



OPEN ACCESS

ORIGINAL ARTICLE

# T-cell repertoires in refractory coeliac disease

Julia Ritter,<sup>1</sup> Karin Zimmermann,<sup>1</sup> Korinna Jöhrens,<sup>1</sup> Stefanie Mende,<sup>1</sup>  
Anke Seegebarth,<sup>1</sup> Britta Siegmund,<sup>2</sup> Steffen Hennig,<sup>3</sup> Kremena Todorova,<sup>4</sup>  
Andreas Rosenwald,<sup>5</sup> Severin Daum,<sup>2</sup> Michael Hummel,<sup>1</sup> Michael Schumann<sup>2,6,7</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/gutjnl-2016-311816>).

<sup>1</sup>Institute of Pathology, Charité—University Medicine, Berlin, Germany

<sup>2</sup>Department of Gastroenterology, Infectious Diseases and Rheumatology, Charité—University Medicine, Berlin, Germany

<sup>3</sup>HS Diagnostics GmbH, Berlin, Germany

<sup>4</sup>Center for Tumor Medicine, Charité—University Medicine, Berlin, Germany

<sup>5</sup>Institute of Pathology, University of Würzburg, and Comprehensive Cancer Center Mainfranken (CCCMF), Würzburg, Germany

<sup>6</sup>Berlin Institute of Health, Berlin, Germany

<sup>7</sup>Berlin-Brandenburg School for Regenerative Therapies, Berlin, Germany

**Correspondence to**

Professor Michael Hummel, Institute of Pathology, Charité—University Medicine, Berlin D-12200, Germany, [michael.hummel@charite.de](mailto:michael.hummel@charite.de)

Received 8 March 2016

Revised 1 December 2016

Accepted 8 December 2016

Published Online First

10 February 2017

**ABSTRACT**

**Objective** Refractory coeliac disease (RCD) is a potentially hazardous complication of coeliac disease (CD). In contrast to RCD type I, RCD type II is a precursor entity of enteropathy-associated T-cell lymphoma (EATL), which is associated with clonally expanding T-cells that are also found in the sequentially developing EATL. Using high-throughput sequencing (HTS), we aimed to establish the small-intestinal T-cell repertoire (TCR) in CD and RCD to unravel the role of distinct T-cell clonotypes in RCD pathogenesis.

**Design** DNA extracted from duodenal mucosa specimens of controls (n=9), active coeliacs (n=10), coeliacs on a gluten-free diet (n=9), RCD type I (n=8), RCD type II (n=8) and unclassified Marsh I cases (n=3) collected from 2002 to 2013 was examined by TCRβ-complementarity-determining regions 3 (CDR3) multiplex PCR followed by HTS of the amplicons.

**Results** On average, 10<sup>6</sup> sequence reads per sample were generated consisting of up to 900 individual TCRβ rearrangements. In RCD type II, the most frequent clonotypes (ie, sequence reads with identical CDR3) represent in average 42.6% of all TCRβ rearrangements, which was significantly higher than in controls (6.8%; p<0.01) or RCD type I (6.7%; p<0.01). Repeat endoscopies in individual patients revealed stability of clonotypes for up to several years without clinical symptoms of EATL. Dominant clonotypes identified in individual patients with RCD type II were unique and not related between patients. CD-associated, gliadin-dependent CDR3 motifs were only detectable at low frequencies.

**Conclusions** TCRβ-HTS analysis unravels the TCR in CD and allows detailed analysis of individual TCRβ rearrangements. Dominant TCRβ sequences identified in patients with RCD type II are unique and not homologous to known gliadin-specific TCR sequences, supporting the assumption that these clonal T-cells expand independent of gluten stimulation.

**INTRODUCTION**

While atrophy of the small intestinal mucosa in patients suffering from coeliac disease (CD) reverses in >95% of patients once a gluten-free diet (GFD) is introduced, there are refractory cases, who—in spite of closely following the diet—continue to suffer from malabsorption of nutrients secondary to small intestinal atrophy.<sup>1 2</sup>

More importantly, two types of refractory CD (RCD) exist. While RCD type I appears to be an aetiologically heterogeneous condition that most

**Significance of this study****What is already known on this subject?**

- High-throughput sequencing (HTS) is a sensitive approach to identify subpopulations of clonal T-cells and to determine the T-cell diversity.
- However, T-cell variability of the small intestinal compartment has not been studied in detail, especially regarding population size and consistency of clonal T-cell subpopulations.
- Refractory coeliac disease (RCD) is a severe complication of coeliac disease, frequently progressing to a fatal enteropathy-associated T-cell lymphoma (EATL).
- Intraepithelial clonal T-cell populations with a specific immunophenotype occur in RCD type II.

**What are the new findings?**

- Analysis of TCR diversity reveals small dominant (clonal) subpopulations that are stable over years even in healthy duodenal mucosa.
- In RCD type II, T-cell diversity is significantly diminished. Dominant clonotypes consisting of up to 73% of all TCR rearrangement occur.
- Patients with highly frequent clonotypes appear to be at-risk for progression to EATL.
- Dominant RCD type II T-cell clones appear not to be gliadin-dependent.

**How might it impact on clinical practice in the foreseeable future?**

- Analysis of TCR diversity and tracking of stable TCR clonotypes might be a prognostic tool to identify patients at-risk for progression to EATL.
- For reasons of high sensitivity of HTS-TCR and its potential to reproducibly identify clones by their specific TCR-CDR3 sequences, TCRβ-HTS might become a valuable diagnostic tool for future treatment studies in RCD type II, since it will likely pick-up even residual clonal lymphocyte populations and determine copy numbers allowing quantitative judgement of treatment efficacy.

often can be treated with an immunosuppressive regimen, the small intestinal mucosa in RCD type II is infiltrated with a population of aberrant intraepithelial T-lymphocytes, which are clonally expanding.<sup>1 3</sup> Recently, the presumable precursor



**To cite:** Ritter J, Zimmermann K, Jöhrens K, et al. *Gut* 2018;**67**:644–653.

of RCD type II-associated aberrant T-cells was identified in patients without CD as an IL-15 receptor(+) CD7(+) CD127(-) CD34(-) cell population.<sup>4</sup> Since patients with RCD type II carry a high risk of developing an enteropathy-associated T-cell lymphoma (EATL) and since T-cell clones found in RCD type II are indistinguishable from EATL T-cell clones by their identical T-cell receptor (TCR) gene rearrangements, RCD type II is regarded as a precursor lesion to EATL.<sup>5-8</sup> Strategies to diagnose RCD type II include immunohistochemistry revealing a loss of CD8 in intraepithelial lymphocytes (IELs) in combination with evidence for clonal expansion of IELs, as uncovered by T-cell receptor  $\beta$  (TCR $\beta$ ) or TCR $\gamma$ -PCR in duodenal samples.<sup>9,10</sup> Alternative diagnostic procedures include flow cytometric immunophenotyping of isolated IELs from duodenal mucosa.<sup>11</sup>

The structure of TCR gene rearrangements in RCD type II is a matter of current research.<sup>12</sup> The TCR acts as a heterodimer, which is in 85% of small intestinal IELs composed of  $\alpha$  and  $\beta$  chains, whereas in 15% of these cells it consists of  $\gamma$  and  $\delta$  chains.<sup>13</sup> Antigen binding of TCR $\alpha/\beta$  is determined by three hypervariable complementarity-determining regions (CDR)1-3, where CDR3 is mainly responsible for recognising processed antigens.<sup>14</sup> These CDRs are generated by somatic recombination among the variable (V), diversity (D present only in TCR $\beta$ ) and joining (J) segments. In addition, non-templated bases are incorporated at the recombination junctions, which generate overall a spectrum of  $>2.5 \times 10^7$  unique TCR receptors.<sup>15</sup> So far, it is well known that clonally expanded mucosal T-cells expressing the same CDR3 region are a hallmark of RCD type II.<sup>9</sup> Since it has been shown from early on that clones found in RCD type II were identical to the clonal structure of EATLs that later arose within the same individual, the assumption came up that RCD type II is a premalignant condition that might require an oncologic rather than an immunosuppressive treatment strategy.<sup>16-19</sup> However, the TCR repertoire in the duodenal mucosa in its molecular heterogeneity with respect to the number and frequency of co-existing clonally expanding T-cells in the healthy duodenum as well as in active CD and diseases with clonal expansion of T-cells as RCD type II needs further characterisation. Moreover it is unknown, if RCD type II aberrant IELs express public TCR sequences, that is, CDR3 segments that are found interindividually in RCD type II and whether they relate to previously published TCR sequences.

Up to now, analysis of the TCR-CDR3 region was dominated by technologies like spectratyping generating information about fragment length rather than sequence identity.<sup>20,21</sup> With the advent of high-throughput sequencing (HTS), it is now possible to determine a huge number of sequences concurrently. In this work, we employed a multiplex TCR $\beta$  PCR assay followed by HTS of the generated amplicons to determine the entire individual spectrum of CDR3 sequences of patients with RCD type II compared with uncomplicated CD and RCD type I. This allows for the first time a precise and quantitative dissection of the individual T-cell repertoire (TCR) of these patients and by comparison to healthy individuals insights into the general CDR3 sequence variability of the human small intestine.

## METHODS

A more comprehensive description of the methods is included in the online supplementary material methods section.

### Study design and sample acquisition

Duodenal mucosal tissue was collected after informed consent during routine endoscopy at the Department of

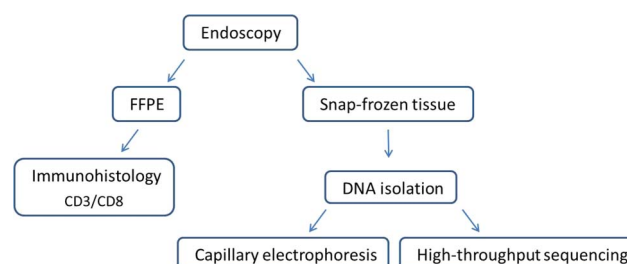
Gastroenterology (Ethics approval EA4/016/14). Endoscopic biopsies were snap-frozen immediately after sampling or processed to formalin-fixed paraffin embedded (FFPE) samples. FFPE samples were immunostained and reviewed by pathologists, whereas the snap-frozen tissue was additionally used for DNA extraction and subsequent molecular analyses. A detailed flow chart of the sample preparation and analysis is given in figure 1.

### Patient characteristics

The Charité coeliac tissue bank was searched for patients with diagnosis of RCD and matched appropriate controls. Fifty-two small intestinal tissue samples suitable for this study were identified that had been collected within the time span from 2002 to 2013. This included a total of 47 patients with either active coeliac disease (ACD), CD on a GFD (CD GFD), RCDs, patients with a Marsh I enteropathy and healthy controls (table 1). Control subjects were scoped for reasons of dyspepsia and did not reveal clinical or histological signs of enteropathy. Patients on a normal diet that revealed on histology a Marsh III enteropathy (ie, crypt hyperplasia or crypt hyperplasia plus villous atrophy), showed positive transglutaminase-IgA serology and later responded to a GFD, were judged as ACD (n=10). Patients with CD GFD (n=9) had been previously diagnosed as coeliacs and had been successfully treated with a strict GFD for at least 1 year. RCD was diagnosed in patients who were refractory to a GFD for  $>1$  year. Differential diagnoses and a failure to adhere to the diet had to be ruled out (RCD type I, n=8; RCD type II, n=8).<sup>2</sup> Exceptions from this rule included one RCD type II case where diagnosis had to be established earlier to introduce timely treatment. Subtyping of RCD in type I and II was not accomplished by flow cytometry as it is now included in the current diagnostic standards, since at the time of collection (2002-2013) this diagnostic modality was not part of clinical routine, but on the basis of TCR clonality (GeneScan analysis<sup>21</sup>) and loss of CD8 expression in  $>50\%$  of CD3-positive intraepithelial T-cells. Patients fulfilling clinical criteria for RCD but with preserved CD8 expression and no evidence for TCR clonality were classified as RCD type I. Patients without CD (n=9) underwent gastroscopy due to unclear upper abdominal symptoms but examination revealed no pathological findings. A small group of patients with a Marsh I enteropathy only (n=3), but without CD (ie, negative coeliac serology/no improvement on a GFD/HLA-DQ2 negative), were classified as a second control group, called 'Marsh I'. In five patients, biopsy specimen of two independent examinations was collected to analyse the reproducibility of data.

### Immunohistochemistry and histology

Duodenal biopsies were reviewed by experienced pathologists and classified according to Marsh<sup>22</sup> and Oberhuber *et al.*<sup>23</sup>



**Figure 1** Flow chart of sample preparation and analysis. FFPE, formalin-fixed paraffin embedded.

**Table 1** Characteristics of patients enrolled in the study

Patient	Sex	Age	Group	Histology (Marsh)	IEL count*	CD8 staining†	TCR-βF1 staining‡	TCRβ clonality§	TCRγ clonality§	Tg-IgA serology	HLA-DQ¶	Died of EATL	No. of reads**
1	F	49	ACD	IIIC	100	Positive	–	Not clonal	Not clonal	Positive	DQ2.5-ho	–	1 386 048
2	M	60	ACD	IIIA	100	Positive	–	Not clonal	Not clonal	Positive	nd	–	1 375 338
3A	F	61	ACD	IIIB	52	Positive	–	Not clonal	Not clonal	Positive	DQ2.5	–	1 492 114
3B	F	65	CD GFD	0	<25	nd	–	Not clonal	Not clonal	Negative		–	1 229 854
4A	M	64	ACD	IIIA	65	Partially negative	–	A: clonal; B/C: not clonal	Clonal	Positive	DQ2.5-ho	–	1 398 559
4B	M	69	ACD	IIIA	61	Partially negative	–	A: clonal; B/C: not clonal	Clonal	Positive		–	1 109 972
5A	M	61	ACD	IIIA	52	Positive	–	Not clonal	Not clonal	Positive	DQ2.5	–	3 838 298
5B	M	61	ACD	IIIA	45	Positive	–	Not clonal	Not clonal	Positive		–	2 500 660
6	M	59	ACD	IIIB	80	Partially negative	–	B: clonal; A/C: not clonal	Not clonal	Positive	DQ2.5	–	2 914 869
7	M	55	ACD	IIIA	25	Positive	–	B: clonal; A/C: not clonal	Not clonal	Positive	DQ2.5/2.2	–	1 166 697
8	M	82	ACD	IIIA	58	Positive	–	Not clonal	Not clonal	Positive	DQ2.5/2.2	–	1 403 191
9	F	36	ACD	IIIA	70	Positive	–	Not clonal	Not clonal	Positive	DQ2.5	–	1 056 191
10A	F	51	ACD	IIIA	43	Positive	–	Not clonal	Not clonal	Positive	DQ2.5	–	1 489 344
10B	F	52	CD GFD	0	<25	nd	–	Not clonal	Not clonal	Negative		–	1 819 643
11	F	54	CD GFD	I	40	Positive	–	B: clonal; A/C: not clonal	Clonal	Negative	DQ2.5	–	1 096 650
12	F	34	CD GFD	I	60	nd	–	Not clonal	Not clonal	Negative	DQ2.5-ho	–	1 429 008
13	F	44	CD GFD	I	50	Positive	–	B: clonal; A/C: not clonal	Clonal	Negative	DQ8.5	–	2 120 940
14	F	47	CD GFD	I	38	Positive	–	Not clonal	Not clonal	Negative	DQ8	–	1 389 764
15	F	44	CD GFD	0	<25	nd	–	Not clonal	Not clonal	Negative	DQ2.5	–	4 882 201
16	M	62	CD GFD	I-II	52	nd	–	Not clonal	Not clonal	Negative	DQ6	–	1 232 456
17	F	36	CD GFD	I	67	Positive	–	Not clonal	Not clonal	Negative	DQ2.5	–	1 767 305
18	F	29	CD GFD	IIIA††	67	Positive	–	Not clonal	Not clonal	Positive	DQ2.5/8	–	1 076 997
19	F	32	CD GFD	IIIA††	47	nd	–	B: clonal; A/C: not clonal	Not clonal	Negative	DQ2.5	–	2 867 422
20	F	42	Control	0	<25	nd	–	A/B: not clonal; C clonal	Not clonal	Negative	DQ6	–	1 587 356
21	F	50	Control	0	<25	nd	–	Not clonal	Not clonal	Negative	DQ2.5-ho	–	1 661 085
22	F	20	Control	0	<25	nd	–	Not clonal	Not clonal	Negative	DQ2.5-ho	–	2 001 279
23	F	30	Control	0	<25	nd	–	Clonal	Clonal	Negative	DQ6	–	1 169 960
24	F	22	Control	0	<25	nd	–	Not clonal	Not clonal	Negative	DQ6	–	1 370 690
25	M	43	Control	0	<25	nd	–	Not clonal	Not clonal	Negative	DQ6	–	564 245
26	F	20	Control	0	<25	nd	–	A: biallelic clonal; B/C: not clonal	Not clonal	Negative	DQ2.5	–	1 480 608
27	F	36	Control	0	<25	nd	–	Not clonal	Not clonal	Negative	DQ2.5	–	1 439 428
28	F	60	Control	0	<25	nd	–	Not clonal	Not clonal	Negative	DQ2.5	–	2 148 355
29	M	51	RCD type I	IIIA	35	Positive	Positive	Not clonal	Not clonal	Negative	DQ6	–	1 289 891
30	F	60	RCD type I	IIIB	80	Positive	Positive	Not clonal	Not clonal	Positive	DQ2.5	–	888 615
31	M	49	RCD type I	IIIB	69	Positive	Partially negative	Not clonal	Not clonal	Negative	DQ8	–	1 711 745
32	F	45	RCD type I	IIIA	32	Positive	Positive	Not clonal	Not clonal	Negative	DQ2.5-ho	–	1 283 412
33	F	47	RCD type I	IIIA	51	Positive	Partially negative	Not clonal	Not clonal	Negative	DQ7	–	1 022 807
34	F	75	RCD type I	IIIB	25	Positive	negative	Not clonal	Not clonal	Negative	DQ8	–	839 327
35	M	73	RCD type I	IIIA	70	Positive	Positive	Not clonal	Not clonal	Negative	DQ7	–	1 382 335
36‡‡	F	71	RCD type I	IIIB	81	Negative	Positive	Not clonal	Clonal	Negative	DQ2.5	–	1 018 656
37	M	49	RCD type II	IIIB	54–63	Negative	Partially negative	A/B: not clonal; C clonal	Not clonal	Positive	DQ2.5	–	975 336
38	M	67	RCD type II	IIIA	74	Negative	Positive	Clonal	Clonal	Positive	DQ2.5-ho	Yes	4 150 734
39	F	41	RCD type II	IIIC	>25	Negative	Partially negative	Clonal	Clonal	Negative	DQ2.5-ho	Yes	4 261 385
40	F	69	RCD type II	IIIC	nd	nd	Partially negative	Clonal	Biallelic clonal	nd	DQ2.5	Yes	2 283 715
41	M	75	RCD type II	IIIB	120	Negative	Partially negative	Clonal	Clonal	Positive	DQ2.5-ho	Yes	6 012 299

Continued

Table 1 Continued

Patient	Sex	Age	Group	Histology (Marsh)	IEL count*	CD8 staining†	TCR-βF1 staining‡	TCRβ clonality§	TCRγ clonality§	Tg-IgA serology	HLA-DQ¶	Died of EATL	No. of reads**
42	F	82	RCD type II	IIIC	52	Positive	negative	Biallelic clonal	Not clonal	Positive	DQ2.5/8	–	4 045 900
43A§§	F	51	RCD type II	IIIB	<25	Negative	negative	Not clonal	Not clonal	Positive	DQ2.5-ho	–	1 053 659
43B	F	56	RCD type II	IIIA	68	Negative	negative	Not clonal	Not clonal	Negative			1 408 767
44	M	62	RCD type II	IIIC	87	Negative	Positive	Biallelic clonal	Biallelic clonal	Negative	DQ8	Yes	4 965 878
45	M	56	Marsh I	I	80	Positive	–	Not clonal	Biallelic clonal	Negative	DQ5	–	958 161
46	F	53	Marsh I	I	47	Partially negative	–	Not clonal	Not clonal	Negative	DQ2.2	–	1 693 995
47	F	45	Marsh I	I	72	Positive	–	A: biallelic clonal; B/C: not clonal	Biallelic clonal	Negative	DQ8	–	1 659 303

Patients labelled as 'A' and 'B' were analysed twice using duodenal mucosa sampled at distinct gastroscopies.

\*Counted per 100 enterocytes.

†Positive: >50% of CD3-T-cells are CD8+.

‡Positive: IHC detection of TCR on CD3-T-cells.

§Determined by GeneScan (BIOMED-2 TCRβ primer sets A/B and C; TCRγ primer sets A and B), detailed TCRβ GeneScan results are provided in the online supplementary material I.

¶DQ2.5, heterozygous for DQ2.5, DQ2.5-ho, homozygous for HLA-DQ2.5, DQ2.5/2.2, heterozygous for DQ2.5 and DQ2.2.

\*\*Number of analysed sequence reads, a mean of  $1.9 \times 10^6$  reads per sample guaranteed a 130-fold coverage of single TCRβ sequences (100 ng DNA input  $\approx$  14,500 T-cells).

††Significant improvement on a GFD. Further disease course showed complete remission.

‡‡Classification as RCD type I was made on the basis of lack of clonality of TCR-β GeneScan and positivity of TCR-βF1 immunostaining.

§§Patient 43 had previously been TCRγ-monoclonal and thus been categorised as RCD type II and treated accordingly.

ACD, active coeliac disease; CD, coeliac disease; EATL, enteropathy-associated T-cell lymphoma; GFD, gluten-free diet; IEL, intraepithelial lymphocytes; IHC, immunohistochemistry; nd, not determined; RCD, refractory coeliac disease; TCR, T-cell receptor; Tg-IgA, transglutaminase-IgA serology.

Quantification of CD3-positive T-cells was analysed by immunofluorescence using an anti-CD3 antibody. The average T-cell content was  $12.1\% \pm 1.2\%$  (range 5.6%–19.6%; online supplementary figure S1).

### Sequencing of TCRβ gene rearrangements and clonotype analysis

For HTS of TCRβ gene rearrangements, amplicons covering the CDR3 and proportions of the respective Vβ-segments and Jβ-segments were generated as described previously (see also supplementary material methods section for details).<sup>24</sup> Clonotypes consist of clusters of identical or almost identical (maximum 1 bp difference) sequences of TCRβ gene rearrangements. For each patient, the percentage distribution and the composition (diversity) of the clonotypes was assessed. To identify identical TCRβ rearrangements, intersection between clonotypes was determined and displayed as a heatmap. For detection of public clonotypes, we compared in silico translated peptide CDR3 sequences of our data set with published CDR3 sequences of patients with CD without allowing amino acid mismatches.<sup>25 26</sup> Results of T-cell diversity analysis were not corrected for percentage of T-cells found in the samples, since this was not inter-related (see online supplementary figure S2).

### Statistics

Non-parametric analysis between data groups were performed by the Mann-Whitney U test using GraphPad PRISM V.6 software (GraphPad Software, La Jolla, California, USA). Linear correlation between T-cell diversity and patient age was tested by calculation of the Pearson's correlation coefficient. The biodiversity of each sample was assessed by using the Shannon diversity index defined as:

$$\sum (p_i \times \log_2(p_i)),$$

where  $p_i = n/N$  is the proportion of clonotypes  $n$  ( $n$ =number of reads per clonotype) relative to the total number of clonotypes

$N$  ( $N$ =number of reads per sample).<sup>27</sup> A  $p$  value  $<0.05$  was considered to be statistically significant.

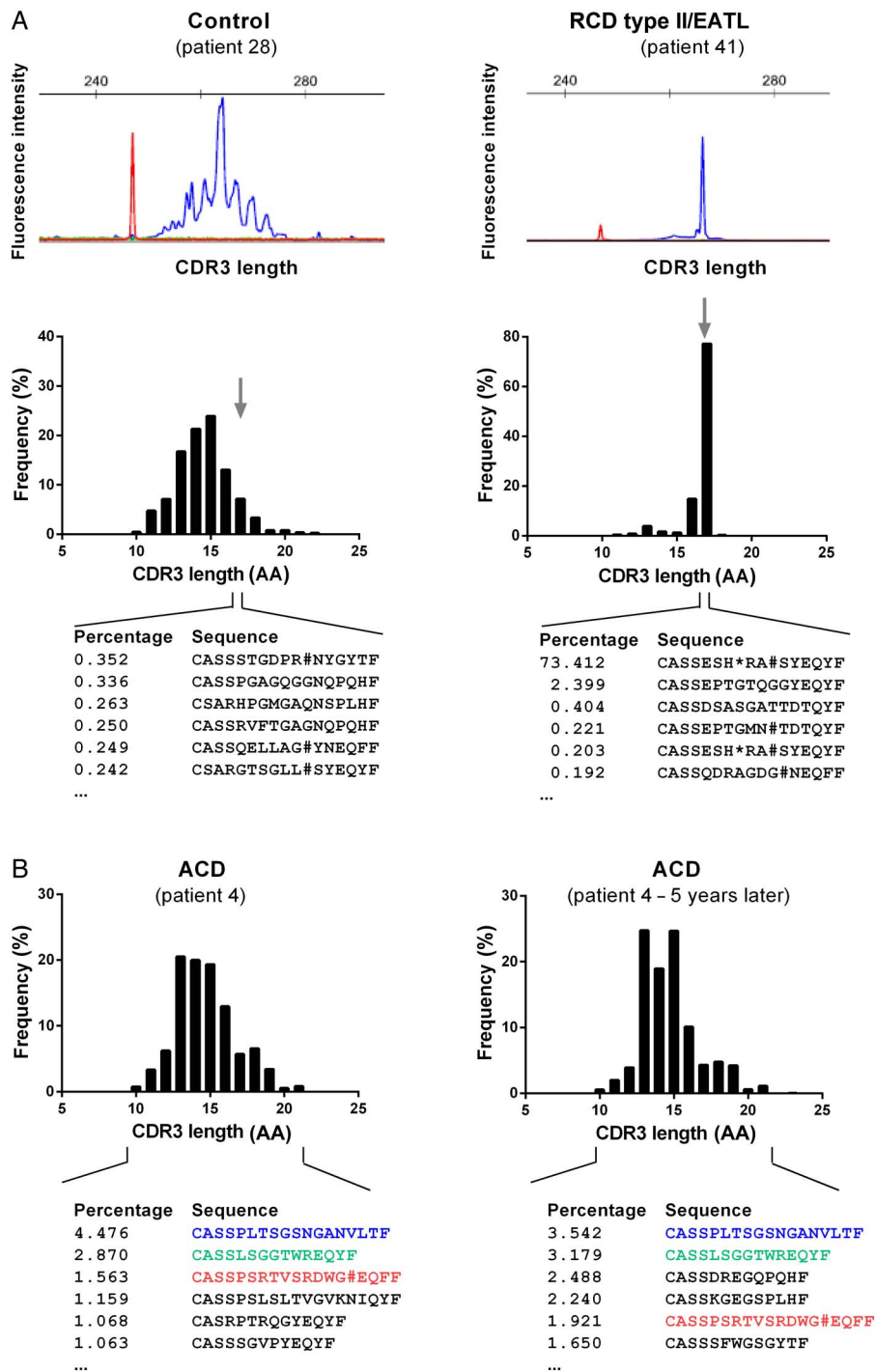
## RESULTS

### TCRβ repertoire diversity in healthy duodenum and various states of coeliac disease

Fifty-two duodenal biopsies of 47 patients were included in the study. Classification was done according to the criteria described in the Methods section yielding samples for ACD ( $n=12$ ), RCD type I ( $n=8$ ), RCD type II ( $n=9$ ), CD GFD ( $n=11$ ), Marsh I ( $n=3$ ) and controls ( $n=9$ ). DNA from all cases was subjected to multiplex TCRβ PCR to amplify all available gene rearrangements. Amplicons were subsequently sequenced by HTS. For comparability of HTS data among the different groups, the same amount of genomic DNA (100 ng) was employed for each TCRβ multiplex PCR assay and the number of T-cells was assessed by immunostaining for CD3. A mean of approximately  $1.9 \times 10^6$  joined sequence reads per sample was available for further analysis.

Size distribution comparison of TCRβ gene rearrangements generated by HTS and by conventional capillary electrophoresis (GeneScan) based on BIOMED-2-derived TCRβ amplicons revealed very similar patterns resulting in an identical interpretation of the results in all cases (figure 2A). In contrast to capillary electrophoresis, HTS is able to resolve each individual TCRβ gene rearrangement on the basis of its DNA sequence. To compare our findings with the previously reported preferred V-segment usage in CD-associated T-cells, we examined the V-segment and J-segment usage of HTS-derived TCRβ gene rearrangements among the various patient groups (see online supplementary figure S3A). The analysis was focused on the top 20 clonotypes per individual and compared the RCD type II group with all others to raise the question, if RCD type II clones were distinct in their V-segment usage. We detected a borderline significant association with TRBV6 and TRBV7 ( $p=0.045$  and  $p=0.050$ , respectively, online supplementary figure S3B). However, statistics turned insignificant once

**Figure 2** T-cell repertoire (TCR) $\beta$  repertoires in patients with coeliac disease (CD). (A) Representative complementarity-determining regions 3 (CDR3) length spectrum of TCR $\beta$  sequences as determined by capillary gel electrophoresis (GeneScan, BIOMED-2 primer combinations, upper panel) or high-throughput sequencing (HTS)-based TCR $\beta$  multiplex analysis of a control individual and a representative patient with refractory coeliac disease (RCD) type II (lower panel). For the digital HTS-based CDR3 length spectrum sequences of the most frequent TCR $\beta$  clonotypes with 17 amino acids (AA) in length are shown. TCR $\beta$  CDR3 size is defined as all AA starting from the conserved 5' cysteine in the V-segment and ending at the conserved 3' phenylalanine in the J-segment. In red the size marker is delineated. (B) Two independent digital CDR3 length distribution of TCR $\beta$  sequences of the same patient with active coeliac disease (ACD) sampled in an interval of 5 years. The most frequent TCR sequences were shown and identical CDR3 sequences were coloured.



multiple testing was applied. Moreover, when statistics were done group-specifically or when all clonotypes (>0.03%) were included, this borderline significant finding was not confirmed (see online supplementary figure S3A). Thus, we could not definitely confirm the previously reported distinct usage of V-segments in RCD type II.<sup>28</sup> We then further focused V-segment usage analysis on the subgroup of RCD type II and RCD type II/EATL and found some indicators of distinct V-segment usage of RCD type II and RCD type II/EATL (see online supplementary figure S4). This was supported by the fact that the hierarchical clustering algorithm almost perfectly grouped RCD type II and RCD type II/EATL. Interestingly, this grouping was presumably caused by usage of the T-cell receptor beta V (TRBV) genes 19, 2, 15, 3 and 29. With respect to the

limited number of cases, larger cohorts will be needed for more confirmative assumptions.

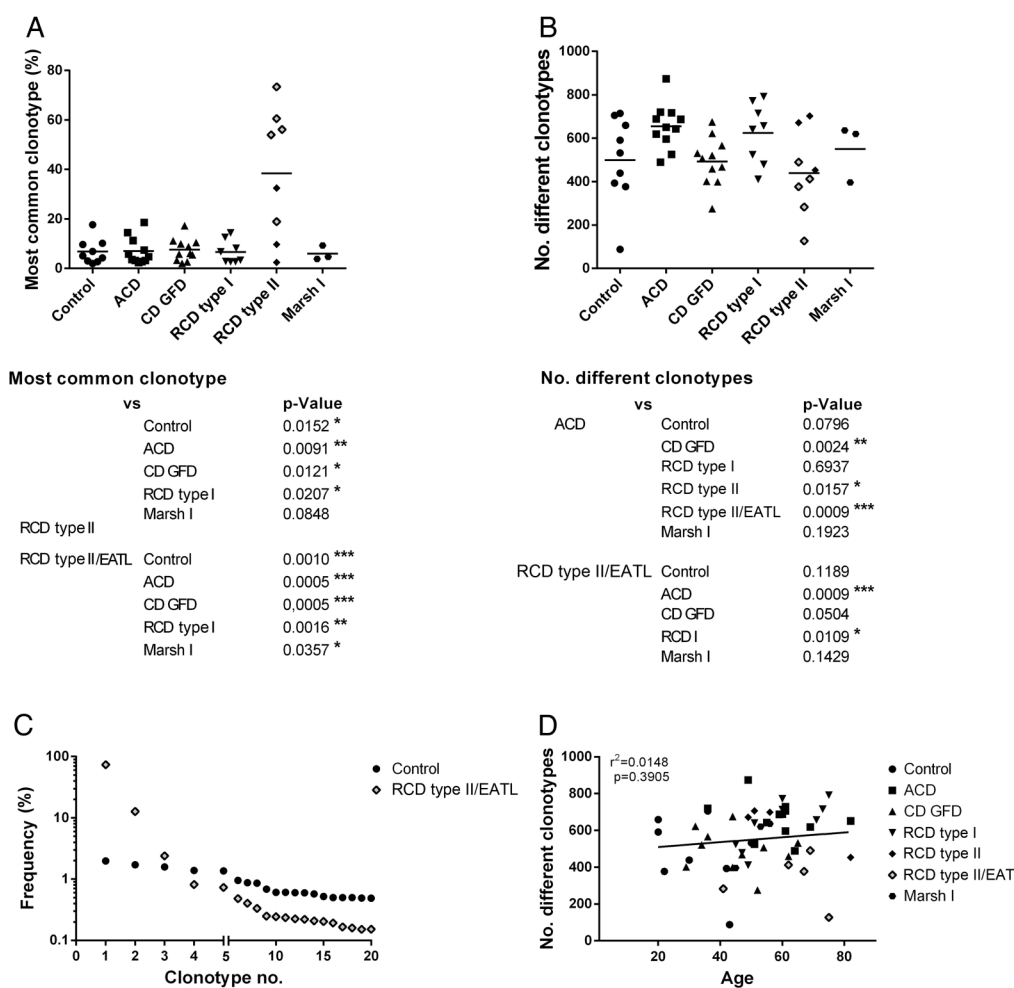
The current clinical approach to classify RCD based on T-cell clonality uses the occurrence of reproducible dominant peaks after GeneScan analysis, whereas HTS enables the detection of clonality by precise quantification of the number of sequence reads in clonotypes and the clonotype distribution as a feature of TCR diversity. In the control group, the mean frequency of the most frequent clonotype was 6.8% (range from 2.0%–17.7%). In contrast, the frequency distribution found in RCD type II was highly scattered, revealing dominant clonotypes with frequencies up to 73.4% and as low as 2.4% (mean: 42.6%). According to this observation, we introduced an additional classification of patients with RCD type II, which distinguished

patients with RCD type II who later developed an EATL (RCD type II/EATL) from those that did not (RCD type II). Of note, patients who were assigned to the RCD type II/EATL group had no clinical characteristics of EATL at the time of endoscopy when samples used for our study were acquired. Strikingly, 80% of patients with RCD type II (4/5) progressing to an EATL had frequencies of the most common clonotype >50%, resulting in a highly significant difference if compared with any other CD group (ACD  $p=0.0005$ ; CD GFD  $p=0.0005$ ; RCD type I  $p=0.0016$ ; Marsh I  $p=0.0357$ ) as well as to controls ( $p=0.0010$ ; figure 3A). In line, significantly lower numbers of clustered clonotypes were detectable in RCD type II/EATL as compared with ACD ( $p=0.0009$ ) and RCD type I ( $p=0.0109$ ; figure 3B). Patients with ACD showed a significantly higher number of clonotypes compared with patients with CD GFD ( $p=0.0024$ ). In figure 3C, an RCD type II case is compared with a control case with respect to the frequency of the 20 most common T-cell clones, highlighting the strikingly different

distribution of clonotypes. To ascertain that differences in frequencies of T-cell clones do not merely reflect the age of the patient as suggested previously for peripheral blood T-cells, we plotted the number of individual clonotypes indicative for T-cell diversity against the age of the patients—irrespective of CD diagnosis.<sup>29</sup> Interestingly, no correlation was found (figure 3D). Similarly, no correlation between transglutaminase-IgA titre and the frequency of the most dominant clonotypes was detected (see online supplementary figure S5). The finding of significantly higher clonotype frequencies in RCD type II as well as RCD type II/EATL cases and the accompanied lower TCR $\beta$  diversity could be approved by the quantitative assessment of biodiversity via the Shannon diversity index (figure 4).<sup>27</sup>

### TCR $\beta$ repertoire variety and stability among patients with CD

In contrast to assays restricted to fragment length, HTS-based data are based on individual TCR $\beta$  sequences providing an unlimited resolution of the entire TCR of a given sample. Thus,



**Figure 3** T-cell repertoire (TCR) $\beta$  diversity in patients with coeliac disease (CD). (A) Dot plot of the most common clonotypes in all small intestinal mucosal samples included in this study. Each symbol illustrates the most common clonotype of a single sample in per cent. Group mean values are delineated by a horizontal bar. The average frequency of the most common clonotype in patients with refractory coeliac disease (RCD) type II who progressed to enteropathy-associated T-cell lymphoma (EATL) (RCD type II (EATL)) differs significantly to all other groups ( $p=0.036$ , Marsh I). (B) Diversity dot plot of TCR sequences in patients with CD. Each symbol represents the numbers of unique clonotypes of each biopsy. Patients with acute coeliac disease (ACD) presented significantly higher numbers of clonotypes in comparison with the CD gluten-free diet (GFD) (\*\* $p<0.005$ ) or RCD type II/EATL group (\*\* $p<0.001$ ). However, in patients with RCD type II/EATL significantly lower numbers of clonotypes were found as compared with patients with RCD type I (\* $p<0.05$ ). For the other groups, no further significant differences could be detected. (C) Clonotype distribution of T-cells in a representative control and a representative patient with RCD type II. Each dot represents the occurrence of a distinct TCR $\beta$  clonotype in per cent of overall clonotypes. In RCD type II, a small number of clonotypes is found at very high frequencies at the expense of the remaining sequences. (D) Dot plot of T-cell clonotype numbers per sample. No change of individual clonotype numbers with increasing donor age irrespective of CD diagnosis could be detected.

the presence of individual TCRβ sequences allows to compare sequences interindividually within our study group as well as to compare sequences with published (therefore called *public*) CDR3 protein sequences.

First, to address the question of stability of the TCR in small intestinal biopsy samples and reproducibility of this technique, we analysed the TCR composition of multiple small intestinal biopsies sampled from the same individual (patient #3, #4, #5, #10 and #43) at different time points (time in between endoscopies from two consecutive days up to 5 years). Thereby, we were able to track defined TCR rearrangements in samples over time. Figure 2B exemplifies the stability of individual T-cell clones over a period of 5 years. This corresponded to a Shannon diversity index that stayed fairly constant in mucosa samples that had been examined repeatedly (see online supplementary figure S6). This implied that (i) TCR diversity data from individual mucosal biopsies are meaningful and that

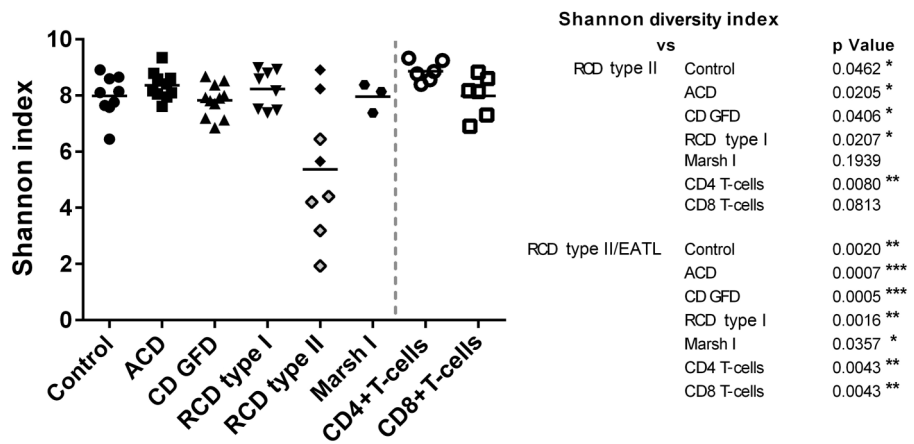
(ii) in the small intestine a set of dominant clonotypes can persist over a long period of time.

Second, to study the variety and re-occurrence of TCRβ sequences within or among CD groups, we compared all clonotypes of individual cases with all other clonotypes of our cohort and displayed the resulting overlaps as a heatmap (figure 5). Interestingly, no disease-specific identical clonotypes or TCRβ motifs were found within the various CD groups. Likewise, no specific TCRβ sequence was found to be identical among all groups.

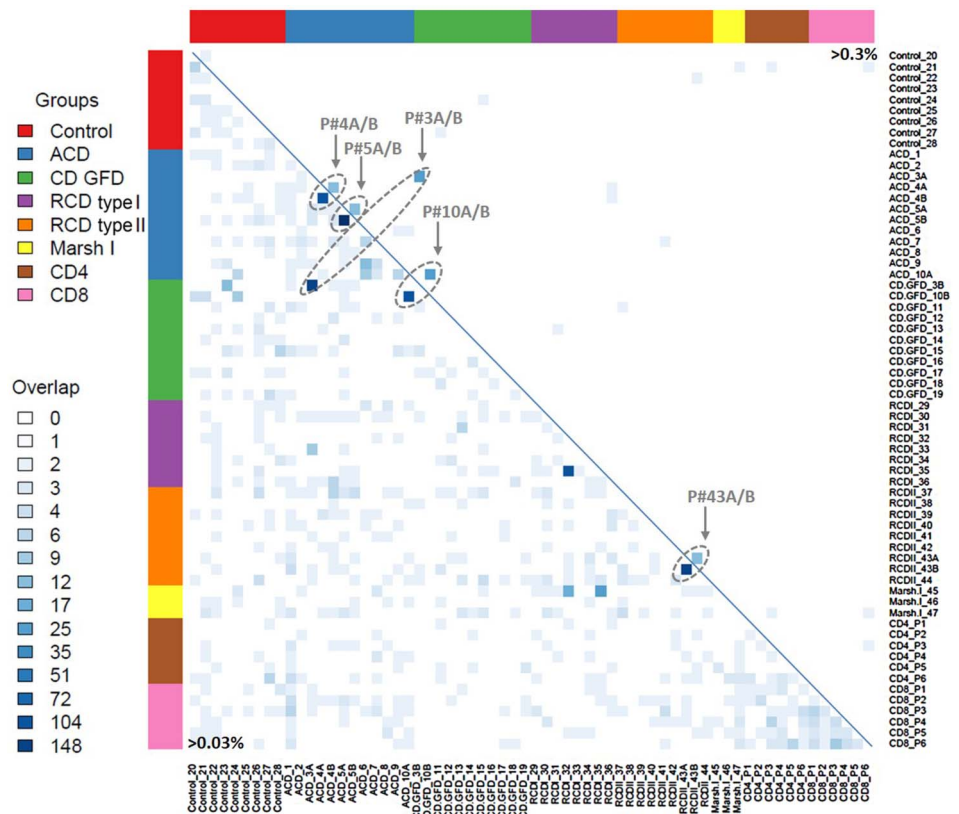
**Identification of public TCRs**

To identify previously published CD-related TCRβ sequences in our data set (designated as ‘public TCRs’), an alignment analysis of the corresponding CDR3 peptide sequences was performed (figure 6).<sup>30</sup> First, the highest frequency of public CDR3 sequences was detected in the ACD group, which appears

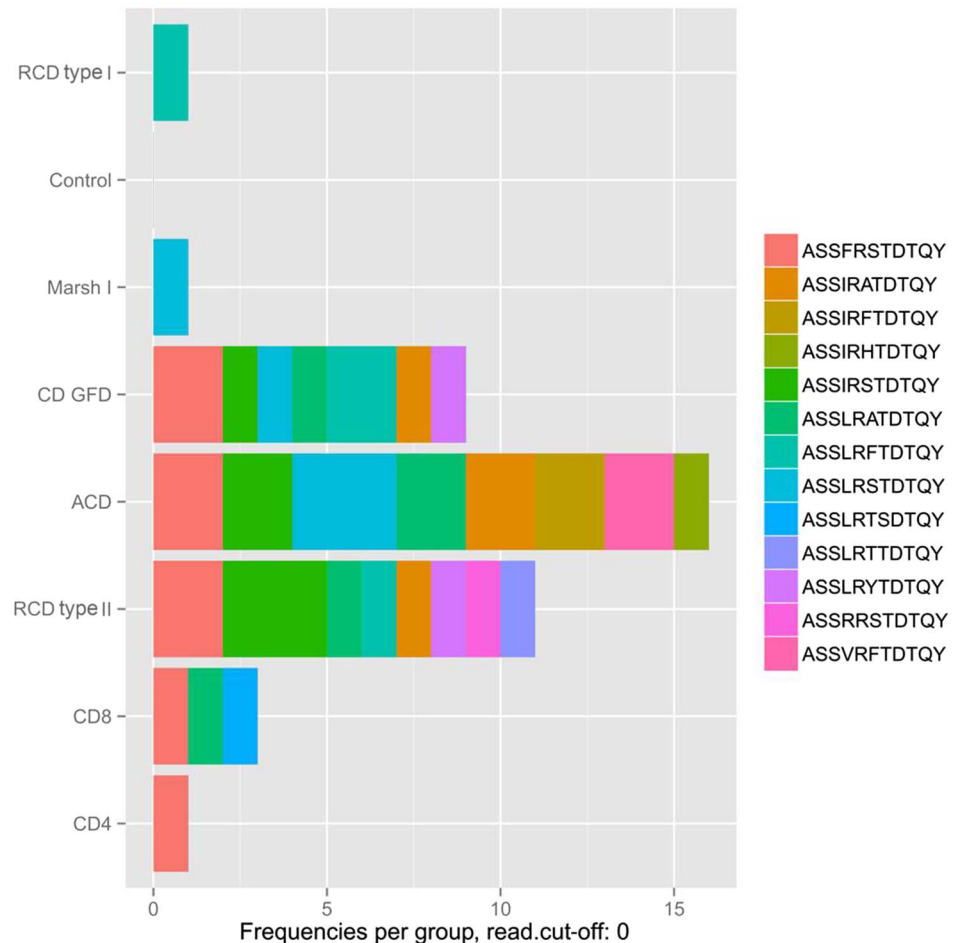
**Figure 4** Dot plot illustrating T-cell repertoire (TCR) diversity of each patient by delineation of corresponding Shannon diversity indices and comparison with peripheral blood CD4+ and CD8+ T-cells. Patients with refractory coeliac disease (RCD) type II/enteropathy-associated T-cell lymphoma (EATL) represents a significant reduced diversity in accordance with any other coeliac disease (CD) subgroup ( $p > 0.05$ ).



**Figure 5** Heatmap of T-cell repertoire (TCR)β sequence overlap between patients with coeliac disease (CD). The size of intersection in TCRβ-complementarity-determining regions 3 (CDR3) sequences (amino acid level) between all pairwise patients is displayed without clustering of columns or rows. The number of identical clonotypes between two biopsies was encoded in a grey scale. The lower triangular matrix displays the overlap of TCRβ-CDR3 sequences between samples with frequencies over 0.03% (35 361 sequences in total). The upper triangular matrix illustrates the identical overlap including all TCRβ-CDR3 sequences over 0.3% (2445 sequences in total). Overlapping sequences between two biopsies of one patient are highlighted. ACD, active coeliac disease; GFD, gluten-free diet.



**Figure 6** Public T-cell repertoire (TCR) $\beta$  sequences of small intestinal T-cells and peripheral blood CD4<sup>+</sup> and CD8<sup>+</sup> T-cells. Alignment analysis of the complementarity-determining regions 3 (CDR3) peptide sequence revealed homology with previously published public, gliadin-specific CDR3 sequences. The numbers of public sequences are plotted group-specifically. Of 39 CDR3 sequences that were evaluated, 13 were identified in our data set, some of which occurred in multiple subjects (number of individuals with confirmed public TCR-CDR3 sequence: control=0, ACD=5, CD GFD=4, RCD type I=1, RCD type II=4, Marsh I=1, CD4<sup>+</sup> T-cells=1, CD8<sup>+</sup> T-cells=3).<sup>25 26</sup> The percentage of reads corresponding to the detected clonotypes range at the low distribution end with a maximal occurrence of 0.036 read per cent. A detailed overview of the detected public TCR $\beta$  sequences and their individual occurrence are given in online supplementary table S1. ACD, active coeliac disease; CD, coeliac disease; RCD, refractory coeliac disease; GFD, gluten-free diet.



plausible, since this finding presumably reflects the active gliadin response. Interestingly, a small number of public TCRs was also found in the peripheral blood of healthy donors suggesting that healthy individuals rarely reveal a presumably limited immune response to gliadin peptides that is not further progressing to an enteropathy (figure 6). Since public TCR sequences in subjects with Marsh I and RCD type I occurred at similar levels if compared with healthy controls, it can be assumed that immune pathology of these two entities is different from CD. In contrast, TCR sequences from ACD, CD-GFD or patients with RCD type II included public TCR sequences that previously had been associated with CD. It is worth noting that the most dominant of these public sequences had a frequency of <0.04% and approaches the previously reported frequency range of gliadin-specific CD4<sup>+</sup> T-cells, which was determined by a totally unrelated approach, that is, tetramer staining of isolated intestinal T-cells (0.1%–1.2% of CD4<sup>+</sup> T-cells).<sup>31</sup> Moreover, this indicates that none of the public CDR3 peptide sequences had expanded to a dominant clonotype in any of the CD groups (see online supplementary table S1)—this also includes public, virus-specific TCR rearrangements (see online supplementary figure S7 and supplementary table S2).<sup>32 33 34</sup> Instead, it is justified to assume that gliadin-independent T-cell clones dominate the duodenal mucosa in RCD. Since all analysis were carried out using a cut-off value (frequency >0.03%), analysis was repeated without this cut-off. This did not result in the detection of RCD type II clonotypes in the other CD types.

## DISCUSSION

CD is an autoimmune, T-cell-driven disease that causes mucosal atrophy in the small intestine. Interestingly, the course of disease is heterogeneous. While the majority of CD is curable by a strict GFD, two types of refractory CD (RCD type I and II) exist. Whereas RCD type I can be controlled by immunosuppressive treatment in most cases, RCD type II, characterised by clonally related intraepithelial T-cells, progresses in up to 50% to an EATL within 5 years.<sup>5 6</sup> Since the prognosis of a patient having progressed to overt T-cell lymphoma is often fatal, it is essential to reliably recognise patients at-risk early in their disease course in order to introduce timely immunosuppressant in RCD type I or cytoreductive treatment in RCD type II. Identification of clonal T-cell populations in the small intestinal mucosa of patients with refractory CD is currently performed either by conventional immunohistology, analysis of TCR gene rearrangement or by flow cytometric analysis of isolated intestinal lymphocytes after immunostaining.<sup>2 3</sup>

Previous attempts to decipher the TCR repertoire in CD were hampered by the fact that merely a size distribution of TCR $\beta$  rearrangements (GeneScan analysis) or a distribution and usage of rearranged TCR V $\beta$ -segments (spectratyping) could be determined. With the recent development of HTS technologies that allow parallel analysis of huge amounts of heterogeneous sequence mixtures, these limitations have been overcome. By combination of the TCR $\beta$  multiplex PCR for simultaneous amplification of all TCR $\beta$  rearrangements in a given sample and



subsequent HTS, the plethora of PCR products can be resolved to the level of individual sequences. Moreover, HTS generated copy numbers of identical sequences contributed to objectify quantitative statements in molecular pathology. In our study, DNA extracted from small intestinal mucosa biopsies was analysed by a previously established TCR $\beta$ -HTS approach in order (i) to compare patterns of CD-specific and RCD-specific TCR $\beta$  gene rearrangements with those of individuals with healthy mucosa, (ii) to better understand the impact of dominant CDR3 sequences on the disease course and (iii) to identify the abundance of TCR $\beta$  motifs previously shown to be associated with the coeliac immune reaction.<sup>24</sup>

### T-cell diversity in small intestinal mucosa

The first question we addressed refers to the frequency distribution of the TCR $\beta$  rearrangements found in small intestinal tissue. The frequency of the most dominant TCR $\beta$  rearrangement in control tissue was 6.8% on average (range 2.0% and 17.7%). This contrasts with the frequency of dominant TCR $\beta$  rearrangements in peripheral blood T-cells of healthy donors, where an average of 0.8% of the purified CD4+ fraction and 4.7% in the purified CD8+ fraction were described as highest frequencies.<sup>24</sup> This is in accordance with previous studies using more indirect measures for T-cell variability, which suggested the existence of an oligoclonal TCR in the small intestine of humans and mice.<sup>35–36</sup> Particularly studies in young children and in mice preweaning and postweaning have revealed that an initially polyclonal TCR becomes oligoclonal, presumably by gut microbiota-dependent expansion of specific clones that undergo proliferation, by clonal deletion, or by selective clonal emigration.<sup>37–38</sup> In our study,  $5 \pm 2$  TCR gene rearrangements within an individual sample had a frequency of  $>1\%$ , pointing to the presence of multiple dominant T-cell populations within the mucosa. Interestingly, all of these apparently oligoclonal TCR sequences were unique when compared with other healthy samples and in the study participants, who were studied repeatedly, clonal T-cell populations revealed a remarkable stability.

### Reduced TCR diversity in RCD type II; predictive function of TCR $\beta$ -HTS

The major objective of our study was the comparison of the TCR repertoires of RCD type II (as this is regarded as pre-EATL) with RCD type I as well as active and inactive CD (ie, patients on a GFD). For RCD type II, one would expect the TCR repertoire to be reduced, since this condition is known to be associated with dominant TCR rearrangements.<sup>9–12, 39</sup> This was confirmed for RCD type II with strikingly dominant TCR sequences, that is, individual TCR $\beta$  clonotypes found at frequencies as high as 70% and a significantly reduced number of distinct TCR $\beta$  gene rearrangements. This was also reflected by a significant reduction in the Shannon diversity index, which is a well-established measure of diversity.<sup>27</sup>

By distinction of the RCD type II group with or without progression from EATL, significant higher frequencies of dominant TCR rearrangements were detectable in patients with progressive disease. This is exemplified by the finding that all patients of our cohort (with the exception of one) with a TCR $\beta$  clonotype frequency  $>50\%$  showed progression to an overt T-cell lymphoma. This observation might be especially relevant, since a biomarker that has predictive function is currently lacking. Conventional diagnostics including GeneScan and CD8/CD3 immunohistochemistry as well as flow cytometry are established parameters in the diagnosis of RCD, but display no prognostic value regarding clinical outcome.

### Identity of TCR $\beta$ sequences in RCD type II

The presence of individual and highly frequent T-cell clones in patients with RCD type II poses the question, if these TCRs represent private or public sequences. Using the HTS-generated TCR sequence information, we present evidence that these dominant T-cell clones are private and unique for each patient with RCD type II. Of note, no disease group-specific, public sequences were found even at frequencies as low as 0.05%.

Moreover, no previously published 'coeliac' TCR sequences (CDR3 sequences from gliadin-restricted T-cells) were detected within the subset of dominant clonotypes. In fact, we were able to identify gliadin-dependent TCR rearrangements, however, exclusively within clonotypes of very low frequencies. On the one hand, this exemplifies the extraordinary sensitivity of TCR $\beta$ -HTS, since it was capable to identify T-cell populations at an extremely high level of scarcity. On the other hand, this finding appears to be at variance to previous studies regarding the frequency of these clonotypes. However, in previous studies, in vitro propagated T-cells were analysed, whereas our data reflect the unbiased in vitro situation present in primary tissue specimens.<sup>26–30</sup> Furthermore, gliadin-dependent sequences detectable in our samples were present most often in patients with ACD, whereas RCD type II and inactive CD showed less overlaps with public TCR sequences. The rareness of matching events ( $<0.04\%$  of public TCR clonotypes) is in concordance with previously published frequencies and also reflects the fact that unstimulated T-cells (ie, T-cells not expanded with gliadin in vitro) were tested. Since public sequences were never detectable in the major clonotypes of patients with RCD type II, it is justified to conclude that these T-cell clones are not derived from gliadin-restricted T-cells and that these sequences appear to be unique (ie, *private*) within the RCD type II group. In this regard, data from Tjon *et al*<sup>40</sup> support the idea of a non-functional TCR in RCD type II, as they realised that TCR-CD3 complexes from RCD type II aberrant lymphocytes are subject to a defective intracellular trafficking, thereby rendering the receptor in an intracellular compartment, which is inaccessible for conventional TCR stimuli. Furthermore, it was shown for ACD that CD3+ NKG2C+ intraepithelial cytotoxic T-cells oligoclonally expand leading to a high percentage of T-cells expressing the same TCRBV segments.<sup>41</sup> The dominant TRBV segments detected were unique for each patient, which is in agreement with the 'privacy' of TCR sequences identified in our study.

### Validity of TCR $\beta$ -HTS data

The reliability of data generated by our highly sensitive TCR $\beta$ -HTS approach is indicated by the following considerations: (i) independent mucosal samples of the same individual display a remarkable consistency of the patient-associated T-cell diversity (Shannon diversity index) and (ii) identical dominant T-cell populations persist with a reproducible frequency over many years in the respective patients. Previous studies were unable to ascertain this finding at the sequence level but drew similar conclusions based on the length distribution of the dominant PCR products as revealed by GeneScan or spectratyping analyses.

Finally, we want to sum up potentials and major advantages of using the TCR $\beta$ -HTS in RCD as proposed in this study:

1. TCR $\beta$ -HTS reveals the identity of clonal small intestinal lymphocytes in RCD type II and thereby has indicated a role for gliadin-independent lymphocytes in RCD type II.
2. As a prognostic tool, TCR $\beta$ -HTS might identify those patients with RCD type II, who are at an immediate risk to develop EATL.

3. Secondary to its high sensitivity on the one hand and its potential to specifically and reproducibly assign clones to TCR-CDR3 sequences on the other hand, TCR $\beta$ -HTS might become a valuable diagnostic tool for future treatment studies in RCD type II. While the conventional GeneScan analysis might fail to detect a pretherapeutically identified clone after effective treatment secondary to its disappearance in the 'background noise' (and thereby might falsely label the post-treatment situation 'successful elimination of the clone'), HTS will likely pick up even residual lymphocyte populations expressing the respective TCR and determine a copy number allowing quantitative judgement of treatment efficacy.

**Contributors** Study concept and design: JR, BS, AR, SD, MH, MS; acquisition of data: JR, SM, AS, SH, MH; analysis and interpretation of data: JR, KZ, KJ, SD, MH, MS; manuscript preparation: JR, BS, AR, SD, MH, MS, statistical analysis: JR, KZ.

**Funding** Moniek Crammeh, Hans-Henning Müller, Erika Berg, Claudia Heldt, Ulrike Dethlefs for active support. We received funding from the Deutsche Zöliakie Gesellschaft, Dr. Schär and the Clinical Scientist programme of the Charité. The sponsors were neither involved in study design nor the collection, analysis or interpretation of data.

**Competing interests** None declared.

**Ethics approval** EA4/016/14.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

## REFERENCES

- 1 Daum S, Cellier C, Mulder CJ. Refractory coeliac disease. *Best Pract Res Clin Gastroenterol* 2005;19:413–24.
- 2 Malamut G, Meresse B, Cellier C, *et al*. Refractory coeliac disease: from bench to bedside. *Semin Immunopathol* 2012;34:601–13.
- 3 Nijeboer P, van Wanrooij RL, Tack GJ, *et al*. Update on the diagnosis and management of refractory coeliac disease. *Gastroenterol Res Pract* 2013;2013:518483.
- 4 Schmitz F, Tjon JM, Lai Y, *et al*. Identification of a potential physiological precursor of aberrant cells in refractory coeliac disease type II. *Gut* 2013;62:509–19.
- 5 Daum S, Ipczynski R, Schumann M, *et al*. High rates of complications and substantial mortality in both types of refractory sprue. *Eur J Gastroenterol Hepatol* 2009;21:66–70.
- 6 Malamut G, Afchain P, Verkarre V, *et al*. Presentation and long-term follow-up of refractory coeliac disease: comparison of type I with type II. *Gastroenterology* 2009;136:81–90.
- 7 Rubio-Tapia A, Kelly DG, Lahr BD, *et al*. Clinical staging and survival in refractory coeliac disease: a single center experience. *Gastroenterology* 2009;136:99–107; quiz 352–3.
- 8 Verbeek WH, Schreurs MW, Visser OJ, *et al*. Novel approaches in the management of refractory coeliac disease. *Expert Rev Clin Immunol* 2008;4:205–19.
- 9 Daum S, Weiss D, Hummel M, *et al*. Frequency of clonal intraepithelial T lymphocyte proliferations in enteropathy-type intestinal T cell lymphoma, coeliac disease, and refractory sprue. *Gut* 2001;49:804–12.
- 10 Perfetti V, Brunetti L, Biagi F, *et al*. TCRbeta clonality improves diagnostic yield of TCRgamma clonality in refractory coeliac disease. *J Clin Gastroenterol* 2012;46:675–9.
- 11 Verbeek WH, von Blomberg BM, Scholten PE, *et al*. The presence of small intestinal intraepithelial gamma/delta T-lymphocytes is inversely correlated with lymphoma development in refractory coeliac disease. *Am J Gastroenterol* 2008;103:3152–8.
- 12 Tack GJ, van Wanrooij RL, Langerak AW, *et al*. Origin and immunophenotype of aberrant IEL in RCDII patients. *Mol Immunol* 2012;50:262–70.
- 13 Meresse B, Cerf-Bensussan N. Innate T cell responses in human gut. *Semin Immunol* 2009;21:121–9.
- 14 Davis MM. The evolutionary and structural 'logic' of antigen receptor diversity. *Semin Immunol* 2004;16:239–43.
- 15 Arstila TP, Casrouge A, Baron V, *et al*. A direct estimate of the human alpha beta T cell receptor diversity. *Science* 1999;286:958–61.
- 16 Carbone F, Grollet-Bioul L, Brouet JC, *et al*. Are complicated forms of coeliac disease cryptic T-cell lymphomas? *Blood* 1998;92:3879–86.
- 17 Brousse N, Verkarre V, Patey-Mariaud de Serre N, *et al*. Is complicated coeliac disease or refractory sprue an intestinal intra-epithelial cryptic T-cell lymphoma? *Blood* 1999;93:3154–5.
- 18 Bagdi E, Diss TC, Munson P, *et al*. Mucosal intra-epithelial lymphocytes in enteropathy-associated T-cell lymphoma, ulcerative jejunitis, and refractory coeliac disease constitute a neoplastic population. *Blood* 1999;94:260–4.
- 19 Cellier C, Delabesse E, Helmer C, *et al*. Refractory sprue, coeliac disease, and enteropathy-associated T-cell lymphoma. French Coeliac Disease Study Group. *Lancet* 2000;356:203–8.
- 20 Holtmeier W, Hennemann A, May E, *et al*. T cell receptor delta repertoire in inflamed and noninflamed colon of patients with IBD analyzed by CDR3 spectratyping. *Am J Physiol Gastrointest Liver Physiol* 2002;282:G1024–34.
- 21 van Dongen JJ, Langerak AW, Brüggemann M, *et al*. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia* 2003;17:2257–317.
- 22 Marsh MN. Clinical and pathological spectrum of coeliac disease. *Gut* 1993;34:1740; author reply 1.
- 23 Oberhuber G, Caspary WF, Kirchner T, *et al*. [Study Group of Gastroenterological Pathology of the German Society of Pathology. Recommendations for coeliac disease/sprue diagnosis]. *Z Gastroenterol* 2001;39:157–66.
- 24 Ritter J, Seitz V, Balzer H, *et al*. Donor CD4T cell diversity determines virus reactivation in patients after HLA-matched allogeneic stem cell transplantation. *Am J Transplant* 2015;15:2170–9.
- 25 Broughton SE, Petersen J, Theodossis A, *et al*. Biased T cell receptor usage directed against human leukocyte antigen DQ8-restricted gliadin peptides is associated with coeliac disease. *Immunity* 2012;37:611–21.
- 26 Qiao SW, Christophersen A, Lundin KE, *et al*. Biased usage and preferred pairing of alpha- and beta-chains of TCRs specific for an immunodominant gluten epitope in coeliac disease. *Int Immunol* 2014;26:13–19.
- 27 Shannon CE. The mathematical theory of communication. 1963. *MD Comput* 1997;14:306–17.
- 28 Dahal-Koirala S, Risnes LF, Christophersen A, *et al*. TCR sequencing of single cells reactive to DQ2.5-glia- $\alpha$ 2 and DQ2.5-glia- $\omega$ 2 reveals clonal expansion and epitope-specific V-gene usage. *Mucosal Immunol* 2016;9:587–96.
- 29 Britanova OV, Putintseva EV, Shugay M, *et al*. Age-related decrease in TCR repertoire diversity measured with deep and normalized sequence profiling. *J Immunol* 2014;192:2689–98.
- 30 Qiao SW, Råki M, Gunnarsen KS, *et al*. Posttranslational modification of gluten shapes TCR usage in coeliac disease. *J Immunol* 2011;187:3064–71.
- 31 Bodd M, Råki M, Bergseng E, *et al*. Direct cloning and tetramer staining to measure the frequency of intestinal gluten-reactive T cells in coeliac disease. *Eur J Immunol* 2013;43:2605–12.
- 32 Lim A, Trautmann L, Peyrat M-A, *et al*. Frequent contribution of T cell clonotypes with public TCR features to the chronic response against a dominant EBV-derived epitope: application to direct detection of their molecular imprint on the human peripheral T cell repertoire. *J Immunol* 2000;165:2001–11.
- 33 Miconnet I, Coste I, Beermann F, *et al*. Cancer vaccine design: a novel bacterial adjuvant for peptide-specific CTL induction. *J Immunol* 2001;166:4612–19.
- 34 Price DA, Brenchley JM, Ruff LE, *et al*. Avidity for antigen shapes clonal dominance in CD8+ T cell populations specific for persistent DNA viruses. *J Exp Med* 2005;202:1349–61.
- 35 Balk SP, Ebert EC, Blumenthal RL, *et al*. Oligoclonal expansion and CD1 recognition by human intestinal intraepithelial lymphocytes. *Science* 1991;253:1411–15.
- 36 Blumberg RS, Yockey CE, Gross GG, *et al*. Human intestinal intraepithelial lymphocytes are derived from a limited number of T cell clones that utilize multiple V beta T cell receptor genes. *J Immunol* 1993;150:5144–53.
- 37 Holtmeier W, Witthöft T, Hennemann A, *et al*. The TCR-delta repertoire in human intestine undergoes characteristic changes during fetal to adult development. *J Immunol* 1997;158:5632–41.
- 38 Probert CS, Williams AM, Stepankova R, *et al*. The effect of weaning on the clonality of alpha beta T-cell receptor T cells in the intestine of GF and SPF mice. *Dev Comp Immunol* 2007;31:606–17.
- 39 Cellier C, Patey N, Mauvieux L, *et al*. Abnormal intestinal intraepithelial lymphocytes in refractory sprue. *Gastroenterology* 1998;114:471–81.
- 40 Tjon JM, Verbeek WH, Kooy-Winkelaar YM, *et al*. Defective synthesis or association of T-cell receptor chains underlies loss of surface T-cell receptor-CD3 expression in enteropathy-associated T-cell lymphoma. *Blood* 2008;112:5103–10.
- 41 Meresse B, Curran SA, Ciszewski C, *et al*. Reprogramming of CTLs into natural killer-like cells in coeliac disease. *J Exp Med* 2006;203:1343–55.