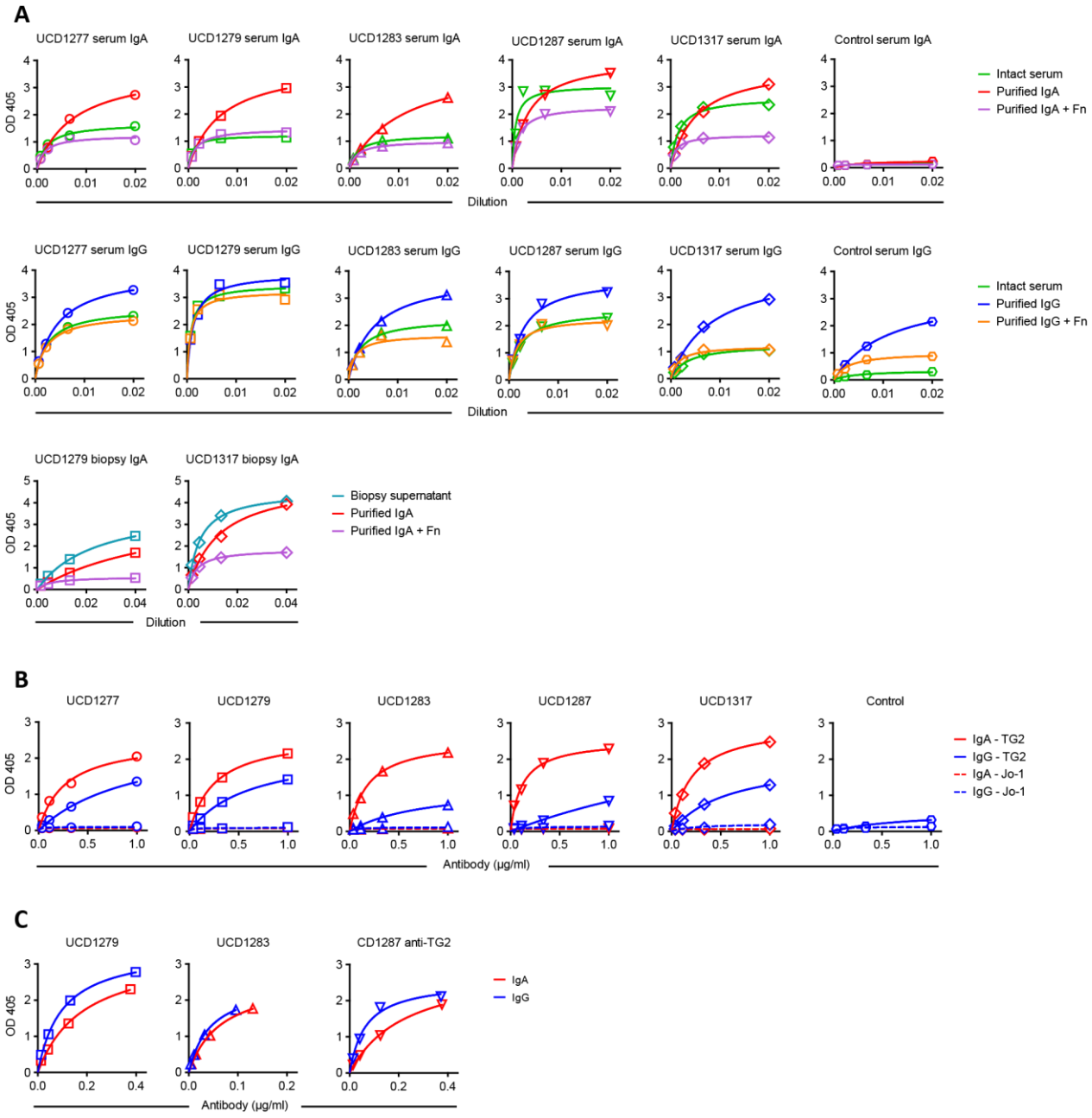
Cell Reports, Volume *20*

Supplemental Information

Strong Clonal Relatedness between Serum and Gut IgA despite Different Plasma Cell Origins

Rasmus Iversen, Omri Snir, Maria Stensland, José E. Kroll, Øyvind Steinsbø, Ilma R. Korponay-Szabó, Knut E.A. Lundin, Gustavo A. de Souza, and Ludvig M. Sollid

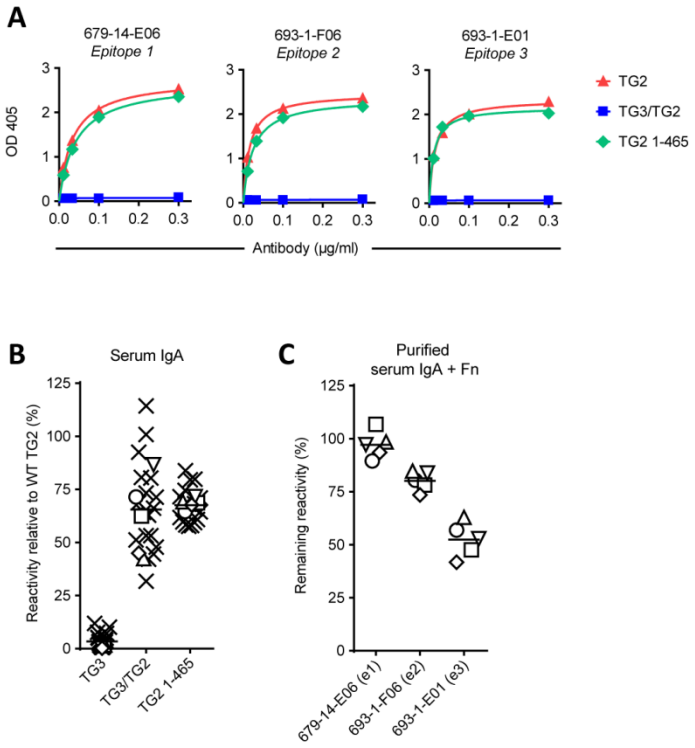
Figure S1. Assessment of purified anti-TG2 IgA or IgG reactivity by ELISA. Related to Figures 2 and 5



(A) Total IgA or IgG purified from celiac disease patient or control sera or from supernatants of cultured gut biopsies were incubated in TG2-coated microtiter plates in the presence or absence of fibronectin (Fn, 5 µg/ml in the highest concentration). Whole sera or culture supernatants were added in corresponding dilutions for comparison, and bound IgA or IgG was detected using isotype-specific goat anti-human antibodies. Supplementation of purified antibodies with Fn decreases their TG2 reactivity and make them mimic the behavior of whole sera, indicating that binding between TG2 and serum Fn blocks a substantial number of antibody binding sites. For the control subject, TG2 reactivity is present in the purified IgG fraction but not in whole serum or in purified IgA, indicating that purification of IgG alters the binding properties of some antibody molecules. The reason for this behavior of IgG is most likely

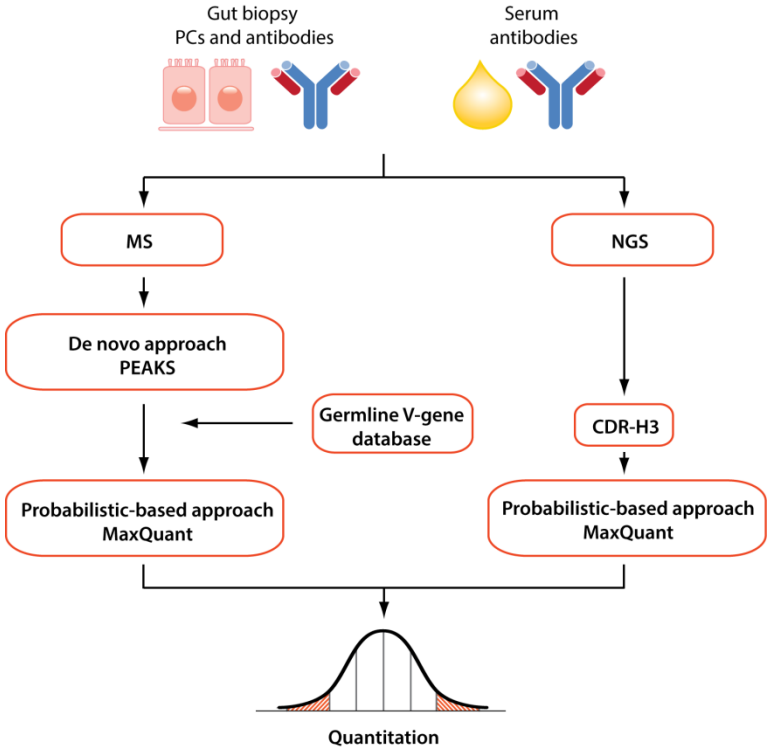
that exposure to low pH during purification alters the structure of some IgG molecules, allowing them to interact with new antigens (Dimitrov et al., 2013).
(B) Binding of affinity-purified anti-TG2 antibodies isolated from total serum IgA or IgG preparations. Histidyl tRNA synthetase (Jo-1) was included as an irrelevant antigen to test the specificity of the purified antibodies. In all cases, the purified IgG antibodies have lower apparent affinity than the IgA antibodies. Notably, TG2-binding IgG, but not IgA, could be isolated from non-celiac control serum, but these antibodies had very low affinity.
(C) TG2 binding of affinity-purified IgA and IgG isolated from whole sera without initial separation into total IgA and IgG fractions. When purifying the TG2-binding fraction from whole sera, there is no clear difference between IgA and IgG affinities.

Figure S2. Reactivity of antibodies with different regions of TG2 assessedby ELISA. Related to Figure 2



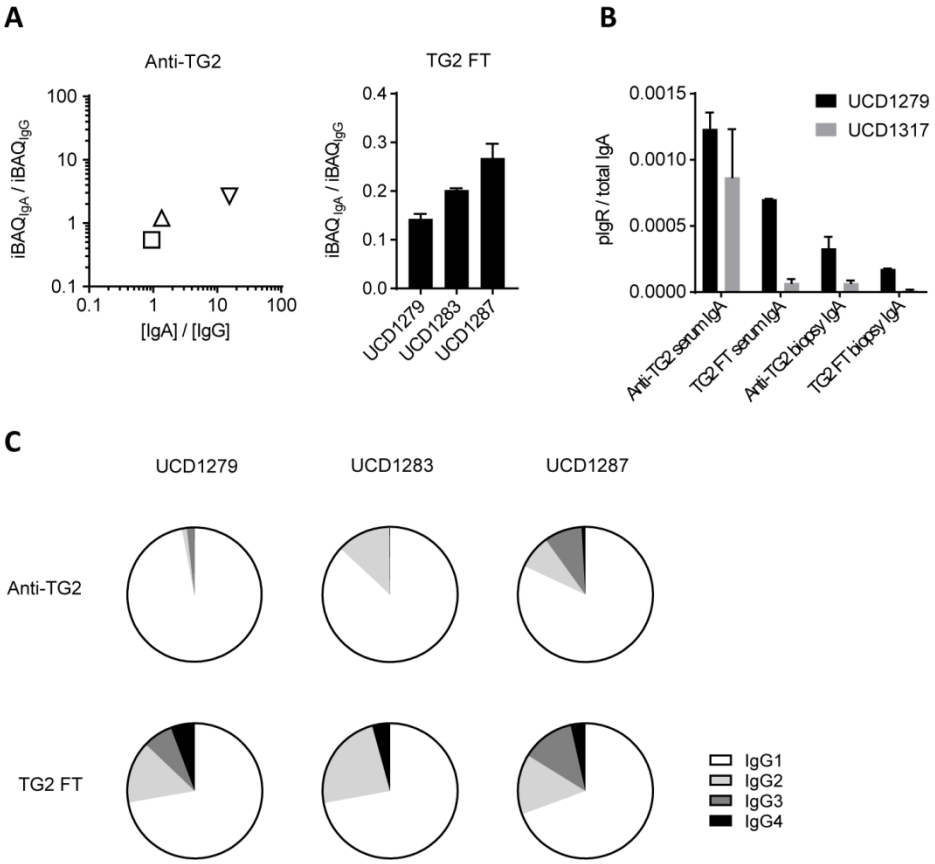
(A) Saturation binding curves for anti-TG2 mAbs targeting different epitopes to TG2 domain variants as determined by ELISA. All three mAbs lose reactivity when the N-terminal domain of TG2 is replaced with the N-terminal domain of TG3 (TG3/TG2) but show intact reactivity when the two C-terminal domains are removed (TG2 1-465). (B) Reactivity of IgA in whole serum samples from adult celiac disease patients (n = 23) with TG2 domain variants as shown for purified serum IgA in Figure 2A. No reactivity was observed against full-length TG3, indicating that binding observed with TG3/TG2 is mediated solely by TG2 epitopes located outside of the N-terminal domain. (C) Ability of the indicated anti-TG2 mAbs to compete with serum IgA for TG2 binding in the presence of added fibronectin (Fn, 1 µg/ml). Total serum IgA was purified from celiac disease patients (n = 5), and the reactivity is given relative to the signal obtained in the absence of mAbs but with added Fn. In the presence of Fn, addition of epitope-1 mAb does not have an effect on binding of serum IgA (compare with Figure 2B). Horizontal lines indicate means.

Figure S3. Overview of the quantitative proteomics approach and software used for antibody analysis. Related to Table 1 and Figures 3 and 4



Antigen-specific antibodies were either isolated from cultured gut biopsies or serum samples of celiac disease patients and analyzed by LC-MS/MS. Peptides derived from V-gene segments and CDR-H3 were identified by two separate procedures. Heavy and light chain V-gene peptides were identified using a database consisting of germline amino acid sequences combined with mutated peptides discovered by de novo MS sequencing. In order to identify CDR-H3 peptides, antigen-specific PCs were sorted from gut biopsies, and heavy chain V-region sequences were obtained by NGS.

Figure S4. Quantification of antibody class distributions by LC-MS/MS. Related to Figure 6



(A) The ratio between IgA and IgG antibodies was determined based on iBAQ values obtained with MaxQuant software in affinity-purified anti-TG2 or TG2 flow-through (FT) fractions obtained from whole sera of three celiac disease patients. For the TG2-binding fraction (left panel), the IgA:IgG ratio determined by LC-MS/MS (y-axis) was compared to the ratio obtained by measuring the IgA and IgG concentrations by ELISA (x-axis). At higher IgA:IgG ratios, the LC-MS/MS method appears to overestimate the contribution of IgG. The ratios obtained from the TG2 FT fractions (right panel) are within the expected range (Gonzalez-Quintela et al., 2008). Error bars indicate SD based on duplicate LC-MS/MS runs.

(B) Quantification of the amount of polymeric Ig receptor (pIgR) associated with the TG2-binding or TG2 FT fractions of purified total IgA from sera or gut biopsies of two celiac disease patients. The quantification is based on iBAQ values and is given relative to the summed iBAQ values of the IgA1 and IgA2 constant regions in each sample. This ratio should give an indication of the contribution of secretory IgA to each population. Error bars represent SD based on duplicate (UCD1279) or triplicate (UCD1317) LC-MS/MS runs.

(C) The distribution of IgG between individual subclasses based on iBAQ values measured in affinity-purified anti-TG2 and TG2 FT fractions obtained from whole sera of celiac disease patients. In all cases, the IgG1 subclass is overrepresented among TG2-specific antibodies compared to the TG2 FT fraction.

Table S1. Identification of Ig proteins in affinity-purified gut biopsy anti-TG2 and TG2 flow-through IgA of a representative celiac disease patient (UCD1317) by LC-MS/MS. Related to Figure 3

	Anti-TG2		TG2 flow-through	
	Germline only	Germline + de novo	Germline only	Germline + de novo
Protein groups identified[a]	121	147	116	144
Proteins per group (avg.)	2.72	2.35	2.91	2.36
Peptides per group (avg.)	4.03	41.4	3.55	48.2
Unique peptides per group (avg.)	1.73	5.57	1.82	6.75

[a] Each protein group comprises all proteins in the database that share the same identified peptides. The germline only database contains all Ig V-gene segments and constant regions available from http://www.imgt.org. Entries in this database were supplemented with peptide sequences harboring single-amino acid substitutions identified by de novo sequencing to obtain the germline + de novo database. Typically, individual alleles could not be distinguished and were placed in the same protein group, whereas different V-gene segments often could be separated. Due to the large degree of similarity between V-gene segments, many peptides were shared between groups, but each group is represented by at least one unique peptide.

Table S2. List of V-gene segments overrepresented among purified anti-TG2 or anti-DGP serum IgA of celiac disease patients. Related to Figure 3

	Anti-TG2[a]		Anti-DGP[a]
	5 of 5 patients	4 of 5 patients	2 of 2 patients
IGHV1-69		**X**	
IGHV3-7		X	**X**
IGHV3-15			**X**
IGHV3-30			X
IGHV3-43		**X**	
IGHV3-74			**X**
IGHV5-51	**X**		
IGKV1-5	**X**		
IGKV1-12	**X**		
IGKV1-16		**X**	
IGKV1-17	**X**		
IGKV1-27		**X**	
IGKV1-39	**X**		
IGKV3-20	**X**		
IGKV4-1			**X**
IGLV1-44	**X**		
IGLV1-47	**X**		
IGLV2-11		X	
IGLV2-14	X		
IGLV4-69			**X**
IGLV5-45	**X**		

[a] V-gene segments that were significantly more abundant in antigen-specific compared to non-specific serum IgA fractions across celiac disease patients are indicated. Symbols in bold indicate V-gene segments previously found to be used by TG2-specific (Di Niro et al., 2012; Snir et al., 2015; Roy et al., 2017) and DGP-specific (Steinsbø et al., 2014) gut PCs.

Table S3. Anti-TG2 CDR-H3 peptides detected by LC-MS/MS in affinity-purified anti-TG2 or flow-through fractions isolated from IgA of serum or gut biopsy secretions. Related to Figure 4 and Table 1.
See Excel file

Table S4. Anti-DGP CDR-H3 peptides detected by LC-MS/MS in affinity-purified anti-DGP or flow-through fractions isolated from serum IgA. Related to Figure 4 and Table 1.
See Excel file

Table S5. Anti-TG2 CDR-H3 peptides detected by LC-MS/MS in affinity-purified anti-TG2 or flow-through fractions isolated from whole sera. Related to Figure 5.
See Excel file

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

ELISAs. Recombinant human proteins were coated in microtiter plates at 3 µg/ml in TBS. For comparison of TG2 and DGP reactivity, biotinylated TG2 (1 µg/ml) or biotinylated DGP (biotin-GSGSGS-PLQPEQPFP, obtained from GL Biochem, 50 nM) was attached to streptavidin-coated plates (Thermo) in TBS containing 0.1% (v/v) Tween 20 (TBST). For comparison of reactivity to different TG2 variants, the enzymes were associated with the irreversible inhibitor Ac-P(DON)LPF-NH_2 (Zedira) to ensure that they had equal and homogenous conformations prior to coating (Iversen et al., 2014). To this end, the TG2 variants were incubated with 0.5 mM of the inhibitor for 20 min at room temperature before the reaction was started by addition of $CaCl_2$ to a final concentration of 5 mM and further incubation for 30 min. Incubation with antibodies were carried out at 37°C in 3% (w/v) BSA/TBST followed by detection of bound IgA or IgG using alkaline phosphatase-conjugated goat anti-human IgA (Sigma) or goat anti-human IgG (Southern Biotech). To compare antibody reactivity against different TG2 variants or in the presence or absence of competitor, initial titration of all sera or purified antibodies was carried out, and a concentration falling within the linear range of the reference assay (using wild-type TG2 in the absence of competitor) was in each case picked for determination of relative binding levels based on changes in OD value. Competition between polyclonal antibodies and anti-TG2 mAbs was assessed by incubating coated TG2 with 10 µg/ml anti-TG2 IgG1 mAbs for 30 min at 37°C, before polyclonal antibody preparations were added without removing the mAbs, and incubation was continued for 1 h. The level of bound polyclonal IgA was subsequently detected as above. Concentrations of IgA and IgG in TG2-binding fractions purified from whole sera were determined by incubation with goat anti-human IgA (Sigma) or rabbit anti-human IgG (Dako), each coated at 5 µg/ml, and detection of bound antibodies using alkaline phosphatase-conjugated goat anti-human IgA or rabbit anti-human IgG (Abcam). Obtained signals were compared to a standard curve obtained using purified total serum IgA or IgG.

Antibody purification. Total IgG was purified from 20 ml of serum or plasma by diluting the sample in 20 mM NaP_i, pH 7.4 and passing it over a protein G column (GE Healthcare). Bound IgG was eluted with 0.1 M glycine-HCl, pH 2.5 and immediately neutralized with 1 M Tris-HCl, pH 9. IgA was purified from the protein G flow-through fraction using peptide M-agarose (Invivogen) followed by elution of bound IgA as described for purification of IgG. IgA secreted from duodenal biopsies was purified from culture supernatants in the same way. To isolate TG2-binding IgA and IgG, we mixed up to 10 ml of purified antibodies or 1 ml of whole serum diluted ten times in TBS with 400 µl 50% (v/v) streptavidin-agarose (Novagen) that had been pre-associated with 100 µg biotinylated TG2 in a Poly-Prep chromatography column (Bio-Rad), followed by incubation for 1 h under constant rotation at room temperature. A new column with fresh TG2 protein was prepared for each sample. The flow-through fraction was collected, and a small part to be analyzed by LC-MS/MS was desalted using a Zeba Spin column (Thermo). The column containing TG2-bound antibodies was centrifuged at 1700 rpm for 1 min and washed three times with 1 ml TBS and three times with 1 ml water followed by elution three times with 200 µl 20 mM HCl. The column was centrifuged between each washing and elution step to ensure proper clearing of liquid. A small part of the eluate to be used in binding studies was immediately neutralized with NaOH and diluted in TBS. DGP-binding antibodies were purified from the flow-through fraction of the anti-TG2 purification step in the same way as described above, using streptavidin-agarose pre-incubated with 10 µM biotinylated DGP.

LC-MS/MS. For liquid chromatography separation, we used an EasySpray column (C18, 2 µm beads, 100 Å, 75 µm inner diameter) (Thermo) capillary of 25 cm bed length and a flow rate of 0.3 µl/min. The applied solvent gradient was from 5% (v/v) B to 30% B over 60 min, followed by washing with 90% B for 20 min. Solvent A was aqueous 0.1% (v/v) formic acid, and solvent B was acetonitrile with 0.1 % formic acid. The column temperature was kept at 60°C. The mass spectrometer was operated in the data-dependent mode to automatically switch between MS and MS/MS acquisition. Survey full scan MS spectra (from m/z 400 to 1200) were acquired in the Orbitrap with resolution R = 70,000 at m/z 200 (after accumulation to a target of 3,000,000 ions in the quadrupole). The method used allowed sequential isolation of the most intense multiply-charged ions, up to ten, depending on signal intensity, for fragmentation on the HCD cell using high-energy collision dissociation at a target value of 100,000 charges or maximum acquisition time of 100 ms. MS/MS scans were collected at 17,500 resolution at the Orbitrap cell. Target ions already selected for MS/MS were dynamically excluded for 30 s. General mass spectrometry conditions were: electrospray voltage, 2.1 kV; no sheath and auxiliary gas flow, heated capillary temperature of 250°C, normalized HCD collision energy 25%. Ion selection threshold was set to 5e4 counts. Isolation width of 3.0 Da was used.

De novo peptide sequencing. Peptide sequences were identified using PEAKS Studio (Zhang et al., 2012) with the following parameters: methionine oxidation as variable modification and cysteine carbamidomethylation as fixed modification, trypsin without proline restriction as enzyme option with three allowed miscleavages. Precursor and fragment ion error tolerances were set to 20 ppm and 0.05 Da respectively. The SPIDER (Sequence Tag Homology/Mutated Peptides) function in PEAKS was used to find mutations. Peptide confidence cutoff was set to FDR 1% and ALC (Average local confidence) 80% in order to only include peptides with high-quality MS spectra. We then developed an in-house Perl script which collected de novo-sequenced peptides from the PEAKS output and added them to the aligned protein entry in the germline database. Peptides were added N-terminally to the reference germline sequence in order to keep tryptic site structure. The modified germline database was then used for a regular, probabilistic-based peptide search, using MaxQuant (version 1.5.2.8)(Cox and Mann, 2008) for improved protein identification, sequence coverage and label-free quantitation. Methionine oxidation and cysteine carbamidomethylation were included as variable and fixed modifications, respectively. First search error window was set to 20 ppm and mains search error was set to 6 ppm at MS level. Trypsin without proline restriction was used as enzyme option with two allowed miscleavages. Minimal unique peptides were set to 1, and FDR allowed was 1% for peptide and protein identification. Generation of reversed sequences was selected to assign FDR rates. The levels of individual V-gene segments in antigen-specific and non-specific fractions were compared by *t*-test analysis using Perseus software (Tyanova et al., 2016).

Isolation of TG2- and DGP-reactive PCs and sequencing of heavy chain V regions. PCs reactive with TG2 or DGP were stained using multimers as previously described (Di Niro et al., 2012; Steinsbø et al., 2014; Snir et al., 2015). In short, biotinylated TG2 and DGP were multimerized using PE- and APC-labeled streptavidin, respectively, at a 4:1 molar ratio. Single-cell suspensions from duodenal biopsies were stained with the two multimers and the following cellular lineage markers: IgA-FITC (Southern Biotech), CD3- and CD14-Brilliant Violet 570 (BioLegend), CD19-pacific blue (BD Biosciences) and CD27-PE-Cy7 (eBioscience). PCs were defined as live, large CD14/CD3-negative and with high expression of CD27. IgA PCs binding either TG2 or DGP were sorted into 25 µl PBS supplemented with 10 mM DTT and 2 U/µl RNAsin (Promega) and kept at -70°C until cDNA synthesis. For cDNA synthesis, 13.5 µl of sorted PCs extracts were incubated with 1.4 µM indexed IGHJ-rev primer (see primer list below), 0.45% (v/v) NP-40, 0.5 µl 40 U/µl and DEPC-treated water in a total volume of 25 µl at 65°C for 5 min. Next, we added 10 µl 5x RT buffer, 3.5 µl 100 mM DTT, 4.3 µl 10 mM dNTP-Mix (Promega), 0.7 µl 40 U/µl RNAsin, 0.85 µl SuperScript III (Invitrogen) and 4.5 µl DEPC-treated water, and cDNA was synthesized using the following temperature steps: 42°C for 10 min, 25°C for 10 min, 50°C for 60 min and 94°C for 5 min followed by storage at −20°C. Second-strand cDNA was synthesized using AmpliTaq Gold polymerase (Applied Biosystems) with indexed IGHV1-6 framework region (FR) 2 (van Dongen et al., 2003). Six random nucleotides were included in IGHJ and IGHV1-6 primers and were used as unique molecular identifiers (UMIs). Second-strand synthesis was carried out using the temperature steps: 95°C for 7.5 min, 52°C for 2 min, and 72°C for 10 minutes. Double-stranded cDNA was further purified using AMPure XP (Beckman Coulter) at a 1:1 ratio according to the manufacturer's instructions. Next, the second part of the Illumina adapter was connected using R1 and R2 primers. PCR was performed using Qiagen Multiplex PCR at 95°C for 15 min, x25 (95°C for 30 seconds, 60°C for 45 seconds, 72°C for 90 seconds) and 72°C for 10 minutes. Primers were previously described (Snir et al., 2015). Final *IGHV* amplicon libraries were purified and concentrated using AMPure XP and further extracted from agarose gel. Paired-end sequencing of 2x300 bp was performed using Illumina MiSeq at the Norwegian Sequencing Centre in Oslo, Norway (http://www.sequencing.uio.no). Raw sequencing data was processed using pRESTO (Vander Heiden et al., 2014). Unique sequences for which at least two copies were present were further analyzed using IMGT and Immunoglobulin Analysis Tool (IgAT) software (Rogosch et al., 2012).

Primer sequences

PCR	Name	Sequence (5´-3´)
1st	IGHV1-FR2	GGCATTCCTGCTGAACCGCTCTTCCGATCTNNNNNN(Index)CTGGGTGCGACAGGCCCCTGGACAA
1st	IGHV2-FR2	GGCATTCCTGCTGAACCGCTCTTCCGATCTNNNNNN(Index)TGGATCCGTCAGCCCCCAGGGAAGG
1st	IGHV3-FR2	GGCATTCCTGCTGAACCGCTCTTCCGATCTNNNNNN(Index)GGTCCGCCAGGCTCCAGGGAA
1st	IGHV4-FR2	GGCATTCCTGCTGAACCGCTCTTCCGATCTNNNNNN(Index)TGGATCCGCCAGCCCCCAGGGAAGG
1st	IGHV5-FR2	GGCATTCCTGCTGAACCGCTCTTCCGATCTNNNNNN(Index)GGGTGCGCCAGATGCCCGGGAAAGG
1st	IGHV6-FR2	GGCATTCCTGCTGAACCGCTCTTCCGATCTNNNNNN(Index)TGGATCAGGCAGTCCCCATCGAGAG
1st	IGHJ-Rev	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNN(Index)CTTACCTGAGGAGACGGTGACC
2nd	R1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC
2nd	R2	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTC

List of primers that were used for amplification of *IGHV*. IGHV1-FR2 – IGHV6-FR2 (forward) and IGHJ-Rev (reverse) were used for first PCR. Blue and red color indicate partial read (R)1 and partial R2 Illumina adapters, respectively. Six random nucleotides (NNNNNN) were used as unique molecular identifiers (UMIs), followed by 6 nucleotides indices and gene specific sequences. The remaining parts of Ilumina adapters were incorporated by using R1 and R2 primers in the second PCR.

SUPPLEMENTAL REFERENCES

Dimitrov, J.D., Planchais, C., Roumenina, L.T., Vassilev, T.L., Kaveri, S.V., and Lacroix-Desmazes, S. (2013). Antibody polyreactivity in health and disease: statu variabilis. J. Immunol. *191*, 993-999.

Gonzalez-Quintela, A., Alende, R., Gude, F., Campos, J., Rey, J., Meijide, L.M., Fernandez-Merino, C., and Vidal, C. (2008). Serum levels of immunoglobulins (IgG, IgA, IgM) in a general adult population and their relationship with alcohol consumption, smoking and common metabolic abnormalities. Clin. Exp. Immunol. *151*, 42-50.

Rogosch, T., Kerzel, S., Hoi, K.H., Zhang, Z., Maier, R.F., Ippolito, G.C., and Zemlin, M. (2012). Immunoglobulin analysis tool: a novel tool for the analysis of human and mouse heavy and light chain transcripts. Front. Immunol. *3*, 176.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731-740.

van Dongen, J.J., Langerak, A.W., Bruggemann, M., Evans, P.A., Hummel, M., Lavender, F.L., Delabesse, E., Davi, F., Schuuring, E., Garcia-Sanz, R., et al. (2003). 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 *17*, 2257-2317.

Vander Heiden, J.A., Yaari, G., Uduman, M., Stern, J.N., O'Connor, K.C., Hafler, D.A., Vigneault, F., and Kleinstein, S.H. (2014). pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. Bioinformatics *30*, 1930-1932.

Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., Zhang, W., Zhang, Z., Lajoie, G.A., and Ma, B. (2012). PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. Mol. Cell. Proteomics *11*, M111 010587.