iPSC-based modeling of helicase deficiency reveals impaired cell proliferation and increased apoptosis after NK cell lineage commitment

Seungmae Seo[1], Sagar L. Patil[1,a], Yong-Oon Ahn[1], Jacqueline Armetta[1], Everardo Hegewisch-Solloa[1], Micah Castillo[2], Nicole C. Guilz[1], Achchhe Patel[3], Barbara Corneo[3], Malgorzata Borowiak[4], Preethi Gunaratne[2], and Emily M. Mace[1]

[1]Department of Pediatrics, Vagelos College of Physicians and Surgeons, Columbia University Irving Medical Center, New York NY 10032
[2]Department of Biology and Biochemistry, University of Houston, Houston, TX, USA, 77204
[3]Columbia Stem Cell Initiative, Columbia University Irving Medical Center, New York, NY, USA, 10032
[4]Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland
[a]Current affiliation: Department of Bioanalysis, Immunogenicity, and Biomarkers, GlaxoSmithKline, Collegeville, PA, USA, 19426

Cell proliferation is a ubiquitous process required for organismal development and homeostasis. However, individuals with partial loss-of-function variants in DNA replicative helicase components often present with immunodeficiency due to specific loss of natural killer (NK) cells. Such lineage-specific disease phenotypes raise questions on how the proliferation is regulated in cell type-specific manner. We aimed to understand NK cell-specific proliferative dynamics and vulnerability to impaired helicase function using iPSCs from individuals with NK cell deficiency (NKD) due to hereditary compound heterozygous *GINS4* variants. We observed and characterized heterogeneous cell populations that arise during the iPSC differentiation along with NK cells. While overall cell proliferation decreased with differentiation, early NK cell precursors showed a short burst of cell proliferation. GINS4 deficiency induced replication stress in these early NK cell precursors, which are poised for apoptosis, and ultimately recapitulate the NKD phenotype.

Replication stress | NK cell | immunodeficiency | cell cycle | CMG helicase | GINS4

Correspondence: ***em3375@cumc.columbia.edu***

DRAFT

Introduction

Human natural killer (NK) cells are cytotoxic innate lymphocytes that play critical roles in controlling viral infection and tumor immunity (1–4). $CD56^{+}CD3^{-}$NK cells comprise 10% of peripheral blood mononuclear cells (PBMCs) (5–7). Individuals with significantly reduced frequencies of NK cells in their peripheral blood (natural killer cell deficiency, NKD) frequently experience recurring viral infection and hospitalization (8). While monogenic causes of isolated NKD are rare, variants in components of the CDC45-MCM-GINS (CMG) helicase complex, including *MCM4*, *MCM10*, *GINS1*, and *GINS4*, have previously been reported to cause NKD (9–16). Clinical phenotypes apart from NKD and susceptibility to viral infections are heterogeneous, ranging from cell type specific deficiencies such as neutropenia to more general manifestations including growth retardation and short stature (17). Other lymphocytes, including T and B cells, are largely unaffected, and the mechanism underlying the specific deficiency of the innate lymphocytes in these individuals is unknown.

The observation that NK cells are particularly susceptible to CMG helicase variants is surprising, given the ubiquitous requirement for these proteins. DNA replication is a highly coordinated and regulated process that relies on the tightly regulated function of the replisome (18). In S phase, CDC45, the MCM2-7 hexamer, and the GINS1-4 tetramer form the CMG helicase complex that unwinds double-stranded DNA (19). Before formation of the replisome, excess MCM hexamers are loaded onto chromatin, of which only 5-10% are used in S phase (20, 21). This phenomenon is called the MCM paradox, and the extra MCM hexamers bound to dormant origins serve as a backup pool that can be activated to rescue stalled or collapsed replication forks under stress, providing a mechanism for cells to withstand mild helicase deficiencies in the presence of sufficient amounts of MCM proteins (22, 23). The regulation of the expressions and functions of the other CMG helicase components remain to be investigated.

Since previously reported helicase variants affect NK cell maturation, it is crucial to model NK cell development to understand their pathogenesis. A stepwise model describes human NK cell differentiation in six stages from CD34+ multipotent progenitor (NK progenitors, NKPs) (24, 25). As $CD34^{+}$ progenitors develop into mature NK cells, they likely migrate from bone marrow to secondary lymphoid tissues where they can remain or exit into the bloodstream. The lineage potential of the maturing NK progenitors decreases from all hematopoietic lineage to lymphocytes, then to an NK/ILC precursor, and then to lineage-committed conventional NK cells (stage 4) (24–26). Previous studies have shown that $CD34^{+}$ cells with helicase variants have impaired NK cell maturation in in vitro differentiation assays, which mirrors patient phenotypes in vivo (10, 12). Other studies have relied upon primary cell lines and PBMCs for studying functional defects in NK cells, thus limiting understanding of the full spectrum of NKD pathogenesis (9–12).

Induced pluripotent stem cells (iPSCs) provide a unique opportunity to study the pathogenesis of NKDs caused by congenital variants and investigate the developmental process of

other cell types. Previous studies have reported protocols for differentiating NK cells from iPSCs (27–30). While these studies demonstrate the generation of functional, mature NK cells, they lack a detailed understanding of the developmental process these cells undergo. In this work, we adapt an embryoid body (spin-EB) based differentiation method on autologous stroma or a xenogenic stroma, OP9, to understand the steps of NK cell differentiation, specification, and maturation (27).

Results

Generation and validation of GINS4 variant iPSCs . To better understand the effect of helicase deficiencies on human NK cell development and characterize the trajectory of NK cell differentiation from human pluripotent stem cells, we performed in-depth phenotypic analysis of NK cell differentiation using iPSC lines from a previously described family with compound heterozygous variants in *GINS4* (12). iPSCs were generated from PBMCs from two affected siblings with compound heterozygous *GINS4* variants (II.1, II.2) and their apparently healthy parents each carrying a single variant (I.1, I.2) (Fig. 1A) (12). As previously described using EBV-transformed B lymphoblastoid cell lines from peripheral blood, GINS4 protein expression in iPSC from siblings II.1 and II.2 was significantly decreased (Fig. 1B). Using CRISPR-Cas9, we corrected both alleles to generate an isogenic control line (II.1-Cor) from the more severely clinically affected individual (II.1) and confirmed the restoration of protein expression by Western blot (Fig. 1C). All iPSC lines were validated for pluripotency and normal karyotypes. We adapted a method for NK cell differentiation using the spin-EB method, which has previously been shown to support NK cell differentiation with the generation of autologous stroma (27). We additionally sorted CD34$^+$ cells from EBs and differentiated these on irradiated OP9 stromal cells to test NK cell differentiation on non-autologous xenogenic stroma (Fig. 1D) (31). In all conditions, cells were cultured in previously defined NK cell differentiation cytokines with IL-3 (7 days only), IL-7, IL-15, SCF, and FLT3L for 35 days (32). Cells were collected weekly and assessed by flow cytometry and RNA-seq (Fig. 1D).

NK cell differentiation from iPSCs generates multiple hematopoietic lineages. Stages of NK cell differentiation can be defined by combinatorial expression of CD34, CD45, CD56, CD94, CD117, CD16, killer Ig-like receptors (KIR), granzymes, and perforin measured by flow cytometry (25) (Table S1). We generated flow cytometry data from 6 timepoints of spin-EB differentiation using 6 iPSC lines, namely the 4 members of the proband's family (I.1, I.2, II.1, and II.2), II.1-Cor, and an unrelated control (C.1). Flow cytometry data from all timepoints were concatenated, normalized, and visualized following dimensionality reduction (Fig. 2A). Cluster identities were assigned based on relative expression of markers (Fig. 2A, right). On day 6 of differentiation from iPSC, we consistently observed the emergence of CD34$^+$ endothelial and hematopoietic precursors and CD34$^-$ stromal cell precursors (Fig. S1A). CD34 expression decreased as the cells sequentially gained CD45, CD56, and granzyme B, indicating NK cell functional maturation (Fig. 2A, Fig. S1B). The emergence of CD45+ cells was primarily detected at day 14, followed by commitment to the NK/ILC lineage indicated by appearance of CD56$^+$CD45$^+$ cells at day 28 (Fig. 2A, S1A). NK cell maturation marked by CD94, CD117, perforin, and granzyme B expression was prominent by days 35 and maintained through day 42 (25).

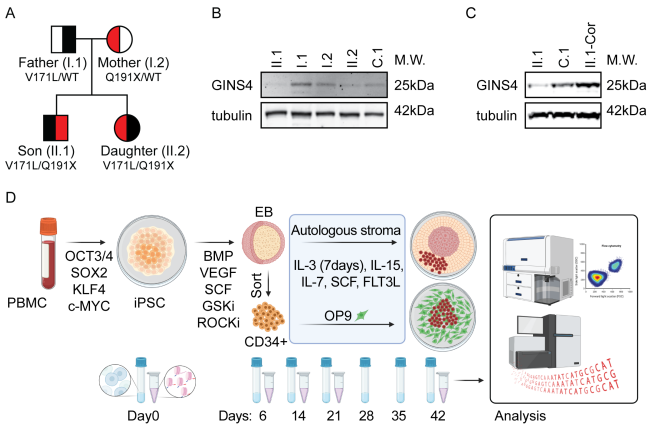

Fig. 1. Generation and validation of *GINS4* variant iPSCs A) Pedigree showing *GINS4* variants in the proband and his family (described in (12)). B) Western blot showing GINS4 levels in the iPSCs generated from the family members and an unrelated control with tubulin for a loading control. C) Western blot showing GINS4 levels in the II.1, C.1, and an isogenic control (II.1-Cor) lines. D) Experimental schematic summarizing the 42-day protocol for NK cell differentiation and sample collection.

We additionally included Eomes, NKp44, CD294 and CD103 in our analysis of CD45$^+$ hematopoietic lineage cells to further distinguish NK cells from other ILC lineages and used Ki-67 to identify proliferating cells (33–35) (Fig. 2B, Table S1C). These analyses identified a small cluster of ILC-like cells present at days 28 and 35 marked by high expression of NKp44, CD294 and CD103 (Fig. 2B). We also identified proliferative CD45$^+$ cells primarily at days 14 and 21, and a small population of proliferative NK cell precursors (CD45$^+$CD56$^+$Ki-67^{high}) present at day 28. At later time points, most cells were Stage 2-3 (CD117$^+$CD94$^-$) or Stage 4-5 (CD117$^{+/-}$CD94$^+$Perforin$^{+/-}$) NK cells (25) (Fig. 2B). Further phenotyping of the CD45$^+$CD16$^+$ non-NK cell clusters transiently present between days 14 and 21 showed that these cells were MPO$^+$CD45^{dim}PU.1$^+$CD16b$^+$CD15$^+$ granulocytes and MPO$^-$CD45$^+$PU.1$^+$M-CSFR$^+$CD32$^+$CD14$^+$ monocytes (Fig. 2C). However, these populations were not present by day 42 of differentiation cultures, and the majority of CD45$^+$ cells at days 35 and 42 were mature NK cells expressing CD56 and granzyme B (Fig. S1C). In summary, throughout the 42 days of differentiation, spin-EB cultures recapitulated the trajectory of human NK cells with coincident generation of other hematopoietic lineage cells, leading to generation of a functionally mature NK cell population as previously described (24, 25, 28).

While NK cells were generated from all control iPSC lines, we found experimental and cell line-specific variation in the

differentiation efficiency. Similar differences in efficiency have previously been reported in differentiation to other lineages using iPSCs (36, 37). To better define the variability of NK cell differentiation from multiple healthy donors, we repeated the NK cell differentiation using six control lines, including the parents of the GINS4 proband (I.1, I.2), two iPSC lines from healthy donor PBMCs (C.1, C.2), one iPSC line from fibroblasts (C.3), and the embryonic stem cell H9 line. We used the mean frequencies of NK cell progenitors expressing CD45, CD56, or granzyme B from control experiments on days 35 and 42 to categorize the efficiency of NK cell generation and visualize the relative efficiency of each line. Individual experiments were scored as inefficient (1: <55% $CD45^+$), moderate (2: >55% $CD45^+$ but <39% $CD56^+$ when gated on $CD45^+$ cells), efficient (3: >39% $CD56^+$ cells when gated on $CD45^+$ cells but <32% granzyme B^+ when gated on $CD45^+$ $CD56^+$), or proficient (4: >32% granzyme B^+ when gated on $CD45^+CD56^+$) (scored in Fig. 2D). In 70% of experiments, we observed efficient or proficient NK cell generation (Fig. 2D). C.1 and I.2 iPSC lines had the highest efficiency (100%, and 89%), followed by I.1 (55%) and C.2 (63%) iPSC lines (Fig. 2E). C.3, an iPSC line derived from fibroblasts, and the H9 embryonic stem cell line showed lower efficiency in NK cell generation when compared to the other four iPSC lines derived from PBMCs (50%) (Fig. 2E). Nonetheless, each control line generated NK cells (scores 3-4) more than 50% of the time under spin-EB differentiation conditions on autologous stroma.

To test the effect of variability introduced by autologous stroma, we differentiated $CD34^+$ cells isolated from day 6 EBs on OP9 stromal cells. Samples were collected between days 21 and 42. As with spin-EB differentiation on autologous stroma, $CD45^+CD16^+$ non-NK cells emerged first on days 21 and 28. Subsequently, $CD117^+$ NK precursors (stage 2-3) appeared on day 28, and granzyme B^+ mature NK cells on day 35 (Fig. 3A). Relative to spin-EB differentiation, the OP9 culture supported further maturation of stage 5 NK cells, namely the increased generation of $CD16^+$ and KIR^+ cells (Fig. 3A, right). Stage 5 NK cells appeared on day 35, one week earlier than their observed appearance on autologous stroma (Fig. 2A), but were not sustained, leaving only a few cells detected by day 42. Differentiation efficiency was measured 24 times in six control iPSC or ES lines (Fig. 3B). On OP9, once the cells committed to $CD56^+$ NK cells, the majority (66.2%, compared to 32.0% in spin-EB culture) of them matured to granzyme B^+ NK cells, resulting in more efficient NK cell generation using our classification described in Fig. 2. As with spin-EB cultures, we again observed experimental variation in the relative frequencies of subsets generated, and the PBMC-derived iPSCs were more efficient at NK cell generation than the fibroblast-derived iPSC (C.3) or ES cell line (H9) (Fig. 3C). Given differences in efficiencies in the blood and fibroblast derived lines, we chose to pair iPSC lines generated from the peripheral blood of the affected siblings (II.1 and II.2) with a blood-derived control line, (C.1) for future experiments.

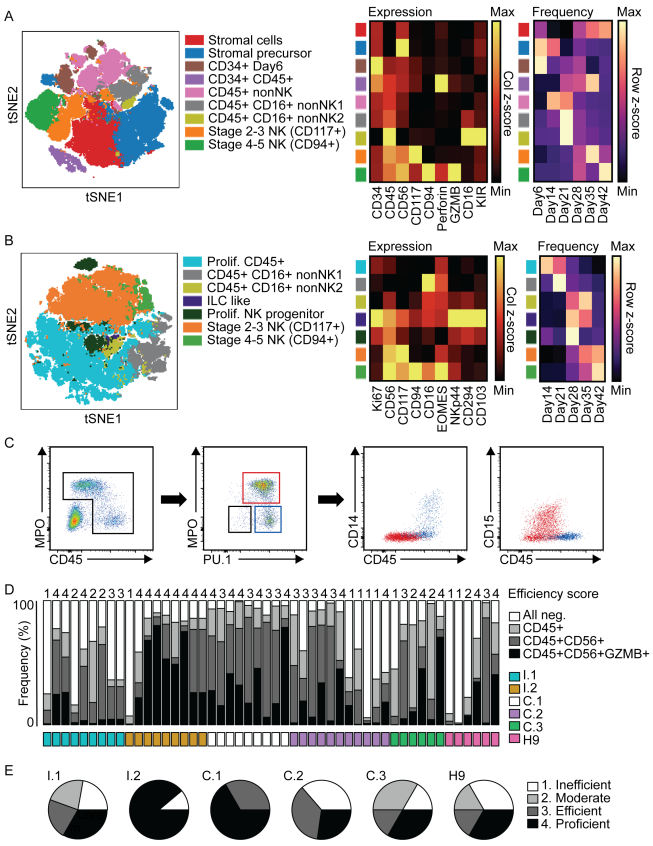

Fig. 2. NK cell differentiation from iPSCs on autologous stroma gives rise to various hematopoietic populations NK cells were differentiated from iPSC lines using the spin-EB method. Flow cytometry was performed to monitor the progression of NK cell differentiation then data were concatenated, normalized, and visualized. A) tSNE plot summarizing the flow cytometry data from six cell lines: *GINS4* variant carriers: I.1, I.2, *GINS4* variant lines: II.1, II.2, an unrelated healthy control (C.1), and an isogenic control (II.1-Cor) at 6 weekly time points. Heatmaps showing relative protein expression of the nine clusters (left) and temporal changes of each cluster (right). B) tSNE plot summarizing the flow cytometry with markers for further dissecting $CD45^+$ populations in six lines and all time points. Heatmaps show gene expression patterns of the seven clusters (left), and the temporal dynamics of each cluster (right). C) Representative flow cytometry results showing the myeloid identity of non-NK cell $CD45^+$ populations on days 14 and 21. Gating strategy shows the selection of $CD45^{dim}$ and $CD45^+$ populations. $MPO^+CD45^{dim}PU.1^+$ cells in red box (middle) are represented as red dots in CD14 and CD15 plots (right), and $MPO^-CD45^+PU.1^+$ cells in blue box (middle) are represented as blue dots in CD14 and CD15 plots (right). D) Stacked bar graph showing the proportion of cells with an NK cell phenotype at days 35 or 42 in each control line. n=6-11 technical replicates. Differentiation efficiency score is shown above each bar based on the frequency of cells with an NK cell phenotype ($CD45^+CD56^+$ or $CD45^+CD56^+GZMB^+$). E) Summary of the differentiation efficiency in each control line showing the normal variation in NK cell differentiation from iPSCs on autologous stroma n=6-11 technical replicates.

***GINS4* variant lines have limited potential for NK cell maturation.** To understand the effect of *GINS4* variants on NK cell development, we performed NK cell differentiation from iPSC using the II.1 and II.2 lines and the isogenic control line (II.1-Cor). There were no apparent cell cycle defects in iPSCs from affected siblings (Fig. 4A), despite their significantly decreased expression of GINS4 protein (Fig. 1). When differentiated on autologous stroma using the spin-EB method, cells from the II.1 line generated NK cells efficiently or proficiently only 30% of the time, while the efficiency in the II.1-cor line was 80% (Fig. 4B). The efficiency of mature NK cell generation from the II.2 line, despite having the

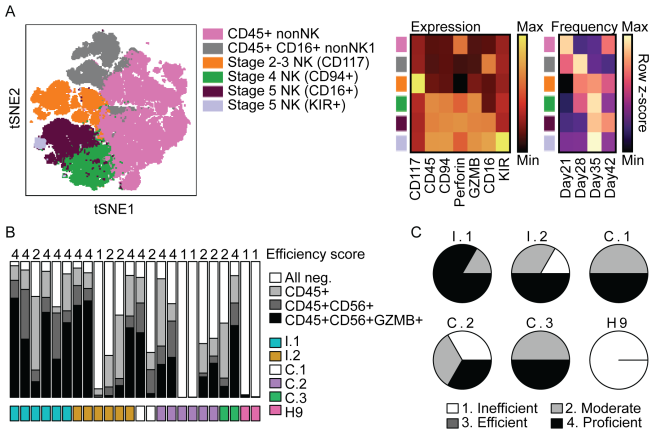

Fig. 3. NK cell differentiation from iPSCs on OP9 reveals further maturation of NK cells NK cells were differentiated from 6 iPSCs by sorting CD34$^+$ cells at day 6 and co-culturing with OP9 stroma. A) tSNE plot summarizing flow cytometry data from six lines (I.1, I.2, II.1, II.2, C.1, and II.1-Cor) at 4 weekly time points beginning on day 21 (14 days after CD34$^+$ cell seeding). Heatmaps showing relative protein expression patterns of the six clusters (left) and temporal changes of each cluster (right). B) Stacked bar graph showing the proportion of cells with an NK cell phenotype at days 35 or 42 in each control line. n=2-6 technical replicates. Differentiation efficiency score is shown above each bar based on the frequency of the NK lineage cells (CD45$^+$CD56$^+$ or CD45$^+$CD56$^+$GZMB$^+$). C) Summary of the differentiation efficiency in each control line showing the normal variation in NK cell differentiation from iPSCs on OP9 (n=2-6 technical replicates).

same *GINS4* variants as II.1, was reproducibly intermediate between the II.1 and control lines (II.2 line: 56% efficient; control lines: 70% efficient) (Fig. 4B, Fig2E). These differences in differentiation efficiencies reflect the variable severity of the clinical phenotypes of the siblings, with II.1 having more severe immunodeficiency and growth retardation than II.2 (12).

As in the control lines, the frequency of CD45$^+$ cells in the II.1 line increased significantly on days 14 and 21, representing the emergence of MPO$^+$CD45^{dim}PU.1$^+$CD16b$^+$CD15$^+$ granulocytes and MPO$^-$CD45$^+$PU.1$^+$M-CSFR$^+$CD32$^+$CD14$^+$ monocytes (Fig. 4C, D). However, the CD45$^+$ population was not sustained through days 28 to 42, when CD56 and granzyme B expressing NK cells appeared in control line experiments (Fig. 4D). NK cell cytotoxicity assays performed on day 42 confirmed that the mature NK cells derived from a healthy donor (C.1) and II.1-Cor line were functionally competent, while CD45$^+$ cells from the II.1 line failed to kill K562 target cells (Fig. 4E). On OP9 stromal cells, the NK cell differentiation efficiency was <30% for both II.1 and II.2 lines, while II.1-Cor line showed 75% efficiency (Fig. 4F). II.1 and II.2 cells remained in the most primitive CD45$^+$ non-NK subset, and the cell viability was low throughout the time of culture (Fig. 4G). This result suggests that while the impairment in the NK cell lineage is inherent to hematopoietic cells, interactions with autologous or non-autologous stromal cell lines can also affect lineage potential and the relative efficiency of NK cell generation.

Next, we tested the impact of *GINS4* variants on the production of other hematopoietic lineages. We performed CFU assays on the II.1 line and control lines. In a condition that modeled primitive hematopoietic differentiation with a transient WNT inhibitor (IWP2) treatment (29, 30, 38). CD43$^+$CD34$^-$ precursors from all lines generated >90% CFU-E and 5% CFU-G colonies (Fig. 4H, left). Definitive differentiation was also modeled using transient GSK inhibition with CHIR99021 (29, 38). Again, the CD34$^+$CD45$^+$ precursors from all cell lines generated a variety of colonies with no significant difference in frequency or absolute numbers between cell lines (Fig. 4H, right). We also performed T cell differentiation through the double positive stage and found no significant difference in efficiency of double positive T cell in all lines tested, consistent with no reported T cell deficiency in the proband or his sibling (II.1, II.2) (12) (Fig. 4I). Taken together, these experiments show that loss-of-function *GINS4* variants specifically restrict NK cell generation from pluripotent stem cells, with a greater impairment in NK cell maturation in cells from the clinically most affected individual (II.1).

Transcriptional signatures of differentiating iPSCs align with in vivo developmental and hematopoietic cell identities. To better understand the trajectory of human NK cell development from pluripotent stem cells, we performed bulk RNA-seq on sorted CD34$^+$ cells on day 6, and separated CD45$^+$ hematopoietic and CD45$^-$ non-hematopoietic/mesodermal cells from days 14, 21, and 42. We performed a Bayesian non-negative matrix factorization method using Coordinated Gene Activity in Pattern Sets (CoGAPS) to define the transcriptional modules that shape NK cell emergence (39). Of the 15 CoGAPS patterns defined from all timepoints, 9 were time- and cell type-specific patterns that associated with hematopoietic or mesodermal signatures (Fig. 5A). The other six were high expression in all samples or single sample specific patterns with no specific biological signatures. Genes associated with pluripotency, including *POU5F1*, *LIN28A*, and *EPCAM*, defined a transcriptional signature of pluripotent (day 0) samples, named 'pluripotent 1' pattern (Table S2) and reflected by GO term enrichment for pluripotency pathways (Fig. 5B). The other pluripotent pattern was enriched in tRNAs, without specific GO term enrichment (Table S3). In the CD34$^+$ day 6 pattern, genes linked to vascular development and hematopoiesis, including CD34, *CDH5*, *ECSCR*, were highly ranked (Table S2, Fig. 5B). CoGAPS patterns for CD45$^+$ cells at subsequent time points showed enrichment for immune-related genes, and immune related GO signatures were predominant in CD45$^+$ cells past day 14 (Fig. 5B). NK cell associated genes, namely *GNLY*, *GZMA*, *GZMB*, *GZMK*, *NKG7*, and *PRF1*, dominated the transcriptional signature in day 42 CD45$^+$ cells (Table S2, Fig. 5B). CD45$^-$ samples consistently had mesodermal transcriptional signatures, although the identity of these signatures also changed through time. These patterns defined a dynamic yet orderly progression of differentiation from pluripotency to CD45$^+$ hematopoietic lineages along with the CD45$^-$ non-hematopoietic mesodermal cells.

As the CoGAPS patterns indicate, our experiment shows that iPSCs exit pluripotency to commit to a mesoderm fate followed by the induction of hemo-endothelial progen-

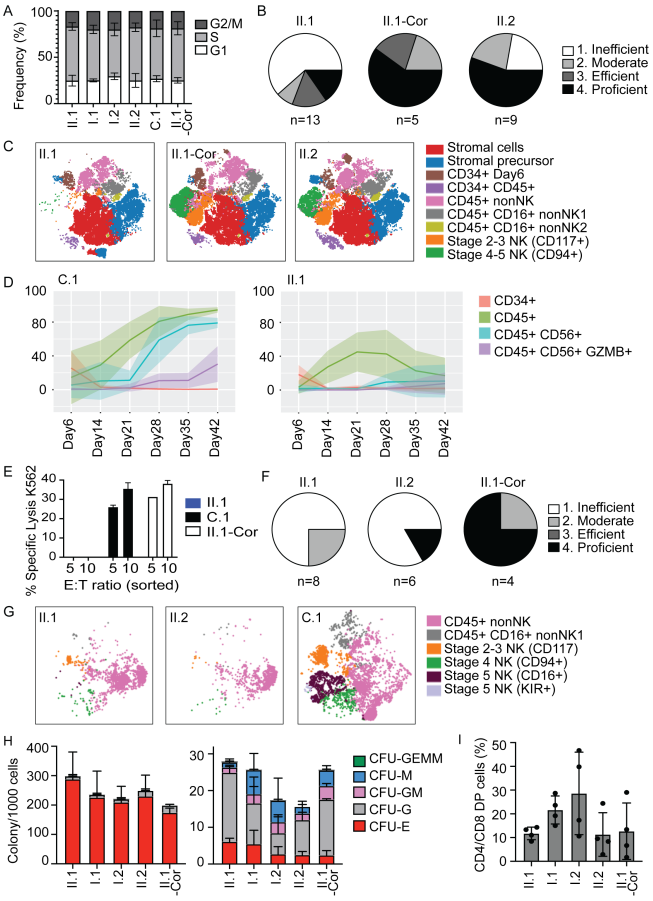

Fig. 4. *GINS4* variant lines show limited potential for NK cell lineage NK cells were differentiated from proband (II.1) or control iPSCs using the spin-EB method or co-culture of sorted CD34$^+$ cells with OP9 stroma and differentiation was monitored at weekly time points by flow cytometry. A) Cell cycle profiles of six iPSC lines measured by DAPI and EdU staining by flow cytometry. n=3-5 technical replicates, error bars standard deviation (SD). B) Summary of the differentiation efficiency in the II.1, II.1-Cor, and II.2 lines showing the NK cell differentiation efficiencies on autologous stroma. n=5-13 technical replicates. C) tSNE plot of the flow cytometry data combined for all timepoints showing the cluster distribution in II.1, II.1-Cor and II.2 lines. D) Line graphs showing the frequency of cells expressing hematopoietic lineage and NK cell specific markers (C.1, n= 9; II.1, n=13 technical replicates, ribbon: SD). E. Bar graph showing K562 target cell killing from sorted CD45$^+$ cells on day 42 performed by Cr51 assay (n=3 technical replicates, error bars: SD). F) Summary of the differentiation efficiency in the II.1, II.2, and II.1-Cor lines showing NK cell differentiation efficiencies on OP9 (n=4-8 technical replicates). G) tSNE plot of the flow cytometry data with markers further dissecting CD45$^+$ populations combined from all timepoints of differentiation from II.1, II.1-Cor and II.2 lines. H) Colony forming assays showing the lineage potential of each cell line when modeling primitive hematopoiesis (left) or definitive hematopoiesis (right) (n=3-9 technical replicates, error bars: SD). I) Frequency of CD4$^+$ CD8$^+$ double positive T cells arising from each cell line (n=4 technical replicates, error bars: SD).

itors and then production of hematopoietic lineage cells. This sequence of lineage commitments resembles embryonic hematopoiesis. Therefore, to understand how in vitro differentiation trajectories reflect early development, we projected the 8 timepoint- and condition-specific differentiation patterns, namely Pluripotent1, CD34$^+$ Day6, CD45$^+$ Day14, CD45$^+$ Day21, CD45$^+$ Day42, CD45$^-$ Day14, CD45$^-$ Day21, and CD45$^-$ Day42 patterns, to mouse early embryo scRNAseq data (40). As expected, the Pluripotent1 pattern correlated with epiblast cells, and the CD34$^+$ day 6 pattern correlated with endothelium and haemato-endothelial progenitors (Fig. 5C). The CD45$^+$ cell signatures moved from endothelium to blood progenitor clusters through time. CD45$^-$ cell signatures changed significantly, spanning mesenchyme and heart at day 14, extraembryonic endoderm at day 21, and back to mesenchyme on day 42 (Fig. 5C).

To further dissect the CD45$^+$ cell signatures beyond the blood progenitor stage, we adapted the summarized human bone marrow cell atlas scRNA-Seq data cataloging 35 cell types from 8 individuals (41). In a tSNE plot, 18 subtypes of CD34$^+$ progenitor cells clustered closely in the center, while more differentiated cell types formed four clustered around the periphery (Fig. 5D). Projection of our day 6 CD34$^+$ pattern showed they were aligned with CD34$^+$ progenitor populations, specifically HSCs and lymphoid biased cells (pre-B cycling, CLP, HSC, HSC cycling, pre-B, and unknown lymphoid) (Fig. 5D). The CD45$^+$ cell signature on day 14 aligned with more specified CD34$^+$ progenitor clusters, including CD34$^+$ megakaryocyte progenitor and CD34$^+$ granulocyte progenitors. On day 21, the CD45$^+$ signature projected to granulocytes, with the most enrichment in the neutrophil, immature neutrophil, and monocyte subsets, confirming the transient presence of myeloid cells that precede NK/ILC generation detected by flow cytometry (Fig. 1). Finally, the day 42 signature aligned most closely with NK cells (Fig. 5D). CD45$^-$ signatures from our data consistently aligned with the stromal cell cluster, further confirming their non-hemogenic identities. Together, these projections relate the heterogeneous cell populations that emerges in iPSC-derived culture to in vivo hematopoietic populations, further confirming our characterization based on flow cytometry phenotypes.

Next, we sought to understand how the *GINS4* variants affect transcriptional dynamics during NK cell emergence. As with our flow cytometry data, we observed variation within healthy individuals in the relative expression of transcriptional signatures, especially later in differentiation (Fig. 5E). When considering CoGAPS signatures, the II.1 line samples fell within the normal variation until day 42, at which point the CD45$^+$ population diverged from the other lines (Fig. 5E). II.1 CD45$^+$ samples showed the lowest expression of the day 42 CD45$^+$ signature, with no expression of mature NK cell markers *GNLY*, *GZMA*, *GZMB*, *GZMK*, *IL32*, or *NCAM1*. Instead, the II.1 CD45$^+$ sample was transcriptionally more like day 42 CD45$^-$ cells of other cell lines, with high expression of *CDH2*, *GATA4*, *INHBA*, and *FN1* (Fig. 5E). As expected, based on our flow cytometry data, re-expression of GINS4 in the II.1-Cor line rescued these transcriptional phenotypes. The II.2 line efficiently generated NK cells in this experiment and the NK cells were transcriptionally indistinguishable from the NK cells made in the control lines, aligning with the C.1 and II.1-Cor CD45$^+$ samples in the day 42 CD45$^+$ pattern. This result suggests that the GINS4 variants specifically affects the emergence of NK cells, rather than the phenotype of mature NK cells.

***GINS4* variant impair cell cycle progression of committed NK cell progenitors.** To better understand the role of GINS4 during NK cell differentiation, we measured GINS4 protein expression through differentiation from iPSCs. West-

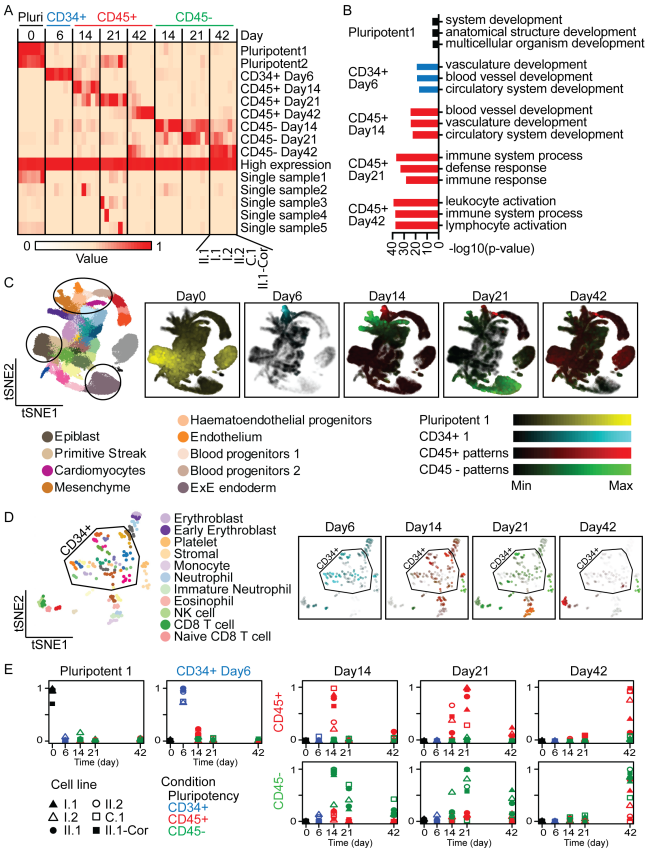

Fig. 5. Transcriptional signatures of differentiating iPSCs align with the in vivo developmental and hematopoietic cell identities Bulk RNA-Seq was performed on iPSCs, sorted CD34$^+$ cells from day 6 and sorted CD45$^+$ or CD45- cells at days 14, 21 and 42. Data were analyzed with non-normalized matrix factorization (CoGAPS). A) Heatmap representing the 15 CoGAPS patterns detected from all samples. Each pattern is named based on the condition with the highest expression of the signature. B) GO terms in hematopoietic lineage related patterns (3 picked from top 6 GO terms with the lowest p-values). C-D. Projection of the time point and condition specific eight CoGAPS patterns into (C) early mouse development scRNA-seq data (40) or (D) human bone marrow atlas data (41). Opacity and color represent higher enrichment of each pattern. E) Eight time point and condition specific CoGAPS patterns coded by sample and timepoint. Cell lines are represented by shapes and conditions are represented by color.

ern blot analysis confirmed significantly decreased GINS4 protein expression in II.1 and II.2 lines, showing the greatest difference from the control lines at day 6 due to the high expression of GINS4 at days 6 and 14 in controls (Fig. 6A, B) (12). During differentiation, GINS4 protein levels decreased in control lines, consistent with a decrease in transcription measured by RNA-seq (Fig. 6C). As a result of this decrease in expression coinciding with differentiation, by day 21 GINS4 protein levels did not differ significantly between the control and *GINS4* lines (Fig. 6A, B).

Since *GINS4* is a DNA helicase component (19) and critical for S phase progression, we sought to better understand the cell cycle dynamics associated with the impaired generation of NK cells in the II.1 line. RNA-seq data showed transcriptional patterns of S phase genes (42, 43) that were similar to *GINS4*, namely the conserved pattern of high expression in pluripotency and CD34$^+$ day 6 cells, which then decrease through differentiation in both CD45$^+$ and CD45$^-$ cells (Fig. 6D). While decreasing in general, the CD45$^+$ population still showed more expression in these genes than the CD45$^-$ population at all time points (Fig. 6D).

We next sought to investigate the effect of *GINS4* variants on cell cycle progression. Our previous studies have identified mild cell cycle impairment in B cell lines derived from II.1 and II.2, namely delayed re-entry to cell cycle detectable by BrdU-EdU pulse labeling (12). Consistent with previous reports and our transcriptional data, pluripotent cells were highly proliferative, and labeling with EdU defined a high frequency of pluripotent cells in S phase (Fig. 4A). Despite the highly proliferative nature of iPSCs, we found no significant impairment in cell cycle progression between patient-derived and healthy donor or isogenic control lines. Similarly on day 6, the frequencies of CD34$^+$ cells were comparable between control and *GINS4* variant lines and within CD34$^+$ or CD34$^-$ cells the frequencies of EdU$^+$ populations were comparable (Fig. S2), suggesting no cell cycle impairment by this point.

To further dissect the cell cycle dynamics during NK cell generation, we measured Ki-67 levels during differentiation. Consistent with the rapid emergence of CD45$^+$ myeloid cells at days 14 and 21, a significant frequency of CD45$^+$CD56$^-$ cells were Ki-67$^+$ at day 14 (Fig. 6E). CD45$^+$ cells from the healthy donor, C.1, and isogenic control line, II.1-Cor, underwent significant proliferation at day 28, especially in the CD45$^+$CD56$^+$ cell population (Fig. 6E, F). On the other hand, the CD45$^+$ cells from the II.1 line did not show an increase in Ki-67$^+$ cells between days 21 and 28. In the control lines, this proliferation of CD56$^+$CD45$^+$ cells decreased from day 28 through days 35 and 42, supporting our transcriptional data suggesting that proliferation decreases following an initial expansion of precursors early in NK cell differentiation. While these data were obtained from in vitro cultures, Ki-67 staining of NK cell developmental subsets from tonsil has also identified NK cell progenitors (stages 1 and 2) as consistently being most proliferative relative to later stages of NK cell maturation (44).

Next, we performed single cell transcriptional profiling (scRNAseq) on all cells from the II.1 and C.2 lines at day 28 of spin-EB differentiation. 2,530 cells from II.1 and 532 cells from C.2 passed quality control thresholds. While the overall resolution of the data was relatively low, Seurat analysis revealed 6 clusters conserved between healthy donor and control lines (Fig. 6G)(45). Based on the highly expressed genes and the GO analysis from each cluster, we identified stromal cells with vascular, heart and muscle related phenotypes that were similar to those identified by bulk RNA-Seq in Fig. 5. Cluster 6 was enriched for genes associated with cell division and an NK cell related transcriptional signature (Fig. 6G, Table S5). Within this cluster, we detected *CD52*, a mature lymphocyte marker, *PRF1*, and *GZMA* expressing cells in the C.2 line but not the II.1 line, indicative of lineage specification and maturation occurring only in the control line cells. Instead, the cluster 6 cells from the II.1 line expressed cell cycle enriched signature defined by genes including *CENPF*, *CENPE*, *TOP2A*, and *MKI67* (Fig. S3). Assessing the predicted cell cycle states by gene expression revealed that clus-

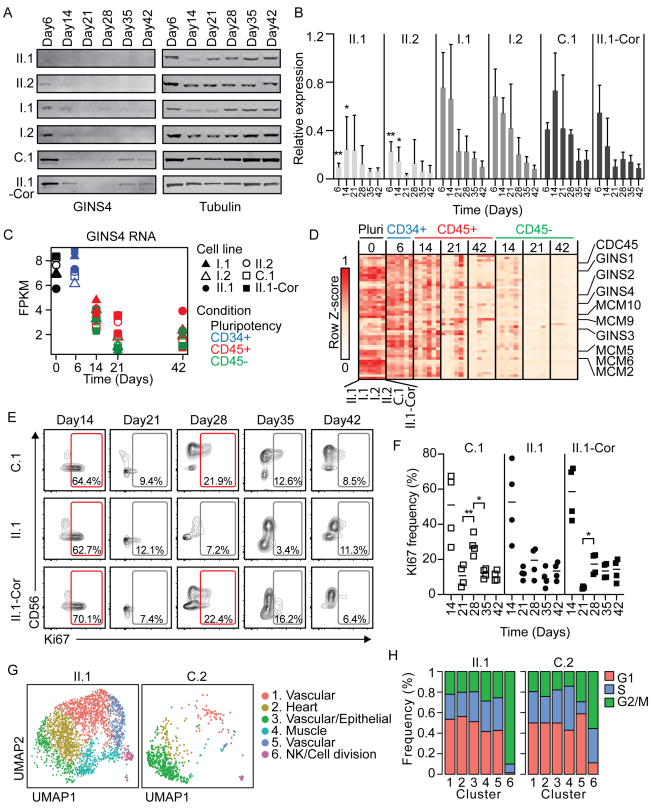

Fig. 6. Committed NK progenitors from the GINS4 proband have impaired cell cycle progression Expression of the cell cycle related genes were monitored by western blot, flow cytometry, and RNAseq during NK cell differentiation from iPSCs A) Representative western blots showing GINS4 protein levels at weekly timepoints of NK cell differentiation in six lines (*GINS4* family, C.1, and II.1-Cor). B) Quantification of GINS4 protein levels from blots shown in (A) (n=6-8 technical replicates, *<0.05, **<0.01 p-value represents II.1 or II.2 compared to all control lines combined, error bars: SD). C) *GINS4* RNA expression from RNAseq data with each sample coded by color and condition. D) Heatmap of transcriptional levels of S-phase related genes from all cell lines at timepoints and conditions as indicated (42, 43). E) Representative flow cytometry plots showing Ki-67 and CD56 expression in CD45$^+$ cells of C.1, II.1, and II.1-Cor lines at weekly time points through NK cell differentiation. F) Quantification of the frequency of CD45$^+$Ki-67$^+$ populations from flow cytometry shown in (E). n=4 technical replicates, *<0.05, **<0.01, horizontal lines indicate average. G) UMAP of scRNAseq showing 6 clusters from II.1 and C.2 lines on day 28. H) Predicted frequencies of cell cycle phase from analysis of scRNASeq data shown in (G).

ter 6 cells had a signature suggesting accumulation in G2/M phase. Within cluster 6, II.1 line cells showed even higher percentage (90.1%) of cells predicted to be in G2/M compared to the C.2 cluster 6 cells (55.6%, Fig. 6H). Together, these results demonstrate that lack of GINS4 specifically affects proliferation and inhibits expansion of NK cell precursors observed at the day 28 timepoint of differentiation.

***GINS4* variants lead to apoptosis in response to proliferation after NK cell lineage commitment .** To confirm if the NK cell progenitors from the II.1 line accumulate in G2/M phase of the cell cycle, we measured the DNA content of the cells every 24 hours between days 21 and 28, when the first CD56$^+$ NK cells start to emerge (Fig. 7A). CD45$^+$CD56$^+$ NK lineage cells in both II.1 and the isogenic control (II.1-Cor) lines, corresponding to cluster 6 of the scRNA-seq (Fig. 6G, H), had a greater frequency of G2/M cells when compared to the non-hematopoietic CD45$^-$ or mixed hematopoietic CD45$^+$CD56$^-$ populations (Fig. 7A, B). This flow cytometric approach revealed that all populations of II.1 cells accumulate in G2/M to a greater extent when compared to the II.1-Cor cells (Fig. 7B). The CD45$^+$CD56$^-$ and CD45$^+$CD56$^+$ populations had 2.9±0.5 and 2.7±1.3-fold increase, respectively, while the change in CD45$^-$ population was the most subtle, with a 1.5±0.14-fold increase. The biggest difference in G2/M frequency was observed in the day 28 samples of CD45$^+$CD56$^+$, cells where the II.1 line had 5.6-fold more cells in G2/M, suggesting that a more severe cell cycle phenotype occurs as more NK cells arise with time. Together with the Ki-67 measurements and the scRNA-seq data, these results confirm that the NK cell precursors of the II.1 line are generated but accumulate in the G2/M phase without being able to proliferate and progress to mature NK cells.

To further address the question of 'why NK cells', we hypothesized that NK cells may be more sensitive to programmed cell death after accumulating at G2/M phase of the cell cycle. In the bone marrow cell atlas scRNA-Seq data with 35 cell types, NK cells showed the highest accumulative expression of pro-apoptotic genes, followed by monocytes, platelets, and CD8 T cells (Fig. 7C, Fig. S4A, gene list: Table S4). NK cells also highly expressed anti-apoptotic genes, in addition to platelets, dendritic cells and monocytes, suggesting that they may be poised for apoptosis while the expression of anti-apoptotic genes keeps the apoptosis in check (Fig. S4B). In our day 28 scRNAseq cluster 6, II.1 cells showed more pro-apoptotic gene expression than control (C.2) cells, whereas they showed comparable anti-apoptotic gene expression (Fig. S4C).

To test if this increased apoptotic signature is causing the NKD phenotype, we measured phospho-p53 (Ser15) levels in cells between days 21 and 28 as an early apoptotic marker (46). CD13, CD14, and CD15 (Lin) antibodies were used to distinguish neutrophils or monocytes from early NK progenitors in the CD45$^+$CD56$^-$ population. As demonstrated by FSC, CD45$^+$CD56$^-$Lin$^-$ cells were smaller in size than the Lin$^+$ cells, suggesting that they have entered lymphoid lineage (Fig. 7D). Within the CD45$^+$CD56$^-$Lin$^+$ population, the frequency of the phospho-p53 (Ser15)$^+$ population did not differ between the II.1 and II.1-Cor lines (Fig. 7E). In contrast, CD45$^+$CD56$^-$Lin$^-$ cells from II.1 had a higher frequency of phospho-p53 (Ser15)$^+$ cells than those from the isogenic control (II.1-Cor). This difference became more pronounced as NK precursor cells matured to express CD56 (Lin$^-$CD56$^+$), where the median phospho-p53 (Ser15)$^+$ frequency was 12.2% in the II.1 and 7.6% in the II.1-Cor lines, respectively. This result clearly shows increased apoptotic activity in the II.1 cells that are committed to the NK cell lineage, even before CD56 expression. In summary, these experiments between days 21 and 28 demonstrate that the NKD phenotype of the II.1 line arises from cells experiencing replication stress shortly after undergoing NK cell lineage commitment. As these cells are impaired in their ability to proliferate, they subsequently undergo apoptosis.

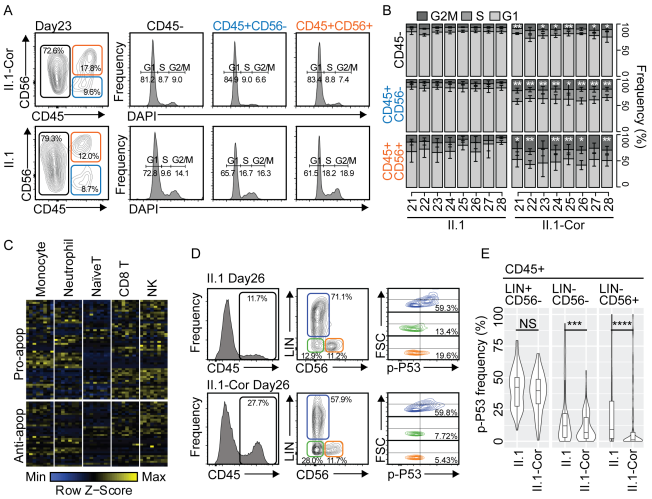

Fig. 7. *GINS4* variants lead to apoptosis in response to proliferation after NK cell lineage commitment. Cell cycle and apoptotic signal were monitored by flow cytometry during NK cell differentiation from iPSCs using the spin-EB method. A) Representative flow cytometry data showing CD45 and CD56 population gating and DAPI histograms for DNA content of each gated population. B) Stacked bar graph showing cell cycle profiles quantified from DAPI histograms as shown in (A). n=6-8 technical replicates, *<0.05, **<0.005, p-values represent G2/M frequencies of II.1 lines compared to II.1-Cor lines from the same time point, error bars: SD. C) Heatmap showing summary expression of apoptosis genes (KEGG APOPTOSIS, M8492) in select samples from the human bone marrow atlas data (41). D) Representative flow cytometry data showing CD45, Lin (CD13, CD14, and CD15), and CD56 population gating, and phospho-p53 (Ser15) and FSC scatter plots for each gated populations (n=6-8 technical replicates). E) Frequency of phospho-p53 (Ser15)$^+$ cells in each population from all timepoints combined (n=6-8 technical replicates, ***<0.01, ****<0.001, box plots mark median, Q1 and Q3).

Discussion

Previous studies of NK cell differentiation from iPSCs have focused on the generation of functional and mature NK cells and the efficiency of differentiation (27–29, 38, 47). While these studies highlight important milestones for generating iPSC-NK cells that can be used for cancer immunotherapy, iPSC-derived NK cells also offer the potential for understanding NK cell development in the context of NKD (48). Here, we dissect the emergence of heterogeneous hematopoietic populations through the differentiation of NK cells from multiple iPSC lines using the previously described spin-EB method (27). Our work adds developmental context to understanding the ontogeny of NK cells and defines how helicase gene variants can impair NK cell development to induce NKD (12, 27).

Comparison with human bone marrow cell atlas data showed that CD34$^+$ cells from iPSCs were transcriptionally similar to multiple CD34$^+$ progenitor populations, including pre-B cycling cells, CLPs, and HSCs (41). Such a diffuse signature may indicate heterogeneity within the in vitro CD34$^+$ population, that the CD34$^+$ cells we have generated have an ambiguous identity with lymphoid biased lineage potentials, or that these data represent the difference between embryonic hematopoietic progenitors modeled by iPSCs and adult hematopoietic progenitors from the bone marrow.

Unlike CD34$^+$ precursors, CD45$^+$ day 21 or day 42 signatures mapped specifically to granulocytes or NK cells, and we confirmed these identities with protein expression and functional assays. The early emergence of myeloid cells and the later emergence of lymphoid cells parallels the order of cell emergence during primitive and definitive hematopoiesis and after hematopoietic stem cell transplant (49). During ischemic organ injury or systemic infection, HSCs expand and relocate to undergo emergency myelopoiesis (50, 51). The bias towards myelopoiesis under stress suggests that the myeloid lineage may be a default pathway for hematopoietic progenitors, and the limited capacity for T cell generation from iPSCs has been previously documented.

Other studies have shown that primitive hematopoiesis and definitive hematopoiesis can be distinctly modelled from iPSCs by regulating WNT signaling (29, 30, 38). We have adapted these methods and observed distinct lineage potentials in the colony forming assay. Using WNT agonist during the first 6 days of differentiation, our protocol is more similar to the definitive hematopoiesis protocol, yet we observe the myeloid lineage cells first. Understanding in vivo embryonic hematopoiesis in more detail and developing protocols that accurately and distinctly model the primitive, definitive, and adult hematopoiesis from iPSCs will significantly advance the usage of iPSCs for modeling normal development and disease.

We identified a proliferative stage of NK cell development prior to the emergence of stage 2-3 NK cells, and the *GINS4* variant lines showed G2/M accumulation as the cells failed to proliferate at this timepoint. This cell cycle defect was only observed in the CD45$^+$ cells at day 28, suggesting that despite the conserved expression and utility of the CMG helicase, the impact of replication stress generated by *GINS4* variants is cell type specific. This proliferative timepoint likely corresponds to proliferation of NK precursors at an analogous stage of differentiation (44), and the late timing of this event in our cultures further underscores the delayed onset of NK cell generation following the wave of myeloid cells observed from days 14-28.

Complete knockout mouse models for *MCM4*, *MCM10*, *GINS1*, and *GINS4* are embryonic lethal, and individuals with NKD variants have partial preservation of protein function (9–12, 17, 52–55). Furthermore, the study of *GINS1* variants revealed that the extent of other cell types aside from NK cells that were affected negatively correlated with the estimated residual CMG activity (11). Together, these data suggest that different cell types require different levels of CMG helicase activity and that NK cells are primarily affected because they are especially vulnerable to such loss of helicase expression or function. Investigating the other cell types affected in the additional common clinical phenotypes, such as adrenal insufficiency and neutropenia, may uncover similar sensitivities (9, 11, 56).

The proband in our study (II.1) and his sister (II.2) have the same *GINS4* genotype and show decreased NK cell frequencies in PBMC, yet the clinical phenotype of II.1 is more severe, with reported viral infections in utero and childhood growth deficiency (12). When differentiated on autologous stroma, the II.2 line showed a better efficiency of generating NK cells (55%) than the II.1 line (15%). However, on OP9

stroma, both cell lines were similarly inefficient in generating NK cells, and the rescue of the phenotype of II.1 by CRISPR correction of *GINS4* variant alleles demonstrates that modifying variants are not the source of this variability.

Analyses of CD45$^-$ autologous stromal cells from all donors show significant changes in transcriptional signatures at different time points, suggesting a dynamic stromal-hematopoietic interaction that supports NK cell differentiation differently from the OP9-hematopoietic interaction. While the nature of stroma generated in this system has not been well characterized, these results show that cell-cell interactions or epigenetic regulation outside of genetic variation could play a critical role in improving the survival of NK cell progenitors even under replication stress and contribute to donor variability. Alternatively, environmentally induced epigenetic factors could play a role in determining severity of the cellular and clinical phenotypes of siblings with shared genotypes. In support of this hypothesis, we observed higher efficiency with PBMC-derived iPSC lines than fibroblast-derived iPSC or ES lines. As previously reported for iPSCs from different sources, our PBMC-iPSC lines may carry residual hematopoietic cell epigenetic memories that bias their hematopoietic potential (57–59). As a result, we have exclusively used PBMC-derived iPSCs in our experiments comparing multiple cell lines. Among the PBMC-derived control lines, we still observed variations in differentiation efficiency, validating previous reports of the effect of individual variation in lineage potential (36, 37). A method for generating iPSCs to minimize effects of residual epigenetic memory has been proposed (60), and 3D models that resemble early human embryo are being developed (61–63). Together, these technical advances will allow the more accurate modeling of embryonic hematopoiesis.

Regardless of limitations, iPSC models provide the opportunity to study cell type-specific human diseases which are often difficult to understand due to the restricted availability of primary cells. In this study, we provide the first in-depth transcriptional and phenotypic characterization of the trajectory of human NK cell differentiation from pluripotent stem cells. In doing so, we have defined the unique sensitivity of proliferative lineage committed NK cell precursors to impaired cell cycle progression. This finding explains the unexpected, isolated NK cell deficiency resulting from partial loss-of-function helicase variants and defines differential regulation of replisome function as a cell biological mechanism underlying disease (12).

Methods

Cell culture. Human iPSCs were cultured in feeder-free condition with Stemflex (Gibco, A3349401) or mTeSR plus (Stemcell Technologies, 100-0276) at 37°C with 5% CO2. Cells were dissociated to single cells with Accutase (Sigma, A6964), plated at a density of 1 X 10^6 cells/well in a Matrigel-coated 6 well plates (BD, 354277; Fisher Scientific, 140675). For the first 24 hours after passaging, cells were cultured with 5μM ROCK inhibitor (Y27632; Stemcell Technologies, 72304) and passaged every five days in culture. OP9 cells (ATCC CRL-2749) were cultured in nongelatinized culture flasks (Fisher Scientific, 10-126-28) in α-MEM (Fisher Scientific, 12561056) with 20% heat-inactivated FBS (Gemini, 900-108) and 1% penicillin/streptomycin (100 U/ml; Fisher Scientific, 15140163). For co-culture differentiation experiments, OP9 cells were seeded into 96-well flat-bottomed plates (Genesee, 25-109). Once confluent, the cells were subjected to mitotic inactivation by irradiation at 300 Gy prior to co-culture with precursors.

Generation of iPSC lines. Human iPSCs were generated from peripheral blood mononuclear cells from *GINS4* affected individuals (12) or unrelated adult healthy donor controls using StemRNA Gen Reprogramming Kit (Stemgent) as previously described (10). Cells were validated for pluripotency and periodically karyotyped and tested for mycoplasma.

Generation of isogenic iPSC line. *GINS4* target editing was performed by using CRISPR Cas-9 as previously described (64). Briefly, target-specific 20bp sgRNA was designed using a freely available webtool (https://design.synthego.com/) and commercially synthesized from Integrated DNA Technology (IDT). Single-stranded donor DNA was designed and commercially synthesized from (IDT). For electroporation, RNP complex was prepared with Lonza P3 Primary Cell 4D-Nucleofector X Kit according to the manufacturer's protocol (Buffer 75 μl, Supplement A 18 μl, sgRNA 3 μl (15 μg), Cas9 10 μg, ssODN 3 μl (15μg)), and the reaction mix was incubated for 25 minutes at RT. iPSCs were dissociated with Accutase, and 10^6 cells were mixed with RNP complex. The mixture was electroporated using CA137 program in Amaxa 4D nucleofector (Lonza). Cells were incubated for survival, and low-density single cell seeding was performed. Screening was done by colony picking and sanger sequencing of target regions. Edited clones were picked and further characterized.

NK cell differentiation. NK cell differentiation from iPSC was performed as described in Knorr et al. (27). Briefly, 9000 cells per well were plated on 96 well polystyrene round bottomed microwell plates (Fisher Scientific, 262162) for EB formation in APEL-2 media (Stemcell Technologies, 5275) containing 10μM ROCK inhibitor, 40ng/ml SCF (Peprotech, 300-07), 20ng/ml BMP (RD, 314-BP), 20ng/ml VEGF (RD, 293-VE), and 1μM GSK inhibitor (CHIR99021; Stemcell Technologies, 72052). EBs were cultured for six days.

For autologous stroma supported differentiation, day 6 EBs were transferred to gelatin (Millipore, ES-006-B) coated 24 well plates (4EB/well; Fisher Scientific, 08-772-1C) with NK cell differentiation basal media supplemented with 5ng/ml IL-3 (1st week only; Peprotech, 200-03), 10ng/ml IL-15 (Peprotech, 200-15), 10ng/ml Flt3L (Peprotech, 300-19), 20ng/ml SCF, and 20ng/ml IL-7 (Peprotech, 200-07). NK cell differentiation basal media is made with 55% DMEM (Fisher Scientific, 11965118), 28% Hems F12 (Fisher Scientific, 10-080-CV), 15% human AB serum (GeminiBio,

100-512), 1% L-Glutamax (Fisher Scientific, 35050079), 1% Penicillin/streptomycin (100U/ml), 25uM BME (Fisher Scientific, 21985023), 50 uM Ethanolamine (Sigma, E0135-100ML), 25 µg/mL Ascorbic acid (Sigma, A5960-25G), and 5 ng/mL sodium selenite (Sigma, S5261-25G). After transfer, the cells from EB propagate and differentiate to NK cells in the 24 well plate for 35 days with weekly half media change. For OP9 supported differentiation, day 6 EBs were dissociated with TrypLE (Fisher Scientific, 12605028), and $CD34^+$ cells were sorted (BD FACS Aria II). 2 X 10^3 $CD34^+$ cells were seeded on each well of 96 well plates with irradiated OP9 (ATCC) stroma. For differentiation, the same differentiation basal media and cytokine supplements as differentiation supported by autologous stroma were used for 35 days with weekly half media change. T cell differentiation was performed using STEMDiff T cell Kit (Stemcell Technologies, 100-0194) according the manufacturer's protocol.

CFU assay. Methods were adapted from Dege et al. (30) for CFU assays. 9000 cells were seeded in each well of round bottom 96 well plates for EB formation in APEL-2 media with 10 ng/ml BMP4. On day1, bFGF (final concentration 5 ng/ml; RnD, 233-FB-025) was added. 18 hours later, final concentrations of either 3 µM CHIR99021 for definitive hematopoietic differentiation or 3 µM IWP2 (Selleckchem, S7085) and 1 ng/ml activin A (Stemcell Technologies, 78001) for primitive hematopoietic differentiation were added to the media for 30 hours. Then, for three days, the EBs were placed in APEL-2 media with 15ng/ml VEGF and 5ng/ml bFGF. On day6, final concentrations of 100ng/ml SCF, 2IU/ml EPO (RnD, 287-TC-500), 10ng/mL IL-6 (Peprotech, 200-06), 5ng/ml IL-11 (Peprotech, 200-11), and 25ng/ml IGF-1 (Peprotech, 100-11) were added to the culture media. EBs were cultured in 37°C, 5% CO2, and 5% O2 condition.

On day 8, CHIR99021 treated EBs were dissociated with TrypLE and $CD34^+/CD43^-/CD73^-/CD184^-$ population was sorted and 10^4 cells/well were plated on low 96-well low-adherent cell culture plates (Fisher Scientific, 07-200-603) in HE media (APEL-2 with 5ng/ml VEGF, 5ng/ml bFGF, 100ng/ml SCF, 2IU/ml EPO, 10ng/ml IL-6, 5ng/ml IL-11, 25ng/ml IGF-1, 30ng/ml TPO (Peprotech, 300-18), 10ng/ml BMP4, 20ng/ml SHH (RnD, 1845-SH-025), 10µg/ml angiotensin II (RnD, 1158/5), 100µM losartan potassium (RnD, 3798/50)). Once the cells aggregate into clumps, they were transferred to Matrigel coated 24 well plate in HE media. On day 20, $CD34^+/CD45^+$ population was isolated with FACS sort.

On day 8, IWP2 treated EBs were put on day8 media (APEL-2 with 100ng/ml SCF, 2UI/ml EPO, 10ng/ml IL-6, 5ng/ml IL-11, 25ng/ml IGF-1, 30ng/ml TPO, 10ng/ml Flt-3L, and 30ng/ml IL-3) for two days. Then on day 10, $CD43^+/CD34^-$ population was isolated with FACS. Sorted cells were plated in Methocult optimum media (2-3 X 10^3 cells/well; Stemcell Technologies, H4034) according to manufacturer's protocol. Colonies were counted in 7 - 10 days.

Flow cytometric analysis and sorting. Flow cytometry was performed at stated time points using antibodies listed in Supplementary Table1. Cells were dissociated with TrypLE, passed through a 35 µm filter to remove clumps, then incubated with surface receptor antibodies and a fixable viability dye. After a wash, cells were fixed and permeabilized using CytoFix/CytoPerm (Fisher Scientific, BDB554714) or Foxp3/Transcription factor staining buffer (Thermo Fisher, 00-5523-00) according to manufacturer's protocol then cells were incubated with intracellular antibodies. For DAPI staining, the stained cells were resuspended in a DAPI solution containing (0.1% BSA, 0.5µg/ml DAPI, 0.02% NaN3, and 250µg/ml RNase A). For EdU staining, cells were incubated with 10µM EdU for 2h before harvesting, and the signals were detected using Click-iT™ EdU Alexa Fluor 488 Flow Cytometry Assay Kit (Thermo Fisher, C10420) according to manufacturer's protocol. Data were acquired on a Bio-Rad ZE5 Cell Analyzer and analyzed using FlowJo 10 (BD Biosciences) or OMIQ (Dotmatics). tSNE was performed on OMIQ using all tested markers. For sorting, dissociated cells were incubated with surface receptor antibodies. After a wash, cells were resuspended in PBS with propidium iodide (1:200), and the desired populations were sorted on a BD Aria II cytometer with a 100 µm nozzle. Sorted cells were cultured directly after isolation.

RNA-seq. Dissociated cells were stained with anti-CD34 or anti-CD45 antibody and PI. PI^-CD34^+, PI^-CD34^-, PI^-CD45^+, and PI^-CD45^- cells were sorted by flow cytometry with BD Aria II cytometer. 2 10 X 104 cells were lysed, and RNA was collected with NucleoSpin RNA XS kit (Takara, 740902) according to the manufacturer's protocol. RNA quality control was performed using the Agilent 2100 Bioanalyzer System. Libraries were generated with SMARTer stranded RNA-Seq kit (Takara, 634839) and sequenced on NovaSeq 6000 (Illumina). The FASTQ files were generated using FastQC (v0.11.7) and trimmed reads (Trimmomatic 0.38) were mapped to UCSC hg38 with HISAT2 (v2.1.0, Bowtie2 2.3.4.1), splice-aware aligner. The sequencing depth is between 58 and 80 million mappable reads. Transcript is assembled by StringTie (v2.1.3b) and abundance of genes are calculated as read counts and FPKM. CoGAPS (v3.16.0) was run on log2 RNA-seq FPKM data using default parameters as previously described, for a range of k patterns (k=15 selected) and uncertainty as 10% (39, 65). CoGAPS patterns were projected to other datasets using the defaut projectR function in the projectR package (v1.12.0) (66).

scRNA-Seq. Day28 spin-EB culture was dissociated with TrypLE, and the cells were collected without sorting. Libraries are prepared using Chromium Next GEM Single Cell 3' GEM, Library Gel Bead Kit (v3, 10X genomics) and sequenced. The reads were aligned to UCSC hg38 and counted using Cell Ranger (v5.0.1). UMAP dimension reduction, clustering, and cell cycle profiling was done using Seurat (v4.3.0.1).

Western blotting. Samples were collected and sonicated in RIPA buffer (Fisher Scientific, 89900) with protease inhibitor (Fisher Scientific, 78443). The protein concentration was measured with Pierce BCA protein assay kit (Fisher Scientific, 23225). The same amount of protein was loaded in each well. The samples were run on 12% Bis-Tris gels and transferred onto nitrocellulose membranes (Fisher Scientific, LC2000). The membranes were blocked and incubated with primary antibody overnight at 4C. After secondary antibody incubation for 30min at room temperature, the blots were imaged with Li-COR and analyzed with ImageJ.

ACKNOWLEDGEMENTS

We would like to thank the staff of the Columbia Stem Cell Initiative Flow Cytometry Core Facility, under the leadership of Michael Kissner, at Columbia University Irving Medical Center for their contributions to the work presented in this manuscript. Work presented here was partially supported by NIH-NIAID R01AI137073 to EMM. SS was the recipient of a NYSTEM training grant. This work was also supported by McNair Medical Institute funding and Polish National Science Center Funding (OPUS-UMO-2020/37/B/NZ3/0917) to M.B. We thank Dr. Jordan Orange for sharing *GINS4* iPSC lines that were generated with the support of NIH-NIAID R01AI120989.

Bibliography

1. J. R. Ortaldo, R. K. Oldham, G. C. Cannon, and R. B. Herberman. Specificity of natural cytotoxic reactivity of normal human lymphocytes against a myeloid leukemia cell line. *J Natl Cancer Inst*, 59(1):77–82, 1977. ISSN 0027-8874 (Print) 0027-8874 (Linking). doi: 10.1093/jnci/59.1.77.
2. R. B. Herberman, M. E. Nunn, and D. H. Lavrin. Natural cytotoxic reactivity of mouse lymphoid cells against syngeneic acid allogeneic tumors. i. distribution of reactivity and specificity. *Int J Cancer*, 16(2):216–29, 1975. ISSN 0020-7136 (Print) 0020-7136 (Linking). doi: 10.1002/ijc.2910160204.
3. F. Sendo, T. Aoki, E. A. Boyse, and C. K. Buafo. Natural occurrence of lymphocytes showing cytotoxic activity to balb/c radiation-induced leukemia rl male 1 cells. *J Natl Cancer Inst*, 55 (3):603–9, 1975. ISSN 0027-8874 (Print) 0027-8874 (Linking). doi: 10.1093/jnci/55.3.603.
4. R. Kiessling, E. Klein, and H. Wigzell. "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*, 5(2):112–7, 1975. ISSN 0014-2980 (Print) 0014-2980 (Linking). doi: 10.1002/eji.1830050208.
5. L. Moretta, E. Ciccone, M. C. Mingari, R. Biassoni, and A. Moretta. Human natural killer cells: origin, clonality, specificity, and receptors. *Adv Immunol*, 55:341–80, 1994. ISSN 0065-2776 (Print) 0065-2776 (Linking). doi: 10.1016/s0065-2776(08)60513-1.
6. T. M. Ellis and R. I. Fisher. Functional heterogeneity of leu 19"bright"+ and leu 19"dim"+ lymphokine-activated killer cells. *J Immunol*, 142(8):2949–54, 1989. ISSN 0022-1767 (Print) 0022-1767 (Linking).
7. T. Hercend, J. D. Griffin, A. Bensussan, R. E. Schmidt, M. A. Edson, A. Brennan, C. Murray, J. F. Daley, S. F. Schlossman, and J. Ritz. Generation of monoclonal antibodies to a human natural killer clone. characterization of two natural killer-associated antigens, nkh1a and nkh2, expressed on subsets of large granular lymphocytes. *J Clin Invest*, 75(3):932–43, 1985. ISSN 0021-9738 (Print) 0021-9738 (Linking). doi: 10.1172/JCI111794.
8. E. M. Mace and J. S. Orange. Emerging insights into human health and nk cell biology from the study of nk cell deficiencies. *Immunol Rev*, 287(1):202–225, 2019. ISSN 1600-065X (Electronic) 0105-2896 (Linking). doi: 10.1111/imr.12725.
9. L. Gineau, C. Cognet, N. Kara, F. P. Lach, J. Dunne, U. Veturi, C. Picard, C. Trouillet, C. Eidenschenk, S. Aoufouchi, A. Alcais, O. Smith, F. Geissmann, C. Feighery, L. Abel, A. Smogorzewska, B. Stillman, E. Vivier, J. L. Casanova, and E. Jouanguy. Partial mcm4 deficiency in patients with growth retardation, adrenal insufficiency, and natural killer cell deficiency. *J Clin Invest*, 122(3):821–32, 2012. ISSN 1558-8238 (Electronic) 0021-9738 (Print) 0021-9738 (Linking). doi: 10.1172/JCI61014.
10. E. M. Mace, S. Paust, M. I. Conte, R. M. Baxley, M. M. Schmit, S. L. Patil, N. C. Guilz, M. Mukherjee, A. E. Pezzi, J. Chmielowiec, S. Tatineni, I. K. Chinn, Z. C. Akdemir, S. N. Jhangiani, D. M. Muzny, A. Stray-Pedersen, R. E. Bradley, M. Moody, P. P. Connor, A. G. Heaps, C. Steward, P. P. Banerjee, R. A. Gibbs, M. Borowiak, J. R. Lupski, S. Jolles, A. K. Bielinsky, and J. S. Orange. Human nk cell deficiency as a result of biallelic mutations in mcm10. *J Clin Invest*, 130(10):5272–5286, 2020. ISSN 1558-8238 (Electronic) 0021-9738 (Print) 0021-9738 (Linking). doi: 10.1172/JCI134966.
11. J. Cottineau, M. C. Kottemann, F. P. Lach, Y. H. Kang, F. Vely, E. K. Deenick, T. Lazarov, L. Gineau, Y. Wang, A. Farina, M. Chansel, L. Lorenzo, C. Piperoglou, C. S. Ma, P. Nitschke, A. Belkadi, Y. Itan, B. Boisson, F. Jabot-Hanin, C. Picard, J. Bustamante, C. Eidenschenk, S. Boucherit, N. Aladjidi, D. Lacombe, P. Barat, W. Qasim, J. A. Hurst, A. J. Pollard, H. H. Uhlig, C. Fieschi, J. Michon, V. P. Bermudez, L. Abel, J. P. de Villartay, F. Geissmann, S. G. Tangye, J. Hurwitz, E. Vivier, J. L. Casanova, A. Smogorzewska, and E. Jouanguy. Inherited gins1 deficiency underlies growth retardation along with neutropenia and nk cell deficiency. *J Clin Invest*, 127(5):1991–2006, 2017. ISSN 1558-8238 (Electronic) 0021-9738 (Print) 0021-9738 (Linking). doi: 10.1172/JCI90727.
12. M. I. Conte, M. C. Poli, A. Taglialatela, G. Leuzzi, I. K. Chinn, S. A. Salinas, E. Rey-Jurado, N. Olivares, L. Veramendi-Espinoza, A. Ciccia, J. R. Lupski, J. C. Aldave Becerra, E. M. Mace, and J. S. Orange. Partial loss-of-function mutations in gins4 lead to nk cell deficiency with neutropenia. *JCI Insight*, 7(21), 2022. ISSN 2379-3708 (Electronic) 2379-3708 (Linking). doi: 10.1172/jci.insight.154948.
13. C. R. Hughes, L. Guasti, E. Meimaridou, C. H. Chuang, J. C. Schimenti, P. J. King, C. Costigan, A. J. Clark, and L. A. Metherell. Mcm4 mutation causes adrenal failure, short stature, and natural killer cell deficiency in humans. *J Clin Invest*, 122(3):814–20, 2012. ISSN 1558-8238 (Electronic) 0021-9738 (Print) 0021-9738 (Linking). doi: 10.1172/JCI60224.
14. F. Bernard, C. Picard, V. Cormier-Daire, C. Eidenschenk, G. Pinto, J. C. Bustamante, E. Jouanguy, D. Teillac-Hamel, V. Colomb, I. Funck-Brentano, V. Pascal, E. Vivier, A. Fischer, F. Le Deist, and J. L. Casanova. A novel developmental and immunodeficiency syndrome associated with intrauterine growth retardation and a lack of natural killer cells. *Pediatrics*, 113(1 Pt 1):136–41, 2004. ISSN 1098-4275 (Electronic) 0031-4005 (Linking). doi: 10.1542/peds.113.1.136.
15. C. Eidenschenk, E. Jouanguy, A. Alcais, J. J. Mention, B. Pasquier, I. M. Fleckenstein, A. Puel, L. Gineau, J. C. Carel, E. Vivier, F. Le Deist, and J. L. Casanova. Familial nk cell deficiency associated with impaired il-2- and il-15-dependent survival of lymphocytes. *J Immunol*, 177(12):8835–43, 2006. ISSN 0022-1767 (Print) 0022-1767 (Linking). doi: 10.4049/jimmunol.177.12.8835.
16. C. Eidenschenk, J. Dunne, E. Jouanguy, C. Fourlinnie, L. Gineau, D. Bacq, C. McMahon, O. Smith, J. L. Casanova, L. Abel, and C. Feighery. A novel primary immunodeficiency with specific natural-killer cell deficiency maps to the centromeric region of chromosome 8. *Am J Hum Genet*, 78(4):721–7, 2006. ISSN 0002-9297 (Print) 1537-6605 (Electronic) 0002-9297 (Linking). doi: 10.1086/503269.
17. N. C. Guilz, Y. O. Ahn, S. Seo, and E. M. Mace. Unwinding the role of the cmg helicase in inborn errors of immunity. *J Clin Immunol*, 43(5):847–861, 2023. ISSN 1573-2592 (Electronic) 0271-9142 (Linking). doi: 10.1007/s10875-023-01437-3.
18. A. Riera, M. Barbon, Y. Noguchi, L. M. Reuter, S. Schneider, and C. Speck. From structure to mechanism-understanding initiation of dna replication. *Genes Dev*, 31(11):1073–1088, 2017. ISSN 1549-5477 (Electronic) 0890-9369 (Print) 0890-9369 (Linking). doi: 10.1101/gad.298232.117.
19. A. Costa and J. F. X. Diffley. The initiation of eukaryotic dna replication. *Annu Rev Biochem*, 91:107–131, 2022. ISSN 1545-4509 (Electronic) 0066-4154 (Linking). doi: 10.1146/annurev-biochem-072321-110228.
20. O. Hyrien, K. Marheineke, and A. Goldar. Paradoxes of eukaryotic dna replication: Mcm proteins and the random completion problem. *Bioessays*, 25(2):116–25, 2003. ISSN 0265-9247 (Print) 0265-9247 (Linking). doi: 10.1002/bies.10208.
21. R. C. Alver, G. S. Chadha, and J. J. Blow. The contribution of dormant origins to genome stability: from cell biology to human genetics. *DNA Repair (Amst)*, 19(100):182–9, 2014. ISSN 1568-7856 (Electronic) 1568-7864 (Print) 1568-7856 (Linking). doi: 10.1016/j.dnarep.2014.03.012.
22. A. M. Woodward, T. Gohler, M. G. Luciani, M. Oehlmann, X. Ge, A. Gartner, D. A. Jackson, and J. J. Blow. Excess mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. *J Cell Biol*, 173(5):673–83, 2006. ISSN 0021-9525 (Print) 1540-8140 (Electronic) 0021-9525 (Linking). doi: 10.1083/jcb.200602108.
23. X. Q. Ge, D. A. Jackson, and J. J. Blow. Dormant origins licensed by excess mcm2-7 are required for human cells to survive replicative stress. *Genes Dev*, 21(24):3331–41, 2007. ISSN 0890-9369 (Print) 1549-5477 (Electronic) 0890-9369 (Linking). doi: 10.1101/gad.457807.
24. A. G. Freud, A. Yokohama, B. Becknell, M. T. Lee, H. C. Mao, A. K. Ferketich, and M. A. Caligiuri. Evidence for discrete stages of human natural killer cell differentiation in vivo. *J Exp Med*, 203(4):1033–43, 2006. ISSN 0022-1007 (Print) 0022-1007 (Linking). doi: 10.1084/jem.20052507.
25. S. D. Scoville, A. G. Freud, and M. A. Caligiuri. Modeling human natural killer cell development in the era of innate lymphoid cells. *Front Immunol*, 8:360, 2017. ISSN 1664-3224 (Print) 1664-3224 (Linking). doi: 10.3389/fimmu.2017.00360.
26. A. G. Freud and M. A. Caligiuri. Human natural killer cell development. *Immunol Rev*, 214: 56–72, 2006. ISSN 0105-2896 (Print) 0105-2896 (Linking). doi: 10.1111/j.1600-065X.2006.00451.x.
27. D. A. Knorr, Z. Ni, D. Hermanson, M. K. Hexum, L. Bendzick, L. J. Cooper, D. A. Lee, and D. S. Kaufman. Clinical-scale derivation of natural killer cells from human pluripotent stem cells for cancer therapy. *Stem Cells Transl Med*, 2(4):274–83, 2013. ISSN 2157-6564 (Print) 2157-6564 (Linking). doi: 10.5966/sctm.2012-0084.
28. Y. Li, D. L. Hermanson, B. S. Moriarity, and D. S. Kaufman. Human ipsc-derived natural killer cells engineered with chimeric antigen receptors enhance anti-tumor activity. *Cell Stem Cell*, 23(2):181–192 e5, 2018. ISSN 1875-9777 (Electronic) 1875-9777 (Linking). doi: 10.1016/j.stem.2018.06.002.
29. C. M. Sturgeon, A. Ditadi, G. Awong, M. Kennedy, and G. Keller. Wnt signaling controls the specification of definitive and primitive hematopoiesis from human pluripotent stem cells. *Nat Biotechnol*, 32(6):554–61, 2014. ISSN 1546-1696 (Electronic) 1087-0156 (Print) 1087-0156 (Linking). doi: 10.1038/nbt.2915.
30. C. Dege and C. M. Sturgeon. Directed differentiation of primitive and definitive hematopoietic progenitors from human pluripotent stem cells. *J Vis Exp*, (129), 2017. ISSN 1940-087X (Electronic) 1940-087X (Linking). doi: 10.3791/55196.
31. L. Herrera, J. M. Salcedo, S. Santos, M. A. Vesga, F. Borrego, and C. Eguizabal. Op9 feeder cells are superior to m2-10b4 cells for the generation of mature and functional natural killer cells from umbilical cord hematopoietic progenitors. *Front Immunol*, 8:755, 2017. ISSN 1664-3224 (Print) 1664-3224 (Electronic) 1664-3224 (Linking). doi: 10.3389/fimmu.2017.00755.
32. F. Cichocki and J. S. Miller. In vitro development of human killer-immunoglobulin receptor-positive nk cells. *Methods Mol Biol*, 612:15–26, 2010. ISSN 1940-6029 (Electronic) 1064-3745 (Linking). doi: 10.1007/978-1-60761-362-6_2.
33. M. D. Hazenberg and H. Spits. Human innate lymphoid cells. *Blood*, 124(5):700–9, 2014. ISSN 1528-0020 (Electronic) 0006-4971 (Linking). doi: 10.1182/blood-2013-11-427781.
34. M. Parodi, H. Favoreel, G. Candiano, S. Gaggero, S. Sivori, M. C. Mingari, L. Moretta, M. Vitale, and C. Cantoni. Nkp44-nkp44 ligand interactions in the regulation of natural killer cells and other innate lymphoid cells in humans. *Front Immunol*, 10:719, 2019. ISSN 1664-3224 (Electronic) 1664-3224 (Linking). doi: 10.3389/fimmu.2019.00719.

35. R. Verma, J. Z. Er, R. W. Pu, J. Sheik Mohamed, R. A. Soo, H. M. Muthiah, J. K. C. Tam, and J. L. Ding. Eomes expression defines group 1 innate lymphoid cells during metastasis in human and mouse. *Front Immunol*, 11:1190, 2020. ISSN 1664-3224 (Electronic) 1664-3224 (Linking). doi: 10.3389/fimmu.2020.01190.
36. H. Kilpinen, A. Goncalves, A. Leha, V. Afzal, K. Alasoo, S. Ashford, S. Bala, D. Bensaddek, F. P. Casale, O. J. Culley, P. Danecek, A. Faulconbridge, P. W. Harrison, A. Kathuria, D. McCarthy, S. A. McCarthy, R. Meleckyte, Y. Memari, N. Moens, F. Soares, A. Mann, I. Streeter, C. A. Agu, A. Alderton, R. Nelson, S. Harper, M. Patel, A. White, S. R. Patel, L. Clarke, R. Halai, C. M. Kirton, A. Kolb-Kokocinski, P. Beales, E. Birney, D. Danovi, A. I. Lamond, W. H. Ouwehand, L. Vallier, F. M. Watt, R. Durbin, O. Stegle, and D. J. Gaffney. Common genetic variation drives molecular heterogeneity in human ipscs. *Nature*, 546(7658): 370–375, 2017. ISSN 1476-4687 (Electronic) 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/nature22403.
37. V. Volpato, J. Smith, C. Sandor, J. S. Ried, A. Baud, A. Handel, S. E. Newey, F. Wessely, M. Attar, E. Whiteley, S. Chintawar, A. Verheyen, T. Barta, M. Lako, L. Armstrong, C. Muschet, A. Artati, C. Cusulin, K. Christensen, C. Patsch, E. Sharma, J. Nicod, P. Brownjohn, V. Stubbs, W. E. Heywood, P. Gissen, R. De Filippis, K. Janssen, P. Reinhardt, J. Adamski, I. Royaux, P. J. Peeters, G. C. Terstappen, M. Graf, F. J. Livesey, C. J. Akerman, K. Mills, R. Bowden, G. Nicholson, C. Webber, M. Z. Cader, and V. Lakics. Reproducibility of molecular phenotypes after long-term differentiation to human ipsc-derived neurons: A multi-site omics study. *Stem Cell Reports*, 11(4):897–911, 2018. ISSN 2213-6711 (Electronic) 2213-6711 (Linking). doi: 10.1016/j.stemcr.2018.08.013.
38. C. Dege, K. H. Fegan, J. P. Creamer, M. M. Berrien-Elliott, S. A. Luff, D. Kim, J. A. Wagner, P. D. Kingsley, K. E. McGrath, T. A. Fehniger, J. Palis, and C. M. Sturgeon. Potently cytotoxic natural killer cells initially emerge from erythro-myeloid progenitors during mammalian development. *Dev Cell*, 53(2):229–239 e7, 2020. ISSN 1878-1551 (Electronic) 1534-5807 (Linking). doi: 10.1016/j.devcel.2020.02.016.
39. E. J. Fertig, J. Ding, A. V. Favorov, G. Parmigiani, and M. F. Ochs. Cogaps: an r/c++ package to identify patterns and biological process activity in transcriptomic data. *Bioinformatics*, 26 (21):2792–3, 2010. ISSN 1367-4811 (Electronic) 1367-4803 (Print) 1367-4803 (Linking). doi: 10.1093/bioinformatics/btq503.
40. B. Pijuan-Sala, J. A. Griffiths, C. Guibentif, T. W. Hiscock, W. Jawaid, F. J. Calero-Nieto, C. Mulas, X. Ibarra-Soria, R. C. V. Tyser, D. L. L. Ho, W. Reik, S. Srinivas, B. D. Simons, J. Nichols, J. C. Marioni, and B. Gottgens. A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature*, 566(7745):490–495, 2019. ISSN 1476-4687 (Electronic) 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/s41586-019-0933-9.
41. S. B. Hay, K. Ferchen, K. Chetal, H. L. Grimes, and N. Salomonis. The human cell atlas bone marrow single-cell interactive web portal. *Exp Hematol*, 68:51–61, 2018. ISSN 1873-2399 (Electronic) 0301-472X (Linking). doi: 10.1016/j.exphem.2018.09.004.
42. M. L. Whitfield, L. K. George, G. D. Grant, and C. M. Perou. Common markers of proliferation. *Nat Rev Cancer*, 6(2):99–106, 2006. ISSN 1474-175X (Print) 1474-175X (Linking). doi: 10.1038/nrc1802.
43. E. Z. Macosko, A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, J. J. Trombetta, D. A. Weitz, J. R. Sanes, A. K. Shalek, A. Regev, and S. A. McCarroll. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, 161(5):1202–1214, 2015. ISSN 1097-4172 (Electronic) 0092-8674 (Print) 0092-8674 (Linking). doi: 10.1016/j.cell.2015.05.002.
44. E. Hegewisch-Solloa, J. E. Melsen, H. Ravichandran, A. F. Rendeiro, A. G. Freud, B. Mundy-Bosse, J. C. Melms, S. E. Eisman, B. Izar, E. Grunstein, T. J. Connors, O. Elemento, A. Horowitz, and E. M. Mace. Mapping human natural killer cell development in pediatric tonsil by imaging mass cytometry and high-resolution microscopy. *bioRxiv*, page 2023.09.05.556371, 2023. doi: 10.1101/2023.09.05.556371.
45. Y. Hao, S. Hao, E. Andersen-Nissen, 3rd Mauck, W. M., S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert, and R. Satija. Integrated analysis of multimodal single-cell data. *Cell*, 184(13):3573–3587 e29, 2021. ISSN 1097-4172 (Electronic) 0092-8674 (Print) 0092-8674 (Linking). doi: 10.1016/j.cell.2021.04.048.
46. J. Loughery, M. Cox, L. M. Smith, and D. W. Meek. Critical role for p53-serine 15 phosphorylation in stimulating transactivation at p53-responsive promoters. *Nucleic Acids Res*, 42 (12):7666–80, 2014. ISSN 1362-4962 (Electronic) 0305-1048 (Print) 0305-1048 (Linking). doi: 10.1093/nar/gku501.
47. J. Zeng, S. Y. Tang, L. L. Toh, and S. Wang. Generation of "off-the-shelf" natural killer cells from peripheral blood cell-derived induced pluripotent stem cells. *Stem Cell Reports*, 9(6): 1796–1812, 2017. ISSN 2213-6711 (Electronic) 2213-6711 (Linking). doi: 10.1016/j.stemcr.2017.10.020.
48. B. H. Goldenson, P. Hor, and D. S. Kaufman. ipsc-derived natural killer cell therapies - expansion and targeting. *Front Immunol*, 13:841107, 2022. ISSN 1664-3224 (Electronic) 1664-3224 (Linking). doi: 10.3389/fimmu.2022.841107.
49. K. Weiskopf, P. J. Schnorr, W. W. Pang, M. P. Chao, A. Chhabra, J. Seita, M. Feng, and I. L. Weissman. Myeloid cell origins, differentiation, and clinical implications. *Microbiol Spectr*, 4(5), 2016. ISSN 2165-0497 (Electronic) 2165-0497 (Linking). doi: 10.1128/microbiolspec.MCHD-0031-2016.
50. J. L. Schultze, E. Mass, and A. Schlitzer. Emerging principles in myelopoiesis at homeostasis and during infection and inflammation. *Immunity*, 50(2):288–301, 2019. ISSN 1097-4180 (Electronic) 1074-7613 (Linking). doi: 10.1016/j.immuni.2019.01.019.
51. I. Mitroulis, L. Kalafati, G. Hajishengallis, and T. Chavakis. Myelopoiesis in the context of innate immunity. *J Innate Immun*, 10(5-6):365–372, 2018. ISSN 1662-8128 (Electronic) 1662-811X (Print) 1662-811X (Linking). doi: 10.1159/000489406.
52. H. J. Lim, Y. Jeon, C. H. Jeon, J. H. Kim, and H. Lee. Targeted disruption of mcm10 causes defective embryonic cell proliferation and early embryo lethality. *Biochim Biophys Acta*, 1813 (10):1777–83, 2011. ISSN 0006-3002 (Print) 0006-3002 (Linking). doi: 10.1016/j.bbamcr.2011.05.012.
53. N. Shima, A. Alcaraz, I. Liachko, T. R. Buske, C. A. Andrews, R. J. Munroe, S. A. Hartford, B. K. Tye, and J. C. Schimenti. A viable allele of mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet*, 39(1):93–8, 2007. ISSN 1061-4036 (Print) 1061-4036 (Linking). doi: 10.1038/ng1936.
54. M. Ueno, M. Itoh, L. Kong, K. Sugihara, M. Asano, and N. Takakura. Psf1 is essential for early embryogenesis in mice. *Mol Cell Biol*, 25(23):10528–32, 2005. ISSN 0270-7306 (Print) 1098-5549 (Electronic) 0270-7306 (Linking). doi: 10.1128/MCB.25.23.10528-10532.2005.
55. T. Mohri, M. Ueno, Y. Nagahama, Z. Y. Gong, M. Asano, H. Oshima, M. Oshima, Y. Fujio, and N. Takakura. Requirement of sld5 for early embryogenesis. *PLoS One*, 8(11):e78961, 2013. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0078961.
56. K. M. Knapp, D. E. Jenkins, R. Sullivan, F. L. Harms, L. von Elsner, C. W. Ockeloen, S. de Munnik, Emhf Bongers, J. Murray, N. Pachter, J. Denecke, K. Kutsche, and L. S. Bicknell. Mcm complex members mcm3 and mcm7 are associated with a phenotypic spectrum from meier-gorlin syndrome to lipodystrophy and adrenal insufficiency. *Eur J Hum Genet*, 29(7):1110–1120, 2021. ISSN 1476-5438 (Electronic) 1018-4813 (Print) 1018-4813 (Linking). doi: 10.1038/s41431-021-00839-4.
57. K. Kim, A. Doi, B. Wen, K. Ng, R. Zhao, P. Cahan, J. Kim, M. J. Aryee, H. Ji, L. I. Ehrlich, A. Yabuuchi, A. Takeuchi, K. C. Cunniff, H. Hongguang, S. McKinney-Freeman, O. Naveiras, T. J. Yoon, R. A. Irizarry, N. Jung, J. Seita, J. Hanna, P. Murakami, R. Jaenisch, R. Weissleder, S. H. Orkin, I. L. Weissman, A. P. Feinberg, and G. Q. Daley. Epigenetic memory in induced pluripotent stem cells. *Nature*, 467(7313):285–90, 2010. ISSN 1476-4687 (Electronic) 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/nature09342.
58. K. Kim, R. Zhao, A. Doi, K. Ng, J. Unternaehrer, P. Cahan, H. Huo, Y. H. Loh, M. J. Aryee, M. W. Lensch, H. Li, J. J. Collins, A. P. Feinberg, and G. Q. Daley. Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat Biotechnol*, 29(12):1117–9, 2011. ISSN 1546-1696 (Electronic) 1087-0156 (Print) 1087-0156 (Linking). doi: 10.1038/nbt.2052.
59. Y. Ohi, H. Qin, C. Hong, L. Blouin, J. M. Polo, T. Guo, Z. Qi, S. L. Downey, P. D. Manos, D. J. Rossi, J. Yu, M. Hebrok, K. Hochedlinger, J. F. Costello, J. S. Song, and M. Ramalho-Santos. Incomplete dna methylation underlies a transcriptional memory of somatic cells in human ips cells. *Nat Cell Biol*, 13(5):541–9, 2011. ISSN 1476-4679 (Electronic) 1465-7392 (Print) 1465-7392 (Linking). doi: 10.1038/ncb2239.
60. S. Buckberry, X. Liu, D. Poppe, J. P. Tan, G. Sun, J. Chen, T. V. Nguyen, A. de Mendoza, J. Pflueger, T. Frazer, D. B. Vargas-Landin, J. M. Paynter, N. Smits, N. Liu, J. F. Ouyang, F. J. Rossello, H. S. Chy, O. J. L. Rackham, A. L. Laslett, J. Breen, G. J. Faulkner, C. M. Nefzger, J. M. Polo, and R. Lister. Transient naive reprogramming corrects hips cells functionally and epigenetically. *Nature*, 620(7975):863–872, 2023. ISSN 1476-4687 (Electronic) 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/s41586-023-06424-7.
61. B. Oldak, E. Wildschutz, V. Bondarenko, M. Y. Comar, C. Zhao, A. Aguilera-Castrejon, S. Tarazi, S. Viukov, T. X. A. Pham, S. Ashouokhi, D. Lokshtanov, F. Roncato, E. Ariel, M. Rose, N. Livnat, T. Shani, C. Joubran, R. Cohen, Y. Addadi, M. Chemla, M. Kedmi, H. Keren-Shaul, V. Pasque, S. Petropoulos, F. Lanner, N. Novershtern, and J. H. Hanna. Complete human day 14 post-implantation embryo models from naive es cells. *Nature*, 2023. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/s41586-023-06604-5.
62. B. A. T. Weatherbee, C. W. Gantner, L. K. Iwamoto-Stohl, R. M. Daza, N. Hamazaki, J. Shendure, and M. Zernicka-Goetz. Pluripotent stem cell-derived model of the post-implantation human embryo. *Nature*, 2023. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/s41586-023-06368-y.
63. R. M. Karvas, J. E. Zemke, S. S. Ali, E. Upton, E. Sane, L. A. Fischer, C. Dong, K. M. Park, F. Wang, K. Park, S. Hao, B. Chew, B. Meyer, C. Zhou, S. Dietmann, and T. W. Theunissen. 3d-cultured blastoids model human embryogenesis from pre-implantation to early gastrulation stages. *Cell Stem Cell*, 30(9):1148–1165 e7, 2023. ISSN 1875-9777 (Electronic) 1875-9777 (Linking). doi: 10.1016/j.stem.2023.08.005.
64. A. Patel, G. Iannello, A. G. Diaz, D. Sirabella, V. Thaker, and B. Corneo. Efficient cas9-based genome editing using crispr analysis webtools in severe early-onset-obesity patient-derived ipscs. *Curr Protoc*, 2(8):e519, 2022. ISSN 2691-1299 (Electronic) 2691-1299 (Linking). doi: 10.1002/cpz1.519.
65. G. L. Stein-O'Brien, J. L. Carey, W. S. Lee, M. Considine, A. V. Favorov, E. Flam, T. Guo, S. Li, L. Marchionni, T. Sherman, S. Sivy, D. A. Gaykalova, R. D. McKay, M. F. Ochs, C. Colantuoni, and E. J. Fertig. Patternmarkers gwcogaps for novel data-driven biomarkers via whole transcriptome nmf. *Bioinformatics*, 33(12):1892–1894, 2017. ISSN 1367-4811 (Electronic) 1367-4803 (Print) 1367-4803 (Linking). doi: 10.1093/bioinformatics/btx058.
66. G. Sharma, C. Colantuoni, L. A. Goff, E. J. Fertig, and G. Stein-O'Brien. projectr: an r/bioconductor package for transfer learning via pca, nmf, correlation and clustering. *Bioinformatics*, 36(11):3592–3593, 2020. ISSN 1367-4811 (Electronic) 1367-4803 (Print) 1367-4803 (Linking). doi: 10.1093/bioinformatics/btaa183.

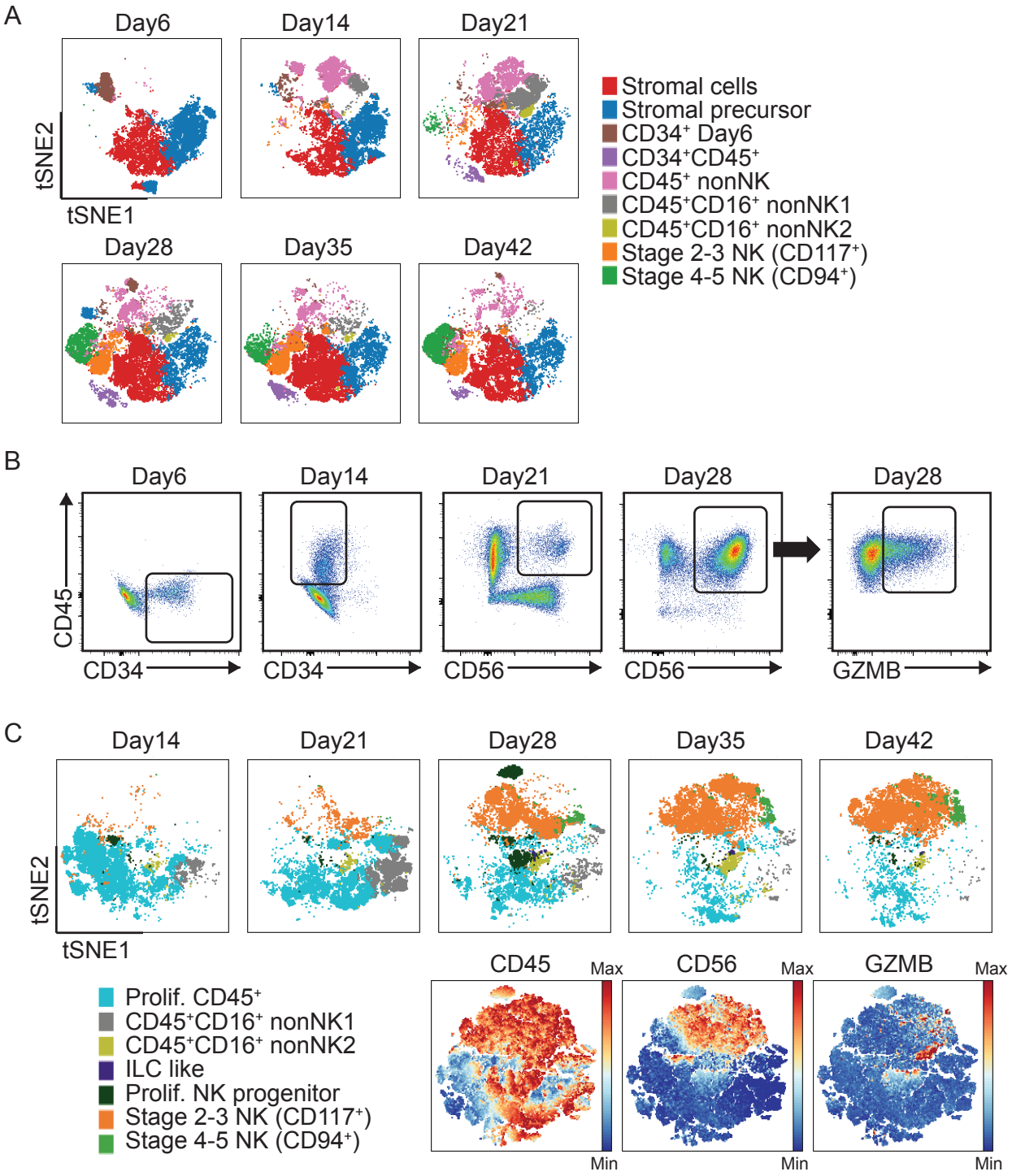

Figure S1. NK cell differentiation from the iPSCs on autologous stroma using the spin-EB method. NK cells were differentiated from iPSC lines using the spin-EB method. Flow cytometry was performed to monitor the progression of NK cell differentiation then data were concatenated, normalized, and visualized. A) tSNE plot summarizing flow cytometry data from six lines (I.1, I.2, II.1, II.2, C.1, and II.1-Cor) visualized by time points. B) Representative flow cytometry plots demonstrating sequential expression of CD34, CD45, CD56 and granzyme B (GZMB) throughout NK cell differentiation. C) tSNE plot summarizing the flow cytometry data of CD45$^+$ population from six lines (GINS4 family, C.1, and II.1-Cor) visualized by time points. Expression of CD45, CD56 and granzyme B (GZMB) shows the process of NK cell maturation in the tSNE combined from all time-points.

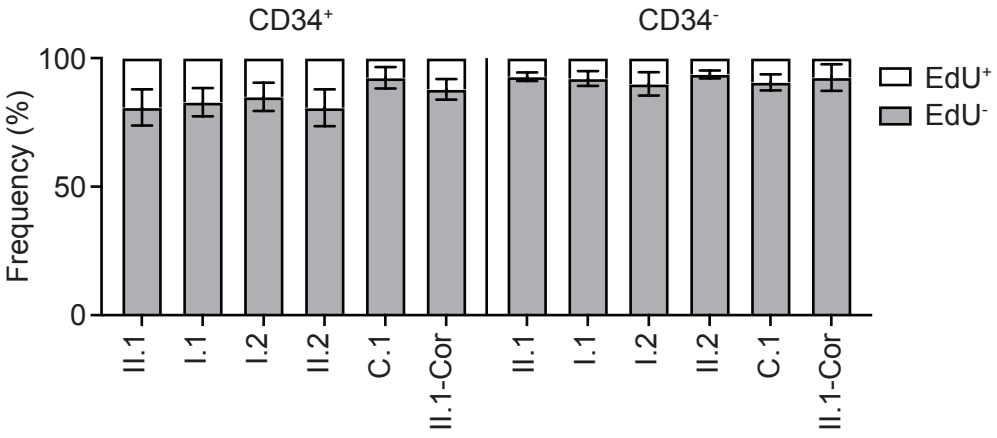

Figure S2. The effect of GINS4 variants on cell cycle of day 6 cells. Day6 cells from the EBs were stained with CD34 and EdU for assessing frequency of cells in the S phase by flow cytometry. Frequency of EdU^+ populations in $CD34^+$ or $CD34^-$ populations from six cell lines (n=4 technical replicates, error bars=SD).

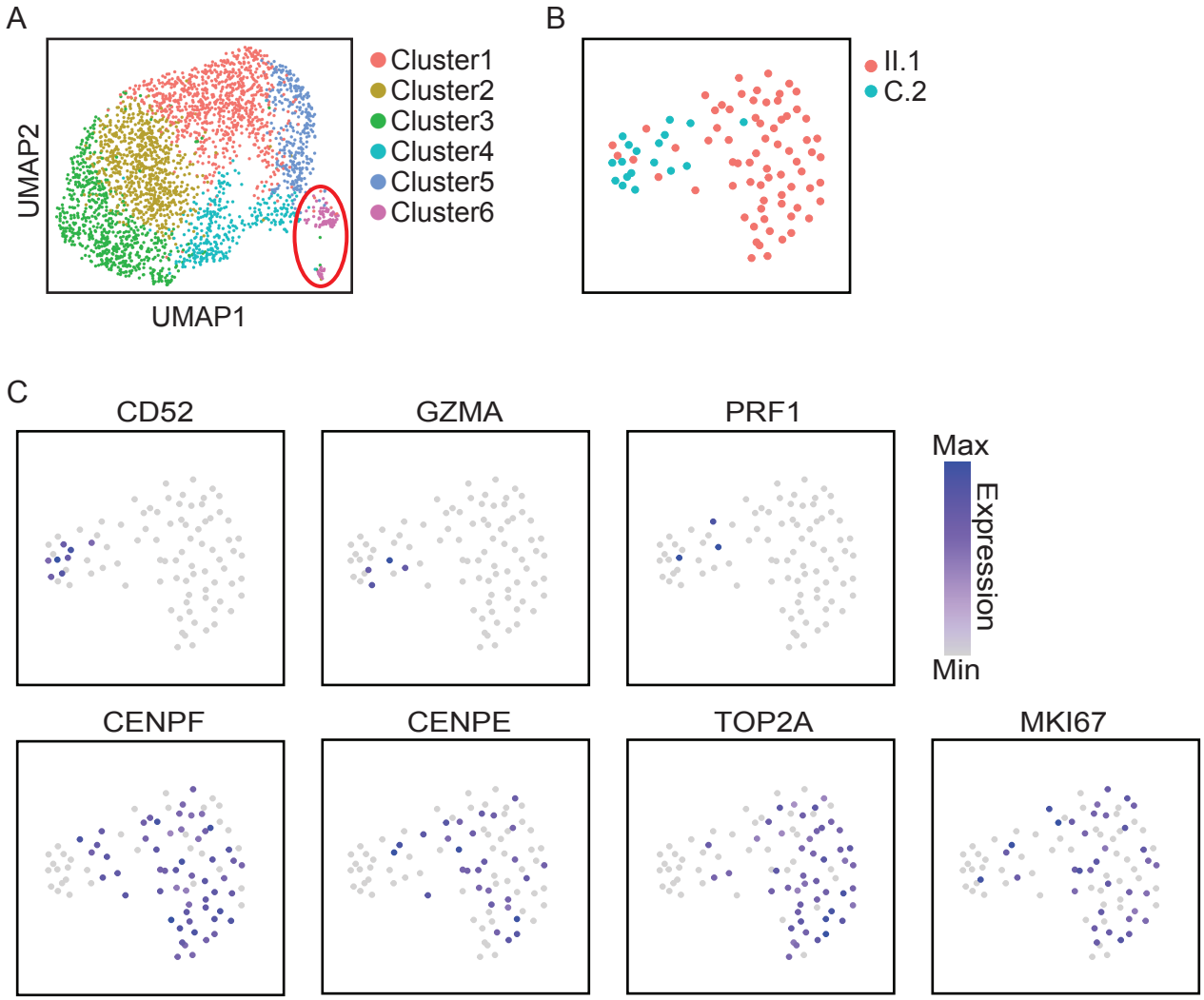

Figure S3. NK progenitor cells from II.1 line on day 28 show distinct transcriptional signatures. scRNAseq was performed on day28 cells from the II.1 and C.2 lines. Data were processed, normalized, and analyzed with Seurat (45). A) UMAP showing the scRNAseq result from lines II.1 and C.2 on day 28. Red circle (cluster 6) indicates cells committed to NK lineage. B. UMAP of cluster 6 cells showing distinct transcriptional signatures between proband (II.1) and control (C.2) cells. C. Gene expression levels of NK cell or cell cycle related genes in cluster 6 cells overlaid on UMAP of combined (II.1 and C.2) scRNA-Seq data.

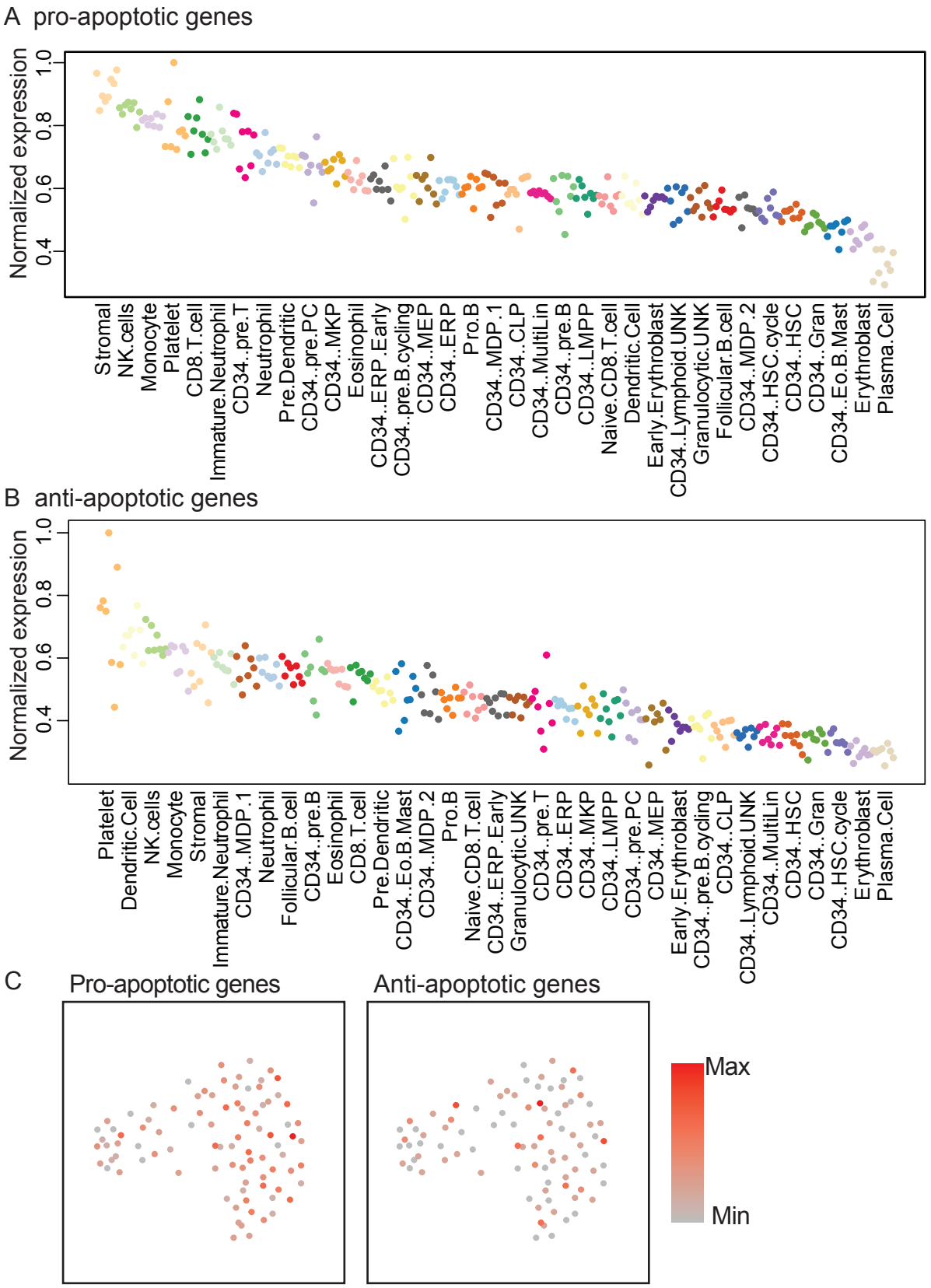

Figure S4. Apoptotic gene expression in human bone marrow atlas and day 28 cluster 6 cells. We curated the apoptosis genes (KEGG_APOPTOSIS, M8492) into pro- and anti-apoptosis and calculated accumulative expression in the bone marrow atlas data set (41) and our day 28 scRNAseq data cluster6. A-B. Cumulative levels of (A) pro-apoptotic or (B) anti-apoptotic genes in the human bone marrow atlas data ordered by their mean expression. C) Cumulative levels of pro-apoptotic (left) and anti-apoptotic (right) genes in day 28 cluster 6 cells from scRNA-Seq data.