

# Comparative promoter analysis in vivo: identification of a dendritic cell-specific promoter module

Stephanie L. Edelmann,<sup>1</sup> \*Peter J. Nelson,<sup>2</sup> and \*Thomas Brocker<sup>1</sup>

<sup>1</sup>Institute for Immunology, Ludwig-Maximilians-University, Munich, Germany; and <sup>2</sup>Medizinische Poliklinik, Ludwig-Maximilians-University, Munich, Germany

**Dendritic cells (DCs) are important immune cells. This study focused on transcriptional networks active in murine DCs, but DCs are difficult to study using conventional molecular techniques. Therefore, comparative promoter analysis was used to identify evolutionarily conserved features between the murine CD11c and DC-STAMP promoters. A promoter framework consisting of 4 transcription factor binding sites was identified that included signal transducer and activator of tran-**

**scription, homeodomain transcription factors, and 2 members of the Brn POU domain factors family. This promoter module was functionally verified by in vivo promoter analysis and site-directed mutagenesis. Hematopoietic stem cells were engineered by lentiviral vectors and expression of green fluorescent protein reporter was monitored in primary hematopoietic cell types that develop without further manipulation in irradiated recipient mice. The verified promoter module**

**was then modeled and used in a bioinformatics-based search for other potential coregulated genes in murine DCs. A promoter database search identified 2 additional genes, *Ppof2* and *Ppof1*, which have a similar promoter organization and are preferentially expressed in murine DCs. The results define a regulatory network linked to development of murine DCs. (*Blood*. 2011;118(11):e40-e49)**

## Introduction

Dendritic cells (DCs) play key roles in shaping the immune response.<sup>1,2</sup> They are highly specialized in uptake and presentation of antigen. In their steady state, DCs migrate from peripheral tissues to the draining lymph node and present self-antigen to naive T cells, inducing tolerance. On encounter of pathogen, DCs change their phenotype (undergo maturation), which leads to a more pronounced migratory behavior and an increase in stimulatory capacity.<sup>3</sup>

Plasticity is a pervasive feature of DC biology. DC subsets from different tissues show differential morphology, phenotypes, and functions. An additional distinction can also be made between conventional DCs (cDCs) and plasmacytoid DCs (pDCs), which are mainly involved in viral immune responses. cDCs also include a subset (CD8<sup>+</sup> in mice) specialized in antigen cross-presentation.<sup>4</sup>

Networks of genes linked by common control mechanisms may help orchestrate homeostasis, as well as the response of DCs to external stimuli. The prediction and analysis of these networks represent a major challenge in understanding immunology. Transcriptional mechanisms have been very difficult to study in DCs because of their propensity to quickly change phenotypes in response to physical manipulation and their limited numbers. Gene expression is controlled by a series of processes that include locus control regions, chromatin rearrangement, methylation, gene enhancers, etc. These regulatory events ultimately end at the promoter level to effect gene transcription. Promoters integrate this information to achieve the tissue and signal specific control of gene expression required for complex biologic processes, such as DC biology and function.

Within promoters, transcription factors (TF) binding to specific sites (TFBSs) drive gene transcription. In concert with DNA polymerase and associated proteins, the TFs form a unique three-dimensional protein complex on the promoter. This structure, or initiation complex, restricts the relative order, orientation, and spacing of TFBSs. Because of this, sequences important for TF binding and transcriptional control tend to be conserved over evolution.<sup>5</sup> Importantly, evolutionary convergence often leads to the conservation in the organization in both orientation and distance of TFBSs within the promoters of genes whose products must functionally interact in the same biologic context.<sup>6,7</sup> A recognizable shared common context of transcription factor binding sites between coregulated or orthologous promoters helps ensure the transcriptional synchronization of those genes that are linked to the development of immune effector function.

In the present study, we sought to characterize transcriptional networks active in murine DCs. To this end, a bioinformatics based comparison of CD11c and DC-STAMP promoters led to the identification of an evolutionarily conserved promoter framework consisting of 4 TFBSs. The importance of this putative transcriptional control region was verified by in vivo promoter analysis using sequential deletion constructs of the mouse CD11c promoter. Analysis with a minimal promoter/enhancer-trap system and site-directed mutagenesis of TFBSs further supported the functional significance of the identified structure. A promoter model based on these elements was then used in a proactive promoter database search. This led to the identification of 2 novel genes with a similar promoter organization, which are preferentially expressed

Submitted March 11, 2011; accepted May 27, 2011. Prepublished online as *Blood* First Edition paper, June 9, 2011; DOI 10.1182/blood-2011-03-342261.

\*P.J.N. and T.B. contributed equally to this study.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2011 by The American Society of Hematology

in DCs, and thus may be linked in a regulatory network underlying the functional maturation of DCs.

## Methods

### Mice

C57BL/6 mice were maintained and bred in the animal facility of the Institute for Immunology. They were used at 6-12 weeks of age in accordance with the guidelines of the local ethical committee, and all animal experiments were approved by the animal committee of the state of Bavaria.

### Computational promoter analysis

Sequence information of promoter regions was retrieved from EIDorado/Gene2Promoter software (Genomatix Software GmbH). Without further experimental information to integrate, a promoter is generally assumed as 500 bp upstream and 100 bp downstream of the transcriptional start site, but for some analyses we enlarged the regions further upstream to -1400 or -1900 bp. The location of transcriptional start sites is depicted according to EIDorado and is derived from CAGE tags and individual cDNAs. For definition of TFBSs, the Matrix family library (Version 8.2; Genomatix) was used. Thereby, various data defining the binding sequence are used to calculate a weight matrix to avoid redundancy. Based on similarities in binding pattern and functionality, the matrices are clustered into families.<sup>8</sup> The version of the vertebrate library we used contained 727 matrices grouped in 170 families. Using Frameworker 5.5 software (Genomatix Software GmbH), different promoter regions were compared and screened for a set of TFs with a defined order and orientation, a so-called "framework." Default settings were used except for the maximum distance variance between 2 elements, which was increased from 10-20 bp (CD11c mouse and human with DC-STAMP mouse) or 30 bp (CD11c rat). In addition, the search sometimes had to be limited to the respective TF families (CD11c rat and DC-STAMP rat). The CD11c/DC-STAMP promoter model was refined by adjusting the distance ranges to the values from the CD11c rat search result (FasM; Genomatix). Subsequently, the model was used to perform a mouse promoter database search with default settings (ModelInspector; Genomatix).

### Lentiviral vectors

The lentiviral backbone is based on the FUGW<sup>9</sup> vector, which was always digested with Pac I and Age I to remove the original Ubiquitin C promoter. The mouse CD11c promoter sizes 2000 bp and 1500 bp correspond to the restriction sites *Hind*III and *SexA* I within the 5-kb promoter, respectively. Initially, the 1000-bp fragment was amplified by polymerase chain reaction (PCR) with 5'-ATTTGCGGCCGCTAGCACCCAGTCTTTGCT-3' and 5'-TCGCGACTGCAGCCCACTGGAGAA-3' primers and cloned into a different vector with *Not* I and *Nru* I. From there, it was isolated with *Not* I, the end was filled up by Klenow enzyme, and with Age I. This fragment was ligated to the FUGW, which has been treated with Pac I, Klenow enzyme, and Age I. To generate the smaller mouse CD11c promoter-green fluorescent protein (GFP) constructs, different promoter fragments were amplified using standard PCR. The following primers (MWG Biotechnology) were used: 900 5'-CCTTAATTAACAATGCTTACCCACCCCTC-3', 750 5'-CCTTAATTAACAGTTTTAGTATTCTTGTACCTTGG-3', 500 5'-GCTATTAATTAATATGTTGAGCAAATGACTAAT-3', 400 5'-GCTATTAATTAATGCTTACTTCTTAGTCTACTTCCA-3', and the same reverse primer 5'-GCATACCGTGCAGCTGGAGAACAGAAGCA-3' for all of the constructs. The minimal SV40 promoter was amplified from the pGL3-Promoter vector (Promega) using 5'-CCTTAATTAAGCGATC-TGCATCTCAATT-3' and 5'-GCATACCGGTGCCAAGCTTTTTG-CAAAAGC-3'. For the combination of CD11c promoter fragments with the minimal SV40 promoter, the fragments were amplified by PCR (CD11c<sup>500-400bp</sup> 5'-CCGCTCGAGCGGTATGTTGAGCAAATGAC-3' and 5'-GGAAGATCTTCTGATCCATGTAGGAGC-3', CD11c<sup>750-574bp</sup> 5'-CTAGCTAGCATTGCTTCTGAAATTCAG-3' and 5'-GAAGATCTGAG-TAAAAGCAGATGG-3', CD11c<sup>750-400bp</sup> 5'-CTAGCTAGCATTGCTTCT-GAAATTCAG-3' and 5'-GGAAGATCTTCTGATCCATGTAGGAGC-

3') and cloned first into the pGL3-Promoter vector using the following restriction enzymes: *Xho*I and *Bgl* II (CD11c<sup>500-400bp</sup>) or *Nhe* I and *Bgl* II (CD11c<sup>750-574bp</sup>, CD11c<sup>750-400bp</sup>). In the next step, the CD11c fragments were isolated together with the minimal SV40 promoter: *Sma* I and *Hind*III for CD11c<sup>500-400bp</sup> or *Nhe* I and *Hind*III for CD11c<sup>750-574bp</sup> and CD11c<sup>750-400bp</sup>. All these fragments were completely blunt-ended and ligated into a blunt-ended FUGW. Specific point mutations were introduced into the CD11c<sup>1000bp</sup> promoter fragment using the GeneTailor Site-Directed Mutagenesis System (Invitrogen) and the primers indicated in supplemental Table 2 (available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article). The point mutations were designed by the SequenceShaper program (Genomatix) with settings to avoid the generation of other TFBSs. All CD11c-GFP constructs were validated by sequencing (Sequiserve).

### Generation and titration of lentiviral stocks

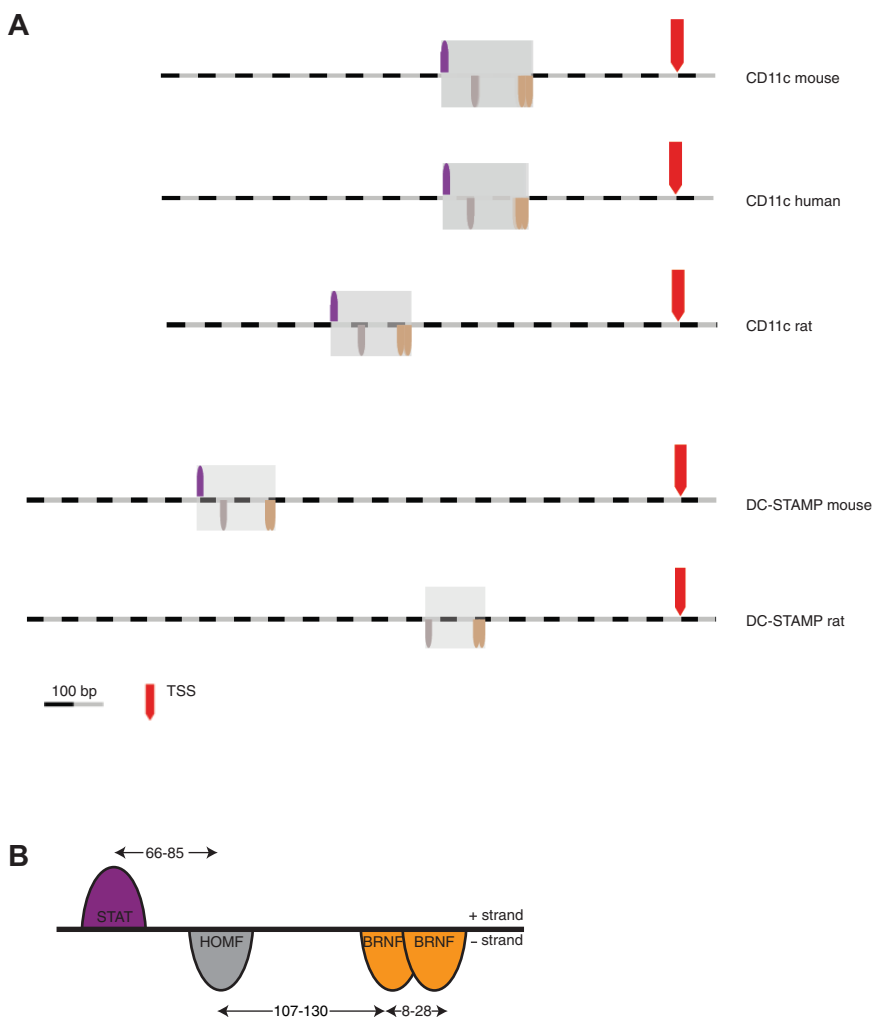
For virus production, 293FT cells (Invitrogen) were transfected using the standard calcium phosphate method with 20 μg vector-DNA, 15 μg of pCMVΔR8.2 and 10 μg of pMD2G (VSV-G). Supernatants were collected for 3 consecutive days starting one day after transfection, filtered (0.45-μm filter; Nalgene Nunc), and concentrated using Centricon filter devices (Plus-70; Millipore). Aliquots were snap-frozen and stored at -80°C. To determine viral titers, NIH3T3 cells were transduced with various virus dilutions using spin infection (300g, 2 hours at 32°C) in the presence of polybrene (8 μg/mL; Sigma-Aldrich). After 4 hours at 37°C, the virus was removed and the cells were incubated for 2 more days. DNA was then isolated (DNeasy Blood & Tissue Kit; QIAGEN) and analyzed by quantitative PCR (Lightcycler FastStart DNA Master<sup>PLUS</sup> SYBR Green, Roche Diagnostics) on a Lightcycler Carousel-based system (Roche Diagnostics). Viral integration (forward 5'-TGAAAGCG-AAAGGGAACCA-3' and reverse 5'-CCGTGCGCGCTTCAG-3') was shown per cell (BDNF forward 5'-ACGACATCACTGGCTGACAC-3' and BDNF reverse 5'-CATAGACATGTTTGGCGGCATC-3'). Standard curves were generated with serial dilutions of plasmids containing the relevant template DNA and absolute quantification was used to calculate the viral titers.

### Generation of bone marrow chimeras

Recipient mice were lethally irradiated with 2 separate doses (2 × 550 cGy) using a Cesium source (Gammacell 40, AECL) and supplied with neomycin (1.2 g/L; Sigma-Aldrich) containing drinking water for 5 weeks. Chimeras were analyzed 8-10 weeks after bone marrow transfer. Donor mice were injected intravenously with 5-fluorouracil (150 mg/kg body weight; Invivo-gen) and after 4 days bone marrow was isolated. This stem cell-enriched bone marrow was depleted of erythrocytes (Mouse Erythrocyte Lysing Kit; R&D Systems) and cultured in serum-free medium (stemline II hematopoietic stem cell expansion medium; Sigma-Aldrich) with 1% penicillin/streptomycin (Invitrogen). The cells were stimulated with a cytokine mixture (Miltenyi Biotec) containing murine IL-3 (10 ng/mL), murine stem cell factor (50 ng/mL), and human IL-6 (50 ng/mL). At day 3 of culture, cells were spin-infected (300g, 2 hours at 32°C) with cell-free stocks of lentivirus (multiplicity of infection between 0.2 and 0.5) in the presence of protamine sulfate (4 μg/mL; Sigma-Aldrich). After 4 more hours, incubation at 37°C the virus was removed and 1 to 3 × 10<sup>6</sup> cells per recipient mouse were injected intravenously the next day.

### Flow cytometry

Single-cell suspensions of organs were prepared by enzymatic digestion with Liberase CI (0.42 mg/mL) and DNase I (0.2 mg/mL), both from Roche Diagnostics) for 25 minutes at 37°C followed by mechanical dispersion. Cells were stained for 20 minutes at 4°C with monoclonal antibodies (from BD Biosciences or eBioscience, unless stated otherwise) against the following antigens: CD3 (145-2C11), CD4 (GK1.5), CD8α (53-6.7), CD11b (M1/70), CD11c (HL3), CD19 (1D3), CD86 (GL1), IA<sup>b</sup> (AF6-120.1), NK1.1 (PK136), B220 (RA3-6B2), Gr-1 (RB6-8C5), F4/80 (BM8), and neutrophils (7/4) from AbD Serotec. Polyclonal goat anti-mouse antibodies against κ- and λ-chain were purchased from



**Figure 1. CD11c and DC-STAMP share a common promoter framework.** (A) Identification of a regulatory framework by computational comparison of the CD11c promoter with the DC-STAMP promoter across species. TSS indicates transcription start site. (B) Detailed representation of the CD11c/DC-STAMP promoter framework, which contains 4 elements over a length of 250 bp. The transcription factor families and their respective orientation on the + and - strand are displayed. The indicated distance ranges result from the refinement of the framework in the rat CD11c promoter.

Southern Biotechnology. Flow cytometry was performed on a FACSCalibur or FACSCanto II instrument (BD Biosciences) and analyzed with FlowJo software (TreeStar).

### Quantitative RT-PCR

DCs were generated from bone marrow according to the protocol of Inaba et al,<sup>10</sup> and aliquots of  $2 \times 10^6$  cells were taken at days 0, 4, 7, and 8. The cells were snap frozen in liquid nitrogen and stored at  $-80^\circ\text{C}$ . Total RNA was isolated using the RNeasy Mini Kit (QIAGEN), and cDNA synthesis was done using the SSIII First-Strand Synthesis SuperMix for quantitative RT-PCR (Invitrogen). The TaqMan Assay was performed using the LightCycler TaqMan Master Kit (Roche Diagnostics) and the Universal ProbeLibrary Set mouse (Roche Diagnostics) on a CFX96 Real Time System (Bio-Rad) using the primers and probes listed in supplemental Table 3. Expression levels were normalized to ubiquitin C and relative quantification was calculated using the  $\Delta\Delta\text{C}_T$  method.

## Results

### Comparative promoter analysis of CD11c and DC-STAMP promoters across species-identified conserved promoter structures

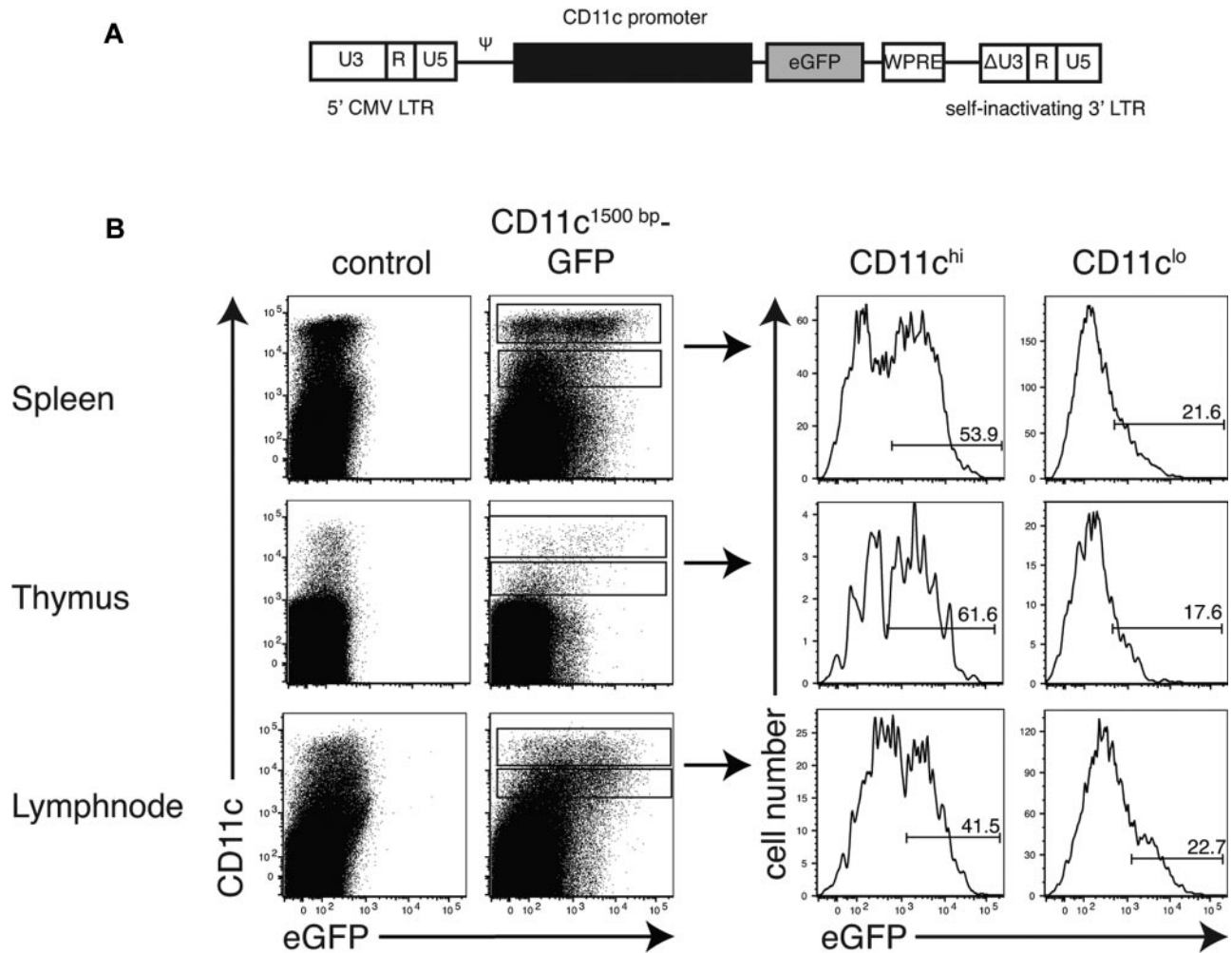
Ordered sets of TFBSs conserved in distance and orientation are referred to as a promoter module or promoter framework.<sup>6,7</sup> A promoter module is defined as a functionally characterized frame-

work, whereas a promoter framework describes a conserved set of transcription factor binding sites identified between promoters and thus a putative module that can be tested for functionality.<sup>6,7</sup>

For computational analysis of conserved promoter frameworks linked to DC function, we chose the CD11c and DC-STAMP genes expressed in mouse DCs.<sup>11,12</sup> The immediate upstream regions of these 2 genes were compared using bioinformatics. To this end, 1500 bp ( $-1400/+100$  bp) of the mouse and human CD11c promoters, and 2000 bp ( $-1900/+100$  bp) of the mouse DC-STAMP promoter were analyzed using the FrameWorker analysis program (Genomatix). Two highly similar promoter frameworks were identified as shared between the 2 promoters comprising 4 TFBSs that differ only in position 3. In the next step, orthologous promoters in the rat were screened to determine whether the frameworks were also conserved over evolution. One of the frameworks was found to be better conserved in the rat DC-STAMP promoter. We chose to focus on this framework, shown in Figure 1A, for the subsequent verification studies.

The resultant framework spans  $\sim 250$  bp and contains binding sites for: signal transducer and activator of transcription (STAT), homeodomain transcription factors (HOMF), and 2 members of the Brn POU domain factors (BRNF) family (Figure 1B).

The regulatory region identified corresponds to one of the few regions conserved between mouse and human CD11c (data not



**Figure 2. The pCD11c<sup>1500bp</sup> drives expression selectively in DCs.** (A) Schematic representation of the lentiviral-based SIN-vector containing a fragment of the mouse CD11c promoter to control expression of enhanced GFP (eGFP). CMV indicates cytomegalovirus;  $\Psi$ , packaging signal; SIN, self-inactivating; LTR, long terminal repeat; WPRE, woodchuck hepatitis virus post-transcriptional regulatory element; and  $\Delta$ U3, deletion in the U3 region. (B) Analysis of the CD11c<sup>1500bp</sup> promoter activity in different organs. HSCs from C57BL/6 mice were modified by transduction with a lentiviral vector presented in panel A and injected into lethally irradiated C57BL/6 recipient mice. After 8 weeks, single-cell suspensions of spleen, thymus, and lymph nodes (mandibular, axillary, and subiliac) from these chimeras were stained for CD11c and MHC-II, and GFP expression was analyzed by flow cytometry. For detailed quantification, we gated on CD11c<sup>hi</sup> and CD11c<sup>lo</sup> cells and histograms were shown. Data are representative of 2 independently performed experiments with 4 or 5 mice per group.

shown, DiAlignTF, Genomatix) and overlaps with a region that has been proposed to be a potential transcriptional enhancer in the human CD11c promoter.<sup>13,14</sup>

#### In vivo evaluation of candidate promoter sequences using transgenic engineering of hematopoietic stem cells

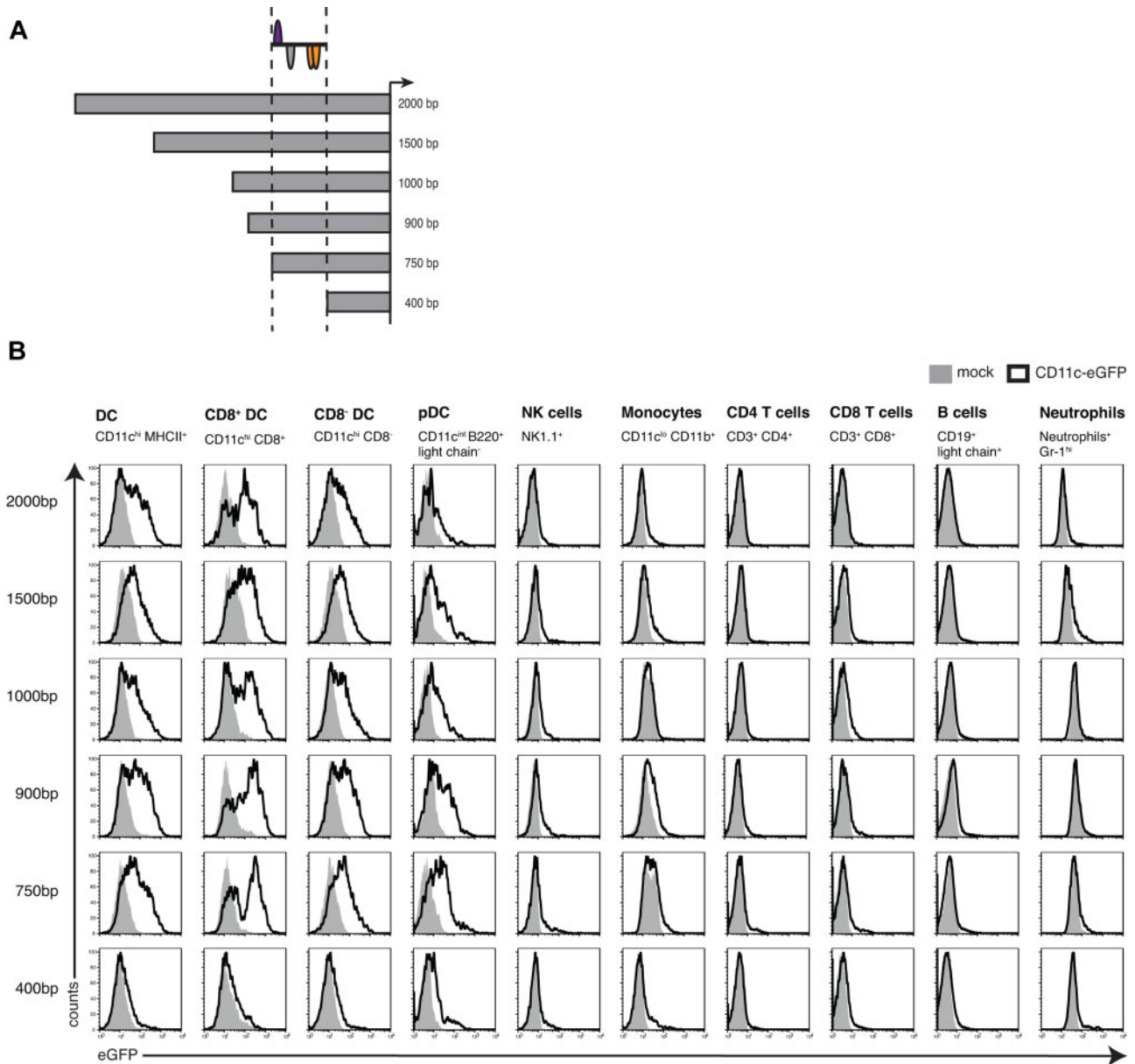
Because of their sensitivity to physical manipulation, and restricted cell numbers, DCs are difficult or impossible to study using “conventional” molecular biology approaches, such as transient transfection or electrophoretic mobility shift assays. To validate the promoter framework experimentally, the transcriptional activity of the mouse CD11c promoter was analyzed in a physiologic setting. To this end, fragments of the CD11c promoter (pCD11c) region were cloned into a self-inactivating lentiviral vector upstream of a GFP reporter gene (Figure 2A). The resulting viruses were then used to genetically modify hematopoietic stem cells (HSCs) of C57BL/6 mice, which were subsequently injected into irradiated recipient mice. In the bone marrow chimeras, promoter activity was examined in organs and primary leukocyte cell types on reconstitution of the hematopoietic system. We have previously

shown that DCs resulting from lentiviral engineered HSCs retain DC phenotypes and functions.<sup>14,15</sup>

Using a CD11c promoter fragment of  $-1500$  bp (pCD11c<sup>1500bp</sup>), GFP expression could be detected preferentially in DC populations from spleen, thymus, and lymph nodes (Figure 2B). To avoid multiple insertions per cell, low viral titers were used (multiplicity of infection between 0.2 and 0.5) for transduction of HSCs. At this low viral titer,  $\sim 50\%$  of all CD11c<sup>+</sup> cells did express GFP. The intensity of the GFP signal seen generally corresponded to the surface expression of CD11c, as CD11c<sup>low</sup> cells showed also less GFP expression (Figure 2B right). This demonstrated that the promoter fragment tested included the necessary regulatory information to allowing tissue-specific expression of CD11c.

#### Sequential deletion analysis of the mouse CD11c promoter and evaluation in vivo was used to identify a minimal optimal promoter sequence

To characterize the sequences required for the functional CD11c promoter region, the 5' end of the promoter region was sequentially truncated and the resulting constructs (Figure 3A) were analyzed



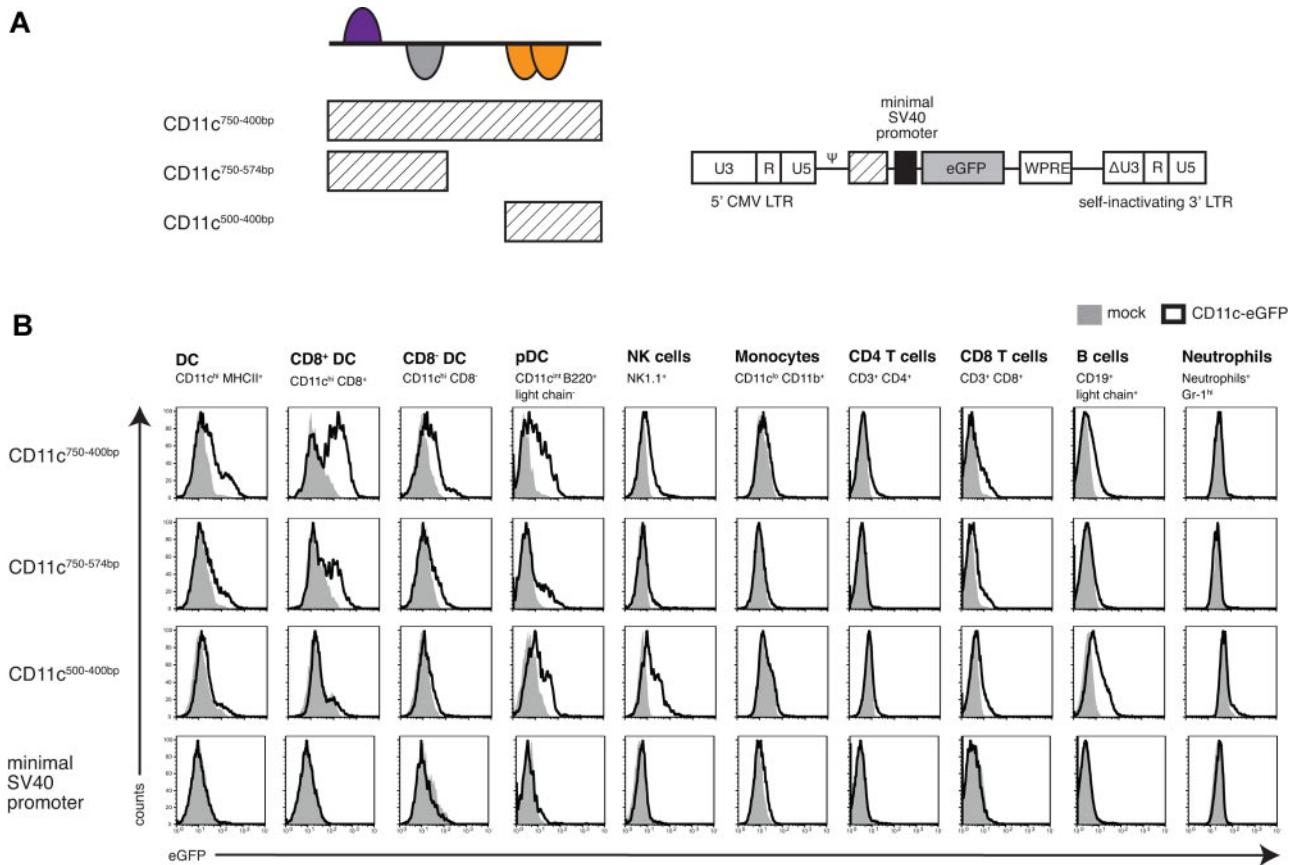
**Figure 3. Sequential deletion analysis of the CD11c promoter in vivo.** (A) Schematic illustration of pCD11c deletion constructs with the respective localization of the CD11c/DC-STAMP promoter model. (B) Promoter activity of different sizes of the mouse CD11c promoter measured by flow cytometry. Spleen cells from bone marrow chimeras generated with HSCs transduced with lentiviral constructs carrying the indicated pCD11c segment were analyzed for GFP expression. The different cell types were identified based on the expression of the indicated markers. Overlays were generated after gating on the relevant population using cells from C57BL/6 mice as negative controls. For every promoter construct, the data shown are representative of 2 independently performed experiments with 4 or 5 mice per group.

for their potential to drive gene expression in cell types of the immune system (detailed gating strategy in supplemental Figure 1). The series of CD11c promoter deletions between  $-2000$  and  $-750$  bp were found to be comparable in their gene expression profiles (Figure 3B). GFP expression could be observed in all DC populations: the CD8<sup>+</sup> DC subset produced the highest levels of the transgene compared with CD8<sup>-</sup> DCs or pDCs. The latter were lowest for GFP, but this is in accordance with their intermediate expression of CD11c. In contrast, all other cell types tested expressed only negligible amounts of GFP. In T and B cells, expression was absent, and only very low promoter activity was detected in monocytes, neutrophils, and NK cells (for statistical analysis, see supplemental Figure 2). Strikingly, expression in cDCs was lost completely in the pCD11c<sup>400bp</sup> fragment, whereas

some residual activity in pDCs remained. The 350-bp region located between positions  $-750$  and  $-400$  bp includes the CD11c/DC-STAMP-promoter framework and represents an important functional region driving CD11c gene transcription.

#### Analysis of the conserved CD11c/DC-STAMP promoter framework

To further characterize the transcriptional potential of the CD11c/DC-STAMP-framework, we next tested whether the relevant region was still capable of driving DC-specific expression when taken out of its original genomic context. To this end, the pCD11c<sup>750-400bp</sup> region was subcloned into a lentiviral backbone upstream of the SV40 minimal promoter using GFP as the reporter gene (Figure 4A). The SV40 minimal promoter alone did not



**Figure 4. Segmental analysis of the newly defined promoter model.** (A) Schematic illustration of the different parts of the CD11c/DC-STAMP promoter model and their respective location. pCD11c<sup>750-400bp</sup>, pCD11c<sup>750-574bp</sup>, and pCD11c<sup>500-400bp</sup> were cloned upstream of a SV40 minimal promoter into a lentiviral vector. (B) In vivo promoter analysis of the 3 pCD11c fragments indicated and the minimal SV40 promoter alone. Spleen cells from bone marrow chimeras were analyzed for GFP expression by flow cytometry. Histogram overlays of cells from C57BL/6 mice as negative controls are shown. Data are representative of 3 independently performed experiments (n = 4 or 5).

show any *in vivo* promoter activity, but the presence of the pCD11c<sup>750-400bp</sup> region resulted in GFP expression preferentially in DCs (Figure 4B), which was comparable with that seen in the longer CD11c promoter fragments (Figure 3, 750-2000 bp). This demonstrates that the 350-bp region has characteristics of a DC-specific enhancer.

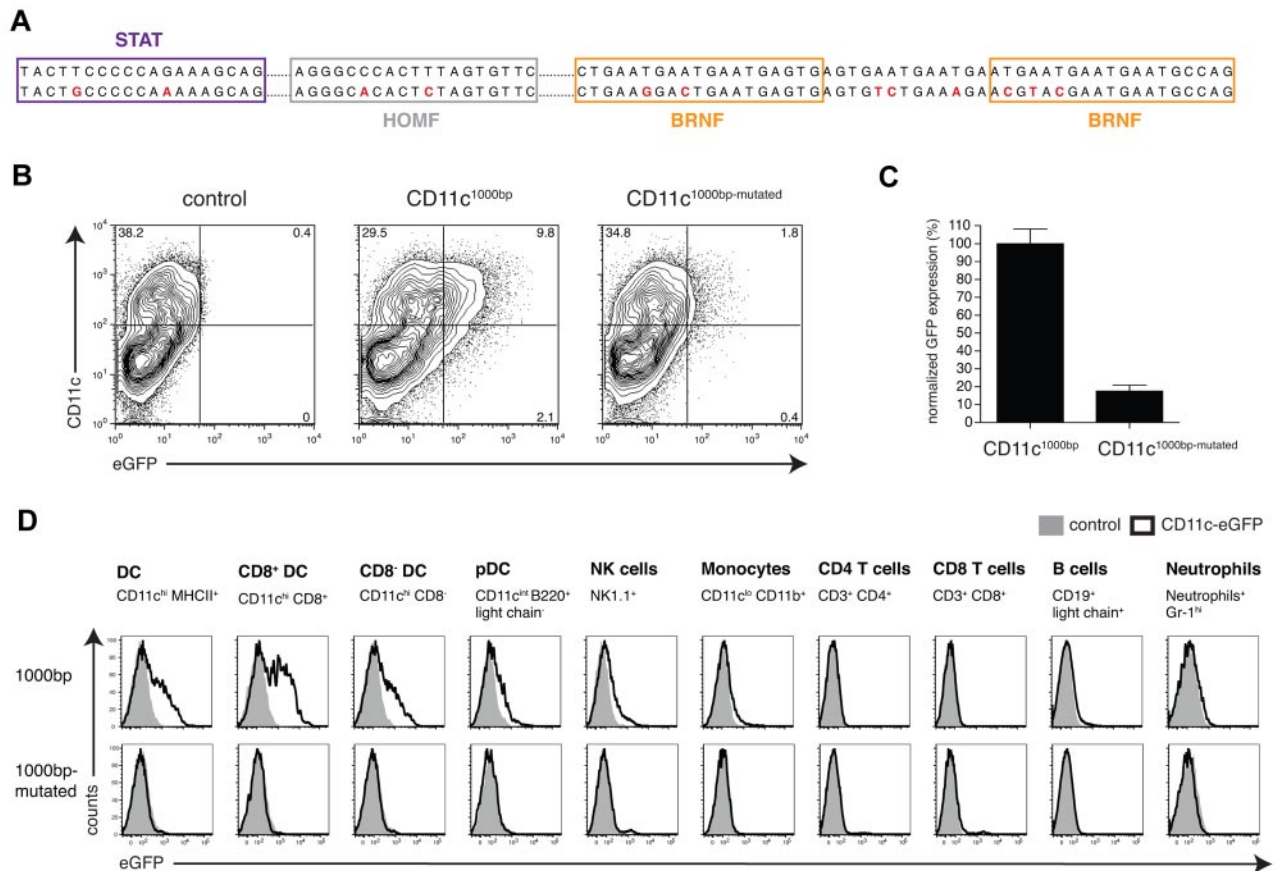
The complete framework was required for optimal tissue-specific expression. When the region was separated into 2 parts, where each half contained 2 of the 4 TFBSs (Figure 4A), the subregions did not function efficiently in reporter-gene assays. Compared with the complete promoter framework, the more 5' sequence (pCD11c<sup>750-574bp</sup>) was capable of inducing limited reporter gene expression in DCs (Figure 4B). In contrast, the more 3' region (pCD11c<sup>500-400bp</sup>) could not drive any expression in cDCs but showed some activity in other cell types, such as pDCs, NK cells, or B cells (Figure 4B). The results suggest that the region containing the binding sites for the STAT family and the homeodomain TFs is of central importance. However, optimal promoter activity is seen only if the whole region is used, allowing all elements to work together.

To further verify the importance of the 4 TFBSs, the individual binding sites in the CD11c/DC-STAMP promoter framework were selectively mutated in the context of the pCD11c<sup>1000bp</sup> promoter construct (Figure 5A). The SequenceShaper tool (Genomatix) was used to design the point mutations. The program ensures that the binding site will be eliminated and that

no new bindings will be generated as a result of the mutation. Elimination of the 4 potential binding sites in the context of the pCD11c<sup>1000bp</sup> expression vector required the mutation of 12 nucleotides of the -1000-bp promoter region (Figure 5A). These mutations resulted in a loss of promoter activity in DCs both *in vitro* (Figure 5B-C) and *in vivo* (Figure 5D).

#### Identification of coregulated genes using a combinatorial approach

Promoter modules can also be used in bioinformatics-based searches to identify potential coregulated genes that share conserved promoter structures. The CD11c/DC-STAMP model consisting of 4 TFBSs with defined spacing and orientation was generated using FasM (Genomatix) and then used to search for novel genes that share the same organization of regulatory elements within their promoters. A database of 72 900 mouse promoters (Genomatix murine database) was screened using ModelInspector (Genomatix). Only 49 promoter sequences (0.07% of all sequences, supplemental Table 1) were found to contain the CD11c/DC-STAMP model. The results were then filtered using a transcriptomic database to verify their potential expression in murine DC populations. To this end, a platform derived from the "immunologic genome project" ([www.immgen.org](http://www.immgen.org)) was applied.<sup>16</sup> In this approach, gene expression data from specific cell types across the immune system were compared and the correlation of



**Figure 5. Site-directed mutagenesis of the CD11c<sup>1000bp</sup> promoter sequence.** (A) The original CD11c<sup>1000bp</sup> promoter sequence (top row) was modified by site-directed mutagenesis. The predicted TFBSs are represented by boxes, and the mutated bases are highlighted in red (bottom row). Because of the repetitive sequence structure in the BRNF binding region, additional mutations had to be introduced. (B) Flow cytometric analysis of promoter activity in DCs generated in vitro. At day 1 or 2 of the granulocyte-macrophage colony-stimulating factor culture, bone marrow cells were transduced with lentiviral vectors (CD11c<sup>1000bp</sup>-GFP and CD11c<sup>1000bp-mutated</sup>-GFP) at a multiplicity of infection of 1, and GFP expression was analyzed at day 7 of the culture. The control was only treated with polybrene. (C) Results from 2 independently performed experiments are normalized (GFP expression of the CD11c<sup>1000bp</sup> construct set to 100%). (D) In vivo comparison of the CD11c<sup>1000bp</sup> and the CD11c<sup>1000bp-mutated</sup> promoters by flow cytometric analysis of splenocytes from bone marrow chimeras, which were generated and analyzed analogously to Figure 3B. Data are representative of 2 independently performed experiments with 3 mice per group.

gene expression was calculated by principal component analysis. Using this database, 2 genes derived from the promoter search were specifically linked to CD11c: *Pfkl* and *Ppef2*. As detailed in the “immunologic genome project,” *Pfkl* expression is strongly correlated with CD11c expression, whereas *Ppef2* shows correlation to *Pfkl* (supplemental Figures 3 and 4). In addition, the CD11c/DC-STAMP-model was partially conserved within the *Pfkl* promoter regions across species (Figure 6A): 3 of 4 elements are present in the rat promoter. In the human ortholog, the 2 elements determined to be of highest importance for promoter activity (STAT and HOMF, Figure 4B) were conserved (Figure 6A). Interestingly, the CD11c/DC-STAMP-model is rotated by 180 degrees in the *Ppef2* promoter (Figure 6B), a phenomenon commonly associated with promoter modules.<sup>6,7</sup>

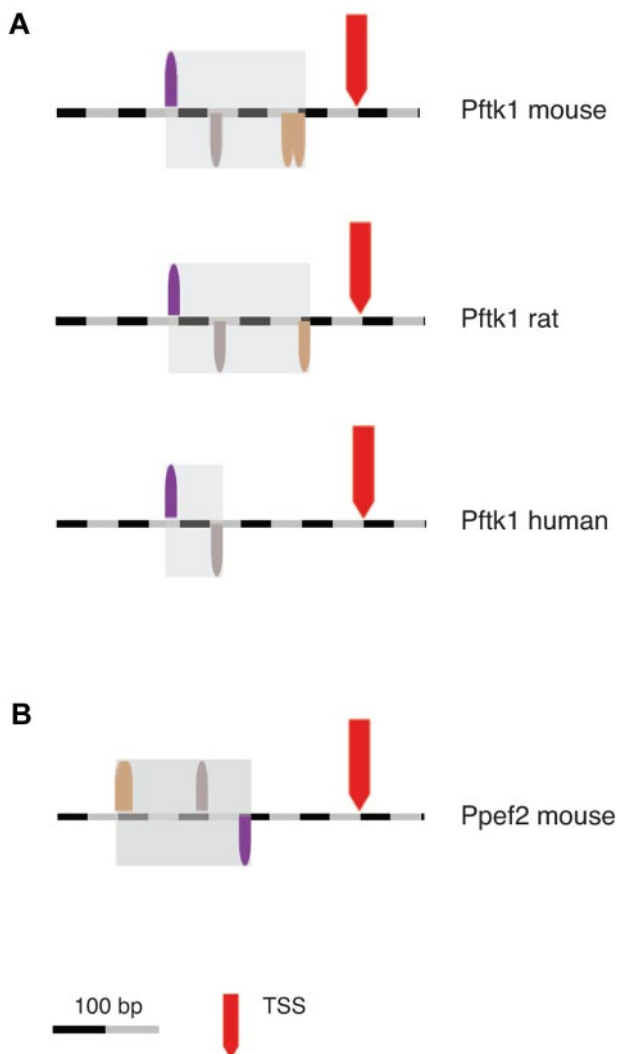
To verify the expression of the 2 genes in DCs, mRNA isolated from sorted primary cells was analyzed by real-time PCR. Both genes were expressed in DCs, particularly in the CD8<sup>+</sup> subset, but not in CD8 T cells from spleen (Figure 7A). To further investigate the potential coregulation of these genes, samples from bone marrow precursor cells were analyzed at different time points during in vitro DC differentiation, where the percentage of DCs increases with duration of the granulocyte-macrophage colony-

stimulating factor culture. Accordingly, the expression of CD11c and DC-STAMP increased proportionally but showed down-regulation after lipopolysaccharide (LPS) stimulation.<sup>12</sup> *Ppef2* showed a gene expression pattern that paralleled those seen with CD11c and DC-STAMP. *Pfkl*, which is already expressed in bone marrow cells, remained constant during the development of DCs (Figure 7B). However, *Pfkl* mRNA levels did decrease on DC stimulation by LPS, as seen with the CD11c, DC-STAMP, and *Ppef2* genes (Figure 7B).

## Discussion

A 5.5-kb fragment containing the mouse CD11c 5' untranslated region has been previously shown to target transgene expression to DCs.<sup>17</sup> Although this approach was successfully adopted by many different laboratories<sup>18-22</sup> and the human CD11c promoter has been studied extensively,<sup>23</sup> to date there is no information about the functional regulatory elements present in the promoter of the mouse ortholog.

Comparison of CD11c and DC-STAMP-promoters using a bioinformatics approach identified a conserved promoter framework, which focused attention on a 250-bp region shared between



**Figure 6. Identification of *Ppef2* and *Pftk1* as novel candidate genes.** The ModelInspector program (Genomatix) was used to scan DNA sequences for the CD11c/DC-STAMP promoter model. First, a database of mouse promoters of annotated genes was screened. Then orthologous promoters, which were identified using a comparative genomics tool (EIDorado, Genomatix), were analyzed by the same method. Location of the CD11c/DC-STAMP promoter model in *Pftk1* promoters across species (A) and in the mouse *Ppef2* promoter model (B). In the *Ppef2* promoter, the 2 BRNF binding sites are so close together (10 bp), that they appear as one.

the 2 promoters (Figure 1). By in vivo promoter-reporter gene analysis using lentiviral vectors and hematopoietic stem cell transplantation (Figure 2), we confirmed that this region in the CD11c promoter drives DC-specific expression in cells that develop from the transplanted engineered hematopoietic stem cells (Figure 3). Recent findings of Ni et al support these results as they successfully used a 700-bp fragment of the CD11c promoter in a vaccination approach against tumor antigens.<sup>24</sup> The promoter module identified here is present within this 700-bp region. The fact that gene expression driven by pCD11c constructs between  $-2000$  and  $-750$  bp results in similar expression profiles (Figure 3) argues against other positive or negative regulatory elements within this region. Other cell types, including pDCs, monocytes, macrophages, and NK cells, display low expression of CD11c. In the studies reported here, limited promoter activity was observed in monocytes and splenic macrophages (Figure 3; and data not shown). The CD11c promoter was found to drive some transgene expression in

NK cells (Figure 3). Similar results have been shown in transgenic mice using the 5.5-kb promoter.<sup>21</sup> Importantly, this low expression level can still be enough to shape the NK phenotype.<sup>25</sup> pDCs display intermediate expression levels but are nevertheless resistant to diphtheria toxin if the corresponding receptor is expressed under the control of the CD11c promoter.<sup>26</sup> These findings emphasize that the type of application and transgene have a strong impact on the outcome of CD11c-driven gene expression.

The TF families composing the CD11c/DC-STAMP promoter module are not unique to DCs. The HOMO and BRNF families display a wide tissue distribution, whereas the STAT family shows a more restricted activity to cells of the hematopoietic system (MatBase, Genomatix). The tissue specificity seen in this setting is dictated by the higher organization of the elements.<sup>6,7</sup> The BRNF family members tend to form homodimers.<sup>27</sup> This also probably occurs in the CD11c/DC-STAMP promoter model, as 2 members of this family bind in close proximity (Figure 1B).

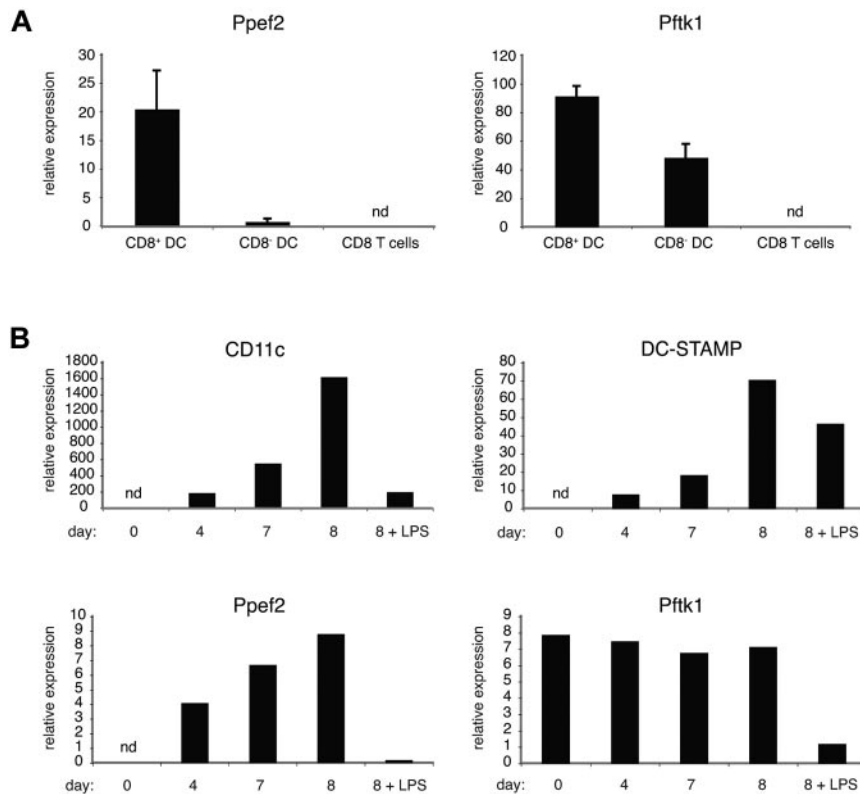
The results suggest a prominent role for STAT family members in the regulation of CD11c, as deletion resulted in a complete loss of DC-specific expression (Figures 4-5). This finding is supported by the observation that IL-4, which signals via STAT6<sup>28</sup> and is added in some culture conditions to generate DCs, enhances the expression of DC-STAMP and CD11c.<sup>12,29</sup>

After refinement of the CD11c/DC-STAMP promoter model, it was used to screen a mouse promoter database (supplemental Table 1). In combination with gene constellation analysis (Immgen), the results led to the identification of 2 novel candidate genes (Figure 6; supplemental Figure 3) *Pftk1* and *Ppef2*, which show a linked expression during DC development (Figure 7).

These results suggest that CD11c, DC-STAMP, *Pftk1*, and *Ppef2* are part of a larger promoter network that helps orchestrate gene expression during the development of murine DCs. Analysis of the potential functional significance of *Pftk1* and *Ppef2* expression in the context of DC biology is currently underway. *Ppef2* is a serine/threonine-protein phosphatase known to interact with calmodulin via its N-terminus.<sup>30</sup> *Ppef2* is also linked to the control of cytokine production in response to LPS<sup>31</sup> but has at present no known role in DC biology. *Pftk1* encodes the cyclin-dependent kinase-14, but its function is at present unclear. A potential role for *Pftk1* in cell cycle progression and cell proliferation has been suggested.<sup>32</sup> *Pftk1* is also expressed in postmitotic cells<sup>33</sup> and thus may have functions beyond the cell cycle. *Pftk1* mRNA expression does not follow the same kinetics as CD11c and DC-STAMP in DCs (Figure 7B) as it is also present in bone marrow cells (as suggested by public microarray data; GEO profiles, National Center for Biotechnology Information). Therefore, it is probable that additional regulatory elements may be active in the *Pftk1* promoter in this tissue setting (bone marrow). Interestingly, *Brn3a*, a member of the BRNF family, has been shown to play a role in transcriptional regulation of the *Pftk1* gene, although in a different tissue setting.<sup>34</sup>

Although the overall functional role of this promoter network remains to be demonstrated, analysis of the 2 initial components in DC biology (CD11c and DC-STAMP) suggests that easy definition of phenotypes may be problematic. CD11c represents one of the most important markers of murine DCs in use today, yet to date, no clear DC phenotype has been demonstrated for CD11c knockout mice.<sup>35,36</sup> DC-STAMP was originally isolated from DC cultures, but knockout mice have principally shown an osteoclast fusion phenotype<sup>37</sup> and only much later a potential DC knockout phenotype.<sup>38</sup> Overall, these findings suggest that, although a specific function for individual genes can be difficult to define in a given





**Figure 7. Expression of *Ppef2* and *Pftk1* in DCs.** (A) Gene expression profiling of ex vivo spleen cells. After density enrichment, DCs were sorted on a BD FACSAria instrument based on their expression of CD11c, MHCII, and CD8. CD8 T cells were identified as CD3<sup>+</sup>, CD8<sup>+</sup>, and CD4<sup>-</sup> cells. Gene expression levels were calculated relative to ubiquitin C. Error bars represent SD between 2 independently performed experiments. (B) Gene expression profiling of DCs generated in vitro. Real-time PCR analysis of RNA from samples isolated at days 0, 4, 7, and 8 during DC culture. At day 7, parts of the cultures were stimulated overnight with LPS (2 μg/mL). Shown is relative gene expression after normalization to ubiquitin C. Data are representative of 2 independently performed experiments. nd indicates not detectable.

biologic setting, regulatory promoter networks may have applications beyond their individual parts.

Interestingly, all of the pCD11c constructs between -750 and -2000 bp showed higher gene expression in CD8<sup>+</sup> DCs (Figure 3). This is also found in transgenic mice using the 5.5-kb promoter.<sup>22</sup> The promoter module maintained this characteristic when used to drive the SV40 minimal promoter (Figure 4). A similar trend was observed for DC-STAMP<sup>14</sup> and the newly identified genes *Pftk1* and *Ppef2* at the mRNA level (Figure 7A; and GEO profiles, National Center for Biotechnology Information). This argues that the promoter module described here is not only specific for DCs but also shows higher activity in CD8<sup>+</sup> DCs. As this subset is important for cross-presentation,<sup>39,40</sup> and thereby for the generation of optimal cytotoxic immune responses, this feature makes the promoter module an attractive tool for future engineering of vaccination vectors.

In conclusion, common promoter frameworks between CD11c and DC-STAMP promoters were found using comparative analysis and verified by in vivo promoter analysis, identifying a DC-specific promoter module. This information was subsequently used to detect genes previously not associated with DC biology. These hierarchical features effectively define a sub-network that orchestrates the expression of coregulated genes in specific DC development. The power of this combinatorial approach will help to face the challenge of dissecting complex transcriptional networks.

## References

- Coquerelle C, Moser M. DC subsets in positive and negative regulation of immunity. *Immunol Rev*. 2010;234(1):317-334.
- Pulendran B, Tang H, Manicassamy S. Programming dendritic cells to induce T(H)2 and tolerogenic responses. *Nat Immunol*. 2010;11(8):647-655.
- Reis e Sousa C. Dendritic cells in a mature age. *Nat Rev Immunol*. 2006;6(6):476-483.
- Shortman K, Heath WR. The CD8<sup>+</sup> dendritic cell subset. *Immunol Rev*. 2010;234(1):18-31.
- Qiu P. Recent advances in computational promoter analysis in understanding the transcriptional regulatory network. *Biochem Biophys Res Commun*. 2003;309(3):495-501.
- Werner T, Fessele S, Maier H, Nelson PJ. Computer modeling of promoter organization as a tool to study transcriptional coregulation. *FASEB J*. 2003;17(10):1228-1237.
- Cohen CD, Klingenhoff A, Boucherot A, et al. Comparative promoter analysis allows de novo identification of specialized cell junction-associated proteins. *Proc Natl Acad Sci U S A*. 2006;103(15):5682-5687.
- Cartharius K, Frech K, Grote K, et al. MatInspector

## Acknowledgments

The authors thank C. Ried for excellent assistance with the mutagenesis project, and A. Bol, W. Mertl, and the whole mouse facility team for animal care.

This work was supported by the Deutsche Forschungsgemeinschaft (SFB 456, T.B.; and SFB 571, P.J.N. and T.B.) and the European commission (FP6 EU-HEVAR, T.B.).

This manuscript includes part of the doctoral work of S.L.E.

## Authorship

Contribution: T.B. planned and supervised experiments and wrote the manuscript, P.J.N. provided assistance with bioinformatics and promoter modeling and wrote the manuscript, and S.L.E. performed experiments.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Thomas Brocker, Institute for Immunology, Ludwig-Maximilians-University, Goethestrasse 31, D-80336 Munich, Germany; e-mail: brocker@lmu.de.

- and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*. 2005;21(13):2933-2942.
9. Lois C, Hong EJ, Pease S, Brown EJ, Baltimore D. Germine transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science*. 2002;295(5556):868-872.
  10. Inaba K, Inaba M, Romani N, et al. Generation of large numbers of dendritic cells from mouse bone marrow cultures supplemented with granulocyte/macrophage colony-stimulating factor. *J Exp Med*. 1992;176(6):1693-1702.
  11. Hartgers FC, Vissers JL, Looman MW, et al. DC-STAMP, a novel multimembrane-spanning molecule preferentially expressed by dendritic cells. *Eur J Immunol*. 2000;30(12):3585-3590.
  12. Eleveld-Trancikova D, Triantis V, Moulin V, et al. The dendritic cell-derived protein DC-STAMP is highly conserved and localizes to the endoplasmic reticulum. *J Leukoc Biol*. 2005;77(3):337-343.
  13. Lopez-Cabrera M, Nueda A, Vara A, Garcia-Aguilar J, Tugores A, Corbi AL. Characterization of the p150,95 leukocyte integrin alpha subunit (CD11c) gene promoter: identification of cis-acting elements. *J Biol Chem*. 1993;268(2):1187-1193.
  14. Dresch C, Edelmann SL, Marconi P, Brocker T. Lentiviral-mediated transcriptional targeting of dendritic cells for induction of T cell tolerance in vivo. *J Immunol*. 2008;181(7):4495-4506.
  15. Edelmann SL, Marconi P, Brocker T. Peripheral T cells re-enter the thymus and interfere with central tolerance induction. *J Immunol*. 2011;186(10):5612-5619.
  16. Heng TS, Painter MW. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol*. 2008;9(10):1091-1094.
  17. Brocker T, Riedinger M, Karjalainen K. Targeted expression of major histocompatibility complex (MHC) class II molecules demonstrates that dendritic cells can induce negative but not positive selection of thymocytes in vivo. *J Exp Med*. 1997;185(3):541-550.
  18. Jung S, Unutmaz D, Wong P, et al. In vivo depletion of CD11c+ dendritic cells abrogates priming of CD8+ T cells by exogenous cell-associated antigens. *Immunity*. 2002;17(2):211-220.
  19. Probst HC, Lagnel J, Kollias G, van den Broek M. Inducible transgenic mice reveal resting dendritic cells as potent inducers of CD8+ T cell tolerance. *Immunity*. 2003;18(5):713-720.
  20. Lindquist RL, Shakhar G, Dudziak D, et al. Visualizing dendritic cell networks in vivo. *Nat Immunol*. 2004;5(12):1243-1250.
  21. Caton ML, Smith-Raska MR, Reizis B. Notch-RBP-J signaling controls the homeostasis of CD8- dendritic cells in the spleen. *J Exp Med*. 2007;204(7):1653-1664.
  22. Keller AM, Schildknecht A, Xiao Y, van den Broek M, Borst J. Expression of costimulatory ligand CD70 on steady-state dendritic cells breaks CD8+ T cell tolerance and permits effective immunity. *Immunity*. 2008;29(6):934-946.
  23. Corbi AL, Lopez-Rodriguez C. CD11c integrin gene promoter activity during myeloid differentiation. *Leuk Lymphoma* 1997;25(5-6):415-425.
  24. Ni J, Nolte B, Arnold A, Fournier P, Schirmacher V. Targeting anti-tumor DNA vaccines to dendritic cells via a short CD11c promoter sequence. *Vaccine*. 2009;27(40):5480-5487.
  25. Laouar Y, Sutterwala FS, Gorelik L, Flavell RA. Transforming growth factor-beta controls T helper type 1 cell development through regulation of natural killer cell interferon-gamma. *Nat Immunol*. 2005;6(6):600-607.
  26. Sapozhnikov A, Fischer JA, Zaft T, Krauthgamer R, Dzionek A, Jung S. Organ-dependent in vivo priming of naive CD4+, but not CD8+, T cells by plasmacytoid dendritic cells. *J Exp Med*. 2007;204(8):1923-1933.
  27. Cook AL, Sturm RA. POU domain transcription factors: BRN2 as a regulator of melanocytic growth and tumorigenesis. *Pigment Cell Melanoma Res*. 2008;21(6):611-626.
  28. Hou J, Schindler U, Henzel WJ, Ho TC, Brasseur M, McKnight SL. An interleukin-4-induced transcription factor: IL-4 Stat. *Science*. 1994;265(5179):1701-1706.
  29. Staeger H, Brauchlin A, Schoedon G, Schaffner A. Two novel genes FIND and LIND differentially expressed in deactivated and Listeria-infected human macrophages. *Immunogenetics*. 2001;53(2):105-113.
  30. Kutuzov MA, Solov'eva OV, Andreeva AV, Bennett N. Protein Ser/Thr phosphatases PPEF interact with calmodulin. *Biochem Biophys Res Commun*. 2002;293(3):1047-1052.
  31. Matsuzawa A, Saegusa K, Noguchi T, et al. ROS-dependent activation of the TRAF6-ASK1-p38 pathway is selectively required for TLR4-mediated innate immunity. *Nat Immunol*. 2005;6(6):587-592.
  32. Shu F, Lv S, Qin Y, et al. Functional characterization and characterization of expression of human PFTK1 as a cyclin-dependent kinase. *Proc Natl Acad Sci U S A*. 2007;104(22):9248-9253.
  33. Besset V, Rhee K, Wolgemuth DJ. The identification and characterization of expression of Pftaire-1, a novel Cdk family member, suggest its function in the mouse testis and nervous system. *Mol Reprod Dev*. 1998;50(1):18-29.
  34. Quina LA, Wang S, Ng L, Turner EE. Brn3a and Nurr1 mediate a gene regulatory pathway for habenula development. *J Neurosci*. 2009;29(45):14309-14322.
  35. Hume DA. Macrophages as APC and the dendritic cell myth. *J Immunol*. 2008;181(9):5829-5835.
  36. Wu H, Rodgers JR, Perrard XY, et al. Deficiency of CD11b or CD11d results in reduced staphylococcal enterotoxin-induced T cell response and T cell phenotypic changes. *J Immunol*. 2004;173(1):297-306.
  37. Yagi M, Miyamoto T, Sawatani Y, et al. DC-STAMP is essential for cell-cell fusion in osteoclasts and foreign body giant cells. *J Exp Med*. 2005;202(3):345-351.
  38. Sawatani Y, Miyamoto T, Nagai S, et al. The role of DC-STAMP in maintenance of immune tolerance through regulation of dendritic cell function. *Int Immunol*. 2008;20(10):1259-1268.
  39. Pooley JL, Heath WR, Shortman K. Cutting edge: intravenous soluble antigen is presented to CD4 T cells by CD8- dendritic cells, but cross-presented to CD8 T cells by CD8+ dendritic cells. *J Immunol*. 2001;166(9):5327-5330.
  40. Schnorrer P, Behrens GM, Wilson NS, et al. The dominant role of CD8+ dendritic cells in cross-presentation is not dictated by antigen capture. *Proc Natl Acad Sci U S A*. 2006;103(28):10729-10734.