

Published in final edited form as:

*Science*. 2023 September 15; 381(6663): eadh0301. doi:10.1126/science.adh0301.

## Conserved $\gamma\delta$ T cell selection by BTNL proteins limits progression of human inflammatory bowel disease

Robin J Dart<sup>#1,2,3,\*</sup>, Iva Zlatareva<sup>#1,2</sup>, Pierre Vantourout<sup>1,2</sup>, Efstathios Theodoridis<sup>1,2</sup>, Ariella Amar<sup>4</sup>, Shichina Kannambath<sup>5</sup>, Philip East<sup>6</sup>, Timothy Recaladin<sup>7</sup>, John C Mansfield<sup>8,9</sup>, Christopher A Lamb<sup>8,9</sup>, Miles Parkes<sup>10</sup>, Peter M Irving<sup>3</sup>, Natalie J Prescott<sup>4</sup>, Adrian C Hayday<sup>1,2,\*</sup>

<sup>1</sup>Peter Gorer Dept of Immunobiology, King's College London at Guy's Hospital Campus, London, United Kingdom

<sup>2</sup>Immunosurveillance Laboratory, The Francis Crick Institute, London, UK

<sup>3</sup>Department of Gastroenterology, Guy's and St Thomas' Foundation Trust, London, UK

<sup>4</sup>Department of Medical and Molecular Genetics, King's College London, London, UK

<sup>5</sup>NIHR BRC Genomics Centre, King's College London, London, UK

<sup>6</sup>Bioinformatics and Biostatistics, The Francis Crick Institute, London, UK

<sup>7</sup>GammaDelta Therapeutics Ltd., London, UK

---

This work is licensed under a [BY 4.0 International license](#).

\* Corresponding author. robin.dart@kcl.ac.uk (R.J.D.); adrian.hayday@kcl.ac.uk (A.C.H.).

### Author Contributions:

Conceptualization - R.J.D., I.Z., P.V., N.J.P., A.C.H.

Methodology - R.J.D., I.Z., P.V., S.K., P.E. N.J.P.

Software - S.K., P.E.

Validation - R.J.D., I.Z., P.V., E.T., A.A., N.J.P., A.C.H.

Formal analysis - R.J.D., I.Z., S.K., P.E.

Investigation - R.J.D., I.Z., P.V., E.T., A.A., T.R., J.C.M., C.A.L.,

Resources - R.J.D., I.Z., A.A., P.V., J.C.M., C.A.L., M.P., P.M.I., N.J.P. A.C.H.

Data Curation - R.J.D., I.Z., S.K., P.E., J.C.M., C.A.L., M.P., N.J.P., A.C.H.

Writing - original draft preparation R.J.D., I.Z., P.V., A.C.H.

Writing - review and editing - all authors Visualisation - R.J.D., I.Z., P.V., P.E., N.J.P.

Supervision - P.V., P.M.I., N.J.P., A.C.H.

Project administration - R.J.D., I.Z., A.C.H.

Funding Acquisition - R.J.D., N.J.P., A.C.H.

**Competing interests:** R.J.D., I.Z., P.V., N.J.P. and A.C.H. are inventors on patent application WO2019053272A1 submitted by King's College London that covers the use of "compositions and methods for enhancing gamma delta T cells in the gut". R.J.D., I.Z., P.V., and A.C.H. are inventors on patent application WO2019234136A1 submitted by King's College London and Gamma Delta Therapeutics that covers the use of "BTNL3/8 targeting constructs for delivery of payloads to the gastrointestinal system". P.V. and A.C.H. are inventors on patent application WO2021171002A1 and WO2021171003A1 submitted by Gamma Delta Therapeutics entitled "Ex vivo gamma delta T cell populations". R.J.D. has received speaking/consulting fees from Janssen, Takeda, Tillotts. All unrelated to this manuscript., C.A.L. declares research support and/or fees for development and delivery of non-promotional education from Janssen, Dr Falk, Takeda, Abbvie, AstraZeneca, Eli Lilly, Orion, Pfizer, Roche, Sanofi Aventis, Ferring, UCB, Biogen and Genentech Inc. All unrelated to this manuscript. P.M.I. declares research support and/or fees for development and delivery of non-promotional education from AbbVie, Arena, Boehringer-Ingelheim, BMS, Celgene, Celltrion, Falk Pharma, Ferring, Genentech, Galapagos, Gilead, Hospira, MSD, Janssen, Lilly, Pfizer, Pharmacosmos, Prometheus, Roche, Sandoz, Samsung Bioepis, Takeda, Tillotts, Topivert, Sapphire Medical, Shire, VH2, Vifor Pharma, Warner Chilcott. All unrelated to this manuscript. A.C.H. is a scientific consultant for Prokarium and eGenesis and Takeda, and receives research funds from Takeda Pharmaceuticals. All other authors report no conflict of interest.

<sup>8</sup>Translational & Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK

<sup>9</sup>Department of Gastroenterology, Newcastle upon Tyne Hospitals NHS Foundation Trust, Royal Victoria Infirmary, Newcastle upon Tyne, UK

<sup>10</sup>Department of Medicine, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK

# These authors contributed equally to this work.

## Abstract

Murine intraepithelial  $\gamma\delta$  T cells include unique, tissue-protective cells selected by epithelial Butyrophilin-like (BTNL) heteromers. Interrogating whether this biology is conserved in humans, we characterized the colonic  $\gamma\delta$  T cell compartment, identifying a diverse repertoire that includes a phenotypically distinct subset co-expressing T cell receptor V $\gamma$ 4 and the epithelium-binding integrin CD103. This subset was disproportionately diminished and dysregulated in inflammatory bowel disease (IBD), whereas on-treatment CD103<sup>+</sup> $\gamma\delta$  T cell restoration was associated with sustained IBD remission. Moreover, CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cell dysregulation and loss were also displayed by humans with germline BTNL3/BTNL8 hypomorphism, which we identified as a risk factor for penetrating Crohn's disease. Thus, BTNL-dependent selection and/or maintenance of unique, tissue-intrinsic  $\gamma\delta$  T cells appears as an evolutionarily conserved axis limiting progression of a complex, multifactorial, tissue-damaging disease of increasing global incidence.

---

Central and peripheral repertoire selection shape adaptive T cell immunity, purging repertoires of autoreactive T cells while enriching for those suited to recognising diverse pathogens. Thus, polymorphisms in major histocompatibility complex (MHC proteins) that present antigen to T cells can perturb repertoire selection and show causal associations with autoimmune and auto-inflammatory diseases (1).

However, many T cells, particularly  $\gamma\delta$  T cells, are not MHC restricted (2). In mice, for example, intraepithelial  $\gamma\delta$  T cell compartments of limited T cell receptor (TCR) diversity are selected for and/or maintained by TCR-dependent interactions with organ-specific Butyrophilin-like (BTNL) proteins expressed by epithelial cells (3–5). The consequent  $\gamma\delta$  T cell repertoires appear to play major roles in maintaining and restoring tissue homeostasis, with  $\gamma\delta$  T cell deficient mice showing increased susceptibility to hapten-elicited atopic dermatitis and to chemically induced colitis (6–10). Moreover, a recently reported anti-inflammatory role for mouse intestinal  $\gamma\delta$  T cells was manifest in their protection of Paneth cells dysregulated by an IBD-associated polymorphism in *ATG16L1* (11). Hence, there are potentially profound implications of the identification in humans of colonic, TCR $\gamma\delta$ <sup>+</sup> intraepithelial lymphocytes (IEL), a subset of which expresses V $\gamma$ 4 chains that can directly bind heteromers of BTNL3+BTNL8 displayed by colonic epithelium (12, 13).

Nonetheless, to our knowledge, there is to date no evidence that BTNL3+BTNL8 acts as a selection/maintenance factor for human intestinal IEL in a manner akin to murine IEL selection; no clarity as to how BTNL-selected cells might differ from other gut T cells; and no resolution of whether or not they are even the major gut  $\gamma\delta$  T cell subset. Likewise, despite reports that BTNL3 and BTNL8 expression can be substantially reduced

in IBD tissue samples, as well as in celiac disease (14, 15), no evidence exists linking the BTNL3+BTNL8 - V $\gamma$ 4 T cell axis to the regulation of IBD and/or its treatment.

## Results

### Human gut $\gamma\delta$ T cells are diverse, with distinct phenotypes

We first sought to elucidate the composition of the human colonic  $\gamma\delta$  T cell compartment, so that we could more fully understand the nature of dysregulation in IBD. Thus, we established a clinical workflow permitting the analysis of intraepithelial  $\gamma\delta$  T cells from colonic endoscopic biopsies from 173 adults, among whom the majority of immunological data was obtained from a core cohort of: [a] non-IBD controls (“Ctrl”) (n=34); [b] uninflamed gut from patients with confirmed IBD (“IBD”) (n=42); [c] inflamed lesions from patients with IBD (“IBDI”) (n=27); and from 7 individuals, from whom paired inflamed and uninflamed tissue was obtained at the time of endoscopy. Comprehensive donor information can be found in table S1, through which individual donors can be linked to each data-set presented.

First our samples confirmed that the colon is enriched in V $\gamma$ 1<sup>+</sup> cells relative to V $\gamma$ 2<sup>+</sup> cells that dominate the peripheral blood (2), as depicted by an illustrative flow cytometry analysis (fig. S1A) and confirmed by deep sequencing of productively rearranged TCR VDJ gene segment transcripts, among which *TRDV1* (encoding V $\gamma$ 1) accounted for 60%-80% of reads (fig. S1B). By flow cytometry, most gut V $\gamma$ 2<sup>neg</sup> cells reacted to an antibody detecting V $\gamma$ 2, V $\gamma$ 3 or V $\gamma$ 4 chains, and the vast majority were usually V $\gamma$ 1<sup>+</sup> (fig. S1C); hence colonic V $\gamma$ 2<sup>neg</sup>  $\gamma\delta$  T cells that are examined in several experiments below *de facto* comprise primarily V $\gamma$ 2/3/4<sup>+</sup>V $\gamma$ 1<sup>+</sup> T cells.

Given that BTNL-selected V $\gamma$ 7 cells dominate the murine gut  $\gamma\delta$  IEL compartment (3), the reactivity of gut T cells to anti-V $\gamma$ 2/3/4 might have reflected a dominance of cells expressing TCRV $\gamma$ 4 that mediates responsiveness to BTNL3+BTNL8 (13). However, deep sequencing revealed that *TRGV4* reads often accounted for only ~20%-35% of reads, with comparable representation of productively rearranged *TRGV2* and *TRGV5* gene segments (Fig. 1A). Moreover, high donor-to-donor variation existed for the frequencies of *TRGV4* and *TRGV2* gene rearrangements, that were inversely correlated, as reflected by paired samples (fig. S1D). There was also somewhat variable representation of *TRGV9*, the product of which (V $\gamma$ 9) often, albeit not exclusively, pairs with V $\delta$ 2. However, unlike *TRGV2*, the expression levels of this and other *TRGV* genes did not correlate inversely with *TRGV4*. To conclude, whereas the dominant gut  $\gamma\delta$  T cell type expressed V $\delta$ 1, transcripts encoding BTNL-reactive V $\delta$ 4 were commonly not the most highly expressed of the V $\gamma$  chain mRNAs.

We next noted that V $\gamma$ 2<sup>+</sup>, V $\gamma$ 3<sup>+</sup> and V $\gamma$ 4<sup>+</sup> cells that collectively accounted for over 75% of the gut V52<sup>neg</sup>  $\gamma\delta$  T cell compartment (median 85.10% [95% CI 62.5-90.9; n=14]) (fig. S1C) could be segregated according to whether they expressed integrin  $\alpha$ E $\beta$ 7 (a.k.a. CD103), which was true for most cells (median 89.5% [95% CI 83.5-93.8; n=43]) (Fig. 1B), or did not express it. CD103 is typically associated with IEL and tissue-resident memory (T<sub>RM</sub>) CD8<sup>+</sup>  $\alpha$  $\beta$  T cells (16), with its ligand, E-cadherin, expressed by differentiated

epithelial cells (17). To assess the potential significance of CD103 expression by V $\gamma$ 2/3/4 lymphocytes, we compared transcriptomes of paired CD103<sup>+</sup> and CD103<sup>neg</sup>V $\gamma$ 2/3/4<sup>+</sup> T cell preparations from 4 control donors (Fig. 1C; table S2). The most differentially and significantly over-expressed genes in CD103<sup>+</sup>V $\gamma$ 2/3/4<sup>+</sup> cells across all donors included those encoding: *TIGIT*, a negative regulator upregulated by BTNL-mediated selection of mouse TCR $\alpha\beta$ <sup>+</sup> IEL (3); *KIR2DL4*, an activating NK cell receptor (18); *CCR9*, associated with homing to gut epithelium (19); *GZMA*, associated with cytolytic potential (20); *IL18R*, encoding a receptor for which the ligand (IL18) promotes IFN $\gamma$  production (21); *ENTPDI* (CD39), that suppresses inflammation by hydrolysing ATP (22); and *HAVCR2* (TIM3), associated with chronic TCR cell stimulation (23) (Fig. 1C). Additionally, at low expression levels, CD103<sup>+</sup>  $\gamma\delta$  cells differentially expressed *FGF9*, a wound-healing factor produced by murine dermal  $\gamma\delta$  T cells (24). These specific genes combined with others to compose significant overall similarities of CD103<sup>+</sup>V $\gamma$ 2/3/4<sup>+</sup> cells to human intestinal intraepithelial T cells (25) (Fig. 1D, E); to mouse V $\gamma$ 7<sup>+</sup> IEL (3); and to human, CD69<sup>+</sup>CD8<sup>+</sup>lung T<sub>RM</sub> cells (26) (that are primarily intraepithelial) (Fig. 1E), as assessed from publicly available datasets. Conversely, overlap was less with total intestinal TCR $\alpha\beta$ <sup>+</sup>CD8<sup>+</sup>CD103<sup>+</sup> T<sub>RM</sub> cells (27) (fig. S1E), considered further below.

CD103<sup>neg</sup>V $\gamma$ 2/3/4<sup>+</sup> T cells, by contrast, over-expressed genes commonly associated with circulating T cells (Fig. 1C; table S2), specifically: *SELL* (CD62L) and *CCR7* (28); integrins *ITGB2* (CD18) and *ITGAX1* (CD11c) that mediate heterotypic cell-cell interactions (29); *CD5* and *CD6*, encoding proteins associated with cell signalling thresholds (30, 31); *KLF2* and *LEF1*, commonly co-expressed in naïve, memory and/or quiescent CD8<sup>+</sup> T cells (32) with *LEF1* implicated in suppressing IL17 production by mouse  $\gamma\delta$  T cells (33); and *GZMK*, associated with tissue-infiltrating human CD8<sup>+</sup> $\alpha\beta$  T cells of reduced cytolytic but high cytokine-producing potential (34) (Fig. 1C). Indeed, there was an overall significant similarity to tissue-infiltrating human synovial CD8<sup>+</sup> $\alpha\beta$  T cells (34); to human lung T cells lacking CD69 (26); and to human intestinal CD103<sup>neg</sup>CD8<sup>+</sup> T cells (27) (Fig. 1E).

In validating transcriptomic differences by flow cytometry, we noted as expected that neither CD103<sup>+</sup> nor CD103<sup>neg</sup> cells were homogeneous, with approximately two-thirds of CD103<sup>+</sup> cells expressing TIGIT and/or TIM3, while CD103<sup>neg</sup> cells were mosaic for TIGIT and also for the costimulatory receptor ICOS (Fig. 1F; fig. S1F). Despite this, there was a consensus CD103<sup>+</sup> cell phenotype, that was CD69<sup>+</sup>CD101<sup>+</sup>ICOS<sup>+</sup>CD18<sup>neg</sup>, whereas most CD103<sup>neg</sup> cells were CD101<sup>neg</sup>CD5<sup>+</sup>CD18<sup>+</sup>CD31<sup>+</sup>TIM3<sup>neg</sup> (Fig. 1F, G; fig. S1F). In conclusion, CD103 segregated V $\gamma$ 2/3/4<sup>+</sup> T cells into qualitatively distinct subsets (Fig. 1G). Moreover, whereas CD103<sup>neg</sup>V $\gamma$ 2/3/4<sup>+</sup> cells closely resembled human intestinal CD103<sup>neg</sup>TCR $\alpha\beta$ <sup>+</sup>CD8<sup>+</sup> cells (Fig. 1E, H), CD103<sup>+</sup>V $\gamma$ 2/3/4<sup>+</sup> cells appeared unique, being clearly distinguishable from human tissue-resident CD103<sup>+</sup>TCR $\alpha\beta$ <sup>+</sup>CD8<sup>+</sup> cells, e.g., by low expression of CD5 and CD18 (Fig. 1I). Also of note, CD103<sup>+</sup>V $\gamma$ 2/3/4<sup>+</sup> cells often expressed several NK receptors, including NKp44, NKG2C, and NKp46 that were mostly expressed by neither CD103<sup>neg</sup>V $\gamma$ 2/3/4<sup>+</sup> cells nor CD103<sup>+</sup>TCR $\alpha\beta$ <sup>+</sup>CD8<sup>+</sup> cells (Fig. 1F, I; fig. S1F). In contrast, all subsets expressed comparably high levels of NKG2D (fig. S1F), that can mediate innate responsiveness and co-stimulate TCR-mediated adaptive responses (35).

## Selective responsiveness of human gut $\gamma\delta$ T cells

We next asked how gut CD103<sup>+</sup> cells might respond to selecting and activating stimuli, respectively. As previously described (3), cells stained by an anti-V $\gamma$ 2/3/4 antibody showed partial yet significant TCR downregulation (reduced gMFI of CD3) when co-cultured with HEK293T cells transduced to express BTNL3+BTNL8 (293T.L3L8) versus empty-vector transduced HEK293T cells (293T.EV) (fig S2A). There was considerable inter-individual variation (fig. S2A, B), but as expected, BTNL3+BTNL8 responsiveness (assessed by TCR downregulation) correlated with the representation of productive *TRGV4* rearrangements and correlated inversely with *TRGV2* rearrangements (fig. S2C). As evidenced by flow cytometry, TCR levels on CD103<sup>+</sup> cells were higher relative to those on CD103<sup>neg</sup> cells and were disproportionately downregulated by BTNL-exposure (Fig. 2A).

Of note, CD103 staining was unaffected by BTNL3+BTNL8 engagement (Fig. 2B), whereas many cells showing TCR downregulation upregulated the co-stimulatory receptor, TNFRSF9 (4-1BB) (Fig. 2C). Predictably, no response to BTNL3+BTNL8 was shown by V $\gamma$ 2<sup>+</sup> or TCR $\alpha\beta$ <sup>+</sup> T cells in the same cultures (Fig. 2C). *TNFRSF9* transcript upregulation was also noted in an RNA sequencing study that was conducted to better understand the impact of BTNL-mediated selection on human V $\gamma$ 2/3/4<sup>+</sup>V $\gamma$ 2<sup>neg</sup> T cells (Fig. 2D). Specifically, RNA was compared from 4 independent donors' colonic T cells co-cultured with either 293T.L3L8 cells or 293T.EV cells (fig. S2D), revealing that in addition to *TNFRSF9*, there was upregulation of several genes including *XCL1*, *XCL2* and *GZMB* in contrast to which *CCR6* was downregulated (Fig. 2D). Collectively, these gene expression changes strongly phenocopy the response of murine intraepidermal V $\gamma$ 5<sup>+</sup> IEL responding to Skint1, their counterpart BTNL selecting element (36). Thus, CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> gut  $\gamma\delta$  T cells respond strongly and in an evolutionarily conserved fashion to BTNL3+BTNL8 engagement of their TCR, with conspicuous changes in expression of coregulatory receptors, chemokines and other effector potentials.

This notwithstanding, BTNL engagement of V $\gamma$ 2<sup>neg</sup> cells induced neither CD107a, a surrogate marker for induced cytolytic activity (fig. S2E), nor appreciable amounts of cytokines, e.g., TNF $\alpha$ , IFN $\gamma$ , IL2 (fig. S2F, G). Furthermore, gut V $\gamma$ 2<sup>neg</sup> cells were strikingly hyporesponsive to PMA + ionomycin (P+I), a potent activator of signal transduction downstream of the TCR, by contrast to overt TNF $\alpha$ , IFN $\gamma$ , and IL2 production induced in gut TCR $\alpha\beta$ <sup>+</sup> cells in the same cultures (fig. S2F, G). Indeed, V $\gamma$ 2/3/4<sup>+</sup> cells were the least responsive of gut T cells, and by paired analysis, CD103<sup>+</sup>V $\gamma$ 2/3/4<sup>+</sup> cells were less responsive than CD103<sup>neg</sup>V $\gamma$ 2/3/4<sup>+</sup> cells. By contrast, CD103 did not functionally segregate intestinal TCR $\alpha\beta$ <sup>+</sup> cells (Fig. 2G, H). Finally, whereas IL17A production was observed for ~10% of gut TCR $\alpha\beta$ <sup>+</sup>CD4<sup>+</sup> cells, it was only rarely displayed by V $\gamma$ 2/3/4<sup>+</sup> cells (Fig. 2I). In short, the  $\gamma\delta$  compartment of healthy human gut includes CD103<sup>+</sup> V $\gamma$ 4V $\gamma$ 1<sup>+</sup> cells which respond strongly to TCR-mediated selection signals, provided by BTNL3+BTNL8, but which are recalcitrant to potent activating stimuli, specifically P+I.

## Gut V $\gamma$ 4<sup>+</sup> cells are unique

Focussing now on CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> T cells, we sought to elucidate the potential impact of BTNL3+BTNL8 on the signature phenotypes of human gut TCR $\gamma$ S<sup>+</sup> T cells. To facilitate



this, we used phage-display to obtain V $\gamma$ 4-specific antibodies (see Methods), initially screening for reactivity to recombinant V $\gamma$ 4, but not to recombinant V $\gamma$ 2 (the V-region most similar to V $\gamma$ 4). To identify V $\gamma$ 4-specific antibodies unaffected by V $\gamma$  chain usage, and to epitope-map them, we screened for reactivity against human JRT3 cells transduced with various wild-type, mutated, and chimeric TCRs and, based on this, chose to use two antibodies (G4\_18 and G4\_9) (fig. S3A).

In control samples of colonic V $\gamma$ 2<sup>neg</sup> T cells, the mean frequency staining for V $\gamma$ 4<sup>+</sup> cells was 47.7% (Fig. 3A), although as was observed for *TRGV4* transcript reads (see Fig. 1, above), there was high inter-individual variation, with V $\gamma$ 4<sup>+</sup> lymphocytes often failing to compose the majority of  $\gamma$ S cells. The fact that the median value for V $\gamma$ 4<sup>+</sup> cells of 48.3% [95% CI 29.8-60.1; n=23] exceeds the median *TRGV4* reads (24.26% [95% CI 18.8-30; n=13]) most likely reflects the respective denominators, i.e., V $\gamma$ 2<sup>neg</sup> T cells versus whole gut biopsies used for the transcript reads, and greater purity of the flow cytometry samples. Investigating further, it was the CD103<sup>+</sup>V $\gamma$ 2<sup>neg</sup> cell subset that was relatively enriched in V $\gamma$ 4<sup>+</sup> cells (Fig. 3A), consistent with which, a greater majority of V $\gamma$ 4<sup>+</sup> cells were CD103<sup>+</sup>, by comparison with V $\gamma$ 2/3<sup>+</sup>, V $\gamma$ 5/8/9<sup>+</sup> and V $\gamma$ 2<sup>+</sup> cells (Fig. 3B), strongly suggesting a greater propensity of V $\gamma$ 4<sup>+</sup> cells to occupy intraepithelial niches.

Next, to assess whether intestinal V $\gamma$ 4<sup>+</sup> T cells displayed unique traits that might be attributable to BTNL-mediated interactions, we used an anti-V $\gamma$ 4 antibody (G4\_18; above) together with existing antibodies to purify cells for RNASeq from four donors. To maximise potential differences, we first compared V $\gamma$ 4<sup>+</sup> with V $\gamma$ 5/8/9<sup>+</sup> cells, and then examined differentially expressed genes across the broader range of  $\gamma$  $\delta$  T cells. In fact, there were very few significant differences, but attesting to the dataset's validity, *TRGV4* was the most prominently enriched RNA in V $\gamma$ 4<sup>+</sup> cells, while *TRGV5* was significantly under-represented (table S3). Beyond *TRGV4*, the most differentially expressed RNA was *FCER1G*, encoding Fc $\epsilon$ R1 $\gamma$ , a protein that can pair with and/or substitute for CD3 $\zeta$  in TCR-induced signal transduction (37, 38), particularly in mouse intestinal IEL (39), and that can transduce signals from NK receptors (see below). We noted that RNA encoding *API5*, recently implicated in protecting Paneth cells (11), was not significantly overexpressed in V $\gamma$ 4<sup>+</sup> cells relative to other  $\gamma$  $\delta$  T cells.

Whereas the transcriptomic analysis was undertaken on four donors, validation was sought using flow cytometry on many more independent donors. Illustrative plots presented with graphs of multiple donor data validated that Fc $\epsilon$ R1 $\gamma$  was highly enriched in gut CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cells versus cells expressing any other V $\gamma$  regions, except for a single donor in which it was highly expressed by CD103<sup>+</sup>V $\gamma$ 2/3<sup>+</sup> V $\gamma$ 4<sup>neg</sup> cells (Fig. 3C; fig S3B). Strikingly, there was no expression of Fc $\epsilon$ R1 $\gamma$  in any subset of CD103<sup>neg</sup> $\gamma$  $\delta$  T cells (Fig. 3C), nor in any TCRV $\gamma$ -subtype in PBMC, including V $\gamma$ 4<sup>+</sup> cells (fig. S3C). Hence, Fc $\epsilon$ R1 $\gamma$  is in essence a signature phenotype of cells defined by CD103 expression, implying epithelial association; by intestinal or more general extra-lymphoid localisation; and by specific usage of TCRV $\gamma$ 4 that uniquely facilitates BTNL3+BTNL8-responsiveness. Interestingly, this same pattern proved applicable to some NK receptors, e.g., NKp46 and NKG2C (Fig. 3D, E; fig. S3D, E). Conversely, intestinal CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cells showed

significantly lower expression of an adhesion molecule, CD31 (*PCAMI*) (Fig. 3F; fig. S3F), by contrast to significantly higher expression of its ligand, CD38 (fig. S3G).

Likewise, whereas *CD18* and *CD5* were less frequently expressed by CD103<sup>+</sup> versus CD103<sup>neg</sup>  $\gamma\delta$  cells (see Fig 1I, above), they were expressed at disproportionately low levels by CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cells (Fig. 3G, H). Thus, by several criteria, V $\gamma$ 4<sup>+</sup>CD103<sup>+</sup> gut T cells displayed a unique phenotype by comparison to CD103<sup>+</sup> gut T cells expressing other V $\gamma$  chains (Fig. 3I), thereby attesting to a profound impact of interactions with BTNL3+BTNL8. At the same time, gut CD103<sup>neg</sup>V $\gamma$ 4<sup>+</sup> cells were more similar to CD103<sup>neg</sup> cells expressing other V $\gamma$  chains (Fig. 3J), which suggests that either there are modifiers of the impact of BTNL-mediated selection on V $\gamma$ 4<sup>+</sup> cells, or that selection of CD103<sup>neg</sup>V $\gamma$ 4<sup>+</sup> cells was incomplete, possibly because the cells were in transition to the epithelium.

### Disproportionate CD103<sup>+</sup>V $\gamma$ 4 cell loss in inflammatory bowel disease

We next investigated how the gut  $\gamma\delta$  T cell compartment, and within it the unique CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> subset, might be affected in patients with IBD, by sampling inflamed (IBDI) and macroscopically uninfamed regions (IBU). Whether sampling inflamed Crohn's Disease (CDI) or inflamed ulcerative colitis (UCI),  $\gamma\delta$  cell frequencies in IBDI were significantly less than in outwardly healthy gut (Fig. 4A), and this was also observed for uninfamed regions of UC (Fig. 4A). While the representation of V $\gamma$ 2<sup>+</sup> cells among colonic T cells (ordinarily the minor subset [see fig S1A, above]) was also reduced in CDI, UC and UCI (fig. S4A), there were more overt reductions of V $\gamma$ 2<sup>neg</sup> cell frequencies, again being significant in each of UC, UCI, and CDI (Fig. 4B). Furthermore, whereas V $\gamma$ 2<sup>neg</sup> cells remained mostly V $\gamma$ 1<sup>+</sup> (Fig. 4C), the frequencies of V $\gamma$ 4<sup>+</sup> cells in IBDI were reduced, in both CDI and UCI (Fig. 4D; note that red circles denote 8 donors with distinct BTNL genotypes, to be considered below). Moreover, significant disease-associated reductions in the frequency of V $\gamma$ 4<sup>+</sup> cells in the CD103<sup>+</sup>V $\gamma$ 2<sup>neg</sup> compartment were evident when paired inflamed sites were compared to uninfamed sites from the same donors (Fig. 4E; CD sample, open circle; and UC samples, filled circles). This was further corroborated by deep sequencing which showed significant, selective reductions in *TRGV4* relative to increased *TRGV2* within the V $\gamma$ 2/3/4<sup>+</sup> compartment (Fig. 4F; CD samples, open circles; UC samples, filled circles).

In going from controls through IBD to IBDI, the overall V $\gamma$  repertoire and the V $\gamma$ 2, V $\gamma$ 3, and V $\gamma$ 4 sub-repertoires all showed increasing TCR sequence diversity as measured by Shannon entropy, reaching significance in the control versus IBDI groups, even when there was no aggregate change in representation, i.e., *TRGV3* (above) (Fig. 4G). This lack of TCR focussing strongly implied that the V $\gamma$  repertoire shifts that included the relative loss of V $\gamma$ 4<sup>+</sup> cells, were not driven by clonotypic antigens, unlike the clonal dynamics reported for V $\gamma$ 1<sup>+</sup> cells in other settings, e.g., response to cytomegalovirus (40, 41). This nonclonal, innate-like response of total  $\delta\gamma$  and the V $\gamma$ 2, V $\gamma$ 3, and V $\gamma$ 4 sub-repertoires was validated for IBDI by D50, an independent measure of diversity (fig. S4B).

We therefore considered that the repertoire changes might reflect cells being recruited to the gut de novo, consistent with which there was an increase in proportion of V $\gamma$ 2<sup>neg</sup> cells that were CD103<sup>neg</sup> and that co-expressed CD45RA and CD27 which jointly demarcate

poorly differentiated cells. (fig. S4C, D) (42). Indeed, there was more generally a significant reduction from ~90% to ~55% of  $V\gamma 2/3/4^+V\gamma 2^{\text{neg}}$  cells expressing CD103 in CDI, and from ~90% to ~30% in UCI (Fig. 4H), albeit with inter-individual variation, and this was well illustrated for  $V\gamma 4^+V\gamma 2^{\text{neg}}$  cells by comparing paired inflamed and uninflamed regions from the same individuals sampled at the same endoscopy (fig. S4E, F). Additionally, in a subset of donors where the data were available, the residual CD103<sup>+</sup>  $V\gamma 2/3/4^+$  T cells in UCI and CDI showed atypically high frequencies of CD5<sup>+</sup> cells (Fig. 4I) which in the healthy colon is a phenotype more often associated with CD103<sup>neg</sup> $V\gamma 2/3/4^+$  cells as considered above (see Fig. 1F). Hence, the main dysregulation in IBD was a shift away from the signature frequency and phenotype of CD103<sup>+</sup> $V\gamma 4^+$  cells.

Because we had observed that among T cells in the healthy colon the CD103<sup>+</sup> $V\gamma 2/3/4^+V\gamma 2^{\text{neg}}$  subset was the most refractory to strong activation, we considered that the relative increases in CD103<sup>neg</sup> cells in IBDI donors as presented above (see Fig. 2E, F) might translate into precocious cytokine production. Indeed, donors with CDI, UC and UCI showed significantly higher frequencies of  $V\gamma 2/3/4^+$  cells producing TNF $\alpha$  and, for UCI, IFN $\gamma$  in response to P+I (Fig. 4J). There were no significant increases in IL17A-producing  $V\gamma 2/3/4^+$  cells (fig. S4G). Moreover, of several gut T cell subsets sampled, only  $V\gamma 2/3/4^+$  cells showed significantly increased frequencies of cytokine producing cells in IBDI, although it is important to note that in each case, those frequencies were substantially less than the baseline frequencies of cytokine-producing  $V\gamma 2^+$  and  $\alpha\beta$  T cells in both the healthy and diseased colon (Fig. 4J; fig. S4G, H, I). Hence, increased inflammatory cytokine production by  $V\gamma 2/3/4^+$  cells may be more profound in its reflection of the cells' altered phenotype rather than any substantial contribution to local cytokine levels. Indeed, paralleling their increased effector responses,  $V\gamma 2/3/4^+$  cells showed significantly decreased responses to 293T.L3L8 (Fig. 4K), consistent with the reductions in  $V\gamma 4^+$  cells. We concluded that in many individuals with IBD, particularly in inflamed lesions, there was a disproportionate diminution and alteration of the unique colonic CD103<sup>+</sup> $V\gamma 4^+$  cell compartment, which is in turn associated with increased responses of the residual cell pool to an activating stimulus and reduced responses to a selecting stimulus. Therefore, we focussed next on the prospect that a major deleterious factor in IBD was the reduction in signature CD103<sup>+</sup> $V\gamma 4^+$  cells. This would echo the severe exacerbation of murine colitis in the absence of  $\gamma\delta$  T cells that ordinarily promote tissue repair and homeostasis (8).

### Inflammatory bowel disease-associated cytokines dysregulate CD103<sup>+</sup> $V\gamma 4^+$ T cells

Initially we asked what factors might induce the depletion and/or altered phenotypes of intestinal  $V\gamma 4^+CD103^+$  cells observed in IBDI. Whilst those altered phenotypes were suggestive of newly infiltrating CD103<sup>neg</sup>CD45RA<sup>+</sup>CD27<sup>+</sup>  $\gamma\delta$  cells (see fig. S4C, D), we also hypothesised that the IBD milieu might contribute. It is challenging to quantitate gene expression in diseased versus control whole gut preparations because cell compositions and phenotypes can change dramatically. Nonetheless, the data revealed significantly increased transcripts encoding the cytokines IL12, IL18, IL23, and TNF $\alpha$  in IBDI (fig. S5A; note that although CD samples (open circles) and UC samples (closed circles) were pooled as IBD or IBDI, their data points completely overlapped). These cytokines are germane to IBD because of their proinflammatory nature and because antibodies neutralising TNF $\alpha$ ,



IL12/IL23, and IL23 are all efficacious in IBD (43–45). Additionally, V $\gamma$ 4 cells express RNAs encoding receptors for IL12, IL18, and TNF $\alpha$  (fig. S5B). Thus, we cultured primary gut T cells in IL2+IL15 for 7 days, additionally supplementing with IL12+IL18, or TNF $\alpha$ , or IL1 $\beta$ +IL23 (which might induce changes in V $\gamma$ 4 cells indirectly *via* cells expressing the cognate receptors).

Strikingly, IL12+IL18 selectively shifted V $\gamma$ 4<sup>+</sup> cells away from the signature healthy gut phenotype, showing instead significant reductions in the frequencies of cells expressing CD103 and TIGIT (Fig. 5A, B) and significantly reduced expression levels of NKp46 (Fig. 5C). Complementing these changes, there was an increased frequency of cells expressing CD18 (Fig. 5D) that is ordinarily associated with CD103<sup>neg</sup> cells (see Fig. 1F). Consistent with this overall phenotypic change, V $\gamma$ 2/3/4<sup>+</sup> cells showed less TCR downregulation in response to BTNL3+BTNL8, although it is also the case that their TCR expression was lower at the outset (Fig. 5E). Conversely, no changes were observed in response to IL1 $\beta$ +IL23 (Fig. 5A-E). Interestingly, TNF $\alpha$  exposure did not reduce the frequencies of cells expressing CD103 or NKp46 but it significantly reduced CD103 and NKp46 cell surface expression levels, and it also increased both the frequencies of cells expressing CD18 and its surface expression levels (Fig. 5F, G, H), while having no clear impact on TIGIT (Fig. 5I). Hence, one can conclude that the signature CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> gut T cell phenotype might not only be displaced by infiltrating cells (see fig. S4C, D), but that it is vulnerable to the selective actions of cytokines characteristic of the IBD milieu. Therefore, the prospect exists that the signature CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> gut T cell phenotype might also be disrupted by other factors of the IBD milieu that we have not tested, including reported reductions in BTNL3+BTNL8 expression (15, 46).

### BTNL3+BTNL8 are selection/maintenance factors for intestinal V $\gamma$ 4<sup>+</sup> T cells

There are several challenges in investigating the potential impacts on CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> gut T cells of reported reductions in BTNL3+BTNL8 expression in whole IBD gut. Specifically, as was considered above in the context of cytokine expression, there are changes in cell composition and activation states, added to which there has been no measure of whether surface BTNL3+BTNL8 expression by colonocytes adjacent to V $\gamma$ 4<sup>+</sup> T cells is *de facto* reduced. Therefore, to unequivocally assess such a situation, we turned to a well-established copy number variation (CNV) polymorphism in which deletion causes an in-frame fusion of the BTNL8 N-terminus to the BTNL3 C-terminus (Fig. 6A) (47). Homozygosity for this allele occurs at ~9% in Europeans. The allele could be identified by a standard PCR assay (fig. S6A) but given that CNV $\gamma$  can be challenging to detect by routine genetic analyses, we also identified a surrogate SNP rs72494581 to facilitate reliable identification of the CNV in DNA microarrays (fig. S6B, C; Methods). From this analysis, CNV homozygotes derived their annotation, “CC”, compared to “TT”, denoting homozygosity for the high frequency allele (full length *BTNL8* and *BTNL3*) and “CT”, denoting heterozygotes (fig. S6C).

When the “C” allele was cloned, FLAG-tagged, and expressed in HEK293T cells, its protein product (BTNL8\*3) showed poor cell surface expression (fig. S6D, bottom left). We previously reported BTNL8 co-expression can rescue poor surface BTNL3 expression (3, 48), but by contrast, BTNL8\*3 was ineffective at this (fig. S6D, 2<sup>nd</sup> row). Moreover,

whereas BTNL3 can enhance surface BTNL8 expression (fig. S6D, 3<sup>rd</sup> row, left and centre), it was less effective at rescuing BTNL8\*3 expression (fig. S6D, 4<sup>th</sup> row, left and centre). Added to this inefficient cell surface expression, BTNL8\*3 lacks the BTNL3 N-terminal V-region that directly engages human V $\gamma$ 4 (13), and reflecting this, BTNL8\*3+BTNL8 could not drive surface V $\gamma$ 4V $\gamma$ 1 TCR downregulation (Fig. 6B). Thus, in relation to  $\gamma\delta$  T cell regulation, BTNL8\*3 is encoded by a severely hypomorphic allele.

Because BTNL1+BTNL6 drive selection of the mouse intestinal V $\gamma$ 7<sup>+</sup>  $\gamma\delta$  cell compartment (3), BTNL1-deficient mice display severe, albeit incomplete losses of V $\gamma$ 7<sup>+</sup> IEL, while in BTNL6-deficient mice, V $\gamma$ 7<sup>+</sup> IEL are reduced by ~50% (4). Moreover, once the compartment forms, neither BTNL1 nor BTNL6 is necessary in the short-term to maintain V $\gamma$ 7<sup>+</sup> IEL, but they do maintain several parameters of the cells' phenotype (4). In this regard, persons homozygous for BTNL8\*3 offered an opportunity to ask whether analogous mechanisms regulated the human gut  $\gamma\delta$  T cell compartment. Indeed, frequencies of V $\gamma$ 4<sup>+</sup> gut T cells were significantly, albeit incompletely, reduced in healthy individuals with a CC genotype (Fig. 6C, D), that were also those denoted by red circles among the control cohort depicted in Fig 4D (above: Note, however, that even when removing *a posteriori* those individuals homozygous for CC from the overall comparison, there was still a significant reduction in V $\gamma$ 4<sup>+</sup> cell frequencies in IBDI (fig S6E)). As well as constituting a smaller fraction of the V $\gamma$ 2<sup>neg</sup> compartment in persons homozygous for the CC genotype, the residual V $\gamma$ 4<sup>+</sup> T cells were also phenotypically altered. Thus, whereas for donors with CT and TT genotypes, most V $\gamma$ 4<sup>+</sup> T cells were CD103<sup>+</sup>, this association of V $\gamma$ 4 and CD103 was lost in homozygotes for BTNL8\*3, who displayed comparable frequencies of CD103<sup>+</sup> and CD103<sup>neg</sup> V $\gamma$ 4 cells (Fig. 6E).

The severed association of V $\gamma$ 4 and CD103 phenocopied the shift from CD103<sup>+</sup> to CD103<sup>neg</sup> cells among V $\gamma$ 4<sup>+</sup>V $\gamma$ 2<sup>neg</sup> cells in IBDI (shown above in fig. S4E). Indeed, we identified four IBD patients with the CC genotype who also showed reduced frequencies of V $\gamma$ 4<sup>+</sup> T cells (also denoted in red in Fig 4D, above) and the shift from CD103<sup>+</sup> to CD103<sup>neg</sup> V $\gamma$ 4 cells (Fig. 6E). Moreover, CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> lymphocytes from CC donors further phenocopied IBDI in that they showed significantly higher CD5 expression (Fig. 6F). However, because most IBD patients we studied carried either the TT or CT phenotype, we could conclude that there were two distinct settings showing comparable diminution and phenotypic alteration of colonic CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> T cells: IBDI, irrespective of BTNL genotype, and BTNL loss-of-function, irrespective of disease status. Importantly, the reduction in frequency and altered phenotypes of CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cells in persons homozygous for BTNL8\*3 jointly establish that humans, like mice, deploy BTNL proteins to select and/or maintain cognate gut  $\gamma\delta$  T cells with a unique phenotype. That neither BTNL-deficient mice nor humans homozygous for BTNL8\*3 showed complete loss of V $\gamma$ 7<sup>+</sup> and V $\gamma$ 4<sup>+</sup> gut T cells, respectively, is discussed below.

### BTNL3+BTNL8 hypomorphism is a disease modifier

A genetic basis for V $\gamma$ 4 cell dysregulation in outwardly healthy individuals offered the opportunity to investigate whether this might predispose to disease. In a pilot snapshot assessment in our clinical practise, we noted that among 64 patients with CD, the frequency

of CC exceeded that among 115 control donors, and that the CC frequency was significantly greater among 52 patients with UC (Fig. 6G). We therefore investigated the frequency of CC homozygotes in 12,767 individuals with IBD comprising a test cohort (n=4517) and a replication cohort (n=8250) provided by the UK Bio-resource, against population-based controls from the 1958 British Birth Cohort (n=6767) (49). Notwithstanding a trend in UC, the CC genotype showed no significant association with disease incidence (Fig. 6H). By contrast, when disease phenotype was assessed in 6308 patients with CD (Test 1817; Replication 4491), there was a significant association of CC frequency with severe “B3” disease which penetrates through the bowel wall forming intra-abdominal abscesses or fistulae (11.9%, 146/1225), compared with inflammatory (B1) or stricturing (B2) disease which remains confined within the intestinal lumen (8.7%, 455/5083) (Fig. 6I) (50). We conclude that for those with CD, BTNL3 loss-of-function and/or consequent dysregulation of gut  $V\gamma 4^+CD103^+$  cells predispose toward penetrating disease and is therefore a disease modifying factor. The lack of a similar severity classification based on objective disease behavioural parameters precluded a similar analysis of UC.

In identifying the association with penetrating CD, we considered several potentially confounding issues. First, the test-set was an older cohort from tertiary referral centres, London and Newcastle, collected between 1999 and 2017 wherein B3 disease is represented at 24%, whilst the replication-set is the NIHR UK IBD bioresource cohort started in 2016, wherein only 16% are classified as B3, possibly reflecting improving therapeutic approaches, including anti-cytokine antibodies, more aggressive top-down strategies, early surgery, and standardisation of care through guideline development. Moreover, to mitigate the potential effect of disease duration on findings, we increased the stringency of our study by only including patients with 5-year follow up in the B1 cohort, although clearly this could exclude some patients who might later have transitioned to B3 disease (51). These considerations notwithstanding, our genetic analysis was still able to identify a significant association of BTNL3+BTNL8 /  $V\gamma 4^+$  T cell deficiency with penetrating CD.

### **Inflammatory bowel disease treatment response associates with $\gamma\delta$ T cell reversion**

For the concluding part of our study, we investigated potential associations of gut  $\gamma\delta$  cells with responses to treatment. Unlike the characteristic skip lesions of CD, UC is characterised by confluent inflammation: thus, in persons with extensive colitis, one can be confident that samples of uninflamed colonic mucosa that lie toward the rectal side of the extent of disease will have previously been inflamed. Thus, we examined such areas of macroscopically healed mucosa from regions known to be previously affected and found that some such regions displayed recovery of the healthy gut phenotype by the criteria of the frequency of  $V\gamma 2/3/4^+$  cells among  $V\gamma 2^{neg}$  cells; the proportion of those  $V\gamma 2/3/4^+$  cells expressing CD103; and the TCR downregulation response to BTNL3+BTNL8 (Fig. 6J). Thus, in seeming contrast to reports of gut  $\gamma\delta$  T cell dysregulation in celiac disease (14), the nonclonal switch in the signature  $\gamma\delta$  T cell compartment observed in inflamed IBD lesions seemed to be reversible, evoking the wound healing process to which  $\gamma\delta$  T cells can contribute as reported in murine colitis (8-10, 52).

The capacity of histological assessments of healing to define IBD remission, particularly in Crohn's disease, has remained unclear (53). We therefore examined whether the phenotypic renormalisation of the colonic  $\gamma\delta$  T cell compartment might be associated with sustained remission. To this end, we assessed the status of 19 patients who at time of sampling were in complete mucosal and clinical remission on treatment (table S4). The patients were classified as " $V\gamma 2/3/4^+CD103^{high}$ " if the percentage of their  $V\gamma 2/3/4^+$  cells expressing CD103 scored within the 95% confidence intervals of the median value of control CD103 expression by  $V\gamma 2/3/4^+$  cells (i.e.,  $>83.5\%$  [see Fig 1B, above]), and contrasted them with  $V\gamma 2/3/4^+CD103^{low}$  individuals who fell below that range. Of 19 patients, 7 (3 CD; 4 UC) scored as  $V\gamma 2/3/4^+CD103^{high}$  whereas 12 patients (8 CD; 4 UC) were  $V\gamma 2/3/4^+CD103^{low}$ . When their disease progression was compared over the next five years, either until relapse or the end of the study, only one relapse was observed in the  $V\gamma 2/3/4^+CD103^{high}$  group, whereas 9/12  $V\gamma 2/3/4^+CD103^{low}$  patients relapsed (hazard ratio = 6.25 (1.68-23.32  $p=0.035$ )) (Fig. 6K). Conversely, 18 of the 19 patients collectively showed no reliable association of remission / relapse with the phenotype of colonic  $CD103^+CD8^+ \alpha\beta$  T cells (for one of the 19 patients,  $CD8^+$  T cell data could not be recovered for technical reasons) (Fig. 6L).

## Discussion

By investigating the human colonic T cell compartment, we have combined human genetics with tissue immunological studies to establish that the BTNL-dependent selection and phenotypic maintenance of a subset of gut-intrinsic  $\gamma\delta$  T cells is evolutionarily conserved. Although the selected  $CD103^+V\gamma 4^+$  cells often do not dominate the  $\gamma\delta$  compartment, which is in contrast to mice, the cells' phenotype is unique. Moreover, persons homozygous for hypomorphic  $BTNL8^*3$  displayed significant reductions in  $V\gamma 4^+$  cell frequencies and the residual cells were phenotypically altered, both of which phenotypes closely phenocopy the situation in mice with deficiencies in  $BTNL1$  or  $BTNL6$  (3, 4). In the settings of  $BTNL$  deficiency, the incomplete loss of  $V\gamma 4^+$  cells in humans and of  $V\gamma 7^+$  cells in mice is likely attributable to the potential of other interactions, e.g., clonotypic antigen reactivities, to promote the amplification and survival of some human  $V\gamma 4^+$  and mouse  $V\gamma 7^+$  T cells, respectively.

The combined loss and phenotypic perturbation of  $CD103^+V\gamma 4^+$  cells was also the single most overt  $\gamma\delta$  cell dysregulation in inflamed IBD. Whereas this dysregulation may in part be a consequence of pathology (i.e., it is a disease biomarker), causative contributions could be surmised from the association of CD severity with homozygosity for  $BTNL3/8$  hypomorphism. Because UC currently does not benefit from the same objectively defined disease behavioural phenotypes, we could not subject it to the same analysis. Nonetheless, throughout our study patients with UC and CD showed conspicuously similar  $\gamma\delta$  cell dysregulation, and patients with either disease who displayed a renormalised  $CD103^+\gamma\delta$  cell compartment showed significantly higher post-treatment remission rates than those with less obvious renormalisation, when assessed five years later. Limitations of this part of our study include the cohort size and the lack of paired longitudinal samples. Thus, we hope that a larger study will now be undertaken, since validation of our findings would pertain to two key points.

First, that routine mucosal CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cell status assessment may offer a practical prognostic indicator. Whereas histological remission in UC is currently offering improved indicators of disease outcome, 50% of patients in the recent ERICA study who showed histological remission progressed to relapse, highlighting their ongoing, undetected dysregulation (54). Thus, a focus on local  $\gamma\delta$  T cell dynamics may offer disease-stage specific molecular signatures akin to those proving beneficial in managing bowel cancers (55).

The second key point is the probability that CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cell renormalisation is likely to be of physiologic benefit. While the nature of that benefit remains to be elucidated, the cells' wound-healing and anti-inflammatory potentials seem evident in their expression of several potentially relevant genes including *API5* (11), and *FGF9* (24). A capacity of CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cells to promote epithelial repair would be another conserved feature, given that the exacerbated pathology in  $\gamma\delta$ -deficient mice experiencing chemically induced colitis is ascribed to defective wound healing (9). From this perspective, the lack of significant benefit versus placebo observed during Phase 3 UC maintenance trials of etrolizumab, that targets CD103, might in part reflect its antagonising homeostatic CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cells (56, 57). Irrespective of the basis of their wound-healing potential, our study has shown that BTNL-selected CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cells are largely unique in terms of their sensory machinery permitting them to patrol the epithelium (58). That machinery is composed of the TCR; Fc $\gamma$ RI $\gamma$  that can mediate signalling from the TCR and from NK receptors; and several NK receptors, particularly NKp46 that can recognise cells undergoing endoplasmic reticulum (ER) stress which is a feature of the colonic epithelium in IBD (59).

It is formally the case that BTNL3/8 loss-of-function might influence IBD progression via a mechanism independent of CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cell regulation. Nonetheless, the only clearly established biochemical properties of BTNL3+8 heteromers, and of their mouse counterparts, is to engage intestinal  $\gamma\delta$  TCR V $\gamma$  chains. Interestingly, direct evidence germane to the importance of CD103<sup>+</sup>V $\gamma$ 4<sup>+</sup> cell regulation may in the near future be provided by analyses of IBD incidence and progression in individuals carrying recently described *TRGV4* alleles encoding chains that show reduced responsiveness to BTNL3+BTNL8 (60).

In this regard, a further two points merit clarification. First, because V $\gamma$ 4<sup>+</sup> cells appear highly sensitive to BTNL3+BTNL8 engagement, the reduced levels of BTNL3+BTNL8 expression reported for IBD (15, 46) are not necessarily causal to V $\gamma$ 4<sup>+</sup> cell dysregulation, by contrast to either BTNL loss-of-function in donors homozygous for BTNL8\*3 or the IBD-associated cytokine milieu. Indeed, upstream disease-associated factors contributing to V $\gamma$ 4<sup>+</sup> cell dysregulation would seem consistent with the cells being more germane to disease progression than to disease incidence. Second, it was reported that whereas decreased BTNL3 and ablated BTNL8 expression in celiac disease could renormalise, this did not correlate with renormalisation of either intestinal V $\gamma$ 1<sup>+</sup> cells or *TRGV4* reads (14), which is also consistent with other factors contributing to the cells' dysregulation. That we have observed renormalisation in a cohort of patients with treated IBD may simply reflect disease-specific and treatment-specific factors distinguishing IBD from celiac disease, including but not limited to the fact that celiac disease is largely associated with



increases in  $\gamma\delta$  T cell frequencies whereas we show IBD to be associated with  $\gamma\delta$  T cell depletions.

Finally, we note that BTNL is not a hit in GWAS for disease susceptibility in either publicly available or our own datasets. BTNL variants have also not previously shown association with disease progression but neither the imputation of our variant of interest nor clinical classifications used in previous international meta-analyses can be assumed to be sufficiently reliable to assess this. By comparison to ~300 loci associated with IBD susceptibility, only few polymorphisms disposing to IBD prognosis have been identified; specifically, *FOXO3*, *XACT*, *IGFBP1*, and *HLA*, which were identified via GWAS of patients with indolent and clinically more severe CD (61), rather than stricturing versus penetrating disease behaviours used here. While the mechanistic impact of the few progression-associated genes has yet to be fully elucidated, such genes and pathways may offer useful prognostic indicators by comparison to the arguably disappointing performance of genes disposing to susceptibility (62, 63). Hence, the identification of the BTNL- $\gamma\delta$  axis as a disease modifier may offer new insights into understanding and tracking IBD progression. Furthermore, improved tissue repair, by contrast to anti-inflammatory agents, might achieve long-term, drug-free disease remission, which is of particular importance given that IBD is common and is an increasing public health issue in more recently industrialised countries where prevalence continues to rise (64, 65).

## Materials and Methods

### Human Samples

**Human colonic samples**—Up to 12 endoscopic biopsies were obtained from the colon of adult donors undergoing diagnostic colonoscopy after informed consent and in compliance with local ethical approval (REC number 07/H0803/237 or 16/LO/0642). The location and macroscopic appearances of the biopsy site were carefully noted.

Non-IBD controls (Ctrl) were defined as persons undergoing colonoscopy for non-IBD indications where no inflammatory intestinal disease was found and when either none or minimal distant pathology, e.g., minor diverticulosis or sporadic adenomas, were found. In patients who underwent colonoscopy to investigate chronic diarrhoea, histology was taken to exclude microscopic or collagenous colitis. Patients with coeliac disease or other significant intestinal co-morbidity were excluded.

Donors with inflammatory bowel disease (IBD) were identified as either patients with a known history, or symptoms suggestive of IBD prior to endoscopy and prospectively consented. Sampling was performed either in macroscopically uninfamed regions (IBD, Crohn's disease uninfamed (CD) or ulcerative colitis uninfamed (UC)) or macroscopically infamed regions (IBDI, CDI or UCI). Demographic details are provided (table S1).

Remission was defined according to strict criteria. Endoscopic remission was defined using validated endoscopic scores: Ulcerative Colitis Endoscopic Index of Severity  $\leq 1$  and Simple Endoscopic Score for Crohn's Disease  $< 3$  (53) or Rutgeert's score i0-1 (66). Patients were also in clinical remission, with no evidence of ongoing biochemical or

radiological inflammation. Relapse was defined as symptomatic, endoscopic, radiological, or biochemical relapse leading to treatment escalation, surgery or admission.

Samples obtained at endoscopy were placed directly into Hanks balanced salt solution (HBSS). Up to 4h post collection they were placed in cell culture or frozen in 90% FCS and 10% DMSO.

**Blood samples**—Blood samples were obtained from cone blood from the NHS Blood and Transplant service (London, UK). Peripheral blood mononuclear cells (PBMC) were isolated using the Ficoll-Paque method and frozen in 10% DMSO 50% FCS RPMI.

This study was conducted adhering to the principles of the Declaration of Helsinki.

### DNA and RNA isolation from whole tissue

A single colonic biopsy was placed in RNAlater (ThermoFisher) at the time of acquisition. AllPrep DNA/RNA Mini kit (Qiagen) was used to extract mRNA and genomic DNA from donor tissue following manufacturer's instructions. Qiagen TissueLyser II was used to homogenise tissue samples.

### TCR deep sequencing

Total RNA extracted from colonic biopsies was sent for  $\gamma\delta$  TCR chain (*TRDV/TRGV*) deep-sequencing using the IlluminaMiSeq platform with short-read 100/150 PER primers (iRepertoire, USA). Shannon's entropy and D50 were calculated as previously described (67).

### qPCR

cDNA was synthesized using Superscript-II (Invitrogen) according to manufacturer's instructions. Gene expression was analysed using the PowerUp SYBR Green Master Mix (Invitrogen) following manufacturer's instructions. Reactions were run on ViiA7 Real-time PCR machine (Applied Biosystems). Target gene expression was normalised to *RSP9* expression using the  $2^{-\Delta CT}$  method. *RPS9* primer sequences were from Bamias et al (68). *TNF* primer sequences were from Rajput et al (69).

L3qPCR\_Ex4F 5' - TGTTGTCATGGGGATGATAATTG-3'

L3qPCR\_Ex8R 5'- CATACCACCCTACATTTTGTCCC-3'

L8qPCR\_Ex4F 5'- GGCATTGTTGGACTGAAGATTTTC-3'

L8qPCR\_Ex8R 5'- CGCCACCTTTTATTGTGTCTC-3'

RSP9\_F 5' - TCCTTTTCCACTTCGTCTGAG-3'

RSP9\_R 5' - AAGCTGATCGGCGAGTATG-3'

TNF\_F 5' -ATGAGCACTGAAAGCATGATCC-3'

TNF\_R 5' -GAGGGCTGATTAGAGAGAGGTC-3'

Other primers:

Quantitect IL-12A primers (Qiagen #QT00000357)

Quantitect IL-18 primers (Qiagen #QT00014560)

Quantitect IL-23A primers (Qiagen #QT00204078)

### Intestinal lymphocyte isolation

**Explant method**—Human colonic lymphocytes were isolated as previously described (3) and used after a short-term crawl out (36-48h) without the addition of exogenous cytokines, or after 5-7 days of culture in the presence of cytokines when cells were used in co-culture assays or cytokine skewing experiments (see below).

**Digestion**—Whole colonic biopsies were digested using Miltenyi's Tumour Dissociation Kit (Miltenyi Biotec) according to manufacturer's instructions for the isolation of lymphocytes.

### Primary cell culture conditions

Samples used in co-culture assays and cytokine skewing experiments were isolated using the explant method and cultured for 5 days in complete media supplemented with IL2 (100 U/mL, Novartis Pharmaceutical UK) and IL15 (10 ng/mL, Biolegend). 1ml of medium was aspirated every second day and replaced with complete medium containing 2x concentrated cytokines. In cytokine skewing experiments, media was also supplemented with IL2 (100 U/mL), IL15 (10 ng/mL), IL1 $\beta$  (10 ng/mL, Peprotech) and IL18 (50 ng/mL, Medical and Biological Laboratories); or with IL2 (100 U/mL), IL15 (10 ng/mL), IL1 $\beta$  (10 ng/mL Biolegend), IL23 (10 ng/mL, Biolegend); or with IL2 (100 U/mL), IL15 (10 ng/mL), TNF $\alpha$  (50 ng/ml, Biogen).

### Flow cytometry

Flow cytometry was performed using the following antibodies, coupled to the indicated fluorochromes (clones indicated in brackets). Antibodies were purchased from Biolegend unless otherwise stated in brackets. Viability dye Aqua was from Invitrogen. Viability dye Zombie NIR was from Biolegend. Ultra-LEAF™ Purified Human IgG<sub>1</sub> Isotype Control antibody (QA16A12) was from Biolegend. The following antibodies were used: V $\alpha$ 2-AF700 (B6); V $\alpha$ 2-PerCpCy5.5 (B6);  $\gamma$ 2TCR-PE.Cy7 (IMMU510; Beckman Coulter); V $\alpha$ 1-FITC (TS8.2; ThermoFisher); V $\alpha$ 1-PE (REA173; Miltenyi); V $\alpha$ 1-APC (REA173; Miltenyi); CD3-BV786 (OKT3); CD3-BV510 (OKT3); CD8 $\alpha$ -BV510 (HIT8a); CD8 $\alpha$ -PerCpCy5.5 (HIT8a); CD8 $\alpha$ -FITC (HIT8a); CD103-FITC (Ber-ACT8); CD103-BV605 (Ber-ACT8); CD103- PerCpCy5.5 (Ber-ACT8); CD103-BV421 (Ber-ACT8); CD5-BV421 (L17F12); CD18-PE (1B4/18); 41BB-BV605 (4B4-1); IFN $\gamma$ -BV421 (4S.B3); TNF $\alpha$ -APC (MAb11); IL17A- BV605 (BL168); CD101-PerCpCy5.5 (BB27); TIGIT-PerCpCy5.5 (A15153D); TIGIT-PerCpCy5.5 (741182; BD Biosciences); Nkp30-BV711 (P30-15); Nkp44-BV605 (p44-8); Nkp46-PE (9E2); Nkp46-BV421 (9E2); CD49a-APC (TS2/7); CD69-BV510 (FN50); CD38- BV605 (HB-7); ICOS-BV711 (C398.4A); NKG2C-BV421 (134591; BD Biosciences); CCR9-APC (L053E8); CD31-BV421 (WM59); CD96-BV421 (NK92.39); NKG2D-BV605 (1D11); DNAM1-BV605 (DX11; BD Biosciences); CCR7-APC (G043H7); CCR7-BV421(G043H7); NKG2A-BV711 (131411; BD Biosciences);

TIM3-BV421 (F38-2E2); Fc $\gamma$ RI $\gamma$ -FITC (FCABS400F; Merck); CD107a-PE (LAMP1); IL2-PE (MQ1-17H12); CD45RA-BV421 (HI100); CD27-BV786 (O323). The biotin-conjugated antibody against V $\gamma$ 2/3/4 (23D12) (70) was detected by conjugation to BV650 or PE-streptavidin (Biolegend). The anti-V $\gamma$ 4 antibodies (G4\_9 and G4\_18; Iontas) were human IgG $_1$  isotype and goat anti-human IgG (H+L)-AF647 (polyclonal; ThermoFisher) was used to detect unconjugated clones. In addition, clone G4\_18 was conjugated to AF647 using the Lightning Link Fast conjugation kit (Abcam) following manufacturer's instruction with the exception that the incubation period with modifier agent was extended to overnight at 4°C. Primary cells were fixed with BD CellFIX (BD Biosciences) for 20 min at room temperature prior to acquisition.

For cytokine stains, lymphocytes were stimulated with 10 ng/ml PMA (phorbol 12-myristate 13-acetate - Merck) and 1  $\mu$ g/ml ionomycin (Merck) in the presence of 20  $\mu$ g/ml BFA (Merck) for 4 h at 37°C 5% CO $_2$ . Following surface staining and cell fix, lymphocytes were permeabilized with Intracellular Staining Permeabilization Wash Buffer (Biolegend) for 20 min at room temperature. The relevant conjugated monoclonal antibodies were added to the permeabilization buffer to the cells and incubated for another 20 min. The cells were washed and acquired.

When examining degranulation in the co-culture assay, conjugated anti-CD107a-PE antibody was added at the start of a 6 h co-culture assay and surface staining was performed at the end of the assay.

Experiments were acquired on BD LSRFortessa X-20 or BD FACSCanto II. Flow cytometry data analysis was performed on FlowJo (Version 10).

### Generation of anti-human V $\gamma$ 4 antibodies

Anti-V $\gamma$ 4 TCR antibodies were generated by Iontas and provided by Adaptate Biotherapeutics. They were derived using phage display following industry standard techniques (71). A discovery strategy was devised to exert selection pressure towards the V $\gamma$ 4V $\gamma$ 1 and V $\gamma$ 4V $\gamma$ 2 TCR complexes. To this end, phage display selections were performed with recombinant TCR antigens in solution phase. Binders were screened in soluble scFv format by subcloning the entire selection outputs in a suitable scFv expression vector. Hits were assessed by specific binding to V $\gamma$ 4V $\gamma$ 1 and V $\gamma$ 4V $\gamma$ 2 TCR antigens and not V $\gamma$ 2V $\gamma$ 1, V $\gamma$ 2V $\gamma$ 2, V $\gamma$ 8V $\gamma$ 1 TCR antigens. Unique binder sequences were ranked according to their relative affinity to V $\gamma$ 4 in a scFv capture ELISA. A final selection of 24 binders with low sequence liability, varying estimated affinities and CDR3 diversity were chosen for IgG conversion. Of the 24 candidates, 2 were validated to bind V $\gamma$ 4 $^+$  TCR specifically, with good staining indexes and with minimal background signal (fig. S3A).

### Cell lines

HEK293T cells were from the ATCC and were maintained in DMEM (supplemented with 10% FBS and 100 U/mL Penicillin-Streptomycin). Human TCR $\beta^-$  cell line JRT3 was a gift from Dr Salah Mansour, University of Southampton and was cultured in RPMI 1640 (supplemented with 10% FBS and 100 U/mL Penicillin-Streptomycin).

Cells transduced with empty vector (EV) or BTNL3 and BTNL8 have previously been described (3) and were cultured in DMEM (10% FBS, 100 U/mL Penicillin-Streptomycin) supplemented with 1 µg/mL puromycin (Sigma-Aldrich) and 500 µg/mL hygromycin (Thermo Fisher).

### Plasmids, cloning, transfection and lentiviral transduction

Plasmids encoding hu17 (V $\gamma$ 4V $\gamma$ 1), hu17.V $\gamma$ 2 (V $\gamma$ 2V $\gamma$ 1), hu17.V $\gamma$ 3 (V $\gamma$ 3V $\gamma$ 1), hu17.V $\gamma$ 3-V $\gamma$ 4<sup>HV4</sup>(V $\gamma$ 3V $\gamma$ 1), hu17.V $\gamma$ 3-V $\gamma$ 4<sup>CDR2-HV4</sup>(V $\gamma$ 3V $\gamma$ 1), hu20 (V $\gamma$ 4V $\gamma$ 1) and huPB (V $\gamma$ 9V $\gamma$ 2) have previously been described (12, 13). The plasmid encoding hu20/huPB (V $\gamma$ 4V $\gamma$ 2) was generated by subcloning the hu20 gamma chain using NcoI/XbaI and the huPB delta chain using XhoI/NotI into the self-inactivating lentiviral vector pCSIGPW after removal of the IRES-GFP and CMVp-Puro<sup>R</sup> cassettes (3, 48).

The self-inactivating lentiviral vector pCSIGPW (SFFV promoter – Multiple Cloning Site [MCS] – IRES-GFP – CMV promoter – Puromycin<sup>R</sup>), and plasmids encoding FLAG-BTNL3 and FLAG-BTNL8 were previously described (3). The FLAG-BTNL8\*3 construct was derived from FLAG-BTNL8 and FLAG-BTNL3 plasmids by using a unique DraIII restriction site at the beginning of the B30.2 domains of BTNL8 and BTNL3.

All plasmids were transduced and transfected following a previously established protocol (12, 13).

### Co-culture assays

5x10<sup>5</sup> HEK293T cells, transduced with either EV (293T.EV), BTNL3 and BTNL8 (293T.L3L8) or BTNL8 and BTNL8\*3 (293T.L8L8\*3) were co-cultured with 2x10<sup>5</sup> primary human lymphocytes harvested after 5 days of culture in the presence of IL2 and IL15 as previously described (12). Cells were co-cultured overnight at 37°C, 5% CO<sub>2</sub> in 96-well round bottom plates with complete medium without supplementary cytokines. Cells were subsequently analysed by flow cytometry.

### RNA sequencing

**RNA sequencing of human intestinal CD103<sup>+neg</sup> V $\gamma$ 2/3/4<sup>+</sup> cells**—Intestinal lymphocytes were isolated using the Miltenyi tumour dissociation kit as described above. Cells were sorted directly into TRIzol n=1 or PBS n=2 and sent to GENEWIZ (Bishop's Stortford, UK) for RNA extraction, cDNA was prepared with an ultralow input. Single-end 150 bp sequencing was undertaken on an Illumina HiSeq. The final sample (Sample A) was FACS sorted directly into lysis buffer and RNA isolated via the PicoPure kit (ThermoFisher) as per manufacturer's instructions. cDNA was prepared with Ovation RNA-Seq System V2 (NuGEN) and libraries were constructed with Ultralow Library System V2 (NuGEN) and sequenced single-end sequence read 75 bp undertaken on an Illumina HiSeq 2500. Single-end reads were aligned to human genome GRCh38 from Ensembl using STAR (v2.5.2a). Gene-level quantifications using annotation release 86 from Ensembl were generated using RSEM v1.3.0. To identify differential expression raw gene level counts were imported into DESeq2 (72). Genes with a mean normalized count below zero across all samples were removed prior to statistical analysis.



The data have been deposited in NCBI's Gene Expression Omnibus (GEO) under accession number GSE218058 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE218058>) (73).

**RNA sequencing of human intestinal V $\gamma$ 2/3/4<sup>+</sup> cells co-cultured with 293T.EV or 293T.L3L8 cells.**—Intestinal lymphocytes were isolated using the explant grid culture method for two weeks in IL2 and IL15. Cells were then harvested and co-cultured overnight with 293T.EV or 293T.L3L8. V $\gamma$ 2/3/4<sup>+</sup> cells from each condition were sorted directly into lysis buffer and RNA was extracted using the RNEasy MicroPlus kit (Qiagen) following manufacturer's instructions. RNA was sent for library preparation and single-end sequencing on Illumina HiSeq 2500 at the Advanced Sequencing Facility (The Francis Crick Institute).

75 base-pair single-end reads were aligned to human genome GRCh38 from Ensembl using STAR (v2.5.2a). Gene-level quantifications using annotation release 86 from Ensembl were generated using RSEM v1.3.0. To identify differential expression raw gene level counts were imported into DESeq2(72). Genes with a mean normalized count below zero across all samples were removed prior to statistical analysis. Differential expression between the V $\gamma$ 2/3/4<sup>+</sup> cells co-cultured with 293T.EV and 293T.L3L8L3+L8 groups was identified by taking into account the paired structure within the replicate groups. Using an FDR of 0.01 37 culture-condition-dependent gene expression effects were identified.

The data have been deposited in NCBI's GEO under the accession number GSE224412 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE224412>) (73).

**RNA sequencing of human intestinal  $\gamma$ 6 T cells**—11 biopsies from control donors were digested and live V $\gamma$ 4<sup>+</sup>V $\gamma$ 2/3/4<sup>+</sup> $\gamma$ 6TCR<sup>+</sup>CD3<sup>+</sup>V $\gamma$ 2<sup>neg</sup> and V $\gamma$ 4<sup>neg</sup>V $\gamma$ 234<sup>neg</sup> $\gamma$ 6TCR<sup>+</sup>CD3<sup>+</sup>V $\gamma$ 2<sup>neg</sup> cells were sorted directly into extraction buffer from the PicoPure RNA Isolation Kit (ThermoFisher). RNA was isolated immediately, adhering to manufacturer's instruction. Total RNA was stored at -80°C freezer until submission. Libraries were prepared by the Advanced Sequencing facility at the Francis Crick Institute using the NEBNext Ultra II RNA Library prep kit. Libraries were sequenced by the BRC Genomics Facility at Guy's hospital on Illumina NextSeq 2000.

All software for the data analysis was run with default settings, unless otherwise indicated. The quality of the sequencing reads was examined using FastQC (v0.11.4). Raw sequencing reads (100-nt, paired-end) were trimmed using TrimGalore (v0.4.4). Traces of ribosomal DNA and mitochondrial DNA were removed using the Bowtie2 (v2.2.5) (74). Reads were aligned to the human reference genome GRCh38 using STAR (v2.5.3a) with default setting (75). Duplicate reads were removed using the MarkDuplicates function of the Picard tools (v2.17.11). Reads were annotated using the Partek E/M (Partek software) gene count. Differentially expressed genes (DEG) were identified using DESeq2 (v3.5) (72). All software for the data analysis was run with default settings, unless otherwise indicated. The quality of the sequencing reads was examined using FastQC (v0.11.4). Raw sequencing reads (100-nt,paired-end) were trimmed using TrimGalore (v0.4.4). Traces of ribosomal DNA and mitochondrial DNA were removed using the Bowtie2 (v2.2.5) (74). Reads were aligned to the human reference genome GRCh38 using STAR (v2.5.3a) with

default setting (75). Duplicate reads were removed using the MarkDuplicates function of the Picard tools (v2.17.11). Reads were annotated using the Partek E/M (Partek software) gene count. Differentially expressed genes (DEG) were identified using DESeq2 (v3.5) (72).

The data have been deposited in NCBI's Gene Expression Omnibus (GEO) and are accessible through GEO Series accession number GSE218446 and at the following link: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE218446> (73).

**Gene Set Enrichment Analysis (GSEA)**—GSEA pre-ranked analysis was performed using DEGs between  $V\gamma 2/3/4$  CD103<sup>+</sup> (positive values) and CD103<sup>neg</sup> (negative values) with several published datasets (Fig. 1D, E; fig. S1E) (76).

For human datasets the published gene signature of the cells under investigation was used. To compare the published murine  $V\gamma 7^+$  data, murine genes were converted to human orthologues using <https://biit.cs.ut.ee/gprofiler/gost> (77). Three published gene signatures were large (3, 25, 26), hence the top 500 significantly DEG were taken as the gene signature.

### DNA samples for genotyping experiments

The study included 12,767 unrelated UK patients with IBD. The ethnicity of both cases and controls was limited to White European ancestry as these alleles are underrepresented in other ethnicities (47). The study included a test set of 4517 IBD patients consisting of patients recruited after ethical review and obtaining of informed consent from Guy's and St. Thomas' Hospitals London, United Kingdom and the Royal Victoria Infirmary, Newcastle, United Kingdom (REC 12/YH/0172). The replication set included 8250 IBD patients incorporated into the NIHR IBD bioresource. The diagnosis of IBD was made by established criteria of clinical, radiologic, and endoscopic analysis and from histology reports. Patients were classified according to the Montreal Classification into disease location (L1-3) and disease behavior (B1-3). Duration of disease was defined as onset of symptoms or date of diagnosis to date of inclusion in the database. B1 cases were included only if they had at least 5 years documented follow up to avoid incorporating patients who had not yet progressed to other phenotypes. B2 and B3 disease were recorded regardless of disease duration. Only patients with CD and sufficient phenotypic and disease duration data were included in the final analysis of genetic association with disease behaviour (test  $n=1817$ , replication  $n=4491$ ).

The study also included 6767 population controls, obtained from the 1958 British Birth Cohort, including all subjects born between 3<sup>rd</sup> and 9<sup>th</sup> March 1958 in England, Scotland, and Wales (National Child Development Study: <http://www.cls.ioe.ac.uk>) (49).

### BTNL8\*3 CNV genotyping by PCR

All locally sourced gDNA obtained from donors was used at 10 ng/reaction. PCR was performed using Q5 High Fidelity DNA polymerase (NEB) following manufacturer's instructions and in a single tube reaction. The following primers were used:

L8\_F 5'-GACTTTTCACCCCCAAAACC-3'

L8\_R 5'-CTCATGTCCACCTGAGATTCTG-3'

L3\_R 5'-GTAAGCCCTCTATTTGACTTTGTG-3'

Predicted product sizes were 1602 bp for BTNL8\*3 and 1310 bp for BTNL8. PCR product was separated by 1% agarose gel.

### Identification of rs72494581 as a surrogate SNP of the BTNL8\*3 CNV

We cloned and sequenced the PCR products generated by amplification of the genomic region resulting from the CNV using primers L8\_F and L3\_R (sequences above), which generates a 1602 bp product (fig. S6A). After aligning the obtained sequences from 6 different CNV homozygous donors (GUT1-6) to the reference sequences of the genomic regions of BTNL8 (L8) and BTNL3 (L3), as well as the predicted sequence of the CNV allele (L8\*3), we confirmed the location of the non-allelic homologous recombination region suggested by (47) (fig. S6B). Of note, this region starts with a sequence that is near-identical to a recombination hot spot sequence reported by Myers et al. (78).

The CNV alleles from all 6 donors all included a T>C SNP (rs72494581, not reported by Aigner et al (47)). We then compared the results of the RT-PCR method using L8\_F, L8\_R and L3\_R primers (sequences above and fig S6A) using gDNA from donors that had previously been genotyped with a custom TaqMan-based genotyping assay for rs72494581 (fig. S6C). Given the relatively limited number of samples tested, we cannot guarantee that rs72494581 is in complete LD in all cases but given its proximity to the putative break point (<200bp) and the perfect match between the RT-PCR and TaqMan data (fig S6A, C), this SNP is potentially the best marker of the CNV allele.

### BTNL8\*3 CNV genotyping by TaqMan assay

Custom TaqMan SNP genotyping assay for rs72494581 was ordered from ThermoFisher and used as per manufacturer's instructions with TaqMan Genotyping Master Mix (ThermoFisher) and a minimum of 4 ng of each gDNA sample per reaction. Post-amplification allelic discrimination was carried out using the ABI7900HT (TaqMan plate reader, end-point read). Case control association test was performed by Chi-squared test. Primer sequences were:

rs72494581\_F 5' -CTCGTCCCACCTGTGCAA-3'

rs72494581\_R 5' -GAAACAGTATATTCAGTGGTGGACACA-3'

### KASP genotyping

A KASP assay for rs72494581 was custom designed and validated against known samples. High throughput genotyping was outsourced and undertaken by KASP genotyping (LGC Biosearch Technologies, Hoddesdon, UK). Clinical data relating to patients tested for rs72494581 was provided by the NIHR IBD Bioresource.

### Statistical analysis

With the exception of sequencing analysis (see above), all other statistics were performed in GraphPad Prism v9. Statistical information is included in the figure legends. In all

figures horizontal bars denote the median and error bars denote the minimum and maximum datapoints.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

This work was supported by the NIHR Newcastle Biomedical Research Centre, the NIHR Guy's and St Thomas Biomedical Research Centre and the NIHR Biomedical Research Centre at Cambridge University Hospital NHS Foundation Trust and the University of Cambridge. We acknowledge the help of the IBD Bioresource, supported by the NIHR, Medical Research Council, Crohn's and Colitis UK, Open Targets, and the Wellcome Trust Sanger Institute. We thank O. Polyakova, A. McGinley, D. Wesch, and D. Kabelitz for provision of antibodies and J. Sanderson and the staff and patients of Guy's and St Thomas Foundation Trust for support related to clinical sampling. We thank D. Gibbons, T. Raine, J. Spencer and D. McKenzie for their critical review of the manuscript and their insights. The print summary figure was created with BioRender.com.

## Funding

The Francis Crick Institute, which receives its core funding from Cancer Research UK (CRUK) (FC001003), the UK Medical Research Council (FC001003), and the Wellcome Trust (FC001003). Clinical research training fellowship from the Wellcome Trust (108745/Z/15/Z) (R.J.D.); A studentship from the National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London (I.Z); The Wellcome Trust (094491/Z/10/Z) (N.J.P.) and (106292/Z/14/Z and 100156/Z/12/Z) (A.C.H.).

## Data and materials availability

All data are available in the manuscript or the supplementary materials or are deposited as follows: sequencing data generated for this study is available on NCBI GeneExpression Omnibus (GEO) under accession numbers GSE218058, GSE224412 and GSE218446. All reagents described in this paper derived for these studies will be shared upon request with other researchers in academia. Any MTAs will be organised upon receipt of request to the corresponding authors.

## References

1. Fernando MMA, et al. Defining the Role of the MHC in Autoimmunity: A Review and Pooled Analysis. *PLOS Genetics*. 2008; 4 e1000024 [PubMed: 18437207]
2. Hayday, AC. Paul's Fundamental Immunology. Wolters Kluwer/Lippincott Williams Wikins; Philadelphia: 2022. 1054–1101. chap. 35
3. Di Marco Barros R, et al. Epithelia Use Butyrophilin-like Molecules to Shape Organ-Specific gammadelta T Cell Compartments. *Cell*. 2016; 167: 203–218. e217 [PubMed: 27641500]
4. Jandke A, et al. Butyrophilin-like proteins display combinatorial diversity in selecting and maintaining signature intraepithelial gammadelta T cell compartments. *Nature communications*. 2020; 11 3769
5. Barbee SD, et al. Skint-1 is a highly specific, unique selecting component for epidermal T cells. *Proceedings of the National Academy of Sciences*. 2011; 108: 3330–3335.
6. Girardi M, et al. Resident skin-specific gammadelta T cells provide local, nonredundant regulation of cutaneous inflammation. *The Journal of experimental medicine*. 2002; 195: 855–867. [PubMed: 11927630]
7. Muñoz-Ruiz M, et al. Tissue-intrinsic  $\gamma\delta$  T cells critically regulate Tissue-Resident Memory CD8 T cells. *bioRxiv*. 2022. 2022.2001.2019.476598

8. Kuhl AA, et al. Aggravation of intestinal inflammation by depletion/deficiency of gammadelta T cells in different types of IBD animal models. *Journal of leukocyte biology*. 2007; 81: 168–175. [PubMed: 17041003]
9. Chen Y, Chou K, Fuchs E, Havran WLR. Boismenu, Protection of the intestinal mucosa by intraepithelial gamma delta T cells. *Proceedings of the National Academy of Sciences*. 2002; 99: 14338–14343.
10. Inagaki-Ohara K, et al. Mucosal T cells bearing TCRgammadelta play a protective role in intestinal inflammation. *Journal of immunology*. 2004; 173: 1390–1398.
11. Matsuzawa-Ishimoto Y, et al. The gammadelta IEL effector API5 masks genetic susceptibility to Paneth cell death. *Nature*. 2022; 610: 547–554. [PubMed: 36198790]
12. Melandri D, et al. The  $\gamma\delta$  T cell receptor combines innate with adaptive immunity by utilizing spatially distinct regions for agonist-selection and antigen responsiveness. *Nature immunology*. 2018.
13. Willcox CR, et al. Butyrophilin-like 3 Directly Binds a Human Vgamma4(+) T Cell Receptor Using a Modality Distinct from Clonally-Restricted Antigen. *Immunity*. 2019; 51: 813–825. e814 [PubMed: 31628053]
14. Mayassi T, et al. Chronic Inflammation Permanently Reshapes Tissue-Resident Immunity in Celiac Disease. *Cell*. 2019; 176: 967–981. e919 [PubMed: 30739797]
15. Lebrero-Fernandez C, et al. Altered expression of Butyrophilin (BTN) and BTN-like (BTNL) genes in intestinal inflammation and colon cancer. *Immun Inflamm Dis*. 2016; 4: 191–200. [PubMed: 27957327]
16. Cerf-Bensussan N, et al. A monoclonal antibody (HML-1) defining a novel membrane molecule present on human intestinal lymphocytes. *European journal of immunology*. 1987; 17: 1279–1285. [PubMed: 3498635]
17. Cepek KL, et al. Adhesion between epithelial cells and T lymphocytes mediated by E-cadherin and the alpha E beta 7 integrin. *Nature*. 1994; 372: 190–193. [PubMed: 7969453]
18. Cantoni C, et al. p49, a putative HLA class I-specific inhibitory NK receptor belonging to the immunoglobulin superfamily. *European journal of immunology*. 1998; 28: 1980–1990. [PubMed: 9645380]
19. Wermers JD, McNamee EN, Wurbel MA, Jedlicka P, Rivera-Nieves J. The chemokine receptor CCR9 is required for the T-cell-mediated regulation of chronic ileitis in mice. *Gastroenterology*. 2011; 140: 1526–1535. e1523 [PubMed: 21300065]
20. Voskoboinik I, Whisstock JC, Trapani JA. Perforin and granzymes: function, dysfunction and human pathology. *Nature Reviews Immunology*. 2015; 15: 388–400.
21. Okamura H, Kashiwamura S, Tsutsui H, Yoshimoto T, Nakanishi K. Regulation of interferon-gamma production by IL-12 and IL-18. *Curr Opin Immunol*. 1998; 10: 259–264. [PubMed: 9638361]
22. Deaglio S, et al. Adenosine generation catalyzed by CD39 and CD73 expressed on regulatory T cells mediates immune suppression. *Journal of Experimental Medicine*. 2007; 204: 1257–1265. [PubMed: 17502665]
23. Monney L, et al. Th1-specific cell surface protein Tim-3 regulates macrophage activation and severity of an autoimmune disease. *Nature*. 2002; 415: 536–541. [PubMed: 11823861]
24. Gay D, et al. Fgf9 from dermal gammadelta T cells induces hair follicle neogenesis after wounding. *Nature medicine*. 2013; 19: 916–923.
25. Atlasy N, et al. Single cell transcriptomic analysis of the immune cell compartment in the human small intestine and in Celiac disease. *Nature communications*. 2022; 13 4920
26. Kumar BV, et al. Human Tissue-Resident Memory T Cells Are Defined by Core Transcriptional and Functional Signatures in Lymphoid and Mucosal Sites. *Cell Rep*. 2017; 20: 2921–2934. [PubMed: 28930685]
27. Fitz Patrick MEB, et al. Human intestinal tissue-resident memory T cells comprise transcriptionally and functionally distinct subsets. *Cell Rep*. 2021; 34 108661 [PubMed: 33472060]



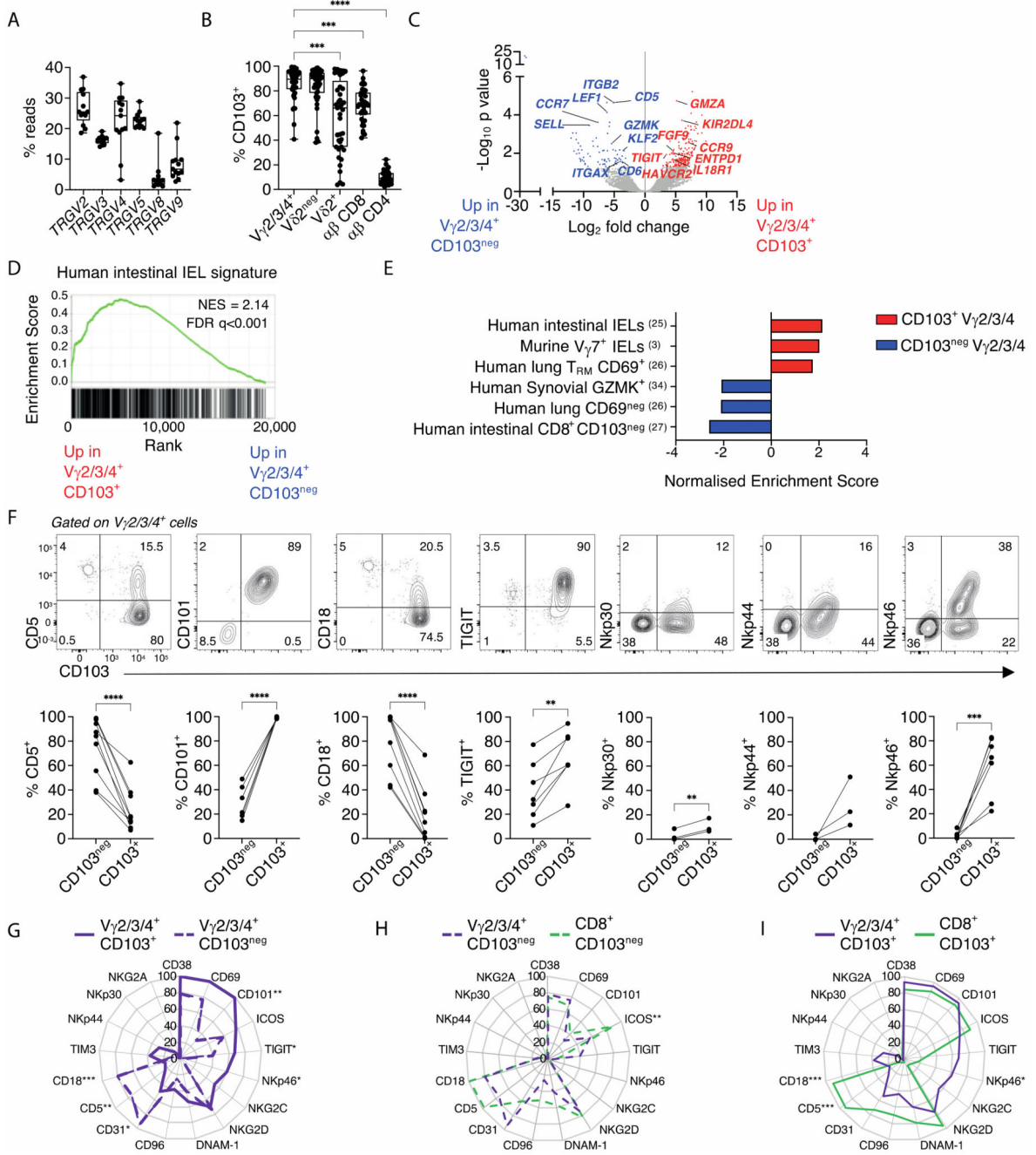
28. Sallusto F, Lenig D, Forster R, Lipp MA, Lanzavecchia, Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature*. 1999; 401: 708–712. [PubMed: 10537110]
29. Blackford J, Reid HW, Pappin DJ, Bowers FS, Wilkinson JM. A monoclonal antibody, 3/22, to rabbit CD11c which induces homotypic T cell aggregation: evidence that ICAM-1 is a ligand for CD11c/CD18. *European journal of immunology*. 1996; 26: 525–531. [PubMed: 8605916]
30. Azzam HS, et al. CD5 expression is developmentally regulated by T cell receptor (TCR) signals and TCR avidity. *The Journal of experimental medicine*. 1998; 188: 2301–2311. [PubMed: 9858516]
31. Orta-Mascaro M, et al. CD6 modulates thymocyte selection and peripheral T cell homeostasis. *The Journal of experimental medicine*. 2016; 213: 1387–1397. [PubMed: 27377588]
32. Kuo CT, Veselits MLJM, Leiden, LKLF: A transcriptional regulator of single-positive T cell quiescence and survival. *Science (New York, NY)*. 1997; 277: 1986–1990.
33. Malhotra N, et al. A network of high-mobility group box transcription factors programs innate interleukin-17 production. *Immunity*. 2013; 38: 681–693. [PubMed: 23562159]
34. Jonsson AH, et al. Granzyme K(+) CD8 T cells form a core population in inflamed human tissue. *Science translational medicine*. 2022; 14 eabo0686 [PubMed: 35704599]
35. Groh V, et al. Costimulation of CD8 $\alpha$  T cells by NKG2D via engagement by MIC induced on virus-infected cells. *Nature immunology*. 2001; 2: 255–260. [PubMed: 11224526]
36. McKenzie DR, et al. Normality sensing licenses local T cells for innate-like tissue surveillance. *Nature immunology*. 2022; 23: 411–422. [PubMed: 35165446]
37. Rodewald HR, Arulanandam AR, Koyasu S, Reinherz EL. The high affinity Fc epsilon receptor gamma subunit (Fc epsilon RI gamma) facilitates T cell receptor expression and antigen/major histocompatibility complex-driven signaling in the absence of CD3 zeta and CD3 eta. *J Biol Chem*. 1991; 266: 15974–15978. [PubMed: 1714902]
38. Rodewald HR, et al. A population of early fetal thymocytes expressing Fc gamma RII/III contains precursors of T lymphocytes and natural killer cells. *Cell*. 1992; 69: 139–150. [PubMed: 1532536]
39. Liu CP, et al. Abnormal T cell development in CD3-zeta $^{-/-}$  mutant mice and identification of a novel T cell population in the intestine. *The EMBO journal*. 1993; 12: 4863–4875. [PubMed: 8223495]
40. Dechanet J, et al. Implication of gammadelta T cells in the human immune response to cytomegalovirus. *The Journal of clinical investigation*. 1999; 103: 1437–1449. [PubMed: 10330426]
41. Prinz I, et al. Donor V $\gamma$ 1+  $\gamma$  $\delta$  T cells expand after allogeneic hematopoietic stem cell transplantation and show reactivity against CMV-infected cells but not against progressing B-CLL. *Exp Hematol Oncol*. 2013; 2: 14. [PubMed: 23663319]
42. Davey MS, et al. Clonal selection in the human Vdelta1 T cell repertoire indicates gammadelta TCR-dependent adaptive immune surveillance. *Nature communications*. 2017; 8 14760
43. Rutgeerts P, et al. Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis. *New England Journal of Medicine*. 2005; 353: 2462–2476. [PubMed: 16339095]
44. Sandborn WJ, et al. Ustekinumab induction and maintenance therapy in refractory Crohn's disease. *The New England journal of medicine*. 2012; 367: 1519–1528. [PubMed: 23075178]
45. D'Haens G, et al. Risankizumab as induction therapy for Crohn's disease: results from the phase 3 Advance and Motivate induction trials. *Lancet (London, England)*. 2022; 399: 2015–2030. [PubMed: 35644154]
46. Parikh K, et al. Colonic epithelial cell diversity in health and inflammatory bowel disease. *Nature*. 2019; 567: 49–55. [PubMed: 30814735]
47. Aigner J, et al. A common 56-kilobase deletion in a primate-specific segmental duplication creates a novel butyrophilin-like protein. *BMC genetics*. 2013; 14: 61. [PubMed: 23829304]
48. Vantourout P, et al. Heteromeric interactions regulate butyrophilin (BTN) and BTN-like molecules governing gammadelta T cell biology. *Proceedings of the National Academy of Sciences*. 2018; 115: 1039–1044.
49. Power CJ, Elliott, Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epidemiol*. 2006; 35: 34–41. [PubMed: 16155052]

50. Silverberg MS, et al. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol*. 2005; 19 (Suppl A) 5a–36a.
51. Lo B, et al. Changes in Disease Behaviour and Location in Patients With Crohn's Disease After Seven Years of Follow-Up: A Danish Population-based Inception Cohort. *Journal of Crohn's colitis*. 2018; 12: 265–272.
52. Jameson J, et al. A role for skin gammadelta T cells in wound repair. *Science (New York, NY)*. 2002; 296: 747–749.
53. Turner D, et al. STRIDE-II: An Update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) Initiative of the International Organization for the Study of IBD (IOIBD): Determining Therapeutic Goals for Treat-to-Target strategies in IBD. *Gastroenterology*. 2021; 160: 1570–1583. [PubMed: 33359090]
54. Rath T, et al. Intestinal Barrier Healing Is Superior to Endoscopic and Histologic Remission for Predicting Major Adverse Outcomes in Inflammatory Bowel Disease: The Prospective ERICA Trial. *Gastroenterology*. 2023; 164: 241–255. [PubMed: 36279923]
55. Le DT, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *The New England journal of medicine*. 2015; 372: 2509–2520. [PubMed: 26028255]
56. Vermeire S, et al. Etrolizumab for maintenance therapy in patients with moderately to severely active ulcerative colitis (LAUREL): a randomised, placebo-controlled, double-blind, phase 3 study. *The Lancet Gastroenterology Hepatology*. 2022; 7: 28–37. [PubMed: 34798037]
57. Peyrin-Biroulet L, et al. Etrolizumab as induction and maintenance therapy for ulcerative colitis in patients previously treated with tumour necrosis factor inhibitors (HICKORY): a phase 3, randomised, controlled trial. *Lancet Gastroenterol Hepatol*. 2022; 7: 128–140. [PubMed: 34798039]
58. Hoytema van Konijnenburg DP, et al. Intestinal Epithelial and Intraepithelial T Cell Crosstalk Mediates a Dynamic Response to Infection. *Cell*. 2017; 171: 783–794. e713 [PubMed: 28942917]
59. Kaser A, et al. XBP1 links ER stress to intestinal inflammation and confers genetic risk for human inflammatory bowel disease. *Cell*. 2008; 134: 743–756. [PubMed: 18775308]
60. Corcoran M, et al. Archaic humans have contributed to large-scale variation in modern human T cell receptor genes. *Immunity*. 2023; 56: 635–652. e636 [PubMed: 36796364]
61. Lee JC, et al. Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn's disease. *Nature genetics*. 2017; 49: 262–268. [PubMed: 28067912]
62. Cleyne I, et al. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *Lancet (London, England)*. 2016; 387: 156–167. [PubMed: 26490195]
63. Colombel JF, Narula NL. Peyrin-Biroulet, Management Strategies to Improve Outcomes of Patients With Inflammatory Bowel Diseases. *Gastroenterology*. 2017; 152: 351–361. e355 [PubMed: 27720840]
64. Jones G-R, et al. IBD prevalence in Lothian, Scotland, derived by capture-recapture methodology. *Gut*. 2019; 68: 1953–1960. [PubMed: 31300515]
65. Ng SC, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet (London, England)*. 2017; 390: 2769–2778. [PubMed: 29050646]
66. Rutgeerts P, et al. Predictability of the postoperative course of Crohn's disease. *Gastroenterology*. 1990; 99: 956–963. [PubMed: 2394349]
67. Joseph M, et al. Global patterns of antigen receptor repertoire disruption across adaptive immune compartments in COVID-19. *Proceedings of the National Academy of Sciences of the United States of America*. 2022; 119 e2201541119 [PubMed: 35943978]
68. Bamias G, et al. Comparative study of candidate housekeeping genes for quantification of target gene messenger RNA expression by real-time PCR in patients with inflammatory bowel disease. *Inflammatory bowel diseases*. 2013; 19: 2840–2847. [PubMed: 24141710]
69. Rajput S, Volk-Draper LD, Ran S. TLR4 Is a Novel Determinant of the Response to Paclitaxel in Breast Cancer. *Molecular Cancer Therapeutics*. 2013; 12 1676 [PubMed: 23720768]

70. Kabelitz D, et al. New monoclonal antibody (23D12) recognizing three different V gamma elements of the human gamma delta T cell receptor. 23D12+ cells comprise a major subpopulation of gamma delta T cells in postnatal thymus. *Journal of immunology*. 1994; 152: 3128–3136.
71. Schofield DJ, et al. Application of phage display to high throughput antibody generation and characterization. *Genome Biol*. 2007; 8 R254 [PubMed: 18047641]
72. Love MI, Huber WS. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15: 550. [PubMed: 25516281]
73. Edgar R, Domrachev MAE. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. 2002; 30: 207–210. [PubMed: 11752295]
74. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; 9: 357–359. [PubMed: 22388286]
75. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013; 29: 15–21. [PubMed: 23104886]
76. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102: 15545–15550. [PubMed: 16199517]
77. Raudvere U, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res*. 2019; 47: W191–W198. [PubMed: 31066453]
78. Myers S, Freeman C, Auton A, Donnelly P, McVean G. A common sequence motif associated with recombination hot spots and genome instability in humans. *Nature genetics*. 2008; 40: 1124–1129. [PubMed: 19165926]

### One-Sentence Summary

A unique human gut  $\gamma\delta$  T cell subset, the perturbation of which is a risk factor for poor inflammatory bowel disease outcome.

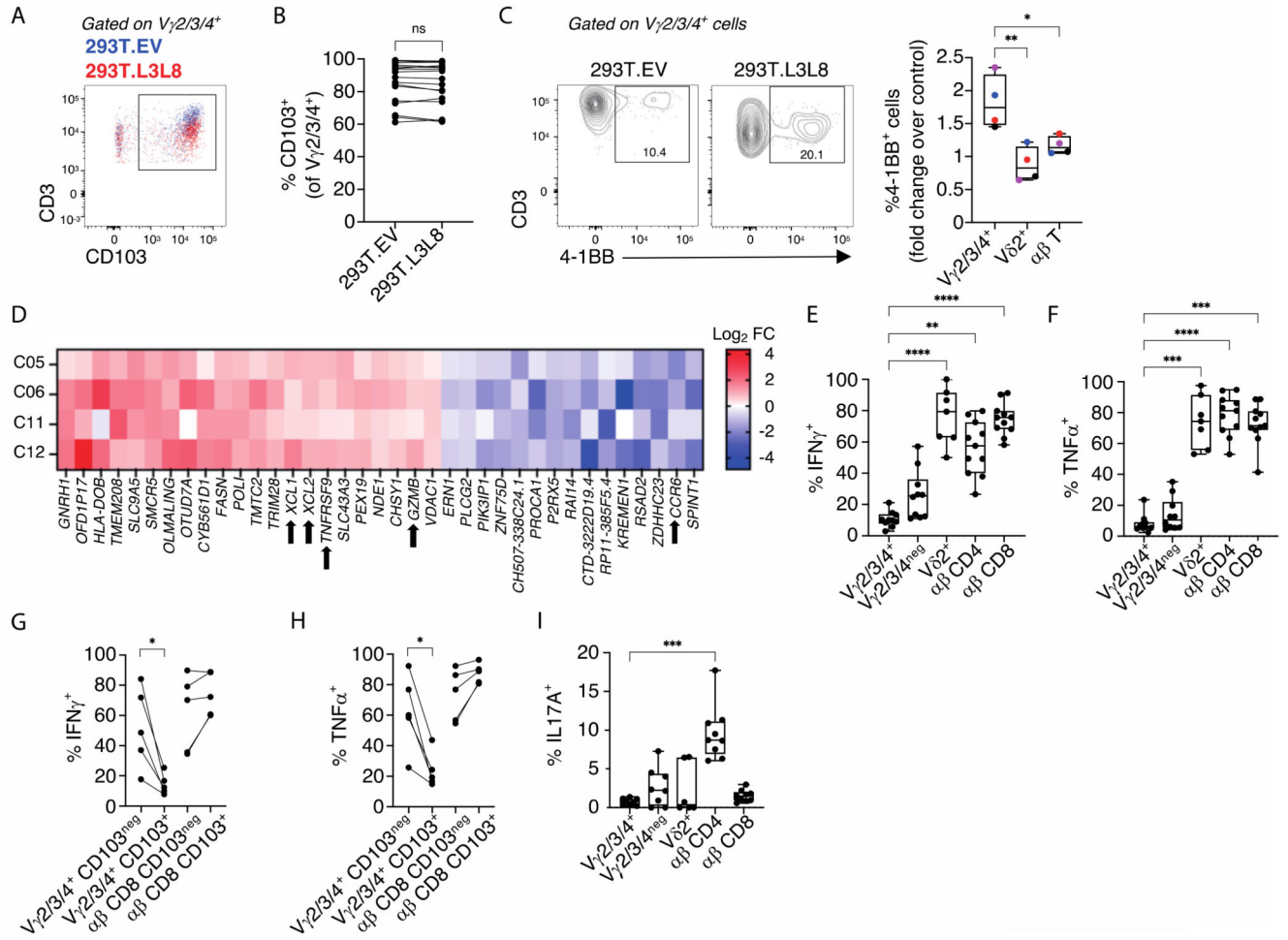


**Figure 1. Human gut  $\gamma\delta$  T cells are diverse and predominantly intraepithelial cells with distinct phenotypes.**

(A). *TRGV* gene usage in control colonic biopsies measured by mRNA TCR deep sequencing (n=13). (B) Flow cytometric analysis of CD103 expression on the indicated  $\gamma\delta$  and  $\alpha\beta$  T cell subsets. (n=38-43). Median, IQR and min-max range shown. Symbols represent individual participants. Kruskal-Wallis test with Dunn's correction against +  $V\gamma 2/3/4^+$  cells was used for analysis, \*\*\* $p$ <0.001; \*\*\*\* $p$ <0.0001. (C). Volcano plot of + neg + differentially expressed genes between CD103+ and CD103neg  $V\gamma 2/3/4^+$  bulk sorted

cells + from control donors (n=4, C36, C37, C39, C88). Red – significantly upregulated in CD103<sup>+</sup>; blue – significantly upregulated in CD103<sup>neg</sup>; grey – non-significant. Corrected for multiple comparisons (FDR<0.05). **(D)**. Gene Set Enrichment Analysis (GSEA) of the human + + CD103 V $\gamma$ 2/3/4 signature from C. in a published human natural intraepithelial lymphocyte (IEL) signature dataset (25). NES – normalized enrichment score. **(E)**. Summary graph of GSEA for CD103<sup>+</sup> V $\gamma$ 2/3/4 (red) and CD103<sup>neg</sup> V $\gamma$ 2/3/4<sup>+</sup> (blue) gene signatures in published datasets, (FDR<0.01). **(F)**. Flow cytometric analysis of the expression of the indicated + neg+ markers (y axis) on CD103<sup>+</sup> and CD103<sup>neg</sup> V $\gamma$ 2/3/4<sup>+</sup> cells. Representative flow plots (top) and summary data (bottom) are shown (n=3 -9). Connected symbols represent individual participants. Paired *t*-tests, \*\**p*<0.01; \*\*\**p*<0.001; \*\*\*\**p*<0.0001 **(G-I)**. Radar plot of the + + + neg cell surface phenotype of (G) V $\gamma$ 2/3/4<sup>+</sup> CD103<sup>+</sup> versus V $\gamma$ 2/3/4<sup>+</sup> CD103<sup>neg</sup> T cells; (H) V $\gamma$ 2/3/4<sup>+</sup> CD103<sup>neg</sup> versus CD8<sup>+</sup> CD103<sup>neg</sup>  $\alpha\beta$  T cells; (I) V $\gamma$ 2/3/4<sup>+</sup> CD103<sup>+</sup> versus CD8<sup>+</sup> CD103<sup>+</sup>  $\alpha\beta$  T cells. (n=3-9 individuals per group) Plotted values represent mean of included participants. Multiple paired *t*-tests with Bonferroni-Dunn correction, \**p*<0.05; \*\**p*<0.01; \*\*\**p*<0.001; \*\*\*\**p*<0.0001.+

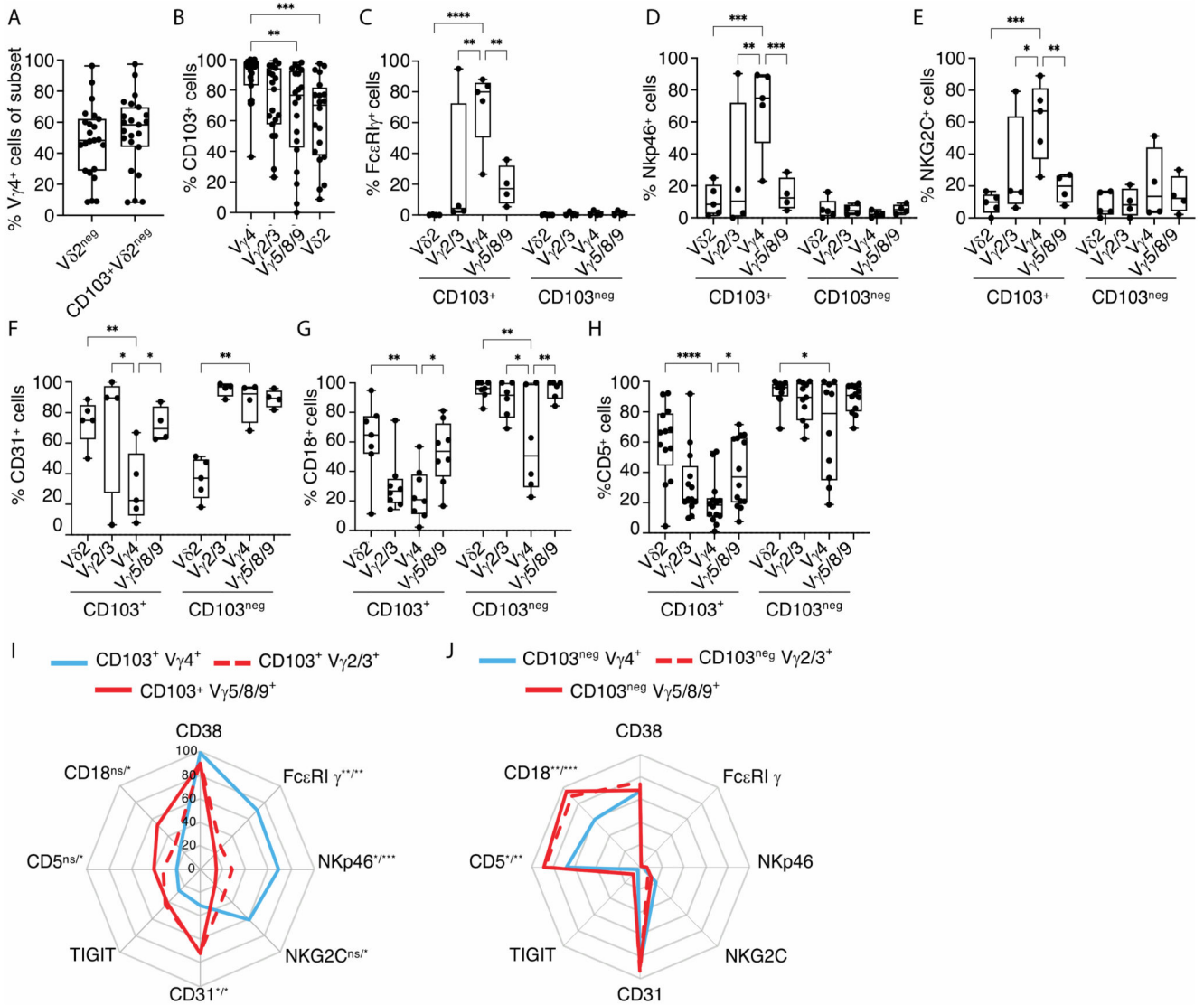




**Figure 2. V $\gamma$ 2/3/4<sup>+</sup> cells make functional responses to BTNL3+BTNL8 stimulation but do not make classical effector cytokine responses.**

(A). Example flow plot of CD3 expression neg on CD103<sup>neg</sup> and CD103<sup>+</sup> V $\gamma$ 2/3/4<sup>+</sup> cells following overnight co-culture with 293T.EV (blue) or 293T.L3L8 cells (red). (B). Flow cytometric analysis of CD103 expression on intestinal V $\gamma$ 2/3/4<sup>+</sup> cells from control donors after co-culture as in A (n=18). Connected symbols represent individual participants. Paired *t*-test, ns – not significant (C). Flow cytometric analysis of 4-1BB expression on indicated colonic cell subsets (x axis) after co-culture as in A. Example flow plots (left) and summary graph (right) shown (n=4). Median, IQR and min-max range shown. Coloured symbols represent matched frequencies from individual participants. Repeated measures one-way ANOVA with Dunnett’s correction against V $\gamma$ 2/3/4<sup>+</sup> cells was used for analysis, \**p*<0.05; \*\**p*<0.01. (D). Significantly differentially expressed genes (x axis) in V $\gamma$ 2/3/4<sup>+</sup> cells co-cultured as in A. Data expressed as Log<sub>2</sub> fold change (FC) for each donor (y axis). (E,F). Flow cytometric analysis of (E) IFN $\gamma$  and (F) TNF $\alpha$  expression by the indicated colonic cell types (x axis) from control donors (n=7 -11) after 4h P+I stimulation. Median, IQR and min-max range shown. Symbols represent individual + participants. Kruskal-Wallis test with Dunn’s correction against V $\gamma$ 2/3/4<sup>+</sup> cells was used for analysis, \*\**p*<0.01; \*\*\**p*<0.001; \*\*\*\**p*<0.0001. (G,H). Flow cytometric analysis of (G) ++ TNF $\alpha$  and (H)

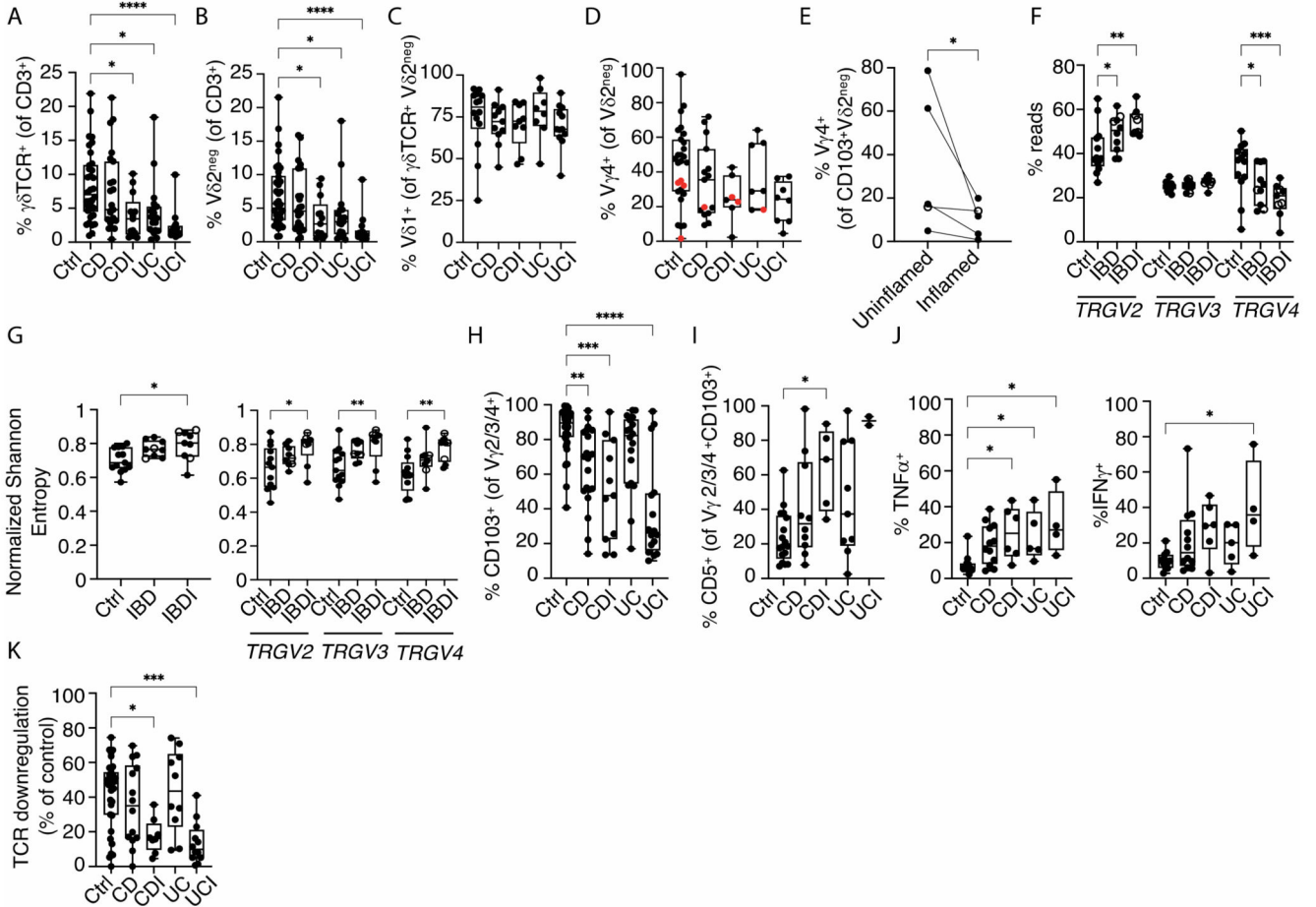
IFN $\gamma$  production by V $\gamma$ 2/3/4<sup>+</sup>  $\gamma\delta$  and CD8<sup>+</sup>  $\alpha\beta$  T cells (x axis) (n=5) and split by their CD103 expression (legend) following 4h of P+I stimulation. Connected symbols represent individual participants. Multiple paired *t*-tests with Bonferroni-Dunn correction, \**p*<0.05. **(I)**. Flow cytometric analysis of IL17A expression as in E and F (n=6-9). Median, IQR and min-max range shown. Symbols represent individual participants. Kruskal-Wallis + with Dunn's correction against V $\gamma$ 2/3/4<sup>+</sup> cells was used for analysis, \*\*\**p*<0.001.



**Figure 3. Vγ4 cells are phenotypically distinct from other γδ cells**

(A). Flow cytometric analysis of % Vγ4<sup>+</sup> cells (y axis) in the indicated γδ subsets (x axis) from control biopsies (n=23). (B). Flow cytometric analysis of CD103 expression on the indicated γδ subsets (x axis) from control biopsies (n=19-23). Kruskal-Wallis test with Dunn’s correction against Vγ4<sup>+</sup> cells was used for analysis, \*\**p*<0.01; \*\*\*\**p*<0.0001. (C-H). Flow cytometric analysis of (C) % FcεRI<sup>+</sup> (n=4-5); (D) % Nkp46<sup>+</sup> (n=4-5); (E) % NKG2C<sup>+</sup> (n=4-5); (F) % CD31<sup>+</sup> (n=4-5); (G) % CD18<sup>+</sup> (n=6-8); (H) % CD5<sup>+</sup> (n=10-14) cells on the indicated γδ subsets (x axis, CD103<sup>+</sup> and CD103<sup>neg</sup>) from control gut tissue. Two-way ANOVA with Dunnett correction against Vγ4<sup>+</sup> cells was used for analysis, \**p*<0.05; \*\**p*<0.01; \*\*\**p*<0.001; \*\*\*\**p*<0.0001. (A-H) Median, IQR and min-max range shown. Symbols represent individual participants. (I, J). Radar plot of (I) the surface phenotype of CD103<sup>+</sup> Vγ4<sup>+</sup> versus CD103<sup>+</sup> Vγ2/3<sup>+</sup> versus CD103<sup>+</sup> Vγ5/8/9<sup>+</sup> γδ T cells; and (J) CD103<sup>neg</sup> Vγ4<sup>+</sup> versus CD103<sup>neg</sup> Vγ2/3<sup>+</sup> versus CD103<sup>neg</sup> Vγ5/8/9<sup>+</sup> γδ T cells from control donors (n=4-14). Plotted values represent mean of included participants. Two-

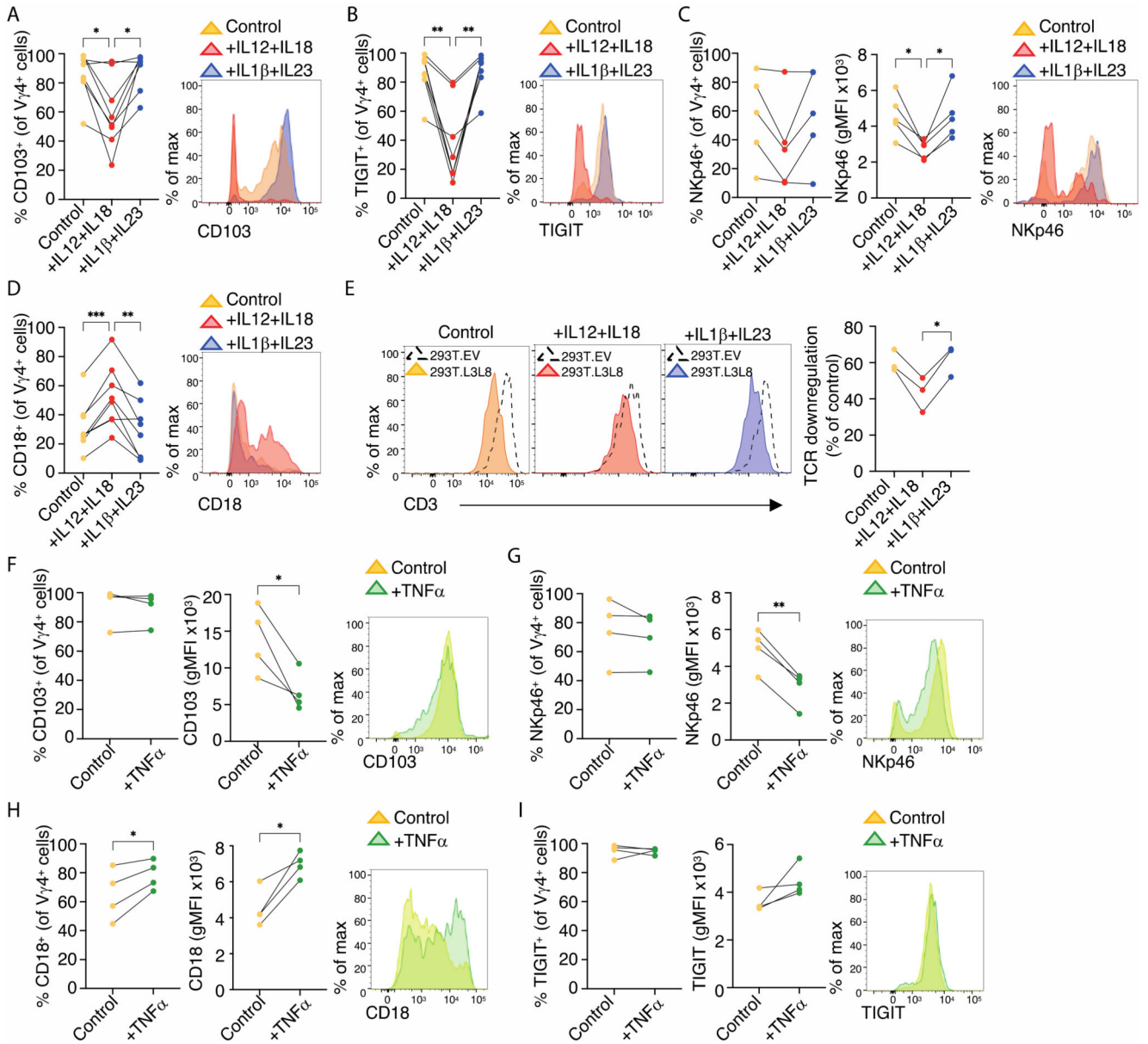
way ANOVA + with Dunnett correction against V $\gamma$ 4 cells was used for analysis, \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; ns – not significant. Significance displayed as “versus V $\gamma$ 2/3<sup>+</sup> / versus + V $\gamma$ 5/8/9<sup>+</sup>”.



**Figure 4. Phenotypic and clonotypic changes of the intestinal  $\gamma\delta$  T cells compartment in IBD.** (A-C). Flow cytometric analysis of (A) %  $\gamma\delta$ TCR<sup>+</sup>, and (B) % V $\delta$ 2<sup>neg</sup> cells of total T lymphocytes isolated from control (n=34), CD (n=23), CDI (n=13), UC (n=19) and UCI (n=14) donors; and (C) % V $\delta$ 1<sup>+</sup> of V $\delta$ 2<sup>neg</sup>  $\gamma\delta$  T cells isolated from control (n=14), CD (n=11), CDI (n=9), UC (n=8) and UCI (n=12) donors. Kruskal-Wallis test with Dunn’s correction against control donors was used for analysis, \* $p$ <0.05; \*\*\*\* $p$ <0.0001. (D). Flow cytometric analysis of % V $\gamma$ 4<sup>+</sup> cells of V $\delta$ 2<sup>neg</sup> colonic  $\gamma\delta$  T cells from control (n=27), CD (n=15), CDI (n=7), UC (n=7), UCI (n=8) biopsies. Red dots – BTNL genotype discussed subsequently. (E). Flow cytometric analysis of % V $\gamma$ 4 cells of CD103<sup>+</sup> V $\delta$ 2<sup>neg</sup> colonic  $\gamma\delta$  T cells from paired uninflamed and inflamed tissue from IBD donors sampled at the same endoscopy (n=5). Connected symbols represent individual participants. Open circles – CD donor; Filled circles – UC donors. Ratio paired  $t$ -test, \* $p$ <0.05. (F). *TRGV2*, *TRGV3*, *TRGV4* gene usage as proportion of total *TRGV2/3/4* reads from TCR deep sequencing of whole biopsies mRNA (Ctrl n=13; IBD n=9, IBDI n=9). Open circles - CD donors; filled circles - UC donors. Twoway ANOVA with Dunnett correction against control donors was used for analysis, \* $p$ <0.05; \*\* $p$ <0.01. (G). Total  $\gamma\delta$ TCR (left) and *TRGV2/3/4* (right) diversity measured by normalized Shannon entropy derived from TCR deep sequencing of whole biopsies mRNA from control (n=13), IBD (n=9) and IBDI (n=9) donors. Open circles – CD donors; filled circles – UC donors. Two-way ANOVA with Dunnett correction

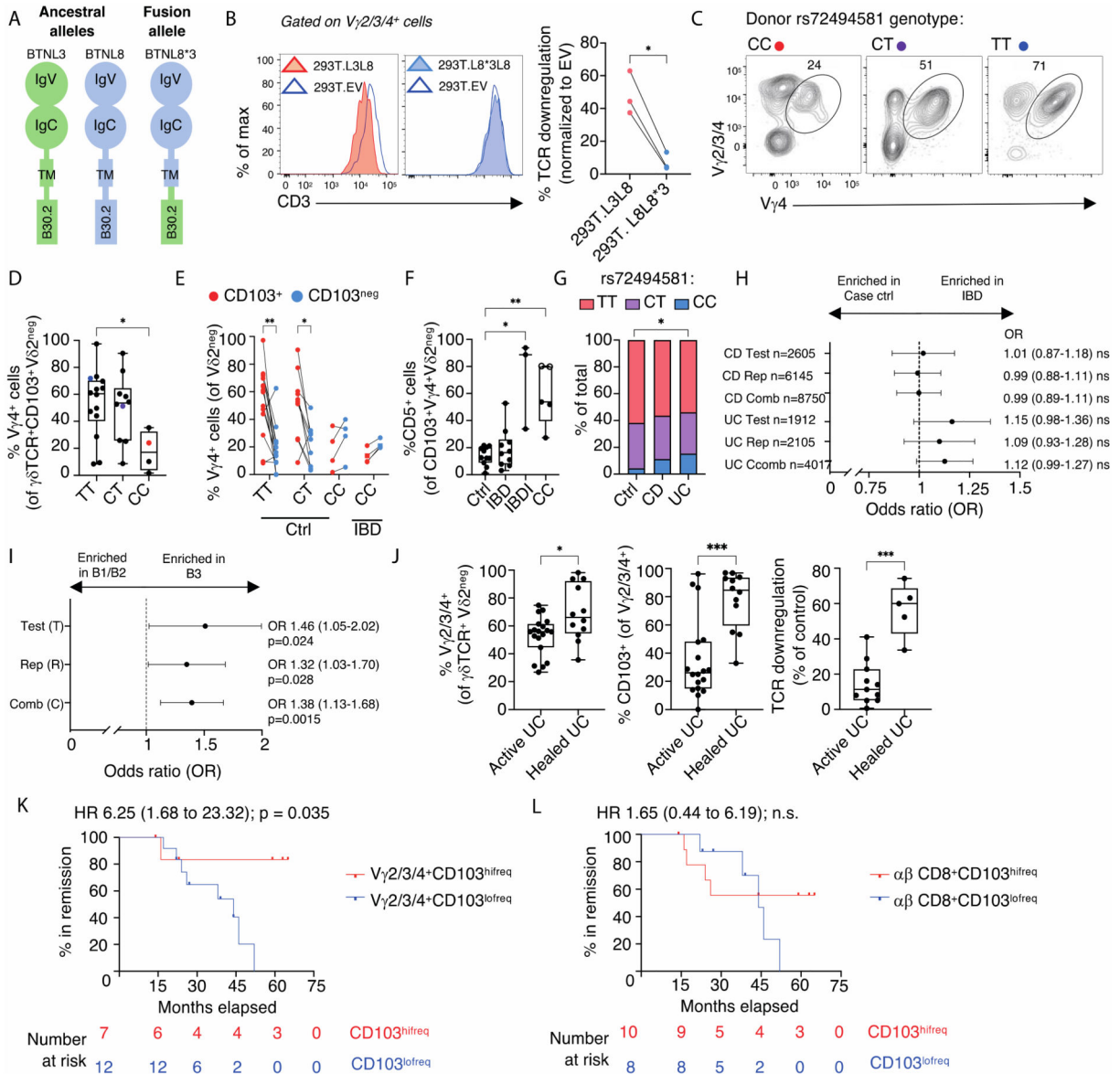
against control donors was used for analysis, \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ . **(H)**. Flow cytometric analysis of CD103 expression on V $\gamma$ 2/3/4<sup>+</sup> colonic  $\gamma\delta$  T cells derived from control (n=43), CD (n=22), CDI (n=11), UC (n=18), UCI (n=17) donors. Kruskal-Wallis test with Dunn's correction against control donors was used for analysis, \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ . **(I)**. Flow cytometric analysis of CD5 expression on CD103<sup>+</sup> V $\gamma$ 4<sup>+</sup> colonic  $\gamma\delta$  T cells derived from control (n=17), CD (n=10), CDI (n=5), UC (n=9), UCI (n=2) donors. Kruskal-Wallis test with Dunn's correction against control donors was used for analysis, \* $p < 0.05$ . **(J)**. Flow cytometric analysis of TNF $\alpha$  (left) and IFN $\gamma$  (right) expression by V $\gamma$ 2/3/4<sup>+</sup> colonic lymphocytes obtained from control (n=11), CD (n=12), CDI (n=6), UC (n=5) and UCI (n=4) donors following 4h of P+I stimulation. Kruskal-Wallis test with Dunn's correction against control donors was used for analysis, \* $p < 0.05$ . **(K)**. Flow-cytometric analysis of %TCR downregulation by human colonic V $\gamma$ 2/3/4<sup>+</sup> lymphocytes derived from control (n=34), CD (n=14), CDI (n=8), UC (n=10) and UCI (n=12) donors, after overnight co-culture with 293T.EV or 293T.L3L8 cells. Kruskal-Wallis test with Dunn's correction against control donors was used for analysis, \* $p < 0.05$ ; \*\*\* $p < 0.001$ . All box and whisker plots: median, IQR and min-max range shown. Symbols represent individual participants.+





**Figure 5. The cytokine milieu can affect the phenotype of intestinal  $V\gamma 4^+$  cells.** (A-D). Flow cytometric analysis of (A) CD103 (n=8), (B) TIGIT (n=8), (C) NKp46 (n=5), and (D) CD18 (n=8) expression on  $V\gamma 4^+$  colonic lymphocytes cultured for 7 days in IL2+IL15 (control); IL2+IL15+IL12+IL18 and IL2+IL15+IL1 $\beta$ +IL23. Summary data (left) and representative histograms (right) shown. Connected symbols represent individual participants. Repeated measures one-way ANOVA with Dunnett's correction against IL12+IL18 treatment was used for analysis, \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ . (E). Flow cytometric analysis of  $V\gamma 2/3/4^+$  colonic lymphocytes cultured for 7 days as in A-D and subsequently co-cultured overnight with 293T.EV and 293T.L3L8 cells (n=3). Representative flow plots of CD3 expression and summary graph (far right) of percent TCR downregulation relative to 293T.EV control are shown. Repeated measures one-way

ANOVA with Bonferroni correction, \* $p < 0.05$ . **(F-I)**. Flow cytometric analysis of (F) CD103, (G) NKp46, (H) CD18 + and (I) TIGIT expression on V $\gamma$ 4 colonic lymphocytes cultured for 7 days in IL2+IL15 (control) and IL2+IL15+TNF $\alpha$  (n=4). Summary data (left, centre) and representative histograms (right) shown. gMFI - geometric mean fluorescent intensity. Connected symbols represent individual participants. Paired *t*-test, \* $p < 0.05$ ; \*\* $p < 0.01$ .



**Figure 6. A genetic influence on the V $\gamma$ 4-BTNL axis.**

(A). Graphical representation of full length BTNL3, BTNL8 and BTNL8\*3 fusion protein structure. TM – transmembrane. (B). Flow cytometric analysis of human colonic V $\gamma$ 2/3/4<sup>+</sup> cells co-cultured overnight with 293T.EV, 293T.L3L8 and 293T.L8L8\*3 (n=3). Example flow plots of CD3 expression and summary graph (far right) of %TCR downregulation relative to 293T.EV control are shown. Connected symbols represent individual participants. Paired *t*-test, \**p*<0.05. (C). Example neg flow cytometry plots of % V $\gamma$ 4<sup>+</sup> cells of CD103<sup>+</sup> V $\gamma$ 2<sup>neg</sup>  $\gamma\delta$  T cells obtained from control biopsies. rs72494581 status indicated above plots. Coloured dots correspond to data points neg displayed in panel D. (D). Flow cytometric analysis of % V $\gamma$ 4<sup>+</sup> cells of CD103<sup>+</sup> V $\gamma$ 2<sup>neg</sup>  $\gamma\delta$  T cells according to their rs72494581 SNP

genotype in control donors. CC (n=4), CT (n=10), TT (n=14). Median, IQR and min-max range shown. Symbols represent individual participants. Kruskal-Wallis test with Dunn's correction against control donors was used for neg analysis, \* $p < 0.05$ . **(E)**. Flow cytometric analysis of %  $V\gamma 4^+$  cells (y axis) of  $CD103^+ V\gamma 2^{neg}$  (red) or  $CD103^{neg} V\gamma 2^{neg}$  (blue) populations in control donors with different rs72494581 genotype (x axis) and IBD donors with the CC genotype. Square – UC donor. CC Ctrl (n=4), CC IBD (n=4), CT Ctrl (n=10), TT Ctrl (n=14). Connected symbols represent individual participants. Wilcoxon matched-pair test, \* $p < 0.05$ ; \*\* $p < 0.01$ . **(F)**. Flow cytometric analysis of %  $CD5^+$  cells of  $V\gamma 4^+ CD103^+ V\delta 2^{neg}$  lymphocytes from control (n=11), IBD (n=10) and IBDI (n=3) TT and CT donors, and control (n=3) and uninflamed IBD (n=2, open circles) CC donors. Median, IQR and min-max range shown. Symbols represent individual participants. Kruskal-Wallis test with Dunn's correction against control donors was used for analysis, \* $p < 0.05$ ; \*\* $p < 0.01$ . **(G)**. Frequency of the rs72494581 genotype within the local cohort (ctrl n=115 ; CD n=64 ; UC n=52). Fisher's exact test, \* $p < 0.05$ . **(H)**. Odds ratios (OR) for development of CD or UC compared to 1950s birth cohort data. OR (95% confidence intervals) presented according to the minor allele (CC) and risk of developing disease. **(I)**. OR for development of B3 versus B1/B2 CD. OR (95% confidence intervals) presented according to the minor allele (CC) and risk of developing disease complications (T n=1817; R n=4491; C n=6308). H,I. Test - test set; Rep – replication set; Comb – combined set. **(J)**. Flow cytometric analysis of: percent  $V\gamma 2/3/4^+$  cells (left); %  $CD103^+ V\gamma 2/3/4^+$  cells (middle); and % TCR downregulation by  $\gamma\delta$  lymphocytes after overnight co-culture with 293T.EV or 293T.L3L8 cells (right) in the same donors. Colonic biopsies were obtained from inflamed (n=11-19) and previously inflamed (healed) areas (n=5-12). Median and inter-quartile range shown (box) with minimum and maximum values (whiskers). Symbols represent individual participants. Mann-Whitney U test, \* $p < 0.05$ ; \*\*\* $p < 0.001$ . **(K,L)**. Hazard ratio (HR) analysis of IBD patients in remission with intestinal (K)  $CD103^{highfreq} V\gamma 2/3/4^+$  cells (CD n=3; UC n=4) and  $CD103^{lowfreq} V\gamma 2/3/4^+$  cells (CD n=8; UC n=4); and (L)  $CD103^{highfreq} CD8^+ \alpha\beta$  T cells (CD n=4; UC n=6) and  $CD103^{lowfreq} CD8^+ \alpha\beta$  T cells (CD n=6; UC n=2) in colonic biopsies at time of sampling. Subjects followed until the end of the study or relapse. Log rank test, Hazard ratio log rank, ns – not significant.