NUDCD3 deficiency disrupts V(D)J recombination to cause SCID and Omenn syndrome

Authors: Rui Chen[1]†, Elena Lukianova[2]†, Ina Schim van der Loeff[1,3], Jarmila Stremenova Spegarova[1], Joseph D.P. Willet[1], Kieran D. James[4], Edward J. Ryder[2], Helen Griffin[1], Hanna IJspeert[5], Akshada Gajbhiye[6], Fredric Lamoliatte[6]‡, Jose L. Marin-Rubio[6]§, Lisa Woodbine[7], Henrique Lemos[1], David J. Swan[1]¶, Valeria Pintar[1], Kamal Sayes[1], Elias R. Ruiz-Morales[2], Simon Eastham[2]#, David Dixon[6], Martin Prete[2], Elena Prigmore[2], Penny Jeggo[7], Joan Boyes[8], Andrew Mellor[1], Lei Huang[1], Mirjam van der Burg[5]**, Karin R. Engelhardt[1], Asbjørg Stray-Pedersen[9], Hans Christian Erichsen[10], Andrew R. Gennery[1,3], Matthias Trost[6], David J. Adams[2], Graham Anderson[4], Anna Lorenc[2], Gosia Trynka[2,11], Sophie Hambleton[1,3]*

Affiliations:

[1]Translational and Clinical Research Institute, Newcastle University; Newcastle upon Tyne, NE2 4HH, UK.

[2]Wellcome Sanger Institute, Wellcome Genome Campus; Hinxton, CB10 1SA, UK.

[3]Great North Children's Hospital, Newcastle upon Tyne Hospitals NHS Foundation Trust; Newcastle upon Tyne, NE1 4LP, UK.

[4]Institute of Immunology and Immunotherapy, University of Birmingham; Birmingham, B15 2TT, UK.

[5]Department of Immunology, Erasmus University Medical Center; Rotterdam, 3000 CA, The Netherlands.

[6]Biosciences Institute, Newcastle University; Newcastle upon Tyne, NE2 4HH, UK.

[7]Genome Damage and Stability Centre, University of Sussex; Brighton, BN1 9RQ, UK.

[8]Faculty of Biological Sciences, University of Leeds; Leeds, LS2 9JT, UK.

[9]Norwegian National Unit for Newborn Screening, Division of Pediatric and Adolescent Medicine, Oslo University Hospital; Oslo, 0424, Norway.

[10]Division of Pediatric and Adolescent Medicine, Oslo University Hospital and Institute of Clinical Medicine, University of Oslo; Oslo, 0424, Norway.

[11]Open Targets, Wellcome Genome Campus; Hinxton, CB10 1SA, UK.

*Corresponding author. Email: sophie.hambleton@ncl.ac.uk

†These authors contributed equally.

‡Present address: MRC PPU, School of Life Sciences, University of Dundee; Dundee, DD1 5EH, UK.

§Present address: Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center; New York, NY, 10065, USA.

¶Present address: School of Medicine, University of Sunderland; Sunderland, SR1 3SD, UK.

#Present address: School of Cellular and Molecular Medicine, University of Bristol; Bristol, BS8 1TD, UK.

**Present address: Departments of Pediatrics and Immunology, Leiden University Medical Center; Leiden, 2300 RC, The Netherlands.

Abstract: Inborn errors of T cell development present a pediatric emergency in which timely curative therapy is informed by molecular diagnosis. In 11 affected patients across four consanguineous kindreds, we detected homozygosity for a single deleterious missense variant in the gene NudC domain–containing 3 (*NUDCD3*). Two infants had severe combined immunodeficiency with the complete absence of T and B cells (T-B- SCID), whereas nine showed classical features of Omenn syndrome (OS). Restricted antigen receptor gene usage by residual T lymphocytes suggested impaired V(D)J recombination. Patient cells showed reduced expression of NUDCD3 protein and diminished ability to support RAG-mediated recombination in vitro, which was associated with pathologic sequestration of RAG1 in the nucleoli. Although impaired V(D)J recombination in a mouse model bearing the homologous variant led to milder immunologic abnormalities, NUDCD3 is absolutely required for healthy T and B cell development in humans.

One-Sentence Summary: Developing T and B cells show a conserved requirement for NUDCD3 to support recombinase-mediated shuffling of antigen receptor genes.

INTRODUCTION

The molecular dissection of inborn errors of immunity provides powerful insights into genes and pathways of nonredundant importance to the human immune system. Nowhere is this truer than in severe combined immunodeficiency (SCID), a failure of T cell development that leaves affected infants exceptionally vulnerable to infection (*1*). Genes associated with SCID highlight the critical role of T cell receptor rearrangement and signaling, alongside survival signals via the interleukin 7 receptor (IL-7R) during T lymphocyte development (*2*). Omenn syndrome (OS) is a related disorder in which an incomplete developmental block allows the dysregulated expansion of oligoclonal T cells, causing inflammation of the skin and viscera as well as SCID-like susceptibility to infection (*3*). Because these conditions are both treatable as well as life-threatening, many nations have recently introduced newborn screening for T cell immunodeficiency (*4*). In affected infants, a specific molecular diagnosis may guide precision medicine such as enzyme replacement, gene therapy, or hematopoietic stem cell transplantation (HSCT) (*1*).

OS is particularly associated with disorders of V(D)J recombination, the complex process whereby antigen receptor loci are rearranged to produce functional T cell receptor and immunoglobulin genes (*5*). Although this process is absolutely required for the generation of a diverse adaptive immune repertoire, the accompanying introduction and recombinatorial repair of DNA double-strand breaks implies a moment of extreme risk for genome integrity. It is therefore anticipated that V(D)J recombination occurs through specialized molecular machinery, including the T cell– and B cell–specific recombination-activating gene (RAG) recombinases, whose expression and activity are tightly controlled (*6*-*8*).

RESULTS

A homozygous, hypomorphic *NUDCD3* variant in infants with SCID/OS

In this study, we sought to extend knowledge of lymphocyte development by studying patients with classical OS or T-B- SCID who lacked pathogenic variants in known disease genes, such as those encoding the recombinases RAG1 and RAG2 or components of the DNA-repair machinery including Artemis (*2*, *5*). The occurrence of multiple familial cases within four independent, consanguineous kindreds from the same ethnic (South Asian) background suggested there was autosomal recessive inheritance of a shared ancestral morbid allele (Fig. 1A and tables S1 and S2). Both family 1 (*9*) and family 2 (*10*) have been previously described. By integrative analysis of whole-exome sequencing data, we identified a single homozygous missense variant in the gene *NUDCD3* (NudC domain–containing 3), which segregated with disease in all four affected kindreds but was absent from the gnomAD and Genome Asia databases, in keeping with a rare disease allele (Fig. 1, A and B, and fig. S1) (*11*, *12*). This gene appeared to be intolerant of loss-of-function variation with a LOEUF score (loss-of-function observed/expected upper-bound fraction) of 0.21 and no instances of homozygous predicted null variants in the gnomAD database (fig. S1) (*12*). The missense variant c.155G>A introduced a charge change by substituting aspartate for a glycine residue at position 52 within the highly conserved N-terminal domain of NUDCD3 (Fig. 1C). Programs including Combined Annotation Dependent Depletion (CADD; score 29.9), Polyphen2 (score 1.0), and VARITY-ER (score 0.964) predicted that this variant is highly damaging (fig. S1 and table S3) (*13*-*15*).

NUDCD3 is ubiquitously expressed in vertebrates and belongs to the NudC family of p23 domain–containing proteins that exhibit cochaperone and intrinsic chaperone activity (*16*). This protein family emerged at the same time as beta-propeller structures within increasingly complex multidomain and multisubunit proteins (*17*, *18*). Family members NUDC and NUDCL2 have recently been shown to act as cochaperones in Hsp90-related refolding of client proteins such as the glucocorticoid receptor and cohesin, respectively (*19*, *20*).

NUDCD3 was specifically linked to dynein intermediate chain stability and viability through study of *NUDCD3*-knockdown cells in vitro (*21*), whereas its overexpression impaired cytokinesis (*22*). The only prior literature linking NUDCD3 to the immune system reported the appearance of RAG1 among its interactome when assessed by a LUMIER screen of potential client proteins shared with Hsp90 (*18*). The same screen noted a predominance for NUDCD3 of binding partners containing kelch-like domains, such as RAG2, which was, however, not tested in this study (*18*). In addition, our reanalysis of data from developing human embryos confirmed the expression of *NUDCD3* transcripts in thymus and bone marrow (*23, 24*).

NUDCD3 protein expression in patient dermal fibroblasts was reduced relative to control despite equivalent mRNA expression, indicating hypomorphic behavior of the variant allele (Fig. 1D and fig. S2). However, NUDCD3^{G52D} was distributed between the cytoplasm and nucleus in a manner indistinguishable from wild-type (WT) protein, whether endogenously expressed or transfected into cells in tagged form (Fig. 1E and fig. S3).

Patient cells divided normally and did not show significantly impaired resistance to gamma-irradiation. Moreover, *NUDCD3* knockdown did not affect the radiation sensitivity of A549 cells, indicating their preserved ability to repair DNA double-strand breaks (fig. S4, A and B). However, *NUDCD3* is an essential gene in many cell lines, suggesting that homozygosity for a complete null allele would not be viable (*25*).

To compare the global properties of mutant and WT NUDCD3, we prepared mild detergent lysates of transfected cells, separated protein complexes by size-exclusion chromatography, and probed the resulting fractions for NUDCD3 following SDS-PAGE under denaturing conditions (fig. S4C). Both WT and variant NUDCD3 participated in protein complexes of diverse size to a similar extent. However, coimmunoprecipitation experiments in which distinctively tagged mutant or WT *NUDCD3* was cotransfected showed significantly reduced homodimerization potential for the G52D variant (Fig. 1F). Thus, the G52D variant does not grossly alter the physical properties of NUDCD3 but disrupts a specific function of the protein within the immune system, potentially one requiring its dimerization.

NUDCD3^{G52D} patients exhibit defective V(D)J recombination

To gain further insight into pathomechanism, we undertook single-cell studies of cryopreserved peripheral blood mononuclear cells from patients with OS associated with either the *NUDCD3* variant or pathogenic defects of *RAG1* or *RAG2* (table S4 and S5). Compared to healthy controls, drastic abnormalities in the distribution of circulating lymphocyte subsets were readily observed in OS patients by flow cytometry, including the complete lack of B cells and naïve T cells (Fig. 2A). Single-cell RNA sequencing (scRNA-seq) revealed widely divergent distribution across the 39 cell subpopulations annotated (32 lymphoid and seven myeloid), including within each of the mutation groups (Fig. 2, B and C). We detected γδ T cells, for example, in only eight of 11 Omenn patients. However, in two of these individuals, γδ T cells constituted more than 20% of total T cells, a much higher proportion than in healthy controls. When present, the CD8$^+$ and γδ T cell compartments were abnormally skewed towards weakly cytotoxic (e.g., lacking *NKG7*, granzymes, granulysin, and chemokines) and non-naïve (e.g., high expression of *IL32* and absence of *CCR7*) phenotypes with low regulatory potential towards T cell and natural killer (NK) cell effectors (Fig. 2C and fig. S5). CD4$^+$ T cells were enriched for central memory (CM) and effector memory (EM) phenotypes (characterized by high expression of *LGALS1*, *NEAT1*, and *GZMA* as well as cytoskeletal and *HLA* genes) (Fig. 2C and fig. S5). Such T cell abnormalities were prominent in both *NUDCD3*- and *RAG*-mutated OS patients, as were associated alterations of the myeloid compartment. The latter included enrichment for classical monocytes distinguished by a cytotoxic, proinflammatory, and chemotactic

(*CCL2*) transcriptional signature (fig. S6).

Analysis of T cell receptor (TCR) gene usage within our single-cell transcriptomic data revealed reduced diversity and expanded clonotype size in OS patients, as expected (Fig. 2D and fig. S7A). We confirmed the previously noted skewing of the *RAG1/2*-deficient TCR repertoire towards the use of *TRA* gene segments near the 3′ end of the Vα locus and the 5′ end of the Jα locus (Fig. 2E) (*26*), a pattern that was reproduced in the T cells of *NUDCD3*^{G52D} patients (Fig. 2E). These proximal *TRA* gene segments are preferentially recombined at the start of the DP stage of thymocyte development. If the resulting TCRα chain, alongside the preexisting TCRβ chain, is incapable of self-peptide–MHC recognition, more distal segments are used in subsequent recombination rounds, until a productive recombination results in the downregulation of *RAG1* expression (*27*). The predominance of proximal Vα and Jα gene segment usage therefore implied that *NUDCD3*^{G52D} T cells were unable to execute repeated rounds of V(D)J recombination during their development, just as in RAG Omenn T cells. We confirmed the reduced diversity of TCRβ chain usage in OS patients without obvious skewing across the *TRB* locus (fig S7, A and B) consistent with prior literature (*28*, *29*).

These data suggested that there was a primary defect in V(D)J recombination associated with homozygosity for the *NUDCD3*^{G52D} variant. We therefore tested the integrity of RAG-dependent recombination by cotransfecting patient and control fibroblasts with *RAG1*, *RAG2*, and a synthetic substrate for RAG recombination, which could be detected by PCR of the rearranged product. In this preliminary screen, where the expression of each transfected component could not be assessed at single-cell level, *NUDCD3*^{G52D} patient cells nonetheless appeared highly defective in supporting RAG-dependent recombination (Fig. 2F). We consequently turned to murine systems in which to model this effect in a more physiological context.

NUDCD3 plays a conserved role in mouse V(D)J recombination

To test whether endogenous V(D)J rearrangement was similarly NUDCD3-dependent, we took advantage of a mouse pre-B cell line, 103/BCL2, in which this process can be triggered by a shift in temperature and detected by quantitative PCR for kappa-deleting excision circles (KRECs) (*30*). Cells in which *Nudcd3* expression had been partially knocked down by siRNA transfection were indeed defective for V(D)J recombination (Fig. 2G).

As these data indicated a conserved role for NUDCD3 in murine V(D)J recombination, we used CRISPR/Cas9 technology to engineer an in vivo mouse model bearing the homologous G52D missense variant on the C57BL/6 background (Fig. 3A). Mice homozygous for the *Nudcd3*^{G52D} variant (referred to below as *Nudcd3*^{Hom}) were born at sub-Mendelian frequency, were significantly smaller than littermates and were sterile but appeared otherwise healthy (Fig. 3, B and C). Expression of NUDCD3 protein was substantially reduced in *Nudcd3*^{Hom} tissues compared with WT tissues (Fig. 3D and fig. S8A). Thymic cellularity was significantly lower than WT despite normal corticomedullary architecture (Fig. 3E and fig. S8B). We observed normal proportions of B and T cells in the periphery and a normal ratio of CD4 to CD8 T cells (Fig. 3, F to I, and figs. S8, C to F, and S9, A and B). Nonetheless, the ratio of CD44^{hi} memory T cells to naïve T cells was consistently increased in the periphery of *Nudcd3*^{Hom} compared to WT littermates (fig. S8, G and H). Kappa light chain was overexpressed relative to lambda light chain in peripheral B cells of *Nudcd3*^{Hom} mice, which would be expected in the context of impaired V(D)J recombination (Fig. 3, J and K, and fig. S9C) (*31*, *32*). Immunoglobulin production was preserved in *Nudcd3*^{Hom} mice, with slightly increased IgM and IgG2b, and markedly increased IgE as is also seen in mice bearing *Rag1*$^{R972Q/R972Q}$ and *Rag2*$^{R229Q/R229Q}$ variants and in human OS patients (Fig. 3L and fig. S9D) (*3*, *33-35*).

Developing thymocytes progress through an orderly series of stages as they undergo successive

rounds of V(D)J recombination, express the resulting TCR, and traverse positive and negative selection checkpoints (*36*). To assess developmental progression, we analyzed the distribution of thymocytes by surface marker expression, noting a highly statistically significant excess of $CD4^{-}CD8^{-}$ double-negative (DN) thymocytes in *Nudcd3*^{Hom} mice (Fig. 4, A and B, and fig. S10). Subdivision of the DN compartment showed that this expansion occurred at the DN3 ($CD25^{hi}CD44^{lo}$) stage (Fig. 4, A and C, and fig. S10, A to D), specifically the $CD27^{lo}$ DN3a phase, which precedes pre-TCR signaling (Fig 4, A and D) (*37*). Consistent with a developmental block, the absolute number of DN3a thymocytes in *Nudcd3*^{Hom} mice exceeded that in WT, despite a much smaller thymus overall (Fig. 4D and Fig. 3E). Mutant mice also showed a relative expansion of the earliest DP subset (DP1), in which alpha-chain rearrangement occurs (Fig. 4, A and E, and fig. S10, E to G). Although expanded, the $CD25^{+}$ DN3 compartment proliferated less as assessed by Ki67 staining (Fig. S10, H and I), whereas the frequency of $Ki67^{+}$ cells within DP1 was not significantly reduced (Fig. S10J). Subsequent development of CD4 and CD8 single-positive (SP) lymphocytes in *Nudcd3*^{Hom} mice appeared normal (fig. S11, A to G) although *Nudcd3*^{Hom} mice did exhibit a trend towards an increased proportion of thymic derived regulatory T (Treg) cells and markedly reduced invariant natural killer T (iNKT) cells, as previously described both in patients with *RAG*-OS and in a mouse model of this syndrome (fig. S11, H to K) (*34*, *38*). Thus, there is an incomplete block in *Nudcd3*^{Hom} thymocyte development at those stages where V(D)J recombination occurs. The expansion of these compartments (DN3a and DP1) implies a direct effect of *Nudcd3*^{Hom} on V(D)J recombination, rather than a primary impairment of survival that shortens the time available for V(D)J recombination.

To explore this effect in greater detail, we isolated splenocytes and bulk-sequenced their rearranged TCR-encoding genes *Tra* and *Trb*. In comparison with WT, TCR clonotypes were less diverse in *Nudcd3*^{Hom} mice and individual clonotypes tended to be more expanded (Fig. 4F and fig. S12, A and B), out of proportion to the lower fraction of naïve cells. Echoing human *NUDCD3*^{G52D} patients, *Nudcd3*^{Hom} mice showed strongly skewed *Tra* gene usage towards V segments from the 3′ end of the V locus and J segments from the 5′ end of the J locus and the absence of a similar effect for *Trb* or *Ig* gene segments (Fig. 4G and fig. S12, E to H) (*26*). This skewing likely explained the lack of iNKT cells (fig. S11, J and K, and fig. S12C), which strongly favor the use of distal *Tra* segments (*39*). Additionally, *Nudcd3*^{Hom} mice had fewer nonproductive *Trb* chain rearrangements than WT (Fig S12D). These exist only if an unproductive recombination is followed by a successful second round of recombination, ensuring T cell survival. Thus, both the pattern of TCR gene segment usage and the lack of nonproductive rearrangements suggest a lower frequency of repeated recombinations of *Tra/Trb* loci in *Nudcd3*^{Hom} mice as in humans.

In parallel, we analyzed B cell development through enumeration of successive Hardy fractions in mouse bone marrow by flow cytometry (fig. S13A). There was a relative accumulation of $IgM^{-}IgD^{-}$ (fraction D) pre-B cells and a trend towards a reduction in mature $IgM^{+}IgD^{+}$ B cells (fraction F) but no effect on pro-B cell populations (fractions A–C) (fig. S13, A to D). Thus, both T and B cell development are perturbed in *Nudcd3*^{G52D} mice although to a lesser degree than that observed in human *NUDCD3*^{G52D} patients or in published mouse models of hypomorphic RAG deficiency, especially in the case of B cells (*33*-*35*).

To examine whether this alteration of lymphocyte development was cell-intrinsic, we prepared mixed bone-marrow chimeras and compared the ability of *Nudcd3*^{Hom} and *Nudcd3*^{WT} cells to contribute to the B and T cell compartments of a WT host. Lethally irradiated $mT^{+}CD45.1^{+}$ WT recipient mice were injected intravenously with equal amounts of $CD45.1^{+}$ WT and $CD45.2^{+}$ *Nudcd3*^{Hom} or *Nudcd3*^{WT} whole bone marrow cells (fig. S13, E and F). *Nudcd3*^{Hom} cells failed to develop normally when in competition with WT cells, as demonstrated by their underrepresentation throughout T cell development (Fig. 4H). The reduction at DN1 implied *Nudcd3*^{Hom} had a possible background effect on early thymocyte fitness relative to WT, which was not strong enough however to prevent the

relative expansion of the subsequent (DN2–DN3a) stages. However, further marked skewing of the DN3b compartment towards WT cells (Fig. 4H) indicated defective developmental transition of *Nudcd3*^{Hom} thymocytes from DN3a to DN3b, confirming the impaired DN3a–DN3b progression observed in steady-state *Nudcd3*^{Hom} mice. This skewing was exacerbated at the transition to the DP stage, again echoing the phenotype seen in the steady-state *Nudcd3*^{Hom} model. By contrast, the contribution of *Nudcd3*^{Hom} cells to the developing B cell compartment appeared consistent throughout development until a reduction in mature IgM^{+}IgD^{+} B cells (fraction F) (fig. S13G). Nonetheless, the percentage of Igλ^{+}CD45.2^{+} *Nudcd3*^{Hom} cells was very low (fig. S13H), consistent with defective light chain rearrangement despite the normal proportion of *Nudcd3*^{Hom} cells in fractions D–E.

Although we do not formally exclude an additional contribution from altered thymic function, the mixed-bone marrow chimera experiments confirm the cell-intrinsic nature of both B and T cell abnormalities in the *Nudcd3*^{Hom} mouse. This is in keeping with the curative potential of hematopoietic stem cell transplantation demonstrated in two of the infants in our case series, who both achieved long-term B and T cell reconstitution with independence of immunoglobulin replacement therapy (table S2). Thus, NUDCD3^{G52D} appears to produce an incomplete developmental block with its major impact on V(D)J recombination in both species, one that differs in severity between humans and mice and between developing B and T cells.

Failure of RAG2-dependent disaggregation of RAG1 in NUDCD3-deficient cells

To further investigate this impact on V(D)J recombination, we next harvested *Nudcd3*^{WT}, *Nudcd3*^{Het}, or *Nudcd3*^{Hom} mouse bone marrow and generated pre-B cell lines containing an inverted GFP recombination reporter (Fig. 5A). Upon triggering by v-Abl inhibition, WT pre-B cells successfully recombined the reporter construct to generate green fluorescence. By contrast, *Nudcd3*^{Hom} pre-B cells showed markedly reduced recombination activity (Fig. 5A). Thus, a primary defect in V(D)J recombination appears to arise from the *Nudcd3* p.G52D mutation, rather than an indirect effect through impaired cell survival.

To gain insight into the molecular mechanism of this defect, we first considered the possibility that the G52D variant might disrupt a normal interaction between NUDCD3 and RAG1 and/or RAG2 such as a cochaperone function required for stabilization and/or localization. By immunofluorescence microscopy, we found no evidence of altered subcellular distribution of either RAG protein when overexpressed alone in mutant versus WT cells, suggesting that the nuclear import of folded RAGs was intact (Fig. 5B). We did however note that the distribution of NUDCD3 was altered by RAG expression in 293T cells, becoming increasingly nuclear (Fig. S14A). Despite this, neither coimmunoprecipitation followed by immunoblotting nor mass-spectrometric (MS) analysis showed a direct interaction between NUDCD3 and endogenous RAGs in a RAG-expressing T lymphoblastoid cell line, although we confirmed other previously reported protein–protein interactions (Fig. 5C; table S6; and data S1) (*18*, *40*). The interactomes of WT and G52D NUDCD3 were largely overlapping and included several previously identified binding partners of RAGs such as heat shock proteins, leaving open the possibility of an indirect interaction.

Immunofluorescence imaging demonstrated the previously described reciprocal nuclear localization of individually expressed RAG proteins, either inside (RAG1) or outside (RAG2) nucleoli (Fig. 5B) (*41*). Recent reports have emphasized that the nucleolar retention of RAG1 restrains RAG recombinase activity (*40*) and that cell cycle–dependent RAG2 coexpression enables the redistribution of its partner from nucleoli (*42*). We therefore investigated the impact of NUDCD3^{G52D} upon RAG2-dependent RAG1 redistribution by cotransducing primary fibroblasts with fluorescently tagged RAG1 and RAG2. Control fibroblasts showed partial colocalization of RAG proteins outside nucleoli, whereas RAG1 was predominantly confined to the nucleoli of dual-transduced patient

fibroblasts (Fig. 5, D and E, and fig. S14, B and C). We confirmed that the nucleolar retention of RAG1 was truly NUDCD3-dependent by performing a similar experiment in HEK293T cells in which *NUDCD3* had been partially knocked down by siRNA transfection prior to transient overexpression of the fluorescently tagged RAGs (Fig. 5, F and G, and fig. S14, D and E). The mislocalization of RAG1 in *NUDCD3*-deficient *RAG2*-cotransfected cells was accompanied by an overall increase in RAG1 protein expression as assessed by immunoblotting (fig. S15A). We were unable to examine the subnuclear localization of endogenous RAGs in a more physiological context because appropriate antibodies for fluorescence microscopy are currently unavailable. However, immunoblotting of *Nudcd3*^{Hom} thymic lysates revealed overexpression of RAG1, in keeping with its nucleolar accumulation in the absence of WT NUDCD3 (fig. S15B). Thus, the redistribution of RAG1 from nucleoli, known to be necessary for its recombinase function, is not only RAG2-dependent but also requires a specific activity of NUDCD3 that the G52D allele lacks.

DISCUSSION

These studies establish *NUDCD3* as a disease gene for T-B- SCID and OS. Its integrity is required for efficient B and T cell development, specifically for RAG-dependent V(D)J recombination, and consequently the generation of antigen receptor diversity upon which adaptive immunity depends. The close phenocopy between human NUDCD3^{G52D}- and RAG-related OS suggests that these disorders would behave similarly in response to treatment. Indeed, two of the eight *NUDCD3*-mutated patients who progressed to HSCT survived, whereas the remaining patients succumbed to a variety of infectious and inflammatory complications (table S1). The generally poor outcome of this cohort highlights the life-threatening nature of SCID/OS and the importance of early diagnosis and therapy. Infants with suspected T cell immunodeficiency, including those identified through newborn TREC screening, should be screened for pathogenic variants of the *NUDCD3* gene.

The focus of this work was to understand the role of NUDCD3 within the immune system because of the severity of the G52D-associated immunological phenotype, the absence of consistent extraimmune features, and the long-term survival of two affected individuals after HSCT. However, given its widespread expression and behavior in essentiality screens, it is likely that human NUDCD3 performs other functions that are less sensitive to the G52D substitution and/or overall NUDCD3 expression level. The fact that knock-in mice were small and sterile may well reflect the compromise of certain core functions of murine NUDCD3, perhaps in keeping with the mouse homolog's lower protein expression. It is not surprising that there would be differences in the degree of redundancy and sensitivity to individual missense mutations between mouse and human homologs. Further work will be required to clarify the NUDCD3 interactome and the extent to which NUDCD3's functions can be fulfilled by alternative cochaperones among different cell types and species.

The molecular mechanism by which RAG2 enables the egress of RAG1 from nucleoli has so far been elusive. First described as a hub for ribosome biogenesis, nucleoli have been proposed to play an important part in nuclear protein quality control based on the nucleolar colocalization of metastable nuclear proteins with chaperones such as Hsp70 following heat stress that can enable their refolding (*43*, *44*). The observed nucleolar retention of RAG1 may therefore reflect its tendency towards aggregation as well as its possession of a specific arginine- and histidine-rich motif previously determined to be a nucleolar retention signal (*40*). Nonetheless, a regulatory aspect to this arrangement is implied by the observations that forced nucleolar exclusion of RAG1 favors increased recombination activity (*40*) and that decoupling RAG2 expression from the cell cycle leads to genomic instability and lymphoid tumorigenesis (*8*). A parsimonious model would see NUDCD3 cooperating with RAG2 and heat shock proteins to enable RAG1 to refold and exit the nucleolus, a cochaperone function reminiscent of that executed by other NudC family members with respect to

distinct client proteins (*19*, *20*). NUDCD3^{G52D}, possibly because of its impaired homodimerization, is unable to fulfil this cochaperone function efficiently, leading to the insufficient withdrawal of RAG1 from nucleoli and diminished recombinase activity. The milder impact of this variant in mice, especially on the B cell compartment, implies greater redundancy with respect to this key immunological function, possibly supported by alternative cochaperones. Understanding this complex biology promises additional perspectives on V(D)J recombination and related disorders.

MATERIALS AND METHODS

Human subjects

Patients were cared for by Newcastle upon Tyne Hospitals NHS Foundation Trust or Oslo University Hospitals. Written informed consent was provided by all human subjects or their legal guardians in accordance with the 1975 Helsinki principles for enrolment in research protocols that were approved by the Newcastle and North Tyneside Research Ethics Committee 1, UK (REC reference 16/NE/0002) or by the regional ethical committee for medical and health research ethics in Norway (REC South-East).

Genetic Analysis

DNA was obtained from patient peripheral blood mononuclear cells or dermal fibroblast cultures using Qiagen's DNeasy Kit. After library preparation and target enrichment using Agilent SureSelect Human All Exon Kit (version 2 or 5), whole-exome sequencing (WES) was performed on an Illumina platform. Alignment, variant calling, and annotation were performed by standard methods based on Genome Analysis Toolkit best practice and ANNOVAR. Pathogenic variants in known disease-causing genes were first excluded by querying against a virtual panel based on the contemporary IUIS classification (*2*). Exome data from family 1 (individuals 1.1 and 1.2) were cross-referenced with family 2 (proband 2.1) to identify shared rare homozygous variants, of which *NUDCD3* c.155G>A was the sole instance. Exome data from patient 3.1 were obtained separately but revealed the same homozygous variant. Variant genotypes of cases and their family members were confirmed by Sanger sequencing (Fig. 1, A and B) using the following sequencing primers: Fwd 5′-GGCAACGTCCAGGATTTCC-3′; Rev 5′-TCCTCCTTTTCCTGGTGTCC-3′.

Plasmid construction

NUDCD3 was cloned into pDONR207, pCR3-C-HA and pCR3-C-Myc by Gateway cloning. Site-directed mutagenesis on pDONR207-NUDCD3 to generate G155A variant was performed by Agilent QuikChange II XL Kit according to manufacturer's instruction using the following primers: Fwd 5′-CGGGCGGGAAGTCCATGCGGTCC-3′ and Rev 5′-GGACCGCATGGACTTCCCGCCCG-3′. HA-tagged WT/G155A mutant NUDCD3 were cloned into pCDH-EF1α-MCS*-T2A-GFP for mass spectrometry purpose using pDONR221-NUDCD3-2×HA purchased from Invitrogen GeneArt Gene Synthesis as templates. *RAG1*- or *RAG2*-containing pDONR223 plasmids were purchased from Horizon Discovery. Depending on the purposes of either transient transfection or lentiviral transduction, relevant coding sequences were subcloned into pLenti6/V5-DEST (Invitrogen), pDEST-N-eGFP (Addgene), pDEST-N-mCherry (Addgene) by Gateway LR reaction. Lentiviral transfer plasmids pLenti-RAG1-eGFP and pLenti-RAG2-mCherry were created by substituting V5-tag in pLenti6-RAG1/RAG2-V5 with eGFP and mCherry CDS cloned from pDEST-N-eGFP, pDEST-N-mCherry using BstBI and AgeI restriction sites. The primers for cloning eGFP and mCherry CDS were: Fwd 5′-GGGCCCGCGGTTCGAAATGGTGAGCAAGGGCGAG-3′ and Rev 5′-TACTAACCGGTACGTTACTTGTACAGCTCGTCC-3′. Retroviral transfer plasmids pMSCV-v-Abl and pMGINV were kindly provided by Dr B.-R. Chen and Professor B. Sleckman (University

of Alabama, Birmingham). *Bcl2* retroviral plasmid was created by replacing the expression cassette in TCR OTI-2A.pMIG II (Addgene #52111) with murine *Bcl2* CDS (NM_009741.5) using BglII and NotI restriction sites.

Immunoblotting

Relevant cells were washed with PBS and lysed using radio-immunoprecipitation assay buffer (RIPA) (150 mM sodium chloride, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulphate, and 50 mM Tris pH7.4) supplemented with cOmplete™ Protease Inhibitor Cocktail and PhosSTOP™ Phosphatase inhibitor Cocktail (Roche, Switzerland). Lysates were centrifuged at 14,000*g* for 10 min and cleared supernatants were denatured at 70°C for 15 min with 10% dithiothreitol (DTT) and 1X NuPAGE LDS Sample Buffer (Thermo Fisher Scientific, USA). Samples were then loaded on to 4 to 12% Bis-Tris gel alongside prestained protein ladder (PageRuler Plus, Thermo Fisher Scientific, USA) for gel electrophoresis in 1X NuPAGE MOPS SDS Running Buffer (Invitrogen, USA). Either an equal volume of lysate was loaded or the protein concentration was measured with the BCA protein assay (Thermo Fisher Scientific, USA) prior to loading. Proteins were transferred to 0.45-µm polyvinyl difluoride (PVDF) membranes (Millipore, USA) at 20V using 1X NuPAGE Transfer Buffer (Invitrogen, USA) in 20% methanol. Membranes were blocked for 60 min using either 5% bovine serum albumin in tris-buffered saline with 0.1% Tween 20 (TBS-T) or 5% milk prior to immunostaining. Membranes were incubated overnight with relevant primary antibodies followed by the appropriate secondary antibody. Membranes were developed with Immobilon ECL substrate (Millipore, USA) and chemiluminescent images were visualized with the LI-COR Odyssey (LI-COR, USA) with LI-COR Image Studio software version 5.2.5. Primary and secondary antibodies used for immunoblotting and their normal dilutions are listed in data S2.

Size-exclusion chromatography

HEK293T cells were seeded in 100-mm diameter Petri dishes at 4×10^6 cells per plate followed by polyethylenimine (PEI, Sigma) transfection with 8 µg of WT or G52D mutant NUDCD3 plasmids (both in pEBB backbone). Transfected cells were lysed in a lysis buffer with 50 mM HEPES-KOH, 150 mM or 500 mM NaCl, 2 mM EDTA, 5% glycerol, and 0.5% Triton X-100, pH 7.5. Lysates were then applied to a Superose 12 column (GE Healthcare) and eluted in PBS at a rate of 0.5 ml/min. Fractions were TCA precipitated with recombinant human insulin (Sigma-Aldrich) as a carrier, and pellets were resuspended in 100 µl of 1X LDS sample buffer (Thermo Fisher) with 1 mM DTT. Samples were then heated and analyzed by immunoblotting as described above.

Coimmunoprecipitation

Cells were lysed in IP buffer composed of 25 mM Tris pH 7.4, 1 mM EDTA, 150 mM NaCl, 1% Nonidet P-40, 1 mM sodium orthovanadate, and 10 mM sodium fluoride, with Roche Complete proteinase inhibitor. The cell lysates were centrifuged at 17,000*g* at 4°C for 10 min. Soluble fractions were precleared for 1 hour at 4°C with Protein G Sepharose 4 beads (Fast Flow, GE Healthcare) that were blocked with 1% BSA IP buffer for 1 hour. Precleared cell lysates were immunoprecipitated overnight with blocked beads that were incubated with antibody for 1 hour. Protein G Sepharose beads were washed three times in IP buffer and then boiled with 4X LDS buffer at 95°C for 10 min to elute the absorbed immunocomplexes. Pulldown samples were analyzed by immunoblotting as described above. Immunoprecipitation for mass spectrometry is described below.

Lentiviral and retroviral transduction

Lentiviruses were produced using Invitrogen ViraPower ™ Lentiviral Systems (Thermo Fisher) or by cotransfection of psPAX2, pCMV-VSV-G (kindly provided by D. Young) and the relevant lentiviral transfer plasmid into HEK293T cells using PEI. Viral particle-contained supernatant was collected 48-96 hours post transfection and filtered with a 0.45-µm sterile filter. Viral particles were

purified by mixing the supernatant with Lenti-X™ Concentrator (TaKaRa 631231) according to manufacturer's instruction, the mixture was centrifuged at 1500*g* at 4°C for 45 min before the pellet was resuspended in 1% of the original volume of supernatant. Patient or healthy donor dermal fibroblasts or MOLT4 cells were spinoculated in six-well plates for 1.5 hours at 845*g*, with target or null control viral particles serially diluted in a total volume of 0.5 ml of DMEM/RPMI containing hexadimethrine bromide (8 mg/ml) (Polybrene, Sigma). Cells were rested in virus-containing medium for 8 hours and then incubated in fresh DMEM/RPMI for 48 hours, when they were selected in DMEM containing blasticidin (4 mg/ml) (Fisher Scientific), or transduced MOLT4 cells were sorted according to GFP intensity. Blasticidin-containing medium was refreshed every 72 hours.

Retroviruses were produced by cotransfecting HEK293T cells with pCL-Eco (Addgene #12371) and relevant retroviral transfer plasmid using PEI. Viral particle-containing supernatant was collected 48 hours post transfection and filtered with a 0.45-μm sterile filter. Viral particles were purified by centrifuging at 16,639*g* for 4–5 hours at 4°C. Mouse mononuclear bone marrow (BM) cells were isolated from BM harvested from femus and tibia by Histopaque 1083 using manufacturer's instructions. These mononuclear BM cells were cultured in BM media (RPMI supplemented with 15% FBS, 5 μM 2-mercaptoethanol (Sigma), 2 mM L-glutamine, 100 U/ml penicillin, and 0.1 mg/ml streptomycin (Sigma)). A total of 5×10^5 mononuclear BM cells were spinoculated in 48-well plates for 75 min at 1600*g* with viral particles diluted in a total volume of 0.2 ml of BM media containing 1 mg/ml of Synperonic® F 108 (Sigma) and 10 ng/ml of murine IL-7 (BioLegend). Transduced cells were cultured in 1 ml of BM media that was refreshed every 48-72 hours.

Coimmunoprecipitation mass spectrometry (CoIP/MS)

A total of 2×10^8 MOLT4 cells transduced with empty vector (EV), HA-tagged WT or G52D NUDCD3 were lysed in 1% NP-40 buffer by sonication. PureProteome Protein A/G Mix magnetic beads were crosslinked with 20 μg of anti-HA antibody (Abcam, ab9110) using 25 mM DMP (dimethyl pimelimidate, Thermo Fisher Scientific) in 200 mM TEA buffer (triethanolamine, Sigma) and quenched with addition of pH 7.4 Tris. The immunoprecipitation was performed by loading precleared lysate to conjugated beads at 4°C overnight before washing and elution with 2.5% SDS in 100 mM AmBic for three times. Samples were digested into peptides by ProtiFi S-Trap™ Micro Spin Column kit according to the manufacturer's recommended protocol. Proteins were reduced with 20 mM TCEP (Tris(2-carboxyethyl)phosphine hydrochloride) (Sigma) at 37°C for 30 min and alkylated with 20 mM iodoacetamide (Sigma) in the dark for 30 min. A ratio of 1:10 (w:w) Trypsin TPCK Treated (Worthington-Biochem) was used to digest the samples for 2 hours at 47°C. Eluted peptides were dried down and resuspended in 5% formic acid. Peptides were analyzed by nanoflow-LC-MS/MS using a Fusion Lumos Tribrid Orbitrap mass spectrometer (Thermo Scientific) coupled to a Dionex Ultimate 3000 with an active gradient of 60 min, following the conditions published previously (*45*).

All discovery proteomics RAW mass spectra were analyzed using MaxQuant (version 1.6.5.0) (*46*) and searched against a SwissProt *Homo sapiens* database (*47,48*), as previously described (*45*). The statistical analysis was done using the R package Limma with a Benjamini–Hochberg FDR threshold of *P*-value of <0.05 (*49*).

Immunofluorescence and microscopy

Fibroblasts transduced with lentiviruses encoding V5- or eGFP- or mCherry tagged RAG1 or RAG2, or HeLa cells transfected with plasmids encoding HA- or Myc-tagged WT/G52D NUDCD3, or HEK293T cells transfected with non-targeting siRNA/*NUDCD3* siRNA (Thermo Fisher, Silencer® Select s23709, s23711, 4390843) using DharmaFECT 1 (Horizon) followed by plasmids encoding eGFP-tagged RAG1 and mCherry tagged RAG2 in μ-Slide eight-well ibiTreat chamber slide (ibidi,

Germany) were fixed with 4% formaldehyde for 15 min and washed with phosphate buffered saline (PBS). Slides were blocked with 10% goat serum and permeabilized with 0.1% Triton X-100 in PBS and then incubated with relevant antibodies at 4°C overnight. Slides were then washed with 0.1% Tween-20 in PBS and PBS and incubated with appropriate secondary antibodies for 1 hour at room temperature, followed by nuclear staining with 0.2 µg/ml of 4′,6-diamidino-2-phenylindole (DAPI) (Thermo Fisher Scientific, USA). Slides were mounted with ProLong™ Glass Antifade Mountant (Thermo Fisher Scentific, USA) and imaged on a Leica SP8-STED confocal microscope with a Leica HC PL APO 63X/f1.40 oil CS2 lens. The image acquisition was carried out by Leica LAS X software (version 3.3.0.16799). Confocal images were deconvolved and analyzed by Huygens Essential/Professional software (version 23.04). For the analysis of nucleolar localization, Leica LAS X software was used to measure average raw fluorescence intensity of region of interest (ROI) using Line Profile and Stack Profile tools and results are presented as a simple ratio. The primary and secondary antibodies used are listed in data S2.

Mouse tissues fixed in 4% formaldehyde in PBS were paraffin-embedded and sectioned at 8 µm and H&E stained by the Laboratories of the Integrated Laboratory Medicine Directorate, Newcastle upon Tyne Hospitals NHS Foundation Trust. Slides were imaged using an Olympus BX43 with an Olympus PLN Plan Achromat 10X/f0.25 lens, SC50 camera, and cellSens Standard software (version 1.16, build 15404 software).

Reverse transcription quantitative PCR (RT-qPCR)

RNA was extracted by lysing cells in TRIzol® reagent (Thermo Fisher Scientific). Chloroform was added to 1/5 total sample volume was added and the sample was centrifuged for 15 min at 12,000*g* at 4°C. The top aqueous phase was transferred to a new tube and 2.5 V of aqueous phase of 75% ethanol, 0.1 V of 3 M sodium acetate (pH 5.2), and 20 µg of glycogen were added to the aqueous solution and incubated at −80°C overnight before centrifuging for 30 min at 12,000*g* at 4°C. RNA pellets were washed twice with 75% ethanol and dried at room temperature before resuspending in 30 µl of RNase-free water by heating at 55°C for 10 min. RT-qPCR primers and related probes were designed by Roche Universal ProbeLibrary System Assay Design. The primer sequences and related probes were as follows otherwise specified in other methods sections: NUDCD3, Fwd 5′-ACCTTTGACTACCACCAGAAGC-3′, Rev 5′-TCAGCATCTCATGGACTTTCA-3′ and UPL probe #21; 18s rRNA, Fwd 5′-CCGATTGGATGGTTTAGTGAG-3′, Rev 5′-AGTTCGACCGTCTTCTCAGC-3′ and UPL probe #81; GAPDH, Fwd 5′-TGGTATCGTGGAAGGACTCA-3′, Rev 5′-GCCATCACGCCACAGTTT-3′ UPL probe #87; HPRT1, Fwd 5′-TGACCTTGATTTATTTTGCATACC-3′, Rev 5′-CGAGCAAGACGTTCAGTCCT-3′ UPL probe #73. Equal amounts of RNA were subjected to qPCR with a Verso 1-step RT-qPCR Mix (Thermo Fisher Scientific) and AriaMx Real-time PCR System (Agilent Technologies) according to the manufacturer's instructions.

Radiosensitivity assays

A549 cells were transfected with 200 nM non-targeting siRNA (Thermo Fisher, 4390843) or siRNA targeting *NUDCD3* (Thermo Fisher, Silencer® Select s23709, s23710, S2371) using a Neon electroporator (Thermo Fisher) according to the manufacturer's protocol for 48 hours. Transfected A549 cells were then trypsinized and γ-irradiated by using a ^{137}Cs source at a dose rate of 0.04 Gy/s. Irradiated cells and untreated control cells were split in six-well plates and cultured for 21 days. Wells were exposed to 100 µl of MTT (5 mg/ml) in PBS for 2 hours after 21-day culture. Media then were removed and 200 µl isopropanol:HCl (100:1) was added to dissolve MTT. 100µl of supernatant (containing MTT in isopropanol:HCl) was transferred to a 96-well plate and analyzed at 570 nm to determine cell viability.

Primary dermal fibroblasts (passage 5–9) were trypsinized and irradiated at a dose rate of 0.09 Gy/s

before plating onto feeder cells prepared 24 hours earlier. Colony survival was analyzed at 3 weeks.

RAG recombination assay

Recombination assays were performed as described previously (*50*). Briefly, patient fibroblasts (5×10^5) were transfected with 2 μg of pEBB-RAG1, 2 μg of pEBB-RAG2, and 1 μg pDVG93 substrate plasmids using the Amaxa nucleofection system (Lonza), following the manufacturer's instructions. Cells were transferred to a prewarmed six-well plate containing 2 ml of complete DMEM and incubated for 48 hours. Extrachromosomal DNA was isolated using a modified Hirt prep and eluted in 20 μl of H_2O. After overnight DpnI digestion at 37°C, a qRT-PCR assay was used to detect the amount of substrate plasmid recombination using the primers DG89 (5′-CACAGCAGCGGCCATATCGAAGGTCG-3′) and DG147 (5′-TACATTGAGCAACTGACTGAAATGCC-3′), and the probe FM23 (5′-[FAM]CTCCATTTTAGCTTCCTTAGCTCCTG[TAMRA]-3′). This signal was normalized to total levels of transfected pDVG93 that were detected using the primers DpnI-U (5′-GAGCGTCAGACCCCGTAGAA-3′) and DpnI-L (5′-TAGCTCTTGATCCGGCAAACA-3′), and the probe T-DpnI (5′-[FAM]TTTCTGCGCGTAATCTGCTGCTTGCA[TAMRA]-3′). Assays were performed in a reaction volume of 25 μl using 2X TaqMan MasterMix (Thermo Fisher), 300 nM each primer, 100 nM probe, and 0.1 mg/ml of BSA using the following program: 2 min at 50°C, 10 min at 95°C, 50 cycles of 15 s at 95°C, and 1 min at 60°C.

Single-cell RNA sequencing (scRNA-seq) sample processing and analysis

Cryopreserved peripheral blood mononuclear cells in cryovials were thawed at 37°C, transferred into a 15-ml tube with 10 ml of prewarmed culture media RPMI-1640 medium (Sigma Aldrich, R0883) supplemented with 10% (v/v) FCS (Gibco, 10270-106), followed by centrifugation at 500*g* for 5 min. The cell pellet was resuspended in 100 μl of cold FACS buffer (PBS + 2% (v/v) FCS + 2 mM EDTA) with mixture of flow cytometry antibodies. Cells were incubated on ice for 30 min in the dark, washed and resuspended in FACS buffer. Viability dye 7AAD was added into cell suspension before cell sorting and analysis on BD FACSAria™ Fusion (Becton Dickinson). Sorted cell buckets (i) $CD3^+/CD19^+$ and (ii) $CD3^-CD19^-$ were washed. After cell counts were determined, cells were loaded onto each channel of a Chromium chip before encapsulation on the Chromium Controller (10x Genomics). Three sample pools (CTRL_1, RAG2_1, NUDCD3_1) contained 12,000 cells within each cell bucket. One sample pool (RAG1_1 and RAG2_2) was loaded twice for each cell bucket having 12,000 cells in each load.

For one sample pool (RAG2_1 and NUDCD3_1), 24,000 cells were loaded for each cell bucket. Two sample pools (RAG1_2 and RAG1_6, and CTRL_3 and NUDCD3_2) were loaded twice for each cell bucket having 16,000 cells in each load. Other sample pools (RAG1_3, RAG1_4, RAG1_5, NUDCD3_3) had 16,000 cells in each load for each cell bucket.

The single-cell sequencing libraries were generated using the Single Cell 5′ V1 kit, as per the manufacturer's protocol. The 5′ gene expression libraries (for CTRL_1, RAG2_1, NUDCD3_1, and RAG2_1 and NUDCD3_1 for each bucket separately) were sequenced using the Illumina HiSeq 4000 platform, with 10x read parameters: 26 bp read 1, 98 bp read 2, 8 bp index 1, 0 bp index 2. The 5′ gene expression libraries for the rest of samples (RAG1_1 and RAG2_2, RAG1_2 and RAG1_6, CTRL_3 and NUDCD3_2, RAG1_3, RAG1_4, RAG1_5, NUDCD3_3) were sequenced using the Illumina HiSeq 4000 platform with read parameters: 28 bp read 1, 91 bp read 2, 8 bp index 1, 0 bp index 2.

The T cell receptor repertoire was profiled for sorted $CD3^+/CD19^+$ cells along with 5′ gene expression, and these libraries were sequenced using the Illumina HiSeq 4000 platform with read parameters: 150 bp read 1, 150 bp read 2, 8 bp index 1, 0 bp index 2.

The antibodies used for cell sorting and immunophenotyping are listed in data S2. Flow cytometry data were analyzed by FlowJo V10 (BD Biosciences).

scRNA-seq data analysis

Raw scRNA-seq data coupled with TCR data of two healthy controls from publicly available datasets were used in the analysis of human TCR sequences (*51*, *52*). Raw scRNA-seq data from all 15 samples, including healthy controls, RAG1, RAG2 and NUDCD3 patients were processed using the Cell Ranger Single-Cell Software Suite (v3.0.2, 10x Genomics). Reads were first assigned to cells and then aligned to the human genome using STAR (*53*), with the hg38 build of the human genome (GRCh38_15_plus_hs38d1 + ensembl_90_transcriptome) as a reference for alignment. Results from RNA quantification in Cell Ranger were imported into R (v3.8.1) and analyzed using Seurat (v3) (*54*).

For doublet detection, DoubletFinder (*55*), DoubletDetection (*56*) and Solo (*57*) were applied. Cells were marked as doublets and removed from further analysis if any two algorithms predict a cell to be a doublet. For eight samples that were pooled together for sequencing in pairs (RAG2_1 and NUDCD3_1, RAG1_1 and RAG2_2, RAG1_2 and RAG1_6, CTRL_3 and NUDCD3_2), cellSNP v0.1.6 and vireoSNP v0.1.3 were used, to additionally deconvolute doublets based on genotypes and whole exome sequencing data (*58*, *59*).

For quality control, normalization and data scaling, cells that expressed fewer than 300 genes, or detected more than 10% UMI from mitochondrial transcripts, or fewer than 200 UMIs were first excluded. Data were normalized for library size and log-transformed using default Seurat parameters with NormalizeData function. A publicly available list of cell-cycle genes (*60*) was used to perform cell-cycle scoring and to assign cells to their respective stage of the cell cycle using Seurat CellCycleScoring function. The data were then normalized with Seurat's SCTransform with default parameters, with regressing out the following covariates: number of expressed genes per cell, number of UMIs per cell, proportion of transcripts mapping to mitochondrial genes and cell cycle score.

For the eight samples (RAG2_1, NUDCD3_1, RAG1_1, RAG1_2, RAG1_6, RAG2_2, CTRL_3, NUDCD3_2) that were sequenced in two batches, sequencing runs were integrated with Seurat::IntegrateData, using all genes as integration anchors. Batch-corrected counts provided in the integrated slot of the Seurat data object were used for further analysis.

Next, for each cell bucket ($CD3^+$/$CD19^+$ and $CD3^-CD19^-$) separately we integrated all cells from all 15 samples together with Seurat::IntegrateData. The integration was performed on 5000 highly variable anchor genes (excluding mitochondrial transcripts, genes coding TRA and TRB, ribosomal proteins, *AL138963*, *AL133415* and *MALAT1* as they correlate with cell viability and TCR specificity rather than with cell type). PCA with Seurat::RunPCA and uniform manifold approximation and projection (UMAP) with Seurat::RunUMAP, using the first 25 PCS were computed. Then cells were clustered using first 25 PCs with Shared Nearest Neighbour (SNN) modularity optimization-based clustering algorithm implemented Seurat::FindNeighbors and Seurat::FindClusters, testing for optimal resolution parameter (0.1-1.5, step=0.1).

After the first round of cell clustering, each cell from the lymphoid bucket was annotated as either T cell or B cell, whereas each cell from the myeloid bucket was annotated as monocyte, NK cell, dendritic cell or other (Fig. 2 and figs. S5 and S6). T and B cells were identified based on expression of T cell markers (*CD3D*, *CD3E*, and *CD3G*), and B cell markers (*CD19*, *CD74A*, and *MS4A1*). T cell clusters were further annotated into CD4 T cell subsets ($CD4^+$), CD8 T cells ($CD8A^+$ and $CD8B^+$), γδ T cells ($TRDV^+$ and $TRGV^+$) and mucosal-associated invariant T (MAIT) cells ($TRAV1\text{-}2^+$ and $TRAJ12/20/33^+$). Monocytes were differentiated from NK and dendritic cells based on CD14 and CD16 markers expression, whereas NK cells were characterized by high expression of *NCAM1*,

NKG7, and *GNLY* coupled with expression of chemokines (*CCL4* and *CCL5*) and granzymes (*GZMK* and *GZMB*). Dendritic cells were annotated based on *FCER1A*, *CD74*, and *CD1C* expression.

To increase the accuracy of cell-type annotation, a second round of dimensionality reduction and clustering was performed, in which each of the six cell populations (all lymphoid cells, CD4 T cells, CD8 T cells, γδ T cells, all myeloid cells, and monocytes) were analyzed separately. Resulting clusters were annotated using known cell type specific gene markers and cluster-specific genes that were identified using the Wilcoxon rank-sum test (figs. S5 and S6). The Mann–Whitney *U* test was used to compare frequencies of the annotated cell subtypes between patients based on the mutation gene group: (i) *NUDCD3* and (ii) *RAG1*- and *RAG2*-deficient and healthy individuals. Owing to the small sample sizes and broad similarity between (i) and (ii), we grouped all OS patients together to perform pairwise comparisons with healthy controls to assess for statistical differences in cell subtype proportions. A non-parametric test was used for pairwise comparisons between patient groups and controls as normality of distribution cannot be established due to extremely small sample size. To account for multiple testing FDR correction was applied to all *P*-values obtained in pairwise comparisons between OS patients and healthy controls.

For single-cell TCR analysis, raw TCR FASTQ reads were processed with a patched version of cellranger v3.0.2 provided by 10x Genomics. We used GRCh38-alts-ensembl-3.1.0 V(D)J reference to obtain TRA, TRB and clonotype assignment. Shannon entropy (H) was estimated for TRA and TRB chains separately, for productive sequences only, as

$$H = -\sum_{i=1}^{S} p_i \log p_i$$

where p_i was a proportion of clonotype i, and S was a number of clonotypes in repertoire.

In the tree plots, each sample is represented as a set of rectangles, corresponding to unique clonotypes. Rectangle size is proportional to the number of cells with this clonotype. Colors are applied to help distinguish different clonotypes within a sample and rectangles of the same color do not represent the same clonotype.

KREC assay

103/BCL-2 pre-B cells were transfected with *Nudcd3*-targeting or non-targeting siRNA (Thermo Fisher, Silencer® Select s102125, s102126, s102127) using the mouse B cell nucleofector kit in conjunction with the Amaxa nucleofection system (Lonza), according to manufacturer's instructions. Transfected cells were cultured in antibiotic free media in the wells of 12-well plates at 33°C for 48 hours to allow sufficient knockdown. Plates were then transferred to 39°C or retained at 33°C, for a further 48 hours. After this incubation period cells were harvested and total DNA and total RNA were extracted using the QIAamp DNA mini kit (QIAGEN) and ReliaPrep RNA miniprep system (Promega) respectively, according to manufacturer's instructions. KREC were detected using a qPCR assay with a 900 nM final primer concentration and a 100 nM final probe concentration. The KREC assay signal (Fwd 5′-CTCCAATAAGTCACCCTTTCCTTGT-3′, Rev 5′-GGAGTGGATTCAGGACACTGCT-3′ and probe 5′-[FAM]CCAGTTTCTGCACGGGCAGTCAGTTAG[TAMRA]-3′) was normalized to the signal from a genomic albumin assay (Fwd 5′-TCACCTTTCCTATCAACCCCA-3′, Rev 5′-CGAAACACACCCCTGGAAAA-3′ and probe 5′-[FAM]TCTCCTCCTCCTCTTCGTCT[TAMRA]-3′). RNA from the same cells was used to synthesize cDNA using the M-MLV reverse transcriptase according to the manufacturer's instructions (Promega). Template cDNA (5 µl) was used in a final volume of 20 µl, using 2X SYBR green master mix (no ROX) (Bioline) and 400 nM of each primer. Assays for murine *Nudcd3* (Fwd 5′-TTCGGCTTTCTCTACCGCAA-3′ and Rev 5′-CCTTGGCCTCTTCCTCCTTT-3′) and the housekeeping gene *Hprt1* (Fwd 5′-CAAACTTTGCTTTCCCTGGT-3′ and Rev 5′-

TCTGGCCTGTATCCAACACTTC-3′) were performed using the following program: 2 min at 95°C, 40 cycles of 5 s at 95°C, 10 s at 60°C, 20 s at 72°C, and a melt-curve step to check for multiple PCR products.

Mouse strains

The *Nudcd3*$^{em1(IMPC)Wtsi}$ mice on a C57BL/6N6NTac/USA background (referred to as *Nudcd3*^{Hom} in the homozygous state) were generated by the Wellcome Trust Genome Campus (Hinxton, UK) by CRISPR-Cas9 editing using the gRNA sequence 5′-CGCCACCCTTCGGACCGCAT-3′, and template DNA sequence 5′-GCCCGCCGCCTCTCTCCACCTCACCTGCAGCACCAGGGCCTGTGCGGCCCCGGGCGGGAAGTCCATGCGGTCCGAAGGGTGGCGCAGCAGGCGGTAGAAGTCGGTCTTGCGGT-3′ (Fig 3A). Heterozygous *Nudcd3*^{Hom} (referred to hereafter as *Nudcd3*^{Het}) mice were subject to the International Mouse Phenotyping Consortium (IMPC) phenotyping process (*61*). All mice were housed in barrier facilities at Comparative Biology Centre (CBC), Newcastle University, UK. All experimental procedures were approved by CBC Newcastle University local authorities and were performed in accordance with UK Home Office regulations. All mice were housed in groups of seven maximum with a 12-hour light–dark cycle and provided food and water ad libitum. *Nudcd3*^{Het} mice were bred to generate homozygous *Nudcd3*^{Hom} mice. Progeny mice were genotyped from ear notches taken at P14. DNA was isolated using a DNeasy Blood & Tissue Kit (QIAGEN) and the "Purification of Total DNA from Animal Tissues (Spin-Column Protocol)". DNA was amplified by PCR using the following primers: Fwd 5′-CAATCAGCCGGTGTGAGGCG-3′ and Rev 5′-GGGTTTGTGATTCTCTCTGG-3′ in the following reaction volumes: 10 µl of 5X MyTaq Reaction Buffer, 2 µl of each primer, 2 µl of template, 0.5 µl of MyTaq™ HS DNA Polymerase (5000 U/ml) (meridian BIO-21112), and 33.5 µl of water for a total volume of 60 µl. The following PCR program was used: 1 min at 95°C, 37 cycles of 15 s at 95°C, 15 s at 57°C, 30 s at 72°C, 5 min at 72°C, and 10 min at 10°C on an Alpha Cycler 1 PCRmax machine. Sanger sequencing was performed by Eurofins Genomics using the following primer: 5′-CCAGAGAGAATCACAAACCC-3′.

The congenic strain of BoyJ mice used in mixed bone marrow chimera experiments (see below) was Ly5.1 Mouse (B6.SJL-PtprcaPepcb/BoyCrl) purchased from Charles River. Host recipients used in mixed chimeras were a double-fluorescent Cre-reporter mouse with universal expression of membrane-targeted tdTomato (mT) prior to Cre excision and membrane-targeted EGFP (mG) after Cre excision (sourced from Jackson Laboratory) (*62*). These mice were housed at the University of Birmingham Biomedical Services Unit. All experimental procedures were approved by the Birmingham Animal Welfare and Ethical Review Body and were performed in accordance with UK Home Office regulations.

Flow cytometry (mouse)

Mice were weighed prior to being sacrificed at 5–8 weeks of age by injecting 100 µl of Euthatal (200 mg/ml) (pentobarbitone injectable, Dipharma) intraperitoneally. Blood was then collected using cardiac puncture and diluted with 100 µl of 5000 IU/ml heparin. Death was confirmed by cervical dislocation. Female and male animals were pooled in the analysis except when comparing weight and thymic cellularity. For these readouts, only 5–12-week-old males were analyzed. Some younger (3 weeks) and older animals (40 weeks) from pilot studies were also included. The reduced Mendelian ratios of homozygous animals born to heterozygous breeding pairs meant that one litter of 6–8 mice often only included one homozygous pup. Animals were analyzed in triplicate alongside littermate controls (*Nudcd3*^{WT}, *Nudcd3*^{Het}, or *Nudcd3*^{Hom}) where possible. If two litter triplicates were 5–8 weeks old simultaneously they were phenotyped on the same day. Details of independent replicates of each experiment can be found in data S3. Thymi, spleens, and clavicular, axillary and mesenteric lymph nodes were isolated by blunt dissection and transferred to cold IMDM with 0.1% BSA (v/v).

Femurs were isolated by cutting just above the hip and below the knee joint. Thymocytes and lymphocytes isolated by mechanically dissociating the thymi, spleens, and lymph nodes either between two sections of nylon mesh (70-µm diameter) or between two glass slides. Single-cell suspensions were counted on a Tali Image Cytometer (ThermoFisher Scientific, USA) or using counting beads (Spherotech AccuCount beads ACBP-100-10). Blood was lysed in ACK lysis buffer (dH_2O with 150 mM NH_4Cl, 10 mM $KHCO_3$, and 0.1 mM EDTA) for 3 min at room temperature. Single cells were washed and resuspended in cold FACS buffer (PBS (GIBCO) with 5 to 10% FCS and 0.5% sodium azide or PBS (GIBCO) with or without $CaCl_2$ and $MgCl_2$) at a concentration of $0.5–5\times10^6$ cells per 100 µl with specific surface antibodies on a rocker in a dark cold room (4°C) for at least 30 min. For analysis of DN and ETP populations, 10^7 cells were stained with 100 µl of antibody mix. Cells were either stained with a viability dye prior to surface staining (fixable Dead Cell Stain Kit Zombie Aqua from Invitrogen™) or just prior to acquisition (DAPI, 1:2000). The antibodies used are listed in data S2. Foxp3 and Ki67 staining were performed using an intracellular Foxp3 kit purchased from eBioscience (00-5523-00). Streptavidin-BV786 (BD Horizon) (1:100) was used to reveal staining with biotinylated antibodies. iNKT were identified using unloaded PBS57-loaded CD1d tetramers (obtained from the National Institutes of Health Tetramer Facility). Samples were analyzed using the BD FACSymphony™ A5 or the BD LSRFortessa™ X20 running BD FACSDiva software followed by analysis in FlowJo V10 (BD Biosciences) and Graphpad Prism 8 or 9. Comparison of population frequency (Freq. of Parent) or cell number (Freq. of Total × organ cellularity counted as above) was performed using ANOVA (Bonferroni and alpha threshold set at 0.05) unless otherwise specified. Only statistically significant differences are marked in the figures.

Generation of mixed bone marrow (BM) chimeras

Host recipients used in mixed chimeras were a double-fluorescent Cre reporter mouse with universal expression of membrane-targeted tdTomato (mT) prior to Cre excision and membrane-targeted EGFP (mG) after Cre excision (sourced from Jackson Laboratory) (*62*) to allow the identification of host cells (mT+) (fig. S10K). Recipient mice were lethally irradiated (2×500 rad) and reconstituted intravenously with 5×10^6 T cell–depleted adult BM preparations, made from CD45.2 congenically marked *Nudcd3*^{WT} or *Nudcd3*^{Hom} cells mixed at a 1:1 ratio with CD45.1 congenically marked WT cells isolated from BoyJ mice. Depletion of T cells was performed using anti CD3-PE and anti-PE microbeads (Miltenyi Biotec) according to the manufacturer's instructions. Mice were sacrificed 28 days post-reconstitution, and tissues were analyzed by flow cytometry described above. The antibodies used in addition to the panels described above are listed in data S2.

Mouse TCR/BCR repertoire by 5′ RACE and analysis

RNA was extracted from approximately 20 mg of sliced and macerated mouse spleen that was snap-frozen in liquid N_2 upon isolation using RNeasy® Plus Mini kit (QIAGEN, Germany) following manufacturer's protocol. RNA concentration and quality (RNA Integrity Number, RIN) were determined by Agilent RNA 6000 Nano Kit using Agilent 2100 Bioanalyzer. Mouse TCR and BCR libraries were prepared using SMARTer Mouse TCR a/b Profiling Kit and SMARTer® Mouse BCR IgG H/K/L Profiling Kit (Takara Bio USA, Inc.) following the manufacturer's protocol. Concentration and quality of DNA libraries were determined by Agilent DNA 1000 Kit using Agilent 2100 Bioanalyzer. TCR and BCR libraries was pooled before being sequenced by Illumina MiSeq Platform using MiSeq Reagent Kit v3 (2×300 bp, 600 cycles). Bioinformatic analysis is described as follows.

Receptors were reconstructed and quantified from raw FASTQ reads with MIXCR (v3.0.3, Built-in V/D/J/C library: repseqio.v1.5) with default settings. TCRs with CDR3 containing stop codon or frameshift mutation were assigned nonproductive status. Diversity estimates were obtained as for human sequences, with productive sequences only.

Mouse immunoglobulin ELISA

Mouse whole blood was left undisturbed at room temperature for 30 min to clot, then centrifuged at 1000*g* for 10 min at 4°C. The resulting serum supernatant was collected for ELISA. The serum was diluted and ELISAs were performed using the following kits: Abcam Mouse IgA ELISA Kit (ab157717), Abcam Mouse IgG1 ELISA Kit (ab133045), Invitrogen Mouse IgG2a ELISA Kit (88-50420-22), Invitrogen Mouse IgG2b ELISA Kit (88-50430-22), Invitrogen Mouse IgG3 ELISA Kit (88-50440-22), Invitrogen Mouse IgM ELISA Kit (88-50470-22), Invitrogen Mouse IgE ELISA Kit (88-50460-22) and according to manufacturer's instructions. The results were read using Tecan Sunrise Microplate Reader.

Mouse pre-B cell in vitro recombination assay

Heterogenous mononuclear BM cells were first infected with Bcl2 retrovirus or coinfected by Bcl2 and v-Abl retroviruses as described above. Transduced cells were initially cultured in BM media supplemented with 10 ng/ml IL-7 (BioLegend) for 7 days. v-Abl pre-B cells were then maintained without IL-7 supplement and infected with pMGINV reporter virus for 72 hours and subsequently treated with 5 μM imatinib (Sigma) for 48 hours to trigger recombination. GFP-positive cells were evaluated using flow cytometry.

Statistics

Statistical analysis and verification of normal distribution were conducted using GraphPad Prism software except in the case of single cell transcriptomic analysis (see below). For comparisons between groups, a one-sample *t* test, an unpaired Student's *t* test or a one-way or two-way analysis of variance (ANOVA) with Bonferroni post hoc multiple comparisons test were used as appropriate and as described in the figure legends. Data (biological replicates) are presented as the 25th–75th percentile with line at median or means±SEM as described in the figure legends. $P<0.05$ is considered as significant. In all figures, *$P<0.05$, **$P<0.01$, ***$P<0.001$, and **** $P<0.0001$.

For single-cell transcriptomic analysis, the Mann–Whitney *U* test to compare frequencies of annotated cell subtypes between patients and healthy individuals was conducted in Python (v3.10) using stats.mannwhitneyu function from SciPy (v1.8.0) library. The Wilcoxon rank-sum test to identify cluster specific genes during single cell data analysis was conducted in R (v3.8.1) with Seurat (v3) package using FindMarkers or FindConservedMarkers functions applied to Seurat data objects.

Supplementary Materials

Figs. S1 to S15

Tables S1 to S6

Data S1 to S5

REFERENCES AND NOTES

1. A. Fischer, L. D. Notarangelo, B. Neven, M. Cavazzana, J. M. Puck, Severe combined immunodeficiencies and related disorders. *Nat Rev Dis Primers* **1**, 15061 (2015).
2. S. G. Tangye, W. Al-Herz, A. Bousfiha, C. Cunningham-Rundles, J. L. Franco, S. M. Holland, C. Klein, T. Morio, E. Oksenhendler, C. Picard, A. Puel, J. Puck, M. R. J. Seppanen, R. Somech, H. C. Su, K. E. Sullivan, T. R. Torgerson, I. Meyts, Human Inborn Errors of Immunity: 2022 Update on the Classification from the International Union of Immunological Societies Expert Committee. *J Clin Immunol*, (2022).
3. A. Villa, L. D. Notarangelo, C. M. Roifman, Omenn syndrome: inflammation in leaky severe combined immunodeficiency. *J Allergy Clin Immunol* **122**, 1082-1086 (2008).
4. J. M. Puck, Newborn screening for severe combined immunodeficiency and T-cell lymphopenia. *Immunol Rev* **287**, 241-252 (2019).
5. J. P. de Villartay, Congenital defects in V(D)J recombination. *Br Med Bull* **114**, 157-167 (2015).
6. C. Liu, Y. Zhang, C. C. Liu, D. G. Schatz, Structural insights into the evolution of the RAG recombinase. *Nat Rev Immunol* **22**, 353-370 (2022).
7. G. Teng, D. G. Schatz, Regulation and Evolution of the RAG Recombinase. *Adv Immunol* **128**, 1-39 (2015).
8. L. Zhang, T. L. Reynolds, X. Shan, S. Desiderio, Coupling of V(D)J recombination to the cell cycle suppresses genomic instability and lymphoid tumorigenesis. *Immunity* **34**, 163-174 (2011).
9. A. R. Gennery, E. Hodges, A. P. Williams, S. Harris, A. Villa, B. Angus, A. J. Cant, J. L. Smith, Omenn's syndrome occurring in patients without mutations in recombination activating genes. *Clin Immunol* **116**, 246-256 (2005).
10. A. Stray-Pedersen, H. S. Sorte, P. Samarakoon, T. Gambin, I. K. Chinn, Z. H. Coban Akdemir, H. C. Erichsen, L. R. Forbes, S. Gu, B. Yuan, S. N. Jhangiani, D. M. Muzny, O. K. Rodningen, Y. Sheng, S. K. Nicholas, L. M. Noroski, F. O. Seeborg, C. M. Davis, D. L. Canter, E. M. Mace, T. J. Vece, C. E. Allen, H. A. Abhyankar, P. M. Boone, C. R. Beck, W. Wiszniewski, B. Fevang, P. Aukrust, G. E. Tjonnfjord, T. Gedde-Dahl, H. Hjorth-Hansen, I. Dybedal, I. Nordoy, S. F. Jorgensen, T. G. Abrahamsen, T. Overland, A. G. Bechensteen, V. Skogen, L. T. N. Osnes, M. A. Kulseth, T. E. Prescott, C. F. Rustad, K. R. Heimdal, J. W. Belmont, N. L. Rider, J. Chinen, T. N. Cao, E. A. Smith, M. S. Caldirola, L. Bezrodnik, S. O. Lugo Reyes, F. J. Espinosa Rosales, N. D. Guerrero-Cursaru, L. A. Pedroza, C. M. Poli, J. L. Franco, C. M. Trujillo Vargas, J. C. Aldave Becerra, N. Wright, T. B. Issekutz, A. C. Issekutz, J. Abbott, J. W. Caldwell, D. K. Bayer, A. Y. Chan, A. Aiuti, C. Cancrini, E. Holmberg, C. West, M. Burstedt, E. Karaca, G. Yesil, H. Artac, Y. Bayram, M. M. Atik, M. K. Eldomery, M. S. Ehlayel, S. Jolles, B. Flato, A. A. Bertuch, I. C. Hanson, V. W. Zhang, L. J. Wong, J. Hu, M. Walkiewicz, Y. Yang, C. M. Eng, E. Boerwinkle, R. A. Gibbs, W. T. Shearer, R. Lyle, J. S. Orange, J. R. Lupski, Primary immunodeficiency diseases: Genomic approaches delineate heterogeneous Mendelian disorders. *J Allergy Clin Immunol* **139**, 232-245 (2017).
11. K. Consortium GenomeAsia, The GenomeAsia 100K Project enables genetic discoveries across Asia. *Nature* **576**, 106-111 (2019).
12. K. J. Karczewski, L. C. Francioli, G. Tiao, B. B. Cummings, J. Alfoldi, Q. Wang, R. L. Collins, K. M. Laricchia, A. Ganna, D. P. Birnbaum, L. D. Gauthier, H. Brand, M. Solomonson, N. A. Watts, D. Rhodes, M. Singer-Berk, E. M. England, E. G. Seaby, J. A. Kosmicki, R. K. Walters, K. Tashman, Y. Farjoun, E. Banks, T. Poterba, A. Wang, C. Seed, N. Whiffin, J. X. Chong, K. E. Samocha, E. Pierce-Hoffman, Z. Zappala, A. H. O'Donnell-

Luria, E. V. Minikel, B. Weisburd, M. Lek, J. S. Ware, C. Vittal, I. M. Armean, L. Bergelson, K. Cibulskis, K. M. Connolly, M. Covarrubias, S. Donnelly, S. Ferriera, S. Gabriel, J. Gentry, N. Gupta, T. Jeandet, D. Kaplan, C. Llanwarne, R. Munshi, S. Novod, N. Petrillo, D. Roazen, V. Ruano-Rubio, A. Saltzman, M. Schleicher, J. Soto, K. Tibbetts, C. Tolonen, G. Wade, M. E. Talkowski, C. Genome Aggregation Database, B. M. Neale, M. J. Daly, D. G. MacArthur, The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434-443 (2020).

13. I. A. Adzhubei, S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova, P. Bork, A. S. Kondrashov, S. R. Sunyaev, A method and server for predicting damaging missense mutations. *Nat Methods* **7**, 248-249 (2010).
14. M. Kircher, D. M. Witten, P. Jain, B. J. O'Roak, G. M. Cooper, J. Shendure, A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* **46**, 310-315 (2014).
15. Y. Wu, R. Li, S. Sun, J. Weile, F. P. Roth, Improved pathogenicity prediction for rare human missense variants. *Am J Hum Genet* **108**, 1891–1906 (2021).
16. Q. Fu, W. Wang, T. Zhou, Y. Yang, Emerging roles of NudC family: from molecular regulation to clinical implications. *Sci China Life Sci* **59**, 455-462 (2016).
17. M. E. Rebeaud, S. Mallik, P. Goloubinoff, D. S. Tawfik, On the evolution of chaperones and cochaperones and the expansion of proteomes across the Tree of Life. *Proc Natl Acad Sci U S A* **118**, (2021).
18. M. Taipale, G. Tucker, J. Peng, I. Krykbaeva, Z. Y. Lin, B. Larsen, H. Choi, B. Berger, A. C. Gingras, S. Lindquist, A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell* **158**, 434-448 (2014).
19. M. M. Biebl, F. Delhommel, O. Faust, K. M. Zak, G. Agam, X. Guo, M. Muhlhofer, V. Dahiya, D. Hillebrand, G. M. Popowicz, M. Kampmann, D. C. Lamb, R. Rosenzweig, M. Sattler, J. Buchner, NudC guides client transfer between the Hsp40/70 and Hsp90 chaperone systems. *Mol Cell* **82**, 555-569 e557 (2022).
20. Y. Yang, W. Wang, M. Li, Y. Gao, W. Zhang, Y. Huang, W. Zhuo, X. Yan, W. Liu, F. Wang, D. Chen, T. Zhou, NudCL2 is an Hsp90 cochaperone to regulate sister chromatid cohesion by stabilizing cohesin subunits. *Cell Mol Life Sci* **76**, 381-395 (2019).
21. T. Zhou, W. Zimmerman, X. Liu, R. L. Erikson, A mammalian NudC-like protein essential for dynein stability and cell viability. *Proc Natl Acad Sci U S A* **103**, 9039-9044 (2006).
22. Y. Cai, Y. Yang, M. Shen, T. Zhou, Inhibition of cytokinesis by overexpression of NudCL that is localized to the centrosome and midbody. *Cell Res* **19**, 1305-1308 (2009).
23. C. Suo, E. Dann, I. Goh, L. Jardine, V. Kleshchevnikov, J. E. Park, R. A. Botting, E. Stephenson, J. Engelbert, Z. K. Tuong, K. Polanski, N. Yayon, C. Xu, O. Suchanek, R. Elmentaite, C. Dominguez Conde, P. He, S. Pritchard, M. Miah, C. Moldovan, A. S. Steemers, P. Mazin, M. Prete, D. Horsfall, J. C. Marioni, M. R. Clatworthy, M. Haniffa, S. A. Teichmann, Mapping the developing human immune system across organs. *Science* **376**, eabo0510 (2022).
24. J. E. Park, R. A. Botting, C. Dominguez Conde, D. M. Popescu, M. Lavaert, D. J. Kunz, I. Goh, E. Stephenson, R. Ragazzini, E. Tuck, A. Wilbrey-Clark, K. Roberts, V. R. Kedlian, J. R. Ferdinand, X. He, S. Webb, D. Maunder, N. Vandamme, K. T. Mahbubani, K. Polanski, L. Mamanova, L. Bolt, D. Crossland, F. de Rita, A. Fuller, A. Filby, G. Reynolds, D. Dixon, K. Saeb-Parsy, S. Lisgo, D. Henderson, R. Vento-Tormo, O. A. Bayraktar, R. A. Barker, K. B. Meyer, Y. Saeys, P. Bonfanti, S. Behjati, M. R. Clatworthy, T. Taghon, M. Haniffa, S. A. Teichmann, A cell atlas of human thymic development defines T cell repertoire formation. *Science* **367**, (2020).
25. https://depmap.org/portal/gene/NUDCD3?tab=overview.

26. A. Berland, J. Rosain, S. Kaltenbach, V. Allain, N. Mahlaoui, I. Melki, A. Fievet, C. Dubois d'Enghien, M. Ouachee-Chardin, L. Perrin, N. Auger, F. E. Cipe, A. Finocchi, F. Dogu, F. Suarez, D. Moshous, T. Leblanc, A. Belot, C. Fieschi, D. Boutboul, M. Malphettes, L. Galicier, E. Oksenhendler, S. Blanche, A. Fischer, P. Revy, D. Stoppa-Lyonnet, C. Picard, J. P. de Villartay, PROMIDISalpha: A T-cell receptor alpha signature associated with immunodeficiencies caused by V(D)J recombination defects. *J Allergy Clin Immunol* **143**, 325-334 e322 (2019).
27. D. Brandle, C. Muller, T. Rulicke, H. Hengartner, H. Pircher, Engagement of the T-cell receptor during positive selection in the thymus down-regulates RAG-1 expression. *Proc Natl Acad Sci U S A* **89**, 9529-9533 (1992).
28. Y. N. Lee, F. Frugoni, K. Dobbs, I. Tirosh, L. Du, F. A. Ververs, H. Ru, L. Ott de Bruin, M. Adeli, J. H. Bleesing, D. Buchbinder, M. J. Butte, C. Cancrini, K. Chen, S. Choo, R. A. Elfeky, A. Finocchi, R. L. Fuleihan, A. R. Gennery, D. H. El-Ghoneimy, L. A. Henderson, W. Al-Herz, E. Hossny, R. P. Nelson, S. Y. Pai, N. C. Patel, S. M. Reda, P. Soler-Palacin, R. Somech, P. Palma, H. Wu, S. Giliani, J. E. Walter, L. D. Notarangelo, Characterization of T and B cell repertoire diversity in patients with RAG deficiency. *Sci Immunol* **1**, (2016).
29. X. Yu, J. R. Almeida, S. Darko, M. van der Burg, S. S. DeRavin, H. Malech, A. Gennery, I. Chinn, M. L. Markert, D. C. Douek, J. D. Milner, Human syndromes of immunodeficiency and dysregulation are characterized by distinct defects in T-cell receptor repertoire development. *J Allergy Clin Immunol* **133**, 1109-1115 (2014).
30. Y. Y. Chen, L. C. Wang, M. S. Huang, N. Rosenberg, An active v-abl protein tyrosine kinase blocks immunoglobulin light-chain gene rearrangement. *Genes Dev* **8**, 688-697 (1994).
31. S. Takeda, Y. R. Zou, H. Bluethmann, D. Kitamura, U. Muller, K. Rajewsky, Deletion of the immunoglobulin kappa chain intron enhancer abolishes kappa chain gene rearrangement in cis but not lambda chain gene rearrangement in trans. *EMBO J* **12**, 2329-2336 (1993).
32. M. van der Burg, T. Tumkaya, M. Boerma, S. de Bruin-Versteeg, A. W. Langerak, J. J. van Dongen, Ordered recombination of immunoglobulin light chain genes occurs at the IGK locus but seems less strict at the IGL locus. *Blood* **97**, 1001-1008 (2001).
33. K. Khiong, M. Murakami, C. Kitabayashi, N. Ueda, S. Sawa, A. Sakamoto, B. L. Kotzin, S. J. Rozzo, K. Ishihara, M. Verella-Garcia, J. Kappler, P. Marrack, T. Hirano, Homeostatically proliferating CD4 T cells are involved in the pathogenesis of an Omenn syndrome murine model. *J Clin Invest* **117**, 1270-1281 (2007).
34. V. Marrella, P. L. Poliani, A. Casati, F. Rucci, L. Frascoli, M. L. Gougeon, B. Lemercier, M. Bosticardo, M. Ravanini, M. Battaglia, M. G. Roncarolo, M. Cavazzana-Calvo, F. Facchetti, L. D. Notarangelo, P. Vezzoni, F. Grassi, A. Villa, A hypomorphic R229Q Rag2 mouse mutant recapitulates human Omenn syndrome. *J Clin Invest* **117**, 1260-1269 (2007).
35. L. M. Ott de Bruin, M. Bosticardo, A. Barbieri, S. G. Lin, J. H. Rowe, P. L. Poliani, K. Ching, D. Eriksson, N. Landegren, O. Kampe, J. P. Manis, L. D. Notarangelo, Hypomorphic Rag1 mutations alter the preimmune repertoire at early stages of lymphoid development. *Blood* **132**, 281-292 (2018).
36. U. Koch, F. Radtke, Mechanisms of T cell development and transformation. *Annu Rev Cell Dev Biol* **27**, 539-562 (2011).
37. T. Taghon, M. A. Yui, R. Pant, R. A. Diamond, E. V. Rothenberg, Developmental and molecular characterization of emerging beta- and gammadelta-selected pre-T cells in the adult mouse thymus. *Immunity* **24**, 53-64 (2006).
38. P. Matangkasombut, M. Pichavant, D. E. Saez, S. Giliani, E. Mazzolari, A. Finocchi, A. Villa, C. Sobacchi, P. Cortes, D. T. Umetsu, L. D. Notarangelo, Lack of iNKT cells in

patients with combined immune deficiency due to hypomorphic RAG mutations. *Blood* **111**, 271-274 (2008).
39. A. Bendelac, P. B. Savage, L. Teyton, The biology of NKT cells. *Annu Rev Immunol* **25**, 297-336 (2007).
40. R. M. Brecht, C. C. Liu, H. A. Beilinson, A. Khitun, S. A. Slavoff, D. G. Schatz, Nucleolar localization of RAG1 modulates V(D)J recombination activity. *Proc Natl Acad Sci U S A* **117**, 4300-4309 (2020).
41. E. Spanopoulou, P. Cortes, C. Shih, C. M. Huang, D. P. Silver, P. Svec, D. Baltimore, Localization, interaction, and RNA binding properties of the V(D)J recombination-activating proteins RAG1 and RAG2. *Immunity* **3**, 715-726 (1995).
42. T. Gan, Y. Wang, Y. Liu, D. G. Schatz, J. Hu, RAG2 abolishes RAG1 aggregation to facilitate V(D)J recombination. *Cell Rep* **37**, 109824 (2021).
43. P. Banski, M. Kodiha, U. Stochaj, Chaperones and multitasking proteins in the nucleolus: networking together for survival? *Trends Biochem Sci* **35**, 361-367 (2010).
44. F. Frottin, F. Schueder, S. Tiwary, R. Gupta, R. Korner, T. Schlichthaerle, J. Cox, R. Jungmann, F. U. Hartl, M. S. Hipp, The nucleolus functions as a phase-separated protein quality control compartment. *Science* **365**, 342-347 (2019).
45. J. L. Marin-Rubio, R. E. Peltier-Heap, M. E. Duenas, T. Heunis, A. Dannoura, J. Inns, J. Scott, A. J. Simpson, H. J. Blair, O. Heidenreich, J. M. Allan, J. E. Watt, M. P. Martin, B. Saxty, M. Trost, A Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Assay Identifies Nilotinib as an Inhibitor of Inflammation in Acute Myeloid Leukemia. *J Med Chem* **65**, 12014-12030 (2022).
46. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372 (2008).
47. UniProt Consortium, UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* **51**, D523-D531 (2023).
48. https://www.uniprot.org/uniprotkb?query=Human&facets=reviewed%3Atrue%2Cmodel_organism%3A9606 (downloaded 18 October 2018).
49. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
50. M. van der Burg, N. S. Verkaik, A. T. den Dekker, B. H. Barendregt, I. Pico-Knijnenburg, I. Tezcan, J. J. vanDongen, D. C. van Gent, Defective Artemis nuclease is characterized by coding joints with microhomology in long palindromic-nucleotide stretches. *Eur J Immunol* **37**, 3522-3528 (2007).
51. D. Dixon, Effects of STAT3 gain of function on myeloid cells in human peripheral blood mononuclear cells. (2020).
52. S. Spencer, S. Kostel Bal, W. Egner, H. Lango Allen, S. I. Raza, C. A. Ma, M. Gurel, Y. Zhang, G. Sun, R. A. Sabroe, D. Greene, W. Rae, T. Shahin, K. Kania, R. C. Ardy, M. Thian, E. Staples, A. Pecchia-Bekkum, W. P. M. Worrall, J. Stephens, M. Brown, S. Tuna, M. York, F. Shackley, D. Kerrin, R. Sargur, A. Condliffe, H. N. Tipu, H. S. Kuehn, S. D. Rosenzweig, E. Turro, S. Tavare, A. J. Thrasher, D. I. Jodrell, K. G. C. Smith, K. Boztug, J. D. Milner, J. E. D. Thaventhiran, Loss of the interleukin-6 receptor causes immunodeficiency, atopy, and abnormal inflammatory responses. *J Exp Med* **216**, 1986-1998 (2019).
53. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

54. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck, 3rd, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821 (2019).
55. C. S. McGinnis, L. M. Murrow, Z. J. Gartner, DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329-337 e324 (2019).
56. A. Gayoso, J. Shor, JonathanShor/DoubletDetection: doubletdetection v3.0 *Zenodo*, http://doi.org/10.5281/zenodo.4359992 (2020).
57. N. J. Bernstein, N. L. Fong, I. Lam, M. A. Roy, D. G. Hendrickson, D. R. Kelley, Solo: Doublet Identification in Single-Cell RNA-Seq via Semi-Supervised Deep Learning. *Cell Syst* **11**, 95-101 e105 (2020).
58. X. Huang, Y. Huang, Cellsnp-lite: an efficient tool for genotyping single cells. *Bioinformatics*, (2021).
59. Y. Huang, D. J. McCarthy, O. Stegle, Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference. *Genome Biol* **20**, 273 (2019).
60. I. Tirosh, B. Izar, S. M. Prakadan, M. H. Wadsworth, 2nd, D. Treacy, J. J. Trombetta, A. Rotem, C. Rodman, C. Lian, G. Murphy, M. Fallahi-Sichani, K. Dutton-Regester, J. R. Lin, O. Cohen, P. Shah, D. Lu, A. S. Genshaft, T. K. Hughes, C. G. Ziegler, S. W. Kazer, A. Gaillard, K. E. Kolb, A. C. Villani, C. M. Johannessen, A. Y. Andreev, E. M. Van Allen, M. Bertagnolli, P. K. Sorger, R. J. Sullivan, K. T. Flaherty, D. T. Frederick, J. Jane-Valbuena, C. H. Yoon, O. Rozenblatt-Rosen, A. K. Shalek, A. Regev, L. A. Garraway, Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189-196 (2016).
61. L. Abeler-Dorner, A. G. Laing, A. Lorenc, D. S. Ushakov, S. Clare, A. O. Speak, M. A. Duque-Correa, J. K. White, R. Ramirez-Solis, N. Saran, K. R. Bull, B. Moron, J. Iwasaki, P. R. Barton, S. Caetano, K. I. Hng, E. Cambridge, S. Forman, T. L. Crockford, M. Griffiths, L. Kane, K. Harcourt, C. Brandt, G. Notley, K. O. Babalola, J. Warren, J. C. Mason, A. Meeniga, N. A. Karp, D. Melvin, E. Cawthorne, B. Weinrick, A. Rahim, S. Drissler, J. Meskas, A. Yue, M. Lux, G. X. Song-Zhao, A. Chan, C. Ballesteros Reviriego, J. Abeler, H. Wilson, A. Przemska-Kosicka, M. Edmans, N. Strevens, M. Pasztorek, T. F. Meehan, F. Powrie, R. Brinkman, G. Dougan, W. Jacobs, Jr., C. M. Lloyd, R. J. Cornall, K. J. Maloy, R. K. Grencis, G. M. Griffiths, D. J. Adams, A. C. Hayday, High-throughput phenotyping reveals expansive genetic and structural underpinnings of immune variation. *Nat Immunol* **21**, 86-100 (2020).
62. M. D. Muzumdar, B. Tasic, K. Miyamichi, L. Li, L. Luo, A global double-fluorescent Cre reporter mouse. *Genesis* **45**, 593-605 (2007).
63. Y. Perez-Riverol, J. Bai, C. Bandla, D. García-Seisdedos, S. Hewapathirana, S. Kamatchinathan, D. J. Kundu, A. Prakash, A. Frericks-Zipper, M. Eisenacher, M. Walzer, S. Wang, A. Brazma, J. A. Vizcaíno, The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* **50**, D543-d552 (2022).
64. D. Taliun, D. N. Harris, M. D. Kessler, J. Carlson, Z. A. Szpiech, R. Torres, S. A. G. Taliun, A. Corvelo, S. M. Gogarten, H. M. Kang, A. N. Pitsillides, J. LeFaive, S. B. Lee, X. Tian, B. L. Browning, S. Das, A. K. Emde, W. E. Clarke, D. P. Loesch, A. C. Shetty, T. W. Blackwell, A. V. Smith, Q. Wong, X. Liu, M. P. Conomos, D. M. Bobo, F. Aguet, C. Albert, A. Alonso, K. G. Ardlie, D. E. Arking, S. Aslibekyan, P. L. Auer, J. Barnard, R. G. Barr, L. Barwick, L. C. Becker, R. L. Beer, E. J. Benjamin, L. F. Bielak, J. Blangero, M. Boehnke, D. W. Bowden, J. A. Brody, E. G. Burchard, B. E. Cade, J. F. Casella, B. Chalazan, D. I. Chasman, Y. I. Chen, M. H. Cho, S. H. Choi, M. K. Chung, C. B. Clish, A. Correa, J. E. Curran, B. Custer, D. Darbar, M. Daya, M. de Andrade, D. L. DeMeo, S. K.

Dutcher, P. T. Ellinor, L. S. Emery, C. Eng, D. Fatkin, T. Fingerlin, L. Forer, M. Fornage, N. Franceschini, C. Fuchsberger, S. M. Fullerton, S. Germer, M. T. Gladwin, D. J. Gottlieb, X. Guo, M. E. Hall, J. He, N. L. Heard-Costa, S. R. Heckbert, M. R. Irvin, J. M. Johnsen, A. D. Johnson, R. Kaplan, S. L. R. Kardia, T. Kelly, S. Kelly, E. E. Kenny, D. P. Kiel, R. Klemmer, B. A. Konkle, C. Kooperberg, A. Kottgen, L. A. Lange, J. Lasky-Su, D. Levy, X. Lin, K. H. Lin, C. Liu, R. J. F. Loos, L. Garman, R. Gerszten, S. A. Lubitz, K. L. Lunetta, A. C. Y. Mak, A. Manichaikul, A. K. Manning, R. A. Mathias, D. D. McManus, S. T. McGarvey, J. B. Meigs, D. A. Meyers, J. L. Mikulla, M. A. Minear, B. D. Mitchell, S. Mohanty, M. E. Montasser, C. Montgomery, A. C. Morrison, J. M. Murabito, A. Natale, P. Natarajan, S. C. Nelson, K. E. North, J. R. O'Connell, N. D. Palmer, N. Pankratz, G. M. Peloso, P. A. Peyser, J. Pleiness, W. S. Post, B. M. Psaty, D. C. Rao, S. Redline, A. P. Reiner, D. Roden, J. I. Rotter, I. Ruczinski, C. Sarnowski, S. Schoenherr, D. A. Schwartz, J. S. Seo, S. Seshadri, V. A. Sheehan, W. H. Sheu, M. B. Shoemaker, N. L. Smith, J. A. Smith, N. Sotoodehnia, A. M. Stilp, W. Tang, K. D. Taylor, M. Telen, T. A. Thornton, R. P. Tracy, D. J. Van Den Berg, R. S. Vasan, K. A. Viaud-Martinez, S. Vrieze, D. E. Weeks, B. S. Weir, S. T. Weiss, L. C. Weng, C. J. Willer, Y. Zhang, X. Zhao, D. K. Arnett, A. E. Ashley-Koch, K. C. Barnes, E. Boerwinkle, S. Gabriel, R. Gibbs, K. M. Rice, S. S. Rich, E. K. Silverman, P. Qasba, W. Gan, N. T.-O. f. P. M. Consortium, G. J. Papanicolaou, D. A. Nickerson, S. R. Browning, M. C. Zody, S. Zollner, J. G. Wilson, L. A. Cupples, C. C. Laurie, C. E. Jaquish, R. D. Hernandez, T. D. O'Connor, G. R. Abecasis, Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* **590**, 290-299 (2021).

65. J. Vandesompele, K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, F. Speleman, Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**, RESEARCH0034 (2002).
66. P. Lobachevsky, L. Woodbine, K. C. Hsiao, S. Choo, C. Fraser, P. Gray, J. Smith, N. Best, L. Munforte, E. Korneeva, R. F. Martin, P. A. Jeggo, O. A. Martin, Evaluation of Severe Combined Immunodeficiency and Combined Immunodeficiency Pediatric Patients on the Basis of Cellular Radiosensitivity. *J Mol Diagn* Vol. **17**, 560-575 (2015).
67. Y. Xing, X. Wang, S. C. Jameson, K. A. Hogquist, Late stages of T cell maturation in the thymus involve NF-kappaB and tonic type I interferon signaling. *Nat Immunol* **17**, 565-573 (2016).
68. Y. Choi, A. P. Chan, PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **31**, 2745-2747 (2015).
69. Y. Itan, L. Shang, B. Boisson, M. J. Ciancanelli, J. G. Markle, R. Martinez-Barricarte, E. Scott, I. Shah, P. D. Stenson, J. Gleeson, D. N. Cooper, L. Quintana-Murci, S. Y. Zhang, L. Abel, J. L. Casanova, The mutation significance cutoff: gene-level thresholds for variant predictions. *Nat Methods* **13**, 109-110 (2016).
70. J. M. Schwarz, C. Rodelsperger, M. Schuelke, D. Seelow, MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* **7**, 575-576 (2010).
71. L. Wiel, C. Baakman, D. Gilissen, J. A. Veltman, G. Vriend, C. Gilissen, MetaDome: Pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. *Hum Mutat* **40**, 1030-1038 (2019).
72. N. L. Sim, P. Kumar, J. Hu, S. Henikoff, G. Schneider, P. C. Ng, SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res* **40**, W452-457 (2012)

Acknowledgments: For their contributions to data generation and processing, we thank colleagues within the core facilities teams at both Newcastle University (Bioimaging, Flow Cytometry, Genomics, Comparative Biology Centre) and the Wellcome Sanger Institute (Sequencing, Cellular

Genetics, Informatics). We thank J. Nsengimana for expert review of statistical approaches. We also gratefully acknowledge S. Teichmann and M. Haniffa for sharing raw data generated in their groups at the Wellcome Sanger Institute and Newcastle University Translational and Clinical Research Institute. We thank L. Bossini-Castillo and S. Navaneethan for their contributions to data generation and analysis. We are grateful to patients and their families, clinical teams and the laboratories of the Integrated Laboratory Medicine Directorate, Newcastle upon Tyne Hospitals NHS Foundation Trust.

Funding: This research was funded in part by the Wellcome Trust. Grants supporting this work were: Wellcome Trust grant 207556/Z/17/Z (S.H.), Wellcome Multiuser Equipment grant 212947/Z/187 (M.T. and S.H.), Wellcome Trust grants WT206194 and 220540/Z/20/A (G.T.), Wellcome Trust grants WT098051 and WT206194 (D.J.A.), Sir Jules Thorn Charitable Trust Biomedical Award 12/JTA (S.H.), Medical Research Council (UK) grant MR/T029765/1 (G.A.), and National Human Genome Research Institute grant 5U54HG006542 through the Baylor-Hopkins Center for Mendelian Genomics (A.S.P.).

Author contributions: Clinical care: H.C.E., A.S.P., A.R.G., and S.H. Conceptualization: R.C., I.S.V.D.L., J.D.P.W., D.J.S., K.R.E., A.S.P., A.R.G., G.T., and S.H. Data curation: R.C., E.L., I.S.V.D.L., J.S.S., J.D.P.W., A.G., F.L., J.L.M.R., E.R.R.M., S.E., K.R.E., A.S.P., and A.L. Formal analysis: R.C., E.L., I.S.V.D.L., J.S.S., J.D.P.W., K.D.J., H.G., E.R.R.M., K.R.E., and A.L. Funding acquisition: A.S.P., M.T., D.J.A., G.A., G.T., and S.H. Investigation: R.C., I.S.V.D.L., J.S.S., J.D.P.W., K.D.J., H.I., A.G., F.L., L.W., H.L., V.P., and K.S. Methodology: R.C., I.S.V.D.L., E.J.R., and D.J.A. Resources: E.J.R., H.L., D.D., M.P., E.P., J.B., A.J.M., L.H., M.V.D.B., A.R.G., D.J.A., and S.H. Supervision: P.J., A.J.M., M.V.D.B., M.T., D.J.A., G.A., A.L., G.T., and S.H. Visualization: R.C., E.L., I.S.V.D.L., J.S.S., J.D.P.W., K.D.J., H.G., J.L.M.R., K.R.E., and A.L. Writing—original draft: R.C., E.L., I.S.V.D.L., J.S.S., A.L., G.T., and S.H. Writing—review and editing: all authors.

Competing interests: E.J.R. is an employee at LGC ASSURE. The authors declare that they have no other competing interests.

Data and materials availability: The mouse strain created for this work, on the C57BL/6NTac/USA background, can be obtained from the European Mouse Mutant Archive, accession number 13836 (full allele name: *Nudcd3*$^{em1(IMPC)Wtsi}$). All data and code with the exception of patient DNA sequencing data, are included within this paper or are deposited in data repositories. Exome sequencing data are only available to vetted researchers (and not to the broader public) via a data transfer agreement (DTA). The human scRNA sequencing data have been deposited to ArrayExpress with accession number E-MTAB-12283. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (*63*) partner repository with the dataset identifier PXD035840. The mouse TCR/BCR sequencing data have been deposited to European Nucleotide Archive with the project identifier PRJEB55545. All other data are available in the main text or the supplementary materials. Tabulated underlying data for all figures can be found in data S4. All raw immunoblot data exported from LI-COR Image Studio software version 5.2.5 as Image Studio Zip File format can be found in data S5. For the purpose of open access, the author has applied a CC BY public copyright license to any author-accepted manuscript version arising from this submission.

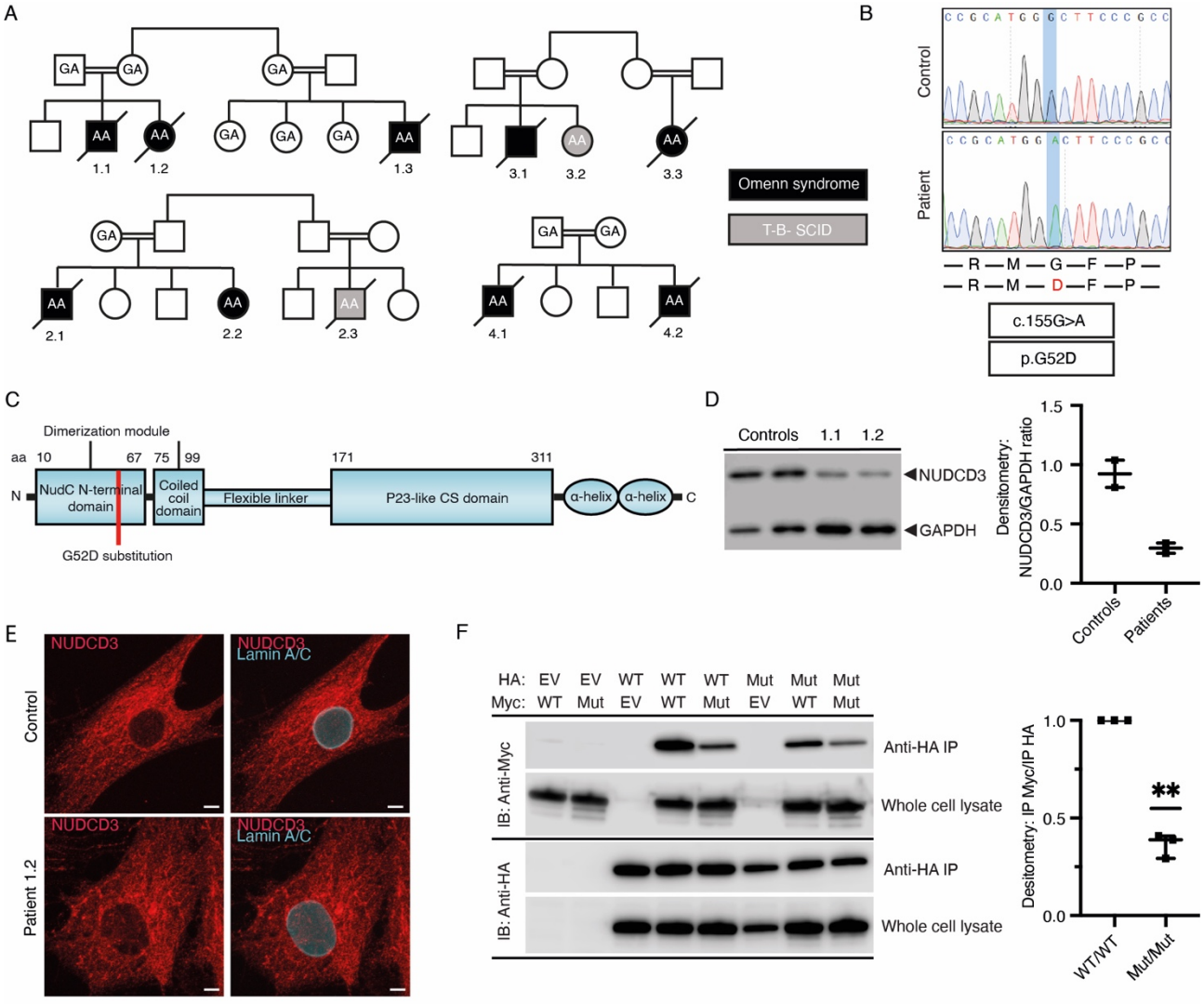
A
GA
AA
1.1
1.2
1.3
2.1
2.2
2.3
3.1
3.2
3.3
4.1
4.2
Omenn syndrome
T-B- SCID
B
Control
Patient
R M G F P
R M D F P
c.155G>A
p.G52D
C
Dimerization module
aa 10 67 75 99 171 311
N
NudC N-terminal domain
Coiled coil domain
Flexible linker
P23-like CS domain
α-helix
α-helix
C
G52D substitution
D
Controls 1.1 1.2
NUDCD3
GAPDH
Densitometry: NUDCD3/GAPDH ratio
1.5
1.0
0.5
0.0
Controls
Patients
E
Control
Patient 1.2
NUDCD3
NUDCD3
Lamin A/C
F
HA: EV EV WT WT WT Mut Mut Mut
Myc: WT Mut EV WT Mut EV WT Mut
IB: Anti-Myc
IB: Anti-HA
Anti-HA IP
Whole cell lysate
Anti-HA IP
Whole cell lysate
Desitometry: IP Myc/IP HA
1.0
0.5
0.0
**
WT/WT
Mut/Mut

Fig. 1. An autosomal recessive variant in *NUDCD3* causes SCID/OS. (**A**) Pedigrees showing patients affected by OS (black) and T-B- SCID (gray) with letters indicating genotype. (**B**) Representative Sanger sequencing chromatogram of patient as compared to reference. (**C**) Schematic representation of NUDCD3 protein domains with the location of the G52D substitution highlighted in red. (**D**) Immunoblot showing reduced protein expression of NUDCD3 in primary patient dermal fibroblasts compared to controls (top) with corresponding densitometry analysis (bottom) (n=2, normalized NUDCD3/GAPDH ratio, each data point represents average from each individual). (**E**) Comparable distribution of NUDCD3 (red) in representative immunofluorescence micrographs of healthy control and patient fibroblasts costained with nuclear membrane marker lamin A/C (cyan) (scale bar: 5 μm). (**F**) Impaired dimerization of $NUDCD3^{G52D}$ by comparison with $NUDCD3^{WT}$ upon coimmunoprecipitation with alternatively tagged $NUDCD3^{WT}$ or $NUDCD3^{G52D}$ in HEK293T cells transfected with plasmids encoding corresponding proteins (EV, empty vector; WT, wild-type; Mut, G52D variant; IP immunoprecipitation; IB, immunoblot) (top) and quantified by densitometry (below) (Myc IP/HA IP, ratio to WT/WT, n=3, one-sample t test. Statistical significance was defined as $^{**}P<0.01$).

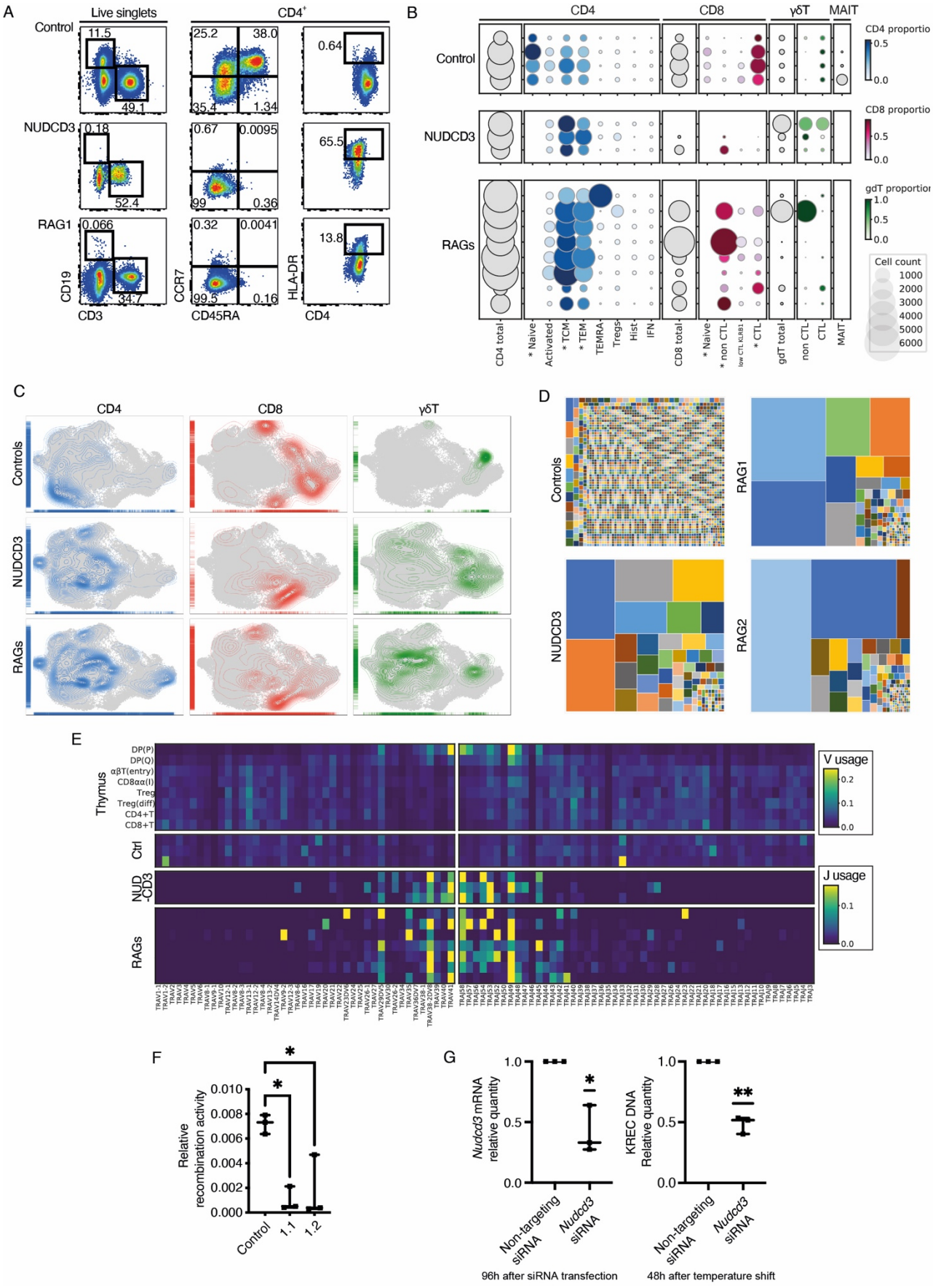
A
Live singlets
CD4+
Control
NUDCD3
RAG1
CD19
CD3
CCR7
CD45RA
HLA-DR
CD4
B
CD4
CD8
γδT
MAIT
Control
NUDCD3
RAGs
CD4 proportio
CD8 proportio
gdT proportio
Cell count
CD4 total
* Naive
Activated
* TCM
* TEM
TEMRA
Tregs
Hist
IFN
CD8 total
* Naive
* non CTL
low CTL KLRB1
* CTL
gdT total
non CTL
CTL
MAIT
C
CD4
CD8
γδT
Controls
NUDCD3
RAGs
D
Controls
RAG1
NUDCD3
RAG2
E
Thymus
DP(P)
DP(Q)
αβT(entry)
CD8αα(I)
Treg
Treg(diff)
CD4+T
CD8+T
Ctrl
NUD-CD3
RAGs
V usage
J usage
F
Relative recombination activity
Control
1.1
1.2
G
Nudcd3 mRNA relative quantity
KREC DNA Relative quantity
Non-targeting siRNA
Nudcd3 siRNA
96h after siRNA transfection
48h after temperature shift

Fig. 2. Impaired V(D)J recombination underlies OS in infants with NUDCD3^{G52D}. (**A**) Representative flow cytometry plots showing a lack of B cells ($CD19^+$) and naïve $CD4^+$ T cells ($CD45RA^+CCR7^+$) and increase of activated $CD4^+$ T cells ($HLA\text{-}DR^+$) in representative NUDCD3 and RAG1 OS patients compared to a healthy control. (**B**) Proportions of T cell subpopulations determined by scRNA-seq with dot size reflecting the number of cells, and the color intensity corresponding to the proportion of all T cells per individual (one individual per row). Asterisks indicate a statistically significant difference in proportions for the comparison between healthy controls and OS patients (NUDCD3 and RAG patients grouped together) with False Discovery Rate (FDR) corrected, Mann–Whitney *U* test. (**C**) UMAP representation of CD4, CD8, and γδ T cells within T cell compartment across healthy individuals, as well as OS patients with RAG or NUDCD3 variants. (**D**) T cell clonotype diversity in representative OS patients and healthy control. Each tree map illustrates all unique clonotypes from one individual, each rectangle is one clonotype and its size is proportional to the number of cells with this clonotype. Colors are chosen randomly to distinguish different clonotypes within a sample and the same color does not represent shared clonotypes. Representative plots for a healthy individual and OS patients with indicated underlying mutations. (**E**) *TRAV* and *TRAJ* gene usage distribution in OS patients, healthy controls and across the stages of human thymus development from (*24*). Each row corresponds to one individual. Genes are ordered according to their genomic position (left 5′ end to right 3′ end). **(F)** NUDCD3^{G52D} patient fibroblasts (1.1, 1.2) support reduced recombination of the substrate plasmid pDVG93 upon cotransfection with *RAG1* and *RAG2* compared to healthy control fibroblasts (n=3) (levels of recombined pDVG93 expressed relative to total levels of transfected pDVG93). **(G)** Reduction in KREC production in mouse pre-B 103/BCL2 cells transfected with siRNA targeting *Nudcd3* prior to a culture temperature shift from 33°C to 39°C. DNA and RNA were harvested for qPCR, evaluating levels of *Nudcd3* mRNA transcripts relative to *Hprt1* (left) and levels of KREC relative to genomic *Alb* (right) (n=3 independent experiments). In (F and G), data are shown as box and whisker plots and comparison between groups was made by ANOVA/Bonferroni correction and one-sample *t* test, respectively. Statistical significance was defined as *P<0.05 and **P<0.01.

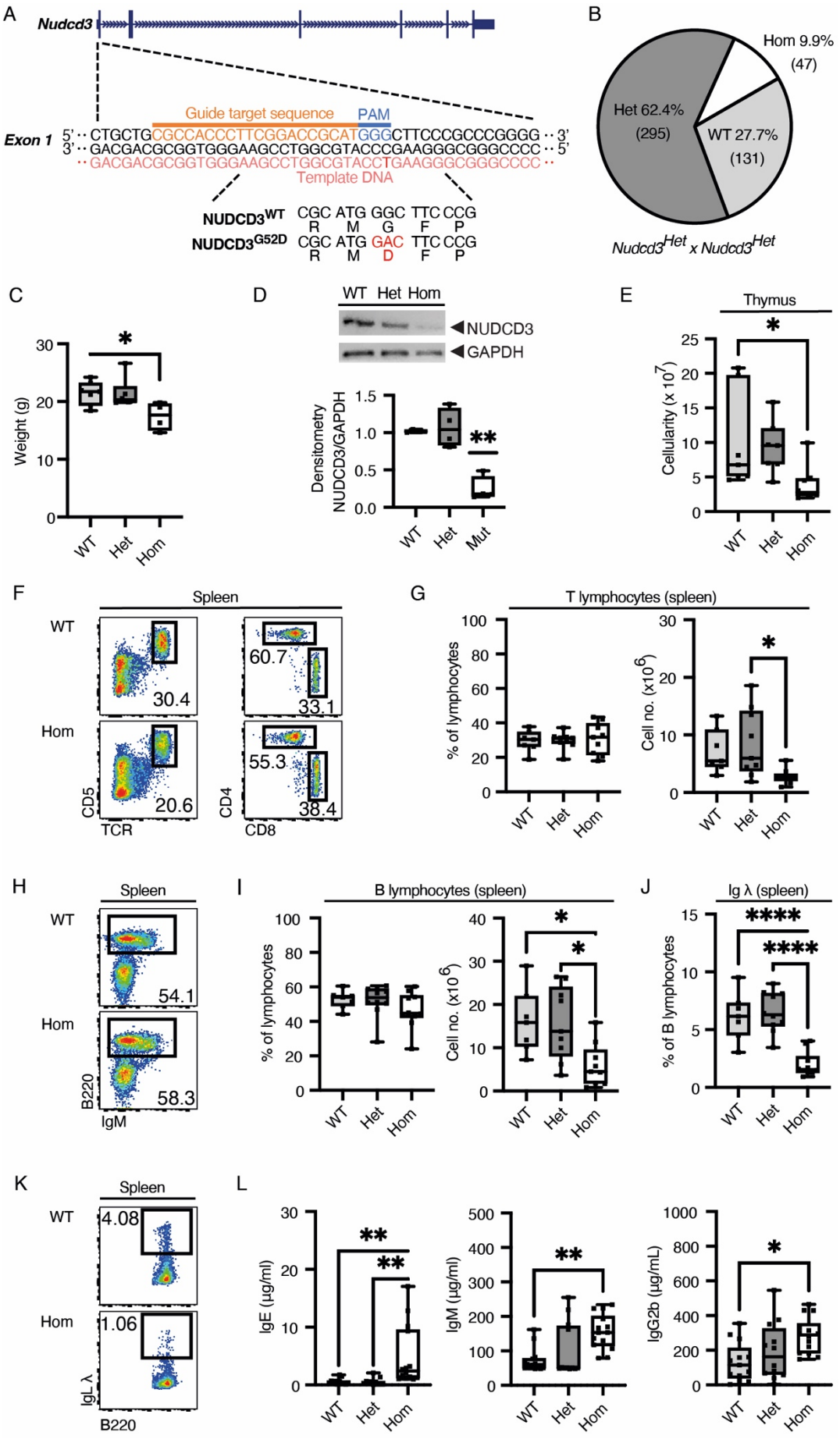
A
Nudcd3
Guide target sequence
PAM
Exon 1
5'·· CTGCTGCGCCACCCTTCGGACCGCATGGGCTTCCCGCCCGGGG ·· 3'
3'·· GACGACGCGGTGGGAAGCCTGGCGTACCCGAAGGGCGGGCCCC ·· 5'
·· GACGACGCGGTGGGAAGCCTGGCGTACCTGAAGGGCGGGCCCC ··
Template DNA
NUDCD3WT CGC ATG GGC TTC CCG
R M G F P
NUDCD3G52D CGC ATG GAC TTC CCG
R M D F P
B
Hom 9.9%
(47)
Het 62.4%
(295)
WT 27.7%
(131)
Nudcd3Het x Nudcd3Het
C
Weight (g)
WT
Het
Hom
D
WT Het Hom
NUDCD3
GAPDH
Densitometry
NUDCD3/GAPDH
WT
Het
Mut
E
Thymus
Cellularity (x 10^7)
WT
Het
Hom
F
Spleen
WT
Hom
30.4
20.6
60.7
33.1
55.3
38.4
CD5
TCR
CD4
CD8
G
T lymphocytes (spleen)
% of lymphocytes
Cell no. (x10^6)
WT
Het
Hom
H
Spleen
WT
Hom
54.1
58.3
B220
IgM
I
B lymphocytes (spleen)
% of lymphocytes
Cell no. (x10^6)
WT
Het
Hom
J
Ig λ (spleen)
% of B lymphocytes
WT
Het
Hom
K
Spleen
WT
Hom
4.08
1.06
IgL λ
B220
L
IgE (µg/ml)
IgM (µg/ml)
IgG2b (µg/mL)
WT
Het
Hom

Fig. 3. Generation of a knock-in mouse model of NUDCD3^{G52D}. (**A**) CRISPR-Cas9 guide target sequence introducing the missense mutation c.155G>A (p.G52D) to mouse *Nudcd3.* (**B**) Reduced Mendelian ratios of mice born to 19 *Nudcd3*^{Het} breeding pairs (*n*=473 pups). (**C**) *Nudcd3*^{Hom} mice (*n*=5) are smaller than *Nudcd3*^{WT} (*n*=8) and *Nudcd3*^{Het} (*n*=7) littermate controls, aged 5-8 weeks. (**D**) Immunoblotting shows reduced expression of NUDCD3 protein in splenocytes from *Nudcd3*^{Hom} (*n*=4) mice compared to littermate controls (*n*=4) using a GAPDH loading control. (**E**) Thymic cellularity is reduced in *Nudcd3*^{Hom} (*n*=8) mice compared to littermate controls (WT *n*=6 and Het *n*=8) aged 5-14 weeks. **(F)** Comparable frequency of splenic T lymphocytes (CD5^{hi}TCR^{hi}) within FSC/SSC lymphocyte gate (left) and CD4 and CD8 T lymphocytes (of T lymphocytes) (right), T lymphocytes quantified in (**G**) between genotypes (WT *n*=7, Het *n*=10, Hom *n*=10). **(H)** Comparable frequency, within FSC/SSC lymphocyte gate, of splenic B lymphocytes (B220^{hi}IgM^{hi}), quantified in (**I**) between genotypes. (**J**) Reduced frequency of lambda light chain expressing B lymphocytes (IgLλ^{hi}) in the spleens of *Nudcd3*^{Hom} mice compared to *WT* and *Het* littermates. (**K**) B cell lambda light chain usage (IgLλ^{hi}). (**L**) Increased IgE (left), IgM (middle), and IgG2 (right) in serum from *Nudcd3*^{Hom} mice compared to *Nudcd3*^{WT} and *Nudcd3*^{Het} littermates (*n*=7). Data, pooled from multiple experiments (see data S3), are shown as box and whisker plots with *Nudcd3*^{WT} (light gray), *Nudcd3*^{Het} (dark gray), and *Nudcd3*^{Hom} (white). Comparisons between groups were made by ANOVA/Bonferroni except in (D), where one-sample *t* test was applied. Statistical significance was defined as **P*<0.05, ***P*<0.01, and *****P*<0.0001.

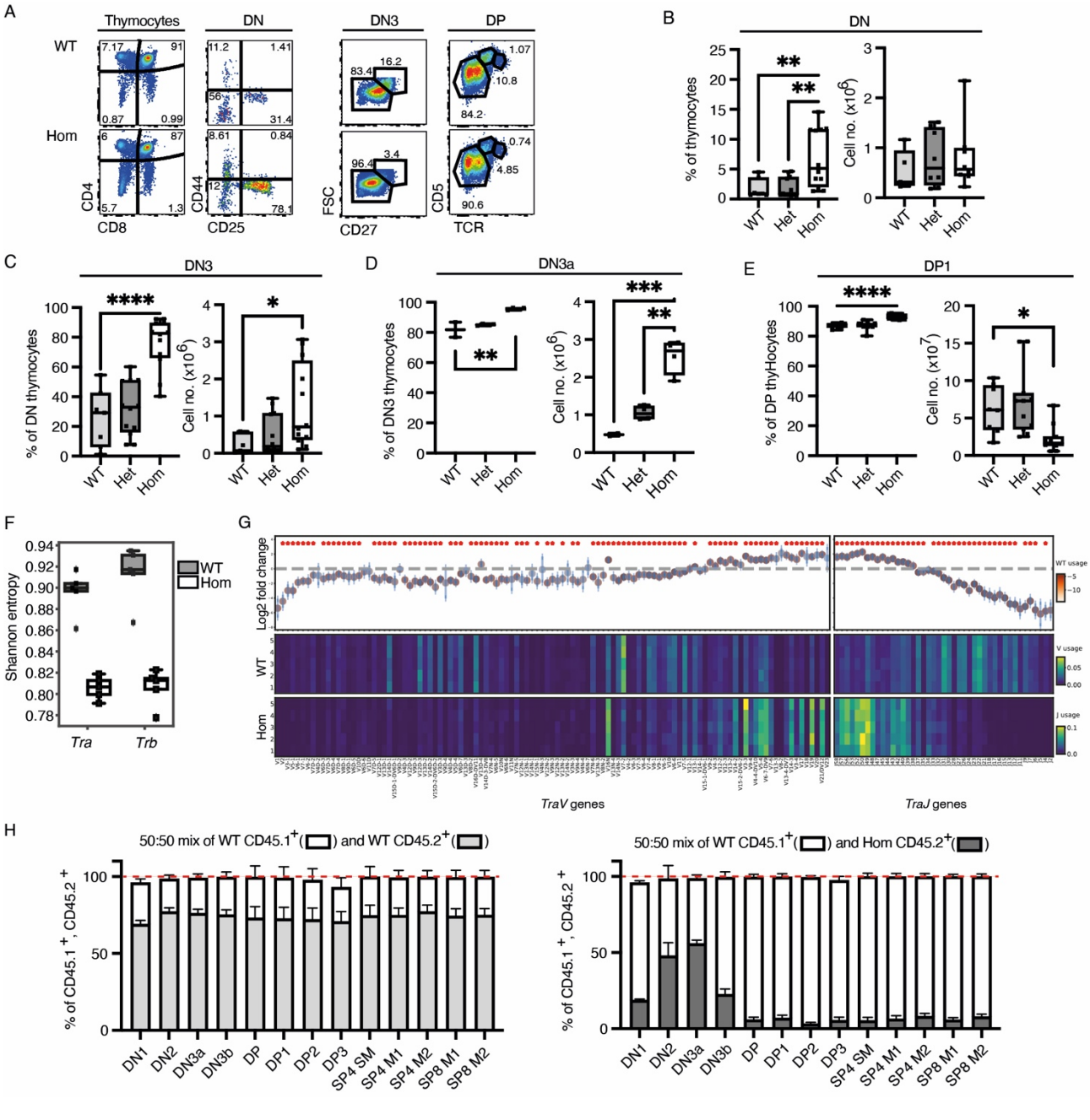
A
Thymocytes
DN
DN3
DP
WT
Hom
CD4
CD8
CD44
CD25
FSC
CD27
CD5
TCR
B
DN
% of thymocytes
Cell no. (x10^6)
C
DN3
% of DN thymocytes
Cell no. (x10^6)
D
DN3a
% of DN3 thymocytes
Cell no. (x10^6)
E
DP1
% of DP thyHocytes
Cell no. (x10^7)
WT
Het
Hom
F
Shannon entropy
Tra
Trb
G
Log2 fold change
WT usage
V usage
J usage
TraV genes
TraJ genes
H
50:50 mix of WT CD45.1+( ) and WT CD45.2+( )
50:50 mix of WT CD45.1+( ) and Hom CD45.2+( )
% of CD45.1 +, CD45.2 +
DN1
DN2
DN3a
DN3b
DP
DP1
DP2
DP3
SP4 SM
SP4 M1
SP4 M2
SP8 M1
SP8 M2

Fig. 4. Impaired V(D)J recombination in *NUDCD3*^{Hom} mice. (**A**) Accumulation of DN ($CD4^{-}CD8^{-}$), DN3 ($CD44^{-}CD25^{+}$), DN3a (CD27lo) and DP1 ($CD5^{lo}TCR^{hi}$) thymocytes in *Nudcd3*^{Hom} mice compared to littermate controls. Increased frequency (left) of DN (**B**), DN3 (**C**), DN3a (**D**) and DP1 (**E**) thymocytes with corresponding enumeration (right) (gating is as per fig. S10A; WT n=7, Het n= 12, Hom n=12 except for D where WT n=2, Het n= 4, Hom n=4). In (B-E), data are shown as box and whisker plots with *Nudcd3*^{WT} (light gray), *Nudcd3*^{Het} (dark gray) and *Nudcd3*^{Hom} (clear). Comparisons between groups were made by ANOVA/Bonferroni and statistical significance was defined as *P<0.05, **P<0.01, and ****P<0.0001. (**F**) Diversity of TCRα and TCRβ (as Shannon entropy) in WT and *Nudcd3*^{Hom} mice (FDR corrected P<0.01, Mann–Whitney U test). (**G**) *TraV* and *TraJ* usage in WT and *Nudcd3*^{Hom} mice. Lower panel organized as in Fig. 2E and the top panel shows $\log_2$-fold change between *Nudcd3*^{Hom} and WT mice. Asterisk indicates a statistically significant difference with FDR corrected P<0.01, Mann–Whitney U test. (**H**) The proportion of thymocyte subsets in lethally irradiated mice reconstituted for 28 days with equal mixtures of *Nudcd3*^{WT} (n=7) or *Nudcd3*^{Hom} (n=6) $CD45.2^{+}$ and WT $CD45.1^{+}$ bone marrow. Bars (mean with standard error of the mean) shown with $CD45.2^{+}$ (filled) and $CD45.1^{+}$ (white).

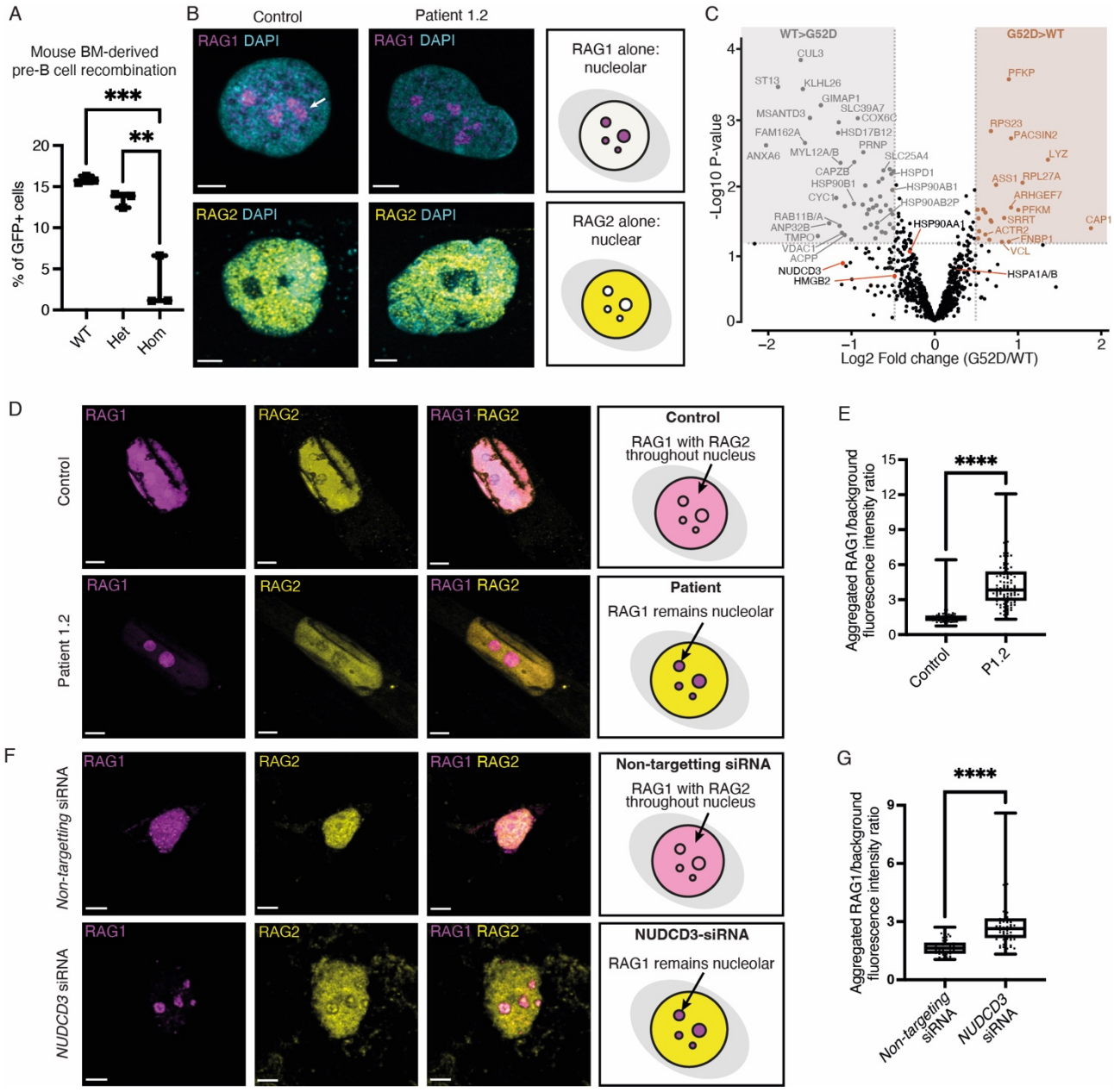
A
Mouse BM-derived pre-B cell recombination
***
**
% of GFP+ cells
20
15
10
5
0
WT
Het
Hom
B
Control
Patient 1.2
RAG1 DAPI
RAG2 DAPI
RAG1 alone: nucleolar
RAG2 alone: nuclear
C
WT>G52D
G52D>WT
-Log10 P-value
Log2 Fold change (G52D/WT)
CUL3
ST13
KLHL26
GIMAP1
SLC39A7
MSANTD3
COX6C
FAM162A
HSD17B12
ANXA6
MYL12A/B
PRNP
SLC25A4
CAPZB
HSPD1
HSP90B1
HSP90AB1
CYC1
HSP90AB2P
RAB11B/A
ANP32B
TMPO
HSP90AA1
VDAC1
ACPP
NUDCD3
HMGB2
HSPA1A/B
PFKP
RPS23
PACSIN2
LYZ
ASS1
RPL27A
ARHGEF7
PFKM
SRRT
ACTR2
FNBP1
VCL
CAP1
D
RAG1
RAG2
RAG1 RAG2
Control
Patient 1.2
Control
RAG1 with RAG2 throughout nucleus
Patient
RAG1 remains nucleolar
E
Aggregated RAG1/background fluorescence intensity ratio
****
Control
P1.2
F
Non-targetting siRNA
NUDCD3 siRNA
Non-targetting siRNA
RAG1 with RAG2 throughout nucleus
NUDCD3-siRNA
RAG1 remains nucleolar
G
Aggregated RAG1/background fluorescence intensity ratio
****
Non-targeting siRNA
NUDCD3 siRNA

Fig. 5. Sequestration of RAG1 in nucleoli of NUDCD3-deficient cells. (**A**) In vitro recombination reporter assay in pre-B cell lines generated from *Nudcd3*^{WT}, *Nudcd3*^{Het} and *Nudcd3*^{Hom} mice. GFP positivity was used as a marker of RAG-dependent recombination (right) (n=3 experiments) and comparisons between groups were made by ANOVA/Bonferroni. (**B**) Normal intranuclear distribution of V5-tagged RAG1 (magenta) and RAG2 (yellow) individually expressed in patient fibroblasts, imaged by confocal microscopy with counterstain nuclear marker DAPI (cyan) (scale bar: 5 μm) and represented in schematic (below). (**C**) Volcano plot comparing differentially interacting proteins between WT and G52D NUDCD3, ascertained by mass spectrometry with significance cut-offs indicated at $\log_2$-fold change ≥0.5 and P<0.05 ($-\text{Log}_{10}$ P>1.3). (**D**) Aberrant nucleolar localization of RAG1 in patient fibroblasts cotransduced with *RAG1* and *RAG2*. Representative confocal microscopy images pseudocolored for RAG1 (magenta) and RAG2 (yellow) (scale bar: 5 μm) with schematic representation of data. (**E**) Quantification of RAG1 signal intensity in nucleoli compared to the rest of the nucleus, shown in box and whisker plots (n=65 control and n=99 *NUDCD3*^{G52D} P1.2 cotransduced fibroblasts analyzed). (**F**) Impaired RAG2-dependent redistribution of RAG1 from nucleoli of *NUDCD3*-knockdown HEK293T cells. Representative confocal microscopy images colored as in (D) (scale bar: 5 μm) with schematic representation of data. (**G**) Quantification of nucleolar signal intensity relative to rest of nucleus as in (E) (n=43 nontargeting control and n=61 *NUDCD3* knockdown cells analyzed). Further images for D and F are in fig. S14, B and D. Comparison between groups was made by ANOVA/Bonferroni in (A) and unpaired t test in (E) and (G), and statistical significance was defined as **P<0.01, ***P<0.001, and ****P<0.0001.