nature immunology

# Expression profiling of constitutive mast cells reveals a unique identity within the immune system

Daniel F Dwyer[1,2], Nora A Barrett[1,2,4], K Frank Austen[1,2,4] & The Immunological Genome Project Consortium[3]

**Mast cells are evolutionarily ancient sentinel cells. Like basophils, mast cells express the high-affinity receptor for immunoglobulin E (IgE) and have been linked to host defense and diverse immune-system-mediated diseases. To better characterize the function of these cells, we assessed the transcriptional profiles of mast cells isolated from peripheral connective tissues and basophils isolated from spleen and blood. We found that mast cells were transcriptionally distinct, clustering independently from all other profiled cells, and that mast cells demonstrated considerably greater heterogeneity across tissues than previously appreciated. We observed minimal homology between mast cells and basophils, which shared more overlap with other circulating granulocytes than with mast cells. The derivation of mast-cell and basophil transcriptional signatures underscores their differential capacities to detect environmental signals and influence the inflammatory milieu.**

© 2016 Nature America, Inc. All rights reserved.

The Immunologic Genome (ImmGen) Project is a consortium of immunologists and computational biologists who seek to determine the gene-expression patterns that characterize the mouse immune system through rigorously standardized cell-isolation protocols and data-analysis pipelines[1]. Tissue-resident mast cells and circulating basophils are granulocytes traditionally associated with type 2 inflammation and host defense against helminthic infection[2]. Here we assessed the gene-expression profiles associated with these populations and placed them within the broader context of the immune system using the power of the ImmGen compendium.

Mast cells are evolutionarily ancient cells that date back at least as far as urochordates[3,4], which predates the emergence of adaptive immunity. Mast cells are morphologically distinct tissue-resident sentinel cells densely packed with secretory granules containing preformed mediators, including histamine, TNF, serotonin and a broad range of mast-cell-specific serine proteases bound to a proteoglycan core with heparin glycosaminoglycans[5]. Granule release following mast-cell activation is accompanied by the generation of pro-inflammatory leukotrienes, prostaglandins, chemokines and cytokines[5,6]. This array of mediators is central to the mast cell's sentinel function in mediating host resistance to bacteria, multicellular parasites and xenobiotic venoms[7–9]. Mast cells can be activated through pattern-recognition receptors[9] or tissue damage[10,11] and express FcεR1 and Fcγ receptors, which allows them to respond to targets of the adaptive immune system[2].

Mast cells are found in two main peripheral tissue compartments. Mucosal mast cells, absent in T cell–deficient humans and mice[12], arise from bone marrow (BM)-derived agranular mast-cell progenitors. These progenitors constitutively home to the intestinal mucosa[13] and are further recruited to the intestine[14] and lung[15] during T cell–mediated inflammation, which directs their maturation into granulated mucosal mast cells[16]. In contrast to mucosal mast cells, connective-tissue mast cells are constitutively present in most connective tissues[17] and are seeded during embryogenesis by circulating progenitors derived from the fetal liver[18]. BM-transfer experiments in adult mice have shown poor engraftment of donor-derived mast cells in connective tissues relative to their recruitment to mucosal sites[19], which suggests that the connective-tissue mast-cell compartment is maintained through longevity or self-renewal rather than replacement by BM-derived precursor cells. While studies have indicated that mast-cell expression of proteases[16,20] and receptors[21] is heterogeneous and is regulated by the tissue microenvironment, the full degree of mast-cell heterogeneity across different tissues is unknown.

Compared with mast cells, basophils are smaller circulating cells with multi-lobular nuclei and fewer, smaller cytoplasmic granules containing histamine and a restricted protease profile[22,23]. Basophils infiltrate peripheral tissue during allergic inflammation[24] and, like mast cells, express FcεR1. Signaling through FcεR1 induces basophil degranulation, accompanied by the rapid generation of leukotrienes and cytokines, including interleukin 4 (IL-4) and IL-13 (refs. 25,26). Unlike connective-tissue mast cells, circulating basophils are short-lived, with a half-life of several days in the periphery[27], and are actively replenished from a progenitor cell[28]. Due to their FcεR1 expression and mediators produced, mast cells and basophils have been believed to be closely related.

The contribution of mast cells to inflammation and immunity has been studied in mouse strains with mutations in the gene encoding the stem-cell-factor receptor c-Kit, which are mast cell deficient, in mice lacking mast-cell-specific proteases and in mice with Cre-recombinase-mediated deletion of mast cells or mast-cell-associated proteins[2,29]. In some cases, newer genetic approaches have supported previous findings, confirming important roles for mast

[1]Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston, Massachusetts, USA. [2]Harvard Medical School, Boston, Massachusetts, USA. [3]Full list of members and affiliations appears at the end of the paper. [4]These authors jointly supervised this work. Correspondence should be addressed to N.A.B. (nbarrett@partners.org) or K.F.A. (fausten@research.bwh.harvard.edu).

Received 18 August 2015; accepted 23 March 2016; published online 2 May 2016; doi:10.1038/ni.3445

cells in IgE-dependent local and systemic anaphylaxis[29], uric-acid-crystal-induced arthritis[30], sensitization to food allergen[31] and resistance to animal venom[32]. In other models, such as contact hypersensitivity[33], the Cre-mediated deletion of cells expressing mast cell protease 5 has contradicted early findings obtained with c-Kit-mutant strains by establishing a pro-inflammatory role for mast cells in sensitization to contact allergens. Such discrepant findings might reflect differences in protocols, the influence of *Kit* mutation beyond the mast-cell compartment or differential deletion of mast cell subsets in these strains. Additionally, some mast-cell-associated proteins, such as carboxypeptidase A3, used to direct Cre expression for the generation of the mast-cell-deficient Cre-Master and Hello Kitty strains, have been detected in basophils[34], which are reduced in number in these strains. Thus, defining the genes and pathways uniquely or dominantly expressed in mast cells relative to their expression in other immune cells might clarify mast-cell functions, identify targets for Cre-mediated disruption and provide candidate loci for the generation of novel strains with mast-cell-specific expression of Cre.

Here we isolated constitutive connective-tissue mast cells from five distinct anatomical locations (the skin, the tongue, the esophagus, the trachea and the peritoneal cavity) and isolated basophils from two locations (the spleen and peripheral blood). Our data showed that the mast-cell transcriptome was distinct, with mast cells clustering independently from all other lymphoid and myeloid-cell populations analyzed. We found that basophils were transcriptionally closest to eosinophils and shared unexpectedly few distinct transcripts with mast cells. We demonstrated the unique transcriptional signatures of mast cells and basophils and found a small signature shared by the two cell populations. Among the mast-cell populations studied, we identified substantial heterogeneity in gene expression and found evidence for previously unappreciated connective-tissue mast-cell turnover in the periphery in the absence of tissue inflammation.

## RESULTS

### Mast cells are transcriptionally distinct among immunocytes

Mast cells were sorted on the basis of co-expression of FcεR1α and c-Kit (CD117) from the peritoneal cavity, the ear (where they reside in the dermis), the tongue (where they reside in the muscular layer), the trachea (where they reside in the submucosa and serosal tissue) and the esophagus (where they reside in the submucosa proximal to the stomach) (**Supplementary Fig. 1**). Mast cells constituted between 0.05% and 10% of $CD45^+$ cells in each compartment (**Fig. 1a**). Basophils were sorted on the basis of co-expression of FcεR1α and integrin CD49b from the spleen and peripheral blood, where they comprised 0.1% of $CD45^+$ cells (**Fig. 1a**). The gating strategy used for the isolation of mast cells (**Supplementary Fig. 2a**) and basophils (**Supplementary Fig. 2b**) was validated through histochemical staining, which indicated that the cells isolated were morphologically mast cells and basophils (**Fig. 1b**). Cells underwent enrichment to high purity through multiple rounds of sorting (**Supplementary Fig. 3**), and final purity was assessed with parallel samples (**Supplementary Table 1**). RNA extracted from sorted mast cells and basophils was assessed by microarray and compared with results in the ImmGen database obtained for immunocytes, including blood eosinophils; peritoneal macrophages and B-1a cells; and splenic dendritic cells, neutrophils, $CD4^+$ T cells, $CD8^+$ T cells, γδ T cells, B-2 cells, natural killer (NK) cells and NKT cells.

Hierarchical clustering using the top 15% of genes with the most variable expression showed that the five mast-cell populations sorted clustered separately from all other lymphoid and myeloid cells analyzed (**Fig. 1c**). Lymphoid cells and myeloid cells clustered independently, as expected, and the myeloid cluster was further divided into one group containing granulocytes (eosinophils, neutrophils and both basophil populations) and a second containing macrophages and dendritic cells (**Fig. 1c**). The distinction of mast cells among immunocytes was based on both high expression of a distinct set of genes and low expression of many other transcripts associated with other cell types. Basophils had high expression of a smaller cluster of genes that had little overlap with the transcripts for which mast cells showed enrichment (**Fig. 1c**). Principal-component analysis further highlighted the distinction of mast cells from the other cell populations profiled, with mast cells from different tissues grouping closely with each other and distantly from other myeloid and lymphoid cells (**Fig. 1d**).

The transcriptional relationships among mast cells, basophils and the other cell populations analyzed were quantified through Euclidean distance measurements (**Fig. 2a**), calculated using the top 15% of transcripts with the most variable expression (**Fig. 1c**). Among mast cells, the mast-cell subsets from the trachea, esophagus and tongue were the most similar, and the mast-cell subsets from the skin and peritoneum were the most different (**Fig. 2a**). Mast cells as a whole were closest to basophils and eosinophils and furthest from neutrophils (**Fig. 2a**). Basophils from the blood and spleen were very similar to each other and were closest to eosinophils (**Fig. 2a**). The distance between basophils and mast cells was similar to the distance between basophils and neutrophils (**Fig. 2a**). Pairwise comparison of dermal mast cells and blood basophils revealed differential expression of 2,563 transcripts at an arbitrary level of twofold or greater. Skin mast cells had higher expression of 1,428 transcripts, relative to their expression in blood basophils, whereas blood basophils had higher expression of 1,135 transcripts, relative to their expression in skin mast cells (**Fig. 2b**); this further underscored their transcriptional differences. In contrast, pairwise comparison of blood eosinophils and blood basophils revealed differential expression of 1,372 transcripts at a level of twofold or higher. Blood eosinophils had higher expression of 503 transcripts, relative to their expression in blood basophils, whereas blood basophils had higher expression of 869 transcripts, relative to their expression in blood eosinophils (**Fig. 2c**). Thus, tissue-resident mast-cell populations expressed a gene program that distinguished them from other immunocytes.

### Transcriptional signature of tissue-resident mast cells

We next identified a transcriptional signature of 128 genes whose expression was twofold or greater in mast cells relative to their expression in all other cells analyzed (**Fig. 3a**). Functional analysis with the PANTHER pathway-classification system revealed that the mast-cell signature showed the most significant enrichment for genes encoding products in the category of 'serine proteases', relative to the abundance of transcripts encoding products in other functional categories (**Table 1**). This group included transcripts for many canonical mast-cell proteases, as well as *Plau*, which encodes a urokinase-type plasminogen activator, *Adamts9*, which encodes a metalloprotease, and *C2*, which encodes complement component C2 of the classical C3 convertase (**Table 1**). Mast cells also showed enrichment for the expression of *Ctsg*, which encodes cathepsin G; expression of this gene was more than fivefold higher in mast cells than in neutrophils (**Fig. 3b**). Additional pathways that showed enrichment in the mast-cell signature included 'sulfur metabolism', which contained transcripts encoding enzymes important for heparin-sulfate biosynthesis, 'polysaccharide metabolism' and 'transferases' (**Table 1**). The last category included *Hpgds*, which encodes hematopoietic prostaglandin $D_2$ synthase, a factor important for synthesis of the mast-cell inflammatory product prostaglandin $D_2$ (**Table 1**).

© 2016 Nature America, Inc. All rights reserved.

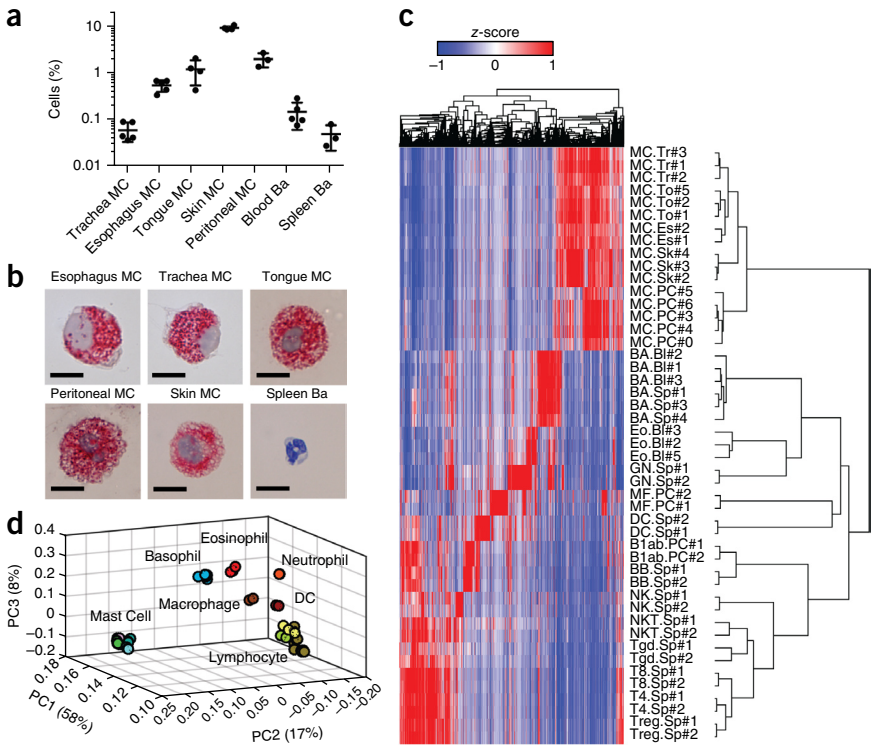



**Figure 1** Identification of mast cells as distinct from other assayed cell populations. (**a**) Frequency of mast cells (MC) and basophils (Ba) among CD45$^+$ cells in various digested tissues (horizontal axis). Each symbol represents an individual biological replicate; small horizontal lines indicate the mean (± s.d.). (**b**) Chloracetate-esterase staining identifying mast cells sorted from trachea, esophagus tongue, ear skin and peritoneal lavage fluid, and toluidine blue staining identifying basophils sorted from spleen (to confirm gating strategy used for the isolation of cell populations). Scale bars, 10 µm. (**c**) Hierarchical clustering of cell populations (right margin) using the top 15% transcripts with the most variable expression; populations include mast cells from the trachea (MC.Tr), tongue (MC.To), esophagus (MC.Es), skin (MC.Sk) and peritoneum (MC.PC), basophils from the blood (BA.Bl) and spleen (BA.Sp), blood eosinophils (Eo.Bl), splenic neutrophils (GN.Sp), peritoneal macrophages (MF.PC), splenic dendritic cells (DC.Sp), peritoneal B-1a cells (B1ab.PC), and splenic B-2 cells (BB.Sp), natural killer cells (NK.Sp), NKT cells (NKT.Sp), γδ T cells (Tgd.Sp), CD8$^+$ T cells (T8.Sp), CD4$^+$ T cells (T4.Sp) and regulatory T cells (Treg.Sp). Bar height is inversely correlated to homology between linked populations; numbers (#) indicate biological-replicate designations of the ImmGen database. (**d**) Principal-component analysis of the populations in **c**, using the top 15% transcripts with the most variable expression. Numbers in parentheses indicate frequency of transcripts described by each principal component. Data are pooled from three (peritoneal mast cell and splenic basophil), four (skin and tongue mast cell) or five (tracheal and esophageal mast cell; blood basophil) independent experiments (**a**), one experiment (**b**) or three (skin, tongue, and tracheal mast cells; splenic and blood basophils), five (peritoneal mast cells) or two (esophageal mast cells) independent experiments (**c**,**d**).

Five of the genes encoding products in the 'signal-transduction' pathway encoded members of the Mrgpr family of G-protein-coupled receptors: *Mrgpra4*, *Mrpgrb1*, *Mrgprb2*, *Mrgprx1* and *Mrgprx2*. One of these, *Mrgprb2*, has been described as the homolog to the human gene *MRGPRX2*. The protein encoded mediates mast-cell activation in response to a broad array of stimuli ranging from wasp venom to several pharmaceutical compounds associated with IgE-independent pseudoallergic reactions in patients[35]. In addition to the five Mrgpr-encoding transcripts in the mast cell signature, *Mrgprb8* and *Mrgprb13* had high expression specifically in skin mast cells, while *Mrgpra6*

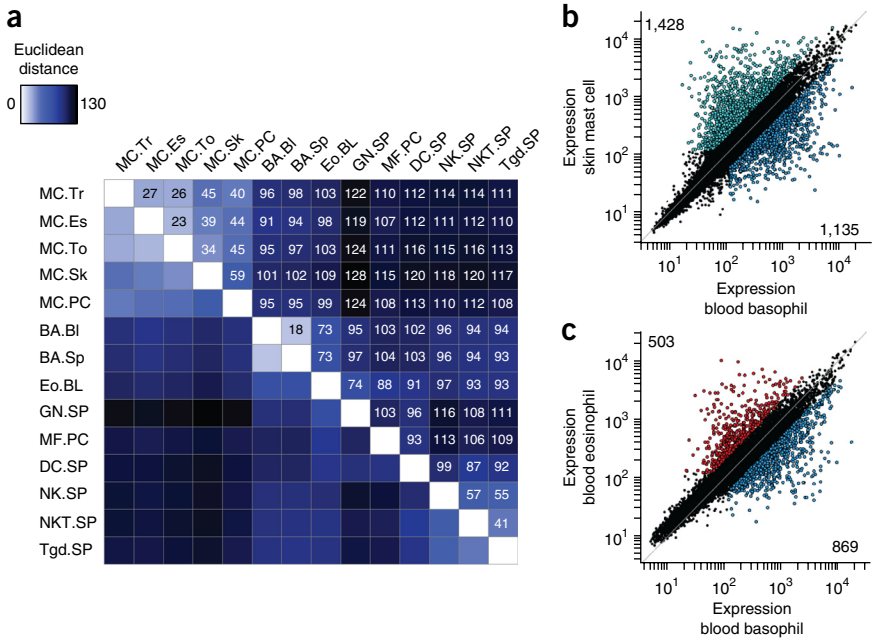



**Figure 2** Characterization of mast cells as transcriptionally distinct from basophils. (**a**) Euclidean distance matrix indicating degree of similarity between selected cell populations calculated using the top 15% transcripts with the most variable expression (identified in **Fig. 1c**); numbers in plot indicate Euclidean distance. (**b**) Gene expression in skin mast cells and blood basophils; dot colors indicate transcripts with expression values greater than 120 and expression twofold or higher in skin mast cells (aqua) or blood basophils (dark blue) than in the other population; numbers in plots indicate total genes expressed differentially in skin mast cells (top left) or blood basophils (bottom right). (**c**) Gene expression in blood eosinophils and blood basophils; dot colors indicate transcripts with expression values greater than 120 and expression twofold or higher in blood eosinophils (red) or blood basophils (dark blue) than in the other population; numbers in plots indicate total genes expressed differentially in blood eosinophils (top left) or blood basophils (bottom right). Data are pooled from three (skin, tongue and tracheal mast cells; splenic and blood basophils; blood eosinophils), five (peritoneal mast cells) or two (all other populations) independent experiments.

© 2016 Nature America, Inc. All rights reserved.

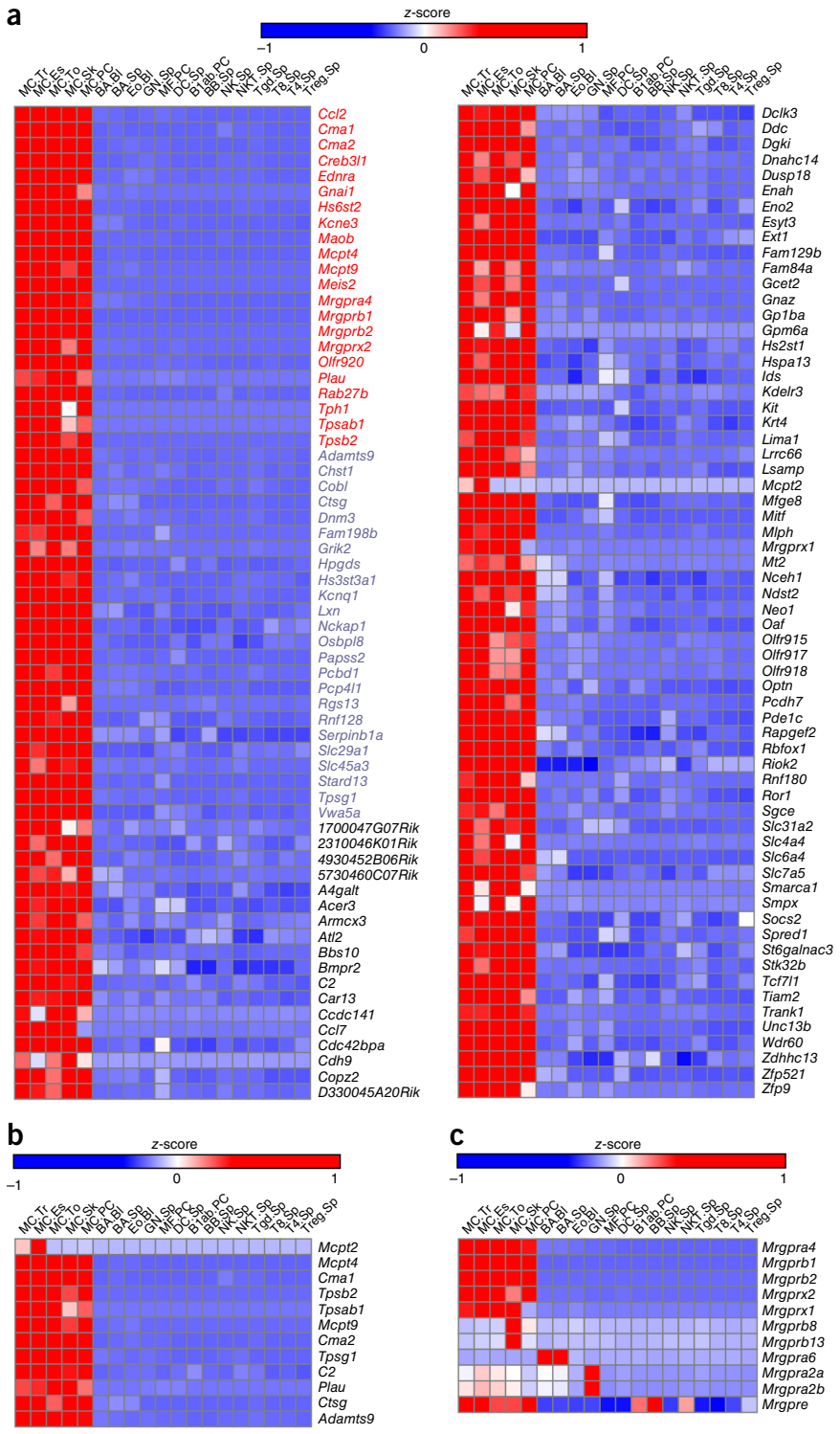



**Figure 3** Derivation of the mast-cell transcriptional signature. (**a**) Mast-cell-specific gene signature derived on the basis of transcript expression twofold or greater in all mast cell populations relative to the expression in all other cell populations analyzed ($P < 0.05$ (*t*-test)); colors of gene symbols along right margin indicate expression fivefold higher (blue) or tenfold higher (red) than that of all non–mast-cell populations. (**b**) Protease-encoding transcripts showing specific enrichment in the mast-cell signature. (**c**) Expression of Mrgpr-encoding transcripts across cell populations analyzed. Data are pooled from three (skin, tongue and tracheal mast cells; splenic and blood basophils; blood eosinophils), five (peritoneal mast cells) or two (all other populations) independent experiments.

had high expression in basophils (**Fig. 3c**). *Mrgpra2a* and *Mrgpra2b* were expressed predominantly by neutrophils, as previously reported[36], but were also detected in all mast-cell populations, and *Mrgpre* was detected in B cells and NKT cells, in addition to being detected in mast cells (**Fig. 3c**). Thus, the unique mast-cell transcriptional program included a broader degree of proteases, biosynthetic enzymes and receptors of the Mrgpr family than previously appreciated.

## Distinct and shared gene expression

A basophil transcriptional signature of 66 transcripts was similarly calculated on the basis of twofold-or-greater expression in both basophil populations relative to the expression in all other cell populations analyzed, including mast cells (**Fig. 4a**). The basophil signature contained a single protease-encoding transcript, *Mcpt8*. The basophil signature also included transcripts from several genes encoding chemokines (*Ccl3*, *Ccl4* and *Ccl9*), growth factors (*Hgf* and *Bmp4*) and adhesion proteins (*Cdh1* and *Itga1*). This suggested mechanisms through which the basophil might interact with and influence the local environment.

To better understand the relationship between mast cells and basophils, we derived a shared signature on the basis of twofold-higher expression of transcripts in all basophil and mast-cell subsets than in any other population analyzed. This analysis revealed a small shared transcriptional signature consisting of only 24 genes (**Fig. 4b**), many of which have previously been characterized in mast cells and basophils. These included *Cd200r3*, which encodes an activating receptor; *Fcer1a* and *Ms4a2*, which encode the α- and β chains, respectively, of the high-affinity receptor for IgE; *Slc24a3*, which encodes a $Ca^{2+}$ transporter; and *Gata2*, which encodes a transcription factor that directs the differentiation and function of both cell types[37]. The protease-encoding transcript *Cpa3* was also present in the shared signature, consistent with published reports of high expression of this transcript by basophils[38], in addition to mast cells.

Mast cells and basophils are well-known sources of histamine[2]. Consistent with that, the mast cell–basophil shared profile identified here included the transcript encoding Slc18a2, a solute transporter involved in loading histamine into secretory vesicles (**Fig. 4b**). Further analysis of the monoamine-biosynthetic pathways indicated that both mast cells and basophils had high expression of transcripts encoding

© 2016 Nature America, Inc. All rights reserved.

**Table 1 Functional pathway-enrichment analysis of mast-cell signature genes**

| Pathway | Genes | | | | *P* value |
|---|---|---|---|---|---|
| Serine protease (MF00216) | *Ctsg* | *Cma1* | *C2* | *Cma2* | $1.3 \times 10^{-6}$ |
| | *Mcpt2* | *Mcpt4* | *Mcpt9* | *Plau* | |
| | *Tpsab1* | *Tpsb2* | *Tpsg1* | | |
| Other transferase (MF00140) | *Ndst2* | *Chst1* | *Hs3st3a1* | *Hs2st1* | $2.3 \times 10^{-6}$ |
| | *Hs6st2* | *Hpgds* | | | |
| Other polysaccharide metabolism (BP00009) | *Ndst2* | *St6galnac3* | *Ext1* | *Hs3st3a1* | $1.9 \times 10^{-3}$ |
| | *Hs2st1* | *Ids* | | | |
| Sulfur metabolism (BP00101) | *Papss2* | *Ndst2* | *Hs3st3a1* | *Hs2st1* | $3.8 \times 10^{-3}$ |
| | *Ids* | | | | |
| Carbohydrate metabolism (BP00001) | *Ndst2* | *St6galnac3* | *A4galt* | *Chst1* | $1.1 \times 10^{-2}$ |
| | *Eno2* | *Ext1* | *Hs3st3a1* | *Hs2st1* | |
| | *Ids* | *Slc45a3* | | | |
| Signal transduction (BP00102) | *Cdc42bpa* | *Mrgprb1* | *Mrgprb2* | *Mrgprx1* | $2.6 \times 10^{-2}$ |
| | *Mrgprx2* | *Ndst2* | *Rab27b* | *Rapgef2* | |
| | *Stard13* | *Tiam2* | *Bmpr2* | *Cdh9* | |
| | *Ccl2* | *Ccl7* | *Dgki* | *Dusp18* | |
| | *Ednra* | *Grik2* | *Gp1ba* | *Gnai1* | |
| | *Gnaz* | *Hs3st3a1* | *Kit* | *Lrrc66* | |
| | *Mfge8* | *Neo1* | *Pde1c* | *Plau* | |
| | *Mrgpra4* | *Pcdh7* | *Ror1* | *Rgs13* | |
| | *Sgce* | *Socs2* | *Tph1* | | |

Identifiers in parentheses (left column) indicate pathway-module designations in the PANTHER pathway-classification system. *P* values, modification of Fisher's exact test.

a histidine transporter (*Slc3a2*) and histidine decarboxylase (*Hdc*) (**Fig. 4c**). Mast cells further expressed transcripts encoding an L-tryptophan transporter (*Slc7a5*), tryptophan hydrolase (*Tph1*) and DOPA decarboxylase (*Ddc*) (**Fig. 4c**). The mast-cell signature also included *Maob*, which encodes a monoamine oxidase (**Fig. 4c**), consistent with published reports[39]. Mast cells and basophils both expressed transcripts encoding a histamine receptor (*Hrh4*) and serotonin re-uptake transporter (*Slc6a4*) (**Fig. 4c**), while basophils expressed transcript encoding a serotonin receptor (*Htr1b*) (**Fig. 4c**).

Several transcription factors had higher expression in either mast cells or basophils than in other immunocytes. The mast-cell signature included *Creb3l1*, *Mitf*, *Smarca1* and *Zfp9* (**Fig. 4d**). Of these, to our knowledge, only *Mitf* has been previously reported in mast-cell biology, as the transcription factor encoded regulates expression of the genes encoding c-Kit and mast cell proteases[40]. The basophil signature included *Sncaip*, *Cebpa*, *Supt3h* and *Nfil3* (**Fig. 4d**). Of these, only C/EBP-α (encoded by *Cebpa*) has been previously reported to have an important role in basophil biology[28], in which it directs the commitment of progenitor cells to the basophil lineage. GATA-2 (encoded by *Gata2*) was the only transcription factor in the signature shared by mast cells and basophils (**Fig. 4d**). The diverse transcription factors, cell-surface receptors and inflammatory-cell proteins expressed by mast cells and basophils extended our earlier cluster analysis (**Fig. 1c**) and pairwise analysis (**Fig. 2b**) and indicated that these cell types are not closely related in function.

## Comparison of mast-cell and basophil signatures across species

Next we used a FANTOM consortium data set that defined the resting transcriptome of human dermal mast cells and blood basophils[41] to evaluate the mast-cell and basophil signatures across species. Human skin mast cells showed significant enrichment for expression of the mouse mast-cell signature, with 55 of the 82 mast-cell-signature genes in both data sets having twofold higher expression in human skin mast cells than in human blood basophils (**Fig. 5**). The transcripts conserved across species included those encoding proteases, hematopoietic prostaglandin $D_2$ synthase, members of the Mrgpr family, and c-Kit (**Supplementary Table 2**). Other transcripts conserved across species encoded products with less-well-defined roles; these included *Maob* and *Gnai1* (which encodes a G protein) (**Fig. 5**).

In contrast, human basophils did not show significant enrichment for the mouse basophil signature, with only 10 of the 44 signature genes present in both data sets having expression twofold higher in human-blood basophils than in human skin mast cells (**Supplementary Table 3**). Among the transcripts conserved in human and mouse basophils were those encoding the chemokines CCL3 and CCL4 (**Fig. 5**), suggestive of a shared role for basophils across species in recruiting other leukocytes to sites of inflammation. Human mast cells showed enrichment (relative to its expression in human basophils) for the signature shared by mouse mast cells and basophils, with transcripts such as *Cpa3* having expression 7.6-fold higher and *Gata2* having expression 5.2-fold higher in human mast cells than in human basophils (**Supplementary Table 4**), which again demonstrated the conserved nature of the mast-cell transcriptional program across species.

## Tissue-specific genetic programs among mast-cell populations

Next we assessed the diversity among mast-cell subsets through pairwise comparison. Because peritoneal mast cells were the only mast-cell population derived from undigested tissue, we first assessed the effect of digestion enzymes on mast-cell transcription. Enzymatic treatment of peritoneal mast cells increased the expression of 137 genes by twofold or more relative to their expression in untreated cells, including 17 genes whose expression increased five- to tenfold and 7 genes (such as *Ccl3*, *Il13* and the transcription-factor-encoding gene *Egr2*) whose expression increased more than tenfold (**Supplementary Fig. 4a**). Enzymatic digestion decreased the expression of 26 genes by twofold or more relative to their expression in untreated cells, including one transcript (*Myl1*, which encodes the myosin light-chain protein) whose expression decreased five- to tenfold (**Supplementary Fig. 4b**). None of these genes were mast-cell signature genes, and hierarchical clustering of mast cells with enzymatically treated peritoneal mast cells again demonstrated that peritoneal mast cells were the most transcriptionally distinct subset (**Supplementary Fig. 5**). However, because a subset of genes underwent transcriptional alteration by such treatment, we used enzymatically treated peritoneal mast cells for subsequent comparison with other mast cell populations.

Mast cells from the tongue showed high homology with mast cells from either the trachea or the esophagus, with only 110 genes showing a difference in expression of twofold or greater in mast cells from the tongue relative to their expression in mast cells from the trachea (**Fig. 6a**), and only 122 genes showing a difference in expression in mast cells from the tongue relative to their expression in mast cells from the esophagus (**Fig. 6b**). In contrast, mast cells from the tongue and peritoneum had differential expression of 612 transcripts (**Fig. 6c**), and mast cells from the peritoneum and skin had differential expression of 957 genes (**Fig. 6c**), which indicated these two mast-cell subsets were the most distinct in this comparison.

We next analyzed transcripts with a difference in expression of fourfold or more in single mast-cell subset relative to their expression in all

© 2016 Nature America, Inc. All rights reserved.

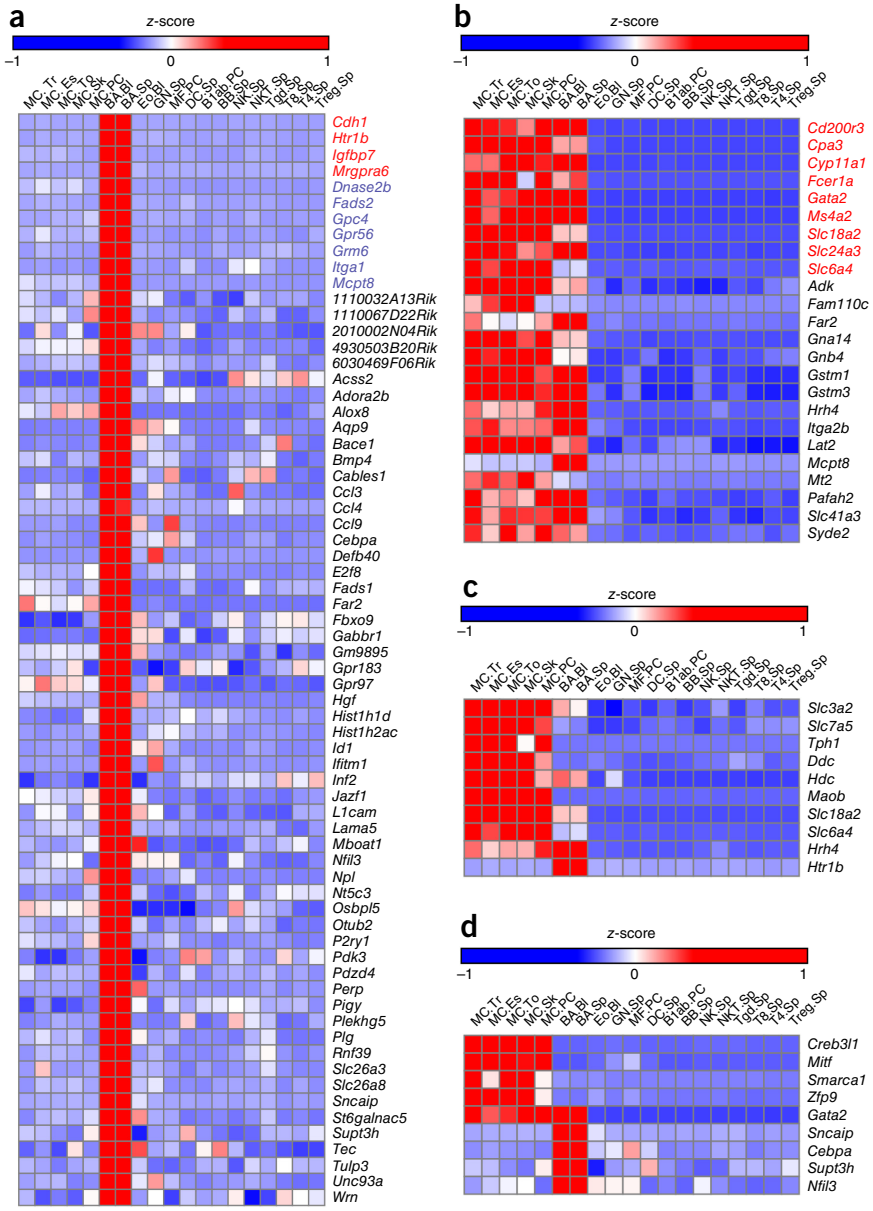



**Figure 4** Distinct and shared transcriptional expression patterns of basophils and mast cells. (**a**) Basophil-specific gene signature derived on the basis of transcript expression twofold or greater in both basophil populations relative to the expression in other cell populations analyzed ($P < 0.05$ (*t*-test)); colors of gene symbols along right margin indicate expression fivefold higher (blue) or tenfold higher (red) than that of all non-basophil cell populations. (**b**) Gene signature shared by mast cells and basophils, derived on the basis of transcript expression twofold or greater in all mast-cell and basophil populations relative to the expression in all other cell populations analyzed ($P < 0.05$ (*t*-test)); red (in gene symbols along right margin) indicates expression tenfold higher than that of all non–mast-cell or non-basophil populations. (**c**) Expression of transcripts encoding products involved in monoamine biosynthesis and neurotransmitter receptors, in mast cells or basophils; all transcripts other than *Hdc* were included in either the mast-cell-specific signature or the signature shared by mast cells and basophils. (**d**) Expression of transcripts encoding transcription factors in the distinct and shared mast-cell and basophil gene signatures. Data are pooled from three (skin, tongue and tracheal mast cells; splenic and blood basophils; blood eosinophils), five (peritoneal mast cells) or two (all other populations) independent experiments.

© 2016 Nature America, Inc. All rights reserved.

other mast-cell populations. Consistent with the transcriptional similarity of mast cells from the trachea, esophagus and tongue, mast cells from the tongue had no transcripts with a difference in expression of fourfold or more relative to their expression in the other mast-cell subsets (data not shown). The expression of five transcripts was at least fourfold higher in mast cells from the esophagus than in other mast cells; these included the protease-encoding transcripts *Mcpt1* (whose expression was limited to this subset) and *Mcpt2* (**Fig. 6d**). Tracheal mast cells had fourfold higher expression of a single transcript, *Lipf* (**Fig. 6e**). No transcript had expression fourfold lower in mast cells from the trachea or esophagus relative to its expression in other mast cell populations (data not shown). The expression of three transcripts, including *Itgb2* (which encodes integrin $\beta_2$ (ItgB2)) and *Bmp2* (which encodes a bone-morphogenic protein), was fourfold higher in peritoneal mast cells than in other mast cells (**Fig. 6f**). The expression of ten transcripts, including *Cd59a* (which encodes a membrane-attack-complex inhibitor) and *Olr1* (which encodes an oxidized lipoprotein receptor), was more than fourfold lower in peritoneal mast cells than in other mast cells (**Fig. 6f**).

Skin mast cells showed fourfold higher expression of 28 genes and fourfold lower expression of 18 genes relative to their expression in the other mast cell subsets (**Fig. 6g**). In addition to having higher expression of *Mrgprb8* and *Mrgprb13*, skin mast cells showed increased expression of transcripts encoding the metalloproteases ADAMTS1 and ADAMTS5, the cytokine and mast-cell growth factor IL-3 and the transcription factor SOX7. Skin mast cells also showed enhanced expression of *CD59a*, which suggested substantially different expression of this gene in mast cells from the skin and those from the peritoneum. Transcripts with expression fourfold lower or more in skin mast cells relative to their expression in other subsets of mast cells included *Cd34* (which encodes a canonical mast-cell marker) and *Alox5* and *Alox5ap* (which encode 5-lipoxygenase and 5-lipoxygenase-activating protein, respectively).

In support of the transcriptional data, analysis by flow cytometry indicated that the adhesion molecule CD34 was expressed on all mast-cell subsets except for skin mast cells, CD59a expression was highest on skin mast cells and was undetectable on peritoneal mast cells, and ItgB2 expression was detected only on peritoneal mast cells (**Fig. 6h**). Enzymatically treated peritoneal mast cells showed no decrease in surface staining of either CD34 or ItgB2 relative to surface staining of these proteins on their untreated counterparts (**Supplementary Fig. 6**).

Because the skin and peritoneal mast-cell populations showed the greatest degree of differential gene expression, we compared these populations by gene-set-enrichment analysis (GSEA). Among the terms of the Gene Ontology (GO) Consortium most 'enriched' in

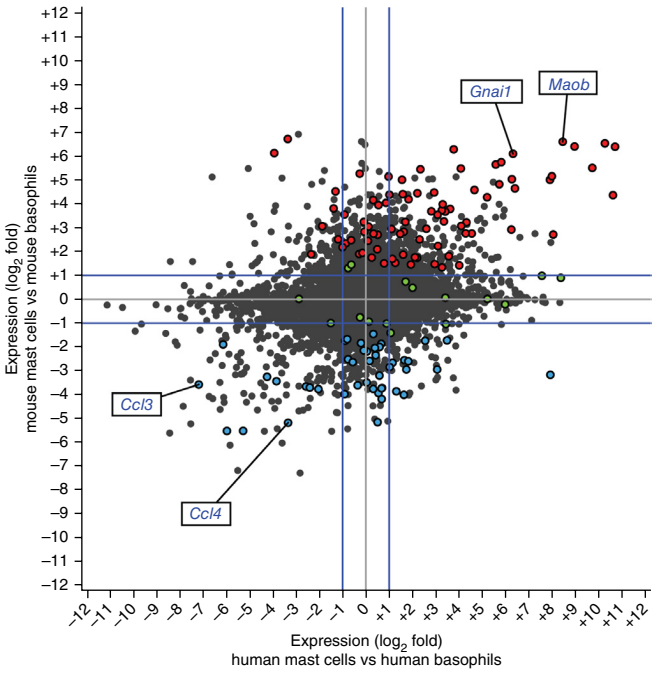



**Figure 5** Enrichment for expression of the mouse mast-cell signature in human mast cells. Expression of all 10,773 transcripts represented in both the ImmGen consortium data set (mouse) and the FANTOM consortium data sets (human), in human skin mast cells versus human blood basophils (horizontal axis) or in mouse skin mast cells versus mouse blood basophils (vertical axis); blue lines indicate cutoff for a twofold difference in expression. Human mast cells show significant enrichment for expression of the mouse mast-cell signature (82 transcripts, red) ($P = 5.5 \times 10^{-16}$) and the signature shared by mast cells and basophils (17 transcripts; green) ($P = 0.0028$, hypergeometric cumulative distribution upper tail); human blood basophils do not show enrichment for expression of the mouse basophil signature (44 transcripts; blue) ($P = 0.33$). Data are from three independent experiments (mouse mast cells and mouse basophils) or three independent donors (human mast cells and human basophils)[41].

peritoneal mast cells were 'Mitosis' and 'M phase' (**Fig. 7a**), which suggested that peritoneal mast cells might be undergoing cellular turnover. Thus, we evaluated peritoneal-mast-cell expression of Ki67, a nuclear protein present during mitosis but rapidly degraded during the G-0 phase. Ki67 staining was greater in peritoneal mast cells than in skin mast cells, which also expressed Ki67 (**Fig. 7b**). In total, 16% of peritoneal mast cells were positive for Ki67, compared with only 4% of skin mast cells (**Fig. 7c**), which indicated a much higher rate of mitosis in the peritoneal-mast-cell population and notable Ki67 expression in both populations in the absence of inflammation.

## DISCUSSION

Heparin-containing mast-cell-like cells are found as far back as urochordates[3], and although mast cells were first identified over 100 years ago, their contribution to immunological defense and disease has been poorly defined. Here we have provided a comprehensive transcriptional analysis of mouse mast cells in comparison with 14 other lymphoid and myeloid cell populations. We identified mast cells as the most transcriptionally distinct cell type, as they clustered independently

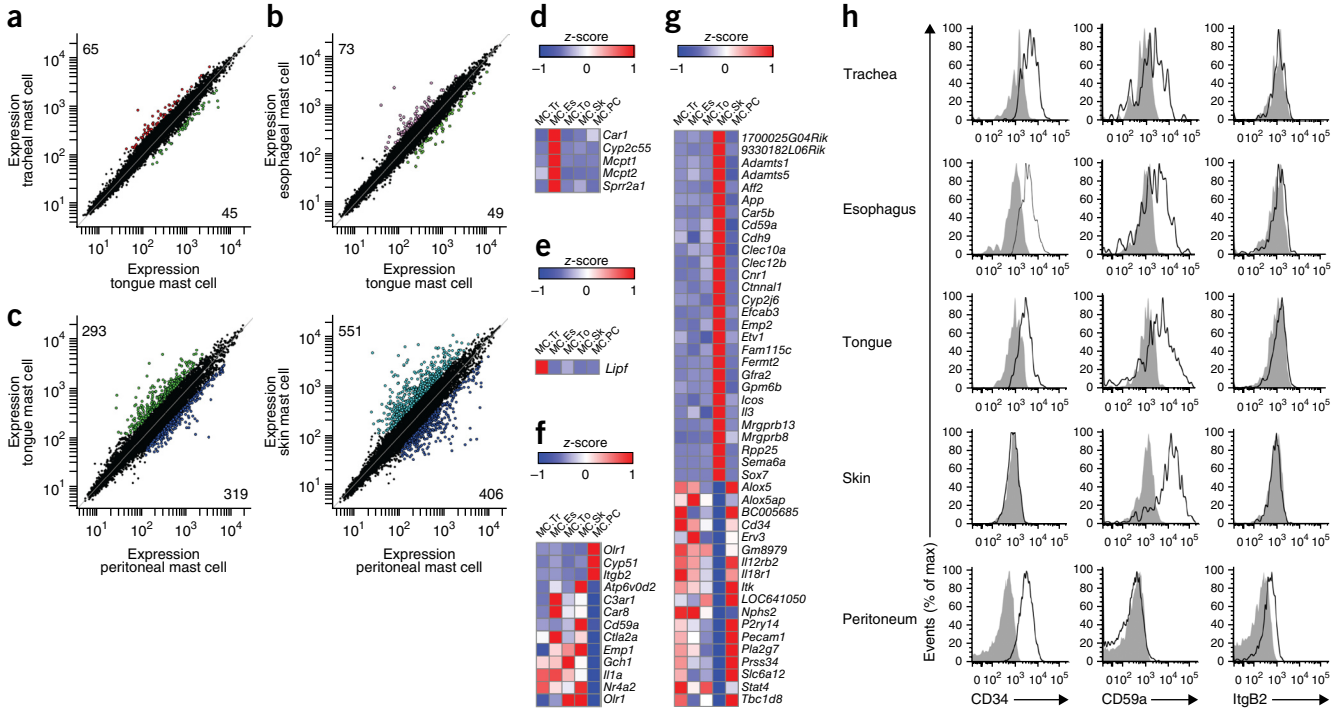



**Figure 6** Tissue-specific mast-cell gene expression. (**a**–**c**) Comparison of gene expression in various mast-cell subsets (axes); dot colors indicate transcripts with expression values greater than 120 and twofold or higher expression in tongue mast cells (light green), tracheal mast cells (red), esophageal mast cells (pink), peritoneal mast cells (dark blue) or skin mast cells (aqua) than in the other population plotted; numbers in plots indicate total differentially expressed genes along the vertical axis (top left) or horizontal axis (bottom right). (**d**–**g**) Transcripts with expression at least fourfold higher or lower in mast cells from the esophagus (**d**), trachea (**e**), peritoneum (**f**) or skin (**g**) than in any other mast cell population. (**h**) Flow cytometry of mast cells from various tissues (left margin) assessing the expression (black lines) of CD34, CD59a and ItgB2 (to confirm differential gene expression suggested by transcript data). Gray shaded curves, isotype-matched control antibody. Data are pooled from three (skin, tongue and tracheal mast cells), five (peritoneal mast cells) or two (esophageal mast cells) independent experiments (**a**–**g**) or are representative of three independent experiments (**h**).

© 2016 Nature America, Inc. All rights reserved.

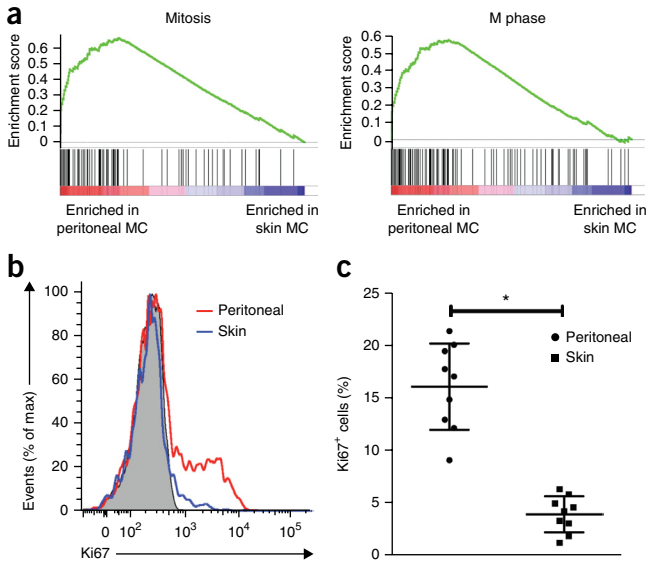



**Figure 7** Transcriptional analysis indicates peritoneal-mast-cell turnover. (**a**) GSEA identification of GO Terms 'Mitosis' and 'M phase' as showing significant enrichment in peritoneal-cavity mast cells treated with digestion enzymes relative to their abundance in skin mast cells treated similarly (nominal *P* value, <0.001; false-discovery rate Q value, <0.005). (**b**) Intracellular Ki67 expression in peritoneal and skin mast cells (key). Gray shaded curve, isotype-matched control antibody. (**c**) Frequency of Ki67+ mast cells in peritoneum and skin. Each symbol represents an individual mouse; small horizontal lines indicate the mean (± s.d.). $^*P = 0.0000062$ (two-tailed unpaired *t*-test with Welch's correction). Data are representative of QQ experiments (**a**) or three independent experiments with a total of $n = 9$ mice (**b**) or are pooled from three independent experiments with a total of $n = 9$ mice (**c**).

from all other populations, including basophils. We described a shared mast-cell transcriptional signature and further recognized tissue-specific regulation of the mast-cell transcriptome. We found that mast cells expressed genes encoding proteins involved in sensing and responding to environmental cues; this should provide a framework for understanding their sentinel function.

Mast cells from various tissues shared a transcriptional signature of 128 genes, among which those encoding serine proteases were a substantial contributor. Mast cells also showed enrichment for the expression of genes encoding products in metabolic pathways required for the generation of a broad range of other preformed mediators, including histamine, serotonin and heparin sulfate. Furthermore, mast cells expressed transcripts encoding products that allow the acute generation of eicosanoids such as prostaglandin $D_2$ and the rapid production of cytokines and chemokines. Together these findings indicate an ability to generate a unique repertoire of mediators. Human mast cells also showed substantial enrichment for expression of the mouse mast-cell signature, suggestive of evolutionary pressures to retain a core mast-cell functionality. These highly conserved genes included well-known mast-cell genes, such as those encoding proteases, and *Hpgds*, as well as several that are poorly understood in the context of mast cells, including *Maob*, *Gnai1* and genes encoding members of the Mrgpr family.

The array of Mrgpr-family members expressed in mast cells was broader than previously appreciated. Originally discovered in sensory neurons[42], eight members of this family were expressed in skin mast cells and six were expressed in the other mast-cell populations. Further analysis showing expression of *Mrgpra6* in basophils and of *Mrgpra2a* and *Mrgpra2b* in neutrophils suggested that the Mrgpr family might have an important role in the innate immune system. MRGPRX2 has been shown to mediate mast cell degranulation in response to the classical mast-cell-activating compound 48/80 in human-cord-blood-derived mast cells[43] and the transformed human LAD2 mast-cell line[44]. The mouse homolog of MRGPRX2, Mrgprb2, mediates degranulation in response to wasp venom, 48/80 and a diverse array of other basic compounds, including therapeutic agents that induce IgE-independent pseudoallergic reactions in humans[35]. Thus, members of this family might have a critical role in mediating the innate activation of mast cells in response to both pharmacological agents and as-yet-unidentified native ligands.

The low homology between mouse mast cells and mouse basophils observed in this study is similar to that previously observed for human cells, as is the closer relationship between basophils and eosinophils[41]. While mouse mast cells and basophils shared expression of transcripts encoding several activating receptors and histamine-biosynthetic enzymes, basophils lacked the diversity of proteases seen in mast cells and expressed different combinations of soluble mediators and receptors. Thus, our transcriptional analysis of mast cells and basophils suggested that these cells have independent roles in regulating homeostasis and host defense, rather than serving similar roles in different tissue compartments. The basophil signature included *Ccl3*, *Ccl4* and *Ccl9*. Human basophils showed enrichment for the expression of two of these transcripts, *Ccl4* and *Ccl4*, relative to their expression in human mast cells, which suggested a conserved role for the encoded chemokines in directing cellular recruitment. However, the poor conservation between human basophils and mouse basophils in their expression of other genes in the basophil signature might reflect evolutionary pressures drove divergence of this cell type in these species.

Comparative analysis of mast-cell populations revealed considerable tissue-specific gene expression, consistent with mast-cell maturation in peripheral tissue and with studies demonstrating regulation of mast cells by neighboring fibroblasts[21,45]. Unlike other mast-cell populations, peritoneal mast cells are not embedded in the tissue but instead line the serosal gut wall. We observed that they showed 'enrichment' for transcriptional pathways associated with cellular turnover, which led to the finding that a substantial fraction of peritoneal mast cells stained positively for Ki67. Thus, the profound transcriptional differences between peritoneal mast cells and other mast-cell compartments might reflect both cell maturation and differential signaling from neighboring cells. Notably, Ki67 staining was also detectable at low levels in skin mast cells, which suggested that local proliferation might have a role in the renewal and maintenance of this compartment.

In conclusion, we found that mast cells were extraordinarily distinct at the transcriptional level. Their core signature showed enrichment for the expression of genes encoding a diverse array of proteases and factors involved in biosynthetic pathways, which would allow the generation of a broad range of mediators, and included several novel gene families whose product function is not yet understood. Analysis of mast-cell heterogeneity revealed three distinct connective-tissue mast-cell subsets and varying capacity for *in situ* proliferation in the absence of tissue inflammation. These findings provide a framework for better definition of the role of these evolutionarily ancient cells in homeostasis, host defense and disease.

## METHODS

Methods and any associated references are available in the online version of the paper.

© 2016 Nature America, Inc. All rights reserved.

**Accession codes.** GEO: microarray data, GSE37448.

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*

ACKNOWLEDGMENTS
We thank the other members of the ImmGen Consortium, especially C. Benoist and T. Shay, for discussions; the core ImmGen team, especially A. Rhodes and K. Rothamel, for technical assistance; and A. Chicoine for assistance with the isolation of cells. Supported by the US National Institutes of Health (R24AI072073 to the ImmGen Consortium; R01 HL120952 and U19AI095219 to N.A.B.; AI095219 to N.A.B.; and T32 AI007306 to D.F.D.) and the Steven and Judy Kaye Young Innovators Award (N.A.B.).

AUTHOR CONTRIBUTIONS
D.F.D. wrote the manuscript, conceived of and conducted experiments, and analyzed the data; N.A.B. and K.F.A. wrote the manuscript and supervised the experimental design; and The ImmGen Project Consortium contributed to data collection and assisted in experimental design.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Heng, T.S. & Painter, M.W. Immunological Genome Project Consortium. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008).
2. Voehringer, D. Protective and pathological roles of mast cells and basophils. *Nat. Rev. Immunol.* **13**, 362–375 (2013).
3. Cavalcante, M.C. *et al.* Occurrence of heparin in the invertebrate styela plicata (*Tunicata*) is restricted to cell layers facing the outside environment. An ancient role in defense? *J. Biol. Chem.* **275**, 36189–36196 (2000).
4. Wong, G.W. *et al.* Ancient origin of mast cells. *Biochem. Biophys. Res. Commun.* **451**, 314–318 (2014).
5. Metcalfe, D.D., Baram, D. & Mekori, Y.A. Mast cells. *Physiol. Rev.* **77**, 1033–1079 (1997).
6. Razin, E. *et al.* IgE-mediated release of leukotriene C4, chondroitin sulfate E proteoglycan, β-hexosaminidase, and histamine from cultured bone marrow-derived mouse mast cells. *J. Exp. Med.* **157**, 189–201 (1983).
7. Metz, M. *et al.* Mast cells can enhance resistance to snake and honeybee venoms. *Science* **313**, 526–530 (2006).
8. Schneider, L.A., Schlenner, S.M., Feyerabend, T.B., Wunderlin, M. & Rodewald, H.R. Molecular mechanism of mast cell mediated innate defense against endothelin and snake venom sarafotoxin. *J. Exp. Med.* **204**, 2629–2639 (2007).
9. Abraham, S.N. & St. John, A.L. Mast cell-orchestrated immunity to pathogens. *Nat. Rev. Immunol.* **10**, 440–452 (2010).
10. Abonia, J.P. *et al.* Mast cell protease 5 mediates ischemia-reperfusion injury of mouse skeletal muscle. *J. Immunol.* **174**, 7285–7291 (2005).
11. Younan, G. *et al.* The inflammatory response after an epidermal burn depends on the activities of mouse mast cell proteases 4 and 5. *J. Immunol.* **185**, 7681–7690 (2010).
12. Gurish, M.F. & Austen, K.F. Developmental origin and functional specialization of mast cell subsets. *Immunity* **37**, 25–33 (2012).
13. Crapper, R.M. & Schrader, J.W. Frequency of mast cell precursors in normal tissues determined by an in vitro assay: antigen induces parallel increases in the frequency of P cell precursors and mast cells. *J. Immunol.* **131**, 923–928 (1983).
14. Friend, D.S. *et al.* Mast cells that reside at different locations in the jejunum of mice infected with *Trichinella spiralis* exhibit sequential changes in their granule ultrastructure and chymase phenotype. *J. Cell Biol.* **135**, 279–290 (1996).
15. Jones, T.G. *et al.* Antigen-induced increases in pulmonary mast cell progenitor numbers depend on IL-9 and CD1d-restricted NKT cells. *J. Immunol.* **183**, 5251–5260 (2009).
16. Xing, W., Austen, K.F., Gurish, M.F. & Jones, T.G. Protease phenotype of constitutive connective tissue and of induced mucosal mast cells in mice is regulated by the tissue. *Proc. Natl. Acad. Sci. USA* **108**, 14210–14215 (2011).
17. Gersch, C. *et al.* Mast cells and macrophages in normal C57/BL/6 mice. *Histochem. Cell Biol.* **118**, 41–49 (2002).
18. Hayashi, C., Sonoda, T., Nakano, T., Nakayama, H. & Kitamura, Y. Mast-cell precursors in the skin of mouse embryos and their deficiency in embryos of Sl/Sld genotype. *Dev. Biol.* **109**, 234–241 (1985).
19. Kitamura, Y., Shimada, M., Hatanaka, K. & Miyano, Y. Development of mast cells from grafted bone marrow cells in irradiated mice. *Nature* **268**, 442–443 (1977).
20. Friend, D.S. *et al.* Reversible expression of tryptases and chymases in the jejunal mast cells of mice infected with *Trichinella spiralis*. *J. Immunol.* **160**, 5537–5545 (1998).
21. Kurashima, Y. *et al.* The enzyme Cyp26b1 mediates inhibition of mast cell activation by fibroblasts to maintain skin-barrier homeostasis. *Immunity* **40**, 530–541 (2014).
22. Dvorak, A.M. *et al.* Ultrastructural identification of the mouse basophil. *Blood* **59**, 1279–1285 (1982).
23. Poorafshar, M., Helmby, H., Troye-Blomberg, M. & Hellman, L. MMCP-8, the first lineage-specific differentiation marker for mouse basophils. Elevated numbers of potent IL-4-producing and MMCP-8-positive cells in spleens of malaria-infected mice. *Eur. J. Immunol.* **30**, 2660–2668 (2000).
24. Siracusa, M.C., Perrigoue, J.G., Comeau, M.R. & Artis, D. New paradigms in basophil development, regulation and function. *Immunol. Cell Biol.* **88**, 275–284 (2010).
25. Gibbs, B.F. *et al.* Purified human peripheral blood basophils release interleukin-13 and preformed interleukin-4 following immunological activation. *Eur. J. Immunol.* **26**, 2493–2498 (1996).
26. Warner, J.A. *et al.* Differential release of mediators from human basophils: differences in arachidonic acid metabolism following activation by unrelated stimuli. *J. Leukoc. Biol.* **45**, 558–571 (1989).
27. Ohnmacht, C. & Voehringer, D. Basophil effector function and homeostasis during helminth infection. *Blood* **113**, 2816–2825 (2009).
28. Arinobu, Y. *et al.* Developmental checkpoints of the basophil/mast cell lineages in adult murine hematopoiesis. *Proc. Natl. Acad. Sci. USA* **102**, 18105–18110 (2005).
29. Galli, S.J. *et al.* Approaches for analyzing the roles of mast cells and their proteases in vivo. *Adv. Immunol.* **126**, 45–127 (2015).
30. Reber, L.L. *et al.* Contribution of mast cell-derived interleukin-1β to uric acid crystal-induced acute arthritis in mice. *Arthritis Rheumatol.* **66**, 2881–2891 (2014).
31. Burton, O.T. *et al.* Immunoglobulin E signal inhibition during allergen ingestion leads to reversal of established food allergy and induction of regulatory T cells. *Immunity* **41**, 141–151 (2014).
32. Marichal, T. *et al.* A beneficial role for immunoglobulin E in host defense against honeybee venom. *Immunity* **39**, 963–975 (2013).
33. Dudeck, A. *et al.* Mast cells are key promoters of contact allergy that mediate the adjuvant effects of haptens. *Immunity* **34**, 973–984 (2011).
34. Feyerabend, T.B. *et al.* Cre-mediated cell ablation contests mast cell contribution in models of antibody- and T cell-mediated autoimmunity. *Immunity* **35**, 832–844 (2011).
35. McNeil, B.D. *et al.* Identification of a mast-cell-specific receptor crucial for pseudo-allergic drug reactions. *Nature* **519**, 237–241 (2015).
36. Ericson, J.A. *et al.* ImmGen Consortium. Gene expression during the generation and activation of mouse neutrophils: implication of novel functional and regulatory pathways. *PLoS One* **9**, e108553 (2014).
37. Li, Y., Qi, X., Liu, B. & Huang, H. The STAT5-GATA2 pathway is critical in basophil and mast cell differentiation and maintenance. *J. Immunol.* **194**, 4328–4338 (2015).
38. Voehringer, D., Shinkai, K. & Locksley, R.M. Type 2 immunity reflects orchestrated recruitment of cells committed to IL-4 production. *Immunity* **20**, 267–277 (2004).
39. Vitalis, T. *et al.* Developmental expression pattern of monoamine oxidases in sensory organs and neural crest derivatives. *J. Comp. Neurol.* **464**, 392–403 (2003).
40. Kitamura, Y., Morii, E., Jippo, T. & Ito, A. Effect of MITF on mast cell differentiation. *Mol. Immunol.* **38**, 1173–1176 (2002).
41. Motakis, E. *et al.* FANTOM consortium. Redefinition of the human mast cell transcriptome by deep-CAGE sequencing. *Blood* **123**, e58–e67 (2014).
42. Dong, X., Han, S., Zylka, M.J., Simon, M.I. & Anderson, D.J. A diverse family of GPCRs expressed in specific subsets of nociceptive sensory neurons. *Cell* **106**, 619–632 (2001).
43. Tatemoto, K. *et al.* Immunoglobulin E-independent activation of mast cell is mediated by Mrg receptors. *Biochem. Biophys. Res. Commun.* **349**, 1322–1328 (2006).
44. Kashem, S.W. *et al.* G protein coupled receptor specificity for C3a and compound 48/80-induced degranulation in human mast cells: roles of Mas-related genes MrgX1 and MrgX2. *Eur. J. Pharmacol.* **668**, 299–304 (2011).
45. Levi-Schaffer, F., Austen, K.F., Gravallese, P.M. & Stevens, R.L. Coculture of interleukin 3-dependent mouse mast cells with fibroblasts results in a phenotypic change of the mast cells. *Proc. Natl. Acad. Sci. USA* **83**, 6485–6488 (1986).

**The complete list of authors is as follows:**
**Daniel F Dwyer[1,2], Nora A Barrett[1,2], K Frank Austen[1,2], Edy Y Kim[1,2], Michael B Brenner[1,2], Laura Shaw[5], Bingfei Yu[5], Ananda Goldrath[5], Sara Mostafavi[6], Aviv Regev[7] Andrew Rhoades[8], Devapregasan Moodley[8],**

© 2016 Nature America, Inc. All rights reserved.

Hideyuki Yoshida[8], Diane Mathis[8], Christophe Benoist[8], Tsukasa Nabekura[9], Viola Lam[9], Lewis L Lanier[9], Brian Brown[10], Miriam Merad[10], Viviana Cremasco[11], Shannon Turley[11], Paul Monach[12], Michael L Dustin[13], Yuesheng Li[14], Susan A Shinton[14], Richard R Hardy[14], Tal Shay[15], Yilin Qi[16], Katelyn Sylvia[16], Joonsoo Kang[16], Keke Fairfax[17], Gwendalyn J Randolph[17], Michelle L Robinette[17], Anja Fuchs[18] & Marco Colonna[17]

[5]Division of Biological Sciences, University of California San Diego, La Jolla, California, USA. [6]Computer Science Department, Stanford University, Stanford, California, USA. [7]Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. [8]Division of Immunology, Department of Microbiology & Immunobiology, Harvard Medical School, Boston, Massachusetts, USA. [9]Department of Microbiology & Immunology, University of California San Francisco, San Francisco, California, USA. [10]Icahn Medical Institute, Mount Sinai Hospital, New York, New York, USA. [11]Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts, USA and Department of Cancer Immunology, Genentech, San Francisco, California, USA. [12]Department of Medicine, Boston University, Boston, Massachusetts, USA. [13]Skirball Institute of Biomolecular Medicine, New York University School of Medicine, New York, New York, USA. [14]Fox Chase Cancer Center, Philadelphia, Pennsylvania, USA. [15]Department of Life Sciences, Ben-Gurion University of the Negev, Be'er Sheva, Israel. [16]Department of Pathology, University of Massachusetts Medical School, Worcester, Massachusetts, USA. [17]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, Missouri, USA. [18]Department of Surgery, Washington University School of Medicine, St. Louis, Missouri, USA.

© 2016 Nature America, Inc. All rights reserved.

## ONLINE METHODS

**Mice.** All cells used for transcriptional and flow cytometric analyses were obtained from male 6-week-old C57BL/6J mice and tissue used for histology was obtained from male 6- to 10-week-old C57BL/6J mice from the Jackson Laboratory. Mice were housed (four mice per cage) in specific pathogen-free facilities at the Dana Farber Cancer Institute (DFCI) under a 12 h light/12 h dark cycle. The use of all mice for these studies was in accordance with institutional guidelines with review and approval by the Animal Care and Use Committee of DFCI.

**Cell isolation and sorting.** Cells were purified according to the standardized ImmGen standard operations protocol (http://www.immgen.org/Protocols/ImmGen%20Cell%20prep%20and%20sorting%20SOP.pdf) using the indicated antibodies (identified below) with modifications for increased digestion time as noted below. Peritoneal cell suspensions were obtained by lavage of the peritoneal cavity with 7 ml HBSS containing 1 mM EDcell TA. Single-cell suspensions were obtained from tongue, esophagus, and trachea by finely mincing tissue between two scalpel blades and incubating for 30 min at 37 °C with 600 U/ml collagenase IV (Worthington), 0.1% dispase (Gibco) and 20 µg/ml DNAse 1 (Roche) in RPMI supplemented with 10% FBS at 500 RPM. Ear digests were obtained using modifications of a previously described protocol[46]. Briefly, dorsal and ventral halves of the ear were separated and incubated for 20 min in HBSS with 2.5 µg/ml dispase at 300 RPM to separate the epidermis. After pulling away the epidermis, remaining tissue was finely minced between two scalpel blades and incubated for 30 min with 600 U/ml collagenase IV and 20 µg/ml DNAse 1 in RPMI supplemented with 10% FBS at 500 RPM. Spleen suspensions were obtained through mechanical disruption of the spleen followed by erythrocyte lysis using ACK buffer (Sigma). Following lysis, lymphocytes were depleted using Dynal beads directed against B220 and Thy1.2 (Invitrogen). Blood was obtained through cardiac puncture and erythrocytes were depleted using a 44%/67% Percol gradient (Sigma). Mast cells were identified as $CD45^+CD11b^-CD11c^-CD19^-CD4^-CD8^-Fc\varepsilon R1\alpha^+CD117^+$. Basophils were identified as $CD3^-CD19^-NK1.1^-CD117^-Fc\varepsilon R1\alpha^+CD49b^+$. Cells were sorted at the Brigham and Women's Human Immunology Flow Core using a BD FACSAria Fusion cell sorter. For surface marker and intracellular analysis, data was acquired on a BD FACSCanto II and analyzed with FlowJo software (Treestar). The following monoclonal antibodies (clone, concentration) were used: anti-FcεR1α (MAR-1, 1:250), anti-CD117 (2B8, 1:250), anti-CD45 (30-F11, 1:250), anti-CD11b (M1/70, 1:250), anti-CD11c (N418, 1:250), anti-CD19 (6D5, 1:250), anti-CD4 (GK1.5, 1:250), anti-CD8 (53-6.7, 1:250), anti-CD49b (DX5, 1:250), anti-NK1.1 (PK136, 1:250), anti-CD34 (MEC14.7, 1:250), anti-CD59b (mCD59.3, 1:250), anti-ItgB2 (M18/2, 1:250), and isotype-matched control monoclonal antibodies (mAbs) were obtained from BioLegend. Anti-IgE (23G3 1:250), anti-Ki67 (SolA15, 1:100), isotype-matched control mAbs, and FoxP3 staining buffer set used for Ki67 staining were obtained from eBioscience.

**Cytospins and microscopy.** For histochemical evaluation of mature mast cells in peripheral tissues, tissue sections were fixed overnight in 4% paraformaldehyde and embedded in glycolmethacrylate. For cytospin evaluation, sorted cells were spun onto charged glass slides and dried overnight. Cut section and cytospins were stained for chloracetate esterase reactivity for the identification of mast cells, and cytospins were stained with toluidine blue for the identification of basophils.

**Cells and animals per microarray replicate.** Mast cells were collected from the skin ($n = 3$, each replicate was 25,000 cells pooled from 8 mice), peritoneal cavity ($n = 5$, each replicate was 30,000 cells pooled from 4 mice), tongue ($n = 3$, each replicate was 10,000 cells pooled from 10 mice), esophagus ($n = 2$, each replicate was 10,000 cells pooled from 24 mice), and trachea ($n = 3$, each replicate was 15,000 cells pooled from 8 mice). Basophils were collected from the blood ($n = 3$, each replicate was 10,000 cells pooled from 5 mice) and spleen ($n = 3$, each replicates was 25,000 cells pooled from 4 mice). Whenever possible, multiple tissues were harvested from each mouse to minimize total number of animals used. Sample sizes were determined based on ImmGen standard protocols targeting a minimum of 10,000 cells per microarray sample.

**Microarray analysis and data evaluation.** Samples were sorted twice and collected directly into TRIzol. RNA was amplified and hybridized to the Affymetrix Mouse Gene 1.0 ST array by ImmGen according to the consortium's standard protocols (https://www.immgen.org/Protocols/Total%20RNA%20Extraction%20with%20Trizol.pdf) with modification. To improve microarray success rate, RNA was treated with heparinase as previously described[47,48]. Briefly, following an initial round of chloroform extraction, RNA was incubated in 5 µm Tris buffer containing 50U of RNAsin plus (Promega) and 0.02 U of heparinase (Sigma) for 2 h at room temperature, and then subjected to a second round of TRIzol extraction. Comparison of peritoneal mast cell RNA treated with heparinase ($n = 3$) or control showed that 4 transcripts among the 21,775 assayed were reduced by a twofold statistically significant ($P < 0.05$) degree, demonstrating minimal impact on detected transcript levels. Data generation and quality-control documentation was also conducted by ImmGen according to the consortium's standard protocols (https://www.immgen.org/Protocols/ImmGen%20QC%20Documentation_ALL-DataGeneration_0612.pdf). Transcripts identified through multiple probes were collapsed based on median values and differential gene expression was characterized using the Multiplot Studio module of GenePattern software (Broad Institute). Tracheal mast cells were found to be enriched for several B cell genes, including immunoglobulin genes, suggesting B cell contamination. Contaminating B cell genes in tracheal mast cells were identified by comparing fold changes in expression between tracheal mast cells and esophagus mast cells to fold changes in expression between esophagus mast cells and splenic B cells. All transcripts with greater than 16-fold increased expression in splenic B cells compared to tongue mast cells also showed increased expression in tracheal mast cells compared to tongue mast cells and were excluded from all pairwise comparisons. Hierarchical clustering for transcripts was conducted using Gene-E (http://broadinstitute.org/cancer/software/GENE-E) based on the top 15% most variable transcripts using Pearson's correlation and cell population clustering was calculated using Spearman's correlation. Euclidean distance matrix and all transcript heat maps were also constructed using Gene-E. Principal component analysis was visualized using MatLab software (MathWorks) using principal components calculated using the PopulationDistances PCA program (S. Davis, Harvard Medical School) based on the top 15% most variable transcripts across all analyzed cell populations. The skin and enzyme-treated peritoneal-mast-cell transcriptomes were further compared using the Gene Set Enrichment Analysis software program (Broad Institute)[49,50] using Gene Ontology Consortium (www.geneontology.org) gene sets.

**Controlling for the effects of collagenase treatment on peritoneal mast cells.** Peritoneal cell suspensions obtained by lavaging the peritoneal cavity with 7 ml HBSS containing 1 mM EDTA were incubated for 30 min at 37 °C with 600 U/ml collagenase IV (Worthington), 0.1% dispase (Gibco) and 20 µg/ml DNAse 1 (Roche) in RPMI supplemented with 10% FBS. Following enzymatic treatment, peritoneal mast cells were either isolated for microarray analysis or stained for cell surface marker expression.

**Derivation of mast-cell and basophil transcriptional signatures.** The mast-cell signature was generated in comparison to all cell populations analyzed. Multiple replicates for each cell population were collapsed based on median values. Transcripts in the mast-cell signature were expressed at least twofold higher in all mast-cell populations than in any non–mast-cell population, including basophils. All transcripts expressed below 120 relative units in more than two mast-cell subsets were excluded, as were all in which there was no statistically significant difference between mast-cell and non–mast-cell expression by student's *t*-test. The mast-cell signature was calculated using non-digested peritoneal mast cells to exclude any genes induced by collagenase and dispase treatment. The basophil signature was calculated similarly, and the shared mast-cell and basophil signature was calculated by determining all transcripts expressed at least two fold higher in both mast cell and basophil than in any non–mast cell and non-basophil. After calculating the signatures, enriched pathways were determined using DAVID software[51,52] based on the PANTHER classification system with $P < 0.05$. Mast-cell- and basophil-specific transcription factors were determined by identifying transcripts in the individual and shared mast-cell and basophil signatures that also

© 2016 Nature America, Inc. All rights reserved.

appeared in the Riken institute transcription factor database (http://genome.gsc.riken.jp/TFdb/).

**Comparison of human and mouse mast cells and basophils.** All 10,773 transcripts identified in both the Affymetrix Mouse 1.0 array and in human cells via CAGE sequencing were visualized on a fold change vs fold change plot. To allow for fold change comparisons in the CAGE sequencing data set, in which numerous transcript levels had a value of zero, a value of 1 was added to each datapoint. Genes found in the mouse mast cell, basophil and combined signatures were then highlighted.

**Statistics.** There was no randomization, blinding, or exclusion of data. Sample size was not predetermined statistically. Significance of PANTHER pathway enrichment was determined using a modified Fisher's exact test in DAVID. Enrichment of human mast cells and basophils for the mouse mast-cell and basophil signatures was evaluated using the hypergeometric cumulative distribution upper tail in Matlab (Mathworks). Differences in intracellular Ki67 levels were evaluated using Prism 6.0 (GraphPad) with a two-tailed unpaired *t*-test with Welch's correction after determining that the samples represented a gaussian distribution using the D'Agostino & Pearson omnibus normality test. $P$ values of $< 0.05$ were considered statistically significant.


46. Bogunovic, M. *et al.* Identification of a radio-resistant and cycling dermal dendritic cell population in mice and men. *J. Exp. Med.* **203**, 2627–2638 (2006).
47. Tsai, M., Miyamoto, M., Tam, S.Y., Wang, Z.S. & Galli, S.J. Detection of mouse mast cell-associated protease mRNA. Heparinase treatment greatly improves RT-PCR of tissues containing mast cell heparin. *Am. J. Pathol.* **146**, 335–343 (1995).
48. Gilchrist, M., MacDonald, A.J., Neverova, I., Ritchie, B. & Befus, A.D. Optimization of the isolation and effective use of mRNA from rat mast cells. *J. Immunol. Methods* **201**, 207–214 (1997).
49. Mootha, V.K. *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
50. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
51. Huang, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
52. Huang, W., Sherman, B.T. & Lempicki, R.A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).

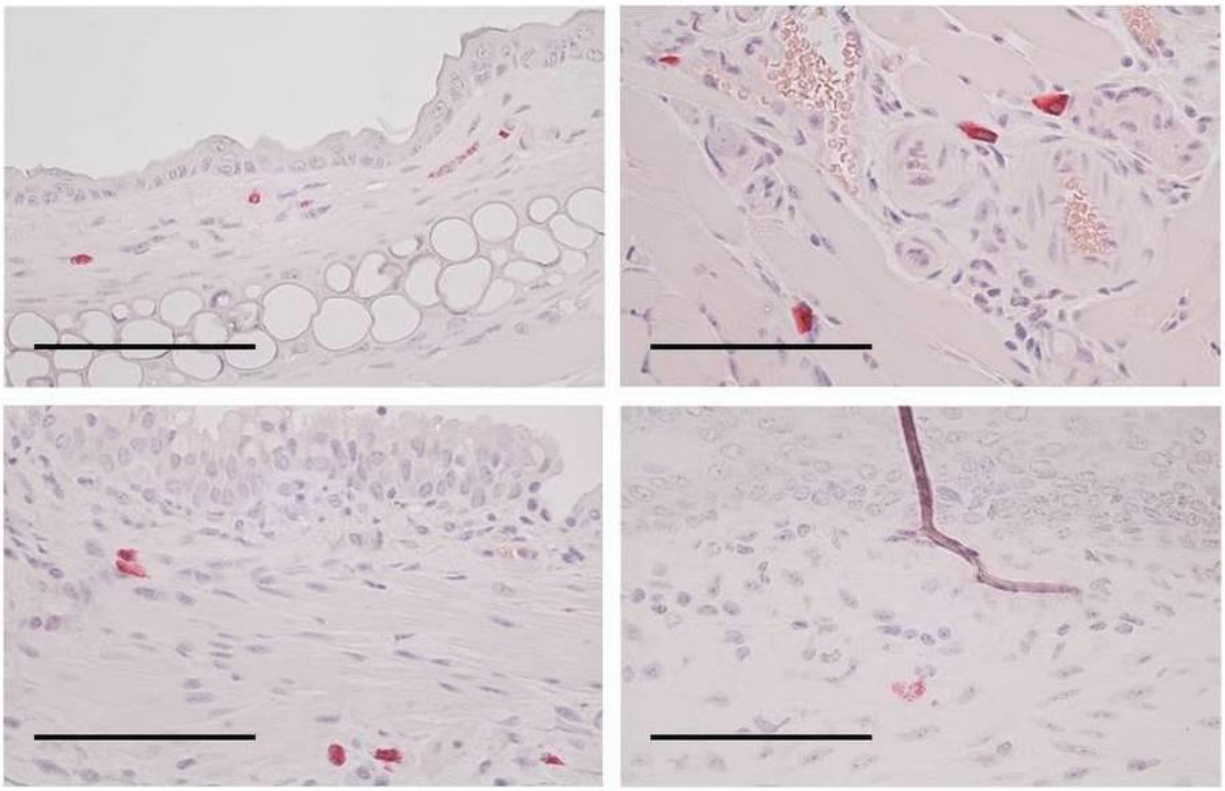

© 2016 Nature America, Inc. All rights reserved.



doi:10.1038/ni.3445

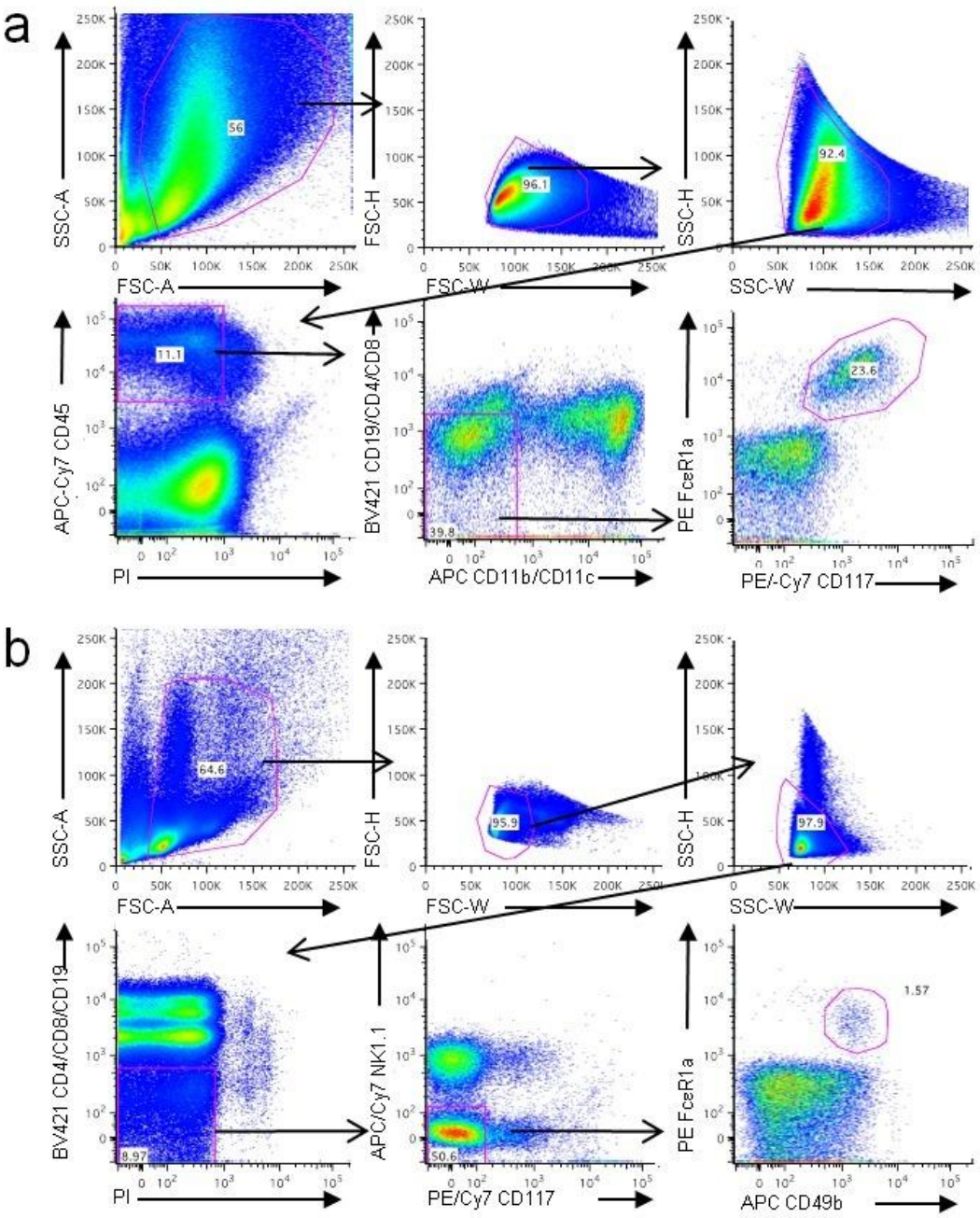
Supplementary Figure 1: Tissue localization of constitutive mast cells.

Chloroacetate esterase reactivity in (clockwise from top left) skin, tongue, trachea and esophagus showing localization of constitutive mast cells in each tissue compartment. Scale bars, 100 µm.



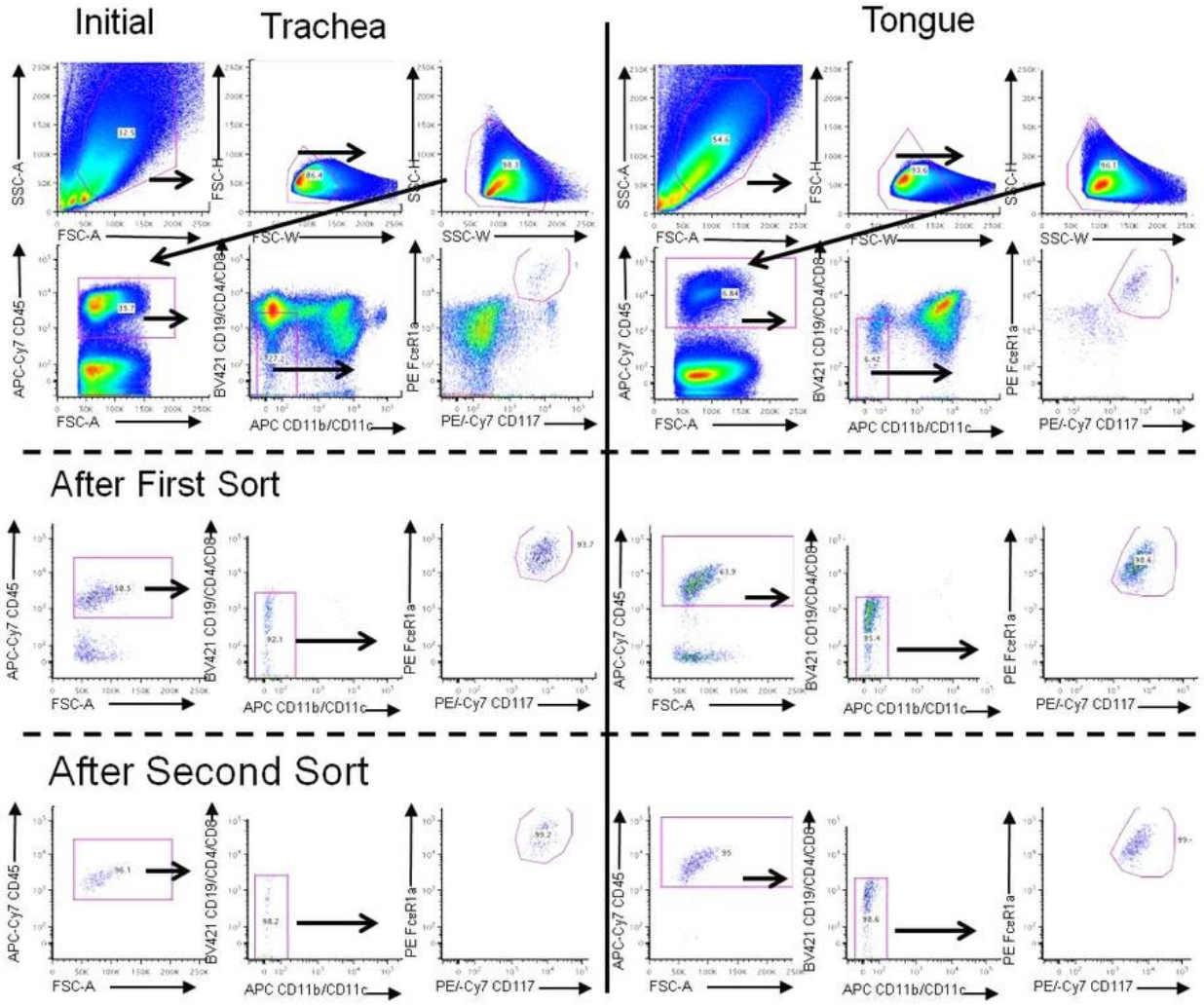
Supplementary Figure 2: Gating strategies for the isolation of mast cells and basophils from peripheral ti

(a) Flow cytometric gating used to identify mast cells. Representative plots show skin mast cells. After excluding doublets, dead cells were excluded using propridium iodide and mast cells were identified as CD45+CD11c–CD11b–CD4–CD8– CD19 – FcεR1α+CD117+. Arrows indicate sequential gates used. (b) Flow cytometric gating used to identify basophils. Representative plots show blood basophils. After excluding doublets, dead cells were excluded using propridium iodide and basophils were identified as CD4–CD8–CD19–NK1.1–CD117–CD49b+FcεR1α+. Arrows indicate sequential gates used.ssue.



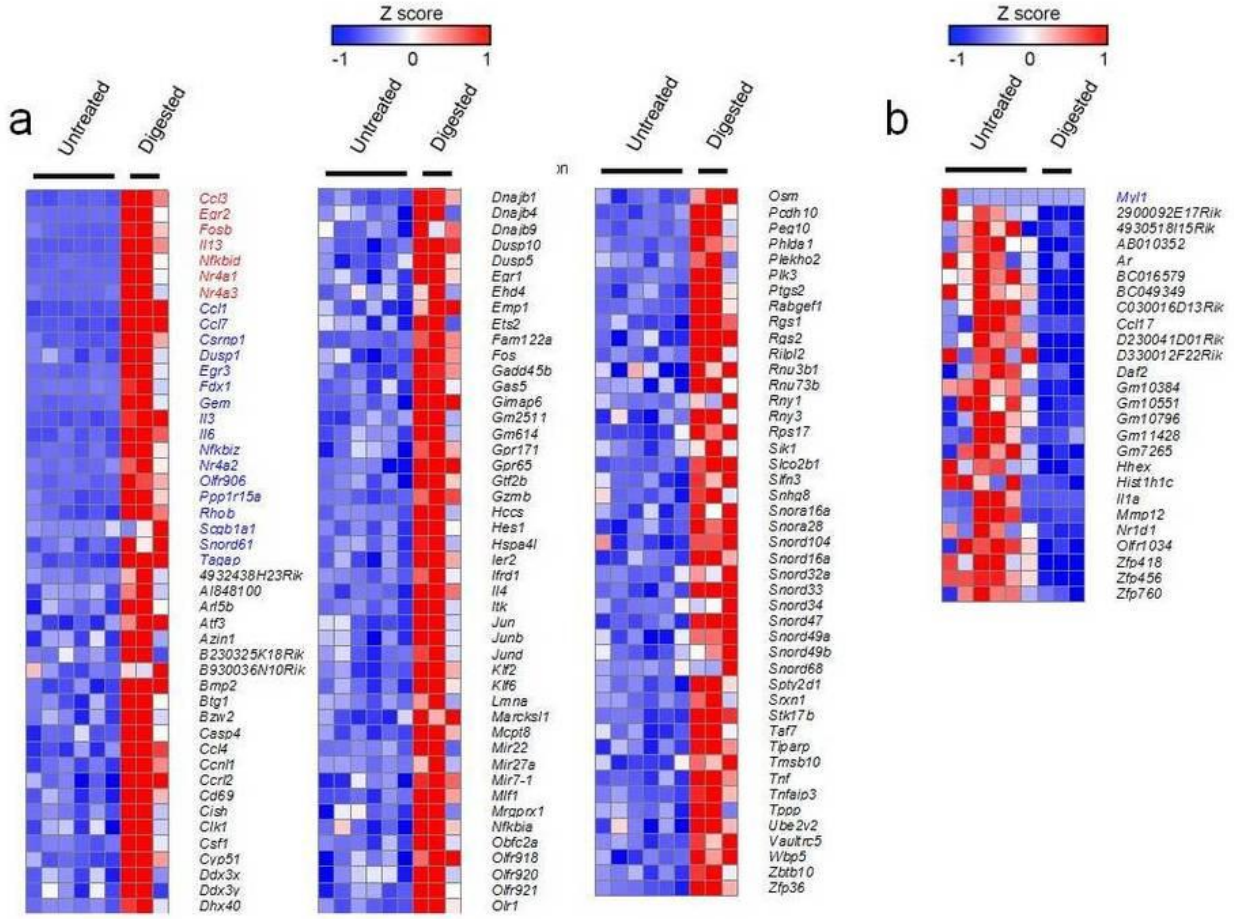
Supplementary Figure 3: Assessment of mast-cell purity following multiple rounds of sorting.

After gating on intact cells and excluding doublets, mast cell purity was determined as a percentage of all intact cells after each round of sorting. mast cells were identified as CD45+CD11c–CD11b–CD4–CD8– CD19 – FcεR1α+CD117+. Arrows indicate sequential gates used. Numbers in each panel indicate the percentage of gated cells from the parental gate.



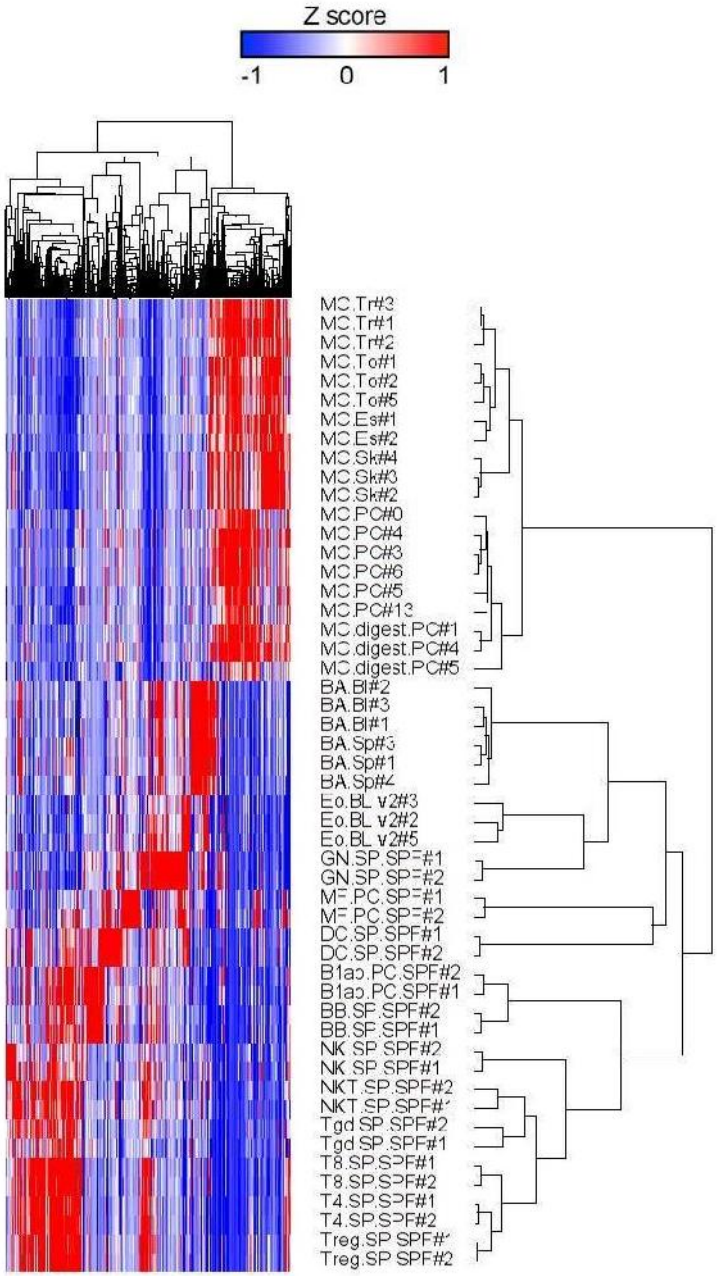
Supplementary Figure 4: Influence of treatment with collagenase and dispase on peritoneal-mast-cell transcription.

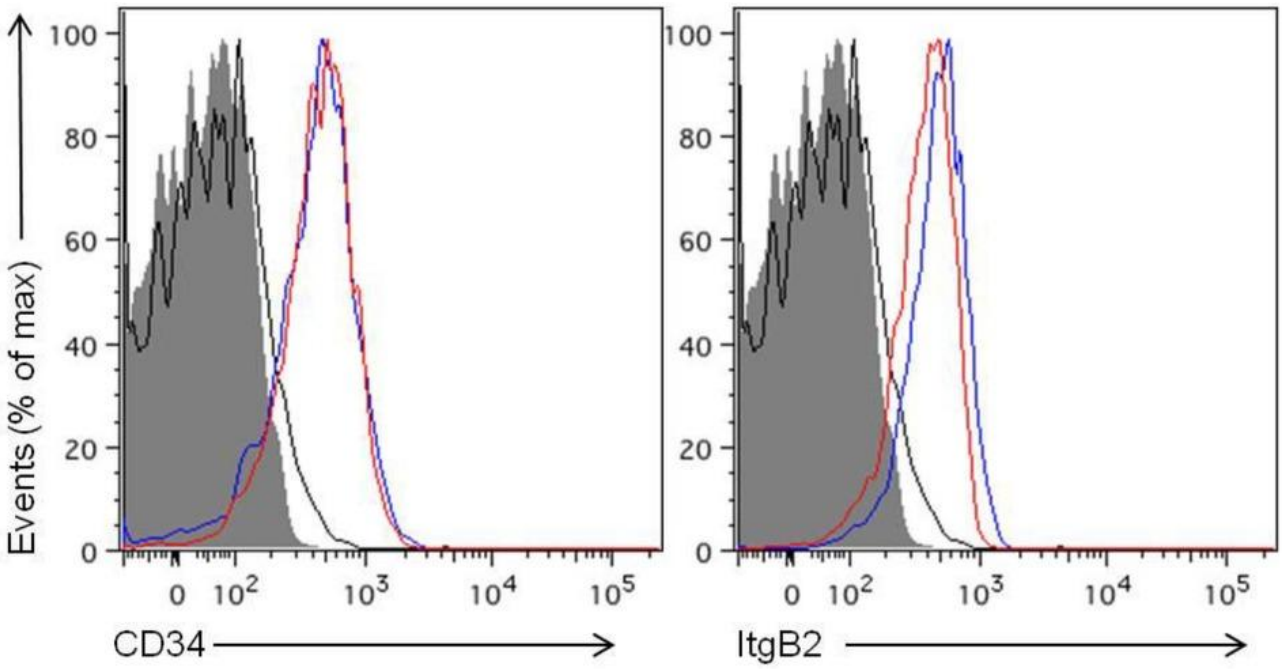
Peritoneal mast cells were subjected to collagenase and dispase treatment for 30 min at 37°C, and microarray results were compared to untreated peritoneal mast cells. (a) Transcripts upregulated by two-fold or more in peritoneal mast cells relative to untreated mast cells following treatment with digestion enzymes. Colored gene symbols indicate fivefold (blue) or tenfold (red) higher expression. (b) Transcripts downregulated by twofold or more in peritoneal mast cells relative to untreated mast cells following treatment with digestion enzymes. Colored gene symbols indicate five-fold (blue) decreased expression. n = 6 untreated replicates and n = 3 enzymatically-treated replicates.



Supplementary Figure 5: Influence of treatment with collagenase and dispase on peritoneal-mast-cell hierarchical clustering.

Peritoneal mast cells were subjected to collagenase and dispase treatment for 30 min at 37°C, and microarray results were compared to all other analyzed immunocyte populations by hierarchical clustering using the top 15% most variable genes. Cell populations denoted using standard ImmGen abbreviations as indicated in Fig.1c. Bar height is inversely correlated to homology between linked populations. Data are from n = 3 independent experiments (skin, tongue, trachea, and enzymatically treated mast cells, spleen and blood basophils), n = 6 independent experiments from peritoneal mast cells, and n = 2 independent experiments from esophagus mast cells.



Supplementary Figure 6: Influence of digestion with collagenase on mast-cell surface-marker staining.

Representative histograms indicating CD34 and ItgB2 surface staining on peritoneal mast cells either untreated (red) or incubated with collagenase IV and dispase I for 30 min at 37°C (blue). Grey solid histogram indicates isotype staining for untreated mast cells, black line indicates isotype staining on enzymatically treated mast cells. Results indicative of n = 3 individual experiments.