

Supplementary Materials for  
**Lymph node–resident dendritic cells drive T<sub>H</sub>2 cell development  
involving MARCH1**

Carlos A. Castellanos *et al.*

Corresponding author: Jeoung-Sook Shin, jeoung-sook.shin@ucsf.edu

*Sci. Immunol.* **6**, eabh0707 (2021)  
DOI: 10.1126/sciimmunol.abh0707

**The PDF file includes:**

Supplementary Methods  
Figs. S1 to S11  
Tables S1 to S3  
References (74, 75)

**Other Supplementary Material for this manuscript includes the following:**

Data files S1 to S4

## Supplemental Methods

### Mice

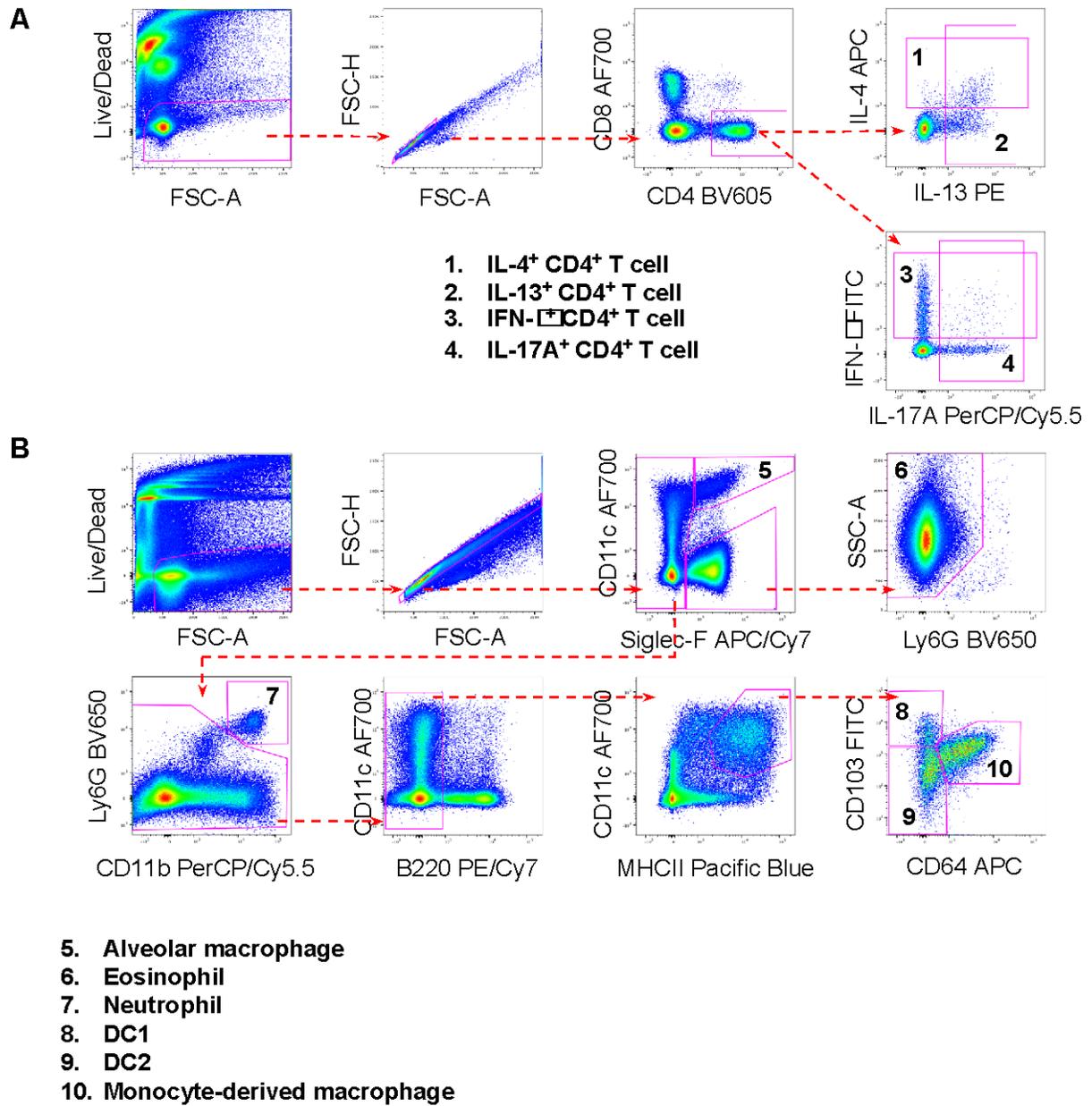
Our lab routinely genotypes MARCH1<sup>fl/fl</sup>Zbtb46<sup>Cre</sup> and MARCH1<sup>fl/fl</sup>MafB<sup>Cre</sup> mice for adventitious germ-line deletion by checking the presence of intact loxp/loxp alleles in ear explants by PCR. We also examine the surface level of MHCII on blood B cells by flow cytometry because B cells are not supposed to be targeted by CD11c<sup>Cre</sup> or MafB<sup>Cre</sup>, but they express MARCH1 and thus increase MHCII surface level when MARCH1 is targeted. The mice that show surface MHCII expression on B cells increased compared to WT mice were excluded from experiments.

### Genetic association with asthma

Since sequencing batch 1 included subjects with asthma cases only (table S1), we could not use the batch variable as a covariate directly in the case/control analysis. Instead, we used the strategies as described below to identify genetic variants associated with sequencing batch and removed them from downstream analysis. In each population, we conducted logistic regression in PLINK to identify genetic variants associated with batch 1 and/or 3 in asthma cases, adjusting for age, sex and the first 5 genetic principal components. In Puerto Ricans, two additional logistic regression analyses were performed: 1) on sequencing batch 1 and 2 in subjects with asthma, adjusting for the same covariates; 2) on sequencing batch 2 and 3 in subjects with and without asthma using asthma status in addition to the same covariates. To minimize false positive discovery in downstream analysis, genetic variants with  $P < 0.05$  in any of the above association analyses were removed from downstream analysis.

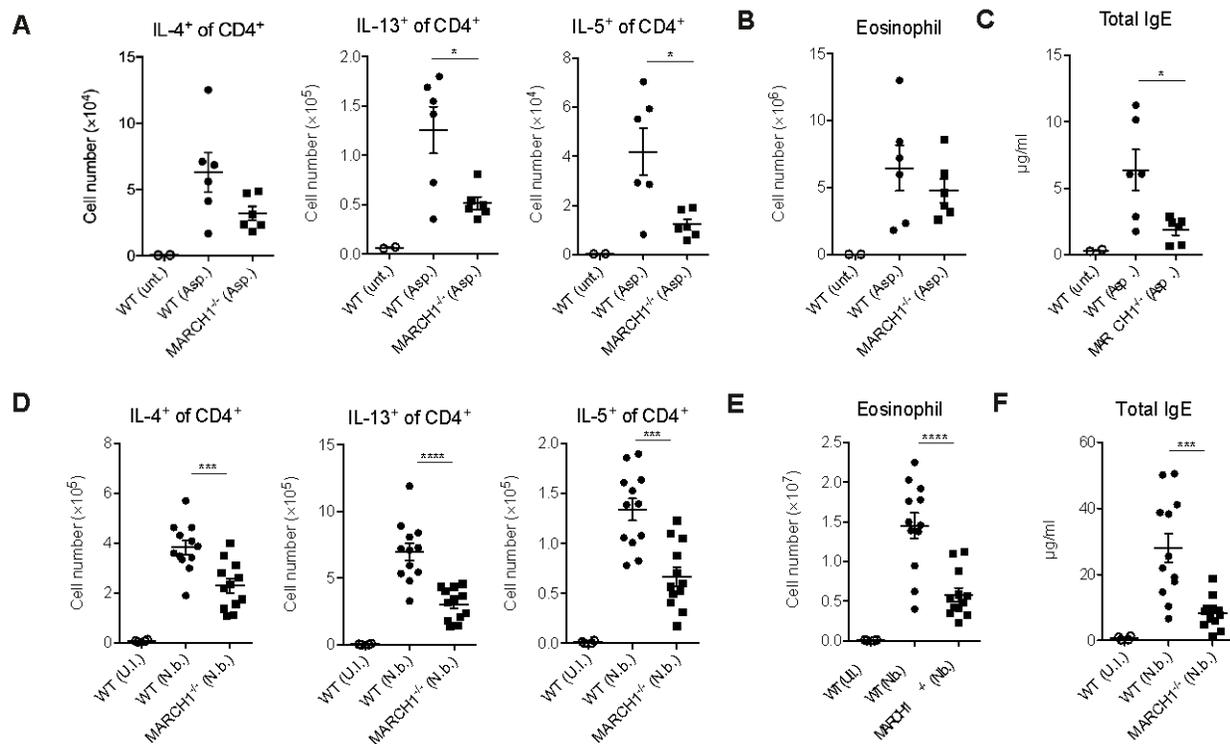
After removing genetic variants potentially associated with sequencing batches, 2,492 common genetic variants (minor allele frequency  $\geq 0.01$ ) in MARCH1 and its 50kb flanking region (hg38 chr4:163,474,298-164,434,050) were used in genetic association analysis on asthma status. Association analysis was performed on each population separately using logistic regression in PLINK and then combined in a trans-ethnic meta-analysis using METAL(74). Age, sex and the first 5 genetic principal components were used as covariates in the discovery

analysis. Sensitivity analysis with and without the smoking status was performed on the top ten associations from the discovery analysis by restricting to subset of samples with non-missing smoking status (table S1). The R (version 3.6.0) package coda (version 0.19-4) was used to estimate the effective number of tests for multiple testing correction in each population and meta-analysis separately. The statistical significance level was calculated by 0.05 divided by the effective number of tests. The effective number of tests for association analyses using African Americans (AA), Puerto Ricans (PR), Mexican Americans (MX) and meta-analysis (meta) were 166, 91, 255 and 122, respectively. Statistical significance levels were calculated by 0.05 divided by the number of effective tests, resulting in  $P$ -value thresholds of:  $3.01 \times 10^{-4}$  for AA,  $5.49 \times 10^{-4}$  for PR,  $1.96 \times 10^{-4}$  for MX and  $4.10 \times 10^{-4}$  for meta-analysis. Suggestive significance levels were calculated by one divided by the number of effective tests, resulting in  $P$ -value thresholds of:  $6.02 \times 10^{-3}$  for AA,  $1.10 \times 10^{-2}$  for PR,  $3.92 \times 10^{-3}$  for MX and  $8.20 \times 10^{-3}$  for meta-analysis. Expression quantitative trait loci (eQTL) from Genotype-Tissue Expression (GTEx) Project version 8 were used for annotations(75). Replication of the top ten trans-ethnic meta-analysis asthma association were performed by querying previously published asthma association studies from Hartford-Puerto Rico Study(40) and Pan-UK Biobank (Phenocode 495, J45, 6152 22127 and 20002) (41).



**Fig. S1. Flow cytometry gating strategy.**

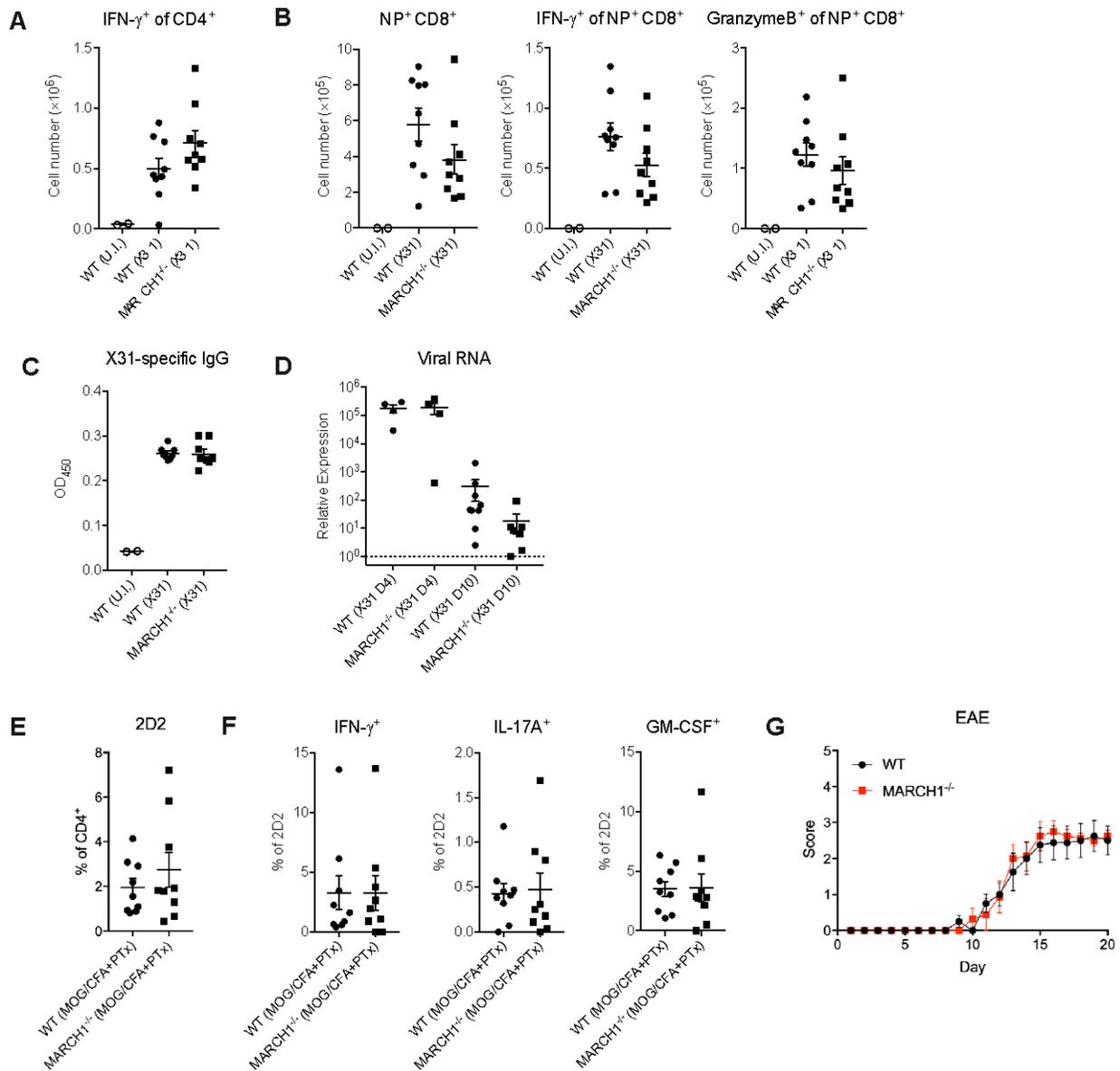
(A and B) Gating strategy to determine cytokine competent CD4<sup>+</sup> T cells (A) and myeloid cells (B) in the lungs. Numbers (1-10) refer to the cell populations labelled on the plots.



**Fig. S2. MARCH1 is required for the development of T<sub>H</sub>2 cell immunity to *Aspergillus fumigatus* extract and parasitic worms.**

(A to C) WT and MARCH1<sup>-/-</sup> mice were given oropharyngeal cellular extract of *Aspergillus fumigatus* (Asp.) on days 0, 7, 8, 9, and 10 and analyzed at day 13. Total numbers of IL-4, IL-13, or IL-5 competent CD4<sup>+</sup> T cells (A) and eosinophils (B) in lungs were determined by flow cytometry. Serum IgE titers (C) were determined by ELISA. Untreated (unt.) mice were used as negative controls. Data in (A to C) are from one experiment with 6 mice per Asp.-treated group. (D to F) WT and MARCH1<sup>-/-</sup> mice were infected with 500 L3 *Nippostrongylus brasiliensis* (N.b.) larvae and analyzed at day 13 post-infection. Total numbers of IL-4, IL-13, or IL-5 competent CD4<sup>+</sup> T cells (D) and eosinophils (E) in lungs were determined by flow cytometry. Serum IgE titers (F) were determined by ELISA. Uninfected (U.I.) mice were used as negative controls. Data in (D to F) are pooled from two independent experiments with 6 mice per N.b.-

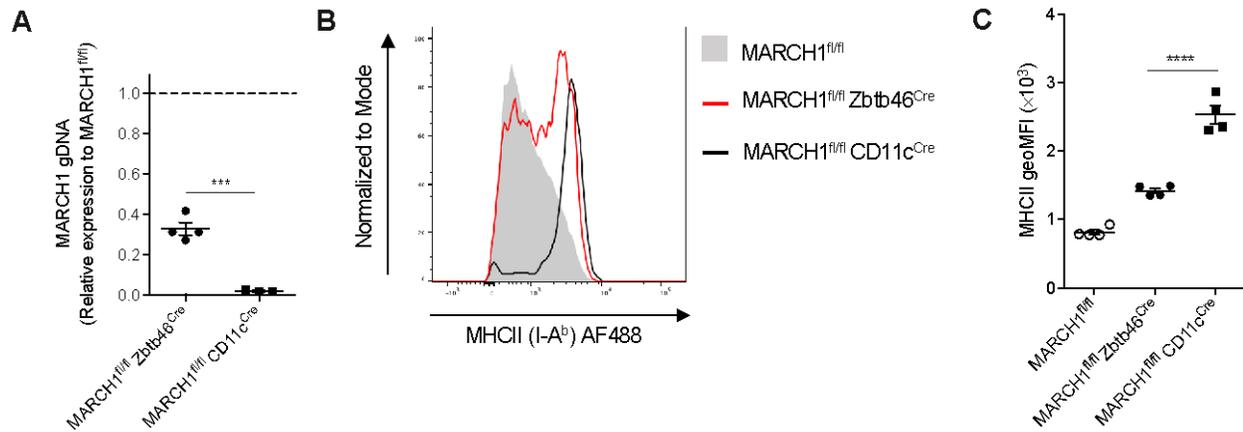
infected group. Data are shown as means  $\pm$  SEM. Statistical significance was determined by one-way ANOVA with Tukey's multiple comparisons test. \* $P < 0.05$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ .



**Fig. S3. MARCH1 is dispensable for the development of TH1, TH17, or cytotoxic CD8<sup>+</sup> T cell immunity to influenza virus or autoimmune self-antigen.**

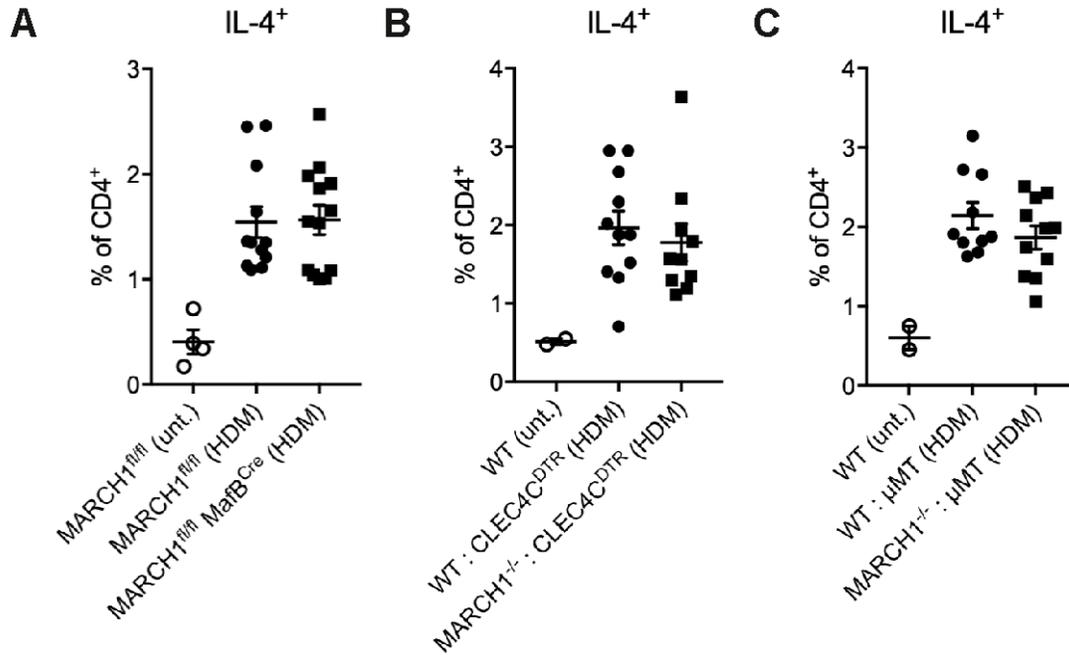
(A) Total numbers of IFN- $\gamma$  competent CD4<sup>+</sup> T cells in lungs of wild-type (WT) or MARCH1<sup>-/-</sup> mice uninfected (U.I.) or infected with  $1 \times 10^5$  TCID<sub>50</sub> of X31 influenza virus. (B) Total numbers of CD8<sup>+</sup> T cells specific for X31 nucleoprotein<sub>366-374</sub> (NP) and the numbers of IFN- $\gamma$  or Granzyme B competent NP-specific CD8<sup>+</sup> T cells in lungs. (C) Optical density (O.D.) values for X31-specific IgG antibodies as determined by ELISA. Data in (A to C) are obtained from one

experiment with 9 mice of each genotype infected with X31 and 2 uninfected control mice. **(D)** Abundance of viral RNA in the lungs. RNA that encodes viral PA protein was quantitated by qRT-PCR at day 4 or 10 after infection. Values were normalized to the mouse with the least viral burden. Data are obtained from one experiment with 4 mice of each genotype at day 4 and 7-9 mice of each genotype at day 10. **(E)** Frequency of adoptively-transferred myelin oligodendrocyte glycoprotein (MOG)-specific 2D2 TCR transgenic T cells in the inguinal lymph node at day 6 after subcutaneous injection of 100 $\mu$ g MOG<sub>35-55</sub> peptide emulsified in Complete Freund's Adjuvant (CFA). Mice were also injected with 200ng of pertussis toxin (PTx) on day 0 and 2. Data are obtained from one experiment with 8 mice of each genotype. **(F)** Frequency of IFN- $\gamma$ , IL-17A, or GM-CSF competent 2D2 T cells. Data are obtained from one experiment with 8 mice of each genotype. **(G)** Clinical scores of experimental autoimmune encephalomyelitis (EAE) development in WT or MARCH1<sup>-/-</sup> mice. Data are obtained from one experiment with 10 mice of each genotype. Data are shown as means  $\pm$  SEM. Statistical significance was not determined by one-way ANOVA with Tukey's multiple comparisons test (**A to D**), unpaired Student's *t*-test (**E and F**) or two-way ANOVA with Fisher's LSD test (**G**).



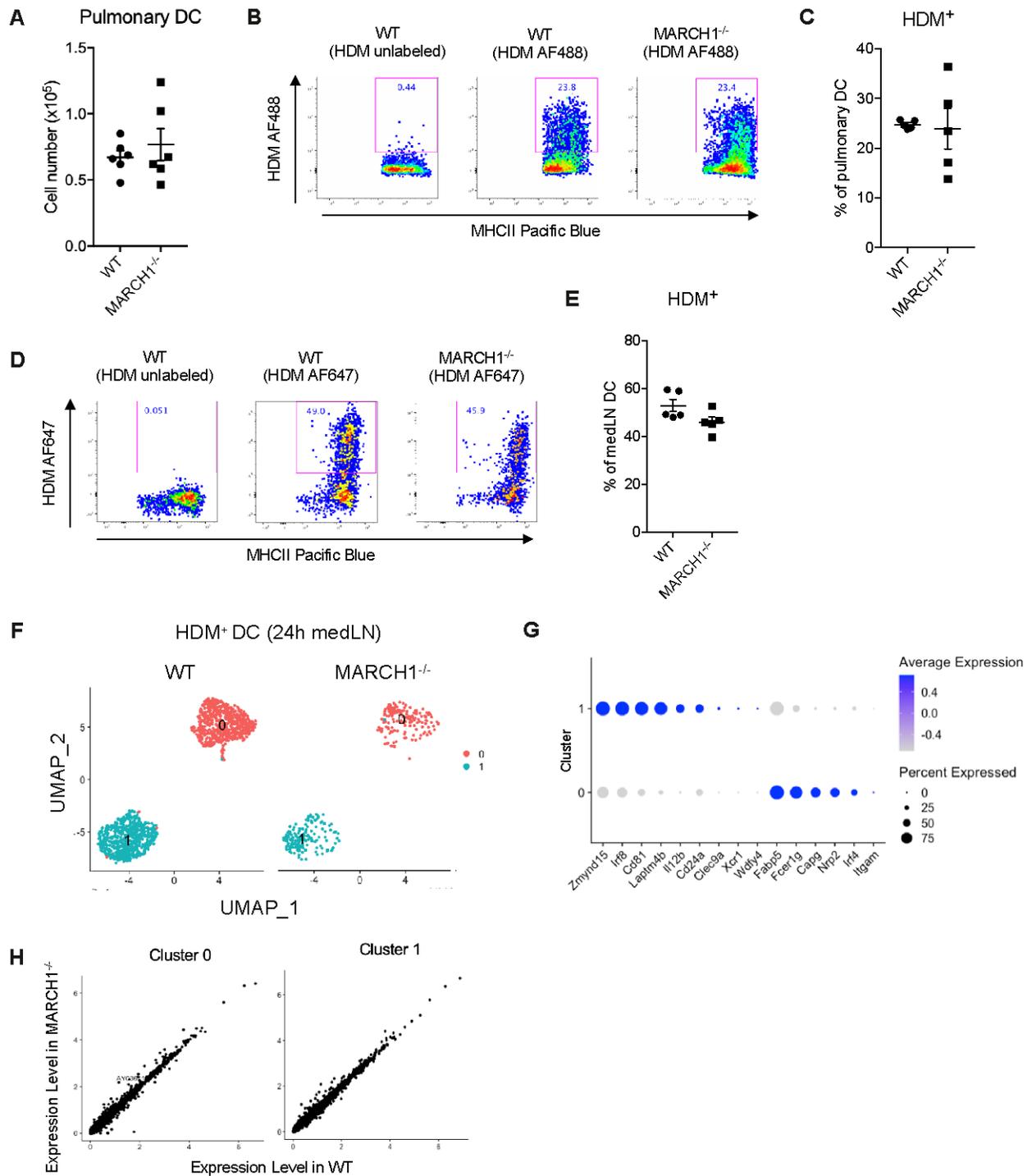
**Fig. S4. MARCH1 deletion efficiency in MARCH1<sup>fl/fl</sup>Zbtb46<sup>Cre</sup> mice.**

(A) Quantification of the genomic DNA encoding MARCH1 loxp site in DCs isolated from spleens of MARCH1<sup>fl/fl</sup>, MARCH1<sup>fl/fl</sup>Zbtb46<sup>Cre</sup>, and MARCH1<sup>fl/fl</sup>CD11c<sup>Cre</sup> mice. Data are presented relative to the amount of the genomic DNA detected in MARCH1<sup>fl/fl</sup> DCs. Data in (A) are from one experiment with 3-4 mice per group. (B and C) Representative histograms showing the MHCII surface expression in splenic DCs (B) and quantification of MHCII surface expression (C) by geometric mean fluorescence intensity (geoMFI). Data in (B and C) are pooled from two experiments with each experiment having 2 mice of each genotype. Data are shown as means  $\pm$  SEM. Statistical significance was determined by Student's *t*-test (A) or one-way ANOVA with Tukey's multiple comparisons test (C). \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ .



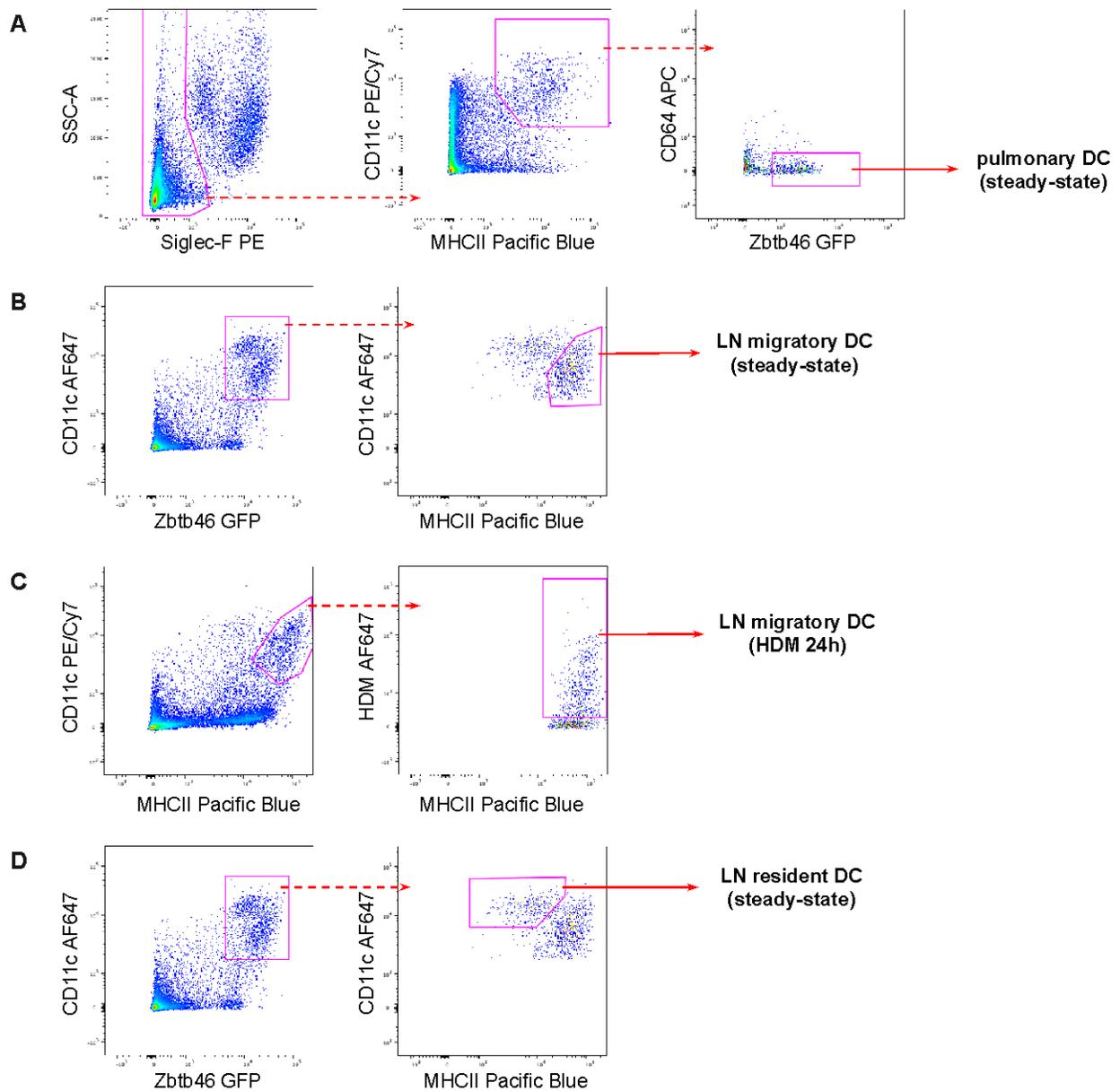
**Fig. S5. MARCH1 expressed by macrophages, plasmacytoid DCs, or B cells does not play a significant role in T<sub>H</sub>2 cell priming.**

(A) Percentages of IL-4 competent CD4<sup>+</sup> T cells in medLNs of MARCH1<sup>fl/fl</sup> and MARCH1<sup>fl/fl</sup>MafB<sup>Cre</sup> mice treated with HDM. Untreated (unt.) mice were used as negative controls. Data are from two independent experiments with 6-7 mice per group. (B) Percentages of IL-4 competent CD4<sup>+</sup> T cells in medLNs of WT:CLEC4C<sup>DTR</sup> or MARCH1<sup>-/-</sup>:CLEC4C<sup>DTR</sup> BM chimeric mice treated with DT and HDM. Untreated (unt.) mice. Data are from one experiment with 10-11 mice per treated group. (C) Percentage of IL-4 competent CD4<sup>+</sup> T cells in medLNs of WT:μMT or MARCH1<sup>-/-</sup>:μMT mixed BM chimeric mice treated with HDM. Untreated (unt.) mice were used negative controls. Data are from one experiment with 10-11 mice per HDM-treated group. Statistical significance was not determined by one-way ANOVA with Tukey's multiple comparisons test. Data are shown as means ± SEM.



**Fig. S6. MARCH1 is dispensable for pulmonary DC homeostasis, capture of HDM allergen, migration to the draining lymph node, or transcriptional reprogramming during migration.**

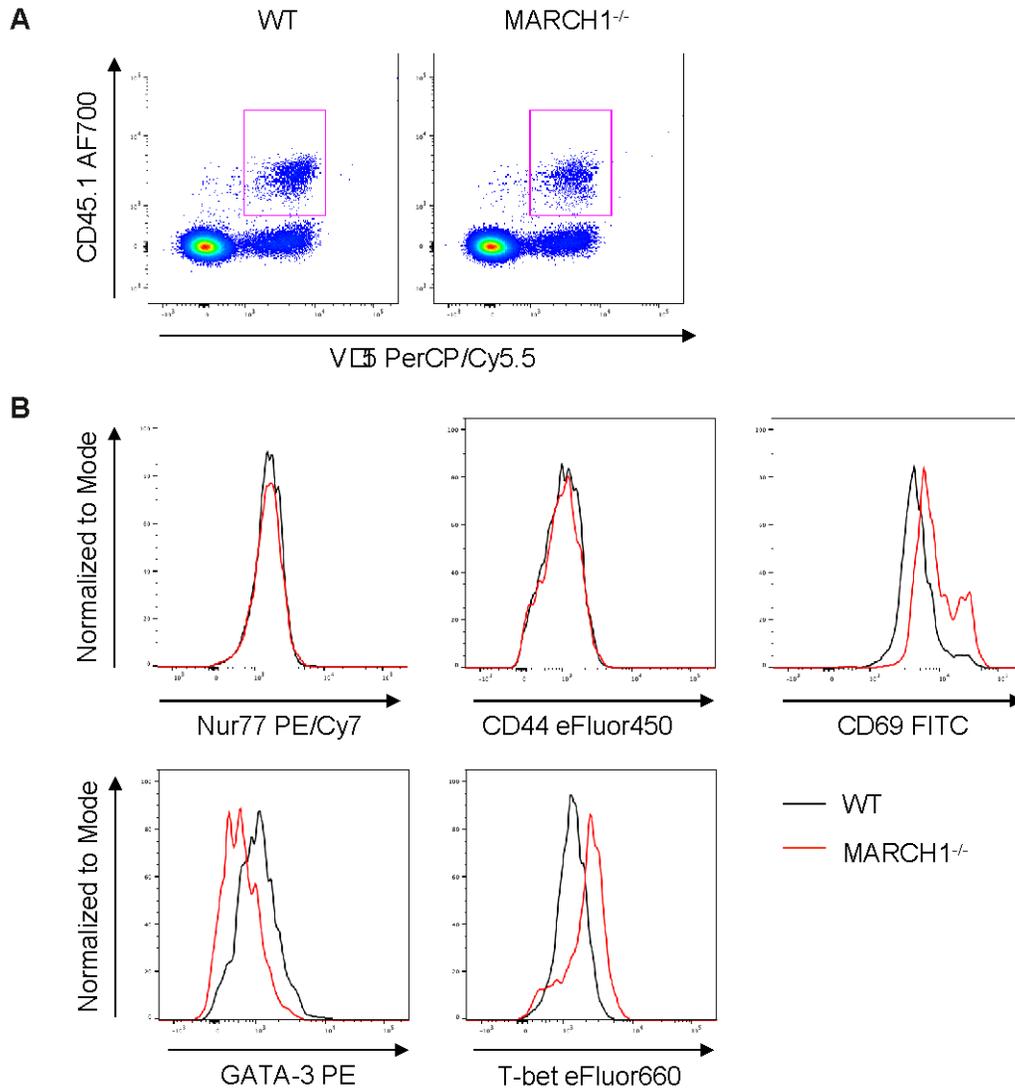
(A) Total numbers of lung DCs (SiglecF<sup>-</sup>B220<sup>-</sup>CD11c<sup>+</sup>MHCII<sup>+</sup>CD64<sup>-</sup> cells) in naïve WT and MARCH1<sup>-/-</sup> mice. Data are pooled from two experiments with 2-4 mice per group. (B and C) WT and MARCH1<sup>-/-</sup> mice were treated with fluorescent HDM (HDM AF488) via oropharyngeal aspiration and 24 h later, pulmonary DCs were analyzed by flow cytometry. A control mouse was treated with non-fluorescent HDM for gating purposes. Shown are representative flow cytometry plots (B) and percentages of pulmonary DCs bearing fluorescent HDM (C). Data are from one experiment with 5 mice per group. (D and E) WT and MARCH1<sup>-/-</sup> mice were treated with fluorescent HDM (HDM AF647) via oropharyngeal aspiration and 24 h later, medLN DCs were analyzed by flow cytometry. A control mouse was injected with non-fluorescent HDM for gating purposes. Shown are representative flow cytometry plots (D) and percentages of medLN DCs bearing fluorescent HDM (E). Data are from one experiment with 5 mice per group. Statistical significance was not determined by unpaired Student's *t*-test. Data are shown as means ± SEM. (F to H) CD45.1/.1 Zbtb46<sup>GFP</sup>:CD45.2/.2 Zbtb46<sup>GFP</sup> MARCH1<sup>-/-</sup> mixed BM chimeric mice were treated with fluorescent HDM (HDM AF647) by oropharyngeal aspiration. HDM<sup>+</sup> DCs derived from WT or MARCH1<sup>-/-</sup> BM (CD45.1<sup>+</sup>CD11c<sup>+</sup>GFP<sup>+</sup>AF647<sup>+</sup> cells and CD45.2<sup>+</sup>CD11c<sup>+</sup>GFP<sup>+</sup>AF647<sup>+</sup> cells, respectively) were sorted from medLN 24 h later and analyzed by single-cell RNA sequencing. UMAP plots show clustering of 1227 WT and 217 MARCH1<sup>-/-</sup> DCs (F) and dot plots show average expression of various transcripts differentially expressed between the two clusters (G). Scatter plots show average expression of individual genes in WT and MARCH1<sup>-/-</sup> DCs of each cluster (H). Note that AY036118 is the only gene differently expressed between WT and MARCH1<sup>-/-</sup> DCs by more than 2-fold (Log<sub>2</sub>FC>1 or Log<sub>2</sub>FC<-1) with adjusted *P*<0.05. Data are from one experiment with cells pooled from 5 mice.



**Fig. S7. FACS gating strategy for purifying pulmonary or medLN DC populations.**

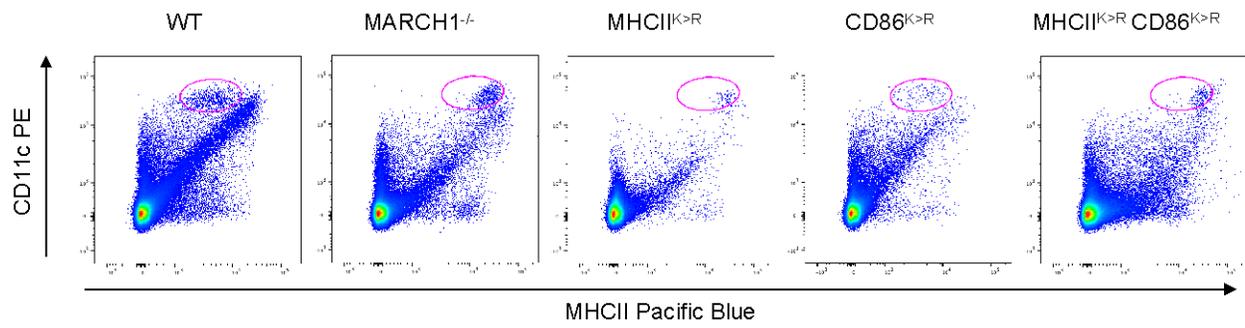
(A to D) Cell suspensions from lungs or medLNs were enriched for DCs by depleting B cells, T cells and neutrophils through negative selection using an EasySep magnet and then stained and sorted with a FACS Aria II. Flow cytometry plots show the gating strategies used for purifying pulmonary DCs at steady-state (A), LN migratory DCs at steady-state (B), LN migratory DCs bearing fluorescent HDM at 24 h after challenge (C), or LN resident DCs at steady-state (D). In

some of the sorts, *Zbtb46*<sup>GFP</sup> reporter mice were used to facilitate the discrimination and isolation of the DCs. The sorts are representative of 3-10 mice and approximately 10,000 DCs were sorted for each population to determine MARCH1 mRNA transcript level by quantitative real-time PCR.



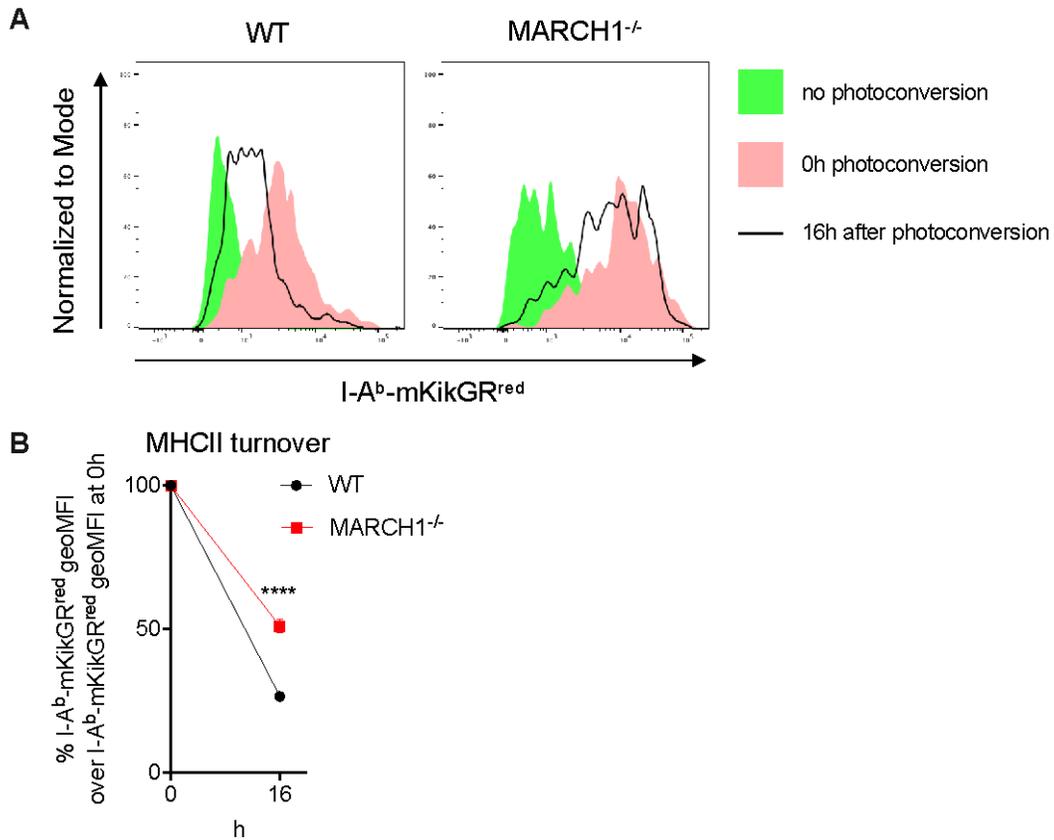
**Fig. S8. Gating of OT-II T cells and extent of expression of activation markers, GATA-3 and T-bet.**

(A and B) Flow cytometry plots of CD4<sup>+</sup> T cells showing the gating of adoptively transferred OT-II cells (A) in the medLN of a WT or a MARCH1<sup>-/-</sup> mouse at 48h after HDM and OVA challenge and histograms showing the extent of expression of Nur77, CD44, CD69, GATA-3 or T-bet (B) in the adoptively transferred OT-II cells. Data are representative from 3 independent experiments with each experiment having 3-5 mice.



**Fig. S9. Gating of LN resident DCs.**

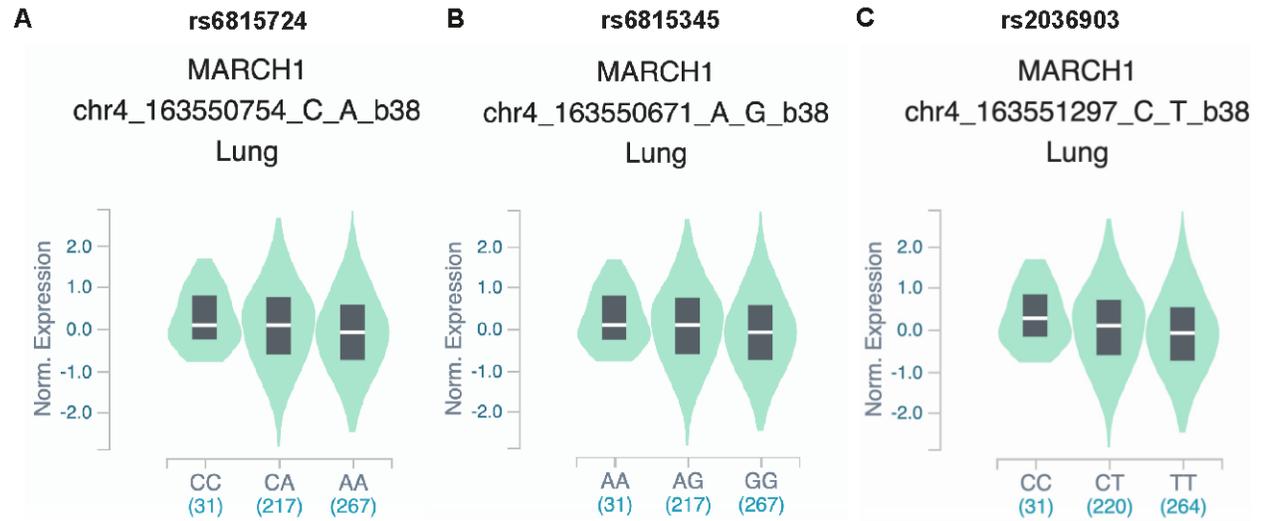
Flow cytometry plots (excluding B220<sup>+</sup> B cells) showing the gating of LN resident DCs in the medLN of WT, MARCH1<sup>-/-</sup>, MHCII<sup>K>R</sup>, CD86<sup>K>R</sup> or MHCII<sup>K>R</sup> CD86<sup>K>R</sup> mice.



**Fig. S10. MARCH1 facilitates MHCII turnover.**

(A) WT or MARCH1<sup>-/-</sup> BMDCs were retrovirally transduced with MSCV-I-Thy1.1 plasmid encoding an I-A<sub>b</sub>-mKikGR fusion protein (the mKikGR is fused to the cytoplasmic tail of MHCII β chain). The cells were exposed to violet light at day 9-10 of culture and analyzed by flow cytometry at the indicated time points. mKikGR<sup>red</sup> fluorescence in DCs is shown in the histograms, which included a “no photoconversion” control. (B) The rate of MHCII turnover was determined by the percent of mKikGR<sup>red</sup> fluorescence left at 16 h post-photoconversion. Data are obtained from one experiment with 2-3 replicates per condition. Data are shown as means ± SEM. Statistical significance was determined by two-way ANOVA with Fisher’s LSD test.

\*\*\*\* $P < 0.0001$ .



**Fig. S11. eQTL violin plots of associations between genotypes of rs6815724, rs6815345, rs2036903 with MARCH1 expression in lung tissue.**

(A) Violin plots of three variants were acquired from the GTEx Project version 8 database and show that protective A allele of rs6815724 (A), G allele of rs6815345 (B), and T allele of rs2036903 (C) are associated with lower MARCH1 expression in lung. Note that these three variants were nominally associated with lower asthma risk in trans-ethnic meta-analysis (see table S2).

	African American (N=1379)		Mexican (N=1353)		Puerto Rican (N=2118)		Overall (N=4860)	
	Control (N=503)	Case (N=876)	Control (N=707)	Case (N=646)	Control (N=936)	Case (N=1182)	Control (N=2146)	Case (N=2704)
<b>age (years)</b>								
Median [Q1, Q3]	16.5 [13.1, 18.9]	13.7 [10.9, 17.0]	14.1 [10.9, 17.5]	12.4 [9.98, 15.3]	13.1 [11.1, 15.8]	12.2 [9.97, 15.2]	14.0 [11.3, 17.6]	12.7 [10.3, 15.9]
<b>Sex</b>								
Female	285 (56.7%)	424 (48.4%)	422 (59.7%)	283 (43.8%)	504 (53.8%)	535 (45.3%)	1211 (56.4%)	1242 (45.9%)
Male	218 (43.3%)	452 (51.6%)	285 (40.3%)	363 (56.2%)	432 (46.2%)	647 (54.7%)	935 (43.6%)	1462 (54.1%)
<b>African (%)</b>								
Median [Q1, Q3]	81.8 [75.4, 87.0]	82.6 [76.2, 86.8]	3.39 [1.97, 4.89]	3.48 [2.29, 4.98]	18.1 [13.0, 25.1]	19.0 [13.8, 26.9]	15.6 [4.89, 41.7]	21.5 [8.23, 75.5]
<b>Indigenous American (%)</b>								
Median [Q1, Q3]	< 1	< 1	58.6 [47.7, 74.6]	56.4 [46.5, 68.2]	10.0 [8.31, 11.9]	10.2 [8.56, 12.0]	11.0 [6.02, 47.5]	9.68 [0.967, 17.4]
<b>European (%)</b>								
Median [Q1, Q3]	17.4 [12.5, 23.7]	16.7 [12.4, 22.9]	36.7 [23.0, 47.8]	39.0 [28.3, 48.5]	71.0 [63.5, 76.4]	69.7 [61.9, 75.5]	49.0 [22.8, 69.7]	46.0 [20.4, 68.2]
<b>Sequencing batch</b>								
Batch 1	0 (0%)	449 (51.3%)	0 (0%)	469 (72.6%)	0 (0%)	440 (37.2%)	0 (0%)	1358 (50.2%)
Batch 2	0 (0%)	0 (0%)	0 (0%)	0 (0%)	474 (50.6%)	567 (48.0%)	474 (22.1%)	567 (21.0%)
Batch 3	503 (100%)	427 (48.7%)	707 (100%)	177 (27.4%)	462 (49.4%)	175 (14.8%)	1672 (77.9%)	779 (28.8%)
<b>Current or ever smoker</b>								
No	308 (61.2%)	608 (69.4%)	591 (83.6%)	582 (90.1%)	718 (76.7%)	793 (67.1%)	1617 (75.3%)	1983 (73.3%)
Yes	73 (14.5%)	71 (8.1%)	76 (10.7%)	35 (5.4%)	27 (2.9%)	20 (1.7%)	176 (8.2%)	126 (4.7%)
Missing	122 (24.3%)	197 (22.5%)	40 (5.7%)	29 (4.5%)	191 (20.4%)	369 (31.2%)	353 (16.4%)	595 (22.0%)

**Table S1. Demographics of GALA II and SAGE participants included in this study.**

Q1/Q3, quantile 1/3. African/Indigenous American/European indicate the corresponding global ancestry proportion.

SNP	BP	Ref	A1	A2	Meta-analysis		AA			MX			PR		
					OR	P	Freq	OR	P	Freq	OR	P	Freq	OR	P
rs6536807	164325654	A	A	G	1.33	2.37E-03 *	0.02	1.06	8.52E-01	0.05	1.20	2.96E-01	0.07	1.45	1.95E-03 *
rs6815724 <sup>e,F</sup>	163550754	C	C	A	1.16	2.62E-03 *	0.26	1.20	4.48E-02	0.15	1.13	2.61E-01	0.27	1.15	4.66E-02
rs6815345 <sup>e,F</sup>	163550671	A	A	G	1.16	3.80E-03 *	0.25	1.19	5.84E-02	0.15	1.13	2.63E-01	0.27	1.15	5.38E-02
rs4691971	164388870	G	G	A	1.29	7.69E-03 *	0.03	1.17	5.26E-01	0.05	1.16	4.09E-01	0.06	1.39	8.66E-03 *
rs72690079	164426854	C	A	C	1.19	8.12E-03 *	0.13	1.25	6.28E-02	0.11	1.12	3.79E-01	0.11	1.19	7.47E-02
rs7666847 <sup>^</sup>	163848106	A	A	G	1.15	8.26E-03	0.24	1.25	1.98E-02	0.24	1.27	9.07E-03	0.13	0.96	6.81E-01
rs4691973 <sup>^,F</sup>	164392771	A	A	C	1.14	9.35E-03	0.28	1.03	7.64E-01	0.16	1.29	2.18E-02	0.25	1.16	4.74E-02
rs2036903 <sup>^,e</sup>	163551297	C	C	T	1.13	1.14E-02	0.28	1.16	1.09E-01	0.15	1.11	3.44E-01	0.27	1.13	8.11E-02
rs1027755	164386782	C	C	T	1.25	1.30E-02	0.04	1.14	5.38E-01	0.06	1.07	6.85E-01	0.06	1.40	6.33E-03 *
rs1027756	164387898	A	A	C	1.24	1.52E-02	0.05	1.03	8.73E-01	0.05	1.11	5.42E-01	0.06	1.43	4.11E-03 *

**Table S2. Top ten genetic associations with asthma in trans-ethnic meta-analysis.**

Variants were sorted by *P*-value in meta-analysis. Odd ratios (OR) and allele frequency (Freq) were expressed relative to allele 1 (A1), which is also the minor allele; BP, hg38 genomic coordinates on chromosome 4; Reference allele (Ref); AA, African Americans; MX, Mexican Americans; PR: Puerto Ricans; \*, suggestively significant asthma association in meta-analysis ( $P < 8.20 \times 10^{-3}$ ). <sup>^</sup>, genetic association with asthma at  $P < 0.05$  in the replication study, Hartford-PR (table S3) or Pan-UK Biobank (supplementary file 2). <sup>F</sup>, flip-flop genetic association with at  $P < 0.05$  in the replication study, Pan-UK Biobank (supplementary file 2). <sup>e</sup>, eQTLs in GTEx (supplementary file 3).

SNP	BP	Ref	A1	A2	Meta-analysis			Replication: Hartford-PR		
					OR	P	*	Freq	OR	P
rs6536807	164325654	A	A	G	1.33	2.37E-03	*	0.08	1.12	5.57E-01
rs6815724 <sup>e, F</sup>	163550754	C	C	A	1.16	2.62E-03	*	0.27	1.05	6.43E-01
rs6815345 <sup>e, F</sup>	163550671	A	A	G	1.16	3.80E-03	*	0.26	1.06	5.85E-01
rs4691971	164388870	G	G	A	1.29	7.69E-03	*	0.08	1.08	6.80E-01
rs72690079	164426854	C	A	C	1.19	8.12E-03	*	0.12	0.79	9.68E-02
rs7666847 <sup>A</sup>	163848106	A	A	G	1.15	8.26E-03		0.15	0.73	1.79E-02
rs4691973 <sup>A, F</sup>	164392771	A	A	C	1.14	9.35E-03		0.27	1.27	2.27E-02
rs2036903 <sup>A, e</sup>	163551297	C	C	T	1.13	1.14E-02		0.27	1.07	5.39E-01
rs1027755	164386782	C	C	T	1.25	1.30E-02		0.08	1.08	6.61E-01
rs1027756	164387898	A	A	C	1.24	1.52E-02		0.08	1.13	4.83E-01

**Table S3. Replication analysis of the top ten associations in the trans-ethnic meta-analysis.**

Variants were sorted by *P*-value in meta-analysis. Odd ratios (OR) and allele frequency (Freq) were expressed relative to allele 1 (A1), which is also the minor allele; BP, hg38 genomic coordinates on chromosome 4; Reference allele (Ref); \*, suggestively significant asthma association in meta-analysis ( $P < 8.20 \times 10^{-3}$ ). <sup>A</sup>, variants with genetic association with asthma at  $P < 0.05$  in the replication study, Hartford-Puerto Rico Study (Hartford-PR) or Pan-UK Biobank (supplementary file 2). <sup>F</sup>, variants with flip-flop genetic association with at  $P < 0.05$  in the replication study, Pan-UK Biobank (supplementary file 2). <sup>e</sup>, eQTLs in GTEx (supplementary file 3).