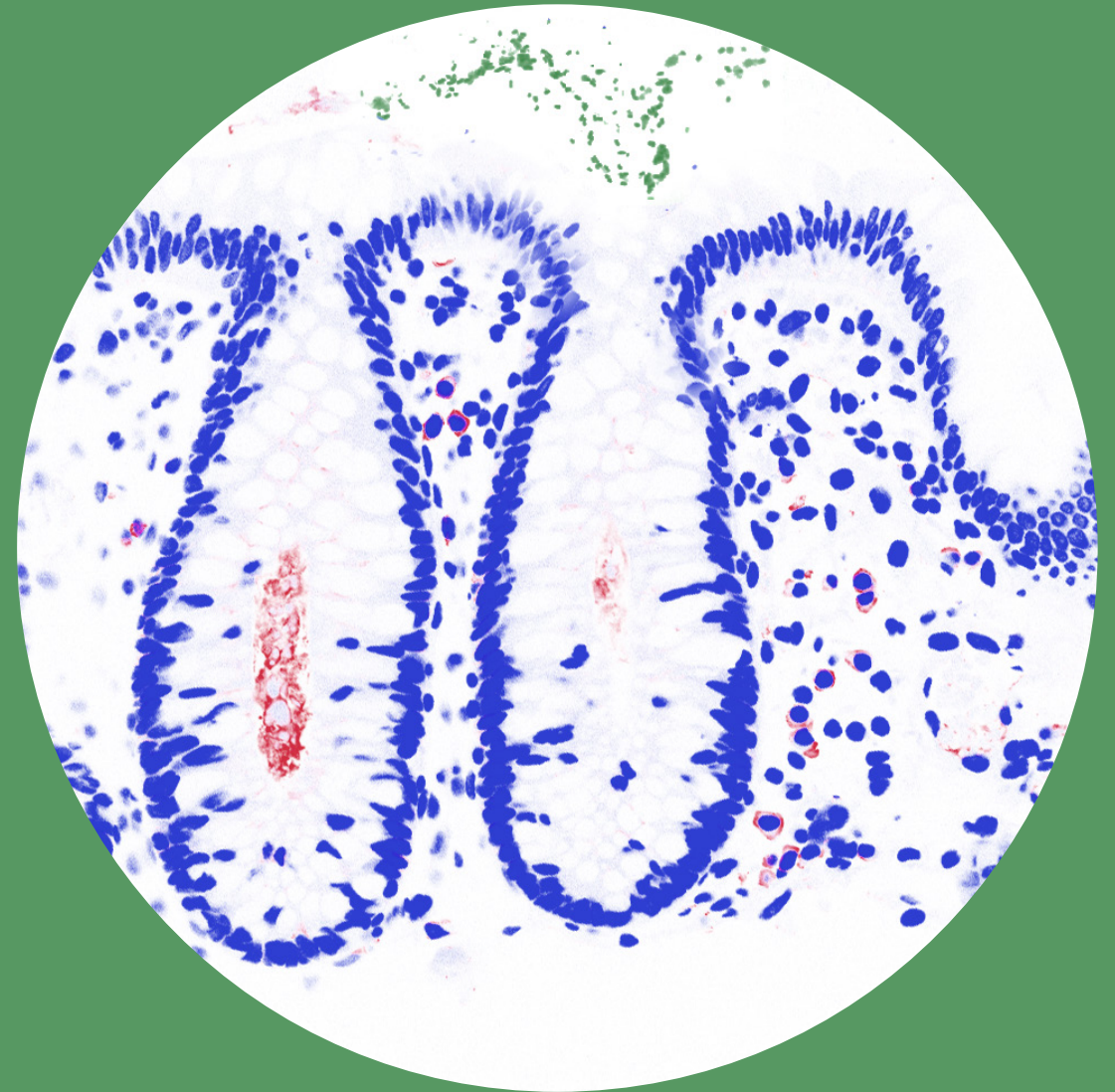
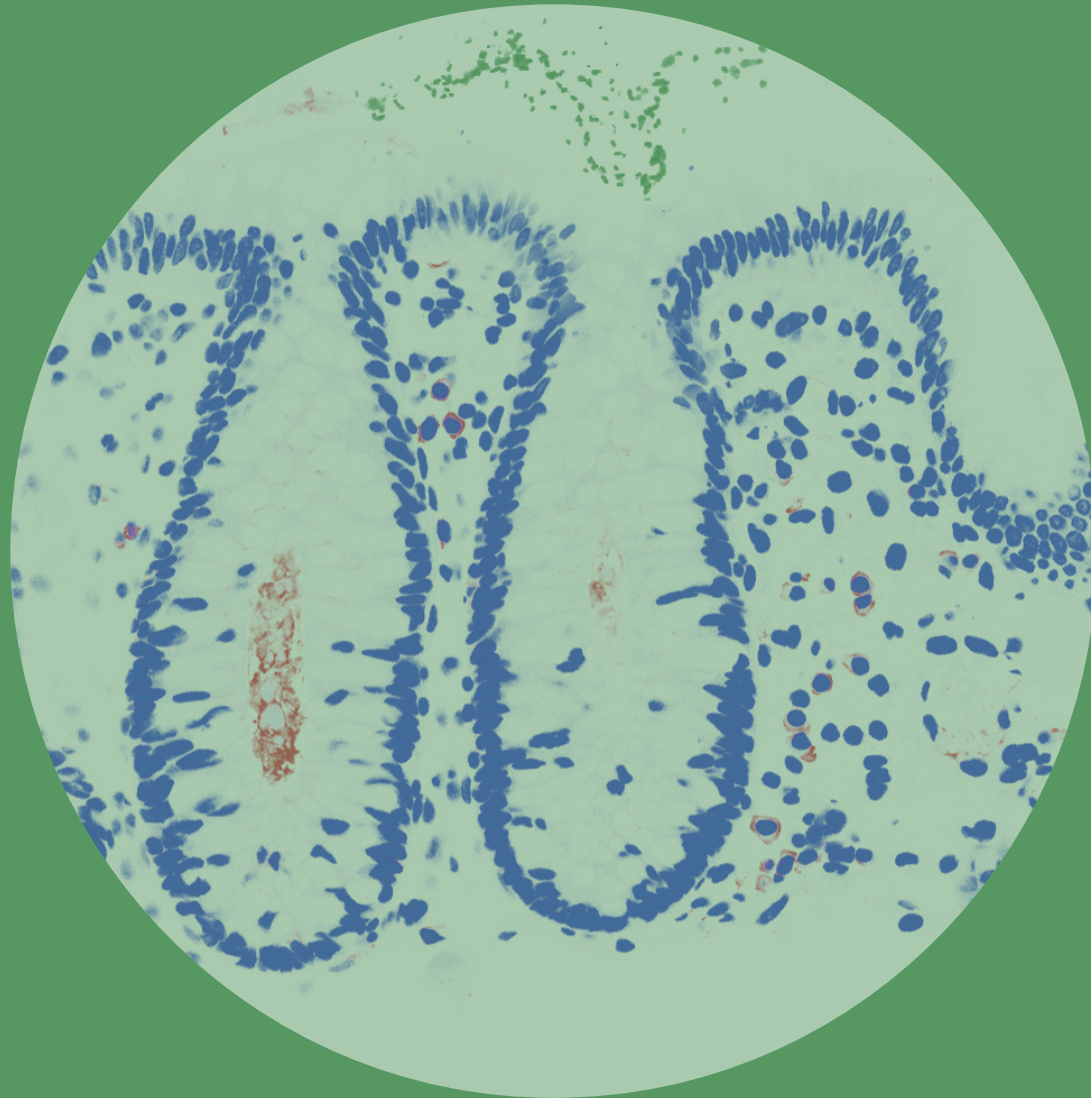


Microbiome and Inflammation in Antibody Deficiency



Microbiome and Inflammation in Antibody Deficiency



Roos-Marijn Berbers

Roos-Marijn Berbers

Microbiome and Inflammation in Antibody Deficiency

Roos-Marijn Berbers

Microbiome and Inflammation in Antibody Deficiency

© Roos-Marijn Berbers, 2021

ISBN: 978-94-6423-354-4

Cover design: R.M. Berbers and Vera van Beek. The cover is a digital adaptation of two microscopy images showing an IgA staining in red (immunofluorescence) and a bacterial 16S rRNA stain in green (fluorescence in situ hybridization), and a nuclear stain in blue (DAPI), on a healthy control colon biopsy.

Lay-out: RON Graphic Power || www.ron.nu

Printing: ProefschriftMaken || www.proefschriftmaken.nl

Financial support for the research presented in this thesis was granted by: het Wilhelmina Kinderziekenhuis fonds, stichting Zeldzame Ziekten, the UMC Utrecht Infection&Immunity boost grant, and the Utrecht Exposome Hub.

The printing of this thesis was financially supported by: CSL Behring, Lamepro, Sanquin and Takeda.

ALL RIGHTS RESERVED. Any authorized reprint or use of this material is prohibited. No part of this thesis may be reproduced, stored, or transmitted in any form or by any means, without written permission of the author or, when appropriate, of the publishers of the included scientific publications.

Microbiome and Inflammation in Antibody Deficiency

Microbioom en Inflammatie in Antistof Deficiëntie

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 21 september 2021 des middags te 12.15 uur

door

Roos-Marijn Berbers

geboren op 20 april 1991
te Amsterdam

Promotoren:

Prof. dr. R.J.L. Willems

Prof. dr. J.M. van Laar

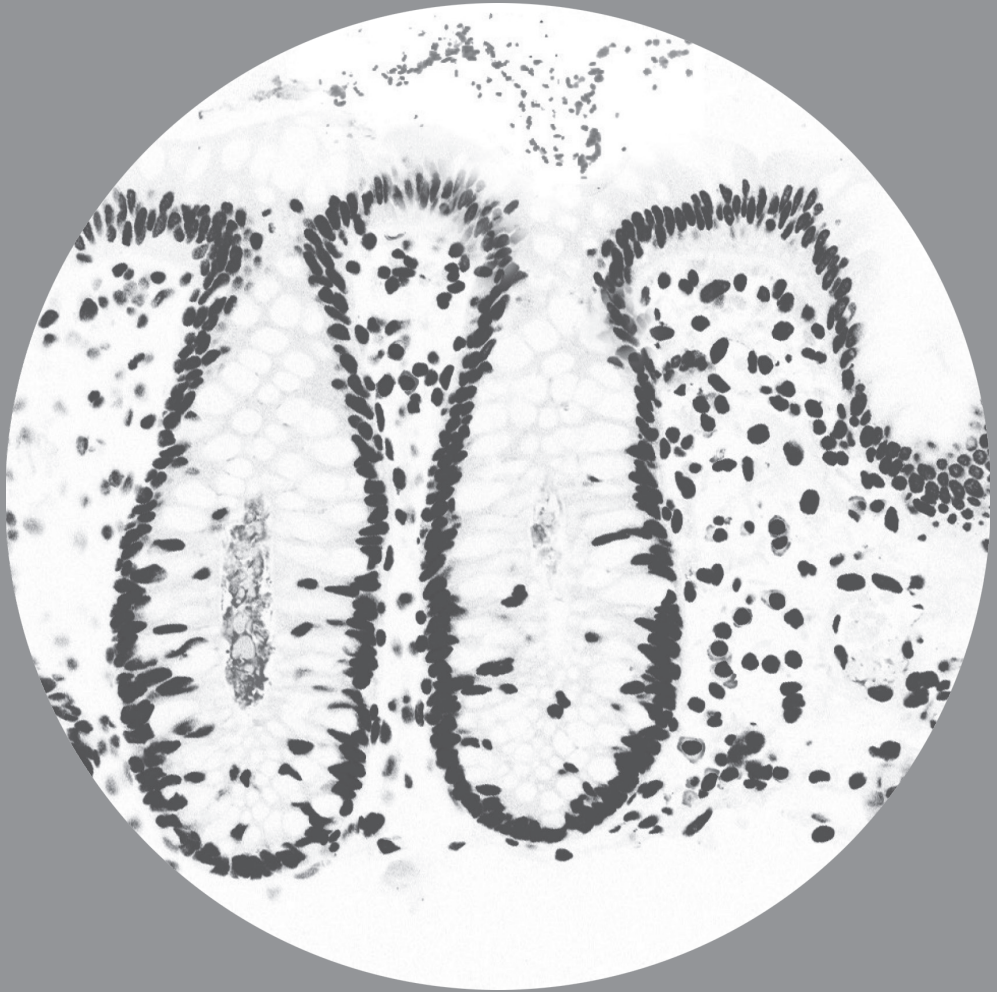
Copromotoren:

Dr. H.L. Leavis

Dr. P.M. Ellerbroek

Table of contents

Chapter 1	General introduction	7
PART I	IMMUNE DYSREGULATION IN CVID	15
Chapter 2	Targeted proteomics reveals inflammatory pathways that classify immune dysregulation in Common Variable Immunodeficiency	17
Chapter 3	Chronically Activated T-cells retain their Inflammatory Properties in Common Variable Immunodeficiency	47
PART II	THE MICROBIOME IN CVID	75
Chapter 4	Microbial Dysbiosis in Common Variable Immune Deficiencies: Evidence, Causes and Consequences	77
Chapter 5	Immunoglobulin A and microbiota in primary immunodeficiency diseases	97
Chapter 6	Low IgA associated with oropharyngeal microbiota changes and lung disease in primary antibody deficiency	113
Chapter 7	Gut Microbial Dysbiosis and <i>Enterococcus gallinarum</i> in Common Variable Immunodeficiency with Immune Dysregulation	141
PART III	ANALYSING MICROBIOMES	179
Chapter 8	Handling compositionality and zeros in sparse microbiome data – is the cure worse than the disease?	181
Chapter 9	General discussion	205
	Appendices	221
	Nederlandse samenvatting	222
	Curriculum Vitae	226
	List of Publications	227
	Dankwoord	228



Chapter 1

General introduction

The primary antibody deficiencies (PADs) are a heterogeneous group of diseases that share a problem with the production of immunoglobulins. One of the most prevalent clinically relevant PADs is a disease called Common Variable Immunodeficiency, or CVID. Despite its name, CVID is a rare disease with a prevalence in Europe of around 4 cases per 100,000 inhabitants¹⁻³, with local variations such as Finland where CVID is more common at 6.9 cases per 100,000⁴. CVID is characterized by low serum levels of immunoglobulin (Ig) G, and low IgA and/or IgM, and classically manifests either in childhood or adulthood with recurrent infections caused by encapsulated bacteria⁵⁻⁷. Treatment with IgG replacement therapy from healthy donors is effective in ameliorating the infection load in CVID⁶, but despite this treatment the mortality of CVID is still higher than the general population, with long-term survival after diagnosis ranging between 60%⁸ and 85%⁶. Most morbidity and mortality in CVID is caused by immune dysregulation complications⁸ which include autoimmune disease, granulomatous-interstitial lung disease (GLILD), enteritis, lymphoproliferative disease, and malignancy. Together, these complications affect an estimated 25% (in children) to 62% (in adults) of CVID patients⁵, and patients with CVID and immune dysregulation (CVIDid) have a poorer prognosis than patients with CVID and infections only (CVIDio)⁹. In one study, the risk of death in CVIDid was 11-fold larger than in CVIDio over a 40-year follow up period⁸.

Immune dysregulation in CVID

The immune dysregulation phenotype in CVID is a heterogeneous group of inflammatory complications that can affect CVID patients throughout their lifetime, and are sometimes already present at diagnosis⁵. The most common clinical immune dysregulation phenotypes^{5,7,8} are autoimmune (thrombo-) cytopenias, GLILD and enteritis. In addition, many patients with immune dysregulation display signs of generalized inflammation such as splenomegaly, lymphadenopathy, and polyclonal lymphocytic infiltration⁵. The etiology of immune dysregulation in CVID is largely unclear. Known risk factors for the development of immune dysregulation in CVID include low naïve CD4 T cells¹⁰, increased peripheral CD21^{low} B cells^{11,12}, and IgA deficiency^{10,12}. Biopsies of affected tissues show a pleiotropic infiltrate, with a predominance of CD4+ T-cells, variable numbers of CD8+ T-cells and B-cells, and an absence of regulatory T-cells^{13,14}.

There is no consensus about how to treat immune dysregulation complications in CVID, with expert opinion-based guidelines varying per country^{14,15}. Generally, CVID-associated immune dysregulation is treated in the same way in which the specific type of immune dysregulation is treated in non-CVID patients. However, this may not be the optimal treatment strategy, which can be illustrated by the observation that a monoclonal antibody that blocks gut-homing integrin $\alpha 4\beta 7$, which is very effective in inflammatory bowel disease, was not effective in a case series of patients with CVID-related enteritis¹⁶.

The cause of CVID and immune dysregulation in CVID is poorly understood. The heritability of CVID is estimated at 20%¹⁷, and despite great efforts to identify genes causing CVID, only a minority of CVID cases can currently be explained by monogenetic diseases^{18,19}. Genes that can result in a CVID phenotype when affected include NFKB1²⁰,

CTLA4²¹ and ICOS²², and most monogenetic cases of CVID are accompanied by immune dysregulation in addition to hypogammaglobulinemia. However, what causes the immune dysregulation in nongenetic CVID is currently unknown. It is possible that a complex genetic landscape predisposes to CVID and immune dysregulation, but the low heritability suggest that environmental factors likely also contribute.

The microbiome and immune dysregulation

The healthy human body is colonised by a vast number²³ of micro-organisms that occupy our skin, our respiratory tract, and our gut. This collection of micro-organisms is often referred to as the microbiota. The human immune system has evolved in response this microbiota, preventing infections and keeping us healthy. Over the past decade, advances in next generation sequencing have made it possible to characterize these microbial communities, and it has become increasingly clear that our microbiota have a profound impact on our immune system and our health²⁴. The bacterial gut microbiota is currently the best researched in human disease, and changes in the fecal microbiota have been linked to diseases ranging from inflammatory bowel disease²⁵ to diabetes²⁶, cancer²⁷, and neurological disorders²⁸. Specific bacterial taxa are thought to have the ability to trigger disease in genetically susceptible individuals (so-called pathobionts), while other bacteria are considered to provide health benefits (“probiotics”). Mechanisms by which the microbiota can influence host immunity include production of immunomodulatory metabolites such as short-chain fatty acids, bacterial products such as lipopolysaccharide (LPS), or direct contact with the epithelium or mucosal immune system²⁴. In addition, a diverse commensal bacterial community can provide colonization resistance against pathogens, by occupying spatial or nutritional niches²⁹.

Also in human autoimmune disease links with the microbiota have been described, and clinical trials investigating microbiome-targeting therapies such as feces transplantation, administration of probiotic bacteria, and antibiotic therapy targeting specific pathobionts for autoimmune diseases are currently underway³⁰. For instance, *Klebsiella pneumoniae* can damage the gut epithelial barrier and cause liver inflammation in patients with primary sclerosing cholangitis³¹, and translocation of *Enterococcus gallinarum* to the liver and lymph nodes is thought to contribute to the formation of auto-antibodies in systemic lupus erythematosus and autoimmune hepatitis³². It has been hypothesised that similar mechanisms may be at play in immune dysregulation in primary immunodeficiency disorders such as CVID^{33,34}.

Research objectives and outline of this thesis

In this thesis, we study the immune system and the microbiome of patients with CVID and immune dysregulation. We are particularly interested in the differences between CVIDio and CVIDid, and investigate the hypothesis that the microbiome contributes to the pathophysiology of immune dysregulation.

In **part 1** of this thesis, we investigate the immune system in CVIDid. In **chapter 2**, we broadly characterize the type of inflammation present in CVID patients with immune

dysregulation by using a targeted proteomics approach to measure cytokines and chemokines in serum. In **chapter 3**, we extend this analysis by investigating the peripheral blood T-cell profile using flow cytometry. We focus on whether immune exhaustion occurs in CVIDid, and assess markers important for regulatory T-cells.

Part 2 of this thesis concerns the microbiome in CVIDid. In **chapter 4**, we review the literature relating to the microbiome in antibody deficiencies, and introduce the hypothesis that the microbiome may contribute to the pathophysiology of immune dysregulation in CVID. We propose that IgA deficiency in CVID may cause microbial dysbiosis, and in **chapter 5** we summarize the currently known effects of IgA deficiency on specific microbial taxa.

Lung damage occurs frequently in clinical antibody deficiencies, despite immunoglobulin replacement therapy and even in the absence of clinical infections^{6,35}. A disturbed microbiota of the respiratory tract may cause subclinical inflammation that contributes to lung damage. To investigate this, we characterise the oropharyngeal microbiota in CVID patients using 16S ribosomal RNA (rRNA) sequencing **chapter 6**, and relate this to the severity of lung damage on computed tomography (CT) scans and IgA production. Next, we investigate the gut microbiota of CVID patients in **chapter 7** to identify bacterial taxa that may be linked to inflammation in CVID with immune dysregulation. We use microscopy, 16S rRNA sequencing, metagenomics shotgun sequencing and culturing techniques to identify the localization of the microbiota, characterize the composition of the microbiota and isolate candidate pro-inflammatory bacterial strains in CVIDid. These bacterial strains were then exposed to immune cells in order to investigate whether they were able to cause inflammation *in vitro*.

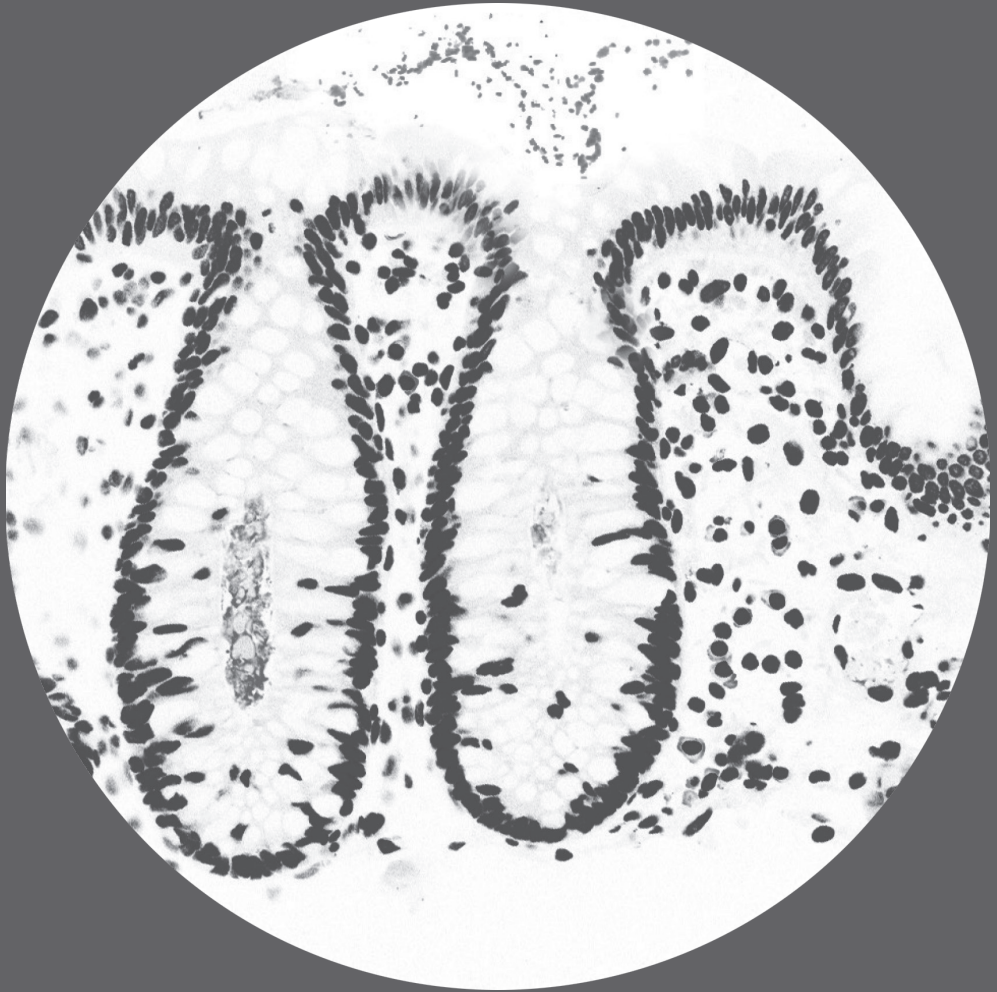
Finally, **part 3** of this thesis addresses statistical analysis of microbiome data. In **chapter 8**, we collaborated with the Julius Center for Biostatistics in order to investigate a problem we encountered during our microbiome data analysis. We assess how a commonly used transformation technique affects microbiome data structure and statistical analysis, using our oropharyngeal 16S rRNA sequencing data from chapter 6.

In **chapter 9** we provide a summary and discussion of our findings. We integrate the results of this thesis in order to place it in the context of clinical immune dysregulation in CVID, and conclude with recommendations for future studies on the role of the microbiome in immune dysregulation in CVID.

REFERENCES

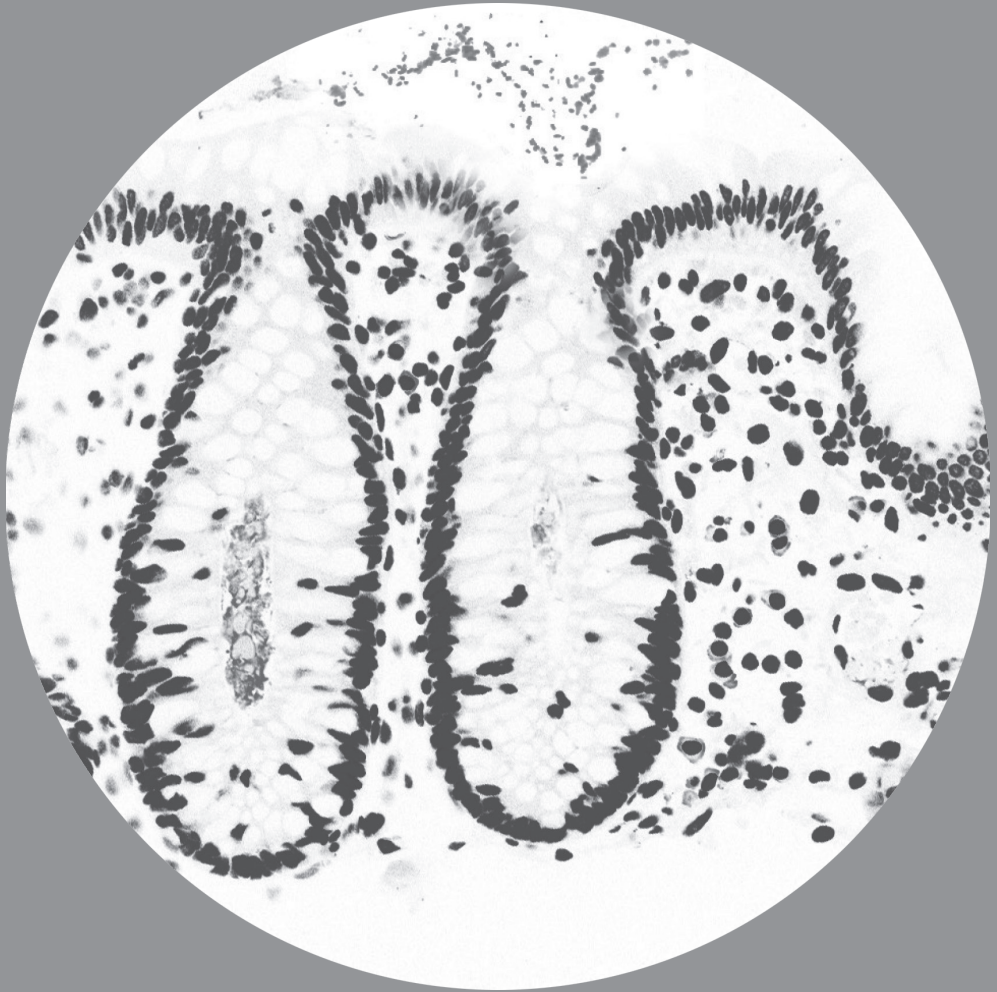
1. CEREDIH. The French national registry of primary immunodeficiency diseases. *Clin. Immunol.* **135**, 264–272 (2010).
2. Marschall, K. *et al.* The Swiss National Registry for Primary Immunodeficiencies: report on the first 6 years' activity from 2008 to 2014. *Clin. Exp. Immunol.* **182**, 45–50 (2015).
3. Westh, L. *et al.* Identification and Characterization of a Nationwide Danish Adult Common Variable Immunodeficiency Cohort. *Scand. J. Immunol.* **85**, 450–461 (2017).
4. Selenius, J. S. *et al.* Unexpectedly high prevalence of common variable immunodeficiency in Finland. *Front. Immunol.* **8**, 1–10 (2017).
5. Janssen, L. M. A., van der Flier, M. & de Vries, E. Lessons Learned From the Clinical Presentation of Common Variable Immunodeficiency Disorders: A Systematic Review and Meta-Analysis. *Front. Immunol.* **12**, (2021).
6. Quinti, I. *et al.* Long-Term Follow-Up and Outcome of a Large Cohort of Patients with Common Variable Immunodeficiency. *J. Clin. Immunol.* **27**, 308–316 (2007).
7. Maarschalk-Ellebroek, L. J., Hoepelman, A. I. M., Van Montfrans, J. M. & Ellebroek, P. M. The spectrum of disease manifestations in patients with common variable immunodeficiency disorders and partial antibody deficiency in a university hospital. *J. Clin. Immunol.* **32**, 907–921 (2012).
8. Resnick, E. S., Moshier, E. L., Godbold, J. H. & Cunningham-Rundles, C. Morbidity and mortality in common variable immune deficiency over 4 decades. *Blood* **119**, 1650–1658 (2012).
9. Odnoletkova, I., Kindle, G., Quinti, I., Grimbacher, B. & Knerr, V. The burden of common variable immunodeficiency disorders : a retrospective analysis of the European Society for Immunodeficiency (ESID) registry data. *Orphanet J. Rare Dis.* **0**, 1–17 (2018).
10. Giovannetti, A. *et al.* Unravelling the complexity of T cell abnormalities in common variable immunodeficiency. *J. Immunol.* **178**, 3932–43 (2007).
11. Warnatz, K. *et al.* Severe deficiency of switched memory B cells (CD27+IgM-IgD-) in subgroups of patients with common variable immunodeficiency: A new approach to classify a heterogeneous disease. *Blood* **99**, 1544–1551 (2002).
12. Hartono, S. *et al.* Predictors of granulomatous lymphocytic interstitial lung disease in common variable immunodeficiency. *Ann. Allergy, Asthma Immunol.* **118**, 614–620 (2017).
13. Rao, N., MacKinnon, A. C. & Routes, J. M. Granulomatous and lymphocytic interstitial lung disease: A spectrum of pulmonary histopathologic lesions in common variable immunodeficiency - Histologic and immunohistochemical analyses of 16 cases. *Hum. Pathol.* **46**, 1306–1314 (2015).
14. Chase, N. M. *et al.* Use of combination chemotherapy for treatment of granulomatous and lymphocytic interstitial lung disease (GLILD) in patients with common variable immunodeficiency (CVID). *J. Clin. Immunol.* **33**, 30–39 (2013).
15. Hurst, J. R. *et al.* British Lung Foundation/United Kingdom Primary Immunodeficiency Network Consensus Statement on the Definition, Diagnosis, and Management of Granulomatous-Lymphocytic Interstitial Lung Disease in Common Variable Immunodeficiency Disorders. *J. Allergy Clin. Immunol. Pract.* **5**, 938–945 (2017).
16. Sifers, T. *et al.* Vedolizumab therapy in common variable immune deficiency associated enteropathy: A case series. *Clin. Immunol.* **212**, (2020).
17. Li, Y. R. *et al.* Genetic sharing and heritability of paediatric age of onset autoimmune diseases. *Nat. Commun.* **6**, 1–10 (2015).
18. Maffucci, P. *et al.* Genetic diagnosis using whole exome sequencing in common variable immunodeficiency. *Front. Immunol.* **7**, (2016).
19. van Schouwenburg, P. A. *et al.* Application of whole genome and RNA sequencing to investigate the genomic landscape of common variable immunodeficiency disorders. *Clin. Immunol.* **160**, 301–314 (2015).
20. Tuijnburg, P. *et al.* Loss-of-function nuclear factor κ B subunit 1 (NFKB1) variants are the most common monogenic cause of common variable immunodeficiency in Europeans. *J. Allergy Clin. Immunol.* **142**, 1285–1296 (2018).
21. Kuehn, H. S. *et al.* Immune dysregulation in human subjects with heterozygous germline mutations in CTLA4. *Science* (80-.). **345**, 1623–1627 (2014).

22. Grimbacher, B. *et al.* Homozygous loss of ICOS is associated with adult-onset common variable immunodeficiency. *Nat. Immunol.* **4**, 261–268 (2003).
23. Sender, R., Fuchs, S. & Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol.* **14**, 1–14 (2016).
24. Belkaid, Y. & Hand, T. W. Role of the Microbiota in Immunity and Inflammation. *Cell* **157**, 121–141 (2014).
25. Sartor, R. B. & Wu, G. D. Roles for Intestinal Bacteria, Viruses, and Fungi in Pathogenesis of Inflammatory Bowel Diseases and Therapeutic Approaches. *Gastroenterology* **152**, 327–339.e4 (2017).
26. Knip, M. & Siljander, H. The role of the intestinal microbiota in type 1 diabetes mellitus. *Nat. Rev. Endocrinol.* **12**, 154–167 (2016).
27. Helmink, B. A., Khan, M. A. W., Hermann, A., Gopalakrishnan, V. & Wargo, J. A. The microbiome, cancer, and cancer therapy. *Nat. Med.* **25**, 377–388 (2019).
28. Cryan, J. F., O’Riordan, K. J., Sandhu, K., Peterson, V. & Dinan, T. G. The gut microbiome in neurological disorders. *Lancet Neurol.* **19**, 179–194 (2020).
29. Buffie, C. G. & Pamer, E. G. Microbiota-mediated colonization resistance against intestinal pathogens. *Nat. Rev. Immunol.* **13**, 790–801 (2013).
30. Marietta, E., Horwath, I., Balakrishnan, B. & Taneja, V. Role of the intestinal microbiome in autoimmune diseases and its use in treatments. *Cell. Immunol.* **339**, 50–58 (2019).
31. Nakamoto, N. *et al.* Gut pathobionts underlie intestinal barrier dysfunction and liver T helper 17 cell immune response in primary sclerosing cholangitis. *Nat. Microbiol.* **4**, 492–503 (2019).
32. Manfredo Vieira, S. *et al.* Translocation of a gut pathobiont drives autoimmunity in mice and humans. *Science* (80-.). **359**, 1156–1161 (2018).
33. Pellicciotta, M. *et al.* The microbiome and immunodeficiencies : Lessons from rare diseases. *J. Autoimmun.* 1–17 (2019) doi:10.1016/j.jaut.2019.01.008.
34. Jorgensen, S. F. *et al.* Altered gut microbiota profile in common variable immunodeficiency associates with levels of lipopolysaccharide and markers of systemic immune activation. *Mucosal Immunol.* **9**, 1455–1465 (2016).
35. Janssen, W. J. . *et al.* IgG trough levels and progression of pulmonary disease in pediatric and adult common variable immunodeficiency disorder patients. *J. Allergy Clin. Immunol.* **140**, 303–306 (2017).



Part I

Immune dysregulation in COVID



Chapter 2

Targeted proteomics reveals inflammatory pathways that classify immune dysregulation in Common Variable Immunodeficiency

Roos-Marijn Berbers
Julia Drylewicz
Pauline M. Ellerbroek
Joris M. van Montfrans
Virgil A.S.H. Dalm
P. Martin van Hagen
Baerbel Keller
Klaus Warnatz
Annick van de Ven
Jaap M. van Laar
Stefan Nierkens
Helen L. Leavis

Journal of Clinical Immunology
(2021) 41:362–373

ABSTRACT

Patients with common variable immunodeficiency (CVID) can develop immune dysregulation complications such as autoimmunity, lymphoproliferation, enteritis and malignancy, which cause significant morbidity and mortality. We aimed to (i) assess the potential of serum proteomics in stratifying patients with immune dysregulation using two independent cohorts and (ii) identify cytokine and chemokine signaling pathways that underlie immune dysregulation in CVID.

A panel of 180 markers was measured in two multicenter CVID cohorts using Olink Protein Extension Assay technology. A total of 37 CVID patients with immune dysregulation (CVIDid), 40 CVID patients with infections only (CVIDio), and 42 healthy controls (HC) were included. A classification algorithm was trained to distinguish CVIDid from CVIDio in the training cohort (CVIDid $n=14$, CVIDio $n=16$), and validated on a second testing cohort (CVIDid $n=23$, CVIDio $n=24$). Differential expression in both cohorts was used to determine relevant signaling pathways.

An elastic net classifier using MILR1, LILRB4, IL10, IL12RB1 and CD83 could discriminate between CVIDid and CVIDio patients with a sensitivity of 0.83, specificity of 0.75, and area under the curve of 0.73 in an independent testing cohort. Activated pathways (fold change >1.5 , FDR-adjusted $p < 0.05$) in CVIDid included Th1 and Th17-associated signaling, as well as IL10 and other immune regulatory markers (LAG3, TNFRSF9, CD83).

Targeted serum proteomics provided an accurate and reproducible tool to discriminate between patients with CVIDid and CVIDio. Cytokine profiles provided insight into activation of Th1 and Th17 pathways and indicate a possible role for chronic inflammation and exhaustion in immune dysregulation. These findings serve as a first step towards the development of biomarkers for immune dysregulation in CVID.

Key words

Immune dysregulation, Common Variable Immunodeficiency (CVID), Cytokines, Biomarkers, Prediction, Primary Immunodeficiency

INTRODUCTION

Common variable immunodeficiency disorder (CVID) is a primary immunodeficiency hallmarked by low serum immunoglobulins and impaired production of specific antibodies in response to vaccinations, resulting in increased risk for recurrent bacterial infections^{1,2}. The cause of CVID is assumed to be multifactorial with estimated heritability around 20%³, and while new monogenetic causes are discovered each year, in the majority of cases a genetic cause remains undefined⁴. Under adequate immunoglobulin replacement therapy (IgRT) severe infections can usually be prevented, but over a third of patients develop additional complications related to immune dysregulation⁴⁻⁶. These complications include autoimmune disease, granuloma formation, lymphoproliferative disease including increased risk of lymphoma, and enteropathy⁷, and together they cause most morbidity and mortality in CVID^{4,5}. There is no consensus about how to treat inflammatory complications in CVID, with expert opinion-based guidelines varying per country^{8,9}. The identification of predictive biomarkers for immune dysregulation, and biomarkers that can be used to monitor therapeutic response are needed to improve clinical care for these patients. In addition, better understanding of the underlying immune mechanisms that cause immune dysregulation can provide new therapeutic targets and aid in better selection of novel targeted immunotherapies such as cytokine blockade or JAK/STAT inhibitors in the clinic.

Known risk factors for the development of immune dysregulation in CVID include low naïve CD4 T cells¹⁰, increased peripheral CD21^{low} B cells^{11,12}, and IgA deficiency^{10,12}. These factors predict long-term risk to develop clinical complications, but do not inform about a current inflammatory state, and are therefore less suitable to monitor short-term disease progression or therapeutic response. For those purposes, the use of serum cytokine and chemokine biomarkers is entering clinical practice in other inflammatory diseases, for example soluble IL2 receptor in hemophagocytic lymphohistiocytosis¹³, CXCL10 in juvenile dermatomyositis¹⁴ and IL18 in adult-onset Still's disease and systemic juvenile idiopathic arthritis¹⁵.

Previous studies in CVID have frequently reported conflicting results about serum cytokines (reviewed by Varzaneh et al.¹⁶). For example, IL10 is often found to be upregulated in CVID as compared to healthy controls^{17,18}, but in a different cohort a decrease of serum IL10 was described¹⁹. Overall, markers associated with an activated myeloid compartment are consistently upregulated in CVID¹⁸⁻²⁰, and the T-helper (Th) serum cytokine profile observed in immune dysregulation in CVID is mostly found to be Th-1 driven^{17,18,21}. Varying findings in these studies highlight the need to consider CVID patients with immune dysregulation (CVIDid) separately from patients with an “infection only” (CVIDio) -phenotype.

As a first step to the identification of biomarkers for immune dysregulation in CVID, we used a targeted proteomic approach in two clinically diverse multicenter CVID cohorts in order to (i) assess the potential of serum proteomics in stratifying patients with immune dysregulation from patients with infections only using two independent cohorts and (ii)

identify cytokine and chemokine signaling pathways that underlie immune dysregulation in COVID.

MATERIALS AND METHODS

Ethics statement

Ethical approval for this study for all Dutch participants was received from the Medical Ethical Committee of the Erasmus Medical Centre in Rotterdam, the Netherlands (METC: NL40331.078). Ethical approval for sampling of patients included from Freiburg, Germany was received from the University of Freiburg Ethics Committee 282/11. Written informed consent was obtained from all patients and controls according to the Declaration of Helsinki.

Study population and sample collection

Patients diagnosed with common variable immunodeficiency disease according to the European Society for Immunodeficiencies criteria¹ aged 7 or older were included during outpatient clinic visits of the University Medical Center Utrecht, the Erasmus Medical Center in Rotterdam and the University Medical Center Groningen, in the Netherlands, and the University Medical Center Freiburg, Germany. Healthy controls (HC) were recruited from household members of patients. Medication use up to 3 months prior to sampling was recorded. Clinical characteristics were collected from electronic patient files.

Targeted proteomics

Serum samples were stored at -80°C within 4 hours of sampling until use. Serum levels of 180 unique inflammation and immune response related proteins (supplementary table S1) were measured using proximity extension immunoassay (PEA; Olink Proteomics, Uppsala, Sweden)²², using the ProSeek Multiplex Inflammation and Immune Response kits.

Briefly, the proximity extension immunoassay technology is based on dual recognition of serum proteins by pairs of antibodies coupled to a cDNA-strand that ligates when brought into proximity by its target. This DNA tag is PCR amplified and detected using a Biomark HD 96 × 96 dynamic PCR array (Fluidigm, San Francisco, USA). After corrections from DNA extension- and interpolation controls, a normalized protein expression value (NPX) is generated on a log₂ scale.

Data analysis and statistics

All data analysis was performed in R (version 3.5.1)²³, and all scripts used have been made publicly available on https://gitlab.com/rberbers/cvid_cytokines_olink.

PEA data was analyzed using NPX values (on a log₂ scale) unless stated otherwise. For proteins that were detected below the lower limit of detection (LLOD), the measured value was replaced by the LLOD value/2. The healthy control samples (training cohort

n=15, testing cohort n=27) were used to correct for batch effects between the first and the second cohort.

PCA was performed using the package *prcomp*, and 3D plots displaying the first three principal components were generated using the packages *rgl* and *car*. Differences between the clinical groups were assessed using pairwise PERMANOVA with correction for false discovery rate (FDR) using the packages *vegan* and *pairwiseAdonis*. Volcano plots were generated using Mann-Whitney U tests with Benjamini-Hochman correction for FDR and \log_2 fold change calculated on the linear scale (2^{NPX}). Differential expression was defined as FDR-adjusted (adj.) $p < 0.05$ and \log_2 fold change > 0.58 (equal to linear fold change of > 1.5).

For the classification model, the following machine learning algorithms were tuned on the training cohort using the package *caret*: random forest (package *randomForest*), elastic net regression (package *glmnet*) and extreme gradient boosting (package *xgboost*), with five-fold repeated cross-validation. All measured biomarkers with the addition of age and sex were included in the training of the algorithms. Prediction performance of the final model on the testing cohort was assessed with area under the curve (AUC) of receiver-operator curves (ROC) using the R-package *pROC*. The sensitivity and specificity for each model was calculated for the threshold with maximum Youden's Index (sensitivity+specificity -1). Regression coefficients of selected variables were standardized by their standard deviation.

Pathway analysis was performed using Ingenuity Pathway Analysis (IPA) software (QIAGEN Inc., <https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis>)²⁴, using differentially expressed proteins detected in the volcano plots and the 180 measured proteins as the reference dataset.

RESULTS

Exploratory analysis of the serum protein profile of CVID with immune dysregulation, CVID with infection only, and healthy controls

In order to evaluate whether serum cytokine profiles could be used to distinguish CVIDio from CVIDid, 180 serum markers (supplementary table S1) were measured in two independent cohorts using a proximity extension assay. Patients were categorized as CVIDid when they had clinical (history of) autoimmune disease, granulomatous-lymphocytic interstitial lung disease (GLILD), granulomatous disease other (non-GLILD), lymphoproliferation, enteritis and/or malignancy. Splenomegaly was also recorded but splenomegaly alone was not sufficient to be categorized as CVIDid. Patients without any of these complications were classified as CVIDio. All CVID patients were on IgRT at the time of sampling.

First, an age- and gender balanced training cohort (table 1) was selected using 45 participants (15 healthy controls (HC), 16 CVIDio and 14 CVIDid) from two academic hospitals in the Netherlands. The most common clinical complication in the CVIDid group

Table 1: Characteristics training cohort. * types of autoimmune disease: monoarthritis, rheumatoid arthritis, coeliac disease, Sjögren's disease, autoimmune (thrombo-)cytopenias, alopecia, vitiligo, myositis. HC: healthy control, CVIDio: CVID with infections only, CVIDid: CVID with immune dysregulation, IQR: interquartile range, AI: autoimmune, GLILD: granulomatous lymphocytic interstitial lung disease, VUS: variant of unknown significance.

	HC	CVIDio	CVIDid
Total N	15	16	14
Age (years), median (IQR)	38 (35-57)	38.5 (28.25-50.5)	41.5 (34-51.75)
Male N (%)	7 (47%)	8 (50%)	7 (50%)
Inclusion site N (%)			
Utrecht, the Netherlands	9 (60%)	11 (69%)	10 (71%)
Rotterdam, the Netherlands	6 (40%)	5 (31%)	4 (29%)
Clinical phenotype N (%)			
AI disease	0	0	9 (64%)*
GLILD	0	0	2 (14%)
Granulomatous disease other	0	0	2 (14%)
Enteritis	0	0	5 (35%)
Lymphoproliferation	0	0	1 (7%)
Malignancy	0	0	2 (14%)
Splenomegaly	0	0	7 (50%)
Medication use during 3 months prior to sampling N (%)			
Antibiotics	0	0	4 (29%)
Immune suppressive therapy	0	0	0
Genetics N (%)			
Genetics done	0	1 (6%)	5 (35%)
Nothing found	0	0	1 (7%)
VUS found	0	0	2 (14%)
Relevant pathogenic mutation found	0	1 (6%)	2 (14%)

was autoimmune disease, (64%), followed by and enteritis (35%). Splenomegaly was also highly prevalent at 50% of CVIDid patients in this cohort. Genetic screening had been performed for clinical care in a minority of patients (6% of CVIDio and 35% CVIDid), yielding one patient with CTLA4 haploinsufficiency, one patient with STAT1 gain-of-function, and two patients with variants of unknown significance (VUS) in the CVIDid group (supplementary table S2); one patient with a VUS in UNC13D and one patient with VUS in PLCG2 and heterozygosity for JAK3. In the CVIDio group, one TNFRSF13B (TACI) mutation was found.

Next, 74 participants (27 HC, 24 CVIDio and 23 CVIDid) from four academic hospitals in the Netherlands and Germany were included in a second independent testing cohort (table 2). In this cohort, there were more males in the CVIDid group (70% in CVIDid vs 33% in HC and 46% in CVIDio), and the CVIDid group was younger (mean age 36.7 in CVIDid, 40.1 in CVIDio, 44.3 in HC). The most common clinical complications in the CVIDid group were autoimmune disease (57%) and enteritis (43%), similar to prevalences in the training cohort. Four patients received immunosuppressive therapy around time

Table 2: Characteristics testing cohort. * types of autoimmune disease: systemic lupus erythematosus-like disease, Sjögren's disease, autoimmune (thrombo-)cytopenias, type 1 diabetes, membranous glomerulonephritis, alopecia, hepatitis, vitiligo. HC: healthy control, CVIDio: CVID with infections only, CVIDid: CVID with immune dysregulation, IQR: interquartile range, AI: autoimmune, GLILD: granulomatous lymphocytic interstitial lung disease, VUS: variant of unknown significance.

	HC	CVIDio	CVIDid
Total N	27	24	23
Age (years), median (IQR)	43 (37-49)	37 (25-57.25)	37 (23-49)
Male N (%)	9 (33%)	11 (46%)	16 (70%)
Inclusion site N (%)			
UMCU	19 (70%)	11 (46%)	13 (57%)
EMC	4 (15%)	7 (29%)	3 (13%)
UMCG	4 (15%)	0	0
Freiburg	0	6 (24%)	7 (30%)
Clinical phenotype N (%)			
AI disease	0	0	13 (57%)
GLILD	0	0	8 (35%)
Granulomatous disease other	0	0	1 (4%)
Enteritis	0	0	10 (43%)
Lymphoproliferation	0	0	6 (26%)
Malignancy	0	0	1 (4%)
Splenomegaly	0	0	10 (43%)
Medication use during 3 months prior to sampling N (%)			
Antibiotics	0	8 (33%)	5 (22%)
Immune suppressive therapy	0	0	4 (17%)
Genetics N (%)			
Genetics done	0	7 (29%)	7 (30%)
Nothing found	0	7 (29%)	2 (9%)
VUS found	0	0	1 (4%)
Relevant pathogenic mutation found	0	0	4 (17%)

of sampling; two patients used TNF- α blockade, and two prednisone (5 mg and 40 mg/day, respectively). In this cohort, 30% of CVIDid and 29% of CVIDio patients had been genetically screened (supplementary table 2), resulting in three TNFRSF13B mutations and one PIK3R1 mutation found in the CVIDid group. No relevant mutations were found in CVIDio.

A total of 180 unique inflammation- and immune response related proteins were quantified in serum of the training and testing cohort. Principal component analysis (PCA) shows distinct clustering of HC, CVIDio and CVIDid patients in the training cohort (figure 1A, CVIDid vs CVIDio adj. $p = 0.003$, CVIDid vs HC adj. $p = 0.002$, CVIDio vs HC adj. $p = 0.002$). One outlier in the CVIDid group was the patient with a heterozygote mutation in JAK3 and a VUS in PLCG2. In the testing cohort (figure 1B), PCA showed significant clustering of CVIDid and CVIDio from HC (CVIDid vs HC adj. $p = 0.003$,

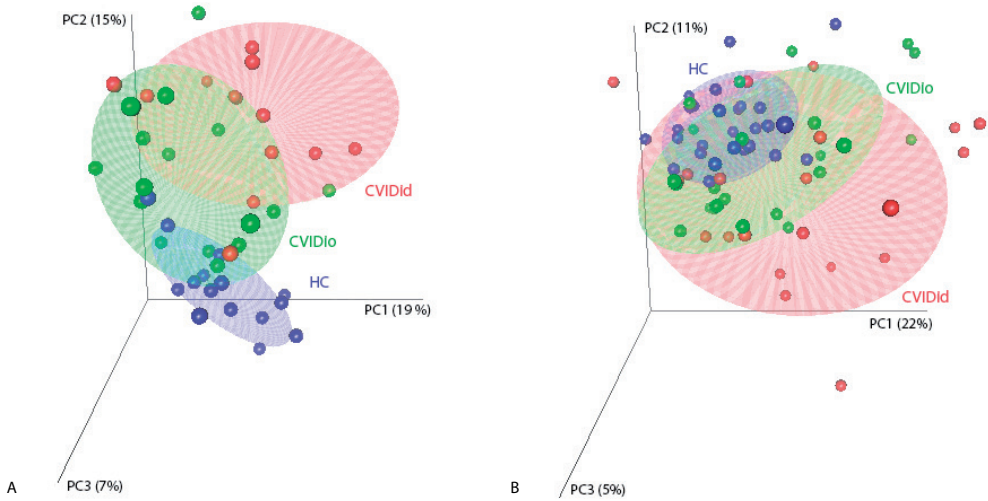


Figure 1: Principal component analysis of first 3 principal components (PC). Ellipses indicate 95% confidence intervals.

A) Training cohort: CVIDid: CVID with immune dysregulation (n=14), CVIDio: CVID with infection only (n=16), HC: healthy controls (n=15). FDR-corrected pairwise PERMANOVA using Euclidean distance: CVIDid vs CVIDio adj. $p=0.003$; CVIDid vs HC adj. $p=0.002$; CVIDio vs HC adj. $p=0.002$.

B) Testing cohort: CVIDid (n=23), CVIDio (n=24), HC (n=27). FDR-corrected pairwise PERMANOVA using Euclidean distance: CVIDid vs CVIDio adj. $p=0.144$; CVIDid vs HC adj. $p=0.003$; CVIDio vs HC adj. $p=0.005$.

CVIDio vs HC adj. $p=0.005$), but there was more overlap between CVIDid and CVIDio in this cohort (CVIDid vs CVIDio adj. $p=0.144$), with larger spread of the CVIDid group.

Machine learning approaches reveal a serum protein signature consisting of MILR1, LILRB4, IL10, IL12RB1 and CD83 to classify immune dysregulation in CVID

In order to assess whether the serum protein profiles could be used to classify immune dysregulation in CVID, three machine learning algorithms (random forest, elastic net and extreme gradient boosting) were trained on the training cohort, and their performance was assessed by using the resulting models to predict which patients had immune dysregulation in the second independent testing cohort. Given the clinical context in which detection of patients at risk for immune dysregulation is desirable, a high sensitivity was deemed more important than a high specificity.

Elastic Net (enet) and Extreme Gradient Boosting (xgb) were the best performing algorithms (supplementary table S3). In order to reduce overfitting on the training cohort only the markers that were selected as the most important variables by both models were selected for the final algorithm: mast cell immunoglobulin-like receptor 1 (MILR1), leukocyte immunoglobulin-like receptor subfamily B member 4 (LILRB4), IL10, IL12 receptor

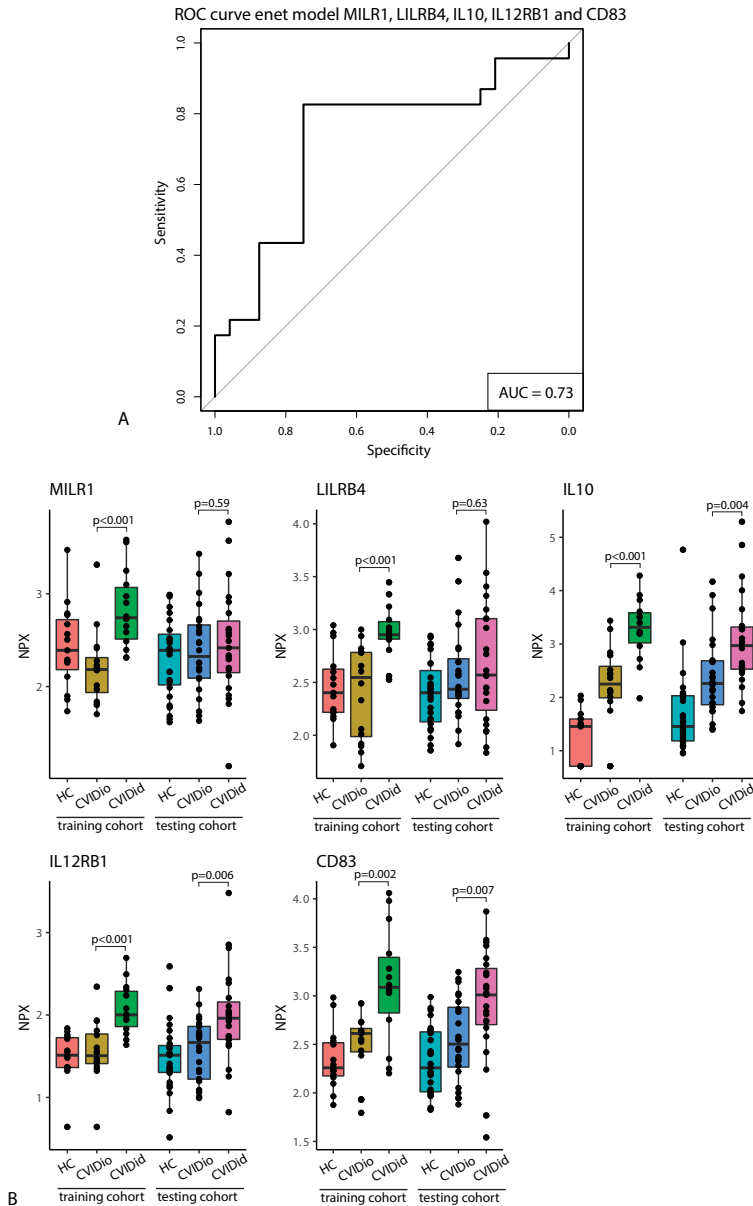
subunit beta 1 (IL12RB1), and CD83 (an immunoglobulin superfamily receptor expressed by mature antigen presenting cells²⁵). The elastic net model using these five proteins yielded the best combination of high area under the curve (AUC) and sensitivity: AUC of 0.73, sensitivity of 0.83 and specificity of 0.75 with threshold selected for maximum Youden's Index (figure 2A and table 3). Despite being trained only on samples collected in the Netherlands, this model performed equally well on the samples from the testing cohort collected in Freiburg (Germany), correctly identifying 12/13 samples. The two patients receiving TNF- α blockade were grouped correctly as CVIDid, but the patient who was sampled under 40mg prednisone was misclassified as CVIDio.

While IL10 (training set $p < 0.001$, testing set $p = 0.004$), IL12RB1 (training set $p < 0.001$, testing set $p = 0.006$) and CD83 (training set $p = 0.002$, testing set $p = 0.007$), were consistently significantly upregulated in CVIDid compared to CVIDio in both the training and the testing cohort (figure 2B), this was not the case for MILR1 and LILRB4. These two were significantly upregulated in the training cohort ($p < 0.001$ for both) but not in the testing cohort ($p = 0.59$ and $p = 0.63$, respectively), so the performance of the model without these markers was assessed in a post-hoc analysis (supplementary table S4). A logistic regression model using only IL10, IL12RB1 and CD83 yielded a higher AUC (0.76) than the original model, but with lower sensitivity (0.79) specificity (0.71) at maximum Youden's Index on the testing cohort.

Immune dysregulation in CVID is characterized by upregulation of T helper 1, T helper 17 and immune regulatory proteins

In order to infer which inflammatory pathways were differentially expressed in immune dysregulation in CVID, the training and testing cohorts were merged. In CVIDid ($n=37$), thirteen proteins were differentially upregulated compared to CVIDio ($n=40$) with FDR-corrected p -value < 0.05 and fold change (FC) > 1.5 (figure 3A, supplementary figure S1A). These markers included cytokine IL10 (adj. $p = 0.001$, $FC = 1.82$) and receptors LAG3 (adj. $p = 0.001$, $FC = 1.85$) and TNFRSF9 (also known as 4-1BB, adj. $p = 0.024$, $FC = 1.73$), all associated with negative regulation of the immune response. Th1 activation was observed in the upregulation of CXCL9 (adj. $p = 0.046$, $FC = 2.05$) and CXCL11 (adj. $p = 0.047$, $FC = 2.10$). In addition, cytokines and chemokines associated with Th17 activation were upregulated, including IL17A (adj. $p = 0.011$, $FC = 2.44$), IL12B (also known as IL12p40, a subunit of IL12 and IL23; adj. $p = 0.044$, $FC = 1.64$) and mucosal tissue homing chemokine CCL20 (adj. $p = 0.043$, $FC = 2.39$). IL-6 production, which would be consistent with activated Th17 cells, was upregulated in both CVIDid (adj. $p < 0.001$, $FC = 2.34$) and CVIDio (adj. $p = 0.001$, $FC = 2.14$) compared to HC, but not between CVIDid and CVIDio (adj. $p = 0.58$, $FC = 1.09$) (data not shown).

Natural killer (NK) cell activation marker KLRD1 (adj. $p = 0.013$, $FC = 1.84$) was upregulated in CVIDid and CVIDio, as well as SH2D1A (adj. $p = 0.024$, $FC = 1.52$), which is involved in NK- T- and B-cell stimulation. Also upregulated in CVIDid were TRANCE (also known as RANK-L, adj. $p = 0.023$, $FC = 1.52$) which induces monocyte chemotaxis, and CCL19 (adj. $p = 0.006$, $FC = 1.93$) which induces lymphocyte homing to secondary lymphoid organs.

**Figure 2**

A: Receiver-Operator Curve (ROC) for the classification of CVID with immune dysregulation (CVIDid) vs CVID with infections only (CVIDio) on the testing cohort (CVIDid $n=23$, CVIDio $n=24$) using the elastic net model using MILR1, LILRB4, IL10, IL12RB1 and CD83 as variables.

B: Distribution of classifying variables selected in the elastic net model. Training cohort healthy controls (HC, $n=15$), CVIDio ($n=16$), CVIDid ($n=14$), testing cohort HC ($n=27$), CVIDio ($n=24$), CVIDid ($n=23$). The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within $1.5 \times$ interquartile range. P-values: Mann-Whitney U test after false discovery rate correction.

Table 3: performance of the enet model using MILR1, LILRB4, IL10, IL12RB1 and CD83 as classifiers. Classification predicted at threshold with maximum Youden's Index. CVIDio: CVID with infections only, CVIDid: CVID with immune dysregulation. lymphocytic interstitial lung disease, VUS: variant of unknown significance.

Predicted	True		
	CVIDid	CVIDio	total predicted
CVIDid	19	6	25
CVIDio	4	18	22
total true	23	24	

One outlier in this analysis was TNF- α , which was not significantly upregulated (adj. $p=0.316$) but had a high fold change (FC=9.41) (supplementary figure S2). This was driven by the two patients who used TNF- α blockade therapy and had high serum levels of TNF- α , an effect that has previously been observed²⁶. These two patients did not have aberrant expression of other proteins and therefore were not excluded from the study (data not shown). TNF- α levels excluding these two patients were not different between CVIDid and CVIDio (adj. $p=0.46$, FC=1.18), but were upregulated in CVIDid (adj. $p=0.0005$, FC=1.45) and CVIDio (adj. $p=0.003$, FC=1.23) compared to HC.

Autoimmune disease, GLILD and splenomegaly in CVID are associated with a distinct serum protein profile

CVID with autoimmune disease ($n=22$) was characterised by upregulation of fifteen proteins as compared to CVID without autoimmune disease ($n=55$) (figure 3B, supplementary figure S1B), with much overlap with the upregulated proteins observed in CVIDid. In this subgroup analysis, cytokines and chemokines associated with Th1 signalling were upregulated: CXCL9 (adj. $p=0.043$, FC=2.41), CXCL10 (adj. $p=0.014$, FC=2.95), CXCL11 (adj. $p=0.014$, FC=2.67), IL18 (adj. $p=0.014$, FC=1.81), and CD28 (adj. $p=0.029$, FC=1.54). Markers of negative immune regulation were also increased in autoimmunity: LAG3 (adj. $p=0.014$, FC=1.82) and TNFRSF9 (adj. $p=0.006$, FC=1.98), and CD83 (adj. $p=0.003$, FC=1.51) which in soluble form is thought to be immunosuppressive²⁷.

In CVID with GLILD ($n=10$) (figure 3C, supplementary figure S1C), negative regulators CD83 (adj. $p=0.026$, FC=1.53) and IL10 (adj. $p=0.035$, FC=2.30) were again upregulated as compared to CVID without GLILD ($n=67$). In addition, LAMP3 (adj. $p=0.026$, FC=1.86) was upregulated, which is associated with DC maturation and is especially highly expressed in type-2 pneumocytes in the lung²⁸. T cell receptor co-receptors CD6 (adj. $p=0.034$, FC=1.84) and CD28 (adj. $p=0.034$, FC=1.75) were also upregulated in GLILD.

In CVID patients with splenomegaly ($n=17$) (figure 3D, supplementary figure S1D), nineteen markers were upregulated compared to CVID without splenomegaly ($n=60$). These included markers of immune suppression, such as TNFRSF9 (adj. $p=0.001$, FC=2.36), CD83 (adj. $p=0.002$, FC=1.50), LAG3 (adj. $p=0.010$, FC=1.99), IL10 (adj. $p=0.006$, FC=1.94), and FC receptor-like protein 3 (FCRL3) (adj. $p=0.001$, FC=2.14)²⁹. Increased Th1-associated proteins were CXCL9 (adj. $p=0.018$, FC=2.20), CXCL11

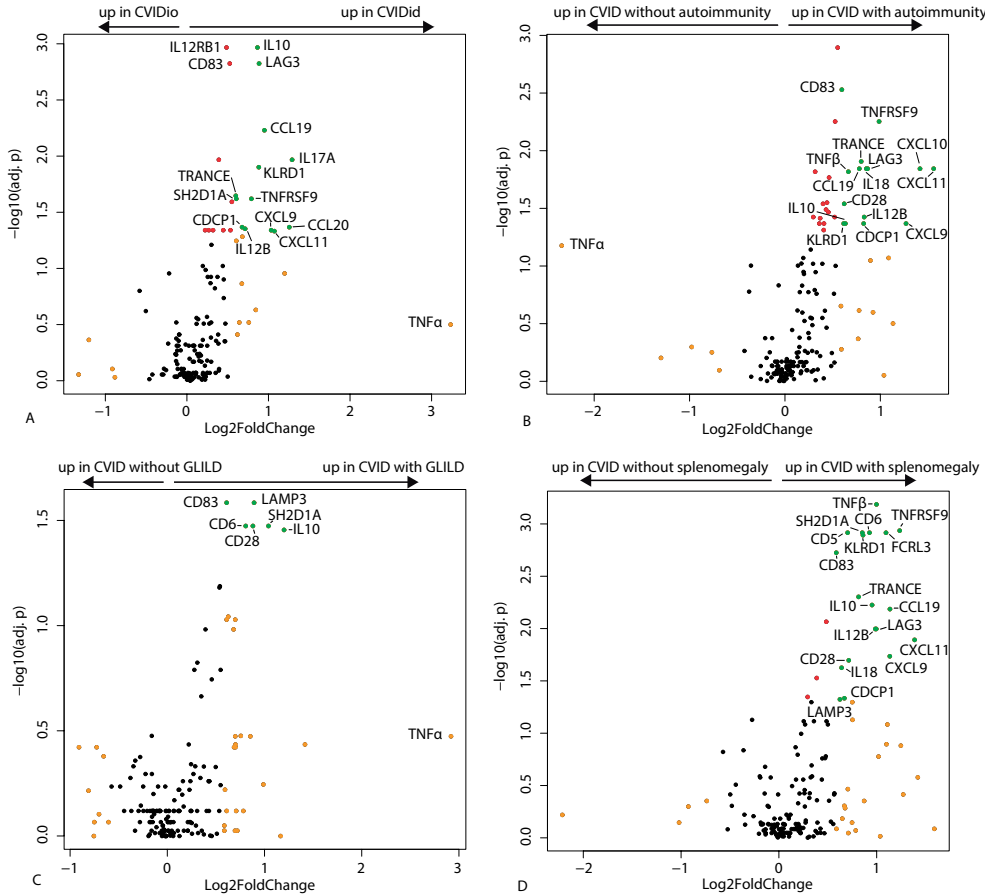


Figure 3: Volcano plots: Green dots indicate markers with log₂ fold change > 0.58 (= fold change 1.5) and false discovery rate (FDR)-adjusted p-value < 0.05. Red dots indicate markers with log₂ fold change < 0.58 and FDR-adjusted p-value < 0.05. Orange dots indicate markers with log₂ fold change > 0.58 and FDR-adjusted p-value > 0.05.

A: Differential expression analysis of proteins upregulated in CVID with immune dysregulation (CVIDid, n=37) as compared to CVID with infection only CVIDio, n=40).

B: Differential expression analysis of proteins upregulated in CVID with autoimmunity (n=22) as compared to CVID without autoimmunity (n=55).

C: Differential expression analysis of proteins upregulated in CVID with granulomatous-lymphocytic interstitial lung disease (GLILD) (n=10) as compared to CVID without GLILD (n=67).

D: Differential expression analysis of proteins upregulated in CVID with splenomegaly (n=17) as compared to CVID without splenomegaly (n=60).

(adj. p=0.013, FC=2.62), IL18 (adj. p=0.024, FC=1.56), CD28 (adj. p=0.020, FC=1.64) and IL12B (adj. p=0.010, FC=1.98). Inflammatory markers KLRD1 (adj. p = 0.001, FC=1.80), TNF-β (adj. p < 0.001, FC=2.00), SH2D1A (adj. p=0.001, FC=1.82), CD5 (adj. p=0.001,

Pathway upregulated proteins immune dysregulation in CVID

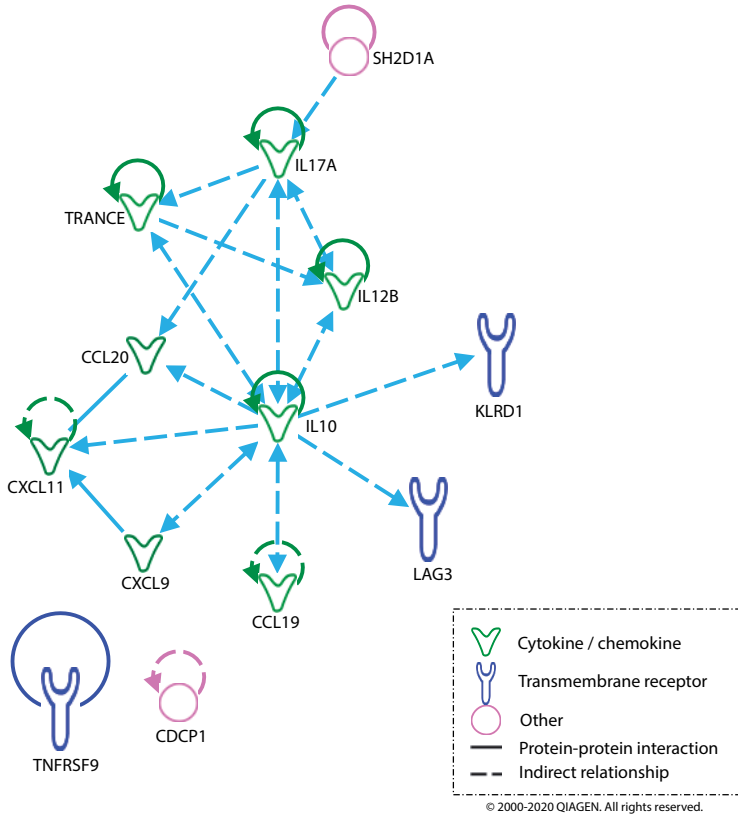


Figure 4: Literature-based pathway analysis of differentially expressed proteins upregulated in CVID with immune dysregulation (CVIDid) as compared to CVID with infections only (CVIDio), using Ingenuity Pathway Analysis software.

$FC=1.63$), CD6 (adj. $p=0.001$, $FC=1.90$) and one of its ligands CDCP1 (adj. $p=0.048$, $FC=1.54$), and TRANCE (adj. $p=0.005$, $FC=1.76$) were also upregulated.

Subgroup analyses of granulomatous disease, lymphoproliferation, enteritis and malignancy did not yield any differentially expressed proteins, possibly due to smaller sample size.

In order to integrate these findings, a literature-based pathway analysis was performed using Ingenuity Pathway Analysis. Of the upregulated proteins in CVIDid, IL10 had the most connections to the other differentially expressed markers (figure 4A), indicating that IL10 may be a keystone regulator in the inflammatory profile of CVIDid. The same was observed for pathway analysis of proteins upregulated in autoimmunity, GLILD, and splenomegaly (data not shown).

DISCUSSION

We demonstrated that an algorithm using serum biomarkers MILR1, LILRB4, IL10, IL12RB1 and CD83 identified by targeted serum proteomics could classify immune dysregulation in CVID in a discovery cohort and in our independent testing cohort, providing a first step towards the development of a screening tool for immune dysregulation in CVID. Of the selected biomarkers, IL10, IL12RB1 and CD83 were consistently upregulated in the testing and the training cohorts, in contrast to MILR1 and LILRB4 which were not reproducible in the testing cohort. This may be due to sensitivity of MILR1 and LILRB4 to batch effects or minute differences in sample handling. However, an algorithm using only IL10, IL12RB1 and CD83 performed almost as well as when MILR1 and LILRB4 were included. This is in accordance with previous findings of upregulated IL10^{17,18,21}, IL12^{18,21} and CD83²¹ in CVID compared to healthy controls.

As a next step towards the application of this screening tool, the dynamics of these serum markers in early disease need to be monitored, as this cohort included only patients with current immune dysregulation or immune dysregulation in remission. If these markers are upregulated before the full clinical phenotype has developed, the algorithm may be used for early detection of disease, and allow for earlier intervention with immunosuppressive therapy. Moreover, the screening tool may be further improved by re-training and testing the algorithm on additional cohorts. To simplify the applicability of the screening tool, further studies may choose to measure the selected biomarkers using more widely available techniques such as enzyme-linked immunosorbent assays (ELISA). Possibly, IL12B can substitute IL12RB1 in the model, as IL12B was also differentially expressed in this study and assays may be more widely available. A limitation of the PEA technology is that the normalized protein expression values were not converted to an absolute concentration but could only be compared between samples in the same run. In this study, the healthy control samples were used to correct for batch effects between the testing and training cohorts, but this is not practical for use in the clinic. A solution would be to include standard curves for the biomarkers of interest, and/or to use a different technique such as ELISA in order to quantify these markers and identify cutoff values for the normal range.

After merging the two cohorts, differentially expressed proteins specific for CVIDid as a whole, and CVID with autoimmunity, GLILD and splenomegaly were identified. In our study, IFN γ -responsive cytokines CXCL9, -10 and -11 were upregulated in CVIDid, which are instrumental in Th1 skewing³⁰. This is in line with previous data showing accumulating support for Th1/T follicular helper (Tfh) 1 skewing in CVIDid patients²¹⁻³¹. Our data also indicates Th17 activation in CVID patients with immune dysregulation, reflected by upregulation of IL17, IL12B (the subunit for both IL12 and IL23) and CCL20 in CVIDid, which are associated with Th17 skewing and consistent with reports in other inflammatory disorders³². Upregulation of IL6, which may be consistent with Th17 and Tfh activation, was observed in both CVIDio and CVIDid compared to HC in the present study. Other Th17-associated cytokines such as IL21 and IL22 were not measured here.

However, these findings are in contrast to previous studies reporting a suppression of IL17/Th17 cells in CVID^{18,33,34}. A possible alternative source for IL17 production in these patients are a population of innate lymphoid cells that produce IL17 and IFN γ that were described in the blood of CVID patients with immune dysregulation³⁵.

In parallel to the Th1/Th17-associated inflammatory cytokines, we observed an increase of immune regulatory proteins. The induction of IL10 often accompanies production of pro-inflammatory cytokines in both myeloid and T helper cells, suggestive of an (in this case insufficient) compensatory feedback loop that limits immune pathology³⁶. While IL10 can be produced by regulatory T cells, there was no co-upregulation of TGF- β in this cohort, suggesting an alternative source for the IL10, such as monocytes. Induction of IL10 due to persistent antigen exposure can result in functional T cell exhaustion³⁷. Chronic antigen exposure in CVID despite IgRT may be related to bacterial translocation from the gastrointestinal tract³⁸.

LAG3, TNFRSF9 (also known as 4-1BB) and CD83 were also upregulated in immune dysregulation in our cohorts and may reflect a chronically activated immune state and immune exhaustion. LAG3 is a co-inhibitory receptor that limits the proliferative capacity of T cells and can confer suppressive properties to other T cells^{39,40}. However, the soluble form of LAG3 has been shown to be immune potentiating and is being investigated as a vaccine adjuvant⁴¹. TNFRSF9 also has complex effects on different cell types^{42,43}, and increased serum levels of soluble TNFRSF9 correlate with disease severity in rheumatoid arthritis and other autoimmune diseases, suggesting that the soluble form may act as a decoy receptor preventing TNFSF9-mediated suppression of T cells⁴⁴. Similar dynamics have been described for antigen presenting cell maturation marker CD83, which is reported to be immunosuppressive in autoimmune diseases in its soluble form²⁵.

Taken together, the upregulation of IL10, LAG3, TNFRSF9 and CD83 in immune dysregulation in CVID may indicate a chronic and refractory immune activated state. Whether functional exhaustion of T cells also occurs in these patients will need to be assessed on a cellular level. One study that investigated this in a mixed cohort of CVID and CVIDio patients reported functional exhaustion of CD4 T cells, which correlated with serum endotoxaemia⁴⁵. Authors of this study did not find upregulation of membrane-bound LAG3, but do report increased surface expression of PD-1⁴⁵. In our study, serum levels of PD-L1 (the ligand for PD-1) were significantly increased in CVID with autoimmunity but did not pass the fold-change criterion (adj. $p=0.029$, $FC=1.32$).

To conclude, this study shows a promising first step in the development of a screening tool for immune dysregulation in CVID using serum proteins IL10, IL12RB1 and CD83 as biomarker. In addition, the immune dysregulation clinical phenotype was associated with increased levels of Th1- and Th17- related serum proteins, and displays a complex immune regulatory profile that includes IL10, LAG3, TNFRSF9 and CD83 signaling. Further research focusing on the dynamics of these biomarkers longitudinally is necessary to evaluate its use as an early detection screening tool for immune dysregulation in CVID. In addition, studying the behavior of these biomarkers under immunosuppressive therapy will indicate whether these markers can be used to monitor therapeutic response.

Funding

This study was funded by the Wilhelmina Children's Hospital Fund.

Acknowledgements

Samples for this project were obtained from the CCI-Biobank, a partner biobank of the University Medical Center Freiburg and Medical Faculty "Center for Biobanking – FREEZE".

Disclosure of potential conflict of interests

PH reports research grants and personal fees from Shire/Takeda and CSL Behring. VD reports research grants and personal fees from Shire/Takeda, Grifols, Actelion, Novartis and CSL Behring. JM reports personal fees from Shire/Takeda. AvdV reports personal fees from CSL Behring. HL reports research grants and personal fees from Shire/Takeda. All other authors report no potential conflict of interest.

Author contributions

HL, RMB, PE, JvM, VD, PvH, AvdV, BK, KW, SN, JvL: study design and implementation. HL, PE, JvM, VD, PvH, AvdV, BK, KW: supported sample acquisition. SN: technical expertise targeted proteomics. JD: statistical support. RMB, HL: data analysis and interpretation. All authors substantially contributed to the acquisition, analysis or interpretation of the data and approved the final manuscript.

REFERENCES

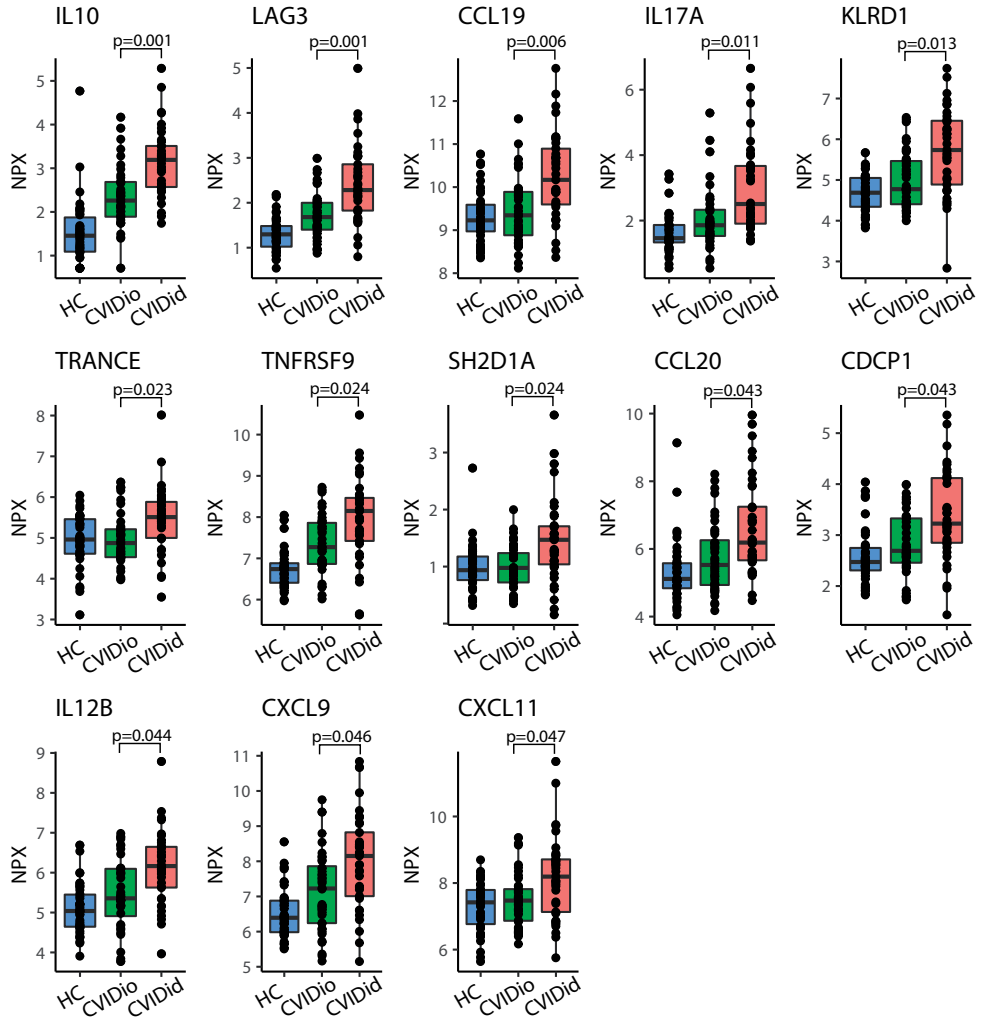
1. Immunodeficiencies ES for. Diagnostic Criteria PID . 2019. Available from: <https://esid.org/Education/Diagnostic-Criteria-PID>
2. Bonilla FA, Barlan I, Chapel H, Costa-Carvalho BT, Cunningham-Rundles C, de la Morena MT, et al. International Consensus Document (ICON): Common Variable Immunodeficiency Disorders. *J Allergy Clin Immunol Pract*. 2016;4:38–59.
3. Li YR, Zhao SD, Li J, Bradfield JP, Mohebnasab M, Steel L, et al. Genetic sharing and heritability of paediatric age of onset autoimmune diseases. *Nat Commun*. 2015;6:1–10.
4. Resnick ES, Moshier EL, Godbold JH, Cunningham-Rundles C. Morbidity and mortality in common variable immunodeficiency over 4 decades. *Blood*. 2012;119:1650–8.
5. Chapel H, Lucas M, Lee M, Bjorkander J, Webster D, Grimbacher B, et al. Common variable immunodeficiency disorders: division into distinct clinical phenotypes. *Blood*. 2008;112:277–87.
6. Gathmann B, Mahlaoui N, Gérard L, Oksenhendler E, Warnatz K, Schulze I, et al. Clinical picture and treatment of 2212 patients with common variable immunodeficiency. *J Allergy Clin Immunol*. 2014;134.
7. Maarschalk-Ellebrouck LJ, Hoepelman AIM, Van Montfrans JM, Ellebrouck PM. The spectrum of disease manifestations in patients with common variable immunodeficiency disorders and partial antibody deficiency in a university hospital. *J Clin Immunol*. 2012;32:907–21.
8. Hurst JR, Verma N, Lowe D, Baxendale HE, Jolles S, Kelleher P, et al. British Lung Foundation/United Kingdom Primary Immunodeficiency Network Consensus Statement on the Definition, Diagnosis, and Management of Granulomatous-Lymphocytic Interstitial Lung Disease in Common Variable Immunodeficiency Disorders. *J Allergy Clin Immunol Pract*. 2017;5:938–45.
9. Chase NM, Verbsky JW, Hintermeyer MK, Waukau JK, Tomita-Mitchell A, Casper JT, et al. Use of combination chemotherapy for treatment of granulomatous and lymphocytic interstitial lung disease (GLILD) in patients with common variable immunodeficiency (CVID). *J Clin Immunol*. 2013;33:30–9.
10. Giovannetti A, Pierdominici M, Mazzetta F, Marziali M, Renzi C, Mileo AM, et al. Unraveling the complexity of T cell abnormalities in common variable immunodeficiency. *J Immunol*. 2007;178:3932–43.
11. Warnatz K, Denz A, Dräger R, Braun M, Groth C, Wolff-Vorbeck G, et al. Severe deficiency of switched memory B cells (CD27+IgM-IgD-) in subgroups of patients with common variable immunodeficiency: A new approach to classify a heterogeneous disease. *Blood*. 2002;99:1544–51.
12. Hartono S, Motosue MS, Khan S, Rodriguez V, Iyer VN, Divekar R, et al. Predictors of granulomatous lymphocytic interstitial lung disease in common variable immunodeficiency. *Ann Allergy, Asthma Immunol. American College of Allergy, Asthma & Immunology*; 2017;118:614–20.
13. Hayden A, Lin M, Park S, Pudek M, Schneider M, Jordan MB, et al. Soluble interleukin-2 receptor is a sensitive diagnostic test in adult HLH. *Blood Adv*. 2017;1:2529–34.
14. Wienke J, Bellutti Enders F, Lim J, Mertens JS, van den Hoogen LL, Wijngaarde CA, et al. Galectin-9 and CXCL10 as Biomarkers for Disease Activity in Juvenile Dermatomyositis: A Longitudinal Cohort Study and Multicohort Validation. *Arthritis Rheumatol*. 2019;71:1377–90.
15. Kudela H, Drynda S, Lux A, Horneff G, Kekow J. Comparative study of Interleukin-18 (IL-18) serum levels in adult onset Still's disease (AOSD) and systemic onset juvenile idiopathic arthritis (SJIA) and its use as a biomarker for diagnosis and evaluation of disease activity. *BMC Rheumatol. BMC Rheumatology*; 2019;3:1–11.

16. Varzaneh FN, Keller B, Unger S, Aghamohammadi A, Warnatz K, Rezaei N. Cytokines in common variable immunodeficiency as signs of immune dysregulation and potential therapeutic targets - A review of the current knowledge. *J Clin Immunol.* 2014;34:524–43.
17. Rezaei N, Aghamohammadi A, Kardar GA, Nourizadeh M, Pourpak Z. T-helper 1 and 2 cytokine assay in patients with common variable immunodeficiency. *J Investig Allergol Clin Immunol.* 2008;18:449–53.
18. Hel Z, Huijbregts RPH, Xu J, Nechvatalova J, Vlkova M, Litzman J. Altered Serum Cytokine Signature in Common Variable Immunodeficiency. *J Clin Immunol.* 2014;34:971–8.
19. Polito R, Nigro E, Pecoraro A, Monaco ML, Perna F, Sanduzzi A, et al. Adiponectin Receptors and Pro-inflammatory Cytokines Are Modulated in Common Variable Immunodeficiency Patients: Correlation With Ig Replacement Therapy. *Front Immunol.* 2019;10:1–8.
20. Ueland T, Frøland SS, Bollerslev J, Aukrust P. Increased levels of biochemical markers of bone turnover in relation to persistent immune activation in common variable immunodeficiency. *Eur J Clin Invest.* 2001;31:72–8.
21. Hultberg J, Ernerudh J, Larsson M, Nilsson-Augustinsson Å, Nyström S. Plasma protein profiling reflects TH1-driven immune dysregulation in common variable immunodeficiency. *J Allergy Clin Immunol.* 2020;1–12.
22. Assarsson E, Lundberg M, Holmquist G, Björkstén J, Thorsen SB, Ekman D, et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS One.* 2014;9.
23. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2019. p. URL <https://www.R-project.org/>.
24. Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics.* 2014;30:523–30.
25. Li Z, Ju X, Silveira PA, Abadir E, Hsu WH, Hart DNJ, et al. CD83: Activation marker for antigen presenting cells and its therapeutic potential. *Front Immunol.* 2019;10:1–9.
26. Berkhout LC, L'Ami MJ, Ruwaard J, Hart MH, Ooijevaar-De Heer P, Bloem K, et al. Dynamics of circulating TNF during adalimumab treatment using a drug-tolerant TNF assay. *Sci Transl Med.* 2019;11.
27. Zinser E, Lechmann M, Golka A, Lutz MB, Steinkasserer A. Prevention and treatment of experimental autoimmune encephalomyelitis by soluble CD83. *J Exp Med.* 2004;200:345–51.
28. Akasaki K, Nakamura N, Tsukui N, Yokota S, Murata SI, Katoh R, et al. Human dendritic cell lysosome-associated membrane protein expressed in lung type II pneumocytes. *Arch Biochem Biophys.* 2004;425:147–57.
29. Kochi Y, Myouzen K, Yamada R, Suzuki A, Kurosaki T, Nakamura Y, et al. FCRL3, an Auto-immune Susceptibility Gene, Has Inhibitory Potential on B-Cell Receptor-Mediated Signaling. *J Immunol.* 2009;183:5502–10.
30. Tokunaga R, Zhang W, Naseem M, Puccini A, Berger MD, Soni S, et al. CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation – A target for novel cancer therapy. *Cancer Treat Rev.*; 2018;63:40–7.
31. Unger S, Seidl M, van Schouwenburg P, Rakhmanov M, Bulashevska A, Frede N, et al. The TH1 phenotype of follicular helper T cells indicates an IFN- γ -associated immune dysregulation in patients with CD21low common variable immunodeficiency. *J Allergy Clin Immunol.* 2018;141:730–40.
32. Yasuda K, Takeuchi Y, Hirota K. The pathogenicity of Th17 cells in autoimmune diseases. *Semin Immunopathol. Seminars in Immunopathology*; 2019;41:283–97.
33. Barbosa RR, Silva SP, Silva SL, Melo AC, Pedro E, Barbosa MP, et al. Primary B-cell deficiencies reveal a link between human IL-17-producing CD4 T-cell homeostasis and B-cell differentiation. *PLoS One.* 2011;6.
34. Edwards ESJ, Bosco JJ, Aui PM, Stirling RG, Cameron PU, Chatelier J, et al. Predominantly Antibody-Deficient Patients With Non-infectious Complications Have Reduced Naive B, Treg, Th17, and Tfh17 Cells. *Front Immunol.* 2019;10.

35. Cols M, Rahman A, Maglione PJ, Garcia-Carmona Y, Simchoni N, Ko HBM, et al. Expansion of inflammatory innate lymphoid cells in patients with common variable immune deficiency. *J Allergy Clin Immunol*. 2016;137:1206-1215.e6.
36. Saraiva M, Saraiva M, Vieira P, Vieira P, Vieira P, O'Garra A, et al. Biology and therapeutic potential of interleukin-10. *J Exp Med*. 2020;217:1-19.
37. Brooks DG, Trifilo MJ, Edelmann KH, Teyton L, McGavern DB, Oldstone MBA. Interleukin-10 determines viral clearance or persistence in vivo. *Nat Med*. 2006;12:1301-9.
38. Jorgensen SF, Troseld M, Kummen M, Anmarkrud JA, Michelsen AE, Osnes LT, et al. Altered gut microbiota profile in common variable immunodeficiency associates with levels of lipopolysaccharide and markers of systemic immune activation. *Mucosal Immunol United States*; 2016;9:1455-65.
39. Huang CT, Workman CJ, Flies D, Pan X, Marson AL, Zhou G, et al. Role of LAG-3 in regulatory T cells. *Immunity*. 2004;21:503-13.
40. Anderson AC, Joller N, Kuchroo VK. Lag-3, Tim-3, and TIGIT: Co-inhibitory Receptors with Specialized Functions in Immune Regulation. *Immunity*. 2016;44:989-1004.
41. Fougeray S, Brignone C, Triebel F. A soluble LAG-3 protein as an immunopotentiator for therapeutic vaccines: Preclinical evaluation of IMP321. *Vaccine*. 2006;24:5426-33.
42. Vinay DS, Kwon BS. 4-1BB (CD137), an inducible costimulatory receptor, as a specific target for cancer therapy. *BMB Rep*. 2014;47:122-9.
43. Fröhlich A, Loick S, Bawden EG, Fietz S, Dietrich J, Diekmann E, et al. Comprehensive analysis of tumor necrosis factor receptor TNFRSF9 (4-1BB) DNA methylation with regard to molecular and clinicopathological features, immune infiltrates, and response prediction to immunotherapy in melanoma. *EBioMedicine*. 2020;52:1-14.
44. Jung HW, Choi SW, Choi JIL, Kwon BS. Serum concentrations of soluble 4-1BB, and 4-1BB ligand correlated with the disease severity in rheumatoid arthritis. *Exp Mol Med*. 2004;36:13-22.
45. Perreau M, Vigano S, Bellanger F, Pellaton C, Buss G, Comte D, et al. Exhaustion of bacteria-specific CD4 T cells and microbial translocation in common variable immunodeficiency disorders. *J Exp Med*. 2014;211:2033-45.

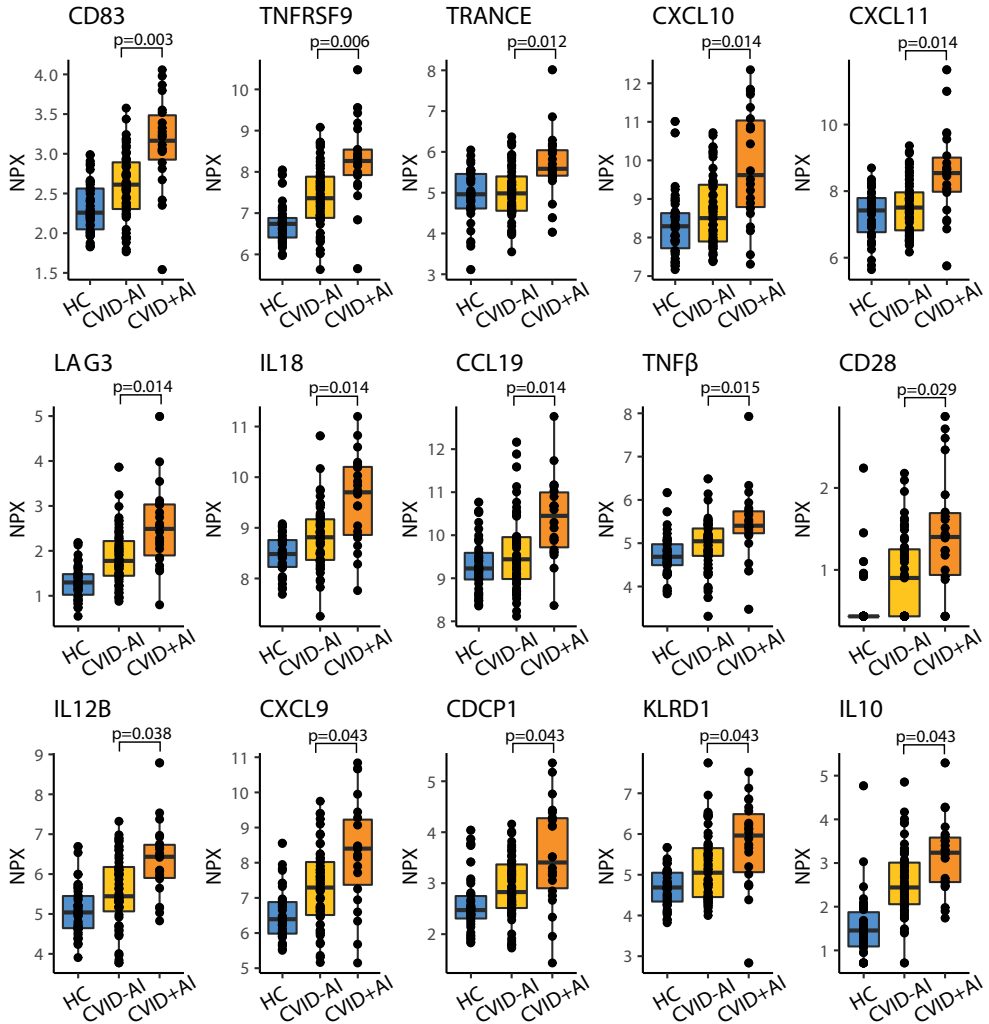
SUPPLEMENTARY DATA SERUM CYTOKINES IN CVID

Differentially expressed proteins in immune dysregulation in CVID



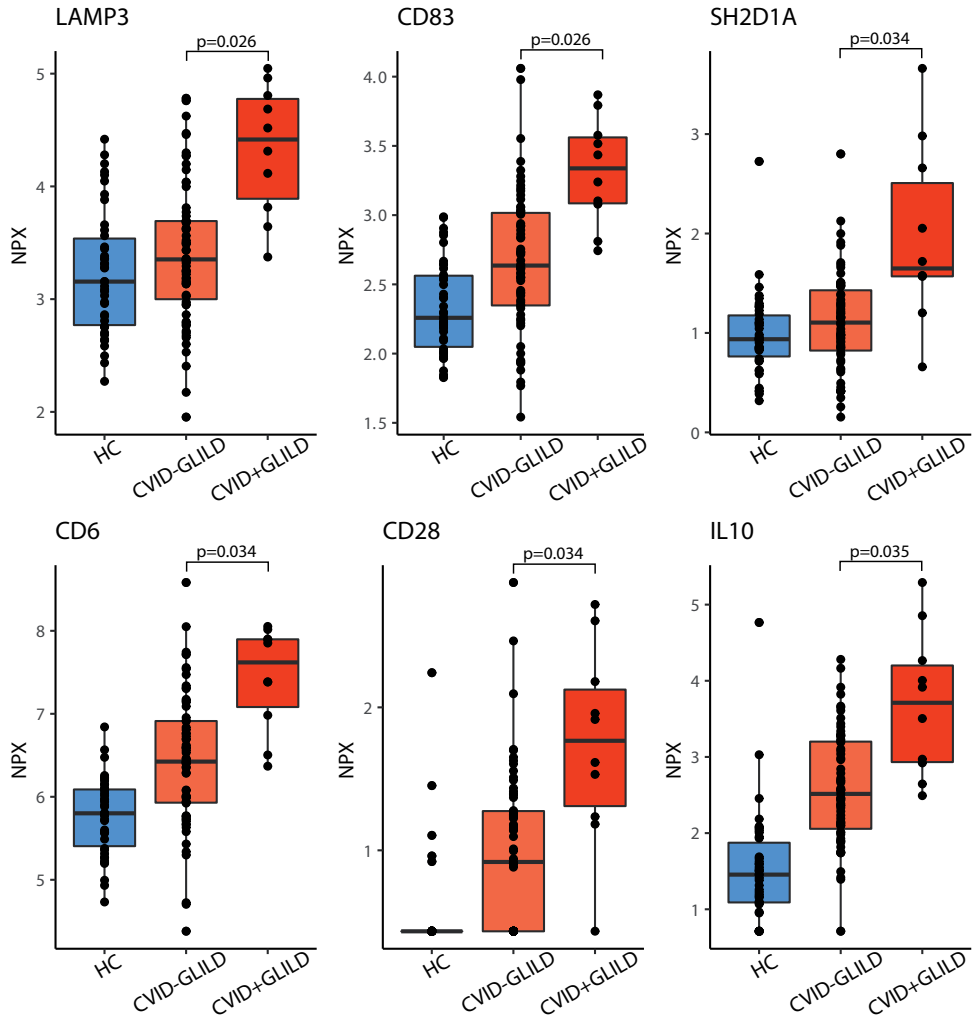
Supplementary figure S1A: Boxplots of differentially expressed proteins (log₂ fold change >0.58 and false discovery rate (FDR)-adjusted p <0.05) in CVID with immune dysregulation (CVIDid, n=37) as compared to CVID with infection only CVIDio, n=40). Healthy controls (HC, n=42) are added as reference. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within 1.5×interquartile range. P-values: Mann-Whitney U test after FDR correction.

Differentially expressed proteins in autoimmunity in CVID



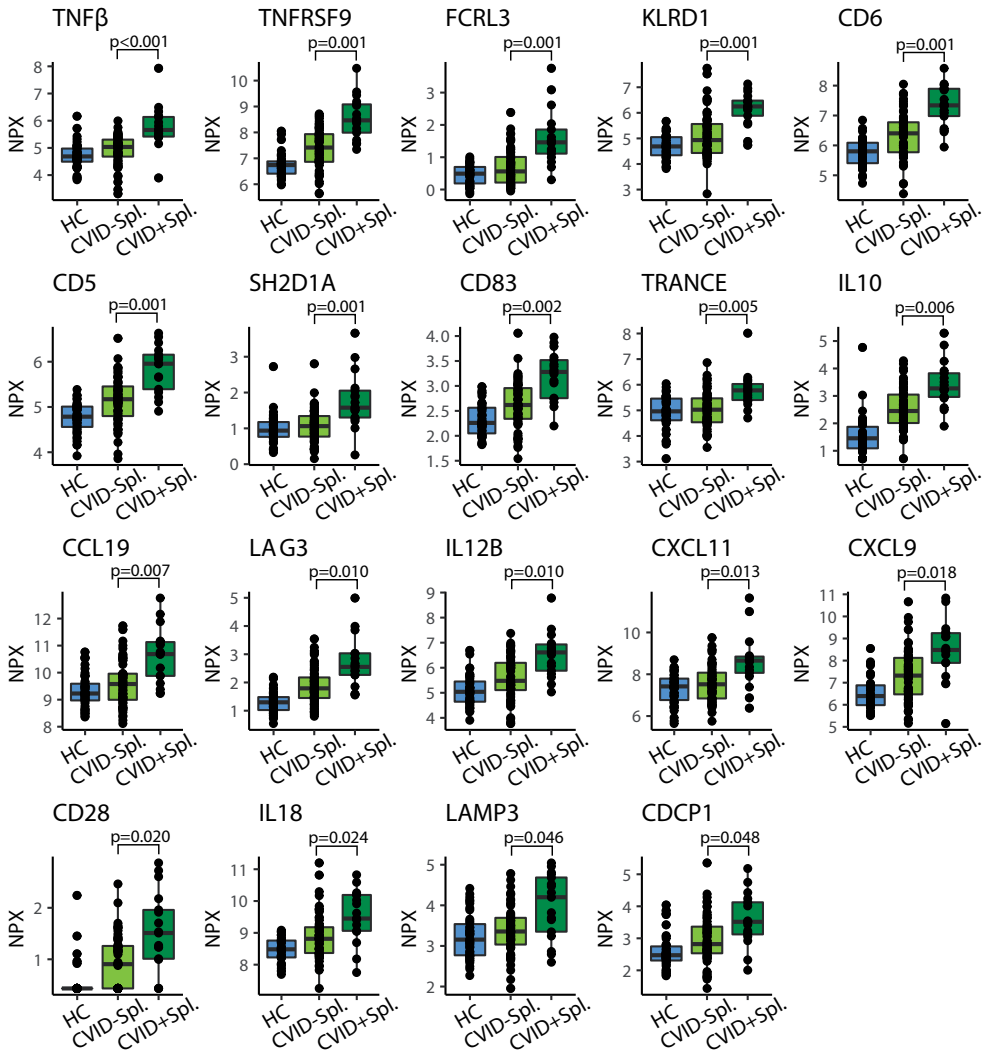
Supplementary figure S1B: Boxplots of differentially expressed proteins (log₂ fold change >0.58 and false discovery rate (FDR)-adjusted $p < 0.05$) upregulated in CVID with autoimmunity ($n=22$) as compared to CVID without autoimmunity ($n=55$). Healthy controls (HC, $n=42$) are added as reference. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within $1.5 \times$ interquartile range. P-values: Mann-Whitney U test after FDR correction.

Differentially expressed proteins in GLILD in CVID

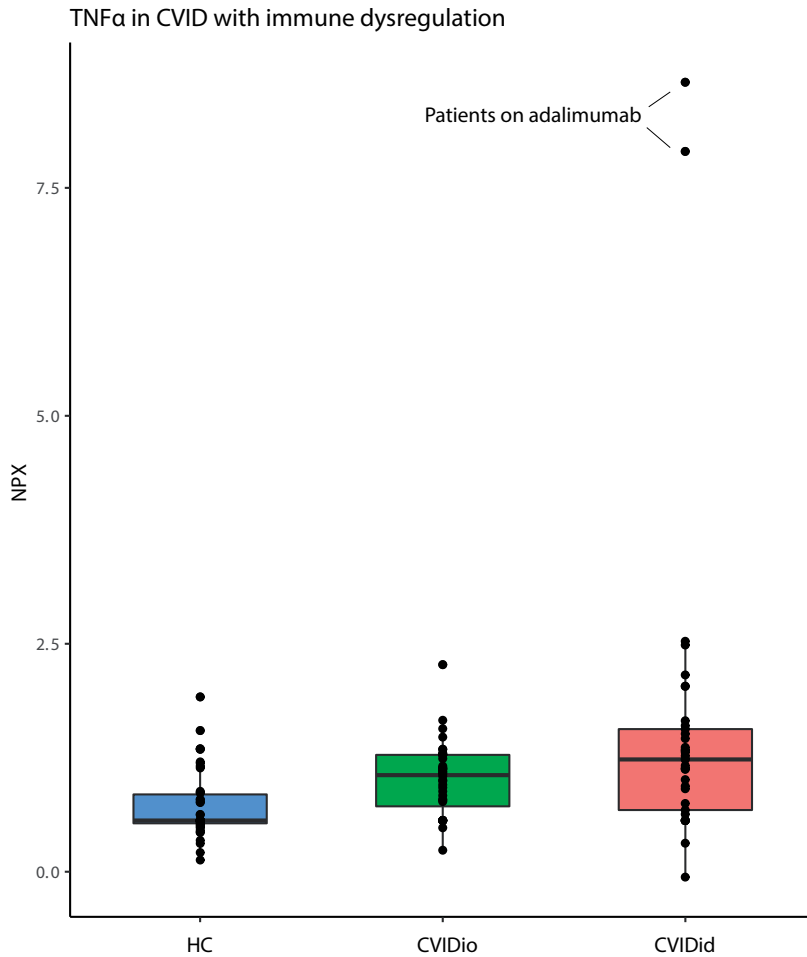


Supplementary figure S1C: Boxplots of differentially expressed proteins (log₂ fold change >0.58 and false discovery rate (FDR)-adjusted p <0.05) upregulated in CVID with granulomatous-lymphocytic interstitial lung disease (GLILD) (n=10) as compared to CVID without GLILD (n=67). Healthy controls (HC, n=42) are added as reference. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within 1.5×interquartile range. P-values: Mann-Whitney U test after FDR correction.

Differentially expressed proteins in splenomegaly in CVID



Supplementary figure S1D: Boxplots of differentially expressed proteins (log2 fold change >0.58 and false discovery rate (FDR)-adjusted $p < 0.05$) upregulated in CVID with splenomegaly (CVID+Spl., $n=17$) as compared to CVID without splenomegaly (CVID-Spl., $n=60$). Healthy controls (HC, $n=42$) are added as reference. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within $1.5 \times$ interquartile range. P-values: Mann-Whitney U test after FDR correction.



Supplementary figure S2: Boxplots of TNF- α in CVID with immune dysregulation (CVIDid, n=37) as compared to CVID with infection only CVIDio, n=40). Healthy controls (HC, n=42) are added as reference. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within 1.5 \times interquartile range. P-values: Mann-Whitney U test after false discovery rate correction.

Supplementary table S1: has been added as a separate file.

Simplified biomarker list:

4E-BP1	CLEC4C	FAM3B	IL-17C	LIF-R	PTH1R
ADA	CLEC4D	FCRL3	IL18	LILRB4	SCF
AREG	CLEC4G	FCRL6	IL-18R1	LY75	SH2B3
ARNT	CLEC6A	FGF-19	IL2	MASP1	SH2D1A
ARTN	CLEC7A	FGF2	IL-20	MCP-1	SIRT2
AXIN1	CNTNAP2	FGF-21	IL-20RA	MCP-2	SIT1
BACH1	CSF-1	FGF-23	IL-22 RA1	MCP-3	SLAMF1
Beta-NGF	CST5	FGF-5	IL-24	MCP-4	SPRY2
BIRC2	CX3CL1	Flt3L	IL-2RB	MGMT	SRPK2
BTN3A2	CXADR	FXYD5	IL33	MILR1	ST1A1
CASP-8	CXCL1	GALNT3	IL4	MMP-1	STAMPB
CCL11	CXCL10	GDNF	IL5	MMP-10	STC1
CCL19	CXCL11	GLB1	IL6	NCR1	TANK
CCL20	CXCL12	HCLS1	IL7	NF2	TGF-alpha
CCL23	CXCL5	HEXIM1	IL8	NFATC3	TNF
CCL25	CXCL6	HGF	IRAK1	NRTN	TNFB
CCL28	CXCL9	HNMT	IRAK4	NT-3	TNFRSF9
CCL3	DAPPI	HSD11B1	IRF9	NTF4	TNFSF14
CCL4	DCBLD2	ICA1	ITGA11	OPG	TPSAB1
CD244	DCTN1	IFN-gamma	ITGA6	OSM	TRAF2
CD28	DDX58	IFNLR1	ITGB6	PADI2	TRAIL
CD40	DFFA	IL-1 alpha	ITM2A	PD-L1	TRANCE
CD5	DGKZ	IL10	JUN	PIK3AP1	TREM1
CD6	DNER	IL-10RA	KLRD1	PLXNA4	TRIM21
CD83	DPP10	IL-10RB	KPNA1	PPP1R9B	TRIM5
CD8A	EDAR	IL-12B	KRT19	PRDX1	TSLP
CDCP1	EGLN1	IL12RB1	LAG3	PRDX3	TWEAK
CDSN	EIF4G1	IL13	LAMP3	PRDX5	uPA
CKAP4	EIF5A	IL-15RA	LAP TGF-beta-1	PRKDC	VEGFA
CLEC4A	EN-RAGE	IL-17A	LIF	PSIP1	ZBTB16

Supplementary table S2: results of genetic screening in training and testing cohort. CVIDid: CVID with immune dysregulation, CVIDio: CVID with infections only, VUS: variants of unknown significance.

VUS found in training cohort	Pathogenic mutations found in training cohort
CVIDid: Heterozygote JAK3, VUS in PLCG2	CVIDio: TNFRSF13B (TAC1)
CVIDid: Heterozygote UNC13D	CVIDid: CTLA4
	CVIDid: STAT1 GOF
VUS found in testing cohort	Pathogenic mutations found in testing cohort
CVIDid: PRKDC	CVIDid: TNFRSF13B (3x)
	CVIDid: PIK3R1

Supplementary table S3A: Performance of random forest, elastic net and extreme gradient boosting algorithms trained on training data. Area under the curve (AUC) based on Receiver-Operator Curves of classification performance on testing cohort. Sensitivity and specificity based on threshold selected for maximum Youden's Index.

	Random Forest	ElasticNet	ExtremeGradientBoosting
AUC	0.71	0.72	0.71
Sensitivity	0.70	0.74	0.78
Specificity	0.71	0.71	0.67

Supplementary table S3B: standardized regression coefficients of the elastic net algorithm

ElasticNet	standardized regression coefficients
MILR1	49.5275225
LILRB4	45.9481322
IL10	39.4471273
IL12RB1	32.5195959
CXCL10	26.8637128
BTN3A2	23.6616556
CD83	22.3648257
IL17A	22.1304286
CCL19	20.07503
CXCL5	19.4975474
CLEC6A	18.9570038
LAG3	18.6745838
IL5	18.0662128
FCRL6	17.8391726
CXCL11	16.1581789
CCL20	14.5387933
IL18	13.3605353
SH2D1A	9.0432834
CD28	8.8901486
CKAP4	2.8303555
CXCL9	1.4464496
ITM2A	0.9145627
KLRD1	0.1589813

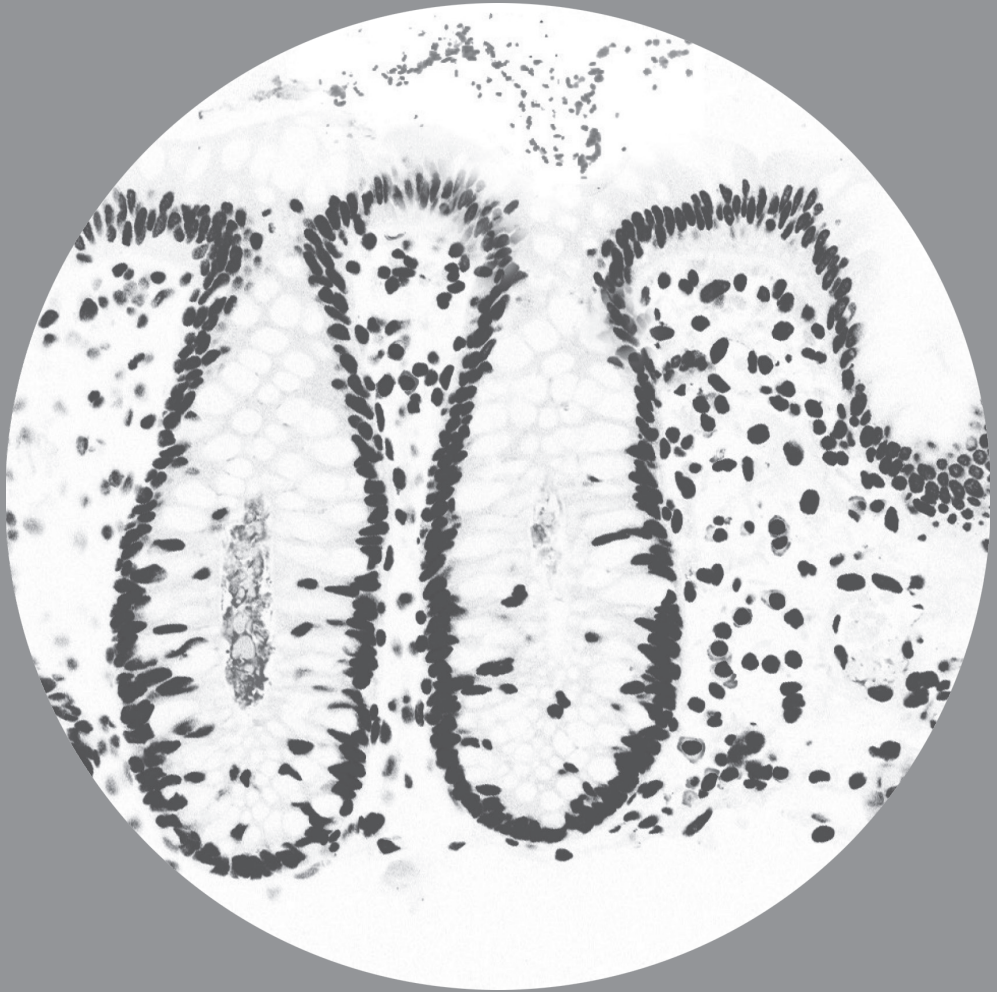
Supplementary table S3C: standardized regression coefficients of the extreme gradient boosting algorithm

ExtremeGradientBoosting	standardized regression coefficients
MILR1	46.19873132
CD83	24.38600922
LILRB4	22.38396687
IL10	16.60761097
CLEC6A	10.40267680
IL12RB1	9.87841521
CCL25	8.86002056
DCBLD2	8.42205461
TRANCE	6.53458466
CCL20	6.44588911
ITM2A	5.74735616
SH2DIA	4.86901403
IL18	4.57302688
PSIP1	4.51630353
BTN3A2	4.39895158
CD244	3.16099943
KLRD1	2.44351124
CXCL10	2.08361928
CD28	2.02518904
FGF.21	1.95971723
NCRI	1.81192499
NFATC3	1.75484145
CXCL11	1.67102412
FCRL6	1.49078755
X4E_BP1	1.27968655
CKAP4	1.27881171
KRT19	1.03650139
MMP.1	1.01803351
HNMT	0.89113541
CCL19	0.84695373
CXADR	0.84156278
IL5	0.82510504
CCL11	0.82376997
Beta.NGF	0.82260374
CD5	0.79428321
BACH1	0.78728905
FGF.19	0.68545008
CLEC4G	0.67419116
TNFB	0.64908048
TRIM5	0.63198289
IL.17C	0.63148997
IL7	0.61091238
TANK	0.60665911

ExtremeGradientBoosting	standardized regression coefficients
CXCL5	0.59876505
TGF.alpha	0.53295115
GALNT3	0.53198514
ZBTB16	0.51442105
CCL28	0.4742683
DNER	0.21160083
IFNLR1	0.20355936
MCP.3	0.20194989
GDNF	0.19465451
PIK3AP1	0.19146449
ITGA6	0.19084172
CDSN	0.18973834
TNF	0.18456823
MASPI	0.18163372
DPP10	0.17539809
MCP.2	0.17471131
HEXIM1	0.17235877
IL10RA	0.16277171
PTH1R	0.16087889
HSD11B1	0.1604582
EIF4G1	0.16014179
EDAR	0.15683554
CST5	0.15533562
PLXNA4	0.14730777
CLEC7A	0.14674432
LAG3	0.13184895
CXCL9	0.12793841
IFN.gamma	0.1270753
CCL4	0.1263282
MGMT	0.1117052
ITGA11	0.09234367
NT.3	0.0868073
CCL3	0.08641593
LAP.TGF.beta.1	0.08193079
CLEC4A	0.07763499
EN.RAGE	0.07434307
TRAIL	0.07028783
LIF	0.06844958
IL8	0.06356907

Supplementary table S4: post-hoc analysis assessing the prediction accuracy of reduced combinations of the classifying model on the testing cohort. AUC: area under the curve

	AUC	sensitivity	specificity
IL10+IL12RB1+CD83	0.765	0.786	0.708
IL10+IL12RB1	0.772	0.739	0.750
IL10+CD83	0.741	0.780	0.708
IL12RB1+CD83	0.748	0.739	0.625
IL10	0.743	0.826	0.625
IL12RB1	0.730	0.870	0.500
CD83	0.728	0.826	0.583



Chapter 3

Chronically Activated T-cells retain their Inflammatory Properties in Common Variable Immunodeficiency

Roos-Marijn Berbers
M. Marlot van der Wal
Joris M. van Montfrans
Pauline M. Ellerbroek
Virgil A.S.H. Dalm
P. Martin van Hagen
Helen L. Leavis*
Femke van Wijk*

* These authors contributed equally to this work

Journal of Clinical Immunology 2021

ABSTRACT

Purpose

Immune dysregulation complications cause significant morbidity and mortality in common variable immunodeficiency (CVID), but the underlying pathophysiology is poorly understood. While CVID is primarily considered a B-cell defect, resulting in the characteristic hypogammaglobulinemia, T-cells may also contribute to immune dysregulation complications. Here, we aim to further characterize T-cell activation and regulation in CVID with immune dysregulation (CVIDid).

Methods

Flow cytometry was performed to investigate T-cell differentiation, activation and intracellular cytokine production, negative regulators of immune activation, regulatory T-cells (Treg), and homing markers in 12 healthy controls, 12 CVID patients with infections only (CVIDio) and 20 CVIDid patients.

Results

Both CD4+ and CD8+ T-cells in CVIDid showed an increased activation profile (HLA-DR+, Ki67+, IFN γ +) when compared to CVIDio, with concomitant upregulation of negative regulators of immune activation PD1, LAG3, CTLA4, and TIGIT. PD1+ and LAG3+ subpopulations contained equal or increased frequencies of cells with the capacity to produce IFN γ , Ki67, and/or GzmB. The expression of PD1 correlated with serum levels of CXCL9, 10 and 11. Treg frequencies were normal to high in CVIDid, but CVIDid Tregs had reduced CTLA-4 expression, especially on CD27+ effector Tregs. Increased migratory capacity to inflamed and mucosal tissue was also observed in CVIDid T-cells.

Conclusion

CVIDid was characterized by chronic activation of peripheral T-cells with preserved inflammatory potential rather than functional exhaustion, and increased tissue migratory capacity. While Treg numbers were normal in CVIDid Tregs low levels of CTLA-4 indicate possible Treg dysfunction. Combined studies of T-cell dysfunction and circulating inflammatory proteins may direct future treatment strategies.

Key words (4-6)

Immune Dysregulation
Common Variable Immunodeficiency (CVID)
T-cells
Immune exhaustion
Regulatory T-cells
Autoimmunity

INTRODUCTION

Common Variable Immunodeficiency (CVID) is characterized by recurrent infections caused by low immunoglobulin (Ig)G, and IgA or IgM¹, for which patients are treated with immunoglobulin G replacement therapy (IgRT)². IgRT has significantly decreased the risk of infectious complications in CVID, but nonetheless over a third of patients develop immune dysregulation complications^{3,4}, resulting in significant morbidity and mortality^{5,6}. A wide range of immune dysregulation phenomena can be observed in CVID, including granulomatous-lymphocytic interstitial lung disease (GLILD), enteritis, autoimmune cytopenias, lymphoproliferation and hematological malignancies^{4,7}. The underlying pathophysiology of immune dysregulation in CVID is currently poorly understood, which complicates diagnostics and treatment^{8,9}.

While the defining hypogammaglobulinemia in CVID is considered to be primarily the result of B-cell dysfunction, several lines of evidence suggest an additional role for T-cells in CVID with immune dysregulation (CVIDid). Biopsies of lung granulomas in CVIDid show predominance of CD4+ T helper (Th) cells^{10,11}, while regulatory T-cells (Tregs) are often absent¹⁰. In peripheral blood of patients with CVIDid, a decreased CD4/CD8 ratio was observed with decrease of naïve T-cells¹², Tregs, Th17 cells, and follicular helper T (Tfh) cells^{13,14}. Moreover, there are indications that CVID T-cells may be functionally exhausted, including reduced capacity to respond to bacterial antigens and increased expression of PD1^{15,16}. Our group and others have previously demonstrated that serum cytokines in CVIDid^{17,18} are shifted towards a Th1 phenotype, and we observed an upregulation of proteins associated with immune regulation – IL10, LAG3, and 4-1BB. Monogenic primary immunodeficiencies caused by mutations in immune regulation genes such as CTLA4¹⁹ and ICOS²⁰ often result in a CVIDid phenotype. However, how the interplay between immune regulation and immune activation results in CVIDid remains poorly understood.

To further study the balance between immune activation and immune regulation in CVIDid, we used flow cytometry to evaluate naïve T-cell subsets, T-cell activation and cytokine production, exhaustion, negative regulators of immune activation, regulatory T-cells, and T-cell homing markers.

METHODS

Ethics statement

Ethical approval for this study for all participants was received from the Medical Ethical Committee of the Erasmus MC University Medical Center in Rotterdam, the Netherlands (METC: 2013-026). Written informed consent was obtained from all patients and controls according to the Declaration of Helsinki.

Study population and sample collection

Patients were diagnosed with CVID according to the European Society for Immunodeficiencies criteria¹ and were included at the outpatient clinics of the UMC Utrecht and

the Erasmus MC University Medical Center Rotterdam, the Netherlands. Patients were eligible if they were aged seven or older, and they and/or their legal guardians signed informed consent. Household members of patients were recruited as healthy controls (HC). Medication use up to three months prior to sampling was recorded.

Sample processing

Peripheral blood mononuclear cells (PBMC) were isolated from blood by ficoll-density centrifugation (GE Healthcare-Biosciences, AB), and frozen at -180°C until use. Cells were subsequently thawed, counted and plated at 1,000,000 live cells per panel per sample. For the panels including intracellular cytokine measurement, cells were first restimulated with 20 ng/mL phorbol 12-myristate 13-acetate (PMA, MilliporeSigma) and 1 $\mu\text{g}/\text{mL}$ ionomycin (MilliporeSigma) for 4 hours at 37°C with addition of Monensin (Golgistop, BD Biosciences, 1:1500) during the last 3.5 hours. Cell death was stained in all panels using Fixable Viability Dye eFluor 506 (eBioscience). Next, cells were incubated with the surface antibodies (Supplementary Table 1) for 20 minutes at 4°C and washed. Cells were then permeabilized with fixation / permeabilization reagent (eBioscience) for 30 minutes at 4°C , washed, and incubated with the intracellular antibodies (Supplementary Table 1). Cells were stored at 4°C until the next day, when they were measured on the LSR Fortessa (BD Biosciences).

Analysis and statistics

For flow cytometric data, median fluorescence intensities (MFI) and percentages of positive cells were analyzed in FlowJo. All statistical analyses and graphic representations were done in R 3.2.0²¹. Continuous variables were compared using the Mann-Whitney rank test, or paired Wilcoxon-Rank test for paired samples. Correlation was calculated using Spearman's correlation. P values below 0.05 were considered statistically significant.

RESULTS

The overall T-cell profile of CVIDid patients differs from CVIDio and HC

To investigate immune activation and regulation in CVID, PBMC were isolated from 12 healthy controls (HC), 12 CVID patients with infections only (CVIDio), and 20 patients with CVIDid (Table 1). Patients were selected from a cross-sectional Dutch primary immunodeficiency cohort when they received IgRT and did not use any immunomodulatory medication during and the last 3 months prior to sampling. Three of the CVIDid patients had a known CVID-associated monogenetic disease (CTLA4 haploinsufficiency, STAT1 gain of function and PIK3R1). Flow cytometry was performed as described in the supplementary information (see also Supplementary Table 1 and Supplementary Figures 1-4). First, pooled flow cytometry data was analyzed in an unsupervised manner using principal component analysis (PCA, Figure 1A). CVIDio and HC clustered closely together and

Table 1: cohort characteristics. HC: healthy control. CVIDio: CVID with infections only. CVIDid: CVID with immune dysregulation. IQR: interquartile range. DMARD: disease modifying anti-rheumatic drug. VUS: variant of unknown significance.

Summary Statistics	HC (n=12)	CVIDio (n=12)	CVIDid (n=20)
Characteristics			
age (median, IQR)	45.50 (40.25, 52.00)	38.50 (29.00, 58.25)	38.50 (35.75, 43.50)
sex (male)	5 (41.67%)	4 (33.33%)	11 (55.00%)
center (Utrecht)	9 (75.00%)	11 (91.67%)	15 (75.00%)
antibiotics	0 (0.00%)	3 (25.00%)	8 (40.00%)
immunosuppressive medication	0 (0.00%)	0 (0.00%)	0 (0.00%)
IgA <0.1 g/L	0 (0.00%)	2 (16.67%)	13 (65.00%)
Immune dysregulation complications			
Pulmonary	0 (0.00%)	0 (0.00%)	8 (40.00%)
Hematological	0 (0.00%)	0 (0.00%)	2 (10.00%)
Gastrointestinal	0 (0.00%)	0 (0.00%)	8 (40.00%)
Rheumatological	0 (0.00%)	0 (0.00%)	6 (30.00%)
Dermatological	0 (0.00%)	0 (0.00%)	4 (20.00%)
Hematological malignancy	0 (0.00%)	0 (0.00%)	2 (10.00%)
Lymphoproliferation (incl splenomegaly)	0 (0.00%)	0 (0.00%)	10 (50.00%)
Other	0 (0.00%)	0 (0.00%)	4 (20.00%)
DMARD-naïve / subclinical disease	NA	NA	12 (60.00%)
Disease in remission	NA	NA	8 (40.00%)
Genetics			
Not done	12 (100.00%)	12 (100.00%)	12 (60.00%)
Nothing found	0 (0.00%)	0 (0.00%)	2 (10.00%)
VUS found	0 (0.00%)	0 (0.00%)	2 (10.00%)
Pathogenic mutations found	0 (0.00%)	0 (0.00%)	3 (15.00%) + 1 TNFRSF13B (TACI) mutation

were distinct from most CVIDid samples. This suggests that most variation in the flow cytometric T-cell data related to immune dysregulation and not to the hypogammaglobulinemia shared by CVID patients. CVIDid patients did not cluster by treatment history or location of autoimmunity (Supplementary Figure 5), suggesting that peripheral blood T-cell skewing may be generalized among CVIDid patients.

CVIDid T-cells are Th-1 skewed and chronically activated

Next, the distribution of CD4⁺ and CD8⁺ T-cell subsets in CVIDid was assessed. Frequencies in CD4⁺, CD8⁺ and naïve/memory subsets were similar to that reported in previous studies^{12,22}: a decreased CD4/CD8 ratio (Figure 1B), and a trend of decreased naïve (CD45RA+CCR7⁺) CD4⁺ T-cells in CVIDid compared to CVIDio (p=0.07) (Figure 1C). In addition, a subset of CVIDid patients showed a high percentage of CD4⁺ T effector memory cells, and CD4⁺ T effector memory cells re-expressing CD45RA (TEMRA), which

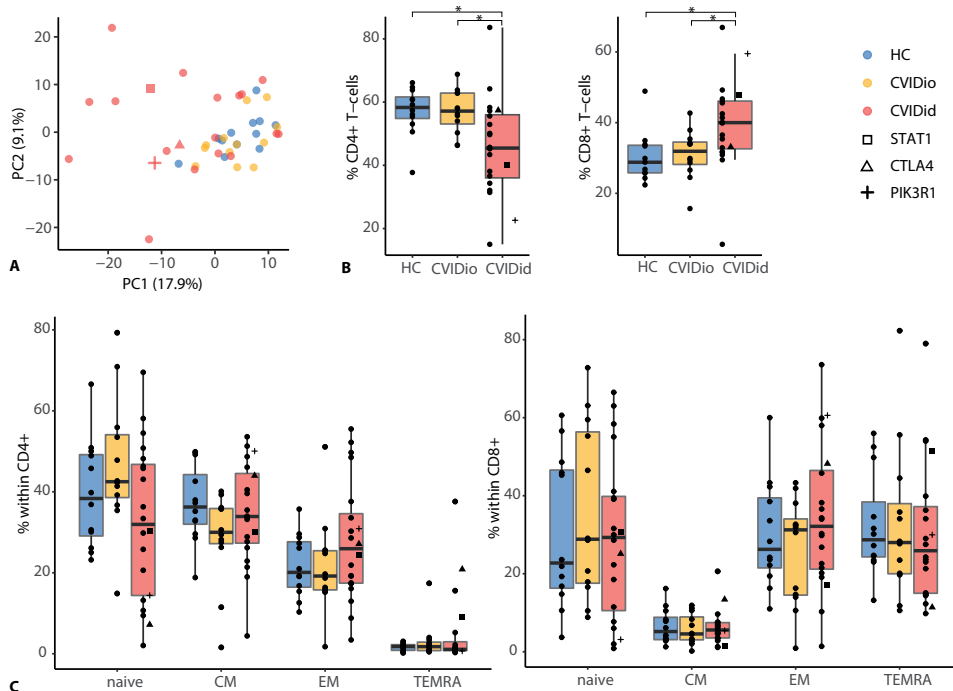


Figure 1: General description of T-cell subsets in CVID. **A**) Principal component analysis of FACS data (all panels combined). **B**) CD4+ and CD8+ T-cells. **C**) naïve (CD45RA+CCR7+), central memory (CM: CD45RA-CCR7+) and effector memory (EM: CD45RA-CCR7-) and terminally differentiated effector memory cells (TEMRA: CD45RA+CCR7-) in CD4+ T-cells. CVIDid = CVID with immune dysregulation (n=20), CVIDio = CVID with infections only (n=12), HC = healthy controls (n=12). Statistics: Mann-Whitney U-test. * p<0.05, ** p<0.01, *** p<0.001.

are associated with chronic activation such as observed in viral infections²³ (Figure 1C). Within the naïve CD4+ T-cells, CD31+ recent thymic emigrants were more abundant in CVIDid than the CD31- central naïve T-cells (Supplementary Figure 6). No differences in CD8+ T-cell distribution were observed (Figure 1C).

Within the effector/memory (CD45RO+) subpopulation, proportions of CD4+ T-cells expressing HLA-DR, Ki67 and IFN γ were significantly increased in CVIDid, while IL17a, IL13, and TNF- α expressing T-cells were not different between CVIDid and CVIDio (Figure 2A). A similar activation pattern was observed in the CD8+ effector memory population for HLA-DR, Ki67 and GzmB (Figure 2B). We previously described¹⁷ serum cytokine and chemokine levels in an overlapping cohort, allowing comparison between soluble serum markers and T-cell characteristics. Pooling the data from HC, CVIDio and CVIDid, we observed that the proportion of IFN γ + Th cells correlated with serum levels of interferon-inducible chemokines CXCL9, 10, and 11 (Figure 2C). While we observed an increase of serum IL17a in CVIDid, there was no corresponding increase of IL17a-producing Th cells, and the frequency of these cells did not correlate with the serum IL17a levels (Supplementary Figure 7).

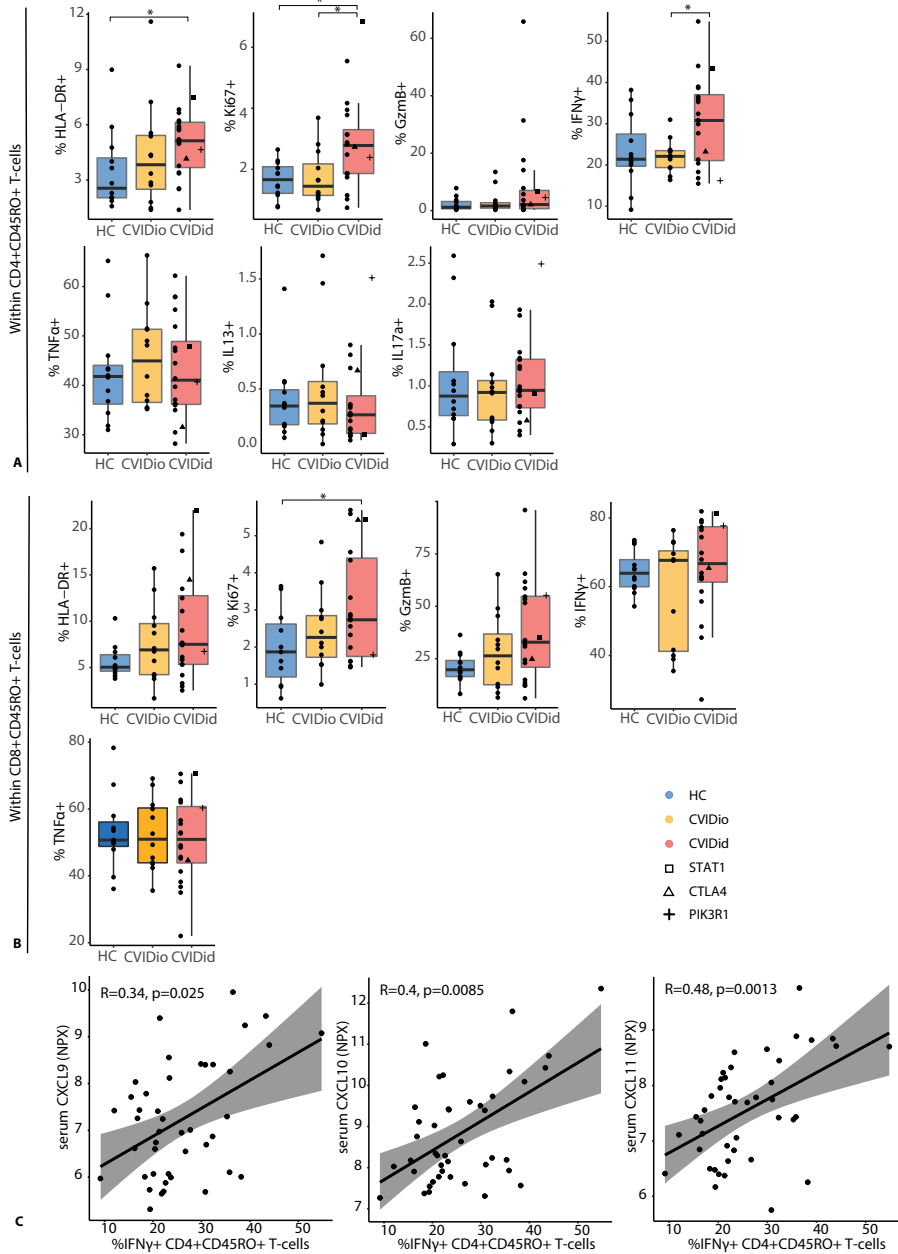


Figure 2: Expression of activation markers HLA-DR, Ki67 and GzmB, and of intracellular cytokines after PMA/ionomycin stimulation: IFN γ , TNF α , IL-13 and IL-17a. **A**) in CD4+CD45RO+ T-cells, **B**) in CD8+ CD45RO+ T-cells. CVIDid = CVID with immune dysregulation (n=20), CVIDio = CVID with infections only (n=12), HC = healthy controls (n=12). Statistics: Mann-Whitney U-test. * p<0.05, ** p<0.01, *** p<0.001. **C**) %IFN γ + CD4+ T-cells correlate with serum levels of CXCL9, CXCL10 and CXCL11. Statistics: Spearman correlation.

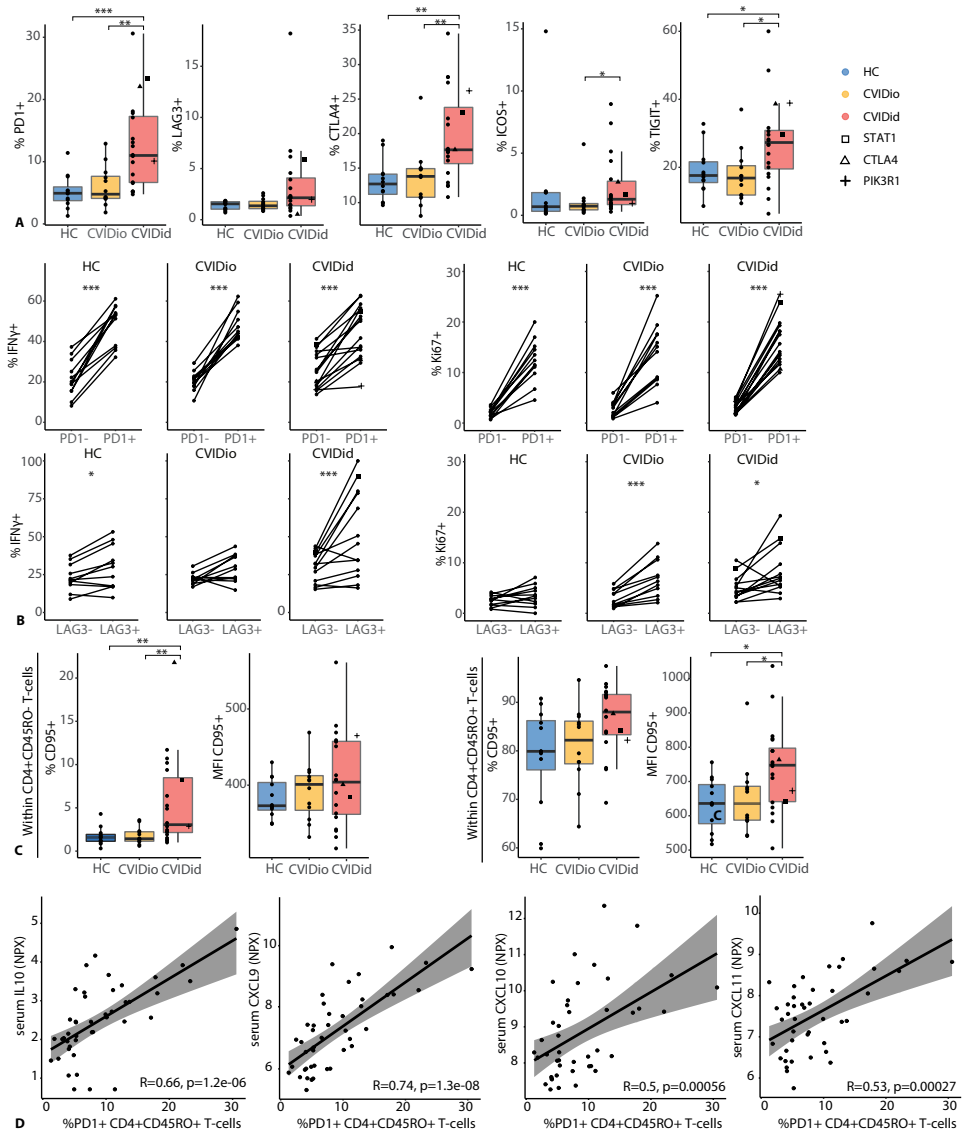


Figure 3: Negative regulators of immune activation in CVID. **A)** proportions of PD1, LAG3, CTLA4, ICOS and TIGIT in CD4+CD45RO+ T-cells **B)** percentage and median fluorescence intensity of CD95 (FAS-L) in naïve (CD45RO-) and effector-memory (CD45RO+) CD4+ T-cells **C)** comparison of IFN γ + and Ki67+ cells in PD1 and LAG3 positive and negative populations. Only samples with >50 events in the PD1/LAG3 positive and PD1/LAG3 negative populations were included. **D)** Spearman correlation between PD1 and IL10, CXCL9, CXCL10 or CXCL11. CVIDid = CVID with immune dysregulation (n=20), CVIDio = CVID with infections only (n=12), HC = healthy controls (n=12). Statistics (A&C): Mann-Whitney U-test. Statistics B: paired Wilcoxon-Rank test. * p<0.05, ** p<0.01, *** p<0.001.

T-cells expressing negative regulators of immune activation retain their inflammatory potential in CVIDid

Next, we investigated whether this immune activation resulted in immune exhaustion in CVIDid, which is known to happen in the context of chronic inflammation²⁴. In CVIDid CD4⁺-cells, we observed an increased proportion of cells expressing negative regulators of immune activation PD1, LAG3, CTLA4, ICOS and TIGIT (Figure 3A). In addition, CVIDid CD4⁺-T-cells expressed higher levels of CD95 (FAS-L), showing that they are terminally differentiated and may be more prone to apoptosis (Figure 3B). In CD8⁺ T-cells, LAG3, CTLA4, ICOS and TIGIT, but not PD1 and CD95, were similarly increased in CVIDid (Supplementary Figure 8).

In order to assess whether the PD1 and LAG3 expressing cells were functionally exhausted, production of IFN γ , GzmB, Ki67 and CD95 was compared between the PD1/LAG3⁺ and PD1/LAG3⁻ populations (Figure 3C and Supplementary Figure 9). In CD4⁺ T-cells, the PD1⁺ and LAG3⁺ subpopulations showed equal or increased expression of IFN γ , Ki67, and/or GzmB as the PD1⁻ and LAG3⁻ subpopulations, indicating retained functional capacity. Expression of PD1 correlated with serum levels of pro-inflammatory CXCL9, 10, and 11 as well as immune regulatory IL10 (Figure 3D), further illustrating the association between PD1 and chronic inflammation. In CD8⁺ T-cells, IFN γ and GzmB expression was also intact in PD1⁺ and LAG3⁺ cells, but Ki67⁺ cells were less frequent in the LAG3⁺ subpopulation, indicating reduced proliferative capacity (Supplementary Figure 10).

CVIDid regulatory T-cells fail to upregulate CTLA4

In addition to negative regulators of immune activation, regulatory T (Treg) cells also contribute to limiting inflammation-related pathology. In contrast to previous studies²⁵, we did not observe a decreased proportion of CD25⁺FOXP3⁺ Tregs in CVIDid, and a subgroup of CVIDid patients had even increased Treg frequencies (Figure 4A). ICOS expression in CVIDid Tregs also indicated normal to increased activation²⁶ of Tregs (Figure 4B), but CTLA4 expression was reduced in CVIDid Tregs, both in cell frequency as MFI in the CTLA4 positive population (Figure 4B). The failure to upregulate CTLA4 in CVIDid Tregs was isolated to the activated and usually highly suppressive²⁷ CD27⁺ Treg population (Figure 4C), and may result in reduced Treg function, as expression of CTLA4 is important for the suppressive function of Tregs²⁸. In addition, a subpopulation of CVIDid patients showed low levels of TIGIT-expressing Tregs, which may impair TIGIT-driven selective suppression of Th1 and Th17 effector cells²⁹.

Increased migratory capacity in CVIDid T-cells

The immune dysregulation of most CVIDid patients occurs not only in the systemic immune system, but also locally in affected tissues, which are often mucosal sites (lung and/or gut). Expression of the mucosal homing marker CCR9 was increased on naïve CD4⁺ and CD8⁺ populations of CVIDid T-cells (Figure 5A), but not on non-naïve (CD45RA⁻) CD4⁺ T-cells (Figure 5B), suggesting that this upregulation was at least not entirely antigen-driven. CVIDid CD8⁺ non-naïve (CD45RA⁻) T-cells (Figure 5C) did also

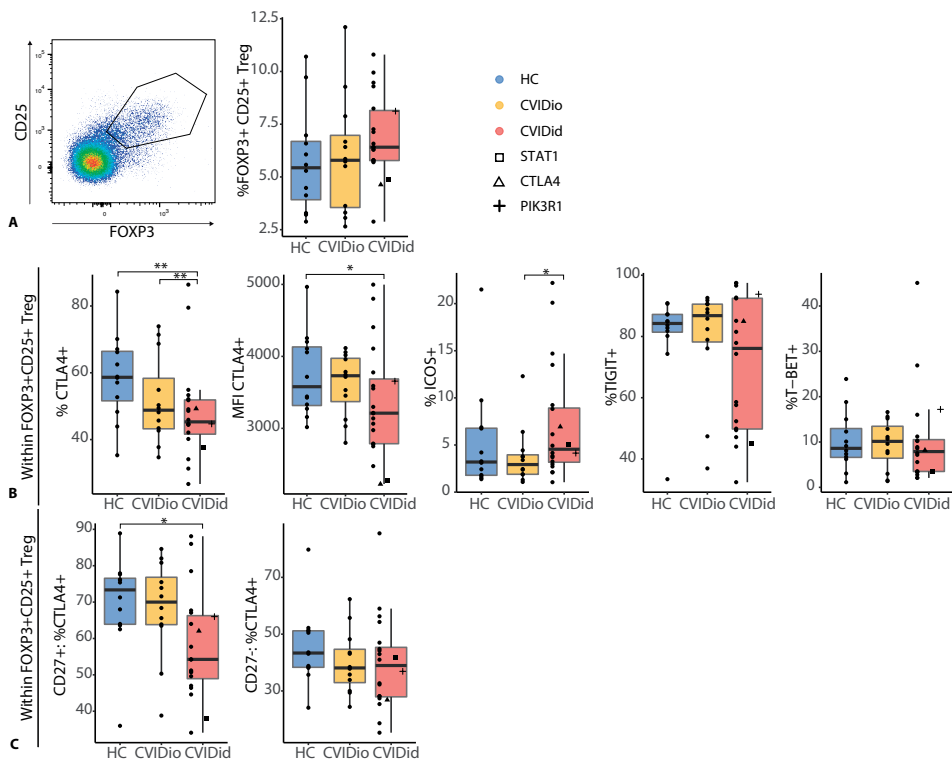


Figure 4: regulatory T-cells in CVID **A)** Gating strategy and proportion of CD25+FOXP3+ T-cells within CD4+ population. **B)** Expression of CTLA4, ICOS, TIGIT and T-BET in the Treg population. **C)** Decreased fraction of CTLA4+ Tregs was confined to the CD27+ population. CVIDid = CVID with immune dysregulation (n=20), CVIDio = CVID with infections only (n=12), HC = healthy controls (n=12). Statistics: Mann-Whitney U-test. * p<0.05, ** p<0.01, *** p<0.001.

show increased expression of CCR9 as well as integrin $\alpha 4\beta 1$, which mediates homing to inflamed tissues, including the lung. CVIDid FOXP3+CD4+ T-cells (Figure 5D) also showed increased migratory capacity, as they were enriched for cells expressing integrin $\alpha 4\beta 1$ and gut-homing marker integrin $\alpha 4\beta 7$.

Patterns of T-cell dysregulation in CVIDid patients with GLILD

Finally, in order to illustrate patterns of T-cell immune dysregulation in individual patients with varying severity of specific immune dysregulation phenotypes, we selected markers most strongly associated with (chronic) immune activation in CD4+ T-cells (HLA-DR, Ki67, IFN γ , and IFN γ in PD1+ and LAG3+, PD1+, LAG3+, TEMRA and CD95) and immune regulation (Tregs, CTLA4 in Tregs, MFI of CTLA4 in Tregs, TIGIT in Tregs, ICOS in Tregs, and CTLA4 in CD4+ T-cells) in CVIDid. We investigated these in five patients with GLILD (Figure 6), ranging from stably mild radiographic GLILD without lung function deterioration (patients 1 and 2), to clinical GLILD that required immunosuppressive treatment shortly after sampling (patients 4 and 5). In patients 1-4, immune activation

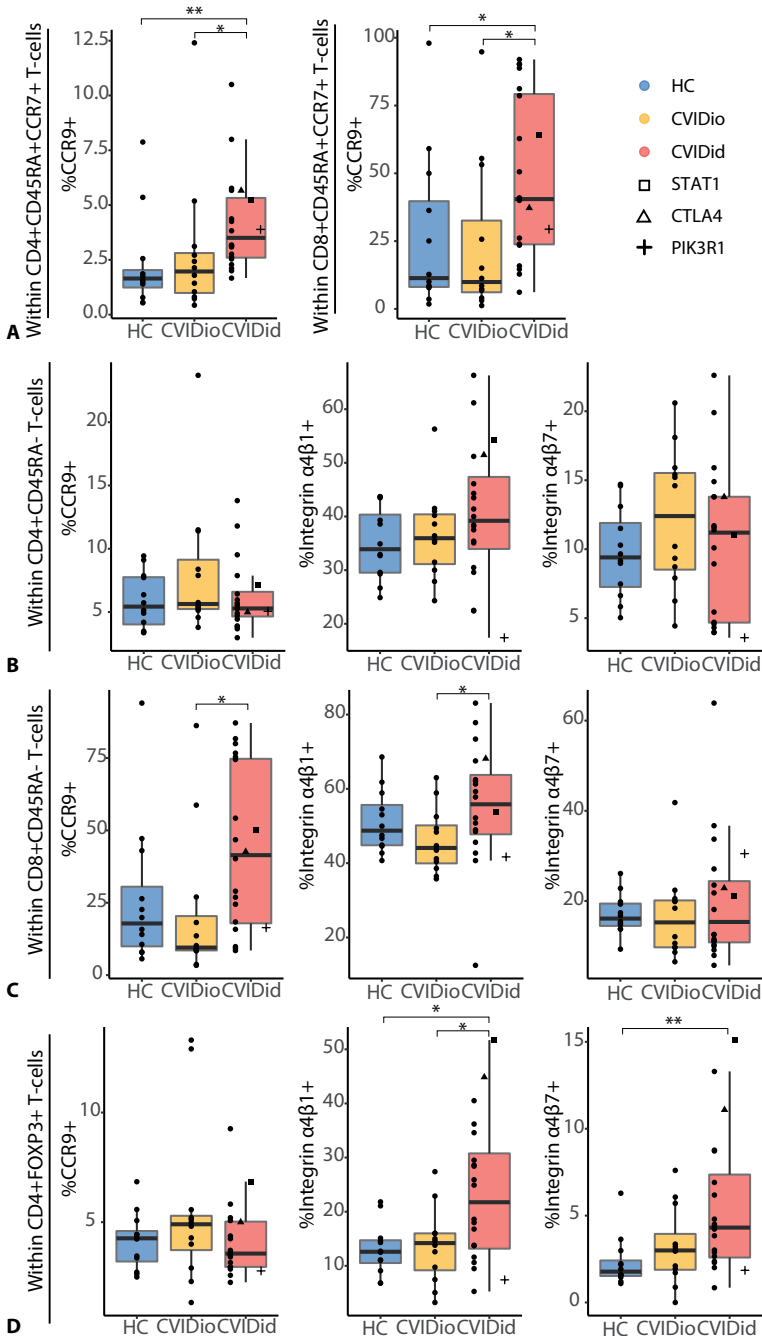


Figure 5: Migratory capacity in CVID: **A**) in CCR7+CD45RA+ naïve CD4 and CD8+ T-cells. **B**) in non-naïve CD45RA- CD4+ T-cells, **C**) in non-naïve CD45RA- CD8+ T-cells, **D**) in FOXP3+ CD4+ T-cells. CVIDid = CVID with immune dysregulation (n=20), CVIDio = CVID with infections only (n=12), HC = healthy controls (n=12). Statistics: Mann-Whitney U-test. * p<0.05, ** p<0.01, *** p<0.001.

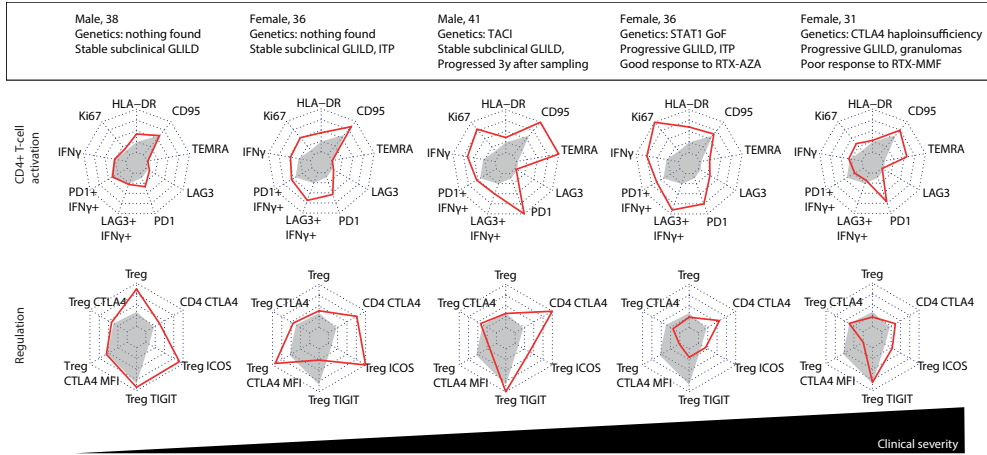


Figure 6: Patterns of CD4+ T-cell activation: % of HLA-DR, Ki67, IFN γ , IFN γ in the PD1+ population and the LAG3+ population, %PD1, %LAG3, %CD45RA+CCR7- (TEMRA) cells and % of CD95. And patterns of T-cell regulation: %CD25+FOXP3+ Treg, %CTLA4 in Treg, MFI of CTLA4 in Treg, %TIGIT in Treg, %ICOS in Treg, %CTLA4 in CD4+ T-cells in five patients with GLILD. Red line indicates individual patient legends, gray shaded area indicates the median for the CVIDio group. Axis ranges are the minimum and the maximum for that marker for the entire cohort.

MFI: median fluorescence intensity. GLILD: granulomatous-lymphocytic interstitial lung disease, VUS: variant of unknown significance. ITP: idiopathic thrombocytopenia purpura. RTX: rituximab. AZA: azathioprine. MMF: mycophenolate mofetil, SCT: stem-cell transplantation. GOF: gain of function.

increased with severity of disease. In patient 3, who at time of sampling had subclinical GLILD that exacerbated three years later, as well as patients 4 and 5, the median fluorescence intensity of CTLA4 on Tregs was decreased, which was not observed for the two subclinical GLILD patients that remained stable (1 and 2). Patient 4 with STAT1 GoF, who required treatment for GLILD shortly after sampling, showed high levels of immune activation and PD1 expression, and low Treg functional markers. Finally, patient 5 with CTLA4 haploinsufficiency demonstrated low CTLA4 expression on Tregs as expected, but also low T-cell activation (HLA-DR, Ki67 and IFN γ) in peripheral blood, despite progressive GLILD and the need for treatment shortly after sampling.

DISCUSSION

In this study, we showed that T-cells in patients with CVIDid were more often activated, proliferating and Th1-skewed than those of patients with CVIDio. In addition, more CVIDid T-cells expressed immune co-inhibitory receptors PD1, LAG3, CTLA4, ICOS and TIGIT, and these cells retained their inflammatory properties. Chronic activation was observed in both the CD4+ and CD8+ compartment. In the Treg compartment we observed low

CTLA₄ expression in CVIDid, while ICOS expression remained intact. Finally, CVIDid T-cells showed increased migratory capacities to mucosal tissues.

High T-cell activation and Th-1 skewing in CVIDid were consistent with previous studies^{12,18}. In addition, we did not observe increased frequencies of IL17-producing T-cells, despite our previous finding of increased IL-17a in serum of the same patients sampled at the same time. This supports the hypothesis for an alternative source of IL17a production in CVIDid, such as type-3 ILCs³⁰.

Previous studies have also reported increased PD1 expression in CVID, and interpreted this as a sign of functional exhaustion and impaired T-cell function^{15,18,31}. However, we observed that the PD1- and LAG3- expressing cells in CVIDid retained the capacity to produce pro-inflammatory cytokines and to proliferate, and thus were not functionally exhausted. Upregulation of negative regulators of co-stimulation has been suggested to be a mechanism to limit inflammation-related damage to tissues in settings of chronic inflammation, while maintaining the ability to respond to pathogens³². In CVIDid, however, it is possible that this compensatory response is insufficient in severe states of immune dysregulation and that these chronically activated cells still contribute to immune dysregulation-related pathology.

In addition to this chronically activated T-cell state, we observed a decreased ability of Tregs to upregulate CTLA₄, while CTLA₄ expression was increased in the whole CD4+ population. Expression of CTLA₄ by Tregs is an important mechanism by which Tregs mediate their suppressive function³³. Clinical CTLA₄ haploinsufficiency often results in a CVIDid phenotype with hypogammaglobulinemia and autoimmune disease, and functional Treg dysfunction has been described¹⁹. In this study, the T-cell profile of the patient with CTLA₄ haploinsufficiency was often not very different from the other non-genetic CVIDid patients. Therefore, the expression of CTLA₄ in Tregs of non-genetic CVIDid patients may be relevant to the overall underlying pathophysiology of CVIDid and warrants further research. A recent study shows that abatacept, a CTLA₄ fusion protein, was safe and effective in the treatment of CVIDid with interstitial lung disease³⁴. As the population of CVIDid patients with low CTLA₄+CD27+Tregs represented a mix of patients with pulmonary inflammation but also other organ-specific autoimmunity (data not shown), abatacept may be efficacious in other CVIDid patients as well. In addition, longitudinal monitoring of CTLA₄ Treg expression in CVIDid may indicate whether it can be used as a biomarker for disease exacerbation and/or therapeutic response.

Despite these overall differences between CVIDid and CVIDio, the heterogeneity within the CVIDid group was substantial. Subgroup analyses of patients with organ-specific autoimmunity did not yield insightful patterns, except that GLILD patients were often more extreme in all observed differences (data not shown). In addition, disease severity did not always reflect immune activation. For example, the patient with CTLA₄ haploinsufficiency did not show an inflammatory state in peripheral blood, while the patient that had the strongest IFN γ signature combined with low immune regulation markers (Supplementary Figure 11) was clinically stable and did not require immunosuppressive therapy. It is possible that the peripheral blood immune phenotype only gives limited

information that is clinically relevant, as T-cells may have migrated to the inflamed tissues in the sickest patients. In addition, this study may be limited by sample size to detect differences between subgroups of CVIDid. One other aspect that this study does not address is differences in absolute T-cell numbers. A recent study showed that while absolute T-cells are often lower in CVID, they do not differ between CVIDio and CVIDid¹³.

To conclude, these results indicate that CVIDid Th cells are highly activated, and that, unlike in classically exhausted cells such as originally described in chronic viral infection and cancer³², PD1 and LAG3 expression in CVIDid CD4+ T-cells reflects chronic activation with preserved inflammatory potential rather than functional exhaustion, similar to previous findings in human auto-immune inflammation²⁴. Combined studies of T-cell dysfunction and circulating inflammatory proteins in peripheral blood may help predict response to T-cell targeted therapies in individual CVIDid patients. Moreover, loss of CTLA4 upregulation in activated CVIDid Tregs may be mechanistically important in maintaining the inflammatory loop in CVIDid, and warrants further research.

Declarations

Funding: This study was financially supported by the Wilhelmina Children's Hospital Fund (Utrecht, the Netherlands).

Conflicts of interest: PH reports research grants and personal fees from Shire/Takeda and CSL Behring. VD reports research grants and personal fees from Shire/Takeda, Griffols, Actelion, Novartis, Pharming and CSL Behring. JM reports personal fees from Shire/Takeda. HL reports research grants from Shire/Takeda. All other authors report no potential conflict of interest.

Availability of data and material: not applicable.

Author's contributions: Study was conceived by HL, FvW and RB. Sample collection was performed by RB, JM, PH, VD and PE. Wet lab work was performed by RB and MvdW. Data analysis and interpretation: HL, FvW and RB. All authors substantially contributed to the acquisition, analysis or interpretation of the data, and all approved the final manuscript.

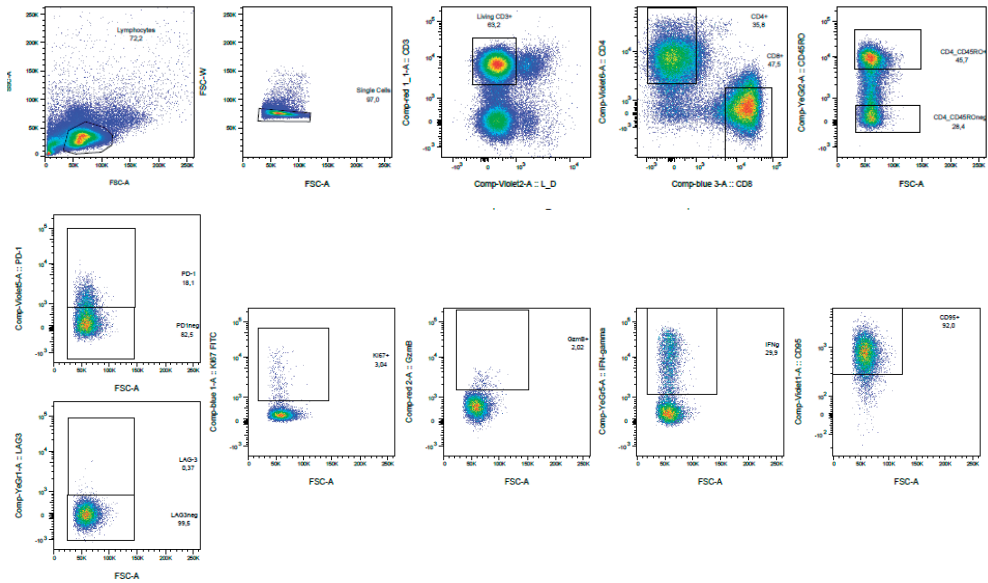
Ethics approval: Ethical approval for this study for all participants was received from the Medical Ethical Committee of the Erasmus MC University Medical Center Rotterdam, the Netherlands (METC 2013-026). Written informed consent was obtained from all patients and controls according to the Declaration of Helsinki.

REFERENCES

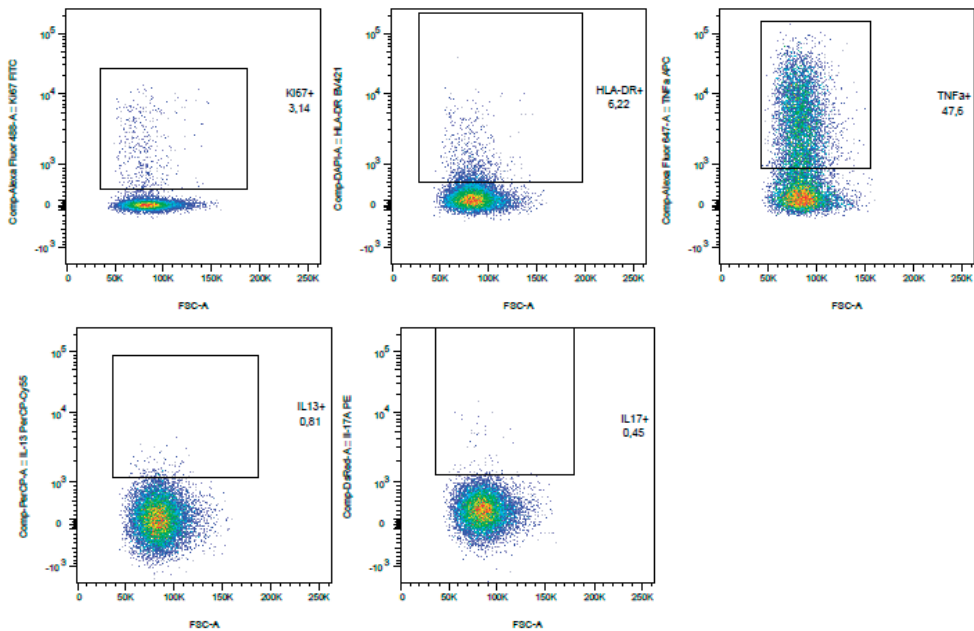
1. Immunodeficiencies ES for. Diagnostic Criteria PID . 2019. Available from: <https://esid.org/Education/Diagnostic-Criteria-PID>
2. Cunningham-Rundles C. How I treat common variable immune deficiency. *Blood*. 2010;116:7–15.
3. Gathmann B, Mahlaoui N, Gérard L, Oksenhendler E, Warnatz K, Schulze I, et al. Clinical picture and treatment of 2212 patients with common variable immunodeficiency. *J Allergy Clin Immunol*. 2014;134: 116–126.
4. Chapel H, Lucas M, Lee M, Bjorkander J, Webster D, Grimbacher B, et al. Common variable immunodeficiency disorders: division into distinct clinical phenotypes. *Blood*. 2008;112:277–87.
5. Resnick ES, Moshier EL, Godbold JH, Cunningham-Rundles C. Morbidity and mortality in common variable immune deficiency over 4 decades. *Blood*. 2012;119:1650–8.
6. Odnoletkova I, Kindle G, Quinti I, Grimbacher B, Knerr V. The burden of common variable immunodeficiency disorders : a retrospective analysis of the European Society for Immunodeficiency (ESID) registry data. *Orphanet J Rare Dis*. *Orphanet Journal of Rare Diseases*; 2018;0:1–17.
7. Maarschalk-Ellebroek LJ, Hoepelman AIM, Van Montfrans JM, Ellebroek PM. The spectrum of disease manifestations in patients with common variable immunodeficiency disorders and partial antibody deficiency in a university hospital. *J Clin Immunol*. 2012;32:907–21.
8. Hurst JR, Verma N, Lowe D, Baxendale HE, Jolles S, Kelleher P, et al. British Lung Foundation/ United Kingdom Primary Immunodeficiency Network Consensus Statement on the Definition, Diagnosis, and Management of Granulomatous-Lymphocytic Interstitial Lung Disease in Common Variable Immunodeficiency Disorders. *J Allergy Clin Immunol Pract* . Elsevier Inc; 2017;5:938–45. Available from: <http://dx.doi.org/10.1016/j.jaip.2017.01.021>
9. Gereige JD, Maglione PJ. Current Understanding and Recent Developments in Common Variable Immunodeficiency Associated Autoimmunity. *Front Immunol*. 2019;10:0–1.
10. Rao N, MacKinnon AC, Routes JM. Granulomatous and lymphocytic interstitial lung disease: A spectrum of pulmonary histopathologic lesions in common variable immunodeficiency - Histologic and immunohistochemical analyses of 16 cases. *Hum Pathol* . Elsevier Inc.; 2015;46:1306–14. Available from: <http://dx.doi.org/10.1016/j.humpath.2015.05.011>
11. Chase NM, Verbsky JW, Hintermeyer MK, Waukau JK, Tomita-Mitchell A, Casper JT, et al. Use of combination chemotherapy for treatment of granulomatous and lymphocytic interstitial lung disease (GLILD) in patients with common variable immunodeficiency (CVID). *J Clin Immunol*. 2013;33:30–9.
12. Giovannetti A, Pierdominici M, Mazzetta F, Marziali M, Renzi C, Mileo AM, et al. Unravelling the complexity of T cell abnormalities in common variable immunodeficiency. *J Immunol* . 2007;178:3932–43. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17339494>
13. Edwards ESJ, Bosco JJ, Aui PM, Stirling RG, Cameron PU, Chatelier J, et al. Predominantly Antibody-Deficient Patients With Non-infectious Complications Have Reduced Naive B, Treg, Th17, and Tfh17 Cells. *Front Immunol*. 2019;10:2593.
14. Unger S, Seidl M, van Schouwenburg P, Rakhmanov M, Bulashevska A, Frede N, et al. The TH1 phenotype of follicular helper T cells indicates an IFN- γ -associated immune dysregulation in patients with CD21low common variable immunodeficiency. *J Allergy Clin Immunol* . Elsevier Ltd; 2018;141:730–40. Available from: <http://dx.doi.org/10.1016/j.jaci.2017.04.041>
15. Perreau M, Vigano S, Bellanger F, Pellaton C, Buss G, Comte D, et al. Exhaustion of bacteria-specific CD4 T cells and microbial translocation in common variable immunodeficiency disorders. *J Exp Med* . 2014;211:2033–45.
16. Le Coz C, Bengsch B, Khanna C, Trofa M, Ohtani T, Nolan BE, et al. Common variable immunodeficiency-associated endotoxemia promotes early commitment to the T follicular lineage. *J Allergy Clin Immunol* . Elsevier Inc.; 2019;144:1660–73.
17. Berbers RM, Drylewicz J, Ellebroek PM, van Montfrans JM, Dalm VASH, van Hagen PM, et al. Targeted Proteomics Reveals Inflammatory Pathways that Classify Immune Dysregulation in Common Variable Immunodeficiency. *J Clin Immunol*. *Journal of Clinical Immunology*; 2020;41:362–373.

18. Hultberg J, Ernerudh J, Larsson M, Nilsson-Augustinsson Å, Nyström S. Plasma protein profiling reflects TH1-driven immune dysregulation in common variable immunodeficiency. *J Allergy Clin Immunol*. 2020;146(2):417-428.
19. Kuehn HS, Ouyang W, Lo B, Deenick EK, Niemela JE, Avery DT, et al. Immune dysregulation in human subjects with heterozygous germline mutations in CTLA4. *Science* (80-). 2014;345:1623-7.
20. Grimbacher B, Hutloff A, Schlesier M, Glocker E, Warnatz K, Dräger R, et al. Homozygous loss of ICOS is associated with adult-onset common variable immunodeficiency. *Nat Immunol*. 2003;4:261-8.
21. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2019. p. URL <https://www.R-project.org/>.
22. Arumugakani G, Wood PMD, Carter CRD. Frequency of treg cells is reduced in CVID patients with autoimmunity and splenomegaly and is associated with expanded CD210 B lymphocytes. *J Clin Immunol*. 2010;30:292-300.
23. Tian Y, Babor M, Lane J, Schulten V, Patil VS, Seumois G, et al. Unique phenotypes and clonal expansions of human CD4 effector memory T cells re-expressing CD45RA. *Nat Commun*. 2017;8:1473.
24. Petrelli A, Mijneer G, Hoytema Van Konijnenburg DP, Van Der Wal MM, Giovannone B, Mocholi E, et al. PD-1+CD8+ T cells are clonally expanding effectors in human chronic inflammation. *J Clin Invest*. 2018;128:4669-81.
25. Fevang B, Yndestad A, Sandberg WJ, Holm AM, Müller F, Aukrust P, et al. Low numbers of regulatory T cells in common variable immunodeficiency: Association with chronic inflammation in vivo. *Clin Exp Immunol*. 2007;147:521-5.
26. Vocanson M, Rozieres A, Hennino A, Poyet G, Gaillard V, Renaudineau S, et al. Inducible costimulator (ICOS) is a marker for highly suppressive antigen-specific T cells sharing features of TH17/TH1 and regulatory T cells. *J Allergy Clin Immunol*. 2010;126(2):280-289.
27. Hornero RA, Georgiadis C, Hua P, Trzupke D, He L, Qasim W, et al. CD70 expression determines the therapeutic efficacy of expanded human regulatory T cells. *Commun Biol*. Springer US; Available from: <http://dx.doi.org/10.1038/s42003-020-1097-8>
28. Wing K, Onishi Y, Prieto-Martin P, Yamaguchi T, Miyara M, Fehervari Z, et al. CTLA-4 control over Foxp3+ regulatory T cell function. *Science* (80-). 2008;322:271-5.
29. Anderson AC, Joller N, Kuchroo VK. Lag-3, Tim-3, and TIGIT: Co-inhibitory Receptors with Specialized Functions in Immune Regulation. *Immunity*. Elsevier Inc.; 2016;44:989-1004.
30. Cols M, Rahman A, Maglione PJ, Garcia-Carmona Y, Simchoni N, Ko HBM, et al. Expansion of inflammatory innate lymphoid cells in patients with common variable immune deficiency. *J Allergy Clin Immunol*. Elsevier Ltd; 2016;137:1206-1215.e6. Available from: <http://dx.doi.org/10.1016/j.jaci.2015.09.013>
31. Stuchlý J, Kanderová V, Vlková M, Heřmanová I, Slámová L, Pelák O, et al. Common Variable Immunodeficiency patients with a phenotypic profile of immunosenescence present with thrombocytopenia. *Sci Rep*. 2017;7:1-13.
32. Speiser DE, Utzschneider DT, Oberle SG, Münz C, Romero P, Zehn D. T cell differentiation in chronic infection and cancer: Functional adaptation or exhaustion? *Nat Rev Immunol*. 2014;14:768-74.
33. Walker LSK. Treg and CTLA-4 : Two intertwining pathways to immune tolerance. *J Autoimmun*. Elsevier Ltd; 2013;45:49-57. Available from: <http://dx.doi.org/10.1016/j.jaut.2013.06.006>
34. Spee-mayer C Von, Echnernach C, Agarwal P, Gutenberger S. Abatacept Use Is Associated with Steroid Dose Reduction and Improvement in Fatigue and CD4-Dysregulation in CVID Patients with Interstitial Lung Disease. *J Allergy Clin Immunol Pract*. Elsevier Inc; 2020;9:760-770.e10. Available from: <https://doi.org/10.1016/j.jaip.2020.10.028>

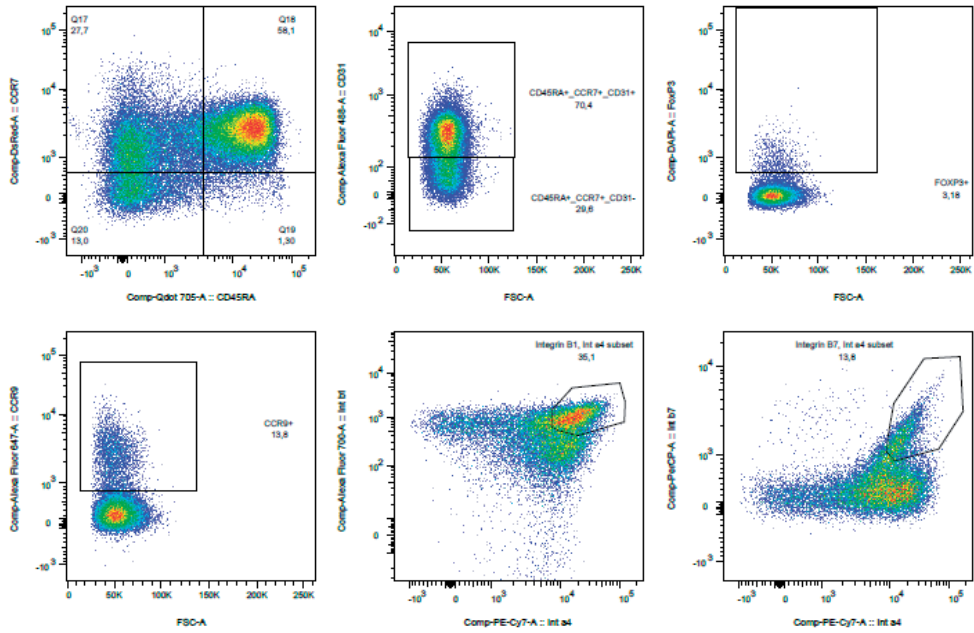
SUPPLEMENTARY MATERIALS



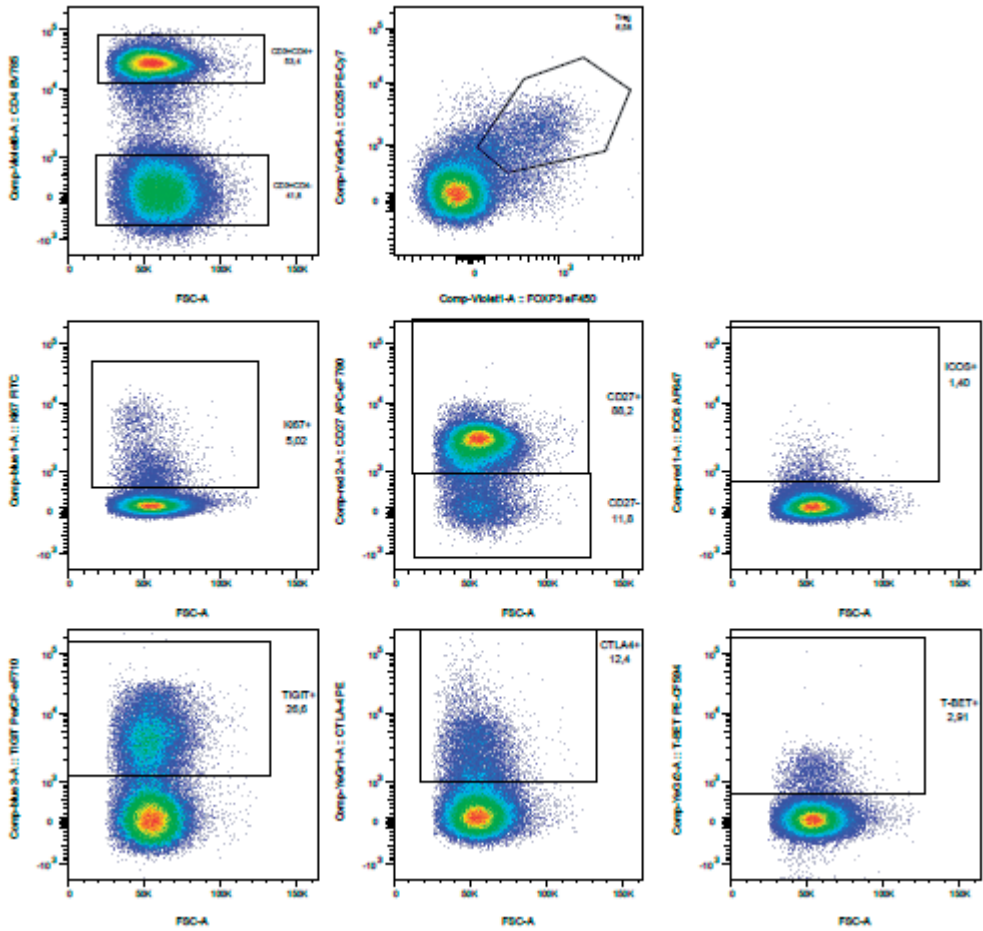
Supplementary Figure 1: Representative gating: general gating strategy and exhaustion panel



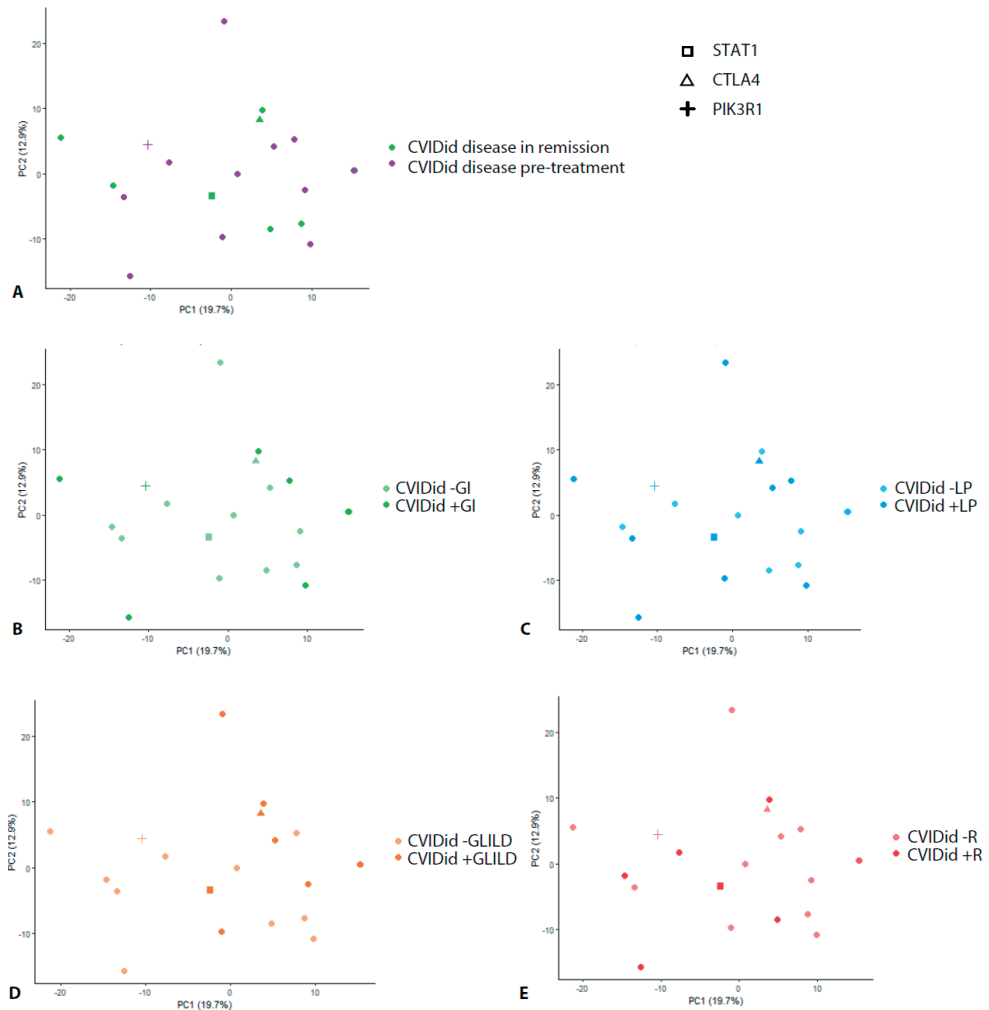
Supplementary Figure 2: Representative gating: gating strategy Th skewing panel



Supplementary Figure 3: Representative gating naïve / homing panel



Supplementary Figure 4: Representative gating Treg panel



Supplementary Figure 5: Principal component analysis of FACS data (all panels combined). CVID with immune dysregulation (CVIDid, n=20) only.

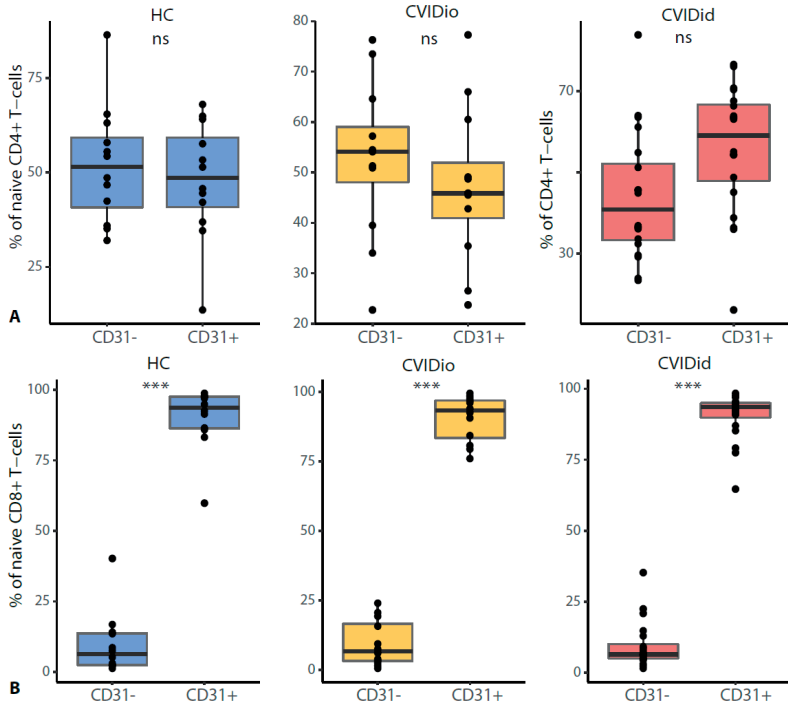
A: disease status: disease in remission (n=8) versus disease pre-treatment (n=12).

B: CVIDid patients with (+GI, n=8) and without (-GI, n=12) gastrointestinal complications.

C: CVIDid patients with (+LP, n=10) and without (-LP, n=10) lymphoproliferation.

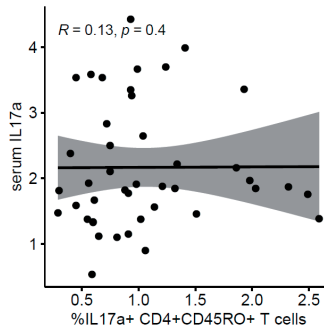
D: CVIDid patients with (+GLILD, n=8) and without (-GLILD, n=12) granulomatous-lymphocytic interstitial lung disease

E: CVIDid patients with (+R, n=6) and without (-R, n=14) rheumatological complications.

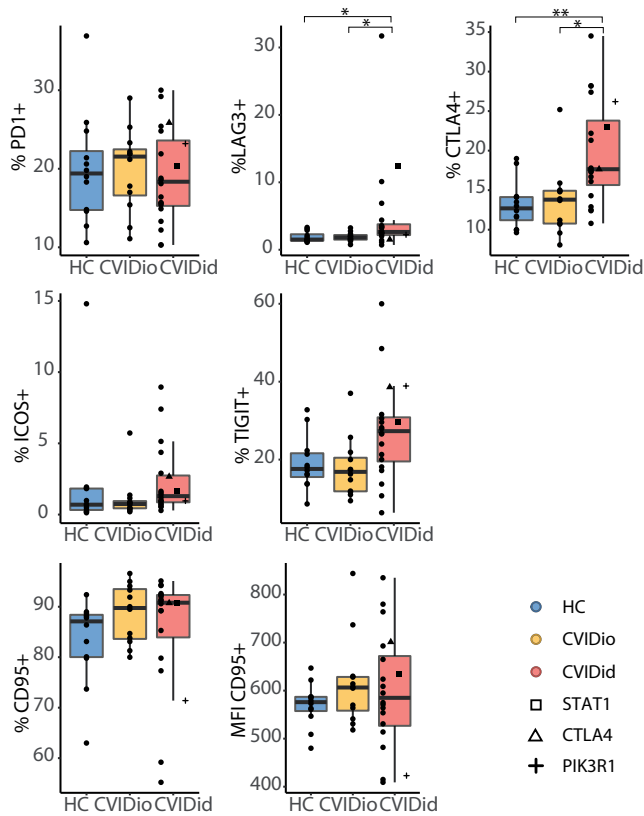


Supplementary Figure 6: CD31 expression within naïve (CD45RA+CCR7+) T-cells. CVIDid = CVID with immune dysregulation (n=20), CVIDio = CVID with infections only (n=12), HC = healthy controls (n=12). Statistics: Mann-Whitney U-test. * p<0.05, ** p<0.01, *** p<0.001.

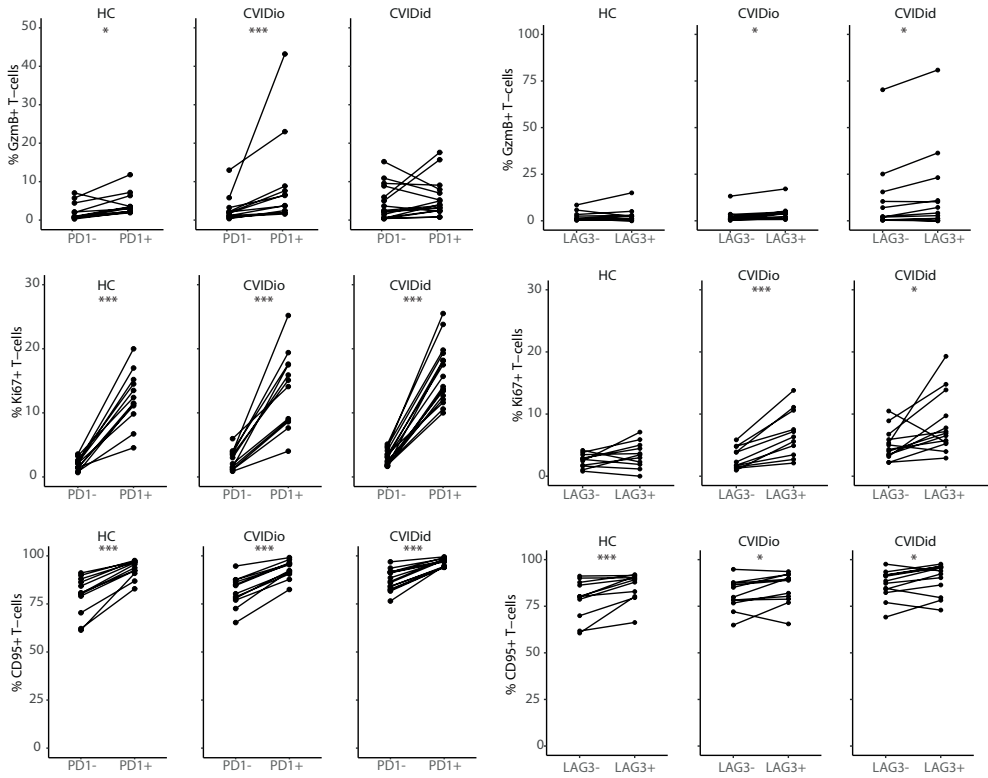
A: naïve CD4+ T-cells
B: naïve CD8+ T-cells



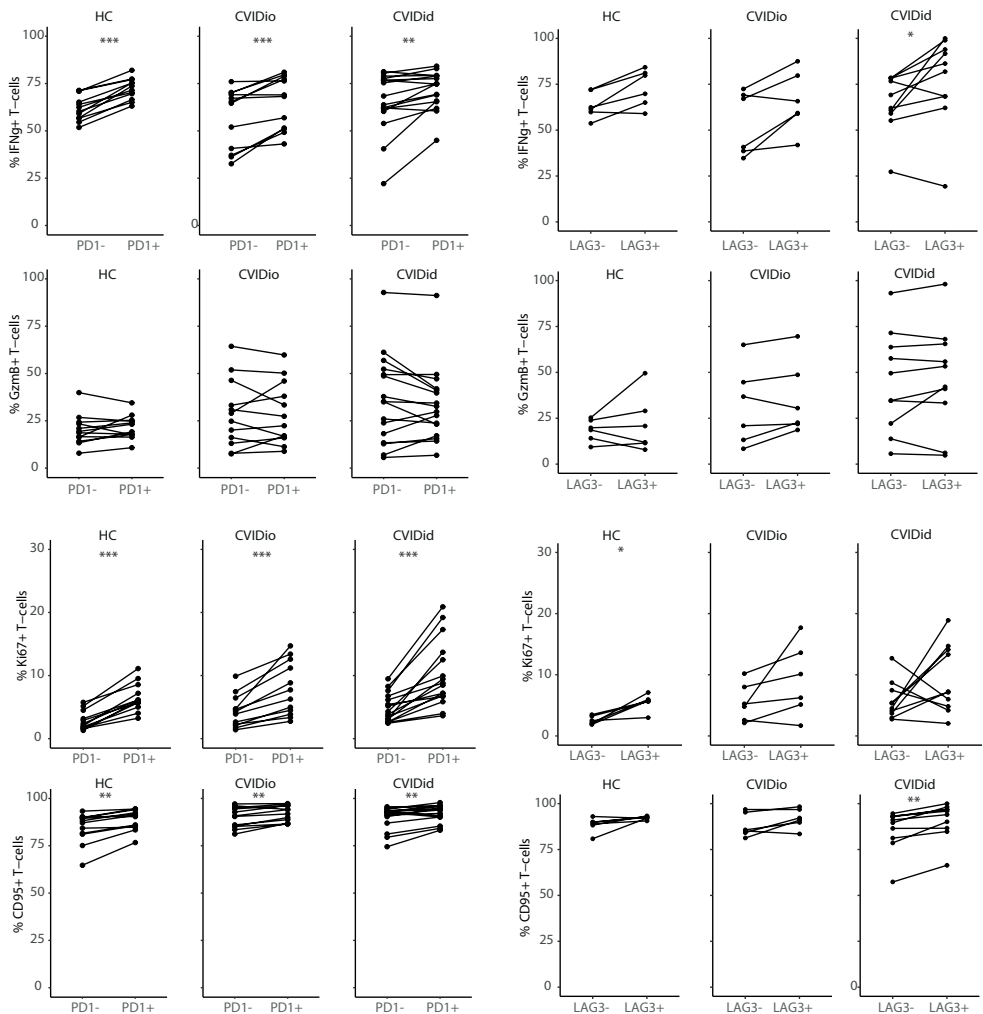
Supplementary Figure 7: Percentage of IL17+ CD4+CD45RO+ T-cells did not correlate with serum levels of IL-17A. Statistics: Spearman correlation.



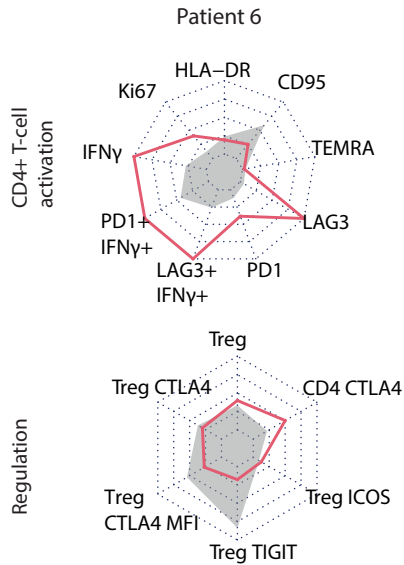
Supplementary Figure 8: Proportions of negative regulators of immune activation PD1, LAG3, CTLA4, ICOS and TIGIT, and apoptosis marker CD95 (FASL) in CD8+CD45RO+ T-cells. CVIDid = CVID with immune dysregulation (n=20), CVIDio = CVID with infections only (n=12), HC = healthy controls (n=12). MFI: median fluorescence intensity. Statistics: Mann-Whitney U-test. * p<0.05, ** p<0.01, *** p<0.001.



Supplementary Figure 9: Comparison of percentage of GzmB+ and CD95+ cells in PD1 and LAG3 positive and negative populations in CD4+CD45RO+ T-cells. Only samples with >50 events in the PD1/LAG3 positive and PD1/LAG3 negative populations were included. CVIDid = CVID with immune dysregulation (n=20), CVIDio = CVID with infections only (n=12), HC = healthy controls (n=12). Statistics: paired Wilcoxon-Rank test. * p<0.05, ** p<0.01, *** p<0.001.



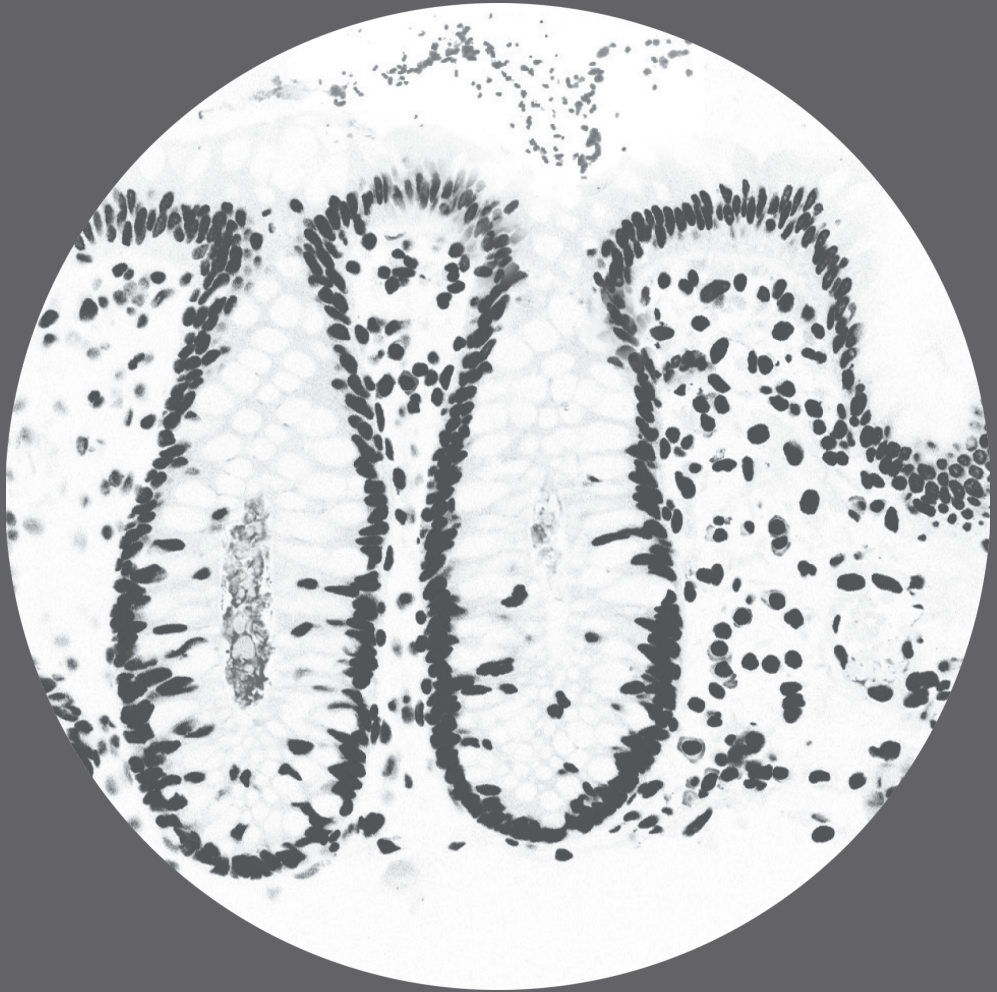
Supplementary Figure 10: Comparison of percentage of IFN γ +, GzmB+, Ki67+ and CD95+ cells in PD1 and LAG3 positive and negative populations in CD8+CD45RO+ T-cells. Only samples with >50 events in the PD1/LAG3 positive and PD1/LAG3 negative populations were included. CVIDid = CVID with immune dysregulation (n=20), CVIDio = CVID with infections only (n=12), HC = healthy controls (n=12). Statistics: paired Wilcoxon-Rank test. * p<0.05, ** p<0.01, *** p<0.001.



Supplementary Figure 11: Patterns of CD4+ T-cell activation and T-cell regulation in a patient with high immune activation markers but mild clinical phenotype. Coeliac disease, Sjögren-like disease, autoimmune gastritis, not requiring immunosuppressive therapy. Genetics were done but no relevant mutations were found. MFI: median fluorescence intensity.

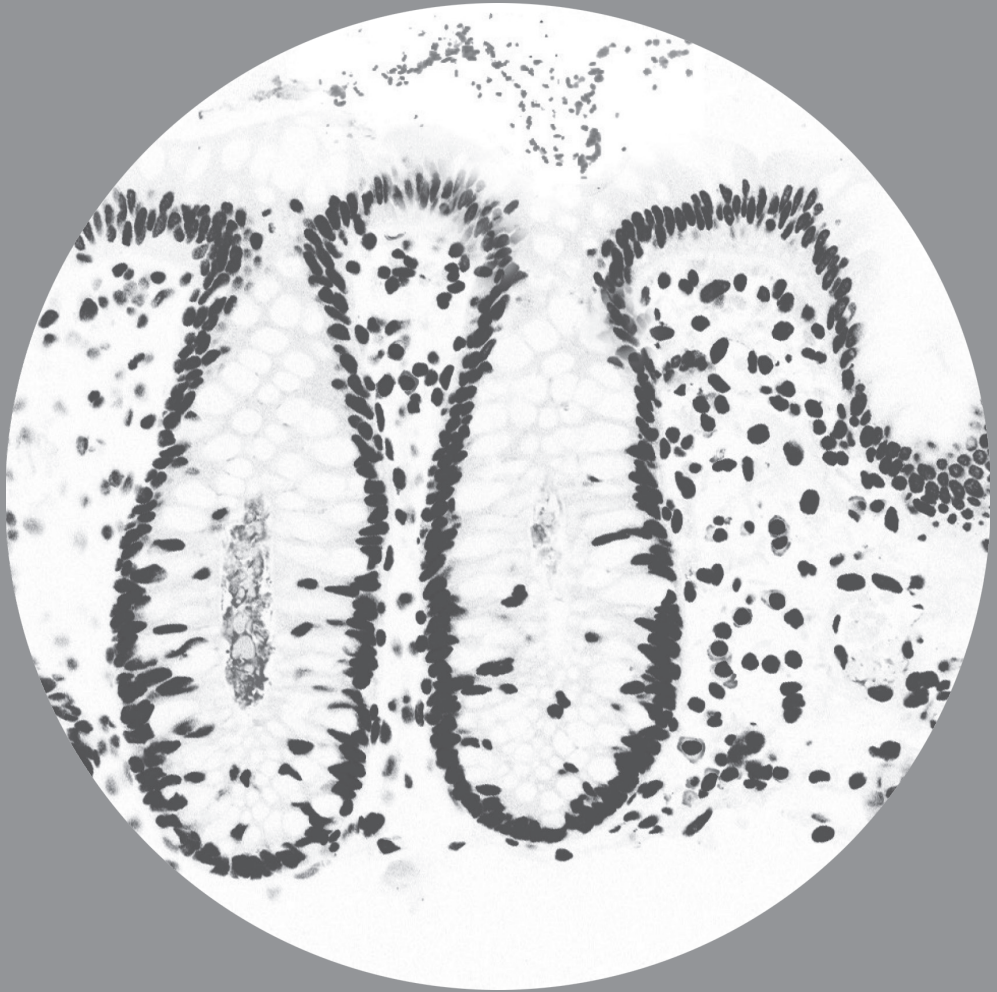
Supplementary Table 1: FACS staining antibody list

Antibody	Fluoro- chrome	Company	Panel
<i>Surface</i>			
CCR7	PE	eBioscience	Homing/naive
CCR9	APC	Biolegend	Homing/naive
CD137 (4-1BB)	APC	BD	Exhaustion
CD25	PE-Cy7	BD	Treg
CD27	APC-eF780	eBioscience	Treg
CD3	AF700	Biolegend	Exhaustion, Treg
CD3	BV605	Biolegend	Th skewing, Homing/Naive
CD31	FITC	BD	Homing/naive
CD4	BV785	Biolegend	Exhaustion, Th skewing, Homing/Naive Treg
CD45RA	BV711	Biolegend	Homing/naive
CD45RO	ECD	Beckman Coulter	Exhaustion, Th skewing
CD8	APC-Cy7	BD	Th skewing, Homing/Naive
CD8a	PerCP-Cy5.5	Biolegend	Exhaustion
CD95 (FAS ligand)	eFluor450	eBioscience	Exhaustion
HLA-DR	BV421	Biolegend	Th skewing
ICOS	AF647	eBioscience	Treg
Int α 4	PE-Cy7	Biolegend	Homing/naive
Int β 1	AF700	Biolegend	Homing/naive
Int β 7	PerCP-Cy5.5	Biolegend	Homing/naive
LAG-3	PE	R&D Systems	Exhaustion
PD-1	BV711	BD	Exhaustion
TIGIT	PerCP-eF710	eBioscience	Treg
TNF	APC	Biolegend	Th skewing
<i>Intracellular</i>			
CTLA-4	PE	BD	Treg
FOXP3	eF450	eBioscience	Homing/naive, Treg
GzmB	APC-Fire750	Biolegend	Exhaustion
IFNy	PE-Cy7	BD	Exhaustion
IL-13	PerCP-Cy5.5	Sony Biotechnology	Th skewing
IL-17A	PE	eBioscience	Th skewing
Ki67	FITC	Dako (Agilent)	Exhaustion, Th skewing, Treg
T-BET	PE-CF594	BD	Treg



Part II

The Microbiome in CVID



Chapter 4

Microbial Dysbiosis in Common Variable Immune Deficiencies: Evidence, Causes and Consequences

Roos-Marijn Berbers
Stefan Nierkens
Jacob M. van Laar
Debby Bogaert
Helen L. Leavis

Trends in Immunology (2017):
38 (3) 206-216

ABSTRACT

Common variable immunodeficiency (CVID) is an immune disorder that not only causes increased susceptibility to infection, but also to inflammatory complications such as autoimmunity, lymphoid proliferation, malignancy and granulomatous disease. Recent findings implicate the microbiome as a driver of this systemic immune dysregulation. Here, we critically review the current evidence for a role of the microbiome in the pathogenesis of CVID immune dysregulation, and describe the possible immunologic mechanisms behind causes and consequences of microbial dysbiosis in CVID. We will integrate this evidence into a model describing a role for the gut microbiota in the maintenance of inflammation and immune dysregulation in CVID, and suggest research strategies to contribute to the development of new diagnostic tools and therapeutic targets.

COMMON VARIABLE IMMUNODEFICIENCY AND THE MICROBIOME

Common variable immunodeficiency (CVID) is the most common symptomatic primary immunodeficiency and classically manifests either in childhood or adulthood with recurrent respiratory tract infections that often result in structural pulmonary damage¹. The diagnostic criteria (see box 1) for this clinically heterogeneous primary immune deficiency include a marked decrease in multiple classes of immunoglobulins, as well as impaired vaccination responses². Alternative diagnostic criteria are also in use³, reflecting the heterogeneity of the disease and complicating the comparability of research. Similar to autoimmune and autoinflammatory diseases, genetic susceptibility loci have been identified (in less than 10% of cases⁴). Although these cases provide genetic insights into disease pathogenesis, the role of environmental factors is possibly critical and still largely unclear⁵.

In addition to immunodeficiency, more than half of all CVID patients develop non-infectious complications such as lymphoproliferative disease, granulomatous disease, malignancies, or autoimmunity (see box 2)⁶, causing significant morbidity and mortality⁷⁻⁸. These complications suggest an underlying state of immune dysregulation that may result from chronic systemic immune activation, which has been suggested to be a consequence of increased microbial translocation in CVID⁹⁻¹².

In the gut, the mucus layer reduces contact between microorganisms, epithelial cells and the immune cells that patrol the gastrointestinal system¹³. The relative compartmentalization allows the immune system to interact with the gut microbiota through regular sampling¹³. This communication between the two compartments occurs bi-directionally: the immune system influences the composition of the microbiota¹⁴, and the gut microbiota directs immune maturation and is necessary for immune tolerance and homeostasis¹⁵.

In case of dysbalance in immune function, the resulting impaired surveillance of the gut could result in increased microbial translocation across the gut barrier. In turn, microbial translocation can cause dysregulation of the systemic immune response, leading to a vicious cycle in which the immune system and the microbiota move further away from homeostasis. In addition, (local) inflammation may further exacerbate gut leakiness¹⁶, causing more translocation and resulting inflammation. The process by which impaired immunosurveillance can lead to systemic immune activation through microbial translocation was first shown in patients with chronic human immunodeficiency virus (HIV) infections¹⁷, but has been suspected to play a role in other immunodeficiency syndromes as well, including CVID⁹⁻¹¹. However, shifts in human microbiota composition mediated by immune deficiencies such as CVID, and the possible role for microbiome in the pathogenesis of this disease are still poorly understood processes. In this review, we discuss by which mechanisms CVID may lead to microbial dysbiosis, and how this can result in immune dysregulation. In this model, an individual prone to develop CVID can progress to disease when impaired immunosurveillance of the gut leads to an unbalanced gut microbiome and/or increased bacterial translocation (see figure 1). This can drive systemic immune activation, leading to immune dysregulation and the associated clinical complications.

Box 1: CVID diagnostic criteria: Diagnostic criteria of CVID as defined by the European Society for Immunodeficiencies (ESID) and the Pan-American Group for Immunodeficiency (PAGID) ²:

Onset of immunodeficiency at greater than 2 years of age
 Absent isohemagglutinins and/or poor response to vaccines
 Defined causes of hypogammaglobulinemia have been excluded

and

Marked decrease of IgG AND IgM or IgA: probable CVID

Marked decrease of IgG OR IgM OR IgA: possible CVID

These criteria are the most widely accepted and are used in the studies discussed in this article. However, there is some discussion in the field as to whether these criteria are up-to-date. The ESID criteria have been revised in 2014 to reflect the immune dysregulation phenotype as part of CVID³.

Alternative criteria have been proposed by Ameratunga et al.³, which are more extensive, providing additional diagnostic tests that can be used to support the diagnosis of CVID. These include reduced memory B cell subsets and/or increased CD21low subsets, IgG3 deficiency, serological evidence of significant autoimmunity (e.g. Coombs test) and sequence variations of genes predisposing to CVID³.

Box 2: CVID immune dysregulation phenotype (prevalence) ^{6,7,79-81}

CVID with one of the following (total: 51-72% of patients with CVID):

Autoimmune disease (23-48%)

- autoimmune hemolytic anemia, autoimmune thrombocytopenia, vitiligo, pernicious anemia, systemic lupus erythematosus, rheumatoid arthritis, antiphospholipid syndrome, anti IgA antibody disease, juvenile idiopathic arthritis, Sjögren's disease, psoriasis, thyroiditis, uveitis, vasculitis

Lymphoproliferative (<20-28%) and granulomatous disease (8-33%)

- systemic: polyclonal lymphoid infiltration, granulomatous inflammation
- lung (granulomatous lymphocytic interstitial lung disease)
- splenohepatomegaly
- gastrointestinal tract (lymphocytic enteropathy)

Malignancy (7%)

- lymphoid malignancy
- Other

EVIDENCE FOR A POTENTIAL ROLE OF THE HUMAN MICROBIOME IN PATHOGENESIS OF CVID

The human microbiome interacts with the systemic immune system broadly in two ways ¹³; 1. the microbial community is regularly sampled by immune cells; 2. bacteria cross the gut epithelium and can expose the systemic immune system to microbial components

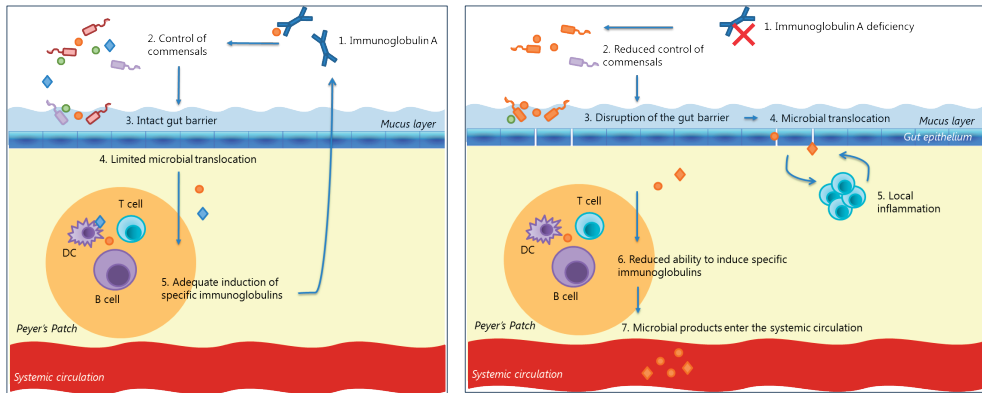


Figure 1: control of the microbiota in health and CVID

Normal control of the microbiota (above):

(1) Presence of IgA in the gut lumen is important in maintaining a diverse and balanced microbiome.^{39–41} (2) IgA promotes expulsion of pathogenic bacteria³⁸ and confers immunological tolerance to commensals^{45,46}. (3) IgA helps keep the gut barrier intact by protecting against infections. (4) In the absence of inflammation and the presence of an intact gut barrier, microbial translocation is limited but still occurs^{16,20} (5) IgA against commensals is induced in Peyer's patches, where live bacteria transported in dendritic cells are presented to T and B cells⁴².

Reduced control of the microbiome in CVID (below): (1) Patients with CVID often have low or absent IgA. (2) In the absence of IgA, the composition of the microbiome may change, possibly allowing more invasive and pro-inflammatory species to proliferate, leaving less room for beneficial anti-inflammatory commensals^{38–41}. (3) IgA deficiency is associated with higher gastrointestinal tract infection-load³⁹, and this may lead to disruption of the gut barrier, allowing more bacteria and bacterial products to cross. (4) Crossing of microbes and their products is called microbial translocation¹⁷. (5) Impaired clearance of microbial products may lead to local inflammation¹⁷. Moreover, IgA can dampen the inflammatory response to beneficial commensals^{45,46}. Without this mechanism, the increased local inflammation may lead to further disruption of the gut barrier, making it more leaky. (6) IgA is important for the sampling of bacteria and their transport to Peyer's patches, where normally specific IgA would be induced⁴². In the absence of IgA, and given the observed immunological impairments in CVID such as B cell dysfunction but also DC hypo-responsiveness and T cell exhaustion, the generation of specific immunoglobulin is reduced. (7) Finally, microbial products may reach the circulation, where they may contribute to immune dysregulation.

(microbial translocation). Bacterial products such as lipopolysaccharide (LPS), peptidoglycan, flagellin, RNA and DNA can activate the immune response through recognition of microbe-associated molecular patterns by the innate immune system^{17,18}. Although immune sampling and microbial translocation co-occur in healthy individuals as well¹⁶, increased translocation as a consequence of diminished barrier function may result in chronic inflammation and immune dysregulation¹⁷. Also, microbial dysbiosis may lead

to overgrowth of pro-inflammatory bacteria or a decrease of anti-inflammatory bacteria which subsequently leads to further imbalance of the immune system¹⁹. We hypothesise that this may contribute to increased (mucosal) inflammation and immune dysregulation. There is evidence that both translocation and dysbiosis are likely to play a role in CVID, as discussed in the following paragraphs.

Microbial Translocation in CVID

Microbial translocation has been shown to regularly occur in healthy individuals¹⁶ and is known to increase in a broad range of diseases, including inflammatory diseases of the gut such as inflammatory bowel disease²⁰, but also chronic infections such as HIV²¹, and hepatitis B and C²². Presence of lipopolysaccharide in the systemic circulation can indicate increased microbial translocation of gram negative bacteria. LPS is therefore commonly used as a marker for endotoxaemia in these studies. Other markers often used in the context of microbial translocation are soluble CD14 and soluble CD25. sCD14 is a marker expressed by monocytes that can directly bind LPS in the presence of LPS-binding protein²¹. sIL2R is a generic marker of T cell activation, and was shown to correlate closely with microbiome dysbiosis¹². Both sCD14 and sCD25 are therefore used as markers of immune activation that have been closely linked to microbial translocation.

When markers for endotoxaemia are compared between CVID patients with and without complications from immune dysregulation, higher levels of LPS and soluble IL-2 receptor (sIL2R) are associated with the former group¹². When looking at CVID as a whole, the largest study conducted (n=104) reports increased levels of LPS compared to healthy controls¹². This result is reproduced in a study that investigates endotoxaemia in untreated CVID patients with serum IgG levels below 4.9 (n=8), while two other studies (n=35 and n=31, respectively) find normal levels of LPS with high levels of sIL2R and sCD14^{9,23}. Notably, the large study reporting increased endotoxaemia consisted for 80% of patients with immune dysregulation symptoms and shows significantly higher LPS levels in this group than in the patients without immune dysregulation symptoms. Normal levels of LPS were found in one study with 55% immune dysregulation, unfortunately the other smaller study did not fully report immune dysregulation symptoms. Other factors that may explain differences in measured levels of LPS are the assays themselves, which are known to be highly variable, making it difficult to compare between different experimental methods (e.g. serum vs plasma, different dilutions). Additional larger studies of treated CVID patients will be needed to assess the true extent of endotoxaemia and whether it is indeed associated with immune dysregulation symptoms. Moreover, since most studies do not investigate markers for gram positive or anaerobic bacteria, their role has been overlooked so far. Components of gram positive bacteria such as lipoteichoic acid and peptidoglycan are known to contribute significantly to immune activation in sepsis²⁴ and, as important activators of the innate immune system²⁵, are worth investigating in CVID.

Microbiome Composition in CVID

The composition of the microbial community in the gut has in recent years become increasingly implicated in the pathogenesis of a variety of immune-mediated diseases. For example, next generation 16S-based sequencing of fecal samples from patients with rheumatoid arthritis revealed that *Prevotella copri* is much more prevalent in patients with new-onset untreated disease than in healthy controls, and this may reflect true causality, since *P.copri* caused increased inflammation in a colitis mouse model²⁶. Using an alternative approach, Palm et al. showed that fecal commensal bacteria eliciting a high IgA gut response are colitogenic in inflammatory bowel disease, which was confirmed when germ-free mice exposed to these bacteria developed colitis more frequently than unexposed mice²⁷.

In addition to identifying specific groups of commensal bacteria that can be proinflammatory, the diversity and stability of the microbiome may play a role in the pathogenesis of a disease. Microbial diversity within a community can be described by alpha diversity, a measure for the amount of different bacteria present in a single sample that takes into account the number of observed operational taxonomic units (OTUs) (richness) and the relative abundances of OTUs (evenness). Microbial diversity between two communities is described by beta diversity. A highly diverse microbiome is generally associated with health, inducing resilience, stability and prevention of overgrowth and invasion by proinflammatory species²⁸. In HIV, microbial dysbiosis has been linked to increased markers of microbial translocation²⁹.

To our knowledge, there is one published study that has investigated the composition of the gut microbiome in CVID, and it was also the first analysis of the gut microbiome of a primary immunodeficiency. Jørgensen et al.¹² compared 44 CVID patients with 45 patients with inflammatory bowel disease and 263 healthy controls. Their findings include reduced alpha diversity in CVID patients, which correlated with increased levels of serum endotoxemia and reduced levels of IgA. This effect could solely be attributed to those patients within the CVID cohort that had complications from immune dysregulation, as LPS levels and alpha diversity were comparable between healthy controls and CVID patients without immune dysregulation. Interestingly, alpha diversity in CVID patients did not differ from that in IBD patients, but there were differences in taxonomic profile between the two diseases. Also, within IBD patients, alpha diversity did not correlate with endotoxemia markers, as was the case in CVID patients. This may suggest that specific bacteria lead to increased microbial translocation in CVID rather than broad differences in diversity. Antibiotic use within the last year was also investigated, but no obvious associations with alpha diversity were found in CVID patients. The specific effect of long-term antibiotic use on the diversity of the human microbiome is still not sufficiently understood, especially not in relation to immune dysregulation. Some studies do describe long-term effects of antibiotic use on the microbiome in immune competent individuals³⁰⁻³² and antibiotic use has to be considered as a potential confounder when studying the microbiome in immunodeficiency.

This evidence suggests a link between the microbiome and immune dysregulation, in particular given the differences between CVID patients with immune dysregulation and CVID patients without immune dysregulation. Additional studies into CVID and the microbiome are ongoing to further investigate this link. Furthermore, there is clear evidence that increased microbial translocation occurs in CVID. Nevertheless, attributing causality to microbial changes remains a challenge. Does CVID or CVID immune dysregulation lead to differences in the microbiome, does the microbiome cause the disease itself, or are any changes in the microbiome simply a bystander effect of another process that leads to immune dysregulation in CVID? Definitive answers to this question will need to come from additional longitudinal studies complemented with mechanistic studies, but indirect evidence that may aid in this issue is discussed below.

IMMUNOLOGICAL CAUSES OF MICROBIOME DYSBIOSIS IN CVID

Immunoglobulin Deficiency

The main hallmark of CVID is hypo(gamma)globulinemia combined with impaired polysaccharide-specific vaccine responses. In a minority of cases, this can be explained by mutations in genes necessary for B cell function³³⁻³⁶. The remaining cases are thought to be caused by complex polygenetic inheritance and interaction with environmental factors³⁷ that eventually lead to a disability to produce immunoglobulin G, and IgA or IgM. IgA is especially of interest in this setting, as secretory IgA plays a key role in shaping and maintenance of the intestinal microbiota (reviewed by Pabst et al.³⁸). IgA deficiency in CVID is measured in serum, and serum levels of IgA are known to correlate with soluble IgA secreted in the lumen of the gut in antibody deficient patients³⁹.

IgA is secreted into the mucus layer of the gastrointestinal tract and limits growth and microbial invasion by blocking bacterial adhesion to the gut epithelial cells and promoting their expulsion from the lumen (reviewed by Mantis et al. ⁴⁰). Presence of IgA is a crucial determinant of the composition of the microbiome⁴¹, illustrated by the finding that mice lacking IgA have much lower microbial diversity in their gut than their wildtype littermates⁴², and that IgA deficiency in mice allows expansion of segmented filamentous bacteria resulting in systemic immune activation⁴³. IgA against commensals is induced in Peyer's patches, where live bacteria transported in dendritic cells are presented to T and B cells⁴⁴. In addition, IgA stimulates the uptake of IgA-bound antigens by M cells, transporting the complexes to Peyer's Patches, thereby forming a positive feedback loop⁴⁵. Without IgA, non-penetrative species do not reach the Peyer's patches, preventing the induction of specific IgA. It has been proposed that polyreactive IgA, produced by innate-like B₁ cells, play an important role in the initial targeting of the microbiome by facilitating the uptake of new bacteria and their presentation to the immune system⁴⁶. The direct effect of IgA in the gut lumen is multifactorial and not yet completely understood. Aside from promoting the clearance of inflammatory bacteria, IgA has also been shown to limit the pro-inflammatory response to the bacteria it coats. This is thought to occur for commensals^{47,48} as well for pathogenic bacteria such as *Shigella* spp. ⁴⁹.

The effects of IgA deficiency on the CVID microbiome can be threefold: Firstly, low IgA is associated with increased susceptibility to gastrointestinal tract viral infection and attendant mucosal inflammation³⁹. High (gastrointestinal-) infection load can directly perturb the normal gut microbiota⁵⁰, and repeated infections can lead to damage to the intestinal epithelium and thereby to increased microbial translocation. Secondly, poor control of the gut microbiome and increased expansion of specific groups of bacteria can lead to systemic immune activation, as described in mice⁴³. Thirdly, due to the anti-inflammatory effects of IgA, loss of IgA could lead to increased local inflammation in response to commensal species. Through these functions, IgA deficiency not only leads to alterations in microbiome composition and increased opportunities for microbial translocation, it can also directly lead to local mucosal inflammation.

Overall, IgA deficiency may be an important factor in the systemic inflammation that leads to immune dysregulation in CVID. In CVID, low IgA is one of the clinical parameters most strongly associated with increased morbidity from inflammation⁵¹. Immunoglobulin substitution therapy restores IgG levels but not IgA or IgM, largely resolving the specific immune deficiency but generally not improving immune dysregulation symptoms⁶. Isolated IgA deficiency is often asymptomatic, but in the context of other genetic factors and/or hypogammaglobulinaemia such as in CVID, it is likely to contribute to a severe imbalance in the microbiota as well as in the immune system. Although many patients with isolated IgA deficiency do not experience any symptoms, it is associated epidemiologically with an increased risk of autoimmunity, including Graves' disease, rheumatoid arthritis, systemic lupus erythematosus and type-1 diabetes (reviewed by Wang et al.⁵²). A possible explanation for these findings is that there is redundancy in the IgA system when it comes to control of the microbiota and concomitant inflammation, and that these cases of autoimmunity illustrate instances where additional genetic or environmental factors lead to loss of redundancy. IgM and IgG also play a role in mucosal immunity^{43,53}, and it is possible that their presence compensates for loss of IgA. In the case of CVID where IgG and sometimes also IgM production are impaired, this compensatory mechanism may be lost. IgG supplementation is not specifically tailored to an individual's microbiome composition, and it is currently unknown how much of supplemented IgG is secreted into the gut lumen.

Impaired Epithelial Gut Barrier

Evidence for physical damage of the gut epithelium in CVID is well documented, and this damage may be a direct result of the primary immune defect which permits mucus invasion and epithelial infection³⁷, as well as be a result of persistent systemic inflammation due to immune dysregulation. Damage to the epithelium can lead to disruption of the physical barrier between the immune system and the microbiota.

Involvement of the gut does not seem to be rare in CVID: many CVID patients suffer from abdominal complaints^{8,54}, and GI tract pathology is highly prevalent⁵⁵. Histological defects observed in CVID include celiac-like disease, reduction of plasma cells and nodular lymphoid hyperplasia⁵⁵. Interestingly, these histological defects are much more

common than abdominal complaints in CVID patients. Plasma cell reduction in the lamina propria, which is the most frequently found histological anomaly, is associated with nonspecific markers of immune activation or microbial translocation, sCD14 and sCD25⁵⁵. Therefore plasma cell aberrations in CVID may cause subclinical enteropathy that contributes to systemic inflammation. As plasma cells are producers of immunoglobulins, this association may well mirror the effects of IgA described above.

IMMUNOLOGICAL CONSEQUENCES OF MICROBIAL DYSBIOSIS IN CVID

Expansion of Pro-inflammatory Innate Immune Cells

CVID patients have an expansion of pro-inflammatory CD14^{bright}CD16⁺ monocytes⁵⁶ in their peripheral blood⁹ in a cohort that primarily consists of patients with immune dysregulation, lacking a non-dysregulatory group as a control. It is plausible that this monocyte expansion causes the high levels of sCD14 seen in microbial translocation, as the monocyte subset expansion correlated with sCD14⁹. In addition, activation of B and T cells correlated with this pro-inflammatory subset⁹, suggesting that microbial translocation may independently lead to monocyte and B and T cell activation, or that the monocytes themselves activate the adaptive compartment.

Moreover this pro-inflammatory monocyte subset has been shown to be expanded in other immune dysregulation syndromes, including sepsis⁵⁷, rheumatoid arthritis⁵⁸, and Sjögren's syndrome⁵⁹. In addition, CVID monocytes were shown to have an increased tendency to fuse, forming giant cells that may lead to granuloma formation⁶⁰. Monocyte activation, which can be associated with microbial translocation⁸, may therefore play an important role in the development of immune dysregulation, perhaps by driving a state of sustained immune activation.

Dendritic Cell Hyporesponsiveness

The total number of circulating CD11c+CD123⁻ myeloid dendritic cells (myDC) was found to be decreased in CVID patients compared to healthy controls in two studies^{61,62}. Poor differentiation from peripheral monocytes to dendritic cells was seen *in vitro*⁶², and the dendritic cells that did mature had reduced capacity to stimulate T helper cell differentiation due to poor CD40 expression and failure to express HLA-DR on their surface^{63,64}. This impaired ability to upregulate maturation markers was not related to any specific genetic defects, and suggests dendritic cell exhaustion or hyporesponsiveness. In support of this, I multiple studies report that CVID plasmacytoid dendritic cells produce less interferon (IFN)- α in response to Toll-like receptors (TLR) 7 and 9 ligation *in vitro*⁶⁵⁻⁶⁷, while others report no differences between CVID and healthy cells⁶¹. It is probable that these discrepant results reflect different states of dendritic cell exhaustion, rather than an intrinsic TLR signalling defect.

Correlation of myDCs and CVID with the immune dysregulation phenotype has been described indirectly. CVID patients with granulomatous disease have fewer circulating

dendritic cells than those without⁶². In addition, the two studies that reported significantly low dendritic cells in CVID had a high percentage of patients with immune dysregulation symptoms in their cohorts^{61,62}, and, in a study describing impaired dendritic cell maturation, all patients suffered from autoimmune disease, granulomatous disease and/or chronic diarrhoea⁶⁸. It is therefore possible that dendritic cell hyporesponsiveness is especially pronounced in patients with immune dysregulation. Dendritic cells play an important role in regulating the systemic immune response as well as the local mucosal response. For instance, dendritic cells are necessary to induce commensal-specific IgA⁴⁴, and reduced IgA responses may contribute to microbial dysbalance in CVID. This illustrates the circular process in CVID in which impaired immunity leads to increased antigenic exposure to the gut microbiome, causing DC hyporesponsiveness, which further impairs mucosal immunity by hampering IgA production.

T cell Activation and Exhaustion

Classical $\alpha\beta$ T helper cell dysregulation in CVID⁶⁹ includes low naïve T helper cell fractions with high T cell activation levels and a limited T-cell receptor (TCR) repertoire⁵¹. Despite increased overall T-cell activation levels, the production of cytokines associated with Th1, Th2 and Th17 cells (among which IFN- γ , IL-2, -9, -13, -17) is much lower than in healthy controls⁷⁰, and T cell-related transcripts are down-regulated on gene expression level⁷¹. A subgroup of patients with extremely low naïve T cell numbers, high T cell activation and TCR repertoire disruption show severe clinical immunodeficiency, generally associated with lymphoproliferative disease^{51,72}.

These defects may reflect a state of T cell exhaustion, as more CVID CD4+ T cells expressed the exhaustion marker PD-1 than healthy control CD4+ T cells^{11,73}. CVID CD4+ T cells had reduced proliferation capacity upon exposure to common bacterial and viral antigens and this effect was restored upon blocking of the PD-1 pathway *in vitro*. PD-1 expression on T cells correlated with serum LPS levels¹¹, suggesting that this may be related to microbial translocation.

Loss of Regulatory T cells

CVID patients have lower numbers of regulatory T (Treg) cells compared to healthy controls in their peripheral blood^{10,74,75}, with reduced suppressive capacity towards autologous cells⁷⁵. Interestingly, low frequencies of Treg cells were also found in a small cohort of patients with an isolated IgA deficiency, and were especially low in those patients who also suffered from autoimmunity⁷⁶. This may mean that the microbial changes that occur in CVID as a result of IgA deficiency lead to reduced Treg cells, or alternatively that IgA deficiency directly leads to reduced Treg cells and that the resulting inflammation causes microbiome alterations. Kawamoto et al. describe a regulatory feedback loop by which a healthy microbiota stimulates regulatory T cell expansion, and formation of germinal centers and IgA, which in turn maintains the diversity and stability of the microbiota⁴². Disruption of this feed-back loop in CVID through a combination of IgA deficiency and a disturbed microbiota can thereby have a profound effect on regulatory T cells. Moreover,

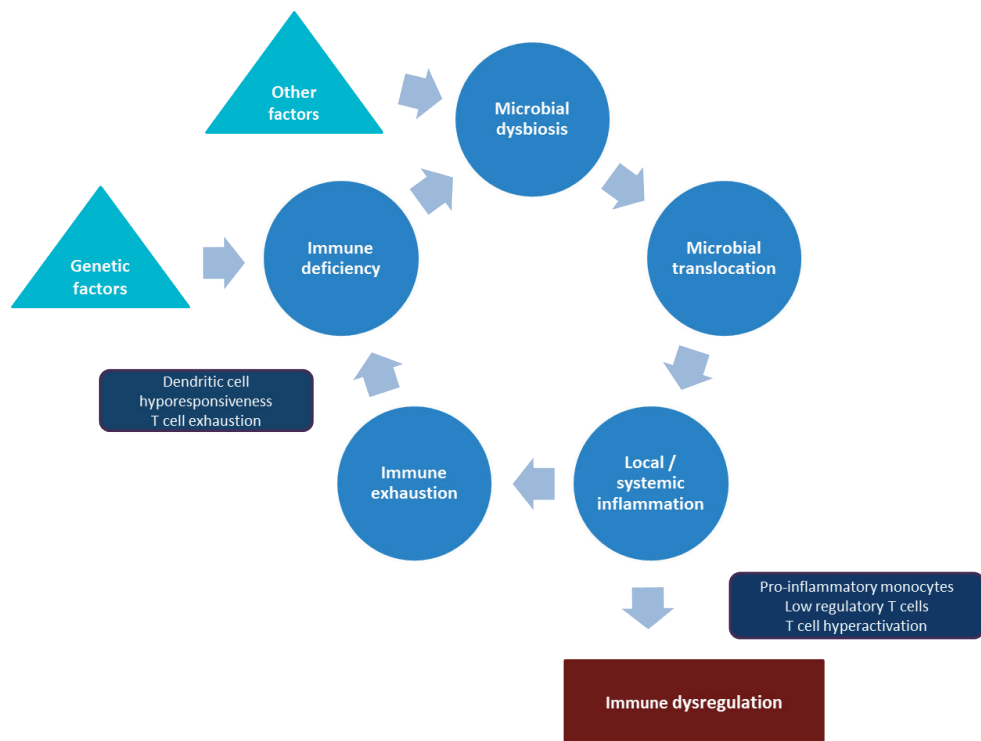


Figure 2: schematic representation of the proposed mechanism : Susceptibility to CVID in combination with other factors may lead to microbial dysbiosis. These other factors have not yet been identified, but suggested influences include infection (bacterial or viral), antibiotic use, other susceptibility factors. The resulting systemic inflammation can lead to immune dysregulation and can further amplify immune deficiency through exhaustion.

specific bacterial species and metabolites such as *Clostridium species*⁷⁷ and distinct short-chain fatty acids⁷⁸ are known to induce regulatory T cells in the gut of mice. Given their central role in the prevention of immune dysregulation, disruption of regulatory T cells may be essential to the development of the dysregulation phenotype in CVID. The suppression of the Treg subset is also not alleviated by immunoglobulin substitution therapy¹⁰, just as the phenotype does not resolve with this therapy.

CONCLUDING REMARKS: FUTURE DIRECTIONS AND CLINICAL OPPORTUNITIES

To summarize (see figure 2), there is evidence of increased microbial translocation and differences in the diversity of the microbiome in CVID. Although a major aim of microbiome research is to distinguish cause and consequence, in CVID continuous interplay between immune system and microbiota demonstrates that addressing the matter this

way may be an oversimplification. Viewing immunodeficiency and microbiome dysbiosis as a vicious cycle leaves room for the complexity of both systems and allows for the consideration of all possible mechanisms, including causes, consequences, and bystander effects. An alternative explanation to the idea that we have presented here is that microbial translocation and/or dysbiosis in CVID are a result of localized inflammation and do not further influence the disease itself. However, the intimate relationship between the microbiome and the immune system, especially considering the impaired resilience of the immune system in CVID, make it unlikely that the one can change without the other.

We have shown that the link between CVID and the microbiome is especially strong in patients with complications from immune dysregulation and have illustrated the several mechanisms by which immune deficiency can lead to changes in the microbiome, which drive systemic inflammation and immune dysregulation. This has yielded several unresolved questions and directions for future research (see Outstanding Questions). Identification of pathogenic microbial species and key pro-inflammatory modulators can help improve CVID patient care in two main ways: by providing new methods to predict at an early stage which patients are susceptible to immune dysregulation, and by indicating novel therapeutic targets that specifically interfere in the relevant (immune-) pathways or the microbiome itself. Given the intimate relationship between the microbiome and the immune system, changes in the microbiome may lead to improvement of inflammatory complication and aspects of the immune deficiency itself, as illustrated by the findings that a balanced microbiome may restore regulatory T cell control and perhaps even some IgA production in the gut.

GLOSSARY

Alpha diversity: A measure for the diversity of a microbial population. Takes into account the number of different OTUs as well as their respective prevalence. In other words, a high diversity represents a large number of species of roughly even abundances.

Beta diversity: A measure for the diversity the microbial population between two separate communities. Beta diversity also takes into account number of different OTUs as well as their abundances. If two communities (for instance two different patient groups) have a high beta diversity, this means there is little overlap between the OTUs they host.

CVID: common variable immunodeficiency disorder. A primary immunodeficiency that is characterized by low IgG, low IgA or IgM, and insufficient vaccination responses. Clinically, untreated patients primarily suffer from recurrent respiratory tract and gastrointestinal tract infections.

HLA-DR: Human Leukocyte Antigen, or a type-II Major Histocompatibility Complex protein, a surface protein that presents antigens to other immune cells.

LPS: Lipopolysaccharide. A cell wall component of Gram-negative bacteria. Upon recognition by innate immune receptors such as toll-like receptor 4, LPS has potent immunostimulatory effects.

Next generation 16S rRNA sequencing: a sequencing approach that uses the gene for the 16S ribosomal subunit as a method for the clustering of sequences of closely related culturable and non-culturable bacteria in operational taxonomic units (OTUs), to study bacterial composition of different samples.

Operational Taxonomic Unit: reflects sequence variation of a specific taxonomic marker gene. As such, it is used as a pragmatic approximation of the identity of a bacterial species, in the absence of the classical criteria for species identification (such as culture characteristics or biochemical approaches).

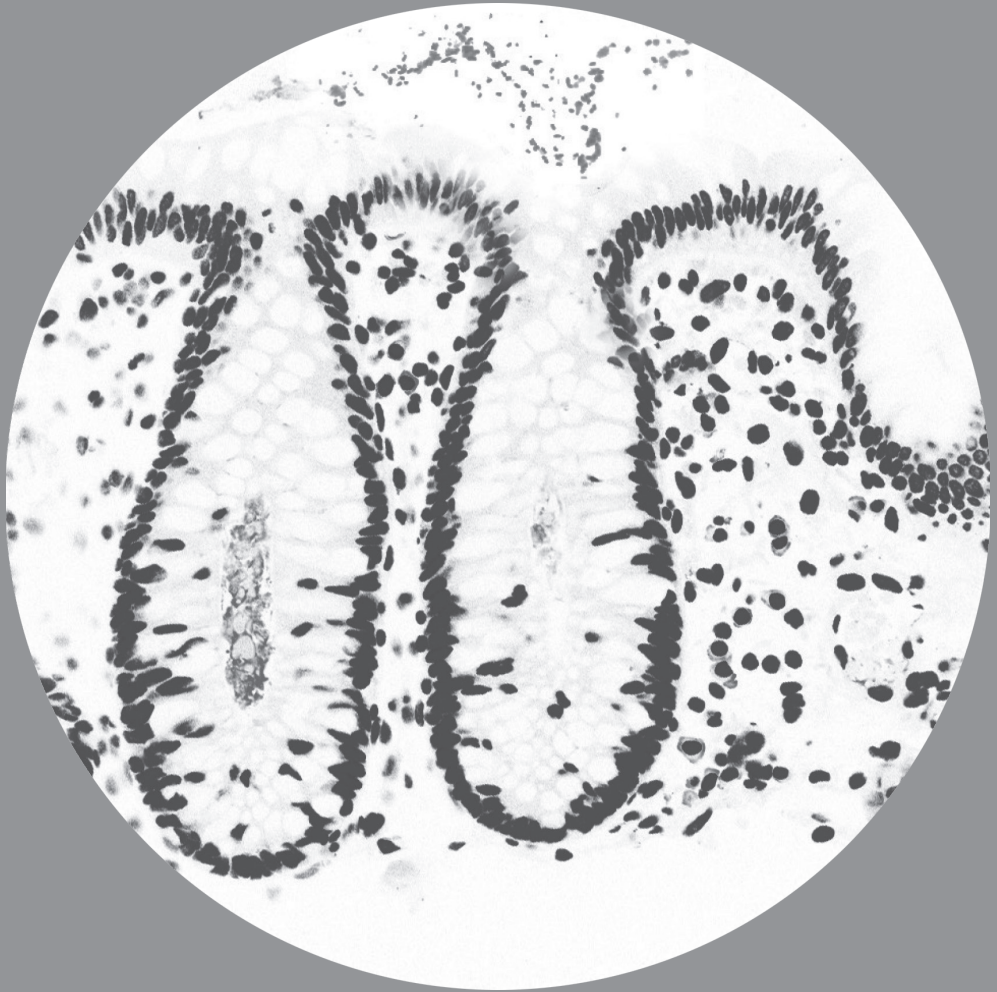
REFERENCES

- 1 Cunningham-Rundles, C. (2010) How I treat common variable immune deficiency. *Blood* 116, 7–15
- 2 Conley, M. *et al.* (1999) Diagnostic criteria for primary immunodeficiencies. *Clin. Immunol.* 93, 190–197
- 3 Ameratunga, R. *et al.* (2014) Comparison of Diagnostic Criteria for Common Variable Immunodeficiency Disorder. *Front. Immunol.* 5, 1–9
- 4 Salzer, U. *et al.* (2012) Common variable immunodeficiency - an update. *Arthritis Res. Ther.* 14, 1–11
- 5 Park, J.H. *et al.* (2011) Perspectives on common variable immune deficiency. *Ann. N. Y. Acad. Sci.* 1246, 41–9
- 6 Maarschalk-Ellebroek, L.J. *et al.* (2012) The spectrum of disease manifestations in patients with common variable immunodeficiency disorders and partial antibody deficiency in a university hospital. *J. Clin. Immunol.* 32, 907–921
- 7 Chapel, H. *et al.* (2008) Common variable immunodeficiency disorders : division into distinct clinical phenotypes. *Blood* 112, 277–287
- 8 Resnick, E.S. *et al.* (2012) Morbidity and mortality in common variable immune deficiency over 4 decades. *Blood* 119, 1650–7
- 9 Barbosa, R.R. *et al.* (2012) Monocyte activation is a feature of common variable immunodeficiency irrespective of plasma lipopolysaccharide levels. *Clin. Exp. Immunol.* 169, 263–72
- 10 Paquin-Proulx, D. *et al.* (2013) IVIg immune reconstitution treatment alleviates the state of persistent immune activation and suppressed CD4 T cell counts in CVID. *PLoS One* 8, e75199
- 11 Perreau, M. *et al.* (2014) Exhaustion of bacteria-specific CD4 T cells and microbial translocation in common variable immunodeficiency disorders. *J. Exp. Med.* 211, 2033–2045
- 12 Jørgensen, S.F. *et al.* (2016) Altered gut microbiota profile in common variable immunodeficiency associates with levels of lipopolysaccharide and markers of systemic immune activation. *Mucosal Immunol.* DOI: 10.1038/mi.2016.18
- 13 Belkaid, Y. and Hand, T.W. (2014) Role of the microbiota in immunity and inflammation. *Cell* 157, 121–141
- 14 Hooper, L. V and Macpherson, A.J. (2010) Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat. Rev. Immunol.* 10, 159–69
- 15 Kabat, A.M. *et al.* (2014) Modulation of immune development and function by intestinal microbiota. *Trends Immunol.* 35, 507–517
- 16 Laugerette, F. *et al.* (2011) Emulsified lipids increase endotoxemia: Possible role in early postprandial low-grade inflammation. *J. Nutr. Biochem.* 22, 53–59
- 17 Brenchley, J.M. *et al.* (2006) Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat. Med.* 12, 1365–71
- 18 Sandler, N.G. and Douek, D.C. (2012) Microbial translocation in HIV infection: causes, consequences and treatment opportunities. *Nat. Rev. Microbiol.* 10, 655–66
- 19 Xavier, R.J.P.D.K. (2007) Unravelling the pathogenesis of inflammatory bowel disease. *Nature* 448, 427–434
- 20 Rojo, Ó.P. *et al.* (2007) Serum lipopolysaccharide-binding protein in endotoxemic patients with inflammatory bowel disease. *Inflamm. Bowel Dis.* 13, 269–277
- 21 Sandler, N.G. *et al.* (2011) Plasma levels of soluble CD14 independently predict mortality in HIV infection. *J. Infect. Dis.* 203, 780–790
- 22 Sandler, N.G. *et al.* (2011) Host response to translocated microbial products predicts outcomes of patients with HBV or HCV infection. *Gastroenterology* 141, 1220–1230.e3
- 23 Litzman, J. *et al.* (2012) Chronic immune activation in common variable immunodeficiency (CVID) is associated with elevated serum levels of soluble CD14 and CD25 but not endotoxaemia. *Clin. Exp. Immunol.* 170, 321–32
- 24 Wang, J.E. *et al.* (2003) Peptidoglycan and lipoteichoic acid in gram-positive bacterial sepsis: receptors, signal transduction, biological effects, and synergism. *Shock* 20, 402–414

- 25 Yoshimura, a *et al.* (1999) Recognition of Gram-positive bacterial cell wall components by the innate immune system occurs via Toll-like receptor 2. *J. Immunol.* 163, 1–5
- 26 Scher, J.U. *et al.* (2013) Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *Elife* 2013, 1–20
- 27 Palm, N.W. *et al.* (2014) Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* 158, 1000–1010
- 28 Lozupone, C. *et al.* (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* 489, 220–30
- 29 Nowak, P. *et al.* (2015) Gut microbiota diversity predicts immune status in HIV-1 infection. *Aids* 29, 2409–18
- 30 Sjölund, M. *et al.* (2003) Long-Term Persistence of Resistant *Enterococcus* Species after Antibiotics to Eradicate *Helicobacter pylori*. *Ann. Intern. Med.* 139, 483–487+I42
- 31 Dethlefsen, L. and Relman, D. a (2011) Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl. Acad. Sci. U. S. A.* DOI: 10.1073/pnas.1000087107
- 32 Jernberg, C. *et al.* (2007) Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *Isme.J.* 1, 56–66
- 33 Van Zelm, M.C. *et al.* (2006) An antibody-deficiency syndrome due to mutations in the CD19 gene. *N. Engl. J. Med.* 354, 1901–1912
- 34 Kuijpers, T.W. *et al.* (2010) CD20 deficiency in humans results in impaired T cell-independent antibody responses. *J. Clin. Invest.* 120, 214–222
- 35 Bossaller, L. *et al.* (2006) ICOS deficiency is associated with a severe reduction of CXCR5+CD4 germinal center Th cells. *J. Immunol.* 177, 4927–4932
- 36 Li, J. *et al.* (2015) Association of CLEC16A with human common variable immunodeficiency disorder and role in murine B cells. *Nat. Commun.* 6, 6804
- 37 Li, J. *et al.* (2016) Understanding the genetic and epigenetic basis of common variable immunodeficiency disorder through omics approaches. *Biochim. Biophys. Acta - Gen. Subj.* DOI: 10.1016/j.bbagen.2016.06.014
- 38 Pabst, O. *et al.* (2016) Secretory IgA in the Coordination of Establishment and Maintenance of the Microbiota. *Trends Immunol.* 37, 287–296
- 39 Van de Ven, a. a J.M. *et al.* (2014) Increased Prevalence of Gastrointestinal Viruses and Diminished Secretory Immunoglobulin a Levels in Antibody Deficiencies. *J. Clin. Immunol.* 34, 962–970
- 40 Mantis, N.J. *et al.* (2011) Secretory IgA's complex roles in immunity and mucosal homeostasis in the gut. *Mucosal Immunol.* 4, 603–11
- 41 Wei, M. *et al.* (2011) Mice carrying a knock-in mutation of *Aicda* resulting in a defect in somatic hypermutation have impaired gut homeostasis and compromised mucosal defense. *Nat. Immunol.* 12, 264–270
- 42 Kawamoto, S. *et al.* (2014) Foxp3+ T Cells Regulate Immunoglobulin A Selection and Facilitate Diversification of Bacterial Species Responsible for Immune Homeostasis. *Immunity* 41, 152–165
- 43 Suzuki, K. *et al.* (2004) Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut. *Proc. Natl. Acad. Sci. U. S. A.* 101, 1981–1986
- 44 Macpherson, A.J. and Uhr, T. (2004) Induction of protective IgA by intestinal dendritic cells carrying commensal bacteria. *Science*. 303, 1662–1665
- 45 Mantis, N.J. *et al.* (2002) Selective adherence of IgA to murine Peyer's patch M cells: evidence for a novel IgA receptor. *J. Immunol.* 169, 1844–1851
- 46 Franssen, F. *et al.* (2015) BALB/c and C57BL/6 Mice Differ in Polyreactive IgA Abundance, which Impacts the Generation of Antigen-Specific IgA and Microbiota Diversity. *Immunity* 43, 527–540
- 47 Peterson, D. a. *et al.* (2007) IgA Response to Symbiotic Bacteria as a Mediator of Gut Homeostasis. *Cell Host Microbe* 2, 328–339
- 48 Peterson, D. a. *et al.* (2015) Characterizing the interactions between a naturally primed immunoglobulin a and its conserved *Bacteroides thetaiotaomicron* species-specific epitope in gnotobiotic mice. *J. Biol. Chem.* 290, 12630–12649
- 49 Boullier, S. *et al.* (2009) Secretory IgA-mediated neutralization of *Shigella flexneri* prevents intestinal tissue destruction by down-regulating inflammatory circuits. *J. Immunol.* 183, 5879–5885

- 50 Hand, T.W. (2016) The Role of the Microbiota in Shaping Infectious Immunity. *Trends Immunol.* 37, 647–658
- 51 Giovannetti, a. *et al.* (2007) Unravelling the Complexity of T Cell Abnormalities in Common Variable Immunodeficiency. *J. Immunol.* 178, 3932–3943
- 52 Wang, N. *et al.* (2011) Selective IgA deficiency in autoimmune diseases. *Mol. Med.* 17, 1383–96
- 53 Yoshida, M. *et al.* (2004) Human neonatal Fc receptor mediates transport of IgG into luminal secretions for delivery of antigens to mucosal dendritic cells. *Immunity* 20, 769–783
- 54 Maarschalk-Ellebroek, L.J. Oldenburg, B. Mombers, I. M. H., Hoepelman, A. I. M. Brosens, L.A.A. *et al.* (2013) Outcome of screening endoscopy in common variable immunodeficiency disorder and X-linked agammaglobulinemia. *Endoscopy* 45, 320–323
- 55 Jørgensen, S.F. *et al.* (2016) A Cross-Sectional Study of the Prevalence of Gastrointestinal Symptoms and Pathology in Patients With Common Variable Immunodeficiency. *Am. J. Gastroenterol.* DOI: 10.1038/ajg.2016.329
- 56 Ziegler-Heitbrock, L. *et al.* (2010) Nomenclature of monocytes and dendritic cells in blood. *Blood* 116, e74–80
- 57 Fingerle, G. and Pforte, A. (1993) The novel subset of CD14+/CD16+ blood monocytes is expanded in sepsis patients. *Blood* 82, 3170–3176
- 58 Rossol, M. *et al.* (2012) The CD14⁺ brightCD16⁺ monocyte subset is expanded in rheumatoid arthritis and promotes expansion of the Th17 cell population. *Arthritis Rheum.* 64, 671–677
- 59 Ciccia, F. *et al.* (2013) IL-34 is overexpressed in the inflamed salivary glands of patients with Sjogren's syndrome and is associated with the local expansion of pro-inflammatory CD14(bright) CD16⁺ monocytes. *Rheumatology (Oxford)*. 52, 1009–17
- 60 Scott-Taylor, T.H. *et al.* (2016) Enhanced formation of Giant cells in common variable immunodeficiency; relation to granulomatous disease. *Clin. Immunol.* DOI: 10.1016/j.clim.2016.11.002
- 61 Taraldsrud, E. *et al.* (2014) Common variable immunodeficiency revisited: normal generation of naturally occurring dendritic cells that respond to Toll-like receptors 7 and 9. *Clin. Exp. Immunol.* 175, 439–48
- 62 Viillard, J.-F. *et al.* (2005) Altered dendritic cell distribution in patients with common variable immunodeficiency. *Arthritis Res. Ther.* 7, R1052–5
- 63 Scott-Taylor, T.H. *et al.* (2006) Defective maturation of dendritic cells in common variable immunodeficiency. *Clin. Exp. Immunol.* 145, 420–7
- 64 Bayry, J. *et al.* (2004) Common variable immunodeficiency is associated with defective functions of dendritic cells. *Blood* 104, 2441–2443
- 65 Cunningham-Rundles, C. *et al.* (2006) TLR9 Activation Is Defective in Common Variable Immune Deficiency. *J. Immunol.* 176, 1978–1987
- 66 Yu, J.E. *et al.* (2009) Toll-like receptor 7 and 9 defects in common variable immunodeficiency. *J. Allergy Clin. Immunol.* 124, 349–56, 356.e1–3
- 67 Yu, J.E. *et al.* (2012) TLR-mediated B cell defects and IFN- α in common variable immunodeficiency. *J. Clin. Immunol.* 32, 50–60
- 68 Bayry, J. *et al.* (2004) Common variable immunodeficiency is associated with defective functions of dendritic cells. *Blood* 104, 2441–2443
- 69 Yazdani, R. *et al.* (2014) Genetic defects and the role of helper T-cells in the pathogenesis of common variable immunodeficiency. *Adv. Biomed. Res.* 3, 2
- 70 Hel, Z. *et al.* (2014) Altered Serum Cytokine Signature in Common Variable Immunodeficiency. *J. Clin. Immunol.* DOI: 10.1007/s10875-014-0099-z
- 71 Park, J. *et al.* (2013) Interferon signature in the blood in inflammatory common variable immune deficiency. *PLoS One* 8, e74893
- 72 Maarschalk-Ellebroek, L.J. *et al.* (2014) CT Screening for Pulmonary Pathology in Common Variable Immunodeficiency Disorders and the Correlation with Clinical and Immunological Parameters. *J. Clin. Immunol.* 34, 642–654
- 73 De Lollo, C. *et al.* (2016) Impaired CD8⁺ T cell responses upon Toll-like receptor activation in common variable immunodeficiency. *J. Transl. Med.* 14, 138
- 74 Genre, J. *et al.* (2009) Reduced frequency of CD4(+)CD25(HIGH)FOXP3(+) cells and diminished FOXP3 expression in patients with Common Variable Immunodeficiency: a link to autoimmunity? *Clin. Immunol.* 132, 215–21

- 75 Arandi, N. *et al.* (2013) Evaluation of CD4+CD25+FOXP3+ regulatory T cells function in patients with common variable immunodeficiency. *Cell. Immunol.* 281, 129–133
- 76 Soheili, H. *et al.* (2013) Evaluation of natural regulatory T cells in subjects with selective IgA deficiency: From senior idea to novel opportunities. *Int. Arch. Allergy Immunol.* 160, 208–214
- 77 Atarashi, K. *et al.* (2011) Induction of Colonic Regulatory T Cells by Indigenous Clostridium Species. *Science.* 331, 337–342
- 78 Arpaia, N. *et al.* (2013) Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature* 504, 451–5
- 79 Cunningham-Rundles, C. (2008) Autoimmune manifestations in common variable immunodeficiency. *J. Clin. Immunol.* 28 Suppl 1, S42–5
- 80 Yong, P.F.K. *et al.* (2011) “A Rose is a Rose is a Rose,” but CVID is Not CVID. *Common Variable Immune Deficiency (CVID), What do we Know in 2011?*, (1st edn) , 111Elsevier Inc.
- 81 Resnick, E.S. *et al.* (2012) Morbidity and mortality in common variable immune deficiency over 4 decades. *Blood* 119, 1650–7



Chapter 5

Immunoglobulin A and microbiota in primary immunodeficiency diseases

Roos-Marijn Berbers
Ingrid A. Franken
Helen L. Leavis

Current opinion in allergy and clinical immunology (2019)
19 (6): 536-570

ABSTRACT

Purpose of review

With the emergence of the microbiota as a potential driver of host inflammation, the role of immunoglobulin (Ig)A is becoming increasingly important. This review discusses the current evidence regarding the effects of clinical IgA deficiency on the microbiota, and the possible role of microbial dysbiosis in driving inflammation in primary immunodeficiency disease (PID) patients.

Recent findings

The gut microbiota has been investigated in selective IgA deficiency and common variable immunodeficiency, revealing an important role for IgA in maintaining gut microbiota homeostasis, with disparate effects of IgA on symbionts and pathobionts. While IgA deficiency is associated with microbial translocation and systemic inflammation, this may be partially compensated by adequate IgG and IgM induction in IgA deficiency but not in common variable immunodeficiency. Therapeutic strategies aimed at correction of the microbiota mostly focus on fecal microbiota transplantation. Whether this may reduce systemic inflammation in PID is currently unknown.

Summary

Clinical IgA deficiency is associated with microbial dysbiosis and systemic inflammation. The evidence for microbiota-targeted therapies in PID is scarce, but indicates that IgA-based therapies may be beneficial, and that fecal microbiota transplantation is safe in patients with antibody deficiency.

Key words

IgA, antibody deficiency, microbiota, inflammation, PID

Key message

- Microbiota profiling in patients with primary immunodeficiencies has revealed an important role for IgA in modulating gut microbiota homeostasis.
- IgA deficiency and microbial dysbiosis are associated with systemic inflammation in PID and endotoxemia.
- IgA-enriched immunoglobulin substitution therapy and microbiota-targeted therapies should be investigated as promising options for PID patients with systemic inflammation.

INTRODUCTION

IgA is considered to be one of the main host factors regulating microbiota homeostasis^{1,2}, but the effects of clinical IgA deficiency on the microbiota and its consequences are still poorly understood. Many patients with primary immunodeficiency diseases (PID) have reduced or absent IgA production, often in combination with other humoral or cellular defects. PIDs that involve IgA deficiency include the predominantly humoral immune defects selective IgA deficiency (sIgAD), common variable immunodeficiency (CVID), X-linked agammaglobulinemia (XLA), hyper IgM syndromes³, and all (severe) combined immunodeficiencies, but also other PID. For instance, low IgA levels are part of the diagnostic criteria for Ataxia-Telangiectasia and Nijmegen breakage syndrome⁴, and IgA deficiency has been reported in DiGeorge syndrome⁵, WHIM⁶ (warts, hypogammaglobulinemia, infections and myelokathexis syndrome), and ADA2 deficiency⁷. While IgG can be supplemented in case of hypogammaglobulinemia, IgA deficiency remains untreated.

In PID, immunodeficiency and autoimmune disease often co-exist⁸. Dysbiosis of the microbiota is increasingly investigated as a potential cause in the pathogenesis of inflammatory disorders such as inflammatory bowel disease (IBD), rheumatoid arthritis, multiple sclerosis and type-1 diabetes⁹. As IgA is an important regulator of the microbiota, microbial dysbiosis in PID may mediate immunodeficiency with autoimmune disease. Clinically, low IgA in CVID is associated with immune dysregulation^{10,11}, and sIgAD is associated with increased prevalence of autoimmune disease^{12,13}. In one study, the increased risk of juvenile idiopathic arthritis, systemic lupus erythematosus, and IBD was found to be 8.9, 8.9 and 5.0 respectively in sIgAD compared to the general population¹².

As the recent developments in the field of microbiota have renewed the interest in the role of IgA in maintaining microbiota homeostasis, special attention has gone out to PIDs with antibody deficiency as a natural model for IgA deficiency (IgAD). Here, we aim to integrate recent findings of microbiota studies in clinical antibody deficiency with experimental data on the functions of IgA in order to specify which bacteria are affected in IgAD and to evaluate the evidence for inflammatory consequences of microbial dysbiosis in PID. In addition, we highlight novel therapeutic strategies that target the microbiota and discuss their potential role in PID treatment.

Role of IgA in orchestrating mucosal immunity

Each individual is colonized with a distinct community of microbiota, which is shaped after birth by intrinsic and extrinsic factors, such as the immune response, host genetics, exposure to microbes, medication and diet. Of all microbial communities hosted by the human body, the gut microbiota harbors the largest and most diverse community of approximately 1.000 bacterial species per individual¹⁴. The intestinal microbiota contributes to the maturation of the mucosal immune system, including the development of gut-associated lymphoid tissues¹⁵ and the secretion of IgA¹⁶. IgA, in turn, is thought to modulate the intestinal microbiota through various different mechanisms, including immune exclusion, neutralization, and modulation of bacterial gene expression¹⁷.

IgA can be induced broadly by two immunological mechanisms, distinguishing T-cell independent and T-cell dependent IgA production¹⁸. T-cell independent IgA does not undergo somatic hypermutation or affinity maturation and has been suggested to be a type of “innate” or naturally occurring antibody that targets mainly commensals and/or self-antigens^{17,19}. This IgA has been shown to be polyreactive, with monoclonal IgAs recognizing various taxonomically distinct bacteria¹⁹. T-cell deficient mice displayed similar levels of IgA-coated bacteria as wildtype mice²⁰, but the microbiota composition of IgAD mice differed²¹. T-cell dependent IgA has higher affinity and avidity than T-cell independent IgA, and is thought to mainly target pathogens^{22,23}. T-independent and T-dependent IgA may have distinct roles in homeostasis and disease²³, which could reflect the diverse findings of IgA-coating studies. In addition, this implies that also PID with a dominant T-cell defect and seemingly intact IgA titers may still result in gut dysbiosis due to impaired induction of T-dependent high affinity IgA.

IgA can exert different functions in different contexts. For instance, IgA can both protect the intestinal barrier from bacterial attachment, and promote bacterial association to the epithelium²⁴. For instance, the symbiotic immunoregulatory² *Bacteroides fragilis* synthesizes phase-variable polysaccharides, not to evade²⁵ but to promote IgA mediated binding to the intestinal epithelium²⁴. Compartmentalization between the mucosa and the gut lumen is reduced in IgA deficiency, illustrating the role of IgA mediating the location of bacteria^{26–28}. As a result, IgA may orchestrate antigen exposure and systemic antibody responses against the intestinal commensal bacteria²⁹. IgA preferentially coats IBD-driving bacteria in mice and humans²⁸, possibly because pro-inflammatory bacteria elicit a stronger IgA response, and/or because impaired barrier function in IBD leads to increased exposure of these bacteria to the immune system. Moreover, polymeric Ig receptor knock-out in mice, which impairs secretion of IgA and IgM to the lumen, is associated with higher translocation of bacteria to mesenteric lymph nodes and an increase in IgA-producing cells in the gut, as well as in serum IgA and IgG reactive to these bacteria²⁹.

IgA coating has differential effects on symbionts and pathobionts

Bacteria targeted by IgA in human studies include both symbionts (resident microbes with no known detrimental effects on the host, may in some cases even be health-promoting) and pathobionts (resident microbes with pathogenic potential), and IgA coating can vary within bacterial families. A complete overview is presented in table 1. The gut microbiota has been investigated in two humoral PID; sIgAD in which serum IgA is decreased, but the other immunoglobulin isotypes are still intact; and CVID in which IgG and either IgM and/or IgA are decreased. In CVID, gut dysbiosis is more severe than in sIgAD, reflected by decreased alpha diversity (within-sample diversity, this reflects the amount and distribution of different taxa within a sample) and distinct beta diversity (between-sample diversity, this reflects how similar samples are to each other) compared to healthy controls^{30,31}.

Integrating studies on IgA coating of fecal bacteria and differential abundance of these bacteria in sIgAD shows that IgA coating appears to have a different effect on symbionts

and pathobionts. In both sIgAD and CVID, IgA-targeted commensals (including many Firmicutes) are predominantly depleted, while IgA-bound pathobionts such as *Escherichia coli* are expanded^{30,31}. Symbionts *Blautia*^{30–32}, *Bifidobacterium*^{28,31} and *Bacteroides fragilis*^{24,25} are coated and reduced in sIgAD, while non-coated *Roseburia*^{28,33} are expanded. Conversely, pathobionts *Prevotella*^{30,32} and *E. coli*^{30,33} are IgA coated and expanded in sIgAD.

In mice, specific IgA reduces the abundance of Gammaproteobacteria relative to Firmicutes and Bacteroides, but IgAD impairs this maturation of the gut and allows for Gammaproteobacteria, including Enterobacteriaceae, to persist and cause intestinal damage³⁴. Especially Enterobacteria such as *Escherichia coli* are found to be IgA-coated^{20,30,33–35} and to be expanded in IgAD^{30,31}, suggesting a controlling function of IgA. Also Prevotellaceae species are coated³² and increased in sIgAD³⁰ in humans, which may propagate intestinal inflammation and promote susceptibility to IBD³². The immunogenic Erysipelotrichaceae are similarly IgA-coated in humans and mice²⁸ and increased in human sIgAD³⁰.

Strikingly, other studies do not observe any change in diversity^{29,30} or composition²⁹ of the sIgAD microbiota. This is for instance elaborately investigated for *Lactobacillus*, which is IgA-coated^{20,28,32,35,37} but does not alter relative abundance in sIgAD³⁸. Loss of *lactobacilli* could be harmful as they may exert anti-inflammatory effects and counteract pathogenic colonization³⁹.

In primary antibody deficiency, gut dysbiosis may contribute to systemic inflammation

In addition to dysbiosis of gut microbiota, bacterial products have also been detected in humoral PID serum. While IgG supplementation does reduce endotoxemia in CVID⁴⁰, increased lipopolysaccharide (LPS) has been observed in patients with inflammatory and autoimmune complications, also reflected by increased sCD14 and sCD25 levels³¹. Although correlation of LPS with IgA is not found in this study, very low serum IgA has been correlated with lower alpha diversity and increased dysbiosis³¹. In another study, mucosal IgA deficiency in CVID has been associated with duodenal inflammation and malabsorption⁴¹, providing a link between IgA deficiency in CVID and epithelial damage that may facilitate microbial translocation. However, microbial translocation has not been shown in sIgAD, with one study reporting no increase in serum LPS levels³⁰. These results suggest that microbiota dysbiosis and local inflammation may be IgA-associated, but that systemic microbial translocation (as reflected by serum LPS level) may be prevented by compensation through IgG or IgM. Indeed, at composition level, Fadlallah et al. report that IgM can compensate sIgAD for some, but not all bacteria. IgM effectively binds Actinobacteria (including *Bifidobacterium*), but only a narrow range of Firmicutes (including *Clostridium* and *Faecalibacterium*) and Proteobacteria (excluding Enterobacteriaceae)³⁰. The authors proceed to show that patients with CVID, who in some instances harbor reduced serum levels of IgM as well as IgA, display a greater decrease in Actinobacteria diversity³⁰. For example, loss of *Bifidobacterium* in sIgAD is exacerbated in the absence of compensatory IgM in CVID³⁰. Similarly, adequate IgM and/or IgG induction in sIgAD may protect from endotoxemia, while this compensatory response is lacking in

Table 1: Findings on IgA-Coating and Differential Relative Abundance in IgA Deficiency, on Taxa Investigated in More Than One Study**Taxonomy**

Phylum	Class	Order	Family	Genus
Actinobacteria	Actinobacteria	Bifidobacteriales	Bifidobacteriaceae	<i>Bifidobacterium</i>
Actinobacteria	Actinobacteria	Bifidobacteriales	Bifidobacteriaceae	<i>Bifidobacterium</i>
Actinobacteria	Coriobacteria	Eggerthellales	Eggerthellaceae	<i>Adlercreutzia</i>
Bacteroidetes	Bacteroidia	Bacteroidales		
Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	<i>Bacteroides</i>
Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	
Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	<i>Prevotella</i>
Bacteroidetes	Bacteroidia	Bacteroidales	Rikenellaceae	
Bacteroidetes	Bacteroidia	Bacteroidales	S24-7	
Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	<i>Enterococcus</i>
Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	<i>Lactobacillus</i>
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	<i>Clostridium</i>
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	SFB
Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	
Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	<i>Dorea</i>
Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	<i>Blautia</i>
Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	<i>Coproccoccus</i>
Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	<i>Roseburia</i>
Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	<i>Faecalibacterium</i>
Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	<i>Oscillospira</i>
Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	<i>Ruminococcus</i>
Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	<i>Ruminococcus</i>
Firmicutes	Clostridia	Clostridiales		
Firmicutes	Clostridia	Clostridiales	Mogibacteriaceae	
Firmicutes	Erysipelotrichia	Erysipelothriachiales	Erysipelotrichaceae	
Firmicutes	Negativicutes	Veillonellales	Veillonellaceae	<i>Veillonella</i>
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	<i>Escherichia</i>
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Morganellaceae	<i>Morganella</i>
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonaceae	<i>Pseudomonas</i>
Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	<i>Stenotrophomonas</i>
Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	<i>Akkermansia</i>

Species	IgA Coating in Humans						IgA Coating in Mice					Abundance in IgA-Deficient Humans		Abundance in IgA-Deficient Mice	
	ref (34)	ref (24)	ref (31)	ref (32)	ref (29)	ref (27)	ref (36)	ref (27)	ref (19)	ref (23)	ref (33)	ref (29)	ref (30)	ref (23)	ref (26)
						IgA+							↓		
<i>bifidum</i>				IgA+	IgA+		IgA+								↑
						IgA-	IgA-								↑
<i>fragilis</i>		IgA+							IgA+						↓
					IgA+					IgA-/+		↑			↓
							IgA-	IgA-	IgA+						↓
							IgA+	IgA+							
<i>faecalis</i>						IgA-			IgA+	IgA-					
						IgA+			IgA+	IgA+					↑
						IgA+	IgA+					↑/↓			
							IgA+	IgA+	IgA+						↑
						IgA+	IgA+	IgA-		IgA-	IgA-		↓		↓
						IgA+			IgA+				↑		
						IgA+	IgA-	IgA+		IgA+		↓	↓		↑
							IgA-					↓			↓
							IgA-	IgA-					↑		
						IgA+	IgA+								
							IgA-		IgA-						
						IgA+	IgA-	IgA+		IgA-			↓		↑
<i>torques</i>						IgA+	IgA+	IgA+							
									IgA-				↓/↑		↑
							IgA-		IgA-						
							IgA+		IgA+				↑		
						IgA+	IgA+								
									IgA+	IgA+		↑	↑		
<i>coli</i>						IgG		IgA+	IgA+				↑		
<i>morganii</i>						IgA+				IgA+					
										IgA+					
										IgA+					
<i>muciniphila</i>										IgA+					

CVID. As CVID patients can have defects in somatic hypermutation⁴², even any residual immunoglobulin production may be functionally impaired. In addition, a second study by Fadlallah et al. recently showed that IgG from healthy donors targets CVID microbiota much less effectively than healthy microbiota⁴³, suggesting that IgG supplementation may be insufficient to limit microbiota dysregulation and/or translocation in CVID.

Increased LPS has been associated with exhaustion markers and reduced T-cell response to bacteria in CVID, suggesting that these T cell defects may be related to microbial translocation⁴⁰. Interestingly, the T cell phenotype observed in CVID is reminiscent of T cell dysregulation in HIV, where a causal role for microbial translocation in systemic immune activation has been shown⁴⁴. Also in a mouse model that carries the same genetic hypomorphic RAG mutation as observed in Omenn syndrome, IgAD resulted in dysbiosis of the intestinal microbiota that has been shown to contribute to Th1 and Th17-mediated inflammation and autoimmunity^{45,46}. sIgAD has also been associated with systemic inflammation (increased levels of sCD14, and Th17 cells) and exhaustion markers (increased PD-1 expression)³⁰. Notably, half of sIgAD patients in this study suffered from clinical autoimmune disease³⁰. The cause for this inflammation is currently unknown, given the lack of reported endotoxemia in this study. Possibly, local inflammation in the gut due to microbial dysbiosis is sufficient to cause systemic immune dysregulation in some patients.

Exploration of therapeutic options

This increasing evidence for IgA-dependent microbial dysbiosis and microbiota-dependent immune dysregulation in PID suggests that patients may also benefit from microbiota-targeted therapeutics. Modulation of the microbiota can be considered from two directions; direct targeting of the bacterial composition through fecal microbiota transplantation (FMT), pre- or probiotics, or through immune-mediated strategies.

In non-PID microbial dysbiosis, the first strategy is most often considered. The most rigorous way to alter the microbiota is through FMT. Most experience with FMT has been the treatment of recurrent *Clostridium difficile* infections, and three studies mention inclusion of PID patients⁴⁷⁻⁴⁹. One study with three CVID and three sIgAD patients reported no safety concerns and successful resolution of *C. difficile* infection⁴⁹ (Dr. M. Fischer, personal communication). There are currently no reports on the effects of FMT on inflammatory or immune dysregulatory symptoms. The question remains whether temporary correction of microbiota will be sufficient to limit dysbiosis and microbial translocation in PID while the immune deficiency persists. Possibly, recurrent treatment would be required. Indeed, one-time FMT in HIV patients resulted in limited engraftment of the donor microbiota⁵⁰, although this may be due to lack of antibiotic pre-treatment in this study.

Antibiotic treatment is also known to alter the gut microbiota. While any antibiotic use over the year before microbiota sampling (both prophylactic and short-term) was not found to alter the microbiota in one CVID study³¹, cumulative effects of long-term antibiotic use, as often occurs in PID patients, cannot be excluded. Jørgensen et al. administered a broad-spectrum antibiotic that only acts locally in the gut with the aim to correct micro-

bial dysbiosis in CVID, but observed no effects on immune dysregulation, serum LPS, sCD15 or sCD25⁵¹. This outcome, together with reports from a mouse study that showed that broad-spectrum antibiotics, including amoxicillin plus clavulanic acid, can repress mucosal IgA production⁵², may suggest that antibiotic monotherapy is not a promising strategy to modulate the gut microbiota. Approaches that target specific bacteria, such as probiotics or phage therapy, require better understanding of the causal role of individual bacteria in causing inflammation in PID.

Alternatively, immunological interventions may be considered. As IgG from healthy donors does not target CVID microbiota very effectively, it has been proposed that patients may benefit more from IgG from asymptomatic sIgAD patients⁴³. Direct supplementation of IgA may also be considered. One case study in which fresh frozen plasma was administered to an XLA patient showed detectable secretory IgA in saliva and increased agglutination of *Haemophilus influenzae*, proving that IgA administered systemically can be delivered to mucosal tissues⁵³. IgA- and IgM enriched immunoglobulin preparations, were shown to be more effective in clearing endotoxemia than regular IgRT^{54–56}. These preparations have been used to treat persistent *Campylobacter jejuni* infections and was well tolerated in two hypogammaglobulinemia patients⁵⁷. An IgA-enriched immunoglobulin preparation was shown to be effective in children with chronic diarrhea when orally administered⁵⁸.

For patients with lung complications, inhalation IgA may be considered. In a mouse study, intranasal IgA was shown to improve *P. aeruginosa* immunity and prevent systemic translocation⁵². A similar approach may be considered for oral immunoglobulin administration. In mice, oral monoclonal IgA targeting *E. coli* was found to abrogate colitis in mice³⁷. The medical nutritional supplement known as serum bovine immunoglobulins (SBI), which contains serum proteins such as albumin and ferritin in addition to IgG, IgA and IgM, was shown to improve gut leakiness and systemic inflammation in HIV enteropathy⁵⁹. Enteral IgG therapy and breast milk (as a source of IgA) have been used as an experimental therapy for chronic norovirus infection in CVID^{60,61}, with mixed results and no data on improvement of immune dysregulation and/or microbial translocation.

CONCLUSION

Here, we have provided an overview of recent developments in our understanding of the role of IgA in maintaining the gut microbiota, alterations in the microbiota composition due to IgAD, subsequent association with local and systemic inflammation and the influence of IgAD in PID. Based on this, we identify IgA as an important potential therapeutic target that may improve inflammation in a variety of PIDs. In addition, we highlight the increased prevalence of immune dysregulation in patients with IgAD, which justifies increased vigilance for autoimmune- or inflammatory complications in PID patients with concurrent IgAD, such as CVID with absent serum IgA.

Further studies are encouraged to further explore microbiota- and IgA-based therapies in models of PID and/or microbial dysbiosis. In addition, longitudinal studies with PID patients with concurrent IgA deficiency would help elucidate whether IgAD precedes or is preceded by microbial dysbiosis and translocation. Other currently unexplored territory includes the respiratory microbiota, which may contribute to lung disease in PID, but to our knowledge has not yet been characterized in patients with (humoral) PID. The contribution of antibiotics to the observed microbial dysbiosis in PID also warrants further investigation. Although no effects of short-term antibiotic use have been observed in CVID⁶¹, long-term antibiotic treatment may be harmful to the gut microbiota composition⁶². If this is indeed the case, limiting the effects of antibiotics on the gut microbiota, for instance through active charcoal treatment⁶³, may be beneficial to PID patients.

IgA-based therapy may also be beneficial as a tool to modulate the microbiota in other inflammatory diseases where the microbiota has been causally implicated. Given the intimate dialogue between the immune system and the microbiota, it is possible that interventions aimed at the microbiota alone may not yield the desired results. Illustrative of this point is the observation that 25% of IBD patients who received FMT for *C. difficile* infection experienced a flare of their IBD⁶⁴, suggesting that “correction” of the microbiota may be less effective and may have unexpected side-effects in patients with concurrent immune dysregulation. However, a study showing that FMT was very effective in resolving refractory autoimmune colitis in patients treated with immune checkpoint inhibitors⁶⁵ indicates that FMT has potential in the treatment of (T-cell mediated) immune dysregulation.

Acknowledgements

The authors have no acknowledgements to report.

Financial support and sponsorship

This work was supported by the Wilhelmina Children’s Hospital Fund and UMC Utrecht Infection & Immunity Strategic Theme grant, Utrecht, NL.

Conflicts of interest

H.L. reports grants and personal fees from Shire/Takeda, outside the submitted work. For the remaining authors none were declared.

REFERENCES

1. Corthésy B. Multi-faceted functions of secretory IgA at mucosal surfaces. 2013;4(July):1–11.
2. Palm NW, de Zoete MR, Flavell RA. Immune-microbiota interactions in health and disease. *Clin Immunol*. 2015;159(2):122–7.
3. Davies EG, Thrasher AJ. Update on the hyper immunoglobulin M syndromes. *Br J Haematol*. 2010;149(2):167–80.
4. Immunodeficiencies ES for. Diagnostic Criteria PID . 2019. Available from: <https://esid.org/Education/Diagnostic-Criteria-PID>
5. Finocchi A, Di Cesare S, Romiti ML, Capponi C, Rossi P, Carsetti R, et al. Humoral immune responses and CD27+ B cells in children with DiGeorge syndrome (22q11.2 deletion syndrome). *Pediatr Allergy Immunol*. 2006;17(5):382–8.
6. Wetzler M, Talpaz M, Kleinerman ES, King A, Huh YO, Gutterman JU, et al. A new familial immunodeficiency disorder characterized by severe neutropenia, a defective marrow release mechanism, and hypogammaglobulinemia. *Am J Med*. 1990;89(5):663–72.
7. Meyts I, Aksentijevich I. Deficiency of adenosine deaminase 2 (DADA2): Updates on the phenotype, genetics, pathogenesis, and treatment. *J Clin Immunol*. 2018;38(5):569–78.
8. Schmidt RE, Grimbacher B, Witte T. Autoimmunity and primary immunodeficiency: Two sides of the same coin? *Nat Rev Rheumatol* . 2018;14(1):7–18.
9. Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* . 2009 May;9(5):313–23.
10. Giovannetti A, Pierdominici M, Mazzetta F, Marziali M, Renzi C, Mileo AM, et al. Unravelling the complexity of T cell abnormalities in common variable immunodeficiency. *J Immunol* . 2007;178(6):3932–43.
11. Resnick ES, Moshier EL, Godbold JH, Cunningham-Rundles C. Morbidity and mortality in common variable immune deficiency over 4 decades. *Blood*. 2012;119(7):1650–8.
12. Ludvigsson JF, Neovius M, Hammarström L. Association between IgA deficiency & other autoimmune conditions: A population-based matched cohort study. *J Clin Immunol*. 2014;34(4):444–51.
13. Jorgensen GH, Gardulf A, Sigurdsson MI, Sigurdardottir ST, Thorsteinsdottir I, Gudmundsson S, et al. Clinical symptoms in adults with selective IgA deficiency: A case-control study. *J Clin Immunol*. 2013;33(4):742–7.
14. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. *Nature*. 2012;489(7415):220–30.
15. Eberl G, Lochner M. The development of intestinal lymphoid tissues at the interface of self and microbiota. *Mucosal Immunol*. 2009 Nov;2(6):478–85.
16. Suzuki K, Meek B, Doi Y, Muramatsu M, Chiba T, Honjo T, et al. Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut. *Proc Natl Acad Sci*. 2004;101(7):1981–6.
17. Bunker JJ, Bendelac A. IgA Responses to Microbiota. *Immunity* . 2018;49(2):211–24. Available from: <https://doi.org/10.1016/j.immuni.2018.08.011>
18. Cerutti A. The regulation of IgA class switching. *Nat Rev Immunol*. 2008;8(6):421–34.
19. * Bunker JJ, Erickson SA, Flynn TM, Henry C, Koval JC, Meisel M, et al. Natural polyreactive IgA antibodies coat the intestinal microbiota. *Science* (80-). 2017;358(6361).
This study investigated the specificity of IgA to microbiota, and showed that the T-cell independent IgA response targets a broad but defined subset of the microbiota
20. Bunker JJ, Flynn TM, Koval JC, Shaw DG, Meisel M, McDonald BD, et al. Innate and Adaptive Humoral Responses Coat Distinct Commensal Bacteria with Immunoglobulin A. *Immunity* . 2015;43(3):541–53.
21. Kawamoto S, Maruya M, Kato LM, Suda W, Atarashi K, Doi Y, et al. Foxp3 + T Cells Regulate Immunoglobulin A Selection and Facilitate Diversification of Bacterial Species Responsible for Immune Homeostasis. *Immunity* . 2014;41(1):152–65.
22. * Moor K, Diard M, Sellin ME, Felmy B, Wotzka SY, Toska A, et al. High-avidity IgA protects the intestine by enchaining growing bacteria. *Nature* . 2017;544(7651):498–502.
High avidity IgA protects against enteric pathogens by enchaining bacteria, thus improving clearance of these pathogens and preventing collateral damage to the host.

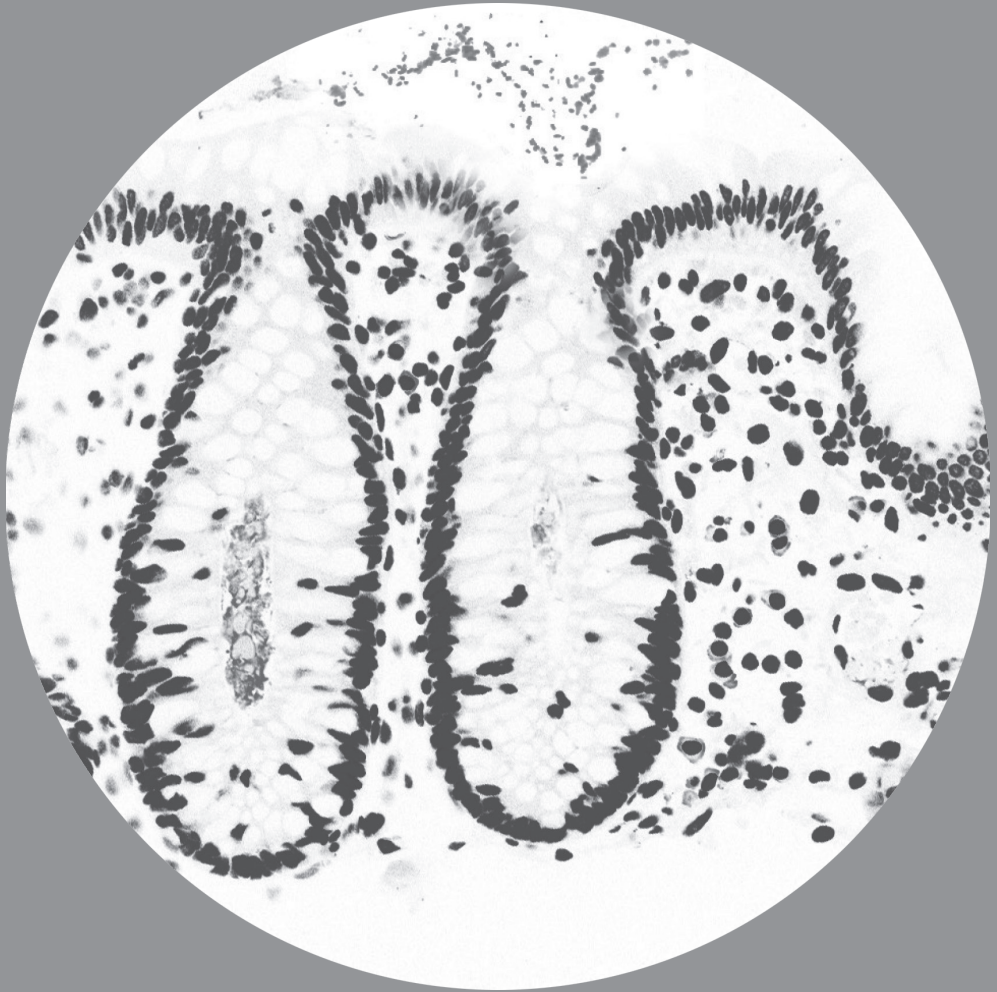
23. Ost KS, Round JL. Communication Between the Microbiota and Mammalian Immunity. *Annu Rev Microbiol.* 2018;72(1):399–422.
24. * Donaldson GP, Ladinsky MS, Yu KB, Sanders JG, Yoo BB, Chou W-C, et al. Gut microbiota utilize immunoglobulin A for mucosal colonization. *Science* (80-). 2018 May 18;360(6390):795–800. **This study showed that select commensals actively elicit IgA coating, which allows them to colonize the enteric mucosa. This led to the insight that IgA can also enhance commensal-host interactions.**
25. Zitomersky NL, Coyne MJ, Comstock LE. Longitudinal analysis of the prevalence, maintenance, and IgA response to species of the order Bacteroidales in the human gut. *Infect Immun.* 2011;79(5):2012–20.
26. Reikvam DH, Derrien M, Islam R, Erofeev A, Grcic V, Sandvik A, et al. Epithelial-microbial cross-talk in polymeric Ig receptor deficient mice. 2012;2959–70.
27. Kubinak JL, Petersen C, Connell RMO, Round JL, Kubinak JL, Petersen C, et al. MyD88 Signaling in T Cells Directs IgA-Mediated Control of the Microbiota to Promote Health Article MyD88 Signaling in T Cells Directs IgA-Mediated Control of the Microbiota to Promote Health. *Cell Host Microbe.* 2015;17(2):153–63.
28. Palm NW, De Zoete MR, Cullen TW, Barry NA, Stefanowski J, Hao L, et al. Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell.* 2014;158(5):1000–10.
29. Sait LC, Galic M, Price JD, Simpfendorfer KR, Diavatopoulos DA, Uren TK, et al. Secretory antibodies reduce systemic antibody responses against the gastrointestinal commensal flora. *Int Immunol.* 2007;19(3):257–65.
30. ** Fadlallah J, El Kafsi H, Sterlin D, Juste C, Parizot C, Dorgham K, et al. Microbial ecology perturbation in human IgA deficiency. *Sci Transl Med.* 2018;10(439):1–16. **The first characterization of the gut microbiota in sIgAD. The authors revealed that sIgAD results in a limited gut dysbiosis, partially because IgM is able to compensate for the lack of IgA. In addition, increased inflammation markers were observed in serum in this cohort.**
31. ** Jorgensen SF, Troseid M, Kummen M, Anmarkrud JA, Michelsen AE, Osnes LT, et al. Altered gut microbiota profile in common variable immunodeficiency associates with levels of lipopolysaccharide and markers of systemic immune activation. *Mucosal Immunol.* 2016 Nov;9(6):1455–65. **The first characterization of the gut microbiota in CVID showed decreased alpha diversity and increased dysbiosis that correlated with plasma endotoxemia and inflammation markers. The decreased alpha diversity was most profound in CVID patients with concurrent deep IgA deficiency.**
32. Auria GD, Peris-bondia F, Dzunková M, Mira A, Collado MC, Latorre A, et al. Active and secreted IgA-coated bacterial fractions from the gut reveal an under-represented microbiota core. *Sci Rep.* 2013;3(3515):1–9.
33. Planer JD, Peng Y, Kau AL, Blanton L V, Ndao IM, Tarr PI, et al. Development of the gut microbiota and mucosal IgA responses in twins and gnotobiotic mice. *Nature.* 2016 Jun;534(7606):263–6.
34. Mirpuri J, Raetz M, Sturge CR, Wilhelm CL, Benson A, Savani RC, et al. Proteobacteria-specific IgA regulates maturation of the intestinal microbiota. *Gut Microbes.* 2013;5(1).
35. Benckert J, Wiedenmann B, Benckert J, Schmolka N, Kreschel C, Zoller MJ. The majority of intestinal IgA + and IgG + plasmablasts in the human gut are antigen- specific. *Journal of Clinical Investigation.* 2011; 121(5): 1946-1955.
36. Elinav E, Strowig T, Kau AL, Henao-mejia J, Thaiss CA, Booth CJ, et al. NLRP6 Inflammasome Regulates Colonic Microbial Ecology and Risk for Colitis. *Cell.* 2011;145(5):745–57.
37. * Okai S, Usui F, Yokota S, Hori-I Y, Hasegawa M, Nakamura T, et al. High-affinity monoclonal IgA regulates gut microbiota and prevents colitis in mice. *Nat Microbiol.* 2016 Sep 4;1(9):16103. **Mouse study showing that pathobiont-specific monoclonal IgA can abrogate inflammatory bowel disease, providing proof of principle for this treatment strategy.**
38. Lönnermark E, Nowrouzian F, Adlerberth I, Ahrné S, Wold A, Friman V. Oral and faecal lactobacilli and their expression of mannose-specific adhesins in individuals with and without IgA deficiency. *Int J Med Microbiol.* 2012;302(1):53–60.
39. Durack J, Kimes NE, Lin DL, Rauch M, McKean M, McCauley K, et al. Delayed gut microbiota development in high-risk for asthma infants is temporarily modifiable by *Lactobacillus* supplementation. *Nat Commun.* 2018;9(1).

40. Perreau M, Vigano S, Bellanger F, Pellaton C, Buss G, Comte D, et al. Exhaustion of bacteria-specific CD4 T cells and microbial translocation in common variable immunodeficiency disorders. *J Exp Med* . 2014;211(10):2033–45.
41. ** Shulzhenko N, Dong X, Vyshenska D, Greer RL, Gurung M, Vasquez-Perez S, et al. CVID enteropathy is characterized by exceeding low mucosal IgA levels and interferon-driven inflammation possibly related to the presence of a pathobiont. *Clin Immunol* . 2018;197(September):139–53. Microbiota characterization of gut biopsies of CVID patients showed that enteropathy is associated with mucosal IgA deficiency, highlighting the association between IgA and intestinal inflammation.
42. Schouwenburg PA Van, Ijspeert H, Pico-knijnenburg I, Patel SY, Burg M Van Der. Identification of CVID Patients With Defects in Immune Repertoire Formation or Specification. *Front Immunol* . 2018;9:2545.
43. ** Fadlallah J, Sterlin D, Fieschi C, Parizot C. Synergistic convergence of microbiota-specific systemic IgG and secretory IgA. *J Allergy Clin Immunol* . 2018;143(4):1575–85. This group showed that healthy donor IgG does not target CVID microbiota as effectively as it does healthy microbiota. The authors suggest that IgG from sIgAD patients may be more beneficial to CVID patients due to better microbiota match.
44. Brenchley JM, Price DA, Schacker TW, Asher TE, Silvestri G, Rao S, et al. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat Med* . 2006;12(12):1365–71.
45. * Rigoni R, Fontana E, Guglielmetti S, Fosso B, D'Erchia AM, Maina V, et al. Intestinal microbiota sustains inflammation and autoimmunity induced by hypomorphic RAG defects . *J Exp Med* . 2016;213(3):355–75. In the mouse equivalent of Omenn syndrome, Rigoni et al show that the intestinal microbiota drives systemic Th1/Th17 mediated inflammation, and that this was linked to IgAD. This provides a potential mechanism for how microbiota may drive immune dysregulation in human PID as well.
46. Rigoni R, Grassi F, Villa A, Cassani B. RAGs and BUGS: An alliance for autoimmunity. *Gut Microbes* . 2016;7(6):503–11.
47. Patron RL, Hartmann CA, Allen S, Griesbach CL, Kosiorek HE, Dibaise JK, et al. Vancomycin Taper and Risk of Failure of Fecal Microbiota Transplantation in Patients with Recurrent Clostridium difficile Infection. *Clin Infect Dis* . 2017;65(7):1214–7.
48. Sacco KA, Pongdee T, Binnicker MJ, Espy M, Pardi D, Khanna S, et al. Presence of immune deficiency increases the risk of hospitalization in patients with norovirus infection. *Diagn Microbiol Infect Dis* . 2018;90(4):300–6.
49. * Fischer M, Kao D, Mehta SR, Martin T, Dimitry J, Keshteli AH, et al. Predictors of Early Failure after Fecal Microbiota Transplantation for the Therapy of Clostridium Difficile Infection: A Multi-center Study. *Am J Gastroenterol* . 2016;111(7):1024–31. Study on FMT treatment for recurrent *C. difficile* infection that included three sIgAD and three CVID patients. Although no data on inflammation was reported, FMT was shown to be safe and effective for *C. difficile* for these patients.
50. Vujkovic-Cvijin I, Rutishauser RL, Pao M, Hunt PW, Lynch S V., McCune JM, et al. Limited engraftment of donor microbiome via one-time fecal microbial transplantation in treated HIV-infected individuals. *Gut Microbes* . 2017;8(5):440–50.
51. ** Jørgensen SF, Macpherson ME, Bjørnstrøm T, Holm K, Kummén M, Rashidi A, et al. Rifaximin alters gut microbiota profile, but does not affect systemic inflammation - a randomized controlled trial in common variable immunodeficiency. *Sci Rep* . 2019;9(1):1–10. After their seminal paper characterizing microbial dysbiosis in CVID , this group attempted to correct the microbiota using antibiotic treatment. No effect on inflammation was observed, suggesting that antibiotics may not be a suitable treatment option.
52. * Robak OH, Heimesaat MM, Kruglov AA, Prepens S, Ninnemann J, Gutbier B, et al. Antibiotic treatment-induced secondary IgA deficiency enhances susceptibility to *Pseudomonas aeruginosa* pneumonia. *J Clin Invest* . 2018;128(8):3535–45. This study investigates the respiratory mucosal IgA response after antibiotic treatment, and shows that in mice, intranasal IgA supplementation may help protect against pathogen colonization.

53. * Langereis JD, Jacobs JFM, de Jonge MI, van Deuren M. Plasma therapy leads to an increase in functional IgA and IgM concentration in the blood and saliva of a patient with X-linked agammaglobulinemia. *J Transl Med* . 2019;17(1):174.
Interesting proof of principle study showing that intravenous IgA suppletion in the form of fresh frozen plasma could be secreted in saliva in an XLA patient.
54. Wand S, Klages M, Kirbach C, Warszawska J, Meybohm P, Zacharowski K, et al. IgM-Enriched Immunoglobulin Attenuates Systemic Endotoxin Activity in Early Severe Sepsis: A Before-After Cohort Study. *PLoS One* . 2016;11(8):e0160907.
55. Behre G, Schedel I, Nentwig B, Wormann B, Essink M, Hiddemann W. Endotoxin concentration in neutropenic patients with suspected gram-negative sepsis: Correlation with clinical outcome and determination of anti-endotoxin core antibodies during therapy with polyclonal immunoglobulin M-enriched immunoglobulins. *Antimicrob Agents Chemother*. 1992;36(10):2139-46.
56. Shmygalev S, Damm M, Knels L, Strassburg A, Wünsche K, Dumke R, et al. IgM-enriched solution BTo86 improves host defense capacity and energy store preservation in a rabbit model of endotoxemia. *Acta Anaesthesiol Scand*. 2016;60(4):502-12.
57. Borleffs J, Schellekens J, Brouwer E, Rozenberg-Arska. Use of an Immunoglobulin M Containing Preparation for Treatment of Two Hypogammaglobulinemic Patients with Persistent Campylobacter jejuni Infection. *Eur J Clin Microbiol Infect Dis*. 1993;12:772-5.
58. Casswall TH, Hammarström L, Bogstedt A, Nord CE, Veress B, Brockstedt U, et al. Oral IgA treatment of chronic non-specific diarrhoea in infants and children. *Gastroenterology*. 2005;108(4):A792.
59. Utay NS, Somasunderam A, Hinkle JE, Petschow BW, Detzel CJ, Somsouk M, et al. Serum Bovine Immunoglobulins Improve Inflammation and Gut Barrier Function in Persons with HIV and Enteropathy on Suppressive ART. *Pathog Immun*. 2019;4(1):124.
60. Florescu DF, Hermsen ED, Kwon JY, Gumeel D, Grant WJ, Mercer DF, et al. Is there a role for oral human immunoglobulin in the treatment for norovirus enteritis in immunocompromised patients? *Pediatr Transplant*. 2011;15(7):718-21.
61. Van De Ven AAJM, Douma JW, Rademaker C, Van Loon AM, Wensing AMJ, Boelens JJ, et al. Pleconaril-resistant chronic parechovirus-associated enteropathy in agammaglobulinaemia. *Antivir Ther*. 2011;16(4):611-4.
62. Jernberg C, Löfmark S, Edlund C, Jansson JK. Long-term impacts of antibiotic exposure on the human intestinal microbiota. *Microbiology*. 2010;156(11):3216-23.
63. De Gunzburg J, Ghoulane A, Ducher A, Le Chatelier E, Duval X, Ruppé E, et al. Protection of the human gut microbiome from antibiotics. *J Infect Dis*. 2018;217(4):628-36.
64. Khoruts A, Rank KM, Newman KM, Viskocil K, Vaughn BP, Hamilton MJ, et al. Inflammatory Bowel Disease Affects the Outcome of Fecal Microbiota Transplantation for Recurrent Clostridium difficile Infection. *Clin Gastroenterol Hepatol* . 2016 Oct;14(10):1433-8.
65. * Wang Y, Wiesnoski DH, Helmink BA, Gopalakrishnan V, Choi K, DuPont HL, et al. Fecal microbiota transplantation for refractory immune checkpoint inhibitor-associated colitis. *Nat Med* . 2018;24(12):1804-8.
Immune checkpoint inhibitor therapy can result in autoimmune side-effects, such as colitis. This study shows that FMT was effective in resolving this colitis, providing evidence for FMT as a potential treatment for T-cell mediated dysregulation. May also be relevant in human CTLA-4 deficiency.

Box 1: microbiota concepts explained

Alpha diversity	within-sample diversity, this reflects the amount and distribution of different taxa within a sample. In gut microbiota, increased alpha diversity is generally associated with better health
Beta diversity	between-sample diversity, this reflects how similar samples are to each other
Symbiont	A resident microbe with no known detrimental effects on the host, may in some cases even be health-promoting
Pathobiont	A resident microbe with pathogenic potential



Chapter 6

Low IgA associated with oropharyngeal microbiota changes and lung disease in primary antibody deficiency

Roos-Marijn Berbers
Firdaus A. A. Mohamed Hoesein
Pauline M. Ellerbroek
Joris M. van Montfrans
Virgil A.S.H. Dalm
P. Martin van Hagen
Fernanda L. Paganelli
Marco C. Viveen
Malbert R.C. Rogers
Pim A. de Jong
Hae-Won Uh
Rob J.L. Willems
Helen L. Leavis

Frontiers in Immunology (2020)
11: 1245

ABSTRACT

Common Variable Immunodeficiency (CVID) and X-linked agammaglobulinemia (XLA) are primary antibody deficiencies characterised by hypogammaglobulinemia and recurrent infections, which can lead to structural airway disease (AD) and interstitial lung disease (ILD). We investigated associations between serum IgA, oropharyngeal microbiota composition and severity of lung disease in these patients.

In this cross-sectional multicentre study we analysed oropharyngeal microbiota composition of 86 CVID patients, 12 XLA patients and 49 healthy controls (HC) using next-generation sequencing of the 16S ribosomal RNA gene. qPCR was used to estimate bacterial load. IgA was measured in serum. High resolution CT scans were scored for severity of AD and ILD.

Oropharyngeal bacterial load was increased in CVID patients with low IgA ($p=0.013$) and XLA ($p=0.029$) compared to HC. IgA status was associated with distinct beta (between-sample) diversity ($p=0.039$), enrichment of (*Alloprevotella*), and more severe radiographic lung disease ($p=0.003$), independently of recent antibiotic use. AD scores were positively associated with *Prevotella*, *Alloprevotella*, and *Selenomonas*, and ILD scores with *Streptococcus* and negatively with *Rothia*.

In clinically stable patients with CVID and XLA, radiographic lung disease was associated with IgA deficiency and expansion of distinct oropharyngeal bacterial taxa. Our findings highlight IgA as a potential driver of upper respiratory tract microbiota homeostasis.

Keywords

Microbiota, Immunoglobulin A, Lung disease, CVID, XLA

INTRODUCTION

The microbiota of the respiratory tract is increasingly recognized as an important driver of respiratory health¹, and has been associated with susceptibility to infection², hypersensitivity reactions such as asthma³, and immune-mediated lung disease such as sarcoidosis⁴. The quantity and composition of the lung microbiota is determined by host defence mechanisms and mucociliary clearance, but originates in the upper respiratory tract, from where it migrates to the lung via micro aspiration or directly via the mucus layer⁵.

Immunoglobulin (Ig)A is thought to be important for the regulation of the microbiota at mucosal surfaces⁶, but the effects of clinical IgA deficiency on respiratory tract microbiota homeostasis in humans remain uninvestigated⁷. Studying the microbiota of patients with primary antibody deficiency such as common variable immunodeficiency (CVID) and X-linked agammaglobulinemia (XLA) can provide insight into the role and importance of the humoral immune system in controlling the microbiota⁸.

The antibody deficiency in CVID is defined as low IgG, and either low IgA or IgM, or both⁹. As a result, some CVID patients have residual IgA production and others are completely IgA deficient, and studying differences between these two CVID subgroups can provide information on the consequences of IgA deficiency, including on the microbiota¹⁰. While CVID can develop later in life and its causes are thought to be multifactorial⁹, XLA is a congenital disease that is the result of a mutation in Bruton's tyrosine kinase, resulting in an early B cell defect and complete humoral immunodeficiency from birth, including IgG, IgA and IgM¹¹.

Despite immunoglobulin G replacement therapy (IgGRT) which limits the recurrence of (respiratory) infections¹², approximately 16-25% of CVID patients develop structural airway disease (AD) and 3-19% develop interstitial lung disease (ILD)¹³⁻¹⁵, causing significant morbidity and mortality in these patients¹³. XLA patients are also treated with IgGRT and remain prone to develop AD, but generally do not develop ILD¹⁶.

AD - which includes bronchiectasis - may seriously compromise pulmonary function and prevention is, therefore, important for a patients' prognosis¹⁷. AD is thought to result from recurrent lower respiratory tract infections, but may also progress in the absence of evident clinical infections¹⁸. ILD in CVID may present as granulomatous lung disease, lymphoid interstitial pneumonia, organizing pneumonia and lymphoproliferative disorders, summarised as granulomatous – lymphocytic interstitial lung disease (GLILD). Causes of ILD in CVID are poorly understood and there is currently no consensus on therapeutic strategies¹⁹. Better understanding of the mechanisms that cause (GL)ILD can contribute to improving clinical care of these patients.

In the gut, low IgA in CVID has been associated with changes of the microbiota, including reduced alpha diversity and expansion of Bacilli and Gammaproteobacteria, which correlated with increased LPS levels in plasma, suggesting increased microbial (product) translocation¹⁰. Whether similar changes occur in the (upper) respiratory tract microbiota in CVID has not been determined.

We hypothesise that IgA deficiency in CVID and XLA may lead to changes in the microbial composition of the respiratory tract, which could contribute to the development of structural lung disease. As a first step in investigating this hypothesis, we characterised the composition of the oropharyngeal microbiota of CVID and XLA patients and correlated this with serum IgA levels and severity of lung disease.

MATERIALS AND METHODS

More detailed materials and methods can be found in the online supplementary material.

Ethics statement

Ethical approval for this study was received from the Medical Ethical Committee of the Erasmus Medical center in Rotterdam, the Netherlands and the University Medical Center Utrecht in Utrecht, the Netherlands (METC: NL40331.078). Written informed consent was obtained from all patients (and in case of minors, their legal guardians) and controls according to the Declaration of Helsinki.

Study population

Patients were age 7 or over, and diagnosed with CVID or XLA according to the European Society for Immunodeficiencies (ESID) criteria²⁰. Partners, friends and family members of patients were included as healthy controls (HC). All CVID and XLA patients received immunoglobulin substitution therapy at time of sampling, with target IgG trough levels of >8.0 g/L. Clinical data was collected from the hospital electronic patient files.

Serum analysis

Serum was collected at time of microbiota sampling, and stored at -80°C until analysis. IgA was measured in serum using a PEG-enhanced immunoturbidimetric method (Atellica CH, Siemens). Very low IgA was defined as serum IgA <0.1 g/L, in order to be consistent with the first gut microbiota study in CVID by Jørgensen et al.¹⁰.

Sample collection and DNA isolation

Oropharyngeal swabs (eSwab, Copan Innovation, Brescia, Italy) were collected by the treating physician or researcher and stored at -80°C the same day. DNA isolation was performed as described by Wyllie et al.²¹.

Bacterial load qPCR

Bacterial load in the oropharyngeal swab samples was estimated using the BactQuant qPCR targeting the bacterial 16S ribosomal RNA (rRNA) gene, as described by Liu et al.²². Primers and probes were ordered from IDT DNA technologies. Forward primer: 5'- CCTACGGDGGCWGCA-3', reverse primer: 5'- GGACTACHVGGGTMTCTAATC-3', probe: (6-FAM/ZEN) 5'-CAGCAGCCGCGGTA-3' (Iowa Black@FQ)²². qPCR was performed using a StepOnePlus RT-PCR system (ThermoFisher).

16S rRNA sequencing and bioinformatics

The 469 basepair V₃ and V₄ hypervariable regions of the 16S rRNA gene were amplified and sequenced using the Illumina MiSeq instrument and Reagent Kit v3 (600-cycle) according to Fadrosch et al.²³. The resulting amplicon pool generated a total of 6.6 million reads. The QIIME2 microbial community analysis pipeline (version 2018.8)²⁴ was used with DADA2²⁵ for sequence variant detection, and SILVA as 16S rRNA reference gene database (SILVA 132)²⁶. Sequencing data has been made available on the European Nucleotide Archive under project code PRJEB34684.

Lung disease scores

High resolution computed tomography (HRCT) volumetric scans were performed in routine diagnostic evaluation every 5 years according to local protocol for COVID and XLA patients. Each scan was scored by a board-certified thoracic radiologist (F.M.H.) for AD and ILD in each lobe using a previously published scoring system^{27,28}. AD was scored as extent and severity of bronchiectasis, airway wall thickening, mucus plugging, tree-in-bud and airtrapping. ILD was scored as extent and severity of opacities, ground glass, septa thickening and lung nodules. The obtained score was normalised by the maximum obtainable score. For one patient who had undergone lobectomy and one with atelectasis of a single lobe, the maximum obtainable score was adapted to exclude the missing lobe. In thirteen cases where expiratory scans were not available, airtrapping could not be evaluated and this element was removed from the score.

Data analysis and statistical methods

All analyses were performed using R 3.2.0²⁹, and made publically available on Gitlab: https://gitlab.com/rberbers/covid_mbiota_oral/. Continuous variables were compared using the Mann-Whitney rank test or a Student's t-test, depending on distribution of the data. Categorical variables were compared using a two-tailed Fisher's exact test. Averages were reported as mean±standard deviation (SD), or as median±interquartile range (IQR). Inverse Simpson index was calculated using the package *vegan*. Principal Component Analysis (PCA) was performed using the *prcomp* function on centered log ratio (CLR) transformed data³⁰. PERMANOVA was used to detect global community differences in PCA using the package *vegan*. Differential abundance testing was performed using ANCOM.2³¹ with Benjamini-Hochberg correction for false discovery rate (FDR) using an alpha of 0.05 as a threshold for significance. All ANCOM analyses were corrected for age and gender.

Correlation between lung scores and microbiota was determined upon the CLR-transformed sequencing data as described above. Linear regression was performed using the function *lm()* and the following model: [lung score] ~ gender + age + [bacterium]

Bootstrapped confidence intervals were generated using the function *boot()* and 1000 iterations. Benjamini Hochberg correction was used to correct for FDR.

RESULTS

Study population characteristics and serum IgA

Oropharyngeal swabs were collected from 86 CVID patients, 49 HC and 12 XLA patients (table 1) in two academic hospitals in Rotterdam and Utrecht, the Netherlands. Household members of patients were included as healthy controls. CVID patients were categorised as having very low IgA (IgA < 0.1 g/L, CVID-IgA, n=36) or residual IgA serum levels (IgA > 0.1 g/L, CVID+IgA, n=50). Serum IgA could be determined in 39/49 HC, and was normal (IgA > 0.7 g/L) for all HC tested. There were fewer males in the HC than in both CVID groups (proportions of male participants 31% HC, 50% CVID+IgA, 58% CVID-IgA), and the HC were older (mean age \pm standard deviation SD, HC 42 \pm 12, CVID+IgA 34 \pm 19 and CVID-IgA 38 \pm 15). Naturally, all XLA patients were male. There was more recent antibiotic use in XLA than in CVID (58% versus 36% and 28% in the CVID groups).

One patient with XLA suffered from juvenile idiopathic arthritis, while 18% of CVID+IgA and 39% of CVID-IgA suffered from autoimmune disease. Serum IgA levels were lower in CVID patients with (n=23) than without (n=63) autoimmune disease (mean \pm SD 0.21 \pm 0.24 g/L and 0.55 \pm 0.60 g/L, respectively, p=0.010, data not shown).

Increased oropharyngeal bacterial load in CVID and XLA

Bacterial load for each oropharyngeal swab sample was estimated by qPCR with 16S rRNA gene-based primers (figure 1). There was a clear association between the median bacterial loads in oropharyngeal swabs and serum IgA deficiency: median bacterial load increased gradually as patients were more profoundly IgA impaired (median \pm interquartile range IQR: HC 3.1x10⁶ \pm 38.4x10⁶, CVID+IgA 15.0x10⁶ \pm 132.2x10⁶, CVID-IgA 33.4x10⁶ \pm 154.3x10⁶, XLA 50.5x10⁶ \pm 172.2x10⁶ copies of the 16S rRNA gene). As bacterial loads may be influenced by antibiotic use, we repeated analyses with patients who did not use antibiotics 3 months prior to sampling (supplementary figure 2), which yielded a broadly similar pattern.

Alpha diversity and community structure differ in absence of IgA

After 16S rRNA amplicon sequencing, 14 samples had insufficient sequencing coverage (<8000 reads per sample) and were excluded from data analysis, leaving 41 HC, 81 CVID (48 CVID+IgA, 33 CVID-IgA) and 11 XLA samples (for an overview, please see figure 1 and table 1 of the supplementary materials). The overall top 10 most abundant bacterial genera were *Streptococcus*, *Actinomyces*, *Veillonella*, *Rothia*, *Prevotella*, *Gemella*, *Leptotrichia*, *Haemophilus*, *Neisseria* and *Megasphaera* (figure 2A).

Alpha diversity, a measure for the richness and evenness of a sample (expressed here as inverse Simpson index, figure 2B), followed the same pattern as for bacterial loads; the lower the IgA, the higher the inverse Simpson index (median \pm IQR: HC 3.28 \pm 3.37, CVID+IgA 3.71 \pm 3.12, CVID-IgA 4.95 \pm 5.28, XLA 6.46 \pm 3.78). Similar results were observed in patients who were not using antibiotics at least three months prior to sampling (supplementary figure 3A).

Table 1 Study population characteristics. HC: healthy control, CVID: common variable immunodeficiency, SD: standard deviation, XLA: X-linked agammaglobulinemia GLILD: granulomatous - lymphocytic interstitial lung disease. Serum IgA below the limit of detection was reported as 0.070 g/L.

	HC	CVID +IgA (IgA >0.1 g/L)	CVID -IgA (IgA <0.1 g/L)	XLA
Total N	49	50	36	12
Age: mean ± SD	42 ± 12	34 ± 19	38 ± 14	22 ± 15
Male % (N)	31% (15/49)	50% (25/50)	58% (21/36)	100% (12/12)
Medication use during 3 months prior to sampling: % (N)				
Antibiotics	0% (0/49)	36% (18/50)	28% (10/36)	58% (7/12)
Immune suppressive therapy	0% (0/49)	12% (6/50)	14% (5/36)	8% (1/12)
Clinical phenotype: % (N)				
Any inflammatory complication	0% (0/49)	30% (15/50)	67% (24/36)	8% (1/12)
Autoimmune disease	0% (0/49)	18% (9/50)	39% (14/36)	8% (1/12)
GLILD (clinical diagnosis)	0% (0/49)	6% (3/50)	8% (3/36)	0% (0/12)
Granulomatous disease other	0% (0/49)	2% (1/50)	6% (2/36)	0% (0/12)
Enteritis	0% (0/49)	14% (7/50)	28% (10/36)	0% (0/12)
Malignancy	0% (0/49)	6% (3/50)	3% (1/36)	0% (0/12)
IgA status:				
Serum IgA levels available (N)	80% (39/49)	100% (50/50)	100% (36/36)	100% 12/12
% serum IgA low (<0.1 g/L)	0% (0/39)	0% (0/50)	100% (36/36)	100% (12/12)
Serum IgA mean ± SD in g/L	2.06 ± 0.78	0.74 ± 0.57	0.07 ± 0.00	0.07 ± 0.00

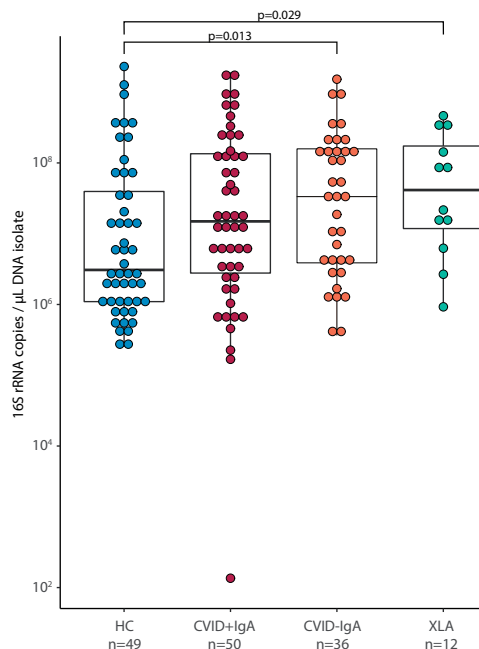


Figure 1 Bacterial load in oropharyngeal swab samples as determined by qPCR for copies of the 16S rRNA gene in DNA isolates from oropharyngeal swabs in healthy controls (HC, n=49), CVID +IgA (n=50), CVID -IgA (n=36) and X-linked agammaglobulinemia (XLA, n=12). The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within 1.5×interquartile range. Statistical test: Mann-Whitney U test.

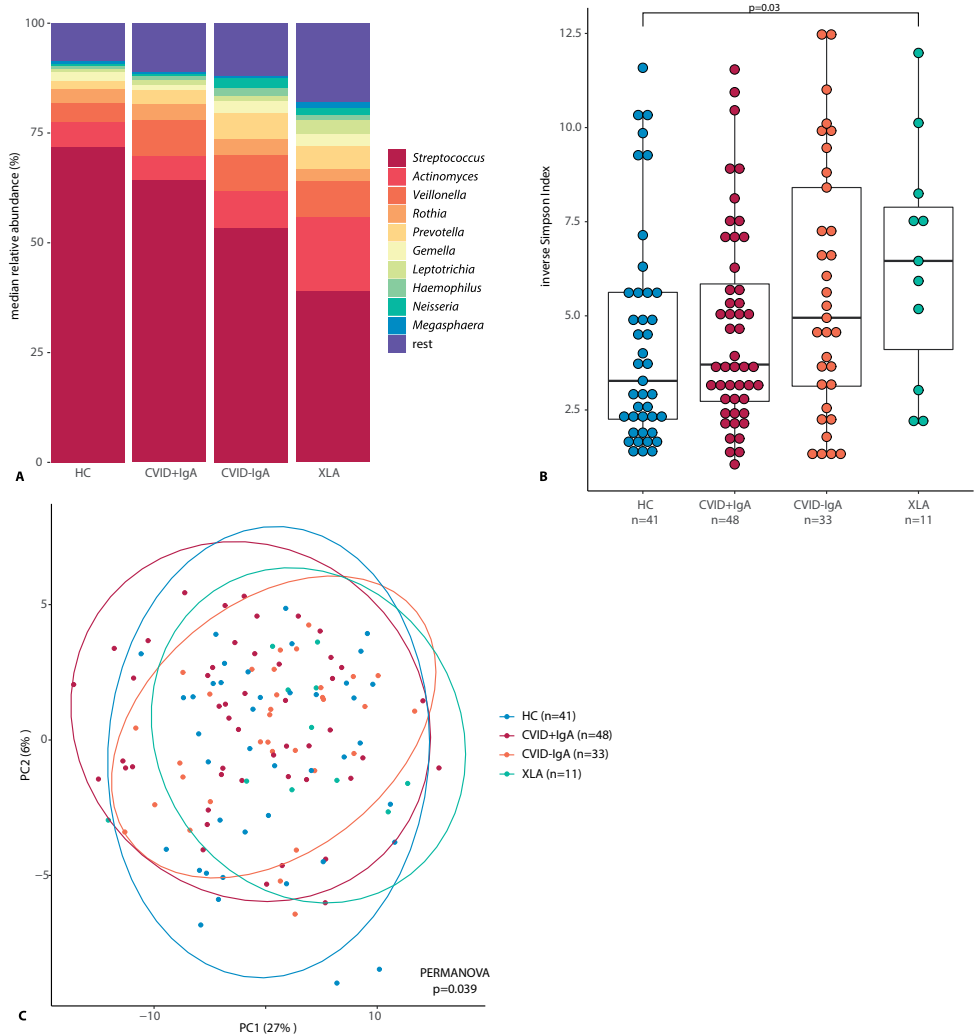


Figure 2

A Median relative abundance of top 10 most abundant genera determined by 16S rRNA sequencing of healthy controls (HC, n=41), CVID +IgA (n=48), CVID -IgA (n=33) and XLA (n=11).

B Alpha diversity of the same samples described in **A** as measured by inverse Simpson's index on 16S rRNA sequencing data. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within 1.5×interquartile range. Statistical test: Mann-Whitney U test.

C Principal component analysis of centered log ratio (CLR)-transformed family level 16S rRNA sequencing data of the same samples described in **A**. Ellipses indicate 95% confidence intervals. PERMANOVA using Euclidean distance on CLR data

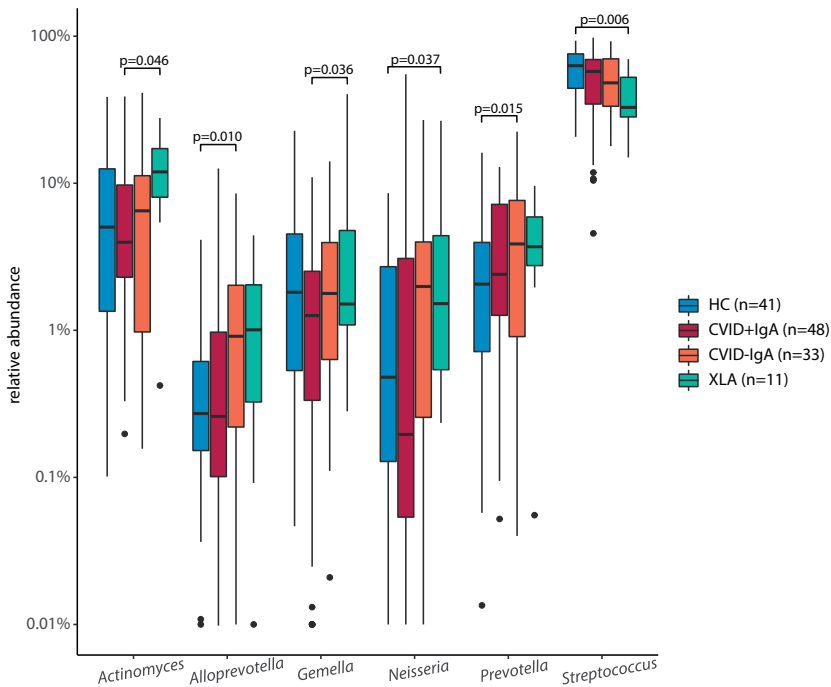


Figure 3 Differentially abundant taxa in 16S rRNA sequencing of healthy controls (HC, n=41), CVID +IgA (n=48), CVID -IgA (n=33) and XLA (n=11).

Statistics: ANCOM corrected for age and gender and Benjamini-Hochberg correction for false discovery rate. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within 1.5×interquartile range.

Also the overall community structure (beta diversity) differed when grouping patients by IgA status (figure 2C; $p=0.039$). After excluding patients with recent antibiotic use (supplementary figure 3B), overall community structure remained significantly different between the groups ($p=0.039$).

Expansion of Prevotellaceae bacteria associated with IgA deficiency

In order to further examine the differences in oropharyngeal microbiota in these patients, we next determined which bacterial genera were differentially abundant (figure 3). Compared to HC, CVID-IgA patients had a higher relative abundance of two genera belonging to the Prevotellaceae family; *Prevotella* and *Alloprevotella* ($p=0.015$ and $p=0.010$, respectively). The same pattern of higher relative abundance of Prevotellaceae bacteria in more profound IgA deficiency was observed in patients without recent antibiotic use (supplementary figure 4). In XLA, relative abundance of *Prevotella* and *Alloprevotella* was similar to CVID-IgA but was not significant compared to HC, probably due to smaller sample size of the XLA group.

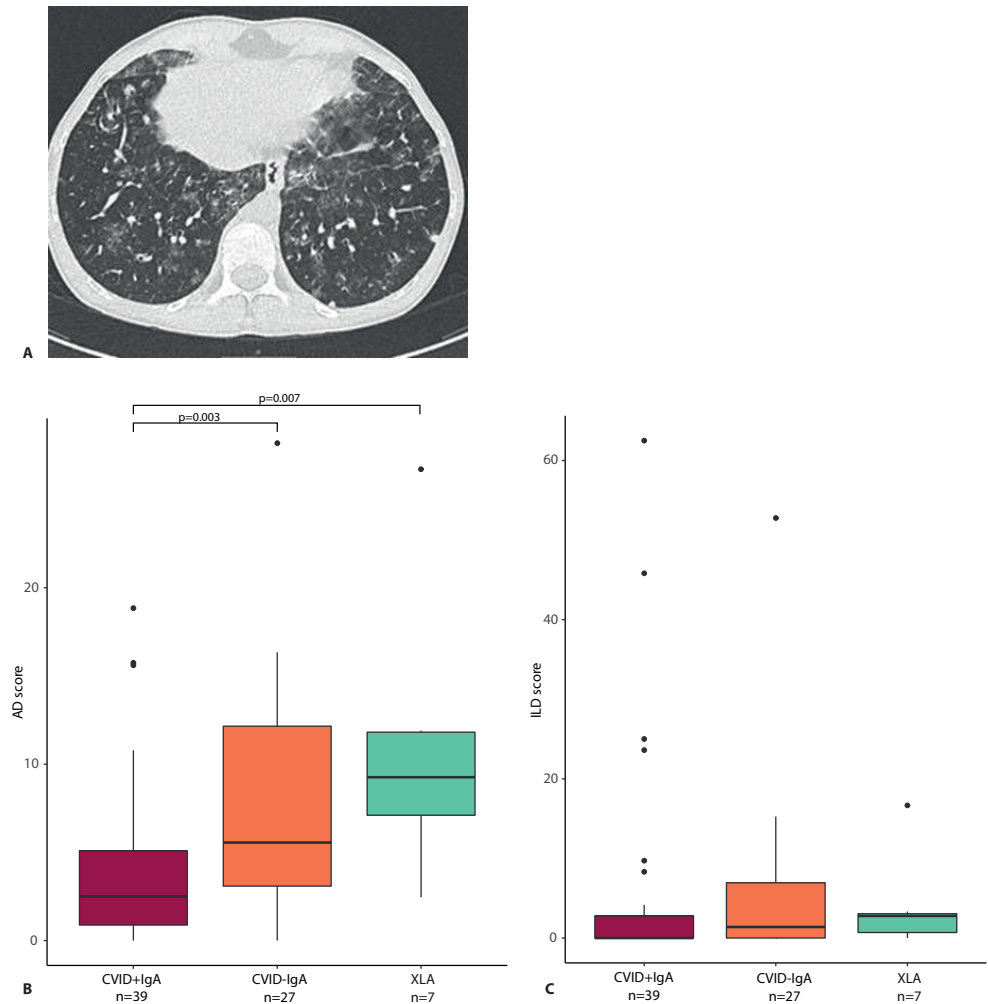


Figure 4

A Axial CT image: This 15-year old male patient with CVID had a total airway disease (AD) score of 13, and an interstitial lung disease (ILD) score of 48. Opacities, ground glass, septa thickening and lung nodules were observed, especially in the lower lobes.

B airway disease (AD) scores in CVID +IgA (n=39), CVID -IgA (n=27), and XLA (n=7).

C interstitial lung disease (ILD) scores in the same samples as in **B**.

The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within 1.5×interquartile range. Statistics: Mann-Whitney U test.

Table 2 Linear regression with airway disease (AD) or interstitial lung disease (ILD) scores as dependent variable, and species level microbial sequencing data as independent variables, using data from 73 patients (39 CVID +IgA, 27 CVID -IgA and 7 XLA). Age and gender were included as covariates. Beta coefficient (beta) was bootstrapped to generate 95% confidence intervals (95%CI). Only taxa with $p < 0.05$ and 95% confidence intervals that do not span zero are reported.

AD score	Beta	p value	95%CI (Beta)	R ²
age	0.057	0.189	-0.005, 0.146	
gender	-0.062	0.966	-2.879, 2.266	
<i>Prevotella</i>	0.769	0.027	0.176, 1.654	0.099
<i>Alloprevotella</i>	0.682	0.029	0.004, 1.262	0.097
<i>Selenomonas</i>	0.650	0.023	0.166, 1.168	0.103
<i>Megasphaera</i>	0.551	0.032	0.142, 1.031	0.095
<i>Fusobacterium</i>	-0.704	0.020	-1.252, -0.048	0.107
<i>Johnsonella</i>	-1.577	0.035	-2.472, -0.949	0.093
ILD score	Beta	p value	95%CI (Beta)	R ²
age	0.054	0.562	-0.142, 0.153	
gender	2.180	0.483	-3.581, 9.215	
<i>Streptococcus</i>	2.173	0.001	0.257, 4.604	0.176
<i>Rothia</i>	-1.856	0.044	-4.465, -0.434	0.082

Other bacteria that were consistently differentially expressed in both the full cohort and the subgroups without recent antibiotic use were *Streptococcus* (decreased in XLA vs HC, $p = 0.006$), and *Actinomyces* (increased in XLA vs CVID+IgA, $p = 0.046$).

Radiographic lung disease correlated with IgA status, *Prevotella*, *Alloprevotella*, *Selenomonas*, and *Streptococcus*

Next, we investigated whether the observed oropharyngeal microbiota differences in IgA deficient patients were associated with radiographic lung disease. From the group of patients in whom microbial community composition of oropharyngeal swabs was analysed, high resolution chest CT scans were available for a total of 73 patients (39 CVID+IgA, 27 CVID-IgA and 7 XLA). Scans were performed for routine clinical follow-up, and mean time between scan and oropharyngeal swab was 1 year and 56 days (± 1 year). For a representative scan, please see figure 4A. Mean AD score in this population was 6.1 ± 6.0 with 27% of patients scoring below 2 and 23% scoring 10 or higher. Mean ILD score was 5.2 ± 11.4 with 64% scoring below 2 and 14% scoring 10 or higher.

AD scores were higher in CVID patients with low IgA than in CVID+IgA (figure 4B; CVID-IgA and XLA vs CVID+IgA $p = 0.003$ and $p = 0.007$, respectively). No significant differences were observed for ILD score versus IgA status (figure 4C). Relationships between AD or ILD scores and bacterial load or alpha diversity were not detected here (supplementary figure 5 A-C).

In linear regression corrected for age and gender including all CVID patients for whom scans were available ($n = 66$, see table 2), *Prevotella* ($p = 0.027$, beta 95%CI 0.176,

1.654), *Alloprevotella* ($p=0.029$, beta 95%CI 0.004, 1.262), *Selenomonas* ($p=0.023$, beta 95%CI 0.166, 1.168) and *Megasphaera* ($p=0.032$, beta 95%CI 0.142, 1.031) positively correlated with AD score, while *Fusobacterium* ($p=0.020$, beta 95%CI -1.252, -0.048) and *Johnsonella* ($p=0.035$, beta 95%CI -2.472, -0.949) correlated negatively (table 2). *Streptococcus* ($p=0.001$, beta 95%CI 0.257, 4.604) positively correlated with ILD scores, and *Rothia* was negatively associated with ILD score ($p=0.044$, beta 95%CI -4.465, -0.434). None of the reported bacteria were significant after correction for false discovery rate, but all were stably detected after bootstrapping of samples to exclude effects due to outliers in the data.

DISCUSSION

In this study we determined bacterial load and community composition of the oropharyngeal microbiota in CVID and XLA patients. We observed that bacterial load, alpha diversity and relative abundance of bacteria from the Prevotellaceae family were consistently increased in patients with more profound IgA deficiency, specifically CVID-IgA and XLA. Moreover, IgA deficiency and expansion of Prevotellaceae bacteria were associated with lung disease in these patients.

While IgA deficiency is one of the diagnostic criteria for CVID, some patients have residual IgA production, and others are almost completely IgA deficient. In XLA, no immunoglobulins are produced at all from birth. Despite IgGRT titrated to protect patients clinically against infection, oropharyngeal bacterial load was increased in CVID and XLA, and we observed a trend of increasing bacterial loads in patients with more profound IgA deficiency (HC < CVID+IgA < CVID-IgA < XLA). While control of bacterial load is most likely a multifactorial process, our findings indicate that IgA may contribute to limiting the total amount of colonising bacteria in the upper respiratory tract in these patients.

An additional consequence of increased bacterial load in CVID and XLA concerns interpretation of relative abundance data³². As 16S rRNA gene based microbial profiling provides only compositional information, the increased total bacterial load in CVID and XLA means that any reported increase of relative abundance in these groups (such as for *Alloprevotella* and *Prevotella*) is expected to reflect a much greater increase in absolute bacterial numbers. The reverse may be the case for bacteria for which a relative decrease was reported (such as for instance in *Streptococcus*).

Concurrent with bacterial load, alpha diversity increased as patients were more profoundly IgA deficient, suggesting that IgA also limits the colonization of more different bacterial taxa. In gut microbiota high alpha diversity is generally associated with health, but this seems not to always apply to respiratory microbiota, as increased alpha diversity of the respiratory tract has been reported in other disease states such as asthma³³. Increased alpha diversity indicates a more complex alteration of the microbial composition rather than the outgrowth of a few pathobionts, which would result in decreased diversity as has been observed in infections⁴ and acute exacerbations of non-CVID/XLA bronchiectasis³⁴.

Prevotellaceae genera were significantly expanded in CVID with low IgA levels and

XLA. *Prevotella* are known mucus degraders³⁵ and have been associated with gut microbiome changes and immune-mediated disease³⁶. These bacteria are reportedly IgA coated in the gut³⁷ and were found to be increased in gut microbiota of patients with IgA deficiency and concurrent Th17 skewing³⁸. In the lung, *Prevotella*-derived outer membrane vesicles are thought to drive inflammation and fibrosis through the induction of Th17 responses upon Toll-like receptor (TLR)-2 activation³⁹. *Prevotella* may also drive inflammation indirectly by degrading mucins, as intact mucins can dampen innate immune responses by shielding bacterial ligands from TLRs⁴⁰. *Prevotella* has been associated with (exacerbations of) non-CVID bronchiectasis in other studies⁴¹.

Relative abundance of *Prevotella* positively correlated with AD scores. This correlation was not significant after correction for false discovery rate, and additional studies will be needed to confirm the link between *Prevotella* (and the other detected bacteria) and airway disease. Increased relative abundance of *Streptococcus* was associated with higher ILD scores. Specific *Streptococcus* species have previously been linked with ILD progression in non-CVID patients in a prospective cohort study⁴². Abundance of *Rothia* was negatively associated with ILD scores, suggesting a potential protective role. *Rothia* is described as part of the core microbiota of the oropharynx⁴³, and has also been found to be associated occurrence of pneumonia in elderly patients².

Strengths of this study are the integration between culture-independent microbiological community profiling, clinical immunology and pulmonary radiology. Limitations of this study include the cross-sectional nature, which does not allow for cause-effect distinctions (does IgA cause microbiota changes and do these cause lung disease, or are these bystander effects of a separate process?). There was a time delay between the oropharyngeal swabs, which were taken specifically for this study, and the chest CT scans, which are performed routinely every five years for clinical care in our clinics. Our group has previously shown that progression of radiographic AD and ILD scores over a 3- to 5-year follow-up period is very limited in CVID patients²⁸, and therefore the mean time between scan and swab of approximately one year in this study was deemed acceptable. We elected to use radiological evaluation in this study as previous work by our group has shown that the use of chest CT scans evaluated by a trained pulmonary radiologist is a superior predictor of early pulmonary abnormalities compared to pulmonary function testing in CVID⁴⁴. A potential confounder in this study is use of antibiotics by CVID patients. Recent antibiotic use was recorded, and sensitivity analyses excluding all patients who had recently used antibiotics did not yield different insights as compared to analyses using the full cohort. However, an effect of long-term antibiotic use in this cohort cannot be excluded.

To conclude, we demonstrated that – despite IgGRT and independent of recent antibiotic use – patients with primary antibody deficiency carry an increased bacterial load in the upper respiratory tract, and have compositional changes of the oropharyngeal microbiota related to IgA deficiency. These compositional changes were associated with the radiographic presence of airway disease and interstitial lung disease. We speculate that IgA deficiency-induced changes in the microbiota composition of the respiratory tract may

cause low-grade inflammation through increased microbial challenge, mucus degradation, and Th-17 skewing, resulting in inflammation-driven airway remodeling, and in some cases even interstitial lung disease.

While the oropharyngeal microbiota has been found to partially overlap with pulmonary microbiota^{45,46}, they are still distinct communities with important differences in bacterial load and composition^{45,46}. Therefore, while our findings regarding oropharyngeal microbiota load and composition may reflect changes in the pulmonary microbiota indirectly through bacterial seeding of the lower respiratory tract from the oropharynx or similar consequences of IgA deficiency, further studies are required to show a direct, temporal and spatial relationship between IgA, microbiota and lung disease.

Direct interrogation of the lung microbiota, for instance through bronchial-alveolar lavage, can provide more information about local processes contributing to the development of lung disease. In addition, IgA-based therapeutic interventions in mouse models of antibody deficiency may help answer cause-effect questions, as well as provide first steps for better treatment in patients.

Conflict of Interest and Funding statement

This study was funded by the Wilhelmina Children's Hospital Fund, the UMC Utrecht Infection & Immunity focus area, and the Utrecht Exosome Hub.

MvH reports research grants and personal fees from Shire/Takeda and CSL Behring. VD reports research grants and personal fees from Shire/Takeda, Griffiths, Actelion, Novartis and CSL Behring. JvM reports personal fees from Shire/Takeda. PdJ reports research grants from Philips Healthcare. HL reports research grants and personal fees from Shire/Takeda.

Author Contributions

Study design and implementation: HL, RB, PE, JvM, VD and MvH. Radiological assessment: FMH, PdJ. Wet lab work was performed by RB and MV. Bioinformatics and statistical expertise: MR, FP, HWU. Data analysis and interpretation: RB, HL, RW, FP. All authors substantially contributed to the acquisition, analysis or interpretation of data for the manuscript and approved the final manuscript.

Acknowledgments

We would like to acknowledge and thank prof. dr. Marc Bonten (dept. of medical microbiology, University Medical Center Utrecht) for his thoughtful review of the manuscript.

REFERENCES

1. de Steenhuijsen Piters WAA, Sanders EAM, Bogaert D. The role of the local microbial ecosystem in respiratory health and disease. *Phil Trans R Soc B* (2015) **370**:20140294.
2. de Steenhuijsen Piters W, Huijskens EGW, Wyllie AL, Biesbroek G, Bergh MR Van Den, Veenhoven RH, Wang X, Bogaert D. Dysbiosis of upper respiratory tract microbiota in elderly pneumonia patients. *ISME J* (2016) **10**:97–108. doi:10.1038/ismej.2015.99
3. Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, Davies J, Ervine A, Poulter L, Pachter L, et al. Disordered Microbial Communities in Asthmatic Airways. *PLoS One* (2010) **5**:e8578. doi:10.1371/journal.pone.0008578
4. Zimmermann A, Knecht H, Häsler R, Zissel G, Gaede KI, Hofmann S, Nebel A, Müller-quernheim J, Schreiber S, Fischer A. Atopobium and Fusobacterium as novel candidates for sarcoidosis-associated microbiota. *Eur Respir J* (2017) **50**:1600746. doi:10.1183/13993003.00746-2016
5. Salisbury ML, Han MK, Dickson RP, Molyneaux PL, Arbor A, Heart N. The microbiome in interstitial lung disease: from pathogenesis to treatment target. *Curr Opin Pulm Med* (2017) **23**:404–410. doi:10.1097/MCP.000000000000399.The
6. Bunker JJ, Bendelac A. IgA Responses to Microbiota. *Immunity* (2018) **49**:211–224. doi:10.1016/j.immuni.2018.08.011
7. Berbers RM, Franken IA, Leavis HL. Immunoglobulin A and microbiota in primary immunodeficiency diseases. *Curr Opin Allergy Clin Immunol* (2019) **19**:563–570. doi:10.1097/ACI.0000000000000581
8. Pellicciotta M, Rigoni R, Liana E, Holland SM, Villa A, Cassani B. The microbiome and immunodeficiencies : Lessons from rare diseases. *J Autoimmun* (2019) **11**–17. doi:10.1016/j.jaut.2019.01.008
9. Bonilla FA, Barlan I, Chapel H, Costa-Carvalho BT, Cunningham-Rundles C, de la Morena MT, Espinosa-Rosales FJ, Hammarström L, Nonoyama S, Quinti I, et al. International Consensus Document (ICON): Common Variable Immunodeficiency Disorders. *J Allergy Clin Immunol Pract* (2016) **4**:38–59. doi:10.1016/j.jaip.2015.07.025
10. Jorgensen SF, Troseid M, Kummen M, Anmarkrud JA, Michelsen AE, Osnes LT, Holm K, Hoivik ML, Rashidi A, Dahl CP, et al. Altered gut microbiota profile in common variable immunodeficiency associates with levels of lipopolysaccharide and markers of systemic immune activation. *Mucosal Immunol* (2016) **9**:1455–1465. doi:10.1038/mi.2016.18
11. Vetrie D, Vořechovský I, Sideras P, Holland J, Davies A, Flinter F, Hammarström L, Kinnon C, Levinsky R, Bobrow M, et al. The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases. *Nature* (1993) **361**:226–233. doi:10.1038/361226a0
12. de Gracia J, Vendrell M, Alvarez A, Pallisa E, Rodrigo M, de la Rosa D, Mata F, Andreu J, Morrel F. Immunoglobulin therapy to control lung damage in patients with common variable immunodeficiency. *Int Immunopharmacol* (2004) **4**:745–753. doi:10.1016/j.intimp.2004.02.011
13. Bates CA, Ellison MC, Lynch DA, Cool CD, Brown KK, Routes JM. Granulomatous-lymphocytic lung disease shortens survival in common variable immunodeficiency Pulmonary disorders. *J Allergy Clin Immunol* (2004) **5**:415–421. doi:10.1016/j.jaci.2004.05.057
14. Resnick ES, Moshier EL, Godbold JH, Cunningham-Rundles C. Morbidity and mortality in common variable immune deficiency over 4 decades. *Blood* (2012) **119**:1650–1658. doi:10.1182/blood-2011-09-377945.The
15. Maarschalk-Ellerbroek LJ, Hoepelman AIM, Van Montfrans JM, Ellerbroek PM. The spectrum of disease manifestations in patients with common variable immunodeficiency disorders and partial antibody deficiency in a university hospital. *J Clin Immunol* (2012) **32**:907–921. doi:10.1007/s10875-012-9671-6
16. Jolles S, Carne E, Brouns M, El-Shanawany T, Williams P, Marshall C, Fielding P. FDG PET-CT imaging of therapeutic response in granulomatous lymphocytic interstitial lung disease (GLILD) in common variable immunodeficiency (CVID). *Clin Exp Immunol* (2017) **187**:138–145. doi:10.1111/cei.12856
17. Hampson FA, Chandra A, Sreaton NJ, Condliffe A, Kumararatne DS, Exley AR, Babar JL. Respiratory disease in common variable immunodeficiency and other primary immunodeficiency disorders. *Clin Radiol* (2012) **67**:587–595. doi:10.1016/j.crad.2011.10.028

18. Quinti I, Soresina A, Spadaro G, Martino S, Donnanno S, Agostini C, Claudio P, Franco D, Pesce AM, Borghese F, et al. Long-Term Follow-Up and Outcome of a Large Cohort of Patients with Common Variable Immunodeficiency. *J Clin Immunol* (2007) 27:308–316. doi:10.1007/s10875-007-9075-1
19. Hurst JR, Verma N, Lowe D, Baxendale HE, Jolles S, Kelleher P, Longhurst HJ, Patel SY, Renzoni EA, Sander CR, et al. British Lung Foundation/United Kingdom Primary Immunodeficiency Network Consensus Statement on the Definition, Diagnosis, and Management of Granulomatous-Lymphocytic Interstitial Lung Disease in Common Variable Immunodeficiency Disorders. *J Allergy Clin Immunol Pract* (2017) 5:938–945. doi:10.1016/j.jaip.2017.01.021
20. Immunodeficiencies ES for. Diagnostic Criteria PID. (2019) Available at: <https://esid.org/Education/Diagnostic-Criteria-PID>
21. Wyllie AL, Chu MLJN, Jansen MD, Ende A Van Der, Bogaert D, Sanders EAM, Trzcin K. Streptococcus pneumoniae in Saliva of Dutch Primary School Children. *PLoS One* (2014) 9:e102045. doi:10.1371/journal.pone.0102045
22. Liu CM, Aziz M, Kachur S, Hsueh P, Huang Y, Keim P, Price LB. BactQuant : An enhanced broad-coverage bacterial quantitative real-time PCR assay. *BMC Microbiol* (2012) 12:56.
23. Fadrosch DW, Ma B, Gajer P, Sengamalay N, Ott S, Brotman RM, Ravel J. An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* (2014) 2:1–7.
24. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* (2019) 37:852–857. doi:10.1038/s41587-019-0209-9
25. Callahan BJ, Mcmurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2 : High-resolution sample inference from Illumina amplicon data. *Nat Methods* (2016) 13:581–587. doi:10.1038/nmeth.3869
26. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Glo FO, Yarza P. The SILVA ribosomal RNA gene database project : improved data processing and web-based tools. *Nucleic Acids Res* (2013) 41:590–596. doi:10.1093/nar/gks1219
27. Ven AAJM Van De, Montfrans JM Van. A CT Scan Score for the Assessment of Lung Disease in Children With Common Variable Immunodeficiency Disorders. *Chest* (2010) 138:371–379. doi:10.1378/chest.09-2398
28. Janssen WJ., Mohamed Hoesein F, van de Ven AJM, Maarschalk-Ellebrouck LJ, van Royen F, de Jong PA, Sanders E, van Montfrans JM, Ellebrouck PM. IgG trough levels and progression of pulmonary disease in pediatric and adult common variable immunodeficiency disorder patients To. *J Allergy Clin Immunol* (2017) 140:303–306. doi:10.1016/j.jaci.2016.11.050
29. R Foundation for Statistical Computing, Vienna A. R: A language and environment for statistical computing. (2018) Available at: <https://www.r-project.org/>
30. Aitchison J. *The statistical analysis of compositional data: Monographs on statistics and applied probability*. Chapman & Hall Ltd., London (1986).
31. Kaul A, Mandal S, Davidov O, Peddada SD. Analysis of Microbiome Data in the Presence of Excess Zeros. *Front Microbiol* (2017) 8:2114. doi:10.3389/fmicb.2017.02114
32. Gloor GB, Macklaim JM, Pawlowsky-glahn V, Egozcue JJ. Microbiome Datasets Are Compositional : And This Is Not Optional. *Front Microbiol* (2017) 8:1–6. doi:10.3389/fmicb.2017.02224
33. Huang YJ, Nelson CE, Brodie EL, Desantis TZ, Baek MS, Liu J, Woyke T, Allgaier M, Bristow J, Wiener-kronish JP, et al. Airway microbiota and bronchial hyperresponsiveness in patients with suboptimally controlled asthma. *J Allergy Clin Immunol* (2011) 127:372–381.e3. doi:10.1016/j.jaci.2010.10.048
34. Tunney MM, Einarsson GG, Wei L, Drain M, Klem ER, Cardwell C, Ennis M, Boucher RC, Wolfgang MC, Elborn JS. Lung Microbiota and Bacterial Abundance in Patients with Bronchiectasis when Clinically Stable and during Exacerbation. *Am J Crit Care Med* (2013) 187:1118–1126. doi:10.1164/rccm.201210-1937OC
35. Wright DP, Rosendale DI, Robertson AM. Prevotella enzymes involved in mucin oligosaccharide degradation and evidence for a small operon of genes expressed during growth on mucin. *FEMS Microbiol Lett* (2000) 190:73–79.

36. Scher JU, Sczesnak A, Longman RS, Segata N, Ubeda C, Bielski C, Rostron T, Cerundolo V, Pamer EG, Abramson SB, et al. Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *Elife* (2013) 2:e01202. doi:10.7554/eLife.01202
37. Palm NW, De Zoete MR, Cullen TW, Barry NA, Stefanowski J, Hao L, Degnan PH, Hu J, Peter I, Zhang W, et al. Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* (2014) 158:1000–1010. doi:10.1016/j.cell.2014.08.006
38. Fadlallah J, El Kafsi H, Sterlin D, Juste C, Parizot C, Dorgham K, Autaa G, Gouas D, Almeida M, Lepage P, et al. Microbial ecology perturbation in human IgA deficiency. *Sci Transl Med* (2018) 10:1–16. doi:10.1126/scitranslmed.aan1217
39. Yang D, Chen X, Wang J, Wu M, Yang D, Chen X, Wang J, Lou Q, Lou Y, Li L, et al. Dysregulated Lung Commensal Bacteria Drive Interleukin-17B Production to Promote Pulmonary Fibrosis through Their Outer Membrane Vesicles. *Immunity* (2019) 50:692–706. doi:10.1016/j.immuni.2019.02.001
40. Putten JPM Van, Strijbis K. Transmembrane Mucins : Signaling Receptors at the Intersection of Inflammation and Cancer. *J Innate Immun* (2017) 9:281–299. doi:10.1159/000453594
41. Rogers GB, Gast CJ Van Der, Cuthbertson L, Thomson SK, Bruce KD, Martin ML, Serisier DJ. Clinical measures of disease in adult non-CF bronchiectasis correlate with airway microbiota composition. *Thorax* (2013) 68:731–737.
42. Han MK, Zhou Y, Murray S, Tayob N, Lama VN, Moore BB, White ES, Kevin R, Huffnagle GB, Martinez FJ. Association Between Lung Microbiome and Disease Progression in IPF: a prospective cohort study. *Lancet Respir Med* (2015) 2:548–556. doi:10.1016/S2213-2600(14)70069-4. Association
43. Ling Z, Kong J, Jia P, Wei C, Wang Y, Pan Z, Huang W, Li L, Chen H, Xiang C. Analysis of Oral Microbiota in Children with Dental Caries by PCR-DGGE and Barcoded Pyrosequencing. *Microb Ecol* (2010) 60:677–690. doi:10.1007/s00248-010-9712-8
44. Maarschalk-Ellerbroek LJ, de Jong PA, Van Montfrans JM, Lammers JWJ, Bloem AC, Hoepelman AIM, Ellerbroek PM. CT Screening for Pulmonary Pathology in Common Variable Immunodeficiency Disorders and the Correlation with Clinical and Immunological Parameters. *J Clin Immunol* (2014) 34:642–654. doi:10.1007/s10875-014-0068-6
45. Bassis CM, Erb-Downward JR, Dickson RP, Freeman CM, Schmidt TM, Young VB, Beck JM, Curtis JL, Huffnagle GB. Analysis of the Upper Respiratory Tract Microbiotas as the Source of the Lung and Gastric Microbiotas in Healthy Individuals. *MBio* (2015) 6:1–10. doi:10.1128/mbio.00037-15
46. Marsh RL, Kaestli M, Chang AB, Binks MJ, Pope CE, Hoffman LR, Smith-Vaughan HC. The microbiota in bronchoalveolar lavage from young children with chronic lung disease includes taxa present in both the oropharynx and nasopharynx. *Microbiome* (2016) 4:1–18. doi:10.1186/s40168-016-0182-1

SUPPLEMENTARY MATERIAL

There are supplementary materials available for this manuscript.

Data Availability Statement

The datasets generated and analyzed for this study can be found in the European Nucleotide Archive under project code PRJEB34684. All R code used to analyze this data can be found at https://gitlab.com/rberbers/cvid_mbiota_oral/.

SUPPLEMENTARY METHODS

Patient inclusion

Patients were included at the outpatient clinics of the departments of clinical immunology, infectious diseases and paediatrics at the University Medical Center in Utrecht, the Netherlands, and the Erasmus Medical Center in Rotterdam, the Netherlands. Medication use in the three months prior to sampling was recorded. Clinical data was collected from the hospital electronic patient files. For laboratory measurements below the detection limit, the cut-off value was replaced with the limit of detection when computing mean serum levels. GLILD reported in the study characteristics was based on clinical assessment of the treating physician.

DNA isolation

DNA isolation was performed as described by Wyllie et al.¹. Briefly, oral swab stored in 200 µL of Amies medium was thawed and bead-beated twice in lysis buffer and phenol. The resulting DNA phase was purified using magnetic beads (LGC Genomics) and eluted into 100 µL of DNA isolate. Negative controls and mock communities (ZymoBIOMICS microbial community (DNA) standard, Zymo research, USA) were used from the beginning of DNA isolation up to the data analysis stage.

Bacterial load qPCR

Total bacterial load was determined using the BactQuant qPCR, as described by Liu et al.², on a StepOnePlus RT-PCR system (ThermoFisher). Briefly, 2 µL of undiluted DNA isolate for each sample was analysed with Taqman qPCR for 16S rRNA. Serial dilutions of a plasmid containing one *E.coli* 16S rRNA copy were used to calculate absolute 16S rRNA concentration in each sample. Primers and probes were ordered from IDT DNA technologies. TaqPath master mix (ThermoFisher) was used with standard cycling conditions on a StepOnePlus RT-PCR system (ThermoFisher). Forward primer: 5'- CCTACGGGDDGG-CWGC A-3', reverse primer: 5'- GGACTACHVGGGTMTCTAATC -3', probe: (6-FAM/ZEN) 5'-CAGCAGCCGCGGTA-3' (Iowa Black®FQ).

Bioinformatics

The 469 basepair V3 and V4 hyper-variable regions of the 16S rRNA gene were amplified and sequenced using the Illumina MiSeq instrument and Reagent Kit v3 (600-cycle) according to Fadrosch et al.³. The resulting amplicon pool generated a total of 6.6 million paired-end reads. These 2x300bp paired-end reads were pre-processed as follows. The first 12bp of each paired-end containing the index sequences, were extracted and afterwards concatenated to dual-index barcodes of 24bp specific for each read-pair and sample. Sequencing reads were de-multiplexed using the 'qiime demux emp-paired' command from the QIIME2 microbial community analysis pipeline (version 2018.8)⁴. The DADA2⁵ pipeline in QIIME2 using the command 'qiime dada2 denoise-paired' (with the options --p-trim-left-f 21 --p-trim-left-r 21 --p-trunc-len-f 275 --p-trunc-len-r 260) was then used for read quality filtering, removal of sequencing primers, paired-end reads merging, generating amplicon sequencing variants (ASVs) and the removal of chimeric sequences. After these steps, and removal of the control samples, a total of 3.2 million sequences were retained, with a mean amount of 19.3k per sample. The obtained ASVs (7351 in total) were aligned to the SILVA 16S rRNA gene database⁶ (SILVA 132) using the command 'qiime feature-classifier classify-sklearn'. Samples with total read count below 8,000 were considered to have insufficient coverage and were removed from further analysis.

Three water contaminants were detected across the 38 negative control samples: *Halomonas* (average read count in negative controls 496; in patient samples 21), *Pseudomonas* (average read count in negative controls 132; in patient samples 44), *Shewanella algae* (average read count in negative controls 424; in patient samples 71). *Pseudomonas* and *Halomonas* were not identified to the species level. In order to correct for the contamination, the genera *Halomonas*, *Pseudomonas* and *Shewanella* were removed from all further analyses. Raw sequencing data will be made available on the European Nucleotide Archive, project code PRJEB34684.

Chest CT scores

HRCT scans were performed for routine diagnostic screening every 5 years. For 74 patients, one scan closest to time of oropharyngeal sampling was scored by a pulmonary radiologist (F.M.H.) for AD and ILD in each lobe using a previously published scoring system^{7,8}. AD was scored as extent and severity of bronchiectasis, airway wall thickening, mucus plugging, tree-in-bud and airtrapping. ILD was scored as extent and severity of opacities, ground glass, septa thickening and lung nodules. The obtained score was normalised by the maximum obtainable score. For one patient who had undergone lobectomy and one with atelectasis of a single lobe, the maximum obtainable score was adapted to exclude the missing lobe. In thirteen cases where expiratory scans were not available, airtrapping could not be evaluated and this element was removed from the score.

Data analysis and statistical methods

All analyses were performed using R 3.2.0⁹, and made publically available on our group's Gitlab page: www.gitlab.com/rberbers/cvid_mbiota_oral. Continuous baseline

parameters, bacterial load, alpha diversity and AD/ILD scores were compared using the Mann-Whitney rank test or Student's t-test depending on distribution of the data. Categorical variables were compared using a two-tailed Fisher's exact test. Alpha diversity was calculated using the inverse Simpson index using the package *vegan*. Principal Component Analysis (PCA) was performed using the *prcomp* function on the centered log ratio (CLR) transformed data¹⁰. Count zero multiplicative replacement (CZM) from the package *zCompositions* was used to replace zeroes prior to CLR transformation. PERMANOVA was used to detect overall differences in PCA using the *adonis* function in *vegan*. Differential abundance testing was performed using ANCOM¹¹ with Benjamini-Hochberg correction for multiple testing using an alpha of 0.05 as a threshold for significance. Variables with percentage zeroes greater than 25% of all samples were excluded from analysis. All ANCOM analyses were corrected for age and gender.

Correlation between lung scores and microbiota were done on the CLR-transformed sequencing data as described above. Linear regression was performed using the function *lm()* and the following model: [lung score] ~ gender + age + [bacterium]

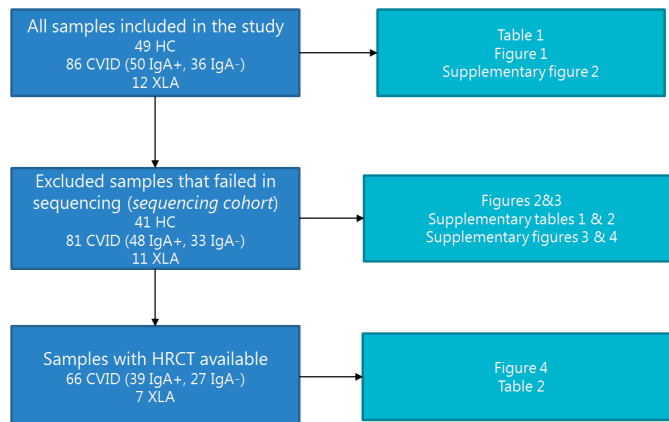
Bootstrapped confidence intervals were generated using the function *boot()* and 1000 iterations. Benjamini Hochberg correction was used to correct for false discovery rate.

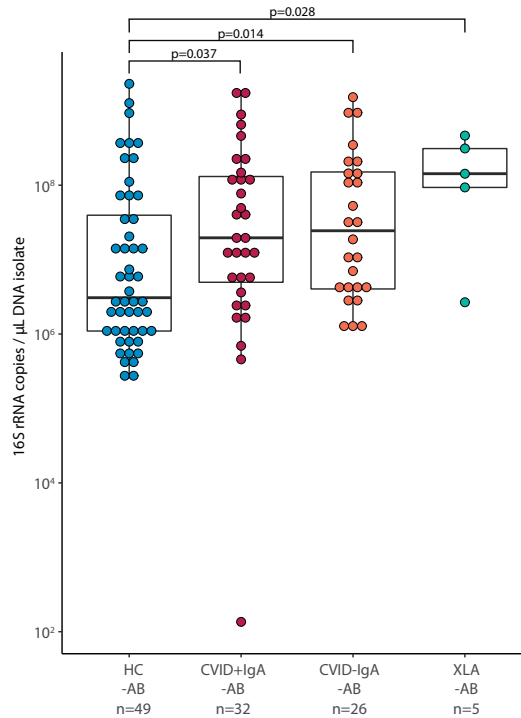
SUPPLEMENTARY REFERENCES

1. Wyllie, A. L. *et al.* Streptococcus pneumoniae in Saliva of Dutch Primary School Children. *PLoS One* **9**, e102045 (2014).
2. Liu, C. M. *et al.* BactQuant : An enhanced broad-coverage bacterial quantitative real-time PCR assay. *BMC Microbiol.* **12**, 56 (2012).
3. Fadrosch, D. W. *et al.* An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* **2**, 1–7 (2014).
4. Bolyen, E. *et al.* QIIME 2 : Reproducible , interactive , scalable , and extensible microbiome data science. *PeerJ Prepr.* (2018).
5. Callahan, B. J. *et al.* DADA2 : High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–587 (2016).
6. Quast, C. *et al.* The SILVA ribosomal RNA gene database project : improved data processing and web-based tools. *Nucleic Acids Res.* **41**, 590–596 (2013).
7. Ven, A. A. J. M. Van De & Montfrans, J. M. Van A CT Scan Score for the Assessment of Lung Disease in Children With Common Variable Immunodeficiency Disorders. *Chest* **138**, 371–379 (2010).
8. Maarschalk-Ellebrouck, L. J. *et al.* CT Screening for Pulmonary Pathology in Common Variable Immunodeficiency Disorders and the Correlation with Clinical and Immunological Parameters. *J. Clin. Immunol.* **34**, 642–654 (2014).
9. R Foundation for Statistical Computing, Vienna, A. R: A language and environment for statistical computing. (2018).at <<https://www.r-project.org/>>
10. Aitchison, J. *The statistical analysis of compositional data: Monographs on statistics and applied probability.* (Chapman & Hall Ltd., London, 1986).
11. Kaul, A., Mandal, S., Davidov, O. & Peddada, S. D. Analysis of Microbiome Data in the Presence of Excess Zeros. *Front. Microbiol.* **8**, 2114 (2017).

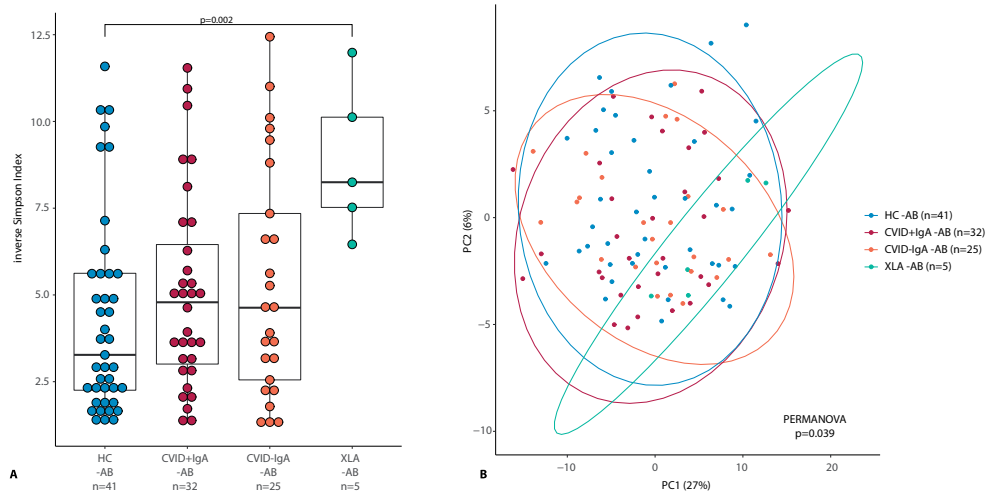
Table S1: study characteristics for the sequencing cohort

	HC	CVID +IgA (IgA >0.1 g/L)	CVID -IgA (IgA <0.1 g/L)	XLA
Total N	41	48	33	11
Age: mean ± SD	41 ± 12	34 ± 19	39 ± 14	23 ± 15
Male % (N)	27% (11/41)	52% (25/48)	61% (20/33)	100% (11/11)
Medication use during 3 months prior to sampling: % (N)				
Antibiotics	0% (0/41)	33% (16/48)	24% (8/33)	55% (6/11)
Immune suppressive therapy	0% (0/41)	13% (6/48)	15% (5/33)	8% (1/12)
Clinical phenotype: % (N)				
Any inflammatory complication	0% (0/41)	31% (15/48)	67% (22/33)	9% (1/11)
Autoimmune disease	0% (0/41)	19% (9/48)	36% (12/33)	9% (1/11)
GLILD (clinical diagnosis)	0% (0/41)	6% (3/48)	6% (2/33)	0% (0/11)
Granulomatous disease other	0% (0/41)	2% (1/48)	6% (2/33)	0% (0/11)
Enteritis	0% (0/41)	15% (7/48)	27% (9/33)	0% (0/11)
Malignancy	0% (0/41)	6% (3/48)	3% (1/33)	0% (0/11)
IgA status:				
Serum IgA levels available (N)	80% (34/41)	100% (48/48)	100% (33/33)	100% 11/11
% serum IgA low (<0.1 g/L)	0% (0/34)	0% (0/48)	100% (33/33)	100% (11/11)
Serum IgA mean ± SD in g/L	2.04 ± 0.79	0.75 ± 0.58	0.07 ± 0.00	0.07 ± 0.00

**Supplementary Figure 1.** Overview of all samples used in each figure and table. HC: healthy control. CVID: common variable immunodeficiency. XLA: X-linked agammaglobulinemia. IgA: immunoglobulin A. HRCT: high-resolution computed tomography.



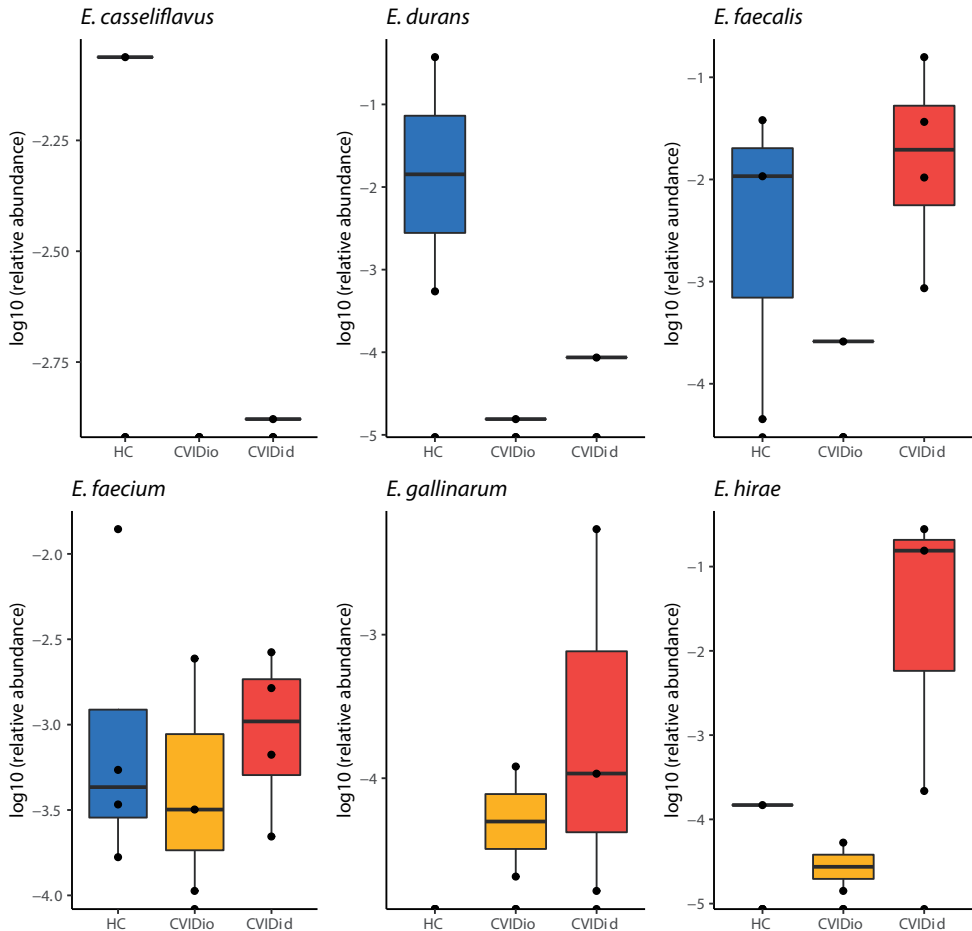
Supplementary Figure 2: Bacterial load in oropharyngeal swab samples as determined by qPCR for copies of the 16S rRNA gene in DNA isolates from oropharyngeal swabs in participants who did not use antibiotics (-AB) up to three months prior to sampling. Healthy controls (HC, n=49), CVID +IgA (n=32), CVID -IgA (n=26) and X-linked agammaglobulinemia (XLA, n=5). The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within $1.5 \times$ interquartile range. Statistical test: Mann-Whitney U test.



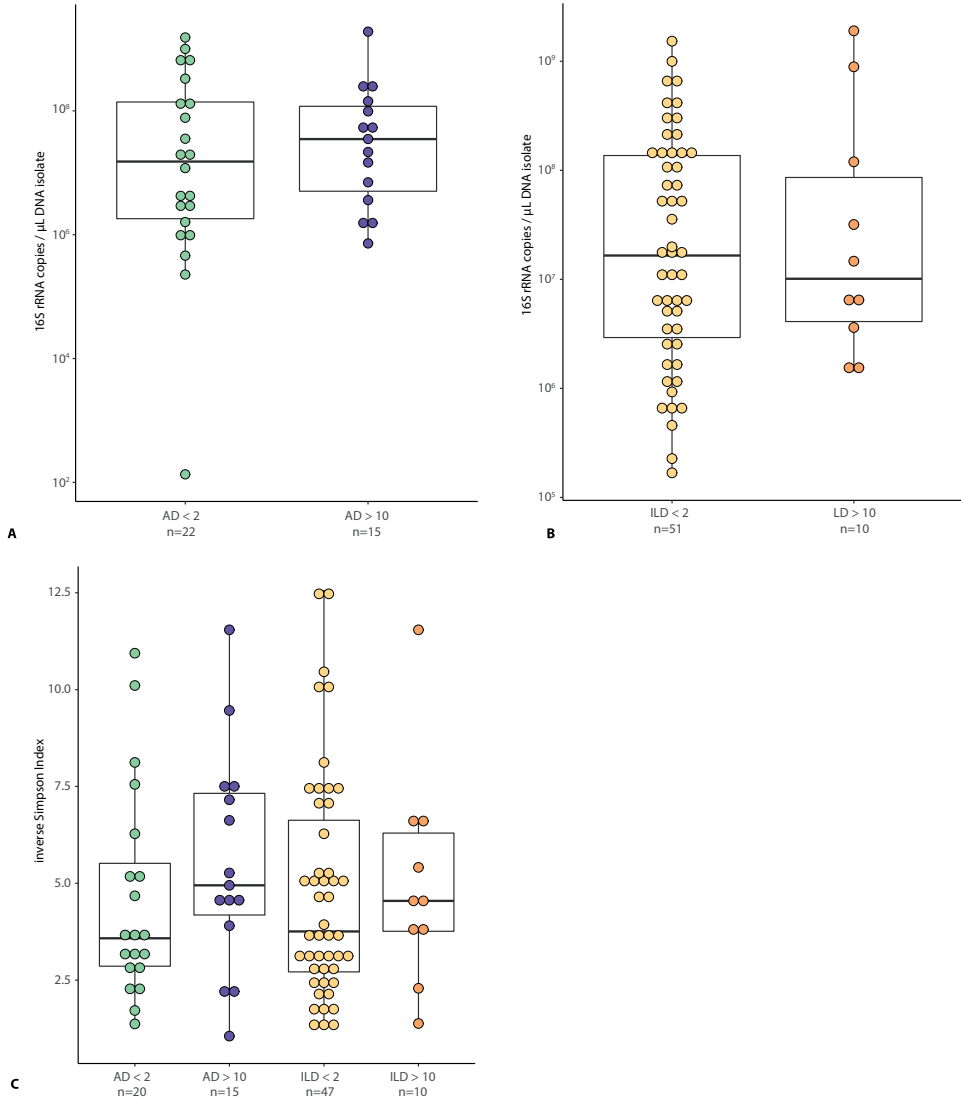
Supplementary Figure 3

A Alpha diversity as measured by inverse Simpson's index on 16S rRNA sequencing data of participants who did not use antibiotics up to 3 months prior to sampling (-AB): healthy controls (HC, n=41), CVID +IgA (n=32), CVID -IgA (n=25) and XLA (n=5). The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within 1.5×interquartile range. Statistical test: Mann-Whitney U test.

B Principal component analysis of centered log ratio (CLR)-transformed family level 16S rRNA sequencing data of the same samples described in **A**. Ellipses indicate 95% confidence intervals. Statistics: PERMANOVA using Euclidean distance on CLR data, $p=0.015$.



Supplementary figure 4 Differentially abundant taxa in 16S rRNA sequencing of healthy controls (HC, n=41), CVID +IgA (n=32), CVID -IgA (n=25) and XLA (n=5) patients who did not use antibiotics (-AB) 3 months prior to sampling. Statistics: ANCOM corrected for age and gender and Benjamini-Hochberg correction for false discovery rate. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within 1.5×interquartile range.



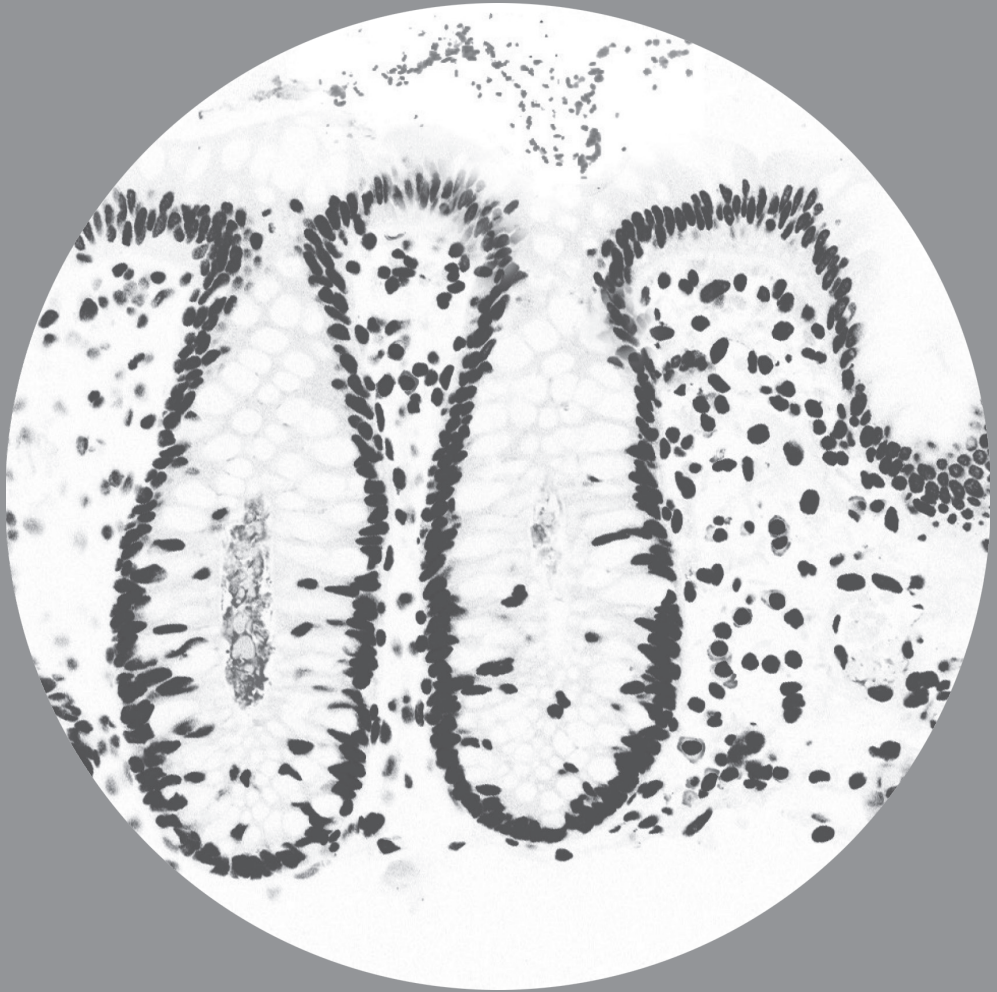
Supplementary figure 5

A Bacterial load in oropharyngeal swab samples as determined by qPCR for copies of the 16S rRNA gene in DNA isolates from oropharyngeal swabs in patients with airway disease score (AD) <2 (n=22) or >10 (n=15).

B Bacterial load in oropharyngeal swab samples as determined by qPCR for copies of the 16S rRNA gene in DNA isolates from oropharyngeal swabs in patients with interstitial lung disease score (ILD) <2 (n=51) or >10 (n=10).

C Alpha diversity of the same samples described in **A and B** as measured by inverse Simpson's index on 16S rRNA sequencing data.

The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within 1.5×interquartile range. Statistical test: Mann-Whitney U test.



Chapter 7

Gut Microbial Dysbiosis and *Enterococcus gallinarum* in Common Variable Immunodeficiency with Immune Dysregulation

Roos-Marijn Berbers
Fernanda L. Paganelli
Joris M. van Montfrans
Pauline M. Ellerbroek
Marco C. Viveen
Malbert R.C. Rogers
Moniek Salomons
Jaap Schuurmans
Martine van Stigt Thans
Remi M.M. Vanmaris
Lodewijk A.A. Brosens
Maria Marlot van der Wal
Virgil A.S.H. Dalm
P. Martin van Hagen
Annick van de Ven
Hae-Won Uh
Femke van Wijk
Rob J.L. Willems
Helen L. Leavis

Submitted

ABSTRACT

Common Variable Immunodeficiency (CVID) is a primary antibody deficiency characterised by hypogammaglobulinemia and recurrent infections. Significant morbidity and mortality is caused by immune dysregulation complications (CVIDid), which affect around one-third of CVID patients and have a poorly understood etiology. Here, we investigate the hypothesis that gut microbial dysbiosis contributes to the inflammation underlying CVIDid.

In this cross-sectional multicenter study, bacterial localization and crypt architecture were analyzed in gut biopsies of 15 CVID patients, 3 patients with X-linked agammaglobulinemia (XLA), and 9 healthy controls (HC). Next, bacterial load and microbiota composition were characterized using 16S rRNA-targeted qPCR and amplicon sequencing, respectively, in stool samples of 42 CVIDid patients, 51 CVID patients with infections only (CVIDio), 48 HC and 11 XLA patients.

The CVID gut microbiota was characterized by expansion of Enterobacteriaceae and bacterial invasion of colonic crypts in CVID and XLA colon biopsies. Increased fecal bacterial load and decreased alpha diversity were observed in CVIDid compared to HC, as well as a distinct beta diversity in CVIDid. Shotgun metagenomic sequencing followed by selective culturing and qPCR revealed a higher prevalence of *Enterococcus gallinarum* in CVIDid relative to CVIDio. Presence of *E. gallinarum* in stool correlated with increased serum IL-17A, IL-10, LILRB4 and Flt3L. When exposed to *E. gallinarum* supernatants, CVID and HC monocytes showed increased production of IL-6 and IL-6/IL-10 ratio.

This study further supports the hypothesis that a dysregulated gut microbiota contributes to systemic inflammation in primary antibody deficiency, and introduces *E. gallinarum* as a potential pathobiont in CVID with immune dysregulation.

Key words

Common Variable Immunodeficiency (CVID); Immune Dysregulation; Gut Microbiota; Pathobionts

INTRODUCTION

Common variable immunodeficiency disorder (CVID) is a primary immunodeficiency hallmarked by low serum immunoglobulins and impaired production of specific antibodies, resulting in an increased risk for infections with polysaccharide encapsulated bacteria^{1,2}. While the infection frequency and severity can be ameliorated with adequate immunoglobulin replacement therapy (IgRT), over a third of CVID patients still develop additional immune dysregulation complications such as autoimmune disease, granuloma formation, enteropathy, and lymphoproliferative disease with an increased risk of lymphoma. This CVID with immune dysregulation (CVIDid) phenotype causes significant morbidity and mortality, resulting in poorer long-term survival compared to CVID with infections only (CVIDio)^{3–6}. The cause of immune dysregulation complications in CVID is unknown, but it has been hypothesized that a dysbiotic gut microbiota may play a role in causing inflammation and clinical complications in CVIDid^{7,8}. A more profound antibody deficiency is observed in patients with X-linked agammaglobulinemia (XLA), which is caused by mutations in the early B-cell development gene Bruton's tyrosine kinase, and results in complete absence of B-cells and immunoglobulins⁹. XLA patients typically do not show the same predisposition for the development of immune dysregulation complications as CVID patients¹⁰.

The gut microbiota plays an important role in the maintenance of immune homeostasis, and has been implicated in the pathogenesis of several autoimmune diseases¹¹. The presence of commensal gut bacteria with a (suspected) causal link to the onset of disease – so-called pathobionts – is thought to contribute to diseases such as rheumatoid arthritis¹², systemic lupus erythematosus (SLE) and autoimmune hepatitis¹³, and mouse models of multiple sclerosis¹⁴. One proposed mechanism by which gut commensals can drive an immune response to autoantigens is through molecular mimicry – if the gut commensal provides an antigenic stimulus that causes cross-recognition of host proteins by the adaptive immune system¹⁵. Alternatively, pathobionts may provide a broad pro-inflammatory stimulus that causes innate immune activation and contributes to autoimmunity in individuals that are susceptible to self-recognition¹³. Identification of the mechanisms by which a suspected pathobiont contributes to disease is crucial in showing a causal role for the gut microbiota in inflammatory disease, and paves the way for possible therapeutic approaches.

The composition of the microbiota is partly regulated by the immune system, with an important role for immunoglobulin A (IgA) in the regulation of the microbiota at mucosal surfaces¹⁶. IgA is thought to not only influence the composition, but also the localization of the microbiota, thereby contributing to immune exclusion¹⁷. IgA deficiency is a known risk factor for the development of immune dysregulation in CVID^{18,19}, and Jørgensen et al.⁸ showed that low IgA was associated with decreased gut microbiota richness in a cohort of CVID patients of whom 80% had immune dysregulation complications. The presence of lipopolysaccharide (LPS) in serum of CVID patients in this cohort was suggestive of microbial (product) translocation and positively correlated with T-cell activation marker

sCD25⁸. These findings support the hypothesis that insufficient IgA may affect mucosal immunity, leading to an altered gut microbiota composition and translocation of microbial products in CVID, thereby contributing to immune dysregulation in CVID. The question whether CVIDid patients also have a specific gut microbiome signature compared to CVIDio patients remained unanswered due to small sample size of the CVIDio group in the Jørgensen study⁸.

Here, we assessed the localization of the microbiota in gut biopsies of CVID and XLA patients, and we characterized the bacterial load and composition of the gut microbiota in CVIDio and CVIDid in a cross-sectional multicenter study. We finally relate the presence of pathobiont species, in particular *Enterococcus gallinarum*, to inflammation markers in patient serum and assess their *in vitro* immunostimulatory capacities.

METHODS

Ethics statement

Ethical approval for this study for all Dutch participants was received from the Medical Ethical Committee of the Erasmus MC University Medical Center in Rotterdam, the Netherlands (METC: 2013-026). Written informed consent was obtained from all patients and controls according to the Declaration of Helsinki.

Study population

Patients aged seven or older diagnosed with CVID according to the criteria of the European Society for Immunodeficiencies criteria¹ were included during outpatient clinic visits of the University Medical Center Utrecht, the Erasmus MC University Medical Center Rotterdam and the University Medical Center Groningen, in the Netherlands. Household members of patients were recruited as healthy controls (HC). Medication use was recorded up to 3 months prior to sampling. All CVID and XLA patients received IgRT at time of sampling, with target IgG trough levels of >7.0 g/L. Clinical characteristics were collected from electronic patient files.

Biopsies

Residual biopsy tissue from endoscopic screening for gastrointestinal malignancy in patients with primary antibody deficiency was obtained from the Pathology Biobank of the UMC Utrecht and permission was granted by the UMC Utrecht Biobank Research Ethics Committee (TCBio 16-493). Biopsies from a control group of patients were age- and gender matched to biopsies from CVID patients and only included if they had no histological abnormalities and the indication for the biopsy was unrelated to a (suspected) inflammatory or immunodeficient condition. All biopsies were formalin fixed and paraffin-embedded. Slides of 4µm were deparaffinated and stained with hematoxylin and eosin (H&E) for histopathological assessment, and used for immunofluorescence or fluorescence *in situ* hybridization (FISH) staining. All biopsies used in this study were also evaluated by

a resident gastrointestinal pathologist of the UMC Utrecht in the context of regular care.

Fluorescence *in situ* hybridization

Tissue slides were de-waxed and hybridized for 2h at 50°C with a universal bacterial EUB-338 probe (Cy3 ~ 5'-GCTGCCTCCCGTAGGAGT- 3' ~ Cy3, IDT DNA technologies) or a non-EUB control probe (Cy3 ~ 5'- ACT CCT ACG GGA GGC AGC -3' ~ Cy3, IDT DNA technologies) in hybridization buffer (0,9M NaCl, 20mM TRIS at pH 7,5, SDS 0.1% wt/vol) with 20% formamide. Biopsies were analyzed with a confocal laser scanning microscope (CLSM, Leica SP5) at 63X magnification with digital zoom for close-up images, using LAS AF software (Leica).

Immunofluorescence staining

IgA staining was performed as described previously by Hendrickx et al²⁰. Briefly, after de-waxing and antigen retrieval (boiling in Na-citrate buffer pH 6 for 20min), staining was performed with primary antibody against IgA (1:100 unlabeled goat anti-human IgA, Southern Biotech, USA) and secondary antibody (1:500 Alexa Fluor 488 donkey anti-goat IgG, Invitrogen/Thermo Fisher Scientific, USA). TO-PRO-3 iodide (1:1000 Molecular Probes/Life technologies, the Netherlands) was added as a nuclear stain. Biopsies were analyzed with a confocal laser scanning microscope (CLSM, Leica SP5), using LAS AF software (Leica) at 40X magnification, with digital zoom for close-up images.

Histology

Hematoxylin and eosin-stained slides were scanned and stored at 40x magnification (Hamamatsu Digital slide scanner, Japan), and assessed by two independent researchers using NDP.view2 software (Hamamatsu, Japan). For each transverse colon biopsy, 5 longitudinal crypts not in the proximity of a Peyer's patch were chosen. For each crypt, the theca area of the eight largest goblet cells was measured and the length of the crypt was measured. Goblet cell theca area was used as a measure of mucus productive capacity of these cells²¹. Biopsies that did not contain 5 suitable longitudinally cut crypts were excluded. Means of the two independent measurements were used for final assessment.

Serum measurements

Serum was collected at time of fecal sampling, and stored at -80°C until analysis. IgA was measured in serum using a PEG-enhanced immunoturbidimetric method (Atellica CH, Siemens). Very low IgA was defined as serum IgA <0.1 g/L, in order to be consistent with the first gut microbiota study in CVID by Jørgensen et al.⁸. Serum cytokines were measured using Olink proximity extension assays (immune response and inflammation panels), as reported previously²².

Fecal sample collection and DNA isolation

Fecal samples were stored at -80°C within 24 hours of production. Total DNA from feces samples was isolated using the QiaAmp DNA stool minikit (Qiagen) according to manu-

facturers' protocol with the addition of a bead-beating step with 0.1mm zirconium beads in the stool lysisbuffer (ASL buffer). DNA was stored at -20°C prior to further analysis.

Bacterial load and species-specific qPCR

All primers and probes were ordered from IDT DNA technologies (Supplementary Table 1). qPCR was performed using a StepOnePlus RT-PCR system (ThermoFisher). Bacterial load in the fecal samples was estimated using the BactQuant qPCR, as described by Liu et al.²³. 16S rRNA copy number calculations were performed using serial dilutions of a plasmid containing the target sequence (IDT DNA technologies). For design of the species-specific qPCR assay, strains originating from human stools were selected from the laboratory collection of the Department of Medical Microbiology, UMC Utrecht (for an overview of all bacterial strains used in this study, see Supplementary Table 2). All qPCR assays were tested for cross-reactivity against the other strains.

16S rRNA sequencing and bioinformatics

The 469 basepair V3 and V4 hypervariable regions of the 16S rRNA gene were amplified and sequenced using the Illumina MiSeq instrument and Reagent Kit v3 (600-cycle) according to Fadrosh et al.²⁴. The resulting amplicon pool generated a total of 10.2 million read-pairs (sample median of 60.9k read-pairs). The QIIME2 microbial community analysis pipeline (version 2018.8)²⁵ was used with DADA2²⁶ for sequence variant detection (with default settings, except for --p-trunc-len-f 275 --p-trunc-len-r 260), and SILVA as 16S rRNA reference gene database (SILVA 132)²⁷. Sequencing data has been made available on the European Nucleotide Archive under project code PRJEB44275. Samples with total reads below 1500 were excluded from all analyses.

Shotgun metagenomic sequencing and bioinformatics

Shotgun metagenomic sequencing of fecal samples was performed with the same DNA extracted for use in 16S rRNA sequencing. Sequence libraries were prepared using the Nextera XT Kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions, using 1ng of total DNA input. Libraries were sequenced by the Utrecht Sequencing Facility on an Illumina NextSeq 500 system with a 300 cycle (2 × 150 bp) NextSeq 500/550 High-Output v2.5 Kit. Illumina sequencing data were quality-assessed and trimmed using Trim Galore (default settings)²⁸.

Shotgun metagenomic sequencing of fecal DNA yielded 550 million paired-end reads (sample median of 13 million reads). Taxonomic classification was performed using the OneCodex platform²⁹ with the OneCodex database. Trimmed and quality-checked whole genome sequencing data was used for *de novo* genome assembly (metaSPAdes³⁰ with default settings and additional flag --only-assembler) and annotation (Prokka³¹ with default settings).

Selective culturing and whole genome sequencing of bacterial strains

For selective culturing of enterococci, fecal samples were first cultured in Enterococcosel

broth³² at 37°C overnight. Next, liquid culture was plated on Slanetz and Bartley agar (Merck Millipore) (for isolation of *E. gallinarum* with the addition of 4µg/mL vancomycin) and cultured for 48h at 37°C. Individual bacterial colonies were identified at the species level using MALDI-TOF mass spectrometry, and stored at -80°C until further use.

Whole genome sequencing of cultured *Enterococcus* strains was performed as follows. DNA was isolated using the Wizard Genomic DNA Isolation Kit (Promega) with the protocol for gram positive bacteria. Sequence libraries for Illumina sequencing were prepared using the Nextera XT Kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Libraries were sequenced by the Utrecht Sequencing Facility on an Illumina NextSeq 500 system with a 300 cycle (2 × 150 bp) NextSeq 500/550 Mid-Output v2.5 Kit. Trimmed and quality-checked whole genome sequencing data was used for *de novo* genome assembly (metaSPAdes³⁰ with default settings and additional flag --only-assembler) and annotation (Prokka³¹ with default settings). Mega X³³ software was used to calculate the neighborhood-joining tree using the *fliC* sequences of *E. gallinarum*.

Monocyte cocultures

Bacterial supernatants were cocultured with patient or HC cells as described by the Zitvogel group³⁴, with some modifications. Whole blood was collected from patients at the same time as serum and fecal sampling, and peripheral blood mononuclear cells were isolated by ficoll-density centrifugation (GE Healthcare-Biosciences, AB), and frozen at -180°C until use. For monocyte-bacterial coculture, PBMCs were thawed and monocytes were isolated using magnetic-activated cell sorting using CD14+ microbeads (Miltenyi Biotech). Per condition, 5000 monocytes were plated in 96 well-plates in Iscove's Modified Dulbecco's Medium (IMDM, ThermoFisher) with 5% glutamine 5% hepes and 10% human AB serum. Bacteria were prepared by inoculating 5mL brain-heart infusion (BHI) medium with enterococci isolated from patient stools, or *E. coli* E783 in 5mL lysogeny broth. After overnight culture at 37°C, bacterial supernatants were filter sterilized at 0.2 µm. For each condition 20µL bacterial supernatant was added to the monocytes and incubated at 37°C and 5% CO₂. The next day, culture supernatants were harvested and frozen at -20°C until further use. Cytokine measurements of IL-6 and IL-10 in culture supernatants were performed using Luminex technology³⁵.

Data analysis and statistical methods

All analyses were performed using R 3.2.0³⁶. Continuous variables were compared using the Mann-Whitney rank test. Categorical variables were compared using a two-tailed Fisher's exact test. Inverse Simpson index was calculated using the package *vegan*. Principal Component Analysis (PCA) was performed using the *prcomp* function on hyperbolic arcsine (asinh) transformed data. PERMANOVA with correction for false discovery rate (FDR) was used to detect global community differences in PCA using the package *vegan*. Differential abundance testing was performed using ANCOM-BC³⁷ with Benjamini-Hochberg FDR correction using an alpha of 0.05 as a threshold for significance. All ANCOM-BC analyses were corrected for age and sex using the following formula: bacterium ~ patient_

group + age + sex. Bacterial taxa with zero reads in more than 90% of samples were excluded from analysis. For subgroup analyses within the immune dysregulation group, structural zero detection was not applied due to small sample size and increased risk of false positives.

RESULTS

Local IgA deficiency, bacterial crypt invasion and altered crypt architecture in CVID

In order to investigate the localization of the microbiota in primary antibody deficiency, we performed 16S rRNA FISH staining on colon biopsies from 15 CVID patients, 3 XLA patients and 9 HC (Supplementary Table 3). No patients experienced clinical signs of enteritis at time of sampling. Luminal bacteria were observed in all biopsies (Figure 1A). In 3 CVID patients and 1 XLA patient, but none of the HC (Figure 1B-D), presence of bacteria deep inside colon crypts was observed. Of the patients with crypt invasion, all had serum IgA < 0.1 g/L, and the biopsies showed lymphoid infiltration. Of the CVID patients, 2 had immune dysregulation complications (one had autoimmune gastritis, and the other granulomatous-lymphocytic interstitial lung disease and splenomegaly). Crypt-invading bacteria were of mixed morphology, suggesting location shift of a heterogeneous population of bacteria rather than of one invasive species.

As IgA is thought to contribute to immune exclusion of the microbiota¹⁷, we next assessed whether serum IgA levels also reflected local IgA production in the gut in CVID by staining the biopsies for secretory IgA. Indeed, in 4/4 CVID patients with serum IgA > 0.1 g/L and 9/9 HC, IgA+ plasma cells were detected (Figure 2A-B), while they were absent in 10/11 CVID patients with serum IgA < 0.1 g/L (Figure 2C). An additional barrier between the gut microbiota and the epithelium is mucus itself. Bacterial crypt invasion has also been reported in a Muc2 (the mucin that composes the mucus barrier in the colon) deficient mouse model³⁸. Similar to the Muc2 deficient mice, we observed crypt elongation ($p=0.019$) (Figure 2D) and a slight increase in size of Goblet cell theca area (Figure 2E) in CVID that was not statistically significant ($p=0.101$). Disturbed mucus and IgA production may contribute to increased contact of bacteria with the epithelium.

Increased fecal bacterial load in CVIDid without use of antibiotic and immunosuppressive therapy

In order to investigate whether the gut microbiota is associated with inflammation in CVID, we assessed if the abundance and composition of fecal microbiota were different in CVIDid compared to CVIDio. Fecal samples were collected from a cohort of 42 CVIDid, 51 CVIDio, 11 XLA patients and 48 HC (Table 1). In this cohort, serum IgA was lower in CVIDid than CVIDio ($p=0.006$) and HC ($p<0.001$), and normal (>0.70 g/L) for all HC (Supplementary Figure 1).

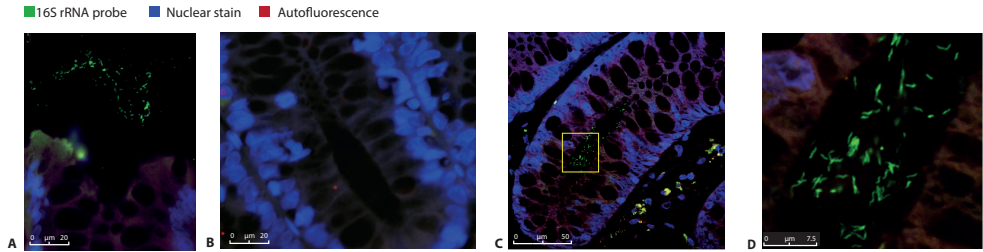


Figure 1: Localization of the microbiota in CVID. Representative 16S rRNA Fluorescence *in-situ* hybridization (FISH) staining. Luminal microbiota (in green) were excited at 488nm; DNA in the colon epithelium cell (in blue) is stained with DAPI at 405nm. Images were taken using a Leica SP5 confocal microscope with a 63x objective.

A: luminal microbiota in a healthy control colon biopsy.

B: colonic crypt without presence of microbiota in a healthy control colon biopsy

C: representative 16S rRNA FISH staining showing bacterial presence in colonic crypts in a CVID patient. Yellow box indicates area enlarged in D.

D: detail of C, showing microbiota invasion in colonic crypts of a CVID patient.

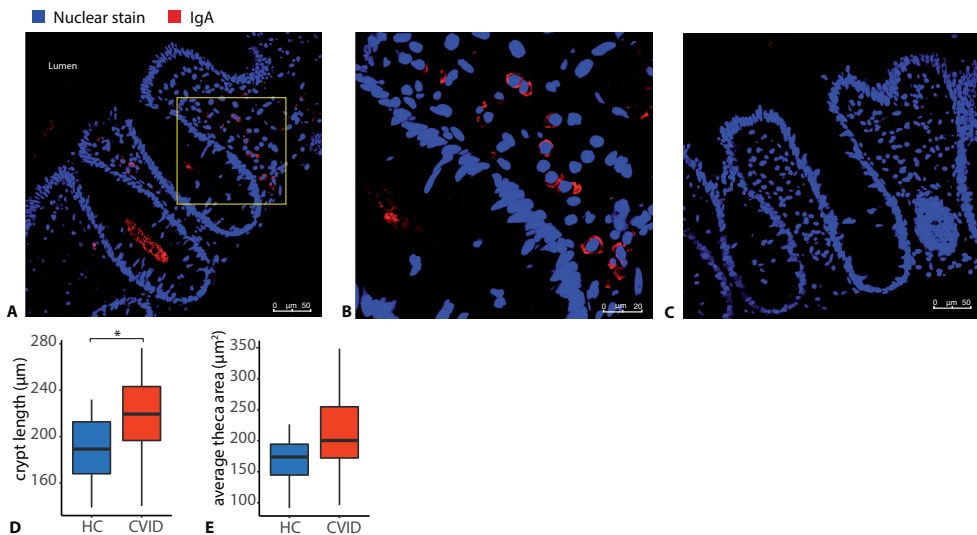


Figure 2: IgA deficiency and crypt architecture in CVID

A: representative IgA staining of a healthy control. Yellow box indicates area enlarged in B. IgA (in red) was excited at 488nm; DNA in the colon epithelium cell (in blue) was stained with TO-PRO-3 iodide at 642 nm. Images were taken using a Leica SP5 confocal microscope with a 40x objective.

B: detail of **A** showing IgA+ plasma cells, using digital zoom.

C: representative IgA staining of a CVID patient with no detectable IgA in serum, and absence of IgA in colon biopsy

D: average crypt length in healthy control (HC, n=9) and CVID (n=15) colon biopsies

E: average Goblet cell theca area in healthy control (HC, n=9) and CVID (n=15) colon biopsies
The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within 1.5 × interquartile range. Statistics: Mann-Whitney U test. * p<0.05, ** p<0.01, *** p<0.001.

Table 1: cohort overview. HC: Healthy control, CVIDio: CVID with infections only, CVIDid: CVID with immune dysregulation, XLA: X-linked agammaglobulinemia. IQR: interquartile range. VUS: variant of unknown significance.

Summary Statistics	HC (N = 48)	CVIDio (N = 51)	CVIDid (N = 42)	XLA (N = 11)
Characteristics				
age - median (IQR)	43.50 (36.00, 49.50)	40 (25.00, 56.00)	40.50 (32.00, 53.75)	16 (12.00, 26.50)
sex (male)	18 (37.50%)	22 (43.14%)	20 (47.62%)	11 (100.00%)
center (Utrecht)	41 (85.42%)	30 (58.82%)	29 (69.05%)	9 (81.82%)
antibiotics	0 (0.00%)	18 (35.29%)	20 (47.62%)	7 (63.64%)
immunosuppressive medication	0 (0.00%)	5 (9.80%)	11 (26.19%)	1 (9.09%)
serum IgA >0.1 g/L	48 (100.00%)	36 (70.59%)	17 (40.48%)	0 (0.00%)
Immune dysregulation complications				
Pulmonary	0 (0.00%)	0 (0.00%)	11 (26.19%)	0 (0.00%)
Hematological	0 (0.00%)	0 (0.00%)	10 (23.81%)	0 (0.00%)
Gastrointestinal	0 (0.00%)	0 (0.00%)	22 (52.38%)	0 (0.00%)
Rheumatological	0 (0.00%)	0 (0.00%)	10 (23.81%)	0 (0.00%)
Dermatological	0 (0.00%)	0 (0.00%)	7 (16.67%)	0 (0.00%)
Lymphoproliferative	0 (0.00%)	0 (0.00%)	18 (42.86%)	0 (0.00%)
Other	0 (0.00%)	0 (0.00%)	6 (14.29%)	0 (0.00%)
Genetics				
Not done	48 (100.00%)	48 (94.12%)	30 (71.43%)	7 (63.64%)
Nothing found	0 (0.00%)	2 (3.92%)	4 (9.52%)	0 (0.00%)
Only VUS found	0 (0.00%)	0 (0.00%)	5 (11.90%)	0 (0.00%)
Pathogenic mutations found	0 (0.00%)	1 (1.96%) §	4 (9.52%) #	4 (36.36%)

§ pathogenic mutations: TAC1

pathogenic mutations: 1 CTLA4 haploinsufficiency, 1 STAT1GoF, 1 PI3KR1, 1 TAC1

The total amount of bacteria in each fecal sample was quantified by determining 16S rRNA gene load using qPCR. While there was no difference in bacterial load when all sampled patients were included (Figure 3A), bacterial load was increased in CVIDid compared to HC ($p=0.02$, Figure 3B) when patients who used antibiotic or immunosuppressive therapy 3 months prior to sampling were excluded. In patients who did use medication, bacterial load was similar to that of HC (Supplementary Figure 2A). The increased bacterial load in CVIDid was mostly present in patients with lymphoproliferation (Figure 3C). No difference in bacterial load between CVID with serum IgA <0.1 g/L (CVID-IgA) and CVID with serum IgA >0.1 g/L (CVID+IgA) was observed (Supplementary Figure 2B), regardless of medication use (Supplementary Figure 2C).

Lower alpha diversity in CVIDid

Taxonomic characterization of the fecal microbiota in CVID patients by 16S rRNA amplicon sequencing showed that the overall most abundant bacterial genera in feces (Figure 4A) were *Blautia*, *Faecalibacterium* and a further unidentified genus belonging to the

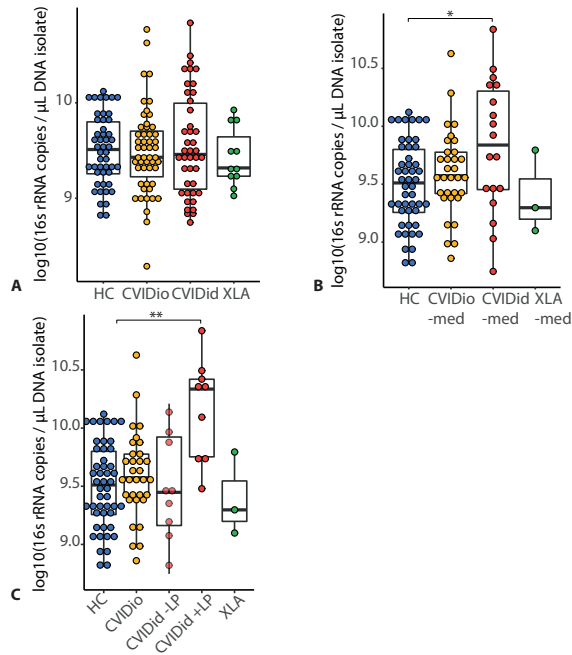


Figure 3: Bacterial load as determined using 16S rRNA qPCR in:

A: Healthy controls (HC n=48), CVID with infections only (CVIDio, n=51) CVID with immune dysregulation (CVIDid, n=42), and X-linked agammaglobulinemia (XLA, n=11).

B: Patients who did not use antibiotic or immunosuppressive treatment up to 3 months prior to sampling (-med). Healthy controls (HC n=48), CVIDio-med (n=32), CVIDid-med (n=18) XLA-med (n=3).

C: Same analysis; splitting up the CVIDid-med group in those without lymphoproliferation (n=9) and those with lymphoproliferation (n=9).

The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within $1.5 \times$ interquartile range. Statistics: Mann-Whitney U test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Lachnospiraceae family. Microbial alpha diversity as expressed by inverse Simpson Index (Figure 4B) was decreased in CVIDid compared to HC ($p = 0.002$), with CVIDio being the intermediate group (CVIDio vs HC $p = 0.004$), regardless of use of antibiotic or immunosuppressive therapy (Supplementary Figure 3A). The same pattern was observed when the CVID group was split according to IgA status, with decreased alpha diversity in the CVID-IgA group (HC vs CVID-IgA $p = 0.002$, Supplementary Figure 3B and C).

Beta diversity showed a small shift of both CVID phenotypes away from the HC (CVIDio vs HC $p_{\text{adj}} = 0.010$, CVIDid vs HC $p_{\text{adj}} = 0.006$) (Figure 4D), regardless of medication use (Supplementary Figure 3D). Beta diversity for CVID+IgA and CVID-IgA showed similar patterns (Supplementary Figure 3E). No differences in beta diversity were found between CVIDid patients with similar immune dysregulation phenotypes as reported in Table 1 (data not shown).

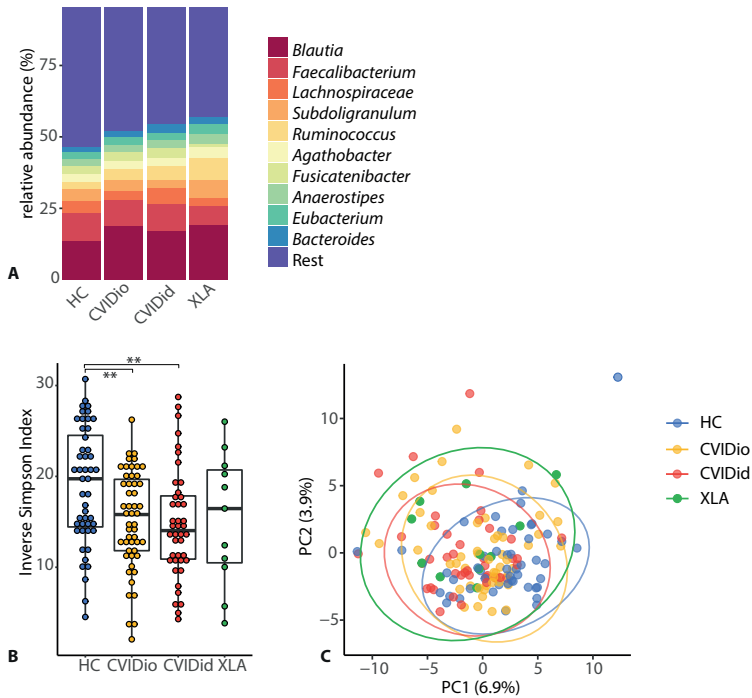


Figure 4: composition of the gut microbiota in CVID as determined using 16S rRNA gene sequencing. Healthy controls (HC n=48), CVID with infections only (CVIDio, n=51) CVID with immune dysregulation (CVIDid, n=42), and X-linked agammaglobulinemia (XLA, n=11).

A: Top-10 most abundant bacterial genera

B: Alpha diversity as calculated using the inverse Simpson index. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within $1.5 \times$ interquartile range. Statistics boxplots: Mann-Whitney U test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

C: Beta diversity shown as principal component analysis (PCA) on genus-level hyperbolic arcsine transformed data. FDR-adjusted PERMANOVA: CVIDid vs HC adj. $p = 0.006$, CVIDio vs HC adj. $p = 0.010$. Ellipses indicate the 95% confidence interval.

Bacteria from the genus *Enterococcus* are associated with CVIDid

Next, we studied the distribution of bacterial taxa between all CVID patients and HC (Table 2). Bacterial phyla Firmicutes (p . adj = 0.002) and Proteobacteria (p . adj = 0.004) were slightly more abundant in CVID compared to HC, while the class of Actinobacteria (p . adj = 0.048) was increased in HC. The bacterial group *Escherichia-Shigella* and its corresponding higher taxonomic levels showed the strongest association with CVID (effect size β -0.433; standard error SE 0.125; adj. $p = 0.011$). *Escherichia-Shigella* was more often present in CVID (% zero in HC 62.50 vs 43.01 in CVID) and at higher abundance (0.22% in HC vs 0.47% in CVID).

Other bacterial genera more frequently found in CVID patients compared to HC were *Eggerthella*, *Alloprevotella*, *Lactococcus*, *Erysipelatoclostridium*, *Veillonella*, *Parasutterella*,

Table 2: Differentially abundant bacteria identified using 16S rRNA sequencing. CVID (n=93) vs healthy control (HC, n=48): coefficients, standard errors (SE), p-values, FDR-adjusted p-values. Median relative abundance of non-zero counts per group. Statistics: ANCOM-BC.

Up in cVID	HC Zero percentage	Median relative abundance HC	CVID Zero percentage	Median relative abundance CVID	Beta CVID vs HC	SE CVID vs HC	p-value CVID vs HC	adj. p-value CVID vs HC
D_0_0_Bacteria;D_1_1_Actinobacteria;D_2_2_Corrobacteriia;D_3_3_Corrobacteriataes;D_4_4_Eggerthellaceae;D_5_5_Eggerthella	83.33	0.14	66.67	0.20	-0.272	0.089	0.002	0.034
D_0_0_Bacteria;D_1_1_Bacteroidetes;D_2_2_Bacteroidia;D_3_3_Bacteroidiales;D_4_4_Prevotellaceae;D_5_5_Alloprevotella	97.92	1.70	82.80	0.31	-0.250	0.070	0.000	0.000
D_0_0_Bacteria;D_1_1_Firmicutes	0.00	89.49	0.00	92.15	-0.213	0.060	0.000	0.002
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Bacilli	16.67	0.57	12.90	0.69	-0.391	0.131	0.003	0.011
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Bacilli;D_3_3_Lactobacillales;D_4_4_Streptococcaceae;D_5_5_Lactococcus	93.75	0.10	82.80	0.07	-0.148	0.053	0.000	0.000
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Clostridia	0.00	85.54	0.00	88.11	-0.199	0.060	0.001	0.005
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Clostridia;D_3_3_Clostridiales;D_4_4_Eubacteriaceae	93.75	0.02	87.10	0.07	-0.145	0.054	0.000	0.000
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Clostridia;D_3_3_Clostridiales;D_4_4_Family XIII;D_5_5_Eubacterium 1	87.50	0.08	75.27	0.08	-0.192	0.067	0.004	0.035
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Clostridia;D_3_3_Clostridiales;D_4_4_Lachnospiraceae;D_5_5_Ruminococcus 1	18.75	0.80	47.31	1.35	-0.336	0.120	0.005	0.035
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Clostridia;D_3_3_Clostridiales;D_4_4_Lachnospiraceae;D_5_5_Blaustia	2.08	13.87	0.00	17.22	-0.218	0.068	0.001	0.023
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Clostridia;D_3_3_Clostridiales;D_4_4_Lachnospiraceae;D_5_5_Sellimonas	93.75	0.08	70.97	0.30	-0.368	0.087	0.000	0.000
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Clostridia;D_3_3_Clostridiales;D_4_4_Lachnospiraceae;D_5_5_Tuzzerella	95.83	1.00	81.72	0.27	-0.231	0.078	0.000	0.000
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Clostridia;D_3_3_Clostridiales;D_4_4_Lachnospiraceae;D_5_5_Tuzzerella 3	89.58	0.37	76.34	0.26	-0.229	0.082	0.005	0.035
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Clostridia;D_3_3_Clostridiales;D_4_4_Ruminococcaceae;D_5_5_Caprioproducens	93.75	0.06	86.02	0.08	-0.137	0.053	0.000	0.000
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Clostridia;D_3_3_Clostridiales;D_4_4_Ruminococcaceae;D_5_5_Flavonifractor	89.58	0.20	74.19	0.21	-0.248	0.084	0.003	0.035
D_0_0_Bacteria;D_1_1_Firmicutes;D_2_2_Clostridia;D_3_3_Clostridiales;D_4_4_Ruminococcaceae;D_5_5_Oscillibacter	70.83	0.12	58.06	0.13	-0.233	0.088	0.008	0.048

Table 2: Continued

	Zero percentage HC	Median relative abundance HC	Zero percentage CVID	Median relative abundance CVID	Beta CVID vs HC	SE CVID vs HC	p-value CVID vs HC	adj. p-value CVID vs HC
Up in CVID								
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Ruminococcales;D_5__Ruminococcaceae;UG:004	54.17	0.36	33.33	0.29	-0.311	0.107	0.004	0.035
D_0__Bacteria;D_1__Firmicutes;D_2__Erysipelotrichia;D_3__Erysipelotrichales;D_4__Erysipelotrichaceae;D_5__Erysipelotrichaceae	87.50	0.03	68.82	0.10	-0.247	0.068	0.000	0.010
D_0__Bacteria;D_1__Firmicutes;D_2__Negativicutes;D_3__Selenomonadales;D_4__Veillonellaceae;D_5__Veillonella	93.75	0.18	82.80	0.33	-0.220	0.090	0.000	0.000
D_0__Bacteria;D_1__Proteobacteria	14.58	0.30	13.98	0.47	-0.348	0.108	0.001	0.004
D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria	31.25	0.23	20.43	0.47	-0.492	0.132	0.000	0.002
D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Betaproteobacteriales;D_4__Burkholderiaceae;D_5__Parasutterella	85.42	0.29	70.97	0.24	-0.225	0.080	0.005	0.035
D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Enterobacteriales	60.42	0.24	37.63	0.56	-0.372	0.135	0.006	0.025
D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Enterobacteriales;D_4__Enterobacteriaceae	60.42	0.24	37.63	0.56	-0.462	0.132	0.000	0.007
D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Enterobacteriales;D_4__Enterobacteriaceae;D_5__Escherichia-Shigella	62.50	0.22	43.01	0.47	-0.433	0.125	0.001	0.011
Up in HC								
D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria	4.17	2.12	25.81	1.51	0.312	0.131	0.017	0.048
D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Bifidobacteriales	4.17	2.12	25.81	1.51	0.522	0.125	0.000	0.000
D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Bifidobacteriales;D_4__Bifidobacteriaceae	4.17	2.12	25.81	1.51	0.431	0.121	0.000	0.007
D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Bifidobacteriales;D_4__Bifidobacteriaceae;D_5__Bifidobacterium	4.17	2.12	25.81	1.51	0.433	0.121	0.000	0.010
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Prevotellaceae;D_5__Prevotella 9	62.50	1.91	88.17	1.87	0.383	0.145	0.008	0.048
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Peptostreptococcales;D_5__Romboutsia	16.67	1.40	39.78	1.18	0.346	0.124	0.005	0.035
D_0__Bacteria;D_1__Firmicutes;D_2__Negativicutes;D_3__Selenomonadales	2.08	1.57	6.45	0.84	0.269	0.098	0.006	0.025

and 11 bacteria belonging to the order Clostridiales: *Ruminococcus*, *Blautia*, *Sellimonas*, *Tyzzerella*, *Caproiciproducens*, *Flavonifractor*, *Oscillibacter* and a genus from the Ruminococcaceae UCG-004 group. After exclusion of patients with recent antibiotic use, all associated bacteria except *Alloprevotella* and *Ruminococcus* remained similarly associated with CVID although some lost statistical significance after FDR-correction (Supplementary Table 4).

Comparison of the composition of the fecal microbiota between CVIDid and CVIDio patients revealed that *Enterococcus* was exclusively detected in feces of CVIDid patients and absent in feces of CVIDio patients (Table 3), also after exclusion of patients with recent antibiotic use (Supplementary table 5). Bacterial genera that were associated with CVIDid but lost statistical significance after correction for multiple testing were *Eggerthella*, *Bacteroides*, *Alloprevotella*, *Lachnoclostridium*, *Flavonifractor* and *Veillonella*, while the genera belonging to the Lachnospiraceae family CAG-56 and *Coprococcus* were associated with CVIDio (Table 3). When comparing CVID+IgA to CVID-IgA, *Enterococcus* was significantly associated with CVID-IgA as a structural zero, in addition to *Eubacterium* (Supplementary table 6).

***Enterococcus gallinarum* as a differentially abundant pathobiont in immune dysregulation in CVID**

Since the presence of bacterial species from the genus *Enterococcus* have previously been associated with inflammation^{39,40} and pathophysiology of autoimmune disease^{13,41}, we next assessed whether the enterococci present in CVIDid in our cohort play a similar role. In order to investigate which *Enterococcus* species might be more abundant in the fecal microbiota in CVIDid, shotgun metagenomic sequencing of 4 HC, 4 CVIDio and 4 CVIDid was performed. Overall, the most prevalent enterococci were *Enterococcus casseliflavus*, *Enterococcus durans*, *Enterococcus faecium*, *Enterococcus faecalis*, *Enterococcus gallinarum*, and *Enterococcus hirae* (Supplementary Figure 4). Of these, only *E. gallinarum* and *E. hirae* were enriched in CVIDid. *E. casseliflavus* and *E. durans* were not prevalent in CVID, and therefore were excluded from further analysis.

To confirm presence of specific enterococcal species in the whole cohort, a parallel approach of bacterial culturing and qPCR was used. Species-specific qPCR assays were developed for *E. faecium*, *E. faecalis*, *E. gallinarum* and *E. hirae*, and tested against the other species to ascertain that no cross-reactivity occurred (data not shown). As enterococci can easily acquire high levels of antibiotic resistance, samples of patients who had used antibiotics 3 months prior to sampling were excluded in order to prevent selection bias based on antibiotic use. Selective culturing combined with qPCR (Table 4) showed that *E. faecium* and *E. faecalis* were overall the most prevalent enterococci in CVIDio (*E. faecium* 60.61% and *E. faecalis* 48.48%) and CVIDid patients (*E. faecium* 68.18% and *E. faecalis* 63.64%) alike. Although all *Enterococcus* species showed increased prevalence in CVIDid patients, *E. gallinarum* was most strongly associated with the CVIDid phenotype with a 3.75-fold detection increase in CVIDid (22.73%) when compared to CVIDio patients (6.06%), although this difference did not reach statistical significance ($p=0.103$), followed

Table 3: Differentially abundant bacteria identified using 16S rRNA sequencing. CVID with immune dysregulation (CVIDid, n=42) versus CVID with infections only (CVIDio, n =51): coefficients, standard errors (SE), p-values, FDR-adjusted p-values. Median relative abundance of non-zero counts per group. Statistics: ANCOM-BC. Black: p-adj.<0.05. Grey: p-adj.>0.05.

	CVIDio Zero percentage	Median relative abundance CVIDio	Zero percentage CVIDid	Median relative abundance CVIDid	Beta CVIDid - CVIDio	SE CVIDid - CVIDio	p-value CVIDid - CVIDio	adj. p-value CVIDid - CVIDio
Up in cVIDid								
D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Enterococcaceae	100	NA	76.190	0.383	-0.811	0.296	0.000	0.000
D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Enterococcaceae;D_5__Enterococcus	100	NA	76.190	0.383	-0.919	0.297	0.000	0.000
D_0__Bacteria;D_1__Actinobacteria;D_2__Coriobacteriia;D_3__Coriobacteriales;D_4__Eggerthellaceae;D_5__Eggerthella	76.471	0.211	54.762	0.198	-0.794	0.324	0.014	0.364
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidiales;D_4__Bacteroidaceae;D_5__Bacteroides	3.922	2.162	0.000	3.285	-0.576	0.291	0.048	0.542
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Prevotellaceae;D_5__Alloprevotella	74.510	0.369	92.857	0.211	0.687	0.297	0.042	0.849
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Lachnospiraceae;D_5__Lachnospiridium	33.333	0.474	21.429	0.547	-0.900	0.450	0.045	0.542
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Ruminococcaceae;D_5__Flavonifractor	84.314	0.177	61.905	0.208	-0.891	0.325	0.006	0.208
D_0__Bacteria;D_1__Firmicutes;D_2__Negativicutes;D_3__Selenomonadales;D_4__Veillonellaceae;D_5__Veillonella	90.196	0.193	73.810	0.389	-0.820	0.389	0.035	0.507
Up in cVIDio								
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Lachnospiraceae;D_5__CAG-56	60.784	0.516	88.095	0.467	1.090	0.354	0.002	0.111
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Lachnospiraceae;D_5__Coproccoccus 2	31.373	0.551	42.857	0.428	1.080	0.492	0.029	0.486

Table 4:

Presence of <i>Enterococcus</i> species detected with culturing and/or qPCR	CVIDio - antibiotics	CVIDid - antibiotics	Fisher's exact test
patient without antibiotic usage (n)	33	22	
<i>E. faecium</i> n (%)	20 (60.61%)	15 (68.18%)	0.7753
<i>E. faecalis</i> n (%)	16 (48.48%)	14 (63.64%)	0.4074
<i>E. gallinarum</i> n (%)	2 (6.06%)	5 (22.73%)	0.1025
<i>E. hirae</i> n (%)	5 (15.15%)	6 (27.27%)	0.3164

by *E. hirae*, with a 1.8-fold detection increase in CVIDid (27.27%) compared to CVIDio (15.15%) ($p=0.316$).

Enterococcus gallinarum* is associated with inflammatory cytokines *in vivo* and *in vitro

In order to investigate a potential link between carriage of *E. gallinarum* and inflammation we compared previously measured²² serum cytokine levels between patients with and without *E. gallinarum* in their stool (n=5 CVIDid, n=2 CVIDio, n=2 HC with *E. gallinarum* vs n= 23 CVIDid, n=27 CVIDio, n=28 HC without *E. gallinarum*). We observed increased levels of pro-inflammatory IL-17a, hematopoietic growth factor Flt3L, immune-regulatory IL-10 and chronic activation marker LILRB4 in *E. gallinarum*+ individuals (Figure 5), supporting a relationship with inflammation and (chronic) immune activation. In addition, TWEAK (BAFF-L), ITGA11 and CLEC4G were decreased in *E. gallinarum*+ individuals.

Next, we tested the immunostimulatory capacity of *E. gallinarum* and *E. hirae* strains isolated from CVIDid patients *in vitro*. Bacterial supernatants from *E. gallinarum* and *E. hirae* were cocultured with primary monocytes isolated from CVIDid (n=3), CVIDio (n=3) or HC (n=3). IL-6 and IL-10 were measured to reflect pro- or anti-inflammatory cytokine production by monocytes, respectively, as these cytokines were previously found increased in CVID(id)^{22,42}. Stimulation with *E. gallinarum* resulted in higher pro-inflammatory IL-6 production than after stimulation with *E. coli* or *E. hirae*, especially in HC monocytes (Figure 6A). *E. gallinarum* stimulation also resulted in production of immune regulatory IL-10, which also occurred after stimulation with *E. coli* (Figure 6B). The ratio between IL-6 and IL-10 levels, which reflects the balance between pro- and anti-inflammatory signals, was high in both *E. gallinarum* and *E. hirae*, and low for *E. coli* (Figure 6C). Overall, *E. gallinarum* caused the most consistent IL-6 upregulation in combination with an increased IL-6/IL-10 ratio, indicating that *E. gallinarum* can induce a pro-inflammatory state in monocytes. We also observed that HC monocytes produced more IL-6 and IL-10 following stimulation than when compared to cells from CVIDid patients regardless of the bacterial stimulus, with CVIDio being the intermediate group (Supplementary Figure 5A).

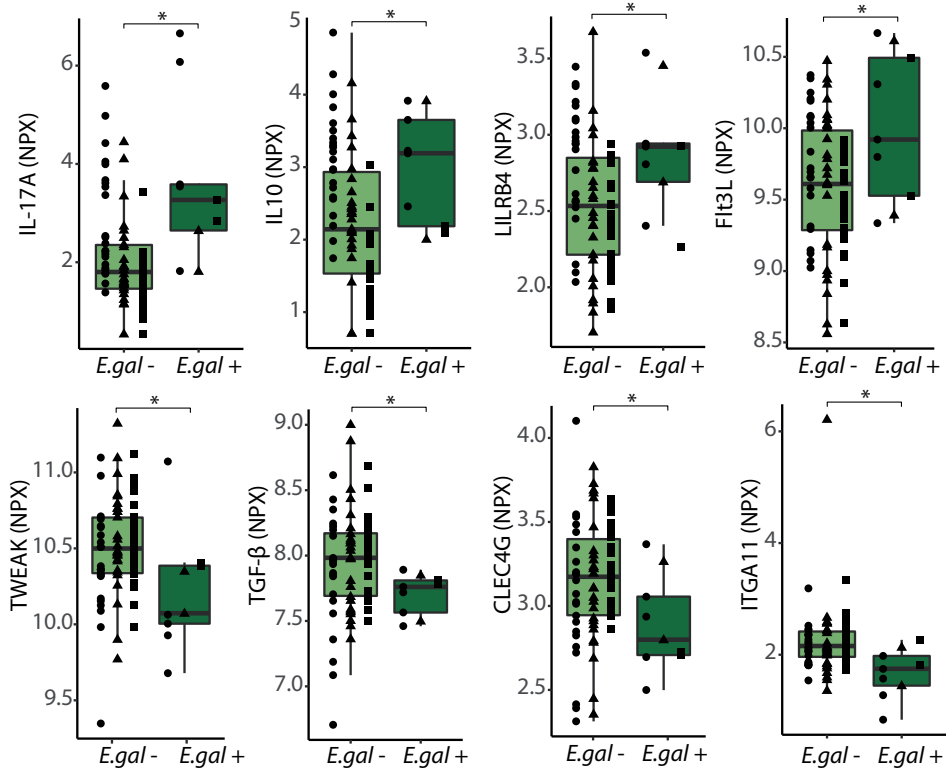


Figure 5: Serum cytokines as measured using targeted proteomics. NPX: normalized protein expression. E.gal+: n= 5 CVIDid, n=2 CVIDio, n=2 HC with *E.gallinarum* in their stools at time of sampling. E.gal-: n= 23 CVIDid, n=27 CVIDio, n=28 HC without *E.gallinarum* in their stools at time of sampling. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within $1.5 \times$ interquartile range. P-values: Mann-Whitney U-test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

DISCUSSION

Recent microbiome studies have supported the presence of gut microbial dysbiosis in CVID patients^{8,43,44}, but the role of specific gut bacteria in the occurrence of immune dysregulation in CVID was thus far unknown. In this study, we demonstrated that increased contact of the gut microbiota with the colon epithelium occurs in some CVID patients, and that the microbiota of CVIDid patients was characterized by an increased bacterial load, decreased alpha diversity, distinct beta diversity, and enrichment with enterococcal species. We identified *E. gallinarum* as a candidate pathobiont in CVIDid patients by showing its association with pro-inflammatory cytokines in serum of patients, as well as its ability to stimulate monocytes *in vitro*. *E. gallinarum* has been previously linked with other immune dysregulation syndromes such as SLE¹³, autoimmune hepatitis¹³, and primary sclerosing cholangitis⁴¹.

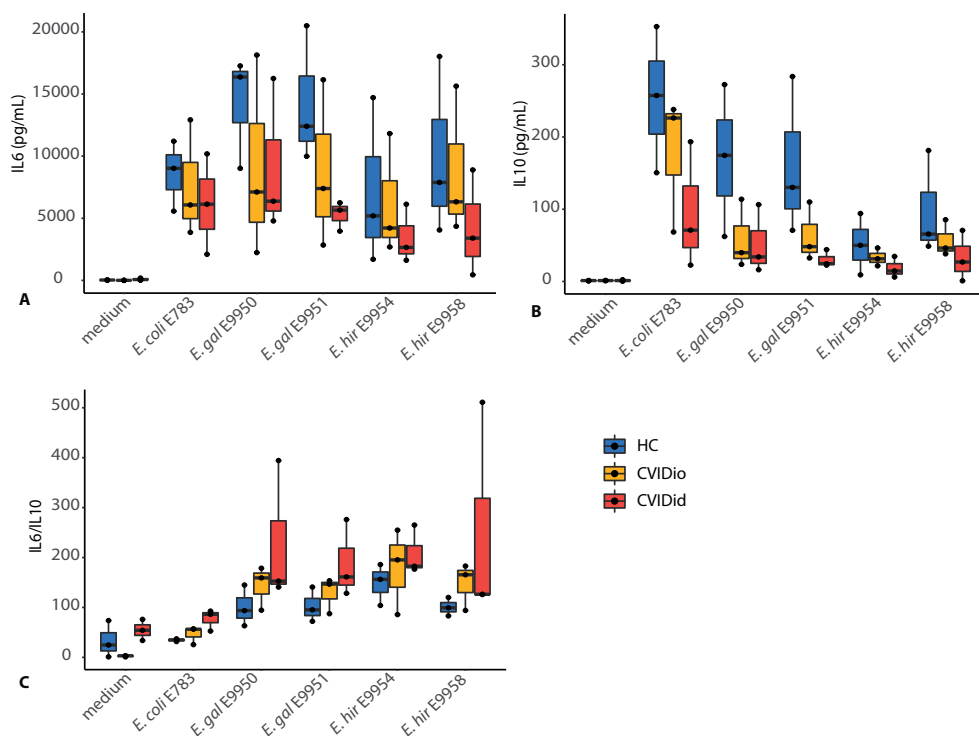


Figure 6:

Cytokine production after *in-vitro* stimulation of monocytes from 3 HC, 3 CVIDio and 3 CVIDid with bacterial supernatants from *E.coli* E783, *E.gallinarum* E9950, *E.gallinarum* E.9951, *E.hirae* E9954 and *E.hirae* E9958. A: IL-6 production, B: IL-10 production, C: IL6/IL10 ratio.

The bacterial crypt invasion we observed in some IgA-deficient CVID and XLA patients suggests that there is increased contact between the gut microbiota and the gut epithelium. This may trigger local inflammation³⁸, and contribute to systemic inflammation if it results in the previously reported endotoxemia in CVID(id) patients^{8,45-47}. Increased contact of the microbiota with the colonic epithelium has been reported in inflammatory bowel diseases²¹ as well as colonic cancer⁴⁸, but to our knowledge this has not previously been described in the context of antibody deficiency. Biopsies only provide local snapshots of processes occurring in the gut, and therefore the true incidence of bacterial crypt invasion in CVID may be higher than we were able to observe here.

The histological findings in gut biopsies of CVID patients in our study were comparable to that observed in MUC2 deficient mice with crypt elongation and bacterial crypt invasion³⁸ and it is possible that defects in mucus production may also occur in CVID. Unfortunately, the fixation method of the biopsies available for this study were not suitable to directly investigate the mucus layer in order to further investigate that question. In addition, IgA may play a role in maintaining sterility of the inner mucus layer. An association between mucosal IgA levels and CVID gut inflammation was previously reported⁴⁴,

and mechanistic studies have provided examples of how interactions between IgA and gut commensals influence the localization of bacteria in the gut⁴⁹.

Our 16S rRNA sequencing findings confirm and expand on previous studies investigating the CVID microbiome compared to HC. Decreased alpha diversity and distinct beta diversity have consistently been observed^{8,43,50}, as well as an association of decreased *Bifidobacterium*⁸ and increased Enterobacteriaceae^{8,43} and *Eggerthella*³⁹ in CVID. Fiedorová et al.⁴³ also reported enrichment of *Enterococcus* in CVID, but do not specify whether this was more present in patients with immune dysregulation symptoms. Differences in detection of minority species such as enterococci by DNA-based approaches between studies may be influenced by the specific DNA isolation protocol used, especially in the case of Gram-positive bacteria, which have a strong cell wall that is resistant to many lysis methods⁵¹. Additionally, selective culturing yielded a far larger prevalence of *Enterococcus* species in our cohort compared to 16S rRNA sequencing, especially of the dominant taxa *E. faecalis* and *E. faecium*.

We next showed increased prevalence of *Enterococcus* species in the stools of CVIDid patients, which was most pronounced for *E. gallinarum*. The presence of *Enterococcus* species in the CVIDid microbiota may be caused by several factors. First, IgA may play a role in preventing colonization of enterococci. While little is known about *E. gallinarum* in the context of IgA deficiency, other *Enterococcus* species have been better researched. Our group has previously shown that agglutination of *E. faecium* with secretory IgA was necessary to maintain spatial segregation of *E. faecium* from the intestinal epithelium in mice²⁰, and human IgA+ memory B-cells have shown high binding frequencies for *E. faecalis*⁵². Secondly, CVID patients frequently require antibiotic therapy, especially before they are diagnosed and adequate immunoglobulin replacement therapy is started. Antibiotic use may lead to generalized gut microbial dysbiosis, thereby allowing colonization with pathobiont species⁵³, especially those that easily acquire antibiotic resistance such as enterococci^{54,55}.

Carriage of *E. gallinarum* was associated with serum markers we previously found to be increased in CVIDid, such as IL-17A, IL-10 and LILRB4²². In our study, exposure of monocytes to supernatant of *E. gallinarum* strains from CVIDid patients resulted in production of IL-6, and an increased IL-6/IL-10 ratio, indicating monocyte activation. Of note, monocytes isolated from CVIDid patients produced lower levels of IL-6 and IL-10 than monocytes from HC, regardless of the bacterial stimulus. Monocyte hyporesponsiveness to Toll-like receptor (TLR)-mediated activation has been previously described in settings of endotoxemia^{56–58}. Whether this also occurs in CVIDid will need to be confirmed in additional studies with larger sample size.

There are several mechanisms through which monocyte activation by *E. gallinarum* could occur. Vieira et al.¹³ demonstrated triggering of the aryl hydrocarbon receptor (AhR) by *E. gallinarum* in SLE and autoimmune hepatitis patients, which resulted in IL-6 production. However, we were unable to abrogate monocyte activation with an AhR antagonist *in vitro* (data not shown), and AhR activation was also shown to induce Th17 skewing, which does not usually occur in CVIDid patients^{59,60}. Lauté-Caly et al.³⁹ showed that the flagellin

of an *E. gallinarum* strain isolated from healthy humans strongly stimulated monocytes, dendritic cells, and human intestinal epithelial cells through TLR5 signaling. The *fliC* (flagellin) allele of this commensal strain was identical to that of 3/3 *E. gallinarum* strains we were able to culture from patients with CVIDid, and 1/4 strains from CVIDio patients, while the other 3/4 CVIDio and 2/2 HC *E. gallinarum* strains encoded different *fliC* alleles harboring several mutations (Supplementary Figure 6). However, whether flagellin-TLR5 signaling also contributes to clinical inflammation in CVIDid remains to be proven with additional mechanistic studies.

In summary, our results support the hypothesis that in CVID, increased contact of the microbiota with the host epithelium and microbial dysbiosis contribute to immune dysregulation. We hypothesize that microbial dysbiosis, possibly in combination with IgA deficiency and/or antibiotic use, may facilitate colonization of pathobionts such as *E. gallinarum*, and that sustained exposure to *E. gallinarum* bacterial products may lead to innate immune system activation, which can provide an extra stimulus to an already autoimmunity-prone immune system⁶¹. Further studies are necessary to confirm the role of IgA deficiency and pathobionts such as *E. gallinarum* in the pathogenesis of immune dysregulation in CVID. Therapeutic targeting of gut pathobionts may be a promising future outlook in the prevention and treatment of immune dysregulation in CVID.

Acknowledgements

We would like to thank dr. Hermie Harmsen from the University of Groningen for sharing his FISH protocols with us, and Domenico Castigliero from the pathology biobank at the UMC Utrecht for his unfailing support in cutting gut biopsies. Finally, our thanks go out to Prof. Laurence Zitvogel and Mme Caroline Flament at the Institute Gustave-Roussy in Paris, France, who welcomed us into their lab to demonstrate their method of bacterial – monocyte – T-cell cocultures.

Funding

This study was funded by the Wilhelmina Children's Hospital Fund, the Dutch Foundation for Rare Diseases (Stichting Zeldzame Ziekten Fonds) and the Infection and Immunity research stimulus grant (UMC Utrecht).

Disclosures

PH reports research grants and personal fees from Shire/Takeda and CSL Behring. VD reports research grants and consulting fees from Shire/Takeda, Griffols, ACtelion, Novartis and CSL Behring. JM reports personal fees from Shire/Takeda. AvdV reports personal fees from Lamepro B.V.. HL reports research grants from Shire/Takeda. All other authors report no potential conflict of interest.

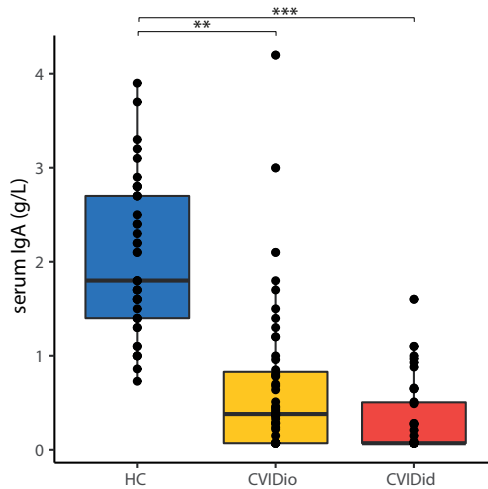
REFERENCES

1. Immunodeficiencies, E. S. for. Diagnostic Criteria PID. <https://esid.org/Education/Diagnostic-Criteria-PID> (2019).
2. Bonilla, F. A. *et al.* International Consensus Document (ICON): Common Variable Immunodeficiency Disorders. *J. Allergy Clin. Immunol. Pract.* **4**, 38–59 (2016).
3. Chapel, H. *et al.* Common variable immunodeficiency disorders: division into distinct clinical phenotypes. *Blood* **112**, 277–287 (2008).
4. Resnick, E. S., Moshier, E. L., Godbold, J. H. & Cunningham-Rundles, C. Morbidity and mortality in common variable immune deficiency over 4 decades. *Blood* **119**, 1650–1658 (2012).
5. Gathmann, B. *et al.* Clinical picture and treatment of 2212 patients with common variable immunodeficiency. *J. Allergy Clin. Immunol.* **134**, (2014).
6. Maarschalk-Ellebreek, L. J., Hoepelman, A. I. M., Van Montfrans, J. M. & Ellebreek, P. M. The spectrum of disease manifestations in patients with common variable immunodeficiency disorders and partial antibody deficiency in a university hospital. *J. Clin. Immunol.* **32**, 907–921 (2012).
7. Berbers, R. M., Nierkens, S., van Laar, J. M., Bogaert, D. & Leavis, H. L. Microbial Dysbiosis in Common Variable Immune Deficiencies: Evidence, Causes, and Consequences. *Trends Immunol.* **38**, 206–216 (2017).
8. Jorgensen, S. F. *et al.* Altered gut microbiota profile in common variable immunodeficiency associates with levels of lipopolysaccharide and markers of systemic immune activation. *Mucosal Immunol.* **9**, 1455–1465 (2016).
9. Vetrie, D. *et al.* The gene involved in X-linked agammaglobulinemia is a member of the src family of protein-tyrosine kinases. *Nature* **361**, 226–233 (1993).
10. Lougaris, V. *et al.* Long-term follow-up of 168 patients with X-linked agammaglobulinemia reveals increased morbidity and mortality. *J. Allergy Clin. Immunol.* 1–9 (2020) doi:10.1016/j.jaci.2020.03.001.
11. Dehner, C., Fine, R. & Kriegel, M. A. The microbiome in systemic autoimmune disease: Mechanistic insights from recent studies. *Curr. Opin. Rheumatol.* **31**, 201–207 (2019).
12. Scher, J. U. *et al.* Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *Elife* **2**, e01202 (2013).
13. Manfredo Vieira, S. *et al.* Translocation of a gut pathobiont drives autoimmunity in mice and humans. *Science (80-.)*. **359**, 1156–1161 (2018).
14. Lee, Y. K., Menezes, J. S., Umesaki, Y. & Mazmanian, S. K. Proinflammatory T-cell responses to gut microbiota promote experimental autoimmune encephalomyelitis. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4615–4622 (2011).
15. Fluckiger, A. *et al.* Cross-reactivity between tumor MHC class I-restricted antigens and an enterococcal bacteriophage. *Science (80-.)*. **369**, 936–942 (2020).
16. Bunker, J. J. & Bendelac, A. IgA Responses to Microbiota. *Immunity* **49**, 211–224 (2018).
17. Huus, K. E., Petersen, C. & Finlay, B. B. Diversity and dynamism of IgA–microbiota interactions. *Nat. Rev. Immunol.* **0123456789**, (2021).
18. Giovannetti, A. *et al.* Unravelling the complexity of T cell abnormalities in common variable immunodeficiency. *J. Immunol.* **178**, 3932–43 (2007).
19. Hartono, S. *et al.* Predictors of granulomatous lymphocytic interstitial lung disease in common variable immunodeficiency. *Ann. Allergy, Asthma Immunol.* **118**, 614–620 (2017).
20. Hendrickx, A. P. A. *et al.* Antibiotic-driven dysbiosis mediates intraluminal agglutination and alternative segregation of enterococcus faecium from the intestinal epithelium. *MBio* **6**, 1–11 (2015).
21. Johansson, M. E. V. *et al.* Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. *Gut* **63**, 281–291 (2014).
22. Berbers, R. M. *et al.* Targeted Proteomics Reveals Inflammatory Pathways that Classify Immune Dysregulation in Common Variable Immunodeficiency. *J. Clin. Immunol.* 362–373 (2020) doi:10.1007/s10875-020-00908-1.
23. Liu, C. M. *et al.* BactQuant: An enhanced broad-coverage bacterial quantitative real-time PCR assay. *BMC Microbiol.* **12**, 56 (2012).

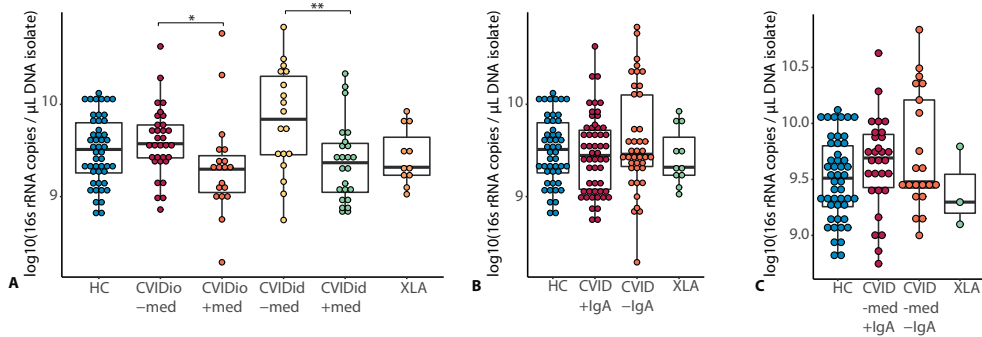
24. Fadrosh, D. W. *et al.* An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* **2**, 1–7 (2014).
25. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
26. Callahan, B. J. *et al.* DADA2 : High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–587 (2016).
27. Quast, C. *et al.* The SILVA ribosomal RNA gene database project : improved data processing and web-based tools. *Nucleic Acids Res.* **41**, 590–596 (2013).
28. Babraham Bioinformatics. Trim Galore. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
29. Minot, S., Krumm, N. & Greenfield, N. One Codex: A Sensitive and Accurate Data Platform for Genomic Microbial Identification. *bioRxiv* 027607 (2015) doi:10.1101/027607.
30. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. MetaSPAdes: A new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).
31. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
32. Et.al., C. J. E. L. Enterococcosel (ECS) agar/broth with or without vancomycin. in *Progress in Industrial Microbiology* vol. 37 465–468 (2003).
33. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
34. Daillère, R. *et al.* Enterococcus hirae and Barnesiella intestinihominis Facilitate Cyclophosphamide-Induced Therapeutic Immunomodulatory Effects. *Immunity* **45**, 931–943 (2016).
35. De Jager, W., Prakken, B. J., Bijlsma, J. W. J., Kuis, W. & Rijkers, G. T. Improved multiplex immunoassay performance in human plasma and synovial fluid following removal of interfering heterophilic antibodies. *J. Immunol. Methods* **300**, 124–135 (2005).
36. R Foundation for Statistical Computing, Vienna, A. R: A language and environment for statistical computing. (2018).
37. Lin, H. & Peddada, S. Das. Analysis of compositions of microbiomes with bias correction. *Nat. Commun.* **11**, 1–11 (2020).
38. Hansson, G. C. & Johansson, M. E. V. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Gut Microbes* **1**, 51–54 (2010).
39. Lauté-Caly, D. L. *et al.* The flagellin of candidate live biotherapeutic Enterococcus gallinarum MRx0518 is a potent immunostimulant. *Sci. Rep.* **9**, 1–14 (2019).
40. Stein-Thoeringer, C. K. *et al.* Lactose drives Enterococcus expansion to promote graft-versus-host disease. *Science (80-.)*. **366**, 1143–1149 (2019).
41. Nakamoto, N. *et al.* Gut pathobionts underlie intestinal barrier dysfunction and liver T helper 17 cell immune response in primary sclerosing cholangitis. *Nat. Microbiol.* **4**, 492–503 (2019).
42. Hultberg, J., Ernerudh, J., Larsson, M., Nilsson-Augustinsson, Å. & Nyström, S. Plasma protein profiling reflects TH1-driven immune dysregulation in common variable immunodeficiency. *J. Allergy Clin. Immunol.* 1–12 (2020) doi:10.1016/j.jaci.2020.01.046.
43. Fiedorová, K. *et al.* Bacterial but not fungal gut microbiota alterations are associated with common variable immunodeficiency (CVID) phenotype. *Front. Immunol.* **10**, (2019).
44. Shulzhenko, N. *et al.* CVID enteropathy is characterized by exceeding low mucosal IgA levels and interferon-driven inflammation possibly related to the presence of a pathobiont. *Clin. Immunol.* **197**, 139–153 (2018).
45. Le Coz, C. *et al.* Common variable immunodeficiency-associated endotoxemia promotes early commitment to the T follicular lineage. *J. Allergy Clin. Immunol.* **144**, 1660–1673 (2019).
46. Romberg, N. *et al.* Patients with common variable immunodeficiency with autoimmune cytopenias exhibit hyperplastic yet inefficient germinal center responses. *J. Allergy Clin. Immunol.* **143**, 258–265 (2019).
47. Perreau, M. *et al.* Exhaustion of bacteria-specific CD4 T cells and microbial translocation in common variable immunodeficiency disorders. *J. Exp. Med.* **211**, 2033–2045 (2014).

48. Raskov, H., Kragh, K. N., Bjarnsholt, T., Alamili, M. & Gögenur, I. Bacterial biofilm formation inside colonic crypts may accelerate colorectal carcinogenesis. *Clin. Transl. Med.* **7**, 2–5 (2018).
49. Donaldson, G. P. *et al.* Gut microbiota utilize immunoglobulin A for mucosal colonization. *Science (80-.)*. **360**, 795–800 (2018).
50. van Schewick, C. M. *et al.* Altered Microbiota, Impaired Quality of Life, Malabsorption, Infection, and Inflammation in CVID Patients With Diarrhoea. *Front. Immunol.* **11**, 1–13 (2020).
51. Hynes, W. L., Ferretti, J. J., Gilmore, M. S. & Segarra, R. A. PCR amplification of streptococcal DNA using crude cell lysates. *FEMS Microbiol. Lett.* **73**, 139–142 (1992).
52. Berkowska, M. A. *et al.* Circulating Human CD27 – IgA + Memory B Cells Recognize Bacteria with Polyreactive Igs . *J. Immunol.* **195**, 1417–1426 (2015).
53. Buffie, C. G. & Pamer, E. G. Microbiota-mediated colonization resistance against intestinal pathogens. *Nat. Rev. Immunol.* **13**, 790–801 (2013).
54. Ubeda, C. *et al.* Vancomycin-resistant Enterococcus domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans. *J. Clin. Invest.* **120**, 4332–4341 (2010).
55. Berkell, M. *et al.* Microbiota-based markers predictive of development of Clostridioides difficile infection. *Nat. Commun.* **12**, 2241 (2021).
56. de Jong, P. R. *et al.* STAT3 regulates monocyte TNF-alpha production in systemic inflammation caused by cardiac surgery with cardiopulmonary bypass. *PLoS One* **7**, 1–9 (2012).
57. Saha, D. C., Astiz, M. E., Eales-Reynolds, L. J. & Rackow, E. C. Lipopolysaccharide-and superantigen-modulated superoxide production and monocyte hyporesponsiveness to activating stimuli in Sepsis. *Shock* **38**, 43–48 (2012).
58. Flier, S. *et al.* Monocyte hyporesponsiveness and toll-like receptor expression profiles in coronary artery bypass grafting and its clinical implications for postoperative inflammatory response and pneumonia : An observational cohort study. *Eur. J. Anaesthesiol.* **32**, 177–188 (2015).
59. Edwards, E. S. J. *et al.* Predominantly Antibody-Deficient Patients With Non-infectious Complications Have Reduced Naive B, Treg, Th17, and Tfh17 Cells. *Front. Immunol.* **10**, (2019).
60. Barbosa, R. R. *et al.* Primary B-cell deficiencies reveal a link between human IL-17-producing CD4 T-cell homeostasis and B-cell differentiation. *PLoS One* **6**, (2011).
61. Roskin, K. M. *et al.* IgH sequences in common variable immune deficiency reveal altered B cell development and selection. *Sci. Transl. Med.* **7**, (2015).

SUPPLEMENTARY MATERIAL



Supplementary Figure 1: serum IgA in healthy controls (HC, n=48), CVID with infections only (CVIDio, n=51) and CVID with immune dysregulation (CVIDio n=42). The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within $1.5 \times$ interquartile range. P values boxplots: Mann-Whitney U test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.



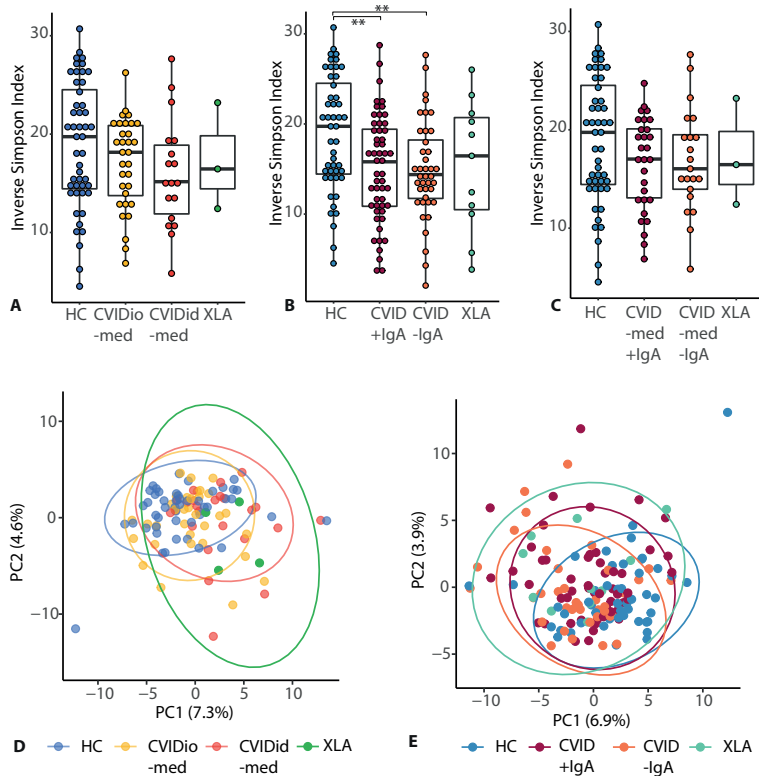
Supplementary Figure 2: Bacterial load as determined using 16S rRNA qPCR in:

A: HC (n=48), CVIDio -med (n=32), CVIDio +med (n=19), CVIDid -med (n=18), CVIDid +med (n=24), XLA (n=11).

B: HC (n=48), CVID +IgA (n=53), CVID -IgA (n=40), XLA (n=11).

C: HC (n=48), CVID -med +IgA (n=29), CVID -med -IgA (n=21), XLA -med (n=3)

HC: healthy control, CVIDio: CVID with infections only, CVIDid: CVID with immune dysregulation, XLA: X-linked agmmaglobulinemia. Med: patients who did (+) or did not (-) use antibiotics or immunosuppressive therapy up to 3 months prior to sampling. IgA: patients with serum IgA higher (+) or lower (-) than 0.1 g/L. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within $1.5 \times$ interquartile range. P values boxplots: Mann-Whitney U test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.



Supplementary Figure 3:

Alpha diversity as calculated using the inverse Simpson index. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within $1.5 \times$ interquartile range. P values boxplots: Mann-Whitney U test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

A: HC (n=48), CVIDio -med (n=32), CVIDid -med (n=18), XLA-med (n=3).

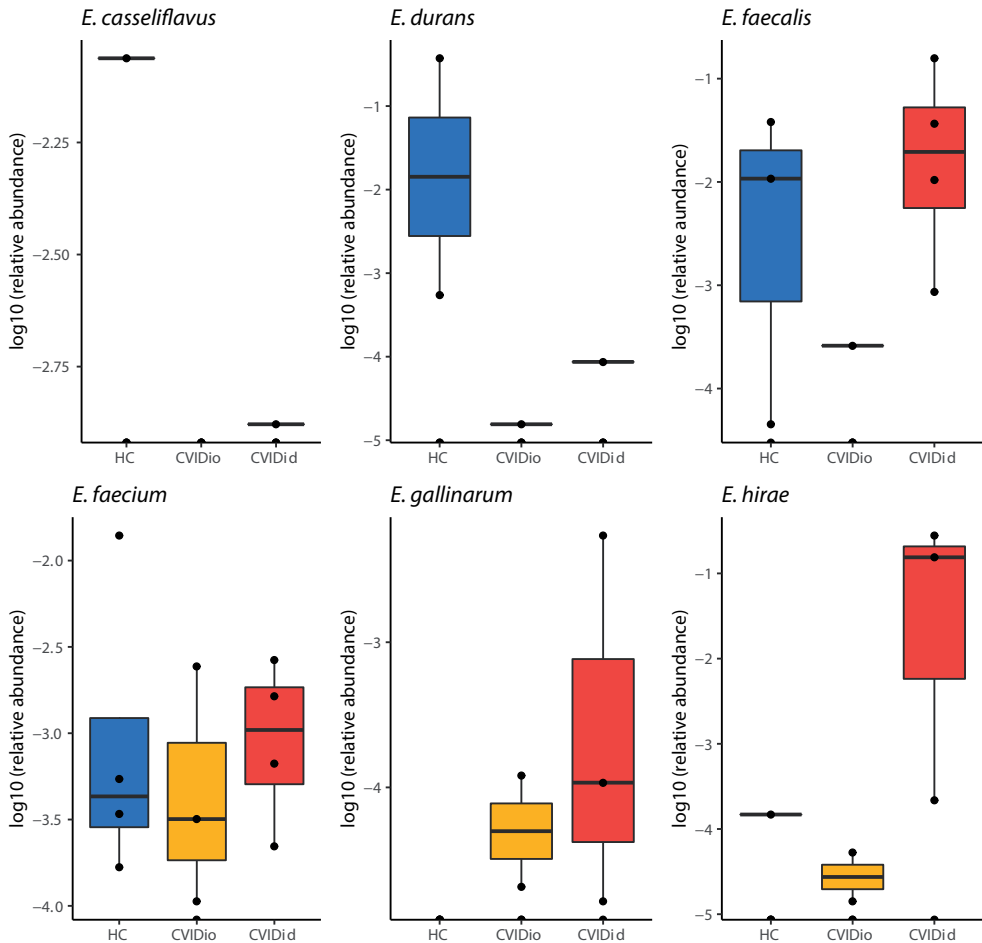
B: HC (n=48), CVID +IgA (n=53), CVID-IgA (n=40), XLA (n=11).

C: HC (n=48), CVID -med +IgA (n=29), CVID -med -IgA (n=21), XLA -med (n=3)

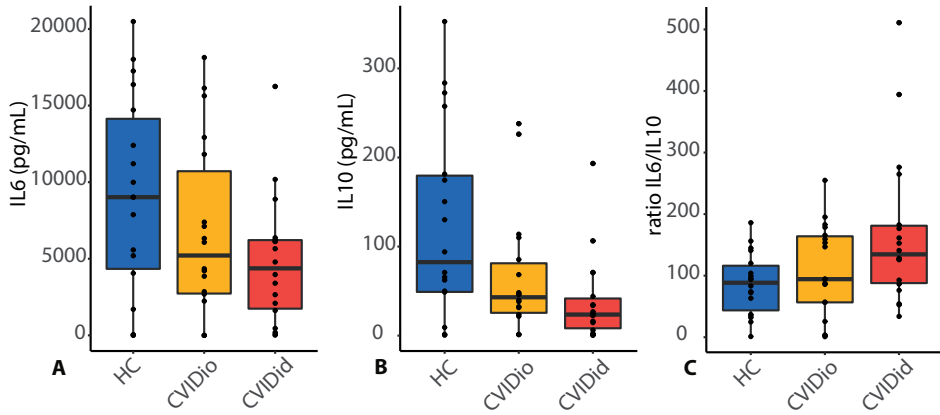
Beta diversity shown as principal component analysis (PCA) on genus-level hyperbolic arc-sine transformed data. Ellipses indicate the 95% confidence interval.

D: HC (n=48), CVIDio -med (n=32), CVIDid -med (n=18), XLA-med (n=3). FDR-adjusted PERMANOVA CVIDio-med vs CVIDid-med $p_{\text{adj}}=0.009$. CVIDid vs HC $p_{\text{adj}}=0.006$

E: HC (n=48), CVID +IgA (n=53), CVID-IgA (n=40), XLA (n=11). FDR-adjusted PERMANOVA: CVID+IgA vs HC $p_{\text{adj}}=0.006$, CVID-IgA vs HC $p_{\text{adj}}=0.006$

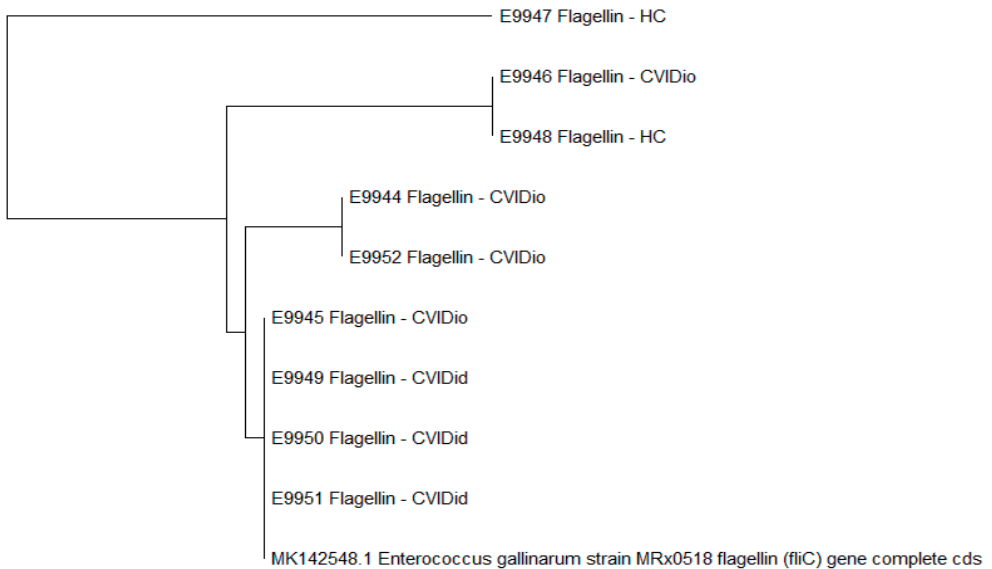


Supplementary Figure 4: Relative abundance of the most prevalent enterococcal species determined using shotgun metagenomic sequencing in 4 HC, 4 CVIDio and 4 CVIDid. The horizontal line inside the box represents the median. The whiskers represent the lowest and highest values within $1.5 \times$ interquartile range.



Supplementary Figure 5:

Cytokine production after *in-vitro* stimulation of monocytes from 3 HC, 3 CVIDio and 3 CVIDid with bacterial supernatants from *E.coli* E783, *E.gallinarum* E9950, *E.gallinarum* E.9951, *E.hirae* E9954 and *E.hirae* E9958 (pooled data). A: IL-6 production, B: IL-10 production, C: IL-6/IL10 ratio.



Supplementary Figure 6: Neighborhood-joining tree comparing the *fliC* alleles of *E. gallinarum* strains cultured from HC, CVIDio and CVIDid patients to the strain described in Lauté-Caly et al.

Supplementary Table 1: qPCR primer-probe sequences, ordered from IDT DNA technologies. All sequences reported from 5'-3', all probes flanked by 6-FAM/ZEN (5') and Iowa Black®Fluorescence Quencher².

<i>BactQuant</i>	
Target gene	16S rRNA
Forward primer	CCTACGGGDGGCWGCA
Probe	CAGCAGCCGCGGTA
Reverse Primer	GGACTACHVGGTMTCTAATC
<i>Enterococcus faecium</i>	
Target gene	<i>ddl</i>
Forward primer	TGCCTGGTGAAGTCGTAAGAAG
Probe	CGAAATGCAGATTCCAGCCGAAGTG
Reverse Primer	AGCTAACTTCGCGTACTCTTG
<i>Enterococcus faecalis</i>	
target gene	<i>ddl</i>
Forward primer	GCACGTGAAATTGAAGTAGCC
Probe	TGGTGAAGTGGTGAAGATGTCGCT
Reverse Primer	GAACATGCGCTGGGATTTG
<i>Enterococcus gallinarum</i>	
Target gene	<i>sodA</i>
Forward primer	TTTGATTCGGTGCCTGAAGA
Probe	AACGGTGGTGGTCATGCAAATCAC
Reverse Primer	AGCATTGGTGCCAAGATTC
<i>Enterococcus hirae</i>	
target gene	<i>ddl</i>
Forward primer	CAGTTCTTTCAGCGTTTTTCAGTC
Probe	AGGGCCTTTCACCCATTGACCTTC
Reverse Primer	CTTTGCTCGTTGGTTTCTCTG

Supplementary Table 2: Overview of bacterial strains used in this study

Identifier	Species	Isolation source	Reference	Use in this study
E06019	<i>E. faecium</i>	Stool hospitalized patient	Arredondo-Alonso et al. ³	Validation species-specific RT-PCR
E06007	<i>E. faecium</i>	Stool hospitalized patient	Wagner et al. ⁴	Validation species-specific RT-PCR
E07089	<i>E. gallinarum</i>	Stool hospitalized patient	In-house strain collection	Validation species-specific RT-PCR
E06960	<i>E. gallinarum</i>	Stool hospitalized patient	In-house strain collection	Validation species-specific RT-PCR
E00277	<i>E. hirae</i>	Stool human community	In-house strain collection	Validation species-specific RT-PCR
E00134	<i>E. hirae</i>	Stool human community	In-house strain collection	Validation species-specific RT-PCR
E04246	<i>E. faecalis</i>	Stool hospitalized patient	In-house strain collection	Validation species-specific RT-PCR
E04665	<i>E. faecalis</i>	Stool hospitalized patient	In-house strain collection	Validation species-specific RT-PCR
E783	<i>E. coli</i>	Bloodstream infection	In-house strain collection	Bacterial - immune cell cocultures
E09950	<i>E. gallinarum</i>	Stool CVIDid	This study	Bacterial - immune cell cocultures, <i>fljC</i> analysis
E09951	<i>E. gallinarum</i>	Stool CVIDid	This study	Bacterial - immune cell cocultures, <i>fljC</i> analysis
E09954	<i>E. hirae</i>	Stool CVIDid	This study	Bacterial - immune cell cocultures, <i>fljC</i> analysis
E09958	<i>E. hirae</i>	Stool CVIDid	This study	Bacterial - immune cell cocultures, <i>fljC</i> analysis
E09947	<i>E. gallinarum</i>	Stool HC	This study	<i>fljC</i> analysis
E09946	<i>E. gallinarum</i>	Stool CVIDio	This study	<i>fljC</i> analysis
E09948	<i>E. gallinarum</i>	Stool HC	This study	<i>fljC</i> analysis
E09944	<i>E. gallinarum</i>	Stool CVIDio	This study	<i>fljC</i> analysis
E09952	<i>E. gallinarum</i>	Stool CVIDio	This study	<i>fljC</i> analysis
E09945	<i>E. gallinarum</i>	Stool CVIDio	This study	<i>fljC</i> analysis
E09949	<i>E. gallinarum</i>	Stool CVIDid	This study	<i>fljC</i> analysis

Supplementary table 3: Characteristics biopsy cohort. HC: healthy control, CVID: Common Variable Immunodeficiency, XLA: X-linked agammaglobulinemia.

	HC	CVID	XLA
n	9	15	3
Age (years, median)	33	38	28
Sex (male)	2 (22%)	6 (40%)	3 (100%)
Antibiotic use	0	7 (47%)	0
Serum IgA < 0.1 g/L	n.d.	11 (73%)	3 (100%)
Histology colon biopsies			
No pathology	9	4 (26%)	1 (33%)
Lymphoid aggregates	0	11 (73%)	1 (33%)
Focal signs of chronic inflammation	0	1 (7%)	1 (33%)
Acute inflammation or infection	0	0	0
Immune dysregulation symptoms			
Pulmonary	0	3 (30%)	0
Hematological	0	2 (13%)	0
Gastrointestinal	0	2 (13%)#	0
Rheumatological	0	1 (7%)	0
Lymphoproliferative	0	7 (47%)	0

#: both autoimmune gastritis

Supplementary table 4: Differentially abundant bacteria identified using 16S rRNA sequencing. CVID without use of antibiotics or immunosuppressive therapy 3 months prior to sampling (n=50) vs healthy control (HC, n=48): coefficients; standard errors (SE), p-values, FDR-adjusted p-values. Median relative abundance of non-zero counts per group. Statistics: ANCOM-BC. Black: p-adj.<0.05. Grey: p-adj.<0.05 but p-adj.>0.05.

	Beta CVID vs HC	SE CVID vs HC	p-value CVID vs HC	adj. p-value CVID vs HC	Beta age	SE age	p-value age	adj. p-value age	Beta sex	SE sex	p-value sex	adj. p-value sex
<i>Eubacteriaceae</i>	-0.156	0.067	0.000	0.000	-0.003	0.006	0.887	0.852	0.369	0.205	0.072	0.439
<i>Lactococcus</i>	-0.138	0.058	0.000	0.000	-0.012	0.005	0.008	0.175	-0.019	0.171	0.912	0.936
<i>Sellimonas</i>	-0.301	0.089	0.000	0.000	-0.016	0.009	0.068	0.445	-0.120	0.290	0.678	0.893
<i>Tyzzerella</i>	-0.197	0.085	0.000	0.000	0.002	0.008	0.830	0.974	-0.444	0.251	0.077	0.332
<i>Anaerotruncus</i>	-0.152	0.062	0.000	0.000	-0.001	0.005	0.804	0.965	-0.039	0.172	0.820	0.913
<i>Caproiciproducens</i>	-0.119	0.059	0.000	0.000	-0.021	0.006	0.000	0.016	-0.166	0.180	0.356	0.674
<i>Veillonella</i>	-0.204	0.110	0.000	0.000	-0.012	0.012	0.300	0.727	-0.577	0.353	0.102	0.401
<i>Escherichia-Shigella</i>	-0.468	0.144	0.001	0.062	0.005	0.013	0.700	0.921	-0.316	0.430	0.463	0.793
<i>Enterobacteriaceae</i>	-0.442	0.145	0.002	0.082	0.011	0.013	0.382	0.794	0.037	0.455	0.952	0.942
<i>Tyzzerella_3</i>	-0.285	0.098	0.004	0.128	0.000	0.009	0.979	0.999	-0.456	0.294	0.121	0.427
<i>D_0_Bacteria;D_1__Proteobacteria</i>	-0.321	0.116	0.006	0.040	0.024	0.012	0.035	0.122	0.104	0.351	0.767	0.858
<i>D_0_Bacteria;D_1__Proteobacteria;</i> <i>D_2__Gammaproteobacteria;D_3__Enterobacteriales</i>	-0.401	0.146	0.006	0.070	0.013	0.013	0.331	0.516	0.201	0.433	0.643	0.911
<i>Ruminococcaceae UCG-004</i>	-0.312	0.115	0.007	0.128	-0.019	0.010	0.067	0.445	0.636	0.330	0.054	0.279
<i>D_0_Bacteria;D_1__Proteobacteria;</i> <i>D_2__Gammaproteobacteria</i>	-0.348	0.131	0.008	0.087	0.020	0.012	0.110	0.286	0.259	0.392	0.509	0.885
<i>Parasutterella</i>	-0.234	0.088	0.008	0.128	-0.003	0.008	0.724	0.921	0.041	0.263	0.876	0.931
<i>Negativibacillus</i>	-0.249	0.095	0.009	0.128	0.013	0.009	0.148	0.567	0.087	0.273	0.749	0.894
<i>D_0_Bacteria;D_1__Actinobacteria;</i> <i>D_2__Actinobacteria;D_3__Bifidobacteriales</i>	0.354	0.137	0.010	0.070	-0.042	0.013	0.002	0.024	-0.219	0.414	0.596	0.911
<i>Eggerthella</i>	-0.258	0.102	0.011	0.128	-0.021	0.010	0.034	0.384	0.345	0.311	0.267	0.590
<i>Prevotella_9</i>	0.392	0.156	0.012	0.128	-0.027	0.014	0.047	0.419	-0.966	0.455	0.034	0.224
<i>uncultured</i>	-0.278	0.112	0.013	0.128	0.002	0.010	0.857	0.980	0.117	0.329	0.722	0.894
<i>D_0_Bacteria;D_1__Actinobacteria;D_2__Actinobacteria</i>	0.339	0.137	0.013	0.087	-0.044	0.013	0.001	0.012	-0.175	0.411	0.670	0.885
<i>Marvinbryantia</i>	0.259	0.106	0.015	0.128	0.005	0.009	0.576	0.863	0.647	0.303	0.033	0.224

	Beta CVID vs HC	SE CVID vs HC	p-value CVID vs HC	adj. p-value CVID vs HC	Beta age	SE age	p-value age	adj. p-value age	Beta sex	SE sex	p-value sex	adj. p-value sex
<i>Ruminococcus 2</i>	-0.389	0.162	0.016	0.128	-0.001	0.015	0.974	0.999	-0.128	0.472	0.786	0.894
<i>Eubacterium 1</i>	-0.188	0.078	0.016	0.128	-0.003	0.007	0.633	0.912	0.095	0.234	0.687	0.893
<i>Ruminococcus 3</i>	0.327	0.138	0.018	0.128	-0.013	0.013	0.325	0.727	-0.425	0.425	0.318	0.635
<i>Oscillibacter</i>	-0.232	0.101	0.022	0.138	-0.009	0.009	0.326	0.727	-0.374	0.294	0.203	0.486
<i>Bifidobacteriaceae</i>	0.312	0.137	0.023	0.206	-0.043	0.013	0.001	0.044	-0.383	0.416	0.358	0.769
<i>Blautia</i>	-0.180	0.079	0.023	0.140	-0.008	0.007	0.225	0.675	0.233	0.215	0.279	0.593
<i>Flavonifractor</i>	-0.211	0.094	0.025	0.141	-0.009	0.009	0.288	0.727	-0.117	0.283	0.680	0.893
<i>Prevotellaceae</i>	0.365	0.166	0.028	0.206	-0.014	0.014	0.327	0.794	-0.967	0.494	0.050	0.439
<i>Coriobacteriales Incertae Sedis</i>	-0.249	0.114	0.029	0.206	0.007	0.010	0.499	0.828	0.348	0.345	0.312	0.769
<i>Eubacterium 3</i>	-0.228	0.106	0.031	0.163	-0.010	0.009	0.259	0.727	0.112	0.306	0.714	0.894
<i>Eubacterium 4</i>	-0.211	0.098	0.032	0.163	0.007	0.010	0.523	0.856	-0.040	0.306	0.897	0.931
<i>D_0_Bacteria;D_1_Firmicutes</i>	-0.131	0.061	0.033	0.115	-0.002	0.006	0.773	0.901	0.083	0.172	0.629	0.858
<i>Erysipelatoclostridium</i>	-0.152	0.072	0.033	0.163	-0.007	0.007	0.303	0.727	-0.403	0.223	0.070	0.320
<i>Paraprevotella</i>	0.197	0.095	0.039	0.181	-0.007	0.008	0.374	0.744	-0.957	0.289	0.001	0.071
<i>D_0_Bacteria;D_1_Tenerificutes;D_2_Mollificutes</i>	0.157	0.077	0.040	0.143	-0.006	0.008	0.448	0.627	0.126	0.235	0.591	0.911
<i>D_3_Mollificutes RF39</i>												
<i>D_0_Bacteria;D_1_Firmicutes;D_2_Negativificutes</i>	0.199	0.097	0.041	0.143	-0.012	0.009	0.175	0.343	-0.106	0.280	0.705	0.911
<i>D_3_Selenomonadales</i>												
<i>Bifidobacterium</i>	0.283	0.138	0.041	0.181	-0.048	0.013	0.000	0.016	-0.614	0.418	0.142	0.448
<i>Deffluviitaleaceae UCG-011</i>	-0.158	0.068	0.042	0.181	-0.006	0.006	0.330	0.727	0.063	0.199	0.752	0.894
<i>D_0_Bacteria;D_1_Firmicutes;D_2_Negativificutes</i>	0.196	0.097	0.043	0.150	-0.014	0.009	0.106	0.286	-0.091	0.278	0.743	0.885
<i>D_0_Bacteria;D_1_Tenerificutes;D_2_Mollificutes</i>	0.164	0.082	0.046	0.150	-0.004	0.008	0.663	0.861	0.255	0.243	0.295	0.885

Supplementary table 5: Differentially abundant bacteria identified using 16S rRNA sequencing. CVID with immune dysregulation (CVIDid, n=18) versus CVID with infections only (CVIDio, n=32). Patients did not use antibiotics or immunosuppressive therapy 3 months prior to sampling; coefficients, standard errors (SE), p-values, FDR-adjusted p-values. Median relative abundance of non-zero counts per group. Statistics: ANCOM-BC.

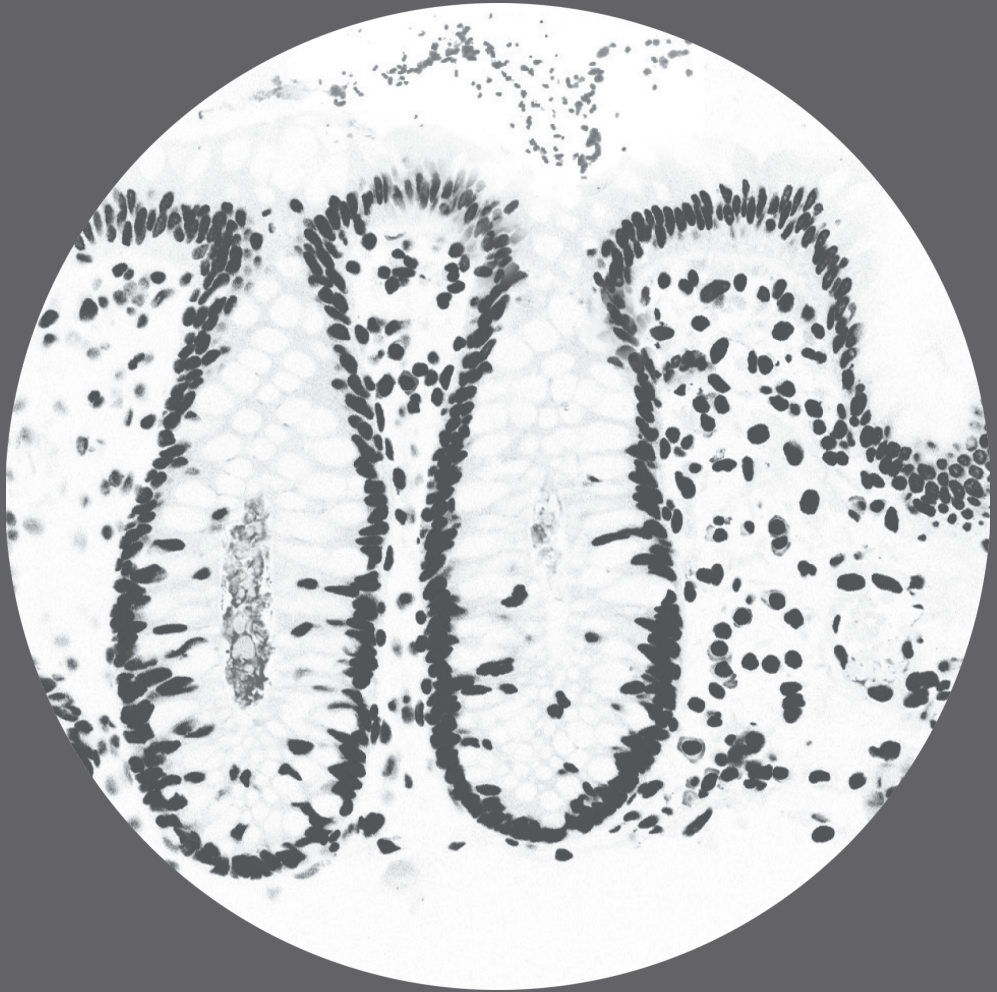
	Beta CVIDid	SE CVIDid	p-value CVIDid	adj. p-value CVIDid	Beta age	SE age	p-value age	adj. p-value age	Beta sex	SE sex	p-value sex	adj. p-value sex
<i>Enterococcaceae</i>	-0.998	0.407	0.000	0.000	-0.011	0.011	0.305	0.840	-0.082	0.365	0.823	0.933
<i>Peptococcaceae</i>	-0.039	0.258	0.000	0.000	0.008	0.008	0.332	0.840	0.221	0.249	0.375	0.708
<i>uncultured.2</i>	-0.085	0.278	0.000	0.000	0.004	0.007	0.607	0.840	-0.266	0.274	0.330	0.661
<i>Pasteurellaceae</i>	-0.269	0.303	0.000	0.000	-0.005	0.008	0.545	0.840	-0.181	0.252	0.473	0.759
<i>D_0__Bacteria;D_1__Proteobacteria;</i>	-0.016	0.297	0.000	0.000	0.005	0.008	0.487	0.852	-0.052	0.284	0.854	0.903
<i>D_2__Alphaproteobacteria;D_3__Rhodospirillales</i>												
<i>D_0__Bacteria;D_1__Proteobacteria;</i>	-0.200	0.320	0.000	0.000	-0.003	0.009	0.708	0.905	0.033	0.270	0.903	0.903
<i>D_2__Gammaproteobacteria;D_3__Pasteurellales</i>												
<i>D_0__Bacteria;D_1__Proteobacteria;</i>	-0.053	0.293	0.000	0.000	0.004	0.007	0.555	0.879	-0.265	0.280	0.344	0.825
<i>D_2__Alphaproteobacteria</i>												
<i>Enterococcus</i>	-0.102	0.431	0.000	0.000	-0.013	0.011	0.245	0.792	-0.168	0.373	0.210	0.440
<i>Eubacterium 2</i>	-0.243	0.388	0.000	0.000	0.009	0.009	0.303	0.836	-0.204	0.380	0.591	0.776
<i>Anaerotruncus</i>	-0.192	0.273	0.000	0.000	0.003	0.006	0.654	0.907	-0.117	0.261	0.653	0.816
<i>Veillonella</i>	-0.122	0.574	0.000	0.000	-0.013	0.017	0.461	0.880	-0.559	0.553	0.312	0.537
<i>gut metagenome.1</i>	-0.101	0.246	0.000	0.000	0.002	0.006	0.739	0.919	-0.653	0.237	0.006	0.066
<i>Haemophilus</i>	-0.286	0.296	0.000	0.000	-0.007	0.008	0.393	0.880	-0.568	0.231	0.014	0.082

Supplementary table 6: Differentially abundant bacteria identified using 16S rRNA sequencing. CVID with IgA <0.1 g/L (n=40) versus CVID with IgA >0.1g/L (n=53): coefficients, standard errors (SE), p-values, FDR-adjusted p-values. Median relative abundance of non-zero counts per group. Statistics: ANCOM-BC.

	Beta CVID-IgA - CVID+IgA	SE CVID-IgA - CVID+IgA	p-value CVID-IgA - CVID+IgA	adj. p-value CVID-IgA - CVID+IgA	Beta age	SE age	p-value age	adj. p-value age	Beta sex	SE sex	p-value sex	adj. p-value sex
<i>Enterococcaceae</i>	-0.471	0.303	0	0	-0.003	0.008	0.670	0.938	0.090	0.286	0.753	0.988
<i>Enterococcus</i>	-0.595	0.3057	0	0	-0.004	0.008	0.577	0.982	-0.049	0.284	0.863	0.927
<i>Eubacterium 2</i>	-0.866	0.328	0	0	0.004	0.007	0.633	0.982	0.152	0.294	0.605	0.816

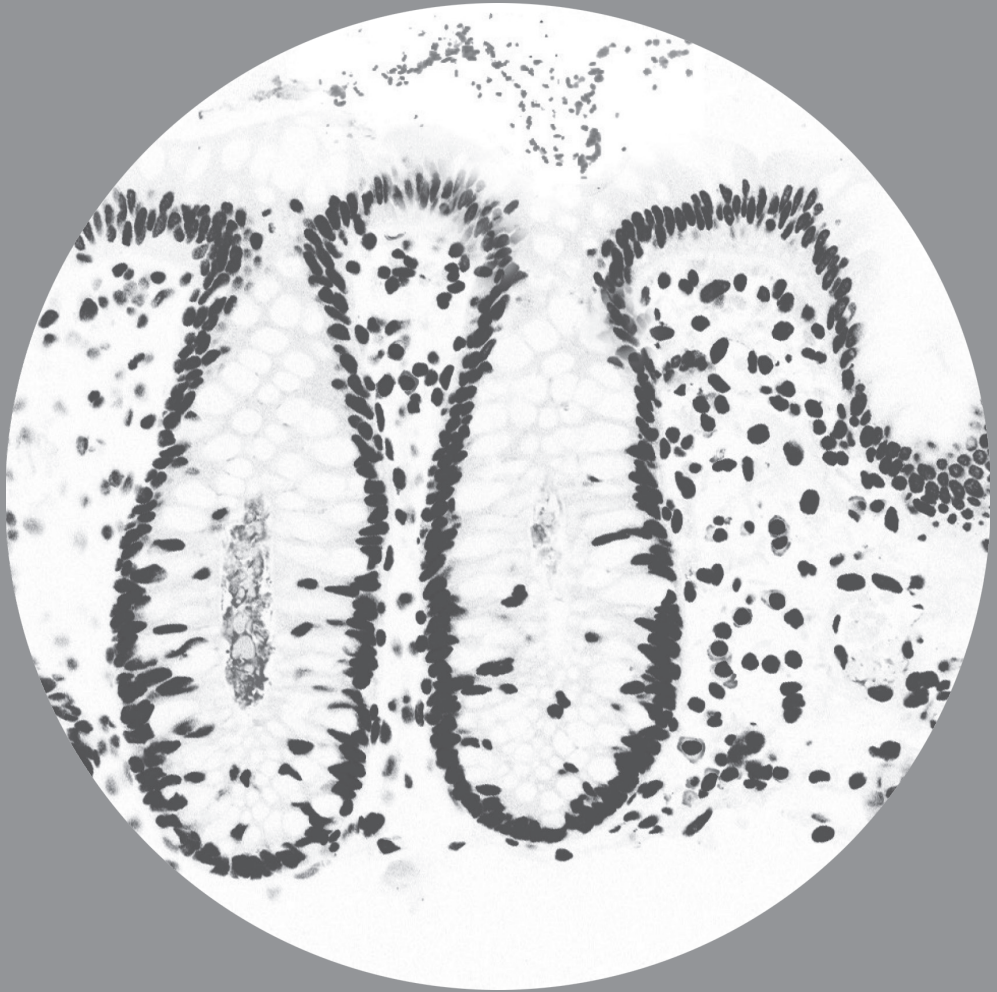
REFERENCES SUPPLEMENTARY MATERIAL

1. Lauté-Caly, D. L. *et al.* The flagellin of candidate live biotherapeutic *Enterococcus gallinarum* MRx0518 is a potent immunostimulant. *Sci. Rep.* **9**, 1–14 (2019).
2. Liu, C. M. *et al.* BactQuant : An enhanced broad-coverage bacterial quantitative real-time PCR assay. *BMC Microbiol.* **12**, 56 (2012).
3. Arredondo-Alonso, S. *et al.* Plasmids shaped the recent emergence of the major nosocomial pathogen *Enterococcus faecium*. *MBio* **11**, 1–17 (2020).
4. Wagner, T. M. *et al.* *Enterococcus faecium* TIR-Domain Genes Are Part of a Gene Cluster Which Promotes Bacterial Survival in Blood. *Int. J. Microbiol.* (2018).



Part III

Analysing Microbiomes



Chapter 8

Handling compositionality and zeros
in sparse microbiome data – is the
cure worse than the disease?

Berbers, RM
Gu, Z
El-Bouhaddani, S
Sprockel, Colino
Sven Kleine Bardenhorst
Paganelli, FL
Willems, RJL
Leavis, HL
Uh, HW

Submitted

ABSTRACT

Statistical analysis of microbiome data is complicated by the compositional, sparse, and high-dimensional nature of the data. In order to overcome these problems, transformations are often applied, but little is known about how they affect the correlation structure of the data and how this influences the outcome of statistical analyses. We applied commonly used microbiome data transformations (compositional, centered log-ratio transformation (CLR) with zero handling methods, and inverse hyperbolic sine) to an existing oropharyngeal 16S ribosomal RNA sequencing dataset, and analysed their effects on principal component analysis (PCA), variable selection, prediction, and network analysis.

Zero replacement (pseudocount, count multiplicative replacement, or Bayesian-multiplicative imputation) prior to CLR transformation introduced artificial correlations within the microbiome dataset. This resulted in larger explained variance in PCA. For variable selection and prediction performance, an ensemble approach integrating single point analysis, elastic net and random forest learners was applied. Although three bacterial taxa were stably detected across the machine learning approaches, the overall composition of bacterial taxa associated with the outcome varied with use of different data transformations. Case-control prediction accuracy and statistical power were improved after all data transformations compared to untransformed compositional data.

To conclude, CLR data transformation with zero handling introduced artificial correlation in a sparse microbiome dataset. As these data transformations resulted in differences in variable selection, we recommend caution in the use of transformation methods that require zero replacement.

Keywords

Centered log ratio (CLR), microbiome, data analysis, sparsity, compositional, next-generation sequencing

INTRODUCTION

Next-Generation Sequencing (NGS) is a powerful tool to determine the composition of complex microbial communities, providing a rich source of information about natural ecosystems that may profoundly influence our living environment and health. A typical research question about microbiome data concerns feature selection – which bacteria are different between two conditions (e.g. health and disease)?

Answering these questions, however, is often challenging due to the high-dimensionality, sparsity, and compositionality of microbiome data. High-dimensionality is caused by the vast diversity of bacterial taxa and increases as lower taxonomic levels are used. Furthermore, lower taxonomic ranking, which gives higher biological resolution, also increases the proportion of taxa that are absent in samples, yielding very sparse datasets. At lower taxonomic levels, sparsity can be as high as ~90% in human oral microbial communities¹. Finally, the compositional structure of microbiome data poses a statistical challenge. The total read count of NGS microbiome datasets (such as 16S rRNA sequencing data) cannot be related to the absolute number of bacteria in the input sample². Therefore, these datasets are usually converted to relative abundance values, resulting in compositional data that sum up to a fixed value. This may lead to inferior performance of traditional methodology if applied directly to these constrained data³. One of the consequences of data compositionality is the introduction of spurious correlations among variables, as already recognised by Pearson in 1897⁴. As it has been frequently posed that the false positive rate resulting from spurious correlations can become unacceptably high in microbiome studies^{3,5}, novel data analysis methods for microbiome data are increasingly accounting for the compositionality problem^{5,6}.

Many novel analysis methods rely on mathematical transformations to overcome constraints on compositional data. Aitchison proposed the centered log-ratio (CLR) transformation⁷, where the dataset is log-transformed and then row centered. A major drawback of CLR in highly sparse microbiome data is that logarithms cannot handle zeros. A commonly used solution is to add a small pseudo-count (e.g., +1) to the raw counts before making the data compositional⁸. An alternative approach replaces the original zeros back after taking log transformation (implemented in “compositions” R package⁹). The zeros in microbiome data are often considered a consequence of the detection limit¹⁰. In this regard, it is reasonable to substitute the zeros by a fixed fraction of the detection threshold (implemented in R package “zCompositions”). Apart from these simple replacement approaches, Martín-Fernández et al.¹¹ proposed a Bayesian-multiplicative imputation method to impute the zeros in the Bayesian framework (implemented in zCompositions). Though less common, researchers also use the raw count data without making them compositional. To reduce the skewness in microbiome count data, the inverse hyperbolic sine (asinh) transformation can be used¹².

Despite developments in the field to validate methods that account for the compositionality problem and modelling of zero counts, the practical consequences of zero replacement in combination with CLR transformations on feature selection are under-studied.

Here, we aim to investigate the impact of different zero-handling and transformation methods on (i) the correlation structure of microbiome data, (ii) dimension reduction analysis, (iii) the results of feature selection, (iv) prediction performance of statistical and machine learning, in particular, ensemble learning, and finally (v) network analysis to detect interaction of bacteria. These methods will be applied to a previously published clinical 16S rRNA dataset that compares the oropharyngeal microbiome of healthy people to that of patients with Common Variable Immunodeficiency (CVID), which leads to deficiency of IgG and IgA¹³. The sample size is relatively small ($n=41$ healthy controls: HC vs $n=37$ CVID with serum $\text{IgA} < 0.1 \text{g/L}$: CVID-IgA), and the microbiome data is sparse and heterogeneous. This dataset therefore accurately reflects real-world statistical challenges often encountered in clinical microbiome studies, more so than simulated data or datasets where very large differences can be expected to. In order to determine a ground truth of which bacteria were associated with the outcome, an ensemble machine learning approach was used. Bacterial taxa that were independently detected by different statistical methods were considered robustly associated with the outcome.

METHODS

All analyses were performed using R 3.6.3 (2020-02-29) and publicly available on Github: https://github.com/zhujiegu/Microbiome_ML.

Datasets

Previously published microbiome data derived from 16S rRNA sequencing of oropharyngeal swabs of patients with CVID was used for all subsequent analyses¹³. Sequencing data was made available on the European Nucleotide Archive under project code PRJEB34684. This dataset contains the following samples; HC ($n=41$), CVID with serum IgA levels $> 0.1 \text{g/L}$ (CVID+IgA, $n=48$), and CVID-IgA ($n=37$), and patients with X-linked agammaglobulinemia (XLA, $n=11$). Genus-level data was used ($P=139$ taxa). For supervised analyses, only the HC and the CVID-IgA groups were used. For the unsupervised network analysis, the whole cohort was used ($n=137$).

Data transformations

The methods that were applied to the dataset in order to investigate the influence of zero handling and data transformation are summarized in Table 1. Supplementary Figure 1 depicts the workflow of these methods.

Correlation and dimension reduction

Correlation structures (Spearman) of each transformed dataset were visualized using the package “ggcorrplot”. Dimension reduction was performed using Principal Component Analysis (package “FactoMineR”), plotting the first 2 PCs. Ellipses indicate confidence regions.

Table 1: Methods for replacement of zeros in Microbiome data. CLR: centered log ratio. Czm: count zero multiplicative. Gbm: geometric Bayesian multiplicative.

Method	Zero handling	CLR transformation	Notes
comp	No	No	Relative abundance (compositional data)
asinh	No	No	Inverse hyperbolic sine (asinh) transformation on count data ¹²
CLRO	No; zeroes are placed back after CLR	Yes	R package <i>compositions</i> ⁹
CLR1	Add a pseudo-count (count +1) to zero abundance	Yes	CLR transformation: compositional → log transformation → row centering
CLRczm	Multiplicative simple replacement	Yes	R package <i>zCompositions</i> ¹⁰
CLRgbm	Bayesian-multiplicative imputation	Yes	R package <i>zCompositions</i> ¹⁰

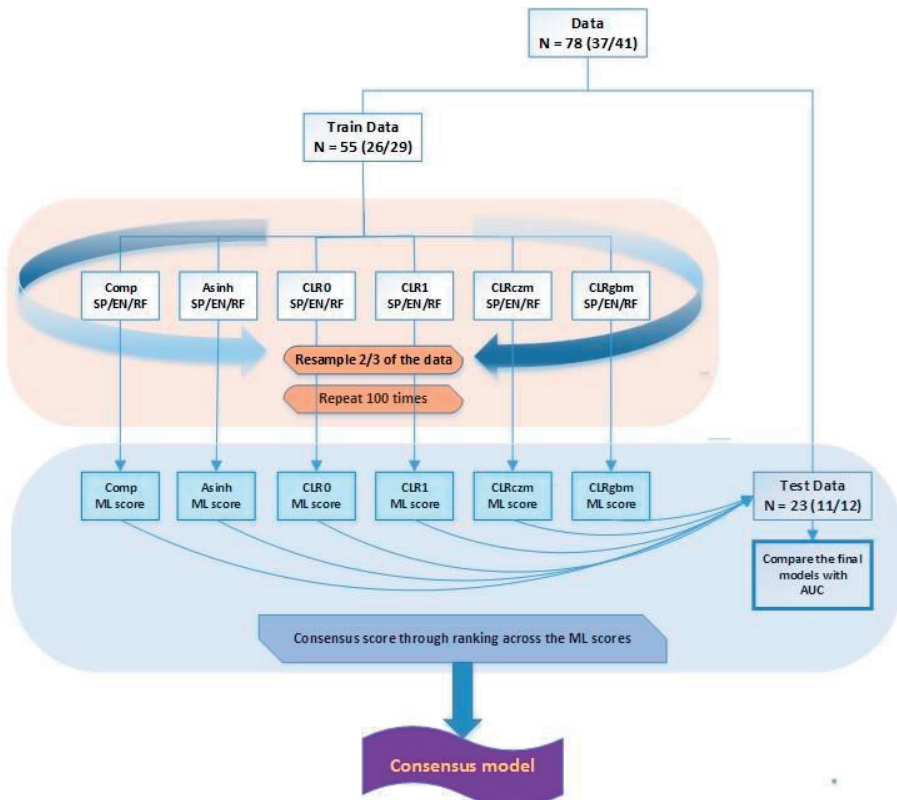


Figure 1: Workflow of ensemble learning. CLR: centered log ratio. N=sample size (n case : n control). Comp: compositional, asinh: inverse hyperbolic sine, CLRO: CLR with zero placed back, CLR1: CLR with pseudocount 1, CLRczm: CLR with count zero multiplicative replacement, CLRgbm: CLR with geometric Bayesian multiplicative zero replacement, ML: machine learning, SP: single point, EN: elastic net, RF: random forest, AUC: area under the curve.

Ensemble learning strategy

Ensemble methods using multiple learning algorithms were applied to obtain robust predictive performance. Figure 1 depicts the workflow. The dataset was randomly split into training and test sets (70% vs 30%), preserving the case-control ratio. Three statistical and machine learning methods (ML) were applied to the training dataset: single point (SP) analysis using univariate logistic regression, logistic Elastic Net (EN), and Random Forest (RF). Details about the implementation of these methods (e.g., hyperparameter tuning, R package, etc.) can be found in the supplementary material. All methods were fitted in 2/3 of the training set, and the bacteria were selected and ranked using appropriate criteria (p-values for SP, beta coefficient for EN, and mean decrease in Gini-index for RF). This procedure was repeated 100 times. For each method, individual ranking was averaged across 100 repetitions (ML score). The final score was calculated by averaging the three ML scores for each of six zero handling methods. The prediction performance of the final models with 5 top-ranking bacteria was evaluated using area under the curve (AUC) in the held-out test set. Finally, a consensus score was constructed as the average ranking of the taxon over all 6 methods. The three top-ranking bacteria were included in the consensus model together with covariate age.

Correlation networks

For estimating networks of bacteria, weighted gene co-expression network analysis (WGCNA) based on pairwise correlations between bacteria was used¹⁴. As network analysis is very sensitive to sparsity and sample size, bacteria with a zero count in more than 90% of samples were removed, resulting in 59 bacteria. To compute correlations of compositional microbiome data, three methods were considered: Compositions, CLR₁, and SparCC¹⁵. SparCC (Sparse Correlations for Compositional data) aims to estimate the true correlations between the log-transformed components. Since the output of SparCC varies strongly across repetitions, the correlation was estimated 10 times, and averaged correlation was used to construct the bacterial network.

RESULTS

CLR after pseudocount or zero imputation introduces an artificial correlation structure

We first investigated the impact of compositionality and data transformation methods on the correlation patterns within the data. The overall correlation structure of the count data (Figure 2a) and compositional data (Figure 2b) was very similar, indicating that compositionality did not introduce spurious correlations here. Next, the six selected zero replacement and compositionality handling methods (Table 1) were applied to the highly sparse dataset (70% of taxa have a zero proportion that is larger than 75%, (Supplementary Figure 2) in order to assess how these methods impacted the correlation structure.

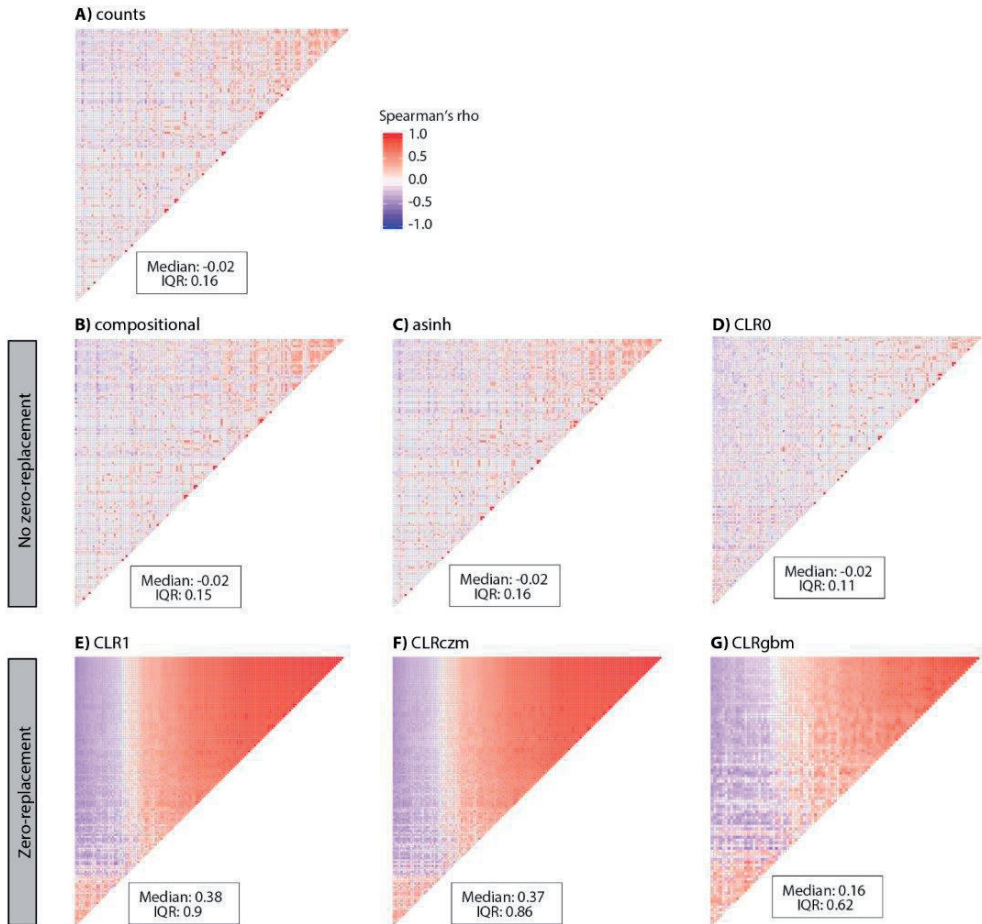


Figure 2: Heatmap showing Spearman's correlation between bacterial genera of (a) raw counts, (b) compositional, (c) inverse hyperbolic sine (asinh), (d) CLR with zero placed back (CLR0), (e) CLR with pseudocount 1 (CLR1), (f) CLR with count zero multiplicative replacement (CLRczm), and (g) CLR with geometric Bayesian multiplicative zero replacement (CLRgbm). Heatmap colour indicates Spearman's rho. IQR: interquartile range.

The overall correlation patterns of the methods without replacement of zero: compositional, asinh, and CLR0 (Figure 2 b-d), were similar, except a marked smaller statistical dispersion of CLR0 (IQR = 0.11 vs. 0.15 and 0.16 for compositional and asinh, respectively). In contrast, methods with zero replacement prior to CLR transformation CLR1, CLRczm, and CLRgbm (Figure 2 e-g) introduced artificial correlations. The fixed value replacement of zeros (CLR1 and CLRczm) gave much larger median (CLR1 0.38, CLRczm 0.37) and IQR (CLR1 0.9, CLRczm 0.86) values. The CLRgbm transformation using Bayesian imputation of the zeros instead of replacing them with a fixed number suffered slightly less from the spurious correlation (median 0.16, IQR 0.62) (Figure 2g).

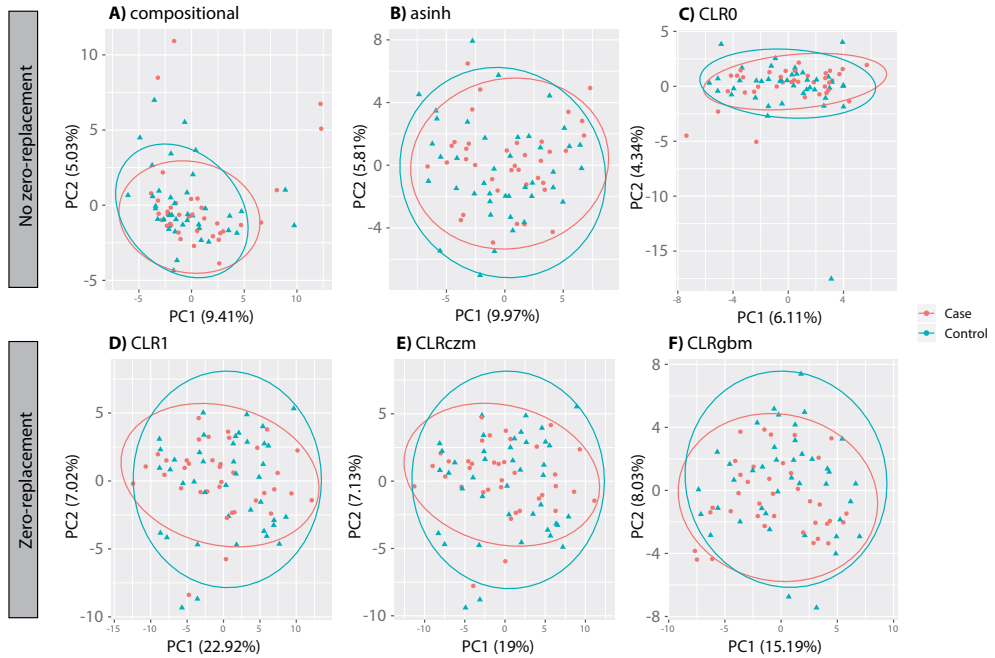


Figure 3: Principal component (PC) analysis presenting the first two principal components (PCs) for each of six data transformation methods: Comp: compositional, asinh: inverse hyperbolic sine, CLRO: CLR with zero placed back, CLR1: CLR with pseudocount 1, CLRczm: CLR with count zero multiplicative replacement, CLRgbm: CLR with geometric Bayesian multiplicative zero replacement. Ellipses indicate 95% confidence region.

The introduction of artificial correlations after zero replacement occurs because two zeros in a dataset can not be detected as being correlated. As soon as the zero is replaced by a small count, and especially if this small count is equal across all zeros, however, a strong artificial correlation will be introduced. Whether the correlation structure introduced by CLR transformation is spurious or is a better reflection of the true biological data structure cannot be determined here, but the correlations introduced by the zero replacement methods that are necessary for CLR are considered mostly artificial.

Zero replacement results in larger variance explained in PCA

In order to assess the influence of the transformation methods on commonly used dimension reduction methods, we performed PCA on each transformed dataset (Figure 3). The methods with zero replacements induced higher correlation resulting in larger variance explained (Figure 3 a-c). Notably, CLRO, in which the zeros are left untouched before CLR transformation but are replaced afterwards, thereby distorting the rank of variables, showed the least variance explained by two PCs (6.11 and 4.33%). Although visually group separation was not observed in any of the plots showing the first two PCs, there was a sig-

nificant difference between the CVID and HC groups for the zero-replacement methods and asinh (PERMANOVA: comp $p = 0.06$, asinh $p = 0.05$, CLRo $p=0.10$, CLRl $p=0.03$, CLRczm $p=0.04$, CLRgbm $p=0.03$). Overall, the spurious correlation introduced by zero replacement methods resulted in an artificially increased explained variance, lower PERMANOVA p -value, while the CLRo method produced nonrobust outliers in PCA.

Feature selection and prediction performance using ensemble learning

As feature selection and outcome prediction are common research questions in the microbiome field, we next assessed how each of the transformation methods performed using machine learning (ML) based on SP, EN, and RF. In order to determine which bacterial taxa were robustly associated with the outcome, the results of the three ML methods were integrated in an ensemble approach by averaging the rank of each taxon across the transformation methods (Table 2) (for individual ML ranks, see Supplementary Table 1). Overall, a non-specified genus belonging to the order *Lactobacillales* (mean rank 4.4), *Prevotella* (mean rank 5.5) and *Alloprevotella* (mean rank 6.4) were consistently found among the top-ranking bacteria. While *Prevotella* and *Alloprevotella* were stably detected by all zero handling methods, the more sparse taxon *Lactobacillales* (zero in 42.31% of samples compared to 0% and 3.85% for *Prevotella* and *Alloprevotella*, respectively) ranked lower in CLRo (rank of *Lactobacillales* in CLRo = 11.6). Apart from these three dominantly detected taxa, the transformation method used introduced more variation in taxa ranking, e.g. *Gemella* was ranked 5th most associated with the outcome when using compositional data, while it ranked between 13-16.7 in the transformation methods that use CLR.

When assessing the performance of the individual ML algorithms (Supplementary Table 1), RF performed less well in ranking *Lactobacillales*, *Prevotella* and *Alloprevotella* than SP and EN, especially when combined with the CLRl (RF rank 8.5, 10.5, 9.0 respectively) and CLRczm (RF rank 5.3, 8.8, 12.3, respectively) methods.

Univariate and multivariate logistic regression were used to investigate the association of *Lactobacillales*, *Prevotella* and *Alloprevotella* with CVID (see Table 3). While all three robust taxa were significantly associated with CVID in a univariate model when using compositional data, transformations gave more power to detect these taxa, especially in the multivariate model. In the multivariate model, *Alloprevotella* was often not significantly associated with the outcome, possibly due to collinearity with the other two taxa in the model.

In order to assess the influence of each transformation model on prediction performance, a multiple logistic regression model using 5 taxa with the highest ML score for each transformation method (see Supplementary Table 1) was applied to the held-out test set to predict which sample belonged in which group (healthy or CVID). The prediction performance in the final model was lowest for compositional data (AUC = 0.811) and was improved by all tested transformation methods (AUCs asinh, CLRo, CLRl, CLRczm and CLRgbm ranged between 0.89- 0.92).

Table 2: Ensemble machine learning (ML) score presented as the ranking of the bacteria for all 6 transformation methods. Smaller rank indicates stronger association with the outcome. Zero %: percentage of samples in which this taxon had zero counts. Abundance %: median relative abundance in non-zero samples. ML-score: mean of rank from single-point, elastic net and random forest. Mean rank: averaged ML score for each taxon. Comp: compositional, asinh: inverse hyperbolic sine , CLRO: CLR with zero placed back, CLRI: CLR with pseudocount 1, CLRczm: CLR with count zero multiplicative replacement, CLRgbm: CLR with geometric Bayesian multiplicative zero replacement

	Zero-CVID	Zero-HC	Abundance		ML score		Ensemble learning score					
			CVID	HC	comp	asinh	CLRO	CLRI	CLRczm	CLRgbm	score	
<i>Lactobacillales</i>	24.32%	58.54%	0.37%	0.30%	4.7	2	11.6	3.8	2.7	1.9	4.4	
<i>Prevotella</i>	0.00%	0.00%	3.66%	2.06%	3.3	5.2	7.8	6.6	5.5	4.4	5.5	
<i>Alloprevotella</i>	5.41%	2.44%	1.00%	0.28%	5.8	6.3	6	6.7	7.6	5.9	6.4	
<i>Gemella</i>	0.00%	0.00%	1.99%	1.82%	5	8.3	16.7	14.8	13.7	13	11.9	
<i>Neisseria</i>	2.70%	4.88%	2.18%	0.66%	12.6	11	12.9	14.8	14.3	10.9	12.7	
<i>Abiotrophia</i>	75.68%	58.54%	0.07%	0.12%	14	8.5	19.6	13.3	14.6	10.7	13.4	
<i>Streptococcus</i>	0.00%	0.00%	48.16%	63.00%	7.2	8.6	11.1	19.5	18.1	17.7	13.7	
<i>Streptobacillus</i>	72.97%	95.12%	0.16%	0.25%	17	14.1	29.2	10.8	10.6	9.7	15.3	
<i>Peptostreptococcus</i>	54.05%	56.10%	0.04%	0.06%	20	16.2	19.2	12.7	12.2	16.4	16.1	
<i>Rothia</i>	0.00%	0.00%	2.79%	2.95%	12	14.7	9.1	20.1	21.6	19.3	16.1	

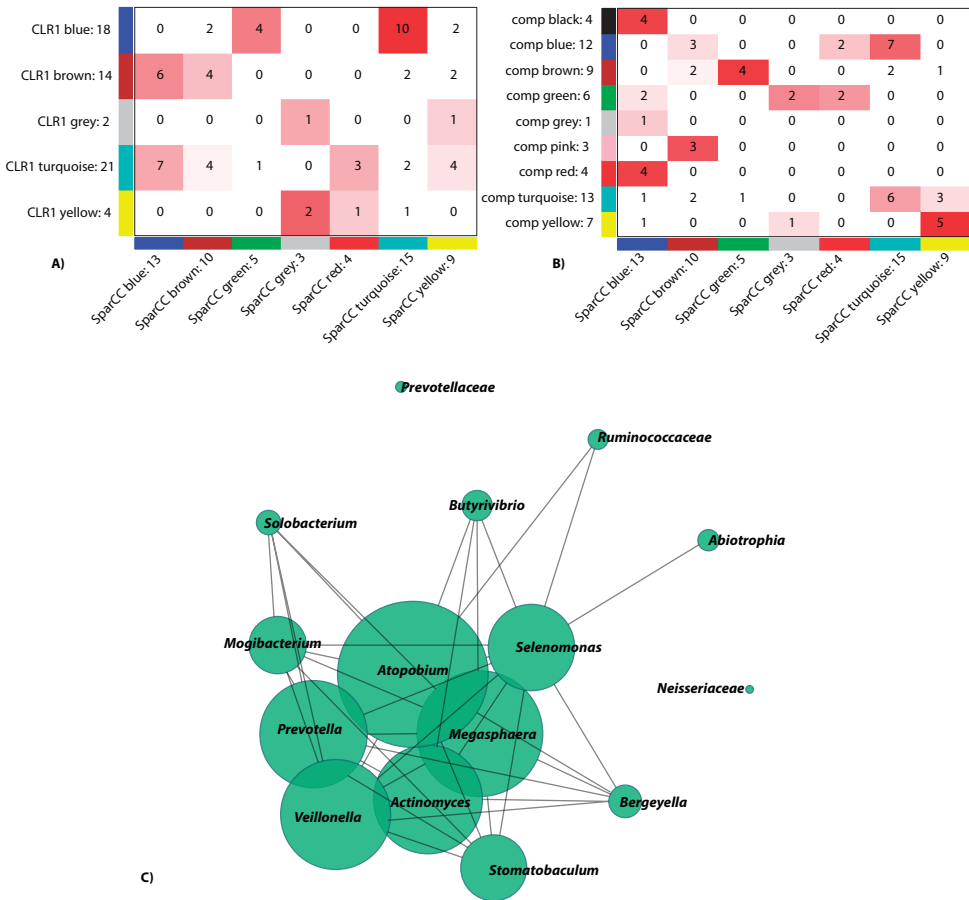


Figure 4: Correspondence of network modules between (A) comp and SparCC and (B) CLR1 and SparCC. Turquoise module of SparCC (C) showing the edges with the weights larger than 0.02. The larger the node, the higher impact of the bacterium in the module.

Investigating bacterial interactions by constructing correlation networks

The assumed introduction of spurious correlation in compositional data could be especially problematic when considering network analysis, which is entirely based on the correlation structure. In addition to the most commonly used compositional data and CLR1, SparCC was also tested, a method that aims to estimate the true correlation in compositional data. SparCC (median -0.01, IQR 0.23) resulted in less newly introduced correlation than CLR1 (median 0.01, IQR 0.45) (Supplementary Figure 3).

WGCNA network analysis yielded networks of 9, 5 and 7 modules for compositional, CLR1 and SparCC data, respectively. There was most overlap between CLR1-Blue (10/18 bacteria) and SparCC-Turquoise (10/15 bacteria) (Figure 4). These modules included *Prevotella*, *Megasphaera*, *Actinomyces*, *Atopobium* and *Veillonella* as the most interconnected, or hub, bacteria (Figure 4c, Supplementary Table 2). The majority of modules identified

Table 3: Summary of univariate and multivariate logistic regression of outcome (COVID/HC) on the three robustly detected bacterial genera, using six data transformation methods: Comp: compositional, asinh: inverse hyperbolic sine, CLR0: CLR with zero placed back, CLRI: CLR with pseudocount 1, CLRczm: CLR with count zero multiplicative replacement, CLRgbm: CLR with geometric Bayesian multiplicative zero replacement.

	Univariate				Multivariate			
	Bacteria	Effect size	Std. err	p-value	Effect size	Std. err	p-value	
Comp	<i>Lactobacillales</i>	-1.16	0.57	0.04	-1.62	0.66	0.01	
	<i>Prevotella</i>	-0.73	0.3	0.01	-1.03	0.46	0.03	
	<i>Alloprevotella</i>	-0.96	0.38	0.01	-0.13	0.53	0.81	
CLRasinh	<i>Lactobacillales</i>	-0.84	0.26	1.20E-03	-1.37	0.36	1.20E-04	
	<i>Prevotella</i>	-0.55	0.26	0.04	-1.24	0.41	2.20E-03	
	<i>Alloprevotella</i>	-0.52	0.27	0.05	0.2	0.32	0.55	
CLR0	<i>Lactobacillales</i>	-0.19	0.24	0.43	-0.31	0.29	0.29	
	<i>Prevotella</i>	-0.42	0.25	0.09	-0.46	0.32	0.15	
	<i>Alloprevotella</i>	-0.71	0.27	0.01	-0.56	0.28	0.04	
CLRI	<i>Lactobacillales</i>	-0.84	0.26	1.40E-03	-1.55	0.39	7.60E-05	
	<i>Prevotella</i>	-0.59	0.27	0.03	-1.47	0.44	8.70E-04	
	<i>Alloprevotella</i>	-0.59	0.27	0.03	0.08	0.32	0.81	
CLRczm	<i>Lactobacillales</i>	-0.84	0.26	1.30E-03	-1.54	0.38	6.50E-05	
	<i>Prevotella</i>	-0.61	0.27	0.02	-1.46	0.43	7.00E-04	
	<i>Alloprevotella</i>	-0.54	0.27	0.05	0.11	0.31	0.73	
CLRgbm	<i>Lactobacillales</i>	-0.89	0.27	1.10E-03	-1.71	0.43	6.80E-05	
	<i>Prevotella</i>	-0.58	0.27	0.03	-1.6	0.48	8.20E-04	
	<i>Alloprevotella</i>	-0.64	0.27	0.02	0.01	0.33	0.98	

had very limited overlap with modules from any of the methods used (Figure 4), illustrating that use of transformation methods greatly influences the outcome of WGCNA network analysis, and making it difficult to draw conclusions as to which method is the most robust. However, as SparCC does not require zero replacement, it may be preferred over methods such as CLR_I and other zero replacement methods that introduce artificial correlation within the dataset.

DISCUSSION

Here, we applied six different data transformation methods to a previously published microbiome dataset in order to investigate the impact of the compositionality problem – and the methods commonly applied to deal with it – on explorative data analysis, feature selection, prediction, and network analysis. The compositionality problem was described by Pearson, who proved that compositionality of data leads to a spurious correlation structure. In order to deal with this, the CLR transformation is commonly used, in combination with a variety of zero-handling methods.

We showed that CLR transformation methods with zero replacement resulted in a much more inter-correlated dataset than the original count or compositional data. The introduction of artificial correlation occurred when zeros were replaced by a fixed constant, be it a pseudocount +1 (CLR), an approximation of the detection limit (czm), and to a lesser extent by zeros that were imputed (gbm). CLR_O, in which zeros are handled after the CLR transformation, and compositional and inverse hyperbolic sine transformed data, in which zero replacement is not necessary, did not lead to the introduction of artificial correlation. CLR_O, however, is problematic on its own, because values smaller than one become negative after log transformation, resulting in a changed ranking of the data when zeros are placed back after CLR transformation. To what extent the compositional nature of microbiome data produces problematic spurious correlations, and whether these may be overcome using CLR transformation remain open research questions. However, the artificial correlation introduced by zero-replacement may pose an unacceptable side-effect to CLR transformation.

To see how this affects commonly used data analysis strategies in the microbiome field, we first performed principal component analysis. We observed that zero-handling methods artificially increased the variation explained by the first two principal components. CLR_O resulted in a much smaller cluster, with several outlier points that were not observed in the other methods. As dimension reduction methods are sometimes used to detect samples with contamination of sequencing errors, application of this method could be problematic when it produces nonrobust outliers.

Secondly, we investigated machine learning methods and assessed variable selection and prediction performance. Overall, compositional data resulted in selection of similar bacteria as the CLR-transformed data and while *Lactobacillales*, *Prevotella* and *Alloprevotella* were stably detected, the choice of transformation method influenced which other

bacteria were associated with the outcome. In prediction performance CLR or asinh transformation had a positive effect on the AUC.

These analyses illustrate that different statistical approaches can lead to different results, which is a recognized problem in the microbiome field that complicates data reproducibility and comparability. In our dataset, the ensemble machine learning approach proved useful to counteract sparsity and small sample size to detect robust associations between our outcome and bacterial taxa. Here, all transformation methods produced comparable results, with the exception of CLR₀. All zero replacement methods and asinh made existing effects more significant in the multivariate approach. Random forest was less robust for feature selection on our dataset compared to single point and elastic net, which may be due to smaller statistical power of the nonparametric method in data with small sample size. Larger sample size will improve the robustness of microbiome data analysis methods, but the artificial correlation structure introduced by zero replacement methods may remain problematic.

Finally, we observed that network analysis was yielded comparable results between methods that account for compositionality of the data such as CLR₁ and SparCC. These methods resulted in detection of more hubs, which convey useful information about bacterial interactions. However, a ground truth was difficult to ascertain in this analysis, and it remains elusive which transformation method performed best. As network analyses rely heavily on the correlation structure of the dataset, transformation methods such as SparCC that do not require zero replacement are recommended.

There are several limitations to this study. We elected to not use simulated data for this study, as data simulation is often based on statistical assumptions or distributions that do not necessarily reflect the structure of microbiome data. By using a real-life dataset, we cannot know which bacterial associations truly exist, and cannot compare the transformation methods to a gold standard. We tried to combat this by using an ensemble approach that is based on the assumption that associations that are detected by a variety of different methods are more likely to be robust. However, we cannot exclude the possibility that we have missed associations with high biological relevance, or have detected false positives in the process. In addition, the underlying structure of microbiome data may vary widely between studies, and especially between microbiota niches. Application of the methods tried here are likely to yield different results on different types of data.

Overall, we observed that the use of untransformed compositional data was less problematic in the introduction of a spurious correlation structure than anticipated based on the literature. As zero-handling methods introduced artificial correlations in the dataset, they should be used with care. For multivariate methods and network analysis, all tested data transformation methods (with the exception of CLR₀) improved statistical power. We recommend use of other methods that reduce the skewness of microbiome data without requiring zero replacement until more is known about the effects of zero replacement in combination with CLR transformation. Further studies are needed to quantify the extent to which the compositional nature of microbiome data results in spurious findings in real datasets, and whether alternative transformation methods may be used to overcome these

problems. Better understanding of how compositionality and sparsity affect microbiome data analysis will improve the interpretation of microbial community studies, and pave the way towards the implementation of standardized methods in the field.

Conflicts of interest

Authors report no conflicts of interest.

Funding

The research leading to these results has received funding and support from the European Union's Horizon 2020 research and innovation programme IMforFUTURE under H2020-MSCA-ITN grant agreement number 721815.

REFERENCES

1. Paulson JN, Colin Stine O, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods*. 2013;10(12):1200–2.
2. Gloor GB, Macklaim JM, Pawlowsky-glahn V, Egozcue JJ. Microbiome Datasets Are Compositional : And This Is Not Optional. *Front Microbiol*. 2017;8(November):1–6.
3. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, et al. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome*. 2017;5(1):1–18.
4. Pearson K. Mathematical contributions to the Theory of Evolution—on a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proc R Soc L*. 1897;489–98.
5. Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, et al. Best practices for analysing microbiomes. *Nat Rev Microbiol* . 2018;16(7):410–22. Available from: <http://dx.doi.org/10.1038/s41579-018-0029-9>
6. Lin H, Peddada S Das. Analysis of compositions of microbiomes with bias correction. *Nat Commun* . 2020;11(1):1–11. Available from: <http://dx.doi.org/10.1038/s41467-020-17041-7>
7. Aitchison J. *The Statistical Analysis of Compositional Data*. 2003rd ed. Caldwell, New Jersey: The Blackburn Press; 1986.
8. Costea PI, Zeller G, Sunagawa S, Bork P. A fair comparison. *Nat Methods*. 2014;11(4):359.
9. van den Boogaart KG, Tolosana-Delgado R. “compositions”: A unified R package to analyze compositional data. *Comput Geosci*. 2008;34(4):320–38.
10. Palarea-Albaladejo J, Martín-Fernández JA. ZCompositions - R package for multivariate imputation of left-censored data under a compositional approach. *Chemom Intell Lab Syst* . 2015;143:85–96. Available from: <http://dx.doi.org/10.1016/j.chemolab.2015.02.019>
11. Martín-Fernández JA, Hron K, Templ M, Filzmoser P, Palarea-Albaladejo J. Bayesian-multiplicative treatment of count zeros in compositional data sets. *Stat Modelling*. 2015;15(2):134–58.
12. Fukuyama J, Rumker L, Sankaran K, Jeganathan P, Dethlefsen L, Relman DA, et al. Multidomain analyses of a longitudinal human microbiome intestinal cleanout perturbation experiment. *PLoS Comput Biol*. 2017;13(8):e1005706.
13. Berbers R, Hoesein FAAM, Ellerbroek PM. Low IgA Associated With Oropharyngeal Microbiota Changes and Lung Disease in Primary Antibody Deficiency. 2020;11(June):1–11.
14. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9.
15. Friedman J, Alm EJ. Inferring Correlation Networks from Genomic Survey Data. *PLoS Comput Biol*. 2012;8(9):1–11.

SUPPLEMENTARY METHODS

Single point analysis (SP)

A univariate logistic model was fitted on each taxon i ($= 1, \dots, 139$) including the covariate *age*. The bacteria with the nominal p-value < 0.05 were selected. This procedure was repeated 100 times, and in each run significant bacteria were recorded. The final ranking of the bacteria was achieved by the frequency of the bacteria selected from 100 runs. When facing a tie, the smaller rank was assigned to both.

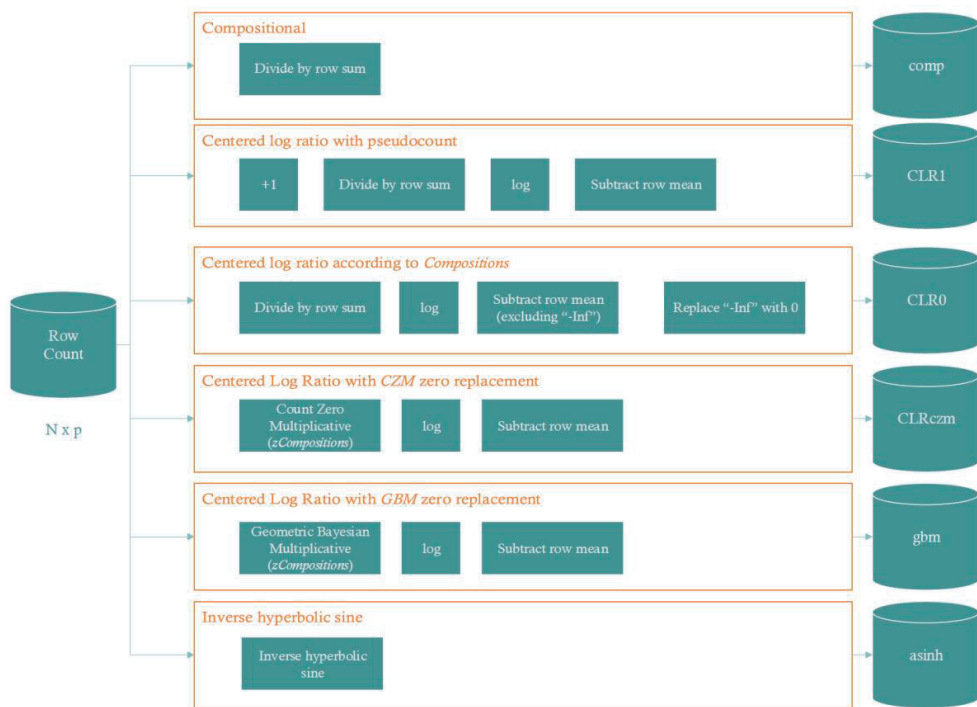
Elastic Net (EN)

The selection algorithm based on elastic net³ was implemented in order to deal with possible correlations between the parameters and to select relevant (or important) variables. This method is a combination of lasso (for selection of bacteria) and ridge (for shrinkage) logistic regression, and encourages grouping effect. EN was fitted on the training set (with covariate *Age*) using the R package “glmnet”. For fine-tuning of the hyperparameters alpha (controlling the ratio between L1 and L2 penalties) and lambda (controlling the sparsity level) we used 5-fold cross-validation.. In each run the taxa were ranked by the size of the coefficients. The final ranking of the taxon was achieved by averaging the rankings over 100 repetitions.

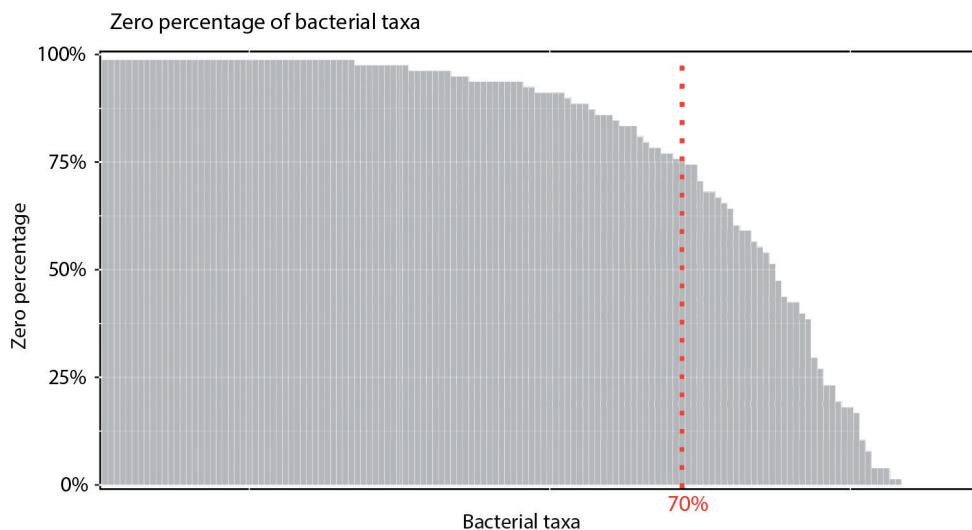
Random Forest (RF)

Since the RF algorithm¹⁴ does not require distributional assumption, it is possibly more suitable to analyze (non-normal) microbiota data¹⁵. Also, by applying RF, one can capture non-linear association between the Bacteria and CVID. The variable importance (Gini index) was used to rank the strength of the association. The final ranking of the taxon was achieved by averaging the rankings over 100 repetitions.

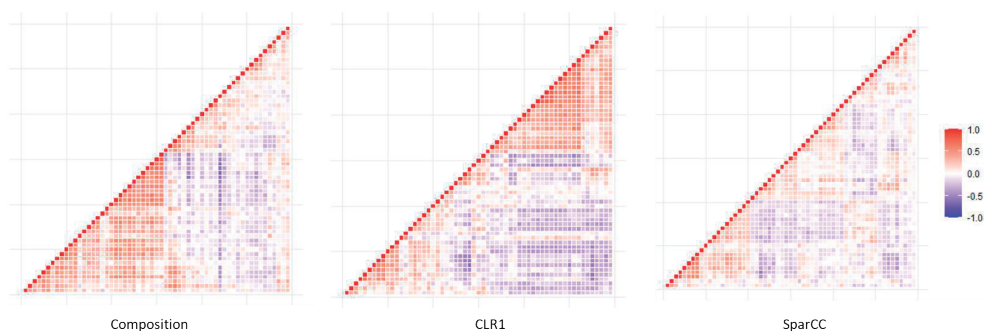
SUPPLEMENTARY FIGURES AND TABLES



Supplementary Figure 1: workflow overview of the transformation methods. Centering the rows: subtract the row mean. Rows are samples (N), columns are bacteria (p). CLR: centered log ratio. Comp: compositional, asinh: inverse hyperbolic sine, CLR0: CLR with zero placed back, CLR1: CLR with pseudocount 1, CLRczm: CLR with count zero multiplicative replacement, CLRgbm: CLR with geometric Bayesian multiplicative zero replacement.



Supplementary Figure 2: Sparsity of the oropharyngeal 16S rRNA dataset. Percentage of samples (y-axis) for which each bacterial genus (x-axis) has zero count.



Supplementary Figure 3: Heatmap showing Spearman's correlation of 59 bacteria used in network analysis. A) untransformed compositional data; Spearman's rho median = 0.06 ; IQR = 0.32, B) CLR1-transformed data; Spearman's rho median = 0.01 ; IQR = 0.45, C) SparCC-transformed data; Spearman's rho median = -0.01 ; IQR = 0.23. CLR1: Centered Log Ratio with pseudocount 1. IQR: interquartile range.

Supplementary Table 1: Results of ensemble learning: the highest ranking bacteria based on the ML score for each zero replacement method are presented, together with validation results using the held-out test dataset. ML: machine learning. SP: single point. EN: elastic net. RF: random forest. AUC: area under the curve. Comp: compositional, asinh: inverse hyperbolic sine, CLRO: CLR with zero placed back, CLR1: CLR with pseudocount 1, CLRczm: CLR with count zero multiplicative replacement, CLRgbm: CLR with geometric Bayesian multiplicative zero replacement.

Taxon		Variable selection performance				Prediction performance
		in Training dataset across 100 repetitions				of the Final model based on ML score in Test dataset
Methods		SP average rank	EN average rank	RF average rank	ML score	AUC
		comp	<i>Prevotella</i>	1	3.6	5.3
	<i>Lactobacillales</i>	5	6.3	2.8	4.7	
	<i>Gemella</i>	3	4.7	7.2	5	
	<i>Alloprevotella</i>	4	7.4	6	5.8	
	<i>Streptococcus</i>	2	11.8	7.9	7.2	
asinh	<i>Lactobacillales</i>	1	1.9	3.1	2	0.894
	<i>Prevotella</i>	2	7.4	6.1	5.2	
	<i>Alloprevotella</i>	4	8.3	6.7	6.3	
	<i>Gemella</i>	5	10	9.8	8.3	
	<i>Abiotrophia</i>	2	9.1	14.6	8.5	
CLRO	<i>Alloprevotella</i>	2	10	6.1	6	0.902
	<i>Atopobium</i>	1	8.6	13.5	7.7	
	<i>Prevotella</i>	5	12.5	5.8	7.8	
	<i>Lautropia</i>	4	4.5	15.1	7.9	
	<i>Granulicatella</i>	2	13.1	10.6	8.6	
CLR1	<i>Lactobacillales</i>	1	1.9	8.5	3.8	0.909
	<i>Prevotella</i>	4	5.2	10.5	6.6	
	<i>Alloprevotella</i>	3	7.4	9.9	6.7	
	<i>Streptobacillus</i>	8	5.1	19.3	10.8	
	<i>Peptostreptococcus</i>	10	9.3	18.7	12.7	
CLRczm	<i>Lactobacillales</i>	1	1.9	5.3	2.7	0.917
	<i>Prevotella</i>	2	5.8	8.8	5.5	
	<i>Alloprevotella</i>	2	8	12.9	7.6	
	<i>Streptobacillus</i>	4	4.8	23.1	10.6	
	<i>Peptostreptococcus</i>	10	9.6	16.9	12.2	
CLRgbm	<i>Lactobacillales</i>	1	1.7	3	1.9	0.909
	<i>Prevotella</i>	3	4.7	5.5	4.4	
	<i>Alloprevotella</i>	2	7	8.6	5.9	
	<i>Streptobacillus</i>	8	4.4	16.8	9.7	
	<i>Corynebacterium</i>	9	8.5	13.5	10.3	

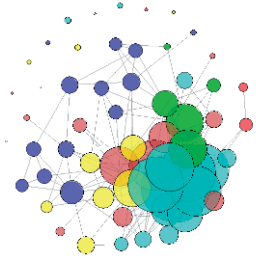
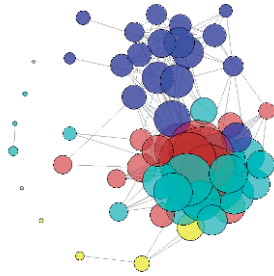
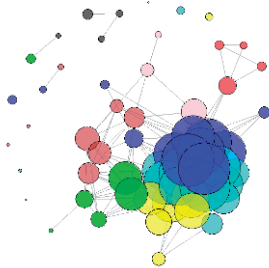
Supplementary table 2: Weighted correlation networks using compositional, CLRI, and SparCC data. Overlapping bacteria between the CLRI blue module and the SparCC Turquoise module are indicated in bold face.

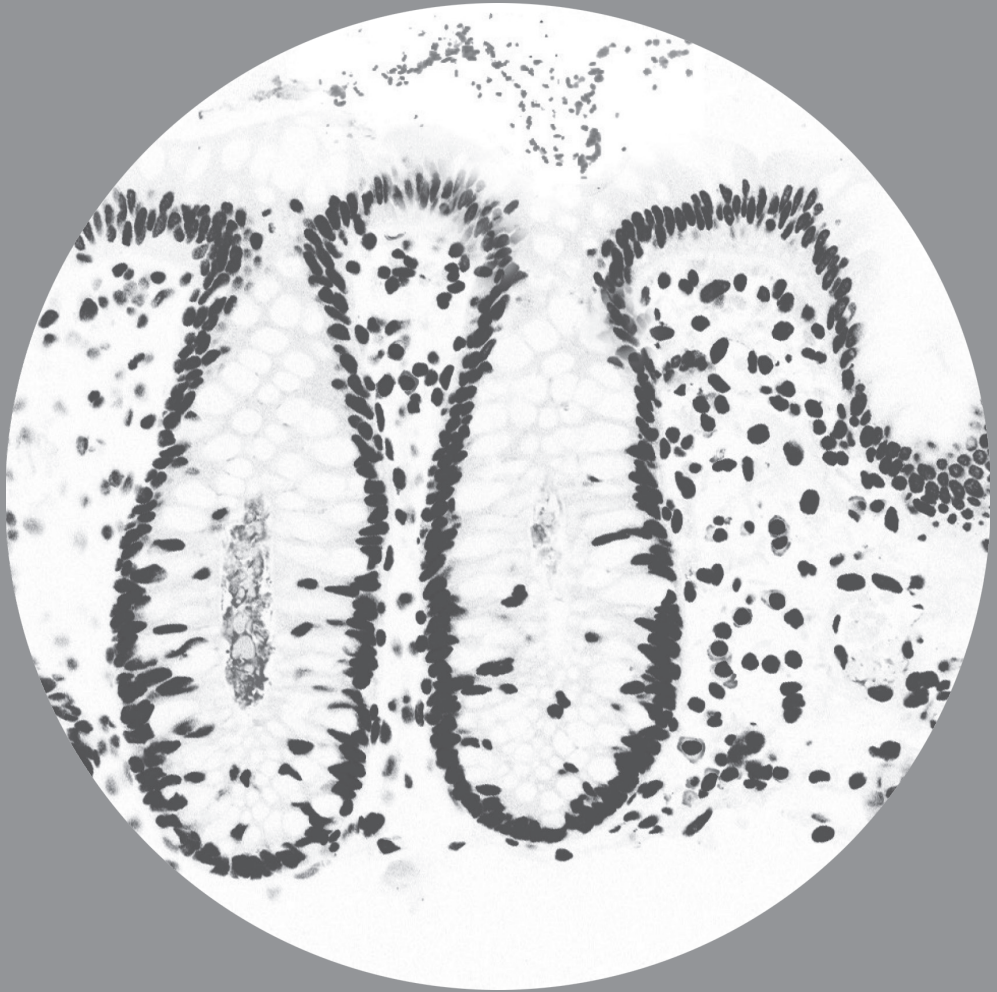
Compositional		CLRI		SparCC	
Module (#nodes)	Bacteria	Module (#nodes)	Bacteria	Module (#nodes)	Bacteria
Turquoise (13)	<i>Alloprevotella</i> , <i>Prevotellaceae</i> , <i>Mogibacterium</i> , <i>Eubacterium_nodatum_group</i> , <i>Butyrivibrio_2</i> , <i>Lachnoanaerobaculum</i> , <i>Stomatobaculum</i> , <i>Lachnospiraceae</i> , <i>Ruminococcaceae_UCG-014</i> , <i>Solobacterium_uncultured.1</i> , <i>Leptotrichia</i> , <i>Kingella</i>	Turquoise (21)	<i>F0332</i> , <i>Bifidobacterium</i> , <i>Bifidobacteriaceae</i> , <i>Rothia</i> , <i>Bergeyella</i> , <i>Chloroplast</i> , <i>Granulicatella</i> , <i>Lactococcus</i> , <i>Catonella</i> , <i>Oribacterium</i> , <i>Lachnospiraceae</i> , <i>Dialister_uncultured.1</i> , <i>Fusobacterium</i> , <i>Streptobacillus</i> , <i>Elkenella</i> , <i>Kingella</i> , <i>Neisseriaceae</i> , <i>Aggregatibacter</i> , <i>Treponema_2</i> , <i>Corynebacterium_all</i>	Turquoise (15)	<i>Actinomyces</i> , <i>Atopobium</i> , <i>Prevotellaceae</i> , <i>Bergeyella</i> , <i>Abiotrophia</i> , <i>Mogibacterium</i> , <i>Butyrivibrio_2</i> , <i>Stomatobaculum</i> , <i>Ruminococcaceae_UCG-014</i> , <i>Solobacterium</i> , <i>Megasphaera</i> , <i>Veillonella</i> , <i>Neisseriaceae</i> , <i>Prevotella_all</i> , <i>Selenomonas_all</i>
Blue (12)	<i>Actinomyces</i> , <i>Rothia</i> , <i>Atopobium</i> , <i>Abiotrophia</i> , <i>Granulicatella</i> , <i>Streptococcus</i> , <i>Megasphaera</i> , <i>Veillonella</i> , <i>Actinobacillus</i> , <i>Aggregatibacter</i> , <i>Prevotella_all</i> , <i>Selenomonas_all</i>	Blue (18)	<i>Actinomyces</i> , <i>Atopobium</i> , <i>Porphyromonas</i> , <i>Capnocytophaga</i> , <i>Gemella</i> , <i>Lactococcus</i> , <i>Mogibacterium</i> , <i>Butyrivibrio_2</i> , <i>Lachnoanaerobaculum</i> , <i>Stomatobaculum</i> , <i>Ruminococcaceae_UCG-014</i> , <i>Megasphaera</i> , <i>Veillonella</i> , <i>Leptotrichia</i> , <i>Neisseria</i> , <i>Haemophilus</i> , <i>Prevotella_all</i> , <i>Selenomonas_all</i>	Blue (13)	<i>Bifidobacterium</i> , <i>Bifidobacteriaceae</i> , <i>Alloprevotella</i> , <i>Chloroplast</i> , <i>Lactobacillus</i> , <i>Peptostreptococcus</i> , <i>Dialister</i> , <i>Fusobacterium</i> , <i>Phyllobacterium</i> , <i>Bradyrhizobium</i> , <i>Sphingomonas</i> , <i>Achromobacter</i> , <i>Corynebacterium_all</i>

Supplementary table 2: Continued

Compositional		CLRI		SparCC	
Module (#nodes)	Bacteria	Module (#nodes)	Bacteria	Module (#nodes)	Bacteria
Brown (9)	<i>F0332, Porphyromonas, Capnocytophaga, Bergeyella, Lautropia, Eikenella, Neisseria, Neisseriaceae, Haemophilus</i>	Brown (14)	<i>Alloprevotella, Campylobacter, Abiotrophia, Streptococcus, Eubacterium_nodatum_group, Peptostreptococcus, Solobacterium, uncultured_bacterium.5, Phyllobacterium, Bradyrhizobium, Sphingomonas, Achromobacter, Lautropia, Actinobacillus</i>	Brown (10)	<i>F0332, Gemella, Streptococcus, Lactobacillales, Eubacterium_nodatum_group, Streptobacillus, Lautropia, Kingella, Actinobacillus, Aggregatibacter</i>
Yellow (7)	<i>Campylobacter, Catonella, Oribacterium, Candidatus_Saccharimonas, uncultured_bacterium.5, Saccharimonadates, Corynebacterium_all</i>	Yellow (4)	<i>Prevotellaceae, Tannerella, Parvimonas, Peptococcus</i>	Yellow (9)	<i>Campylobacter, Catonella, Lachnoaerobaculum, Oribacterium, Lachnospiraceae, Leptotrichia, Candidatus_Saccharimonas, uncultured_bacterium.5, Eikenella</i>
Green (6)	<i>Tannerella, Parvimonas, Peptococcus, Peptostreptococcus, Fusobacterium, Treponema_2</i>			Green (5)	<i>Porphyromonas, Capnocytophaga, uncultured.1, Neisseria, Haemophilus</i>
Red (4)	<i>Phyllobacterium, Bradyrhizobium, Sphingomonas, Achromobacter</i>			Red (4)	<i>Rothia, Tannerella, Granulicatella, Treponema_2</i>

Compositional Module (#nodes)	Bacteria	CLRI Module (#nodes)	Bacteria	SparCC Module (#nodes)	Bacteria
Black(4)	<i>Bifidobacterium</i> , <i>Bifidobacteriaceae</i> , <i>Lactobacillus</i> , <i>Dialister</i>		Bacteria		Bacteria
Pink (4)	<i>Gemella</i> , <i>Lactobacil-</i> <i>lales</i> , <i>Streptobacillus</i>		Bacteria		Bacteria





Chapter 9

General discussion

Common Variable Immunodeficiency (CVID) is a primary antibody deficiency that is characterized by disturbed production of immunoglobulins. This defect in the adaptive humoral immune response leads to increased susceptibility to infections, mostly bacterial infections of the respiratory and gastrointestinal tract. Immunoglobulin G (IgG) replacement therapy (IgRT) from healthy donors has greatly improved the infection load and survival of CVID and other primary antibody deficiency patients^{1,2}. A subgroup of CVID patients, however, suffers from additional immune dysregulation complications that are not ameliorated with IgRT³. These complications affect between 60%³ and 85%¹ of CVID patients, and represent a wide range of autoimmune entities that can occur in any organ system. Immune dysregulation phenotypes seen in CVID include autoimmune cytopenias, granulomatous-lymphocytic interstitial lung disease (GLILD) and enteritis³⁻⁵. Despite this heterogeneity in presentation, patients with autoimmunity in one organ system are also at risk to develop other immune dysregulation complications⁶, and therefore we work under the hypothesis that they are caused by a similar underlying mechanism.

The occurrence of immune dysregulation is not the only ‘variable’ aspect of CVID. The age of onset of CVID can anywhere over two years of age, with peaks between 10-19 years of age and 30-39 years of age^{1,7}. This suggests that in addition to a genetic component, an environmental stimulus contributes to, and may occur prior to the development of clinical antibody deficiency. The same can be said for immune dysregulation in CVID, which can occur at or even before presentation of antibody deficiency, or only develop decades later¹.

Over the past two decades, advances in next-generation sequencing have provided an increasing body of evidence that suggests a pertinent role for the gut microbiota in the pathophysiology of human autoimmune disease⁸. In this thesis, we investigated the hypothesis that the microbiota comprise (one of) the environmental stimuli that contribute to immune dysregulation in CVID. We approached this question by evaluating the immunological profile (**part 1** of this thesis) and microbiological profile (**part 2** of this thesis) in a multicenter cross-sectional cohort in which we obtained serum, plasma, peripheral blood mononuclear cells (PBMCs), oropharyngeal swabs and fecal samples from each patient. We compare CVID with infections only (CVIDio) with CVID with immune dysregulation (CVIDid). In **part 3** of this thesis we address a problem we encountered during the statistical analysis of our microbiome datasets.

KEY FINDINGS OF THIS THESIS

Part 1: the immunology of CVIDid

- The serum cytokine profile of CVIDid patients not only shows strong pro-inflammatory signals suggestive of Th1 skewing, but also upregulation of immune regulatory proteins of the innate and the adaptive immune system (**chapter 2**). We showed that these regulatory proteins are potential biomarkers for immune dysregulation in

CVID. However, we could not answer the question whether expression of immune regulation markers also dampened potentially pathogenic T-cell responses.

- The T-cells that express immune regulatory markers classically associated with immune exhaustion still retain their inflammatory potential (**chapter 3**). This answered some of the questions raised in **chapter 2**, and show that, despite their chronically activated state, these T-cells may still be involved in the pathophysiology of CVIDid. In addition, we found clues that regulatory T-cells may be dysfunctional in CVIDid.

Part 2: the microbiota in CVIDid

- The composition of the oropharyngeal microbiota in CVID is associated with IgA deficiency and extent of lung damage (**chapter 6**). CVID patients with very low serum IgA had increased bacterial load, increased alpha diversity and enrichment with potentially inflammatory bacteria belonging to the Prevotellaceae family. These bacteria also correlated with more severe structural lung damage on CT-scans.
- In the colon of CVID patients, the gut microbiota can invade crypts that are clear of bacteria in healthy controls, thereby increasing the contact between bacteria and the gut epithelium and immune system (**chapter 7**).
- The gut microbiota of CVIDid patients has increased bacterial load, decreased alpha diversity, and enrichment with *Enterococcus* species (**chapter 7**). Especially *Enterococcus gallinarum* was more present in CVIDid, and was associated with increased levels of pro-inflammatory cytokines in serum of patients carrying *E. gallinarum* in feces, and in *in vitro* co-cultures of the bacterium with monocytes. This indicates that pathobionts such as *E. gallinarum* may provide an inflammatory stimulus in some CVIDid patients.

Part 3: statistical analysis of microbiome data

- Microbiome datasets are often highly dimensional and sparse, and the abundance of bacteria can only be interpreted as a percentage within a sample. We discovered that one of the most frequently used solutions to these problems, the centered log ratio transformation, imposes an artificial correlation structure on the data because it cannot handle sparsity (**chapter 8**). This can give false positive results in microbiota data analysis.

THE IMMUNOLOGY BEHIND IMMUNE DYSREGULATION IN CVID

Innate immunity in CVIDid

The innate or inborn immune system is the first responder of the human immune system and consists of cells and molecules (such as neutrophils, NK cells, and the complement system) that specialize in directly damaging pathogens, presenting antigen to and activating the adaptive immune system (such as dendritic cells and monocytes), and cells

that do both (such as macrophages)¹. In CVID, most is known about antigen presenting cells (APCs). CVID patients have a different distribution of monocyte subsets, notably an expansion of pro-inflammatory CD14^{bright}CD16⁺ monocytes as compared to healthy controls and XLA patients^{9,10}. Expansion of this cell subset is also associated with other autoimmune diseases such as in rheumatoid arthritis, where expansion of CD14^{bright}CD16⁺ monocytes is associated with disease severity¹¹. Overall, however, CVID monocytes and dendritic cells often show low activation markers and cytokine production¹²⁻¹⁴. These findings were reported in CVID as a whole, so whether monocyte dysfunction also or specifically occurs in CVIDid is unknown.

Some of the serum cytokines and chemokines we found to be upregulated in CVIDid in **chapter 2** are produced in large quantities by monocytes. These include CXCL9, 10 and 11, which are pro-inflammatory and promote T helper (Th)-1 skewing. The immune regulatory cytokine IL-10 was most strongly and consistently associated with the immune dysregulation phenotype. IL-10 is often produced in concert with pro-inflammatory cytokines such as TNF- α and IL-6, and is thought to provide a negative feedback loop limiting inflammation¹⁵ in the context of chronic antigen exposure¹⁶. The exact role of IL-10 in autoimmune disease remains complex, as IL-10 was also shown to induce pro-inflammatory CD14^{bright}CD16⁺ monocyte expansion in rheumatoid arthritis¹¹. CD83, which regulates activation of APCs, was also increased in serum of CVIDid patients, and is thought to be immunosuppressive in soluble form¹⁷. In **chapter 7**, we stimulated monocytes from healthy controls, CVIDio and CVIDid patients with *E. coli*, *E. gallinarum*, and *E. hirae*, and observed that monocytes from CVIDid produced lower IL-6 and IL-10 than HC monocytes regardless of the bacterial stimulus. These experiments need to be repeated with larger sample size, but they may indicate that CVIDid monocytes have a higher threshold for activation than in HC. Monocyte hyporesponsiveness can be caused by systemic exposure to bacterial products such as lipopolysaccharide (LPS)¹⁸⁻²⁰, which has been reported in CVIDid^{21,22}.

Adaptive immunity in CVIDid

The adaptive immune system provides tailored antigen-specific responses. In CVID, antigen production by B-cells is impaired, resulting in low serum antibody concentrations, poorly protective vaccine responses, and low frequencies of isotype-switched B-cells^{23,24}. The cause of B-cell deficiency in CVID is unknown, and likely differs between patients²⁴. Despite these B-cell problems, however, autoantibodies have been detected in serum of CVIDid patients²⁵, and can contribute to immune dysregulation complications such as autoimmune cytopenias²⁶. Expansion of CD21^{low} B-cells, of which a large proportion are autoreactive²⁷, is strongly associated with CVIDid^{28,29}, and also been described in other immune dysregulation syndromes such as systemic lupus erythematosus³⁰, rheumatoid arthritis²⁷, and human immunodeficiency virus infection³¹. Possibly, the contribution of

1 In biology, dichotomies almost never reflect reality. In fact most cells of the immune system can present antigen (even B-cells), and many cells that were thought to be part of the adaptive immune system have innate-like properties, while innate cells turn out to have memory and can be trained.

B-cells to the development of autoimmune disease in CVID also explains why patients with XLA, who have no residual B-cell function, are at lower risk to develop autoimmunity than CVID patients^{32,33}.

In biopsies of immune dysregulation-affected organs in CVID, a predominance of T-cells has been observed^{34,35}, suggesting a role for T-cells in the pathophysiology of CVID. Serum cytokines in CVID suggest Th-1 skewing (**chapter 2**), which we confirmed looking directly at T-cell profiles in **chapter 3**. Both Th (CD4+) and cytotoxic (CD8+) T-cells had high expression of activation markers and cytokine production. While we found increased levels of IL-17A in serum of CVID patients, there was no co-occurring enrichment of IL-17 expressing Th (Th-17) cells. It is possible that Th-17 cells are present in tissues of CVID patients, but this has not been observed in biopsies³⁶. An alternative source for IL-17A production in CVID are type-3 innate lymphoid cells (ILCs), which are enriched in peripheral blood of CVID patients³⁷.

In serum, immune regulatory markers LAG3, PD-L1 and 4-1BB were increased in CVID (**chapter 2**). These markers have often been described in the context of immune exhaustion, which can occur after chronic antigen stimulation and results in nonresponsive T-cells³⁸. Exhaustion markers have been reported previously in CVID³⁹⁻⁴¹, and Perreau et al.⁴⁰ showed that PD-1 expression resulted in decreased ability of CD4+ T-cells to respond to bacterial stimuli. We confirmed increased expression of LAG3, PD-1 and other negative regulators of costimulation CTLA4, ICOS and TIGIT on CVID T-cells in our cohort in **chapter 3**. However, the cells expression LAG3 and PD-1 retained their ability to proliferate and to produce pro-inflammatory cytokines, suggesting that they were not functionally exhausted, in contrast to the findings of Perreau et al.⁴⁰. Given the large heterogeneity of CVID, it may be that the patients sampled by Perreau et al. simply had a different phenotype than ours. However, there are several differences between the two studies that might explain why they do observe signs of functional exhaustion, while we do not. The assay used by Perreau et al. relied on presentation of bacterial antigens and subsequent activation of T-cells by APCs. This is a complex process, and other factors than T-cell exhaustion such as monocyte production of PD-L1 may also influence the assay. Our study setup was much simpler, stimulating T-cells with phorbol myristate acetate (PMA) and ionomycin, which gives a T-cell receptor (TCR)- independent stimulus and gives a more isolated picture of the *ex vivo* T-cell activation state. In **chapter 7**, we also stimulated CVID CD4+ T-cells with CD28-CD3 (TCR) beads, and found no differences between CVID or healthy control T-cells.

Whether T-cells are exhausted or not is relevant from two perspectives; first, if antigen exposure results in functionally exhausted T-cells, this may exacerbate the already existing immune deficient state, leading to a vicious cycle of increased antigen stimulation. Second, if T-cells are not functionally exhausted but only chronically activated, they may contribute to the pathology of the immune dysregulation phenotype. Given the overrepresentation of T-cells in immune dysregulation biopsies, the efficacy of T-cell targeted therapy in CVID, and our findings in **chapter 3**, there is sufficient evidence pointing in this direction. However, it is possible that in addition to this mechanism, defects in the

interaction between APCs and T-cells still result in a hyporesponsive T-cell response contributing to exacerbated antigen exposure.

THE MICROBIOTA IN IMMUNE DYSREGULATION IN CVID

In **part 2** of this thesis we describe both the oropharyngeal and the fecal microbiota of CVID patients. While other groups have characterized the gut microbiota in CVID, we were the first, and at time of writing this discussion still the only group to do describe the oropharyngeal microbiota in primary antibody deficiencies.

Oropharyngeal microbiota

Lung disease has a profound impact on the quality of life, morbidity and mortality in CVID^{42,43}, and can be caused broadly by two processes; structural airway disease such as bronchiectasis, and inflammatory interstitial lung disease such as GLILD. Structural airway disease can be the result of pulmonary infections, but also progresses in the absence of clinical infection¹. We therefore hypothesized that a dysregulated respiratory microbiota may cause subclinical inflammation in the lung that contributes to structural airway disease and perhaps also interstitial lung disease in CVID. During analysis of the oropharyngeal microbiome data in **chapter 6**, we noticed that grouping patients by their IgA status showed larger differences in bacterial load, alpha and beta diversity and differentially expressed bacteria than grouping them according to their immune dysregulation status. This can probably be explained by the stronger association of the oropharyngeal microbiota with structural airway disease, which we do not consider to be an immune dysregulation phenotype, than that with GLILD. In addition, it indicates that IgA may be an important regulator of the oropharyngeal microbiota. It is possible that GLILD in CVID is more related to systemic immune dysregulation than local processes in the lung, and, therefore, may be more strongly linked to dysregulated gut microbiota. Nevertheless, therapeutic targeting of the respiratory microbiota may ameliorate the development of structural airway disease, which would be very beneficial for patients. Therapeutic targeting of the microbiota will be discussed further on in this discussion.

Fecal microbiota

The fecal microbiota has been better studied in CVID, but while it has been hypothesized that the microbiota may contribute to immune dysregulation in CVID^{21,44}, previous studies were often underpowered to compare CVIDid to CVIDio. In **chapter 4**, we investigated the current literature supporting a role for the microbiota in the pathophysiology of CVIDid. The first study describing the fecal microbiota in CVID was written by Jorgensen et al.²¹, who found a reduced alpha diversity and differentially abundant bacteria in CVID, in a cohort consisting of 44 CVID patients, of whom 35 had CVIDid and 9 CVIDio. Therefore, it is unclear whether the presented differences were related to the shared immunodeficiency or specifically associated with immune dysregulation. The bacteria associated

with CVID included known pathobionts from the Enterobacteriaceae family that were increased in CVID, while taxa that are considered to be probiotic such as *Bifidobacterium* were decreased in CVID. Low alpha diversity was most pronounced in patients with complete IgA deficiency in serum, and with immune dysregulation. Interestingly, LPS levels in plasma of an expanded cohort were found to be increased in the CVIDid group (n=79) compared to CVIDio (n=25). Increased serum LPS is a marker for translocation of bacterial products, which is mostly assumed to come from the gut in the absence of clinical infections elsewhere.

Reduced alpha diversity in CVID associated with complications and IgA deficiency was also reported by Fiedorova et al.⁴⁵ (CVIDio n=15, CVIDid n=12), and was especially pronounced in CVID patients with diarrhea⁴⁶. Enrichment of Enterobacteriaceae was also consistently found to be associated with CVID⁴⁵. Finally, Shulzhenko et al.⁴⁷ investigated the microbiota in duodenal biopsies in CVID patients with enteropathy. They found that these patients were often profoundly IgA deficient, and identified *Acinetobacter baumannii* as a possible pathobiont in CVID with enteropathy. It is difficult to compare these findings to other studies, as the mucosa-associated duodenal microbiota is expected to be distinct from the fecal, and because the DNA extraction and sequencing methods used were very different in this study.

We set out to complement the existing body of literature with a cohort that allowed comparison of CVIDio with CVIDid. In **chapter 7**, we confirm the previously observed changes between CVID and HC, but also show that the CVIDid subgroup was associated with increased bacterial load, decreased alpha diversity, and enrichment with potential pathobionts belonging to the genus *Enterococcus*. These findings were also associated with IgA deficiency in serum, which was to be expected given the enrichment of IgA deficiency in the CVIDid group. We studied differences in the distribution of enterococcal species using shotgun metagenomics sequencing complemented with targeted PCR and bacterial culturing, and showed that especially *E. gallinarum* was enriched in CVIDid. Several enterococci have previously been associated with inflammation, including *E. gallinarum*, *E. hirae*, *E. faecium* and *E. faecalis*, all of which were to some extent more prevalent in CVIDid. *E. gallinarum* was most strongly associated with the CVIDid phenotype, pro-inflammatory cytokines in serum, and monocyte activation *in vitro*.

All this is not to claim that all of CVIDid is caused only by *E. gallinarum*. We were able to detect this pathobiont in 22.7% of CVIDid patients, and it may contribute to maintaining inflammation in CVIDid, but other pathobionts are likely to play a role in other CVIDid patients. *E. hirae*, which was also enriched in CVIDid, also showed monocyte activation in our assay, and *E. hirae* is thought to provide an inflammatory stimulus in cancer patients that can boost the immune system to overcome the local anti-inflammatory tumor environment⁴⁸. Taken together, 36% of CVIDid patients carried *E. gallinarum* and/or *E. hirae* in their gut. The enrichment of Enterobacteriaceae that was consistently found across studies in CVID may include strains of *E. coli* that are highly inflammatory⁴⁹. The same may be true for *E. faecium*, which shows great genetic diversity between strains⁵⁰. All in all, carriage of pathobionts is usually not sufficient to cause disease, but probably

requires other factors such as genetic background of the host, and a trigger that results in exposure of the pathobiont to the immune system.

While the endotoxemia observed in CVIDid indicates that there is bacterial (product) translocation, the mechanism as to how this happens is unknown. One hint we observed was bacterial invasion of colonic crypts in **chapter 7**. The inner mucus layer, which also fills the colonic crypts, is considered to be sterile in healthy individuals⁵¹, but this is not the case in diseases such as inflammatory bowel disease⁵² and colorectal cancer⁵³. IgA deficiency may contribute to maintaining spatial segregation between the epithelium and the microbiota⁵⁴, but IgA deficiency alone does not cause invasion of the inner mucus layer by bacteria⁵⁵. The structure of the mucus itself is most important for this⁵⁵, and mice deficient in *Muc2* develop bacterial crypt invasion and crypt elongation⁵¹, which we also observed in CVID colon biopsies. Further studies are needed to directly investigate whether the mucus layer is impaired in CVID, or whether other factors might result in bacterial crypt invasion.

IgA deficiency

Over the past years, the field of mucosal immunology has taken flight⁵⁶, and the primary immunodeficiencies have not been overlooked. Primary immunodeficiencies have historically been very important in our understanding of basic human immunology⁵⁷, and they may make their next contribution by allowing the study of the effects of defects in specific components of the immune system on the microbiota – something that can otherwise only be done in experimental (animal) models. The primary antibody deficiencies are the ideal model disease to study the impact of antibody deficiency on the microbiota, and in particular that of IgA. Most of our current knowledge about IgA derives from mouse models, but there are key differences between mouse and human immunology (and microbiology!) that limit the extent to which these findings also apply to human disease.

In **chapter 5**, we summarized the current literature about IgA coating by gut bacteria in order to create an overview of which bacterial taxa may be affected by IgA deficiency. Previous studies from Fadlallah et al.^{58,59} have investigated the gut microbiome in patients with selective IgA deficiency (sIgAD). While there was no impact on overall alpha diversity, they observed a mild dysbiosis that consisted of an expansion of pathobionts and a reduction of beneficial symbionts⁵⁸. In contrast to CVID, sIgAD patients have intact IgG and IgM production, and these immunoglobulins appear to largely take over the function of IgA. In the gut, many bacteria that were IgA coated in HC were IgM coated in sIgAD⁵⁸, and the serum IgG of patients with sIgAD targets the microbiota much more than that of HC⁵⁶. This indicates that patients with CVID that are not only IgA but also IgG and often IgM impaired also lack these compensatory mechanisms that may limit microbial dysbiosis (IgM) and endotoxemia (IgG). Interestingly, sIgAD patients are often asymptomatic, and despite an increased risk of autoimmune disease relative to the general population, they are far less prone to develop immune dysregulation than patients with CVID.

Indeed, patients with CVID and low serum IgA levels are more prone to develop CVIDid⁶⁰, and in our studies had a disturbed oropharyngeal and fecal microbiota, more lung

damage, sporadic bacterial crypt invasion, and *Enterococcus* carriage. In the oropharynx and in the feces alike, IgA deficiency was associated with increased bacterial load. Alpha diversity, however, differed between the two sites despite large overlap between the two cohorts, and sampling occurring at the same time. In the respiratory tract (**chapter 6**), IgA deficiency was associated with an increased alpha diversity, which indicates a larger variety of bacteria that is equally distributed, while in feces (**chapter 7**), IgA deficiency was associated with decreased alpha diversity, indicating fewer different bacteria, and/or outgrowth of a few dominant taxa. In feces, higher alpha diversity is associated with health, but this is less clear for the respiratory microbiota. The gut and the respiratory tract are two rather different tissues. The gut epithelium is protected by a thick mucus layer, and the microbiota is necessary for a variety of metabolic processes that are essential to the main function of the gut – to absorb nutrients. The lung mucus layer is much thinner and in constant movement⁶¹, and while the microbiota there might provide colonization resistance to pathogens, it is not thought to contribute to the main function of the lung – the exchange of gases. While Muc2 is the dominant mucus type in the gut, the lung is mostly coated with Muc5AC and Muc5B⁶². It is possible that IgA-Muc interactions tailor the effect of IgA at different mucosal sites. In the gut, IgA appears to anchor bacteria and Muc2 together^{55,63}, forming a layer on top of the outer mucus layer. How IgA interacts with other types of mucus such as those expressed in the respiratory tract remains to be investigated.

OVERARCHING PERSPECTIVES

The immune system and the microbiota are so closely intertwined that discussing them separately is perhaps a somewhat artificial exercise. An undoubtedly oversimplified overview of the interactions between the immune system and the microbiota that may be at play in CVID is given in Figure 1. To summarize, the combination of gut microbial dysbiosis and loss of immune tolerance in CVID is a particularly unlucky one. The antibody deficiency in CVID itself can lead to impaired maintenance of microbiota homeostasis, and this effect is amplified by the often frequent need for antibiotic therapy. Dysbiosis of the microbiota can allow colonization with pathobionts⁶⁴, which in turn can impair the gut barrier function, for instance by degrading mucus⁶⁵ or by forming pores in the gut epithelium⁶⁶, and thereby also contribute to increased contact between the microbiota and the immune system, or even translocation of bacterial products). The resulting chronic antigenic exposure can provide a strong stimulus to an already dysregulated immune system. CVID B-cells can produce antibodies that recognize both gut commensals and self-antigens²⁶, and endotoxemia is associated with expansion of T follicular helper cells²², which are thought to interfere with regulatory T-cell (Treg) function²⁶.

Additional factors that may be involved in microbiota-immune interactions in CVID include epigenetic mechanisms^{67,68}, other environmental factors that affect development of the microbiota such as birth mode and diet⁶⁹, other involved immune cells such as innate-like lymphoid cells³⁷, regulatory B-cells⁷⁰, interactions between Tregs, microbiota

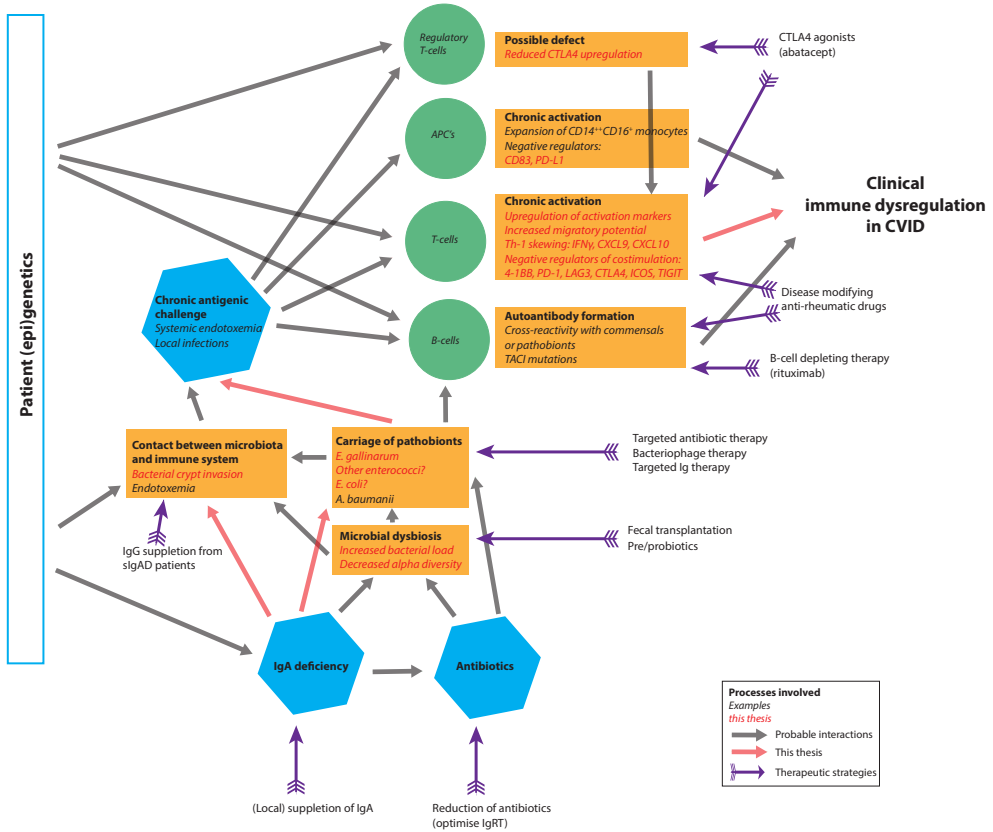


Figure 1: Integration of the possible mechanisms at play in the interaction between microbiome and immunity in CVID. CVID: common variable immunodeficiency. IgRT: immunoglobulin replacement therapy. slgAD: selective IgA deficiency.

and IgA production⁷¹, and other microbiota signaling mechanisms such as production of short-chain fatty acids⁷² and other metabolites⁷³, and this list is by no means meant to be exhaustive.

Despite the complexity of microbiota-immune interactions in CVID, improved understanding of these mechanisms provides novel therapeutic strategies. Immune dysregulation in CVID is now mostly treated using immune system-targeting strategies such as B-cell depletion using rituximab, T-cell suppression using purine synthesis inhibitors and other disease modifying anti-rheumatic drugs^{35,74-75}, and experimental treatments such as CTLA4 agonism using abatacept⁷⁶. However, microbiota-targeting therapies may become more important in the future. Supplementing IgRT with IgA in addition to IgG seems an obvious solution, but in practice this is complicated by anti-IgA antibodies in CVID patients. In one study, 11% of CVID patients had anti-IgA antibodies which predisposed to anaphylactoid reactions upon intravenous immunoglobulin replacement therapy with residual IgA concentrations⁷⁷. Local supplementation of IgA may therefore be more prac-

tical⁷⁸⁻⁸⁰. Alternatively, the composition of the gut microbiota may be modulated by fecal transplantation or administration of pre- or pro-biotics⁸¹. Direct targeting of pathobiont species may also be possible through use of small-spectrum antibiotics, or even dietary interventions in specific cases. For example, lactose was shown to drive gut colonization with *Enterococcus* in patients with graft versus host disease⁸². Novel interventions to target pathobiont species may include therapeutic anti-bacterial immunoglobulins⁸³ or even bacteriophages⁸⁴.

To conclude, immune dysregulation in CVID is a complex and heterogeneous disease that comprises profound chronic activation of the immune system and dysbiosis of the gut microbiota. In CVID, immune deficiency can result in microbial dysbiosis, including presence of specific pathobionts that are associated with autoimmune disease, which exacerbates the deregulation of the immune system and results in clinical disease. Better understanding of the specific mechanisms through which the gut microbiota affects the immune system in CVID and how this leads to clinical disease will be important for the development of new therapeutic interventions for CVID. In addition, study of the microbiota in primary antibody deficiencies such as CVID can contribute to our understanding of how the immune system shapes the microbiota and vice versa.

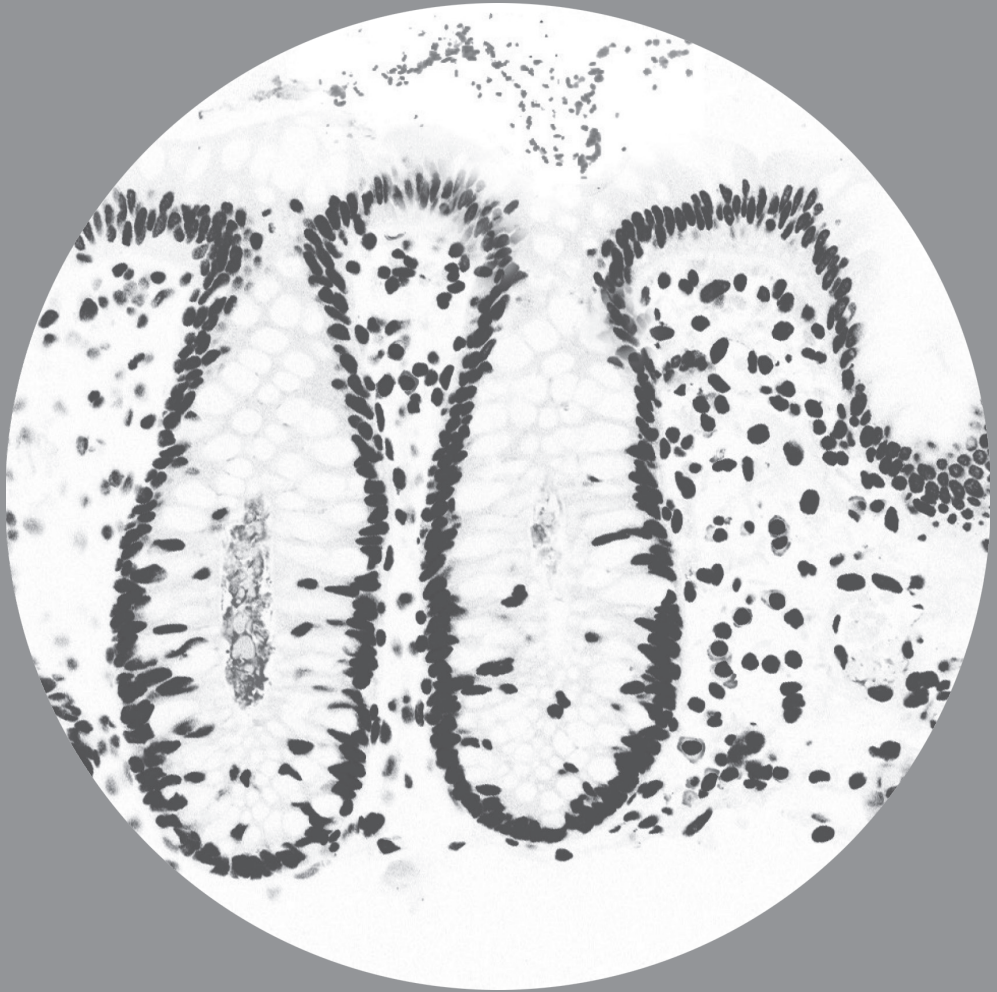
REFERENCES

1. Quinti, I. *et al.* Long-Term Follow-Up and Outcome of a Large Cohort of Patients with Common Variable Immunodeficiency. *J. Clin. Immunol.* **27**, 308–316 (2007).
2. Quinti, I. *et al.* Effectiveness of immunoglobulin replacement therapy on clinical outcome in patients with primary antibody deficiencies: Results from a multicenter prospective cohort study. *J. Clin. Immunol.* **31**, 315–322 (2011).
3. Resnick, E. S., Moshier, E. L., Godbold, J. H. & Cunningham-Rundles, C. Morbidity and mortality in common variable immune deficiency over 4 decades. *Blood* **119**, 1650–1658 (2012).
4. Maarschalk-Ellebrouck, L. J., Hoepelman, A. I. M., Van Montfrans, J. M. & Ellebrouck, P. M. The spectrum of disease manifestations in patients with common variable immunodeficiency disorders and partial antibody deficiency in a university hospital. *J. Clin. Immunol.* **32**, 907–921 (2012).
5. Janssen, L. M. A., van der Flier, M. & de Vries, E. Lessons Learned From the Clinical Presentation of Common Variable Immunodeficiency Disorders: A Systematic Review and Meta-Analysis. *Front. Immunol.* **12**, (2021).
6. Gathmann, B. *et al.* Clinical picture and treatment of 2212 patients with common variable immunodeficiency. *J. Allergy Clin. Immunol.* **134**, (2014).
7. Westh, L. *et al.* Identification and Characterization of a Nationwide Danish Adult Common Variable Immunodeficiency Cohort. *Scand. J. Immunol.* **85**, 450–461 (2017).
8. Dehner, C., Fine, R. & Kriegel, M. A. The microbiome in systemic autoimmune disease: Mechanistic insights from recent studies. *Curr. Opin. Rheumatol.* **31**, 201–207 (2019).
9. Barbosa, R. R. *et al.* Monocyte activation is a feature of common variable immunodeficiency irrespective of plasma lipopolysaccharide levels. *Clin. Exp. Immunol.* **169**, 263–272 (2012).
10. Siedlar, M. *et al.* Preparations of intravenous immunoglobulins diminish the number and proinflammatory response of CD14+CD16++ monocytes in common variable immunodeficiency (CVID) patients. *Clin. Immunol.* **139**, 122–132 (2011).
11. Tsukamoto, M. *et al.* CD14brightCD16+ intermediate monocytes are induced by interleukin-10 and positively correlate with disease activity in rheumatoid arthritis. *Arthritis Res. Ther.* **19**, 1–10 (2017).
12. Scott-Taylor, T. H., Green, M. R. J., Eren, E. & Webster, A. D. B. Monocyte derived dendritic cell responses in common variable immunodeficiency. *Clin. Exp. Immunol.* **138**, 484–490 (2004).
13. Hong, R., Agrawal, S., Gollapudi, S. & Gupta, S. Impaired pneumovax-23-Induced monocyte-derived cytokine production in patients with common variable immunodeficiency. *J. Clin. Immunol.* **30**, 435–441 (2010).
14. Cunningham-Rundles, C. & Radigan, L. Deficient IL-12 and dendritic cell function in common variable immune deficiency. *Clin. Immunol.* **115**, 147–153 (2005).
15. Saraiva, M. *et al.* Biology and therapeutic potential of interleukin-10. *J. Exp. Med.* **217**, 1–19 (2020).
16. Brooks, D. G. *et al.* Interleukin-10 determines viral clearance or persistence in vivo. *Nat. Med.* **12**, 1301–1309 (2006).
17. Li, Z. *et al.* CD83: Activation marker for antigen presenting cells and its therapeutic potential. *Front. Immunol.* **10**, 1–9 (2019).
18. de Jong, P. R. *et al.* STAT3 regulates monocyte TNF-alpha production in systemic inflammation caused by cardiac surgery with cardiopulmonary bypass. *PLoS One* **7**, 1–9 (2012).
19. Saha, D. C., Astiz, M. E., Eales-Reynolds, L. J. & Rackow, E. C. Lipopolysaccharide-and superantigen-modulated superoxide production and monocyte hyporesponsiveness to activating stimuli in Sepsis. *Shock* **38**, 43–48 (2012).
20. Flier, S. *et al.* Monocyte hyporesponsiveness and toll-like receptor expression profiles in coronary artery bypass grafting and its clinical implications for postoperative inflammatory response and pneumonia : An observational cohort study. *Eur. J. Anaesthesiol.* **32**, 177–188 (2015).
21. Jorgensen, S. F. *et al.* Altered gut microbiota profile in common variable immunodeficiency associates with levels of lipopolysaccharide and markers of systemic immune activation. *Mucosal Immunol.* **9**, 1455–1465 (2016).
22. Le Coz, C. *et al.* Common variable immunodeficiency-associated endotoxemia promotes early commitment to the T follicular lineage. *J. Allergy Clin. Immunol.* **144**, 1660–1673 (2019).

23. Warnatz, K. *et al.* Severe deficiency of switched memory B cells (CD27+IgM-IgD-) in subgroups of patients with common variable immunodeficiency: A new approach to classify a heterogeneous disease. *Blood* **99**, 1544–1551 (2002).
24. Bonilla, F. A. *et al.* International Consensus Document (ICON): Common Variable Immunodeficiency Disorders. *J. Allergy Clin. Immunol. Pract.* **4**, 38–59 (2016).
25. Romberg, N. *et al.* CVID-associated TACI mutations affect autoreactive B cell selection and activation. *J. Clin. Invest.* **123**, 4283–4293 (2013).
26. Romberg, N. *et al.* Patients with common variable immunodeficiency with autoimmune cytopenias exhibit hyperplastic yet inefficient germinal center responses. *J. Allergy Clin. Immunol.* **143**, 258–265 (2019).
27. Isnardi, I. *et al.* Complement receptor 2/CD21- human naive B cells contain mostly autoreactive unresponsive clones. *Blood* **115**, 5026–5036 (2010).
28. Wehr, C. *et al.* The EUROclass trial: defining subgroups in common variable immunodeficiency. *Blood* **111**, 77–85 (2008).
29. Moratto, D. *et al.* Combined decrease of defined B and T cell subsets in a group of common variable immunodeficiency patients. *Clin. Immunol.* **121**, 203–214 (2006).
30. Wehr, C. *et al.* A new CD21 low B cell population in the peripheral blood of patients with SLE. *Clin. Immunol.* **113**, 161–171 (2004).
31. Moir, S. *et al.* HIV-1 induces phenotypic and functional perturbations of B cells in chronically infected individuals. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10362–10367 (2001).
32. Jolles, S., Carne, E. & Brouns, M. FDG PET-CT imaging of therapeutic response in granulomatous lymphocytic interstitial lung disease (GLILD) in common variable immunodeficiency (CVID). 138–145 (2016) doi:10.1111/cei.12856.
33. Lougaris, V. *et al.* Long-term follow-up of 168 patients with X-linked agammaglobulinemia reveals increased morbidity and mortality. *J. Allergy Clin. Immunol.* 1–9 (2020) doi:10.1016/j.jaci.2020.03.001.
34. Rao, N., MacKinnon, A. C. & Routes, J. M. Granulomatous and lymphocytic interstitial lung disease: A spectrum of pulmonary histopathologic lesions in common variable immunodeficiency - Histologic and immunohistochemical analyses of 16 cases. *Hum. Pathol.* **46**, 1306–1314 (2015).
35. Chase, N. M. *et al.* Use of combination chemotherapy for treatment of granulomatous and lymphocytic interstitial lung disease (GLILD) in patients with common variable immunodeficiency (CVID). *J. Clin. Immunol.* **33**, 30–39 (2013).
36. Mannon, P. J. *et al.* Excess IL-12 but not IL-23 Accompanies the Inflammatory Bowel Disease Associated With Common Variable Immunodeficiency. *Gastroenterology* **131**, 748–756 (2006).
37. Cols, M. *et al.* Expansion of inflammatory innate lymphoid cells in patients with common variable immune deficiency. *J. Allergy Clin. Immunol.* **137**, 1206–1215.e6 (2016).
38. Petrelli, A. *et al.* PD-1+CD8+ T cells are clonally expanding effectors in human chronic inflammation. *J. Clin. Invest.* **128**, 4669–4681 (2018).
39. Hultberg, J., Ernerudh, J., Larsson, M., Nilsson-Augustinsson, Å. & Nyström, S. Plasma protein profiling reflects TH1-driven immune dysregulation in common variable immunodeficiency. *J. Allergy Clin. Immunol.* 1–12 (2020) doi:10.1016/j.jaci.2020.01.046.
40. Perreau, M. *et al.* Exhaustion of bacteria-specific CD4 T cells and microbial translocation in common variable immunodeficiency disorders. *J. Exp. Med.* **211**, 2033–2045 (2014).
41. Stuchlý, J. *et al.* Common Variable Immunodeficiency patients with a phenotypic profile of immunosenescence present with thrombocytopenia. *Sci. Rep.* **7**, 1–13 (2017).
42. Bates, C. A. *et al.* Granulomatous-lymphocytic lung disease shortens survival in common variable immunodeficiency Pulmonary disorders. *J. Allergy Clin. Immunol.* **5**, 415–421 (2004).
43. Hampson, F. A. *et al.* Respiratory disease in common variable immunodeficiency and other primary immunodeficiency disorders. *Clin. Radiol.* **67**, 587–595 (2012).
44. Berbers, R. M., Nierkens, S., van Laar, J. M., Bogaert, D. & Leavis, H. L. Microbial Dysbiosis in Common Variable Immune Deficiencies: Evidence, Causes, and Consequences. *Trends Immunol.* **38**, 206–216 (2017).

45. Fiedorová, K. *et al.* Bacterial but not fungal gut microbiota alterations are associated with common variable immunodeficiency (CVID) phenotype. *Front. Immunol.* **10**, (2019).
46. van Schewick, C. M. *et al.* Altered Microbiota, Impaired Quality of Life, Malabsorption, Infection, and Inflammation in CVID Patients With Diarrhoea. *Front. Immunol.* **11**, 1–13 (2020).
47. Shulzhenko, N. *et al.* CVID enteropathy is characterized by exceeding low mucosal IgA levels and interferon-driven inflammation possibly related to the presence of a pathobiont. *Clin. Immunol.* **197**, 139–153 (2018).
48. Daillère, R. *et al.* Enterococcus hirae and Barnesiella intestinihominis Facilitate Cyclophosphamide-Induced Therapeutic Immunomodulatory Effects. *Immunity* **45**, 931–943 (2016).
49. Mirsepassi-Lauridsen, H. C., Vallance, B. A., Krogfelt, K. A. & Petersen, A. M. Escherichia coli pathobionts associated with inflammatory bowel disease. *Clin. Microbiol. Rev.* **32**, 1–16 (2019).
50. Leavis, H. L. *et al.* Insertion sequence-driven diversification creates a globally dispersed emerging multiresistant subspecies of E. faecium. *PLoS Pathog.* **3**, 0075–0096 (2007).
51. Hansson, G. C. & Johansson, M. E. V. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Gut Microbes* **1**, 51–54 (2010).
52. Johansson, M. E. V. *et al.* Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. *Gut* **63**, 281–291 (2014).
53. Raskov, H., Kragh, K. N., Bjarnsholt, T., Alamili, M. & Gögenur, I. Bacterial biofilm formation inside colonic crypts may accelerate colorectal carcinogenesis. *Clin. Transl. Med.* **7**, 2–5 (2018).
54. Huus, K. E., Petersen, C. & Finlay, B. B. Diversity and dynamism of IgA–microbiota interactions. *Nat. Rev. Immunol.* **0123456789**, (2021).
55. Rogier, E. W., Frantz, A. L., Bruno, M. E. C. & Kaetzel, C. S. Secretory IgA is concentrated in the outer layer of colonic mucus along with gut bacteria. *Pathogens* **3**, 390–403 (2014).
56. Daniel, N., Lécuyer, E. & Chassaing, B. Host / microbiota interactions in health and diseases — Time for mucosal microbiology ! *Mucosal Immunol.* (2021) doi:10.1038/s41385-021-00383-w.
57. Milner, J. D. & Holland, S. M. The cup runneth over: Lessons from the ever-expanding pool of primary immunodeficiency diseases. *Nat. Rev. Immunol.* **13**, 635–648 (2013).
58. Fadlallah, J. *et al.* Microbial ecology perturbation in human IgA deficiency. *Sci. Transl. Med.* **10**, 1–16 (2018).
59. Fadlallah, J., Sterlin, D., Fieschi, C. & Parizot, C. Synergistic convergence of microbiota-specific systemic IgG and secretory IgA. *J. Allergy Clin. Immunol.* **143**, 1575–1585 (2018).
60. Giovannetti, A. *et al.* Unravelling the complexity of T cell abnormalities in common variable immunodeficiency. *J. Immunol.* **178**, 3932–43 (2007).
61. Salisbury, M. L. *et al.* The microbiome in interstitial lung disease: from pathogenesis to treatment target. *Curr Opin Pulm Med.* **23**, 404–410 (2017).
62. Thai, P., Loukoianov, A., Wachi, S. & Wu, R. Regulation of airway mucin gene expression. *Annu. Rev. Physiol.* **70**, 405–429 (2008).
63. Hendrickx, A. P. A. *et al.* Antibiotic-driven dysbiosis mediates intraluminal agglutination and alternative segregation of enterococcus faecium from the intestinal epithelium. *MBio* **6**, 1–11 (2015).
64. Ubeda, C. *et al.* Vancomycin-resistant Enterococcus domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans. *J. Clin. Invest.* **120**, 4332–4341 (2010).
65. Wright, D. P., Rosendale, D. I. & Roberton, A. M. Prevotella enzymes involved in mucin oligosaccharide degradation and evidence for a small operon of genes expressed during growth on mucin. *FEMS Microbiol. Lett.* **190**, 73–79 (2000).
66. Nakamoto, N. *et al.* Gut pathobionts underlie intestinal barrier dysfunction and liver T helper 17 cell immune response in primary sclerosing cholangitis. *Nat. Microbiol.* **4**, 492–503 (2019).
67. Jørgensen, S. F., Fevang, B. & Aukrust, P. Autoimmunity and Inflammation in CVID: a Possible Crosstalk between Immune Activation, Gut Microbiota, and Epigenetic Modifications. *J. Clin. Immunol.* **39**, 30–36 (2019).
68. Rodríguez-Cortez, V. C. *et al.* Monozygotic twins discordant for common variable immunodeficiency reveal impaired DNA demethylation during naïve-to-memory B-cell transition. *Nat. Commun.* **6**, (2015).

69. Bokulich, N. A. *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* **8**, 1–14 (2016).
70. Kofod-Olsen, E. *et al.* Altered fraction of regulatory B and T cells is correlated with autoimmune phenomena and splenomegaly in patients with COVID. *Clin. Immunol.* **162**, 49–57 (2016).
71. Kawamoto, S. *et al.* Foxp3 + T Cells Regulate Immunoglobulin A Selection and Facilitate Diversification of Bacterial Species Responsible for Immune Homeostasis. *Immunity* **41**, 152–165 (2014).
72. Mohammed, A. D. *et al.* Gut antibody deficiency in a mouse model of covid results in spontaneous development of a gluten-sensitive enteropathy. *Front. Immunol.* **10**, 1–19 (2019).
73. Macpherson, M. E. *et al.* Gut Microbiota-Dependent Trimethylamine N-Oxide Associates With Inflammation in Common Variable Immunodeficiency. *Front. Immunol.* **11**, (2020).
74. Cunningham-Rundles, C. How I treat common variable immune deficiency. *Blood* **116**, 7–15 (2010).
75. Hurst, J. R. *et al.* British Lung Foundation/United Kingdom Primary Immunodeficiency Network Consensus Statement on the Definition, Diagnosis, and Management of Granulomatous-Lymphocytic Interstitial Lung Disease in Common Variable Immunodeficiency Disorders. *J. Allergy Clin. Immunol. Pract.* **5**, 938–945 (2017).
76. Spee-mayer, C. Von, Echternach, C., Agarwal, P. & Gutenberger, S. Abatacept Use Is Associated with Steroid Dose Reduction and Improvement in Fatigue and CD4-Dysregulation in COVID Patients with Interstitial Lung Disease. *J. Allergy Clin. Immunol. Pract.* **9**, 760-770.e10 (2020).
77. Horn, J. *et al.* Anti-IgA antibodies in Common Variable Immunodeficiency (COVID): Diagnostic workup and therapeutic strategy. *Clin. Immunol.* **122**, 156–162 (2007).
78. Vonarburg, C. *et al.* Topical application of nebulized human IgG, IgA and IgAM in the lungs of rats and non-human primates. *Respir. Res.* **20**, 1–16 (2019).
79. Brodzki, N. Add-on or alone? Inhaled nebulized immunoglobulin reduces upper airway infections: 24 months of real-life experience. *Immunotherapy* **12**, 389–394 (2020).
80. Tran, A. C. *et al.* Mucosal Therapy of Multi-Drug Resistant Tuberculosis With IgA and Interferon- γ . *Front. Immunol.* **11**, 1–12 (2020).
81. Pamer, E. G. Fecal microbiota transplantation: Effectiveness, complexities, and lingering concerns. *Mucosal Immunol.* **7**, 210–214 (2014).
82. Stein-Thoeringer, C. K. *et al.* Lactose drives Enterococcus expansion to promote graft-versus-host disease. *Science (80-.)*. **366**, 1143–1149 (2019).
83. Wagner, E. K. & Maynard, J. A. Engineering therapeutic antibodies to combat infectious diseases. *Curr. Opin. Chem. Eng.* **19**, 131–141 (2018).
84. Gelman, D. *et al.* Combined bacteriophages and antibiotics as an efficient therapy against VRE Enterococcus faecalis in a mouse model. *Res. Microbiol.* **169**, 531–539 (2018).



Appendices

Nederlandse samenvatting

Curriculum Vitae

List of Publications

Dankwoord

Nederlandse samenvatting (voor niet-ingewijden)

De ziekte *Common Variable Immunodeficiency* (CVID) is een stoornis van het afweersysteem waarbij er verminderde aanmaak van afweerstoffen (de zogeheten immuunglobulines) is. Door dit gebrek aan immuunglobulines krijgen deze patiënten vaak ernstige bacteriële infecties, met name van de luchtwegen (zoals longontsteking) en het maag-darmstelsel. Deze infecties kunnen grotendeels worden voorkomen door immuunglobuline G (IgG) van gezonde donoren aan patiënten met CVID te geven. Ondanks deze behandeling krijgt ongeveer een derde van CVID patiënten complicaties die te maken hebben met dysregulatie van het immuunsysteem. Dit resulteert in allerlei uitingen van auto-immuniteit, vaak in de longen (interstitiële long ziekte), de darmen (darmontsteking zoals bijvoorbeeld bij de ziekte van Crohn), of in het bloed (afbraak van rode bloedcellen of bloedplaatjes). We noemen deze complicaties bij elkaar ook wel CVID met immuun dysregulatie (CVIDid), en zij veroorzaken veel ziekte en sterfte bij CVID patiënten.

Deze CVIDid patiënten worden dus ziek omdat hun afweersysteem niet goed reageert op bacteriën, maar paradoxaal genoeg wel het eigen lichaam aanvalt. De oorzaak van CVID en CVIDid is grotendeels onbekend, en in slechts 10% van de gevallen kan CVID worden verklaard door een genetische mutatie. In dit proefschrift onderzoeken we de hypothese dat de bacteriën die in en op ons lichaam leven – de zogeheten bacteriële microbiota – een rol spelen in het veroorzaken van immuun dysregulatie in CVID.

De microbiota is de verzameling van micro-organismen die wij rondragen in onze darmen, in onze luchtwegen en op onze huid. Veel van deze bacteriën zijn niet of moeilijk te kweken in het lab, maar sinds de ontwikkeling van DNA-sequencing technologie is steeds meer over hen bekend. De gezonde microbiota bevat voor het grootste deel niet-ziekteverwekkende bacteriën ('commensalen'), en draagt bij aan de ontwikkeling van ons afweersysteem en bescherming tegen bacteriën die wel ziekteverwekkend zijn ('pathogenen'). Veranderingen in de microbiota van de darm zijn inmiddels gevonden bij veel verschillende soorten ziekten, waaronder auto-immuun ziekten. In **hoofdstuk 4** wordt uitgelegd waarom wij denken dat ook in immuun dysregulatie bij CVID de microbiota een belangrijke rol kan spelen: in CVID is niet alleen IgG verlaagd, maar vaak ook IgA en/of IgM. Met name IgA wordt uitgescheiden in slijmvliezen zoals de darm en de luchtwegen, en reguleert de microbiota. Het verlaagde of afwezige IgA in CVID zou, in combinatie met andere factoren, kunnen leiden tot een verstoorde microbiota, wat tot een toename van potentieel ziekte verwekkende bacteriën ('pathobionten') kan leiden, en daarmee immuun dysregulatie kunnen veroorzaken.

Deel 1: het afweersysteem in CVIDid

In **deel 1** van dit proefschrift onderzoeken we het afweer systeem in CVIDid. In **hoofdstuk 2** meten we in het bloed van twee cohorten van CVID patiënten signaalstoffen (cytokines en chemokines) van het afweersysteem, en we vergelijken deze tussen patiënten met CVID

Did en patiënten met CVID met alleen infecties (*infection only*, CVIDio). We laten zien dat signaalstoffen IL-10, IL12RB1 en CD83 kunnen voorspellen wie in het tweede cohort CVIDid en wie CVIDio heeft. Deze stoffen zouden in de toekomst gebruikt kunnen worden als biomarkers voor CVIDid, bijvoorbeeld om CVIDid snel op te kunnen sporen. Verder vergelijken we CVIDid met CVIDio in de twee cohorten bij elkaar, waarbij veel stoffen duiden op sterke activatie van T-cellen in het bloed (CXCL9, CXCL10, IL-17a, IL-12). T-cellen hebben de capaciteit om cellen die niet in het eigen lichaam horen (zoals bacteriën) te herkennen en te doden. Verder kunnen zij B-cellen aansturen om immuunglobulines te produceren die ook bijdragen aan het doden van lichaamsvreemde cellen of stoffen. Als T-cellen het eigen lichaam aanzien als lichaamsvreemd kan dit leiden tot auto-immuunziekten. Daarom zijn er ook veel signaalstoffen om T-cel activatie zo nodig te kunnen remmen. Deze stoffen zien we sterk verhoogd in het bloed van CVIDid patiënten (LAG3, 4-1BB, PD-L1). De regulatie lijkt dus wel in gang te zijn gezet, maar desondanks hebben deze patiënten nog steeds auto-immuun ziekten.

In **hoofdstuk 3** kijken we of de patronen die we vonden in **hoofdstuk 2** ook terug te vinden zijn op de T-cellen zelf. De T-cellen van CVIDid patiënten laten inderdaad sterke activatie en productie van signaalstoffen (cytokines) zien. We vinden ook de regulerende eiwitten terug op de T-cellen (met name LAG3 en PD1 en anderen zoals CTLA4, ICOS en TIGIT), maar de T-cellen die LAG3 en PD1 tot expressie brengen zijn even goed in staat om te delen en signaalstoffen te produceren als T-cellen die deze markers niet hebben. Dit bevestigt dat ondanks de expressie van regulerende eiwitten, de T-cellen in CVIDid toch functioneel zijn en mogelijk bij kunnen dragen aan de ziekte. Verder valt op dat regulatorische T-cellen, een groep cellen die specifiek andere T-cellen remmen, in CVIDid een eiwit missen dat belangrijk is om hun functie goed uit te kunnen voeren (CTLA4). Expressie van regulerende eiwitten wordt vaak gezien in de context van langdurige (chronische) activatie van het immuunsysteem. Deze activatie kan komen door chronische infecties, en waarschijnlijk ook door een verstoorde microbiota.

Deel 2: de microbiota in CVIDid

Deel 2 van dit proefschrift gaat over de samenstelling van de microbiota van CVID patiënten. In **hoofdstuk 4** gaan we in op de hypothese dat de microbiota een rol kan spelen in het ontstaan van immuun dysregulatie bij CVID, en in **hoofdstuk 5** vatten we samen wat er bekend is over welke bacteriën worden beïnvloed door IgA.

In **hoofdstuk 6** worden de bacteriën in de bovenste luchtwegen van patiënten met CVID onderzocht. Longschade is een van de grootste problemen van patiënten met CVID, en kan komen door structurele schade aan de luchtwegen, bijvoorbeeld door longontstekingen, of door een ontstekingsproces dat door immuun dysregulatie veroorzaakt kan worden. Dit laatste ziektebeeld wordt ook wel GLILD genoemd (*granulomatous-lymphocytic interstitial lung disease*). In de bovenste luchtwegen van patiënten met CVID en laag IgA in het bloed observeerden wij een toename van de totale hoeveelheid bacteriën ten opzichte van gezonde mensen en van CVID patiënten met normaal IgA in het bloed. Verder vonden wij in de groep met laag IgA een toename van de diversiteit van de microbi-

ota (meer verschillende bacteriën), en een toename van specifieke bacteriën *Prevotella* en *Alloprevotella*. Op longscans van deze patiënten hebben we de mate van structurele longschade en van GLILD gekwantificeerd. Patiënten met laag IgA hadden meer longschade, en de structurele longschade was ook geassocieerd met *Prevotella* en *Alloprevotella*. Het verband tussen de microbiota en GLILD was minder sterk. Bacteriën uit de Prevotellaceae familie zijn bekende pathobionten die ook de slijmlaag in de darm en mogelijk ook in de luchtwegen kunnen afbreken. De afwezigheid van IgA in CVID zou kunnen leiden tot toename van de totale hoeveelheid verschillende soorten bacteriën, waaronder (*Allo*)*Prevotella*, en zo kunnen bijdragen aan het ontstaan van longschade in CVID.

In **hoofdstuk 7** duiken we in de darmen van CVID patiënten. In darmbiopten van CVID patiënten laten we zien dat de microbiota soms diep in de crypten van de darm infiltreert, waar in gezonde mensen de crypten helemaal schoon zijn. Dit leidt waarschijnlijk tot toegenomen blootstelling van bacteriën aan het afweersysteem. Waar dit door komt weten we niet precies, het kan zijn dat de slijmlaag van de darm is aangetast, dat het door afwezigheid van IgA wordt veroorzaakt (alle patiënten waarin we dit zagen hadden laag IgA in hun bloed), of dat de patiënten specifieke bacteriën bij zich droegen die invasief zijn.

In de ontlasting hadden CVIDid patiënten een toename van de totale hoeveelheid bacteriën, een afname van de diversiteit van bacteriën, en een licht afwijkende samenstelling van de darmbacteriën vergeleken met CVIDio en gezonde controles. Bij vergelijking van alle CVID patiënten met gezonde controles zagen we een toename van pathobionten zoals de familie Enterobacteriaceae (waartoe *Escherichia coli* behoort), en een afname van commensalen als *Bifidobacterium*. Wanneer CVIDid met CVIDio vergeleken werd zagen we dat bacteriën van het genus *Enterococcus* alleen detecteerbaar waren in CVIDid, en niet in CVIDio. Onze analyses zijn gedaan middels 16S rRNA sequencing, waarmee een globaal overzicht van de microbiota verkregen kan worden, maar waarmee meestal niet exact te achterhalen valt om welke bacteriën het gaat. Met behulp van shotgun metagenomic sequencing, qPCR, en kweken (gelukkig zijn enterokokken wel kweekbaar in het lab) kan dit wel, en bleken met name *Enterococcus gallinarum* en *Enterococcus hirae* met immuun dysregulatie geassocieerd te zijn. Toen wij deze data naast de gegevens uit hoofdstuk 2 legden bleken patiënten met *E. gallinarum* in hun ontlasting ook meer signaalstoffen geassocieerd met immuun activatie in het bloed te hebben. Tot slot hebben we immuuncellen (monocyten) van gezonde controles, CVIDio en CVIDid patiënten blootgesteld aan *E. gallinarum* en *E. hirae*, waarop meer IL-6 (immuun activatie) dan IL-10 (immuun regulatie) werd geproduceerd. *E. gallinarum* is ook in de literatuur al eerder gevonden bij patiënten met autoimmuniteit, en ons onderzoek laat zien dat deze bacterie ook bij CVIDid een rol kan spelen.

Deel 3: statistische analyse van microbiom data

Omdat DNA sequencing van bacteriële ecosystemen zoals het microbiom een relatief nieuwe techniek is, staat ook de statistische analyse van dit type data nog in de kinderschoenen. Door de methode van DNA sequencing kan de aanwezigheid van bacteriën

alleen geïnterpreteerd worden als percentages binnen een sample (dit wordt compositionele genoemd). Daarnaast bevat de data veel nullen, omdat veel bacteriële soorten maar in enkele samples voorkomen en dus afwezig zijn in de rest (dit wordt schaarse genoemd). De compositionele van de data leidt tot problemen als traditionele statistische methoden worden toegepast omdat het veel onjuiste correlatie binnen de dataset kan introduceren. Om dit probleem te omzeilen wordt vaak een data transformatie (de *Centered Log Ratio*, of CLR) toegepast. Echter gebruikt CLR logaritmen, waardoor eerst de nullen in de data weg gewerkt moeten worden. In **hoofdstuk 8** laten we zien dat de meest gebruikte methode om de nullen te vervangen in combinatie met CLR juist nog meer onjuiste correlatie introduceert in de data, en dat in dit geval dus de oplossing mogelijk erger is dan de kwaal. We gebruiken onze data uit **hoofdstuk 6** om dit probleem aan te tonen, en suggereren een andere data transformatie die mogelijk beter functioneert in compositionele en schaarse datasets.

Conclusie

In dit proefschrift hebben we immuun dysregulatie van CVID benaderd vanaf de immunologische en de microbiologische kant. We hebben aangetoond dat het immuunsysteem van CVID patiënten sterk geactiveerd is, ondanks de aanwezigheid van immuun regulerende eiwitten en cellen, en dat dit gepaard gaat met verhoogd contact van de darm microbiota met het immuunsysteem, een verstoorde samenstelling van de microbiota, en de aanwezigheid van *E. gallinarum*, een sterk immuun-activerende bacterie. Deze bevindingen waren geassocieerd met IgA, net als in de bovenste luchtwegen, waar tekort aan IgA gepaard ging met longschade, ook een verstoorde microbiota, en aanwezigheid van *Prevotella*. Tot slot maken we een kleine bijdrage aan de statistische analyse van microbiota data.

Al met al draagt dit proefschrift een klein puzzelstukje bij aan het beter begrijpen van immuun dysregulatie in CVID. Hopelijk zullen verdere ontwikkelingen in het veld kunnen leiden tot gerichte therapie van immuun dysregulatie, ook in CVID. Het moduleren van de microbiota is bijvoorbeeld mogelijk door middel van een feces transplantatie, of door gerichte therapie tegen schadelijke bacteriën. Tot slot geeft het bestuderen van de microbiota in CVID ook inzichten in de belangrijke rol van IgA in het reguleren van de microbiota.

Curriculum Vitae

Roos-Marijn Berbers was born on April 20th 1991 in Amsterdam, the Netherlands. In 2009 she completed secondary education Cum Laude at the Berlage Lyceum in Amsterdam. That year she started her study Liberal Arts and Sciences at the Amsterdam University College where she received her Bachelor's degree Summa Cum Laude in 2012 and her interest in immunology was sparked. This resulted a master's degree in Integrated Immunology at the University of Oxford, St. Cross college, which she obtained in 2013. Her master's thesis research tackled the role of regulatory T cells in the immune dysregulation in PTEN hamartoma syndrome, which was supervised by Dr. Hannah Chen and Prof. Holm Uhlig of the John Radcliffe Hospital in Oxford. Subsequently, she enrolled in the 4-year Selective Utrecht Medical Master (SUMMA) at the university of Utrecht, where she obtained her medical degree in 2017. During the SUMMA programme, she started a research internship in 2014 at the department of Rheumatology and Clinical Immunology with Dr. Helen Leavis, which got out of hand and culminated in a PhD project, the results of which are presented in this thesis. During this PhD, she closely collaborated with the group of Prof. Rob Willems from the department of Medical Microbiology, the group of Prof. Femke van Wijk from the Center for Translational Immunology, and the group of Dr. Hae-Won Uh from the Julius Center for Biostatistics.

Roos lives in Maarssen with Jonas and their daughter Maartje, and is currently working as a resident in Internal Medicine at the St. Antonius Hospital in Nieuwegein.

List of Publications

Publications related to this thesis:

- Berbers, R. M.**, Nierkens, S., van Laar, J. M., Bogaert, D. & Leavis, H. L. Microbial Dysbiosis in Common Variable Immune Deficiencies: Evidence, Causes, and Consequences. *Trends Immunol.* 38, 206–216 (2017).
- Berbers, R. M.**, Franken, I. A. & Leavis, H. L. Immunoglobulin A and microbiota in primary immunodeficiency diseases. *Curr. Opin. Allergy Clin. Immunol.* 19, 563–570 (2019).
- Berbers, R. M.**, Mohamed Hoesein, F.A.A., Ellerbroek, P.M., van Montfrans, J.M., Dalm, V.A.S.H., van Hagen, P.M., Paganelli, F.L., Viveen, M.C., Rogers, M.R.C., de Jong, P.A., Uh, H.W., Willems, R.J.L., Leavis, H.L. Low IgA Associated With Oropharyngeal Microbiota Changes and Lung Disease in Primary Antibody Deficiency. *Front. Immunol.* 11, 1–11 (2020).
- Berbers, R. M.**, Drylewicz, J., Ellerbroek, P.M., van Montfrans, J.M., Dalm, V.A.S.H., van Hagen, P.M., Keller, B., Warnatz, K., van de Ven, A., van Laar, J.M., Nierkens, S., Leavis, H.L. Targeted Proteomics Reveals Inflammatory Pathways that Classify Immune Dysregulation in Common Variable Immunodeficiency. *J. Clin. Immunol.* 362–373 (2020).
- Berbers, R.M.**, van der Wal, M.M., van Montfrans, J.M., Ellerbroek, P.M., Dalm, V.A.S.H., van Hagen, P.M., Leavis, H.L., van Wijk, F. J. Chronically Activated T-cells retain their Inflammatory Properties in Common Variable Immunodeficiency. *Clin. Immunol.* (2021).

Publications not related to this thesis:

- Chen, H. H., Händel, N., Ngeow, J., Muller, J., Hühn, M., Yang, H.T., Heindl, M., **Berbers, R.M.**, Hegazy, A.N., Kionke, J., Yehia, L., Sack, U., Bläser, F., Rensing-Ehl, A., Reifemberger, J., Keith, J., Travis, S., Merckenschlager, A., Kiess, W., Wittekind, C., Walker, L., Ehl, S., Aretz, S., Dustin, M.L., Eng, C., Powrie, F., Uhlig, H.H. Immune dysregulation in patients with PTEN hamartoma tumor syndrome: Analysis of FOXP3 regulatory T cells. *J. Allergy Clin. Immunol.* 139, 607–620.e15 (2017).
- Paganelli, F. L., Luyer, M., Hazelbag, C.M., Uh, H.W., Rogers, M.R.C., Adriaans, D., **Berbers, R.M.**, Hendrickx, A.P.A., Viveen, M.C., Groot, J.A., Bonten, M.J.M., Fluit, A.C., Willems, R.J.L., Leavis, H.L. Roux-Y Gastric Bypass and Sleeve Gastrectomy directly change gut microbiota composition independent of surgery type. *Sci. Rep.* 9, 1–8 (2019).
- van der Houwen, T. B., van Laar, J.A.M., Kappen, J.H., van Hagen, P.M., de Zoete, M.R., van Muijlwijk, G.H., **Berbers, R.M.**, Fluit, A.C., Rogers, M., Groot, J., Hazelbag, C.M., Consolandi, C., Severgnini, M., Peano, C., D’Elios, M., Emmi, G., Leavis, H.L. Behçet’s Disease Under Microbiotic Surveillance? A Combined Analysis of Two Cohorts of Behçet’s Disease Patients. *Front. Immunol.* 11, 1–10 (2020).

Dankwoord

Lieve Helen, dank je wel voor de ontzettend leuke 3 (en een beetje) afgelopen jaren. Vanaf het begin heb je vertrouwen in mij gehad, naar mijn ideeën geluisterd, me vrijheid gegeven om er zelf richting aan te geven, en tegelijkertijd me altijd ondersteund waar nodig. Daarnaast hebben we veel plezier gehad op congressen, samen geborrelt en veel gelachen. Er zijn denk ik weinig copromotoren die zo betrokken zijn als jij bent geweest. We zijn een goed team. Dank je wel!

Beste Rob, toen jij betrokken raakte bij het microbioom onderzoek en mijn promotietraject bracht je een oase van rust met je mee. Je was er altijd om knopen door te hakken, of om een luisterend oor te bieden als ik ergens mee zat. Daarnaast ben je natuurlijk inhoudelijk ook ijzersterk als het om onze favoriete bacteriën gaat!

Beste Pauline, mede dankzij jouw inzet hebben we dit prachtige CVID cohort kunnen opzetten. Daarna hebben we elkaar op congressen beter leren kennen, en weet ik bijvoorbeeld dat we een liefde voor oude huizen delen.

Beste Jaap, ik vond het altijd prettig om de voortang van mijn onderzoek aan je te komen vertellen. Je had altijd goede input over mijn immunologische onderzoekslijnen, en ik denk dat jij ook wel het een en ander over bacteriën van mij hebt geleerd. Daarnaast heb je me veel geadviseerd over de rest van mijn carrière.

Bas en Paul, mijn paranimfjes! Jullie waren mijn maatjes bij de immunologie en microbiologie, respectievelijk. Dit kwam bijeen tijdens een memorabele MMB kerstborrel in de Cambridge bar, waar werd meegezongen (*I believe in a thing called love*) tot we allemaal hees waren. Bas, als jij op een feestje/borrel/meeting verschijnt dan weet je dat het gezellig gaat worden (als je de meeting in kwestie tenminste niet vergeten was). Ik heb je ontzettend hard zien werken om al je projecten in de lucht te krijgen, om van de verbouwing van je huis nog maar niet te spreken, en ik weet zeker dat jouw proefschrift net zo shiny gaat worden als je nieuwe badkamer. Paul, ik ken weinig mensen met een muzieksmaak die zo eclectisch is als de jouwe. Van Taylor Swift tot Rammstein draaien de speakers in het lab altijd op volle toeren. Daarbij stuiter jij op een goeie dag minstens zo hard van de muren, als je tenminste geen verschrikkelijk experiment from hell aan het doen was. Ik weet niet precies hoeveel nieuwe labtechnieken jij hebt opgezet de afgelopen jaren, maar ik denk dat er eigenlijk wel twee Pauls hadden kunnen promoveren op jouw projecten (behalve dan dat ik niet weet of de afdeling dat had aangekund).

Dear Fernanda, you were a huge support to me throughout my PhD. Your door was always open for advice or a chat, whether it was research-related or personal. And of course, being

a veteran of the Helen Leavis and Rob Willems School of PhD, you often had invaluable insights into how to manage things. I had a lot of fun going to Ireland with you and Helen, and I know for a fact that my PhD would not have been the same experience without you.

Dan de zeer belangrijke overige leden van team primaire immuundeficiënties. Joris, als een van de drijvende krachten achter onze landelijke studie ben ook jij vanaf het begin betrokken geweest bij mijn project. Daarbij hield jij ook altijd heel goed in het vizier welke onderzoeksvragen het belangrijkste zijn voor deze patiëntengroep. Als een van de weinigen liet jij je niet tot wanhoop drijven als de monitor weer eens langs moest komen, en ook je gevoel voor humor wist je altijd mee te nemen naar onze besprekingen (ook als het over eindeloze METC amendementen ging). Je hebt ook een goede neus voor nieuwe onderzoeksprojecten, zoals onze samenwerking met Yvonne die toch maar mooi van de grond is gekomen. Yvonne, ik vond het ontzettend leuk om te zien hoe we met zijn allen van een goed idee naar funding naar een echt onderzoeksproject zijn gekomen. Natuurlijk dit alles niet zonder de vastberadenheid van Stephanie, die altijd dingen al heeft uitgevoerd nog voordat je zelf doorhad dat je ze geopperd hebt. Lilly en Mischa, onze microbiom partners in crime! Ik vond het altijd erg leuk om met jullie te sparren over jullie ideeën, nieuwe inzichten, en gezamenlijke struggles. Bas, jij bent natuurlijk al eerder aan bod gekomen, maar jouw persoonlijkheid laat toch nog wel wat ruimte voor een tweede vermelding hier. Ik vond het heel leuk toen ik jou 'subtiel' naar voren kon schuiven toen Joris op zoek was naar een nieuwe PhD student en vroeg of ik nog iemand wist. Jij fleurt elke bespreking altijd op, en draagt echt bij aan ons groepsgevoel. Marianne is daarbij onze nieuwste aanwinst. Jij hebt een van de woeste ideeën van Helen en mij opgepakt en bent onverschrokken aan de slag gegaan met ons huisartsen algoritme. Ondanks dat we elkaar nog maar 1 keer in het echt hebben gezien (Corona...) was je direct onderdeel van onze groep en bood je altijd je hulp aan waar dat nodig was.

Zo'n beetje de gehele microbiologie afdeling heeft het geweten dat Helen en ik een microbiom project wilde opstarten. En wat een fijne afdeling zijn jullie. Marco, toen ik begon aan mijn onderzoek wist ik eigenlijk vrij weinig van bacteriën. Jij hebt me geleerd hoe je écht in het lab moet werken met bacteriën, en was altijd geduldig als ik weer eens per ongeluk iets had aangeraakt wat eigenlijk niet mag. Nu ik weg ben kunnen jij en Paul weer ongestoord heavy metal draaien op het lab! Als de baas Moniek dat natuurlijk goed vindt. Moniek, ook jij ontzettend bedankt voor al je hulp in het lab en dat je nooit hebt geklaagd over mijn vreselijk ingewikkelde plaat lay-outs. Daarnaast moet ik natuurlijk ook Malbert bedanken voor al het bioinformatica werk dat je de afgelopen jaren voor me hebt gedaan. Ik hoop dat ik je niet te veel hoofdpijn heb opgeleverd! Dit laatste geldt ook voor Marcel, ik geloof dat Helen en ik je zo vaak hebben gestalkt in het diergeneeskunde gebouw dat je het maar hebt opgegeven en naar onze afdeling bent verhuisd. Jouw expertise en gezelligheid zijn een fantastische aanwinst voor de onderzoeksgroep. Dat geldt natuurlijk ook voor Guus, Coco en Jiannan die gelukkig met je mee kwamen! Verder mocht ik ook altijd Janetta en Ad lastig vallen met vragen waar ik niet uit kwam, of gewoon voor een praatje,

dank jullie wel daarvoor. To all the PhD, post-doc, and students of the MMB department, and especially my old roomies Hendrik, Axel, Shu, Eva, and Matteo: you're the best! You were always available for technical or emotional support, and had a lot of patience with my ignorance as a lowly medical student. Tot slot Marc, jij hebt mij lang geleden op het spoor van Helen gezet. Ik weet niet of je toen al vermoedde dat je daarna jaren lang aan me vast zou zitten. Je kwam ook regelmatig even peilen hoe het met mij ging op de afdeling als je signalen opving thuis dat het even tegen zat.

Aan de immunologie kant heb ik onmisbare hulp gekregen van Femke en Marlot, en natuurlijk ook van Stefan en Amelia. Femke, ik voelde me direct heel welkom in jouw onderzoeksgroep toen ik daar het laatste jaar van mijn onderzoek mocht aansluiten. Als Helen en ik weer eens met wilde ideeën aankwamen dacht je altijd met ons mee en vertelde je ons precies wat wel en niet haalbaar, danwel überhaupt een goed idee was. Marlot, zonder jouw hulp en tomeloze enthousiasme was mijn FACS experiment nooit van de grond gekomen (en in elk geval was het een stuk minder gezellig geweest!). De warmte van Femke straalt ook door in de rest van de van Wijk groep – dank jullie allemaal voor jullie input als ik weer eens een FACS compensatie probleem had, of mijn eigen data even niet begreep! Stefan, ontzettend bedankt voor je hulp met ons Olink project en alle ad-hoc vragen die Helen en ik je onderweg hebben gesteld ('Stefan weet dit vast, we bellen hem even').

Tijdens het maken van dit proefschrift heb ik veel hulp gekregen van de enthousiaste studenten Remi, Safae, Ingrid, Martine en Jaap. Het valt niet mee om vanuit je geneeskunde studie ineens het lab in te gaan, laat staan als je dan van mij onmogelijke opdrachten krijgt zoals het optimaliseren van DNA isolatie protocollen, qPCR assays, en ingewikkelde microscopie technieken. Jullie hebben je allemaal vol enthousiasme hierop gestort en ik heb heel veel geleerd van jullie.

Ook buiten het UMC heb ik veel hulp gekregen. In het Erasmus MC wil ik heel graag Prof. Martin van Hagen en dr. Virgil Dalm bedanken, die mij op hun poli en in hun lab hebben verwelkomd. Zonder jullie hadden we natuurlijk niet dit mooie cohort kunnen samenstellen. Martin, ontzettend bedankt voor je interesse in mijn project, ik heb altijd veel gehad aan jouw input! Ook vanuit het hoge noorden in het UMCG kwam ondersteuning in de vorm van dr. Annick van de Ven, meestal via hand-overs op het station van Utrecht. Ik vond het ook altijd erg gezellig om jou op congressen te treffen. All the way from Freiburg, Germany, many thanks to Prof. dr. Klaus Warnatz and dr. Bärbel Keller. Thank you for sharing your samples with us and collaborating on our projects. I'm very happy that Klaus was also willing to be part of my defence committee (a statement I hope I won't need to retract after the actual defence).

Lieve Summies, Fieke, Luc, Bernadette, Jelle (jij telt inmiddels ook wel mee als summie), Emma, Jonathan, Tanja, Sophie, Jasmijn, Esther, Wimkees, Amber (Bas is nu echt wel

genoeg aan bod geweest), dank jullie wel voor alle etentjes, borrels, sinterkerstdiners, slechte grappen, slechtere gedichten, en feestjes. Fatima, Fay, het maakt niet uit hoe vaak we elkaar zien (wanneer gaan we weer eens thee drinken??), ik weet dat ik altijd bij je terecht kan. Lindsay and Nick, I still feel very lucky that you chose to move to this little country after we met in Oxford. Thank you for all the emotional support, coffee, drinks, dinners, and fun facts.

Lieve Mathilde, Joris, Thijmen, Miranda en Wiepke, ik voel me altijd ontzettend welkom bij jullie. Dank voor alle gezelligheid, interesse in mijn onderzoek, en kampeeruitjes in de tuin.

Lieve Mama, Sanne en Linde. Er is niet zoveel dat ik hier kan zeggen zonder dat het gelijk enorm cheesy wordt, zeker aangezien Linde de Gilmore Girls speech al heeft gekaapt voor haar scriptie. Ik ben ontzettend trots op jullie en op ons bij elkaar. Ik weet dat ik en Maartje altijd op jullie kunnen rekenen. Dank jullie wel voor alles.

Lieve Jonas. We hadden een wedstrijdje wie als eerste klaar zou zijn met promoveren, en jij hebt natuurlijk glansrijk gewonnen Tussendoor kwam ook Maartje onze chaos verrijken. Ik ben altijd gelukkig als ik weer bij jullie thuis mag komen. Jullie zijn mijn feestje.

EINDE

