



ARTICLE

Single-cell TCR sequencing of gut intraepithelial $\gamma\delta$ T cells reveals a vast and diverse repertoire in celiac diseaseLinn M. Eggesbø¹, Louise F. Risnes^{1,2}, Ralf S. Neumann¹, Knut E. A. Lundin^{1,3}, Asbjørn Christophersen¹ and Ludvig M. Sollid^{1,2}

A hallmark of celiac disease (CeD), a chronic condition driven by cereal gluten exposure, is increase of gut intraepithelial $\gamma\delta$ T cells. This may indicate pathogenic involvement of $\gamma\delta$ T cells and existence of disease-specific $\gamma\delta$ T-cell receptors (TCRs) recognizing defined antigen(s). We performed high-throughput and paired $\gamma\delta$ TCR sequencing of single intraepithelial $\gamma\delta$ T cells of untreated CeD patients ($n = 8$; 1821 cells), CeD patients treated with a gluten-free diet ($n = 5$; 436 cells) and controls ($n = 7$; 1068 cells). We found that CeD patients, both untreated and treated, had larger and more diverse $\gamma\delta$ TCR repertoires, more frequent usage of *TRDV1* and *TRDV3* and different patterns of TCR γ /TCR δ -pairing compared with controls. Although we observed no public CDR3 δ sequences, there were several public CDR3 γ sequences—many of which were shared by not only the CeD patients, but also by the controls. These public CDR3s were characterized by few N/P nucleotide insertions with germline and near-germline configuration, hence being easy to generate. Previous findings of CeD-specific CDR3 motifs were not replicated. Thus, being unable to raise evidence for CeD-specific $\gamma\delta$ TCRs in this first large, paired $\gamma\delta$ TCR single-cell sequencing study, we project challenges for identification of CeD-relevant $\gamma\delta$ TCR ligands.

Mucosal Immunology (2020) 13:313–321; <https://doi.org/10.1038/s41385-019-0222-9>

INTRODUCTION

Celiac disease (CeD) is a prevalent and chronic inflammatory disease of the small intestine that is driven by ingestion of gluten. The disease has a strong HLA-class II association, and clonally expanded CD4⁺ T cells specific for gluten epitopes presented by the disease-associated HLA-DQ2 (DQ2.5 and DQ2.2) or HLA-DQ8 allotypes are found in blood and in gut of CeD patients,^{1,2} but not from controls.^{3,4} The histological manifestation of the CeD lesion is characterized by intraepithelial lymphocytosis (IELs), crypt hyperplasia, and loss of villi.⁵ The IELs in humans are mainly CD8⁺ T cells expressing the $\alpha\beta$ T-cell receptor (TCR) as well as T cells expressing the $\gamma\delta$ TCR.⁶ In untreated CeD, both these types of IELs increase in density with a proportionally larger increase in the $\gamma\delta$ TCR subset. After initiation of a gluten-free diet (GFD), the only available treatment for CeD, the numbers of gluten-specific CD4⁺ T cells in lamina propria and CD8⁺ $\alpha\beta$ TCR IELs rapidly decrease, whereas the elevated fraction of intraepithelial $\gamma\delta$ TCR IELs is reduced at a much slower rate than CD8⁺ $\alpha\beta$ TCR IELs.⁷

There has been a wealth of studies of TCR repertoires of $\alpha\beta$ T cells, but less so on TCR repertoires of $\gamma\delta$ T cells. Most studies of $\gamma\delta$ T cells have focused on the $\gamma\delta$ TCR repertoire from peripheral blood in healthy settings or related to hematopoietic stem cell transplantation.^{8–10} These studies have shown that there is a postnatal reduction in diversity of the $\gamma\delta$ repertoire^{8,10} and that there is a quicker reconstitution of $\gamma\delta$ T cells compared with $\alpha\beta$ T cells after transplantation.⁸ Furthermore, some studies have demonstrated that the V-gene usage of $\gamma\delta$ T cells appear to be tissue-specific, for instance, in peripheral blood there is a biased usage of *TRGV9* and *TRDV2*.^{8,11,12} Studying human colonic $\gamma\delta$ T cells, Di Marco Barros and colleagues found that

the V-gene usage in these T cells are skewed toward usage of *TRGV4* and *TRDV1*.¹³

Studies of the small-intestinal $\gamma\delta$ T-cell repertoire are scarce,^{14–16} and even more so in the context of CeD.^{17,18} A study by Han et al.¹⁹ showed that activated gut-homing $\gamma\delta$ T cells and CD8⁺ $\alpha\beta$ T cells increase in the blood of CeD patients on day 6 after gluten challenge, along with gluten-specific CD4⁺ T cells. For these $\gamma\delta$ T cells, they reported an enrichment of two CDR3 δ motifs among the *TRDV1*⁺ cells. Furthermore, Mayassi et al.²⁰ demonstrated that, in the small intestine of untreated CeD, interferon- γ producing *TRDV1*⁺ $\gamma\delta$ T cells replace the *TRDV1*⁺ $\gamma\delta$ T cells expressing natural cytotoxicity receptors. The authors reported that a TCR motif, an H-J1 motif in CDR3 γ , was highly enriched in untreated CeD, indicating that this H-J1 motif may be linked to recognition of a disease-relevant ligand.

In this study we performed high-throughput TCR sequencing of single $\gamma\delta$ IELs, to investigate the $\gamma\delta$ IEL repertoire in CeD patients and controls. As these cells are located at the site of inflammation in CeD, they should harbor the highest density of disease-relevant $\gamma\delta$ T cells. Our single-cell sequencing method enabled us to sequence the vast majority of all cells sorted from each individual. We found that there are clear differences in the repertoires between CeD patients and controls and that the disease-specific changes are present in well-treated patients. Moreover, we found an increased diversity in the CeD groups compared with controls. We observed many TCR γ sequences being public, i.e., being expressed in multiple subjects, and also a few public sequences that were unique to CeD patients. However, these sequences, as most of the other public TCR γ sequences, had germline or near-germline configuration, suggesting that they are easy to generate

¹K. G. Jebsen Centre for Coeliac Disease Research, University of Oslo, 0424 Oslo, Norway; ²Department of Immunology, University of Oslo and Oslo University Hospital-Rikshospitalet, 0372 Oslo, Norway and ³Department of Gastroenterology, Oslo University Hospital-Rikshospitalet, 0372 Oslo, Norway
Correspondence: Linn M. Eggesbø (l.m.eggesbo@medisin.uio.no) or Ludvig M. Sollid (l.m.sollid@medisin.uio.no)

Received: 11 August 2019 Revised: 1 October 2019 Accepted: 22 October 2019

Published online: 14 November 2019



and unlikely to be CeD-specific. Along with a failure to replicate the above-mentioned CeD-associated CDR3 δ and CDR3 γ motifs, we are unable to provide strong evidence for an existence of CeD-specific $\gamma\delta$ TCRs. Thus, although bold changes in $\gamma\delta$ T-cell repertoires occur as a consequence of CeD, whether this is a primary or secondary phenomenon in the disease remain unclear.

RESULTS

Establishment of a single-cell paired $\gamma\delta$ TCR sequencing platform
In order to characterize the $\gamma\delta$ TCR repertoire with high-throughput sequencing, we established a single-cell multiplexed PCR method based on the same principles as for $\alpha\beta$ TCR sequencing²¹ by modifying previously published *TRDV* primers¹⁹ and designing novel *TRGV* primers (see methods sections for details). We sorted and sequenced single $\gamma\delta$ TCR IELs from CeD patients and controls (gating strategy in Fig. S1). We sorted in total 5083 single $\gamma\delta$ T cells in a 96-well plate fashion. We obtained *TRG* and/or *TRD* sequences from 4334 wells giving a sorting efficiency of 85%. Differences in PCR efficiency of the *TRG* (98%) and *TRD* (79%) genes resulted in a paired sequence information efficiency of 77% (3325 cells). Only wells containing at least one *TRG* and at least one *TRD* sequence were considered valid for clonotype analysis. Of these valid cells, 23% (776/3325) had dual productive TCR γ receptors and 1.2% (39/3325) had dual productive TCR δ receptors. After collapsing of expanded clones into unique T-cell clonotypes, 1542 clonotypes remained.

The intraepithelial TCR γ and TCR δ chain usage is altered in CeD. In concordance with previous reports,^{20,22} we found that untreated CeD patients have increased frequency of intraepithelial $\gamma\delta$ T cells in the small intestine that stay high also in patients who have been on a GFD for many years ($p = 0.008$) (Fig. 1a). Similar to previous reports in healthy subjects,¹³ we also found that V δ 1 (*TRDV1*) is the most prevalent V gene used in the gut, both in CeD patients and in controls (Fig. 1b). Unlike in controls where *TRDV1* and *TRDV2* were more frequently used, *TRDV1* and *TRDV3* were enriched both in treated and untreated CeD patients (Fig. 1b). Notably, we found that the *TRDV* usage was highly similar between the untreated and treated CeD groups. In agreement with previous findings,¹³ we found a preference for *TRGV4* in controls (Fig. 1c). In parallel with altered *TRDV* usage, we found that CeD patients have an altered *TRGV* usage, and this was observed for both CeD groups. Although the *TRDV* and *TRGV* segment usage differed somewhat between donors, the usage was much more similar between the two groups of CeD patients than between CeD patients and controls (Fig. S2A, Fig. S2B). As observed by others,¹³ we also found a preferred pairing between *TRGV4* and *TRDV1* in controls (Fig. 1d, left plot). Interestingly, we found that the bias for this pairing was weaker in treated and untreated CeD, and that there was less chain pairing preference overall (Fig. 1d). Taken together, our findings demonstrate that the V-gene usage is altered in CeD patients regardless of treatment status.

Further, we looked into the J gene usage and found that the *TRDJ* usage was near identical between the three groups, with few inter-individual differences (Fig. S3A, Fig. S3B). We found an enrichment of *TRGJP1* and diminished usage of *TRGJP* in both celiac groups compared with controls (Fig. S3C). As with the *TRDJ* usage, there was overall similar *TRGJ* chain usage profiles at the donor level (Fig. S3D). To further investigate the profile of the $\gamma\delta$ repertoire, we looked at the distribution of CDR3 lengths for both γ and δ sequences (Fig. 1e–f). We found that $\gamma\delta$ IELs from CeD patients tend toward having longer CDR3 δ sequences than the controls, albeit not statistically significant (Fig. 1f). Overall, these data demonstrate a smaller degree of *TRGV4/TRDV1* pairing in $\gamma\delta$ IELs in the gut of untreated CeD patients which also remains in treated CeD.

CeD patients have a more diverse repertoire

Next, we analyzed the clonal expansion of the $\gamma\delta$ IELs from the CeD patients and controls. To visualize the quantitative relation between highly expanded and less expanded clones, we divided the repertoire into three parts, i.e., the 10 most expanded clones (top 10 clones), clones 11:100 and clones 101:1000. The clonal proportion between the three participant groups differed in that the top 10 clones took up less space of the total repertoire from CeD patients, and this was especially the case in untreated CeD (Fig. 2a). When focusing on the top 10 clones within each participant group, we observed that the largest (top 1) clonotype in the untreated CeD repertoire was significantly smaller than the largest clonotype in the control repertoires ($p = 0.0022$) (Fig. 2b). However, this was in large part because some of the controls had a nearly monoclonal $\gamma\delta$ IEL repertoire (Fig. S4). Next, we determined the clonal diversity in the repertoires by estimating Shannon diversity index. Interestingly, the untreated CeD group was shown to have a higher clonal diversity compared with the controls ($p = 0.0006$) (Fig. 2c). A similar trend was observed for the treated CeD group, however, it did not reach significance likely owing to the smaller group of treated patients ($p = 0.07$). These data show that CeD patients have a more diverse intraepithelial $\gamma\delta$ TCR repertoire as compared with controls.

Public $\gamma\delta$ TCR sequences

Sharing of TCR repertoires between individuals may reflect a response toward common antigens. When determining whether any CDR3 γ or CDR3 δ sequences were shared between two or more donors on the amino-acid level, we observed no shared CDR3 δ sequences but a notable fraction of the CDR3 γ sequences, also with identical *TRGV*, were shared (Fig. 3a). This supports the notion that the TCR γ repertoire has a higher degree of public sequences and that the TCR δ repertoire is unique for each individual.⁸ Although there was sequence sharing between all participant groups, most of the shared sequences were observed between either treated or untreated CeD, or within the untreated group alone (Fig. 3a). However, existence of public sequences was not more prominent among CeD patients, as the percentage of shared sequences over the total number of clonotypes sampled from each group, was ~20% in each group (Fig. 3b). Importantly, the degree of sharing will likely be restricted by the size of repertoire interrogated. We thus compared sharing of our sequences with a compilation of other published $\gamma\delta$ TCR with CDR3 amino-acid sequences.^{8–10,20,23,24} Although there was some variation between the individual subjects (Fig. S5), the average degree of sharing was much larger (~40%) (Fig. 3c). Again, the average degree of sharing was similar for each subject group (Fig. 3c).

We further analyzed sharing of sequences in our data set at the nucleotide level. Of 87 public amino-acid TCR γ sequences, 45 of these were also identical at the nucleotide level (supplementary Table 1). Strikingly, these shared nucleotide sequences were hallmarked by having few N/P nucleotide insertions (Fig. 3d), thus often being in germline or near-germline configuration. Therefore, many of the public sequences would consequently be easy to generate.

We found one CDR3 δ sequence that was reported in two other publications.^{10,20} Interestingly, this CDR3 δ sequence is only 10 amino acids long, which is one of the shorter CDR3 δ sequences in our data (see CDR3 δ length distributions depicted in Fig. 1f). Notably, as observed with shared TCR γ sequences, this TCR δ sequence was near-germline in configuration lacking a diversity (D) segment and with only three inserted N/P-nucleotides. Taken together, our data provide evidence that expanded $\gamma\delta$ IELs from the small intestine of CeD patients consist of TCR γ chains that are commonly found in human subjects.

Of general features of the TCRs worth noting, we observed some TCRs of the same individual having identical CDR3 that

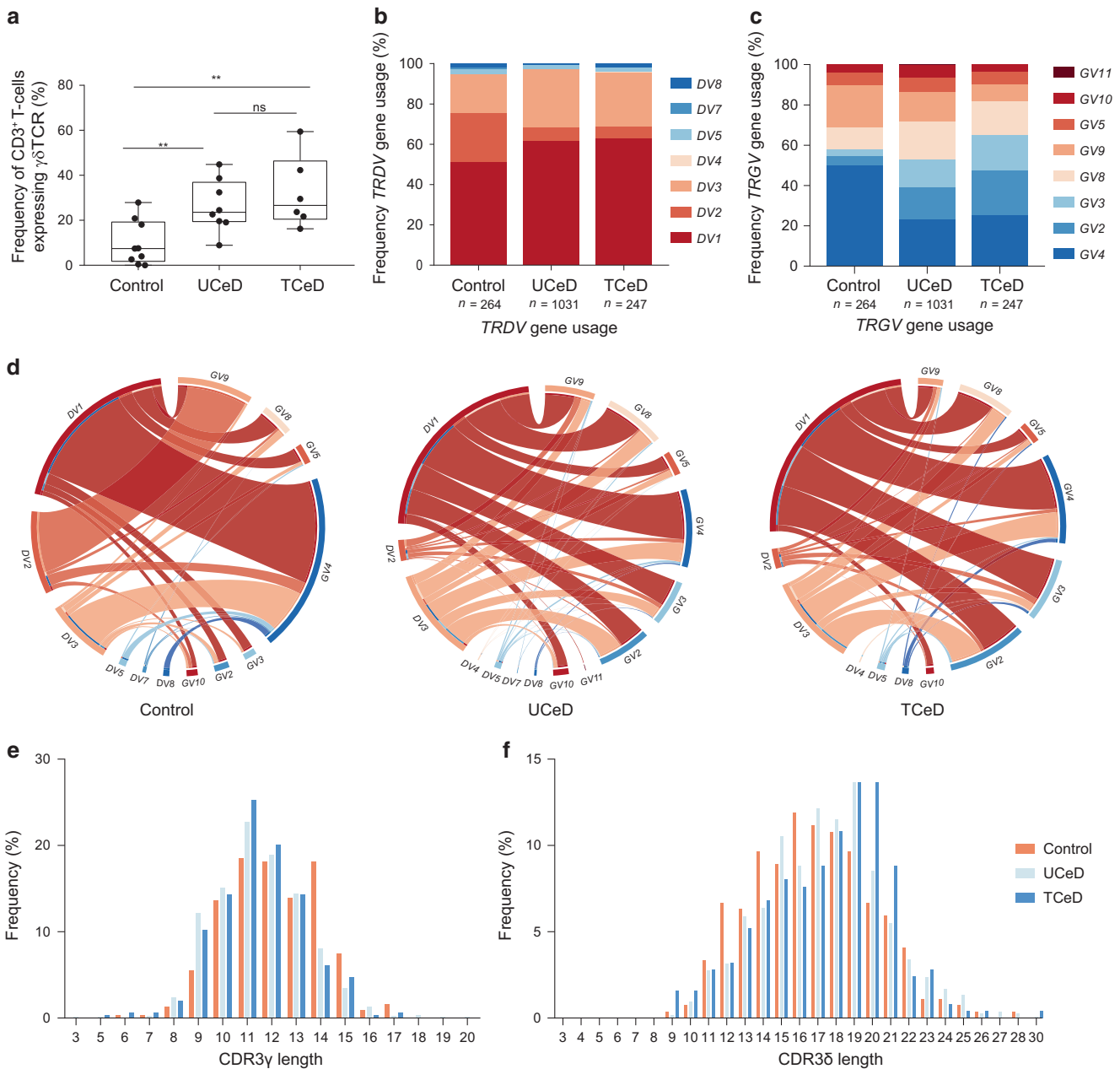


Fig. 1 Altered *TRDV* and *TRGV* chain usage in intraepithelial lymphocytes in CeD. **a** Boxplots depicting frequency of CD3⁺ T cells expressing the $\gamma\delta$ TCR in the duodenum as measured by flow cytometry in controls (controls, $n = 9$), untreated CeD (UCeD, $n = 8$), and treated CeD (TCeD, $n = 5$). **b, c** The frequency of *TRDV* usage **b** and *TRGV* usage **c** in duodenal $\gamma\delta$ T-cell clonotypes. **d** Chain pairing of *TRGV* and *TRDV* are displayed as chord diagrams where ribbons connecting chains indicate frequency of pairing (controls, $n = 264$, UCeD, $n = 1031$ and TCeD, $n = 247$ clonotypes). **e, f** The different lengths of CDR3 γ **e** and CDR3 δ **f** are shown as frequencies (%). For **b–f**, data are presented for controls (controls, $n = 7$), untreated CeD (UCeD, $n = 8$), and treated CeD (TCeD, $n = 5$). P value ** = $P < 0.01$, ns not significant

paired with different TCR δ , and as such these were defined as distinct clonotypes (see supplementary table 2). Further, we observed cases of shared sequences with identical CDR3 γ that occurred in sequences harboring different V genes.

Analysis of CeD-associated sequence motifs

Two previous reports have indicated CeD-associated CDR3 γ and CDR3 δ motifs. Han et al.¹⁹ reported CDR3 δ motifs found among expanded blood $\gamma\delta$ T cells on day six following a gluten challenge, whereas Mayassi et al.²⁰ reported a CDR3 γ motif found among gut *TRDV1*⁺ IELs. We wanted to explore the frequency of these motifs in our data set.

The so-called H-J1 motif reported by Mayassi et al.²⁰ consists of histidine (H) of the CDR3 γ paired with *TRGJ1*. We identified the *TRDV1*⁺ repertoire in our data set and investigated presence of the H-J1 motif in this subset. Owing to the inability to distinguish *TRGJ1*02* from *TRGJ2*01* we also included these as *TRGJ1* chains in the analysis. We found that there was a slight but not statistically significant increase in H-J1 motif expression among the untreated and treated CeD groups compared with controls (Fig. 4a, left). Although there were wide inter-individual differences with regard to the expression of the H-J1 motif in our groups, we found the level of H-J1 expression never exceeded 14% in any CeD patient. This contrasts the findings of Mayassi et al. who reported a

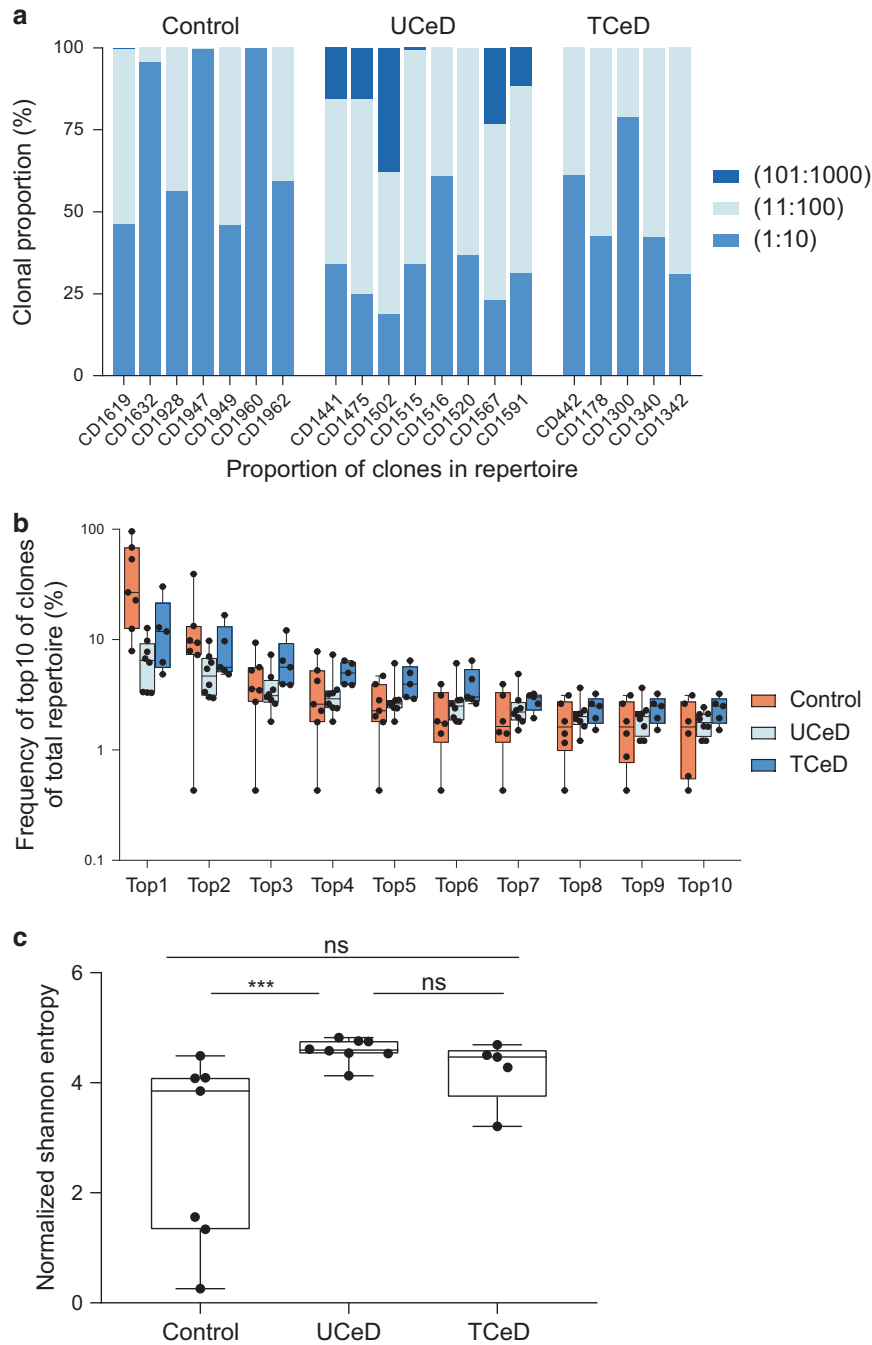


Fig. 2 CeD patients display a more diverse repertoire. **a** The degree of clonal expansion for each group is depicted as clonal proportions of top 1000 clonotypes into three groups, 1:10 (blue), 11:100 (light blue), and 101:1000 (dark blue). **b** Frequency of the top 10 clonotypes of the total clonotypes is shown as boxplots for each group, and each patient within the groups is indicated by dots. **c** Clonal diversity is estimated by normalized Shannon entropy (normalized by calculating the average Shannon entropy of a large number of random subsamples of a uniform size). P value *** = $P < 0.001$, ns not significant

frequency of up to 80% in some individuals. When extending the analysis to include only $TRDV1^-$ clonotypes, the H-J1 motif was found in all groups at low frequencies, not reaching statistical significance for any comparisons (Fig. 4a, middle). We repeated the analysis for the H-J1 motif within all $\gamma\delta$ IELs and found that the motif was not enriched or significant among untreated CeD patients when compared with controls (Fig. 4a, right). We also looked into the presence of this motif in clonally expanded cells of CeD patients. Of note, the few clonotypes harboring the H-J1 motif were not present among the most expanded clonotypes

(Fig. 4b), and importantly, this motif did not appear to be restricted to the CeD patients (Fig. 4a).

The CDR3 δ motifs identified by Han et al.¹⁹ were the sequences CxxxxxxxYWGI (YWGI) and CxxxxPxLGD (PxLGD). In the CeD groups, the YWGI motif was observed among $TRDV1^+$ clones, (Fig. 4c, left). The motif was also observed in controls, but to a lesser degree than that of the CeD groups; albeit not reaching statistical significance. Assessing all $TRDV$ genes, we found that the motif was present in both CeD groups as well as in the controls (Fig. 4c, right). Further, in $TRDV1^+$ cells, the PxLGD motif was

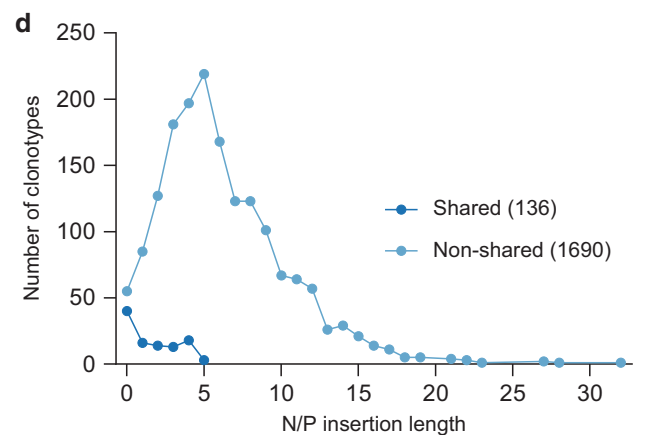
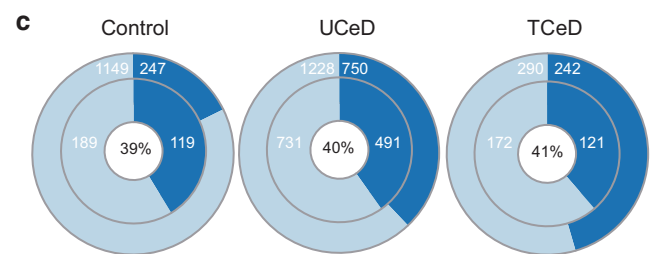
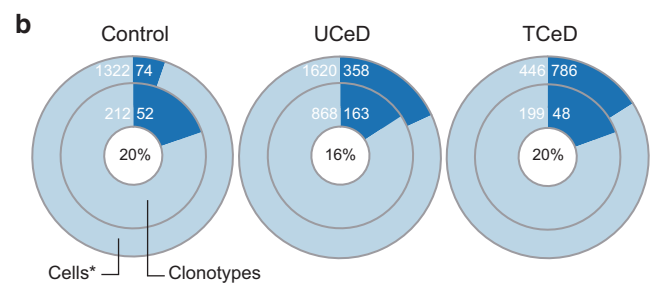
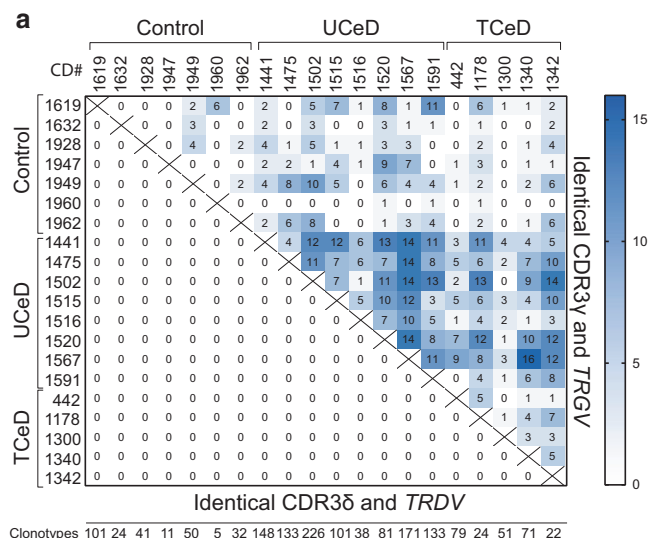


Fig. 3 Sharing of TCR γ and TCR δ sequences. **a** Sharing of TCR γ and TCR δ sequences, within and between the subjects of the three groups, is displayed as a heatmap. A clonotype is defined as a unique pair of TCR γ and TCR δ . Consequently, identical TCR γ sequences are defined as different clonotypes if paired with different TCR δ sequences. **b** Sharing of TCR γ sequences at the amino-acid level, within and between the three groups depicted in doughnut charts. Dark blue indicates shared cells/clonotypes, whereas light blue indicates non-shared cells/clonotypes. The outer circle shows cells*, and the middle circle shows clonotypes. The central circle gives the frequency of shared clonotypes over total number of clonotypes. **c** As in **b**, but showing sharing of TCR γ cells/clonotypes with sequences in previously reported studies. **d** Length of N/P nucleotide insertions in shared and non-shared TCR γ sequences. *A cell with dual TCR γ is counted as two distinct cells, resulting in a higher number than total cells analyzed

H-J1 motif in untreated CeD patients, nor corroborate any significant CeD-specific increase in the YWGI and PxLGD motifs.

Performing a global search for sequences in our material shared only between CeD patients also including for comparison sequences of CeD patients reported Mayassi et al.²⁰ we identified altogether nine sequences (Table 1). Notably, TCRs with these apparently CeD-associated TCR γ sequences in our data set all paired with different CDR3 δ sequences of varying lengths. Moreover, all nine sequences had few N/P nucleotide insertions, similar to the shared nucleotide sequences within our data set (Fig. 3d). This suggests that if tested against an even larger TCR repertoire, they may no longer be CeD specific.

DISCUSSION

The density of both TCR $\alpha\beta$ and $\gamma\delta$ IELs are increased in CeD,²² and although the TCR $\alpha\beta$ IEL population frequency goes down when gluten is removed from the diet, the $\gamma\delta$ IELs do so to a lesser degree.^{7,25} As there have been relatively few studies addressing the role of $\gamma\delta$ IELs in CeD, we sought to explore the repertoire of these cells in CeD patients. Using single-cell sorting combined with high-throughput sequencing, we sequenced the $\gamma\delta$ TCR genes in three groups of participants and our method allowed for successful sequencing of the majority of single-sorted cells from each patient.

In our study, we confirm the observation of Mayassi et al.²⁰ who found that the prevalent *TRGV4/TRDV1* pairing of the normal state small intestine¹³ is lost in both treated and untreated CeD. Our findings support the notion that in inflammatory conditions, $\gamma\delta$ IELs that are normally present in the gut mucosa are replaced by new $\gamma\delta$ IELs expressing other *TRGV* and *TRDV* gene segments. Although Mayassi et al.²⁰ reported a frequency of up to 80% of $\gamma\delta$ T cells expressing the H-J1 motif and suggested that cells with this motif are driven by gluten, we were not able to corroborate this finding as we found that this motif never exceeded 14% in any CeD patient. This disparity in findings may possibly relate to how the cell sorting was performed and sequencing data were generated. In the Mayassi et al. study, the authors bulk sorted the IELs and generated cDNA introducing a template-switch oligo during the cDNA synthesis followed by subcloning of amplicons and Sanger sequencing of individual plasmids. By this procedure, the sequence coverage may be too small to cover a true representation of a diverse IEL $\gamma\delta$ TCR repertoire. By contrast, our high-throughput, single-cell study design allowed us to sequence most of the sorted cells and enabled us to identify TCR $\gamma\delta$ chain pairing. Importantly, sequencing single cells allowed for more-precise measurement of clonal expansion. Our study points to a larger degree of clonal diversity in untreated CeD as opposed to the controls. Furthermore, the H-J1 motif was not

present in very low frequencies in the untreated CeD group but not at all in the treated CeD group or the controls (Fig. 4d, left). Again, when including all *TRDV* genes, we found the motif to be present in low frequencies, in untreated CeD and in controls, but not in treated CeD patients (Fig. 4d, right). Overall, our data neither replicate the previously reported major enrichment of the

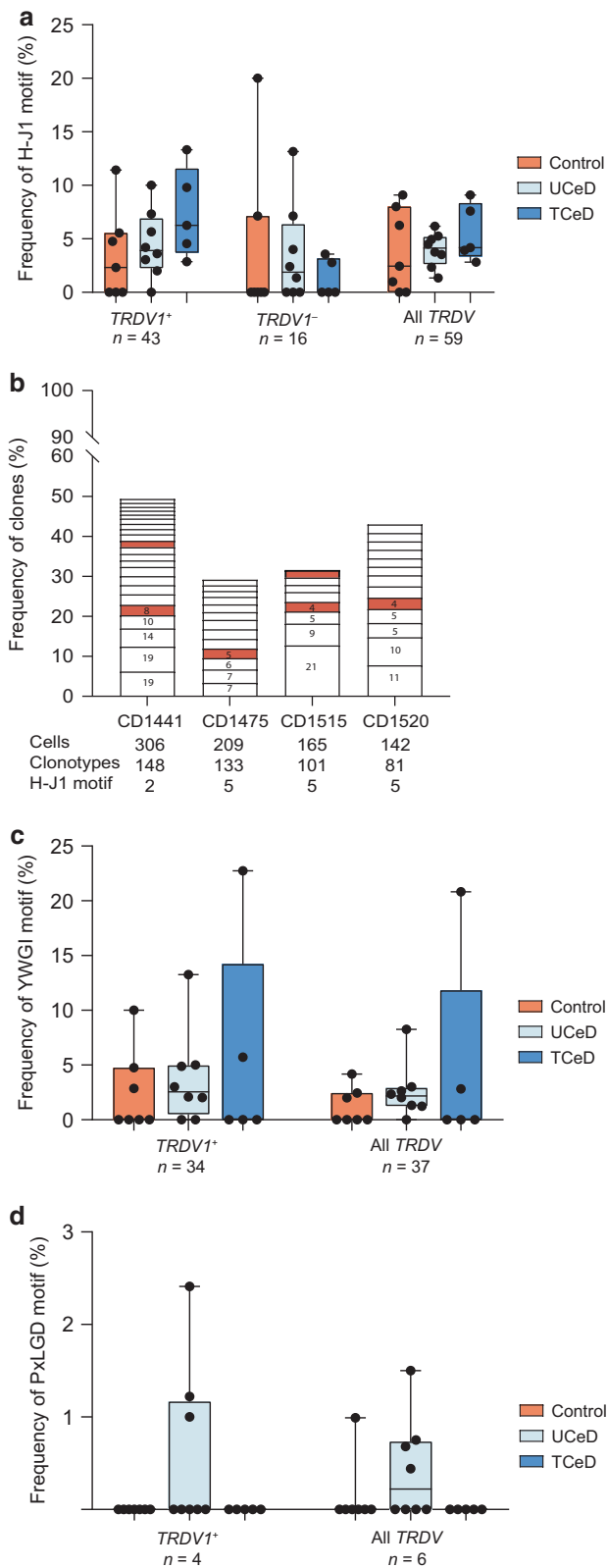


Fig. 4 Reported CDR3 motifs. **a** Frequency of H-J1 motif in the CDR3 γ of $\gamma\delta$ T cells pairing with TRDV1⁺ (left), TRDV1⁻ (middle) or all TRDV (right) segments. **b** Presence of H-J1 motif in clonally expanded cells of four untreated CeD patients is denoted in red. Percentages indicate the frequency of the repertoire that is made up of clonotypes that have an expansion of at least three cells. Motifs observed in clonotypes with an expansion less than three are not visualized. **c** Frequency of YWGI motif in TRDV1⁺ (left), or all TRDV (right) cells. **d** Frequency of the PxLGD motif in TRDV1⁺ (left), or all TRDV (right) cells

suggest that the motifs identified by Han et al. and the H-J1 motif are not CeD specific. Although it was suggested that one or few antigens drive the expansion of $\gamma\delta$ T cells with these reported sequences in CeD,^{19,20} this interpretation seems less likely based on our data.

A previous report, in line with our findings, concluded the CDR3 δ repertoire is mostly private.⁸ The literature is, however, conflicting with regards to how much of the TCR γ repertoire is public.^{8,9} We found that this repertoire is highly public with ~40% of the clonotypes using a public TCR γ sequence. Notably, however, all our shared TCR γ sequences paired with TCR δ chains with unique TCR δ s suggesting that these TCRs have different antigen specificities. Importantly, we observed that many of the public TCR γ sequences had germline or near-germline configuration thus likely explaining why they become public. Interestingly, the most shared TCR γ sequence in our data, (ATWDRPEKL) was listed as private sequence in a previous report.¹⁰ This conflicting finding likely relates to an issue of repertoire size when assigning sequences as public or private, as does the previous contradictory results on the publicness of the TCR γ repertoire. We identified nine public TCR γ sequences only observed in CeD patients. These could represent truly CeD-specific sequences, but their common feature of having germline or near-germline configuration, as observed for the shared nucleotide sequences (Fig. 3d), speaks to a high likelihood that they will be shared by CeD patients and controls alike.

Our study highlights the importance of defining true clonotypes based on paired TCR γ and TCR δ sequences rather than based on CDR3 or unpaired TCR sequences alone. This is so as identical TCR γ chains or TCR γ chains with identical CDR3 γ but different V genes can pair with several different TCR δ chains. This concern is particularly relevant for studies that involve tracking of T-cell clonotypes in time and space.

Our results suggest that the identification of $\gamma\delta$ T cells implicated in the pathogenesis of CeD based on sequence analysis will be difficult. On choosing a successful strategy to accomplish this endeavor, insight can be taken from work on gluten-specific CD4⁺ T cells—cells that are considered to be the drivers of the pathology of CeD.²⁶ Gluten epitopes for these cells were identified by generating T-cell clones from antigen reactive polyclonal lines, and the knowledge of the epitopes were then used to construct HLA-DQ:gluten tetramers thereby allowing direct visualization of the gluten-specific CD4⁺ T cells.²⁷ In CeD patients there is clonal expansion of gluten-specific CD4⁺ T cells including public TCRs. Yet, the low frequency of CD4⁺ T cells specific for one of the dominant gluten T-cell epitopes both in blood (1–10 per million CD4⁺ T cell) and gut mucosa (~1 per 100 CD4⁺ T cell)^{2,28} would make it very difficult to identify the gluten-specific T cells from repertoire analysis alone. A recent analysis combining mass cytometry with HLA-DQ:gluten tetramers revealed that the gluten-specific CD4⁺ T cells occupy a surprisingly distinct phenotype.²⁹ Of note, the t-SNE plots of the lamina propria CD4⁺ T cells also revealed large and general phenotypic differences between CeD subjects and controls, yet the gluten-specific T cells were distinct among the fairly globally and disease-associated change of phenotype among CD4⁺

present among the most expanded clonotypes. Looking into published CDR3 γ sequences, we found this motif present in $\gamma\delta$ T cells of adult blood and cord blood.^{8,10} Similarly, we found that the motifs reported by Han and colleagues¹⁹ are present in low frequencies in both CeD patients and controls. These data

Table 1. CDR3 γ sequences shared only between CeD patients

Donors	Group	CDR3 γ	TRGV	TRGJ	Clone size	Pairs with TRD	TRD VDJ-genes
CD1342	TCD	ALWEGDTTGWFKI	TRGV9	TRGJP1	1	ALGNLGRPWGTYTDKLI	TRDV3, TRDD2, TRDJ1
CD1567	UCD				4	ALGELLGLGDESKLI	TRDV2, TRDD3, TRDJ1
CD1502	UCD	ALWEVHGKKL	TRGV9	TRGJ1 2	2	ALGERRIGHWGIRTDKLI	TRDV1, TRDD3, TRDJ1
CD1515					3	ALGGNLFQLDNWGIDKLI	TRDV1, TRDD3, TRDJ1
CD1515	UCD	ALWEVLGKKIKV	TRGV9	TRGJP	1	AFSFFLLRPTGGYYTDKLI	TRDV1, TRDD2, TRDJ1
CD1591					1	ACDALLGDNTDKLI	TRDV1, TRDD2, TRDJ1
CD1342	TCD	ATWDILEKL	TRGV3	TRGJ1 2	1	AFRSSSLSRNVLGEGYTDKLI	TRDV1, TRDD3, TRDJ1
CD1516	UCD				2	AFIPLPTSWGIRTIYTDKLI	TRDV1, TRDD3, TRDJ1
CD1178	TCD	ATWDPGWFKI	TRGV2	TRGJ1 2	1	ASRSWSPRYWGISPYTDKLI	TRDV1, TRDD3, TRDJ1
CD442			TRGV8	TRGJP1	2	ALGILPTMELPWSPPRHDKLI	TRDV1, TRDD2, TRDJ1
CD1340	TCD	ATWDPNTTGWFKI	TRGV4	TRGJP1	2	AFRAGDYTDKLI	TRDV1, TRDD2, TRDJ1
CD1342			TRGV5		5	ALGESLLGVYTDKLI	TRDV1, TRDD3, TRDJ1
CD1520	UCD	ATWDRDSDWIKT	TRGV8	TRGJP2	1	ALGGSLRWGYGKALTAQLF	TRDV3, TRDD2, TRDJ1
ID113*	TCD				N/A	N/A	TRDV1
CD1591	UCD	AAWDLTTGWFKI	TRGV10	TRGJP1	1	ALGERPLGRYIQLGRLGDKFSWDTRQMF	TRDV1, TRDD2, TRDJ1
ID113*	TCD				N/A	N/A	TRDV1
CD1567	UCD	ATWDDATGWFKI	TRGV4	TRGJP1	1	AFTRFTGGPYTDKLI	TRDV1, TRDD3, TRDJ1
IDS1*					N/A	N/A	TRDV1

*Patients and CDR3 sequences from a recent CeD study²⁰

T cells. In other words, identification of gluten-specific T cells would not be possible by guidance of the global phenotypic changes seen among lamina propria CD4⁺ T cells of CeD. This is relevant for findings of Mayassi et al.²⁰ who found a global phenotypic shift in the gut $\gamma\delta$ T cells of CeD compared with controls. If there are particular $\gamma\delta$ T cells with defined TCR that are involved in the pathogenesis of CeD, it remains unclear whether these will be sharing the global phenotype typically seen in CeD-associated $\gamma\delta$ T cells.

In conclusion, our single-cell sequencing study has revealed that untreated CeD patients have a higher degree of diversity in the repertoire from $\gamma\delta$ TCR IELs and that these changes are also present in CeD patients on a gluten-free diet. Despite these CeD-specific findings, we are unable to firmly conclude on an existence of CeD-associated $\gamma\delta$ T-cell sequences. We identified a few apparently CeD-associated TCR γ sequences, but these may not be disease specific when the TCR γ repertoire is more extensively characterized. Our findings project that an identification of a prototype $\gamma\delta$ TCR specific for CeD, if it does exist, will be a challenge.

METHODS

Human material and sample preparation

All donors gave informed, written consent prior to sampling. The study was approved by the Regional Committee for Medical and Health Research Ethics South-East Norway (2010/2720, 2012/341, and 2013/1237). Patients donated 8–12 biopsy samples for research taken from the D3 position of the duodenum during upper gastroduodenoscopy. Patients that were undergoing gastroduodenoscopy for esophagus-related symptoms, not owing to gut inflammation, were included as controls. List of patient characteristics can be found in Table S3. Duodenal biopsies were collected in ice-cold RPMI-1640 media. The biopsies were transferred to a 2 mM EDTA, 2% fetal calf serum (FCS) in phosphate-buffered saline (PBS) solution and put on a rotating shaker at 37 °C for 2 × 10 min in order to remove the epithelial layer. This IEL fraction was then filtered through a 40- μ m cell strainer prior to cryopreservation of the cells.

Sample preparation and single-cell sorting

Frozen IEL cell suspensions were thawed and filtered through a 70- μ m cell strainer prior to staining with antibodies. The IELs were stained on ice in the dark for 20 min. The following antibodies were used: CD8-PerCP-Cy5.5 (clone SK1, BioLegend), CD8-PerCP (clone SK1, BioLegend), $\gamma\delta$ TCR-PE (clone 5A6.E9, Invitrogen), CD3-eVolve605 (clone OKT3, eBioscience), CD3-Superbright600 (clone OKT3, eBioscience) CD38-PE-Cy7 (clone HIT2, BioLegend), CD38-APC (clone HB7, eBioscience), epithelial antigen-FITC (clone Ber-EP4, DAKO), CD11c-Pacific Blue (clone B-ly6, BD Biosciences), CD14-Pacific Blue (clone M5E2, BioLegend), CD19-Pacific Blue (clone HIB19, BioLegend), CD56-Pacific Blue (clone MEM-188, BioLegend), CD103-APC (clone B-ly7, eBioscience), CD4-APH-H7 (clone SK3, BD Biosciences), CD39-PE-Cy7 (clone eBioA1, eBioscience). After staining, cells were washed with 0.5 mM EDTA, 2% FCS in PBS and sorted into 5 μ l of cell capture buffer (20 mM Tris-HCl pH8, 1% NP-40 in H₂O) in 96-well plates. Sorted cells were frozen on dry ice before being transferred to –80 °C. All cell sorting was performed on an Aria II Cell Sorter (BD Biosciences) at the Flow Cytometry Core Facility at Oslo University Hospital. All flow cytometry data was analyzed with the FlowJo Software (FlowJo LLC). Gating strategy is provided in supplemental Fig. 1.

Single-cell TCR sequencing using multiplex PCR

To obtain paired TCR γ and TCR δ sequences, we performed PCR with multiplexed primers covering all TRGV and TRDV genes. The primers for the TRDV and TRDC regions were adapted from Han et al.¹⁹ and likewise, the TRGV and TRGC primers were adapted from Guo et al.³⁰ The primers were modified to fit the format described in Han et al.²¹ and be compatible with Illumina sequencing. The modified primers are compatible with simultaneous TCR α and TCR β sequencing, using the protocol of Risnes et al.² List of primers used can be found in Table S4.

We sorted single cells into 96-well plates containing 5 μ l capture buffer (20 mM Tris-HCl pH8, 1% NP-40). The plates were stored at –80 °C until cDNA synthesis to facilitate cell lysis. For first-strand cDNA synthesis, we added 5 μ l cDNA mix (1 × SSII RT buffer, 1 mM dNTP, 2.5 mM DDT, 1 μ M oligo d(T) (5'-CTGAATCT₁₆-3'), 1 μ M reverse TRGC (5'-CCCAGAATCGTGTGCT-3') and TRDC (5'-GATGAC



AATAGCAGGATCAAAC-3') primers, 1.5 U/ μ l RNase Inhibitor, 2.5 U/ μ l Superscript II in final 10 μ l reaction volume). The cDNA synthesis was carried out at 42 °C for 50 min followed by an inactivation step at 72 °C for 10 min. The cDNA plates were stored at -20 °C. The first of the three nested PCR steps was carried out in a total volume of 25 μ l using all cDNA (10 μ l), whereas the second and third PCR steps was carried out in a total volume of 10 μ l using 1 μ l PCR template. Each step used KAPA HiFi HotStart ReadyMix (Kapa Biosystems). In the first of the nested PCR reactions the final concentration of each TRGV and TRGC primers was 0.0067 μ M and 0.04 μ M. In the second PCR reaction the concentration was 0.005 μ M and 0.04 μ M for the TRGV and TRGC primers, respectively. For the two first nested PCR reactions the final concentration of each TRDV and TRDC primers was 0.01 μ M and 0.08 μ M. In the final barcoding PCR step, we added 5'-barcoding primers (0.044 μ M) and 1:1 ratio of the 3'-barcoding primers, TRGC and TRDC (0.18 μ M). In addition, Illumina Paired-End primers were added to the master mix (0.5 μ M each). Primer sequences and cycling conditions for all three PCR reactions are provided in the original protocol.²¹ All sequencing was done using the Illumina MiSeq platform, performed at the Norwegian Sequencing Center. All TCR sequence data were deposited in the European Genome-Phenome Archive (accession number EGAS00001003897).

TCR repertoire analysis

Obtained reads from the Illumina sequencing was processed as described² with the exception that the gene-specific primers used were for TRD/TRG instead of TRA/TRB. In brief, we pre-processed using tools from the pRESTO toolkit³¹ followed by submitting the TCR sequences to the IMGT/HighV-Quest online tool.³² This online tool allowed for determination of the V, D, and J genes and alleles as well as the nucleotide sequences of the CDR3 junctions. The resulting output files were imported into an in-house Java program where they were subject to further filtration, as described.² Only productive sequences as determined by IMGT were included and sequences were collapsed if single cells had identical V gene, J gene, and CDR3 nucleotide sequences. Only sequences with at least 50 reads supporting them were included in the analysis. Only cells with a single TRD and TRG or dual chains, with a maximum of three chains in total, were considered for analysis. Cells that had identical TRD and TRG CDR3 nucleotide sequences, along with identical V genes and J genes (within the same individual), were considered belonging to the same clonotype.

Clonotypes with dual and different TCR γ were considered as a separate clonotypes. Different clonotypes were also assigned when identical TCR γ , but different TCR δ were observed. For the analysis of shared sequences, we analyzed a large data set kindly provided by Martin Davey,¹⁰ the sequence data from Mayassi et al.²⁰ as well as sequences of other papers.^{8,9,20,23,24} In this analysis, we cross-referenced our CDR3 γ and CDR3 δ clonotypes with the other data sets.

Treemaps were generated by exporting files from the in-house program and using the R package "Treemapify" version 2.5.3. Chord diagrams were generated using the online circos tool (http://circos.ca/circos_online). Other graphs were made in Graphpad Prism 8 (Graphpad Software Inc).

Statistics

The integrated statistical tools in Graphpad Prism 8 software (Graphpad Software Inc) were used for Figs. 1a, e, f, 2b, c, and 4a, c, d. For non-parametric data, unpaired Mann-Whitney tests were used. Fishers exact test was used contingency tables (determining statistics for chain usage, Fig. 1b). To ensure that diversity estimates were comparable across variably sized data sets, the diversity of each data set was normalized by calculating the average Shannon entropy of a large number of random subsamples of a uniform size.

ACKNOWLEDGEMENTS

We are grateful to all the individuals who provided biological material to this study. We thank S. Furholm, C. Hinrichs, and M.H. Bakke for collecting biological material, the nursing and endoscopy staff at the Endoscopy unit for performing the endoscopic procedures, and we thank Dr. Martin Davey for providing a $\gamma\delta$ TCR sequencing data set. This work has been supported by grants from Stiftelsen Kristian Gerhard Jebsen (project SKGJ-MED-017) and by the South-Eastern Norway Regional Health Authority (projects 2011050 and 2018068).

AUTHOR CONTRIBUTIONS

LMS, LME, AC, and LFR designed the study. LME and LFR collected patient material and acquired the data. LME, LFR, AC, and LMS analyzed the data and wrote the manuscript. RSN developed the bioinformatics tools and revised the manuscript. KEAL provided patient and donor material and revised the manuscript.

ADDITIONAL INFORMATION

The online version of this article (<https://doi.org/10.1038/s41385-019-0222-9>) contains supplementary material, which is available to authorized users.

Competing interests: The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Lundin, K. E. et al. Gliadin-specific, HLA-DQ(α 1*0501, β 1*0201) restricted T cells isolated from the small intestinal mucosa of celiac disease patients. *J. Exp. Med.* **178**, 187–196 (1993).
- Risnes, L. F. et al. Disease-driving CD4⁺ T cell clonotypes persist for decades in celiac disease. *J. Clin. Invest.* **128**, 2642–2650 (2018).
- Molberg, O. et al. Gliadin specific, HLA DQ2-restricted T cells are commonly found in small intestinal biopsies from coeliac disease patients, but not from controls. *Scand. J. Immunol.* **46**, 103–109 (1997).
- Christophersen, A. et al. Healthy HLA-DQ2.5+ subjects lack regulatory and memory T cells specific for immunodominant gluten epitopes of celiac disease. *J. Immunol.* **196**, 2819–2826 (2016).
- Jabri, B. & Sollid, L. M. Tissue-mediated control of immunopathology in coeliac disease. *Nat. Rev. Immunol.* **9**, 858–870 (2009).
- Chang, F., Mahadeva, U. & Deere, H. Pathological and clinical significance of increased intraepithelial lymphocytes (IELs) in small bowel mucosa. *APMIS* **113**, 385–399 (2005).
- Calleja, S. et al. Dynamics of non-conventional intraepithelial lymphocytes-NK, NKT, and $\gamma\delta$ T-in celiac disease: relationship with age, diet, and histopathology. *Dig. Dis. Sci.* **56**, 2042–2049 (2011).
- Ravens, S. et al. Human $\gamma\delta$ T cells are quickly reconstituted after stem-cell transplantation and show adaptive clonal expansion in response to viral infection. *Nat. Immunol.* **18**, 393–401 (2017).
- Arruda, L. C. M., Gaballa, A. & Uhlin, M. Graft $\gamma\delta$ TCR sequencing identifies public clonotypes associated with hematopoietic stem cell transplantation efficacy in acute myeloid leukemia patients and unravels cytomegalovirus impact on repertoire distribution. *J. Immunol.* **202**, 1859–1870 (2019).
- Davey, M. S. et al. Clonal selection in the human V δ 1 T cell repertoire indicates $\gamma\delta$ TCR-dependent adaptive immune surveillance. *Nat. Commun.* **8**, 14760 (2017).
- Dimova, T. et al. Effector V γ 9V δ 2 T cells dominate the human fetal $\gamma\delta$ T-cell repertoire. *Proc. Natl Acad. Sci. USA* **112**, E556–E565 (2015).
- Willcox, C. R., Davey, M. S. & Willcox, B. E. Development and selection of the human V γ 9V δ 2⁺ T-cell repertoire. *Front. Immunol.* **9**, 1501 (2018).
- Di Marco Barros, R. et al. Epithelia use butyrophilin-like molecules to shape organ-specific $\gamma\delta$ T cell compartments. *Cell* **167**, 203–218.e217 (2016).
- Chowers, Y., Holtmeier, W., Harwood, J., Morzycka-Wroblewska, E. & Kagnoff, M. F. The V δ 1 T cell receptor repertoire in human small intestine and colon. *J. Exp. Med.* **180**, 183–190 (1994).
- Holtmeier, W., Chowers, Y., Lumeng, A., Morzycka-Wroblewska, E. & Kagnoff, M. F. The δ T cell receptor repertoire in human colon and peripheral blood is oligoclonal irrespective of V region usage. *J. Clin. Invest.* **96**, 1108–1117 (1995).
- Holtmeier, W., Witthoft, T., Hennemann, A., Winter, H. S. & Kagnoff, M. F. The TCR- δ repertoire in human intestine undergoes characteristic changes during fetal to adult development. *J. Immunol.* **158**, 5632–5641 (1997).

17. Falk, M. C. et al. Predominance of T cell receptor V δ 3 in small bowel biopsies from coeliac disease patients. *Clin. Exp. Immunol.* **98**, 78–82 (1994).
18. Holtmeier, W., Rowell, D. L., Nyberg, A. & Kagnoff, M. F. Distinct δ T cell receptor repertoires in monozygotic twins concordant for coeliac disease. *Clin. Exp. Immunol.* **107**, 148–157 (1997).
19. Han, A. et al. Dietary gluten triggers concomitant activation of CD4⁺ and CD8⁺ $\alpha\beta$ T cells and $\gamma\delta$ T cells in celiac disease. *Proc. Natl Acad. Sci. USA* **110**, 13073–13078 (2013).
20. Mayassi, T. et al. Chronic inflammation permanently reshapes tissue-resident immunity in celiac disease. *Cell* **176**, 967–981.e919 (2019).
21. Han, A., Glanville, J., Hansmann, L. & Davis, M. M. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat. Biotechnol.* **32**, 684–692 (2014).
22. Halstensen, T. S., Scott, H. & Brandtzaeg, P. Intraepithelial T cells of the TcR γ/δ ⁺ CD8⁻ and V δ 1/J δ 1⁺ phenotypes are increased in coeliac disease. *Scand. J. Immunol.* **30**, 665–672 (1989).
23. Cheng, C. et al. Next generation sequencing reveals changes of the $\gamma\delta$ T cell receptor repertoires in patients with pulmonary tuberculosis. *Sci. Rep.* **8**, 3956 (2018).
24. Harden, J. L., Hamm, D., Gulati, N., Lowes, M. A. & Krueger, J. G. Deep sequencing of the T-cell receptor repertoire demonstrates polyclonal T-cell infiltrates in psoriasis. *F1000Research* **4**, 460 (2015).
25. Kutlu, T. et al. Numbers of T cell receptor (TCR) $\alpha\beta$ ⁺ but not of TcR $\gamma\delta$ ⁺ intraepithelial lymphocytes correlate with the grade of villous atrophy in coeliac patients on a long term normal diet. *Gut* **34**, 208–214 (1993).
26. Sollid, L. M. & Jabri, B. Triggers and drivers of autoimmunity: lessons from coeliac disease. *Nat. Rev. Immunol.* **13**, 294–302 (2013).
27. Jabri, B. & Sollid, L. M. T cells in celiac disease. *J. Immunol.* **198**, 3005–3014 (2017).
28. Christophersen, A. et al. Tetramer-visualized gluten-specific CD4⁺ T cells in blood as a potential diagnostic marker for coeliac disease without oral gluten challenge. *Eur. Gastroenterol. J.* **2**, 268–278 (2014).
29. Christophersen, A. et al. Distinct phenotype of CD4⁺ T cells driving celiac disease identified in multiple autoimmune conditions. *Nat. Med.* **25**, 734–737 (2019).
30. Guo, X. Z. et al. Rapid cloning, expression, and functional characterization of paired $\alpha\beta$ and $\gamma\delta$ T-cell receptor chains from single-cell analysis. *Mol. Ther. Methods Clin. Dev.* **3**, 15054 (2016).
31. Vander Heiden, J. A. et al. pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. *Bioinformatics* **30**, 1930–1932 (2014).
32. Alamyar, E., Duroux, P., Lefranc, M. P. & Giudicelli, V. IMGT(R) tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. *Methods Mol. Biol.* **882**, 569–604 (2012).

