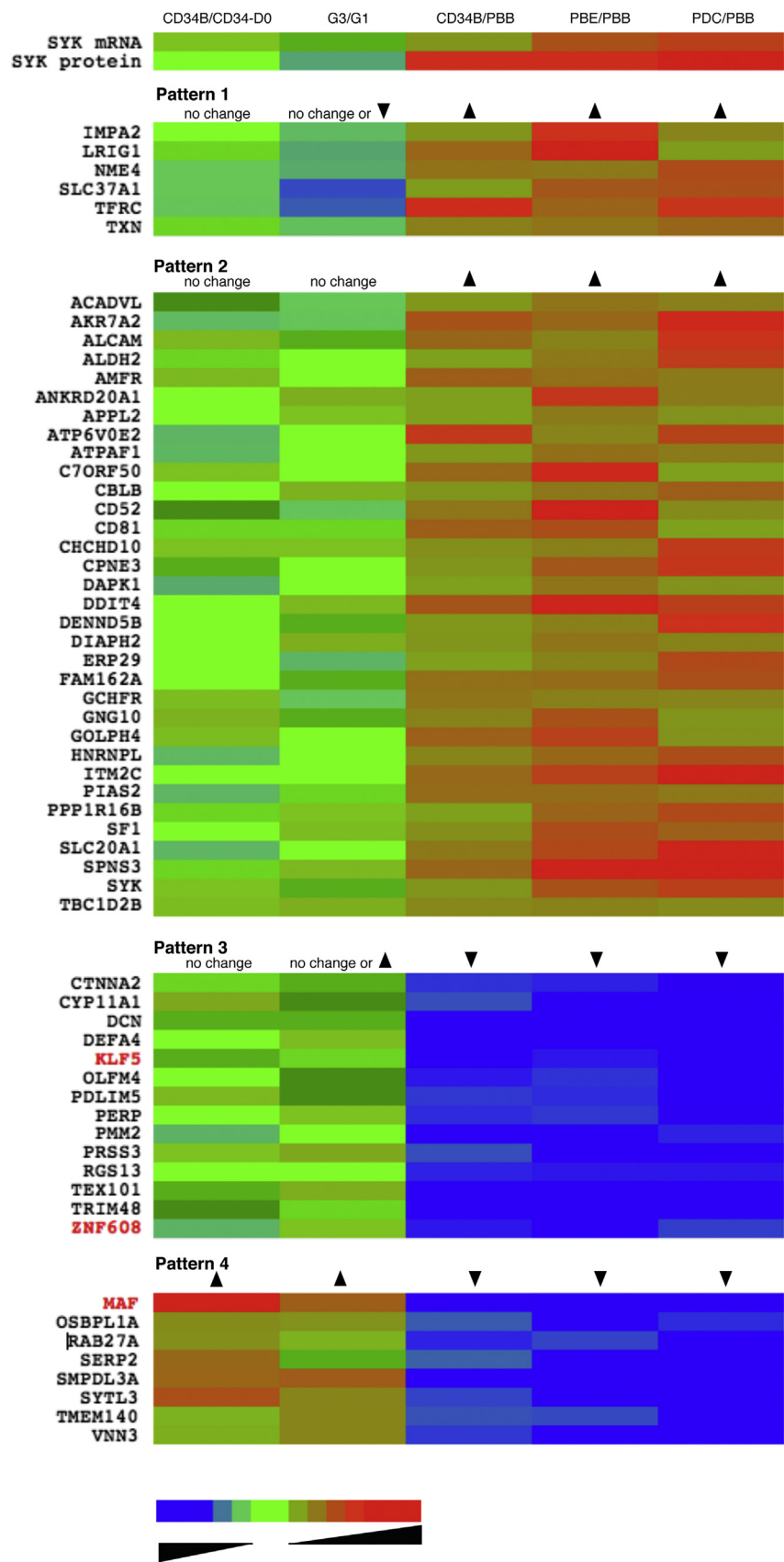


FIG 4. Pattern matching SYK expression with transcript behavior across different cell types displayed as a heat map (see this article's Online Repository at www.jacionline.org for a tabulated presentation of the

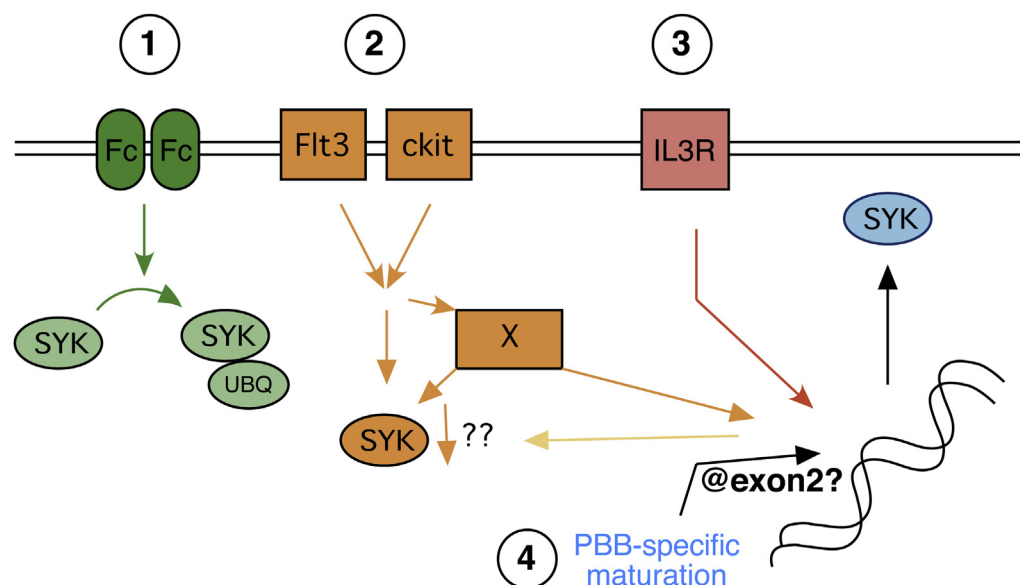


FIG 5. Synopsis of possible mechanisms regulating SYK expression in basophils. There are 4 proposed mechanisms based on available information for regulation of SYK expression.

mRNA stability. *SYK* mRNA expression shows a reasonable correlation with protein expression ($R^2 = 0.29$),⁴ also suggesting a dependence on transcription (with no evidence for mRNA stability differences). However, the current studies reveal that there is at least 1 mechanism, exposure to Flt3L, SCF, or both, operating independently of a mechanism that regulates SYK mRNA levels. Further work will be needed to illuminate the precise mechanisms.

The Flt3 and c-Kit receptors together contribute to down-regulation of SYK but not through downregulation of *SYK* mRNA expression. These 2 receptors have been shown in some cellular contexts to interact with c-CBL and therefore might mediate a loss of SYK expression through mechanisms similar to FcεRI (although this requires further study). The G1/G3 results also again raise the possibility of a microRNA that regulates SYK translation without degradation of the mRNA. It is also possible that Flt3L and SCF act indirectly to alter SYK generation (represented by the “X” factor in the cartoon or the arrow leading to change indirectly through the nucleus). IL-3 is known to induce an increase in SYK mRNA expression (with more modest effects on protein synthesis^{5,13}), suggesting another regulatory pathway that acts on transcription. In other words IL-3 increases SYK mRNA expression, with only minor effects on SYK protein levels, whereas Flt3L/SCF results in no change in SYK mRNA expression, with marked effects on SYK protein levels. Finally, the ATAC-seq results suggest some regulation at the level of DNA organization.

If there is transcriptional regulation of SYK, which has been suggested in previous studies, the pattern matching shown in Fig 4 and Table E3 suggests 3 transcription factors that might participate in this process. A bioinformatics-based screening identified several SYK promoter transcription factors and cis-regulatory elements in the regions identified in the ATAC-seq. Even with some stringency, there are, not surprisingly, many potential transcription factors that could regulate this gene. Some of the possibilities are notable for their connections to other pieces of information. Five merit some discussion and are further explored in this article’s Online Repository at www.jacionline.org.

The tabulated patterns for transcription factors suggest a connection with c-JUN, expression of which is high in CD34B cells and PBEs and low in basophils. Indeed, basophils do not express c-JUN protein. Perhaps linked to c-JUN are 2 transcription factors with patterns of expression that are tightly matched with the expression pattern of SYK in the cells studied (Fig 4). For the pattern displayed by c-MAF, the prediction is that c-MAF is a repressor of transcription. It is not uncommon to find an inverted relationship for c-MAF and c-JUN in the regulation of a gene. This is an interesting possibility because c-MAF is also a transcription factor important in IL-4 secretion,²¹ a cytokine that is a hallmark of basophil secretion.²² It might be relevant that T cells also use c-MAF for *IL4* gene transcription and T cells also suppress SYK expression, as noted above. Also, CD34B cells express IL-4 poorly relative to PBBs,

same data and Table E3). By using microarray results for 5 types of cells (PBBs, PBEs, peripheral blood PDCs, and CD34-derived basophils [CD34B]) developed with the G1 or G3 protocols, transcripts were analyzed for the following 4 patterns of expression. The short gene name is shown to the left of the heat map, and columns in the heat map are colored for the ratio of the presence of transcripts for the comparison shown at the top of the column. Arrows in patterns indicate whether the ratio for a particular comparison is greater than 1.0, less than 1.0, or relatively unchanged. Colors represent the log fold-change of the transcript for the particular comparison ranging from 0.01 to 100 (blue to red), with green representing changes near 1.0-fold. No change refers to ratios between 0.75- and 1.33-fold.

a result consistent with the poor expression of c-MAF in PBBs and CD34B cells. However, as seen in Fig 4 and Table E3, c-MAF mRNA expression is increased in G3 cells relative to G1 cells, whereas mRNA expression of SYK was unchanged. There are canonical c-MAF response element or c-MAF cis-elements within the range of ATAC-seq peaks “A,” 1, and 2. See Figs E5 and E6 for transcription factor cis-element motifs relevant to c-MAF.

The transcription factor that most closely associates with the pattern of SYK expression is KLF5. There is little known about this transcription factor, except that its target cis-element is a GC box. There are 2 such DNA sequences near the TSS for SYK (see Fig E5). The pattern for KLF5 is consistent with a repressive element. Similarly, ZNF608 shows an inverse pattern consistent with repression of SYK expression (see Figs E5 and E6). It is interesting that of all the open regions identified in basophils and B cells, only the regions in basophils show canonical sequences known to bind c-MAF, KLF5, or ZNF608 (in the latter case the motifs for ZNF608 can only be guessed from other zinc finger transcription factors^{23,24} because there has been insufficient study of ZNF608).

A potentially interesting peak is labeled “3” in Fig 2. It should be noted that the MACS2 algorithm as used does not generate a peak call for a peak of this size but is present in all 3 basophil tracks and not in B cells or CD34B cells. As noted above, this region is located around the start of exon 2 of the SYK gene. Splicing of the most proximal stretch (but most distal of the TSS) of the 5' UTR of mature SYK mRNA occurs in this region, and therefore it might represent a region that provides an opportunity for regulation. For example, there is a canonical c-MAF response element located in this region (see Fig E6). Additional study would be needed to draw stronger conclusions regarding its presence or linkage to SYK expression and why it is more evident in mature PBBs and not CD34B cells or B cells (given that SYK transcription would require splicing in this region). The very strong peaks associated only with B cells are interesting, although given the similar expression of SYK in B cells and CD34B cells, it is not clear how these regions are influencing B cells to produce SYK levels greater than those in PBBs but only similar those in CD34B cells.

These studies also provide a more complete perspective on the CD34B cell model of basophil development. It has been known for some time that the cultures were imperfect in producing a cell with all the characteristics of mature PBBs. These results expand on the prior impressions with transcriptome and ATAC-seq information; both indicate that there are many similarities to PBBs, but the model remains imperfect. However, the studies also add a new tool, G3 cultures, to explore regulation of SYK expression. These studies identify a potentially unique region of regulation in the SYK gene in PBBs and expose 3 potential transcription factors, c-MAF, KLF5, and ZNF608, expression of which might modulate expression of SYK and might be linked to basophil differentiation.

We thank Valerie Alexander for her excellent technical assistance and Dr John Schroder for supplying samples of 8-week CD34-derived human mast cells for comparisons and PDCs for mRNA.

Key messages

- The tyrosine kinase SYK, a critical early signal transduction element in IgE-mediated secretion, is regulated by 2 pathways.
- Flt3L and SCF downregulate SYK expression during basophil development from CD34⁺ stem cell progenitors.
- From ATAC-seq results, the SYK gene in PBBs is accessible for transcription, despite the low level of SYK expression in these cells, with 2 regions in the gene that are uniquely present when compared with B cells in which SYK is expressed at 10 times greater levels.
- The transcription factors KLF5, ZNF608, and c-MAF were identified as potential repressor elements in the low expression of SYK in PBBs.

REFERENCES

1. Vilarino N, MacGlashan D Jr. Transient transfection of human peripheral blood basophils. *J Immunol Methods* 2005;296:11-8.
2. MacGlashan DW Jr, Undem BJ. Inducing an anergic state in mast cells and basophils without secretion. *J Allergy Clin Immunol* 2008;121:1500-6.
3. Havard S, Scola AM, Kay LJ, Ishmael SS, MacGlashan DW Jr, Peachell PT. Characterization of syk expression in human lung mast cells: relationship with function. *Clin Exp Allergy* 2011;41:378-88.
4. Ishmael SS, MacGlashan DW Jr. Syk expression in peripheral blood leukocytes, CD34⁺ progenitors, and CD34-derived basophils. *J Leukoc Biol* 2010;87:291-300.
5. MacGlashan DW Jr. Relationship between Syk and SHIP expression and secretion from human basophils in the general population. *J Allergy Clin Immunol* 2007;119:626-33.
6. Ishmael S, MacGlashan D Jr. Early signal protein expression profiles in basophils: a population study. *J Leukoc Biol* 2009;86:313-25.
7. Nguyen KL, Gillis S, MacGlashan DW Jr. A comparative study of releasing and nonreleasing human basophils: nonreleasing basophils lack an early component of the signal transduction pathway that follows IgE cross-linking. *J Allergy Clin Immunol* 1990;85:1020-9.
8. Kepley CL, Youssef L, Andrews RP, Wilson BS, Oliver JM. Syk deficiency in non-releaser basophils. *J Allergy Clin Immunol* 1999;104:279-84.
9. Lavens-Phillips SE, MacGlashan DW Jr. The tyrosine kinases, p53/56lyn and p72syk are differentially expressed at the protein level but not at the mRNA level in non-releasing human basophils. *Am J Respir Cell Mol Biol* 2000;23:566-71.
10. MacGlashan DW Jr, Lavens-Phillips S. Characteristics of the free cytosolic calcium timelag following IgE-mediated stimulation of human basophils: significance for the non-releasing basophil phenotype. *J Leukoc Biol* 2001;69:224-32.
11. MacGlashan D, Miura K. Loss of syk kinase during IgE-mediated stimulation of human basophils. *J Allergy Clin Immunol* 2004;114:1317-24.
12. MacGlashan D. Autoantibodies to IgE and FcεpsilonRI and the natural variability of spleen tyrosine kinase expression in basophils. *J Allergy Clin Immunol* 2019;143:1100-7.e11.
13. MacGlashan D Jr. Expression profiling of human basophils: modulation by cytokines and secretagogues. *PLoS One* 2015;10:e0126435.
14. MacGlashan D Jr. Stability of Syk protein and mRNA in human peripheral blood basophils. *J Leukoc Biol* 2016;100:535-43.
15. Bedi R, Du J, Sharma AK, Gomes I, Ackerman SJ. Human C/EBP-epsilon activator and repressor isoforms differentially reprogram myeloid lineage commitment and differentiation. *Blood* 2009;113:317-27.
16. Denburg JA, Tanno Y, Bienenstock J. Growth and differentiation of human basophils, eosinophils, and mast cells. In: *Mast cell differentiation and heterogeneity*. New York: Raven Press; 1986. pp. 71-83.
17. Otsuka H, Dolovich J, Befus AD, Telizyn S, Bienenstock J, Denburg JA. Basophilic cell progenitors, nasal metachromatic cells, and peripheral blood basophils in ragweed-allergic patients. *J Allergy Clin Immunol* 1986;78:365-71.
18. Denburg JA, Dolovich J, Harnish D. Basophil mast cell and eosinophil growth and differentiation factors in human allergic disease. *Clin Exp Allergy* 1989;19:249-54.

19. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* 2012;22:1760-74.
20. MacGlashan DW Jr, Ishmael S, Macdonald SM, Langdon JM, Arm JP, Sloane DE. Induced loss of Syk in human basophils by non-IgE-dependent stimuli. *J Immunol* 2008;180:4208-17.
21. Lai CY, Lin SY, Wu CK, Yeh LT, Sytwu HK, Miaw SC. Tyrosine phosphorylation of c-Maf enhances the expression of IL-4 gene. *J Immunol* 2012;189:1545-50.
22. MacGlashan D Jr, White JM, Huang SK, Ono SJ, Schroeder JT, Lichtenstein LM. Secretion of IL-4 from human basophils. The relationship between IL-4 mRNA and protein in resting and stimulated basophils. *J Immunol* 1994;152:3006-16.
23. Yang S, Chen J, Guo Y, Lin H, Zhang Z, Feng G, et al. Identification of prognostic biomarkers for response to radiotherapy by DNA microarray in nasopharyngeal carcinoma patients. *Int J Oncol* 2012;40:1590-600.
24. Vandevenne M, Jacques DA, Artuz C, Nguyen CD, Kwan AH, Segal DJ, et al. New insights into DNA recognition by zinc fingers revealed by structural analysis of the oncoprotein ZNF217. *J Biol Chem* 2013;288:10616-27.

Online Supplement for:

Identifying Regulatory Pathways of SYK Expression in Human Basophils

by

Xia Peng

Mingming Zhao

Li Gao

Ranjan Sen

Donald MacGlashan, Jr.

Methods

Materials

The following were purchased: PIPES, BSA, EDTA, D-glucose, erythrosin B (Sigma-Aldrich, St. Louis, MO, USA); and StemPro-34 serum-free medium with nutrient supplement, GlutaMAX, and penicillin-streptomycin (Thermo fisher, Waltham, MA, USA); Percoll (Pharmacia, Piscataway, NJ, USA); rhIL-3 (BioSource, Camarillo, CA, USA), Flt-3 Ligand Protein and SCF (R&D system, Oakville, ON, Canada) ; 4D10 anti-human SYK Ab (Santa Cruz Biotechnology, Inc., Santa Cruz, CA, USA), anti-c-kit/CD117 (STEM CELL), anti-IL-3R/CD123 (9F5) PE-conjugated, anti-FcεRI antibody CRA-1 (eBioscience, San Diego, CA, USA), anti-CD203c (clone NP4D6, BioLegend, San Diego, CA, USA), PE anti-hu CD193 (CCR3, eBio5E8-G9-B4) (eBioscience, San Diego, CA, USA), anti-CD32b (Ab10523 Xencor, Inc. Monrovia, CA, USA (ref)), anti-CD63 antibody (clone H5C6, BD BioSciences, San Jose, CA, USA); anti-mouse IgG1-Alexa Fluor 647, anti-mouse IgG2b-Alexa Fluor 647 and anti-mouse IgG2a-Alexa Fluor 488 (Molecular Probes, Invitrogen, Carlsbad, CA, USA), human IgG (Cappel Laboratories, Malvern, Pa); human CD34 MicroBead Kit (Miltenyi Biotec, Auburn, CA, USA).

PAG buffer was prepared with 25 mM PIPES, 110 mM NaCl, 5 mM KCl, 0.1% glucose, and 0.003% HSA. PAGCM was PAG supplemented with 1 mM CaCl₂ and 1 mM MgCl₂. PAG-EDTA (ethylenediamine N, N, N', N'- tetra-acetic acid) consisted of PAG supplemented with 4 mM EDTA, elutration buffer, PAG containing 0.25% BSA.

Phosphate buffered saline (PBS; APPENDIX 2), Lysis buffer, Molecular biology-grade IGEPAL CA-630 (Sigma-Aldrich, cat. no. I8896), TD (2 × reaction buffer from Nextera kit; Illumina, cat. no. FC-121-1030), TDE1 (Nextera Tn5 Transposase from Nextera kit; Illumina, cat. no. FC-121-1030), Nuclease-free H₂O (available from various molecular biology suppliers), Qiagen MinElute PCR Purification Kit, 25 μM PCR Primer 1 [custom-synthesized by Integrated DNA Technologies (IDT); sequences provided in Buenrostro et al. (2013)], 25 μM Barcoded PCR Primer 2 (custom-synthesized by Integrated DNA Technologies (IDT); sequences provided in ¹, NEBNext High-Fidelity 2 × PCR Master Mix (New England Biolabs, cat. no. M0541), 100 × SYBR Green I (Invitrogen, cat. no. S-7563), 5% TBE polyacrylamide gel, 100-bp DNA ladder (New England Biolabs; optional).

Cell Purification

Peripheral blood basophils were enriched to purities >98% through the combined use of 2-step Percoll gradients and negative selection using the basophil reagents from StemCell Technologies and columns from Miltenyi, as described in previous publications ². Peripheral blood eosinophils were enriched > 99% using the Miltenyi reagents and columns as previously described (after harvesting from the pellet of the same 2-step Percoll gradients used for basophils) ³. Peripheral blood plasmacytoid dendritic cells purified >99% with Miltenyi reagents and columns (after harvesting the pDCs from the upper layer of the same 2-step Percoll gradients used for basophils) ³. B-cells were purified after single –step Percoll gradients and Miltenyi reagents and columns for B-cells.

Generation of CD34-derived basophil from CD34⁺- progenitors

Two sources of CD34⁺ cells were used in these studies. For most of the studies, residual leukocytes from plateletpheresis were enriched by a combination of Percoll gradient separation (53%/62% Percoll) and positive selection. The upper layer of the gradient separation was found to be enriched in CD34⁺ cells (personal communication from Dr. John Schroeder). Magnetic positive selection of CD34⁺ cells (Miltenyi) followed the manufacturers guidelines with slight modifications. Generally, no more than 1×10^9 total cells were incubated for 30 minutes at 4°C with 200 microliters of kit reagent before passage over StemCell columns. In early pilot studies, the CD34⁺ purity after column separation was determined by flow cytometry using a Miltenyi antibody designed to work with the purification reagents. Typically, post-column purities were $42 \pm 13\%$ and yields ranged from 0.2 to 3.0 million CD34⁺ cells (average of $1.0 \pm 0.7 \times 10^6$). For some preliminary studies, CD34⁺ cells were obtained from the lab of Dr. Shelley Heimfeld (Fred Hutchinson Cancer Research Center, Seattle, WA, USA) and IL-3 cultures started without further handling.

The CD34⁺ progenitors were re-suspended at density of $0.1 - 0.2 \times 10^6$ cells/ml in StemPro serum-free medium, which contained the provided nutrient supplement, 2 mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, 5 ng/ml rhIL-3, and different concentrations of Flt3L and SCF. Cells were cultured at 37°C with 5% CO₂. Cell counts and viability were checked on day 14-21 by erythrocin B staining. The cells generated are heterogeneous in size, granularity, and nuclear morphology. Although not presented in the figures, a rough set of categories (large vs. small, strong alcian blue positivity vs. weak, etc) were enumerated to monitor qualitative differences. By-and-large, there were only minor differences in these characteristics for the culture conditions tested. In many cases, a low G centrifugation (40G for 1 minute in a 70 µl volume) was used to remove a significant fraction of dead cells from the harvested cultures prior to use in subsequent assays. Previous studies of these cells established that dead cells do not contribute mRNA, therefore cell

4

lysis procedures for mRNA were based on the live cell counts. However, protein remains present in dead cells, so lysis for protein measurements was based on the total cell count.

Alcian blue staining and histamine content measurement

CD34⁺ derived basophils (CD34B) were detected by Alcian blue staining ⁴ and flow cytometry. Cells stained with blue were recognized as basophils. The cultures were heterogeneous with respect to the presence of alcian blue positivity, as described in previous studies ^{3, 5}. The morphological appearance of the cells ranged from cells that were difficult distinguish from mature peripheral blood basophils (PBB) to weakly-stained with alcian blue that were larger than typical PBB. A significant fraction of the cells were not identifiable by morphology or staining as basophils. For the purposes of accounting for the effects of cytokines in these cultures, the term ‘ABP’ (alcian-blue positive) was used. The degree of positivity varied among cells but a visual threshold of blue staining was required to fit this category of cell. Images of these cells can be found in previous studies ⁶. Cells were also lysed in 1.6% perchloric acid and analyzed for histamine content by automated fluorimetry ⁷. Histamine content is given as ‘pg/ABP cell’. This metric was intrinsically variable due to the variability in assessing alcian-blue positivity or because the cultures are variable.

Analysis of surface markers by flow cytometry

Flow cytometric measurements were performed on a BD FACSCalibur flow cytometer. Cells were fixed in 2% paraformaldehyde for 20 minutes at room temperature, and the cell suspension was stored at 4°C. Cells were blocked with 0.2 mg/ml non-specific human IgG, and labeled with either CD63 (1/250 dilution from manufacturer’s stock solution), or CD203c (1/500 dilution from

manufacturer's stock solution), or CCR3 (1/30 dilution from manufacturer's stock solution). For intracellular CD63 labeling, cells were permeabilized in Medium B (Fix/Perm kit, Caltag Laboratories, S. San Francisco, CA, USA) in the presence of anti-CD63 antibody. When CD32b receptors were detected, the "control" labeling included 20 µg/ml Ab10523, as well as 1 µg/ml Ab10523-Alexa Fluor 647, whereas detection of CD32b included only 1 µg/mL Ab10523-Alexa Fluor 647⁸. The results were expressed as net mean fluorescence intensity, the difference between CD63/CD203c/CCR3/CD32b and isotype-labeled cells.

Analysis of SYK protein expression by intracellular flow cytometry

Measurement of SYK protein expression by flow cytometry using the anti-SYK antibody 4D10 has been described previously and validated with respect to standard Western blotting⁹. Samples were fixed with Medium A (Fix/Perm kit, Caltag Laboratories, S. San Francisco, CA, USA). Cells (100,000–200,000/labeling condition) were then re-suspended in PBS containing 0.04% BSA and blocked with 0.2 mg/ml non-specific human IgG. After labeling with an appropriate gating antibody (anti-CRA-1 antibody (FceRI) and anti-CD123 antibody (IL3R)), cells were permeabilized in Medium B (Fix/Perm kit) in the presence of the anti-SYK antibody 4D10 (2 µg/ml) or a mouse IgG2a isotype control (2 µg/ml). An Alexa647-conjugated anti-mouse IgG2a secondary antibody was used for detection. SYK protein expression is reported as net mean fluorescence intensity, the difference between 4D10 and isotype-labeled cells. Flow data were normalized for variability between samples using CaliBRITE allophycocyanin (APC) calibration beads (BD Biosciences, San Jose, CA, USA).

Western blotting

Cells were lysed with 1X electrophoresis sample buffer (In Vitrogen, Tris-Glycine SDS Sample Buffer, Carlsbad, CA). Total cell counts (live+dead) were used when comparing CD34B cultures. Samples were run in a 15 well 8% tris-glycine gel with molecular weight markers with transblotting to nitrocellulose membranes. Nitrocellulose membranes were developed with 4D10 (anti-SYK) antibody and were analyzed with the KwikQuant digital imager. Previous studies developed a standard for SYK, derived from PBB, that has been calibrated against recombinant SYK^{2, 10} and this standard was included in the western blots to provide a PBB context for the CD34B cells.

Microarray analysis of cellular transcriptomes

RNA was isolated with either the RNeasy extraction kits (Qiagen, Valencia, CA, USA) according to the manufacturer's directions. Total RNA was determined both by the RiboGreen RNA Quantitation Kit (Thermo Scientific-Molecular Probes, Eugene, OR, USA) and by PicoSeries RNA analysis (Pico Series II; Agilent, Santa Clara, CA, USA). The RNA integrity of the samples was also determined with the Agilent Bioanalyzer. Microarray analysis and qPCR were based on total RNA measurements, i.e., comparative samples were equalized for their measured RNA. However, the initial lysis was based on the same number of live cells and analysis of the results considered both normalization for cell number or total RNA to insure that the conclusions were *not* dependent on the method of normalization.

These studies took place over several years and the technology for microarray analysis changed. In preliminary studies, 250 ng of total RNA were analyzed using the human HT12v4 microarray (Illumina, San Diego, CA, USA). Later analysis was done with the Affymetrix Clariom S arrays.

Statistical significance of the changes

Two methods were used to determine thresholds for significant change. The simplest was to calculate the standard deviation for the entire array, i.e., the ratio of two conditions is obtained for each transcript, the data log-transformed, the standard deviation of the ratios of all positive spots determined, and the algorithm iterated once more after removing high and low data (first moment). The resulting global standard deviation was corrected for multiple samples (the Bonferroni correction) to obtain upper and lower thresholds that would be considered statistically significant. This provided the most conservative estimate of thresholds for positivity. When there were replicate experiments, the average standard deviation of replicates was calculated and used to determine upper and lower thresholds for statistically significant changes. This calculation provided the least conservative estimate of thresholds for positivity. To provide an algorithmic estimate that was of intermediate between Bonferroni and replicate thresholds, the Benjamini-Hochberg false discovery rate (FDR) method was calculated.

ATAC-seq Analysis

The samples for ATAC-seq were prepared as described in the literature [ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide]. Briefly, 2×10^5 cells were washed with cold PBS and re-suspended in 50 μ l of cold lysis buffer. The cell pellets were harvested by centrifugation for 10 min at $500 \times g$ (4°C), and placed on ice. The transposition reaction mix was prepared according to the manufacturer's guidelines with adjustments to optimize the fragmentation of the DNA for deep sequencing. Each cell type studied was examined with a range of cell number/reagent ratios to optimize the outcome. Generally, the reaction mixture was 25 μ l TD ($2 \times$ reaction buffer from

Nextera kit), 1 µl TDE1 (Nextera Tn5 Transposase from Nextera kit) and 24 µl nuclease-free H₂O. The nuclei pellet was re-suspended in the transposition reaction mix, and incubated at 37°C for 30 min. Immediately following transposition, DNA was purified using a Qiagen MinElute PCR Purification Kit, and transposed DNA was eluted in 10 µl elution buffer (Buffer EB from the MinElute kit consisting of 10 mM Tris · Cl, pH 8). The transposed DNA fragments were amplified using PCR, and the amplified library was purified using Qiagen MinElute PCR Purification Kit, eluted in 20 µl elution buffer, and separated by gel electrophoresis. The gel was excised in the 200-1000 bp region and DNA eluted and the DNA suspended in 20 µl TE buffer. Sequencing was performed by Illumina Nextseq Sequencing using the high-output 75-cycle reagent kit for single 75bp reads. Sequencing included 150-400 million reads, depending on the number of samples (with their unique bar coding) combined for a particular run. Data analysis workflow included Bowtie2 (2.2.5), Samtools (0.1.19) and peak calling with MACS2 using the hg19 human genome build for reference, and HOMER used for peak annotation. Images of the '.bedgra' files produced from the workflow were generated with the UCSC genome browser.

Not all preparations resulted in higher quality ATACseq results. There were 4 preparations of basophils tested, 2 preparations of B-cells, 2 preparations of CD34-G1 cells and 3 preparations of CD34-G2 cells. For each cell type, one preparation was confirmatory but of insufficient quality to average with the other replicates. Peak areas were measured from the gene browser output and averaged when possible.

Data Preparation

All microarray and ATACseq data relevant to the paper are located in three repositories, The primary repository is the Gene Expression Omnibus (accession # GSExxxxx, to be provided on

publication). Additional data, including an in-house designed application for exploring some of the raw data set, are (will be at publication) available via '<http://www.basophil.net>' and '<http://162.129.217.250/basophilMicroarrays>'. The website includes a variety of files one of which provides a detailed description of the analysis methodology for the microarray results.

Results

Second culture model of basophils

There were three series of experiments exploring how to optimize the final approach. The first series examined two groups, those with IL-3 alone throughout the culture and those using Flt3L and SCF during the first 3 days. The overall cell viability increased by 40% in Group 2 (G2) compared with Group 1 (G1) (Figure E2, panel A). Possibly consistent with the increased viability and the different starting point for G2, there were greater numbers of alcian-blue positive (ABP) cells in the G2 cultures and greater total cell numbers (that somewhat decreased the ABP/live ratio in G2, data not shown).

Two time points for recovery of the cultured cells revealed that ABP cells were generated more rapidly in the G1 as evidenced by the fall-off in recovery between days 14-17 and day 21 although recovery suffered for both cultures by day 21. On day 14-17, 1.6 ± 0.2 million and 2.1 ± 0.6 million ABP cells in G1 and G2 were produced, respectively. On day 21, ABP cells decreased to 0.6 ± 0.1 million in G1, while ABP cells in G2 was reduced only slightly to 1.8 ± 0.7 million, a 1.5-2.5 fold difference with G1 (Figure E2, panels C and D). In addition, G2 basophils were morphologically (by visual assessment) more consistently similar to PBB. These results led to harvesting cultures for the remaining studies on days 14-17.

We also examined several characteristics typically assessed for basophils and found similarity between G1 and G2 cells. On day 14-17, there was no difference in histamine content (1.5 ± 0.2 ng/ABP cells vs. 2.2 ± 0.3 ng/ABP cells) and no difference in FcεRI expression (MFI: 302 ± 51 vs. 311 ± 14). In addition, there were no differences in CD32b expression (MFI: 9.3 ± 0.2 VS 10.8 ± 0.02). No CD63 could be detected on the surface of G1 or G2 cells, but it was present with intracellular staining. There was also no difference in CD203c expression (day 21 MFI: 56 ± 10 vs. 37 ± 7). CCR3 was undetectable on both G1 and G2 cells but also undetectable on purified peripheral blood basophils. However, it could be detected on basophils in whole blood. Although CCR3 is often used to identify PBB in whole blood, we found that simply enriching PBB on a single-step Percoll gradient was sufficient to down-regulate CCR3 to nearly undetectable levels. Therefore, the absence of CCR3 on the CD34B might reflect a similar technical issue with its detection on cells handled *in vitro*.

Despite the similarity in the markers measured in the studies above, SYK expression in G2 cells was reduced. The reduction ranged 7-78% relative to G1 cells. (Figure E2, panel B). This difference offered a potential model of SYK regulation in the developing basophil and we next explored the cytokine conditions that would optimize its reduction. We surmised that the high concentrations of Flt3L and SCF used in the first 3 days of culture were incompletely removed during the single wash step on day 3.

To investigate whether Flt3L and SCF after day 3 altered basophil maturation and SYK expression, cells were cultured with media containing different concentrations of Flt3L and SCF on days following the high concentrations used during the first 3 days. In addition, a control IL-3-only condition was included for comparison to prior results. The concentrations included in day 3-21 cultures ranged from 0.5-8 ng/ml for Flt3L and 0.25-4 ng/ml for SCF (Table E1) and 0 for both cytokines (after the high concentrations for days 0-3 although in this design, the washing was

intentionally more complete). The labeling was similar to **Figure E2**, G1 for the IL-3 alone cultures, and G2-1 through G2-6 for the cultures started with high concentrations of Flt3L and SCF.

The recovery results for G1 were similar to the experiments shown in Figure E2 (with a culture time fixed to days 14-17). As shown in Figure E3, SYK expression was the lowest under G2-1 conditions (Flt3L and SCF at 8 and 4 ng/ml, respectively during days 3-14) SYK was decreased by approximately 50% when comparing G2-1 to G2-6 (no continuous Flt3L and SCF – this differs from the previous experimental series because the cell washing at day 3 was more effective). We concluded that Flt3L and SCF after day 3 was important for SYK down-regulation in CD34+ derived basophils. In addition, G2-1 cells showed the highest ABP cell number, and these cultures produced 3-4 times more ABP cells than G1 or G2-6 (Figure E3).

These results showed that Flt3L and SCF after day 3 decreased SYK expression. The next experimental series was designed to determine the requirement for Flt3L and SCF during days 0-3 and whether both were needed to observe the decrease in SYK expression. Table E2 shows the design of the experiments. Figure E4 shows that the optimal design, from the perspective of producing lower SYK levels, is found with the G2-1 and G3 groups, i.e., both cytokines are necessary throughout the culture although the high levels during the first 3 days (G2-1 group) generates marginally more ABP-cells than a constant lower concentration of Flt3L and SCF (G3 group) and the higher SYK protein expression levels when Flt3L or SCF were used alone. We concluded that stimulation with Flt3L and SCF on day 0-3 is also important for SYK down-regulation in CD34 -derived basophils. In addition, both cytokines contributed to the effect. Based on the other metrics shown in Figure E4, the overall optimal modification to the culture protocol to generate two classes of CD34-derived basophils for side-by-side comparisons of SYK regulation would be G3 group vs. G1.

The inclusion of SCF in these G2/3 cultures raises the possibility that the cells generated have characteristics of mast cells. Therefore, they were examined for SCF receptor (ckit, CD117) expression and compared to 8-week culture-derived human mast cells (derived from the same CD34⁺ sources) produced by consistently high concentrations of SCF and IL-6. Our results showed that the G1 and G3 cells expressed significantly lower c-kit level (MFI: G1 (194 ± 178) G3 (59 ± 11)) than CD34⁺ derived human mast cell (MFI: 5920). The c-kit positive cell ratio was also significantly lower in G1 (22±23 %) and G3 (9±2 %) than that in mast cell (97%). This result is consistent with observations that PBB express low levels of ckit ¹¹. The microarray studies below were examined with mast cell signatures in mind, although there was no direct comparison. The CD34-derived cells, both G1 and G3 cultures, showed no MITF (geneID:4286), SIGLEC8 (geneID:27181), MRGRPX1 (117194), or CMA1 (1215). These cells did express transcripts for TPSAB1 (7177) and CPA3 (1359). Only with CPA3 was there an increase in G3 cells (≈10-fold). However, similar levels of transcripts for TPSAB1 and CPA3 were present in mature basophils.

Transcriptome Analysis

- 1) How variable were CD34B preparations?
- 2) What were the differences between CD34B and PBB.
- 3) Comparisons of CD34B to PBB, PBE, and PDC.
- 4) Comparison of G1- and G3-CD34B.

1) There was considerable heterogeneity among the preparations of CD34B. A useful context is to examine the overall coefficient of variation for all expressed transcripts in PBB and CD34B; an

inter-preparation CV. For PBB, the overall CV is 34% while for CD34B, the CV is 62%, or, from an alternative perspective, 0.32% of expressed transcripts differ by more than 4-fold among different PBB preparations while 1.08% differ >4-fold among CD34B preparations.

2) This study began with a focus on the differences in SYK expression between CD34B and PBB. It is expected that there will be other differences. There were approximately 6700 positive transcripts in CD34B and 8500 in PBB. There were approximately 1700 2.5-fold differences in transcript levels (normalized for RNA level) between G1-CD34B and PBB but given the significant variation among CD34B preparations, a more meaningful expression difference of 7-fold yields 262 differences, or 3-4% of the transcripts. These transcripts are shown in online supplementary Table E5 with a variant in the analysis discussed below and Table E6. In the context of proteins typically associated with the mature basophil, it is notable that mRNA for CCR3, and some defensins were nearly absent in the CD34B cells. CD34B cells also retain some HLA transcripts not present in PBB. However, there are also notable similarities, e.g., the presence of IL3Ralpha (IL3RA), ENPP3 (CD203c), FcεRIbeta (MS4A2) and FcεRIalpha (FCER1A) and histidine decarboxylase (HDC). The full microarray results are available online (see methods).

3) As noted previously, anecdotal observations suggested that culturing CD34B from CD34 progenitors produces some cells with eosinophil-like characteristics early in these cultures. For the preparations of PBB used in these comparisons, eosinophils (and plasmacytoid dendritic cells) were also purified from the same donors. A comparative analysis was performed two ways. For the first analysis, a list of significant differences (6.5 fold) between PBB and PBE was generated (345 transcripts). This list was used to examine the ratios of PBE/CD34B or PBB/CD34B for each list

element. If the CD34B cells were like the PBB, the PBB/CD34B ratios should be not different from 1.0 (or 0.0 on a log scale). The average log-ratio was -0.039 ± 0.040 ($p > 0.05$) while the average for PBE/CD34B was -0.459 ± 0.067 ($p < 0.0001$, $H_0 = 0.0$). This indicated that with this list of transcripts, CD34B were more similar to PBB than PBE (Figure E7). A second approach was to develop a list of significant differences (using the Bonferroni threshold of 22-fold) between PBB and CD34B (PBB/CD34B) and use this list to examine whether the microarray intensities of CD34B were more similar to PBB or PBE. Of the 133 transcripts on this list, 61 (46%) were also different to nearly the same degree in PBE (i.e., PBB = PBE). For the remainder, PBE and PBB differed in expression so it could reasonably be asked whether the CD34B looked more similar to eosinophils and, using K-means clustering, 17% did show expression levels similar to PBE. However, 37% (49/133) were either more similar to PBB than PBE or not similar to either. (Table E7). Using this list of transcripts, the difference between G3 and G1 cultures was examined. Although there were some shifts towards a PBB-like expression in this set of transcripts, many shifted further from PBB (and PBE) or did not shift significantly (see below as well).

4) A comparison of G1 and G3 transcriptomes showed some significant differences. With a comparison of 3 cultures, a 4-fold change was used to generate a list of differences. This list of 350 genes is provided in the online supplement Table E8. The results are not easily understood and an analysis by DAVID algorithms does not reveal clear patterns that fit with current understanding of basophil biology. However, from the perspective of SYK expression, it is notable that SYK mRNA levels were not different between the G1 and G3 groups despite the marked decrease in SYK protein. The average expression intensities in the microarray for each sample were within 2% of each other, indicating similar loading. To verify this result, mRNA was isolated from 3 additional experiments and was analyzed by qPCR. No change was observed in these experiments; table E4

summarizes the characteristics of these cultures and mRNA results. Other than the differences in SYK protein expression, the cultures showed similar gross characteristics. There are many different ways to examine the results from the various experiments but of particular interest is assessing the similarity of CD34B and PBB since the model was created to generate basophil-like cells. In this context, we examined the transition from D0 (CD34+ progenitors) to Day 21 CD34B (CD34B-D21/CD34-D0) and for the same chosen transcripts (that were 3-fold changed) examined the PBB/CD34B ratios. Although it is useful to note transcripts that were progressively more different in the transitions of D0 -> D21 -> PBB (e.g., a transcript drops in the transition D0->D21 and drops further with D21->PBB), only 52 transcripts were found to make significant changes with each transition, $52/619 = 8\%$. A potentially more interesting comparison were transcripts that reversed the direction of change between these two transitions. Of the 619 transitions (3-fold between D0->D21) only 196 demonstrated this behavior. However, low ratios (decreases for the D0->D21 transition) were rarely (6%) reversed during the D21->PBB transition while 43% of the increases (high ratio) for the D0->D21 transition were reversed in the D21->PBB transition. These were of interest because the result suggests that there are transcripts that are incorrectly present (or not present) in the culture-derived preparations that are meant to be a model of basophils. Alternatively, these transcripts may represent genes that are transiently needed during maturation and therefore would not be expected to be represented the same way in the mature PBB (e.g., LTC4S, gene ID 4056, expression is high in G1 cells but very low in PBB and yet it is well documented that PBB are very active secretors of LTC4, suggesting that the protein was synthesized prior to basophils entering the circulation and that mRNA for this protein was down-regulated in the circulating PBB). The list of these transcripts is noted in Table E5. It is instructive to use this short list of transcripts to ask whether the G1 to G3 difference represents further maturation towards the circulating basophil phenotype. As noted in table E3, this did not appear to be the case when taking these 153 genes

together (i.e., there was no correlation between the log(ratios) of G3/G1 vs. PBB/CD34B-G1 for these transcripts, $R = -0.15$, $p=0.066$ for $H_0:R=0.0$).

Patterns in ATACseq

Beyond the specific details about the SYK gene, the ATACseq results generated other observations. For example, for PBB, there were >100,000 peak calls. Many of these were associated with genes that had no representation in the mRNA profile. Even removing these genes from a list of peak calls, the intensity of the peak calls (either all peak cells associated with a gene or the peak calls associated only with the TSS of a gene) showed no relationship with the levels of mRNA globally ($R = 0.11$, $p = 0.066$). Yet there were many specific examples of this relationship, especially when comparing genes in PBB and B-cells where there are identifiable protein expression differences associated with the cell's phenotype (see main report).

Additional Discussion

Relevant to the findings in this study are some of the cis-elements and transcription factors that may regulate SYK gene transcription. One is an Elk-1 binding site and this TF is known to participate in C/EBPa binding, where C/EBPa is thought to be important in basophil development in mice ^{12, 13}. We have found an inverse relationship between the presence of BOB1/OCA-B and SYK expression in PBB ¹⁰ and this transcription factor (and potential cytosolic regulator of SYK protein loss), binds with Oct-1 ¹⁴, a TF that binds to 4 of the 15 highest stringency cis-elements in the SYK promoter. The third example is a newly identified cAMP-response element (CRE) ¹⁵. The bZIP TFs that bind to this element are sensitive to PKA and calcium/calmodulin-dependent kinases ^{16, 17}.

Yet another example is an AP-1(c-JUN) cis-element. The tabulated patterns for transcription factors suggest a connection with c-JUN whose expression is high in CD34B and PBE and low in basophils. Indeed, basophils don't express c-JUN protein (unpublished observations) which is curious because they do show the gene accessible in the ATACseq results. This would suggest this transcription factor may act as an enhancer. As noted previously, $\alpha\beta$ T cells use ZAP-70 instead of SYK but it is also known that TCR stimulation will increase SYK expression from its nearly zero levels. A recent study has shown that c-JUN binds to the SYK promoter at -354 (and possibly -73), that activation of T cells increases c-JUN, and that its binding results in greater SYK transcription

Figure Legends

Figure E1: Schematic of the tested cultures of CD34-derived basophils.

Figure E2: The effect of Flt3L and SCF (day 0-3) on basophil maturation. Cell viability (panel A) and SYK expression (panel B). In panels C & D, two time points were assessed for the G1 and G2 cultures with measurements of viability (panel C) and ABP cell number (panel D). Results were the average of 3 independent experiments.

Figure E3: The effect of Flt3L and SCF after day 3 on CD34⁺ derived-basophil maturation. Progenitors were separated to two groups. G1 represented cells were cultured in 5 ng/ml IL-3 for 14-17 days. The other group of cells were cultured media with 5 ng/ml IL-3 and high concentration of Flt3L (100 ng/ml) and SCF (50 ng/ml) on day 0-3, thereafter cultured in media containing with 5 ng/ml IL-3, and different concentration of Flt3L and SCF for 11-14 days. According to the Flt3L and SCF, cells were sub grouped by G2-1 (8 ng/ml Flt3L and 4 ng/ml SCF), G2-2 (4 ng/ml Flt3L and 2 ng/ml SCF), G2-3 (2 ng/ml Flt3L and 1 ng/ml SCF), G2-4 (1 ng/ml Flt3L and 0.5 ng/ml SCF), G2-5 (0.5 ng/ml Flt3L and 1 ng/ml SCF), G2-6 (0.25 ng/ml Flt3L and 0.5 ng/ml SCF), and G2-7 (0 ng/ml Flt3L and SCF). SYK expression (panel A), ABP cell number (panel B), cell viability (panel C), and histamine content (panel D) for G1, G2-1, G2-2, G2-3, G2-4, G2-5 and G2-6 cells were compared. Results were the average of 2 independent experiments.

Figure E4: Effect of incubation time and cytokines (Flt3L or SCF) on CD34⁺ derived basophils maturation. Progenitors were separated into three groups. The first group of cells was cultured in media with 5 ng/ml IL-3 on day 0-3, and then cultured with 5 ng/ml IL-3 (G1-1), and 8 ng/ml Flt3L

and 4 ng/ml SCF (G1-2) for 11-14 days. The second group of cells was cultured in media with 5 ng/ml IL-3 and high concentration of Flt3L (100 ng/ml) and SCF (50 ng/ml) on day 0-3, thereafter cultured in media containing with 5 ng/ml IL-3, 8 ng/ml Flt3L and/or 4 ng/ml SCF for 11-14 days. Cells were sub-grouped according to the cytokines added to the media by G2-1 (8 ng/ml Flt3L and 4 ng/ml SCF), G2-2 (8 ng/ml Flt3L), and G2-3 (4 ng/ml SCF). The third group of cells were cultured with 5 ng/ml IL-3 and low concentration of Flt3L (8 ng/ml) and SCF (4 ng/ml) for 14-17 days. SYK expression (panel A), alcian-blue-positive (ABP) cell number (panels B), cell viability (panel C), and histamine content (panel D) for G1-1, G1-2, G2-1, G2-2, G2-3, and G3 cells were compared. Results were average of 5 independent experiments, * $p < 0.05$.

Figure E5: ATACseq results for the region (labeled ‘A’ in Figure 2) near the SYK transcription start site (TSS). The underlying graphics diagram the gene organization derived from Genbank. The lower half of the figure presents the sequence data with highlighting explained. Light blue text background represents the annotated exon 1, dark blue text background, the TSS sequence, bold dark grey font, the known 5’UTR and red text, the known coding sequence of exon 2. The orange highlights GCbox sequence motif that may be site for KLF5. Other c-MAF or ZNF motifs are outside the sequence shown in this image. There is a 50 kbp sequence between exon 1 and 2.

Figure E6: ATACseq results for the regions labeled peak ‘2’ and peak ‘3’ (see Figure 2). The continuation of the coding sequence (from the TSS, peak ‘A’) is highlighted in light blue with the 5’UTR in bold and the protein coding sequence in red text. The green and blue highlighted regions upstream of exon 2 illustrate potential c-MAF binding sites to either a c-MARE motif or minimal c-MAF binding motifs. The orange highlights represent potential zinc-finger protein binding

sequences. Peak '3' is intronic without other identifying sequence markers.

Figure E7: Summary of the ATAC results for the 8 identified peaks in the SYK gene (see Figure 2). The ATACseq results were averaged for preparations that were of high quality; PBB (n=3), CB34B-G3 (n=2), CD34B-G1 (n=1), and B-cells (n=1).

Figure E8: Venn diagram showing the shared transcriptome elements. From the differences between PBB (peripheral blood basophils) and PBE (peripheral blood eosinophils), a list of 345 6.5-fold differences were generated and used to then explore the difference between CD34B (G1) cells and PBB or PBE. Similarity of CD34B to either PBB or PBE is measured by the log(average) ratio of the comparisons (see text).

Table E1: Design of the culture conditions for the results shown in figure E3.

Table E2: Design of the culture conditions for the results shown in figure E4.

Table E3: Pattern matching SYK expression with transcript behavior across different cell types. Using microarray results for 5 types of cells [peripheral blood basophils (PBB), peripheral blood eosinophils (PBE), peripheral blood plasmacytoid dendritic cells (PDC), CD34-derived basophils (CD34B) developed with the G1 or G3 protocols] transcripts were analyzed for 4 patterns of expression (numbered 1-4 in each section). The gene ID is listed with the short gene name in the left 2 columns and the ratio of transcript expression listed in the right 5 columns under the comparison made with the raw data. At the top of the table is listed the information developed in this or prior reports for SYK expression at the mRNA or protein levels. The arrows in the patterns indicate whether the ratio for a particular comparison is greater than 1.0, less than 1.0 or relatively

unchanged (where unchanged is roughly, 0.7 – 1.3 fold ratios). A list of transcript ratios that exceeded 2-fold changes was generated from the comparison of PBE and PBB. Using this list, 4 patterns were extracted from the comparisons of PBB to 5 other cell types (CD34-D0, CD34B-G1, CD34B-G3, PBE, PDC, where CD34-D0 = progenitors used to generate CD34B)(see text). The numbers represent the fold-change of the transcript for the particular comparison.

Table E4: Summary of results for the G1 vs. G3 comparison for SYK mRNA and protein expression and culture characteristics. mRNA measurements are typically performed after normalizing for total RNA content of the extracted samples but the total cell number accounting was also maintained to assess the ratio from normalized cell counts. The mRNA ratio is $\log(\text{ratio})$ where 0.00 = no change. Parenthetical values represent the fold change. For these experiments, the number of alcian blue positive (ABP) cells that closely resembled the size and morphology of PBB (smB) were enumerated.

Table E5: Transcripts showing a 7-fold difference (Bonferroni-adjusted difference) between CD34B-G1 cells and PBB (ratio = ratio of CD34B/PBB), ranked from lowest to highest ratios.

Table E6: Transcripts who ratio is one direction (greater or less than 1.0) when comparing CD34 – day 0 cells vs. CD34B-G1 cells but reversed when comparing CD34B-G1 cells to PBB. The third ratio is the expression in G3 vs. G1 (i.e., does G3 move in the direction of PBB?).

Table E7: The transcript intensity for PBB, PBE and CD34B-G1 cells is compared (see text). Using hierarchical clustering, this set of genes was grouped into those where the CD34B is closer in expression to PBE, CD34B is closer in expression to PBB or neither.

Table E8: Transcripts showing a 4-fold difference between CD34B-G1 and G3 cells.

References

1. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 2013; 10:1213-8.
2. MacGlashan DW, Jr. Relationship Between Syk and SHIP Expression and Secretion from Human Basophils in the General Population. *J. Allergy Clin. Immunol.* 2007; 119:626-33.
3. MacGlashan D, Jr. Stability of Syk protein and mRNA in human peripheral blood basophils. *J Leukoc Biol* 2016; 100:535-43.
4. Gilbert HS, Ornstein L. Basophil counting with a new staining method using alcian blue. *Blood* 1975; 46:279-86.
5. Ishmael SS, MacGlashan DW, Jr. Syk expression in peripheral blood leukocytes, CD34+ progenitors, and CD34-derived basophils. *J Leukoc Biol* 2010; 87:291-300.
6. Langdon JM, Schroeder JT, Vonakis BM, Bieneman AP, Chichester K, Macdonald SM. Histamine-releasing factor/translationally controlled tumor protein (HRF/TCTP)-induced histamine release is enhanced with SHIP-1 knockdown in cultured human mast cell and basophil models. *J Leukoc Biol* 2008; 84:1151-8.
7. Siraganian RP. An automated continuous-flow system for the extraction and fluorometric analysis of histamine. *Anal Biochem* 1974; 57:383-94.
8. Macglashan D, Jr., Moore G, Muchhal U. Regulation of IgE-mediated signalling in human basophils by CD32b and its role in Syk down-regulation: basic mechanisms in allergic disease. *Clin Exp Allergy* 2014; 44:713-23.
9. MacGlashan DW, Jr., Ishmael S, Macdonald SM, Langdon JM, Arm JP, Sloane DE. Induced Loss of Syk in Human Basophils by Non-IgE-Dependent Stimuli. *J Immunol* 2008; 180:4208-17.

10. Ishmael S, MacGlashan D, Jr. Early signal protein expression profiles in basophils: a population study. *J Leukoc Biol* 2009; 86:313-25.
11. Columbo M, Horowitz EM, Botana LM, MacGlashan DW, Jr., Bochner BS, Gillis S, et al. The human recombinant c-kit receptor ligand, rhSCF, induces mediator release from human cutaneous mast cells and enhances IgE-dependent mediator release from both skin mast cells and peripheral blood basophils. *J. Immunol.* 1992; 149:599-608.
12. Arinobu Y, Iwasaki H, Gurish MF, Mizuno S, Shigematsu H, Ozawa H, et al. Developmental checkpoints of the basophil/mast cell lineages in adult murine hematopoiesis. *Proc Natl Acad Sci U S A* 2005; 102:18105-10.
13. Rao KN, Smuda C, Gregory GD, Min B, Brown MA. Ikaros limits basophil development by suppressing C/EBP-alpha expression. *Blood* 2013; 122:2572-81.
14. Siegel R, Kim U, Patke A, Yu X, Ren X, Tarakhovsky A, et al. Nontranscriptional regulation of SYK by the coactivator OCA-B is required at multiple stages of B cell development. *Cell* 2006; 125:761-74.
15. Ghosh D, Kis-Toth K, Juang YT, Tsokos GC. CREMalpha suppresses spleen tyrosine kinase expression in normal but not systemic lupus erythematosus T cells. *Arthritis Rheum* 2012; 64:799-807.
16. Jhala US, Canetti G, Srean RA, Kulkarni RN, Krajewski S, Reed J, et al. cAMP promotes pancreatic beta-cell survival via CREB-mediated induction of IRS2. *Genes Dev* 2003; 17:1575-80.
17. Matthews RP, Guthrie CR, Wailes LM, Zhao X, Means AR, McKnight GS. Calcium/calmodulin-dependent protein kinase types II and IV differentially regulate CREB-dependent gene expression. *Mol Cell Biol* 1994; 14:6107-16.

18. Ghosh D, Tsokos GC, Kyttaris VC. c-Jun and Ets2 proteins regulate expression of spleen tyrosine kinase in T cells. *J Biol Chem* 2012; 287:11833-41.

Table E1

Group	G1	G2					
Subgroup	--	G2-1	G2-2	G2-3	G2-4	G2-5	G2-6
Dilution	--	1:12.5	1:25	1:50	1:100	1:200	0
Flt3L Concentration*	--	8	4	2	1	0.5	0
SCF Concentration*	--	4	2	1	0.5	0.25	0

* concentrations in ng/ml

Table E2

Group	G1		G2			G3
Subgroup	G1-1	G1-2	G2-1	G2-2	G2-3	G3
Flt3L* (Day 0-3)	0	0	100	100	100	8
SCF* (Day 0-3)	0	0	50	50	50	4
Flt3L* (after Day 3)	0	8	8	8	0	8
SCF* (after Day 3)	0	4	4	0	4	4

* concentrations in ng/ml

Table E3

Transcript	ID	CD34B/CD34-D0	CD34B-G3/CD34B-G1	CD34B/PBB	PBE/PBB	PDC/PBB
SYK (mRNA)	6850	1.4	1.2	2.4	5.8	7.7
SYK(protein)		↔(≈1.0)	↓(≈0.5)	↑(≈11)	↑(≈11)	↑(≈30)
(1) Pattern		↔	↓↔	↑	↑	↑
IMPA2	3613	0.91	0.65	2.4	10.0	2.9
LRIG1	26018	1.10	0.51	4.4	19.0	2.2
NME4	4833	0.74	0.54	3.7	3.4	6.4
SLC37A1	54020	0.74	0.18	2.2	5.4	6.0
TFRC	7037	0.72	0.22	11.0	4.5	9.4
TXN	7295	1.10	0.68	3.1	3.6	4.3
Transcript	ID	CD34B/CD34-D0	CD34B-G3/CD34B-G1	CD34B/PBB	PBE/PBB	PDC/PBB
(2) Pattern		↔	↔	↑	↑	↑
ACADVL	37	1.30	0.74	2.3	3.6	3.1
AKR7A2	8574	0.65	0.72	5.7	4.2	12.0
ALCAM	214	1.60	1.20	4.3	3.0	9.6
ALDH2	217	1.10	0.95	2.1	3.4	8.2
AMFR	267	1.60	0.89	4.8	3.8	3.3
ANKRD20A1	84210	0.89	1.80	2.1	8.9	3.3
APPL2	55198	0.98	1.40	2.1	3.3	2.5
ATP6V0E2	155066	0.61	0.97	8.8	2.9	7.2
ATPAF1	64756	0.63	0.87	2.3	3.8	3.3
C7ORF50	84310	1.40	0.93	4.2	12.0	2.1
CBLB	868	0.87	1.70	2.4	3.5	4.8
CD52	1043	1.30	0.71	3.5	18.0	2.7
CD81	975	1.10	1.10	4.8	6.1	2.1
CHCHD10	400916	1.40	1.40	2.6	3.1	8.5
CPNE3	8895	1.20	0.82	2.4	5.4	9.1
DAPK1	1612	0.55	0.97	2.1	3.6	2.5
DDIT4	54541	0.75	1.60	5.6	13.0	7.7
DENND5B	160518	0.80	1.20	2.3	3.1	10.0
DIAPH2	1730	1.00	1.80	2.4	3.6	3.0
ERP29	10961	0.88	0.61	2.1	3.0	6.6
FAM162A	26355	0.89	1.20	3.8	4.0	5.9
GCHFR	2644	1.50	0.72	3.6	3.1	2.9
GNG10	2790	1.70	1.20	3.0	5.8	2.4
GOLPH4	27333	1.50	0.81	4.7	7.5	2.5
HNRNPL	3191	0.64	0.78	2.9	4.2	6.1
ITM2C	81618	0.82	1.00	4.1	7.5	30.0
PIAS2	9063	0.62	1.10	4.0	4.0	3.4
PPP1R16B	26051	1.10	1.40	2.1	4.5	6.6
SF1	7536	0.77	1.50	2.6	6.2	4.7
SLC20A1	6574	0.62	1.00	3.4	6.7	13.0
SPNS3	201305	1.10	1.60	4.3	44.0	14.0
SYK	6850	1.40	1.20	2.4	5.8	7.7
TBC1D2B	23102	1.50	1.80	2.8	3.0	2.7

Transcript	ID	CD34B/CD34-D0	CD34B-G3/CD34B-G1	CD34B/PBB	PBE/PBB	PDC/PBB
(3) Pattern		↔	↔↑	↓	↓	↓
CTNNA2	1496	1.1	1.2	0.14	0.12	0.10
CYP11A1	1583	1.9	1.3	0.19	0.08	0.09
DCN	1634	1.2	1.2	0.05	0.04	0.04
DEFA4	1669	1.0	1.5	0.04	0.04	0.04
KLF5	688	1.2	1.1	0.09	0.11	0.02
OLFM4	10562	0.8	1.3	0.11	0.14	0.09
PDLIM5	10611	1.6	1.3	0.15	0.13	0.06
PERP	64065	1.0	1.4	0.13	0.15	0.10
PMM2	5373	0.6	0.9	0.05	0.04	0.12
PRSS3	5646	1.4	1.9	0.19	0.07	0.05
RGS13	6003	0.9	1.0	0.12	0.11	0.11
TEX101	83639	1.2	1.8	0.04	0.03	0.03
TRIM48	79097	1.3	1.1	0.03	0.02	0.02
ZNF608	57507	0.6	1.4	0.11	0.07	0.16

Transcript	ID	CD34B/CD34-D0	CD34B-G3/CD34B-G1	CD34B/PBB	PBE/PBB	PDC/PBB
(4) Pattern		↑	↑	↓	↓	↓
MAF	4094	15.0	4.8	0.10	0.01	0.01
OSBPL1A	114876	2.7	2.5	0.22	0.05	0.13
RAB27A	5873	2.6	1.7	0.12	0.17	0.03
SERP2	387923	4.2	1.2	0.24	0.04	0.04
SMPDL3A	10924	4.4	5.0	0.07	0.04	0.02
SYTL3	94120	6.0	2.9	0.17	0.02	0.03
TMEM140	55281	1.7	2.9	0.20	0.18	0.04
VNN3	55350	1.8	2.9	0.15	0.09	0.09

Table E4

Assay	n	RNA normalized	CellCt normalized	p-value
microArray	2	0.07±0.06 (1.19)	-0.01±0.14 (0.97)	n.s.
qPCR	3	-0.03±0.30 (0.93)	0.04±0.22 (1.09)	n.s.
μA+qPCR	5	0.02±0.08 (1.05)	0.01±0.15 (1.15)	n.s.
SYK(by flow)			-0.25±0.07 (0.56)	0.021
		G1	G3	p-value
viability		66±6%	55±7%	n.s.
%ABP		88±6	84±6%	n.s.
%smB/ABP		29±6	43±9	n.s.

Table E5

GeneName	Gene ID	Ratio	GeneName	Gene ID	Ratio	GeneName	Gene ID	Ratio
GPNUMB	10457	0.00	KLHDC8B	200942	0.09	PLOD2	5352	0.14
PRG2	5553	0.01	MYC	4609	0.09	PRDX4	10549	0.14
RNASE1	6035	0.01	PEBP1	5037	0.09	QDPR	5860	0.14
ADAMDEC1	27299	0.02	RARRES1	5918	0.09	RAB3IL1	5866	0.14
APOE	348	0.02	SERPINF1	5176	0.09	SORBS3	10174	0.14
C1QB	713	0.02	SLC16A10	117247	0.09	TNFRSF21	27242	0.14
C1QC	714	0.02	SLC22A16	85413	0.09	VNN3	55350	6.9
CTSL1	1514	0.02	TFRC	7037	0.09	CTNNA2	1496	7
LGMN	5641	0.02	TM4SF19	116211	0.09	INDO	3620	7.1
PPAP2A	8611	0.02	TSPAN3	10099	0.09	S100P	6286	7.1
PTGDS	5730	0.02	ATP9A	10079	0.1	CCL5	6352	7.2
SCD	6319	0.02	C11ORF17	56672	0.1	CLEC4A	50856	7.2
SLCO2B1	11309	0.02	CCDC50	152137	0.1	EPSTI1	94240	7.3
PLA2G7	7941	0.03	CCL23	6368	0.1	VCAN	1462	7.3
PRG3	10394	0.03	CCNB2	9133	0.1	FLJ22662	79887	7.4
A2M	2	0.04	CHRNA6	8973	0.1	BTN3A3	10384	7.5
BEX1	55859	0.04	EGR1	1958	0.1	TAP2	6891	7.5
CYTL1	54360	0.04	GAS6	2621	0.1	BAIAP2L2	80115	7.7
F13A1	2162	0.04	HLA-DRB3	3125	0.1	PERP	64065	7.7
LTC4S	4056	0.04	HLA-DRB5	3127	0.1	HCP5	10866	7.8
MMP9	4318	0.04	SEPP1	6414	0.1	S100A12	6283	7.9
MS4A4A	51338	0.04	SQLE	6713	0.1	TNFSF10	8743	7.9
RGL1	23179	0.04	TGFBR3	7049	0.1	RGS13	6003	8.4
SPP1	6696	0.04	TOP2A	7153	0.1	RAB27A	5873	8.5
TMEM97	27346	0.04	ZDHHC19	131540	0.1	PNPLA1	285848	8.7
ACP5	54	0.05	ABCC4	10257	0.11	ZNF608	57507	8.7
EMP1	2012	0.05	ATP6V0E2	155066	0.11	IFI44	10561	9
EPDR1	54749	0.05	CCL3L1	6349	0.11	SAMD9L	219285	9.1
FOSB	2354	0.05	CCL3L3	414062	0.11	IFITM3	10410	9.3
FPR3	2359	0.05	CLEC11A	6320	0.11	OLFM4	10562	9.3
GNPDA1	10007	0.05	EGR2	1959	0.11	S100A9	6280	9.3
HLA-DQA1	3117	0.05	FSCN1	6624	0.11	IFIT2	3433	9.5
HS3ST2	9956	0.05	FUCA1	2517	0.11	JAK2	3717	9.7
IL17RB	55540	0.05	IGF1	3479	0.11	MAF	4094	9.7
PLAU	5328	0.05	MS4A7	58475	0.11	ITGB7	3695	9.8
RGS1	5996	0.05	NAP1L1	4673	0.11	CAMP	820	10
TYMS	7298	0.05	NEK2	4751	0.11	HK3	3101	10
AVP	551	0.06	NQO1	1728	0.11	IL4	3565	10
CCL2	6347	0.06	RHOC	389	0.11	SEMA3C	10512	10
CCL3	6348	0.06	SCG2	7857	0.11	CRISPLD2	83716	11
CST7	8530	0.06	SLC44A1	23446	0.11	KLF5	688	11
CYP51A1	1595	0.06	TARP	445347	0.11	LCN2	3934	11
DFNA5	1687	0.06	ANKRD9	122416	0.12	LILRA3	11026	11
ITGB5	3693	0.06	APOC1	341	0.12	UPB1	51733	11
NGFRAP1	27018	0.06	CDC45L	8318	0.12	CAPN12	147968	12
PRSSL1	400668	0.06	CSDA	8531	0.12	KIAA0513	9764	12
SRPX	8406	0.06	CXXC5	51523	0.12	PALLD	23022	12
STAB1	23166	0.06	DPP4	1803	0.12	SORL1	6653	12
ASPM	259266	0.07	EPB41L3	23136	0.12	CORO6	84940	13

CDCA5	113130	0.07	HBG2	3048	0.12	FCN1	2219	13
DHCR7	1717	0.07	HLA-DMB	3109	0.12	IFIT3	3437	13
FABP4	2167	0.07	LAPTM4B	55353	0.12	OAS2	4939	13
FASN	2194	0.07	MARCKS	4082	0.12	MX2	4600	14
GGH	8836	0.07	MSC	9242	0.12	ST6GAL1	6480	14
HLA-DPA1	3113	0.07	NRP1	8829	0.12	FAM65B	9750	15
KIAA0101	9768	0.07	OPN3	23596	0.12	SMPDL3A	10924	15
LGALS3	3958	0.07	SLC12A8	84561	0.12	BTN3A2	11118	16
LIPA	3988	0.07	SNCA	6622	0.12	IL5RA	3568	16
MAOA	4128	0.07	STXBP6	29091	0.12	RARRES3	5920	16
MPO	4353	0.07	TGFB1	7045	0.12	LTB	4050	17
NENF	29937	0.07	TSPAN4	7106	0.12	BTN3A1	11119	19
PHGDH	26227	0.07	C10ORF61	26123	0.13	C20ORF55	83541	19
RNASE3	6037	0.07	CDKN1A	1026	0.13	DCN	1634	19
RPS4Y1	6192	0.07	DDX3Y	8653	0.13	FOLR3	2352	19
SC4MOL	6307	0.07	HIST1H4C	8364	0.13	GNG2	54331	20
UBE2C	11065	0.07	LIF	3976	0.13	CCR3	1232	21
ACSM3	6296	0.08	NDN	4692	0.13	PMM2	5373	21
CD163	9332	0.08	PDE9A	5152	0.13	SELL	6402	21
CDC2	983	0.08	QPRT	23475	0.13	APOBEC3A	200315	22
COLEC12	81035	0.08	SLC1A3	6507	0.13	FLJ14213	79899	22
DAB2	1601	0.08	SPRED1	161742	0.13	IFI44L	10964	22
HTRA1	5654	0.08	SRPK2	6733	0.13	DHRS9	10170	23
RBP1	5947	0.08	STMN1	3925	0.13	XAF1	54739	23
SIGLEC6	946	0.08	TFPI	7035	0.13	DEFA4	1669	24
SLAMF1	6504	0.08	WBP5	51186	0.13	IFIT1	3434	24
SLC38A6	145389	0.08	ANKRD57	65124	0.14	CACNG6	59285	25
TMEM51	55092	0.08	CACHD1	57685	0.14	TEX101	83639	26
ACOT7	11332	0.09	CKAP2L	150468	0.14	TRPM6	140803	27
CDC20	991	0.09	CMBL	134147	0.14	TRIM48	79097	29
ELA2	1991	0.09	ENTPD1	953	0.14	ANKRD55	79722	32
GCGR	2642	0.09	GFI1	2672	0.14	SPRYD5	84767	32
HBD	3045	0.09	HLA-DRB6	3128	0.14	ADAM19	8728	50
HLA-DRB1	3123	0.09	MAFB	9935	0.14	ABLIM1	3983	52
HMMR	3161	0.09	MATK	4145	0.14	SCGB3A2	117156	54
JARID1D	8284	0.09	MOCOS	55034	0.14	MX1	4599	59
JPH1	56704	0.09	NME1	4830	0.14	LOC728358	728358	144
JUN	3725	0.09	PLAUR	5329	0.14	DEFA3	1668	200
						DEFA1	1667	263

Table E6

GeneName	Gene ID	D21/D0	PBB/D2 1	G3/G1	GeneName	Gene ID	D21/D0	PBB/D2 1	G3/G1	GeneName	Gene ID	D21/D0	PBB/D2 1	G3/G1
MX1	4599	0.09	59	73	LILRB5	10990	5.7	0.15	1.7	RBP1	5947	12	0.08	1.2
LTC4S	4056	0.15	17	0.87	SEPP1	6414	5.7	0.10	9.2	SRPX	8406	12	0.06	1.2
RARRES3	5920	0.15	16	1.2	SNX24	28966	5.7	0.22	0.87	ANKRD33	341405	13	0.20	1.0
XAF1	54739	0.19	23	4.4	ADAP2	55803	5.9	0.43	4.6	CECR1	51816	14	0.47	2.9
REC8	9985	0.22	4.1	0.71	ASPM	259266	5.9	0.07	0.54	CTSG	1511	14	0.15	0.83
ZC3H12A	80149	0.27	1.4	1.3	LRPAP1	4043	5.9	0.17	0.99	GRN	2896	14	0.27	2.4
CCNL1	57018	0.30	2.6	1.6	ALDH3B1	221	6.2	1.2	0.79	PLIN2	123	14	0.22	1.4
IL15	3600	0.33	3.0	0.36	CD33	945	6.3	1.3	1.3	SIGLEC6	946	14	0.08	0.99
NFKB2	4791	0.34	1.6	1.0	LYPD6B	130576	6.4	0.47	1.4	CD48	962	15	1.2	0.85
FYN	2534	0.35	3.2	0.42	RARRES1	5918	6.5	0.09	1.6	FPR3	2359	15	0.05	7.2
ABCG1	9619	0.40	1.6	1.2	STAP1	26228	6.5	0.21	1.2	LYZ	4069	15	0.47	4.2
BIRC3	330	0.46	1.5	2.2	CCL13	6357	6.6	0.16	4.4	RGL1	23179	15	0.04	23
FOSB	2354	0.55	0.05	0.89	CTSC	1075	6.6	0.21	1.4	C1orf186	440712	16	2.0	2.0
ITPR3	3710	0.59	6.8	0.80	SMARCD3	6604	6.6	0.28	1.7	CLEC12A	160364	16	4.4	1.5
EGR1	1958	1.1	0.10	1.3	SLC39A11	201266	6.7	0.31	0.49	CPA3	1359	16	0.65	0.92
GIMAP4	55303	2.2	5.3	2.6	ARSB	411	6.8	0.49	0.71	DFNA5	1687	16	0.06	1.6
TSHZ3	57616	2.3	0.65	1.2	NRP1	8829	6.8	0.12	7.6	GNPDA1	10007	16	0.05	2.3
HS3ST1	9957	2.5	0.31	1.2	UGP2	7360	6.8	0.41	1.7	HS3ST2	9956	16	0.05	5.7
FZD2	2535	2.7	0.40	0.97	DBI	1622	6.9	0.19	1.8	HTRA1	5654	16	0.08	3.2
TLR5	7100	2.7	0.8	2.2	HAVCR2	84868	7.0	0.18	2.8	IGF2R	3482	16	0.30	2.6
LPL	4023	3.0	3.8	1.2	BOK	666	7.1	0.22	1.1	CCL2	6347	17	0.06	20
TBC1D9	23158	3.1	0.57	3.0	CD86	942	7.1	0.38	8.9	F13A1	2162	18	0.04	5.2
TLR8	51311	3.3	1.7	6.4	DPP4	1803	7.2	0.12	0.19	STAB1	23166	18	0.06	8.3
ALOX15B	247	3.4	0.25	2.4	TM6SF1	53346	7.2	0.68	1.0	APOE	348	20	0.02	0.56
CLEC5A	23601	3.4	0.41	6.6	GAS6	2621	7.4	0.10	1.8	CST7	8530	20	0.06	0.42
CYP4F12	66002	3.5	2.5	2.0	SLC11A1	6556	7.4	1.30	1.40	MS4A2	2206	20	3.8	1.9
FARP1	10160	3.6	0.22	2.5	HSD17B4	3295	7.5	0.62	1.3	MS4A3	932	22	3.5	0.72
TREM2	54209	3.6	0.24	2.1	MT1F	4494	7.5	0.65	0.95	MS4A4A	51338	22	0.04	7.3
SLAMF8	56833	3.9	0.23	2.0	CAPG	822	7.6	0.51	0.90	SPP1	6696	22	0.04	4.2
DTNA	1837	4.3	0.19	5.7	FCGRT	2217	7.7	0.58	2.8	A2M	2	26	0.04	9.6
MSR1	4481	4.3	0.24	4.9	SERPINF1	5176	7.7	0.09	5.7	CD24	100133941	26	0.16	0.03
RAB3IL1	5866	4.3	0.14	1.0	IGF1	3479	7.9	0.11	7.1	MAFB	9935	26	0.14	3.8
M6PR	4074	4.4	0.64	1.2	TM4SF19	116211	7.9	0.09	1.2	LGMN	5641	30	0.02	12
TCN2	6948	4.4	0.18	6.7	ZDHHC19	131540	7.9	0.10	0.96	MMP9	4318	30	0.04	15
TNFSF13	8741	4.4	0.33	1.0	ADORA3	140	8.0	0.16	2.2	MS4A7	58475	31	0.11	14
GM2A	2760	4.5	0.20	1.6	IL13RA1	3597	8.1	0.33	17	ADAMDEC1	27299	32	0.02	15
RDH10	157506	4.5	0.28	0.25	SCG2	7857	8.2	0.11	2.4	ELANE	1991	32	0.09	248
MT2A	4502	4.8	0.60	1.4	CHRNA6	8973	8.4	0.10	1.9	LY86	9450	33	0.04	28
TESC	54997	4.8	0.16	0.3	COLEC12	81035	8.4	0.08	1.5	PRG3	10394	33	0.03	0.68
GPR160	26996	4.9	0.37	1.2	TDRD9	122402	8.5	0.49	2.3	PTGDS	5730	41	0.02	4.4
SPIRE1	56907	5.0	0.28	1.4	ADRB2	154	8.7	0.32	0.82	SLCO2B1	11309	45	0.02	5.5
ABCC3	8714	5.1	0.18	0.21	ENTPD1	953	8.8	0.14	1.2	C1QC	714	47	0.02	6.3
AIF1	199	5.2	0.44	2.5	AZU1	566	9.1	0.25	36	C1QB	713	51	0.02	7.2
CD68	968	5.2	0.32	2.9	P2RY2	5029	9.3	0.21	0.77	CD163	9332	61	0.08	3.9
ITGB2	3689	5.2	0.87	1.8	TMEM176A	55365	9.8	0.23	1.1	RNASE3	6037	61	0.07	1.7
SIGLEC10	89790	5.4	0.65	0.84	FGL2	10875	9.9	1.1	51	PRG2	5553	89	0.01	0.40
TOP2A	7153	5.4	0.10	0.56	MT1X	4501	9.9	0.18	1.5	RNASE1	6035	95	0.01	5.3
FAM46A	55603	5.5	0.63	1.4	CD36	948	10	0.16	0.34	GPNMB	10457	152	0.00	3.2
HEXB	3074	5.5	0.33	3.3	HMOX1	3162	10	0.25	3.3	RNASE1	6035	95	0.01	5.3
PRKCA	5578	5.5	0.18	1.1	ITGB5	3693	11	0.06	0.93	GPNMB	10457	152	0.00	3.2
CROT	54677	5.6	0.30	1.4	JPH1	56704	12	0.09	0.85	GPNMB	10457	152	0.00	3.2
LTB	4050	5.6	0.50	1.1	LY96	23643	12	0.27	1.4					

Table E7

GeneName	GeneID	Spot Intensity			GeneName	GeneID	Spot Intensity		
		PBB	CD34B	PBE			PBB	CD34B	PBE
CD34B more similar to PBE					CD34B not similar to either (con't)				
ABLIM1	3983	8411	149	168	IL4	3565	10382	938	104
AKR7A2	8574	322	2013	1515	JUN	3725	359	3603	1867
BAIAP2L2	80115	1496	182	135	LTC4S	4056	401	9830	1363
CAPN12	147968	1922	183	211	LIPA	3988	485	9102	1260
CCL23	6368	423	5210	6760	NENF	29937	281	4185	674
CD74	972	1973	13460	14797	NPL	80896	605	4126	1695
CEACAM6	4680	1382	161	152	PALLD	23022	6867	588	137
CEACAM8	1088	4237	307	212	PEBP1	5037	199	2373	399
ENTPD1	953	290	2359	2727	PLAUR	5329	209	1375	542
FCN1	2219	8116	809	1136	RARRES3	5920	5423	402	1410
FOSB	2354	1376	21867	24770	RGS1	5996	450	8330	27888
GNLY	10578	1572	252	236	RPS4Y2	140032	232	1455	278
HLA-DMB	3109	926	12577	7220	SELL	6402	5905	430	4370
HLA-DRA	3122	2582	26366	25824	SEMA3C	10512	6210	513	186
KLF5	688	8722	638	872	SPRYD5	84767	26809	925	268
MAF	4094	20697	2100	2070	SUMO3	6612	734	4867	2717
PDLIM5	10611	2351	315	353	TFRC	7037	959	10057	4797
PERP	64065	1585	199	190	TNFSF10	8743	9383	1158	469
PMM2	5373	5344	248	214	TRPM6	140803	6179	273	1421
RAB27A	5873	5028	635	855	CD34B more similar to PBB				
SMPDL3A	10924	10101	504	303	CD163	9332	356	7094	85
ST6GAL1	6480	851	127	196	CPVL	54504	797	4997	104
ZNF608	57507	1990	217	119	CST3	1471	1523	13868	137
CD34B not similar to either					CYP1B1	1545	501	3949	85
APOBEC3A	200315	7200	251	1584	DAB2	1601	340	3096	129
C20ORF55	83541	15287	808	2446	HAVCR2	84868	612	4330	110
CAMP	820	5881	283	2529	IL17RB	55540	1244	19489	87
CCL3L1	6349	344	2288	895	LTB	4050	7892	401	17363
CLEC4A	50856	5471	864	328	MAFB	9935	503	3635	121
FSCN1	6624	333	2963	699	MARCKS	4082	763	7112	122
HCP5	10866	9575	1497	3255	MPO	4353	783	6966	117
HLA-DPA1	3113	675	15032	1417	MS4A7	58475	530	6611	64
HLA-DMA	3108	898	7586	2874	PRSSL1	400668	250	3864	100
HLA-DRB1	3123	526	5867	1577	SC4MOL	6307	1089	9996	348
HLA-DRB3	3125	235	2769	659	SNCA	6622	239	1831	72
HLA-DRB5	3127	449	6212	899	TGFBI	7045	1471	12818	88
IFI30	10437	375	2493	5992	TGFBR3	7049	1267	7875	127

Table E8

GeneName	GeneID	G3/G1	GeneName	GeneID	G3/G1	GeneName	GeneID	G3/G1
GYPB	2994	0.01	NFIB	4781	0.13	SLC38A5	92745	0.2
GYPE	2996	0.01	PKLR	5313	0.13	ABCC3	8714	0.21
HBG2	3048	0.01	DAG1	1605	0.14	ALAD	210	0.21
F2RL2	2151	0.02	GP6	51206	0.14	ANGPT1	284	0.21
GYPA	2993	0.02	GUCY1B3	2983	0.14	EDA2R	60401	0.21
RHAG	6005	0.02	PDLIM1	9124	0.14	EXO1	9156	0.21
CD24	100133941	0.03	TRIM6	117854	0.14	FAM83D	81610	0.21
HBD	3045	0.03	WFDC1	58189	0.14	ISYNA1	51477	0.21
KLF1	10661	0.03	XK	7504	0.14	KCNK5	8645	0.21
CMTM5	116173	0.04	AMHR2	269	0.15	MPP7	143098	0.21
KCNH1	3756	0.04	CHML	1122	0.15	MTSS1	9788	0.21
OSBP2	23762	0.04	CLEC4D	338339	0.15	NFIX	4784	0.21
ANK1	286	0.05	CMBL	134147	0.15	PPT2	9374	0.21
CA1	759	0.05	CNN3	1266	0.15	RXFP1	59350	0.21
FHL2	2274	0.05	EYS	346007	0.15	ABCG2	9429	0.22
HBA1	3039	0.05	MYH10	4628	0.15	CRABP1	1381	0.22
HBA2	3040	0.05	SCN9A	6335	0.15	ESR2	2100	0.22
OPN3	23596	0.05	SERPINE2	5270	0.15	HHAT	55733	0.22
SPTA1	6708	0.05	STON2	85439	0.15	PCYT1B	9468	0.22
ADD2	119	0.07	ABCB6	10058	0.16	SPAG5	10615	0.22
DMD	1756	0.07	TRIB2	28951	0.16	TAL1	6886	0.22
LEPR	3953	0.07	ABCC4	10257	0.17	TFRC	7037	0.22
ADGRG1	1	0.08	CDH1	999	0.17	TSPAN9	10867	0.22
CFH	3075	0.08	HEMGN	55363	0.17	ZNF385D	79750	0.22
EPB42	2038	0.08	LXN	56925	0.17	ANXA3	306	0.23
HBB	3043	0.08	PDE9A	5152	0.17	CCDC34	91057	0.23
MYL4	4635	0.08	SOS1	6654	0.17	CTDSPL	10217	0.23
RAP1GAP	5909	0.08	TANC1	85461	0.17	FREM1	158326	0.23
C6orf25	80739	0.09	AHSP	51327	0.18	GP1BA	2811	0.23
CXCR1	3577	0.09	LAPTM4B	55353	0.18	HMBS	3145	0.23
APOC1	341	0.1	PRKAR2B	5577	0.18	ITGA2B	3674	0.23
ATP7B	540	0.1	SLC22A17	51310	0.18	MED12L	116931	0.23
FAM171A1	221061	0.1	SLC37A1	54020	0.18	MEX3B	84206	0.23
PKIG	11142	0.1	C2orf88	84281	0.19	MLLT3	4300	0.23
ALAS2	212	0.11	CCL23	6368	0.19	PLOD2	5352	0.23
APBB2	323	0.11	CXADR	1525	0.19	REXO2	25996	0.23
F2R	2149	0.11	DPP4	1803	0.19	WASF3	10810	0.23
MPL	4352	0.11	ESAM	90952	0.19	APOBEC3B	9582	0.24
RAB6B	51560	0.11	GMPR	2766	0.19	FAM69B	138311	0.24
ST6GAL2	84620	0.11	GUCY1A3	2982	0.19	GAS2	2620	0.24
DNAJC6	9829	0.12	IL9R	3581	0.19	HDAC7	51564	0.24
KEL	3792	0.12	LHFP	10186	0.19	KDELRL3	11015	0.24
MMRN1	22915	0.12	PPP1R14A	94274	0.19	KIAA1211	57482	0.24
RCL1	10171	0.12	TRIM16L	147166	0.19	LCN2	3934	0.24
SLC25A37	51312	0.12	ABLM1	3983	0.2	MLLT11	10962	0.24
EHD2	30846	0.13	B4GALT2	8704	0.2	MUC1	4582	0.24
HPSE	10855	0.13	ME3	10873	0.2	NEXN	91624	0.24
LRRC16A	55604	0.13	RAB38	23682	0.2	STEAP3	55240	0.24

Table E8 con't

GeneName	GeneID	G3/G1	GeneName	GeneID	G3/G1	GeneName	GeneID	G3/G1
STOM	2040	0.24	MAF	4094	4.8	CHN2	1124	6
TIMM13	26517	0.24	SELL	6402	4.8	NAIP	4671	6.1
TOM1L1	10040	0.24	TPSD1	23430	4.8	C1QC	714	6.3
TUBB1	81027	0.24	DOCK4	9732	4.9	CD1E	913	6.3
ZMAT5	55954	0.24	HLA-DQA2	3118	4.9	CXCL16	58191	6.3
CDC20	991	0.25	IGSF6	10261	4.9	MAML2	84441	6.3
MYCT1	80177	0.25	MSR1	4481	4.9	CCNA1	8900	6.4
NCKAP1	10787	0.25	UGT2B11	10720	4.9	FAM105A	54491	6.4
PTRF	284119	0.25	CCR1	1230	5	RIN2	54453	6.4
RDH10	157506	0.25	CPM	1368	5	TLR6	10333	6.4
ROBO1	6091	0.25	HLA-DPB1	3115	5	TLR8	51311	6.4
SELP	6403	0.25	KERA	11081	5	ACPP	55	6.5
SLC25A15	10166	0.25	PKIB	5570	5	FGFR2	2263	6.5
THRB	7068	0.25	RASGEF1B	153020	5	CLEC5A	23601	6.6
TMEM71	137835	4	SERPINA1	5265	5	CMKLR1	1240	6.7
CLEC2D	29121	4.1	SMPDL3A	10924	5	SGK1	6446	6.7
DUSP16	80824	4.1	IFI16	3428	5.1	TCN2	6948	6.7
LILRB4	11006	4.1	PID1	55022	5.1	CYTH4	27128	6.8
CD44	960	4.2	SERPINB8	5271	5.1	EMP2	2013	6.8
CD74	972	4.2	SLC22A4	6583	5.1	SERPINB2	5055	6.8
GTF2A1L	11036	4.2	TNFRSF11A	8792	5.1	AFF3	3899	6.9
HLA-DPA1	3113	4.2	TNFSF8	944	5.1	CYP2S1	29785	6.9
IGSF9B	22997	4.2	AOAH	313	5.2	ME1	4199	6.9
LILRB2	10288	4.2	BLNK	29760	5.2	MS4A14	84689	6.9
LYZ	4069	4.2	F13A1	2162	5.2	RAPH1	65059	6.9
P2RY6	5031	4.2	MYLIP	29116	5.2	IL18BP	10068	7
PIK3R5	23533	4.2	SLC37A2	219855	5.2	ITGAX	3687	7
SPP1	6696	4.2	PZP	5858	5.3	IGF1	3479	7.1
CAPN2	824	4.3	RNASE1	6035	5.3	RGS2	5997	7.1
LYST	1130	4.3	TLR2	7097	5.3	C1QB	713	7.2
NHSL2	340527	4.3	ARAP2	116984	5.5	FPR3	2359	7.2
NRP2	8828	4.3	SLCO2B1	11309	5.5	PTPRE	5791	7.2
CCL13	6357	4.4	P2RY13	53829	5.6	MS4A4A	51338	7.3
EVI2A	2123	4.4	PROM1	8842	5.6	SLC6A13	6540	7.3
HLA-DQA1	3117	4.4	TMCC3	57458	5.6	CSF3R	1441	7.6
IQCJ	654502	4.4	CRLF2	64109	5.7	NRG1	3084	7.6
LIPA	3988	4.4	CST3	1471	5.7	NRP1	8829	7.6
PCNXL2	80003	4.4	DTNA	1837	5.7	AMICA1	120425	7.7
PTGDS	5730	4.4	HS3ST2	9956	5.7	TNFRSF1B	7133	7.8
TNFAIP2	7127	4.4	PTPRM	5797	5.7	CLDN10	9071	7.9
XAF1	54739	4.4	SERPINF1	5176	5.7	PLA2G7	7941	7.9
ADAP2	55803	4.6	ASGR2	433	5.8	TPSB2	64499	8.2
FCGR1A	2209	4.6	BCL6	604	5.8	IFI6	2537	8.3
PRDM1	639	4.6	ENPP2	5168	5.8	RAB31	11031	8.3
SLC1A3	6507	4.6	GPR183	1880	5.8	STAB1	23166	8.3
CD209	30835	4.7	HLA-DMB	3109	5.8	MARCH1	55016	8.4
OLFML2B	25903	4.7	HNMT	3176	5.9	NRG4	145957	8.4
GPR18	2841	4.8	IFI44	10561	5.9	CTSH	1512	8.6

Table E8 con't

GeneName	GeneID	G3/G1	GeneName	GeneID	G3/G1	GeneName	GeneID	G3/G1
RBM47	54502	8.7	MRC1	4360	14	IL1R2	7850	24
SBF2	81846	8.7	MS4A7	58475	14	TGFBI	7045	24
VSIG4	11326	9	ADAMDEC1	27299	15	SAMHD1	25939	25
IL1RN	3557	9.1	IL6R	3570	15	LY86	9450	28
SEPP1	6414	9.2	MMP9	4318	15	S100A12	6283	28
OAS1	4938	9.3	CSF1R	1436	16	FCER2	2208	29
TNFSF13B	10673	9.5	FCAR	2204	16	FCER2	2208	29
A2M	2	9.6	EPYC	1833	17	MPEG1	219972	33
SLC2A5	6518	9.8	IL13RA1	3597	17	AZU1	566	36
TPSAB1	7177	10	ABCA1	19	18	IRF8	3394	42
FLT3	2322	12	TFEC	22797	18	CD200R1L	344807	50
GAS7	8522	12	MX2	4600	19	FGL2	10875	51
LGMN	5641	12	CCL2	6347	20	IFI44L	10964	51
SDC2	6383	12	CLEC10A	10462	20	VNN2	8875	64
SORL1	6653	12	EPB41L3	23136	20	MX1	4599	73
CCR3	1231	13	MS4A6A	64231	20	PRTN3	5657	119
CD14	929	13	OAS2	4939	20	S100A9	6280	141
CLEC7A	64581	13	CD180	4064	22	VCAN	1462	148
RAB7B	338382	13	CD1C	911	22	S100A8	6279	165
RNASE6	6039	13	GLIPR1	11010	22	MPO	4353	223
CYBB	1536	14	RGL1	23179	23	ELANE	1991	248