Supplementary material

Kahrs CR, Chuda K, Tapia G, et.al. Enterovirus as trigger of coeliac disease: nested case-control study within prospective birth cohort.

Table of Contents

Supplementary methods 2

Supplementary references 2

Supplementary figures 3

Supplementary figure A. Coeliac disease screening process 3

Supplementary figure B. Enterovirus distribution over the whole period 4

Supplementary figure C. Age variation of enterovirus and adenovirus 6

Supplementary figure D. Seasonal variation of entero- and adenovirus 7

Supplementary tables 8

Supplementary table A. Characteristics of the whole cohort 8

Supplementary table B. Characteristics of the children with coeliac disease 9

Supplementary table C. Enterovirus types observed in this study 10

Supplementary table D. Enterovirus infectious episodes, long infections and high-quantity infections 11

Supplementary table E. Enterovirus, gluten introduction, breastfeeding and age-groups 12

Supplementary table F. Reported symptoms, infections with symptoms and coeliac disease 13

Supplementary table G. Association between reported symptoms and viral infections independent of case status 14

Supplementary table H: Adenovirus types observed in this study 15

This supplementary material has been provided by the authors to give readers additional information about their work.

Supplementary methods

Case definition and exclusion criteria

Out of the 220 children who consented to participate in our study, 17 had been diagnosed with coeliac disease before inclusion. Of the remaining 203 children, 20 had elevated coeliac disease antibodies and were retested, with 11 children having positive antibodies in the confirmatory blood sample. These were referred to a hospital for diagnostic work-up, with 10 children being confirmed as having a coeliac disease diagnosis. The one child where the diagnosis was not confirmed and the nine children with negative antibodies in the confirmatory blood sample were excluded from the current analysis. All these children were only DGP positive in the first sample, none had tTG antibodies.

Case validation and confirmation

In the parental questionnaire at inclusion of the sub-study on coeliac disease the following covariates were included; first degree relatives with coeliac disease, type of symptoms, age at onset of symptoms, age at coeliac disease diagnosis, whether diagnosis was detected by screening or clinical symptoms, biopsy and serology status. In the validation scheme sent to hospitals after end of follow up we asked for additional diagnostic information including symptoms and coeliac disease antibody levels at diagnosis, biopsy status and Marsh grade, and date of diagnosis.

Analysis of fecal samples

All available fecal supernatant samples from cases and controls were subjected to RNA and DNA co-purification using Qiagen spin columns (Qiagen, Hilden, Germany). We co-purified the total RNA and DNA using a procedure derived from the QiaAmp Viral RNA Mini kit protocol (Qiagen, Hilden, Germany). For increasing the throughput, we utilized a 96-well format with the "QiaAmp 96" plates (Qiagen) instead of individual columns, and to minimize the risk of contamination we utilized centrifugation instead of applying vacuum. West Nile virus Armored RNA as an exogeneous internal control (Asuragen, Austin, TX) and carrier RNA (Qiagen, Hilden, Germany) were added in a constant quantity to the lysis buffer, which was used in the first step of the protocol. Each extraction plate contained a minimum of six negative controls (TE buffer instead of a sample) in irregular intervals. No negative control was positive for entero- or adenovirus.

We tested enterovirus by reverse-transcriptase real-time PCR with the QuantiTect Probe RT-PCR Kit (Qiagen), using 900 nmol/L primers and 300 nmol/L probes (sequences described by Honkanen et al).[1] The combination of primers and probes reacts with an equal sensitivity to Enterovirus A—D species (i.e. members of species *Enterovirus A*, *Enterovirus B*, *Enterovirus C* and *Enterovirus D* of the genus *Enterovirus*, family *Picornaviridae*, order *Picornavirales*) but does not react with human rhinoviruses (i.e. members of species *Rhinovirus A-C*, genus *Enterovirus*). Enterovirus positivity cut-off was 10 copies / µL, and the amplification had to occur before the cycle 35. Adenovirus was tested using a previously published real-time PCR assay,[2] and the positivity threshold was likewise 10 copies / µL.

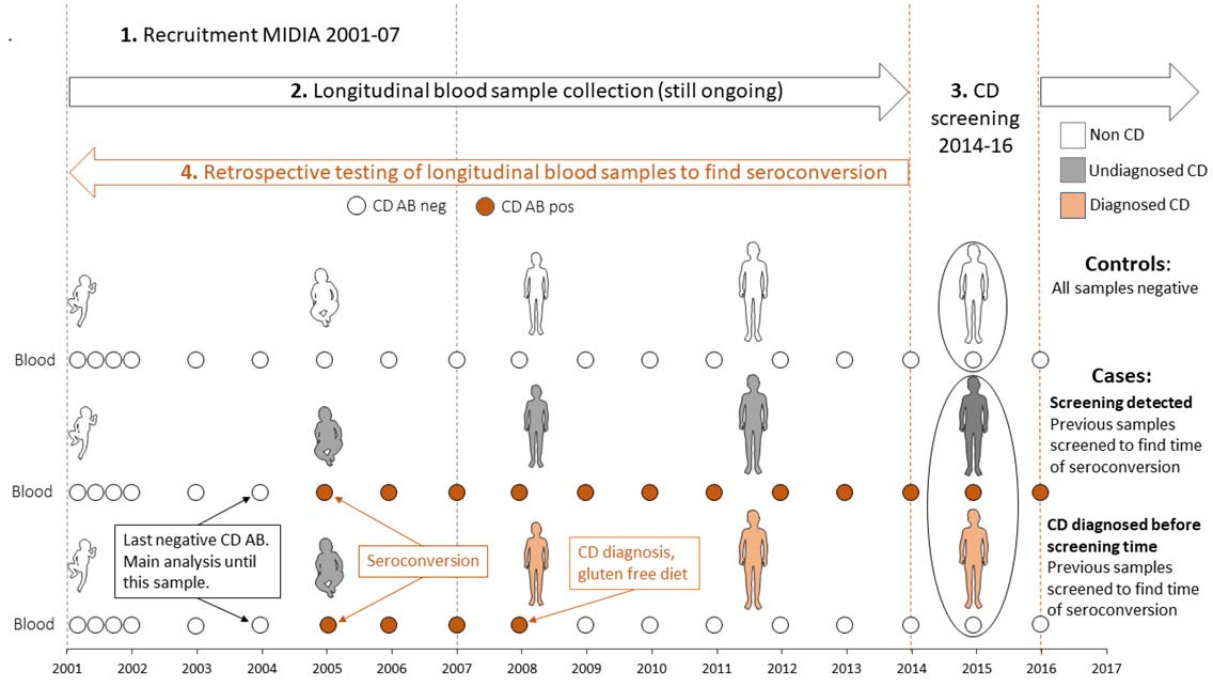
Enterovirus was genotyped using a nested reverse transcriptase PCR targeted to the VP1 gene. Primers were redesigned from the work by Nix et al to accommodate for higher variability of the genotypes,[3] and to make the reaction truly nested as compared to the semi-nested design of the original method. Genotyping of adenovirus was performed by sequencing of hexon gene spanning its hypervariable region 7. The product was amplified using a mixture of 10 forward and 6 reverse primers.

Supplementary references

1. Honkanen H, Oikarinen S, Pakkanen O, et al. Human enterovirus 71 strains in the background population and in hospital patients in Finland. J Clin Virol 2013;56:348-53.
2. Claas EC, Schilham MW, de Brouwer CS, et al. Internally controlled real-time PCR monitoring of adenovirus DNA load in serum or plasma of transplant recipients. J Clin Microbiol 2005;43:1738-44.
3. Nix WA, Oberste MS, Pallansch MA. Sensitive, seminested PCR amplification of VP1 sequences for direct identification of all enterovirus serotypes from original clinical specimens. J Clin Microbiol 2006;44:2698-2704.

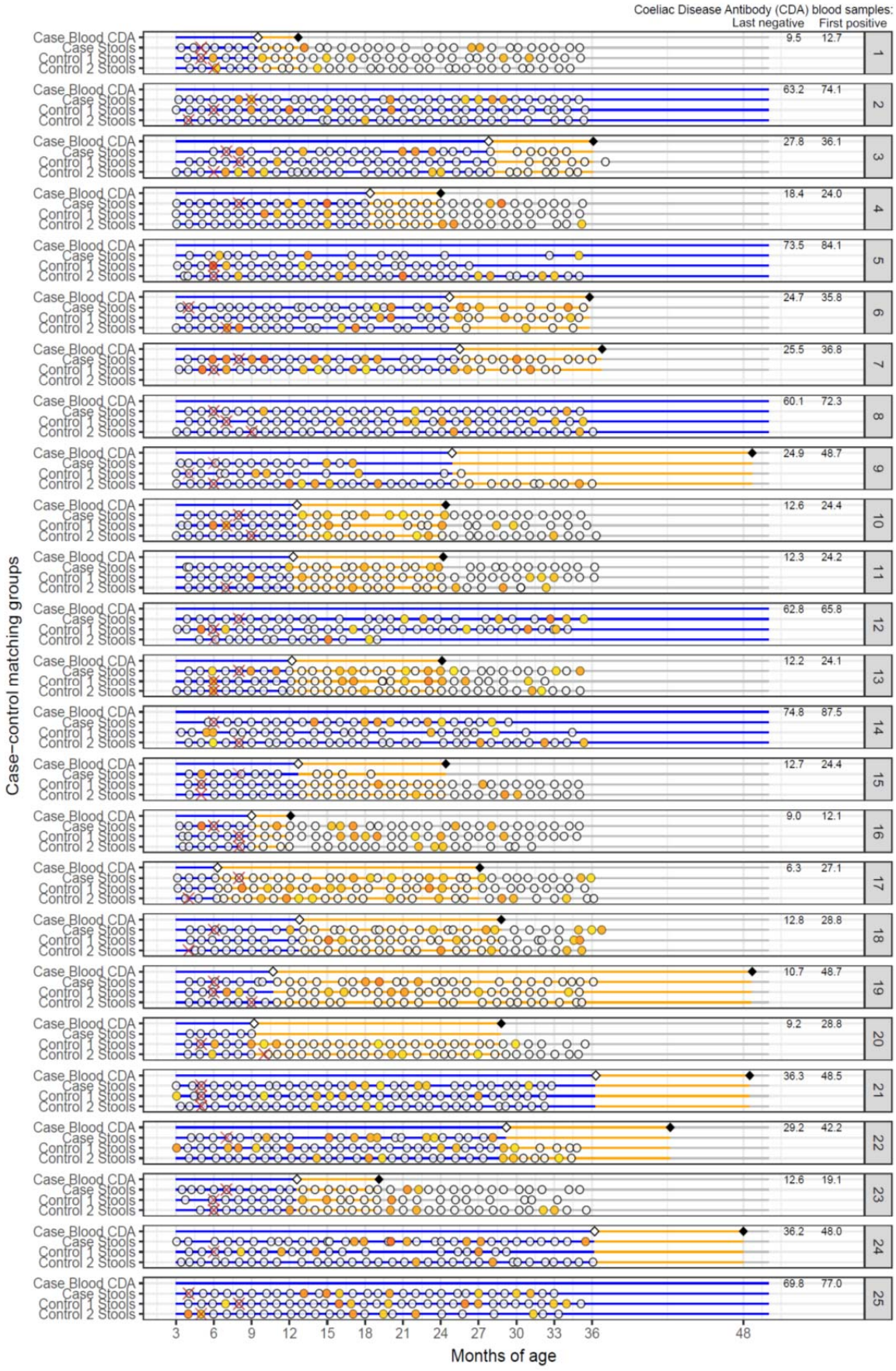
Supplementary figures

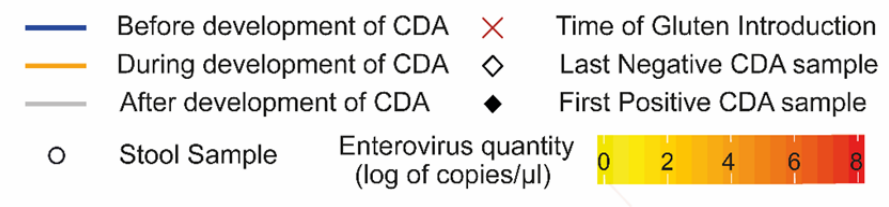
Supplementary figure A. Coeliac disease screening process



CD denotes coeliac disease; AB, antibody. Coeliac disease antibody screening was performed in samples collected from consenting participants during 2014-2016. Those diagnosed after being positive for coeliac disease antibodies on screening in addition to those with known coeliac disease were then screened retrospectively for coeliac disease antibodies in stored samples longitudinally collected from birth to identify time of seroconversion. Children negative at coeliac disease screening during 2014-16 were eligible as potential controls, and confirmed as controls if all previous samples were negative (if previous sample(s) were positive, they were not included as controls). We diagnosed coeliac disease according to the ESPGHAN (The European Society for Paediatric Gastroenterology Hepatology and Nutrition) 2012 criteria.

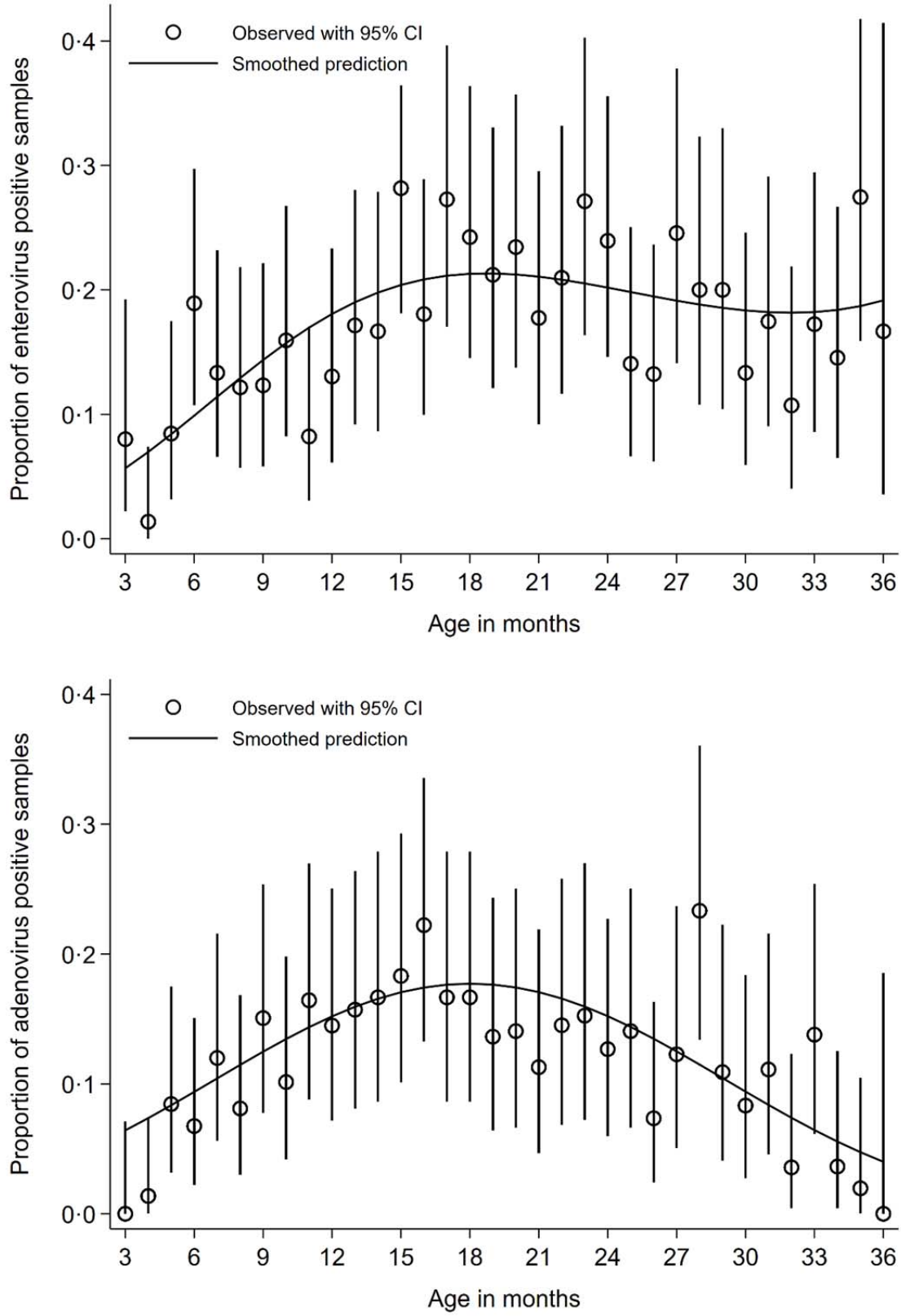
Supplementary figure B. Enterovirus distribution over the whole period





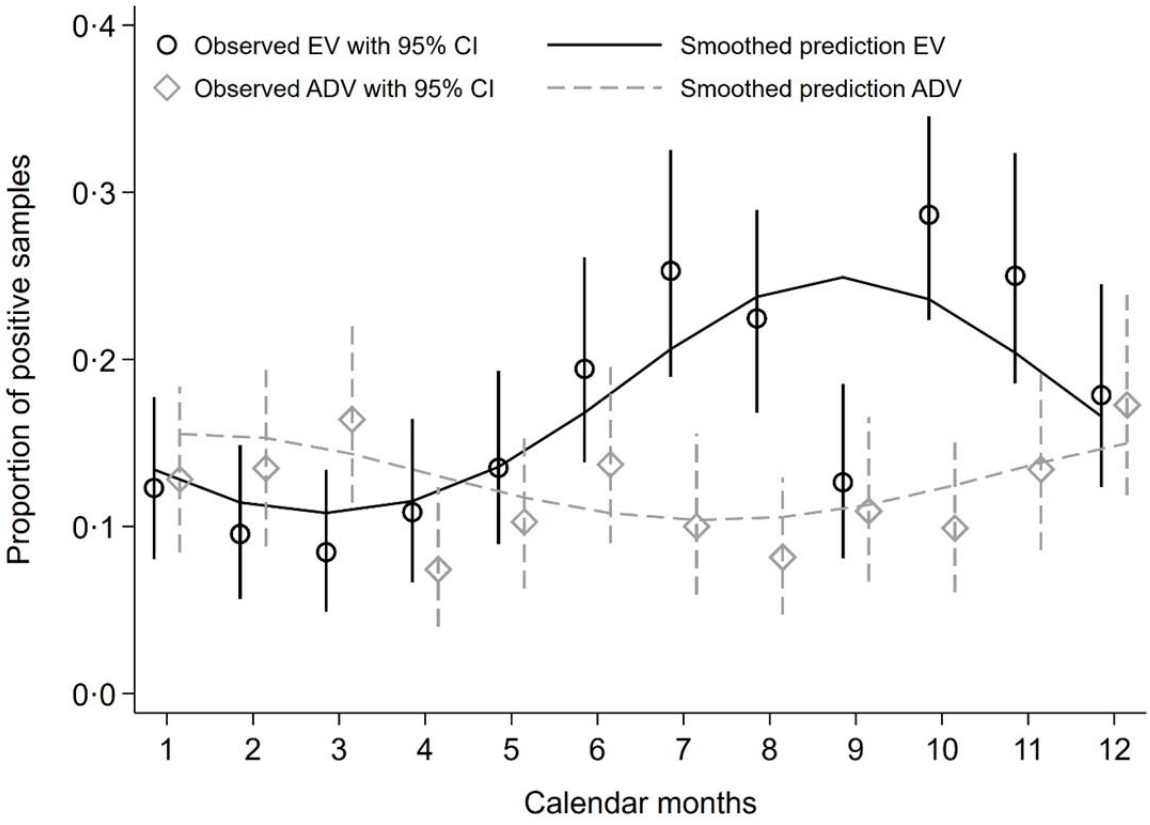
Supplementary figure 1 shows samples from the participants in the study, by month of age and matching set. One control (in matching group 7) was excluded due to missing stool samples. Top line shows the last negative and first positive coeliac disease antibody (CDA) blood sample (marked by a diamond; last negative = white, first positive = black) in cases, followed by stool samples (Enterovirus quantity denoted by colour) from cases and controls. The age in months at last negative and first positive CDA blood sample is shown in columns at the right of the plot. The coloured lines denote the time before (blue line; the main analysis), during (orange line) and after (grey line) development of CDA in cases and the corresponding age in the matched controls. A red X marks time of gluten introduction.

Supplementary figure C. Age variation of enterovirus and adenovirus



The solid line is predicted values from third degree polynomial logistic regression.

Supplementary figure D. Seasonal variation of entero- and adenovirus



EV denotes enterovirus; ADV, adenovirus. Lines are smoothed predictions from regression model with sine and cosine terms for month of sample collection (cosinor model). There was a significant seasonal variation for E (2 degree of freedom likelihood ratio test; P<0.001), but not for adenovirus (P=0.11).

Supplementary tables

Supplementary table A. Characteristics of the whole cohort, children invited to coeliac disease autoantibody screening, and those who consented to screening (participated). Values are numbers (percentages).

Characteristics	Whole cohort (n=912)	Invited (n=501)	Participated (n=220)
Year of birth* 2001 – 2005	431 (47)	241 (48)	103 (47)
Year of birth* 2006 – 2007	481 (53)	260 (52)	117 (53)
Female sex	445 (49)	250 (50)	107 (49)
Other children in household			
None	319 (35)	172 (34)	76 (35)
1	367 (40)	220 (44)	102 (46)
≥2	226 (25)	109 (22)	42 (19)
Family history of coeliac disease†	19 (2)	12 (2)	7 (3)

The "whole cohort" includes children of all consenting parents who were followed up at least once. The "invited" are those who were still actively participating (delivering samples, questionnaires) at the start of the coeliac disease sub-study and were invited to participate in the coeliac disease sub-study, and the "participated" are those who consented to participate.

*Year of birth was distributed as follows in the whole cohort (n, %): 2001 (21, 2.3%); 2002 (57, 6.3%); 2003 (69, 7.6%); 2004 (105, 11.5%); 2005 (179, 19.6%); 2006 (289, 31.7%); 2007 (192, 21.1%)

†Ascertained at time of enrolment. The nested case-control study with screened subjects that were cases and their matched controls, were asked again and have thus more affected first-degree relatives and half-siblings as shown in table 1.

Supplementary table B. Characteristics of the children with coeliac disease (n=25). Values are numbers (percentages) unless stated otherwise.

Ages	Mean (SD) age at development of coeliac disease antibodies, months	
	Last negative sample	30 (23)
	First positive sample	42 (23)
	Mean (SD) age at debut of symptoms, months[*]	73 (38)
	Mean (SD) age at diagnosis, months[†]	87 (34)
Coeliac disease antibodies	Coeliac disease antibody concentration at first positive sample[§]	
	IgA anti-tTG	
	<3 U/mL	5 (22)
	3-70 U/mL	10 (40)
	>70 U/mL	10 (40)
	IgG anti-DGP[‖]	
	<7 U/mL	4 (17)
	7-70 U/mL	15 (65)
	>70 U/mL	4 (17)
	IgA anti-tTG max values	
	<7 U/mL	1 (4)
	7-70 U/mL	3 (12)
	>70 U/mL	21 (84)
	IgG anti-DGP max values	
	<7 U/mL	3 (12)
	7-70 U/mL	12 (48)
	>70 U/mL	10 (40)
Basis for diagnosis	Biopsy (Marsh grade 3)	14 (56)
	Serology according to ESPGHAN 2012 criteria	9 (36)
	Serology and biopsy (uncertain Marsh classification)[‡]	2 (8)
Symptoms	Symptoms before diagnosis[¶]	
	Intestinal	16 (64)
	Extra-intestinal	2 (8)
	None	7 (28)

SD denotes standard deviation; tTG, tissue transglutaminase; DGP, deamidated gliadin peptide.
[*]Missing: 10.
[†]Date of the biopsy taking or date of consultation if diagnosis based on serology.
[‡]Two children with unknown/unclear Marsh-grading were diagnosed based on symptoms and repeated anti-tTG-levels >10 times cut-off level before the ESPGHAN criteria were established in 2012.
[§]IgA anti-tTG >3 or IgG DGP >7.
[‖]Missing: 2.
[¶]All participants with a coeliac disease diagnosis (both those diagnosed prior to the study screening and those diagnosed after the study screening) were asked in retrospect if they noted any symptoms prior to coeliac disease diagnosis and when these symptoms debuted.

Supplementary table C. Enterovirus types observed in this study

	Samples (n)	Percent of total samples[*]	Percent of infections[†]
Enterovirus type not determined[‡]	**62**	**-**	**16.8**
Enterovirus type determined[§]	**308**	**14.4**	**83.2**
Enterovirus A	172	8.1	46.5
CV-A2	41	1.9	11.1
CV-A4	30	1.4	8.1
CV-A5	16	0.7	4.3
CV-A6	24	1.1	6.5
CV-A8	2	0.1	0.5
CV-A10	22	1.0	5.9
CV-A12	1	0.0	0.3
CV-A14	1	0.0	0.3
CV-A16	22	1.0	5.9
EV-A71	15	0.7	4.1
Enterovirus B	139	6.5	37.6
CV-B1	12	0.6	3.2
CV-B2	19	0.9	5.1
CV-B3	11	0.5	3.0
CV-B4	9	0.4	2.4
CV-B5	12	0.6	3.2
CV-A9	17	0.8	4.6
E-3	3	0.1	0.8
E-6	5	0.2	1.4
E-7	2	0.1	0.5
E-9	10	0.5	2.7
E-11	8	0.4	2.2
E-13	1	0.0	0.3
E-18	9	0.4	2.4
E-25	19	0.9	5.1
E-30	6	0.3	1.6
Enterovirus C	4	0.2	1.1
CV-A1	3	0.1	0.8
CV-A22	1	0.0	0.3
PV-2[‖]	1	0.0	0.3
Enterovirus D	0	0	0

[*]2135 stool samples provided enterovirus data of which 370 (17%) were enterovirus positive with 309 (83.5%) having genotype information. Coinfections were common (5.8%).
[†]Number of samples with this genotype divided by number of samples positive for enterovirus (using a cut-off of at least 10 copies/µl).
[‡]Low quantity.
[§]CV denotes coxsackievirus; EV, enterovirus; E, echovirus, PV poliovirus.
[‖]Poliovirus vaccine strain Sabin.

Supplementary table D. Enterovirus infectious episodes, long infections and high-quantity infections[*]

	Cases (n=25)	Controls (n=49)	Unadjusted	Adjusted[†]	
	Positives (n)/ total (N) samples (%)		Odds ratio (95% CI)	Odds ratio (95% CI)	P value
Main analysis	*84/429 (20)*	*129/856 (15)*	*1.37 (1.01 to 1.87)*	*1.49 (1.07 to 2.06)*	*0.02*
Long infections[‡]	22/367 (6)	27/755 (4)	1.72 (0.96 to 3.06)	2.16 (1.16 to 4.04)	0.02
High-quantity infections[§]	28/429 (7)	33/856 (4)	1.73 (0.97 to 3.06)	2.11 (1.24 to 3.60)	0.01
Infectious episodes[‖]	55/400 (14)	95/822 (12)	1.22 (0.85 to 1.74)	1.27 (0.87 to 1.86)	0.21

[*]Before development of coeliac disease antibodies (prior to last negative sample).
[†]Adjusted for sex, age, age squared, season of sample collection, number of siblings, and family history of coeliac disease.
[‡]≥2 positive consecutive monthly samples.
[§]Per infection with high quantity (≥100 000 copies).
[‖]A sequence of consecutively virus-positive fecal samples (a negative stool sample was demanded before defining a new episode).

Supplementary table E. Enterovirus, gluten introduction, breastfeeding and age-groups[*]

	Cases (n=25)	Controls (n=49)	Adjusted[†]	
	Positives (n)/ total (N) samples (%)		Odds ratio (95% CI)	P value
Main analysis	*84/429 (20)*	*129/856 (15)*	*1.49 (1.07 to 2.06)*	*0.02*
Samples 3 to 6 months of age	7/87 (8)	18/174 (10)	0.44 (0.14 to 1.37)	0.16
Samples 6-12 months of age	17/155 (11)	42/302 (14)	0.66 (0.30 to 1.46)	0.30
Samples ≥12 months of age	67/231 (29)	83/463 (18)	1.97 (1.33 to 2.93)	0.001
Prior to gluten introduction	7/64 (11)	9/113 (8)	0.75 (0.21 to 2.63)	0.65
At gluten introduction[‡]	4/37 (11)	13/70 (19)	0.42 (0.11 to 1.61)	0.21
After gluten introduction[§]	66/298 (22)	99/587 (17)	1.52 (1.05 to 2.20)	0.03
While breastfed	18/163 (11)	42/369 (11)	0.78 (0.34 to 1.79)	0.56
After end of breastfeeding[‖]	55/199 (28)	83/455 (18)	1.80 (1.17 to 2.78)	0.01
After end of breastfeeding and gluten introduction	55/195 (28)	71/383 (19)	1.83 (1.19 to 2.81)	0.01

[*]Samples before development of coeliac disease antibodies (prior to last negative sample) independent of case/control/status.
[†]Adjusted for sex, age, age squared, season of sample collection, number of siblings, and family history of coeliac disease.
[‡]Samples at ± 1month of gluten introduction.
[§]Median age 6 months, range 2 to 10.
[‖]Median age 12 months, range 2 to 23.

Supplementary table F. Reported symptoms, infections with symptoms and coeliac disease*

	Cases (n=25)	Controls (n=49)	Adjusted†	
	Positives (n)/ total (N) samples (%)		Odds ratio (95% CI)	P value
Reported symptoms†				
Any symptom	157/430 (37)	302/858 (35)	1.07 (0.68 to 1.67)	0.77
Fever reported	104/440 (24)	181/865 (21)	1.21 (0.77 to 1.90)	0.41
Diarrhoea reported	30/431 (7)	61/858 (7)	1.07 (0.66 to 1.72)	0.79
Common cold reported	93/432 (22)	195/858 (23)	0.99 (0.58 to 1.68)	0.97
Enterovirus positive samples with reported symptoms				
Enterovirus + any symptom	26/419 (6)	44/849 (5)	1.49 (0.88 to 2.52)	0.14
Enterovirus + fever	24/429 (6)	28/856 (3)	2.12 (1.16 to 3.85)	0.01
Enterovirus + diarrhoea	7/420 (2)	6/849 (1)	3.62 (1.14 to 11.49)	0.03
Enterovirus + common cold	13/421 (3)	32/849 (4)	1.00 (0.51 to 1.97)	0.99
Adenovirus positive samples with reported symptoms				
Adenovirus + any symptom	19/383 (5)	48/768 (6)	0.94 (0.44 to 2.01)	0.87
Adenovirus + fever	15/390 (4)	36/775 (5)	0.88 (0.43 to 1.83)	0.74
Adenovirus + diarrhoea	3/383 (1)	9/768 (1)	0.89 (0.23 to 3.49)	0.87
Adenovirus + common cold	10/385 (3)	30/768 (4)	0.78 (0.26 to 2.30)	0.65

*Analysed as main analysis (mixed model), before development of coeliac disease antibodies (prior to last negative sample). Adjusted for sex, age, age squared, season of sample collection, number of siblings, and family history of coeliac disease.

†Self-reported symptoms of the infants as reported by their mothers, coded into yes/no per monthly sample (tolerating ± 15 days from reported date).

Supplementary table G. Association between reported symptoms and viral infections independent of case status[*]

	Reported symptoms[†]/ total samples (n)	Odds ratio (95% CI)	P value
Association enterovirus and simultaneous symptoms[‡]			
Any symptom	696/2119	0.97 (0.77 to 1.22)	0.81
Fever	418/2136	1.10 (0.86 to 1.40)	0.46
Diarrhoea	163/2120	0.90 (0.57 to 1.41)	0.63
Common cold	422/2121	0.97 (0.77 to 1.23)	0.82
Association adenovirus and simultaneous symptoms[§]			
Any symptom	644/1992	1.26 (0.93 to 1.70)	0.13
Fever[‖]	396/2006	1.62 (1.22 to 2.16)	<0.01
Diarrhoea	149/1992	1.11 (0.67 to 1.85)	0.68
Common cold	388/1994	1.11 (0.77 to 1.61)	0.58

[*]All samples analysed using logistic regression.
[†]Number of samples with positive symptoms, tolerating 15 days difference between a sample and date of reported symptoms.
[‡]368 Enterovirus samples with valid symptom data, with the exception of fever (370 samples).
[§]258 Adenovirus samples with valid symptom data.
[‖]Sequencing revealed a possible association between human adenovirus (HAdV)-C5 and fever (OR 2.95 (95% CI 1.45 to 6.03); p<0.01, and HAdV-C2 and fever (OR 1.59 (1.08 to 2.34); p=0.02.

Supplementary table H: Adenovirus types observed in this study*

	N	Percent of total samples†	Percent of infections‡
No adenovirus type determined	**6**	**-**	**2.3**
Adenovirus type determined	**252**	**12.6**	**97.7**
HAdV-C1	102	5.1	39.5
HAdV-C2	109	5.4	42.2
HAdV-B3	17	0.8	6.6
HAdV-A31	8	0.4	3.1
HAdV-F41	9	0.4	3.5
HAdV-C5	24	1.2	9.
HAdV-C57	1	0.0	0.4

*HAdV denotes *human adenovirus*. Adenovirus types were grouped into only specific types (e.g. HAdV-C2), as species *Adenovirus C* dominated in our samples.

†Percent with genotype information of all 2006 samples. Coinfections with two HAdV types were observed in 18 (0.9%) positive samples.

‡Number of samples with this genotype, divided by number of positive samples (using a cut-off of 10 copies/µl).