Host genetic control of Natural Killer cell diversity revealed in the Collaborative Cross

Magali S J Dupont, Vincent Guillemot, Pascal Campagne, Nicolas Serafini, Solenne Marie, Xavier Montagutelli, James P Di Santo, Christian A. J. Vosshenrich

▶ **To cite this version:**

Magali S J Dupont, Vincent Guillemot, Pascal Campagne, Nicolas Serafini, Solenne Marie, et al.. Host genetic control of Natural Killer cell diversity revealed in the Collaborative Cross. Proceedings of the National Academy of Sciences of the United States of America, 2021, 118 (10), ⟨10.1073/pnas.2018834118⟩. ⟨pasteur-03144287⟩

HAL Id: pasteur-03144287

https://pasteur.hal.science/pasteur-03144287v1

Submitted on 9 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

Host genetic control of Natural Killer cell diversity revealed in the Collaborative Cross

Magali S.J. Dupont[a,b,c], Vincent Guillemot[d], Pascal Campagne[d], Nicolas Serafini[a,b], Solenne Marie[a,b], Xavier Montagutelli[e], James P. Di Santo[a,b] and Christian A.J. Vosshenrich[a,b,*]

[a] Innate Immunity Unit, Immunology Department, Institut Pasteur, Paris, France

[b] Institut National de la Santé et de la Recherche Médicale (INSERM) U1223, Paris, France

[c] Université de Paris, Sorbonne Paris Cité, Paris, France

[d] Bioinformatics and Biostatistics Hub, Computational Biology Department, Institut Pasteur, USR 3756 CNRS, Paris, France

[e] Mouse Genetics Laboratory, Department of Genomes and Genetics, Institut Pasteur, Paris, France

* Christian A.J. Vosshenrich

Innate Immunity Unit

NK cell and ILC1 Biology group

Institut Pasteur

25 rue du Docteur Roux

75015 Paris

France

Phone : +-1-4061-3757

Email: christian.vosshenrich@pasteur.fr

ORCiD number Christian A.J. Vosshenrich: 0000-0002-4407-8649

Classification

Major: Biological Sciences; minor: Immunology and Inflammation.

Keywords

NK cells, Innate immunity, Collaborative Cross, Quantitative trait locus mapping.

Author Contributions

M.S.J.D. performed experiments, statistical analyses, QTL mapping and wrote the manuscript; V.G. and P.C. performed statistical analyses, adapted the R package R/QTL2 for the QTL mapping, provided bioinformatical support and edited the manuscript; N.S. and S.M. pre-screened CC mice; X.M. contributed to the design of the study and the bioinformatical analyses and edited the manuscript; J.P.D. contributed to the design of the study, edited the manuscript, obtained the funding; C.V. designed the study, performed experiments and QTL mapping and wrote the manuscript.

Abstract

Natural Killer (NK) cells are innate effectors armed with cytotoxic and cytokine-secreting capacities whose spontaneous anti-tumor activity is key to numerous immunotherapeutic strategies. However, current mouse models fail to mirror the extensive human immune system variation that exists in the normal population which may impact on NK cell-based therapies. We performed a comprehensive profiling of NK cells in the Collaborative Cross (CC), a collection of novel recombinant inbred mouse strains whose genetic diversity matches that of humans, thereby providing a unique and highly diverse small animal model for the study of immune variation. We demonstrate that NK cells from CC lines displayed a breadth of phenotypic and functional variation reminiscent of that reported for humans with regards to cell numbers, key marker expression and functional capacities. We took advantage of the vast genetic diversity of the CC and identified 9 genomic loci through quantitative trait locus mapping driving these phenotypic variations. SNP haplotype-patterns and variant effect analyses identified candidate genes associated with lung NK cell numbers, frequencies of $CD94^+$ NK cells, and expression levels of NKp46. Thus, we demonstrate that the CC represents an outstanding resource to study NK cell diversity and its regulation by host genetics.

Significance Statement

Our work reveals the breadth of NK cell immune variation present in the *Mus musculus* species as represented by the highly diverse Collaborative Cross (CC) mouse resource. We identify unique CC strains that provide an opportunity to study NK cell differentiation and function, and we identify novel genomic loci driving the variation of relevant NK cell parameters, including lung NK cell numbers, frequencies of $CD94^+$ NK cells, and expression levels of NKp46. Finally, our results reveal the similarities of NK cell immune variation between genetically highly diverse human and mouse populations, highlighting the potential of CC mice as pre-clinical model for the development of immunotherapies targeting NK cells.

Main Text

Introduction

Natural Killer (NK) cells are innate cytotoxic lymphoid cells lacking rearranged antigen-specific receptors (1). Upon recognition of infected, cancerous or stressed cells, NK cells can rapidly exert cytolytic activity through granule- and/or cell surface death ligand-mediated mechanisms; they also secrete high concentrations of cytokines IFNγ and TNF that activate pro-inflammatory circuits (2, 3). The innate reactivity of NK cells is crucial for cancer immunosurveillance and for dictating the early immune response to viral or intra-cellular bacterial infections (4, 5). NK cell-based approaches are being exploited in increasing numbers for cancer immunotherapies (https://clinicaltrials.gov/, search term: NK cells and NK cell therapy).

Both mouse and human NK cells consist of heterogeneous subsets that are phenotypically and functionally distinct (6). Analyses of inbred mouse strains including 129S1/SvlmJ, BALB/c, and principally C57BL/6, have led to fundamental discoveries in NK cell biology (7–10). Nonetheless, inbred mice are unable to model the wide range of immune phenotypes observed in the human population under homeostasis or inflammatory conditions as their genetic diversity is much more limited compared to that of humans (11, 12). One example is CD94 that forms heterodimers with NKG2 family members, including NKG2A or C, exhibiting inhibiting or activating functions depending on the partner subunit (13). It is expressed at stable frequencies of around 50% by NK cells from C57BL/6J (B6) and other inbred strains like 129S1/SvlmJ (14), which contrasts with the wide range of frequencies of $CD94^+$ NK cells observed in humans (12). Similarly, NK cell numbers vary little in inbred strains, while in humans extensive interindividual variation has been reported (11, 15).

Novel mouse resources have been generated that harbor genetic diversity matching that of humans. One such model is the Collaborative Cross (CC), a genetically diverse mouse mapping population designed by the Complex Trait Consortium (16). The CC represents a panel of recombinant inbred strains derived from the intercrossing of 5 common laboratory strains and 3 wild-derived strains representing the three main *M. musculus* subspecies (*M.m. domesticus*, *M.m. musculus* and *M.m. castaneus*). Together, the CC strains harbor 45 million segregating polymorphisms and capture 90% of the genetic diversity present in the *Mus musculus* species (16). This far exceeds the 4 million polymorphisms present in most classical inbred strains including B6 (17, 18). CC strains have been shown to display broad variation of

immune parameters (19–21), and studies using CC mice have led to the identification of novel genes associated with susceptibility to pathogens like Salmonella Typhimurium (22) or West Nile Virus (WNV) (23) and of novel mouse models for human infectious diseases (24–26). This demonstrates that the CC is a powerful model to explore the molecular mechanisms underlying immune variation.

We took advantage of the CC resource to characterize NK cell phenotypic and functional diversity. Based on our analysis of more than 220 mice across 32 CC strains, we observed an extensive variation in absolute NK cell numbers, differentiation states, key marker expression and functional capacities *in vitro* and *in vivo*. By exploiting the genetic diversity of the CC, we identified 9 quantitative trait loci (QTL) that drive variations in unique parameters of NK cell biology.

Results

Extensive variation of NK cell numbers in the CC

We assessed the phenotypic diversity of NK cells at homeostasis in adult mice in four different organs from 32 CC strains compared to control B6 mice. More specifically, we focused on the variation of NK cell numbers, differentiation stages as defined by CD27, CD11b, and KLRG1 expression patterns, and expression of other NK cell surface and intracellular markers, including CD94, DNAM-1, NKp46, and the transcriptions factors (TF) Eomes and T-bet.

NK cells were identified through the combined expression of NKp46 that marks NK cells across species (27), and the transcription factors (TF) Eomesodermin (Eomes) and T-bet, which are important regulators of NK cell development (*SI Appendix,* Fig. S1A) (28, 29). Total cell counts per organ varied among the CC strains (*SI Appendix*, Table S1). Absolute numbers as well as frequencies of splenic NK cells varied markedly between CC strains (Fig. 1A and *SI Appendix,* Fig. S1B), however, cell numbers and frequencies did not correlate. Average splenic NK cell numbers varied up to 25-fold ranging from $0.5x10^6$ (CC037) to more than $12.5x10^6$ (CC001; Fig. 1A). In comparison, B6 mice harbored on average $4x10^6$ splenic NK cells (Fig. 1A). We observed a similar strong variation in absolute NK cell numbers and frequencies in lymphoid (mesenteric lymph nodes [mLN]) and non-lymphoid (liver, lung) organs of the same mice (Fig. 1B and *SI Appendix,* Fig. S1B). Similarly, absolute cell numbers of adaptive lymphocytes, including $CD3^+$ $CD5^+$ T cells and $CD19^+$ B cells, varied as well, although to a lesser extent (*SI Appendix,* Fig. S1C). Variations in the absolute numbers of NK and T cells led to a remarkable range in the spleen of T-to-NK cell ratios among the CC strains ranging, from 65:1 (CC037) to 1.25:1 (CC028). As a comparison, B6 displayed a ratio of 10:1 (Fig. 1C and 1D). Similar great ranges of the T-to-NK cell ratios were observed in all organs (Fig. 1C and D and *SI Appendix,* Fig. S2). Of note, some CC strains harbored similar or even higher NK than T cell numbers in liver (CC041, CC042, and CC075) and/or lung (CC007, CC009, CC013, CC027, CC028, CC041, and CC042) (Fig. 1D and *SI Appendix,* Fig. S2). For example, CC043 mice harbored 3-times more NK cells than T cells in the lung and CC042 mice 2.3-times more NK cells than T cells in the liver (Fig. 1D and *SI Appendix,* Fig. S2). However, T cells always prevailed in the spleen and mLN (Fig. 1C and D and *SI Appendix,* Fig. S2). NK cell numbers correlated between spleen, liver and lung (Fig. 1E and F; Pearson's $R>0.5$), which appeared coherent with their recirculation capacity. We did not observe any correlation between NK and T or B cells

numbers in those same organs (Fig. 1E; Pearson's R < 0.5). In contrast, the numbers of NK cells and adaptive lymphocytes were strikingly correlated in mLN (Pearson's R ≥ 0.70), thus suggesting a common factor driving NK, T and B cell counts at this site (Fig. 1E and G).

NK cells from CC mice display diverse patterns of cellular differentiation

NK cell differentiation in peripheral lymphoid and non-lymphoid tissues can be delineated through specific expression patterns of CD27 and CD11b (30–32). The least differentiated NK cells express neither molecule (double negative, DN) and sequentially gain CD27 (CD27 single positive, CD27SP) followed by CD11b expression (double positive, DP) before downregulating CD27 (CD11bSP) (30). In adult B6 mice, the frequency of DN NK cells was very low in the spleen (around 5%, Fig. 2A and *SI Appendix,* Fig. S3A). CD27SP and DP subsets represented around 10% and 20%, respectively, while the mature CD11bSP population was the major subset, accounting for 60 to 70% of total splenic NK cells (Fig. 2A and *SI Appendix,* Fig. S3B-D). In most CC strains, DN and DP NK cells were minor subsets (2-25%) (Fig. 2A and *SI Appendix,* Fig. S3A and S3C). In contrast, the proportion of CD27SP cells varied considerably: in CC001, this subset was marginal (2-4%), while it represented the largest subset in strains CC003, CC061 and CC075, accounting for up to 50% of total splenic NK cells (Fig. 2A and *SI Appendix,* Fig. S3B). Likewise, the proportion of CD11bSP was highly variable, ranging from 85% in CC001 to 25% in CC003 (Fig. 2A and *SI Appendix,* Fig. S3D).

The cadherin receptor KLRG1 is co-expressed with CD11b and thought to mark terminally differentiated NK cells (32–34). However, given the inhibitory activity of KLRG1, it was also suggested that it might be a marker of NK cell senescence or exhaustion (33–35). Around 40% of all splenic NK cells expressed KLRG1 in B6, while this proportion ranged from 10% (CC075) to more than 80% (CC013) across CC strains (Fig. 2B). Considering the inhibitory properties of KLRG1, such variations, observed both in spleen and liver (Fig. 2B), potentially impact functional capacities of NK cells. Moreover, the frequencies of $KLRG1^+$ NK cells were positively correlated with CD11bSP (Fig. 2C), but negatively with DP (not shown) or CD27SP across the CC strains (Fig. 2C). This suggests that, although proportions of mature NK cells subsets were variable, KLRG1 expression was similarly regulated in most CC strains and B6.

NK cells display a high phenotypic diversity in the CC

We further assessed phenotypic NK cell variation through analysis of key surface and intracellular marker expression as mentioned above. CD94 associates with NKG2 family members, forming activating receptors with NKG2C and NKG2E, and inhibitory receptors with NKG2A. Previous studies identified CD94 as marker of more functional NK cells (36), and in humans of memory-like NK cells in the context of HIV infection (37). In B6, the proportion of $CD94^+$ NK cells varies with age, as almost all fetal or newborn NK cells express this protein (38, 39), while this frequency decreases to and remains at around 50% in adult mice (Fig. 3A and 3B) (38). We observed a large and continuous variation of the frequency of $CD94^+$ splenic NK cells across the CC strains (Fig. 3A and 3B). Most CC strains harbored between 30 and 60% $CD94^+$ splenic NK cells. The strains CC075 and CC043 presented the most extreme profiles, with respectively around 10% and 80% of $CD94^+$ NK cells resulting in an almost 7-fold difference between these strains (Fig. 3A and 3B). Expression of CD94 did not correlate with CD27, CD11b, or KLRG1 indicating that CD94 expression was not associated to NK cell differentiation.

DNAM-1 is an activating receptor with an essential role in NK cell-mediated tumor surveillance (40) and NK cell memory formation and maintenance upon MCMV infection (41). More recently it has also been identified as a marker of NK cells with increased functional potential (42). As for CD94 expression, we observed a continuous variation of the frequencies of DNAM-1^+ NK cells in the spleen and mLN across our panel of CC strains (Fig. 3A and 3B). In the spleen, this variation ranged from 25% for CC039 to 85% for CC013, and in the mLN it ranged from 35% for CC061 to more than 90% for CC028. Moreover, while around 50% of B6 NK cells express DNAM-1 in both organs, we observed that a majority of the CC strains were displaying a much larger DNAM-1^+ NK population (Fig. 3A and 3B). DNAM-1 expression by splenic NK cells was neither correlated to the expression of the differentiation markers nor CD94 (*SI Appendix,* Fig. S4A).

NKp46 is an evolutionarily conserved activating receptor that can identify NK cells across multiple species with important roles in immune responses to several types of infections, including influenza or meningococcal infections, as well as cancers (43). Interestingly, we found that NKp46 expression levels were not continuously distributed among CC strains. Rather two groups of mice emerged based on NKp46 expression levels: one group expressing levels similar to B6, and a second group expressing around 2-3-fold higher

levels (colored boxes, Fig. 3C and 3D). NKp46 expression levels were highly correlated across NK cells from the different organs (Fig. 3E and *SI Appendix,* Fig. S4B).

Finally, when analyzing the expression levels of the TF T-bet and Eomes, we noticed a more limited variation as most CC strains expressed them at 0.75 to 1.5-fold of B6 levels. However, some outlier strains like CC060, displayed more divergent T-bet expression levels (*SI Appendix,* Fig. S4 C and E). Moreover, expression of each TF also correlated between NK cells from the different tissues we analyzed (*SI Appendix,* Fig. S4 D and F).

NK cells from the CC strains display functional variation in vitro

One hallmark of NK cell functional responses is the secretion of cytokines, especially IFNγ. For this, we selected CC strains that presented unusual profiles for at least one of the previously described parameters: for instance, CC001 presented a highly mature/exhausted NK cell population as 90% of the splenic NK cells were CD11bSP and KLRG1^{+}. In contrast, CC061 harbored predominantly immature NK cells (CD27SP) in the spleen. NK cells from CC042 expressed high levels of NKp46, and CC037 presented the lowest splenic NK cell count. We assessed the capacity of NK cells from these strains to produce IFNγ upon a 4h-stimulation with the inflammatory cytokines IL-12 and IL-18 or phorbol 12-myristate 13-acetate (PMA) and ionomycin. Both stimuli induced IFNγ production by around 70% of B6 and CC059 NK cells (Fig. 4A). NK cells from two (CC001 and CC042) out of the seven tested CC strains produced significantly less IFNγ under both conditions, while NK cells from CC019 and CC061 showed a reduction of IFNγ production in response to IL-12 and IL-18, but not to PMA and ionomycin. In contrast, NK cells from CC002 mice displayed a specific reduction of IFNγ production in response to PMA and ionomycin, but not IL-12 and IL-18 stimulation (Fig. 4A).

A second functional hallmark of NK cells is the elimination of target cells through granule-mediated release of granzyme B and perforin. At steady-state and upon IL-2 stimulation, only CC059 NK cells contained significantly lower granzyme B levels, while all other CC strains harbored amounts similar to those from B6 (*SI Appendix,* Fig. S5). Moreover, granzyme B expression was restricted to CD11b^{+} NK cells in all strains (Fig. 4B). During a cytotoxic response, lytic granules containing granzyme B are released onto target cells, exposing at the same time the granule marker CD107a at the cell surface. We thus assessed the proportion of degranulating NK cells in the various CC strains by measuring CD107a

surface expression upon stimulation. Around 20% of B6 NK cells degranulated in response to NKp46 stimulation in presence of IL-2 *in vitro* (Fig. 4C). Similar results were obtained using NK cells from CC001, CC002, CC019, and CC037 (Fig. 4C), while significantly more NK cells from CC003, CC042, CC051, CC059, and CC061 mice degranulated under these conditions (Fig. 4C and data not shown). When considering the expression levels of NKp46 in relation to these results, we found that they were correlated to the degranulation capacity of NK cells (Fig. 4D; Pearson's R = 0.73). NK cells expressing low levels of NKp46 (B6, CC001 and CC019) had a reduced degranulation capacity compared to NK cells expressing high levels of NKp46 (CC003, CC042, CC051, and CC059; Fig. 4D). Noteworthy, strains CC002 and CC037 did not fit this model as their NK cells expressed intermediate levels of NKp46 and yet presented the same degranulation capacity as NK cells from B6, suggesting that they might be less responsive to NKp46 stimulation. Likewise, CC061 NK cells expressed low levels of NKp46 while degranulating twice as much as B6 NK cells indicating that CC061 NK cells might be more sensitive to stimulation via NKp46 (Fig. 4D). Of note, both $CD11b^+$ as well as $CD11b^-$ NK cells from CC061 mice degranulated upon stimulation via NKp46 and IL-2, while in B6 mice this was restricted to $CD11b^+$ NK cells (Fig. 4E).

QTL mapping of the CC strains uncovers loci controlling NK cell numbers and phenotypes

Compiled results of the phenotypic analyses are presented in a heatmap with the rank of each CC strain for each parameter (*SI Appendix,* Fig. S6). This heatmap shows a very heterogeneous phenotype distribution with each CC strain displaying a distinctive pattern underscoring the extensive NK cell diversity present in the CC (*SI Appendix,* Fig. S6).

We sought to identify the impact of host genetics on the observed immune variations. Indeed, broad-sense heritability analyses returned scores higher than 0.7 for most measured parameters, with some of them displaying scores higher than 0.9, indicating that host genetics appears to be a prime factor driving the variation of those parameters (Table 1).

We performed QTL mapping using R/QTL2 (44) to reveal genetic loci associated to the variations in the traits we measured. We identified significant (genome-wide $p<0.05$; QTL *NK1-3*; Fig. 5, *SI Appendix,* Figs. S7-S9; Table 2) and suggestive ($p < 0.2$, QTL *NK4-9; SI Appendix,* Fig. S10 and Table 2) QTLs that appeared to drive NK cell numbers in the lung (*NK1*) and mLN (*NK4*), the expression level of NKp46 on NK cells from various organs (*NK2*), the frequencies

of CD94$^+$ NK cells in the spleen (*NK3*), the frequencies of KLRG1$^+$ NK cells in spleen and liver (*NK5)*, the frequencies of CD11bSP (*NK6*) and DP cells (*NK7* and *NK8*) in the spleen, and of the frequency of DNAM-1$^+$ NK cells in the mLN (*NK9*) (Fig. 5 and *SI Appendix,* Figs. S8-S10 and Table 2). We focused on the QTLs that reached genome-wide significance for subsequent candidate gene analysis.

For the number of pulmonary NK cells (Fig. 5A), we identified a QTL on chromosome 13, at around 90 Mbp spanning an 8 Mbp support interval (*NK1*; Fig. 5B and C). Additional analyses demonstrated that *NK1* was specific to NK cell numbers in the lung but not in other organs. Examination of the founder allele effects split CC strains into two groups: those that inherited the NZO allele at this specific locus harbored significantly more pulmonary NK cells than CC strains that inherited another allele (Fig. 5C and D and *SI Appendix,* Fig. S7A). We performed a QTL-wide association analysis resulting in 27 genes within *NK1* carrying SNPs unique to the NZO background, including 16 coding-, 8 non-coding- and 3 unclassified genes (Fig. 5E and *SI Appendix,* Fig. S7B). Only 3 genes presented non-synonymous coding variants: *Msh3*, *Xrcc4* and *Zfyve16* (Fig. 5E; Table 2 and *SI Appendix,* Table S2), all of which were of moderate impact according to *Variant Effect Predictor* analysis (VEP) (45).

QTL mapping analysis of NKp46 expression in hepatic NK cells revealed a QTL at 3.22 Mbp on chromosome 7 with a 3.7 Mbp support interval (*NK2*; *SI Appendix,* Fig. S8 A and B). *NK2* was associated to NKp46 expression levels from NK cells in other organs as well, which was not surprising given the high correlation of NKp46 expression between organs (Fig. 3E). Examination of the founder allele effects revealed that the B6 and WSB alleles were associated with low NKp46 expression levels (*SI Appendix,* Fig. S8C-E). Focusing on genes possessing SNPs unique to the B6 and/or WSB background(s) we identified 14 candidate genes within *NK2*, including 10 coding and 4 non-coding genes (*SI Appendix,* Fig. S8 F and G; Table 2 and *SI Appendix,* Table S3). None of these genes presented coding-variants (*SI Appendix,* Table S3).

Finally, we identified a QTL driving the variation of the frequencies of CD94$^+$ splenic NK cells on chromosome 6 at around 190 Mbp, spanning 9 Mbp (*NK3*; *SI Appendix,* Fig. S9 A and B). In this case, the founder effects were more complex. At first glance, a large CD94-expressing NK cell population appeared to be correlated with the NZO and A/J alleles (*SI Appendix,* Fig. S9C). However, when looking precisely at 133.9 Mbp, we noticed that this high percentage of CD94$^+$ splenic NK cells was not consistent among the CC strains that inherited the A/J allele, and that the NZO allele was only present in one CC strain at this locus (*SI Appendix,* Fig. S9D). On the

contrary, the CAST and WSB alleles were represented in several CC strains at 133.9 Mbp on chromosome 6, and CC strains among these haplotype groups presented more consistent profiles (*SI Appendix,* Fig. S9 D and E). Both, CAST and WSB alleles, appeared to negatively influence the percentage of $CD94^+$ NK cells. Among the genes in *NK3* only 47 genes possessed SNPs specific to the CAST and/or WSB background, of which 36 coding genes and 11 were non-coding genes (*SI Appendix,* Fig. S9 F and G; Table 2 and *SI Appendix,* Table S4). Only one gene, *Clec2d* encoding Clr-b presented a non-synonymous variant (*SI Appendix,* Table S4). Of note, the gene encoding CD94, *Klrd1*, was also among the candidate genes suggesting that variations in the coding or regulatory sequences might influence the frequencies of cells expressing this protein. Comparative analyses revealed more than 98% identity of genomic sequences of *Klrd1* from the eight founder strains (*SI Appendix,* Table S5), yet, none of the coding region or promoter polymorphisms could explain the observed variations. Furthermore, none of the identified SNPs fell into one of the 27 open chromatin regions associated with *Klrd1* expression identified in the Immgen Enhancer Networks database (http://rstats.immgen.org/EnhancerControl/index.html). We could identify two SNPs (chr.6: 129,640,167_T/A and chr.6: 129,640,362_G/A) shared by CAST and WSB that represent intronic variants of *Klrc3* coding for the potential CD94 binding partner NKG2E.

Discussion

The Collaborative Cross resource has been conceived and developed by the mouse genetics community to enable a better understanding of how genes control individual trait variation (16). Previous studies estimated that analysis of around 30 CC strains (with on average 5 mice per strain) would allow uncovering large effect QTLs (46). This suggests that the three significant QTLs we identified - *NK1* (number of NK cells in lung), *NK2* (frequencies of $CD94^+$ splenic NK cells), and *NK3* (expression levels of NKp46 by NK cells from spleen, lung and liver) might strongly affect the corresponding NK cell parameter.

Concerning NK cell numbers, we observed large and continuous variations of the values in the different organs across our panel of CC strains. For instance, NK cell numbers in the lung of the two extreme CC strains differed by a factor of 30. These results mirror previous observations made in humans, where 40 to 60-fold differences in lung NK cell frequencies between healthy donors have been reported (15), hereby highlighting the relevance of the CC to model human immune variation. Two of the three genes presenting non-synonymous variants identified within *NK1, Xrcc4* and *Msh3*, encode proteins involved in DNA repair processes, respectively NHEJ and MMR, and were so far mainly studied for their role in maintaining DNA integrity and preventing cancer onset (47, 48). The third gene, *Zfyve16,* encodes Endofin, an endosomal protein suggested to be involved in the bone morphogenetic proteins (BMP)-Smad signaling pathway through stimulation of Smad phosphorylation (49, 50). Given that i*n vitro* studies of human NK cells suggest that BMPs could play a role in thymic NK differentiation and could favor NK cell function (51, 52), it is conceivable that an Endofin-mediated modification of BMP signaling could affect these processes. However, as *VEP* analysis indicated that all three non-synonymous variants appear to only moderately impact gene functions, it is not excluded that other genes within the QTL might influence the absolute numbers of NK cells in the lung. Moreover, given that non-coding RNAs, including miRNAs and lncRNAs, have been shown to exert crucial roles in the regulation of immune cell development and function (53–55), further analyses should consider the impact of polymorphisms in non-coding genes within the QTL as well. The discovery of a genetic factor influencing NK cell homeostasis in a specific microenvironment like the lung would conceivably be of great value for targeted therapies, if amenable to therapeutic manipulation.

In this context it is interesting to note that a previous report on adaptive and innate immune cell populations in the spleen, including NK cells identified as $CD3^{-}DX5^{+}$ cells, identified two QTLs driving the frequencies of total NK cells and of the $CD27^{+}CD11b^{-}$ subset among these total NK cells in the spleen, respectively (20). They reported a negative association of NZO (and NOD) alleles to the frequencies of splenic NK cells, yet, a positive association of NZO (and WSB) alleles to the frequencies of the $CD27^{+}CD11b^{-}$ subset of NK cells in the spleen (20). Collectively, these findings suggest that the NZO genetic background may exert a distinctive but not exclusive influence on NK cell phenotypes.

In contrast to inbred strains such as C57BL/6, 129S1/SvlmJ, and BALB/c (14, 56) that express CD94 at stable frequencies, we observed a large variation of frequencies of $CD94^{+}$ NK cells in the CC strains (10-80%) that matched the variation observed in humans (ranging from 3 to 82%) (12). *NK2* was mapped to chromosome 6 containing only one non-synonymous variant, *Clec2d* encoding Clr-b. Interactions of Clr-b (or CLEC2D), a broadly expressed C -type lectin-related protein (57) with the inhibitory NK cell receptor NKR-P1B (58) were shown to regulate NK cell functions. However, the frequency of $CD94^{+}$ NK cells was not affected in *Nkrp1b*$^{-/-}$ mice, suggesting that *Clec2d* may not be involved in the variation of this NK cell trait (58). Interestingly, the list of candidate genes encompassed *Klrd1*, encoding CD94. Although the genomic sequences of *Klrd1* were identical at more than 98% among the founder strains, the remaining polymorphisms might explain the observed trait variations. Alternatively, the two intronic variants of *Klrc3* that we identified might influence the prevalence of $CD94^{+}$ NK cells.

With regard to our analysis of NKp46 expression, we revealed two groups of CC strains: one group expressed NKp46 at levels similar to B6, and the other one expressed two to three-fold higher levels. Such distribution pattern has not been reported in human cohorts yet, but our experiments suggest that it could influence NK cell responses to NKp46 ligands. *NK3* was mapped to chromosome 7, but none of the 14 genes within this QTL associated to lower NKp46 expression levels presented coding variants. However, two genes, *Tarm1* and *Lair-1*, appeared interesting as they encode inhibitory receptors expressed by subsets of lymphocytes (59–61), including NK cells, suggesting that inhibitory signals might influence NKp46 expression levels. However, further experimentation is required to test this hypothesis.

Besides serving as a resource for genetic studies, the CC represents a unique tool to assess the degree of immune variation present and tolerated in *Mus musculus* and to discover inter-dependencies of immune traits. This includes the identification of strains with extreme phenotypes or "outlier" strains that depart from generally conserved correlations. Such strains represent novel mouse models for the discovery of the molecular mechanisms involved. For example, we observed a range of distributions in CD27- and CD11b-expressing NK cell subsets in different CC lines. The strains CC001 and CC003 represented the extreme ends of this distribution, with the former harboring almost exclusively terminally differentiated CD11bSP cells, while immature NK cells predominated in the latter. High scores in heritability analyses indicated that the frequencies of these subpopulations were influenced by genetic elements as suggested by a previous report (20). Accordingly, we identified 3 suggestive QTLs (*NK6-8*) associated with the prevalence of the different peripheral subpopulations of NK cells, including DP and CD11bSP. In addition, a fourth suggestive QTL, *NK5*, was associated with the frequencies of KLRG1^{+} NK cells. Further exploration of these strains and QTLs might provide novel insights into the mechanisms of NK cell differentiation.

Unexpectedly, NK cell numbers in the liver and/or lung of certain CC strains outnumbered that of tissue T cells. This observation challenges the general concept that innate cell populations are minor compared to predominant adaptive lymphocytes populations. NK cells and T cells have been shown to mutually influence each other during immune responses (62–68). Those studies are mainly based on mouse models using C57BL/6 mice where T cells prevail over NK cells in spleen, liver and lung, respectively. How corresponding immune responses would unfold in CC mice harboring so different T-to-NK cells ratios remains to be investigated, but has the potential to provide new insights into the question of reciprocally regulation of immune cell dynamics and function in health and disease.

The vast variation of immune traits in the CC was also observed with regard to key NK cell functions such as cytokine secretion and degranulation as well. It is important to note that the variation of functional NK cell-responses in the CC appeared to mirror the breadth of human NK cells responses to similar stimuli (69, 70). Quantitative as well as qualitative differences in the respective signaling pathways involved could be at the origin of these variations, while the differential expression of NK cell activating and inhibitory receptors,

known to dynamically regulate NK cell activity, could also influence some of the observed functional responses of NK cells from the diverse CC strains.

Our observations show that the CC displays ranges of phenotypes rather than discrete, distinct traits, reminiscent of immune phenotypes in large cohorts of healthy individuals (11) and exceeding by far observations made in C57BL/6. While mice are indispensable to explore the role of NK cells in immune responses, the choice of the most adequate mouse model is crucial to enable the extraction of pertinent information and their extrapolation to humans. The restricted allelic diversity of classic mouse inbred strains, which allowed the development of a myriad of transgenic mice and tools to modify the genome with highest precision, ensures limited phenotypic variation. This is advantageous for many experimental approaches; however, it limits their capacity of to model the vast phenotypic and functional variation observed in NK cells from healthy human subjects. As discussed above, our results reveal the similarities of NK cell immune variation between human and CC, which is in stark contrast to the lack of thereof in classical inbred strains. This suggests that CC NK cells might be able to phenocopy the variability in responses to interventions observed in humans.

In conclusion, the CC represents a powerful mouse resource for NK cell research. Its extensive genetic and phenotypic diversity can be used to identify candidate genes through genetic mapping analyses or decipher molecular mechanisms leading to outstanding phenotypes, respectively, of relevant NK cell traits. In addition, the CC represents a promising small animal model to test therapeutic strategies targeting NK cells in pre-clinical settings.

Experimental Procedures

Mice

Collaborative Cross Recombinant Inbred mice were purchased at the Systems Genetics Core Facility at the University of North Carolina, Chapel Hill (UNC) or from the Jackson Laboratories, and bred in the animal facility of the Institut Pasteur under Specific Pathogen Free (SPF) conditions. In order to ensure a pool of individuals as homogeneous as possible to limit non-heritable variability, we only analyzed female mice of 9 to 11 weeks of age bred in the same Specific Pathogen Free (SPF) animal facility and free of prior involvement in other procedures. We included in each experiment at least one C57BL/6J (B6) mouse serving both as experimental control and reference for antibody staining. For the phenotyping experiments, we analyzed 6 mice for 3 strains (CC006, CC037, CC059), 4 mice for 2 strains (CC013 and CC060), 2 mice for 1 line (CC039), and 5 mice for the remaining 26 strains. All mice were randomly allocated to experimental groups. All experiments were approved by the Institut Pasteur Ethics Committee (project dap180067, authorized by the French Ministry of Research in January 2019).

Cell extraction

Single-cell suspensions from spleen, mesenteric Lymph Nodes (mLN), liver, lung and small intestinal lamina propria were obtained as detailed in *SI Appendix*.

Flow Cytometry Analysis

Cells were stained for viability with Fixable Viability Dye eFluor™ 506 (eBioscience) in 1X PBS during 15 minutes on ice. Cells were washed with FACS buffer, then stained for extracellular markers in the presence of anti-FcR (anti-CD16/32, clone 2.4G2, BioXcell) for 45 minutes on ice. Following a washing step with FACS buffer, cells were then fixed and permeabilized according to the Foxp3 Transcription Factor Fixation/Permeabilization Kit (eBioscience) protocol. Cells were then intracellularly stained for 30 minutes at RT with agitation. Antibodies used in this study can be found in *SI Appendix*.

For each CC and each experiment, we calculated the gMFI ratio to B6 of each marker in order to compare expression levels between the CC strains across time.

Data were acquired on a BD LSR Fortessa (BD Biosciences) and analyzed with FlowJo software (TreeStar).

In Vitro Cytokine Secretion Assay

For stimulations $1x10^6$ freshly isolated spleen cells were cultured in round-bottom microtiter plates and stimulated for 4 hours at 37°C, 5% CO_2, in RPMI 10% FCS with the following stimuli: for the degranulation assay, cells were stimulated with immobilized anti-NKp46 (20µg/ml, clone 29A1.4, BioLegend) with or without addition of IL-2 (10µg/ml, PeproTech); for detection of IFNγ cells were stimulated with IL-12 (1µg/ml, R&D) or IL-18 (10µg/ml, R&D) alone or in combination (IL-12 + IL-18) or with PMA (Sigma) and Ionomycin (Sigma). For detection of Granzyme B, cells were stimulated with IL-2. Anti-CD107a antibodies monensin (GolgiStop, BD Biosciences) together with brefeldin A (GolgiPlug, BD Biosciences) were added at the beginning of the assays. At the end, cells were washed and stained for extracellular markers followed by fixation using BD Cytofix/Cytoperm (BD Biosciences), and finally stained intracellularly with antibodies to Granzyme B, IFNγ, and Eomes diluted in BD Cytoperm.

Data Normalization

Numerous variables were recorded and were subsequently transformed according to the type of measurement. Number of cells and gMFIs ratios to B6 were log-transformed, whereas a logit transformation was applied to percentages prior to performing any univariate or multivariate analyses.

ANOVA and multiple comparisons

Comparisons of means among different groups (i.e., strains, treatment, haplotype groups) were done using one-way or two-way ANOVAs followed by multiple comparisons with Welch's

correction or Sidàk's correction. Number of individuals per group is indicated in the Figure legend.

Heritability and QTL mapping

We consider in this study the so-called "broad-sense" heritability: briefly, it consists in measuring the total variance of a trait and compute the part of this variance that is explained by genetic variations (71).

QTL analysis was performed using R/qtl2 (44) with a mixed effect linear model. The analysis typically consists in three steps: (i) computing the probability of a chromosomal segment to be derived from a given ancestor, based on a hidden-Markov model; (ii) calculating lod scores across the genome; (iii) assessing the significance level of LOD scores using ad-hoc permutations, as previously published (44). We performed 10,000 permutations for the parameters we analyzed. Our observations being composed of several observations per CC strain, permutations were performed by permuting genotypes across strains to define statistical thresholds and correct multiple testing of the genetic markers. All our statistical corrections were performed in accordance with previously published studies of multi-parameter phenotyping of the CC (19, 21, 72, 73).

The SNP databases used for the QTL mapping and candidate gene identification are listed in *SI Appendix*.

Gene sequence alignment and Variant Effect Predictor analysis

Gene sequences in FASTA format from the different CC founder strains were retrieved from the MGI database (http://www.informatics.jax.org/genes.shtml) and aligned using BLASTN (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The impact of SNPs on gene function was assessed using the Ensemble VEP analysis tool. For this, we uploaded the list of SNPs from the identified haplotypes to the VEP-web interface (http://www.ensembl.org/info/docs/tools/vep/index.html), selected the *Ensemble/ GENCODE transcripts* database and run the analyses with default settings without filters and restrictions.

Software

All R-based analyses were performed with R version 3.6.2 in Rstudio version 1.3.746. QTL analyses were performed with R/qtl2. Graphical representations were either done with ggplot2 in R, or GraphPad Prism. ANOVAs were performed with GraphPad Prism.

Data availability

All data, protocols, and reagents are available in the main text or SI Appendix.

Declaration of Interests

The authors declare no conflicting interest.

Acknowledgments

We thank Franck Bourgade, Isabelle Lanctin, Pauline Griesmar, Tommy Penel and Jêrome Le Boydre from the central animal facility of the Institut Pasteur for the care with which they raise the CC mice. We thank Dr. Jean Jaubert, Dr. Neetu Gupta, and all the members of the Innate Immunity Unit for helpful discussions. M.S.J.D. was supported by the French Ministry of Higher Education, Research and Innovation. J.P.D. was supported by grants from the Institut Pasteur, the Institut National de la Santé et de la Recherche Médicale (INSERM), and the Ligue Nationale Contre le Cancer (LNCC - Équipe labélisée Ligue Contre le Cancer).

References

1. N. Serafini, C. A. J. Vosshenrich, J. P. Di Santo, Transcriptional regulation of innate lymphoid cell fate. *Nat. Rev. Immunol.* **15**, 415–428 (2015).

2. I. Prager, C. Watzl, Mechanisms of natural killer cell-mediated cellular cytotoxicity. *J. Leukoc. Biol.* **105**, 1319–1329 (2019).

3. T. A. Fehniger, *et al.*, Differential Cytokine and Chemokine Gene Expression by Human NK Cells Following Activation with IL-18 or IL-15 in Combination with IL-12: Implications for the Innate Immune Response. *J. Immunol.* **162**, 4511–4520 (1999).

4. D. H. Raulet, Interplay of natural killer cells and their receptors with the adaptive immune response. *Nat. Immunol.* **5**, 996–1002 (2004).

5. E. Vivier, E. Tomasello, M. Baratin, T. Walzer, S. Ugolini, Functions of natural killer cells. *Nat. Immunol.* **9**, 503–510 (2008).

6. N. D. Huntington, C. A. J. Vosshenrich, J. P. Di Santo, Developmental pathways that generate natural-killer-cell diversity in mice and humans. *Nat. Rev. Immunol.* **7**, 703–714 (2007).

7. K. Kärre, H. G. Ljunggren, G. Piontek, R. Kiessling, Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. *Nature* **319**, 675–678 (1986).

8. S. Kim, *et al.*, Licensing of natural killer cells by host major histocompatibility complex class I molecules. *Nature* **436**, 709–713 (2005).

9. N. C. Fernandez, *et al.*, A subset of natural killer cells achieves self-tolerance without expressing inhibitory receptors specific for self-MHC molecules. *Blood* **105**, 4416–4423 (2005).

10. J. C. Sun, J. N. Beilke, L. L. Lanier, Adaptive immune features of natural killer cells. *Nature* **457**, 557–561 (2009).

11. The Milieu Intérieur Consortium, *et al.*, Natural variation in the parameters of innate immune cells is preferentially driven by genetic factors. *Nat. Immunol.* **19**, 302–314 (2018).

12. L. S. Angelo, *et al.*, Practical NK cell phenotyping and variability in healthy adults. *Immunol. Res.* **62**, 341–356 (2015).

13. A. Gunturi, R. E. Berg, J. Forman, The role of CD94/NKG2 in innate and adaptive immunity. *Immunol. Res.* **30**, 29–34 (2004).

14. M. T. Orr, *et al.*, Development and Function of CD94-Deficient Natural Killer Cells. *PLOS ONE* **5**, e15184 (2010).

15. P. Dogra, *et al.*, Tissue Determinants of Human NK Cell Development, Function, and Residence. *Cell* (2020) https:/doi.org/10.1016/j.cell.2020.01.022 (February 14, 2020).

16. G. A. Churchill, *et al.*, The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat. Genet.* **36**, 1133–1137 (2004).

17. A. Srivastava, *et al.*, Genomes of the Mouse Collaborative Cross. *Genetics* **206**, 537–556 (2017).

18. T. M. Keane, *et al.*, Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289–294 (2011).

19. J. Phillippi, *et al.*, Using the emerging Collaborative Cross to probe the immune system. *Genes Immun.* **15**, 38–46 (2014).

20. R. Collin, L. Balmer, G. Morahan, S. Lesage, Common Heritable Immunological Variations Revealed in Genetically Diverse Inbred Mouse Strains of the Collaborative Cross. *J. Immunol.*, ji1801247 (2018).

21. J. B. Graham, *et al.*, Extensive Homeostatic T Cell Phenotypic Variation within the Collaborative Cross. *Cell Rep.* **21**, 2313–2325 (2017).

22. J. Zhang, *et al.*, Identification of new loci involved in the host susceptibility to Salmonella Typhimurium in collaborative cross mice. *BMC Genomics* **19** (2018).

23. R. Green, *et al.*, Identifying protective host gene expression signatures within the spleen during West Nile virus infection in the collaborative cross model. *Genomics Data* **10**, 114–117 (2016).

24. J. B. Graham, *et al.*, A Mouse Model of Chronic West Nile Virus Disease. *PLoS Pathog.* **12** (2016).

25. A. L. Rasmussen, *et al.*, Host genetic diversity enables Ebola hemorrhagic fever pathogenesis and resistance. *Science* **346**, 987–991 (2014).

26. C. Manet, *et al.*, Genetic Diversity of Collaborative Cross Mice Controls Viral Replication, Clinical Severity, and Brain Pathology Induced by Zika Virus Infection, Independently of Oas1b. *J. Virol.* **94** (2020).

27. T. Walzer, *et al.*, Identification, activation, and selective in vivo ablation of mouse NK cells via NKp46. *Proc. Natl. Acad. Sci.* **104**, 3384–3389 (2007).

28. S. M. Gordon, *et al.*, The Transcription Factors T-bet and Eomes Control Key Checkpoints of Natural Killer Cell Maturation. *Immunity* **36**, 55–67 (2012).

29. M. J. Townsend, *et al.*, T-bet Regulates the Terminal Maturation and Homeostasis of NK and Vα14i NKT Cells. *Immunity* **20**, 477–494 (2004).

30. L. Chiossone, *et al.*, Maturation of mouse NK cells is a 4-stage developmental program. *Blood* **113**, 5488–5496 (2009).

31. Y. Hayakawa, M. J. Smyth, CD27 Dissects Mature NK Cells into Two Subsets with Distinct Responsiveness and Migratory Capacity. *J. Immunol.* **176**, 1517–1524 (2006).

32. N. D. Huntington, *et al.*, NK Cell Maturation and Peripheral Homeostasis Is Associated with KLRG1 Up-Regulation. *J. Immunol.* **178**, 4764–4770 (2007).

33. J. M. Wang, *et al.*, KLRG1 Negatively Regulates Natural Killer Cell Functions through the Akt Pathway in Individuals with Chronic Hepatitis C Virus Infection. *J. Virol.* **87**, 11626–11636 (2013).

34. B. Müller-Durovic, *et al.*, Killer Cell Lectin-like Receptor G1 Inhibits NK Cell Function through Activation of Adenosine 5'-Monophosphate–Activated Protein Kinase. *J. Immunol.* **197**, 2891–2899 (2016).

35. S. H. Robbins, *et al.*, Cutting Edge: Inhibitory Functions of the Killer Cell Lectin-Like Receptor G1 Molecule During the Activation of Mouse NK Cells. *J. Immunol.* **168**, 2585–2589 (2002).

36. J. Yu, *et al.*, CD94 Defines Phenotypically and Functionally Distinct Mouse NK Cell Subsets. *J. Immunol.* **183**, 4968–4974 (2009).

37. Y. Wang, *et al.*, HIV-1-induced cytokines deplete homeostatic innate lymphoid cells and expand TCF7-dependent memory NK cells. *Nat. Immunol.* (2020) https:/doi.org/10.1038/s41590-020-0593-9 (February 24, 2020).

38. K. V. Beneden, *et al.*, Expression of Ly49E and CD94/NKG2 on Fetal and Adult NK Cells. *J. Immunol.* **166**, 4302–4311 (2001).

39. M. Salcedo, *et al.*, Role of Qa-1b-binding receptors in the specificity of developing NK cells. *Eur. J. Immunol.* **30**, 1094–1101 (2000).

40. A. Iguchi-Manaka, *et al.*, Accelerated tumor growth in mice deficient in DNAM-1 receptor. *J. Exp. Med.* **205**, 2959–2964 (2008).

41. T. Nabekura, *et al.*, Costimulatory Molecule DNAM-1 Is Essential for Optimal Differentiation of Memory Natural Killer Cells during Mouse Cytomegalovirus Infection. *Immunity* **40**, 225–234 (2014).

42. L. Martinet, *et al.*, DNAM-1 Expression Marks an Alternative Program of NK Cell Maturation. *Cell Rep.* **11**, 85–97 (2015).

43. A. D. Barrow, C. J. Martin, M. Colonna, The Natural Cytotoxicity Receptors in Health and Disease. *Front. Immunol.* **10** (2019).

44. K. W. Broman, *et al.*, R/qtl2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multiparent Populations. *Genetics* **211**, 495–502 (2019).

45. W. McLaren, *et al.*, Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26**, 2069–2070 (2010).

46. G. R. Keele, W. L. Crouse, S. N. P. Kelada, W. Valdar, Determinants of QTL Mapping Power in the Realized Collaborative Cross. *G3amp58 GenesGenomesGenetics*, g3.400194.2019 (2019).

47. S. N. Andres, *et al.*, A human XRCC4–XLF complex bridges DNA. *Nucleic Acids Res.* **40**, 1868–1878 (2012).

48. F. Palombo, *et al.*, hMutSβ, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Curr. Biol.* **6**, 1181–1184 (1996).

49. Y.-G. Chen, Z. Wang, J. Ma, L. Zhang, Z. Lu, Endofin, a FYVE Domain Protein, Interacts with Smad4 and Facilitates Transforming Growth Factor-β Signaling. *J. Biol. Chem.* **282**, 9688–9695 (2007).

50. J. B. Goh, D. F. Wallace, W. Hong, V. N. Subramaniam, Endofin, a novel BMP-SMAD regulator of the iron-regulatory hormone, hepcidin. *Sci. Rep.* **5**, 1–12 (2015).

51. L. Hidalgo, *et al.*, Expression of BMPRIA on human thymic NK cell precursors: role of BMP signaling in intrathymic NK cell development. *Blood* **119**, 1861–1871 (2012).

52. N. C. Robson, *et al.*, Optimal Effector Functions in Human Natural Killer Cells Rely upon Autocrine Bone Morphogenetic Protein Signaling. *Cancer Res.* **74**, 5019–5031 (2014).

53. Y. G. Chen, A. T. Satpathy, H. Y. Chang, Gene regulation in the immune system by long noncoding RNAs. *Nat. Immunol.* **18**, 962–972 (2017).

54. A. Mehta, D. Baltimore, MicroRNAs as regulatory elements in immune system logic. *Nat. Rev. Immunol.* **16**, 279–294 (2016).

55. W. K. Mowel, *et al.*, Group 1 Innate Lymphoid Cell Lineage Identity Is Determined by a cis-Regulatory Element Marked by a Long Non-coding RNA. *Immunity* **47**, 435-449.e8 (2017).

56. R. E. Vance, A. M. Jamieson, D. Cado, D. H. Raulet, Implications of CD94 deficiency and monoallelic NKG2A expression for natural killer cell development and repertoire formation. *Proc. Natl. Acad. Sci.* **99**, 868–873 (2002).

57. Q. Zhang, *et al.*, Mouse Nkrp1-Clr Gene Cluster Sequence and Expression Analyses Reveal Conservation of Tissue-Specific MHC-Independent Immunosurveillance. *PLOS ONE* **7**, e50561 (2012).

58. M. M. A. Rahim, *et al.*, The mouse NKR-P1B:Clr-b recognition system is a negative regulator of innate immune responses. *Blood* **125**, 2217–2227 (2015).

59. V. Radjabova, *et al.*, TARM1 Is a Novel Leukocyte Receptor Complex–Encoded ITAM Receptor That Costimulates Proinflammatory Cytokine Secretion by Macrophages and Neutrophils. *J. Immunol.* **195**, 3149–3159 (2015).

60. L. Meyaard, *et al.*, LAIR-1, a Novel Inhibitory Receptor Expressed on Human Mononuclear Leukocytes. *Immunity* **7**, 283–290 (1997).

61. X. Tang, *et al.*, Leukocyte-Associated Ig-like Receptor-1–Deficient Mice Have an Altered Immune Cell Phenotype. *J. Immunol.* **188**, 548–558 (2012).

62. G. Gasteiger, *et al.*, IL-2–dependent tuning of NK cell sensitivity for target cells is controlled by regulatory T cells. *J. Exp. Med.* **210**, 1167–1178 (2013).

63. J. Sitrin, A. Ring, K. C. Garcia, C. Benoist, D. Mathis, Regulatory T cells control NK cells in an insulitic lesion by depriving them of IL-2. *J. Exp. Med.* **210**, 1153–1165 (2013).

64. F. Ghiringhelli, *et al.*, CD4+CD25+ regulatory T cells inhibit natural killer cell functions in a transforming growth factor–β–dependent manner. *J. Exp. Med.* **202**, 1075–1085 (2005).

65. F. Bihl, *et al.*, Primed antigen-specific CD4+ T cells are required for NK cell activation in vivo upon Leishmania major infection. *J. Immunol. Baltim. Md 1950* **185**, 2174–2181 (2010).

66. S. N. Waggoner, M. Cornberg, L. K. Selin, R. M. Welsh, Natural killer cells act as rheostats modulating antiviral T cells. *Nature* **481**, 394–398 (2012).

67. P. A. Lang, *et al.*, NK Cells Regulate CD8+ T Cell Mediated Autoimmunity. *Front. Cell. Infect. Microbiol.* **10** (2020).

68. K. Pallmer, *et al.*, NK cells negatively regulate CD8 T cells via natural cytotoxicity receptor (NCR) 1 during LCMV infection. *PLoS Pathog.* **15**, e1007725 (2019).

69. V. Naranbhai, *et al.*, Changes in Natural Killer Cell Activation and Function during Primary HIV-1 Infection. *PLOS ONE* **8**, e53251 (2013).

70. C. W. Pohlmeyer, *et al.*, Identification of NK Cell Subpopulations That Differentiate HIV-Infected Subject Cohorts with Diverse Levels of Virus Control. *J. Virol.* **93** (2019).

71. W. Kruijer, *et al.*, Marker-Based Estimation of Heritability in Immortal Populations. *Genetics* **199**, 379–398 (2015).

72. M. D. Martin, R. Sompallae, C. S. Winborn, J. T. Harty, V. P. Badovinac, Diverse CD8 T Cell Responses to Viral Infection Revealed by the Collaborative Cross. *Cell Rep.* **31**, 107508 (2020).

73. K. E. Noll, *et al.*, Complex Genetic Architecture Underlies Regulation of Influenza-A-Virus-Specific Antibody Responses in the Collaborative Cross. *Cell Rep.* **31**, 107587 (2020).

Figures and Tables

Fig. 1. Extensive variation of NK cell numbers in the CC strains. (A) Absolute number of splenic NK cells was determined by flow cytometry in the indicated strains. Data are represented as mean ± SD. (B) Average NK cell numbers in the mLN, liver and lung, fold-difference between the two extreme strains are highlighted in red, each dot represents the average value of one CC strain. (C) Pyramid-plot of the average splenic NK and T cell numbers in the indicated strains. (D) Ratio of T cells-to-NK cells in the spleen (left) and liver (right) in the indicated strains (log2 scale). (E) Correlation plot (Pearson) of the average NK, T and B cell numbers in the spleen, mLN, liver and lung across all CC strains. Non-significant correlations ($p > 0.05$) are blank. (F) Correlation graphs of the average numbers of splenic against hepatic NK cells (left), splenic NK cells and NK cells in the mLN (right). (G) Correlation graphs of the average numbers of NK cells against T cell (left) and B cell (right) numbers in the mLN across all the CC strains and B6. Regression lines, Pearson's R coefficients and p-values are shown on the graphs. Mean values for the CC strains were calculated with on average 5 mice per CC line (see Methods for more details); B6 with 27 to 29 mice.

Fig. 2. NK cells present diverse patterns of cellular differentiation in the CC strains. (A) Left: representative contour plot of the expression of CD27 vs CD11b (top panel) and KLRG1 vs CD11b (bottom panel) by splenic NK cell from B6, CC003 and CC001; Right: bar graph summarizing percentages of $CD27^{-}CD11b^{-}$, $CD27^{+}CD11b^{-}$, $CD27^{+}CD11b^{+}$, and $CD27^{-}CD11b^{+}$ splenic NK cell subsets for the indicated strains Data are represented as mean ± SD. (B) Left: representative contourplots of Eomes vs KLRG1 expression by splenic and hepatic NK cells from B6, CC075, CC013 and CC001; Right: average percentages of $KLRG1^{+}$ cells among NK cells in the spleen and liver of all CC strains and B6. The fold-difference between the two extreme CC strains is given in red. The extreme strains are identified and color-highlighted according to the colors in the contourplots on the left. B6 is highlighted as a red dot. B cells (light grey) were used as negative population and overlaid on the contourplots. For (A) and (B), numbers on contourplots indicate the percentage of cells within the indicated gates. (C) Correlation graphs of the average percentage of $CD27^{+}$ (top) and percentage of $CD11b^{+}$ (bottom) against the percentage of $KLRG1^{+}$ splenic NK cells across all CC strains. Regression lines, Pearson's R coefficients and p-values are indicated. Mean values for the CC strains were calculated with on average 5 mice per CC line (see Methods for more details); B6 with 27 to 29 individuals.

Fig. 3. NK cells display a high phenotypic variation in the CC strains. (A) Representative contourplots of Eomes vs CD94 expression (top) by splenic NK cells in B6, CC075 and CC043, and expression of Eomes vs DNAM-1 (bottom) by NK cells in the mLN in B6, CC061 and CC028. B cells (light grey) were used as negative population and overlaid on histograms. Numbers on contourplots indicate the percentage of cells within the indicated gates. (B) Average frequencies of NK cells expressing CD94 in spleen and DNAM-1 in mLN and spleen, and of the indicated CC strains. The extreme CC strains are identified and color-highlighted according to the colors in the contourplots in a. B6 is highlighted as a red dot. Fold-differences between the two extreme strains are indicated in red. (C) Representative histogram overlays of NKp46 expression by hepatic NK cells from B6, CC001, and CC042. B cells (light grey) were used as negative population and overlaid on histograms. (D) Ratio of NKp46 expression in the spleen, mLN, lung and liver of all CC strains and B6. The extreme strains are identified and color-highlighted according to the colors in the contourplots in (C). The colored boxes indicate two groups of mice expressing NKp46 at distinct levels in spleen, liver and lung. Deep blue = expression levels are similar to that observed in B6 (ratio ≈1); light blue = expression levels are two- to three-fold higher than in B6. B6 is highlighted as a red dot. (E) Correlation plots of average NKp46 expression in the liver, lung, mLN and spleen across all CC strains. Mean values for the CC strains were calculated with on average 5 mice per CC line (see Methods for more details); B6 with 27 to 29 individuals.

Fig. 4. NK cell from the CC strains present a functional diversity in vitro and in vivo. (A) Left: representative histogram overlays of intracellular IFNγ protein upon IL-12 + IL-18 or PMA + Ionomycin stimulation in splenic NK cells from B6, CC002, CC019 and CC001; Right: bar graph summarizing IFNγ protein expression by splenic NK cells upon IL-12 + IL-18 or PMA + Ionomycin stimulation in the indicated strains. (B) Representative contourplots of Granzyme B vs CD11b expression in splenic NK cells from B6, CC001, CC042 and CC061 following *ex vivo* IL-2 stimulation. Granzyme B gMFI within each indicated gate appears on the plot. (C) Left: Representative histograms of CD107a surface expression upon αNKp46 + IL-2 stimulation by splenic NK cells from B6, CC042, CC002 and CC001; Right: bar graph summarizing CD107a surface expression by splenic NK cells upon αNKp46 + IL-2 from in the indicated strains. For (A) and (C), a two-way ANOVA was performed followed by multiple comparisons, and each CC line was compared to B6 within each group. (D) Correlation graph of the average percentage of $CD107a^+$ NK cells against average NKp46 expression in 9 CC strains and B6 (highlighted in red). Regression line, Pearson's R coefficient and p-value are indicated. (E) Representative contourplots of CD107a versus CD11b expression on splenic NK cells from B6 and CC061 following αNKp46 + IL-2 stimulation. A two-way ANOVA was performed followed by multiple comparisons between strains. ns = non-significant, *: $p<0.05$, **: $p<0.01$, ***: $p< 0.001$, ****: $p<0.0001$.

Fig. 5. QTL driving the variation of NK cell number in the lung. (A) Number of NK cells in the lung of 32 CC strains and B6; the p-value corresponds to a one-way ANOVA. Data are represented as mean ± SD. (B) Genome scan for the number of NK cells in the lung, with a significant peak on chromosome 13 (*p*-value thresholds are indicated). Analysis was performed using the qtl2 package in R, and statistical significance thresholds were determined after 10000 permutations. M: Mitochondria. (C) Top: LOD-score across chromosome 13; bottom: number of NK cell in the lung depending on the inherited haplotype on chromosome 13. Confidence interval of the QTL is highlighted in orange, and the exact position presented in (D) is represented by the black line. (D) Number of NK cells in the lung for each haplotype at 88 Mbp on chromosome 13. A one-way ANOVA was performed followed by multiple comparisons to the NZO group (strains that inherited the NZO allele) with Welch's correction. ns = non-significant, *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, ****: $p<0.0001$. (E) Schematic representation of the strategy used to identify candidate genes and general characteristics of these candidates.

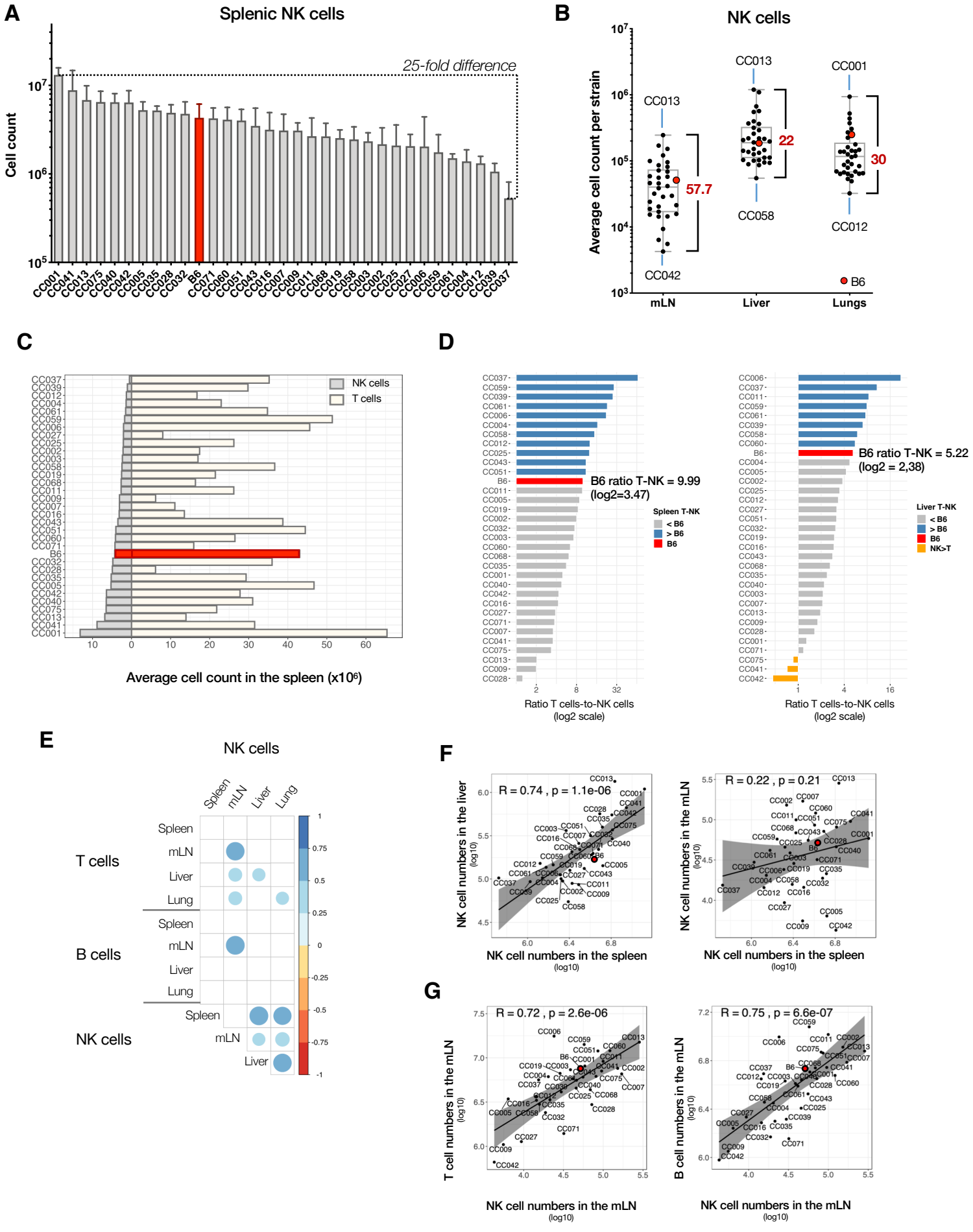


Figure 1. Dupont et al.

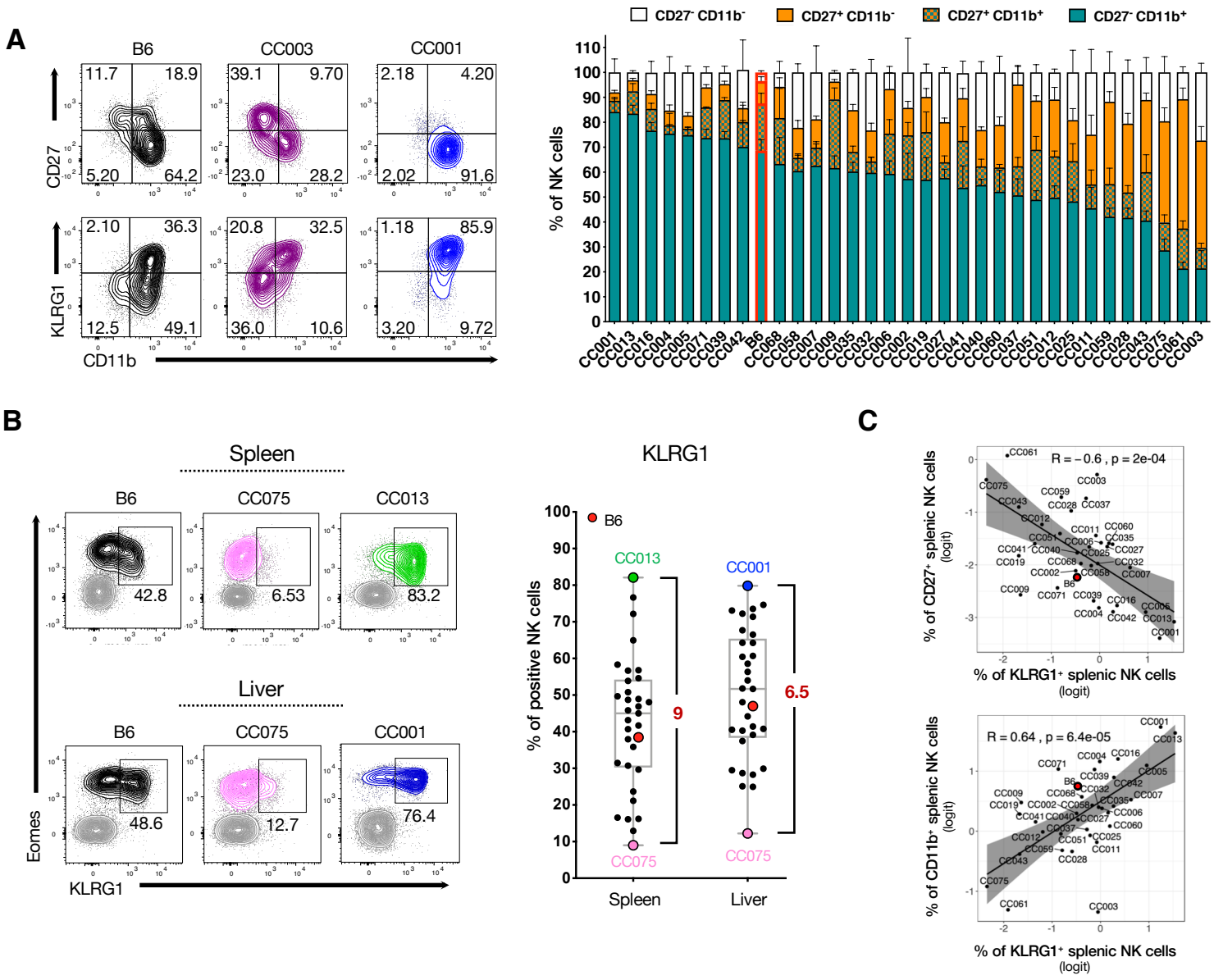


Figure 2. Dupont et al.

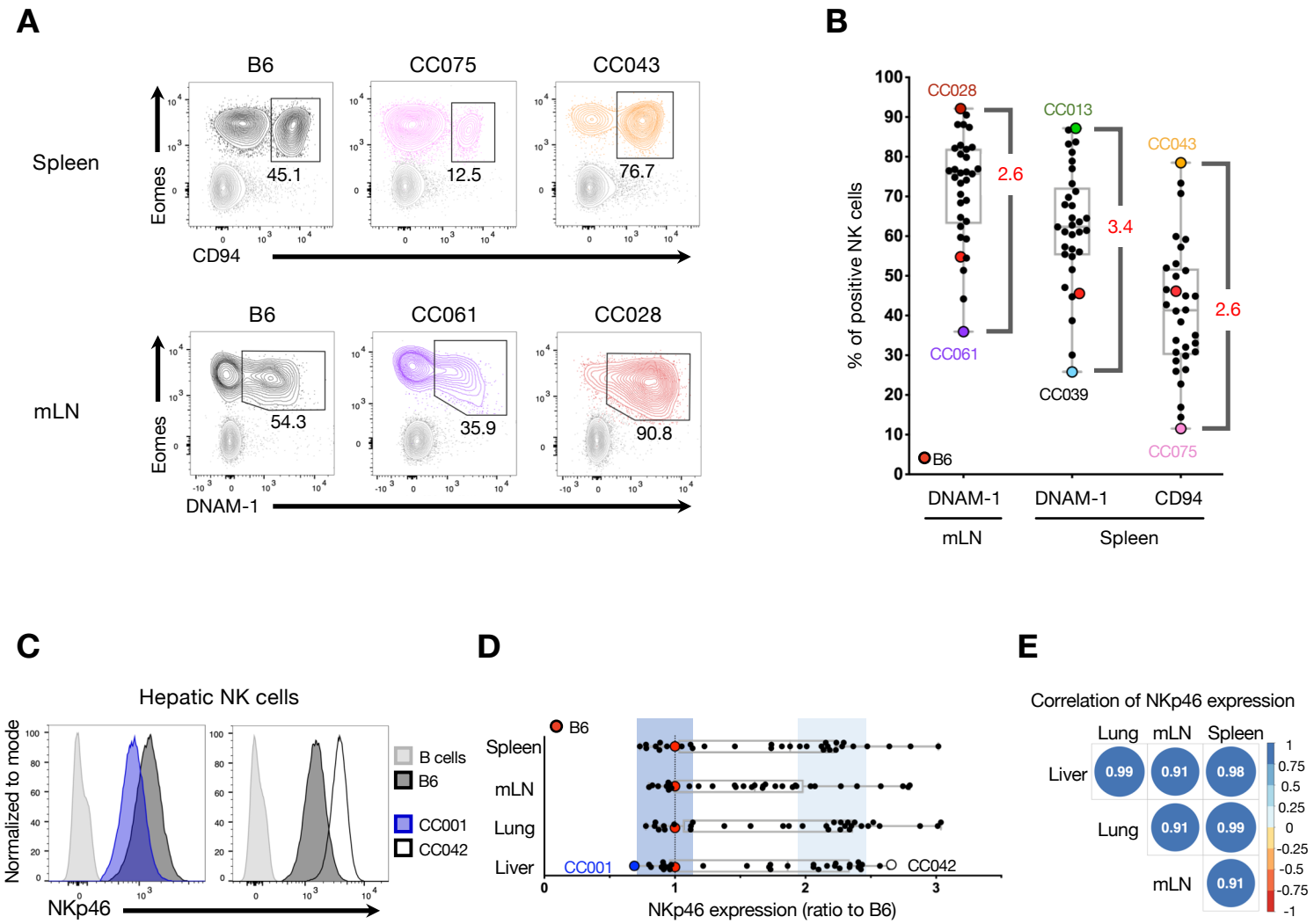


Figure 3. Dupont et al.

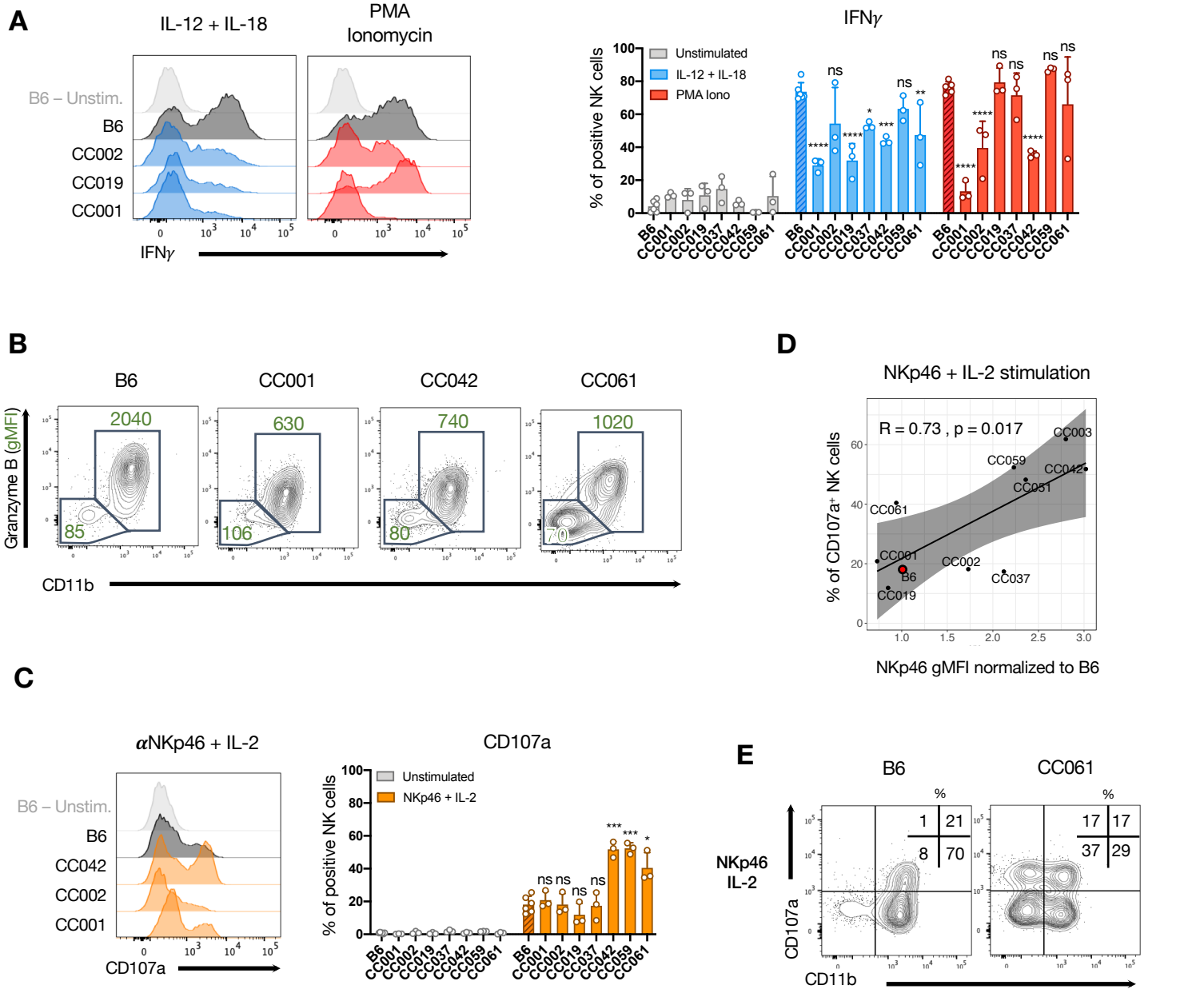


Figure 4. Dupont et al.

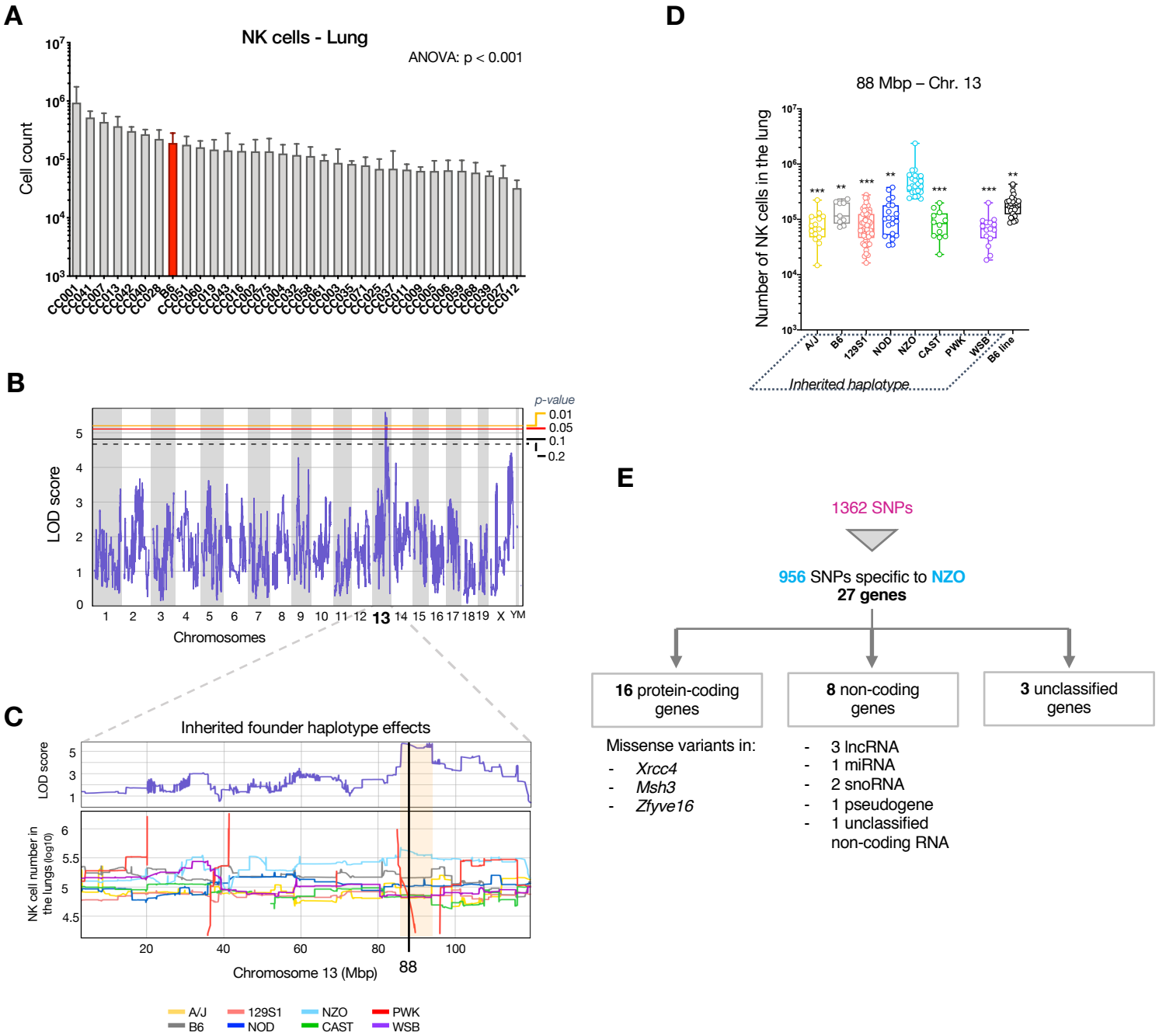


Figure 5. Dupont et al.

Table 1. Broad-sense heritability scores for the indicated traits.

Parameter	Heritability
NK cell numbers	
Spleen	0.67
mLN	0.71
Liver	0.68
Lung	0.67
NK cell differentiation stages - Spleen	
% of CD11bSP	0.84
% of $CD11b^+$ $CD27^+$ (DP)	0.59
% of CD27SP	0.90
% of $CD11b^-$ $CD27^-$ (DN)	0.52
Frequencies of $KLRG1^+$ NK cells	
Spleen	0.83
mLN	0.85
Liver	0.87
Lung	0.86
Frequencies of DNAM-1^+ NK cells	
Spleen	0.82
mLN	0.88
Frequencies of $CD94^+$ NK cells	
Spleen	0.93
NKp46 expression levels (NK cells)	
Spleen	0.94
mLN	0.90
Liver	0.95
Lung	0.94
Eomes expression levels (NK cells)	
Spleen	0.71
mLN	0.75
Liver	0.67
Lung	0.75
T-bet expression levels (NK cells)	
Spleen	0.74
mLN	0.62
Liver	0.75
Lung	0.72

Table 2. Summary table of identified QTLs related to NK cell parameters.

QTL	Phenotype	QTL region	p-value	Founder effect	Number of SNPs	Number of genes	Top candidate genes
Significant peaks							
NK1	NK cell number in the lung	Chr. 13: 85.8-93.93 Mbp	0.0077	NZO high	956	27	*Xrcc4, Zfyve16, Msh3, Atg10*
NK2	NKp46 expression in hepatic NK cells	Chr. 7 : 3.06-6.75 Mbp	0.0018	B6 and WSB low	226	14	*Lair1, Tarm1*
NK3	Frequency of $CD94^+$ splenic NK cells	Chr. 6: 128.22-137 Mbp	0.008	CAST and WSB low	417	47	*Clec2d, Etv6*
Suggestive peaks							
NK4	NK cell number in the mLN	Chr. 4: 102.89-142.91 Mbp	-	-	-	-	-
NK5	Frequency of $KLRG1^+$ NK cells in the liver and spleen	Chr. 19: 3.17-31.23 Mbp	-	-	-	-	-
NK6	Frequency of $CD11b^+$ splenic NK cells	Chr. 7: 36.03-134.99 Mbp	-	-	-	-	-
NK7	Frequency of CD27/CD11b DP splenic NK cells	Chr. 10: 88.23-94.47 Mbp	-	-	-	-	-
NK8		Chr. 13: 48.18-52.75 Mbp	-	-	-	-	-
NK9	Frequency of $DNAM\text{-}1^+$ NK cells in the mLN	Chr. 2: 53.05-76-.25 Mbp	-	-	-	-	-

QTLs were considered as significant when $p < 0.05$, and suggestive when $p < 0.2$.

Supplemental Experimental Procedures

Cell extraction

Following euthanasia, the spleen and mesenteric Lymph Nodes (mLN) were harvested. Liver and lung were perfused with 10 to 20 mL of PBS at room temperature (RT). Spleens were mechanically dissociated on a 100 µm cell strained and Red Blood Cells (RBC) were lysed using RBC Lysis Buffer 10X (BioLegend). mLN, liver and lung were minced and enzymatically digested in RPMI (Gibco) + Fetal Calf Serum (FCS) (Eurobio) 2% with 25 µg/mL of liberase (Roche) and 25 µg/mL of DNAse-I (Roche) during 20 minutes at 37°C with agitation. Following digestion, organs were mechanically dissociated on a 100 µm cell strainer. For the liver and lung, leukocytes were enriched using a Percoll (GE Healthcare) 40-75% density gradient: following 20 minutes 800g centrifugation without break, cells found at the interface between the two phases were collected. Cells were eventually resuspended in fluorescence-activated cell sorting (FACS) buffer (1X PBS, FCS 3%, Hepes 10mM (Fisher Bioreagents), EDTA 2mM (Invitrogen), Penicillin-Streptomycin 100U/mL (Gibco)).

Flow Cytometry Analysis

Detailed list of antibodies used:

Antibodies	source	Identification
Anti-mouse CD3 BV650	BD Biosciences	Cat#: 564378 Clone: 145-2C11 RRID: AB_2738779
Anti-mouse CD5 APCR700	BD Biosciences	Cat#: 565505 Clone:53-7.3 RRID: AB_2739271
Anti-mouse CD11b FITC	eBioscience	Cat#: 11-0112-81 Clone: M1/70 RRID: AB_464934
Anti-mouse CD19 BV786	BD Biosciences	Cat#: 563333 Clone: 1D3 RRID: AB_2738141
Anti-mouse CD27 V450	BD Biosciences	Cat#: 561245 Clone: LG3.A10 RRID: AB_10611853
Anti-mouse CD45 BUV395	BD Biosciences	Cat#: 564279 Clone: 30F11 RRID: AB_2651134
Anti-mouse CD94 PE-Cy7	BioLegend	Cat#: 105509 Clone: 18d3 RRID: AB_2632663
Anti-mouse CD107a FITC	BD Biosciences	Cat#: 553793 Clone: 1D4B RRID: AB_395057

Anti-mouse CD226 PE-Cy7	BioLegend	Cat#: 128811 Clone: 10E5 RRID: AB_2566628
Anti-mouse Eomes eFluor660	eBioscience	Cat#: 50-4875-82 Clone: Dan11mag RRID: AB_2574227
Anti-mouse Granzyme B PE-CF594	BD Biosciences	Cat#: 562462 Clone: GB11 RRID: AB_2737618
Anti-mouse IFNγ PE-Cy7	BD Biosciences	Cat#: 557649 Clone: XMG1.2 RRID: AB_396766
Anti-mouse KLRG1 APC-Cy7	BioLegend	Cat#: 138426 Clone: 2F1/KLRG1 RRID: AB_2566554
Anti-mouse NKp46 PE	eBioscience	Cat#: 12-3351-82 Clone:29A1.4 RRID: AB_1210743
Anti-mouse NKp46 PerCP-eFluor710	eBioscience	Cat#: 46-3351-82 Clone:29A1.4 RRID: AB_1834441
Anti-mouse T-bet PE-Cy7	eBioscience	Cat#: 25-5825-82 Clone: eBio4B10 RRID: AB_11042699
Anti-mouse CD16/CD32	Bio X Cell	Cat#: BE0307 Clone: 2.4G2 RRID: AB_2736987
LEAF purified Anti-mouse NKp46	Biolegend	Cat#137614 Clone: 29A1.4 RRID: AB_11124341

SNP databases used for the QTL mapping and candidate gene identification

For the QTL mapping analyses, GigaMUGA genotypes of the CC strains and founders were retrieved from http://csbio.unc.edu/CCstatus/CCGenomes/#genotypes.

The SNPs on which our GWAS analyses are based were downloaded from:
Broman, Karl (2017): SQLite database of variants in Collaborative Cross founder mouse strains. figshare. Dataset. https://doi.org/10.6084/m9.figshare.5280229.v3

The SQLite database had been created using the following source files available from Sanger (ftp://ftp-mouse.sanger.ac.uk/) or JAX (ftp://ftp.jax.org/SNPtools/variants/):
SNPs: mgp.v5.merged.snps_all.dbSNP142.vcf.gz
Indels: mgp.v5.merged.indels.dbSNP142.normed.vcf.gz
Structural variants: 28strains.REL-1410-SV.sdp.tab.gz

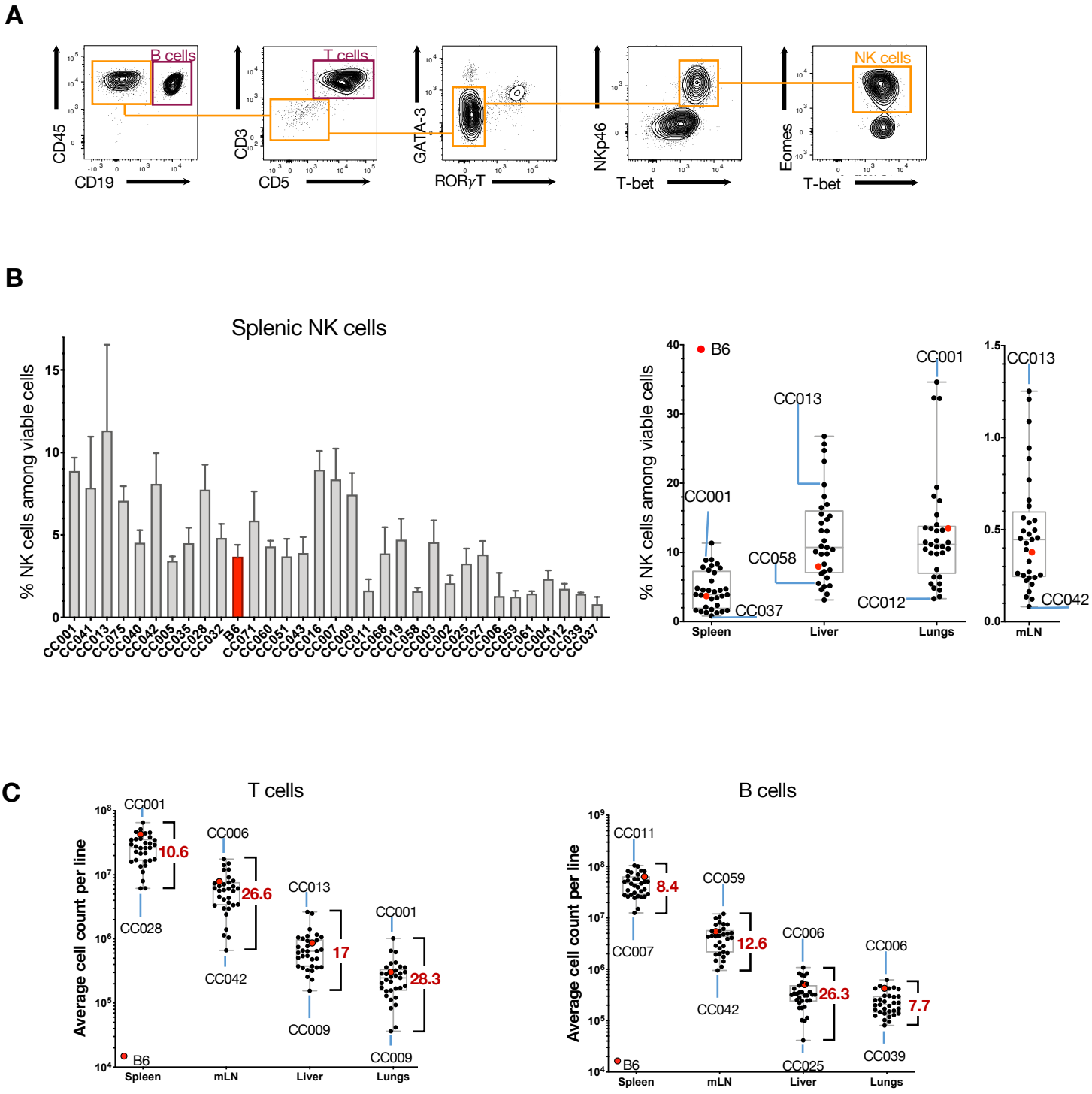


Fig. S1. (A) Gating strategy used to identify B, T and NK cells. Gated on lymphoid, single and live cells: B cells were defined as CD45$^+$ CD19$^+$; T cells as CD45$^+$ CD19$^-$ CD3$^+$ CD5$^+$; NK cells as CD45$^+$ CD19$^-$ CD3$^-$ CD5$^-$ GATA-3$^-$ RORγT$^-$ NKp46$^+$ T-bet$^+$ Eomes$^+$. (B) Left: frequencies of splenic NK cells among viable cells in the indicated lines, right: average NK cell frequencies in spleen, liver, lung and mLN. Data are represented as mean ± SD. (C) Average T (left) and B (right) cell numbers in the spleen, mLN, liver and lung. For (B) and (C), fold-difference between the two extreme lines shown in red alongside the boxplot. One dot represents the average value of one line, with on average 5 mice per CC line (See Methods for more details). B6 is highlighted in red, and average values were calculated with 27 to 29 mice.

Supplemental Figure 1. Dupont et al.

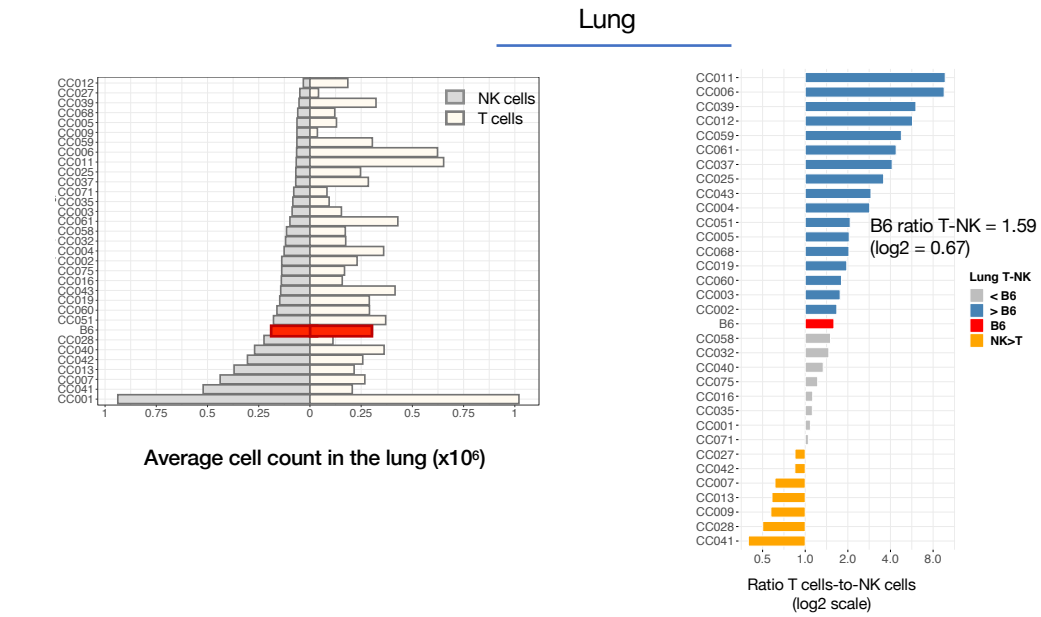
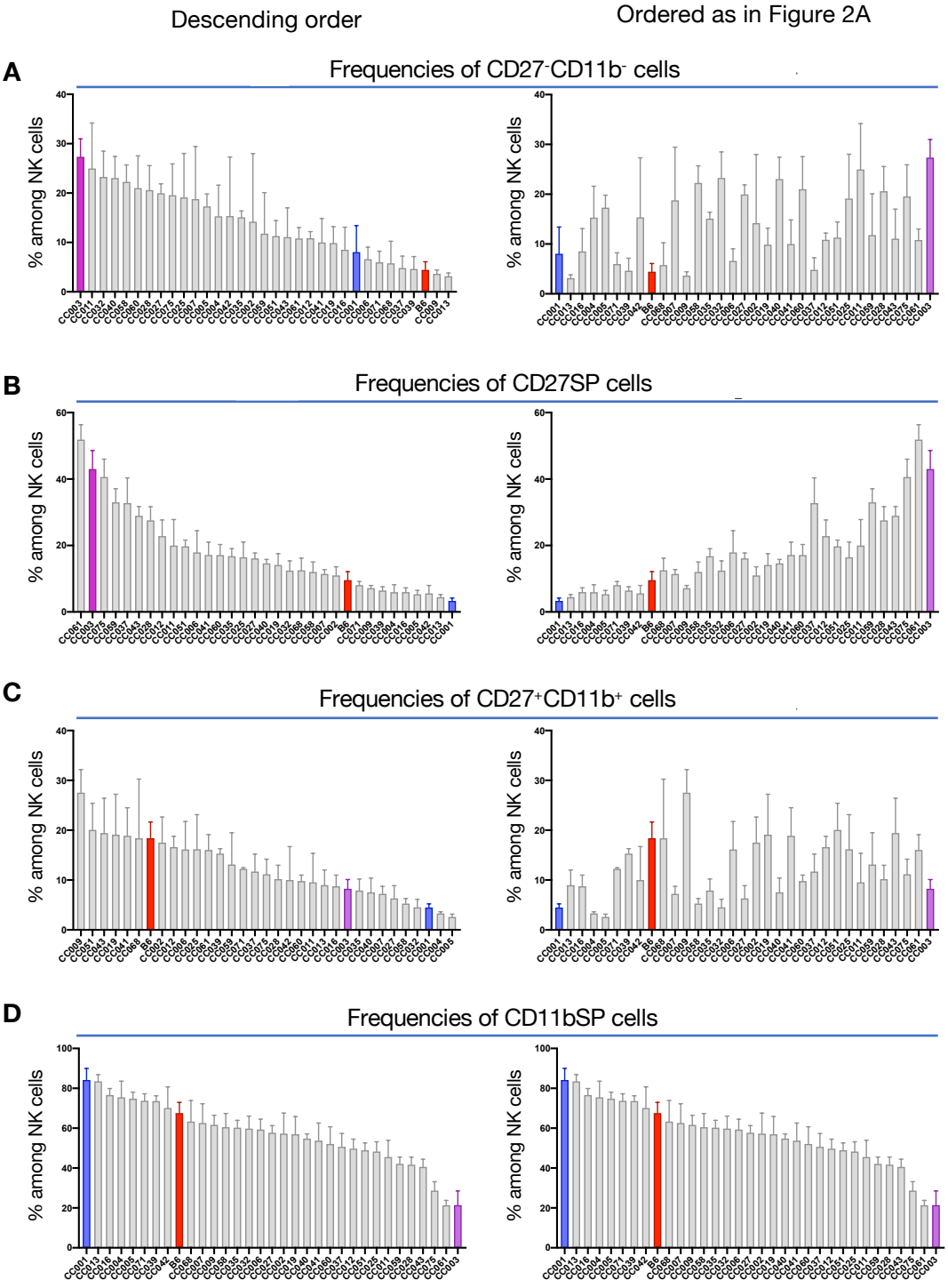


Fig. S2. Pyramid plots of NK and T cell numbers (left) and T-to-NK cell ratio (right) in the lung (top), mLN (middle) and liver (bottom) in the indicated strains (log2 scale). B6 is indicated in red.

Supplemental Figure 2. Dupont et al.



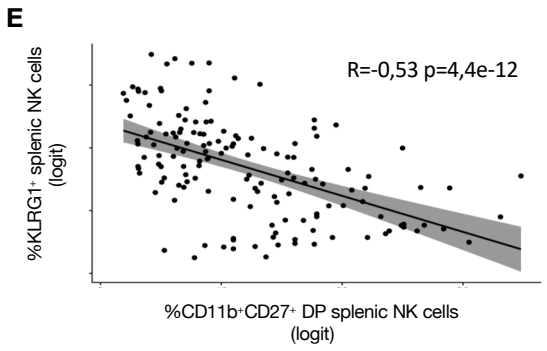


Fig. S3. Frequencies of DN (A), CD27SP (B), DP (C) and CD11bSP (D) splenic NK cells in the indicated lines. On the left panels, lines are arranged in descending order, while they are ordered as in Figure 2A on the right panels. B6 shown in red, CC001 and CC003 are color-coded as in Figure 2A. Data are represented as mean ± SD. (E) Correlation graph of the average percentage of CD11b+CD27+ DP against the percentage of KLRG1+ splenic NK cells across all CC strains. Regression line, Pearson's R coefficient and p-value are indicated.

Supplemental Figure 3. Dupont et al.

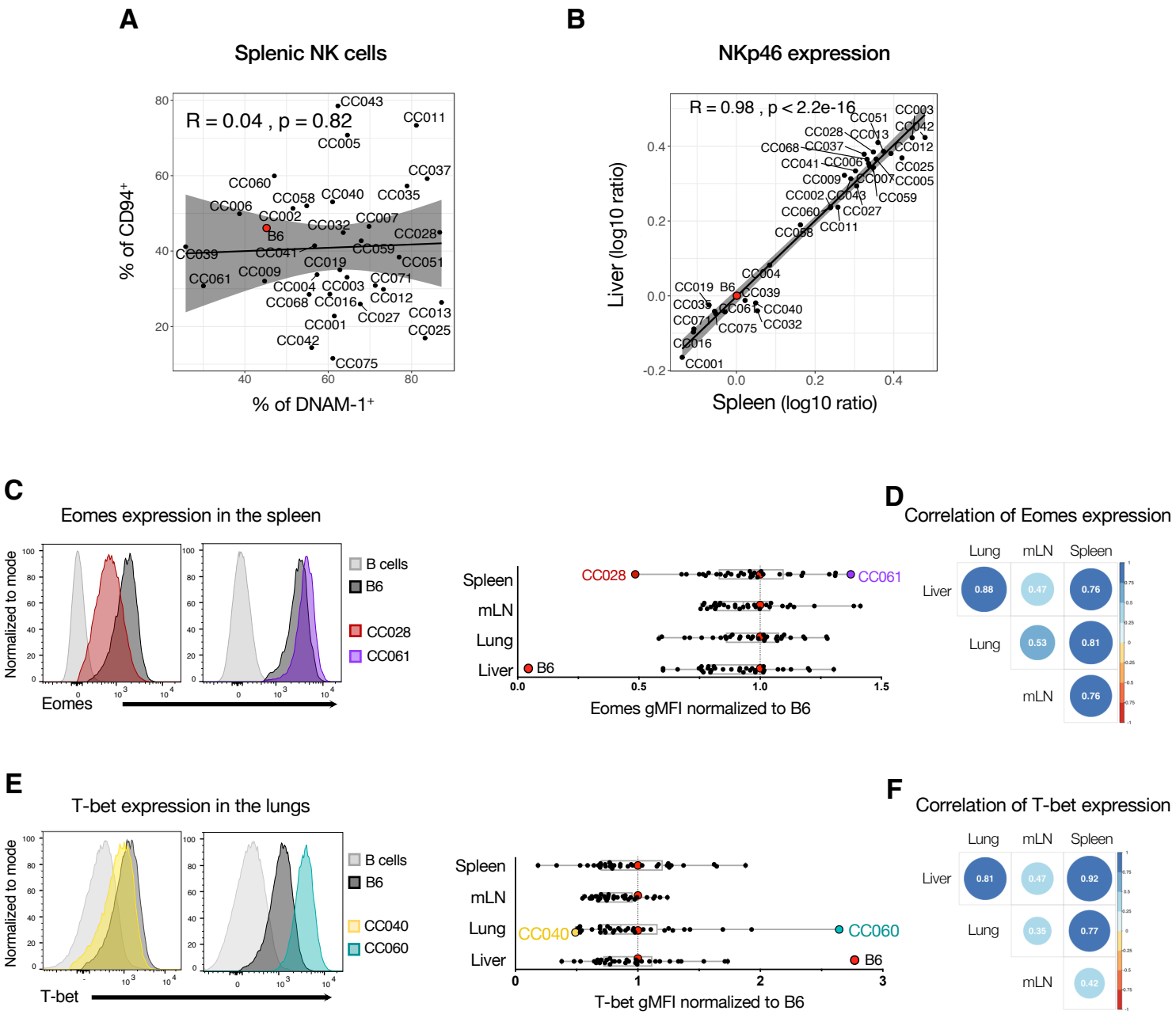


Fig. S4. Correlation graphs of (A) CD94⁺ and DNAM-1⁺ frequencies in splenic NK cells; (B) NKp46 expression (ratio to B6) in the liver against the expression in the spleen. B6 is shown in red. Regression line, Pearson's R coefficient and p-value appear on the graphs. One dot represents the average value of one line, with on average 5 mice per CC line (See Methods for more details); and 27 to 29 mice for B6. (C), (E) Left: representative histogram overlays of Eomes expression in splenic NK cells for B6, CC028 and CC061 (C) or T-bet expression in pulmonary NK cells from B6, CC040, and CC060 (E); Left: ratio of Eomes (C) or T-bet (E) expression in the spleen, mLN, lung and liver of all the CC lines. On the right graph, the extreme strains are identified and color-highlighted according to the colors in the histograms on the left, respectively. B6 is highlighted as a red dot. B cells (light grey) were used as negative population and overlaid on histograms. (D)(F) Respective correlation plots of average Eomes and T-bet expression in the liver, lung, mLN and spleen, across all CC strains.

Supplemental Figure 4. Dupont et al.

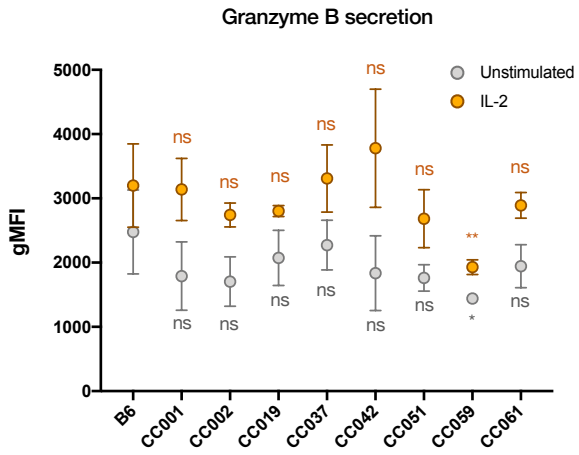


Fig. S5. Granzyme B gMFI in unstimulated or IL-2 stimulated splenic NK cells of 8 CC lines and B6. A two-way ANOVA was performed followed by multiple comparisons. Each CC line was compared to B6 within each group. B6: n = 6, each CC line: n = 3. ns = non-significant, *: p<0.05, **: p<0.01, ***: p<0.001, ****: p<0.0001.

Supplemental Figure 5. Dupont et al.

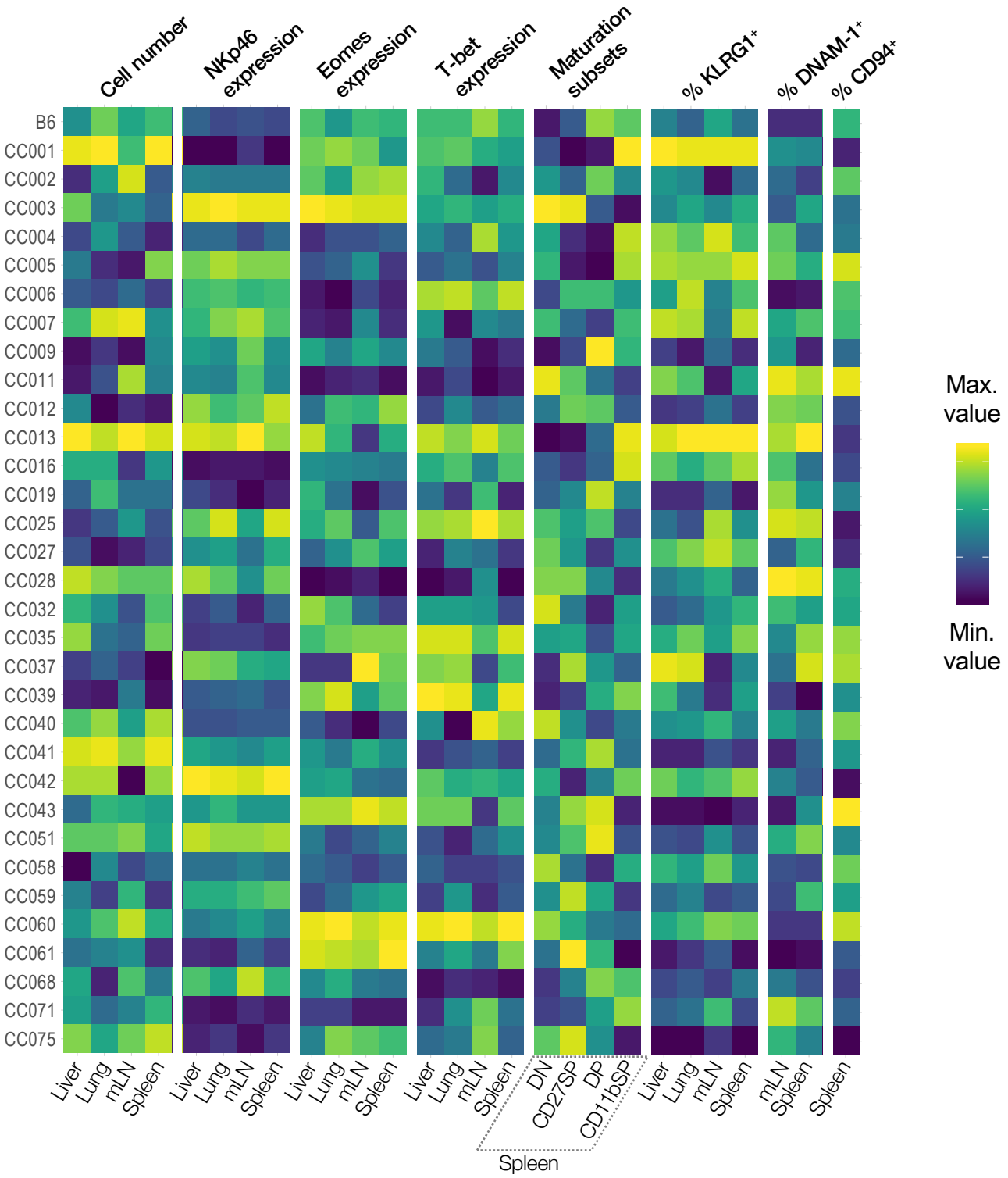


Fig. S6. Heatmap representation of the ranking of each CC strain for each parameter assessed in the indicated organs. Ranking was calculated based on the average value per CC strain.

Supplemental Figure 6. Dupont et al.

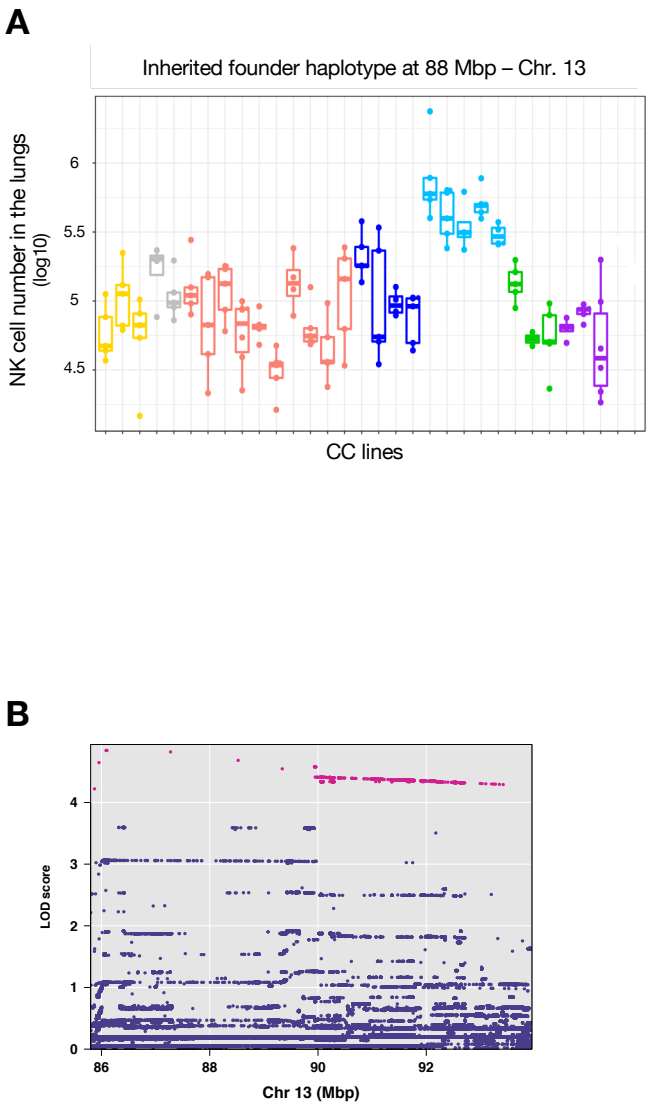
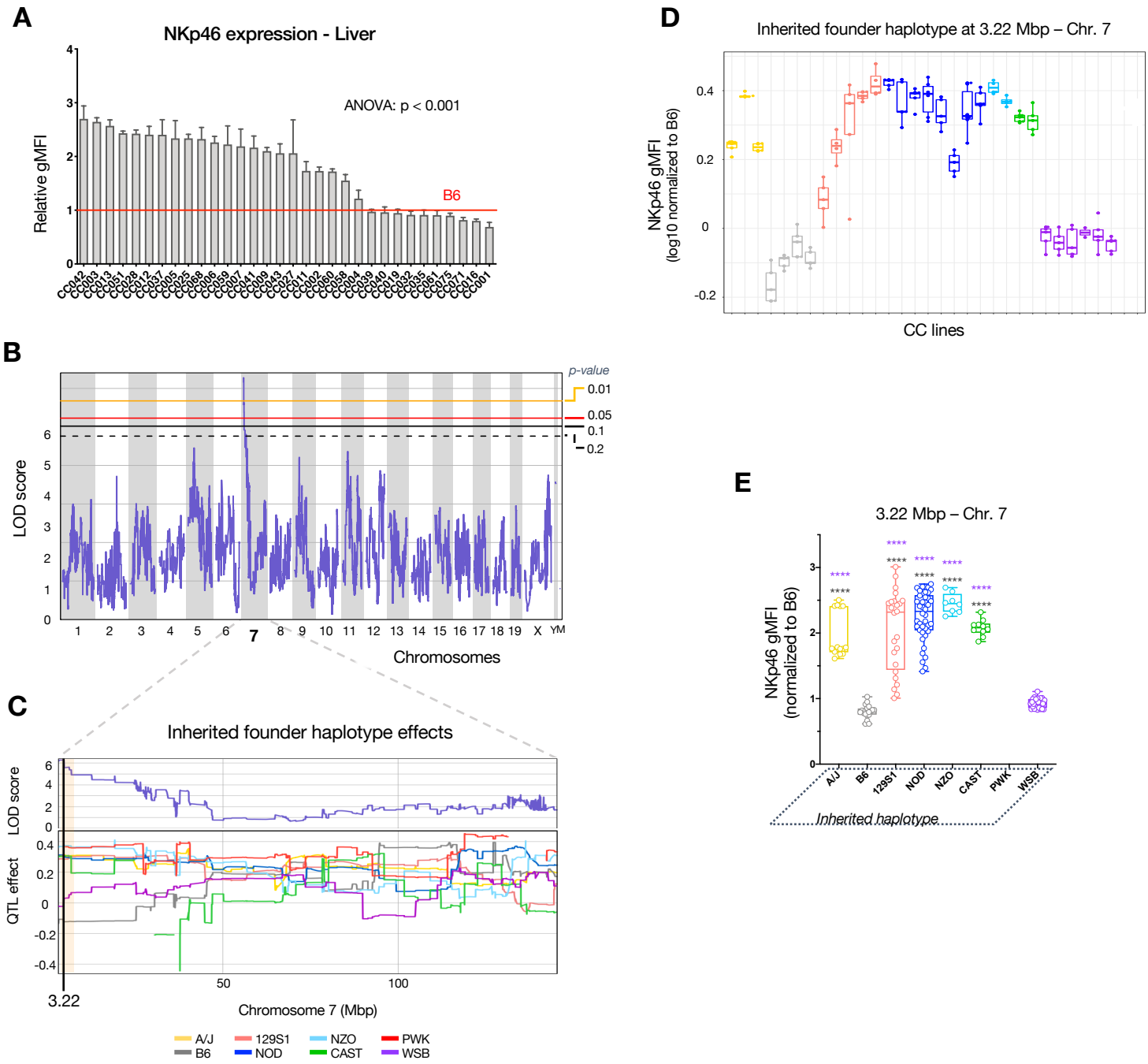
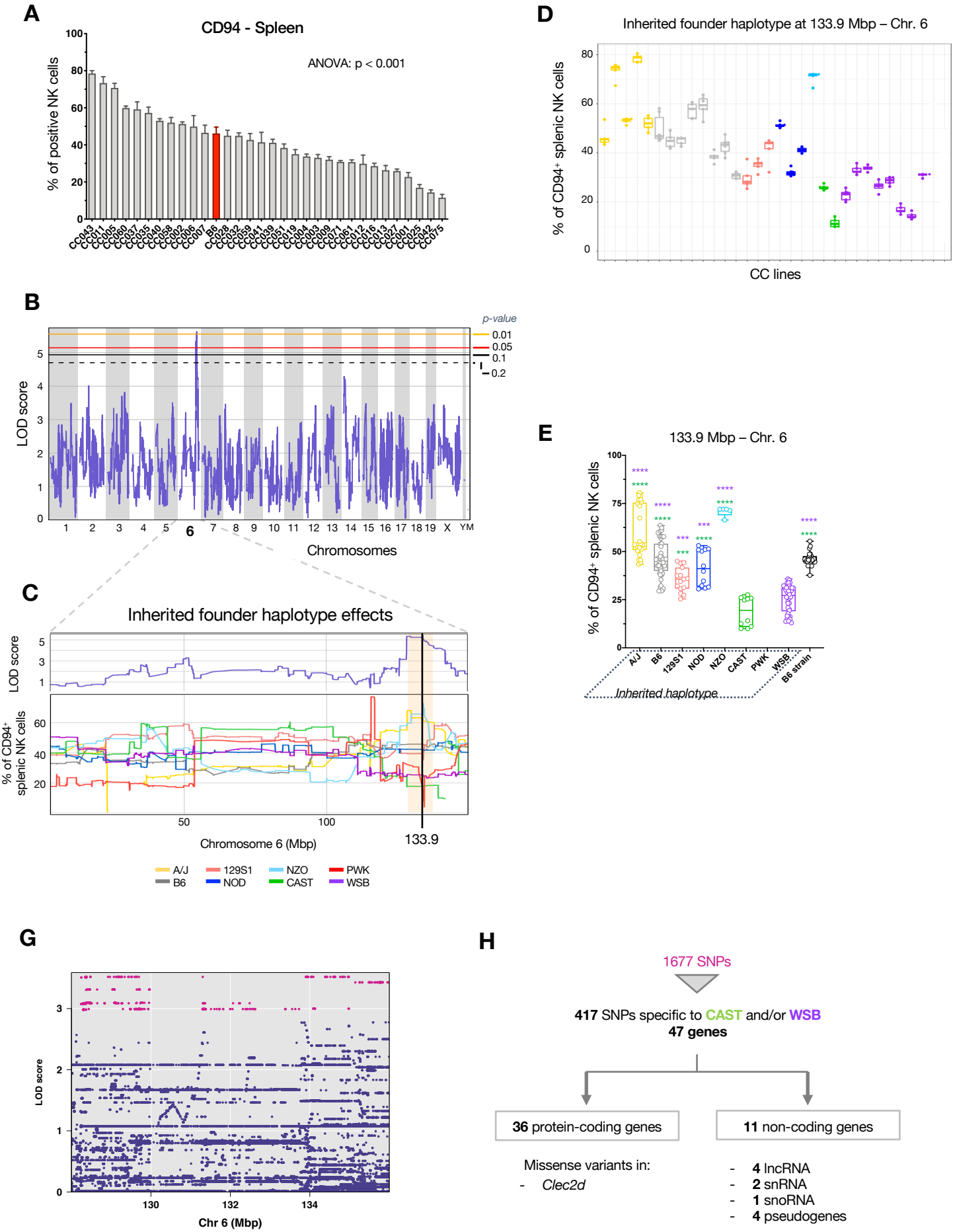


Fig. S7. (A) Number of NK cells in the lung for each CC line according to the haplotype they inherited, and for B6 at 88 Mbp on chromosome 13. (B) QTL-wide association analysis within *NK1* with significant SNPs appearing in pink.



Supplemental Figure 8. Dupont et al.

Fig. S8. (A) NKp46 expression on hepatic NK cells in 32 CC lines; the p-value corresponds to a one-way ANOVA. Data are represented as mean ± SD. (B) Genome scan for NKp46 gMFI (normalized to B6) in hepatic NK cells, with a significant peak on chromosome 7. Analysis was performed using the qtl2 package in R, and statistical significance thresholds were determined after 10'000 permutations. M: Mitochondria. (C) Top: LOD-score across chromosome 7; bottom: NKp46 gMFI (normalized to B6) in hepatic NK cells depending on the inherited haplotype on chromosome 7. Confidence interval of the QTL is highlighted in orange, and the exact position presented in (D) and (E) is represented by the black line. (D) NKp46 expression in hepatic NK cells of the CC lines according to the haplotype they inherited, and for each haplotype represented (E) at 3.22 Mbp on chromosome 7. A one-way ANOVA was performed followed by multiple comparisons between lines with Welch's correction. Results of the comparison of the A/J, 129S1, NOD, NZO, CAST groups to the B6 (grey) or WSB (violet) groups appear on the graph. ns = non-significant, *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, ****: $p<0.0001$. (F) QTL-wide association analysis within *NK2* with significant SNPs appearing in pink. (G) Schematic representation of the strategy used to identify candidate genes and general characteristics of these candidates.



Supplemental Figure 9. Dupont et al.

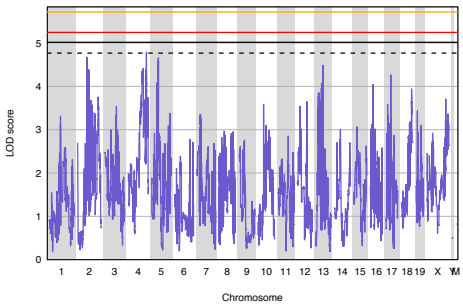
Fig. S9. (A) Percentage of CD94$^+$ splenic NK cells in 32 CC lines and B6; the p-value corresponds to a one-way ANOVA. Data are represented as mean ± SD. (B) Genome scan for percentage of CD94$^+$ splenic NK cells, with a significant peak on chromosome 6. Analysis was performed using the qtl2 package in R, and statistical significance thresholds were determined after 10'000 permutations. M: Mitochondria. (C) Top: LOD-score across chromosome 6; bottom: percentage of CD94$^+$ splenic NK cells depending on the inherited haplotype on chromosome 6. Confidence interval of the QTL is highlighted in orange, and the exact position presented in (D) and (E) is represented by the black line. (D) Percentage of CD94$^+$ splenic NK cells in the CC lines according to the haplotype they inherited, and for the B6 line and each haplotype represented (E), at 133.9 Mbp on chromosome 6. A one-way ANOVA was performed followed by multiple comparisons between lines with Welch's correction. Results of the comparison of the A/J, B6, 129S1, NOD, NZO, B6 strain groups to the CAST (green) or WSB (violet) groups appear on the graph. ns = non-significant, *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, ****: $p<0.0001$. (F) QTL-wide association analysis within *NK3* with significant SNPs appearing in pink. (G) Schematic representation of the strategy used to identify candidate genes and general characteristics of these candidates.

Suggestive QTLs

A

numbers mLN NK cells

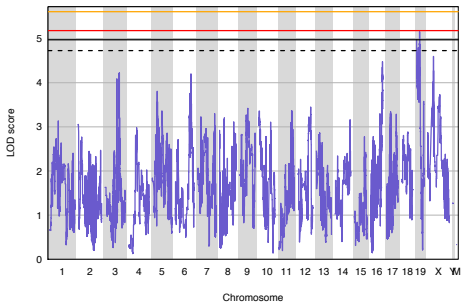
NK4



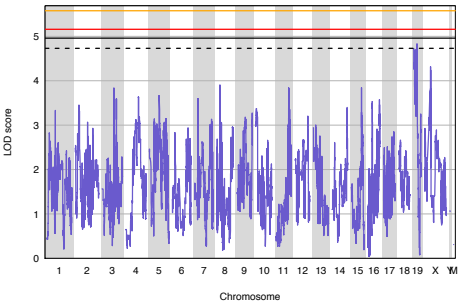
B

% KLRG1⁺ splenic NK cells

NK5



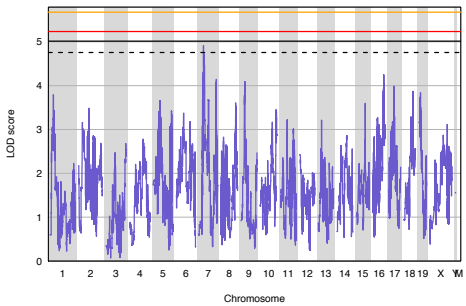
% KLRG1⁺ hepatic NK cells



C

% CD11bSP splenic NK cells

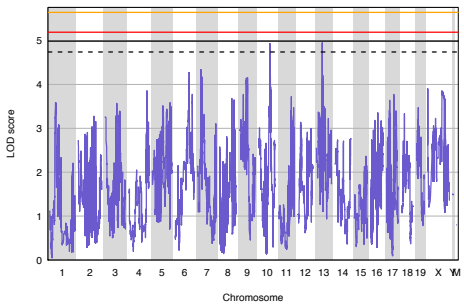
NK6



D

% CD27⁺CD11b⁺ splenic NK cells

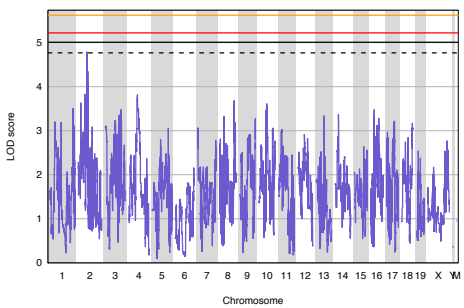
NK7 and *NK8*



E

% DNAM-1⁺ mLN NK cells

NK9



Supplemental Figure 10. Dupont et al.

Fig. S10. Genome scans with suggestive QTLS ($p < 0.2$) for the following traits: (A) *NK4*: absolute number of NK cells in the mLN, (B) *NK5*: frequencies of KLRG1$^+$ NK cells in the spleen (left) and liver (right), (C) *NK6*: frequency of CD11bSP splenic NK cells, (D) *NK7* and *NK8*: frequency of DP splenic NK cells, (E) *NK9*: frequency of DNAM-1$^+$ NK cells in the mLN. Analyses were performed using the qtl2 package in R, and statistical significance thresholds were determined after 10'000 permutations. M: Mitochondria.

Table S1. Total cell number (mean ± SD) calculated for each CC strain in the indicated organs.

Table S2. List of genes within the QTL associated to the variation of lung NK cell numbers and type of associated variants.

Table S3. List of genes within the QTL associated to the variation of NKp46 expression levels by hepatic NK cells and type of associated variants.

Table S4. List of genes within the QTL associated to the variation of the frequencies of $CD94^+$ splenic NK cells and type of associated variants.

Table S5. BLASTN results (query cover and percentage of identity) of *Klrd1* CC founders sequences alignment to C57BL/6J reference sequence.

	Total cell number							
	Liver		Lung		mLN		Spleen	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
B6	2077778	700595	1436923	581822	13754615	5981980	122622222	48038681
CC001	4160000	1820165	2640000	2041568	13700000	7552152	147700000	21147104
CC002	2628000	3524588	1156000	219386	16330000	9399574	104000000	50691469
CC003	2050000	997497	760000	438007	11840000	5292731	52460000	9775377
CC004	1240000	378153	1270000	514296	9220000	1704993	60900000	22968457
CC005	1944000	424358	638000	273715	5356000	2243141	154500000	42333497
CC006	4816667	3769571	1925000	1128605	28366667	10757447	144383333	47346612
CC007	2460000	1028591	3660000	2103093	14030000	11284259	36550000	15772207
CC009	766000	374273	362500	188149	2276000	640023	41600000	6940821
CC011	2075000	1272006	1760000	1016366	16460000	12120561	156900000	32304025
CC012	2260000	816548	994000	387143	8320000	4356260	76800000	17824141
CC013	6792500	1726333	1147500	411613	23862500	5703124	61325000	24705381
CC016	1668000	1107777	1007000	314038	5450000	1371131	35300000	17750352
CC019	1150000	331662	1282500	689841	9820000	5772521	55650000	15608091
CC025	887500	456837	948000	185796	7280000	4433622	59200000	25250050
CC027	1438000	466176	486000	188494	3200000	1151086	53700000	15303594
CC028	3556000	1150187	1432000	407578	7900000	3602777	65700000	19540343
CC032	1910000	559464	868000	209690	4050000	2119670	96800000	19185281
CC035	2550000	776015	683600	135110	5090000	3429898	118000000	18668155
CC037	2233333	1229108	1155000	958473	10900000	10110596	68833333	20764553
CC039	1800000	282843	1130000	183848	6500000	1838478	73500000	13435029
CC040	2040000	507937	2400000	435890	10342500	6289545	146100000	42857905
CC041	2540000	944987	1622000	474521	13194000	8335027	108100000	43839765
CC042	2175000	1085357	1592500	779987	1676000	960146	81820000	31856114
CC043	840000	415933	1156000	942168	9780000	2933769	79480000	49952748
CC051	3140000	1047855	1612000	347448	19820000	5614891	108400000	28494298
CC058	996000	375273	1134000	459924	6050000	3643830	152700000	60572890
CC059	3000000	894427	1102000	529166	26916667	15454891	134583333	48274648
CC060	2025000	776209	1150000	300000	17450000	12689760	96875000	41300474
CC061	2092000	515092	1507500	688970	10080000	3939162	102375000	50161489
CC068	1544000	631411	960000	450555	10600000	5951050	70100000	11221631
CC071	1066000	496266	710000	186011	3000000	1549193	81200000	43875677
CC075	1594000	658164	964000	385006	12670000	6735874	91400000	26168684

Table S1

		Variant type			
Gene name	**Biotype**	**Coding**	**Intron**	**UTR**	**Other**
4833422C13Rik	lncRNA	No	Yes	No	Yes
A830009L08Rik	lncRNA	No	Yes	No	Yes
Gm29540	lncRNA	No	No	No	Yes
Gm27656	miRNA	No	No	No	Yes
Gm20379	predicted coding gene	No	No	No	Yes
Acot12	protein coding	No	Yes	No	Yes
Atg10	protein coding	No	Yes	No	Yes
Atp6ap1l	protein coding	No	Yes	No	No
Cmya5	protein coding	No	Yes	No	No
Dhfr	protein coding	No	No	No	Yes
Fam151b	protein coding	No	Yes	No	Yes
Homer1	protein coding	No	Yes	No	No
Msh3	protein coding	Yes	Yes	Yes	Yes
Rasgrf2	protein coding	No	Yes	No	No
Rps23	protein coding	No	No	No	Yes
Serinc5	protein coding	No	Yes	Yes	Yes
Ssbp2	protein coding	No	Yes	Yes	Yes
Tmem167	protein coding	No	Yes	Yes	Yes
Xrcc4	protein coding	Yes	Yes	No	Yes
Zfyve16	protein coding	Yes	Yes	Yes	Yes
Gm17450	pseudogene	No	No	No	Yes
Gm22778	snoRNA	No	No	No	Yes
Gm24507	snoRNA	No	No	No	Yes
Gm36966	unclassified	No	No	No	Yes
Gm37054	unclassified	No	No	No	Yes
Gm37708	unclassified	No	No	No	Yes
Gm24597	unclassified non-coding RNA	No	No	No	Yes
	Total	11.1% (3/27)	55.5% (15/27)	18.5% (5/27)	85.2% (23/27)

Table S2

		Variant type			
Gene name	**Biotype**	**Coding**	**Intron**	**UTR**	**Other**
9430041J12Rik	lncRNA	No	Yes	No	No
Gm15510	lncRNA	No	Yes	No	Yes
Brsk1	protein coding	No	Yes	No	Yes
Cacng6	protein coding	No	No	No	Yes
Isoc2a	protein coding	No	No	Yes	Yes
Isoc2b	protein coding	No	Yes	No	Yes
Lair1	protein coding	No	Yes	No	Yes
Shisa7	protein coding	No	Yes	No	Yes
Tarm1	protein coding	No	Yes	Yes	Yes
Tmem150b	protein coding	No	No	Yes	Yes
Zfp580	protein coding	No	No	No	Yes
Zfp784	protein coding	No	No	No	Yes
Gm7363	pseudogene	No	No	No	Yes
Gm23741	snRNA	No	No	No	Yes
	Total	0% (0/14)	50% (7/14)	21.5% (3/14)	93% (13/14)

Table S3

		Variant type			
Gene name	**Biotype**	**Coding**	**Intron**	**UTR**	**Other**
2310001H17Rik	lncRNA	No	Yes	No	Yes
Gm10069	lncRNA	No	Yes	No	Yes
Gm15987	lncRNA	No	Yes	No	No
Gm26770	lncRNA	No	No	No	Yes
BC049715	Protein coding	No	Yes	No	Yes
Bcl2l14	Protein coding	No	Yes	Yes	Yes
Borcs5	Protein coding	No	No	No	Yes
Cd69	Protein coding	No	No	No	Yes
Clec1a	Protein coding	No	Yes	No	No
Clec2d	Protein coding	Yes	No	No	Yes
Clec2g	Protein coding	No	Yes	No	Yes
Crebl2	Protein coding	No	Yes	No	Yes
Dusp16	Protein coding	No	Yes	No	Yes
E330021D16Rik	Protein coding	No	Yes	Yes	Yes
Eif4a3l1	Protein coding	No	No	No	Yes
Etv6	Protein coding	No	Yes	No	No
Foxm1	Protein coding	No	Yes	Yes	Yes
Gpr19	Protein coding	No	Yes	No	No
Grin2b	Protein coding	No	Yes	No	No
Gsg1	Protein coding	No	No	No	Yes
Gucy2c	Protein coding	No	Yes	No	Yes
H2aj	Protein coding	No	No	No	Yes
H4f16	Protein coding	No	No	Yes	Yes
Hebp1	Protein coding	No	Yes	No	No
Klrc3	Protein coding	No	Yes	No	No
Klrd1	Protein coding	No	No	No	Yes
Klre1	Protein coding	No	No	No	Yes
Klri1	Protein coding	No	No	No	Yes
Klrk1	Protein coding	No	Yes	No	Yes
Lrp6	Protein coding	No	Yes	No	No
Mansc1	Protein coding	No	Yes	No	Yes
Olr1	Protein coding	No	Yes	No	Yes
Plbd1	Protein coding	No	Yes	No	Yes
Rhno1	Protein coding	No	Yes	Yes	Yes
Smco3	Protein coding	No	No	No	Yes
Styk1	Protein coding	No	Yes	No	No
Tead4	Protein coding	No	Yes	No	No
Tex52	Protein coding	No	No	No	Yes
Tmem52b	Protein coding	No	No	Yes	Yes
Tulp3	Protein coding	No	Yes	No	Yes
Gm15476	pseudogene	No	No	No	Yes
Gm6728	pseudogene	No	No	No	Yes
Klrb1-ps1	pseudogene	No	No	No	Yes
Rpl36a-ps3	pseudogene	No	No	No	Yes
Gm22881	snoRNA	No	No	No	Yes
Gm25882	snRNA	No	No	No	Yes
Gm26160	snRNA	No	No	No	Yes
	Total	2.1% (1/47)	55.3% (26/47)	12.7% (6/47)	78.7% (37/47)

Table S4

Klrd1 reference sequence: C57BL/6J

BLASTN results

Founder Strains	Query cover	Percentage of identity
A/J	62%	99.15%
129S1/SvImJ	62%	99.96%
NOD/LtJ	62%	99.89%
NZO/HiLtJ	95%	98.91%
CAST/EiJ	62%	99.25%
PWK/PhJ	95%	98.20%
WSB/EiJ	93%	99.12%

Table S5