Published in final edited form as:
Sci Transl Med. 2013 October 23; 5(208): 208ra145. doi:10.1126/scitranslmed.3006702.

Genetic and environmental determinants of human NK cell diversity revealed by mass cytometry

Amir Horowitz[1,2,3], **Dara M. Strauss-Albee**[3,4], **Michael Leipold**[2], **Jessica Kubo**[4], **Neda Nemat-Gorgani**[1], **Ozge C. Dogan**[4], **Cornelia L. Dekker**[5], **Sally Mackey**[5], **Holden Maecker**[2], **Gary E. Swan**[6], **Mark M. Davis**[2,3], **Paul J. Norman**[1], **Lisbeth A. Guethlein**[1], **Manisha Desai**[4], **Peter Parham**[1,2,3], and **Catherine A. Blish**[3,4,*]

[1]Department of Structural Biology, Stanford University School of Medicine, Stanford, CA, 94305, USA

[2]Department of Microbiology & Immunology, Stanford University School of Medicine, Stanford, CA, 94305, USA

[3]Stanford Immunology, Stanford University School of Medicine, Stanford, CA, 94305, USA

[4]Department of Medicine, Stanford University School of Medicine, Stanford, CA, 94305, USA

[5]Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, 94305, USA

[6]Center for Health Sciences, SRI International, Menlo Park, CA, 94025, USA

Abstract

Natural Killer (NK) cells play critical roles in immune defense and reproduction, yet remain the most poorly understood major lymphocyte population. Because their activation is controlled by a variety of combinatorially expressed activating and inhibitory receptors, NK cell diversity and function are closely linked. To provide an unprecedented understanding of NK cell repertoire diversity, we used mass cytometry to simultaneously analyze 35 parameters, including 28 NK cell receptors, on peripheral blood NK cells from five sets of monozygotic twins and twelve unrelated donors of defined HLA and killer cell immunoglobulin-like receptor (KIR) genotype. This analysis revealed a remarkable degree of NK cell diversity, with an estimated 6,000-30,000 phenotypic populations within an individual and >100,000 phenotypes in this population. Genetics largely determined inhibitory receptor expression, whereas activation receptor expression was heavily environmentally influenced. Therefore, NK cells may maintain self-tolerance through strictly regulated expression of inhibitory receptors, while using adaptable expression patterns of activating and costimulatory receptors to respond to pathogens and tumors. These findings further suggest the possibility that discrete NK cell subpopulations could be harnessed for immunotherapeutic strategies in the settings of infection, reproduction, and transplantation.

[*]Corresponding Author Catherine Blish, MD, PhD Division of Infectious Diseases and Geographic Medicine Department of Medicine 300 Pasteur Dr., Lane Building, L-134 Stanford, CA 94305-5107 Phone (206) 725-5132 Fax (206) 723-3474 cblish@stanford.edu.

Author contributions. AH and CAB designed and performed the experiments, analyzed the data, and wrote the manuscript; DMS-A analyzed the data, performed ecological analyses, and wrote the manuscript; ML, NN-G, OCD, HM, PJN, LAG performed experiments and contributed to the writing of the manuscript; CAB, JK, and MD performed statistical analyses; GES referred twin pairs for recruitment, CLD and SM conducted the clinical trial to provide samples; GES, CLD and MMD contributed to the writing of the manuscript; AH, PP, and CAB conceived the experiments and wrote the manuscript.

Competing interests
None of the authors have any conflicts of interest to disclose.

Data and materials availability
All data are deposited in ImmPort.

INTRODUCTION

Natural killer (NK) cells were discovered in the late 1970s in mice (1) and humans (2-4). They were first characterized by their ability to rapidly kill tumor cells, and later virus-infected cells (3, 5). Since this initial characterization, the nuanced complexity of NK cell phenotypes and function has been increasingly appreciated (6). The diversity of the NK cell repertoire is determined by the expression of an array of germline-encoded activating and inhibitory receptors (6-8). These receptors interact with MHC class I and class I-like molecules, co-stimulatory ligands, stress-related molecules, and cytokines (9-11). The combinatorial expression of this multitude of receptors on an NK cell results in an abundance of mathematically feasible subpopulations, but this diversity has never been captured in detail.

Early studies using NK cell clones and mRNA expression data (12-14), more recent work incorporating monoclonal antibodies against a variety of NK cell receptors (15, 16), and bone marrow transplant data showing the similarity of recipient to donor inhibitory repertoires (17) all demonstrate that host genetics are a critical determinant of inhibitory receptor expression patterns. Additional non-genetically encoded determinants of diversity may influence activating receptor expression, including NK cell education, epigenetics, epistatic interactions between KIR genes, and viral infections, particularly cytomegalovirus (18-25). These observations emphasize the need for a better understanding of the relationship between genotype, phenotype, and function in the formation of the NK cell repertoire.

Advances in fluorescence cytometry have laid the groundwork for understanding the NK cell repertoire, but the spectral limitations of fluorescence have prevented exploration of its full extent. The recent development of mass cytometry, also known as cytometry by time-of-flight, or CyTOF, provides an opportunity to overcome these limitations. Mass cytometry uses rare metal isotope-conjugated antibodies to simultaneously detect up to 40 cellular markers. Since the metal isotopes are not naturally found in biological systems and have distinct time-of-flight profiles, background is nearly undetectable and the need for compensation across channels is eliminated. Mass cytometry has been used to profile human hematopoiesis (26) and to demonstrate the tremendous diversity within the $CD8^+$ T cell compartment (27).

To provide a framework to better understand the NK cell repertoire, we used mass cytometry in combination with high-resolution genotyping to evaluate the phenotypic heterogeneity of peripheral blood NK cells in 22 healthy individuals, including 5 sets of monozygotic twins. Our study shows a remarkable breadth and diversity in the human NK cell repertoire and the unique role of genetics and the environment in its formation and maintenance.

RESULTS

Mass cytometry for dissecting the phenotypic diversity of NK cells

To better define the phenotypic diversity of human peripheral blood NK cells, we designed a mass cytometry panel of 35 monoclonal antibodies recognizing lineage markers and natural killer receptors (Table S1). As this was the first mass cytometry panel specific for NK cell markers, we compared results between mass cytometry and fluorescence cytometry. All antibodies successfully and comparably distinguished cell populations on both platforms, with representative plots shown in Fig. S1A and S1B. Staining with isotype-control antibodies revealed minimal non-specific background by mass cytometry (Fig. S1C). We also compared NK cells stained with or without pre-treatment with a human Fc receptor

blocking solution, and observed no significant differences in antibody binding (Fig. S1D). Adding to the versatility and applicability of this approach, very similar NK cell staining profiles were observed for fresh and cryopreserved PBMC samples (Fig. S1E and S1F). NK cells identified with a serial gating strategy demonstrated a range of surface receptor expression levels in different individuals (Fig. S2).

Overall diversity of the NK cell receptor phenotype

With this mass cytometry platform, we interrogated the diversity of the NK cell repertoire in a cohort consisting of 12 healthy, unrelated individuals and 5 sets of monozygotic twins of known HLA and KIR genotype (Fig. 1 and S3). Following identification of NK cells with lineage markers, Boolean analysis was used to classify cells as positive or negative for the expression of each of 28 NK cell receptors (Fig. S4). This allowed for the assessment of 2^{28}, or 268,435,456, unique single cell combinations of these receptors. The results revealed a much greater heterogeneity in NK cell phenotypes than has been described previously (Fig. 1). No single phenotype accounted for more than 7% of the total NK cells and only 14 of 28 markers tested were expressed by NK cells exhibiting the top 50 most frequent phenotypes (Fig. 1A). Furthermore, the subsets comprising the top 50 phenotypes accounted for an average of just 15% (range 7-24%; SD 4.8%) of an individual's NK cells.

Although the conserved activating receptors NKG2D and NKp46 are usually described as being expressed on most human NK cells (28, 29), they were expressed on just 28 and 25 of the top 50 phenotypes, respectively. This low expression is partially a result of our Boolean gating strategy, in which only NK cells that were relatively bright for NKG2D and NKp46 were considered positive (Figure S4). Thus, we may have underestimated the frequency of these markers, particularly for cells that express low levels. Nonetheless, these results were concordant with the frequencies of NKG2D and NKp46 in total NK cell populations in this study, where NKG2D and NKp46 were highly expressed on an average of 48% (range 29% – 67%) and 47% (range 28% – 71%) of NK cells, respectively, and with additional studies in which <100% of NK cells expressed these markers (28, 30-33). The conserved inhibitory receptor NKG2A was more prevalent, found on 40 of the top 50 phenotypes and always co-expressed with CD94, its partner polypeptide. In contrast with NKG2A, the highly variable KIRs contributed to relatively few phenotypes. KIR2DL1, KIR2DL5, and KIR3DL1 were not present in the 50 most frequent phenotypes. KIR2DL2/L3/S2, the most prevalent KIR group, was expressed on an average of 30% of an individual's total NK cells, but on only four of the 50 most frequent subpopulations, which together represented an average of only 3.2% (range 0.8 – 6.8%) of KIR2DL2/L3/S2-expressing cells. Thus, the remaining >93% were distributed across a range of phenotypes, each represented by very few cells. These data reveal that a diverse spectrum of low-abundance NK cell phenotypes express inhibitory receptors.

To assess the genetic influences on the phenotypes of NK cells, we compared the frequencies of the top 50 NK cell phenotypes within paired monozygotic twins (Fig. 1B) and pairs of unrelated donors (Fig. 1C). Twin pairs exhibited modest concordance in the frequencies of NK cell subpopulations (Spearman's $\rho = 0.65$), though such concordance was diminished in pairs of unrelated donors (Spearman's $\rho = 0.36$). These data suggest that genetic differences between individuals may influence the repertoire of NK cell surface phenotypes.

NK cell receptor diversity in the expression of inhibitory receptors for HLA class I

We next restricted the Boolean analysis to the six inhibitory NK cell receptors KIR2DL1, KIR2DL2/L3/S2, KIR2DL5, KIR3DL1, CD94/NKG2A and LILRB1, resulting in 64 possible combinatorial cell surface phenotypes. Of these possibilities, the most frequently

detected phenotypes in each individual expressed none of the six inhibitory receptors or only NKG2A (Fig. 2A). These two subsets accounted for an average of 54% (range 41 – 72%) of an individual's NK cells. Most remaining NK cells expressed 1-3 inhibitory receptors. These results are consistent with those obtained using fluorescence cytometry (16, 19, 34, 35).

In analyses parallel to those described in Figs. 1B and 1C, we compared the expression frequency of the inhibitory receptor combinations between paired monozygotic twins (Fig. 2B) and between pairs of unrelated individuals (Fig. 2C). Concordance was extremely high within the pairs of monozygotic twins (Spearman's $\rho = 0.87$), and less dramatic in unrelated individuals (Spearman's $\rho = 0.77$). These results suggest that host genetics determine the inhibitory receptor repertoire, but only weakly affect the expression of other NK cell receptors (Fig. 1C). Together, these data indicate that while the inhibitory repertoire appears to be largely controlled by genetics, environmental and/or stochastic influences play a leading role in determining the overall repertoire of NK cell phenotypes.

Assessing the NK cell repertoire with clustering analysis

Boolean analysis of the mass cytometry data resulted in thousands of distinct NK cell phenotypes defined by combinatorial receptor expression. However, due to the abundance of low frequency populations, we used a clustering algorithm called spanning-tree progression analysis of density-normalized events (SPADE) (26, 36). In SPADE analyses of PBMCs, nodes of phenotypically similar cell clusters corresponding to NK cells, $CD4^+$ T cells, $CD8^+$ T cells, B cells and monocytes/macrophages were readily identified and distinguished (Fig. S5).

Consistent with the Boolean analysis, the distribution of inhibitory receptor expression on NK cells revealed by SPADE was more similar for monozygotic twins (Figs. 3A and S6) than for unrelated individuals (Figs. 3B and S7). For instance, KIR2DL1 was observed on similar subpopulations in twins but not in unrelated individuals. More diversity was observed in the intensity and frequency of expression of activating receptors, including NKG2D, which was expressed at varying intensities and frequencies among different clusters between individuals, even in twins. Additionally, HNK-003 and HNK-007 were both genotypically and phenotypically divergent for KIR2DS4, consistent with HNK-007 lacking a functional copy of the KIR2DS4 gene (Figs. 3B and S3). However, even the genotypically identical Twins 5a and 5b showed differences in KIR2DS4-expressing nodes, with 28 KIRS2DS4-expressing nodes in common but 10 with differential expression. These data further illustrate the impact of environmental factors on the total NK cell repertoire.

A comprehensive phenotype for $CD57^+NKG2C^+$ 'memory-like' NK cells

SPADE analysis showed that five of the unrelated donors (HNK-002, HNK-004, HNK-005, HNK-007, HNK-008) and one pair of twins (Twin-5a and 5b) had large NK cell subpopulations co-expressing CD57 and NKG2C (Fig. 4A) that could represent memory-like NK cells generated in response to CMV infection (24, 25, 37, 38). All of these subjects tested had positive serology for CMV except for HNK-008 (Figure S3). The remaining CMV-seropositive subjects (HNK-001, HNK-012, Twin-4a, Twin4b) also had NKG2C +CD57+ NK cell subpopulations, though these populations were less prominent. The $CD57^+NKG2C^+$ cells expressed CD94 and CD16, but had low expression of CD122 and KIR3DL1, and variable patterns of expression of KIR2DL1, KIR2DL3, NKG2D, NKp30, NKp46, and CD8 (Fig. 4, B and C). In the donors with only the HLA-C1 epitope, $NKG2C^+CD57^+$ NK cells preferentially expressed KIR2DL3, the inhibitory HLA-C1 receptor, while donors with the HLA-C2 epitope preferentially expressed KIR2DL1 (Fig. 4).

NK cell receptor co-expression patterns

We next asked whether any single NK cell receptors were more likely to be expressed with other single receptors on the same NK cell. We therefore calculated the Spearman's rank correlation of each possible receptor pair on each NK cell from all 22 donors (Fig. 5A). Weak negative associations between the expression of CD57 and NKp46, CD122, CD94, and NKG2A and weak positive associations between the expression of NKG2A, CD94, and CD56 were the most dominant features.

Given the relative paucity of strong receptor co-expression patterns, we performed hierarchical clustering of NK cell populations based on surface receptor expression patterns to better understand NK cell population structure (Fig. 5B). The markers CD16, CD56, CD57, CD94, and NKG2A were distinct in their location at the root of the tree, while most other NK cell receptors contributed equally to the overall population structure (Fig. 5B). Further, principal component analyses indicated that NK cell populations expressing CD16, CD56, CD57, NKG2A, and CD94 were distinct and formed separate clusters (Fig. 5C). Overall, these results are consistent with a stochastic expression of most NK cell receptors, with less mature $NKG2A^{+}CD94^{+}$ and more mature $CD16^{+}CD57^{+}$ cells representing the only major distinction.

Quantifying the diversity of the NK cell repertoire

Having observed a stochastic expression of most NK receptors, leading to a broad range of possible combinatorial phenotypes on a single cell, we next sought to quantify this high level of diversity. We reasoned that our system of NK cell phenotypes resembled a collection of species within a habitat, and adapted established ecological methods. We first calculated the inverse Simpson diversity index, which measures the probability that two randomly selected organisms from a population will be of the same species (39-41). This index accounts for both the total abundance and the distribution of species within a population. We first considered each NK cell as an individual organism, each NK cell phenotype as a species, and the group of phenotypes expressing each cell-surface marker as a population. Inverse Simpson indices for each population illustrate the contribution of each marker's cellular distribution to a donor's total phenotypic diversity (Fig. 6A).

Receptors that were highly expressed on NK cells, such as CD56, CD16, CD94, and 2B4, had high average inverse Simpson indices, indicating their contribution to numerous NK cell phenotypes (Fig. 6A). Strikingly, we also observed a broad range within the cohort in the diversity of the NK cell populations expressing these and other receptors. For example, though the average inverse Simpson index of CD56-expressing populations was 659, the minimum was 135 and the maximum was 1017, yielding a range of 882. We expected highly polymorphic KIRs to exhibit broad ranges of inverse Simpson indices, which was true for KIR3DL1 (range 844), KIR2DS4 (range 582), and KIR2DL1 (range 460). However, many non-polymorphic or minimally polymorphic receptors such as CD16 (range 938) and 2B4 (range 691) showed similar or even greater ranges of inverse Simpson indices. Thus, we observed a high degree of variability among donors in the diversity of the NK cell populations on which specific receptors were expressed, which was not wholly dependent on allelic polymorphism.

We next asked whether host genetics played any role in determining this variability by calculating the correlation between the inverse Simpson indices for each receptor in each pair of twins vs. all unrelated individuals. Despite the broad range in receptor diversity across individuals, the inverse Simpson indices for most receptors within twin pairs were remarkably similar, with high intraclass correlation coefficients (>0.6) for all receptors except CD56, 2B4, CD127, CD117, LILRB1, HLA-DR, and CCR7 (Fig. 6B). This suggests

host genetics may be a determinant of the range of NK cell phenotypes on which a given receptor is expressed, with factors in addition to receptor polymorphisms exerting considerable influence on each receptor's diversity.

Next, to understand the range of NK cell diversity in a human population, we calculated the inverse Simpson indices for each individual donor, rather than each receptor (Fig. 6C). Eighteen of the twenty-two individuals had total inverse Simpson indices in the relatively restricted range of 350 to 900, suggestive of an optimal level of human NK cell diversity.

We next asked how many total NK cell phenotypes were likely to be present in each individual. We used the Chao 1 non-parametric species estimator, which projects the total number of species present in a sample by using the abundance of rare species as a predictor of undiscovered species (40). This method predicted between 6,000 and 30,000 distinct NK cell phenotypes within each individual, with a median of 15,000 (Fig. 6D). In this analysis, twins showed a low intraclass correlation of 0.278 (Fig. 6B), suggesting that the combinatorial assortment of receptors in distinct NK cell phenotypes may be more a function of environmental than genetic influences.

Calculating the expanse of the NK cell repertoire

Finally, we calculated how many NK cell phenotypes were likely to be present at a population level. As our experimental analyses of 3,500-35,000 NK cells per donor were unlikely to have captured the total NK cell diversity, we generated a sample-based rarefaction curve to provide an estimate of the total number of NK cell phenotypes as a function of the number of phenotypes sampled (Fig. 7A). This analysis estimated the expansiveness of this cohort's repertoire at 124,651 unique NK cell phenotypes. We also predicted the total number of NK cell species on a population level using three other independent and well-established ecological methods (40). Together, despite their variable estimation approaches and input parameters, all of these calculations yielded remarkably similar estimates of 108,000 to 125,000 total NK cell phenotypes (Fig. 7B).

DISCUSSION

Using multiparametric mass cytometry to examine expression of 28 cell-surface markers, we achieved an incisive dissection of human NK cell repertoires. A vast phenotypic diversity was uncovered, which is influenced by the genetic differences between individual humans as well as by the environmental differences they experience. Genetic differences strongly influence the combinatorial expression patterns of the inhibitory receptors that recognize HLA class I. This repertoire of inhibitory receptors, which are required for maintenance of self-tolerance as well as for NK cell "education" or "licensing", is overlaid by an environmentally influenced diversity in the expression of activating and costimulatory receptors. This complementary combination results in NK cell repertoires that comprise between 6,000 and 30,000 phenotypically distinguishable subpopulations. Beyond such diversity within an individual is further variation between individuals. More than 100,000 NK cell subpopulations were distinguished in the small 'population' of 22 people we studied. These data reveal a potential mechanism by which NK cells could maintain self-tolerance through strictly regulated expression of inhibitory receptors, while using more malleable and adaptable expression patterns of activating and costimulatory receptors to respond to infections and cancers.

Comparing adult monozygotic twins distinguished the effects of genetic and environmental factors. The NK cells from twins had very similar patterns of expression of inhibitory HLA class I receptors, consistent with previous comparisons of unrelated individuals showing that the inhibitory receptor repertoire is influenced by variability in the KIR and HLA class I

genes (1, 12, 13, 16, 19, 34). Although very similar overall, the inhibitory receptor repertoires of twins exhibited some differences, which might be explained by environmental influences, including well-established epigenetic alterations of the KIR genes(19-21).

We observed extensive diversity within populations of cells expressing each combination of inhibitory receptors. For instance, the only other marker consistently co-expressed on NKG2A-expressing cells is CD94, its partner polypeptide. Even co-expression of markers believed to be expressed on all or most NK cells, such as NKp46 and NKG2D, were highly variable. Thus, NKG2A is not expressed by one or a few subpopulations of NK cells, but by thousands. Since both NKG2A and its ligand, HLA-E, are highly conserved (7, 42), this diversity may provide an additional degree of flexibility to the repertoire.

Associated with the paucity of pairwise combinations of receptors in the overall NK cell population (Fig. 5), is our finding that the expression pattern for most of the receptors (including inhibitory KIRs, the C-type lectin receptors, and natural cytotoxicity receptors) is determined solely by the genes encoding the receptor (Fig. 6). This suggests a pre-determined level for both the expression and distribution of these receptors by NK cells. Several other receptors - HLA-DR, CCR7, CD127, CD117, 2B4, and LILRB1 – had diversity patterns more suggestive of environmental influence. These receptors are known to vary in response to activation (HLA-DR and CCR7 (9, 43)), differentiation (CD117 (44)), infection (2B4, LILRB1 (15, 45)), and pregnancy (LILRB1 (46)).

Variation in a receptor's expression within an individual could be explained by genetic polymorphisms that affect expression level (19-21). In particular, KIR polymorphisms greatly influence the avidity and specificity of these interactions, leading to a highly variable repertoire (18, 47). Bolstering this argument is the broad range of expression by KIR3DL1, the most diverse of all KIR (48), and KIR2DS4 (49, 50). CD16, a less polymorphic receptor, is the most diverse of the NK cell receptors in its expression across cellular subpopulations. One interpretation of this finding is that a diversity of NK cells facilitates the critical function of CD16 in mediating antibody-dependent cellular cytotoxicity (ADCC), a bridge between innate and adaptive immunity. Alternatively, the diversity of CD16 could reflect its role as the only receptor demonstrated to be sufficient for NK cell activation (51).

"Null" cells, which express none of the cell-surface markers other than CD56, are the most frequent NK cell subpopulation. They could represent a transitional state capable of upregulating activating or inhibitory receptors. If so, they may be less likely to upregulate expression of inhibitory receptors, since the inhibitory receptor expression patterns are extremely stable over time (52). However, even cells with no inhibitory receptors may play a unique functional role. While these "uneducated" cells lacking receptors for self-HLA are hyporesponsive to HLA-deficient target cells (53), uneducated cells can be effective in ADCC (54).

Our results suggest that the human NK cell repertoire is anchored by the combination of a mature CD16/CD57-expressing population, and a less mature, NKG2A/CD94-expressing population. Stochastic assortment of additional receptors creates tremendous additional diversity within this framework. Expansions of $NKG2C^{+}CD57^{+}$ NK cells that preferentially expressed KIR reactive with self-HLA-C molecules were identified in several healthy, mostly CMV-seropositive, donors. These memory-like cells had a mature phenotype, as previously reported (23, 55) . Despite this subset's presumable expansion in response to a single pathogen, CMV (18), these memory-like NK cells had highly variable expression levels of other markers, further underlining the overall diversity of the NK cell population.

Several limitations of the study should be noted. First, our sample size was modest and serial sampling was not performed, which limited our ability to compare NK cell phenotypes

between individuals of similar KIR/HLA genotype. In addition, the number of NK cells sampled per donor was between ~3,000 and ~20,000, which is well below the theoretical number of 268,435,456 potential phenotypes, limiting our sampling of rare phenotypes. Finally, by using a Boolean gating strategy in which we defined cell populations as positive for a given receptor only if high levels were expressed, we may have underestimated the frequency of some receptors.

NK cells are quick, versatile lymphocytes that function in innate immunity, adaptive immunity and reproduction. In performing these vital functions, we now see how each human individual has an almost incredible diversity of NK cell subsets. Furthermore, the repertoire of these subsets differs greatly from one person, or patient, to another. Among these different subsets are developmental intermediates in pathways of maturation and differentiation, as well as different types of effector NK cell. The latter include effector and memory NK cells, which exhibit, either alone or in combination, some measure of specificity for a pathogen, tumor or allogeneic tissue. The immediate challenge for research is to define the physiological contributions of the different NK cell subsets. The future challenge will be using that knowledge to design specific, effective NK cell mediated therapies against infectious, malignant and reproductive diseases.

MATERIAL AND METHODS

The objective of this study was to profile the baseline diversity of the healthy human NK cell repertoire, providing a framework to understand the roles of NK cells in health and in disease pathogenesis. To this end, we performed an observational study of 22 healthy individuals, including 5 pairs of monozygotic twins, a sample size that was determined both by feasibility and to allow us to sample a representative pool of donors of different KIR and HLA genotypes. This was an observational study; randomization and blinding were not used. All phenotypic analyses were performed *ex vivo* and not run in replicates. Analyses of twins vs. unrelated individuals allowed us to examine the influence of host genetics vs. environment on the NK cell repertoire.

Human subjects and cells

This study was performed in accordance with the Declaration of Helsinki and approved by the Institutional Review Boards of Stanford University and SRI International. Cryopreserved peripheral blood mononuclear cells (PBMC) were obtained from three sources. First, ten generally healthy individuals (5 male, 5 female, aged 24-60, median age 34) were recruited at Stanford University. Second, PBMCs from two healthy anonymous donors were obtained through purchase from the Stanford Blood Center. Third, PBMC samples from 5 sets of monozygotic twin pairs (4 male, 6 female, aged 20-40, median= 23) were collected as part of an ongoing study of influenza vaccination involving participants in the Twin Research Registry (TRR) at SRI International. Full details of the twin registry have been described (56). All subjects gave fully informed and written consent.

Staining, Data Acquisition, and Analysis

PBMCs were thawed and washed with RPMI-1640 (Corning Cellgro) with 10% HI-FBS, 2mM L-Glutamine and antibiotics (penicillin [100 units/ml] and streptomycin [100μg/ml]) (Gibco BRL/Life technologies) and rested 37°C with 5% CO_2 for 4 hours. Two million PBMC were stained for mass cytometry analyses as described (26, 27) using the antibodies listed in Table S1 and data were acquired on a CyTOF® instrument (DVS Sciences). Fluorescence cytometry analyses were acquired on an LSR II (BD Biosciences) using the antibodies in Table S2. FCS files were analyzed using FlowJo software v9.4.8 (Treestar, Inc). Boolean analyses were performed using a custom Python script (http://

www.python.org) to sort cells based upon positive or negative expression of each marker. Spanning-tree progression of density normalized events (SPADE) analyses were performed as described (36).

Antibody Conjugation

Antibodies were purchased from companies specified in Table S1 and labeled using Maxpar-X8 labeling reagent kits (DVS Sciences) according to manufacturer instructions.

KIR and HLA Genotyping and CMV serology

KIR gene-presence and *HLA-A,-B,-C* alleles were determined by PCR-SSOP using a Luminex-100 instrument (Luminex corp.). The assays were performed using LABType® SSO reagents (One Lambda). KIR genotyping at allele-level resolution was determined by pyrosequencing as described (57). DNA from Twin pair 1 was not available for typing, but monozygosity was previously established (56). CMV serology was obtained using the anti-CMV IgG ELISA kit (Gold Standard diagnostics), according to manufacturers suggestions. Plasma samples for CMV testing from the healthy unrelated individuals (HNKs) were obtained approximately 1 year following PBMC sampling.

Quantification of diversity

The inverse Simpson index was calculated using the equation

$$D=\frac{1}{\sum_{i=1}^{S} p_i^2}$$

where S is the total number of species and p_i is the proportional abundance of the ith species.

Sample-based rarefaction

Mean species occurrences of the number of phenotypes observed were plotted and the curve was projected by fitting the asymptotic equation

$$S(n)=\frac{S_{\max} n}{B+n}$$

where n is the number of species included, S_{max} is the total number of species projected, and B is a fitted constant.

Non-parametric species estimators

The Chao 1 estimator is given by the equation

$$S_1=S_{obs}+\frac{F_1^2}{2F_2}$$

where S_{obs} is the number of species observed, F_1 is the number of species observed in exactly one cell, and F_2 is the number observed in exactly two cells.

The Chao 2 estimator, which incorporates occurrence rather than abundance data, is given by the equation

$$S_2 = S_{obs} + \frac{Q_1^2}{2Q_2}$$

where Q_1 is the number of species observed in exactly one donor, and Q_2 is the number observed in exactly two donors.

The second order Jackknife estimator, which reduces sampling bias by systematic recalculation with subsets of the data removed, is given by the equation

$$S_{jack2} = S_{obs} + \left[\frac{Q_1 (2m-3)}{m} - \frac{Q_2 (m-2)^2}{m (m-1)} \right]$$

where m is the total number of samples.

Statistical Analysis

Statistical analyses were performed using Excel (Microsoft Corp), STATA v10 (Stata Corp), Prism v5 (GraphPad Software, Inc.), and the Open Source statistical package R www.r-project.org; ISBN 3-900051-07-0).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Evan Newell, Erin Simonds, Arnav Moudgil, Jamie Haddon, Hugo Hilton, and Paul Bollyky for helpful discussions and/or critical reading of the manuscript. Twins were recruited from the Twin Research Registry at SRI International; we wish to thank Mary McElroy, Lisa Jack, Ruth Krasnow, Jill Rubin, Dina Basin, Lucia Panini, and Marty Ritchey, and funds from SRI's Center for Health Sciences and NIH grants DA011170, DA023063, AI057229, AI090019, and ES022153. We also thank the twins for their contributions to science. We thank the clinical staff of the Stanford-LPCH Vaccine Program including Susan Swope, RN, Cynthia Walsh, RN, Anthony Trela, RN, phlebotomist Michele Ugur, clinical research assistants Ashima Goel, Kyrsten Spann, Raquel Fleischmann, Sushil Batra and Isaac Chang.

Funding

This study was funded by NIH training grant T32 AI07290 (AH), NSF training grant DGE-114740 (DSA), NIH grant AI22039 (PP), NIH U19AI090019 and U19 AI057229 (MMD) NIH/NCRR CTSA award number UL1 RR025744, a New Investigator Award from the University of Washington Center for AIDS Research (P30 AI027757, CAB), and a Beckman Young Investigator Award (CAB).

REFERENCES

1. Kiessling R, Klein E, Wigzell H. "Natural" killer cells in the mouse. I. Cytotoxic cells with specificity for mouse Moloney leukemia cells. Specificity and distribution according to genotype. Eur. J. Immunol. 1975; 5:112–117. [PubMed: 1234049]
2. Kiessling R, Klein E, Wigzell H. "Natural" killer cells in the mouse. I. Cytotoxic cells with specificity for mouse Moloney leukemia cells. Specificity and distribution according to genotype. Eur. J. Immunol. 1975; 5:112–117. [PubMed: 1234049]
3. Herberman RR, Ortaldo JR, Bonnard GD. Augmentation by interferon of human natural and antibody-dependent cell-mediated cytotoxicity. Nature. 1979; 277:221–223. [PubMed: 95205]

4. Ortaldo JR, Bonnard GD, Kind PD, Herberman RB. Cytotoxicity by cultured human lymphocytes: characteristics of effector cells and specificity of cytotoxicity. J Immunol. 1979; 122:1489–1494. [PubMed: 87435]
5. Biron CA, Welsh RM. Activation and role of natural killer cells in virus infections. Med Microbiol Immunol. 1982; 170:155–172. [PubMed: 6176843]
6. Di Santo JP. Functionally distinct NK-cell subsets: Developmental origins and biological implications. Eur. J. Immunol. 2008; 38:2948–2951. [PubMed: 18979515]
7. Colonna M. Specificity and function of immunoglobulin superfamily NK cell inhibitory and stimulatory receptors. 1997; 155:127–133.
8. Raulet DH, Held W, Correa I, Dorfman JR, Wu MF, Corral L. Specificity, tolerance and developmental regulation of natural killer cells defined by expression of class I-specific Ly49 receptors. 1997; 155:41–52.
9. Lanier LLL. Natural Killer Cells: From No Receptors to Too Many. Immunity. 1997; 6:8–8.
10. Ljunggren HGH, Kärre KK. In search of the “missing self”: MHC molecules and NK cell recognition. Immunol Today. 1990; 11:237–244. [PubMed: 2201309]
11. Vivier EE, Ugolini SS. Natural killer cells: from basic research to treatments. Frontiers in immunology. 2011; 2:18–18. [PubMed: 22566808]
12. Valiante NMN, Uhrberg MM, Shilling HGH, Lienert-Weidenbach KK, Arnett KLK, D'Andrea AA, Phillips JHJ, Lanier LLL, Parham PP. Functionally and Structurally Distinct NK Cell Receptor Repertoires in the Peripheral Blood of Two Human Donors. Immunity. 1997; 7:13–13. [PubMed: 9252116]
13. Wagtmann NN, Rajagopalan SS, Winter CCC, Peruzzi MM, Long EOE. Killer cell inhibitory receptors specific for HLA-C and HLA-B identified by direct binding and by functional transfer. Immunity. 1995; 3:9–9. [PubMed: 7621081]
14. Colonna M, Samaridis J. Cloning of immunoglobulin-superfamily members associated with HLA-C and HLA-B recognition by human natural killer cells. Science. 1995; 268:405–408. [PubMed: 7716543]
15. Sternberg-Simon M, Brodin P, Pickman Y, Onfelt B, Karre K, Malmberg K-J, Höglund P, Mehr R. Natural killer cell inhibitory receptor expression in humans and mice: a closer look. Frontiers in immunology. 2013; 4:1–12. [PubMed: 23355837]
16. Yawata M, Yawata N, Draghi M, Partheniou F, Little A-M, Parham P. MHC class I-specific inhibitory receptors and their ligands structure diverse human NK-cell repertoires toward a balance of missing self-response. Blood. 2008; 112:2369–2380. [PubMed: 18583565]
17. Shilling HG. Reconstitution of NK cell receptor repertoire following HLA-matched hematopoietic cell transplantation. Blood. 2003; 101:3730–3740. [PubMed: 12511415]
18. Beziat V, Liu L, Malmberg JA, Ivarsson MA, Sohlberg E, Bjorklund AT, Retiere C, Sverremark-Ekstrom E, Traherne J, Ljungman P, Schaffer M, Price DA, Trowsdale J, Michaelsson J, Ljunggren HG, Malmberg K-J. NK cell responses to cytomegalovirus infection lead to stable imprints in the human KIR repertoire and involve activating KIRs. Blood. 2013 doi:10.1182/blood-2012-10-459545.
19. Schonberg K, Sribar M, Enczmann J, Fischer JC, Uhrberg M. Analyses of HLA-C-specific KIR repertoires in donors with group A and B haplotypes suggest a ligand-instructed model of NK cell receptor acquisition. Blood. 2011; 117:98–107. [PubMed: 20935255]
20. Schonberg K, Fischer JC, Kogler G, Uhrberg M. Neonatal NK-cell repertoires are functionally, but not structurally, biased toward recognition of self HLA class I. Blood. 2011; 117:5152–5156. [PubMed: 21415265]
21. Cichocki F. Epigenetic regulation of NK cell differentiation and effector functions. 2013:1–8.
22. Bjorkstrom NK, Lindgren T, Stoltz M, Fauriat C, Braun M, Evander M, Michaelsson J, Malmberg K-J, Klingstrom J, Ahlm C, Ljunggren HG. Rapid expansion and long-term persistence of elevated NK cell numbers in humans infected with hantavirus. Journal of Experimental Medicine. 2011; 208:13–21. [PubMed: 21173105]
23. Petitdemange C, Becquart P, Wauquier N, Béziat V, Debré P, Leroy EM, Vieillard V, Yokoyama WM. Unconventional Repertoire Profile Is Imprinted during Acute Chikungunya Infection for

Natural Killer Cells Polarization toward Cytotoxicity. PLoS Pathog. 2011; 7:e1002268. [PubMed: 21966274]

24. Lopez-Verges S, Milush JM, Schwartz BS, Pando MJ, Jarjoura J, York VA, Houchins JP, Miller S, Kang S-M, Norris PJ, Nixon DF, Lanier LL. 2011; 108:14725–14732.
25. Foley B, Cooley S, Verneris MR, Pitt M, Curtsinger J, Luo X, Lopez-Verges S, Lanier LL, Weisdorf D, Miller JS. Cytomegalovirus reactivation after allogeneic transplantation promotes a lasting increase in educated NKG2C+ natural killer cells with potent function. Blood. 2012; 119:2665–2674. [PubMed: 22180440]
26. Bendall SC, Simonds EF, Qiu P, Amir E-AD, Krutzik PO, Finck R, Bruggner RV, Melamed R, Trejo A, Ornatsky OI, Balderas RS, Plevritis SK, Sachs K, Pe'er D, Tanner SD, Nolan GP. Single-Cell Mass Cytometry of Differential Immune and Drug Responses Across a Human Hematopoietic Continuum. Science. 2011; 332:687–696. [PubMed: 21551058]
27. Newell EWE, Sigal NN, Bendall SCS, Nolan GPG, Davis MMM. Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes. Immunity. 2012; 36:142–152. [PubMed: 22265676]
28. Bauer SS, Groh VV, Wu JJ, Steinle AA, Phillips JHJ, Lanier LLL, Spies TT. Activation of NK cells and T cells by NKG2D, a receptor for stress-inducible MICA. Science. 1999; 285:727–729. [PubMed: 10426993]
29. Sivori S, Falco M, Moretta L, Moretta A. Extending killer Ig-like receptor function: from HLA class I recognition to sensors of microbial products. Trends Immunol. 2010; 31:289–294. [PubMed: 20630802]
30. Silva A, Andrews DM, Brooks AG, Smyth MJ, Hayakawa Y. Application of CD27 as a marker for distinguishing human NK cell subsets. International Immunology. 2008; 20:625–630. [PubMed: 18326863]
31. Della Chiesa M, Falco M, Podesta M, Locatelli F, Moretta L, Frassoni F, Moretta A. Phenotypic and functional heterogeneity of human NK cells developing after umbilical cord blood transplantation: a role for human cytomegalovirus? Blood. 2012; 119:399–410. [PubMed: 22096237]
32. André P, Castriconi R, Espéli M, Anfossi N, Juarez T, Hué S, Conway H, Romagné F, Dondero A, Nanni M, Caillat-Zucman S, Raulet DH, Bottino C, Vivier E, Moretta A, Paul P. Comparative analysis of human NK cell activation induced by NKG2D and natural cytotoxicity receptors. Eur. J. Immunol. 2004; 34:961–971. [PubMed: 15048706]
33. Milush JM, Long BR, Snyder-Cappione JE, Cappione AJ, York VA, Ndhlovu LC, Lanier LL, Michaelsson J, Nixon DF. Functionally distinct subsets of human NK cells and monocyte/DC-like cells identified by coexpression of CD56, CD7, and CD4. Blood. 2009; 114:4823–4831. [PubMed: 19805616]
34. Bjorkstrom NK, Riese P, Heuts F, Andersson S, Fauriat C, Ivarsson MA, Bjorklund AT, Flodstrom-Tullberg M, Michaelsson J, Rottenberg ME, Guzman CA, Ljunggren H-G, Malmberg K-J. Expression patterns of NKG2A, KIR, and CD57 define a process of CD56dim NK-cell differentiation uncoupled from NK-cell education. Blood. 2010; 116:3853–3864. [PubMed: 20696944]
35. Yawata M. Roles for HLA and KIR polymorphisms in natural killer cell repertoire selection and modulation of effector function. Journal of Experimental Medicine. 2006; 203:633–645. [PubMed: 16533882]
36. Qiu PP, Simonds EFE, Bendall SCS, Gibbs KDK, Bruggner RVR, Linderman MDM, Sachs KK, Nolan GPG, Plevritis SKS. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. Nat Biotechnol. 2011; 29:886–891. [PubMed: 21964415]
37. Guma M. Imprint of human cytomegalovirus infection on the NK cell receptor repertoire. Blood. 2004; 104:3664–3671. [PubMed: 15304389]
38. Gumá M, Budt M, Sáez A, Brckalo T, Hengel H, Angulo A, López-Botet M. Expansion of CD94/NKG2C+ NK cells in response to human cytomegalovirus-infected fibroblasts. Blood. 2006; 107:3624–3631. [PubMed: 16384928]
39. Gotelli, NJ.; Colwell, RK. Biological Diversity: Frontiers in Measurement and Assessment. Magurran, AE.; McGill, BJ., editors. Estimating Species Richness: 2010. p. 39-54.

40. Gotelli NJ, Colwell RK. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol Lett. 2001; 4:379–391.
41. Colwell RK. Biodiversity: Concepts, patterns, and measurement. The Princeton guide to ecology. 2009:257–263.
42. Shum BP, Flodin LR, Muir DG, Rajalingam R, Khakoo SI, Cleland S, Guethlein LA, Uhrberg M, Parham P. Conservation and variation in human and common chimpanzee CD94 and NKG2 genes. J Immunol. 2002; 168:240–252. [PubMed: 11751968]
43. Evans JH, Horowitz A, Mehrabi M, Wise EL, Pease JE, Riley EM, Davis DM. A distinct subset of human NK cells expressing HLA-DR expand in response to IL-2 and can aid immune responses to BCG. Eur. J. Immunol. 2011; 41:1924–1933. [PubMed: 21491418]
44. Grzywacz B, Kataria N, Sikora M, Oostendorp RA, Dzierzak EA, Blazar BR, Miller JS, Verneris MR. Coordinated acquisition of inhibitory and activating receptors and functional properties by developing human natural killer cells. Blood. 2006; 108:3824–3833. [PubMed: 16902150]
45. Jost S, Reardon J, Peterson E, Poole D, Bosch R, Alter G, Altfeld M. Expansion of 2B4+ natural killer (NK) cells and decrease in NKp46+ NK cells in response to influenza. Immunology. 2011; 132:516–526. [PubMed: 21214542]
46. Marlin R, Duriez M, Berkane N, De Truchis C, Madec Y, Rey-Cuille M-A, Cummings J-S, Cannou C, Quillay H, Barré-Sinoussi F, Nugeyre M-T, Menu E. Dynamic Shift from CD85j/ILT-2 to NKG2D NK Receptor Expression Pattern on Human Decidual NK during the First Trimester of Pregnancy. PLoS ONE. 2012; 7:e30017. [PubMed: 22242197]
47. Hilton HG, Vago L, Older Aguilar AM, Moesta AK, Graef T, Abi-Rached L, Norman PJ, Guethlein LA, Fleischhauer K, Parham P. Mutation at Positively Selected Positions in the Binding Site for HLA-C Shows That KIR2DL1 Is a More Refined but Less Adaptable NK Cell Receptor Than KIR2DL3. The Journal of Immunology. 2012; 189:1418–1430. [PubMed: 22772445]
48. Norman PJ, Abi-Rached L, Gendzekhadze K, Korbel D, Gleimer M, Rowley D, Bruno D, Carrington CVF, Chandanayingyong D, Chang Y-H, Crespí C, Saruhan-Direskeneli G, Fraser PA, Hameed K, Kamkamidze G, Koram KA, Layrisse Z, Matamoros N, Milá J, Park MH, Pitchappan RM, Ramdath DD, Shiau M-Y, Stephens HAF, Struik S, Verity DH, Vaughan RW, Tyan D, Davis RW, Riley EM, Ronaghi M, Parham P. Unusual selection on the KIR3DL1/S1 natural killer cell receptor in Africans. Nat Genet. 2007; 39:1092–1099. [PubMed: 17694054]
49. Graef T, Moesta AK, Norman PJ, Abi-Rached L, Vago L, Older Aguilar AM, Gleimer M, Hammond JA, Guethlein LA, Bushnell DA, Robinson PJ, Parham P. KIR2DS4 is a product of gene conversion with KIR3DL2 that introduced specificity for HLA-A*11 while diminishing avidity for HLA-C. Journal of Experimental Medicine. 2009; 206:2557–2572. [PubMed: 19858347]
50. Malmberg K-J, Michaélsson J, Parham P, Ljunggren H-G. Killer Cell Immunoglobulin-like Receptor Workshop: Insights into Evolution, Genetics, Function, and Translation. Immunity. 2011; 35:653–657. [PubMed: 22118518]
51. Bryceson YT. Synergy among receptors on resting NK cells for the activation of natural cytotoxicity and cytokine secretion. Blood. 2006; 107:159–166. [PubMed: 16150947]
52. Shilling HGH, Young NN, Guethlein LAL, Cheng NWN, Gardiner CMC, Tyan DD, Parham PP. Genetic control of human NK cell repertoire. J Immunol. 2002; 169:239–247. [PubMed: 12077250]
53. Anfossi N, André P, Guia S, Falk CS, Roetynck S, Stewart CA, Breso V, Frassati C, Reviron D, Middleton D, Romagné F, Ugolini S, Vivier E. Human NK Cell Education by Inhibitory Receptors for MHC Class I. Immunity. 2006; 25:331–342. [PubMed: 16901727]
54. Tarek N, Le Luduec J-B, Gallagher MM, Zheng J, Venstrom JM, Chamberlain E, Modak S, Heller G, Dupont B, Cheung N-KV, Hsu KC. Unlicensed NK cells target neuroblastoma following anti-GD2 antibody treatment. J Clin Invest. 2012; 122:3260–3270. [PubMed: 22863621]
55. Goodier MR, Mela CM, Steel A, Gazzard B, Bower M, Gotch F. NKG2C+ NK Cells Are Enriched in AIDS Patients with Advanced-Stage Kaposi's Sarcoma. Journal of Virology. 2006; 81:430–433. [PubMed: 17035308]

56. Krasnow RE, Jack LM, Lessov-Schlaggar CN, Bergen AW, Swan GE. The Twin Research Registry at SRI International. Twin research and human genetics : the official journal of the International Society for Twin Studies. 2013; 16:463–470. [PubMed: 23084148]

57. Norman PJ, Abi-Rached L, Gendzekhadze K, Hammond JA, Moesta AK, Sharma D, Graef T, McQueen KL, Guethlein LA, Carrington CVF, Chandanayingyong D, Chang YH, Crespi C, Saruhan-Direskeneli G, Hameed K, Kamkamidze G, Koram KA, Layrisse Z, Matamoros N, Mila J, Park MH, Pitchappan RM, Ramdath DD, Shiau MY, Stephens HAF, Struik S, Tyan D, Verity DH, Vaughan RW, Davis RW, Fraser PA, Riley EM, Ronaghi M, Parham P. Meiotic recombination generates rich diversity in NK cell receptor genes, alleles, and haplotypes. Genome Research. 2009; 19:757–769. [PubMed: 19411600]

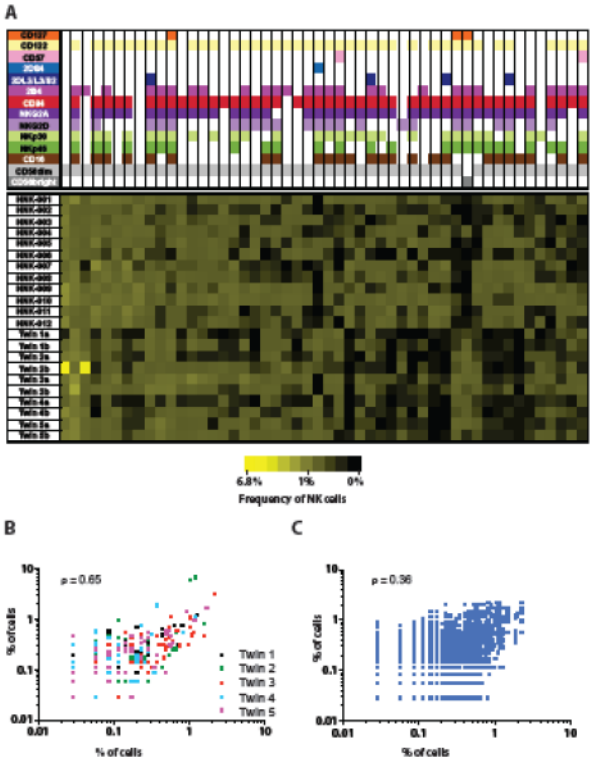

Figure 1. Assessment of NK cell repertoire diversity based on expression of unique receptor profiles

(A) Frequencies of the 50 most abundant NK cell phenotypes based on expression of 28 receptors for 12 unrelated individuals (HNK-001 through HNK-012) and 5 pairs of monozygotic twins (Twins 1a/1b – 5a/5b). Each column represents a phenotype, with colored boxes indicating receptor presence. (B) Comparative correlation of each twin with his/her identical twin of the frequencies of the 50 phenotypes shown in A. Spearman's correlation coefficient is displayed. (C) Comparative correlation of each healthy unrelated individual with the other 11 individuals of the frequencies of the 50 phenotypes shown in A.

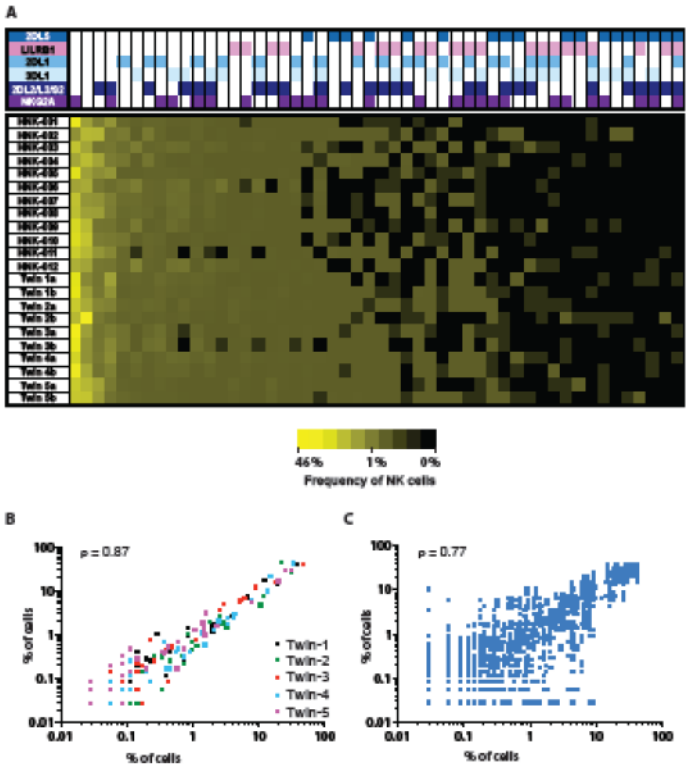

Figure 2. Expression patterns of inhibitory NK cell receptors in unrelated individuals and monozygotic twins

(A) Inhibitory receptor profile of NK cells in 12 unrelated individuals and 5 pairs of monozygotic twins as in Figure 1A. The analysis was restricted to the evaluation of expression profiles of the six inhibitory receptors KIR2DL1, KIR2DL2/L3/S2, KIR2DL5, KIR3DL1, LILRB1, and NKG2A. (B) Comparative correlation of each twin with his/her identical twin of the frequencies of the 50 phenotypes shown in A. Spearman's correlation coefficient is displayed. (C) Comparative correlation of each healthy unrelated individual with the other 11 individuals of the frequencies of the 50 phenotypes shown in A.

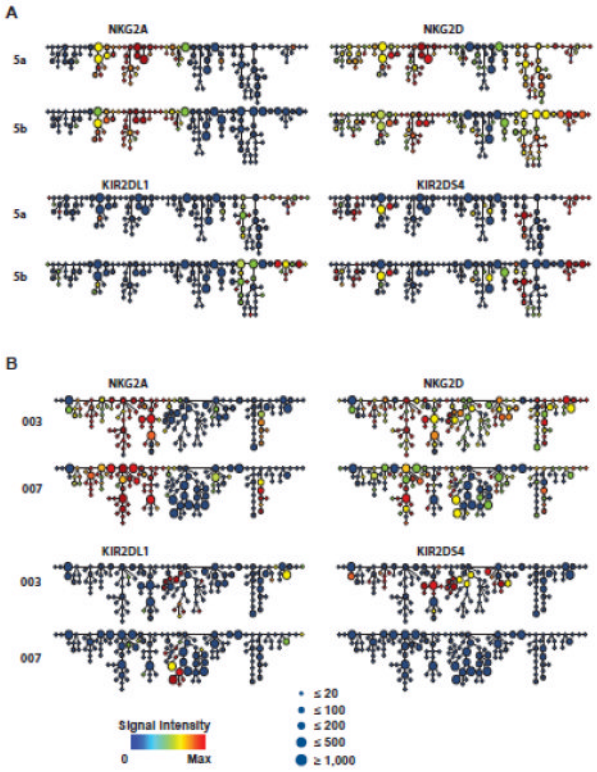

Figure 3. Clustering analysis of NK cells using SPADE reveals diverse receptor expression patterns

(A) Representative SPADE trees of NKG2A, NKG2D, KIR2DL1 and KIR2DS4 for monozygotic twins 5a (top) and 5b (bottom). Node color represents signal intensity of each marker and size represents frequency. (B) Representative SPADE trees of NKG2A, NKG2D, KIR2DL1 and KIR2DS4 for 2 healthy unrelated individuals, HNK-003 (top) and HNK-007 (bottom).

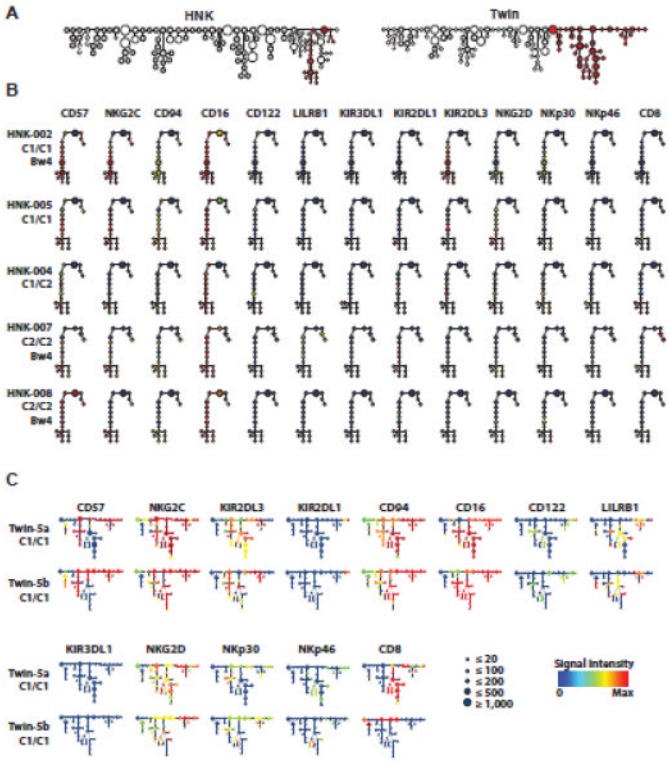

Figure 4. Identification and phenotypic evaluation of CD57^{+}NKG2C^{+} 'memory-like' NK cells

(A) Skeleton structures of a SPADE tree from healthy unrelated individuals (HNK) and monozygotic twins (Twin) with NKG2C-expressing nodes highlighted. (B) Representative SPADE trees of the NKG2C-expressing nodes shown in (A) from 5 unrelated individuals. (C) Representative SPADE trees of NKG2C-expressing nodes for 1 pair of monozygotic twins, as in (B).

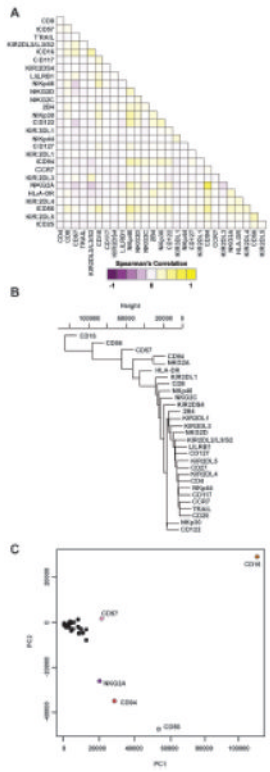

Figure 5. NK cell receptor co-expression patterns and population structure

(A) Spearman's rank correlation matrices of co-expression profiles for each receptor pair for all NK cells from all 22 donors. (B) Hierarchical clustering dendrogram of NK cell receptor profiles from all NK cells from all 22 donors. (C) Principal component analysis of receptor expression patterns from all NK cells from all 22 donors. Dots indicate the contributions of each receptor to principal component 1 (x-axis) vs. principal component 2 (y-axis).

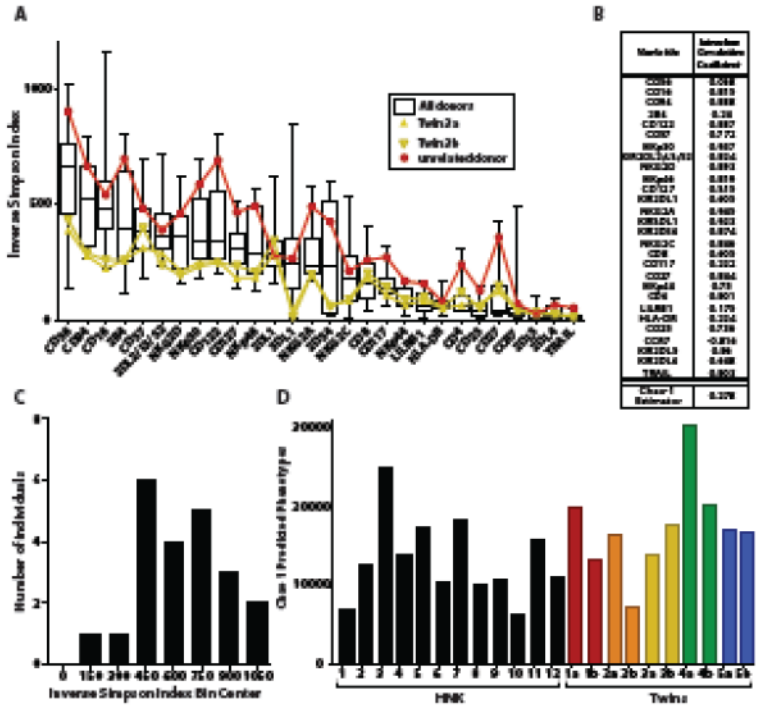

Figure 6. Genetics dictate the expression of most single NK cell receptors, but not the combinatorial assortment of phenotypes

(A) Inverse Simpson for the population of cells expressing each single receptor. Boxplots show the median, 25th and 75th percentiles, and total range of all 22 donors. Yellow lines show a representative twin pair. Red line indicates an unpaired representative twin. (B) Summary of Intraclass Correlation Coefficients (ICC) of the Inverse Simpson Index of twin pairs versus all 22 individuals. (C) Histogram of inverse Simpson indices for the total NK cell population of each donor. (D) Total phenotypes predicted by the Chao 1 non-parametric species estimator (calculated as described in Methods) for all donors.

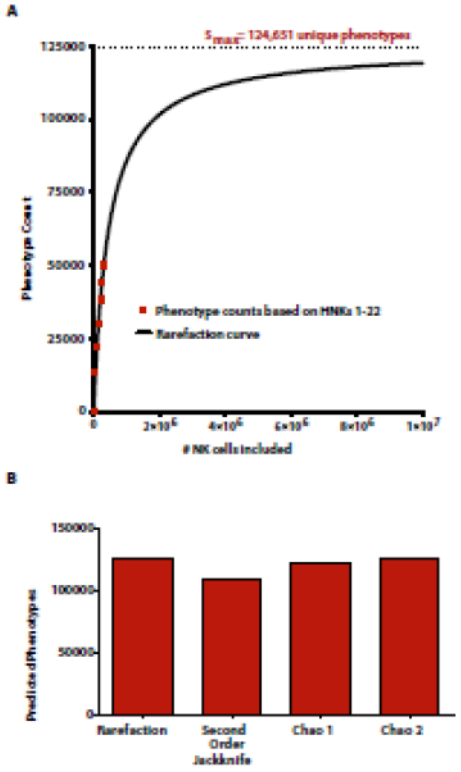

Figure 7. Calculating the expanse of the NK cell repertoire

(A) Sample-based rarefaction predicting the total number of NK cell phenotypes within the cohort of 22 donors. The number of phenotypes observed was calculated for incremental inputs of 50,000 NK cells up to and including the total 304,909 NK cells sampled in all 22 donors, reshuffling the order of donor sampling each of 10 times. These data were used to plot a curve, which when extrapolated to an asymptote provided an estimate of the total NK cell population. (B) Non-parametric species estimators confirm quantification of the total number of NK cell phenotypes.