nature immunology

Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages

Emmanuel L Gautier[1–3], Tal Shay[4], Jennifer Miller[2,5], Melanie Greter[2,5], Claudia Jakubzick[1,2], Stoyan Ivanov[3], Julie Helft[2,5], Andrew Chow[2,5], Kutlu G Elpek[6,7], Simon Gordonov[8], Amin R Mazloom[8], Avi Ma'ayan[8], Wei-Jen Chua[3], Ted H Hansen[3], Shannon J Turley[6,7], Miriam Merad[2,5], Gwendalyn J Randolph[1–3] & the Immunological Genome Consortium[9]

We assessed gene expression in tissue macrophages from various mouse organs. The diversity in gene expression among different populations of macrophages was considerable. Only a few hundred mRNA transcripts were selectively expressed by macrophages rather than dendritic cells, and many of these were not present in all macrophages. Nonetheless, well-characterized surface markers, including MerTK and FcγR1 (CD64), along with a cluster of previously unidentified transcripts, were distinctly and universally associated with mature tissue macrophages. TCEF3, C/EBP-α, Bach1 and CREG-1 were among the transcriptional regulators predicted to regulate these core macrophage-associated genes. The mRNA encoding other transcription factors, such as *Gata6*, was associated with single macrophage populations. We further identified how these transcripts and the proteins they encode facilitated distinguishing macrophages from dendritic cells.

© 2012 Nature America, Inc. All rights reserved.

The team of immunologists and computational biologists of the Immunological Genome (ImmGen) Project share the goal of generating an exhaustive definition of gene-expression and regulatory networks of the mouse immune system through shared resources and rigorously controlled data-generation pipelines[1]. Here we turned our attention to gene-expression and regulatory networks in tissue-resident macrophages. Macrophages are professional phagocytic cells, often long lived, that reside in all organs to maintain tissue integrity, clear debris and respond rapidly to initiate repair after injury or innate immunity after infection[2,3]. Accordingly, macrophages are specialized for the degradation and detoxification of engulfed cargo, and they are potent secretagogues able to develop an array of phenotypes[4]. Macrophages can also present antigens but lack the potency for stimulating T cells observed in dendritic cells (DCs), and they usually fail to mobilize to lymphoid tissues in which naive T cells are abundant. Partially overlapping functions for macrophages and DCs, reflected by overlapping molecular profiles, have for decades fueled some debate over the origins and overall distinction between macrophages and DCs[5].

In the past several years, considerable progress has been made in the identification of precursor cells specific to DCs[6–8]. Moreover, transcription factors have been identified, such as Batf3, that are essential for the development of some DCs but are not required for macrophage specification[9]. Advances have also been made in delineating the development of tissue macrophages. Contrary to the prevalent idea that monocytes are precursors of tissue macrophages, some earlier work contended that tissue macrophages arise from primitive hematopoietic progenitors present in the yolk sac during embryonic development independently of the monocyte lineage[10], and support for that contention has emerged from fate-mapping and genetic models[11,12]. Thus, in the adult, the maintenance of tissue macrophages involves local proliferation, again independently of monocytes and definitive hematopoiesis[10,12]. In this context, the transcription factor MAFB (c-Maf) has been shown to regulate macrophage self-renewal[13]. Some transcription factors that drive the development of specific macrophage types such as osteoclasts[14] or red-pulp macrophages[15] have also been reported. However, much remains to be determined about the transcriptional regulatory pathways that control other types of macrophages or global regulatory pathways that govern macrophages as a group of related cells[3]. The database generated by the ImmGen Project has created a unique resource for the comparison of gene-expression profiles and the identification of regulatory pathways that specify or

[1]Department of Developmental and Regenerative Biology, Mount Sinai School of Medicine, New York, New York, USA. [2]The Immunology Institute, Mount Sinai School of Medicine, New York, New York, USA. [3]Department of Pathology & Immunology, Washington University School of Medicine, St. Louis, Missouri, USA. [4]Broad Institute, Cambridge, Massachusetts, USA. [5]Department of Oncological Sciences and Department of Medicine, Mount Sinai School of Medicine, New York, New York, USA. [6]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA. [7]Department of Cancer Immunology and AIDS, Dana Farber Cancer Institute, Boston, Massachusetts, USA. [8]Department of Pharmacology and Systems Therapeutics & Systems Biology Center New York, Mount Sinai School of Medicine, New York, New York, USA. [9]Full list of members and affiliations appears at the end of the paper. Correspondence should be addressed to G.J.R. (grandolph@path.wustl.edu).

Received 11 April; accepted 9 August; published online 30 September 2012; doi:10.1038/ni.2419

unify macrophage populations from different organs. Our analysis here of the macrophage transcriptome in this context will enable the analysis of networks of genes and their regulators that can be used to better distinguish different types of macrophages and pinpoint the differences between macrophages and DCs.

RESULTS

Tissue macrophage diversity

As part of the ImmGen Project, we sorted several tissue macrophage populations from C57BL/6J mice according to strict, standardized procedures and analyzed these populations by whole-mouse genome miroarray. Strategies for sorting these populations are available at the ImmGen Project website. We began our analysis by examining the gene-expression profiles of resting macrophage populations that have historically been characterized and accepted as true resident tissue macrophages[12]. Although some classic macrophages, such as Kupffer cells of the liver and metallophilic or marginal-zone macrophages of the spleen, proved elusive for definitive identification and/or isolation through sorting by flow cytometry, the following four resting macrophage populations submitted to the ImmGen Project met the criteria of true macrophage populations: peritoneal macrophages; red-pulp splenic macrophages; lung macrophages; and microglia (brain macrophages). Principal-component analysis (PCA) of all genes expressed by the four sorted macrophage populations and several DC populations showed a greater distance between the macrophages than between the DCs (**Fig. 1a**). Pearson correlation values were high for replicates in a given DC or macrophage population according to the quality-control standards of the ImmGen Project; variability within replicates for a single population varied from 0.908 ± 0.048 for microglia to 0.995 ± 0.001 for peritoneal macrophages. Pearson correlations for the gene-expression profiles of various populations of DCs yielded coefficients that ranged from 0.877 (liver CD11b$^+$ DCs versus spleen CD8$^+$ DCs) to 0.966 (spleen CD4$^+$CD11b$^+$ DCs versus spleen CD8$^+$ DCs; mean of all DC populations, 0.931), whereas the correlation coefficients for the tissue macrophages ranged from 0.784 (peritoneal versus splenic red pulp) to 0.863 (peritoneal versus lung) with a mean of 0.812 (**Fig. 1b**). Several thousand mRNA transcripts had a difference in expression of at least twofold in, for example, lung macrophages versus red-pulp splenic macrophages (**Fig. 1c**). This degree of diversity was greater than that observed for DCs of different subsets (CD103$^+$ versus CD11b$^+$) from various organs (**Fig. 1c**). Finally, a dendrogram applied to the various populations showed that DCs clustered more closely than macrophages did (**Fig. 1d**), and this was true whether we considered all gene transcripts in the array (data not shown) or only the top 15% ranked by the cross-population maximum/minimum ratio or coefficient of variation (**Fig. 1d**). Overall, these comparisons indicated considerable diversity among tissue macrophage populations.

Distinct molecular signatures among tissue macrophages

The diversity among the four classical macrophage populations noted above extended to gene families previously associated with macrophage function: those encoding chemokine receptors, Toll-like receptors (TLRs), C-type lectins and efferocytic receptors. For example, in each population, at least one distinct chemokine receptor had much higher expression than the others (**Supplementary Fig. 1a**). The diversity in the expression of TLRs, C-type lectins and efferocytic receptors was also considerable (**Supplementary Fig. 1b–d**). Indeed, only a few of the mRNA transcripts profiled in these categories, including mRNA encoding the Mer tyrosine kinase receptor (MerTK), which is involved in the phagocytosis of apoptotic cells[16], as well as mRNA encoding TLR4, TLR7, TLR8 and TLR13, showed relatively uniform expression across all macrophages compared. Hundreds of mRNA transcripts had a selective difference in expression of at least twofold (higher or lower expression) in only one of the macrophage populations (**Fig. 2a**), and microglia in particular had low expression of hundreds of transcripts that were expressed in other macrophage populations (**Fig. 2a**). Using Ingenuity pathway-analysis software tools, we found enrichment for each specific signature in groups of transcripts encoding molecules with predicted specific functions, including oxidative metabolism in brain macrophages, lipid metabolism in lung macrophages, eicosanoid signaling in peritoneal macrophages and readiness for interferon responsiveness in red-pulp macrophages (**Supplementary Table 1**). Given that

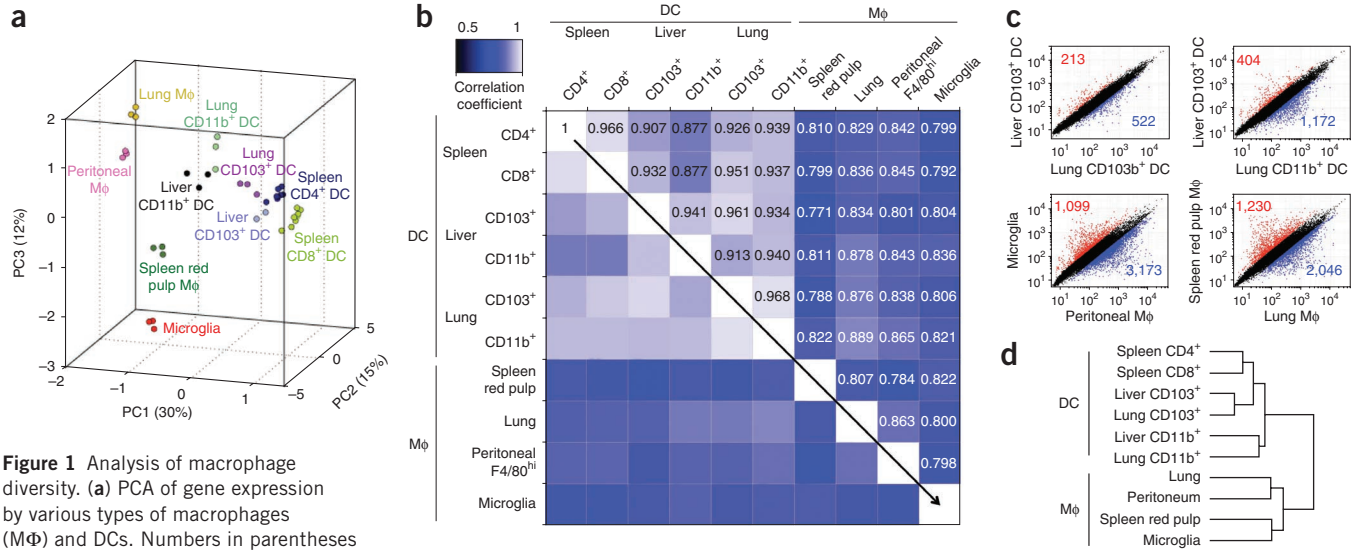
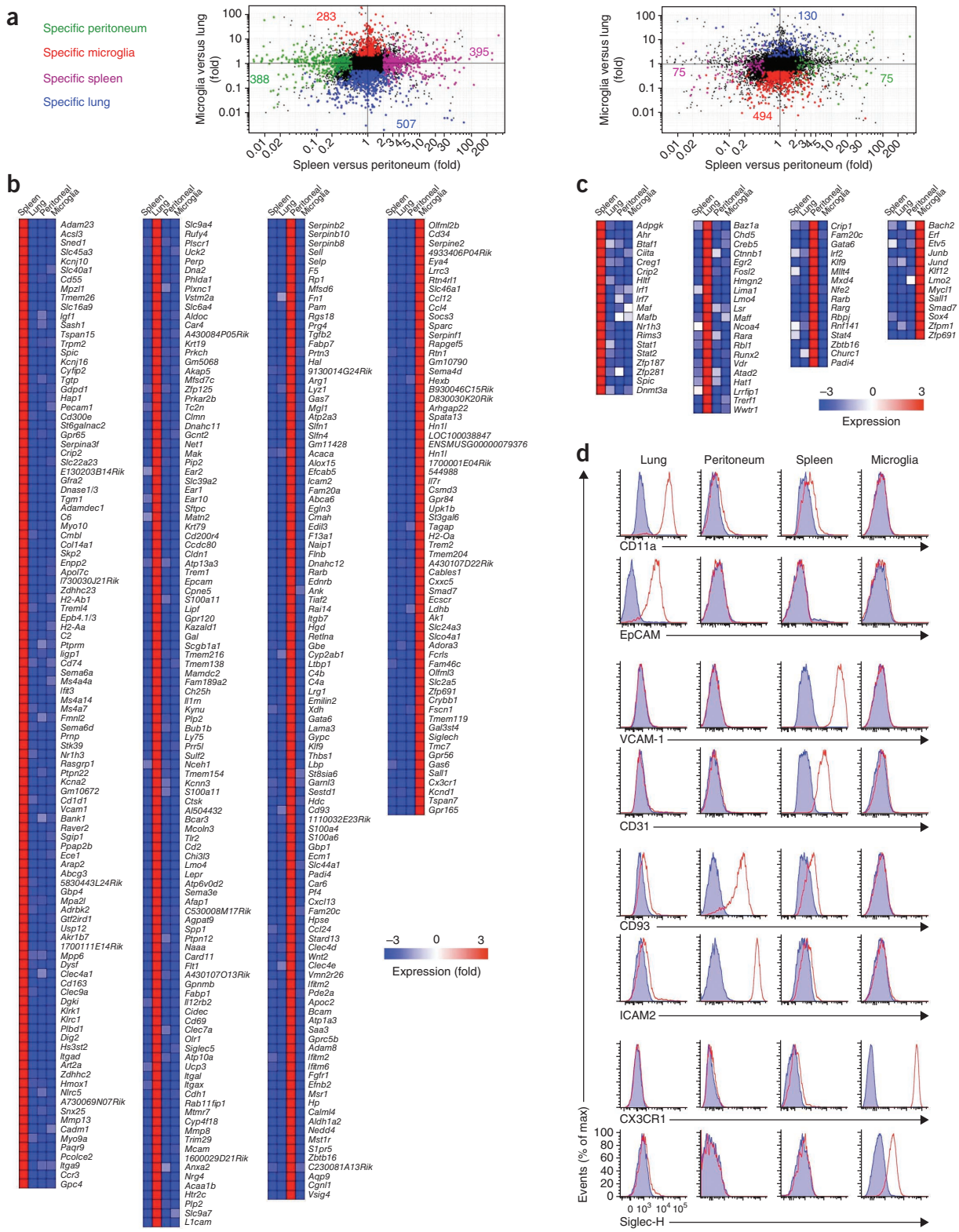


Figure 1 Analysis of macrophage diversity. (**a**) PCA of gene expression by various types of macrophages (MΦ) and DCs. Numbers in parentheses indicate relative scaling of the principal variables. (**b**) Correlation matrix of macrophages and DCs based on all available gene probes. Numbers in plot indicate correlation coefficient value. (**c**) Diversity of gene expression in macrophage populations and DCs. Numbers in plots indicate probes with a minimum change in expression of twofold (red, upregulated; blue, downregulated). (**d**) Hierarchical clustering of macrophages and DCs based on the 15% of genes with the greatest variability. Data are combined from three to seven independent experiments for each population, with cells pooled from three to five mice in each.

© 2012 Nature America, Inc. All rights reserved.



© 2012 Nature America, Inc. All rights reserved.

Figure 2 Unique gene-expression profiles of macrophages from various organs. (**a**) Quantification of mRNA transcripts upregulated twofold or more (left) or downregulated twofold or more (right) in one macrophage population relative to their expression in the remaining three populations. Numbers in plots indicate genes with a minimum change in expression of twofold (colors match key at left). (**b**) Heat map of mRNA transcripts upregulated in each single macrophage population (top) by fivefold or more relative to their expression in the remaining three populations. (**c**) Heat map of mRNA transcripts encoding transcription factors upregulated in only one of the four macrophage populations by twofold or more. (**d**) Flow cytometry analysis of specific cell-surface markers (identified by the gene-expression profiling data) for each macrophage population. Red line, specific antibody; blue shading, isotype-matched control antibody. Data combine results from three or more experiments.

we simultaneously compared the gene-expression profiles of the four macrophage populations, the number of transcripts with expression that was fivefold or more higher or lower in only one macrophage population relative to their expression in all three of the other populations was notable (**Fig. 2b**). We also found that many transcripts had much lower expression in only one population than in the others (**Supplementary Fig. 2**). Several transcription factors had much higher expression in just one of the four macrophage populations (**Fig. 2c**). For example, expression of the gene encoding the transcription factor Spi-C was restricted to splenic red-pulp macrophages, which fit with published work showing that Spi-C has a critical role in controlling the development of these cells[15]. Diversity at the gene-expression level corresponded to that at the protein level. For example, we detected the integrin CD11a (LFA-1) and the adhesion molecule EpCAM on lung macrophages but not on microglia, spleen or peritoneal macrophages; the adhesion molecules VCAM-1 and CD31 (PECAM-1) were selectively displayed by spleen macrophages; the C-type lectin transmembrane receptor CD93 and the adhesion molecule ICAM-2 were expressed by peritoneal macrophages but not the other macrophages; and the chemokine receptor CX3CR1 and the lectin Siglec-H were selectively present in microglia (**Fig. 2d**). Together these data indicated that macrophage populations in different organs expressed many unique mRNA transcripts that would equip them for specialized local functions.

Identification of a core macrophage signature

In the midst of the vast diversity among macrophages from different organs, we next sought to identify a core gene-expression profile that generally unified macrophages and distinguished them from other types of cells of the immune system. Among all hematopoietic cells, the cells anticipated to be most similar to macrophages are DCs[5]. To search for mRNA transcripts that distinguished macrophages from DCs, we compared the four selected prototypical macrophage populations with the most well-defined classic DC populations, including resting $CD8^+$ or $CD4^+$ $CD11b^+$ splenic DCs, $CD103^+$ tissue DCs and various populations of lymph node $CD11c^+$ migratory DCs with high expression of major histocompatibility complex (MHC) class II ($MHCII^{hi}$)[17,18]. Because tissue $CD11b^+$ DCs may be contaminated with macrophages[19], we initially excluded tissue $CD11b^+$ DCs from the comparison. This comparison identified only 14 transcripts that were expressed in all four macrophage populations but were not expressed in DCs (**Table 1**). These included mRNA anticipated to have high expression in macrophages, such as *Fcgr1* (which encodes the immunoglobulin Fc receptor CD64) and *Tlr4*. Two of these molecules, the receptor for the cytokine G-CSF (encoded by *Csf3r*) and the MHC class I–related molecule MR1 (encoded by *Mr1*), which is involved in the activation of mucosa-associated invariant T cells[20], function at least partly at the cell surface. In agreement with the pattern of mRNA expression, we found MR1 protein on spleen and lung macrophages but not on classical DCs (**Supplementary Fig. 3**), which suggested that MR1 on macrophages rather than on DCs may drive the activation of mucosa-associated invariant T cells. Other transcripts identified encode proteins involved in signal transduction, such as the kinase Fert2 (encoded by *Fer* (called '*Fert2*' here)), or in metabolism and lipid homeostasis, such as peroxisomal trans-2-enoyl-CoA reductase (encoded by *Pecr*) and alkyl glycerol monooxygenase (encoded by *Tmem195*), which is the only enzyme that cleaves the O-alkyl bond of ether lipids such as platelet-activating factor, shown to be actively catabolized in association with macrophage differentiation *in vitro*[21]. To that small number of mRNA transcripts, we added probe sets that did not lack expression by DCs but had signal intensity least twofold lower in all single DC populations than the lowest intensity of that same probe set in each macrophage population. Thus, we were able to add 25 more transcripts to that 'macrophage core' list (**Table 1**; mean transcript expression, **Supplementary Table 2**), including those known to be associated with macrophages, such as *Cd14*, *Mertk*, *Fcrg3* (which encodes the immunoglobulin Fc receptor CD16) and *Ctsd* (which encodes cathepsin D).

F4/80 (encoded by *Emr1*) has served as the most definitive marker of macrophages so far[5,12]. However, to identify additional mRNA transcripts widely associated with macrophages with the core list of macrophage-associated genes, including *Emr1*, *Mafb* and *Cebpb*, we found it necessary to adjust the criteria of the approach described above to include transcripts expressed in only three of four macrophage populations because, for example, *Emr1* mRNA had low expression in microglia. Making this adjustment expanded the list of mRNA transcripts associated with macrophages and added another 93 genes (**Table 1**). Additional macrophage-associated genes such as *Mrc1* (which encodes the mannose receptor CD206), *Marco* and *Pparg* were not identified until we 'loosened' the criteria so that only two of four prototypical macrophage populations needed to express a given transcript whose expression was otherwise absent or low on DCs (**Table 2**; transcript expression, **Supplementary Table 3**). *Cd68* mRNA, widely used to identify tissue macrophages, had similar expression in DCs and macrophages and we therefore excluded it from the list. However, as a protein, its expression was still several orders of magnitude higher in macrophages than in DCs of the spleen (**Supplementary Fig. 4**). In summary, the expression of 366 transcripts (**Tables 1** and **2**) was absent from classical DCs or was much lower in classical DCs than in macrophages. However, because of the great diversity among macrophages, expression of only 39 of these transcripts was shared by all tissue macrophages we compared.

Coexpressed genes and predicted transcriptional regulators

The computational biology groups of the ImmGen Project have analyzed the transcriptional program of the entire large database generated by the ImmGen Project (V. Jojic *et al.*, data not shown, and **Supplementary Note 1**). First, mRNA transcripts were clustered into 334 fine modules on the basis of patterns of coexpression. Then the Ontogenet algorithm (developed for the ImmGen Project data set) was applied to identify a regulatory program for each fine module on the basis of its expression pattern, the expression pattern of regulators and the position of the cells on the hematopoietic lineage tree. ImmGen Project modules, including the gene lists in each module, and regulatory program metadata are available online (http://www.immgen.org/ModsRegs/modules.html), and the numbering of the modules there is used here.

When we mapped the list of the 366 mRNA transcripts associated with macrophages according to their placement in various fine modules, 14 modules showed significant enrichment for the macrophage-associated gene signature we identified (**Fig. 3a**). In particular, the 11 genes of module 161 (*A930039a15Rik*, *Akr1b10*, *Blvrb*, *Camk1*, *Glul*, *Myo7a*, *Nln*, *Pcyox1*, *Pla2g15*, *Pon3* and *Slc48a*) were significantly induced in all four macrophage populations used to generate the list of macrophage-associated genes (**Fig. 3a**). Other modules, such as module 165, contained genes significantly induced in several specific groups of macrophages but not in all groups of macrophages (**Fig. 3a**). The 11 genes of module 161 encode molecules involved in redox regulation, heme biology, lipid metabolism and vesicular trafficking (**Supplementary Table 4**). Beyond the comparison to DCs, the genes in module 161, expressed in all macrophages, were not expressed by any other hematopoietic cell types, including granulocytes or any of

© 2012 Nature America, Inc. All rights reserved.

Table 1 Genes upregulated in tissue macrophages relative to their expression in DCs

All MΦ populations	– Peritoneal MΦ	– Lung MΦ	– Microglia	– Splenic red-pulp MΦ
Pecr	***Xrcc5***	***Mafb***	***Hgf***	***Cd151***
Tmem195	***Gm4878***	***Itga9***	***Pilrb2***	***Lonrf3***
Ptplad2	***Slco2b1***	***Cmklr1***	***Mgst1***	***Acy1***
1810011H11Rik	***Gpr77***	***Fez2***	***Klra2***	
Fert2	***Gpr160***	***Tspan4***	***Rnasel***	*C5ar1*
Tlr4	***P2ry13***	***Abcc3***	***Fcgr4***	*Pld1*
Pon3	***Tanc2***	***Nr1d1***	***Rhoq***	*Gpr177*
Mr1	***Sepn1***	***Ptprm***	***Fpr1***	*Arsk*
Arsg		***Ctsf***	***Cd302***	*Plod3*
Fcgr1	*Il1a*	***Tfpi***	***Slc7a2***	*Cd33*
Camk1	*Asph*		***Slc16a7***	*Cebpb*
Fgd4	*Dnase2a*	*Ptgs1*	***Slc16a10***	*Atp6ap1*
Sqrdl	*Slc38a7*	*C1qa*	***Slpi***	*Pros1*
Csf3r	*Siglece*	*Engase*	***Mitf***	*Dhrs3*
	Itgb5	*C1qb*	***Snx24***	*Rnf13*
Plod1	*Rhob*	*C1qc*	***Lyplal1***	*Man2b2*
Tom1	*Mavs*	*Timp2*	***St7***	*Ltc4s*
Myo7a	*Atp13a2*	*Slc11a1*		
A930039A15Rik	*Slc29a1*	*4632428N05Rik*	*Tlr8*	
Pld3	*Slc15a3*	*Sesn1*	*Gbp6*	
Tpp1	*Tmem86a*	*Plxnb2*	*6430548M08Rik*	
Ctsd	*Tgfbr2*	*Apoe*	*C130050O18Rik*	
Pla2g15	*Tnfrsf21*		*Pilra*	
Lamp2			*Pilrb1*	
Pla2g4a			*Lpl*	
MerTK			*Pstpip2*	
Tlr7			*Serpinb6a*	
Cd14			*Slc38a6*	
Tbxas1			*Abcc5*	
Fcgr3			*Lrp1*	
Sepp1			*Pcyox1*	
Glul			*Hmox1*	
Cd164			*Slc17a5*	
Tcn2			*Emr1*	
Dok3			*Hgsnat*	
Ctsl				
Tspan14				
Comt1				
Tmem77				
Abca1				

Genes with higher expression by all four prototypical macrophage populations (far left) or by three of the four populations (lacking (–) one of the four) than in classical or migratory DCs; bolding indicates signal intensity showing lack of expression by DCs; no bolding indicates expression in DCs, but higher expression in macrophages. Data are pooled from three or more experiments.

© 2012 Nature America, Inc. All rights reserved.

npg

the blood monocyte subsets (**Fig. 3b**), which indicated that this list of genes was selectively associated with mature macrophage differentiation in the hematopoietic system.

As a framework for future studies of the transcriptional control of the development, maintenance and function of macrophages, we examined the predicted activators assigned by the Ontogenet algorithm to the modules associated with the macrophage core genes. One example is the activators predicted by Ontogenet algorithm to control the expression of the 11 gene transcripts of module 161 (**Fig. 3b**). Overall, a highly overlapping set of 22 regulators emerged from the 14 macrophage-associated modules (**Fig. 3c**). In particular, TCFE3, C/EBP-α and Bach1 were predicted activators in a majority of these modules (>75%). Other predicted regulators, such as CREG-1 (the cellular repressor of genes stimulated by the transcription factor E1A), were unexpectedly but prominently identified. Among the 22 regulators associated with the 14 modules, 18 were predicted by Ingenuity pathway tools to interact in a regulatory network on the basis of known protein-protein interactions or mutual transcriptional regulation (**Fig. 3d**). These regulators represented five main families of transcriptional factors (**Fig. 3d**). The evaluation score generated for this network had a P value of $\leq 10^{-35}$. Beyond modules of genes that unified the four tissue macrophage populations we studied, several modules were selectively associated with a single macrophage population (**Supplementary Table 5**). In these specific modules, predicted regulators included Spi-C for red-pulp macrophages, which confirmed a regulation already known[15] and thus provided support for the predictive power of the algorithm, and GATA-6 as a regulator of peritoneal macrophages (**Supplementary Table 6**).

Core signatures to identify macrophages

Finally, we used the core signature of resting macrophages defined above to assess mononuclear phagocyte populations that we excluded from our earlier core analysis because of the paucity of information on a given population or controversy in the literature about their origins or functional properties, including whether they should be classified as DCs or macrophages. In the ImmGen Project database, each population has been assigned a classification *a priori* as DC or macrophage. For clarity and for consistency with the database, the names of these populations will be used here (and in **Fig. 4**; glossary, **Supplementary Note 2**). These populations included resting and thioglycollate-elicited mononuclear phagocytes that expressed CD11c and MHC class II (**Supplementary Fig. 5**), skin Langerhans cells, bone marrow macrophages[22], and putative CD11b$^+$ tissue DCs, including those in the liver and gut. All thioglycollate-elicited cells from the peritoneal cavity, even those that coexpressed CD11c and MHC class II, had high expression of genes in the 39-gene macrophage core and in module 161 itself, similar to the prototypic macrophage populations used to generate the core (**Fig. 4a**,**b**); this indicated that these cells were indeed macrophages despite their coexpression of CD11c and MHC class II. However, Langerhans cells and CD11c$^+$MHCII$^+$CD11b$^+$ cells from the liver (CD11b$^+$ liver DCs in the ImmGen Project database) did not have robust expression of the 39-gene macrophage core signature or module 161 alone, nor did bone marrow macrophages (**Fig. 4a**,**b**). CD11c$^+$MHCII$^+$CD11b$^+$CD103$^-$ cells from the intestinal lamina propria and CD11c^{lo}MHCII$^+$CD11b$^+$ cells from the serosa that have been called DCs in many studies expressed genes of the macrophage core signature, including those from module 161 (**Fig. 4a**,**b**), which suggested a strong relationship to macrophages. Accordingly, we call these cells 'CD11b$^+$ gut macrophages' and 'CD11c^{lo} serosal macrophages' here (and on the ImmGen Project website). We clustered those mononuclear phagocytes on the basis of their expression of the 39-gene macrophage core to model their relatedness to each other (**Fig. 4c**). Langerhans cells of the skin and bone marrow macrophages were positioned at the interface between DCs and macrophages, with a distant relationship to classical DCs, but failed to cluster with macrophages (**Fig. 4c**).

As mentioned earlier, nonlymphoid tissue CD11b$^+$ DCs have been suggested to be heterogeneous[19]. Thus, we reasoned that the use of

antibodies to cell-surface proteins identified as macrophage specific by our gene-expression analysis might be used to identify macrophage 'contaminants' in a heterogeneous population. Furthermore, we aimed to determine if the same cell-surface markers might also prove valuable in identifying macrophages universally, including identification in organs beyond those we initially analyzed and/or those in which F4/80 has not proven sufficiently definitive. We selected the lipopolysaccharide receptor CD14, the FcγRI CD64 and the kinase MerTK as

Table 2 Genes upregulated in two of four tissue macrophage populations

Peritoneal + splenic red pulp	Peritoneal + lung	Lung + splenic red pulp	Peritoneal + microglia	Lung + microglia	Microglia + splenic red pulp
Ccl24	***Marco***	***Dmxl2***	***Hnmt***	***Scamp5***	***Lhfpl2***
Gstk1	***P2ry2***	***Dip2c***	***Mtus1***	***Ppp1r9a***	***Osm***
Aspa	***Aifm2***	***Galnt3***	***C3ar1***	***Tppp***	***Mgll***
2810405K02Rik	***Clec4e***	***Niacr1***	***Dagla***	***Abcb4***	***Bhlhe41***
B430306N03Rik	***Plcb1***	***Bckdhb***	***Wrb***	***Kcnj2***	***Ang***
Fcna	***Kcnn3***	***Angptl4***	***Gab1***	***P2ry12***	***D8ertd82e***
Gm5970	***Arhgap24***	***Lrp4***	***Fkbp9***		
Aoah	***Cd93***	***Sh3bgrl2***		*Slc37a2*	*X99384*
Cd5l	***Fundc2***	***Gm5150***	*Rab11fip5*	*Adrb1*	*Serpine1*
Gm4951	***Tspan32***	***Tcfec***	*6230427j02rik*	*Slc16a6*	*Abhd12*
Nr1d1	***Lmbr1***	***Sh2d1b1***	*Scn1b*	*Rab3il1*	*Ms4a6d*
Mlkl	***Adarb1***	***Galnt6***	*Scamp1*	*Mfsd11*	*Cebpa*
Vnn3	***Fzd4***	***Pdgfc***	*Msrb2*	*Flcn*	*Lpcat3*
Igf1	***F7***		*Abca9*	*Tmem63a*	*Manea*
	Ccr1	*6720489N17Rik*	*Plxdc2*	*P2rx7*	*Ctss*
Ptgis	***Hspa12a***	*Pparg*	*Adam15*	*Hpgds*	*Ccl3*
Pitpnc1	***Cav1***	*Megf9*	*Itgam*	*Hpgd*	*Cryl1*
Fam43a	***Nt5e***	*Adcy3*	*Itga6*	*Lpcat2*	*Man1c1*
Itsn1	***1190002a17rik***	*Enpp1*	*Vkorc1*	*Slc7a8*	*Ctns*
Ifi27l1		*Il18*	*1700017b05rik*	*Maf*	*Sgk1*
Rasgrp2	*Cav2*	*Siglec1*	*Smad3*	*Tmem86a*	*Pag1*
Aldh6a1	*Gda*	*Clec4n*	*Smpd1*	*Slc36a1*	*Tgfbr1*
Epb4.1l1	*Frrs1*	*Lgals8*	*Naglu*	*Gna12*	*Clec5a*
Cryzl1	*Tspan5*	*Nceh1*	*Pmp22*	*Adap2*	
Lrp12	*Pdk4*	*Lipa*	*Man2b2*	*Lgmn*	
Cd300ld	*Slc36a4*	*4931406c07rik*	*Tnfrsf1a*	*Hist1h1c*	
Pla2g7	*Fam3c*	*Sirpa*	*Lifr*	*Lair1*	
Cfp	*Ms4a8a*	*Rasgef1b*	*Tlr13*	*Slc40a1*	
Sdc3	*Atoh1*	*Wdfy3*	*Slc25a37*	*Csf1r*	
Dusp7	*Alox5*	*Ermp1*	*Grn*	*P4ha1*	
Tbc1d2b	*Thbd*	*Asah1*		*Iffo1*	
Igsf6	*Gstm1*	*Ear1*		*Dusp6*	
Man2a1	*Cxcl2*	*Ear10*			
Zswim6	*Nhlrc3*	*Ano6*			
Ifnar2	*Fry*	*Mrc1*			
Trf	*F10*	*Camk2d*			
Blvrb	*Sord*	*Gab3*			
Cd38	*Ncf2*	*Syne2*			
Ctsb	*Hexa*	*Axl*			
Tmem87b	*Dram1*	*Tcf7l2*			
Itfg3	*Plaur*	*Ctsc*			
Ninj1	*G6pdx*	*D730040f13rik*			
	Fn1	*Slc15a3*			
	Cybb	*Plk3*			
	Dennd4c	*Hebp1*			
	Mpp1	*Dst*			
	S100a1	*Blvra*			
	Gsr	*Sort1*			
	Abcd2	*Slc12a7*			
	Dab2	*Clec4a3*			
	Ccl6				
	Sepx1				
	Prdx5				
	Dusp3				
	Pgd				
	Gp49a				
	Capg				
	Cndp2				
	Vps13c				
	Adipor2				
	App				
	Atg7				
	Cebpb				

Genes with higher expression by two of four prototypical macrophage populations than in classical or migratory DCs (bolding and debolding as in **Table 1**). Data are pooled from three or more experiments.

© 2012 Nature America, Inc. All rights reserved.

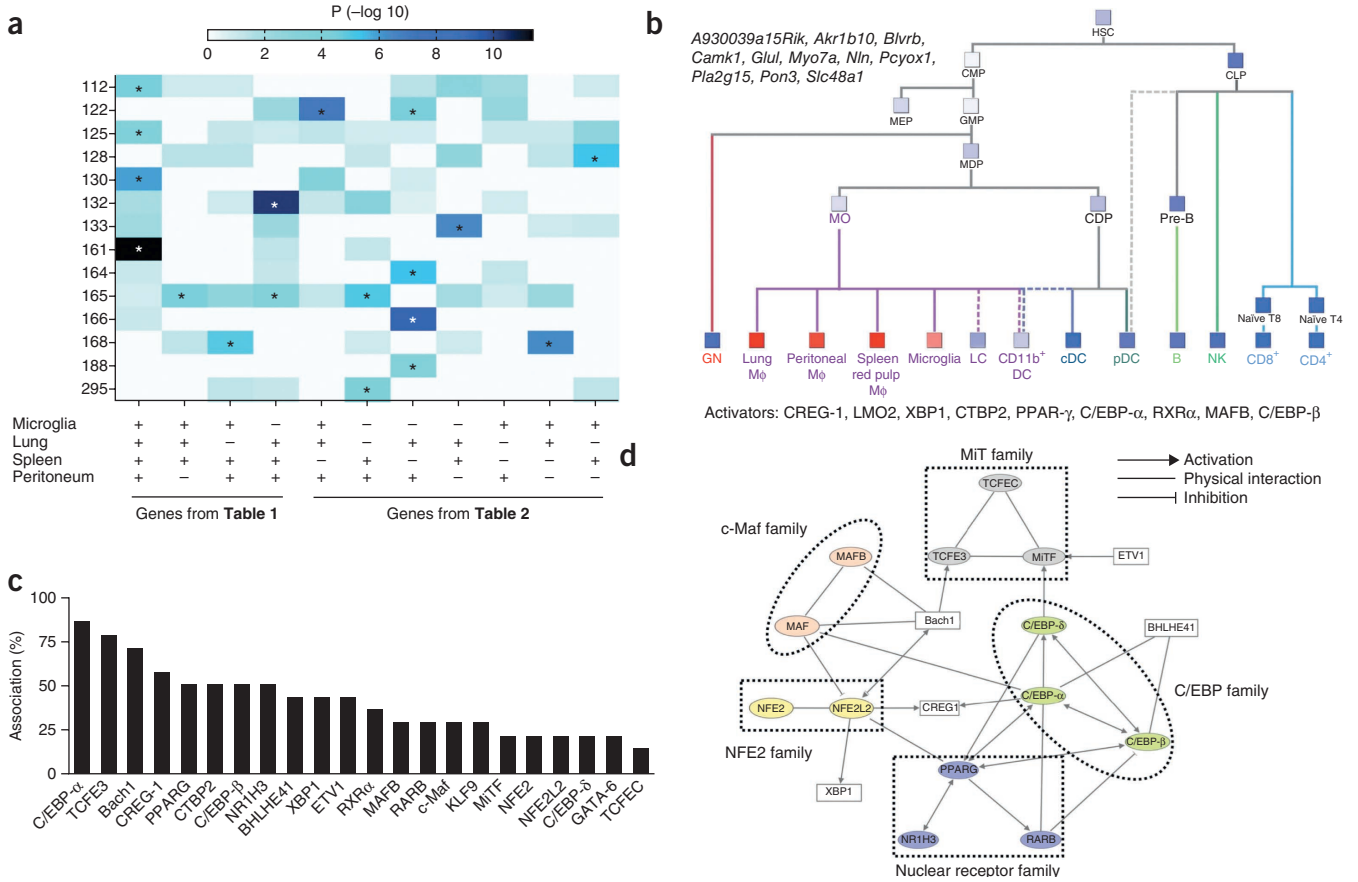


Figure 3 Identification of gene modules enriched for macrophage-related gene signatures and their predicted regulators. (**a**) Heat map of the overlap of ImmGen modules (left margin) and all macrophage-associated gene signatures (**Tables 1** and **2**), including only modules with significant enrichment for at least one signature. Stars indicate significant overlap (hypergeometric test). (**b**) Simplified hematopoietic tree showing expression (red, high; blue, low; purple, intermediate) of genes in module 161 (top left); bottom, predicted positive regulators of the module. HSC, hematopoietic stem cell; CMP, common myeloid progenitor; CLP, common lympoid progenitor; MEP, megakaryocyte-erythrocyte progenitor; GMP, granulocyte-macrophage precursor; MDP, macrophage and DC precursor; MO, monocyte; CDP, common DC precursor; Pre-B, pre-B cell; GN, granulocyte; LC, Langerhans cell; cDC, classical DC; pDC, plasmacytoid DC; B, B cell; NK, natural killer cell; T8 or CD8+, CD8+ T cell; T4 or CD4+, CD4+ T cell. (**c**) Frequency of the association between the 14 modules in **a** and positive regulators (activators; horizontal axis) predicted by the Ontogenet algorithm to regulate two or more of those modules. (**d**) Physical and regulatory interactions between the 18 regulators most frequently represented in the 14 macrophage-associated modules, assessed with Ingenuity pathway-analysis software; straight lines indicate links for established physical interactions; arrows and sideways 'T' symbols indicate known pathways of coactivation and inhibition, respectively. Data are pooled from three or more experiments.

© 2012 Nature America, Inc. All rights reserved.

cell-surface proteins among the group of proteins encoded by the 39 mRNA transcripts with expression deemed to be low or absent in DCs but present in all macrophages and to which high-quality monoclonal antibodies have been generated. Indeed, all of these proteins were expressed on all of the four resident macrophage populations used in our primary analysis (**Fig. 5a**), with lower expression of CD14 than of CD64 or MerTK (**Fig. 5a**). Two of these tissues, spleen and lung, have substantial DC populations. In the spleen, antibodies to MerTK, CD64 or CD14 did not stain CD8$^+$ or CD11b$^+$ DCs (**Fig. 5a**). However, in the lung, in which interstitial pulmonary macrophages are CD11b$^-$, there may still be an underlying heterogeneity of lung CD11b$^+$ DCs that includes a subset of CD11b$^+$ macrophages[19,23,24]. Indeed, CD14, CD64 and MerTK were expressed by a portion of lung CD11b$^+$ DCs but not by CD103$^+$ DCs (**Fig. 5a**). Gating on MerTK$^+$CD64$^+$ cells showed most of these cells were Siglec-F$^+$ lung macrophages, but a small proportion of MerTK$^+$CD64$^+$ cells in the lung were Siglec-F$^-$ cells with high expression of MHC class II (**Fig. 5b**). By our usual gating strategy for lung DCs (**Fig. 5c**), DCs were defined as Siglec F$^-$CD11c$^+$MHCII$^+$ cells. However, the small population of Siglec-F$^-$MerTK$^+$CD64$^+$ cells that may instead have been macrophages (**Fig. 5b**) were partially in the standard DC gate (**Fig. 5c**). Indeed, we were able to separate CD11b$^+$ DCs into CD11b$^+$CD24$^+$CD64^{lo}MerTK$^-$CD14^{int} cells and CD11b$^+$CD24^{lo}CD64$^+$MerTK$^+$CD14^{hi} cells (**Fig. 5d**). Thus, the latter was probably a population of macrophages that segregated together with DCs, through the use of many markers, but were not DCs. Indeed, the CD11b$^+$ DCs were segregated by CD24 expression in the ImmGen Project on the basis of the likelihood that those expressing CD24 were true DCs but those without CD24 were not. Our findings suggested that this possibility was likely and indicated the utility of using markers such as MerTK and CD64 as a panel to facilitate the identification of macrophages versus DCs.

We next turned to two tissues, liver and adipose, not analyzed by the ImmGen Project in terms of gene-expression profiling of macrophages to determine if the use of staining for MerTK and CD64 would facilitate the identification of macrophages in those tissues and distinguish them from DCs. In the liver, we started with a classic approach of plotting F4/80 expression versus CD11c expression. Eosinophils are now recognized as F4/80$^+$ cells with high side

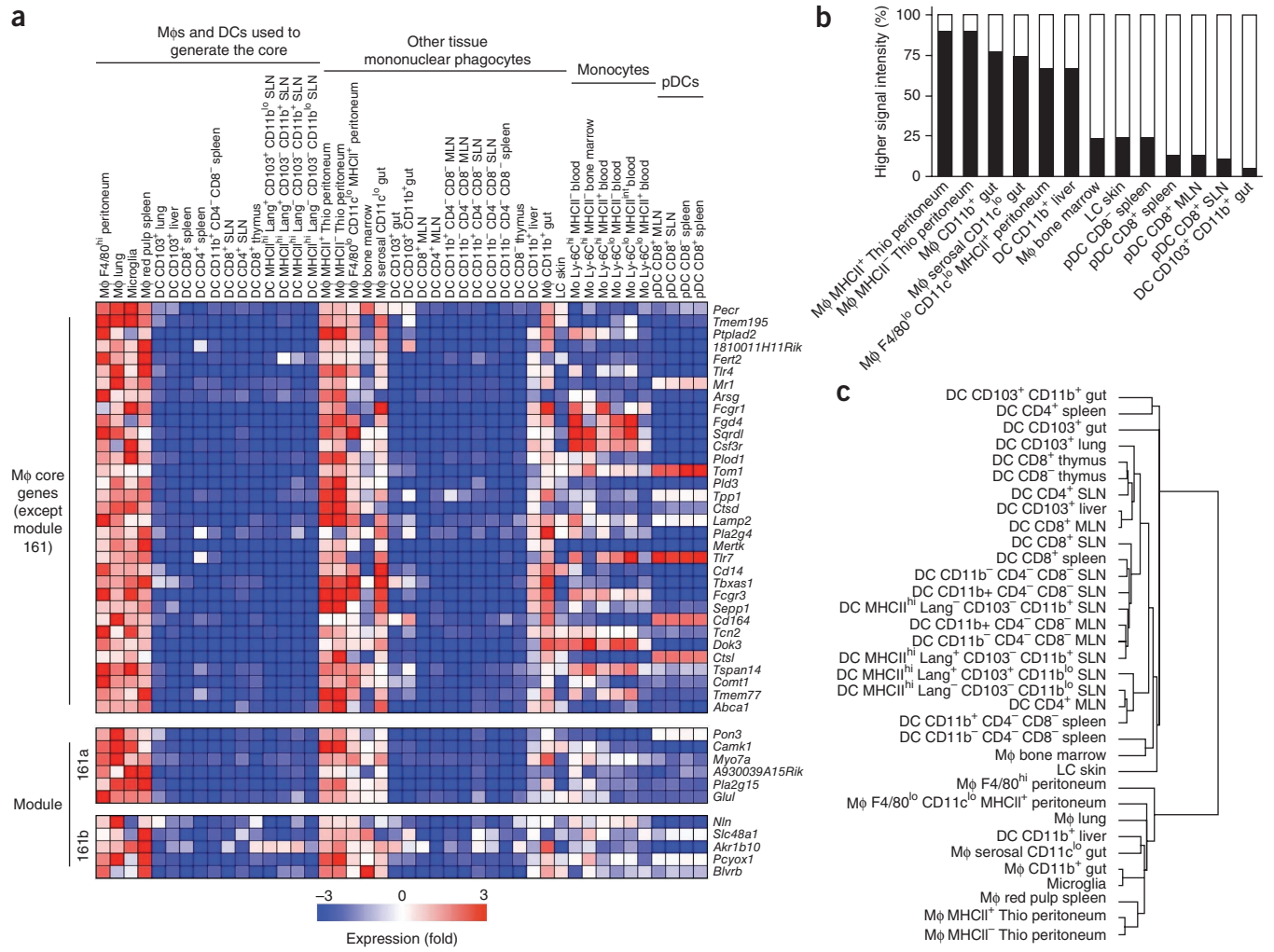


Figure 4 Expression of genes of the macrophage core signature by other populations of mononuclear phagocytes. (**a**) Heat map of 39 genes (right margin) with higher expression in spleen, brain, peritoneal and lung macrophages than in classical or migratory DCs (left); middle, tissue-derived mononuclear phagocytes not included in the generation of this list of genes; right, blood monocytes and plasmacytoid DCs; 161a, genes from module 161 included here; 161b, genes from module 161 that did not meet the criteria for inclusion (**Table 1**). SLN, skin lymph node; MLN, mesenteric lymph node; Thio, thioglycollate-elicited. (**b**) Frequency of signal intensity for the 39 genes in **a** at least twofold higher in various populations (horizontal axis) than in the DC population with the highest expression (filled bars; open bars indicate extension up to 100%). (**c**) Relationships among various mononuclear phagocytes based on their expression of the 39 genes in **a**. Lang, langerin; LC, Langerhans cell. Data are pooled from three or more experiments.

scatter that express Siglec-F universally[25]. Indeed, among macrophages, Siglec-F is observed only on macrophages in the lung[26,27] (as used to identify lung macrophages here; eosinophils did not contaminate lung macrophages, which we separated from eosinophils by their high CD11c expression and relative lack of CD11b expression in the macrophages (**Supplementary Fig. 6**)). In the liver, the abundance of F4/80 on eosinophils overlapped that of another population of F4/80$^+$ cells (those with low side scatter) that were CD11c^{lo} in liver (**Fig. 5e**). Even after excluding eosinophils, we found four gates of cells with various expression of F4/80 and CD11c (**Fig. 5e**). There was high expression of MerTK and CD64 in two of these gates, one composed of cells with the highest expression of F4/80 (gate 2) and another with lower expression of F4/80 (in gate 3). These findings suggested two populations of F4/80^{hi} and F4/80^{lo} liver macrophages that may correspond to the two types of macrophages believed to be present in many organs[12]. The liver CD45$^+$ cells with highest expression of CD11c were MerTK$^-$CD64$^-$ (**Fig. 5e**), which suggested they were liver DCs. Reverse gating showed that all MerTK$^+$CD64$^+$ cells were in one of the two putative macrophage gates (**Fig. 5e**). Gate 1 without eosinophils probably contained blood monocytes, which were not MerTK$^+$. We noted relatively similar results for adipose tissue (**Fig. 5f**), in which the cells with the highest F4/80 expression were MerTK$^+$CD64$^+$ and those with higher CD11c expression and lower F4/80 expression were MerTK$^-$CD64$^-$. In both liver and adipose tissue, expression of MHC class II was high on macrophages and DCs (**Fig. 5e**,**f**). Because F4/80 and CD11c are both expressed by many tissue macrophages and DCs, albeit in amounts that are somewhat different, distinguishing macrophages and DCs on the basis of these traditional markers can be difficult. Staining for MerTK and CD64 offers the advantage of distinct differences in the magnitude of expression in macrophages versus DCs. Thus, we propose that costaining for MerTK and CD64 provides a powerful approach for identifying macrophages universally and selectively in mouse tissues.

© 2012 Nature America, Inc. All rights reserved.

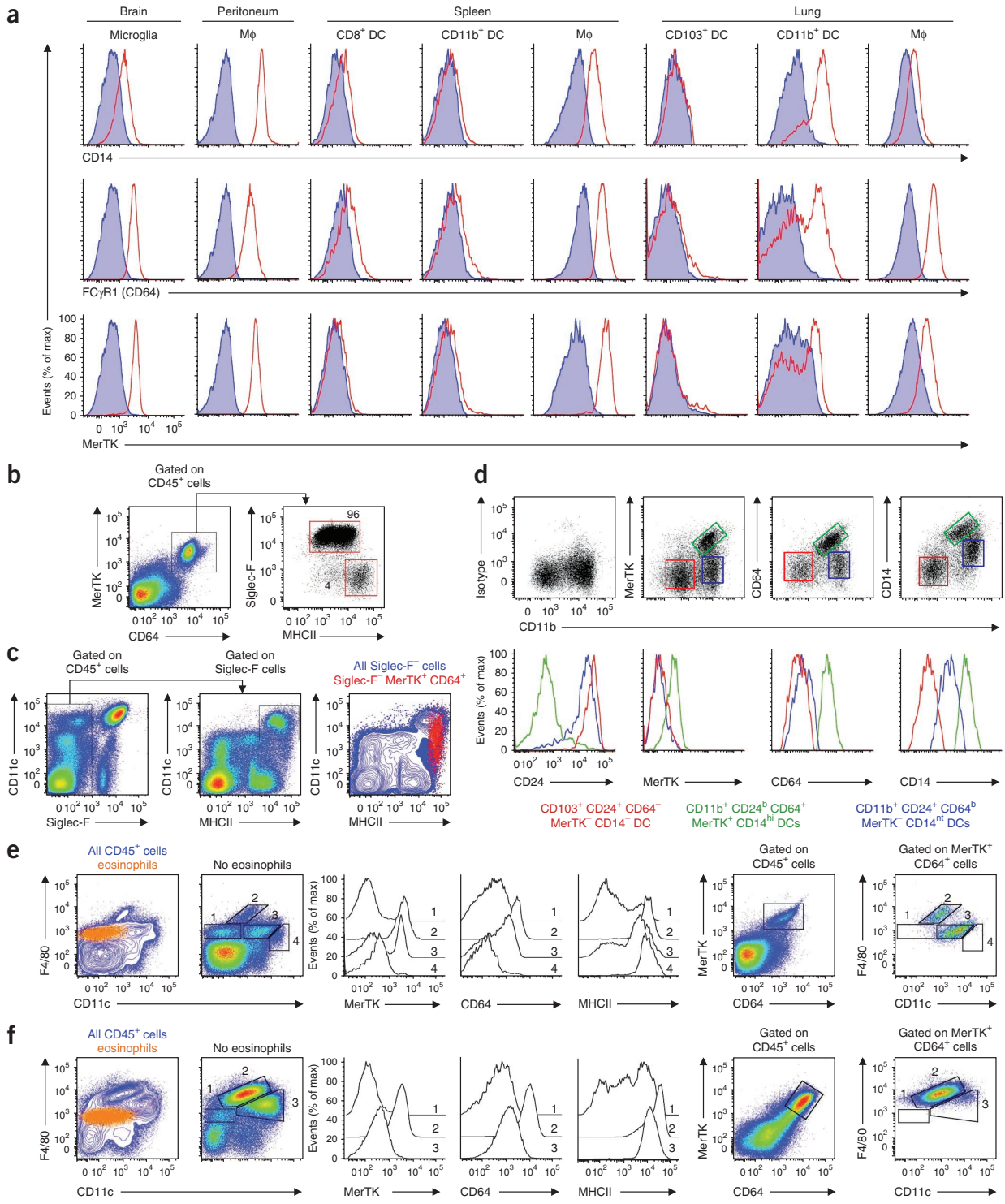


Figure 5 Transcripts of the macrophage core signature, assessed as protein in various tissues. (**a**) Expression of CD14, CD64 and MerTK in macrophages and DCs of brain, peritoneum, spleen and lung. Red line, specific antibody; blue shading, isotype-matched control antibody. (**b**) Flow cytometry of MerTK$^+$CD64$^+$ cells (gated at left), assessing expression of Siglec-F and MHC class II (right). Numbers adjacent to outlined areas indicate percent Siglec-F$^+$MHCII$^+$ cells (top) and Siglec-F$^+$MHCII$^+$ cells (bottom). (**c**) Gating strategy for lung DCs; DCs are CD45$^+$ cells that express CD11c and MHC class II but not Siglec-F. (**d**) Flow cytometry of DCs, gated on lung DCs, assessing expression of CD11b with MerTK, CD64 and CD14 (top), and of CD24, MerTK, CD64 and CD14 (bottom). (**e**,**f**) Flow cytometry of liver (**e**) or adipose (**f**) CD45$^+$ cells stained for F4/80 and CD11c (left), with eosinophils gated (Siglec-F$^+$ and high side scatter; far left) and without eosinophils (second from left), to identify four (**e**) or three (**f**) subsets of cells with differences in their expression of F4/80 and CD11c, followed by analysis of the expression of MerTK, CD64 and MHC class II by those subsets (middle), then reverse gating on MerTK$^+$CD64$^+$ cells (second from right) and analysis of the expression of F4/80 and CD11c by those gated cells (far right). Data are from at least two independent experiments with three mice per group.

© 2012 Nature America, Inc. All rights reserved.

DISCUSSION

The large and unique database and accompanying bioinformatics analysis of the ImmGen Project provide insight into macrophage populations isolated from various organs of mice. A notable initial revelation was that macrophage populations from different organs were considerably diverse, and it is likely that further profiling in macrophages will expand on this diversity. Only a very small group of mRNA transcripts were associated with all macrophages but not DCs. Proteins previously predicted to distinguish macrophages from other cell types, such as F4/80, CD68 and CD115 (c-Fms or CSF1R), did not emerge as the most powerful markers of macrophages. However, many canonical genes did, including those encoding CD14, CD64 (the high-affinity Fcγ receptor I), MerTK (the kinase involved in efferocytosis), cathespin D and the kinase Fert2 (which may have a substantial effect on CD115 signaling but which has not yet been studied in macrophages). The identification of these as being selectively associated with macrophages reinforced the idea of a key role for macrophages in innate immunity, efferocytosis and the clearance of debris, whereas genes encoding molecules associated with antigen presentation and migration to lymphoid tissues were more associated with DCs[28]. However, our data did suggest that macrophages may have a greater role than DCs in the activation of mucosa-associated invariant T cells. On the basis of follow-up protein-expression analysis of MerTK and CD64 in macrophages from six different tissues, we propose that analysis of MerTK and CD64 should serve as a starting point for the identification of macrophages in tissues, as staining for these markers seemed to identify F4/80^{hi} macrophages and other macrophages with somewhat lower F4/80 expression[12] in all tissues. We believe that staining for MerTK and CD64 provides an advantage over traditional staining for F4/80, CD11c and MHC class I but can also be used powerfully in addition to such staining. The expression of F4/80 and CD11c often overlaps in macrophages and DCs in nonlymphoid tissues, but it seems that DCs do not coexpress MerTK and CD64.

Beyond those cell-surface markers closely associated with macrophage identity, we identified other transcripts associated only with macrophages among hematopoietic cells. In particular, module 161 of the ImmGen Project identified a group of genes (*A930039a15Rik*, *Akr1b10*, *Blvrb*, *Camk1*, *Glul*, *Myo7a*, *Nln*, *Pcyox1*, *Pla2g15*, *Pon3* and *Slc48a*) coexpressed across the entire data set of the ImmGen Project and that encode molecules with functions compatible with the function of macrophages, but none of them have previously been considered macrophage markers. Both the genes from this module and their predicted regulators deserve attention in the future.

The Ontogenet algorithm makes it possible to extend the macrophage-associated genes we identified to regulatory programs that may control them. The finding of induction of expression of a single module (330) in red-pulp macrophages relative to its expression all other macrophages and the predictions generated by the algorithm indicating that this module is regulated by Spi-C supported the reliability of the prediction of the regulatory programs by the algorithm, as Spi-C is already known to be required selectively for the development or maintenance of red-pulp macrophages[15]. Additional information has also emerged, such as the association of modules unique to peritoneal macrophages that are predicted to be regulated by GATA-6.

Gene transcripts with high expression in multiple macrophage populations but without high expression in DCs were associated with predicted transcriptional regulatory programs that differed considerably from those identified in DCs[28]. The predicted regulatory programs of modules enriched for macrophage-associated genes included several members of the MiT family of transcription factors recognized as being expressed specifically in macrophages[3], as well as transcription factors not previously associated with macrophages, such as Bach1 and CREG-1. Bach1 has been studied very little in macrophages but has been linked to osteoclastogenesis[29] and is a regulator of heme oxygenase 1 (ref. 30). CREG-1 is a secreted regulator[31,32] associated broadly with differentiation[33] and cellular senescence[34] that has been associated with macrophage-enriched gene modules, although it has never been studied before in the context of macrophage biology, to our knowledge. The Ontogenet algorithm predicted that RXRα is the most prominent key activator of the highly specific and universal macrophage module 161. Future analysis of these predictions should be useful in showing how macrophage identity and function is controlled.

So far, the ImmGen Project has focused mainly on cells recovered from resting, uninfected mice, in which macrophages derive mainly from the yolk sac[12]. Macrophage polarization in the context of infection and inflammation is a topic of great interest that this study has scarcely been able to address beyond finding that monocytes recruited to the peritoneum in response to thioglycollate upregulated the expression of mRNA transcripts observed in resting tissue macrophages, even though it now seems that monocytes are not precursors of resting tissue macrophages as they are of inflammatory macrophages. The foundations laid here suggest that future additions to the ImmGen Project database of macrophages recovered in disease states should add to the understanding of how to manipulate these crucial cells to favor desired outcomes in disease. Given the great diversity of macrophages in different organs, which we anticipate is present even in inflamed organs, such studies may be expected to ultimately generate therapeutic approaches to selectively target macrophages in diseased organs without affecting other cell types.

METHODS

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

Accession codes. GEO: microarray data, GSE15907.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank our colleagues of the ImmGen Project consortium; V. Jojic, J. Ericson, S. Davis and C. Benoist for contributions; eBioscience and Affymetrix for material support of the ImmGen Project; and M. Colonna (Washington University School of Medicine) for monoclonal antibodies (including anti-Siglec-H) and other reagents. Supported by the National Institute of Allergy and Infectious Diseases of the US National Institutes of Health (R24 AI072073 to fund the ImmGen Project, spearheaded by C. Benoist), the US National Institutes of Health (R01AI049653 and R01AI061741 to G.J.R.; P50GM071558-03 and R01DK08854 to A.M.; and5T32DA007135-27 to A.R.M.) and the American Heart Association (10POST4160140 to E.L.G.).

AUTHOR CONTRIBUTIONS

E.L.G. purified macrophage populations, designed and did experiments, analyzed data and wrote the manuscript; G.J.R. designed and supervised experiments, analyzed data and wrote the manuscript; T.S. analyzed data and wrote the manuscript; J.M. designed analytical strategies and analyzed data; M.G., C.J., J.H., A.C. and K.G.E. purified macrophage and DC populations; S.I. did experiments; S.G., A.R.M. and A.M. analyzed data; W.-J.C. and T.H.H. provided reagents and supervised experiments; and S.J.T. and M.M. designed and supervised experiments.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ni.2419.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

© 2012 Nature America, Inc. All rights reserved.
npg

1. Heng, T.S. & Painter, M.W. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008).
2. Gordon, S. & Taylor, P.R. Monocyte and macrophage heterogeneity. *Nat. Rev. Immunol.* **5**, 953–964 (2005).
3. Hume, D.A. Differentiation and heterogeneity in the mononuclear phagocyte system. *Mucosal Immunol.* **1**, 432–441 (2008).
4. Mosser, D.M. & Edwards, J.P. Exploring the full spectrum of macrophage activation. *Nat. Rev. Immunol.* **8**, 958–969 (2008).
5. Geissmann, F., Gordon, S., Hume, D.A., Mowat, A.M. & Randolph, G.J. Unraveling mononuclear phagocyte heterogeneity. *Nat. Rev. Immunol.* **10**, 453–460 (2010).
6. Fogg, D.K. *et al.* A clonogenic bone marrow progenitor specific for macrophages and dendritic cells. *Science* **311**, 83–87 (2006).
7. Onai, N. *et al.* Identification of clonogenic common Flt3^{+}M-CSFR^{+} plasmacytoid and conventional dendritic cell progenitors in mouse bone marrow. *Nat. Immunol.* **8**, 1207–1216 (2007).
8. Liu, K. *et al. In vivo* analysis of dendritic cell development and homeostasis. *Science* **324**, 392–397 (2009).
9. Hildner, K. *et al.* Batf3 deficiency reveals a critical role for CD8α^{+} dendritic cells in cytotoxic T cell immunity. *Science* **322**, 1097–1100 (2008).
10. Takahashi, K. Development and differentiation of macrophages and related cells: Historical review and current concepts. *J. Clin. Exp. Hematop.* **41**, 1–33 (2001).
11. Ginhoux, F. *et al.* Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science* **330**, 841–845 (2010).
12. Schulz, C. *et al.* A lineage of myeloid cells independent of Myb and hematopoietic stem cells. *Science* **336**, 86–90 (2012).
13. Aziz, A., Soucie, E., Sarrazin, S. & Sieweke, M.H. MafB/c-Maf deficiency enables self-renewal of differentiated functional macrophages. *Science* **326**, 867–871 (2009).
14. Teitelbaum, S.L. & Ross, F.P. Genetic regulation of osteoclast development and function. *Nat. Rev. Genet.* **4**, 638–649 (2003).
15. Kohyama, M. *et al.* Role for Spi-C in the development of red pulp macrophages and splenic iron homeostasis. *Nature* **457**, 318–321 (2009).
16. Lemke, G. & Rothlin, C.V. Immunobiology of the TAM receptors. *Nat. Rev. Immunol.* **8**, 327–336 (2008).
17. Ohl, L. *et al.* CCR7 governs skin dendritic cell migration under inflammatory and steady-state conditions. *Immunity* **21**, 279–288 (2004).
18. Ginhoux, F. *et al.* Blood-derived dermal langerin^{+} dendritic cells survey the skin in the steady state. *J. Exp. Med.* **204**, 3133–3146 (2007).
19. Hashimoto, D., Miller, J. & Merad, M. Dendritic cell and macrophage heterogeneity in vivo. *Immunity* **35**, 323–335 (2011).
20. Chua, W.J. *et al.* Endogenous MHC-related protein 1 is transiently expressed on the plasma membrane in a conformation that activates mucosal-associated invariant T cells. *J. Immunol.* **186**, 4744–4750 (2011).
21. Elstad, M.R., Stafforini, D.M., McIntyre, T.M., Prescott, S.M. & Zimmerman, G.A. Platelet-activating factor acetylhydrolase increases during macrophage differentiation. A novel mechanism that regulates accumulation of platelet-activating factor. *J. Biol. Chem.* **264**, 8467–8470 (1989).
22. Chow, A. *et al.* Bone marrow CD169^{+} macrophages promote the retention of hematopoietic stem and progenitor cells in the mesenchymal stem cell niche. *J. Exp. Med.* **208**, 261–271 (2011).
23. Hashimoto, D. *et al.* Pretransplant CSF-1 therapy expands recipient macrophages and ameliorates GVHD after allogeneic hematopoietic cell transplantation. *J. Exp. Med.* **208**, 1069–1082 (2011).
24. Satpathy, A.T. *et al.* Zbtb46 expression distinguishes classical dendritic cells and their committed progenitors from other immune lineages. *J. Exp. Med.* **209**, 1135–1152 (2012).
25. Kim, H.J., Alonzo, E.S., Dorothee, G., Pollard, J.W. & Sant'Angelo, D.B. Selective depletion of eosinophils or neutrophils in mice impacts the efficiency of apoptotic cell clearance in the thymus. *PLoS ONE* **5**, e11439 (2010).
26. Sung, S.S. *et al.* A major lung CD103 (α_E)-β_7 integrin-positive epithelial dendritic cell population expressing langerin and tight junction proteins. *J. Immunol.* **176**, 2161–2172 (2006).
27. Desch, A.N. *et al.* CD103^{+} pulmonary dendritic cells preferentially acquire and present apoptotic cell-associated antigen. *J. Exp. Med.* **208**, 1789–1797 (2011).
28. Miller, J.C. *et al.* Deciphering the transcriptional network of the dendritic cell lineage. *Nat. Immunol.* **13**, 888–899 (2012).
29. Hama, M. *et al.* Bach1 regulates osteoclastogenesis via both heme oxygenase-1 dependent and independent pathways. *Arthritis Rheum.* **64**, 1518–1528 (2012).
30. Sun, J. *et al.* Hemoprotein Bach1 regulates enhancer availability of heme oxygenase-1 gene. *EMBO J.* **21**, 5216–5224 (2002).
31. Veal, E., Eisenstein, M., Tseng, Z.H. & Gill, G. A cellular repressor of E1A-stimulated genes that inhibits activation by E2F. *Mol. Cell. Biol.* **18**, 5032–5041 (1998).
32. Sacher, M. *et al.* The crystal structure of CREG, a secreted glycoprotein involved in cellular growth and differentiation. *Proc. Natl. Acad. Sci. USA* **102**, 18326–18331 (2005).
33. Veal, E., Groisman, R., Eisenstein, M. & Gill, G. The secreted glycoprotein CREG enhances differentiation of NTERA-2 human embryonal carcinoma cells. *Oncogene* **19**, 2120–2128 (2000).
34. Moolmuang, B. & Tainsky, M.A. CREG1 enhances p16^{INK4a}-induced cellular senescence. *Cell Cycle* **10**, 518–530 (2011).

© 2012 Nature America, Inc. All rights reserved.

ImmGen Consortium:

Emmanuel L Gautier[10,11], Claudia Jakubzick[10], Gwendalyn J Randolph[10,11], Adam J Best[12], Jamie Knell[12], Ananda Goldrath[12], Jennifer Miller[10], Brian Brown[10], Miriam Merad[10], Vladimir Jojic[13], Daphne Koller[13], Nadia Cohen[14], Patrick Brennan[14], Michael Brenner[14], Tal Shay[15], Aviv Regev[15], Anne Fletcher[16], Kutlu Elpek[16], Angelique Bellemare-Pelletier[16], Deepali Malhotra[16], Shannon Turley[16], Radu Jianu[17], David Laidlaw[17], Jim Collins[18], Kavitha Narayan[19], Katelyn Sylvia[19], Joonsoo Kang[19], Roi Gazit[20], Brian S Garrison[20], Derrick J Rossi[20], Francis Kim[21], Tata Nageswara Rao[21], Amy Wagers[21], Susan A Shinton[22], Richard R Hardy[22], Paul Monach[23], Natalie A Bezman[24], Joseph C Sun[24], Charlie C Kim[24], Lewis L Lanier[24], Tracy Heng[25], Taras Kreslavsky[13], Michio Painter[25], Jeffrey Ericson[25], Scott Davis[25], Diane Mathis[25] & Christophe Benoist[25]

[10]Icahn Medical Institute, Mount Sinai Hospital, New York, New York, USA. [11]Department of Pathology & Immunology, Washington University, St. Louis, Missouri, USA. [12]Division of Biological Sciences, University of California San Diego, La Jolla, California, USA. [13]Computer Science Department, Stanford University, Stanford, California, USA. [14]Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston, Massachusetts, USA. [15]Broad Institute, Cambridge, Massachusetts, USA. [16]Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts, USA. [17]Computer Science Department, Brown University, Providence, Rhode Island, USA. [18]Department of Biomedical Engineering, Howard Hughes Medical Institute, Boston University, Boston, Massachusetts, USA. [19]Department of Pathology, University of Massachusetts Medical School, Worcester, Massachusetts, USA. [20]Immune Diseases Institute, Children's Hospital, Boston, Massachusetts, USA. [21]Joslin Diabetes Center, Boston, Massachusetts, USA. [22]Fox Chase Cancer Center, Philadelphia, Pennsylvania, USA. [23]Department of Medicine, Boston University, Boston, Massachusetts, USA. [24]Department of Microbiology & Immunology, University of California San Francisco, San Francisco, California, USA. [25]Division of Immunology, Department of Microbiology & Immunobiology, Harvard Medical School, Boston, Massachusetts, USA.

ONLINE METHODS

Mice. Six-week-old male C57BL/6J mice (Jackson Laboratory) were used for sorting and confirmation of results. Mice with sequence encoding green fluorescent protein knocked in to the gene encoding CX3CR1 were from Jackson Laboratories, and *Mr1*-deficient mice[20] were generated, bred and maintained at the Washington University School of Medicine. Mice were housed in specific pathogen–free facilities at the Mount Sinai School of Medicine or Washington University School of Medicine, and experimental procedures were done in accordance with the animal-use oversight committees at these institutions. Most cell populations were sorted from resting mice. However, for thioglycollate-elicited peritoneal macrophages, macrophages were collected from the peritoneal cavity 5 d after the instillation of 1 ml of 3% thioglycollate.

Cell identification and isolation. All cells were purified according to the sorting protocol and antibodies on the ImmGen website (http://www.immgen.org/Protocols/ImmGen Cell prep and sorting SOP.pdf). Cells were sorted directly from mouse tissues and were processed from tissue procurement to a second round of sorting into TRIzol within 4 h with a Beckton-Dickinson FACSAria II. Resting red-pulp macrophages from the spleen were sorted after nonenzymatic disaggregation of the spleen and were identified as F4/80^{hi} cells that lacked B220 but had high expression of CD11c and MHC class II[35,36]; macrophages from the resting peritoneum were collected in a peritoneal lavage and were stained to identify CD115^{hi} cells that were F4/80^{hi}MHCII^{-}; resting pulmonary macrophages were isolated from lungs digested for 15 min with Liberase III and macrophages were identified as Siglec-F^{+}CD11c^{+} cells with low expression of MHC class II[26,27]; and resting brain microglial macrophages were sorted from cells separated by Percoll-gradient centrifugation and digested with Liberase III and were CD11b^{+}CD45^{lo}F4/80^{lo} (ref. 11). Liver and epididymal adipose tissues were digested for 45 min in collagenase D and Liberase III, respectively. Liver cells were further separated on a Percoll gradient, whereas adipocytes were floated for separation from the stromal vascular fraction containing CD45^{+} cells in adipose tissue. The data browser of the ImmGen Project website includes files of flow cytometry plots showing the purification strategies and purity after isolation of these and all other populations. A list of abbreviations used in the ImmGen Project database relevant to macrophages and DCs is in **Supplementary Note 2**.

Microarray analysis, normalization, and data-set analysis. RNA was amplified and hybridized on the Affymetrix Mouse Gene 1.0 ST array by the ImmGen Project consortium with double-sorted cell populations sorted directly into TRIzol (http://www.immgen.org/Protocols/Total RNA Extraction with Trizol.pdf). These procedures followed a highly standardized protocol for data generation and documentation of quality control[37] (http://www.immgen.org/Protocols/ImmGen QC Documentation_ALL-DataGeneration_0612.pdf). A table listing QC data, replicate information, and batch information for each sample is also available on the ImmGen Project website. Data were analyzed with the GenePattern genomic analysis platform (http://www.broadinstitute.org/cancer/software/genepattern/). Raw data were normalized with the robust multi-array algorithm, returning linear values between 10 and 20,000. A common threshold for positive expression at 95% confidence across the data set was determined to be 120 (http://www.immgen.org/Protocols/ImmGen QC Documentation_ALL-DataGeneration_0612.pdf). Differences in gene expression signatures were identified and visualized with the Multiplot module of GenePattern. Probe sets were considered to have a difference in expression with a coefficient of variation of <0.5 and a *P* value of ≥0.05 (Student's *t*-test). Signature transcripts were clustered (centered on the mean) with the Hierarchical Clustering module of GenePattern, employing Pearson's correlation as a metric, and data were visualized with the Hierarchical Clustering Viewer heat-map module. Clustering analyses across the database of the ImmGen Project centered on the most variable gene sets (objectively defined as the top 15% genes, ranked by cross-population maximum/minimum ratio) to avoid 'noise' from genes at background variation. Pearson correlation plots of gene-expression profiles for various cell populations were generated with Express Matrix software. Pathway analysis as well as construction of the transcription factor network were done with Ingenuity Systems Pathway Analysis software. This software calculates a significance score for each network. The score is generated with a *P* value indicative of the likelihood that the assembly of a set of focus genes in a network could be explained by random chance alone.

PCA. Only the 4,417 genes whose variance of expression across all samples from the ten cell types was at least within the 80th percentile of variance were considered for PCA (by MATLAB sofware). Expression normalized by the robust multi-array average method and $\log_2$-transformed was used.

Generation of the core macrophage signature. A macrophage core signature was generated by comparison of gene expression in brain, lung, peritoneal and red-pulp macrophages to that in populations deemed not to be macrophages but authentic DCs. These DCs included CD103^{+} DCs from lung and liver, CD8^{+} DCs from spleen and thymus, CD4^{+}CD11b^{+} DCs from spleen, CD4^{-}CD8^{-}CD11b^{+} DCs from spleen, and all DC populations (resident and migratory) isolated from skin-draining lymph nodes. A first list of possible genes that define macrophages was generated with the median value in the group of macrophages or DCs for each probe set to generate a list of probe sets with median expression that was twofold or more higher in the group of macrophages with a statistical significance of $P < 0.05$ (Student's *t*-test) and a coefficient of variation of <0.5. Then that list of probe sets was filtered for the removal of any probe sets that did not have a normalized intensity value of ≥120 (the threshold for positive expression) in at least two macrophage populations. From the remaining probe sets, we compared the mean expression values of each macrophage and DC population, filtering to identify the lowest mean value in any single macrophage population relative to the highest mean value in any of the DCs. The probe sets with expression at least twofold higher in macrophage with the lowest expression compared with the DC with the highest expression composed the core genes retained (**Table 1**, far left). To account for genes observed in only some macrophage populations but still not expressed in DCs, we also generated lists in which one or two macrophage populations were allowed to be excluded from consideration, but the criteria for comparing the remaining macrophages to the DCs was otherwise as described.

Generation of gene modules and prediction of module regulators. The gene modules, Ontogenet algorithm and regulatory programs have been described (V. Jojic *et al.*, unpublished data; methods described in **Supplementary Note 1**). The normalization of expression data was done as part of the ImmGen Project pipeline, March 2011 release. Data were $\log_2$-transformed. For genes presented on the array with more than one probe set, only the probe set with the highest mean expression was retained. Of those, only the 7,996 probe sets with an s.d. value above 0.5 across the entire data set were used for the clustering. Clustering was done by Super Paramagnetic Clustering[38] with default parameters, which resulted in 80 stable coarse modules of coexpressed genes. Each coarse module was further clustered by hierarchical clustering into more fine modules, which resulted in 334 fine modules.

The Ontogenet algorithm was developed for the ImmGen data set (V. Jojic *et al.*, data not shown, and **Supplementary Note 1**). Ontogenet finds a regulatory program for each coarse and fine module on the basis of regulator expression and the structure of the lineage tree. The regulatory program uses a form of regularized linear regression in which each cell type can have its own regulatory program, but regulatory programs of related cells are 'encouraged' to be similar. This allows switching in the regulatory program but still allows robust fitting given the available data. Fopr visualization of the expression of a module on the lineage tree, the expression of each gene was standardized by subtraction of the mean and division by its s.d. across all data set. Results for replicates were averaged. The mean expression of each module was projected on the tree. Expression values are color coded from minimal (blue) to maximal (red).

Association between the macrophage core signature and gene modules. A hypergeometric test for two groups was used to estimate the enrichment of all ImmGen fine modules for the 11 gene signatures in **Tables 1** and **2**. A Benjamini-Hochberg false-discovery rate of 0.05 or less was applied to the *P* value table of all 11 signatures across all 334 fine modules.

Antibodies used for confirmation studies. Antibody to (anti-) mouse CD11c (N418), anti-CD11b (M1/70), anti-CD24a (30-F1), anti-CD45

© 2012 Nature America, Inc. All rights reserved.

(30-F11), anti-CD14 (Sa2-8), anti-MHCII (M5/114.15.2), anti-F4/80 (BM8), anti-CD8a (53-6.7), anti-CD103 (2E7), anti-CD11a (M17/4), anti-EpCAM (G8.8), anti-VCAM-1 (429), anti-CD31 (390), anti-CD93 (AA4.1), anti-ICAM-2 (3C4 mIC2/4) anti-CD68 (FA-11), and isotype-matched control monoclonal antibodies were from eBioscience. Antibody to mouse MerTK (BAF591) was from R&D Systems. Antibody to mouse FcγRI (X54-5/7.1) and anti-Siglec-F (E50-2440) were from BD Biosciences. Antibody to mouse MR1 has been described[21]. Anti-Siglec-H was a gift from M. Colonna.

35. Nahrendorf, M. *et al.* The healing myocardium sequentially mobilizes two monocyte subsets with divergent and complementary functions. *J. Exp. Med.* **204**, 3037–3047 (2007).
36. Idoyaga, J., Suda, N., Suda, K., Park, C.G. & Steinman, R.M. Antibody to Langerin/CD207 localizes large numbers of CD8α$^+$ dendritic cells to the marginal zone of mouse spleen. *Proc. Natl. Acad. Sci. USA* **106**, 1524–1529 (2009).
37. Narayan, K. *et al.* Intrathymic programming of effector fates in three molecularly distinct gammadelta T cell subtypes. *Nat. Immunol.* **13**, 511–518 (2012).
38. Blatt, M., Wiseman, S. & Domany, E. Superparamagnetic clustering of data. *Phys. Rev. Lett.* **76**, 3251–3254 (1996).

© 2012 Nature America, Inc. All rights reserved.

doi:10.1038/ni.2419