Charles University

Faculty of Science

Study program: Immunology

Bc. Klára Horáčková

Mechanisms of immune dysregulation leading to inflammatory bowel disease

Mechanismy imunitní dysregulace vedoucí k nespecifickému střevnímu zánětu

Diploma thesis

Supervisor: doc. MUDr. Eva Froňková, Ph.D.

Consultant: MUDr. Michael Svatoň

Prague, 2020

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 07.06.2020

…….

Klára Horáčková

Poděkování

Ráda bych poděkovala vedoucí práce doc. MUDr. Evě Froňkové, Ph.D. a konzultantovi práce MUDr. Michaelu Svatoňovi za možnost diplomovou práci pod jejich vedením vypracovat a hlavně za všechen jejich čas, energii a odborné know-how, které mi během tří let práce na projektu poskytli. Také děkuji kolegům bioinformatikům a cytometristům za skvělou atmosféru na pracovišti i ochotu pomoci, kdykoliv to bylo potřeba.

V neposlední řadě děkuji všem zaměstnancům CLIPu za jejich odbornou práci související s náplní diplomové práci, stejně tak dalším pracovníkům zapojeným v rámci klinické části projektu a účastníkům a jejich rodinám za možnost projekt vůbec zrealizovat.

Velké díky také patří mé rodině, příteli a kamarádům za psychickou podporu během celého studia.

Abstrakt

Nespecifický střevní zánět (IBD) je chronické zánětlivé onemocnění gastrointestinálního traktu. Klasický IBD má multifaktoriální charakter s nástupem v dospělosti nebo pozdním dětství. Dětští pacienti s velmi brzkým nástupem (VEO-IBD, do 6 let věku) jsou ale specifickou podskupinou, jejíž onemocnění je často způsobené závažnými mutacemi v genech souvisejících s imunitní homeostázou ve střevech.

Cílem této práce bylo identifikovat kauzální genetické varianty u 20 pacientů s pediatrickým IBD (s nástupem onemocnění od 3 do 154 měsíců) s využitím celoexomového sekvenování.. Vyhodnocovali jsme použití několika bioinformatických postupů pro zpracování celoexomových dat. To zahrnovalo porovnání rozdílů v identifikaci variant pomocí dvou bioinformatických postupů využívajících v prvním případě nástroje VarScan2 nebo ve druhém případě založeném na GATK4. Také jsme srovnávali úspešnost filtrování nalezených kauzálních variant pomocí čtyř virtuálních panelů genů, z nichž jeden panel byl vytvořen speciálně pro účely této diplomové práce. Identifikovali a validovali jsme 5 kauzálních variant ve 4 genech (*DUOX2* složený heterozygot, *FOXP3*, *NLRP3* and *NOD2*), které zastupují 20 % z analyzované kohorty pacientů. Varianta *NOD2* (p.A755V) byla už dříve popsána v literatuře jako kauzální u IBD, ale varianty v *DUOX2* (p.R1216W + p.A1131T), *FOXP3* (p.H400L) a *NLRP3* (p.V200M) jsou popsány poprvé. Navíc jsme navrhli dalších 6 variant v 5 genech jako možné kauzální varianty pro validaci. Tyto informace přispějí k poznání genetické podstaty VEO-IBD. Úspěšnost při identifikaci kauzálních variant byla u bioinformatického postupu využívajícího GATK4 stejná jako u staršího postupu založeném na VarScan2. Tím jsme potvrdili použitelnost nově vytvořeného bionformatického nástroje pro analýzu klinických dat. Určili jsme, že nejúčelnějším přístupem pro identifikaci variant u pacientů s VEO-IBD je použití virtuálního panelu s geny způsobujícími primární imunodeficity a také aktualizovaného panelu s kauzálními geny popsanými v případových studiích pacientů s projevy VEO-IBD.

Z výsledků vyplývá, že část případů VEO-IBD je možné chápat jako střevní projevy primárních imunodeficitů. Nicméně je nutné více funkčně charakterizovat a prozkoumat genetické i negenetické příčiny patogeneze k pochopení tohoto onemocnění.

Klíčová slova: celoexomové sekvenování, nespecifický střevní zánět, primární imunodeficience, imunitní dysregulace

Abstract

Inflammatory bowel disease (IBD) is a complex disorder characterized by chronic inflammation of the gastrointestinal tract. Classical IBD is a multifactorial disease with adulthood or later-childhood onset. However, children with very early onset IBD (VEO-IBD, before 6 years of age) are a specific cohort, whose pathology can be caused by severe genetic defects in genes connected to immune homeostasis in the gut.

We aimed to identify the causal genetic variants in 20 pediatric patients diagnosed with IBD (age of onset from 3 to 154 months) using whole exome sequencing (WES). We evaluated several bioinformatical approaches for WES data analysis. This included a comparison of two methods of variant identification using VarScan2 or GATK4-based tools. Furthermore, we compared 4 gene lists ("virtual panels") for variant filtering, one of which was compiled purposefully for this thesis.

We identified and validated via segregation analysis 5 causal variants in 4 genes (*DUOX2* compound heterozygote, *FOXP3*, *NLRP3* and *NOD2*) accounting for 20 % of the cohort. NOD2 (p.A755V) variant has already been reported in IBD cases, while *DUOX2* (p.R1216W + p.A1131T), *FOXP3* (p.H400L) and *NLRP3* (p.V200M) were newly discovered variants in this context. Moreover, we suggested 6 more variants in 5 genes for further validation, bringing new insight into genetics of VEO-IBD. Finding the same success rate of causal variant calling and filtering, we proved the suitability of the newly developed GATK4-based pipeline for variant analysis in a clinical setting. We found that the most adequate approach for variant filtering in VEO-IBD cases is using virtual panels of genes related to primary immunodeficiencies and up-to-date reported VEO-IBD cases.

In conclusion, we strongly support the hypothesis of many VEO-IBD cases being primary immunodeficiency with gastrointestinal manifestation. However, additional functional tests and research into genetic and non-genetic causes are required to understand molecular mechanisms of the VEO-IBD pathogenesis.

Key words: whole exome sequencing, inflammatory bowel disease, primary immunodeficiency, immune dysregulation

Contents

List of abbreviations 8
INTRODUCTION 10
THEORETICAL PART 12
1. Inflammatory bowel disease 12
1.1. Classification of IBD 12
1.2. Causes of IBD 15
1.3. Very Early Onset IBD 19
2. Molecular mechanisms of VEO-IBD 23
2.1. Epithelial barrier defects 23
2.2. Phagocyte defects 25
2.3. Lymphocyte defects 27
2.3.1. B lymphocyte and antibody defects 28
2.3.2. Effector T lymphocyte defects 29
2.3.3. Regulatory T lymphocyte defects 33
2.4. Systemic inflammatory disorders 35
2.5. Complement defects 38
3. Diagnostics of VEO-IBD 40
3.1. Methodological approaches 40
3.2. Variant filtering criteria 42
3.3. Validation of the variants 45
PRACTICAL PART 47
4. Aims of the thesis 47
5. Materials and methods 48
5.1. Patients 48
5.2. Samples preparation and sequencing 49
5.3. WES data analysis 49
5.3.1. Bioinformatical raw sequencing data processing 49
5.3.2. Variant filtering 50
5.3.3. Identification of the causal variant 55
5.4. Validation of the variants 55
5.5. Evaluation of different bioinformatical approaches 58
5.5.1. Bioinformatical pipelines 58
5.5.2. Gene lists for variant filtering 59

6. Results ... 60
6.1. Identified causal variants in patients ... 60
6.1.1. Patient #2 ... 62
6.1.2. Patient #3 ... 65
6.1.3. Patient #4 ... 66
6.1.4. Patient #8 ... 68
6.1.5. Patient #10 ... 69
6.1.6. Patient #11 ... 70
6.1.7. Patient #13 ... 73
6.1.8. Patient #16 ... 75
6.1.9. Patient #17 ... 77
6.1.10. Patient #19 ... 80
6.2. Evaluation of different bioinformatical approaches ... 82
6.2.1. Comparison of bioinformatical pipelines ... 82
6.2.2. Comparison of gene lists for variant filtering ... 84
7. Discussion ... 87
7.1. Sequencing methods ... 87
7.2. Variant filtering ... 89
7.3. Identification and interpretation of causal variants ... 91
7.4. Understanding of VEO-IBD ... 93
CONCLUSION ... 95
List of references ... 96
Online sources ... 122
List of supplementary files ... 123
Supplementary file 1: Gene lists for variant filtering ... 124
Supplementary file 2: Parameters of identified variants in patients ... 127
Supplementary file 3: Sequences of customized primers ... 128

List of abbreviations

ACMG	The American College of Medical Genetics and Genomics
ADA	Adenosine deaminase
BCR	B cell receptor
BET	Bromodomain and extra-terminal
CADD	Combined annotation dependent depletion
CD	Crohn's disease
CDG	Closest disease-causing genes
CGD	Chronic granulomatous disease
CID	Combined immunodeficiency
CNV	Copy number variations
CVID	Common variable immunodeficiency
EO-IBD	Early onset Inflammatory bowel disease
ESP	Exome sequencing project
ExAC	The Exome Aggregation Consortium
G6PC3	Glucose-6-phosphatase
gnomAD	Genome aggregation database
GWAS	Genome wide association study
HGMD	Human gene mutation database
IBD	Inflammatory bowel disease
IBD-U	IBD unclassified
IC	Indeterminate colitis
IGV	Integrated Genomics Viewer
IUIS	International Union of Immunological Societies
IVA	QIAGEN Ingenuity Variant Analysis
JAK	Janus kinases
LAD-1	Leukocyte adhesion deficiency type 1
MHC	Major histocompatibility complex
MS	Multiple sclerosis
NADPH	Nicotinamide adenine dinucleotide phosphate
NGS	Next generation sequencing
OMIM	Online mendelian inheritance in man
PeCan	Pediatric cancer
PI3K	Phosphatidylinositol-3-kinase
PID	Primary immunodeficiency
PolyPhen-2	Polymorphism phenotyping v2
PRR	Pattern recognition receptors
RA	Rheumatoid arthritis
ROS	Reactive oxygen species
SCID	Severe combined immunodeficiency
SIFT	Sorting intolerant from tolerant
SLE	Systemic lupus erythematous

SNP	Single nucleotide polymorphisms
STAT	Transducers and activators of transcription
T1D	Type 1 diabetes
TCR	T cell receptor
TLR	Toll-like receptors
Treg	T regulatory lymphocytes
UC	Ulcerative colitis
VEO-IBD	Very Early onset Inflammatory bowel disease
WES	Whole exome sequencing
WGS	Whole genome sequencing
yo	Years old

INTRODUCTION

Inflammatory bowel disease (IBD) is a compendious term for a group of chronic immunity-related disorders affecting the gastrointestinal tract. The best-known and described types of IBD are Crohn's disease (CD) and ulcerative colitis (UC). The average prevalence of the diseases in the Czech Republic is around 206/100 000 inhabitants and 236/100 000 in case of UC, respectively, and rising every year (Jarkovsky *et al*., 2017). The increase matches trends from other Western European countries and North America (Kaplan & Ng, 2017; Ng *et al*., 2017) making IBD an emerging health issue. Therefore, deeper investigation of the IBD pathogenesis is necessary for management and prevention of the disorder.

The pathophysiological cause of IBD is heterogenous and in many cases not very well understood. Some of known causal or contributing factors relate to environment and lifestyle - such as smoking, urban living or exposure of specific drugs. Those, however, are typical for adult cohort of IBD patients (Piovani *et al*., 2019). On the other hand, the pediatric patients suffering from IBD are more likely to have a severe genetic defect causing or contributing to the disease. In case of the youngest patients with a diagnosis called very early onset IBD (VEO-IBD), we can follow a pattern of functionally relevant mutations specifically in genes related to immune system regulation such as gene *IL10* encoding interleukin 10 or its receptor subunits *IL10RA* or *IL10RB* (Uhlig *et al*., 2014). The presence of validated innate imune system defects in IBD patients leads to a conclusion that VEO-IBD can be also regarded as a primary immunodeficiency.

The genetic analysis of VEO-IBD patients is nowadays commonly performed by whole exome sequencing aiming to identify the main causal gene (Crowley *et al*., 2018; Fang *et al*., 2019). However, the candidate IBD-causing genes are numerous and their numbers are rising quickly due to a wide availability of genomic analytical methods (Jostins *et al*., 2012; Liu *et al*., 2015, De Lange *et al*., 2017). Therefore, it is crucial to choose an adequate bioinformatical procedure for the sequencing data analysis to be able to identify relevant genes possibly causing the patient's pathological phenotype (Ji *et al*., 2019) followed by data validation via Sanger sequencing and functional assays such as, flow cytometry phenotypizations or Western blotting (Kelsen *et al*., 2015a;

Li *et al*., 2016). Consequently, the completed case studies not only help to adjust patients' treatment, but also contribute to a better understanding of IBD pathogenesis and immune system regulation.

THEORETICAL PART

1. Inflammatory bowel disease

The term “Inflammatory bowel diseases” (IBD) embodies several disorders characterized by complex immunity-impaired phenotype with relapsing intestinal inflammation. The general clinical picture of IBD includes common signs of inflammatory condition - fever, pain and failure to thrive. Moreover, typical symptoms suggesting intestinal involvement such as abdominal pain, watery and bloody diarrhea or even rectal bleeding can be observed (Podolsky, 2002). In addition to that, extraintestinal manifestations are as well common and include musculoskeletal (Falling *et al*., 2019), dermatological (Vavricka *et al*., 2011) or hepatobiliary involvement (Fraga *et al*., 2016). Moreover, rare manifestations such as ocular (Ottaviano *et al*., 2018), renal (Corica & Romano, 2015) or thyroid (Shizuma, 2016) malfunctions were reported in patients with IBD. Both the intestinal, as well as the extraintestinal manifestations can be severe or even life-threatening (Huber *et al*., 2010) and the IBD diagnosis should be taken into consideration as soon as possible to provide a better outcome for the patient (Lee *et al*., 2017a).

1.1. Classification of IBD

The differential diagnosis of IBD and then its subtypes is based on several main criteria. Firstly, it is patient’s and family anamnesis with emphasis on previous history of IBD, celiac disease, colorectal cancer, other autoinflammatory conditions and infections such as intestinal tuberculosis or *Clostridium difficile* infection (Bernstein *et al.*, 2016; Ma *et al*., 2016; Guerri *et al*., 2019). Secondly, it includes physical examination focusing on general assessment of thriving, abdominal and perianal region examination and extraintestinal manifestation signs (Bernstein *et al.*, 2016). Thirdly, the laboratory tests focus on standard hematological, immunological and biochemical as well as IBD-specific parameters. The specific ones include for example calprotectin levels in the stool (van Rheenen *et al*., 2010) or presence of autoantibodies (Seibold *et al.*, 1996; Sura *et al*., 2014). Lastly, imaging and endoscopic techniques take place during the diagnostics of IBD. Results of all above-mentioned diagnostic procedures

assemble a relevant amount of information for proper diagnostics of IBD and suggests its subtype (Bernstein *et al.*, 2016).

Two main subtypes of IBD, Crohn's disease and ulcerative colitis, share many resemblances, but also differ in several parameters. CD and UC share characteristic genetic background, similar overall clinical manifestations and the diagnostic and treatment approach. On the other hand, the IBD subtypes may be distinguished from each other by histological features, localization in the gut, cellular involvement or typical symptoms. One of the most informative features when differentiating between subtypes of IBD is the localization and appearance of the ulceration and the histological assessment of the intestine sample. The Crohn's disease manifests mainly in colon and ileum and the inflammation extends often deeper up to the submucosa. The lesions in CD are commonly disconnected throughout a vast segment of the intestine and accompanied by other complications such as intestinal fistulas and strictures. In contrast, the ulcerative colitis histopathological phenotype is almost exclusively located in colon. The ulcerations in UC are rather continuous and localized only superficially. Bearing in mind clinical, endoscopic and histological features, we can distinguish the IBD subtype in majority of the cases (Podolsky, 2002; Bernstein *et al.*, 2016).

What's more, apart from classical CD and UC, another inflammatory bowel disease subtype called IBD unclassified (IBD-U) or indeterminate colitis (IC), was characterized. Both IBD-U and IC describe the state when we cannot decide in favor of CD or UC due to overlapping phenotype. According to the Montreal World Congress of Gastroenterology in 2005, the terminology of indeterminate colitis should be used only in case of patient who underwent colectomy. The patient should be diagnosed as IBD-U in any other situation when it cannot be clearly distinguished between classical CD and UC (Silverberg *et al*., 2005). Despite being significantly less abundant (Prenzel & Uhlig, 2009), IC is not rare among pediatric patients. Carvalho *et al*. (2006) reports nearly 30% of examined children with undetermined IBD diagnosis. However, approximately one third them has been later reclassified as one of the classical IBD subtypes. This leads to a question, whether IC should be considered a transitive diagnosis based on UC and CD non-exclusion or rather become a separate disorder. The idea of distinguishing a new official IBD diagnostic subtype has been already publicly discussed, however officially it remains unsolved mainly due to a lack of exclusive markers of IC (Guindi & Riddell, 2004; Telakis & Tsironi, 2008; Yu *et al*., 2017).

When evaluating an IBD case, it is important to also consider the age of onset of the disorder. According to the age of onset, the Montreal Classification of IBD is dividing the patients into three groups: younger than 17, 17-40 years old (yo) and older than 40 years of age, as displayed in Tab. 1 (Silverberg *et al*., 2005).

CATEGORY		AGE OF ONSET	CLASSIFICATION
Adult	A3	> 40 yo	Montreal Classification
	A2	17-39 yo	Montreal Classification
Pediatric	A1 A1b	10-16 yo	Montreal Classification Paris Classification
- Early Onset	A1a	<10 yo	Paris Classification
- Very Early Onset		< 6 yo	Paris Classification
- Infantile		< 2 yo	Paris Classification
- Neonatal		< 28 days old	Paris Classification

Tab.1: Classification of IBD based on the age of onset. The categories were set up regarding the genotype and phenotype differences among the age subgroups resulting in creation of the Montreal (Silverberg *et al*., 2005) and Paris (Levine *et al.*, 2011) Classification of IBD

Focusing on the pediatric patients, they statistically represent up to 25% of all IBDs cases (Ruel *et al*., 2013) which makes them an important subgroup. However, the pediatric cases significantly differ between one another. And what's more, the age of 10 years appears to be an important timepoint concerning the distinctive endoscopic phenotype of the IBD (Levine *et al*., 2007). The results of Levine's *et al*. (2007) research concerning the distinctions of age-based pediatric IBD subgroups led to reevaluation of the Montreal Classification (Silverberg *et al*., 2005) for pediatric IBD. Consequently, previously set group A1 (patients < 17 yo) was subdivided into two categories by the Paris Classification of IBD (Levine *et al*., 2011) which newly distinguishes a category A1b including kids from ten to sixteen years old and a group A1a for kids younger than ten years of age (Tab. 1).

Reevaluating the Montreal Classification (Silverberg *et al*., 2005) suitability for pediatric patients, the Paris Classification (Levine *et al.*, 2011) not only further discriminates pediatric IBD into group A1a and A1b (Tab. 1), but also suggests four additional categories within A1a based on more thorough discrepancies in phenotypic features. Namely the categories are early onset (EO-), very early onset (VEO-), infantile and neonatal IBD. Particularly, VEO-IBD is specific in several criteria including distinct manifestation and causes involved in its development.

1.2. Causes of IBD

Nowadays we know that IBD is a multifactorial disease caused by a vast number of possible contributing factors specifically combined in pathogenesis of every patient's case. However, in the past scientists used to consider other causes than the multifactorial mode. The first assumptions included monogenic causal mutations similarly to in that time already known *CYBB* mutation causing X-linked chronic granulomatous disease (Segal *et al*., 1983). This mendelian inheritance mode was investigated by Orholm *et al*. (1993) with results suggesting that CD is caused by a mutation in a major recessive gene and UC by an impaired major gene with dominant character. Shortly after, the theory was reevaluated and other contributing factors apart from the classical mendelian genetics were brought into the consideration (Yang *et al*., 1993). To prove and further understand this multifactorial concept, considerable number of studies has been published describing environmental (Piovani *et al*., 2019), microbiome (Lloyd-Price *et al*., 2019), genetic (Uhlig *et al*., 2014) and epigenetic (Karatzas *et al*., 2014), as well as immunity-related (Rubin *et al*., 2019) involvement in IBD pathogenesis (Fig. 1).

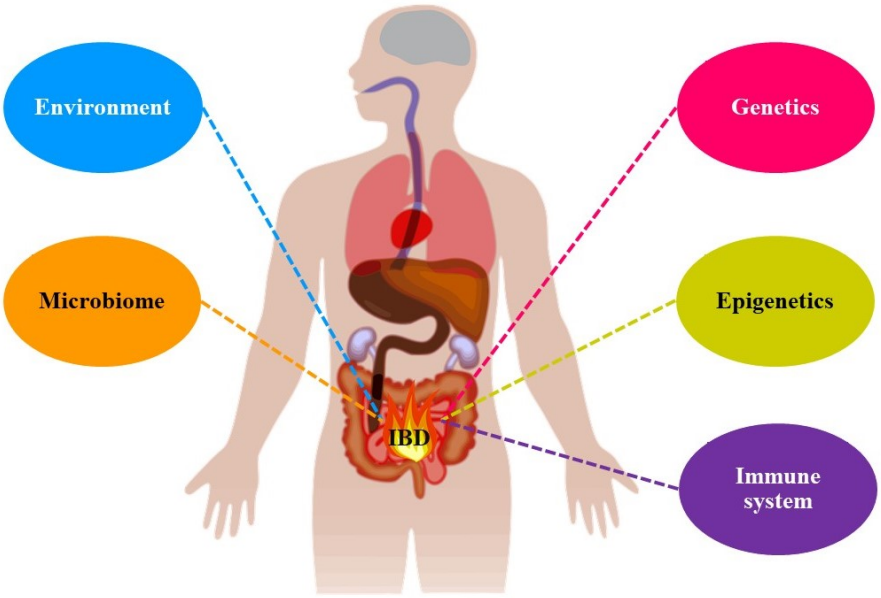


Fig. 1: Factors contributing to development of IBD. IBD is a complex disorder cause by alterations in patient's genome and epigenome structure, his immunological phenotype, as well as his environmental surroundings and lifestyle in addition to his intestinal microbial composition (Figure inspired by Yeshi *et al.*, 2020).

Concerning the environment, several risk factors for IBD development have been identified so far including personal lifestyle habits and general environmental conditions.. Namely the lifestyle-connected contributing factors are smoking (Persson *et al*., 1990), urban-based living (Benchimol *et al*., 2017) and unhealthy lifestyle including diet (such as consumption of additives like carboxymethylcellulose and polysorbate-80) and obesity (Chassaing *et al*., 2015). Those described causes put together with increased incidence of IBD in western developed countries call for classification of IBD as civilization disease (Kaplan & Ng, 2017; Ng *et al*., 2017).

The environmental involvement in IBD pathogenesis includes also a specific drug usage and IBD-unrelated health issues. To evaluate the risk of IBD development in connection with taken medication, meta-analyses were performed that found an increased risk of IBD in women taking oral contraceptives (Cornish *et al*., 2008) and in patients who underwent the treatment with antibiotics, particularly with penicillins or cephalosporins, metronidazole and fluoroquinolones (Theochari *et al*., 2018). Simultaneously, surgical interventions at secondary lymphoid organs like tonsillectomy

(Sun *et al*., 2016) or appendectomy (Kaplan *et al*., 2008) and other health issues such as vitamin D deficiency (Del Pinto *et al*., 2015) were associated with IBD risk as well. Additionally, some infections, for example non-*Helicobacter pylori* enterohepatic *Helicobacter*, have detrimental effect on IBD development (Piovani *et al*. 2019). Interestingly, other infections show rather protective character against IBD development like in case of *Helicobacter pylori* (Castaño-Rodríguez *et al*., 2015). All risk factors mentioned above were recently reevaluated by Piovani *et al*. (2019) and proved to correlate with IBD risk and progression However, so far, we cannot certainly state whether they are the cause, the result or only the bias of inflammatory disorder in the gut.

One of the newest investigated contributing factors of IBD is microbiota and its interaction with the human host. The presence of specific intestinal microflora has been reported, for example an absence of *Bacteroidetes* by de Meij *et al*. (2018) or *Faecalibacterium prausnitzii* by Sokol *et al*. (2009). Also, a lower heterogeneity of microbiota species in IBD patients in comparison to healthy controls has been observed. Comparing the microbiome of UC and CD patients, significant differences can be noticed, supporting a distinct pathogenesis of the subtypes (Sun *et al*., 2019). Summed up, the microbiota analyses of IBD patients show both the differences between healthy controls and patients, as well as the differences within the patients' subtypes. This suggests a possible new tool for diagnostics, assessment of disease progression and potential new targets for treatment (Aschard *et al*., 2019).

Additionally, several studies focused on analyzing the relationships between microbiota and other contributing factors for IBD. When correlated with genetics, microbiota may partially influence the gene expression of some IBD-associated genes like *NOD2* and, thus, contribute to IBD development (Aschard *et al*., 2019). Moreover, in combination with different diets, fecal microbiota significantly influences the progression of IBD in susceptible individuals (Llewellyn *et al*., 2018). All those findings point out the complexity of the disorder and suggests presence of a number of yet unknown functional interactions concerning intestinal microbiome.

Bearing in mind the IBD genetics, several meta-analyses focusing on genes specific either for CD and UC or IBD in general have been performed and a list of genes important for IBD development and progression was assembled. Starting with *NOD2* in

2001 (Hugot *et al*., 2001), the genes associated with IBD have been piling up and up-to-date over 240 susceptibility loci have been identified (Barret *et al*., 2008; Imielinski *et al*., 2009; Franke *et al*., 2010; Anderson *et al*., 2011; Jostins *et al*., 2012; Liu *et al*., 2015; De Lange *et al*., 2017). Many of these loci are associated with canonical immunopathways, suggesting a significant role of immune system defects in IBD development and progression. Surprisingly, only a limited number of them is involved in three major pathways which are targeted by IBD therapeutics (De Lange *et al*., 2017). Examples of genes involved in those therapeutically-targeted signalings are TNFα (Al-Meghaiseeb *et al*.; 2016) in TNF-associated pathway, IL12 and IL23 sharing the same subunit p40 (Oppmann *et al*., 2000) in their signaling, or genes for integrins in the integrin-associated pathway (Sun, 2019).

Some pediatric IBD cases are considered to be monogenic disorders. Over 65 causal genes have been reported (Uhlig *et al*., 2014; Pazmandi *et al*., 2019) so far, including genes essential for immune activation, for example *PLCG2* described in a patient by Neves *et al*. (2018), or immune regulation, like *FOXP3* (Okou *et al*., 2014). These findings suggest that IBD immunogenetics is a complex process and further research is necessary in order to understand mechanisms and causations, and to design fitting treatment methods.

Closely connected with IBD genetics are IBD-associated epigenetic changes. One of the typical epigenetic mechanisms is DNA methylation, usually with repressive effect (McGhee & Ginder, 1979; Harris *et al*., 2018). DNA methylation profile of IBD patients was examined by Katarzas *et al*. (2014) finding a distinct methylation pattern in UC patients. Patients with UC had hypermethylated genes *CXCL14*, *CXCL5*, *GATA3*, *IL17C*, and *IL4R* that are all closely connected with immune system regulation and function. The results of epigenetic analyses could serve as a possible new marker for IBD patients' diagnostics and prognosis assessment (Boyd *et al*., 2018). Moreover, in a novel epigenetic-based treatment option for IBD may be introduced in future. Epigenetic-based treatment trials in cases of other diseases are taking place, namely using Bromodomain and Extra-Terminal motif (BET) inhibitors (Doroshow *et al*., 2017). Mouse model experiments show an amelioration of colitis with sodium butyrate via impaired DNA binding of NF-κB and histone H3 acetylation (Lee *et al*., 2017b) suggesting another possibility for epigenetic management of IBD.

Lastly, immune system plays a crucial role in IBD development. The immune relation is reflected even in the medical classification, where the disease is often put in a group of immune-mediated disorders next to rheumatoid arthritis (RA), systemic lupus erythematous (SLE), multiple sclerosis (MS), type 1 diabetes (T1D) or celiac disease. The immunity-related connection among the diseases is probably based on sharing some susceptible genetic loci. Some of the loci shared among many immunity-mediated disorders including IBD contain genes encoding essential regulating and activating interleukins and their receptors, such as *IL10* or *IL2RA* (Ramos *et al*., 2011).

Apart from impaired interleukin signaling, several other general mechanisms of immune dysregulation leading to IBD development were described. Namely it is impaired signaling and functions of innate immunocytes like phagocytic macrophages (Jones *et al*., 2018) or adaptive immune cells represented by T regulatory lymphocytes (Treg) (Sun *et al*., 2017). Additionally, a disrupted epithelial barrier has been proven to be an underlying condition in IBD patients. (Vivinus-Nébot *et al*., 2014). Overall, it is evident that the immune system and its homeostasis plays a crucial role in IBD development and progression which is reflected in expanding number of immunity-based treatment options as well as in a significant number of recently published research studies on immune aspects of IBD.

1.3. Very Early Onset IBD

The combination of contributing and causal factors like environment or genetics is specific for each patient. However, we can find a pattern of a certain genetic and environmental involvement when comparing pediatric and adult onset IBD. Particularly, very early onset IBD cohort represented by IBD patients younger than 6 years of age is significantly specific due to different ratio of genetic and environmental involvement to pathogenesis (Fig. 2) in comparison to pediatric later onset and adult onset IBD (Uhlig *et al*., 2014).

Concerning the genetic aspect, VEO-IBD cases have a higher prevalence of severe genetic defects. Having analyzed sequencing data of VEO-IBD patients, more than 65 major causal genes have been reported so far (Uhlig *et al*., 2014; Pazmandi *et al*., 2019). Mutations in those genes were found deleterious and functionally correlated with the

phenotype of VEO-IBD patients. The products of the genes are often related to immune system and their malfunction or malproduction leads to systemic immune dysregulation. What's more, the gene defects involved in VEO-IBD development present with high, sometimes even complete penetrance of the mutation, implicating that VEO-IBD can be considered as a manifestation of monogenic primary immunodeficiency (Uhlig *et al*., 2014; Fang *et al*., 2019; Pazmandi *et al*., 2019). In contrast, no study has reported so far the extend of involvement of the environmental factors in VEO-IBD.

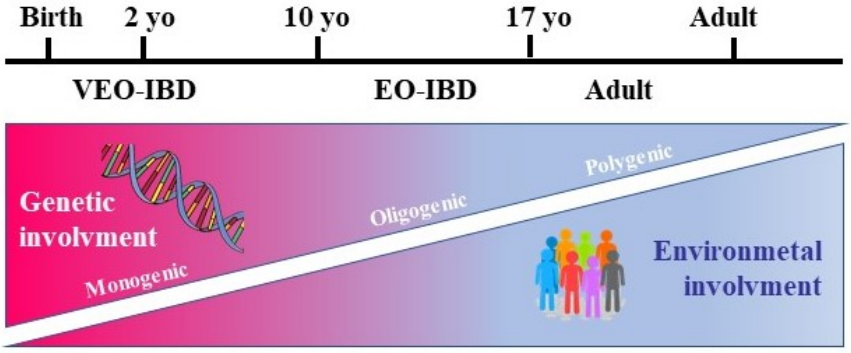


Fig. 2: Ratio of genetic and environmental involvement in pediatric and adult IBD development. The pediatric cases, specifically very early onset IBD, are characterized by a significantly more frequent presence of major genetic defects causing or contributing to the disease, whereas adult IBD cases have a more frequent contribution of environmental and non-genetic factors (Figure inspired by Okou *et al.*, 2014).

As opposed to that, EO-IBD and adult patients usually have oligogenic or polygenic mode of contributing mutations, typically involving genes with lower penetrance and less essential function for immune system (Uhlig & Schwerd, 2016). In addition to genetics, the impact of environmental factors is extensive in older patients (Kaser *et al*., 2010). Specifically, smoking, urban-based, unhealthy lifestyle, oral contraceptives and antibiotics usage, surgeries, infections or vitamin deficiencies are proven to contribute significantly (Piovani *et al*., 2019).

In addition to the ratio of genetic and environmental involvement, the pediatric and adult onset IBD differ also both in clinical manifestation in the intestine and in extraintestinal symptoms. Namely, growth-retardation has been reported in kids with

IBD, particularly in those diagnosed very early in life (Mamula *et al*., 2002). Additionally, the gastrointestinal symptoms of early onset pediatric IBD are more severe and their treatment reactions are more troublesome and less effective than in case of later onset IBD (Ruemmele *et al*., 2006) making pediatric IBD a specific diagnostic challenge.

Summarized, the VEO-IBD is presumed to have a severe genetic load particularly in genes involved in immune system. On the other hand, the involvement of environmental factors in VEO-IBD pathogenesis is minor (Tab. 2 and Fig. 2). Thus, in contrast to classical multifactorial adult IBD, pediatric very early onset IBD might be considered a monogenic primary immunodeficiency (PID). The immunogenetic and immunodeficiency-related characteristics should be taken into account when suspected VEO-IBD patient is evaluated.

	VEO-IBD	**ADULT AND LATER ONSET PEDIATRIC IBD**
Age of onset	< 6 yo (Levine at el., 2011)	> 6 yo Peak in adulthood 20 - 40 yo
Prevalence	0,5:100 000 (Abramson *et al*., 2010a)	200:100 000 (Jarkovsky *et al*., 2017)
Types	CD, UC, IBDU (Abramson *et al*., 2010a) Often accompanied by other extraintestinal symptoms (Uhlig *et al*., 2014)	CD, UC + IBDU rare (Jarkovsky *et al*., 2017; Prenzel & Uhlig, 2009)
Genetics	Monogenic (Uhlig & Schwerd, 2016) Over 65 reported genetic defects (Uhlig *et al*., 2014; Pazmandi *et al*., 2019)	Oligogenic and polygenic (Uhlig & Schwerd, 2016) Over 240 susceptible loci (Barret *et al*., 2008; Imielinski *et al*., 2009; Franke *et al*., 2010; Anderson *et al*., 2011; Jostins *et al*., 2012; Liu *et al*., 2015; De Lange *et al*., 2017)
Environment	Not described yet	Smoking, urban-based, unhealthy lifestyle, oral contraceptives and antibiotics usage, appendectomy, tonsillectomy, non-*Helicobacter pylori* enterohepatic *Helicobacter* infection, vitamin D deficiency (Piovani *et al*., 2019)
Treatment	Poor reaction to conventional treatment (Ruemmele *et al*., 2006)	Good reaction

Tab. 2: Differences between VEO-IBD and adult and later-onset pediatric IBD.

When suspecting VEO-IBD in a patient, several factors based on characteristic manifestations in the patient should be looked for. Similarly to classical adult onset IBD, VEO-IBD diagnostics includes evaluation of patient's history and the clinical phenotype. Concerning the phenotype, Uhlig *et al*. (2014) proposes a memory aid "YOUNG AGE MATTERS MOST" summarizing the typical VEO-IBD symptoms. **Young age** of onset is the main indicator. In addition, family situation concerning **M**ultiple family members and consanguinity should be taken in consideration. Also, personal anamnesis including **A**utoimmunity, failure to **T**hrive and failure of conventional **T**reatment, accompanied by **E**ndocrine issues, **R**ecurrent infections and **S**evere perianal disease might support this diagnosis. Lastly, patients with VEO-IBD might also suffer from disorders such as **M**acrophage activation syndrome and hemophagocytic lymphohistiocytosis, intestinal **O**bstruction and atresia, **S**kin lesions, dental and hair abnormalities and might have present **T**umors. The individual manifestation of these symptoms points to the dysregulated immunity. and, thus, suggests what laboratory tests and analyses should be performed in order to identify and validate causal genetic variant.

2. Molecular mechanisms of VEO-IBD

Considering VEO-IBD a monogenic disease, it is crucial to understand the underlying molecular and signaling mechanism of the dysfunction in order to understand the pathology and, consequently, be able to adjust the patient's treatment. Up-to-date, about 240 susceptible loci (De Lange *et al*., 2017) for IBD development were identified and over 65 genetic defects were reported as monogenic causes for IBD (Uhlig *et al*., 2014; Pazmandi *et al*., 2019). Many of those genes and loci relate to immunity by several possible mechanisms, each of them involving different way of immune system regulation. Immune dysregulation leading to monogenic VEO-IBD (Tab. 5) may occur in case of:

- epithelial barrier defects,
- phagocyte defects,
- lymphocyte defects,
- systemic inflammatory disorders,
- complement deficiencies.

2.1. Epithelial barrier defects

The epithelial barrier in the gut serves as a defense mechanism against microorganisms disallowing them to enter the body. However, sometimes the intestinal barrier disrupts and its permeability increases, making it possible for the pathogens, and also for otherwise symbiotic bacteria cross the barrier and start a local inflammatory response in the gut possibly resulting in IBD development. This phenomenon sometimes called also "leaky gut" occurs when the epithelial barrier is disrupted by defects in its structure, defects in cellular functions including pathological interaction between host and microbiota and defects caused by death of epithelial barrier cells due to innate mutations in genes significant for intestinal homeostasis (Pazmandi *et al*., 2019).

Several genes and its mutations have been identified so far leading to innate epithelial barrier disruptions. An example of gene involved in epithelial barrier structure defect is *TTC7A*, which in mutated state has been found and considered causal for VEO-IBD in multiple cases (Samuels *et al*., 2013; Bigorgne *et al*., 2014). The mechanism behind the pathogenesis is presumably a change in intestinal cellular polarity, growth

and differentiation stemming from upregulated activity of Rho kinase and resulting in disturbed cellular and tissue structure (Bigorgne *et al*. 2014). Similarly, the intestinal tissue structure can be disrupted by *FERMT1* and *COL7A1* mutations due to impaired integrin signaling leading to dystrophy of epithelia including the intestine (Freeman *et al*., 2008; Ashton *et al*., 2016).

Also, cell adhesion in the gut can be impaired by mutation of *EPCAM* leading to atrophy of the villi and consequent enteropathy (Al-Mayouf *et al*., 2009). Additionally, peptidase D/prolidase involved in collagen metabolisms can be dysregulated by mutation in its gene *PEPD,* leading to decreased healing, ulceration and in some cases also very early onset IBD (Kuloglu *et al*., 2015; Rizvi *et al*., 2019). The several ways of intestinal barrier disruption mentioned above display the structural complexity of the barrier itself.

Moreover, the barrier can be also disrupted by dysregulation of antimicrobial-immunity mechanisms. Several paralogs of *CYBB* (also known as *NOX2*, a subunit of NADPH oxidase involved in phagocytosis, described in section *2.2 Phagocyte defects*) were found in other cell types causing matching syndromic symptoms to *NOX2* defects. For example, *NOX1* and *DUOX2* inactivation mutations, which significantly reduced ROS production, has been described in patients with VEO-IBD, suggesting its association with the disease (Hayes *et al*., 2015). Schwerd's *et al*. (2018) study of loss-of-function *NOX1* variants, however, suggests that they rather than being a cause only contribute to IBD development via decreased functioning of ROS.

What's more, immunity-related malfunction of certain intestinal cells caused by mutation in genes like *GUCY2C* or *SLC9A3* can also lead to VEO-IBD. *GUYC2C* is a gene encoding transmembrane metabotropic receptor guanylate cyclase which has functional impact on chloride and water regulation. Fiskerstrand *et al*. (2012), who presented a case of IBD patients with heterozygous gain-of-function variant in *GUCY2C*, suggest that extensive water secretion in the gut might be the cause of chronic diarrhea in this case. Similarly, mutated gene for transmembrane sodium/proton antiporter *SLC9A3* causes congenital secretory sodium diarrhea with prenatal (Holmberg & Perheentupa, 1985) or very early onset (Janecke *et al*., 2015).

Another important part of the barrier are prostaglandins protecting the gastrointestinal mucosa. Loss-of-function mutation in prostaglandin transporter gene *SLCO2A1* can cause chronic enteropathy with ulcerations (Umeno *et al*., 2015). Also, deletion in *PLA2G4A* (Brooke *et al*., 2014) encoding cytosolic phospholipase A2

involved in eicosanoids synthesis can lead to early onset IBD development due to decreased levels of eicosanoids and consequently unprotected gastrointestinal mucosa (Faioni *et al*., 2014).

Closely connected to intestinal barrier defects is also impaired host-microbiota interaction. The microbiota recognition and start-up of a following immune reaction is mediated by pattern recognition receptors (PRR) including Toll-Like Receptors (TLR). For example, TLR4 is activated by LPS and, thus, its function is strongly dependent on LPS presence and degradation. The enzyme responsible for LPS detoxification is alkaline phosphatase encoded by gene *ALPI* in the intestine. The loss-of-function variant of *ALPI* was reported in patients suffering from very early onset IBD presumably due to chronic pro-inflammatory response in the gut (Parlato *et al*., 2018).

Another important family of PRRs is NOD-like receptors. Having been the first genetic defect connected with IBD (Hugot *et al*., 2001), mutations in NOD2 have been thoroughly investigated and reported in a considerable number of cases. The pathophysiological mechanism includes improper clearance of bacteria in the gut due to impaired signaling via NOD2 and downstream NFκB pathway leading to chronic inflammation resulting in IBD phenotype (Sidiq *et al*., 2016). Apart from being a classical multifactorial IBD-associated mutation (Hugot *et al*., 2001), some variants in NOD2 has been also described in patients with monogenic early onset IBD (Girardelli *et al*., 2018). Also, other NOD2-related molecules can be involved in the pathological processes of VEO-IBD. For example, deleterious variant in *TRIM22* (Li *et al*., 2016) can impair NOD2 signaling and consequently lead to IBD with very early onset. The NOD2 signaling pathway is, therefore, an excellent example of IBD-associated pathway defects with diverse manifestation from early onset (Girardelli *et al*., 2018) to adult onset IBD (Hugot *et al*., 2001).

2.2. Phagocyte defects

Professional phagocytes such as neutrophil granulocytes or macrophages are important cells involved in innate immunity response. Phagocytes in the context of IBD pathogenesis not only play a crucial role in keeping the intestinal homeostasis and surveilling for the potential pathogens crossing the intestinal barrier, but as well contribute to the chronic inflammatory response by production of specific cytokines and other soluble or membrane molecules. The impaired function of any kind, as well as

decreased numbers of the cells might consequently lead to immunodeficiency or autoinflammatory condition with IBD symptoms (Pazmandi *et al*., 2019).

One of the possible functional mechanisms behind IBD stemming from phagocyte defects is a chronic granulomatous disease (CGD). CGD is characterized by formation of granulomatous lesions in the location of the inflammation as a result of ineffective pathogen killing. The phagocyted pathogens are normally destroyed by enzymatic reaction called oxidative burst sustained by nicotinamide adenine dinucleotide phosphate (NADPH) oxidase complex via production of reactive oxygen species (ROS). This enzymatic complex and consequently anti-microbial immune response can, however, be defective due to mutation in gene encoding any of the five NADPH oxidase subunits.

NADPH oxidase consists of five subunit and mutations in all subunit-encoding genes are associated with defects of intestinal immunity. The NADPH oxidase subunits are, firstly, the subunit of cytochrome b-245 alpha *CYBA* (Teimourian *et al*., 2008) and beta *CYBB* (Segal *et al*., 1983). Secondly, another three NADPH oxidase subunits are from a family of neutrophilic cytosolic factors and they are encoded by genes *NCF1* (van de Vosse *et al*., 2009), *NCF2* (Muise *et al*., 2012) and *NCF4* (Matute *et al*., 2009). Despite the facts that *CYBB* deficiency is the only one X-linked disorder, it is the most common cause of chronic granulomatosis (Roos *et al*., 2010).

Similarly to NADPH oxidase subunits, mutations in genes for small GTPases involved in the oxidative burst cascade *RAC1* and *RAC2* influence phagocytic functions of neutrophils. Being minimally represented in oxidase complex in neutrophils, *RAC1* has selective effects on neutrophil functions (Heyworth *et al*., 1994). *Rac1* deficient neutrophils in mouse model have impaired chemoattraction and migration, but their oxidative burst is not defective (Glogauer *et al*., 2003). In contrast with *RAC1* selective effects, *RAC2* deficiency has a negative impact on NADPH oxidase activation and causes severe neutrophil immunodeficiency syndrome (Ambruso *et al*., 2000).

In addition to impaired oxidative burst, autoinflammatory phagocyte-related disorder can be also a result of recurrent activation of inflammasome. For example, due to mutations of *NLRC4* the macrophages become chronically activated resulting in very early onset fever, as well as inflammatory manifestation in the intestine and other organs (Canna *et al*., 2014). Similarly, mutation and dysregulation of *CARD8* and *NLRP3* is also associated with IBD, however, not yet in a monogenic way (Mao *et al*., 2018).

Moreover, defective adhesion of the phagocytes to the vessel walls leads to severe immunodefects, too. This pathological mechanism is characteristic for a disorder called leukocyte adhesion deficiency type 1 (LAD-1). LAD-1 syndrome caused by *ITGB2* mutation (also known as *CD18*) negatively influences phagocytes in their adherence and migration to the tissues and consequently in their microbial clearance. This ineffective infection management can lead in case of long-lasting intestinal inflammation, as well, to IBD (Uzel *et al*., 2001).

Lastly, the phagocyte defects can also originate from neutropenia. Severe congenital neutropenia accompanied by other symptoms was reported in patients with *G6PC3* mutations. Glucose-6-phosphataseis an important enzyme involved in glucose metabolism and its deficiency leads to metabolic switch. Mistry *et al*. (2017) suggest that the metabolic changes triggers higher production of inflammatory cytokines leading to autoinflammatory conditions including IBD. In connection of glycogenesis, the deficiency of glucose-6-phosphate translocase encoded by *SLC37A4* is also connected with IBD (Volz *et al*., 2015) highlighting the importance of energetic metabolism for the immune system and its well-being.

2.3. Lymphocyte defects

Genetically-based adaptive immunity defects stemming from lymphocyte deficiencies and functional dysregulation, as well as impaired antibody production, lead to the development of primary immunodeficiencies with possible intestinal involvement. The phenotype, genetics and mechanism of the immunodeficiency can be diverse. Phenotypically, VEO-IBD can present as Common Variable Immunodeficiency (CVID), but also as Severe Combined Immunodeficiency (SCID) or Combined Immunodeficiency (CID) with syndromic features (Bousfiha *et al*., 2018). Underlying genetic defect leading to malfunction of the immune system may affect B lymphocytes and their antibody production, as well as effector T lymphocytes. In addition, regulatory T cells have been thoroughly studied in connection to PID and VEO-IBD development recently, too. Lymphocyte defects are, however, often very versatile and complex disorders and in many cases the IBD manifestation are just one of many symptoms of the disorder (Uhlig *et al*., 2014; Pazmandi *et al*., 2019).

2.3.1. B lymphocyte and antibody defects

Concerning CVID, B cell functions like their activation, proliferation, class switch recombination and antibody production are affected. The causes of those pathogenic mechanisms are mainly polygenic or multifactorial (de Valles-Ibáñez *et al*., 2018), however, some disorders of monogenic origin are known, too. Among the few identified major impact genes related to CVID with intestinal manifestation are *CD40LG* (Levy *et al*., 1997; Wang *et al*., 2014) and *AICDA* (Quartier *et al*., 2004). Mutations of these genes cause hyper IgM syndrome due to impaired B cell activation and antibody class switch. Similarly, hyper IgM syndrome can be also caused by impaired PI3K/Akt/mTOR signaling. For example, gain-of-function mutation of *PIK3CD* (Elgizouli *et al*., 2016; Farachi *et al*., 2018) is involved in immunopathology via impaired phosphatidylinositol-3-kinase (PI3K) mediated BCR signaling. Similar symptoms like in case of *PIK3CD* mutation were also observed in patients with *PTEN* mutation. *PTEN* deleterious variants, however, are of a loss-of-function character as PTEN regulates the PI3K/Akt pathway negatively (Driessen *et al*., 2016).

What's more, the abundance of certain class of antibodies can also occur in case of IgD. Even though IgD levels in plasma are low compared to other immunoglobulin types, hyper IgD syndrome features can be severe. The syndrome is caused by loss-of-function mutation in gene *MVK* encoding melvalonate kinase which is involved in cholesterol metabolism. The typical symptoms include febrile episodes with diarrhea or rash with early onset (Bader-Meunier *et al*., 2011). The good treatment results of melvalonate kinase deficient patients with IL1-targeting drugs implicate IL1 involvement in the disorder pathogenesis (Galeotti *et al*., 2012). Similarly, hyper IgE syndrome caused by *DOCK8* mutation can also manifest with diarrhea and colitis in pediatric patients (Sanal *et al*., 2012).

In addition to hyper IgM and hyper IgD syndromes, agammaglobulinemia is another example of CVID with IBD symptoms. Patients with agammaglobulinemia were reported to have a mutated *BTK* gene (Maekawa *et al*., 2010) resulting in impaired maturation of B cells and, thus, lacking antibody production. Also, *PIK3R1*, which encodes regulatory subunit of PI3K, is associated with agammaglobulinemia, Interestingly, pathogenic mutation of *PIK3R1* has been described in case of homozygous loss-of-function mutation (Conley *et al*., 2012) as well as in case of heterozygous gain-of-function variant (Lucas *et al*., 2014). Those contradictory findings

imply complexity of the PI3K signaling in B cell function and consequently in PID and VEO-IBD development.

Moreover, hypogammaglobulinemia is also a known form of CVID with intestinal manifestation reported in patients with mutation in *TNFRSF13B* (Salzer *et al*., 2009), *CD19* (Kanegane *et al*., 2007) and *IRF2BP2* (Keller *et al*., 2016). Additionally, B cell immunodeficiency due to *TRNT1* (Frans *et al*., 2017), *LRBA* (Alangari *et al*., 2012), *IL21* (Salzer *et al*., 2014), *TTC37* (Hartley *et al*., 2010) and *SKIV2L* (Fabre *et al*., 2012) mutations has been described in patients with IBD. Also, *ICOS* (Grimbacher *et al*., 2003) and its ligand *ICOSLG* (Roussel *et al*., 2018) mutations can lead to immunodefects characterized by decreased generation of antibodies and memory B cells. Lack of antibodies and memory B cells stems from non-activation of B cells due to missing costimulation. The costimulation is impaired either by defects of ICOS on helper T cells or ICOSLG on B cells. The example of ICOS and ICOSLG deficiency displays the need of proper cooperation of all immunocytes and other involved cell types in order to avoid malfunctions of the immune system.

2.3.2. Effector T lymphocyte defects

The development and functions of effector T lymphocytes may be impaired by genetic defect leading to development of severe immunodeficiency with possible IBD manifestation. In addition to T cells, B cells and NK cells might be affected at the same time resulting in combined immunodeficiency with several functional defects. Depending on the phenotype of the disorder, we can distinguish T+/- B+/- NK+/- SCID based on the presence or absence of the cellular type (Bousfiha *et al*., 2018). Because of the severe combined phenotype, all types of SCID are typical of very early onset within first months of life and very serious or even life-threatening phenotype. The danger stems mainly from susceptibility to infections which might be lethal for immunocompromised patients. Thus, it is important to identify the causal genetic variant of the SCID as quickly as possible in order to provide treatment for the patient in time (Fischer *et al*., 2015).

One of the genes involved in B and T lymphocytes maturation, mutation of which can cause SCID, is *RAG1*. *RAG1* is recombination-activating gene important for V(D)J recombination and its deleterious mutations lead to T- B- NK+ SCID with protracted diarrhea. Similarly, symptoms and mechanism match SCID caused by

mutation of *RAG2* (Corneo *et al*., 2001). Also, mutation in another gene involved in V(D)J recombination, *DCLRE1C* (also known as *ARTEMIS*), can lead to very early onset IBD. It results in ARTEMIS deficiency presented as T- B- NK+ SCID (Felgentreff *et al*., 2015). Moreover, DNA-dependent protein kinase catalytic subunit encoded by *PRKDC* and interacting with ARTEMIS during the recombination might be involved in pathogenesis of SCID with IBD, too. This implication is based on a report of an uninvestigated deceased twin, whose brother had validated mutation in *PRKDC* and suffered from immunodeficiency with chronic diarrhea (Mathieu *et al*., 2015). Lastly, LIG4 is both involved in V(D)J recombination, but also plays a role in B cells development. Mutation in *LIG4* can lead to SCID with IBD or B cell-development-related LIG4 syndrome (van der Burg *et al*., 2006). The spectrum of severe defects due to mutated parts of the V(D)J machinery shows its importance for the adaptive immunity functioning.

Similarly, low counts of lymphocytes, especially helper T cells, clinically presenting with immunodeficiency and diarrhea might be present in patients suffering from *FOXN1* mutation (Chou *et al*., 2014). FOXN1 is a master regulator of thymic epithelial cells and, thus, the thymic selection of T cells is impaired by this mutation (Žuklys *et al*., 2016).

The lymphocyte maturation and proliferation are also dependent on metabolism of nucleotides related to enzyme adenosine deaminase (ADA). ADA deficiency causes a build-up of nucleotides, especially deoxyadenosine, which are toxic for cells with high proliferation rate such as T and B lymphocytes. Therefore, T- B- NK- SCID with possible intestinal manifestation can occur due to mutation of *ADA*. As these mutated variants impair the immune system significantly, the onset of the immunodeficiency is commonly within the first months of life (Felgentreff *et al*., 2011). Moreover, PNP deficiency manifests by a similar mechanism and with resembling symptoms (Shanon *et al*., 1988).

Additionally, effector function of T cells is firstly based on activation via MHC-TCR signaling. TCR is associated with protein complex CD3, which consists of several subunits including CD3γ. CD3γ is encoded by *CD3G* gene, mutation of which has been found to cause T+/low B+NK+ SCID with autoimmune enteropathy (Recio *et al*., 2007). Other subunit delta and epsilon of CD3 encoded by *CD3D* (Dadi *et al*., 2003) and *CD3E* (de Saint Basile *et al*., 2004) can also in truncated state cause SCID with very early onset IBD.

What's more, other downstream molecules associated with TCR signaling such as p56 (LCK) encoded by *LCK* (Goldman *et al*., 1998) are known to lead to SCID with VEO-IBD in case of deleterious mutations. Also, intracellular domain of TCR interacts with ZAP-70 during the activation. Therefore, when mutated, *ZAP-70* can cause severe immunodeficiency. Namely, autoimmune syndrome with chronic intestinal inflammation and recurrent infections has been reported in patient with ZAP-70 deficiency (Parry *et al*., 1996). Interestingly, other patients suffered from similar gastrointestinal symptoms due to compound heterozygous mutations of *ZAP-70* - specifically due to a combination of an activating and a hypomorphic allele (Chan *et al*., 2016). This diversity in causal mutations suggests yet undescribed underlying complexity of ZAP-70-associated SCID.

Another important part of effector function of T lymphocytes relates to interleukin 2 . Its receptor consists of three subunits α, β and γ. Subunit α is encoded by gene *IL2RA* also known as *CD25*, duplication of which has been reported in patient with VEO-IBD (Caudy *et al*., 2007; Joosse *et al*., 2018). Additionally, truncated subunit β due to *IL2RB* mutation has been reported in children with IBD (Zhang *et al*., 2019). And what's more, impaired signaling via subunit γ leading to VEO-IBD is also known. SCID resulting from *IL2RG* mutation can present apart from VEO-IBD also with susceptibility to opportunistic infections and failure to thrive (DiSanto *et al*., 1994). The immunodeficiency due to *IL2RG* mutation is more severe than *IL2RA* and *IL2RB* as subunit γ, sometimes also called common gamma chain, is shared with other receptors for interleukins like IL4, IL7 or IL21 making effects of its malfunction broader. In addition to IL2R, IL7R can also be truncated due to mutations leading to SCID with very early onset diarrhea manifestation (Puel *et al*., 1998).

Moreover, several Janus kinases (JAK) and signal transducers and activators of transcription (STAT) are involved in pathogenesis of immunodeficiencies and VEO-IBD mainly via abnormal cytokine-related signaling. For example, STAT1 immunodeficiency is predominantly characteristic by increased susceptibility to infections, especially by opportunistic pathogens. Nonetheless, it can also present with pediatric chronic colitis, reportedly in case of heterozygous mutation of DNA binding domain of *STAT1* (Uzel *et al*., 2013; Thoeni *et al*., 2016). In addition to STAT1 mutation, germline homozygous nonsense as well as somatic gain-of-function mutation of *STAT5b* can manifest similarly (Bernasconi *et al*., 2006; Ma *et al*., 2017). Also, truncate JAK3 was reported in a family with history of severe immunodeficiency

including intestinal manifestation (Barreiros *et al*., 2018). Involvement of several JAK/STAT signaling members in VEO-IBD pathogenesis proclaims the importance of this type of signaling in immunity on a systemic level.

There are other molecules apart from JAK and STAT that affect production and signaling of cytokines, such as DOCK2 (Dobbs *et al*., 2015) which affects interferon, or ADAM17 (Blaydon *et al*., 2011) and OTULIN (Damgaard *et al*., 2016) which are involved in TNFα signaling pathway. Deleterious mutation in coresponding genes might lead to pediatric IBD development with possible manifestation even in first days of life. Also, mutation in *NFAT5* influences production of cytokines in osmotic stress environment and may lead to IBD development (Boland *et al*., 2015). The proposed underlying mechanism behind NFAT5-related malfunction might be salt-induced switch to Th17 development, however, performing more tests, especially in connection to clinical cases, is necessary to validate this theory (Kleinewietfeld *et al*., 2013).

What's more, several other genes have been proposed as causal for VEO-IBD. Recently, a new type of T-, B- and NK- lymphopenia has been described in patients with early onset IBD stemming from mutation in *ANKZF1*. ANKZF1 plays a role in stress response in mitochondria and its loss of function disrupts cell respiration leading to apoptosis in lymphocytes (van Haaften - Visser *et al*., 2017). Another mechanism can involve an ubiquitin ligase ITCH, which can cause multisystem autoimmune disease with gut infiltrates when in truncated state (Lohr *et al*., 2010). Similarly, severe genetic defects affecting predominantly T lymphocytes in patients with VEO-IBD can, as well, stem from centromeric instability due to mutation of *ZBTB24* (Conrad *et al*., 2017) or from calcium channel defects mediated by mutation of STIM1 (Picard *et al*., 2009) and *ORA1* (McCarl *et al*., 2009). Despite proper identification of the deleterious genetic variants, the mechanisms of their dysregulation are not very well understood and should be considered for further investigation in order to get to know IBD pathogenesis more complexly.

Additionally, canonical inflammatory signaling pathway of NFκB, which is connected to mucosal barrier defects via NOD2 or TRIM22 dysregulation, impacts also lymphocyte functions. For example, *IKBKG*, known also as *NEMO or IKKγ*, is a kinase involved in NF-κB signaling. Its knock-out in a mouse model is colitis-inducing (Nenci *et al*., 2007) and the results of the mouse model corresponds with some pediatric IBD reported cases of immunodeficiency with enterocolitis (Cheng *et al*., 2009). In addition to that, immunodeficiency with intestinal manifestation was reported in case of

mutation of *IKBKB*, also known as *IKKB or IKKβ (*Pannicke *et al*., 2013). Considering components of NFκB pathway further downstream, hypermorphic mutation of *NFKB1A*, also called *IKBA (*Courtois *et al*., 2003)*;* as well as mutation of *NFKB1* encoding p50 (Fliegauf *et al*., 2015) and *NFKB2* encoding p52 (Klemann *et al*., 2019) can lead to very early onset IBD within immunodeficiency syndrome. Similarly, NFκB signaling can be impaired by deleterious mutation of *RELA*, sometimes called p65 (Badran *et al*., 2017), and *RELB* with possible association to immunodeficiency and IBD phenotype (Merico *et al*., 2015). Lastly, also NFκB regulator MALT1 is also associated with immunodeficiency and severe gastrointestinal inflammation (McKinnon *et al*., 2014). The number of reported genes related to NFκB, mutation of which can lead to defects in immunity with intestinal manifestation, implies the known importance of NFκB for inflammation and its management.

2.3.3. Regulatory T lymphocyte defects

Regulatory T cells are an important part of suppressive mechanisms of the immunity against self-harm. Their dysregulation, therefore, results often in severe disorders with autoimmune symptoms manifesting for example as diabetes mellitus, inflammatory bowel disease or autoimmune thrombocytopenia, neutropenia and hemolytic anemia. Due to the importance of Treg function in balancing the immune homeostasis, the onset of the causal disorders is very early, leading to life-threating states when untreated.

Treg are characterized by three main markers – they are CD4+ CD25+ and FOXP3+. Deleterious mutation in *FOXP3* leads to development of X-linked Immunodysregulation, Polyendocrinopathy, and Enteropathy called also IPEX syndrome. Interestingly, patients with VEO-IBD due to IPEX can have either significantly decreased Treg numbers because of the FOXP3 mutation (Agakidis *et al*., 2019) or the disorder might stem from truncated-FOXP3-mediated functional dysregulation with normal levels of Treg (Okou *et al*., 2014). The mechanism of FOXP3-related IBD pathology is not understood properly, however, it is suggested that impaired FOXP3-EZH2 protein interaction is involved (Bamidele *et al*., 2019).

Other marker of Treg is CD25, also known as IL2RA. Mutation of corresponding gene *IL2RA* leads to development of IPEX-like syndrome (Caudy *et al*., 2007), which symptomatically resembles the original IPEX syndrome caused by FOXP3 mutation. In

addition to IL2RA, IPEX-like syndromes were identified also due to mutations of genes like *STAT1*, *STAT5b* and *ITCH* (Verbsky *et al*., 2013).

The suppressive function of Treg is mediated by expression of suppressive membrane molecules and production of anti-inflammatory cytokines. One of the membrane suppressive molecules is CTLA4, which is an antagonist of costimulatory molecule CD28. Mutation of CTLA4 causing impairment in the binding to CD80 resulting in a decreased suppressive function, was identified in patients with early onset Crohn's disease and other autoimmune manifestations (Zeissig *et al*., 2015) showing the importance of CTLA4 for immunity homeostasis.

Apart from membrane molecule CTLA4, typical cytokines essential for regulatory function of Treg are IL10 and TGFβ. Loss-of-function mutations of *IL10* and subunits of its receptor *IL10RA* and *IL10RB* (Kotlarz *et al*., 2012) are the typical examples of VEO-IBD genes as their penetrance is nearly complete (Uhlig & Schwerd, 2016). Furthermore, activating mutation in downstream signal transductor gene *STAT3* can lead to severe infantile IBD, too (Haapaniemi *et al*., 2015).

Similarly, mutations in genes encoding suppressive cytokine TGFβ (Kotlarz *et al*., 2018) and its receptor subunits TGFβR1 and TGFβR2 have been identified as a cause of very early onset IBD. Interestingly, mutations in genes for receptor subunits *TGFBR1* and *TGFBR2* are associated with Loeys-Dietz disease typical of craniofacial dysmorphisms and predispositions for aortic aneurysms and dissections. However, reported cases of pediatric IBD patients with deleterious mutation of the same genes imply also association with early onset colitis (Naviglio *et al*., 2014).

Bearing in mind multiple mentioned mechanisms of immunity suppression, molecules involved in regulatory T cell development and function are crucial for keeping the intestinal as well as systemic balance of immunity response against pathogens and other danger signals and on the other hand dealing with autoimmune reactions. Therefore, its involvement in the development of IBD is significant and in many cases the mutations can have such a severe effect that completely penetrant monogenic disorders might occur.

2.4. Systemic inflammatory disorders

Apart from CVID, SCID and other defects with dominantly affected cell type, studies also describe several systemic monogenic immune dysregulations with complex phenotype including intestinal manifestation. One of those disorders is Mediterranean fever with its typical symptoms as fever and local inflammation in different parts of the body including intestine. (Saito at al., 2019). The responsible mutated gene for Mediterranean fever is *MEFV*. Interestingly, the disorder can manifest in autosomal recessive, as well as autosomal dominant mode of inheritance. The classical recessive type is more common and several causal variants of *MEFV* mutation have been reported so far (Uslu *et al*., 2010). On the other hand, the dominant type is more rare and seems to be caused by deleterious mutation reducing the function of pyrin with either pseudo-dominant mode of inheritance or variable penetrance of the mutation (Booth *et al*., 2000) making the evaluation of *MEFV* monogenic causality more complicated.

Also, *PLCG2* mutation can cause systemic autoinflammation displayed as PLCG2-associated antibody deficiency and immune dysregulation with IBD. Phospholipase C gamma-2 (PLCG2) is involved in intracellular signaling important for many immunocytes and other cell types and its impaired function has severe effects on activation of the immune system. Specifically, gain-of-function mutations of PLCG2 can lead to autoinflammatory reaction on skin and in the respiratory and gastrointestinal system starting in the first weeks of life, with poor response to conventional treatment (Neves *et al*., 2018).

Similarly, PSTPIP1 mutation can lead to autoinflammatory syndrome sometimes called PAPA or PAC syndrome stemming from an abbreviation of the typical symptoms - Pyoderma gangrenosum, Acne and ulcerative Colitis. The syndrome predominantly manifests by several disorders of the skin, however, there are also cases of patients with PSTPIP1 mutations presenting with hematological disorders, arthritis or colitis with very early onset (Burlakov *et al*., 2018). Therefore, PAC syndrome should be also considered in case of VEO-IBD patients with extraintestinal skin manifestations.

Another example of multisystem disorder is Behcet-like autoinflammatory syndrome which is caused by mutation in *TNFAIP3*. TNFAIP3 regulates NFκB pathway by deubiquitination of multiple involved signaling molecules. However, in its mutated state NFκB signaling is altered and consequent production of proinflammatory cytokines is increased resulting in chronic inflammatory disease. The symptoms include

early onset arthritis, uveitis and ulcerations, as well present in the gut as reported on a group of patients with infantile IBD by Zheng *et al*. (2018).

Hemophagocytic lymphohistiocytosis is complex disorder and its typical symptoms do not include intestinal inflammation. Nevertheless, several cases of patients suffering from atypical hemophagocytic lymphohistiocytosis with intestinal involvement have been reported. The patients had mutations in *STXBP2* gene, suggesting its importance for intestinal homeostasis (Meeths *et al*., 2010). A significant number of patients diagnosed with hemophagocytic lymphohistiocytosis also have a mutation in *XIAP* and *SH2D1A*, also known as *SAP* (Yang *et al*., 2012). Deleterious mutation in *XIAP* is, however, more often associated with X-linked lymphoproliferative disorder which can also involve intestinal manifestation. The proposed mechanism is NOD2-signaling impairment and thus, host-microbiota interaction defect (Kelsen *et al*., 2015b; Cifaldi *et al*., 2017;). Similarly, X-linked lymphoproliferative disorder can stem from mutation in *SH2D1A* and in case of big deletions can present also with pediatric colitis and gastritis (Booth *et al*., 2011).

Moreover, Niemann-Pick disease is normally characterized by defects in lysosomal lipid storage resulting in neurodegeneration due to mutated *NPC1*. Connected impaired autophagy is influencing also NOD2-related immunological response like in case of *XIAP* deficiency and, therefore, patients with Niemann-Pick disease can also suffer from intestinal inflammation with early onset (Schwerd *et al*., 2017).

Similarly to XIAP, POLA1 deficiency is also X-linked multisystem disorder. Its pathogenesis, however, stems from dysregulation of replication. What's more, it regulates interferon production and, thus, leads to inflammatory disorder with possible intestinal inflammation with early onset (Starokadomskyy *et al*., 2016).

Additionally, MIRAGE syndrome caused by mutation in *SAMD9* manifests typically with Myelodysplastic changes, susceptibility to Infections, growth Restrictions, Adrenal hypoplasia, Genital abnormalities and also chronic early onset Enteropathy (Narumi *et al*., 2016). The symptoms stem from SAMD9 malfunction, which is connected to IFN and TNFα signaling cascade, therefore, it is strongly involved in proinflammatory cascade and its dysregulation (Hershkovitz *et al*., 2011).

Another syndrome, which can present with intestinal pathology, is Hoyeraal-Hreidarsson syndrome. This syndrome typically manifests as congenital dyskeratosis, but intestinal ulcerations have been reported, too. The patients have either RTEL1 (Le Guen *et al*., 2013) or DKC1 (Borggraefe *et al*., 2009) deficiency due to corresponding

gene mutation causing telomeres shortening and consequently mucosal disruptions and bone-marrow-failure-induced immunodeficiency.

Similarly, a few patients suffering from Hermansky-Pudlak syndrome typical of defective membrane organelles like melanosomes, lysosomes and presence of granules in cells have been reported with IBD symptoms. Out of several genes associated up to date with Hermansky-Pudlak syndrome, *HPS1* and *HPS4* mutation can also lead to development of colitis (Hussain *et al*., 2006; Lozynska *et al*., 2018). *HPS6* mutation has been identified in IBD patient with adult onset (Mora& Wolfsohn, 2011) as well as in pediatric patients presenting, however, with ophthalmic symptoms (Hull *et al*., 2016). Concluding, *HPS6* might be a potential causal gene for VEO-IBD, however, further investigation is necessary.

Bare lymphocyte syndrome is a severe disorder caused by a mutation in one of the genes related to major histocompatibility complex (MHC). Bare lymphocyte syndrome I caused by a mutation in gene *TAP1*, *TAP2* or *TAPBP* does not manifest with IBD symptoms, but typically with skin lesions or respiratory tract disorders (de la Salle *et al*., 1999). On the other hand, Bare lymphocyte syndrome II caused by a mutation in one of the genes involved in MHC II antigen presentation - *RFX5*, *RFXAP*, *RFXANK* and *CIITA*, often manifests with very early onset diarrhea and severe immunodeficiency (Aluri *et al*., 2018; El Hawary *et al*., 2018) displaying the expected relevance of MHC presentation to proper immune reaction.

Lastly, Wiskott-Aldrich syndrome, which often manifests with thrombocytopenia and eczema, has been also reported in cases of VEO-IBD. The syndrome is caused by mutation of *WAS* encoding regulator of cytoskeleton WASP (Ohya *et al*., 2017). Alternatively, one case of Wiskott-Aldrich syndrome due to mutation of *WIPF1*, which is gene encoding WASP-interacting protein, has been reported so far (Lanzi *et al*., 2012) suggesting a potential unknown heterogeneity of this disorder. *WAS* and *WIPF1* as well as other above-mentioned causal genes are an example of genetic variants normally associated with different disorder, which should, however, be taken into consideration also in diagnostics of VEO-IBD.

2.5. Complement defects

Complement is a system of multiple proteins involved in the beginning of pathogen recognition and start-up of the following cascade of the innate, as well as adaptive immune response. Defects in complement system often disrupt triggering the immune response and lead to impaired bacterial clearance. Concerning IBD, the impaired interaction with intestinal microbiota leading to infections is a proposed complement-related mechanism contributing to the pathogenesis (Pazmandi *et al*., 2019).

One of the reported defects leading to pediatric IBD and concerning recognition of pathogens via lectin pathway is MASP2 deficiency. MASP2 deficiency is caused by mutation of *MASP2* and can lead to development of ulcerative colitis followed by several other autoimmune disorders (Stengaard-Pedersen *et al*., 2003). On the other hand, *MASP2* was found to be more associated with CD than UC (Bak-Romaniszyn *et al*., 2011) suggesting need for further investigation of this mutation.

Related to MASP2 functions and lectin pathway is ficolin 3, product of *FCN3.* Deficiency of ficolin 3 (also known as H-ficolin) can also lead to early onset immunodeficiency. And what's more, the deficiency has been found in patient with very-early onset necrotizing enterocolitis due to unmanaged intestinal infection stemming from *FCN3* mutation (Schlapbach *et al*., 2011).

Similarly, loss-of-function mutation in *CD55* (also known as *DAF*) has been reported in case of patients with very early onset protein-losing enteropathy (Ozen *et al*., 2017). This genetic association of *CD55* to IBD pathogenesis has been also supported by *Daf1*$^{-/-}$ mouse model (Lin *et al*., 2004). Being an autologous inhibitor of complement cascade, *CD55* mutation is another example of immune dysregulation with possible manifestation in the gut.

Nevertheless the lack of understanding and correlation of complement system defects to IBD development and progression, some studies suggest use of complement inhibitors as IBD treatment based on promising results (Schepp *et al*., 2009). As well, further investigation of complement in IBD patients is needed as Sina *et al*. (2018) imply preferential activation of different pathways in patients with UC (classical pathway) and CD (alternative pathway), which could be potentially useful for clinicians.

The research concerning finding new causal and associated genetic variants to VEO-IBD as well as understanding their mechanisms within immune system is necessary. The information may come from either genome-wide association studies or published case studies. The identification and validation of patient's causal variant provide not only scientific information contributing to the topic, but also serve as a base for the personalized diagnostics and consequently treatment optimization in clinics.

3. Diagnostics of VEO-IBD

3.1. Methodological approaches

Concerning the genetic and genomic analysis, the aim is to identify mutated genes which might cause or contribute to the patient's disorder. Several approaches can be used depending on the aim of the project. For example, genome wide association study (GWAS) analyses are fitting for a population study which is looking for commonly shared, probably contributing variants. On the other hand, targeted gene panel or exome sequencing are more suitable for personalized-diagnostics-based projects searching for rare causal variants (Muise *et al*., 2012). Nonetheless, some variants will not by identified by either of them leaving a space for new high-throughput but costly and technologically-demanding technologies such as whole genome sequencing or even single-cell RNA sequencing (Uhlig & Muise, 2017; Seumois & Vijayanand, 2019).

Considering the possible sequencing technologies, Sanger sequencing used to be a golden standard sequencing method, however, nowadays parallel next generation sequencing (NGS) techniques are used preferentially. The main advantage of NGS is the amount of obtained data in connection to its time-consumption and cost-effectiveness, making it possible to perform high-throughput sequencing-based diagnostics routinely in clinical laboratories. Despite the lower sequencing capacity, Sanger sequencing is still considered more accurate than NGS and, therefore, is often used as a validation of NGS data results in clinics (Mu *et al*., 2016).

Using NGS for VEO-IBD research and diagnostics, targeted gene panel sequencing, whole exome sequencing (WES) or whole genome sequencing (WGS) is performed depending on the specifics of the project as all of them have their advantages and disadvantages (Tab. 3). Targeted panel sequencing is used when aiming to identify causative variants in a set number of candidate genes such as stemming from GWAS or from already published case studies. This approach is useful for diagnostic identification of well-known candidate genes (such as *IL10)* as it considerably simplifies experimental design, data analysis and lowers cost compared to WES and WGS. However, it disallows finding new causal variants in previously unassociated genes. Moreover, the knowledge of candidate genes expands very quickly with publications of new case studies pressuring the scientists to include a considerable number of genes in the panels in order to be able to identify less abundant gene variants.

Therefore, the panel sequencing is not preferentially used in cases of rare genetic mutations such as those causing VEO-IBD (Petersen *et al*., 2017; Uhlig & Muise, 2017).

Bearing in mind clinical research of IBD causative variants, WES and WGS is more suitable approach. Sequencing of the patient's exome allows the scientists to consider variants within over 20 000 protein-coding genes and some other noncoding genes with intron-exon structure. Also, presumably important variants in introns such as splice site surroundings are sequenced during WES to be able to evaluate their impact for coding sequencing and potentially protein expression. However, variants in introns, noncoding and regulatory regions may, as well, influence the gene expression and protein function without being recognized by WES (Meyts *et al*., 2016; Uhlig & Muise, 2017) being the main drawback of this approach for rare variants diagnostics. What's more, when sequencing only exome it is necessary to enrich the exonic sequences. The enrichment, however, can be heterogenous among the genes as some genes such as *IKBKG* or *NCF1* are known to get amplified in a complicated way despite the fact they are protein-coding genes involved in VEO-IBD pathogenesis. Thus, their causal identification by WES might be troublesome in some VEO-IBD cases (Kammermeier *et al*., 2014).

When looking for the most complex picture of patient's genetic profile, whole genome sequencing is the most rewarding approach. WGS can help to identify all variants including non-coding and regulatory parts of the genome. Moreover, compared to WES whole genomic coverage is also more accurate as there is no need for preamplification. However, WGS data obtaining and analysis can be problematic and overdemanding as their processing and analysis are technologically, financially and time-consuming (Meyts *et al*., 2016; Uhlig & Muise, 2017).

Similarly, the RNA sequencing or even single-cell RNA sequencing has been showing great promise lately. Because of RNA analysis, it is able to identify pseudo-genes, different isoforms and alternative splicing variant for maximally 5 000 genetic markers per cell. Also, this approach allows for tissue- and cell-specific expression profiling. However, its cost, complicated data analysis and demands on the volume of biological samples makes it yet inconvenient for clinical purposes (Uhlig & Muise, 2017; Luecken & Theis, 2019; Seumois & Vijayanand, 2019). Therefore, currently targeted panel and WES are the first choice for IBD diagnostic and clinical research use (Fang *et al*., 2018).

Method	Cost	Analyzed genes	Advantages	Disadvantages
Sanger Sequencing	$	1	Low price Very high accuracy	Small number of sequenced genes Identifies variants only in selected gene
Targeted panel by NGS	$$	Max. hundreds	Lower cost than WES, WGS Easier data analysis than WES, WGS	Identifies variants only in selected genes
Whole exome by NGS	$$$	20 000 (all coding)	Lower price than WGS Identifies variants within all coding regions	Identifies variants only in exons Bias due to pre-amplification Demanding data analysis
Whole genome by NGS	$$$$	40 000 (all coding and noncoding)	Identifies variants within whole genome No bias due to pre-amplification compared to WES	High price Demanding data analysis
RNA sequencing	$$$-$$$$$	Max. 5 000	Identifies pseudo-genes, isoforms, splicing variants Can identify cell- and tissue-specific gene expression	Very high price Demanding data analysis Needs huge volume of sample

Tab. 3: Advantages and disadvantages of sequencing techniques suitable for VEO-IBD diagnostics (Seleman *et al*., 2017; Uhlig & Muise, 2017; Willyard, 2018; Luecken & Theis, 2019)

3.2. Variant filtering criteria

When using any next generation sequencing approach and mechanism of sequencing, it is afterwards necessary to process and analyze the data. The amount of raw sequencing data (in FASTQ format) coming from NGS is huge, needing a way to process them effectively. The short fragments of sequences have to be mapped and aligned to the chosen reference human genome (output in BAM format). Followed by variant calling, the final VCF file contains information about the similarities and differences between the reference and patient's genome at every nucleotide position and their annotations. The annotations contain information about genotyping quality score and read depth which are important for subsequent evaluation of the data quality and relevance. The bioinformatical algorithms for mapping and alignment as well as for variant calling may, however, differ, and provide different data sets for interpretation.

Therefore, it is important to establish an adequate data-processing pipeline in order to get relevant data for subsequent filtering and variant identification (Meyts *et al*., 2016).

Filtering is necessary to sort out impactless and incorrectly called variants. Firstly, confidence of the variant calling should be considered in filters. The quality control of the called variants includes:

- **Genotyping (sometime also called call) quality** - refers to the probability, that the non-referential variant is present at that site. It is recommended to use the genotype quality score threshold of 20 and more (Meyts *et al*., 2016).
- **Read depth** - refers to the number of reads (fragments of sequences), that are covering the site. For high-coverage analysis suitable for identification of rare variants read depth > 20 in majority of the analyzed sequences is recommended (Do *et al*., 2012; Meyts *et al*., 2016).
- **Allelic fraction** - refers to the ratio of alleles leading to establishment of zygosity. The value between 20 and 80 % is recommended for calling heterozygote (Do *et al*., 2012).

Secondly, it is necessary to filter out harmless single nucleotide polymorphisms (SNP), copy number variations (CNV) or insertions/deletions from potentially pathogenic variants. The common variants are filtered based on:

- **Allele frequency** - refers to frequency of the allele in a selected population. It is recommended to use the threshold of 1% presence for identification of rare variants (Do *et al*., 2012; Meyts *et al*., 2016).

The data about allele frequency are available for example in The Genome Aggregation Database (gnomAD) (https://gnomad.broadinstitute.org/, accessed on January 21, 2020). The sequencing data and allele frequencies could be also obtained from already ended 1000 Genomes Project (The 1000 Genomes Project Consortium, 2015), The Exome Aggregation Consortium (ExAC) (Lek *et al*., 2016) research, which have been already incorporated into gnomAD, and NHLBI Exome Sequencing Project (ESP) (https://evs.gs.washington.edu/EVS/, accessed on January 21, 2020). As a result, a statistically relevant amount of data is available allowing the scientists to interpret their data in context of whole population as well as in origin-specific cohorts.

Thirdly, several tools are available for prediction of deleterious effects of the variant. The American College of Medical Genetics and Genomics (ACMG) set up a classification of the variants based on their predicted impact. The recognized categories are pathogenic, likely pathogenic, benign, likely benign and uncertain significance (Richards *et al*., 2015). The tools for prediction of the impact are for example:

- **SIFT** (Sorting Intolerant from Tolerant) - is used for assessment of amino acids substitution on the protein function (Ng & Henikoff, 2003) generating result based on the ACMG nomenclature (Richards *et al*., 2015).
- **PolyPhen-2** (Polymorphism Phenotyping v2) - is used for prediction of protein structural and functional changes due to amino acids substitution (Adzhubei *et al*., 2010) generating result based on the ACMG nomenclature (Richards *et al*., 2015).
- **MutationTaster** - is used for evaluating an impact of the variant on produced protein including the intronic and synonymous variants resulting in predicted disease-causing potential (Schwarz *et al*., 2014)
- **CADD** (Combined Annotation Dependent Depletion) - processes a CADD score which is used for measuring the deleteriousness of the variant. The range of CADD score values is from 1 to 99. For example, CADD = 10 is related to a variant which is among top 10 % deleterious variants within the human genome, CAD = 20 relates to a variant within 1 % of the most deleterious etc. (Kircher *et al*., 2014). It is recommended to use CADD score > 15 for identification of rare variants (Meyts *et al*., 2016).

What's more, **dbSNP,** a database of previously described single nucleotide polymorphisms (Sherry *et al*., 1999), can be also used to establish the effect of the variant.

Fourthly, the data can be filtered by a virtual panel of selected genes. Virtual panels are similar to targeted panel sequencing, however, they are applied on whole exome or genome data after the sequencing. The advantage is an easy modification of the virtual panel in context of newly published findings (Meyts *et al*., 2016; Uhlig & Muise, 2017).

Lastly, the list of filtered variants based on above mentioned criteria has to be evaluated manually bearing in mind the biological context of the complete data set. The patient's specific cases should be put together with known objectives such as the allele

frequency and predicted functional impact of the gene variant identified in the patient. Also, localization of the variant within the protein can be checked using for example:

- **Pediatric Cancer Genomic Data Portal** (PeCan, https://pecan.stjude.cloud/, accessed January 21, 2020).

In addition, inheritance mode of the genetic disorder and the disease relevance to patient's diagnosis should be also taken into account. The information may be found in databases such as:

- **Online Mendelian Inheritance in Man** (OMIM, http://omim.org, accessed January 21, 2020)
- **ClinVar** (https://www.ncbi.nlm.nih.gov/clinvar/, accessed January 21, 2020)
- **The Human Gene Mutation Database** (HGMD, http://www.hgmd.cf.ac.uk/, accessed January 21, 2020)

After evaluating the patient's sequencing data by several above-mentioned approaches, the possibly causal variant may be identified. The proposed variant has to be, however, further validated to confirm its presence and functional impact causing or contributing to the pathology (Uhlig & Muise, 2017).

3.3. Validation of the variants

To validate the proposed causal hypothesis, accuracy of the NGS-based findings has to be confirmed. Possible validation methods are PCR-based approaches, cDNA sequencing and Sanger sequencing of the proposedly mutated region. After confirmation of the variant, functional validation of its impact takes place. In case of previously described variant, the literature-based evidence is sufficient for claiming its causation or contribution to the pathology. However, if the specific genetic variant has not yet been described in the literature, it is necessary to validate its impact by functional tests. Functional tests might include cytokine assays, ROS production testing and flow-cytometric immunophenotypization. Only when understanding the genetics in connection to the functional physiological point of view, the causal variant can be properly identified and addressed (Meyts *et al*., 2016; Uhlig & Muise, 2017).

The results of the tests and analyses, namely identification of the causal mutation and its functional validation, help to get to the base of the patient's pathology. Consequently, it is possible to find precision-medicine-based option for patient's treatment such as using specific inhibitors for gain-of-function mutations like in case of PIK3CD syndrome (Notarangelo *et al*., 2017). This analytical approach leading to a new possible targeted treatment is particularly helpful in cases of VEO-IBD patients who poorly react to the conventional treatment methods needing personalized alternative treatment options (Uhlig *et al*., 2014).

PRACTICAL PART

4. Aims of the thesis

The diploma thesis called **Mechanisms of immune dysregulation leading to inflammatory bowel disease** aims to:

- Analyze whole-exome sequencing (WES) data in a cohort of pediatric patients diagnosed with inflammatory bowel disease.
- Identify and validate possible causal variants of the patients' disorder in context of heredity and molecular mechanism of the pathology.
- Compare different bioinformatical approaches for next-generation sequencing (NGS) data processing and filtering and evaluate their use for VEO-IBD variant identification.

5. Materials and methods

5.1. Patients

The whole-exome sequencing was performed in 20 patients diagnosed with pediatric-onset inflammatory bowel disease at Gastroenterological Pediatric Clinic, Second Faculty of Medicine, Charles University and University Hospital Motol, Prague. The cohort included 12 males and 8 females diagnosed with Crohn's disease, ulcerative colitis or inflammatory bowel disease unspecified without known cause of the disease. The patients' age of onset ranged from 3 to 154 months of age (in case of two patients the time of onset was not available) as shown in Tab. 6.

Beforehand the WES analysis, the patients provided family and personal anamnesis and underwent physical and laboratory examination at Gastroenterological Pediatric Clinic of Faculty Hospital Motol. The laboratory tests were performed from peripheral blood and usually covered standard hematological, immunological and biochemical evaluation. When indicated, IBD-specific parameters such as levels of calprotectin in stool were also examined. The information from examinations and tests were available for consideration of case specifics during the causal variant identification.

When a possible causal variant was identified by WES data analysis, parents or other relevant relatives (siblings, grandparents etc.) were asked to provide their peripheral blood sample for validation of the allelic segregation within the family. In some cases, however, the samples have not been provided yet resulting in an inability to validate some proposed causal variants in patients.

The patients' data were anonymized. All samples were used with the consent of patient's legal representative or the participants themselves.

5.2. Samples preparation and sequencing

For the purpose of the whole-exome sequencing, DNA was isolated using QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) from peripheral blood of the patients and their relatives. SureSelectXT Human All Exon V6+UTR kit (Agilent Technologies, Santa Clara, CA) was used for library preparation. The sequencing was performed using Illumina NextSeq 500 (Illumina, San Diego, CA) instrument. The samples were handled in concordance with standard operation procedures by laboratory technicians and researchers of Laboratory of Molecular Genetics of Department of CLIP - Childhood Leukaemia Investigation Prague and Department of Paediatric Haematology and Oncology, Second Faculty of Medicine, Charles University and University Hospital Motol, Prague.

5.3. WES data analysis

5.3.1. Bioinformatical raw sequencing data processing

Having sequenced the exome of patients' samples, FASTQ file was generated. Afterwards, the FASTQ data was processed by two bioinformatical pipelines - VarScan2 pipeline (using bwa, samtools and VarScan2 tools) and GATK4 pipeline (using bwa, BQSR, samtools and GATK4 tools). The differences between the pipelines and their suitability for the analysis were evaluated later and are further described in section *5.5.1 Bioinformatical pipelines.*

Within the pipeline, mapping and alignment of the sequenced fragments to reference genome GRCh37/hg19 took place producing file in BAM format. The BAM file was always accompanied by index file in BAI format. Subsequently, VCF file was automatically produced containing information about variant calling and annotations of the called variants.

5.3.2. Variant filtering

On average, there were hundreds of thousands of variants called within tens of thousands of genes including mainly calls lacking the required quality and known harmless SNPs. Therefore, it was necessary to filter the variants down to tens of variants which can be evaluated individually in context of patient's case as displayed on the right side of the Tab. 4. The filtering was performed by commercially available online tool for variant analysis QIAGEN Ingenuity Variant Analysis (IVA) (https://digitalinsights.qiagen.com/products/qiagen-ingenuity-variant-analysis/, accessed March 25, 2020).

IVA allows the user to set several levels of filtering criteria based on the specifics of his analysis. The filtering criteria can be divided into four sections:

- **Confidence** - criteria concerning quality of the sequencing, alignment and variant calling of the variants,
- **Common variants** - criteria concerning abundance of the variants within population in chosen databases (such as gnomAD),
- **Predicted deleterious** - criteria concerning the functional impact of the variants based on the reports in databases (such as HGMD) or *in-silico* predictions (such as SIFT),
- **Biological context** - criteria concerning the biological and genetic aspects of the investigated disease.

When applying the biological context, you can use customized gene lists and filter the variants by these virtual gene panel. Four different gene lists were used to filter the variants according to their biological context. The gene lists were:

- **Closest Disease-Causing Genes (CDG) gene list** including 425 genes (Supplementary file 1) closely associated with VEO-IBD or with genes described in VEO-IBD pathogenesis (Requena *et al*., 2018),
- **Primary Immunodeficiency (PID) gene list** including 410 genes (Supplementary file 1) coming from official classification of primary immunodeficiencies by International Union of Immunological Societies (IUIS), which was up-to-date at the time of the gene list creation (Bousfiha *et al*., 2018),

- **Monogenic Inflammatory Bowel Disease (IBD) gene list** including 50 genes (Supplementary file 1) reported in cases of monogenic VEO-IBD and summarized by Uhlig *et al*., 2014,
- **Expanded Monogenic Inflammatory Bowel Disease (IBD2)** gene list including 113 genes (Supplementary file 1), which were reported in literature (Tab. 5) as causal in cases of pediatric IBD and summarized for the purposes of this analysis.

The criteria for the IBD patients' analyses were applied as stated in Tab. 4:

Filtering	Criteria	
		300 000
Confidence	Keep only variants which satisfy all of these criteria: • Call quality ≥ 20, • Read depth ≥ 10, • Allele fraction ≥ 30.	130 000
Common Variants	Exclude variants that are observed in any of these populations with an allele frequency of: • ≥ 1 % in the 1000 Genomes Project, • ≥ 1 % in the ExAC, • ≥ 1 % in the gnomAD, • ≥ 1 % in NHLBI ESP exomes, unless an established pathogenic common variant.	5 000
Predicted Deleterious	Keep only variants no more than 20 bases into intron that are: • Classified as Pathogenic or Likely Pathogenic, • Listed in HGMD and ClinVar, • Associated with gain-of-function of a gene in literature, • Associated with loss-of-function of a gene as - frameshift, in-frame indel, or start/stop codon change - missense unless predicted tolerated by SIFT or PolyPhen-2 - predicted deleterious by having CADD score > 15 - splice site loss up to 2 bases into intron as predicted by MaxEntScan	600
Biological Context	Keep only variants that are known or predicted to affect genes listed below: • CDG gene list • PID gene list • IBD gene list • IBD2 gene list	50 variants

Tab. 4: Filtering criteria applied in Ingenuity Variant Analysis with approximate numbers of variants filtered at each step

MECHANISM		GENE	DISORDER	INHERI -TANCE	REFERENCE
Epithelial barrier defects		*ALPI*	Inflammatory bowel disease	AR?	Parlato *et al*., 2018
		COL7A1	Dystrophic epidermolysis bullosa	XLR	Ashton *et al*., 2016; Freeman *et al*., 2008
		DUOX2	Pancolitis	AD?	Hayes *et al*., 2015
		EPCAM	Tufting enteropathy	AR	Al-Mayouf *et al*., 2009
		FERMT1	Kindler syndrome	AR	Ashton *et al*., 2016; Freeman *et al*., 2008
		GUCY2C	GUCY2C gain-of-function diarrhea	AD	Fiskerstrand *et al*., 2012
		NOD2	Inflammatory bowel disease	AR?	Girardelli *et al*., 2018
		NOX1	Pancolitis	XLR?	Hayes *et al*., 2015
		PEPD	Prolidase deficiency	AR	Kuloglu *et al*., 2015; Rizvi *et al*., 2019
		PLA2G4A	Cryptogenic multifocal ulcerating stenosing enteritis	AR	Brooke *et al*., 2014; Faoni *et al*., 2014
		SLC9A3	Secretory sodium diarrhea	AR	Janecke *et al*., 2015
		SLCO2A1	Chronic enteropathy associated with *SLCO2A1*	AR	Umeno *et al*., 2015
		TRIM22	Inflammatory bowel disease	AR?	Li *et al*., 2016
		TTC7A	TTC7A immunodeficiency	AR	Samuels *et al*., 2013; Bigorgne *et al*., 2014
Phagocyte defects		*CYBA*	Chronic granulomatous disease	AR	Teimourian *et al*., 2008
		CYBB	Chronic granulomatous disease	XLR	Segal *et al*., 1983
		G6PC3	Severe congenital neutropenia	AR	Mistry *et al*., 2017
		ITGB2	Leukocyte adhesion deficiency type 1	AR	Uzel *et al*., 2001
		NCF1	Chronic granulomatous disease	AR	van de Vosse *et al*., 2009
		NCF2	Chronic granulomatous disease	AR	Muise *et al*., 2012
		NCF4	Chronic granulomatous disease	AR	Matute *et al*., 2009
		NLRC4	Autoinflammation with infantile enterocolitis	AD	Canna *et al*., 2014
		RAC2	Neutrophil immunodeficiency syndrome	AD	Ambruso *et al*., 2000
		SLC37A4	Glycogen storage disease Ib	AR	Volz *et al*., 2015
Lymphocyte defects	B lypmhocyte and antibody defects	*AICDA*	Hyper IgM syndrome	AR	Quartier *et al*., 2004
		BTK	Agammaglobulinemia	XLR	Maekawa *et al*., 2010
		CD19	CVID	AR	Kanegane *et al*., 2007
		CD40LG	Hyper IgM syndrome	XLR	Levy *et al*., 1997; Wang *et al*., 2014
		DOCK8	Hyper IgE syndrome	AR	Sanal *et al*., 2012
		ICOS	CVID	AR	Grimbacher *et al*., 2003
		ICOSLG	CVID	AR	Roussel *et al*., 2018
		IL21	CVID	AR	Salzer *et al*., 2014
		sBP2	CVID	AD	Keller *et al*., 2016
		LRBA	CVID	AR	Alangari *et al*., 2012
		PIK3CD	Hyper IgM syndrome	AD	Elgizouli *et al*., 2016; Farachi *et al*., 2018
		PIK3R1	Agammaglobulinemia and immunodeficiency	AD, AR	Conley *et al*., 2012; Lucas *et al*., 2014; Farachi *et al*., 2018
		PTEN	CVID	AD	Driessen *et al*., 2016
		SKIV2L	Trichohepatoenteric syndrome	AR	Fabre *et al*., 2012
		TNFRSF13B	CVID, IgA deficiency	AD, AR	Salzer *et al*., 2008
		TRNT1	Sideroblastic anemia with B-cell immunodeficiency, periodic fevers, and developmental delay	AR	Frans *et al*., 2017
		TTC37	Trichohepatoenteric syndrome	AR	Hartley *et al*., 2010

Tab. 5: Summary of monogenic disorders leading to very early onset inflammatory bowel disease - the list of genes was used as IBD2 gene list (AR = autosomal recessive; AD = autosomal dominant; XLR = X-linked recessive

MECHANISM	GENE	DISORDER	INHERI -TANCE	REFERENCE
Effector T cell defects	*ADA*	T- B- NK- SCID	AR	Felgentreff *et al*., 2011
	ADAM17	Neonatal inflammatory skin and bowel disease	AR	Blaydon *et al*., 2011
	ANKZF1	T-, B- and NK-cell lymphopenia	AD?	van Haaften - Visser *et al*., 2017
	CD3D	T- B+ NK+ SCID	AR	Dadi *et al*., 2003
	CD3E	T- B+ NK+ SCID	AR	de Saint Basile *et al*., 2004
	CD3G	T+/low B+ NK+ SCID	AR	Recio *et al*., 2007
	DCLRE1C	T- B- NK+ SCID Omenn syndrome	AR	Felgentreff *et al*., 2015
	DOCK2	Immunodeficiency	AR	Dobbs *et al*., 2015
	FOXN1	T-cell immunodeficiency with congenital alopecia and nail dystrophy	AR	Chou *et al*., 2014
	IKBKB	Immunodeficiency	AD, AR	Pannicke *et al*., 2013
	IKBKG	IKBKG immunodeficiency	XLR	Cheng *et al*., 2009
	IL2RA	Immunodeficiency with lymphoproliferation and autoimmunity IPEX-like syndrome	AR	Caudy *et al*., 2007; Joosse *et al*., 2018
	IL2RB	Immunodeficiency with lymphoproliferation and autoimmunity	AR	Zhang *et al*., 2019
	IL2RG	T- B+ NK- SCID	XLR	DiSanto *et al*., 1994
	IL7R	T- B+ NK+ SCID	AR	Puel *et al*., 1998
	ITCH	Multisystem autoimmune disease with facial dysmorphism IPEX-like syndrome	AR	Lohr *et al*., 2010
	JAK3	T- B+ NK- SCID	AR	Barreiros *et al*., 2018
	LCK	Immunodeficiency	AR	Goldman *et al*., 1998
	LIG4	T- B- NK+ SCID LIG4 syndrome	AR	van der Burg *et al*., 2006
	MALT1	Immunodeficiency	AR	McKinnon *et al*., 2014
	NFAT5	Immunodeficiency Autoimmune enterocolopathy	AD	Boland *et al*., 2015
	NFKB1	CVID	AD	Fliegauf *et al*., 2015
	NFKB2	CVID	AD	Klemann *et al*., 2019
	NFKBIA	Ectodermal dysplasia and immunodeficiency	AD	Courtois *et al*., 2003
	ORAI1	Immunodeficiency	AR	McCarl *et al*., 2009
	PNP	Immunodeficiency due to purine nucleoside phosphorylase deficiency	AR	Shanon *et al*., 1988
	PRKDC	T- B- NK+ SCID	AR	Mathieu *et al*., 2015
	RAG1	T- B- NK+ SCID Omenn syndrome	AR	Corneo *et al*., 2001
	RAG2	T- B- NK+ SCID Omenn syndrome	AR	Corneo *et al*., 2001
	RELA	Chronic mucocutaneous ulceration	AD	Badran *et al*., 2017
	RELB	Immunodeficiency	AR	Merico *et al*., 2015
	STAT1	Immunodeficiency due to STAT1 deficiency IPEX-like syndrome	AR, AD	Uzel *et al*., 2013; Thoeni *et al*., 2016
	STAT5b	Growth hormone insensitivity with immunodeficiency IPEX-like syndrome	AR?, Somatic?	Ma *et al*., 2016
	STIM1	Immunodeficiency	AR	Picard *et al*., 2009
	ZAP70	Immunodeficiency Multisystem autoimmune disease	AR	Parry *et al*., 1996; Chan *et al*., 2016
	ZBTB24	Immunodeficiency-centromeric instability-facial anomalies syndrome	AR	Conrad *et al*., 2017

Tab. 5 - continuing: Summary of monogenic disorders leading to very early onset inflammatory bowel disease - the list of genes was used as IBD2 gene list (AR = autosomal recessive; AD = autosomal dominant; XLR = X-linked recessive

MECHANISM		GENE	DISORDER	INHERI -TANCE	REFERENCE
	Regulatory T cell defects	*CTLA4*	Autoimmune lymphoproliferative syndrome	AD	Zeissig *et al*., 2015
		FOXP3	Immunodysregulation, Polyendocrinopathy, and Enteropathy (IPEX)	XLR	Okou *et al*., 2014; Agakidis *et al*., 2019
		IL10	Early onset Inflammatory Bowel Disease	AR	Kotlarz *et al*., 2012
		IL10RA	Early onset Inflammatory Bowel Disease	AR	Kotlarz *et al*., 2012
		IL10RB	Early onset Inflammatory Bowel Disease	AR	Kotlarz *et al*., 2012
		STAT3	Multisystem autoimmune disease	AD	Haapaniemi *et al*., 2014
		TGFB1	Inflammatory bowel disease, immunodeficiency, and encephalopathy	AR	Kotlarz *et al*., 2018
		TGFBR1	Loeys-Dietz syndrome	AD	Naviglio *et al*., 2014
		TGFBR2	Loeys-Dietz syndrome	AD	Naviglio *et al*., 2014
Complement		*CD55*	Complement hyperactivation, angiopathic thrombosis, and protein-losing enteropathy (CHAPLE)	AR	Ozen *et al*., 2017
		FCN3	Immunodeficiency due to ficolin 3 deficiency	AR	Schlapbach *et al*., 2011
		MASP2	MASP2 deficiency	AR	Stengaard-Pedersen *et al*., 2003
Systemic inflammatory disorders		*DKC1*	Hoyeraal-Hreidarsson syndrome	XLR	Borggraefe *et al*., 2009
		CIITA	Bare lymphocyte syndrome	AR	Aluri *et al*., 2018; El Hawary *et al*., 2018
		HPS1	Hermansky-Pudlak syndrome	AR	Hussain *et al*., 2006
		HPS4	Hermansky-Pudlak syndrome	AR	Hussain *et al*., 2006; Lozynska *et al*., 2018
		HPS6	Hermansky-Pudlak syndrome	AR	Mora *et al*., 2011; Hull *et al*., 2016
		MEFV	Mediterranean fever	AD, AR	Uslu *et al*., 2010, Saito at al., 2019
		MVK	Hyper IgD syndrome	AR	Bader-Meunier *et al*., 2011
		NLRC4	Autoinflammation with infantile enterocolitis	AD	Canna *et al*., 2014
		NPC1	Niemann-Pick disease	AR	Schwerd *et al*., 2017
		PLCG2	Autoinflammation and PLCG2-associated antibody deficiency and immune dysregulation (APLAID)	AD	Neves *et al*., 2018
		POLA1	Pigmentary disorder, reticulate, with systemic manifestations	XLR	Starokadomskyy *et al*., 2016
		PSTPIP1	Pyoderma gangrenosum, acne and ulcerative colitis (PAC) syndrome	AD	Burlakov *et al*., 2018
		RFX5	Bare lymphocyte syndrome	AR	Aluri *et al*., 2018; El Hawary *et al*., 2018
		RFXANK	Bare lymphocyte syndrome	AR	El Hawary *et al*., 2018
		RFXAP	Bare lymphocyte syndrome	AR	Aluri *et al*., 2018
		RTEL1	Hoyeraal-Hreidarsson syndrome	AD, AR	Le Guen *et al*., 2013
		SH2D1A	Lymphoproliferative syndrom	XLR	Booth *et al*., 2010
		STXBP2	Hemophagocytic lymphohistiocytosis	AR	Meeths *et al*., 2010
		TNFAIP3	Behcet-like autoinflammatory syndrome	AD	Zheng *et al*., 2018
		WAS	Wiskott-Aldrich syndrome	XLR	Ohya *et al*., 2017
		WIPF1	Wiskott-Aldrich syndrome	AR	Lanzi *et al*., 2012
		XIAP	Lymphoproliferative syndrom	XLR	Cifaldi *et al*., 2017

Tab. 5 - end: Summary of monogenic disorders leading to very early onset inflammatory bowel disease - the list of genes was used as IBD2 gene list (AR = autosomal recessive; AD = autosomal dominant; XLR = X-linked recessive)

5.3.3. Identification of the causal variant

After the filtering in IVA, the list of filtered variants based on above mentioned criteria needed to be evaluated manually bearing in mind the patient's biological context in connection to the complete sequenced and filtered data set. The objectives of the filtered potentially causal variants were considered checking the allele frequency aiming at variants with very low (preferably under 0.1 %) or unknown frequency and counts of homozygotes and heterozygotes found in gnomAD aiming for zero or very low number. Also predicted or reported functional impact was evaluated looking for previously described SNPs in dbSNP NCBI and pathogenic, likely pathogenic or variants with uncertain significance predicted by SIFT, Polyphen-2 or MutationTaster. Moreover, CADD-score-based deleteriousness of the variant looking for CADD score > 20 was taken into account.

Moreover, localization of the variant within the protein was considered as shown for example in PeCan database. Also, inheritance mode of the genetic disorder and the disease relevance to patient's symptoms, onset and other case specifics, such as family anamnesis, were evaluated using OMIM, ClinVar and HGMD reports. Based on that, we were searching for case reports in literature presenting patients with similar pathology caused by a mutation in a suspected gene. In addition, we also checked the quality of sequencing data to prevent validation of artefacts using Integrated Genomics Viewer (IGV) (Thorvaldsdóttir *et al*., 2013; https://igv.org/, accessed May 28, 2020). Bearing in mind the specifics of the patient's case in relation to present variants, potential causal variants were suggested for further validation.

5.4. Validation of the variants

Validation primarily included confirmation of the variant in patient's sample and its segregation in family. DNA samples of the patient, both parents and when relevant also of other relatives, such as siblings, were sequenced using Sanger sequencing method. The preparation of the samples for the Sanger sequencing included primer design for the patient's mutated region. The primers were designed using reference genome sequence GRCh37/hg19 available in UCSC Genome Browser (https://genome.ucsc.edu/, accessed May 10, 2020) and online tool Primer-BLAST from NCBI (https://www.ncbi.nlm.nih.gov/tools/primer-blast/, accessed May 10, 2020) preferring

primers covering a whole exon if possible. The suggested primer pairs were, then, doublechecked by online tool In-silico PCR from UCSC (https://genome.ucsc.edu/cgi-bin/hgPcr, accessed May 10, 2020) aiming for similar melting temperatures of the forward and reverse primer around 61 °C and avoiding GC rich and repetitive regions. The structure of the primers was also checked in an online tool OligoAnalyzer from IDT (https://www.idtdna.com/pages/tools/oligoanalyzer, accessed May 10, 2020) for the absence of hairpins and self-dimerizing and hetero-dimerizing sequences. The suitable primers were order externally from Sigma Aldrich.

Samples for the Sanger sequencing were prepared by PCR using PCR Master Mix components as stated in Tab. 6 and 1 µl of genomic DNA of approximate concentration 100 ng/µl. Each series of PCR also included negative aqua control.

Reagents	**Volume [µl]**
Water	15,90
10x buffer II (Applied Biosystems)	2,50
MgCl2 solution 25mM (Applied Biosystems)	1,50
dNTPs (Applied Biosystems)	2,00
AmpliTaq Gold™ DNA Polymerase 250 U (Applied Biosystems)	0,10
Primer forward (10 pM/µl)	1,00
Primer reverse (10 pM/µl)	1,00

Tab. 6: PCR Master Mix components

The PCR was performed using 2720 Thermal Cycler (Applied Biosystems) and the standard PCR program consisted of:

- Predenaturation at 95 °C for 7 mins,
- 35 cycles of: Denaturation at 94 °C for 30 secs,
  Annealing at 62 °C for 30 secs,
  Extension at 72 °C for 30 secs,
- Final extension at 72 °C for 10 mins.

In case of problematic amplification, concentration and volumes of Master Mix components and temperature profile of the PCR were optimized for that specific case. When applied, the differences from the standard protocol are stated in section *6. Results* in description of patients' cases.

The PCR results were checked by gel electrophoresis using 2% agarose gel in 1x TAE buffer with addition of 5 µl of GelRed® Nucleic Acid Stain 10,000X in Water (Biotium). 5 µl of the GeneRuler 100 bp Plus DNA Ladder (Thermo Scientific) and 5 µl of the PCR product with 6x DNA Loading Dye (Thermo Scientific), respectively, were loaded into the gel. The electrophoresis run at 90 V for 45 minutes using Bio-Rad Power Pac 300. Subsequently, transilluminator Uvitec FireReader was used to visualize the electrophoretic separation of the amplicons.

PCR amplicons were, subsequently, used for Sanger sequencing, as well as the same forward primers as used for the PCR. The Sanger sequencing premix included reagents as stated in Tab. 7. The Sanger sequencing was performed externally at Eurofins Genomics.

Reagents	**Volume [µl]**
Water	5,00
PCR product	2,50
Primer forward (10 pM/µl)	2,50

Tab. 7: Sanger Sequencing premix components

The results of the Sanger sequencing were evaluated using SnapGene® Viewer 5.0.6. In case of corresponding segregation of the allele within the family, patient's variant was confirmed. In case of a variant already reported in literature as causal in a relevant case, it was stated as causal. In case of a novel unreported variant, the patient's results and samples were forwarded to Laboratory of Flow Cytometry of Department of CLIP - Childhood Leukaemia Investigation Prague and Department of Paediatric Haematology and Oncology, Second Faculty of Medicine, Charles University and University Hospital Motol, Prague, for immunophenotypization and further functional testing.

5.5. Evaluation of different bioinformatical approaches

5.5.1. Bioinformatical pipelines

A bioinformatical pipeline is used for the processing of raw sequencing data aligning the reads, calling and annotating the variants. As the different pipelines process the same raw sequencing data by different tools, the alignment, variant calling and annotation may also differ. Having available two different bioinformatical pipelines for raw sequencing data processing, we evaluated their suitability for the VEO-IBD variant analysis . The two pipelines were:

- VarScan2 pipeline - using bwa, samtools and VarScan2 tools,
- GATK4 pipeline - using bwa, BQSR, samtools and GATK4 tools.

VarScan2-based pipeline was used originally for WES analyses of different types of samples in CLIP Laboratories. A GATK4-based pipeline was recently developed to incorporate newly available bioinformatic tools and required validation on clinical samples. Thus, we compared the results of variant identification using both pipelines with the same samples and under the same conditions.

Raw sequencing data of 19 VEO-IBD patients (Tab. 9) were submitted for the comparative analysis. Patient #1 was not included due to unavailability of the complete data sets at the time of the analysis. The raw sequencing data of the patients were processed twice - once by each pipeline, followed by Ingenuity variant analysis of VarScan2- and GATK4-generated VCF file. The filtering criteria for IVA were set as stated in Tab. 4 and the data of one patient were exported from IVA at the same time point to avoid bias by updates of the online tool.

Subsequently, we manually evaluated the differences in variant calling and filtering using IGV and IVA. Selecting the variants filtered only in one of the pipelines, the reason for the discrepancy was assessed in case of each variant. The possible reasons were either different variant calling (caused by distinct alignment or variant calling) or filtering out of the variant in the Ingenuity Variant Analysis due to different variant parameters given to the it by the pipeline, such as genotype quality, read depth or allelic fraction (all included in the “Confidence” step of filtering). Moreover, the presence of each distinct variant within the gene lists (CDG, PID, IBD and IBD2 as described in section *5.3.2. Variant filtering* and in Supplementary file 1) used for filtering was also

taken into account. The results were statistically evaluated considering the following criteria:

- number of variants called only in VarScan2 (not called in GATK4),
- number of variants called only in GATK4 (not called in VarScan2),
- number of variants filtered only in VarScan2 (filtered out in GATK4),
- number of variants filtered only in GATK4 (filtered out in VarScan2),
- number of variants called and filtered by VarScan2 and GATK4,
- number of positively identified causal variants in VarScan2,
- number of positively identified causal variants in GATK4.

The comparison was made, firstly, for each patient individually. Afterwards, the results were summarized allowing us to assess the similarities and differences of bioinformatical processing of the same raw data by two pipelines using different tools. Based on the findings, we evaluated the suitability of the new GATK4 pipeline for the purposes of genetic diagnostics looking mainly for a wide range of high-quality called variants allowing us to identify the causal variant.

5.5.2. Gene lists for variant filtering

Using four gene lists (CDG, PID, IBD and IBD2 as described in section *5.3.2. Variant filtering* and in Supplementary file 1) for the variant filtering in respect to their biological context, we evaluated their suitability for the analysis. The percentage of positive variant identification was established for each gene list. Based on the results, we concluded the suitability of each gene list for the purposes of VEO-IBD genetic analysis.

6. Results

6.1. Identified causal variants in patients

We analyzed whole-exome sequencing data from 20 patients with VEO-IBD diagnosis (Tab. 9). We found 13 suspected causal variants in 10 out of 20 patients (in 1 case we validated more than one suspected gene and in 2 cases we suspected compound heterozygosity as a cause of the disorder) as displayed in Fig. 3 and Tab. 8

We confirmed 5 variants in 4 genes (*NOD2, NLRP3, FOXP3, DUOX2* compound heterozygote*)* and declined 2 variants (*ERBIN, IRF2BP2*) by Sanger sequencing. We found 6 more potentially causal variants in 5 genes (*MEFV, PRKDC* compound heterozygote*, PSTPIP1, BCL11B, PLCG2*), which, however, were not validated yet by Sanger sequencing due to unavailability of the parental samples.

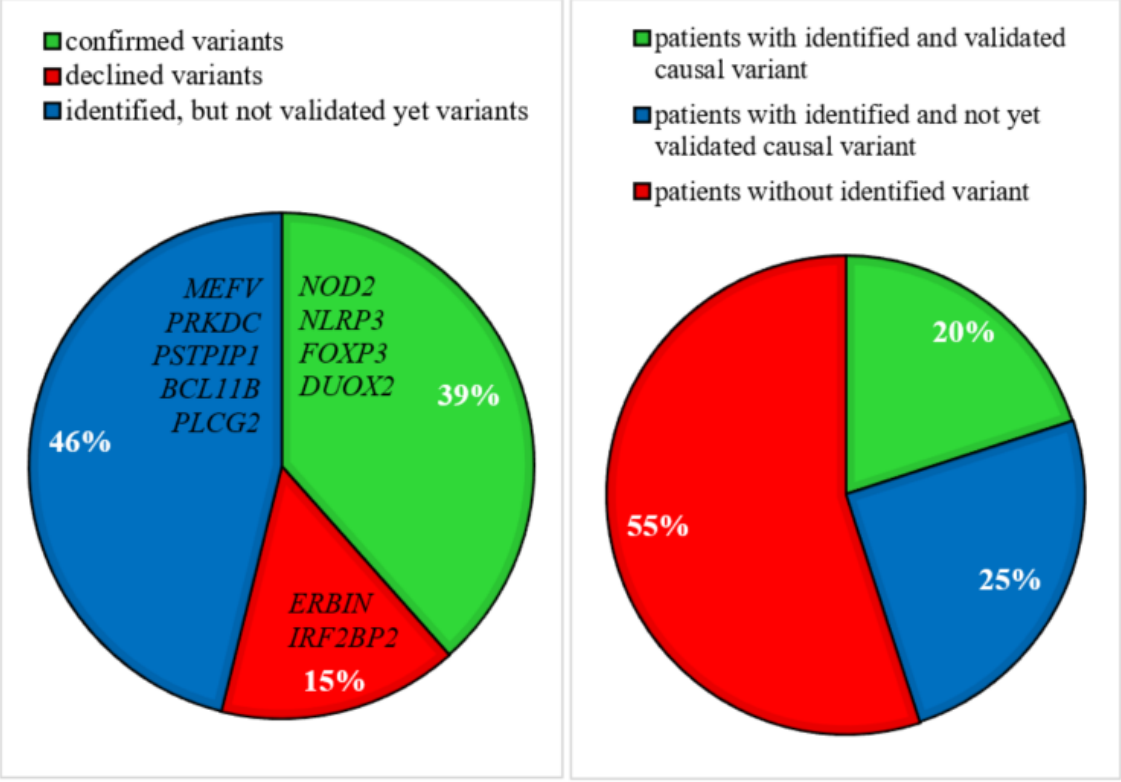


Fig. 3: Results of variant identification in a cohort of 20 VEO-IBD patients
A. Identified and validated variants **B.** Patients with identified and validated causal variant.

Patient no.	Dg.	Onset (months)	Identified variant	Mechanism of dysregulation
#2	CD	5	***ERBIN* (p.D147E)**	Regulatory T lymphocytes
#3	UC	8	***MEFV* (p.V67M)**	Systemic inflammatory disorder
#4	CD	9	***IRF2BP2* (p.T394P)** ***PRKDC* (p.R14R + p.A1393A/p.L1393H)**	B lymphocytes and antibodies Effector T lymphocytes
#8	IBDU	24	***PSTPIP1* (p.Y433*)**	Systemic inflammatory disorder
#10	UC	49	***BCL11B* (p.G728S)**	Regulatory T lymphocytes
#11	CD	52	***NLRP3* (p.V200M)**	Phagocytes
#13	UC	58	***FOXP3* (p.H400L)**	Regulatory T lymphocytes
#16	CD	130	***NOD2* (p.A755V)**	Epithelial barrier
#17	IBDU	131	***DUOX2* (p.R1216W + p.A1131T)**	Epithelial barrier
#19	CD	EOIBD (NA)	***PLCG2* (p.G699S)**	Systemic inflammatory disorder

Tab. 8: Identified variants in a cohort of 20 VEO-IBD patients (confirmed by Sanger sequencing and segregation analysis in the family - **green**; declined by segregation analysis in the family - **red**; suggested, but not validated yet by Sanger sequencing - **blue**)

The overall success rate of identification and validation of the causal variants by whole-exome and Sanger sequencing was 20 % (4 patients out of 20, as shown in Fig. 3B). Focusing on the mechanism of the pathology, we found 2 defects of epithelial barrier (*NOD2*, *DUOX2*) 1 defect of phagocytes (*NLRP3*) and 1 defect of regulatory T cells (*FOXP3*). Detailed reports of the patients' cases are presented in following section 6.*1.1* up to 6.*1.10*.

6.1.1. Patient #2

Patient #2 was a boy diagnosed with CD with no history of IBD in the family. The parents were healthy apart from father's allergy to cow milk. Patient's first symptoms consisting of blood in stool occurred when he was 5 months of age followed by failure to thrive, fever episodes, thrush and hypochromic microcytic anemia. The blood work showed mildly elevated white blood cells, especially monocytes and neutrophils. Later, colonoscopy revealed Crohn-like colitis in his intestines and based on that he was subjected to anti-TNF biological treatment of the IBD.

The WES revealed several variants in genes which are in literature connected to IBD or PID, however, none of them was found potentially causal based on parameters such as frequency in population and inheritance mode. The only suspected variant was found in gene *ERBIN* (p.D147E, predicted loss-of-function, further parameters of the variant are available in Supplementary file 2).

ERBIN, also known as ERBB2 interacting protein (ERBB2IP), is involved in balancing the signaling between STAT3 and Smad3 which is crucial for example for TGF- β expression (Lyons *et al*., 2017). No variant in ERBIN was previously reported as causal in a clinical case of IBD. However, Lyons *et al*. (2015) described a phenotypic convergence of *STAT3* and *ERBIN* loss-of-function heterozygous mutations. Therefore, we hypothesized that defects in ERBIN could have similar clinical phenotype as STAT3-related immunity dysregulations already known for its associated with IBD.

Interestingly, STAT3-dependent immune dysregulation can occur in case of loos-of-function (Minegishi *et al*., 2007) as well as gain-of-function mutations (Flanagan *et al*., 2015). STAT3 or ERBIN inactivating mutations can lead to enhanced production of TGF-β, expression of FOXP3 and elevated counts of Treg. Consequently, the immunopathological phenotype of inactivating mutations involves Th2-mediated atopy and hyper IgE syndrome (OMIM #147060; Lyons *et al*., 2017). On the other hand, the activating mutations are connected to autoimmunity which can manifest by enteropathy due to T reg deficiency and preferential shift to Th1 inflammatory response (OMIM #615952; Flanagan *et al*., 2015). Based on these findings connecting ERBIN indirectly to VEO-IBD and the parameters of the patient's variant, *ERBIN* p.D147E was submitted for Sanger sequencing validation.

Several pairs of primers had to be designed (Supplementary file 3) while optimizing PCR of the mutated site as we have first time unintentionally covered SNP present in

the mother's and the patient's primer binding site. This polymorphism made it complicated to optimize the reaction for all three samples at once resulting in making two final sets of primers and sequencing the patient and the mother with SNP-specific primers and the patient and the father with reference-specific primers.

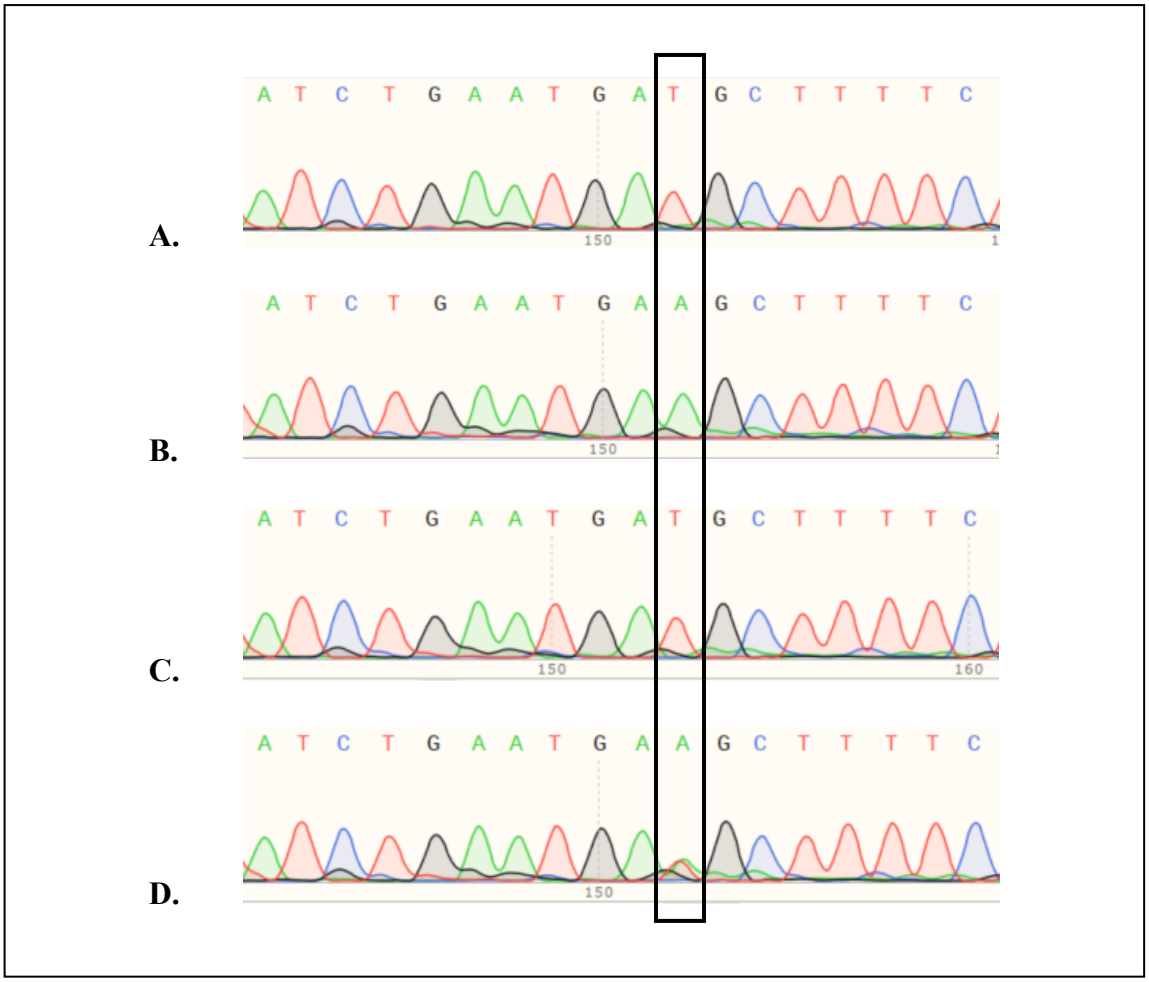


Fig. 4: Sanger sequencing of *ERBIN* in case of Patient #2 - variant p.D147E : wt allele is T, mut allele is A. **A.** Patient (wt allele segregated from mother, sequencing performed with SNP-specific reverse primer), **B.** Patient (mut allele segregated from father, sequencing performed with reference-specific reverse primer) **C.** Mother (wt/wt, sequencing performed with SNP-specific reverse primer) **D.** Father (mut/wt, sequencing performed with reference-specific reverse primer).

Having successfully Sanger sequenced the patient (Fig. 4A and 4B), the mother (Fig. 4C) and the father (Fig. 4D), segregation of the alleles was evaluated finding a discrepancy in predicted causality of the variant. We found that the father was also a heterozygous carrier of the same mutation as the patient, however, did not experience any IBD symptoms. Therefore, we declined the proposed causality of the *ERBIN* variant in the case of Patient #2. Apart from *ERBIN* (p.D147E), no other variant has been selected for validation despite numerous re-analyzations of the WES data, so far.

6.1.2. Patient #3

Patient #3 was a boy diagnosed with UC with no history of IBD or any other related disease in the family. Since he was 8 months old, he has been experiencing blood in the stool. He also suffered from frequent respiratory infections and was found to be anemic and had enlarged spleen. Colonoscopy revealed inflamed tissues in the gastrointestinal tract. His blood work showed no discrepancies apart from elevated levels of calprotectin. Consequently, he was subjected to biological treatment by anti-TNF-based therapy in addition to immunosuppressants.

When analyzing WES data of Patient #3, heterozygous loss-of-function predicted damaging variant in gene *MEFV* (sometimes also called pyrin) p.V67M was found (further parameters of the variant are available in Supplementary file 2). Mutations in *MEFV* cause different types of Familial Mediterranean fever characterized by episodes of fever and abdominal pain due to local inflammation in the gut.

MEFV is involved in regulation of caspase 1 and inflammasome activity. The molecular mechanism of MEFV in the pathogenesis of IBD was previously evaluated in a a mouse model. *Mefv*-/- knockout mice with induced colitis presented with highly expressed *Mefv* in the inflamed gut and, consequently, the levels of inflammatory cytokines such as IL6 were significantly increased. However, IL18 levels dependent on inflammasome activation were decreased proving *MEFV* mutations interference with inflammasome activity in connection to inflammatory bowel disease (Sharma *et al*., 2018). Moreover, there are also available evidence stemming from clinical cases that VEO-IBD can be caused by a considerable number of *MEFV* variants (Uslu *et al*., 2010, Saito at al., 2019).

Based on all the experimental and clinical reports, the validation of the *MEFV* variant in Patient #3 was proposed. This finding would be particularly important for patient's treatment as symptomatic patients with confirmed *MEFV* mutations can be subjected to specialized treatment by colchicine, or, if unresponsive, to IL1 receptor agonist (Roldan *et al*., 2008). The primers for Sanger sequencing of the patient's variant were designed (Supplementary file 3). Regrettably, DNA of the parents has not been collected so far for performing any further tests.

6.1.3. Patient #4

Patient #4 was a boy diagnosed with CD and born to healthy parents with no family history of IBD or immunity-related disorders. He started to experience health issues when he was 3 months old beginning with allergy to cow milk and continuing with intolerance for meat and eggs. At 9 months, he was hospitalized because of respiratory tract infection and enterorrhea leading to suspicion of IBD. Because of the bleeding, he was anemic in contrast with found thrombocytosis and leukocytosis. Immunological screening reveled hypogammaglobulinemia, lower levels of IgA, elevated levels of calprotectin, but no presence of autoantibodies. He was subjected to colonoscopy with tissue histology which confirmed the diagnosis of IBD. Consequently, he underwent several different therapies including mild and strong immunosuppressants and anti-TNF biological treatment with poor reaction and ongoing relapse.

After WES of the Patient #4, there were several potentially deleterious variants found. Firstly, we identified a variant in *IRF2BP2* (p.T394P) with low abundance in the population databases (further parameters of the variant are available in Supplementary file 2). Heterozygous mutation of *IRF2BP2* can cause CVID (OMIM #617765) with hypogammaglobulinemia and chronic diarrhea (Keller *et al*., 2016). Concerning the molecular mechanism, the pathology is caused by defects in IFN signaling as IRF2BP2 is a binding protein and transcriptional corepressor of interferon regulator factor 2 (Childs & Goodburn, 2003).

Bearing in mind the IRF2BP2-caused CVID reports, the variant *IRF2BP2* (p.T394P) was validated by Sanger sequencing. The sequence included, however, a considerable number of repetitive sequences and was very GC-rich. Because of that, it was complicated to amplify the sequence and we had to try several primer pairs as well as different temperature profiles of PCR. The optimal conditions for the PCR became to be the usage of higher magnesium concentration (instead of 1,5 µl we used 2 µl of 25 mM $MgCl_2$) and lower temperature of annealing (instead of 62 °C we used 60 °C). The sequencing (Fig. 5), however, declined the suspected variant in gene *IRF2BP2* as causal since the father of the patient was also heterozygote for the variant (Fig. 5C), but did not experience any IBD- or severe immunity-related symptoms.

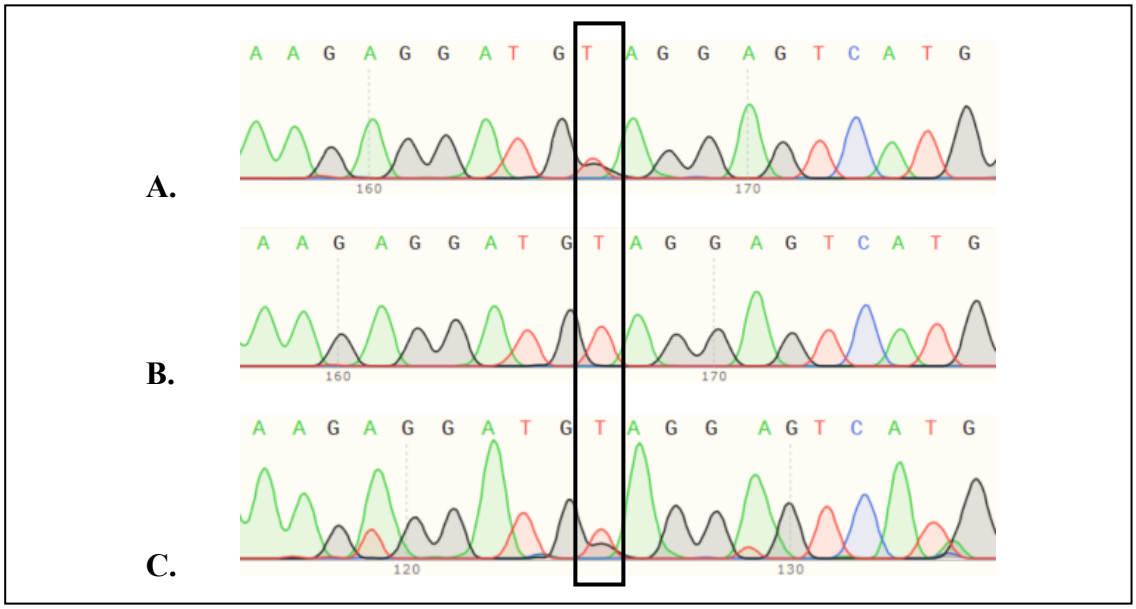


Fig. 5: Sanger sequencing of *IRF2BP2* in case of Patient #4 - variant p.T394P : wt allele is T, mut allele is G. **A.** Patient (mut/wt), **B.** Mother (wt/wt), **C.** Father (mut/wt).

Re-analysis of the WES data revealed another potentially deleterious variant leading to patient's IBD - a compound heterozygous mutation in gene *PRKDC* (p.R14R + p.A1393A/p.L1393H). p.A1393A/p.L1393H is predicted to cause a loss of function of the protein and both of the variants are sporadic in the population. More details about the variants are available in Supplementary file 3.

PRKDC encodes catalytic subunit of DNA-dependent protein kinase involved in non-homologous end joining including VDJ recombination. Because of affected VDJ recombination, the patients present with T-B-NK+ SCID and, thus, are susceptible to infections, fail to thrive and can also have chronic diarrhea. The counts and function of the cells may vary significantly between the patients. Typical is limited diversity of the lymphocytes and antibodies. Mouse model data suggest, as well, involvement of PRKDC in AIRE-dependent processes in thymus, thus, patients with PRKDC-SCID may suffer from autoimmunity-related disorders and test positive for autoantibodies (Abramson *et al*., 2010b; Mathieu *et al*., 2015).

Based on the reports and presence of two variants possibly making together a compound heterozygous mutation, the validation of the *PRKDC* variant in Patient #4 was proposed. The variant was, however, found during reanalysis of the data only recently and has not been validated yet.

6.1.4. Patient #8

Patient #8 was a boy diagnosed with IBD-U with no history of IBD or any other related disease in the family. Since the age of 6 months he has been experiencing episodic diarrhea and from the age of 2 years, there was blood present in his stool. Because of that, he underwent a colonoscopy finding inflammatory changes in the gut. Consequently, he was subjected to corticotherapy and anti-inflammatory treatment.

Having sequenced whole exome of Patient #8, we found a potentially causal variant in *NLRP3*. *NLRP3* is known for its association with several immunodeficiency syndromes and was also considered in connection to IBD, however, was not reported yet as causal in a monogenic way (Mao *et al*., 2018). Apart from *NLRP3*, the Patient #8 has also a potentially pathogenic variant in *PSTPIP1* (p.Y433*, stop gained, further parameters of the variant are available in Supplementary file 2).

PSTPIP1 is proline/serine/threonine phosphatase-interacting protein which is known to cause autosomal dominant PAPA (Pyogenic sterile arthritis, pyoderma gangrenosum and acne) syndrome (OMIM #604416), sometimes also called PAC syndrome (Pyoderma gangrenosum, Acne and ulcerative Colitis). The syndrome predominantly manifests by several disorders of the skin, however, there are also cases of patients with *PSTPIP1* mutations reporting hematological disorder, arthritis or colitis with very early onset (Burlakov *et al*., 2018). PSTPIP1 function is closely connected to MEFV (pyrin) and the proposed mechanism of PSTPIP1 dysfunction is related to pyrin-regulated-inflammasome activity (Waite *et al*., 2009).

Based on the experimental and clinical reports, the validation of the *PSTPIP1* variant in Patient #3 was proposed. The primers for Sanger sequencing of the patient's variant were designed (Supplementary file 3). Regrettably, DNA of the parents has not been collected so far for performing any further tests.

6.1.5. Patient #10

Patient #10 was a girl diagnosed with UC and born to healthy parents. She suffered more frequently from middle ear infections, cataracts and gastroenteritis. She was indicated for further testing in the hospital after one of the gastroenteritis revealing elevated calprotectin levels and presence of pANCA antibodies. Moreover, colonoscopy was performed finding inflammatory deposits in the gut leading to IBD diagnosis. She was also positive for *Clostridium difficile* and because of suspected intestinal dysmicrobia she underwent a fecal transplantation from her mother. During her long-term treatment, the patient did not respond to the conventional immunosuppressive treatment and was chronically relapsing. For so, she was subjected to intensive biological treatment of several kinds including strong immunosuppressants and anti-TNF-based treatment. Currently, she is responding to a biological treatment on an anti-integrin base.

In the WES analysis, no potentially causal variant in a previously IBD-associated gene was found. However, we found a variant in *BCL11B* (p.G728S) of an uncertain significance and unknown frequency within the population. Additional predictive analysis in MutationTaster evaluated this variant as disease causing. Further parameters of the variant are available in Supplementary file 2.

BCL11B mutations are known to cause T-B+NK+ SCID with autosomal dominant inheritance (OMIM #617237) and developmental disorders. What's more, there is a mouse model available testing the function of *Bcl11b* in IBD. Bcl11b was found to cause IBD by immunity dysregulation when removed from T cells during their development. Specifically, IBD occurs due to reduced suppressive activity of the Treg cells by having down-regulated *Foxp3* and *IL10* and, on the other hand, increased expression and production of proinflammatory cytokines (VanValkenburgh *et al*., 2011).

Based on potential severe impact of the variant and mouse model experiments, we subjected the variant *BCL11B* (p.G728S) for validation and further functional testing as there are no clinical reports available describing *BCL11B* mutation in monogenic IBD. The primers for Sanger sequencing of the patient's variant were designed (Supplementary file 3). Regrettably, DNA of the parents has not been collected so far for performing any further tests.

6.1.6. Patient #11

Patient #11 was a girl diagnosed with CD. Her father was under immunological observation due to recurrent infections. Her mother was healthy, however, her maternal grandmother suffered from chronic colitis and diarrhea, ulceration in the stomach and intestine, chronic polyarthritis, allergy and asthma. Her older brother was also monitored because of constipation. The Patient #11 suffered repeatedly from respiratory and middle ear infections during her childhood. At the age of 1 year, she was hospitalized due to gastroenteritis and enterorrhea. At the age of 4 years, she was diagnosed with IBD-U, which was later reclassified as CD. The colonoscopy confirmed the diagnosis of intestinal inflammation, later adding inflammatory changes also in the upper gastrointestinal tract. The laboratory tests showed elevated levels of calprotectin and hypochromic microcytic anemia. Moreover, growth retardation, pain and swelling of the joints and dermatitis were observed suggesting systemic inflammatory disorder. The patient was, however, negative for autoantibodies and did not have a known HLA-B27 haplotype associated with suspected spondyloarthropathy. Because of her severe symptoms, the Patient #11 was subjected to biological treatment by chimeric (with side effects) as well as human recombinant (with better reaction) anti-TNF therapeutics.

Based on the family history and patient's anamnesis, WES was performed. Several non-frequent genetic variants known for its autosomal recessive inheritance were found - namely mutations in *JAK1*, *CASP8*, *TLR3* and *TICAM1*. Apart from those, a predicted gain-of-function variant in gene *NLRP3* (p.V200M) was identified. The variant is not rare in the population, however, the SNP is classified as pathogenic and also MutationTaster predicted it to be disease causing (more details about the NLRP3 variant can be found in Supplementary file 2). What's more, the mutation has already been reported in case report symptomatically resembling Patient #11 and diagnosed with Muckle-Wells syndrome (OMIM #191900) manifesting by episodic fever, rash and pain in the joints (Aróstégui *et al*., 2004).

NLRP3 (also known as cryopyrin) is a protein involved in NLRP3 inflammasome, activation of which leads to production of proinflammatory cytokines. NLRP3 plays a crucial role in mucosal defense against bacterial RNA and its defects can lead to impaired production of IL1β and IL18 (Kanneganti *et al*., 2006), thus, shaping the composition of microbiota and shifting homeostasis in the gut (Hirota *et al*., 2011).

NLRP3 expression is upregulated in patients with active IBD (Ranson *et al*., 2019) and its mutations are associated with IBD pathogenesis (Vilani *et al*., 2009). NLRP3 was not, however, reported yet as causal in case of monogenic IBD (Mao *et al*., 2018). Based on other inflammatory findings in concordance with reports about the variant (Aróstégui *et al*., 2004) and indirect information about IBD significance, we decided to validate the variant *NLRP3* (p.V200M) as causal for patient's immunity-related issues.

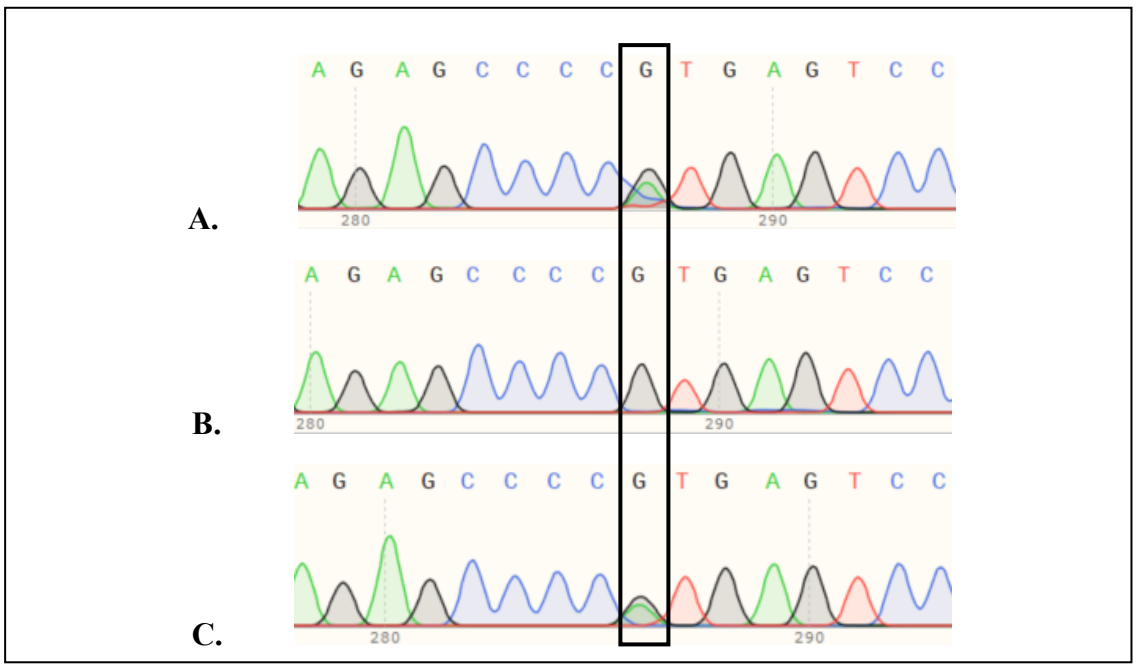


Fig. 6: Sanger sequencing of *NLRP3* in case of Patient #11 - variant p.V200M: wt allele is G, mut allele is A. **A.** Patient (mut/wt), **B.** Mother (wt/wt), **C.** Father (mut/wt).

We amplified the sequence of gene *NLRP3* including suspected variant p.V200M using customized primers (as stated in Supplementary file 3). The subsequent Sanger sequencing revealed a presence of the validated variant in the patient (Fig. 6A) and her father suffering from immunity-related health issues (Fig. 6C). The mother was a healthy wild type homozygote (Fig. 6B). These results, therefore, showed, that *NLRP3* is not the only cause of the pathology in case of Patient #11 as the father presents with significantly less severe symptoms.

Based on the Sanger sequencing, we propose that *NLRP3* (p.V200M) is contributing to Patient's #11 immunity-related pathology with incomplete penetrance, but the cause of her VEO-IBD is multifactorial. The low penetrance of the variant is in concordance with reports (Aróstégui *et al*., 2004) as the father of the reported patient was also an asymptomatic carrier.

Unfortunately, there are no information about IBD in the family of the reported patient with the same variant, and, for so, we cannot clearly state p.V200M as causal for the Crohn's disease in Patient #11. Nevertheless, the NLRP3-related immune dysregulation is influencing the homeostasis of the whole body and can play an important role in patient's intestinal autoimmunity. Bearing in mind the colitis of the patient's grandmother, there are probably additional genetical and non-genetical factors involved, such as diet, housing and other life-style factors (Piovani *et al*., 2019) which are shared in a family and contribute to her IBD development and progression.

6.1.7. Patient #13

Patient #13 was a boy diagnosed with UC and born to healthy parents. His younger sister was also healthy, however, his grandparents had diagnosed several immunity-related disorders. His maternal grandfather suffered from immunity dysregulation and had celiac disease and his paternal grandmother had psoriasis. He has experienced some health issues since the time he was 3 months of age. Starting with atopic dermatitis and suspected allergy to cow milk, he suffered from respiratory tract and middle ear infections more often during his early childhood. What's more, allergies to several allergens such as pollen, animal fur and food like eggs, nuts, beef and soya were found out. He was also experiencing a bloody diarrhea and the laboratory tests revealed elevated levels of calprotectin suggesting pediatric IBD. The blood count was normal except for elevated levels of monocytes. Moreover, at the age of 5, he was notthriving and was diagnosed with celiac disease based on laboratory confirmation. Due to several immunological disorders, he was subjected to immunosuppressive treatment including systemic corticosteroids.

The Patient #13 was indicated for WES, where we found an unknown variant in gene *FOXP3*. The variant p.H400L was predicted to be damaging and causing loss of function of the protein (further parameters of the variant are available in Supplementary file 2). As *FOXP3* is localized on chromosome X, the patient was a hemizygote for this variant and it was expected to be of a dominant-negative effect on the protein function.

FOXP3 is a transcription factor involved in development of Treg (CD4+ CD25+ FOXP3+ cells). Its impact can be studied on a mouse model called Scurfy which is exhibiting similar disorders as patients with FOXP3-based immunodeficiency. The mouse phenotype stems from a lack of functional regulatory T cells, which are not developing physiologically without *FOXP3* expression (Fontenot *et al*., 2003). *FOXP3* mutations cause a severe immunity-related syndrome called IPEX (immunodysregulation, polyendocrinopathy, and enteropathy, X-linked, OMIM #304790) which manifests by multiple autoimmune disorders including chronic diarrhea, eczema, hematological disorders and disorders of endocrine glands such as pancreas and thyroid gland. Patients with IPEX can have either significantly decreased Treg numbers because of the *FOXP3* mutation (Agakidis *et al*., 2019) or the disorder might stem from truncated-FOXP3-mediated functional dysregulation with normal levels of Treg (Okou *et al*., 2014). The patients with IPEX are at risk of premature death

in early childhood. Therefore, it was necessary to validate the variant in Patient #13 quickly to avoid progression to some life-threatening health issue.

Using primers designed for Sanger sequencing of the patient's variant (Supplementary file 3), we prepared DNA samples of the patient, his mother, his father and, additionally, also his younger sister for the validation. The Sanger sequencing confirmed the expected allelic segregation. The patient (Fig. 7A) was a hemizygote for the tested variant. Patient's father (Fig. 7B) was a hemizygote for the wild type allele. Patient's sister (Fig. 7C) was a heterozygous carrier of the mutation. The sample of the patient's mother could not be validated due to unspecific amplification during the PCR. Nonetheless, the sister's genotype leads us to a conclusion that the mother is also a heterozygous carrier of the mutation (mut/wt).

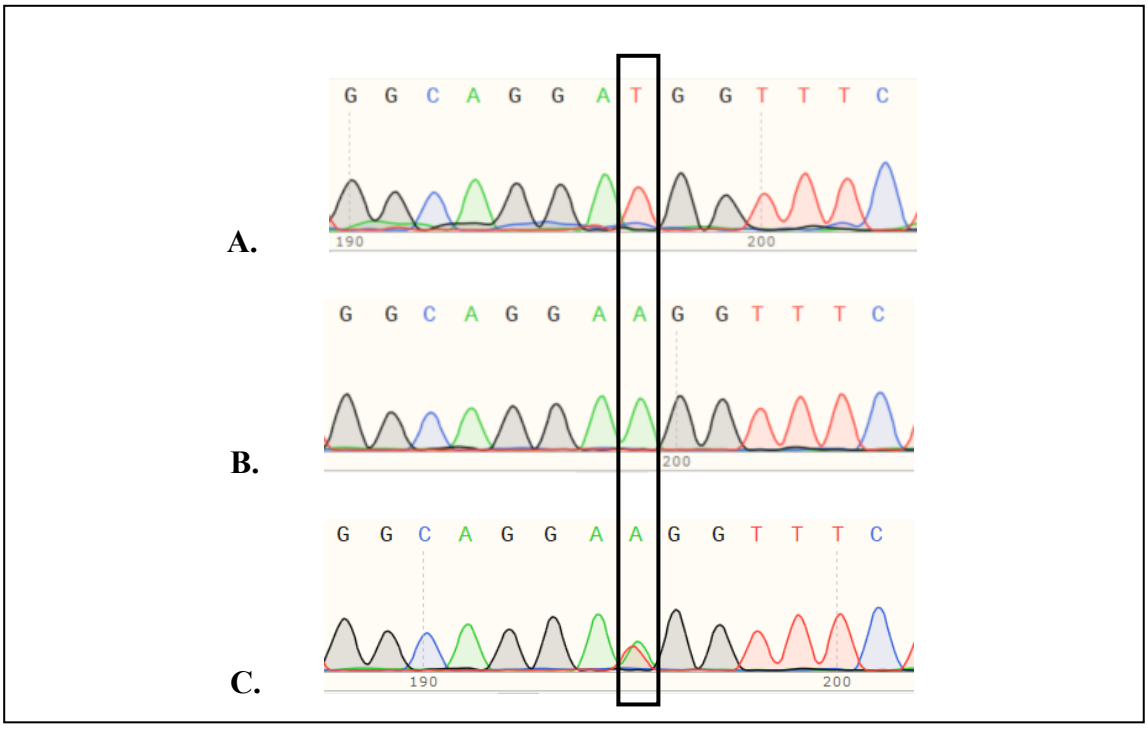


Fig. 7: Sanger sequencing of *FOXP3* in case of Patient #13 - variant p.H400L: wt allele is A, mut allele is T. **A.** Patient (mut/-), **B.** Father (wt/-), **C.** Sister (mut/wt).

Based on the positive confirmation of the variant by Sanger sequencing, the samples and findings were forwarded to Laboratory of Flow Cytometry of Department of CLIP for immunophenotypization and further functional testing. The variant *FOXP3* (p.H400L) has not been reported as causal in the literature so far and, thus, its causality needed to be evaluated more in detail.

6.1.8. Patient #16

Patient #16 was a boy diagnosed with CD. His mother was treated for epilepsy. His father, however, was also diagnosed with CD at the age of 18 years as well as his paternal grandmother. The patient was diagnosed with CD at the age of 10 years, later he had to undergo ileocecal resection, ileo- and colostomy. The colonoscopy revealed inflammatory changes in the intestine, as well as the stomach. He was subjected to immunosuppressive and biological treatment by anti-TNF and anti-p40 (shared subunit of IL12 and IL23) with recurrent relapses.

Despite later onset of the disease, the patient was indicated to WES. When analyzing patient's WES data, we found variants in a few genes connected to primary immunodeficiencies such as *MPO*, *TYK2*, *CD79A*, which cause, however, autosomal recessive disorders. We found also a variant in the gene canonically known in association with IBD - *NOD2* (Hugot *et al*., 2001).

NOD2 (Nucleotide-binding oligomerization domain protein 2) is a pattern recognition receptor associated, apart from IBD, with Blau syndrome. NOD2 is involved in pathogen sensing and activation of immune response for the pathogen clearance in the gut. In case of the mutation, however, the clearance ability is impaired leading to dysbiosis and invasion of the bacteria through the intestinal barrier (Mondot *et al*., 2012). Therefore, the innate immune system is chronically activated on site leading to differentiation of adaptive cells preferentially to Th1 and Th17 branch which, then, contribute to the inflammatory phenotype of the disorder (Podolsky, 2002; Sidiq *et al*., 2016).

Specifically, the patient's variant in *NOD2* was p.A755V (also previously reported as p.A728V; more details about this variant are available in Supplementary File 2), which has already been reported in literature in patients with IBD (Hugot *et al*., 2001). Based on family history of IBD, we decided to validate this variant expecting the same variant present in the father.

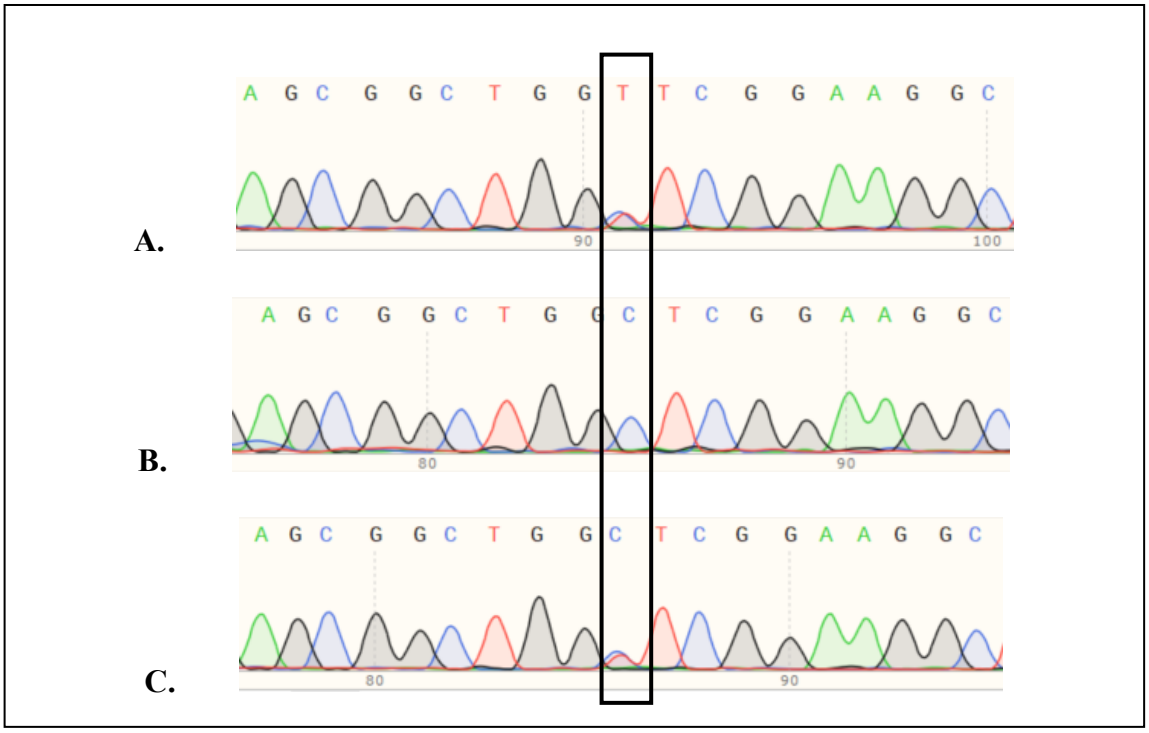


Fig. 8: Sanger sequencing of *NOD2* in case of Patient #16 - variant p.A755V: wt allele is C, mut allele is T. **A.** Patient (mut/wt), **B.** Mother (wt/wt, **C.** Father (mut/wt).

Having designed primers (Supplementary file 3) for the patient's mutated region of *NOD2*, we amplified it and got it sequenced. The Sanger sequencing confirmed our hypothesis as the variant was detected in patient (Fig. 8A) and his father (Fig. 8C). Patient's mother, however, was a homozygote for the wild type allele (Fig. 8B). Since the variant has already been reported in literature (Hugot *et al*., 2001), we could state it as causal in case of Patient #16 and, as well, in case of his father. Bearing in mind the later onset of the disorder and different severity of the phenotype in the patient, his father and probably also his grandmother, the disease is presumable not monogenic in this case. Some environmental factors as described in Piovani *et al*. (2019) are most likely involved in the pathogenesis of IBD in this family.

6.1.9. Patient #17

Patient #17 was a girl diagnosed with IBD-U. Her mother was healthy, but her father also suffered from Crohn's disease and was at that time in complete remission. She has been experiencing atopic eczema since infant period. Later, she also developed asthma and polyvalent allergy. At the age of 11 years, she was diagnosed with IBD based on colonoscopy and histology findings. Multiple laboratory tests were performed revealing thrombocytosis and elevated levels of calprotectin. The bacteriological and serological screening also found presence of opportunistic pathogens *Clostridium difficile* and *Salmonella enteritidis*. In addition to colitis, she was repeatedly hospitalized for pancreatitis and cholecystolithiasis presumably due to irritation from the anti-inflammatory medication. Moreover when evaluating her thriving, she had abnormally short stature suspectedly caused by long-time corticotherapy. Apart from corticotherapy, her inflammatory disease was also treated by other strong immunosuppressants and by biological treatment consisting of anti-integrin and anti-TNF therapeutics.

The inflammatory phenotype with pediatric onset indicated Patient #17 for WES genomic analysis looking for a hereditary cause of her pathology. Analyzing WES data, we found several variants associated with PID and IBD, out of those, however, a duo of *DUOX2* variants (p.R1216W + p.A1131T) stood out. The variants were predicted to be loss-of-function, possibly/probably damaging with high CADD scores around 29 and low frequency in population (more details about the variants are available in Supplementary file 2).

DUOX2 is an oxidase highly expressed in thymocytes and mucosal cells. Thus, its mutation can lead to congenital hypothyroidism (Vigone *et al*., 2005) and immunity dysregulation manifesting in the mucosa of the gut (Hayes *et al*., 2015). The impaired intestinal immunity is related to intestinal barrier function and host-microbiota interaction (Grasberger *et al*., 2015) as ROS produced by DUOX2 are associated with NOD2 immune response and proper pathogen clearance keeping the intestinal barrier intact (Lipinski *et al*., 2009). When the ROS production is decreased by inactivating mutation in *DUOX2*, the microbiota can, consequently, cross the intestinal barrier and induce local inflammation leading to VEO-IBD (Hayes *et al*., 2015).

Interestingly, congenital hypothyroidism caused by *DUOX2* mutation is manifesting with autosomal recessive effect (Vigone *et al*., 2005). On the other hand, reports of DUOX2-related VEO-IBD suggests autosomal dominant mode of inheritance (Hayes *et al*., 2015). Nonetheless the dominancy of the mutation, DUOX2 dysregulation can explain the observed susceptibility to IBD in patients with congenital hypothyroidism (Grasberger *et al*., 2015) as DUOX2 regulates both the intestinal and thyroid homeostasis and physiological functions.

The primers were designed for Sanger sequencing of the patient's variants p.R1216W in exon 28 and p.A1131T in exon 25 (Supplementary file 3). Optimizing PCR amplification for two *DUOX2* variants in Patient #17, we had to alter the temperature profile when amplifying p.A1131T to get clear results in Sanger sequencing. Specifically, the annealing temperature was lowered to 60 °C.

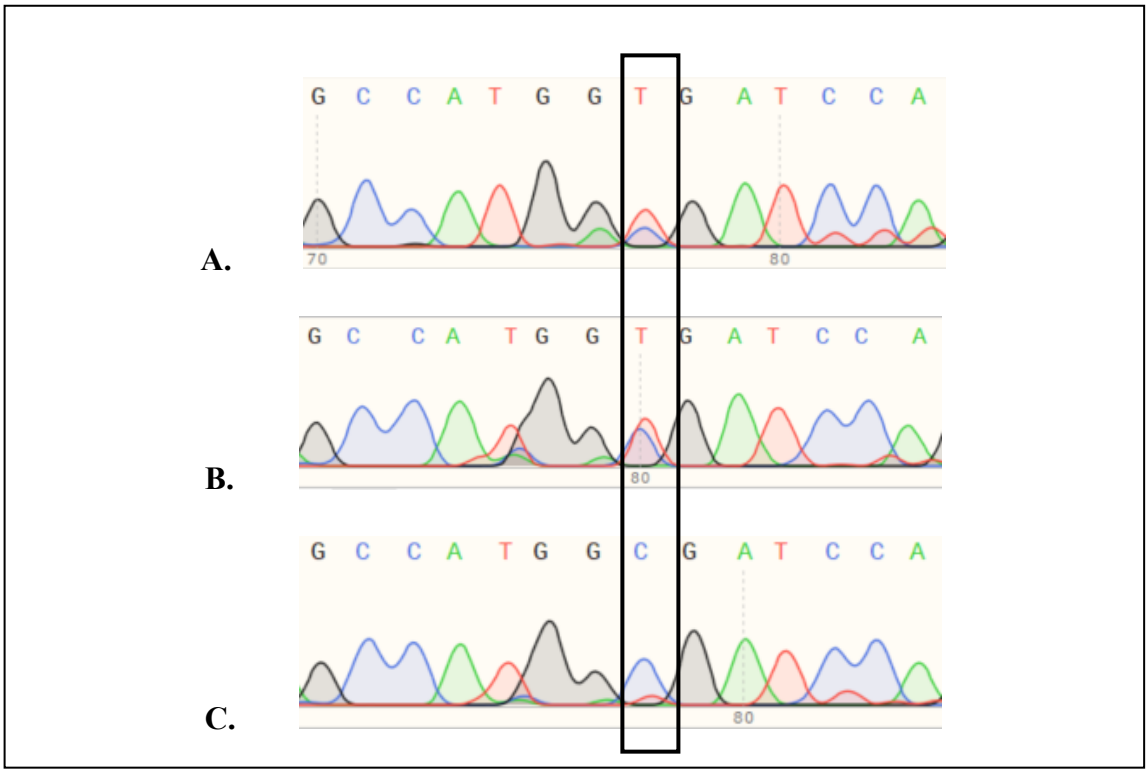


Fig. 9: Sanger sequencing of *DUOX2* in case of Patient #17 - variant p.A1131T: wt allele is T, mut allele is C. **A.** Patient (mut/wt), **B.** Mother (mut/wt), **C.** Father (mut/wt).

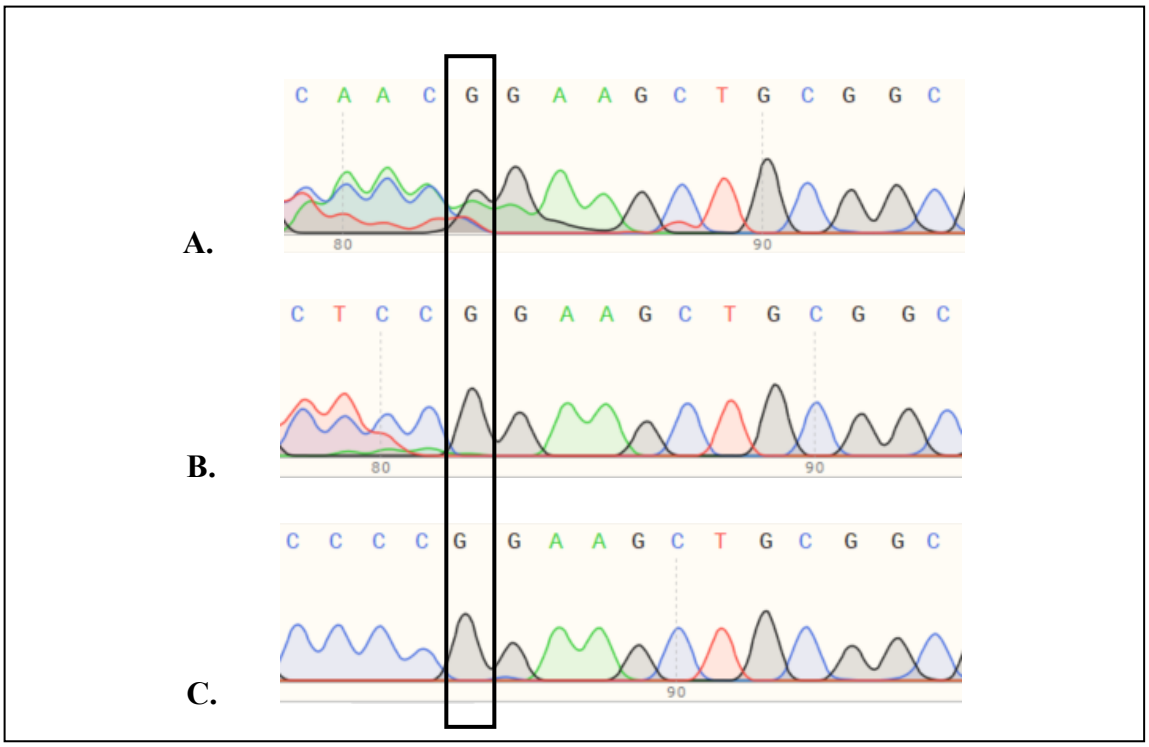


Fig. 10: Sanger sequencing of *DUOX2* in case of Patient #17 - variant p.R1216W: wt allele is G, mut allele is A. **A.** Patient (mut/wt), **B.** Mother (wt/wt), **C.** Father (wt/wt).

We validate the *DUOX2* variants in the patient, her mother and her father. We presumed the patient was a compound heterozygote and, therefore, we expected the parents to be heterozygotes for one of the mutations. The mother (Fig. 9B and 10B) and father (Fig. 9C and 10C), however, were both heterozygotes for the same variant *DUOX2* (p.A1131T). The other allele encoding for protein variant p.R1216W was a *de novo* mutation present only in the patient's sample (Fig. 9A and 10A).

The Sanger sequencing results confirmed the rationale of the patient's compound heterozygosity. The functional impact of the variants *DUOX2* (p.R1216W) and *DUOX2* (p.A1131T), however, needed to be examined further. Especially variant p.R1216W needed additional assessment as it was *de novo* mutation present only in the patients. For so, we forwarded the results to Laboratory of Flow Cytometry of Department of CLIP for immunophenotypization and further functional testing.

6.1.10. Patient #19

Patient #19 was a girl diagnosed with CD. She was born to healthy parents with no family anamnesis of IBD or any other related disease. Apart from IBD, she presented with pain in hip and knee suggesting also other disorders such as chronic recurrent multifocal osteomyelitis, vasculitis or sarcoidosis. Multiple laboratory tests were made showing thrombocytosis, mild anemia, slightly elevated IgM and anti-*Saccharomyces cerevisiae* IgG, significantly elevated levels of CRP and calprotectin. Based on her complex inflammatory phenotype, she was subjected to immunosuppressive treatment and since it did not ameliorate her IBD, she was additionally subjected also to biological treatment by anti-TNF therapeutics.

Having complex inflammatory phenotype with early onset IBD, the Patient #19 was indicated for WES. During the WES data analysis, we found several variants in genes previously described in connection to autosomal dominant diseases, such as *SAMD9*, *STAT3* and *PLCG2*. Out of those, however, only *PLCG2* variant p.G699S matched not only in a biological context, but also in parameters of the variants. The *PLCG2* p.G699S is a missense variant predicted to be probably damaging by PolyPhen-2 and disease casing by MutationTaster. Moreover, its frequency in the population was very low (more details about the variant are available in Supplementary file 2) suggesting its potential causality for the patient's phenotype.

PLCG2 is a gene encoding phospholipase C which is involved in phosphatidylinositol pathway. This signaling is particularly important for B lymphocytes where PLCG2 is involved in the cellular development and BCR down-stream activation. Consequently, the numbers of mature B cells can be decreased and its activation in reaction to antigens defective (Hashimoto *et al*., 2000). When PLCG2 is dysfunctional, autoinflammation, antibody deficiency, and immune dysregulation (APLAID, OMIM #614878) can occur often manifesting with inflammatory infiltrates in skin, joints, eyes and gastrointestinal tract (Zhou *et al*., 2012). The onset of the disease can be early in childhood and the immune dysregulation can manifest predominantly as IBD (Neves *et al*., 2018).

The pathology of the variant, however, was not clear as Zhou *et al*. (2012) and Neves *et al*. (2018) reported variants with hypermorphic/gain-of-function effect, but the prediction for the variant p.G699S was not stating gain of function. The localization of the variant was, however, very close to reported deleterious variant p.S707Y and they

were both localized in the same functional domain of the protein C-terminal Src homology domain suggesting similar impact of the variants.

Based on the reports about PLCG2 dysregulation - mainly due to involvement of gastrointestinal tract and joints matching the Patient #19 case, the variant *PLCG2* p.G699S was chosen for validation. The primers for Sanger sequencing of the patient's variant were designed (Supplementary file 3). Regrettably, DNA of the parents has not been collected so far for performing any further tests.

6.2. Evaluation of different bioinformatical approaches

6.2.1. Comparison of bioinformatical pipelines

Apart from comparison and evaluation of different gene lists, we also tested two bioinformatical pipelines for raw sequencing data processing. We aimed to assess the functional differences and verify the use of new GATK4-based pipeline for the variant analysis. We analyzed 38 whole-exome sequencing data sets coming from 19 patients (as shown in Tab. 9).

Using the exported data from Ingenuity variant analysis (filtering criteria used as stated in Tab. 4), the counts of called and filtered variants by both pipelines were assessed. Additionally, we analyzed unique variants called and filtered only by one of the pipelines more in detail. The reasons for the discrepancy were either that the variant was called differently by the pipeline or that the variant was filtered differently by IVA due to quality parameters assigned by the pipeline.

The results show that 725/945 (77 %) variants were called and filtered by both pipelines in the same way. The rest of the variants were found only in one of the pipeline-derived data (Fig. 11). VarScan2 pipeline produced 828 variants, out of which 88 were called uniquely and 15 were filtered uniquely by VarScan2 pipeline. In comparison, GATK4 pipeline produced 842 variants, out of which 61 were called uniquely and 56 were filtered uniquely by GATK4 pipeline.

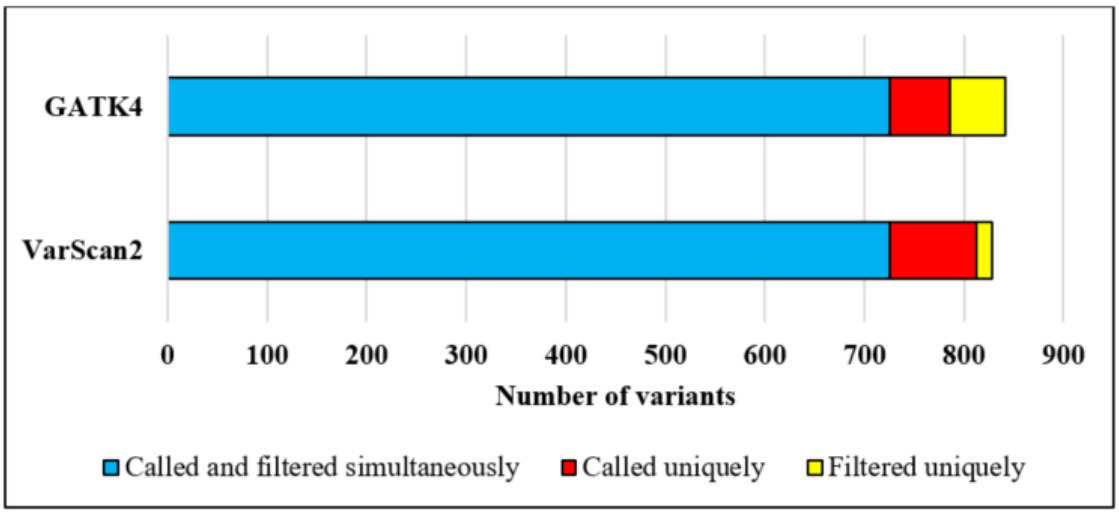


Fig. 11: Differences in variant calling and variant filtering between VarScan2 and GATK4 pipeline divided into three categories: called and filtered simultaneously, called, uniquely and filtered uniquely

In conclusion, VarScan2 generally called variants more widely, while GATK4 in majority of cases considered these variants artefacts of the sequencing. On the other hand, GATK4 scores the variants higher as significant number of them was also called by VarScan2 but was filtered out during the variant analysis due to insufficient quality parameters such as allelic fraction, genotype quality or read depth.

Next we analyzed the sub-counts of differently called and filtered variants by four gene lists (as characterized in section *5.3.2. Variant filtering*). As expected, CDG and PID gene list, which consisted of the highest number of genes reaching over 400, included majority of the variants. On the other hand, less abundant and more specific gene lists IBD and IBD2 differed between the pipeline-derived results only in handful of variants (Fig. 12). Therefore, we found out that IBD-specific gene lists were robust producing very similar results nonetheless the previously used bioinformatical pipeline processing.

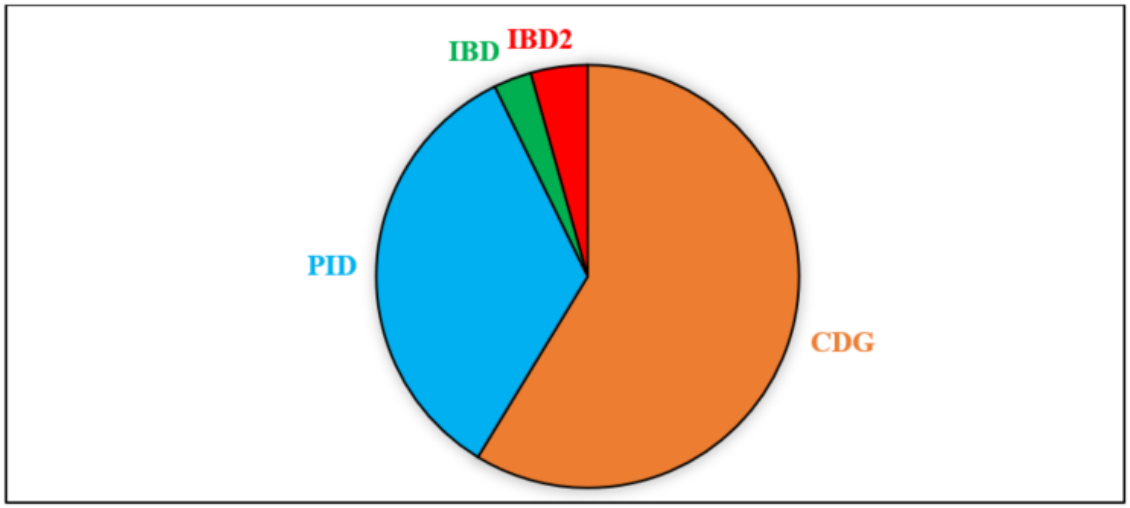


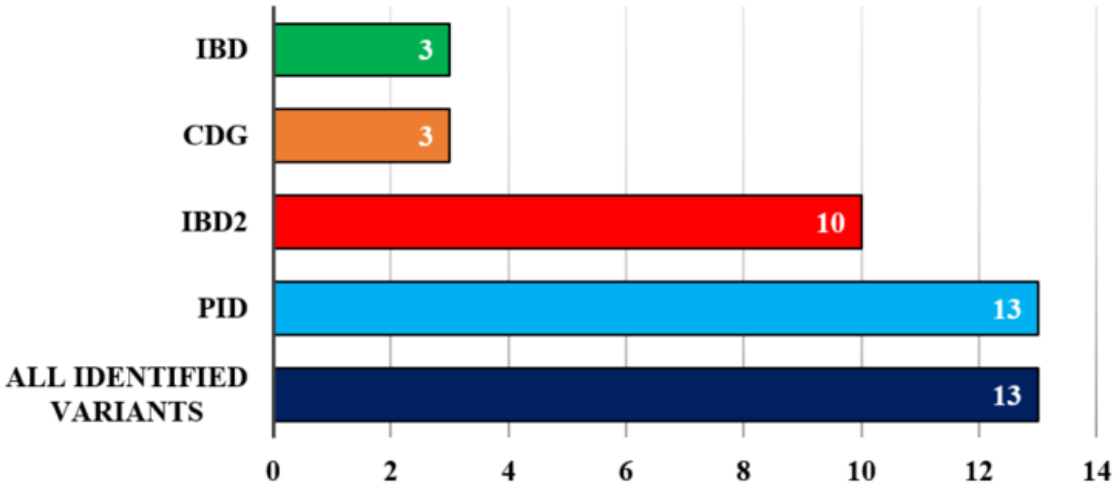
Fig. 12: Distribution of differently called and filtered variants by VarScan-2 and GATK4-based pipelines attributed to four genes lists (CDG, PID, IBD and IBD2) used for variant filtering in IVA

Finally, we evaluated the success rate of the pipelines in identification of the causal variant in our cohort of VEO-IBD patients. All the proposed causal variants were called and filtered by both pipelines - the original VarScan2 as well as by the newly developed GATK4-based pipeline. Thus, the new GATK4 pipeline proved to be suitable for variant analysis and identification.

6.2.2. Comparison of gene lists for variant filtering

Crucial step in the variant analysis was filtering of the variants according to their biological context through four gene lists. The lists were CDG, PID, IBD and IBD2 as described in section *5.3.2. Variant filtering* and in Supplementary file 1. As a result of variant analysis, we identified 13 variants present in 11 genes. The variants were filtered in gene lists with different results providing information for statistical evaluation of the success rate of each list (Tab. 9).

We calculated the percentage of positively filtered variants within each gene list ("virtual gene panel") demonstrating their distinct suitability for the purposes of VEO-IBD genetic analysis (Fig. 13). We found out that all 13 variants (100 %) were filtered in PID gene list. Secondly, IBD2 gene list (customized compilation of the case reports of VEO-IBD) filtered 10/13 (77 %) of the variants, all of which were case-study based. The variants not filtered in IBD2 were either in experimentally reported genes without clinical case reference, so far (ERBIN, BCL11B) or genes associated with IBD, but not in a monogenic way (NLRP3). CDG and IBD gene lists identified only 3/13 (23 %) of the proposed causal variants.



Gene list	CDG	PID	IBD	IBD2
No. of identified variants	3	13	3	10
Rate of positive identification	23 %	100 %	23 %	77 %

Fig. 13: Success rate of four gene list for VEO-IBD variant identification in our cohort

In conclusion, PID gene list is suitable for VEO-IBD genetic analysis of classical monogenic, as well as, specific cases caused by variants in new or generally associated genes. This finding strongly supports the hypothesis of VEO-IBD being an intestinal manifestation of primary immunodeficiency. What's more, IBD2 gene list is also suitable for this analysis, and seems to be relevant in a routine diagnostic way as it filters only variants in genes with clear causal connection to VEO-IBD phenotype. IBD and CDG gene lists both failed to identify majority of the variants proving themselves unsuitable for the genetic analysis of VEO-IBD.

Patient no.	Dg.	Onset (months)	Pipeline	CDG	PID	IBD	IBD2
#1	IBDU	3	VarScan2	N/A	N/A	N/A	N/A
			GAKT4	N/A	N/A	N/A	N/A
#2	CD	5	VarScan2		*ERBIN* p.D147E		
			GAKT4		*ERBIN* p.D147E		
#3	UC	8	VarScan2	*MEFV* p.V67M	*MEFV* p.V67M	*MEFV* p.V67M	*MEFV* p.V67M
			GAKT4	*MEFV* p.V67M	*MEFV* p.V67M	*MEFV* p.V67M	*MEFV* p.V67M
#4	CD	9	VarScan2		*IRF2BP2* p.T394P *PRKDC* p.R14R + p.A1393A/ p.L1393H		*IRF2BP2* p.T394P *PRKDC* p.R14R + p.A1393A/ p.L1393H
			GAKT4		IRF2BP2 p.T394P *PRKDC* p.R14R + p.A1393A/ p.L1393H		IRF2BP2 p.T394P *PRKDC* p.R14R + p.A1393A/ p.L1393H
#5	UC	11	VarScan2				
			GAKT4				
#6	UC	24	VarScan2				
			GAKT4				
#7	IBDU	24	VarScan2				
			GAKT4	*PSTPIP1* p.Y433*	*PSTPIP1* p.Y433*		*PSTPIP1* p.Y433*
#8	IBDU	24	VarScan2	*PSTPIP1* p.Y433*	*PSTPIP1* p.Y433*		*PSTPIP1* p.Y433*
			GAKT4				
#9	UC	36	VarScan2				
			GAKT4		*BCL11B* p.G728S		
#10	UC	49	VarScan2		*BCL11B* p.G728S		
			GAKT4		*NLRP3* p.V200M		
#11	CD	52	VarScan2		*NLRP3* p.V200M		
			GAKT4				
#12	UC	58	VarScan2				
			GAKT4	*FOXP3* p.H400L	*FOXP3* p.H400L	*FOXP3* p.H400L	*FOXP3* p.H400L
#13	UC	58	VarScan2	*FOXP3* p.H400L	*FOXP3* p.H400L	*FOXP3* p.H400L	*FOXP3* p.H400L
			GAKT4				
#14	CD	72	VarScan2				
			GAKT4				
#15	CD	89	VarScan2				
			GAKT4		*NOD2* p.A755V		*NOD2* p.A755V
#16	CD	130	VarScan2		*NOD2* p.A755V		*NOD2* p.A755V
			GAKT4		*DUOX2* p.R1216W + p.A1131T		*DUOX2* p.R1216W + p.A1131T
#17	IBDU	131	VarScan2		*DUOX2* p.R1216W + p.A1131T		*DUOX2* p.R1216W + p.A1131T
			GAKT4				
#18	CD	154	VarScan2				
			GAKT4		*PLCG2* p.G699S	*PLCG2* p.G699S	*PLCG2* p.G699S
#19	CD	EOIBD (N/A)	VarScan2		*PLCG2* p.G699S	*PLCG2* p.G699S	*PLCG2* p.G699S
			GAKT4				
#20	CD	VEOIBD (N/A)	VarScan2				
			GAKT4				

Tab. 9: Results of variant identification in different gene list and bioinformatical pipelines (confirmed by Sanger sequencing and segregation analysis in the family - **green**; declined by segregation analysis in the family - **red**; suggested, but not validated yet by Sanger sequencing - **blue**)

7. Discussion

7.1. Sequencing methods

There are several possible sequencing approaches available when looking for a causal variant in genetic diseases. Whole-genome sequencing, whole-exome sequencing and targeted gene panel sequencing are the most common (specifics of each method are summarized in Tab. 3 in section *3.1 Methodological approaches*). We chose to use WES as it covers all coding sequences of genome which are majorly responsible for functional dysregulation. Also, aiming for a method applicable in routine diagnostics, the significantly lower price and analytical requirements compared to WGS were an important factor.

On the other hand, we are aware of disadvantages and biases related to WES. Mainly, there is a possibility of missing out variants within introns and regulatory regions, which may interfere with gene functions (Meyts *et al*., 2016; Uhlig & Muise, 2017). This phenomenon might be exemplified by report from Starokadomskyy *et al*. (2016), who investigated X-linked reticulate pigmentary disorder (OMIM #301220) caused by POLA1 deficiency sometimes manifesting also by chronic very early onset diarrhea. Firstly, they sequenced their patients by Sanger sequencing and WES not finding any potentially causal variant. Afterwards, they performed WGS finding rare intronic variant, which was predicted to cause alternative splicing and segregated in their patients' families with the disorder.

We acknowledged the limitations of WES concerning variants in non-coding regions. To cover potential splicing changes, we included parts of introns nearby the splicing regions into exom sequencing. However, any deeper intronic or regulatory variants were omitted from our analysis.

Bearing in mind the effectivity of variant identification between the different sequencing approaches, several reported comparisons are available. Comparing results from targeted gene panel sequencing and WES, the set of identified variants differ based on the chosen panel and cohort of patients. Saudi Mendeliome Group (2015) prepared a set of extensive gene panels including over 3 000 known genes previously described in monogenic disorders. They compared the results of causal variant identification using the extensive targeted panels and WES and found about 11 % more disease-causing

variants with WES than with the panels. What's more, another group published even higher difference in success rate of variant identification in favor of WES. They identified 23 % more causal variants by WES than by targeted gene panel sequencing. According to their results, in over one fourth of the cases, the targeted panel sequencing was also more expensive than performing WES (Dillon *et al*,. 2018).

Focusing on the targeted panel sequencing of VEO-IBD patients, Petersen *et al*. (2017) reported no difference in variant identification by targeted panel gene sequencing and WES. Moreover, Charbit-Henrion *et al*. (2018) identified higher percentage of IBD-causing variants with the targeted gene panel sequencing as WES missed three causal deletions in genes included in the panel. The missed variants were large deletions identification of which is a known limitation of WES. The targeted panel sequencing, however, does not allow identification of variants in new genes and needs to be often amended to stay up-to-date with newly reported causal genes making it a suitable method for quick routine, but not investigative research-based analyses.

When comparing results from WES and WGS for the purpose of variant identification, the coverage of the sequences is the crucial issue. WES due to pre-amplification can fail to cover all coding regions, particularly GC-rich sequences and large insertions/deletions are often not covered up-to-required standard for subsequent data analysis. Meienberg *et al*. (2016) reported difference in coverage around 2 % with exceptional higher discrepancies in specific cases such as first exons or specific genes related to pathologies and recommended for sequencing by ACMG (Green *et al*., 2013).

Despite the technical drawbacks of WES, percentual success rate of variant identification by WES and WGS was found to be similar. Genetic analysis made by Carss *et al*. (2017) for identification of causal variants in congenital retinal disease solved 55 % of cases sequenced by WGS and 50 % of cases sequenced by WES. Based on the technical facts and observed data, WES is nowadays a golden standard for causal variant identification, however, WGS is likely to become more common in the future.

7.2. Variant filtering

Crucial part of the variant analysis is variant filtering. When analyzing WGS or WES data sets, virtual gene panels are used to filter the variants in relation to their biological context. Nowadays, there are multiple commercial or open-access gene lists available for different disorders. Also, you can create your own customized gene list for filtering based on the topic of your research. What's more, as the *in silico* virtual gene panels allow us to try many filtering options with no additional experiments and costs, more than one gene list is sometimes incorporated into the variant analysis (Oh *et al*., 2015). To do so, we prepared four gene lists (CDG, PID, IBD and IBD2 as described in section *5.3.2. Variant filtering* and in Supplementary file 1) and evaluated their effectivity for VEO-IBD variant analysis.

The most numerous list we used was Closest Disease-Causing gene (CDG) list (Requena *et al*., 2018). As the author of the CDG server states, it should significantly reduce the number of considered genes in variant analysis of specific genetic diseases. However, our results are not in concordance with their findings as CDG gene list used for VEO-IBD variant analysis allowed us to identify less than 30 % of suggested causal variants.

Bearing in mind the proposed immunological basis of monogenic IBD, a list of genes associated with primary immunodeficiencies was also selected for VEO-IBD genetic analysis. This approach has already been reported, for example by Oh *et al*. (2015), who used a PID-associated gene list when analyzing exomes of Korean children diagnosed with pediatric IBD. Their PID virtual panel stemmed from classification of PID reported by International Union of Immunological Societies (Al-Herz *et al*., 2014). We prepared similar PID gene list using up-dated version of PID classification by IUIS (Bousfiha *et al*., 2018). Our PID gene list filtered 100 % of identified causal variants within our cohort supporting the theory of VEO-IBD being primary immunodeficiency with intestinal involvement.

Despite the positive results of PID gene list filtering, the results might not be completely in concordance with the newest scientific findings. The IUIS revises their reports often having published recently a new up-dated version (Tangye *et al*., 2020). Therefore, we would have to put together a new up-to-date PID gene list for up-coming

and retaken genetic analyses in order to keep up with the advances in genetics of primary immunodeficiencies.

Concerning lists of genes related closely to VEO-IBD, Uhlig *et al*. (2014) made a compilation of causal genetic defects which is widely used as a basis for VEO-IBD genetic analyses (Oh *et al*.; 2015). Their compilation contains 50 genes and we also used it as a core IBD gene list. However, bearing in mind the extensive number of new genetic reports every year, we expected that IBD list based on Uhlig *et al*. (2014) is no longer fully relevant. Our suspicion was confirmed as IBD gene list comprised less than 30 % of identified causal variants in our cohort. Therefore, use of up-dated gene list for VEO-IBD is necessary.

Comparing the available gene lists, regularly edited online open-access gene list for Infantile enterocolitis & monogenic inflammatory bowel disease from Genomics England (https://panelapp.genomicsengland.co.uk/panels/176/, accessed March 30, 2020) consists of 62 verified causal genes. However, going through recent VEO-IBD related reviews, there are significantly more causal genes reported in relation to this diagnosis. Looking into it more in detail, we found over a hundred of genes previously reported with a functional defect leading to development VEO-IBD. Therefore, we constructed an expanded monogenic-IBD-related gene list (IBD2) based on case reports of 113 causal genes (Tab. 5). Expecting high rate of positive identification of variants by IBD2 gene list, we were able to identify 77 % of the causal variant in our cohort when using IBD2. Compared to the original IBD list stemming from Uhlig *et al*. (2014)), IBD2 gene list is significantly more effective providing a suitable reference for performing VEO-IBD genetic analysis. For future use, however, newly reported genes would have to be encompassed regularly to keep it up to date.

Despite encouraging results in PID and IBD2 gene lists, we are aware of the technical draw-back of using gene lists for variant identification. The variants in genes not included into any of the lists will be missed resulting in failure to identify the cause of the patient's disease. Analyzing the variants without the filtration by gene lists, however, would be very time demanding and, thus, not suitable for clinical research and diagnostics.

7.3. Identification and interpretation of causal variants

Despite proper technical performance of WES, the success rate of variant identification in monogenic diseases is lower than 100 %. Generally, the success rate of identification for mendelian disorders is around 25 % in larger cohorts. For example, Yang *et al*. (2013) performed a genetic analysis of 250 patients with suspected congenital, predominantly neurological disorders, making molecular diagnosis of 25 % patients. Similarly, Lee *et al*. (2014). sequenced exome of 814 patients with developmental, neurological, muscular, cardiac and other monogenic disorders. Over-all, they identified genetic cause in 26 % of the patients. In addition, they found a potential causal variant in 28 % of the patients including suspected variants needed further validations such as phenotyping or allele segregation validation by Sanger sequencing.

Exceptionally, the sensitivity of the analysis was higher such as in case of previously mentioned Saudi Mendeliome Group (2015). They found a causal variant in 43 %, however, they claim to have bias in the cohort due to more common consanguinity in the Arabic population. When considering only variants in genes with autosomal dominant or X-linked inheritance, which are normally causal for the disorders, the success rate of identification was 28 %. As this study included 2 357 patients, they also tested disorders closely related to our research such as congenital gastroenterological disorders and primary immunodeficiencies. The rate of positive molecular diagnosis in those categories was 29 % and 37 %, respectively.

As the percentage differs with regards to analyzed disorder as well as the cohort of patients, comparing our findings with similarly big cohort of VEO-IBD would be suitable. Fang *et al*. (2018) performed genetic analysis of 16 patients with VEO-IBD using combination of WES and targeted gene panel sequencing. They identified a causal variant in 9 patients (17 %).

The over-all success rate of identification and validation of the causal variants by whole-exome and Sanger sequencing presented in this thesis was 20 % (4 patients out of 20). Additionally, 6 more variants identified in 5 patients (25 %) were lacking segregation analysis in the families due to unavailability of DNA from parents. The success rate of our WES-based genetic analysis, was therefore, in concordance with available statistical data from other publications.

The composition of specific variants varies vastly between the cohorts. However, common genetic defects among VEO-IBD patients are involving *IL10* and its subunits *IL10RA* and *IL10RB*. The penetrance of causal variants in those genes is statistically nearly complete and deleterious variants cause monogenic IBD with onset during first months of life (Uhlig *et al*., 2016). We have, however, not found any of genetic defects connected to *IL10* in our cohort. The probable reason is the small size of the cohort and low consanguinity in the Czech population as IL10 defects are autosomal recessive.

Another typically reported causal gene in VEO-IBD cases is *FOXP3* causing IPEX syndrome. Mutations of *FOXP3* have nearly complete penetrance and the onset of VEO-IBD-associated IPEX is diverse starting any time during the first decade of life (Uhlig *et al*., 2016). We identified novel variant in *FOXP3* (p.H400L) leading to atypical IPEX with predominant intestinal manifestation in Patient #13. However, further functional validation is necessary for completing the report of this variant.

Similarly, further functional validation has to follow in case of compound heterozygous mutation in *DUOX2* (p.R1216W + p.A1131T) which is also newly reported variant in a patient with VEO-IBD. Recently, p.A1131T was reported as additional mutation in case of a patient with congenital hypothyroidism, but its functional impact was not evaluated (Zhang *et al*., 2020). This finding is, however, in concordance with Patient's #17 small stature, which points to the direction of altered hormonal production including thyroid hormones.

On the other hand, we also found variants in genes with reported low penetrance and later onset. An example of such variant is *NOD2* (p.A755V). This variant is inherited in the family of Patient #16 and causes or contributes to Crohn's disease development in later childhood/early adulthood. The variant p.A755V has already been reported in patients with CD (Hugot *et al*., 2001), thus, needing no additional validation in this case.

Also, NLRP3 genetic defects are of a low penetrance. That is in concordance with family anamnesis and validated allelic segregation in family of Patient #11. Based on a report (Aróstégui *et al*., 2004), the variant p.V200M in Patient #11 is involved in her autoinflammatory condition, specifically suspected Muckle-Wells syndrome (OMIM #191900) with additional intestinal manifestation. Since there are no case reports about VEO-IBD in connection to NLRP3 defects, we cannot clearly conclude the effect of NLRP3 mutation to IBD pathogenesis suggesting other factors involvement in this case.

7.4. Understanding of VEO-IBD

IBD is generally understood as multifactorial disorder caused by interplay among genetics, environmental factors and recently also highly accented microbiome. In cases of VEO-IBD, the genetical defects can play a major role, making it possible to consider the disorder in some cases monogenic. The mutation leading to development of such a severe phenotype has to be of a major impact in a gene relevant for the disease. The effected genes are often encoding molecules involved in immunity homeostasis and functioning, which has either systemic impact or are relevant for intestinal mucosal immunity exemplified by highly (nearly completely) penetrant and rare in population mutations in *IL10* or *FOXP3* (Uhlig *et al*., 2016). In those cases, the mendelian concept of IBD is evident.

On the other hand, there are many genes reported as causal for IBD, but also known for their incomplete penetrance, such as *XIAP* (Uhlig *et al*., 2016). In these cases, the genetic defect of the variant show severe functional impact. When treated with hematopoietic stem cell transplant, the subsequent re-analysis revels normalized physiological phenotype (Cifaldi *et al*., 2017), thus, suggesting monogenicity of the disorder. However, XIAP deficiency is connected to more than IBD phenotype. Patients with the same deleterious *XIAP* mutation were found to suffer from several different conditions (Speckman *et al*., 2013). The causality of incompletely penetrant gene defects is, therefore, unclear, probably involving additional genetic and non-genetic factors, which has not been described yet and more research in this matter is necessary for better understanding.

Lastly, specific variants of lowly penetrant genes are repeatedly found in patients with severe childhood-onset IBD which are suggested to be involved in the pathology. Those are for example some variants in *NOD2* (e.g. c.3020insC reported by Gazouli *et al*. (2010), *IL23R* (e.g. p.R381Q reported by Van Limbergen *et al*., 2007) or *ATG16L1* (e.g. p.T300A reported by Amre *et al*. (2009). Bearing in mind the low penetrance and higher frequency in population, those variants implicate susceptibility rather than causality in majority of the cases showing resemblance to the phenotype of adulthood multifactorial IBD rather than to monogenic primary immunodeficiencies.

The genes causing or contributing to pediatrics IBD are connected to immunity in majority of the cases. The function of immune system, however, depend not only on genetics of the individual, but also on environmental factors. Brodin *et al*. (2015) estimated the variation of the immune system is driven by genetics maximally up to 40 %, meaning the model of monogenic VEO-IBD pathogenesis is oversimplified.

Nevertheless, the simplification is necessary for current clinical research as there are no widely available tools for complex evaluation of patient's phenotype in connection to his genotype and other factors. In the future, new insight could be brought by wide population sequencing projects with high-throughput methods or introduction of single cell RNA sequencing with possible special analysis to the clinics. The IBD outcome could be improved by using safe gene therapy or by manipulating with microbiota to ameliorate the dysbiosis in relevant cases. Those ideas, however, are only a matter of future and, up to date, focusing on identification of at least partial genetic cause by available technologies can significantly improve patients' lives.

CONCLUSION

Very early onset IBD patients are a specific cohort among IBD cases. The patients have often severe phenotype since early childhood and do not respond to conventional treatment well. The proposed mechanism is major defect in genes related to immune system. Therefore, looking for the patient's causal variant can help in treatment optimization and genetic counseling of the family.

We used whole exome sequencing to identify causal variants in 20 pediatric patients with IBD. The required parameters of the causal variant were low frequency in population, predicted deleteriousness and concordance of the phenotype with literature-based case reports. We proposed 13 variants in 11 genes for validation of allelic segregation by Sanger sequencing. We positively validated 5 variants including 4 variants previously not described in IBD (*DUOX2* p.R1216W + p.A1131T, *FOXP3* p.H400L, *NLRP3* p.V200M) and 1 variant (*NOD2* p.A755V) with previously reported in IBD patients. The overall success rate of the WES variant identification was 20 %. In addition, we evaluated several bioinformatical tools for WES data analysis providing evidence of suitability of newly developed GATK4-based pipeline for clinical evaluation. We also compared four gene lists ("virtual panels") for variant filtering, finding PID (primary immunodeficiencies) and IBD2 (up-to-date case-reports-based VEO-IBD) gene lists the most fitting.

Our data suggests, that VEO-IBD cases in our cohort can be in several cases considered primary immunodeficiencies with intestinal manifestation. However, a significant number of patients' analyses did not reveal any causal variant suggesting an involvement of non-genetic factors in the pathogenesis or aberrations in previously unreported genes. Thus, further research in VEO-IBD-related genetics, as well as epigenetics, host-microbiota interactions, immunological involvement and environmental factors is necessary for complex understanding of the pathology.

List of references

Abramson, J., Giraud, M., Benoist, C., & Mathis, D. (2010b). Aire's Partners in the Molecular Control of Immunological Tolerance. *Cell*, *140*(1), 123–135. https://doi.org/10.1016/j.cell.2009.12.030

Abramson, O., Durant, M., Mow, W., Finley, A., Kodali, P., Wong, A., Tavares, V., McCroskey, E., Liu, L., Lewis, J. D., Allison, J. E., Flowers, N., Hutfless, S., Velayos, F. S., Perry, G. S., Cannon, R., & Herrinton, L. J. (2010a). Incidence, prevalence, and time trends of pediatric inflammatory bowel disease in Northern California, 1996 to 2006. *Journal of Pediatrics*, *157*(2), 233-239.e1. https://doi.org/10.1016/j.jpeds.2010.02.024

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*(4), 248–249. https://doi.org/10.1038/nmeth0410-248

Agakidis, C., Agakidou, E., Sarafidis, K., Papoulidis, I., Xinias, I., & Farmaki, E. (2019). Immune dysregulation, polyendocrinopathy, enteropathy, x-linked syndrome associated with a novel mutation of FOXP3 gene. *Frontiers in Pediatrics*, *7*(FEB), 1–7. https://doi.org/10.3389/fped.2019.00020

Alangari, A., Alsultan, A., Adly, N., Massaad, M. J., Kiani, I. S., Aljebreen, A., Raddaoui, E., Almomen, A. K., Al-Muhsen, S., Geha, R. S., & Alkuraya, F. S. (2012). LPS-responsive beige-like anchor (LRBA) gene mutation in a family with inflammatory bowel disease and combined immunodeficiency. *Journal of Allergy and Clinical Immunology*, *130*(2), 1–16. https://doi.org/10.1016/j.jaci.2012.05.043

Al-Herz, W., Bousfiha, A., Casanova, J. L., Chatila, T., Conley, M. E., Cunningham-Rundles, C., Etzioni, A., Franco, J. L., Gaspar, H. B., Holland, S. M., Klein, C., Nonoyama, S., Ochs, H. D., Oksenhendler, E., Picard, C., Puck, J. M., Sullivan, K., & Tang, M. L. K. (2014). Primary immunodeficiency diseases: An update on the classification from the International Union of immunological societies expert committee for primary immunodeficiency. *Frontiers in Immunology*, *5*(APR), 1–33. https://doi.org/10.3389/fimmu.2014.00162

Al-Mayouf, S. M., Alswaied, N., Alkuraya, F. S., Almehaidib, A., & Faqih, M. (2009). Tufting enteropathy and chronic arthritis: A newly recognized association with a novel EpCAM gene mutation. *Journal of Pediatric Gastroenterology and Nutrition*, *49*(5), 642–644. https://doi.org/10.1097/MPG.0b013e3181acaeae

Al-Meghaiseeb, E. S., Al-Robayan, A. A., Al-Otaibi, M. M., Arfin, M., & Al-Asmari, A. K. (2016). Association of tumor necrosis factor-α and -β gene polymorphisms in inflammatory bowel disease. *Journal of Inflammation Research*, *9*, 133–140. https://doi.org/10.2147/JIR.S101225

Aluri, J., Gupta, M., Dalvi, A., Mhatre, S., Kulkarni, M., Hule, G., Desai, M., Shah, N., Taur, P., Vedam, R., & Madkaikar, M. (2018). Clinical, immunological, and molecular findings in five patients with Major histocompatibility complex class II deficiency from India. *Frontiers in Immunology*, *9*(FEB), 1–7. https://doi.org/10.3389/fimmu.2018.00188

Ambruso, D. R., Knall, C., Abell, A. N., Panepinto, J., Kurkchubasche, A., Thurman, G., Gonzalez-Aller, C., Hiester, A., DeBoer, M., Harbeck, R. J., Oyer, R., Johnson, G. L., & Roos, D. (2000). Human neutrophil immunodeficiency syndrome is associated with an inhibitory Rac2 mutation. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(9), 4654–4659. https://doi.org/10.1073/pnas.080074897

Amre, D. K., Mack, D. R., Morgan, K., Krupoves, A., Costea, I., Lambrette, P., Grimard, G., Dong, J., Feguery, H., Bucionis, V., Deslandres, C., Levy, E., & Seidman, E. G. (2009). Autophagy gene ATG16L1 but not IRGM is associated with Crohn's disease in Canadian children. *Inflammatory Bowel Diseases*, *15*(4), 501–507. https://doi.org/10.1002/ibd.20785

Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, L. J., Goyette P, Imielinski M, Latiano A, Lagacé C, Scott R, Amininejad L, B. S., Denson LA, De Vos M, Dubinsky M, Edwards C, Ellinghaus D, Fehrmann RS, F. J., Florin T, Franchimont D, Franke L, Georges M, Glas J, Glazer NL, G. S., Haritunians T, Hayward NK, Hugot JP, Jobin G, Laukens D, Lawrance I, L. M., Levine A, Libioulle C, Louis E, McGovern DP, Milla M, Montgomery GW, M. K., Mowat C, Ng A, Newman W, Ophoff RA, Papi L, Palmieri O, Peyrin-Biroulet L, P., J, Phillips A, Prescott NJ, Proctor DD, Roberts R, Russell R, R. P., Sanderson J, Sans M, Schumm P, Seibold F, Sharma Y, Simms LA, S. M., Steinhart AH, Targan SR, van den Berg LH, Vatn M, Verspaget H, W. T., Wijmenga C, Wilson DC, Westra HJ, Xavier RJ, Zhao ZZ, Ponsioen CY, A. V, Torkvist L, Gazouli M, Anagnou NP, Karlsen TH, Kupcinskas L, S. J., Mansfield JC, Kugathasan S, Silverberg MS, Halfvarson J, Rotter JI, M. C., Griffiths AM, Gearry R, Ahmad T, Brant SR, Chamaillard M, Satsangi J, C. J., Schreiber S, Daly MJ, Barrett JC, Parkes M, Annese V, Hakonarson H, R.-S., & G, Duerr RH, Vermeire S, Weersma RK, R. J. (2011). Meta-analysis identifies 29 additional ulcerative colitis risBaidoo L, Baldassano RN, Barclay M, Bayless TM, Brand S, Büning C, Colombel JF k loci, increasing the number of confirmed associations to 47. *Nat Genet.*, *43*(3), 246–252. https://doi.org/10.1038/ng.764

Aróstegui, J. I., Aldea, A., Modesto, C., Rua, M. J., Argüelles, F., González-Enseñat, M. A., Ramos, E., Rius, J., Plaza, S., Vives, J., & Yagüe, J. (2004). Clinical and genetic heterogeneity among Spanish patients with recurrent autoinflammatory syndromes associated with the CIAS1/PYPAF1/NALP3 gene. *Arthritis and Rheumatism*, *50*(12), 4045–4050. https://doi.org/10.1002/art.20633

Ashton, J. J., Andreoletti, G., Coelho, T., Haggarty, R., Batra, A., Afzal, N. A., Beattie, R. M., & Ennis, S. (2016). Identification of Variants in Genes Associated with Single-gene Inflammatory Bowel Disease by Whole-exome Sequencing. *Inflammatory Bowel Diseases*, *22*(10), 2317–2327. https://doi.org/10.1097/MIB.0000000000000890

Aschard, H., Laville, V., Tchetgen, E. T., Knights, D., Imhann, F., Seksik, P., Zaitlen, N., Silverberg, M. S., Cosnes, J., Weersma, R. K., Xavier, R., Beaugerie, L., Skurnik, D., & Sokol, H. (2019). Genetic effects on the commensal microbiota in inflammatory bowel disease patients. *PLoS Genetics*, *15*(3), 1–25. https://doi.org/10.1371/journal.pgen.1008018

Bader-Meunier, B., Florkin, B., Sibilia, J., Acquaviva, C., Hachulla, E., Grateau, G., Richer, O., Farber, C. M., Fischbach, M., Hentgen, V., Jego, P., Laroche, C., Neven, B., Lequerré, T., Mathian, A., Pellier, I., Touitou, I., Rabier, D., Prieur, A. M., Cuisset, L. & Quartier, P. (2011). Mevalonate kinase deficiency: A survey of 50 patients. *Pediatrics*, *128*(1). https://doi.org/10.1542/peds.2010-3639

Badran, Y. R., Dedeoglu, F., Leyva Castillo, J. M., Bainter, W., Ohsumi, T. K., Bousvaros, A., Goldsmith, J. D., Geha, R. S., & Chou, J. (2017). Human RELA haploinsufficiency results in autosomal-dominant chronic mucocutaneous ulceration. *Journal of Experimental Medicine*, *214*(7), 1937–1947. https://doi.org/10.1084/jem.20160724

Bak-Romaniszyn, L., Szala, A., Sokolowska, A., Mierzwa, G., Czerwionka-Szaflarska, M., Swierzko, A. S., Zeman, K., & Cedzynski, M. (2011). Mannan-binding lectin deficiency in pediatric patients with inflammatory bowel disease. *Scandinavian Journal of Gastroenterology*, *46*(10), 1275–1278. https://doi.org/10.3109/00365521.2011.594087

Bamidele, A. O., Svingen, P. A., Sagstetter, M. R., Sarmento, O. F., Gonzalez, M., Braga Neto, M. B., Kugathasan, S., Lomberk, G., Urrutia, R. A., & Faubion, W. A. (2019). Disruption of FOXP3–EZH2 Interaction Represents a Pathobiological Mechanism in Intestinal Inflammation. *Cmgh*, *7*(1), 55–71. https://doi.org/10.1016/j.jcmgh.2018.08.009

Barreiros, L. A., Segundo, G. R. S., Grumach, A. S., Torgerson, T. R., Ochs, H. D., & Condino-Neto, A. (2018). A novel homozygous JAK3 mutation leading to T-B+NK- SCID in two Brazilian patients. *Frontiers in Pediatrics*, *6*(August), 1–6. https://doi.org/10.3389/fped.2018.00230

Barrett, J. C., Hansoul, S., Nicolae, D. L., Cho, J. H., Duerr, R. H., Rioux, J. D., Brant, S. R., Silverberg, M. S., Taylor, K. D., Barmada, M. M., Bitton, A., Dassopoulos, T., Datta, L. W., Green, T., Griffiths, A. M., Kistner, E. O., Murtha, M. T., Regueiro, M. D., Rotter, J. I., Schumm, L.P., Steinhart, A.H., Targan, S.R., Xavier, R..J., NIDDK IBD Genetics Consortium, Libioulle, C., Sandor, C., Lathrop, M., Belaiche, J., Dewit, O., Gut,. I., Heath, S., Laukens, D., Mni, M., Rutgeerts, P., Van Gossum, A., Zelenika, D., Franchimont, D., Hugot, J. P., de Vos, M., Vermeire, S., Louis, E., Belgian-French IBD Consortium, Wellcome Trust Case Control Consortium, Cardon, L.R., Anderson, C.A., Drummond, H., Nimmo, E., Ahmad, T., Prescott, N. J,, Onnie, C. M., Fisher, S. A., Marchini, J., Ghori, J., Bumpstead, S., Gwilliam, R., Tremelling, M., Deloukas, P., Mansfield, J., Jewell, D., Satsangi, J., Mathew, C. G., Parkes, M., Georges, M., & Daly, M. J. (2008). Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nature Genetics*, *40*(8), 955–962. https://doi.org/10.1038/ng.175

Benchimol, E. I., Kaplan, G. G., Otley, A. R., Nguyen, G. C., Underwood, F. E., Guttmann, A., Jones, J. L., Potter, B. K., Catley, C. A., Nugent, Z. J., Cui, Y., Tanyingoh, D., Mojaverian, N., Bitton, A., Carroll, M. W., Debruyn, J., Dummer, T. J. B., El-Matary, W., Griffiths, A. M., Jacobson, K., Kuenzig, M. E., Leddin, D., Lix, L. M., Mack, D. R., Murthy, S. K., Sánchez, J. N. P., Singh, H., Targownik, L. E., Vutcovici, M., & Bernstein, C. N. (2017). Rural and Urban Residence during Early Life is Associated with a Lower Risk of Inflammatory Bowel Disease: A Population-Based Inception and Birth Cohort Study. *American Journal of Gastroenterology*, *112*(9), 1412–1422. https://doi.org/10.1038/ajg.2017.208

Bernasconi, A., Marino, R., Ribas, A., Rossi, J., Ciaccio, M., Oleastro, M., Ornani, A., Paz, R., Rivarola, M. A., Zelazko, M., & Belgorosky, A. (2006). Characterization of immunodeficiency in a patient with growth hormone insensitivity secondary to a novel STAT5b gene mutation. *Pediatrics*, *118*(5). https://doi.org/10.1542/peds.2005-2882

Bernstein, C. N., Eliakim, A., Fedail, S., Fried, M., Gearry, R., Goh, K. L., Hamid, S., Khan, A. G., Khalif, I., Ng, S. C., Ouyang, Q., Rey, J. F., Sood, A., Steinwurz, F., Watermeyer, G., & LeMair, A. (2016). World Gastroenterology Organisation Global Guidelines Inflammatory Bowel Disease: Update August 2015. *Journal of Clinical Gastroenterology*, *50*(10), 803–818.

Bigorgne, A. E., Farin, H. F., Lemoine, R., Mahlaoui, N., Lambert, N., Gil, M., Schulz, A., Philippet, P., Schlesser, P., Abrahamsen, T. G., Oymar, K., Graham Davies, E., Ellingsen, C. L., Leteurtre, E., Moreau-Massart, B., Berrebi, D., Bole-Feysot, C., Nischke, P., Brousse, N., … De Saint Basile, G. (2014). TTC7A mutations disrupt intestinal epithelial apicobasal polarity. *Journal of Clinical Investigation*, *124*(1), 328–337. https://doi.org/10.1172/JCI71471

Blaydon, D. C., Biancheri, P., Di, W. L., Plagnol, V., Cabral, R. M., Brooke, M. A., Van Heel, D. A., Ruschendorf, F., Toynbee, M., Walne, A., O'Toole, E. A., Martin, J. E., Lindley, K., Vulliamy, T., Abrams, D. J., MacDonald, T. T., Harper, J. I., & Kelsell, D. P. (2011). Inflammatory skin and bowel disease linked to ADAM17 deletion. *New England Journal of Medicine*, *365*(16), 1502–1508. https://doi.org/10.1056/NEJMoa1100721

Boland, B. S., Widjaja, C. E., Banno, A., Zhang, B., Kim, S. H., Stoven, S., Peterson, M. R., Jones, M. C., Su, H. I., Crowe, S. E., Bui, J. D., Ho, S. B., Okugawa, Y., Goel, A., Marietta, E. V, Khosroheidari, M., Jepsen, K., Aramburu, J., López-Rodríguez, C., Sandborn, W. J., Murray, J. A., Harismendy, O. & Chang, J. T. (2015). Immunodeficiency and Autoimmune Enterocolopathy Linked to NFAT5 Haploinsufficiency. *The Journal of Immunology*, *194*(6), 2551–2560. https://doi.org/10.4049/jimmunol.1401463

Booth, C., Gilmour, K. C., Veys, P., Gennery, A. R., Slatter, M. A., Chapel, H., Heath, P. T., Steward, C. G., Smith, O., O'Meara, A., Kerrigan, H., Mahlaoui, N., Cavazzana-Calvo, M., Fischer, A., Moshous, D., Blanche, S., Pachlopnick-Schmid, J., Latour, S., De Saint-Basile, G., Albert, M., Notheis, G., Rieber, N., Strahm, B., Ritterbusch, H., Lankester, A., Hartwig, N. G., Meyts, I., Plebani, A., Soresina, A., Finocchi, A., Pignata, C., Cirillo, E., Bonanomi, S., Peters, C., Kalwak, K., Pasic, S., Sedlacek, P., Jazbec, J., Kanegane, H., Nichols, K. E., Hanson, I. C., Kapoor, N.,

Haddad, E., Cowan, M., Choo, S., Smart, J., Arkwright, P. D. & Gaspar, H. B. (2011). X-linked lymphoproliferative disease due to SAP/SH2D1A deficiency: A multicenter study on the manifestations, management and outcome of the disease. *Blood*, *117*(1), 53–62. https://doi.org/10.1182/blood-2010-06-284935

Booth, D. R., Gillmore, J. D., Lachmann, H. J., Booth, S. E., Bybee, A., Soytürk, M., Akar, S., Pepys, M. B., Tunca, M., & Hawkins, P. N. (2000). The genetic basis of autosomal dominant familial Mediterranean fever. *QJM*, *93*(4), 217–221. https://doi.org/10.1093/qjmed/93.4.217

Borggraefe, I., Koletzko, S., Arenz, T., Fuehrer, M., Hoffmann, F., Dokal, I., Vulliamy, T., Weiler, V., Griese, M., Belohradsky, B. H., & Lang, T. (2009). Severe variant of X-linked dyskeratosis congenita (Hoyeraal-Hreidarsson Syndrome) causes significant enterocolitis in early infancy. *Journal of Pediatric Gastroenterology and Nutrition*, *49*(3), 359–363. https://doi.org/10.1097/MPG.0b013e3181a15b94

Bousfiha, A., Jeddane, L., Picard, C., Ailal, F., Bobby Gaspar, H., Al-Herz, W., Chatila, T., Crow, Y. J., Cunningham-Rundles, C., Etzioni, A., Franco, J. L., Holland, S. M., Klein, C., Morio, T., Ochs, H. D., Oksenhendler, E., Puck, J., Tang, M. L. K., Tangye, S. G., Torgerson, T. R., Casanova, J. L., & Sullivan, K. E. (2018). The 2017 IUIS Phenotypic Classification for Primary Immunodeficiencies. *Journal of Clinical Immunology*, *38*(1), 129–143. https://doi.org/10.1007/s10875-017-0465-8

Boyd, M., Thodberg, M., Vitezic, M., Bornholdt, J., Vitting-Seerup, K., Chen, Y., Coskun, M., Li, Y., Lo, B. Z. S., Klausen, P., Jan Schweiger, P., Pedersen, A. G., Rapin, N., Skovgaard, K., Dahlgaard, K., Andersson, R., Terkelsen, T. B., Lilje, B., Troelsen, J. T., Petersen, A. M., Jensen, K. B., Gögenur, I., Thielsen, P., Seidelin, J. B., Nielsen, O. H., Bjerrum, J. T., & Sandelin, A. (2018). Characterization of the enhancer and promoter landscape of inflammatory bowel disease from human colon biopsies. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-03766-z

Brodin, P., Jojic, V., Gao, T., Bhattacharya, S., Angel, C. J. L., Furman, D., Shen-Orr, S., Dekker, C. L., Swan, G. E., Butte, A. J., Maecker, H. T., & Davis, M. M. (2015). Variation in the Human Immune System Is Largely Driven by Non-Heritable Influences. *Cell*, *160*(1–2), 37–47. https://doi.org/10.1016/j.cell.2014.12.020

Brooke, M. A., Longhurst, H. J., Plagno, V., Kirkby, N. S., Mitchell, J. A., Rüschendorf, F., Warner, T. D., Kelsell, D. P., & MacDonald, T. T. (2014). Cryptogenic multifocal ulcerating stenosing enteritis associated with homozygous deletion mutations in cytosolic phospholipase A2-α. *Gut*, *63*(1), 96–104. https://doi.org/10.1136/gutjnl-2012-303581

Burlakov, V. I., Kozlova, A., Mersiyanova, I., Laberko, A., Viktorova, E., Shvets, O., Yukhacheva, D., Gordeeva, G., Raykina, E., Zakharova, V., Rodina, J., & Shcherbina, A. (2018). Clinical characteristic of a group of patients with pstpip1-associated myeloid-related proteinemia inflammatory syndrome (PAMI). *Annals of the Rheumatic Diseases*, *77*, 514.3-514. https://doi.org/10.1136/annrheumdis-2018-eular.7244

Canna, S. W., de Jesus, A. A., Gouni, S., Brooks, S. R., Marrero, B., Liu, Y., DiMattia, M. A., Zaal, K. J. M., Sanchez, G. A. M., Kim, H., Chapelle, D., Plass, N., Huang, Y., Villarino, A. V, Biancotto, A., Fleisher, T. A., Duncan, J. A., O'Shea, J. J., Benseler, S., Grom, A., Deng, Z., Laxer, R. M., & Goldbach-Mansky, R. (2014). An activating NLRC4 inflammasome mutation causes autoinflammation with recurrent macrophage activation syndrome. *Nature Genetics*, *46*(10), 1140–1146. https://doi.org/10.1038/ng.3089

Carss, K., Arno, G., Erwood, M., Stephens, J., Sanchis-Juan, A., Hull, S., Megy, K., Grozeva, D., Dewhurst, E., Malka, S., Plagnol, V., Penkett, C. J., Stirrups, K., Rizzo, R., Wright, G., Josifova, D., Bitner-Glindzicz, M., Scott, R., Clement, E., Allen, ., Armstrong, R., Brady, A. F., Carmichael, J., Chitre, M., Henderson, R. H. H., Hurst, J., MacLaren, R. E., Murphy, E., Paterson, J., Rosser, E., Thompson, D. A., Wakeling, E., Ouwehand, W. H., Michaelides, M., Moore, A. T., Webster, A. R. &Lucy, R. F. (2017). Comprehensive Rare Variant Analysis via Whole-Genome Sequencing to

Determine the Molecular Pathology of Inherited Retinal Disease. *American Journal of Human Genetics*, *100*(1), 75–90. https://doi.org/10.1016/j.ajhg.2016.12.003

Carvalho, R. S., Abadom, V., Dilworth, H. P., Thompson, R., Oliva-Hemker, M., & Cuffari, C. (2006). Indeterminate colitis: A significant subgroup of pediatric IBD. *Inflammatory Bowel Diseases*, *12*(4), 258–262. https://doi.org/10.1097/01.MIB.0000215093.62245.b9

Castanõ-Rodríguez, N., Kaakoush, N. O., Lee, W. S., & Mitchell, H. M. (2017). Dual role of Helicobacter and Campylobacter species in IBD: A systematic review and meta-analysis. *Gut*, *66*(2), 235–249. https://doi.org/10.1136/gutjnl-2015-310545

Caudy, A. A., Reddy, S. T., Chatila, T., Atkinson, J. P., & Verbsky, J. W. (2007). CD25 deficiency causes an immune dysregulation, polyendocrinopathy, enteropathy, X-linked-like syndrome, and defective IL-10 expression from CD4 lymphocytes. *Journal of Allergy and Clinical Immunology*, *119*(2), 482–487. https://doi.org/10.1016/j.jaci.2006.10.007

Cifaldi, C., Chiriaco, M., Di Matteo, G., Di Cesare, S., Alessia, S., De Angelis, P., Rea, F., Angelino, G., Pastore, M., Ferradini, V., Pagliara, D., Cancrini, C., Rossi, P., Bertaina, A., & Finocchi, A. (2017). Novel X-linked inhibitor of apoptosis mutation in very early-onset inflammatory bowel disease child successfully treated with HLA-haploidentical hemapoietic stem cells transplant after removal of αβ+ T and B cells. *Frontiers in Immunology*, *8*(DEC). https://doi.org/10.3389/fimmu.2017.01893

Conley, M. E., Dobbs, A. K., Quintana, A. M., Bosompem, A., Wang, Y. D., Coustan-Smith, E., Smith, A. M., Perez, E. E., & Murray, P. J. (2012). Agammaglobulinemia and absent B lineage cells in a patient lacking the p85α subunit of PI3K. *Journal of Experimental Medicine*, *209*(3), 463–470. https://doi.org/10.1084/jem.20112533

Conrad, M. A., Dawany, N., Sullivan, K. E., Devoto, M., & Kelsen, J. R. (2017). Novel ZBTB24 Mutation Associated with Immunodeficiency, Centromere Instability, and Facial Anomalies Type-2 Syndrome Identified in a Patient with Very Early Onset Inflammatory Bowel Disease. *Inflammatory Bowel Diseases*, *23*(12), 2252–2255. https://doi.org/10.1097/MIB.0000000000001280

Corica, D., & Romano, C. (2015). Renal involvement in inflammatory bowel diseases. *Journal of Crohn's and Colitis*, *10*(2), 1–10. https://doi.org/10.1093/ecco-jcc/jjv138

Corneo, B., Moshous, D., Güngör, T., Wulffraat, N., Philippet, P., Le Deist, F., Fischer, A., & De Villartay, J. P. (2001). Identical mutations in RAG1 or RAG2 genes leading to defective V(D)J recombinase activity can cause either T-B-severe combined immune deficiency or Omenn syndrome. *Blood*, *97*(9), 2772–2776. https://doi.org/10.1182/blood.V97.9.2772

Cornish, J. A., Tan, E., Simillis, C., Clark, S. K., Teare, J., & Tekkis, P. P. (2008). The risk of oral contraceptives in the etiology of inflammatory bowel disease: A meta-analysis. *American Journal of Gastroenterology*, *103*(9), 2394–2400. https://doi.org/10.1111/j.1572-0241.2008.02064.x

Courtois, G., Smahi, A., Reichenbach, J., Döffinger, R., Cancrini, C., Bonnet, M., Puel, A., Chable-Bessia, C., Yamaoka, S., Feinberg, J., Dupuis-Girod, S., Bodemer, C., Livadiotti, S., Novelli, F., Rossi, P., Fischer, A., Israël, A., Munnich, A., Le Deist, F., & Casanova, J. L. (2003). A hypermorphic IκBα mutation is associated with autosomal dominant anhidrotic ectodermal dysplasia and T cell immunodeficiency. *Journal of Clinical Investigation*, *112*(7), 1108–1115. https://doi.org/10.1172/JCI18714

Crowley, E., Warner, N. E., Fiedler, K., Murchie, R., Khalouei, S., Foong, J., Ramani, A., Church, P., Walters, T. D., Griffiths, A. M., & Muise, A. (2018). 587 - Whole Exome Sequencing of over 1000 Pediatric IBD Patients from a Single Centre on an Expanded Gene Panel Identifies Monogenic Forms of IBD. *Gastroenterology*, *154*(6), S-123. https://doi.org/10.1016/S0016-5085(18)30840-0

Dadi, H. K., Simon, A. J., & Roifman, C. M. (2003). Effect of CD3δ Deficiency on Maturation of α/β and γ/δ T-Cell Lineages in Severe Combined Immunodeficiency. *New England Journal of Medicine*, *349*(19), 1821–1828. https://doi.org/10.1056/NEJMoa031178

Damgaard, R. B., Walker, J. A., Marco-Casanova, P., Morgan, N. V., Titheradge, H. L., Elliott, P. R., McHale, D., Maher, E. R., McKenzie, A. N. J., & Komander, D. (2016). The Deubiquitinase OTULIN Is an Essential Negative Regulator of Inflammation and Autoimmunity. *Cell*, *166*(5), 1215-1230.e20. https://doi.org/10.1016/j.cell.2016.07.019

De La Salle, H., Zimmer, J., Fricker, D., Angenieux, C., Cazenave, J. P., Okubo, M., Maeda, H., Plebani, A., Tongio, M. M., Dormoy, A., & Hanau, D. (1999). HLA class I deficiencies due to mutations in subunit 1 of the peptide transporter TAP1. *Journal of Clinical Investigation*, *103*(5), 9–13. https://doi.org/10.1172/jci5687

De Lange, K. M., Moutsianas, L., Lee, J. C., Lamb, C. A., Luo, Y., Kennedy, N. A., Jostins, L., Rice, D. L., Gutierrez-Achury, J., Ji, S. G., Heap, G., Nimmo, E. R., Edwards, C., Henderson, P., Mowat, C., Sanderson, J., Satsangi, J., Simmons, A., Wilson, D. C., Tremelling, M., Hart, A., Mathew, C. G., Newman, W. G., Parkes, M., Lees, C. W., Uhlig, H., Hawkey, C., Prescott, N. J., Ahmad, T., Mansfield, J. C., Anderson, C. A., & Barrett, J. C. (2017). Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nature Genetics*, *49*(2), 256–261. https://doi.org/10.1038/ng.3760

De Meij, T. G. J., De Groot, E. F. J., Peeters, C. F. W., De Boer, N. K. H., Kneepkens, C. M. F., Eck, A., Benninga, M. A., Savelkoul, P. H. M., Van Bodegraven, A. A., & Budding, A. E. (2018). Variability of core microbiota in newly diagnosed treatment-naïve paediatric inflammatory bowel disease patients. *PLoS ONE*, *13*(8), 1–20. https://doi.org/10.1371/journal.pone.0197649

de Saint Basile, G., Geissmann, F., Flori, E., Uring-Lambert, B., Soudais, C., Cavazzana-Calvo, M., Durandy, A., Jabado, N., Fischer, A., & Le Deist, F. (2004). Severe combined immunodeficiency caused by deficiency in either the delta or the epsilon subunit of CD3. *J Clin Invest*, *114*(10), 1512–1517. https://doi.org/10.1172/JCI200422588.1512

Del Pinto, R., Pietropaoli, D., Chandar, A. K., Ferri, C., & Cominelli, F. (2015). Association Between Inflammatory Bowel Disease and Vitamin D Deficiency. *Inflammatory Bowel Diseases*, *21*(11), 2708–2717. https://doi.org/10.1097/MIB.0000000000000546

Dillon, O. J., Lunke, S., Stark, Z., Yeung, A., Thorne, N., Gaff, C., White, S. M., & Tan, T. Y. (2018). Exome sequencing has higher diagnostic yield compared to simulated disease-specific panels in children with suspected monogenic disorders. *European Journal of Human Genetics*, *26*(5), 644–651. https://doi.org/10.1038/s41431-018-0099-1

Disanto, J. P., Rieux-Laucat, F., Dautry-Varsat, A., Fischer, A., & De Saint Basile, G. (1994). Defective human interleukin 2 receptor γ chain in an atypical X chromosome-linked severe combined immunodeficiency with peripheral T cells. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(20), 9466–9470. https://doi.org/10.1073/pnas.91.20.9466

Do, R., Kathiresan, S., & Abecasis, G. R. (2012). Exome sequencing and complex disease: Practical aspects of rare variant association studies. *Human Molecular Genetics*, *21*(R1), 1–9. https://doi.org/10.1093/hmg/dds387

Dobbs, K., Conde, C. D., Zhang, S. Y., Parolini, S., Audry, M., Chou, J., Haapaniemi, E., Keles, S., Bilic, I., Okada, S., Massaad, M. J., Rounioja, S., Alwahadneh, A. M., Serwas, N. K., Capuder, K., Çiftçi, E., Felgentreff, K., Ohsumi, T. K., Pedergnana, V., Boisson, B., Haskologlu, Ş., Ensari, A., Schuster, M., Moretta, A., Itan, Y., Patrizi, O., Rozenberg, F., Lebon, P., Saarela, J., Knip, M., Petrovski, S., Goldstein, D. B., Parrott, R. E., Savas, B., Schambach, A., Tabellini, G., Bock, C., Chatila, T. A., Comeau, A. M., Geha, R. S., Abel, L., Buckley, R. H., Ikincioğullari, A., Al-Herz, W., Helminen, M., Dogu, F., Casanova, J. L., Boztug, K., & Notarangelo, L. D. (2015). Inherited DOCK2 deficiency in patients with early-onset invasive infections. *New England Journal of Medicine*, *372*(25), 2409–2422. https://doi.org/10.1056/NEJMoa1413462

Doroshow, D. B., Eder, J. P., & LoRusso, P. M. (2017). BET inhibitors: A novel epigenetic approach. *Annals of Oncology*, *28*(8), 1776–1787. https://doi.org/10.1093/annonc/mdx157

Driessen, G. J., IJspeert, H., Wentink, M., Yntema, H. G., van Hagen, P. M., van Strien, A., Bucciol, G., Cogulu, O., Trip, M., Nillesen, W., Peeters, E. A., Pico-Knijnenburg, I., Barendregt, B. H., Rizzi, M., van Dongen, J. J., Kutukculer, N., & van der Burg, M. (2016). Increased PI3K/Akt activity and deregulated humoral immune response in human PTEN deficiency. *Journal of Allergy and Clinical Immunology*, *138*(6), 1744-1747.e5. https://doi.org/10.1016/j.jaci.2016.07.010

El Hawary, R. E., Mauracher, A. A., Meshaal, S. S., Eldash, A., Abd Elaziz, D. S., Alkady, R., Lotfy, S., Opitz, L., Galal, N. M., Boutros, J. A., Pachlopnik Schmid, J., & Elmarsafy, A. M. (2019). MHC-II Deficiency Among Egyptians: Novel Mutations and Unique Phenotypes. *Journal of Allergy and Clinical Immunology: In Practice*, *7*(3), 856–863. https://doi.org/10.1016/j.jaip.2018.07.046

Elgizouli, M., Lowe, D. M., Speckmann, C., Schubert, D., Hülsdünker, J., Eskandarian, Z., Dudek, A., Schmitt-Graeff, A., Wanders, J., Jørgensen, S. F., Fevang, B., Salzer, U., Nieters, A., Burns, S., & Grimbacher, B. (2016). Activating PI3Kδ mutations in a cohort of 669 patients with primary immunodeficiency. *Clinical and Experimental Immunology*, *183*(2), 221–229. https://doi.org/10.1111/cei.12706

Fabre, A., Charroux, B., Martinez-Vinson, C., Roquelaure, B., Odul, E., Sayar, E., Smith, H., Colomb, V., Andre, N., Hugot, J. P., Goulet, O., Lacoste, C., Sarles, J., Royet, J., Levy, N., & Badens, C. (2012). SKIV2L mutations cause syndromic diarrhea, or trichohepatoenteric syndrome. *American Journal of Human Genetics*, *90*(4), 689–692. https://doi.org/10.1016/j.ajhg.2012.02.009

Faioni, E. M., Razzari, C., Zulueta, A., Femia, E. A., Fenu, L., Trinchera, M., Podda, G. M., Pugliano, M., Marongiu, F., & Cattaneo, M. (2014). Bleeding diathesis and gastro-duodenal ulcers in inherited cytosolic phospholipase-A2 alpha deficiency. *Thrombosis and Haemostasis*, *112*(6), 1182–1189. https://doi.org/10.1160/TH14-04-0352

Falling, C., Stebbings, S., Baxter, G. D., Gearry, R. B., & Mani, R. (2019). Musculoskeletal Pain in Individuals with Inflammatory Bowel Disease Reflects Three Distinct Profiles. *Clinical Journal of Pain*, *35*(7), 559–568. https://doi.org/10.1097/AJP.0000000000000698

Fang, Y. H., Luo, Y. Y., Yu, J. D., Lou, J. G., & Chen, J. (2018). Phenotypic and genotypic characterization of inflammatory bowel disease in children under six years of age in China. *World Journal of Gastroenterology*, *24*(9), 1035–1045. https://doi.org/10.3748/wjg.v24.i9.1035

Farachi, S., Werner, L., Konnikova, L., Rea, F., Vardi, I., Romeo, E., Barel, O., De Angelis, P., Dall'Oglio, L., Rechavi, G., Snapper, S., Somech, R., Weiss, B., Cancrini, C., & Shouval, D. (2018). Activating PIK3CD mutations cause severe intestinal lymphonodular hyperplasia and an IBD-like phenotype. *Journal of Crohn's and Colitis*, *12*(supplement_1), S036–S037. https://doi.org/10.1093/ecco-jcc/jjx180.047

Felgentreff, K., Lee, Y. N., Frugoni, F., Du, L., Van Der Burg, M., Giliani, S., Tezcan, I., Reisli, I., Mejstrikova, E., De Villartay, J. P., Sleckman, B. P., Manis, J., & Notarangelo, L. D. (2015). Functional analysis of naturally occurring DCLRE1C mutations and correlation with the clinical phenotype of ARTEMIS deficiency. *Journal of Allergy and Clinical Immunology*, *136*(1), 140-150.e7. https://doi.org/10.1016/j.jaci.2015.03.005

Felgentreff, K., Perez-Becker, R., Speckmann, C., Schwarz, K., Kalwak, K., Markelj, G., Avcin, T., Qasim, W., Davies, E. G., Niehues, T., & Ehl, S. (2011). Clinical and immunological manifestations of patients with atypical severe combined immunodeficiency. *Clinical Immunology*, *141*(1), 73–82. https://doi.org/10.1016/j.clim.2011.05.007

Fischer, A., Notarangelo, L. D., Neven, B., Cavazzana, M., & Puck, J. M. (2015). Severe combined immunodeficiencies and related disorders. *Nature Reviews Disease Primers*, *1*(October). https://doi.org/10.1038/nrdp.2015.61

Fiskerstrand, T., Arshad, N., Haukanes, B. I., Tronstad, R. R., Pham, K. D. C., Johansson, S., Håvik, B., Tnøder, S. L., Levy, S. E., Brackman, D., Boman, H., Biswas, K. H., Apold, J., Hovdenak, N., Visweswariah, S. S., & Knappskog, P. M. (2012). Familial diarrhea syndrome caused by an

activating GUCY2C mutation. *New England Journal of Medicine*, *366*(17), 1586–1595. https://doi.org/10.1056/NEJMoa1110132

Flanagan, S. E., Haapaniemi, E., Russell, M. A., Caswell, R., Allen, L., Franco, E. De, Mcdonald, T. J., Rajala, H., Ramelius, A., Barton, J., Heiskanen, K., Heiskanen-kosma, T., Kajosaari, M., & Nuala, P. (2015). *Europe PMC Funders Group Activating germline mutations in STAT3 cause early-onset multi- organ autoimmune disease*. *46*(8), 812–814. https://doi.org/10.1038/ng.3040.Activating

Fliegauf, M., L. Bryant, V., Frede, N., Slade, C., Woon, S. T., Lehnert, K., Winzer, S., Bulashevska, A., Scerri, T., Leung, E., Jordan, A., Keller, B., De Vries, E., Cao, H., Yang, F., Schäffer, A. A., Warnatz, K., Browett, P., Douglass, J., … Grimbacher, B. (2015). Haploinsufficiency of the NF-κB1 Subunit p50 in Common Variable Immunodeficiency. *American Journal of Human Genetics*, *97*(3), 389–403. https://doi.org/10.1016/j.ajhg.2015.07.008

Fontenot, J. D., Gavin, M. A., & Rudensky, A. Y. (2003). Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. *Nature Immunology*, *4*(4), 330–336. https://doi.org/10.1038/ni904

Fraga, M., Fournier, N., Safroneeva, E., Pittet, V., Godat, S., Straumann, A., Nydegger, A., Vavricka, S. R., Moradpour, D., & Schoepfer, A. M. (2017). Primary sclerosing cholangitis in the Swiss Inflammatory Bowel Disease Cohort Study: Prevalence, risk factors, and long-term follow-up. *European Journal of Gastroenterology and Hepatology*, *29*(1), 91–97. https://doi.org/10.1097/MEG.0000000000000747

Franke, A., McGovern, D. P. B., Barrett, J. C., Wang, K., Radford-Smith, G. L., Ahmad, T., Lees, C. W., Balschun, T., Lee, J., Roberts, R., Anderson, C. A., Bis, J. C., Bumpstead, S., Ellinghaus, D., Festen, E. M., Georges, M., Green, T., Haritunians, T., Jostins, L., Latiano, A., Mathew, C. G, Montgomery, G. W., Prescott, N. J, Raychaudhuri, S., Rotter, J. I., Schumm, P., Sharma, Y., Simms, L. A., Taylor, K. D., Whiteman, D., Wijmenga, C., Baldassano, R. N., Barclay, M., Bayless, T. M., Brand, S., Büning, C., Cohen, A., Colombel, J. F., Cottone, M., Stronati, L., Denson, T., De Vos, M., D'Inca, R., Dubinsky, M., Edwards, C., Florin, T., Franchimont, D., Gearry, R., Glas, J., Van Gossum, A., Guthery, S. L., Halfvarson, J., Verspaget, H. W., Hugot, J. P., Karban, A., Laukens, D., Lawrance, I., Lemann, M., Levine, A., Libioulle, C., Louis, E., Mowat, C., Newman, W., Panés, J., Phillips, A., Proctor, D. D., Regueiro, M., Russell, R., Rutgeerts, P., Sanderson, J., Sans, M., Seibold, F., Steinhart, A. H., Stokkers, P. C. F., Torkvist, L., Kullak-Ublick, G., Wilson, D., Walters, T., Targan, S R., Brant, S. R., Rioux, J. D., D'Amato, M., Weersma, R. K., Kugathasan, S., Griffiths, A. M., Mansfield, J. C., Vermeire, S., Duerr, R. H., Silverberg, M. S, Satsangi, J., Schreiber, S., Cho, J. H., Annese, V., Hakonarson, H., Daly, M. J., & Parkes, M. (2010). Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nature Genetics*, *42*(12), 1118–1125. https://doi.org/10.1038/ng.717

Frans, G., Moens, L., Schaballie, H., Wuyts, G., Liston, A., Poesen, K., Janssens, A., Rice, G. I., Crow, Y. J., Meyts, I., & Bossuyt, X. (2017). Homozygous N-terminal missense mutation in TRNT1 leads to progressive B-cell immunodeficiency in adulthood. *Journal of Allergy and Clinical Immunology*, *139*(1), 360-363.e6. https://doi.org/10.1016/j.jaci.2016.06.050

Freeman, E. B., Köglmeier, J., Martinez, A. E., Mellerio, J. E., Haynes, L., Sebire, N. J., Lindley, K. J., & Shah, N. (2008). Gastrointestinal complications of epidermolysis bullosa in children. *British Journal of Dermatology*, *158*(6), 1308–1314. https://doi.org/10.1111/j.1365-2133.2008.08507.x

Galeotti, C., Meinzer, U., Quartier, P., Rossi-Semerano, L., Bader-Meunier, B., Pillet, P., & Koné-Paut, I. (2012). Efficacy of interleukin-1-targeting drugs in mevalonate kinase deficiency. *Rheumatology (United Kingdom)*, *51*(10), 1855–1859. https://doi.org/10.1093/rheumatology/kes097

Gazouli, M., Pachoula, I., Panayotou, I., Mantzaris, G., Chrousos, G., Anagnou, N. P., & Roma-Giannikou, E. (2010). NOD2/CARD15, ATG16L1 and IL23R gene polymorphisms and childhood-onset of Crohn's disease. *World Journal of Gastroenterology*, *16*(14), 1753–1758. https://doi.org/10.3748/wjg.v16.i14.1753

Girardelli, M., Loganes, C., Pin, A., Stacul, E., Decleva, E., Vozzi, D., Baj, G., De Giacomo, C., Tommasini, A., & Bianco, A. M. (2018). Novel NOD2 Mutation in Early-Onset Inflammatory Bowel Phenotype. *Inflammatory Bowel Diseases*, *24*(6), 1204–1212. https://doi.org/10.1093/ibd/izy061

Glogauer, M., Marchal, C. C., Zhu, F., Worku, A., Clausen, B. E., Foerster, I., Marks, P., Downey, G. P., Dinauer, M., & Kwiatkowski, D. J. (2003). Rac1 Deletion in Mouse Neutrophils Has Selective Effects on Neutrophil Functions. *The Journal of Immunology*, *170*(11), 5652–5657. https://doi.org/10.4049/jimmunol.170.11.5652

Goldman, F. D., Ballas, Z. K., Schutte, B. C., Kemp, J., Hollenback, C., Noraz, N., & Taylor, N. (1998). Defective expression of p561ck in an infant with severe combined immunodeficiency. *Journal of Clinical Investigation*, *102*(2), 421–429. https://doi.org/10.1172/JCI3205

Grasberger, H., Gao, J., Nagao-Kitamoto, H., Kitamoto, S., Zhang, M., Kamada, N., Eaton, K. A., El-Zaatari, M., Shreiner, A. B., Merchant, J. L., Owyang, C., & Kao, J. Y. (2015). Increased Expression of DUOX2 Is an Epithelial Response to Mucosal Dysbiosis Required for Immune Homeostasis in Mouse Intestine. *Gastroenterology*, *149*(7), 1849–1859. https://doi.org/10.1053/j.gastro.2015.07.062

Green, R. C., Berg, J. S., Grody, W. W., Kalia, S. S., Korf, B. R., Martin, C. L., McGuire, A. L., Nussbaum, R. L., O'Daniel, J. M., Ormond, K. E., Rehm, H. L., Watson, M. S., Williams, M. S., & Biesecker, L. G. (2013). ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genetics in Medicine*, *15*(7), 565–574. https://doi.org/10.1038/gim.2013.73

Grimbacher, B., Hutloff, A., Schlesier, M., Glocker, E., Warnatz, K., Dräger, R., Eibel, H., Fischer, B., Schäffer, A. A., Mages, H. W., Kroczek, R. A., & Peter, H. H. (2003). Homozygous loss of ICOS is associated with adult-onset common variable immunodeficiency. *Nature Immunology*, *4*(3), 261–268. https://doi.org/10.1038/ni902

Guerri, S., Danti, G., Frezzetti, G., Lucarelli, E., Pradella, S., & Miele, V. (2019). Clostridium difficile colitis: CT findings and differential diagnosis. *Radiologia Medica*, *124*(12), 1185–1198. https://doi.org/10.1007/s11547-019-01066-0

Guindi, M., & Riddell, R. H. (2004). Indeterminate colitis. *Journal of Clinical Pathology*, *57*(12), 1233–1244. https://doi.org/10.1136/jcp.2003.015214

Haapaniemi, E. M., Kaustio, M., Rajala, H. L. M., Adrichem, A. J. Van, Kainulainen, L., Glumoff, V., Doffinger, R., Kuusanm, H., Heiskanen-kosma, T., Trotta, L., Chiang, S., Kulmala, P., Mustjoki, S., & Kere, J. (2015). *and mycobacterial disease in patients with activating mutations in STAT3*. *125*(4), 639–649. https://doi.org/10.1182/blood-2014-04-570101.The

Harris, C. J., Scheibe, M., Wongpalee, S. P., Liu, W., Cornett, E. M., Vaughan, R. M., Li, X., Chen, W., Xue, Y., Zhong, Z., Yen, L., Barshop, W. D., Rayatpisheh, S., Gallego-Bartolome, J., Groth, M., Wang, Z., Wohlschlegel, J. A., Du, J., Rothbart, S. B., Butter, F., & Jacobsen, S. E. (2018). A DNA methylation reader complex that enhances gene transcription. *Science*, *362*(6419), 1182–1186. https://doi.org/10.1126/science.aar7854

Hartley, J. L., Zachos, N. C., Dawood, B., Donowitz, M., Forman, J., Pollitt, R. J., Morgan, N. V., Tee, L., Gissen, P., Kahr, W. H. A., Knisely, A. S., Watson, S., Chitayat, D., Booth, I. W., Protheroe, S., Murphy, S., de Vries, E., Kelly, D. A., & Maher, E. R. (2010). Mutations in TTC37 Cause Trichohepatoenteric Syndrome (Phenotypic Diarrhea of Infancy). *Gastroenterology*, *138*(7), 2388–2398. https://doi.org/10.1053/j.gastro.2010.02.010

Hashimoto, A., Takeda, K., Inaba, M., Sekimata, M., Kaisho, T., Ikehara, S., Homma, Y., Akira, S., & Kurosaki, T. (2000). Cutting Edge: Essential Role of Phospholipase C-γ2 in B Cell Development and Function. *The Journal of Immunology*, *165*(4), 1738–1742. https://doi.org/10.4049/jimmunol.165.4.1738

Hayes, P., Dhillon, S., O'Neill, K., Thoeni, C., Hui, K. Y., Elkadri, A., Guo, C. H., Kovacic, L., Aviello, G., Alvarez, L. A., Griffiths, A. M., Snapper, S. B., Brant, S. R., Doroshow, J. H., Silverberg, M. S., Peter, I., McGovern, D. P. B., Cho, J., Brumell, J. H., Uhlig, H. H., Bourke, B., Muise, A. M., & Knaus, U. G. (2015). Defects in Nicotinamide-adenine Dinucleotide Phosphate Oxidase Genes NOX1 and DUOX2 in Very Early Onset Inflammatory Bowel Disease. *Cellular and Molecular Gastroenterology and Hepatology*, *1*(5), 489–502. https://doi.org/10.1016/j.jcmgh.2015.06.005

Hershkovitz, D., Gross, Y., Nahum, S., Yehezkel, S., Sarig, O., Uitto, J., & Sprecher, E. (2011). Functional Characterization of SAMD9, a Protein Deficient in Normophosphatemic Familial Tumoral Calcinosis. *Journal of Investigative Dermatology*, *131*(3), 662–669. https://doi.org/10.1038/jid.2010.387

Heyworth, P. G., Bohl, B. P., Bokoch, G. M., & Curnutte, J. T. (1994). Rac translocates independently of the neutrophil NADPH oxidase components p47(phox) and p67(phox). Evidence for its interaction with flavocytochrome b558. *Journal of Biological Chemistry*, *269*(49), 30749–30752.

Hirota, S. A., Ng, J., Lueng, A., Khajah, M., Parhar, K., Li, Y., Lam, V., Potentier, M. S., Ng, K., Bawa, M., McCafferty, D. M., Rioux, K. P., Ghosh, S., Xavier, R. J., Colgan, S. P., Tschopp, J., Muruve, D., MacDonald, J. A., & Beck, P. L. (2011). NLRP3 inflammasome plays a key role in the regulation of intestinal homeostasis. *Inflammatory Bowel Diseases*, *17*(6), 1359–1372. https://doi.org/10.1002/ibd.21478

Holmberg, C., & Perheentupa, J. (1985). Congenital Na+ diarrhea: A new type of secretory diarrhea. *The Journal of Pediatrics*, *106*(1), 56–61. https://doi.org/10.1016/S0022-3476(85)80465-0

Huber, W., Herrmann, G., Schuster, T., Phillip, V., Saugel, B., Schultheiss, C., Hoellthaler, J., Gaa, J., Hartel, M., Schmid, R. M., & Reindl, W. (2010). Lebensbedrohliche Komplikationen von Morbus Crohn und Colitis ulcerosa. *Deutsche Medizinische Wochenschrift*, *135*(14), 668–674. https://doi.org/10.1055/s-0030-1251915

Hugot, J. P., Chamaillard, M., Zouali, H., Lesage, S., Cézard, J. P., Belaiche, J., Almer, S., Tysk, C., O'morain, C. A., Gassull, M., Binder, V., Finkel, Y., Cortot, A., Modigliani, R., Laurent-Puig, P., Gower-Rousseau, C., Macry, J., Colombel, J. F., Sahbatou, M., & Thomas, G. (2001). Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*, *411*(6837), 599–603. https://doi.org/10.1038/35079107

Hull, S., Arno, G., Holder, G. E., Plagnol, V., Gomez, K., Liesner, R., Webster, A. R., & Moore, A. T. (2016). The ophthalmic presentation of Hermansky-Pudlak syndrome 6. *British Journal of Ophthalmology*, *100*(11), 1521–1524. https://doi.org/10.1136/bjophthalmol-2015-308067

Hussain, N., Quezado, M., Huizing, M., Geho, D., White, J. G., Gahl, W., & Mannon, P. (2006). Intestinal disease in hermansky-pudlak syndrome: Occurrence of colitis and relation to genotype. *Clinical Gastroenterology and Hepatology*, *4*(1), 73–80. https://doi.org/10.1016/S1542-3565(05)00858-X

Chan, A. Y., Punwani, D., Kadlecek, T. A., Cowan, M. J., Olson, J. L., Mathes, E. F., Sunderam, U., Fu, S. M., Srinivasan, R., Kuriyan, J., Brenner, S. E., Weiss, A., & Puck, J. M. (2016). A novel human autoimmune syndrome caused by combined hypomorphic and activating mutations in ZAP-70. *Journal of Experimental Medicine*, *213*(2), 155–165. https://doi.org/10.1084/jem.20150888

Charbit-Henrion, F., Parlato, M., Hanein, S., Duclaux-Loras, R., Nowak, J., Begue, B., Rakotobe, S., Bruneau, J., Fourrage, C., Alibeu, O., Rieux-Laucat, F., Lévy, E., Stolzenberg, M. C., Mazerolles, F., Latour, S., Lenoir, C., Fischer, A., Picard, C., Aloi, M., Dias, J. A., Hariz, M. B., Bourrier, A., Breuer, C., Breton, A., Bronski, J., Buderus, S., Cananzi, M., Coopman, S., Crémilleux, C., Dabadie, A., Dumant-Forest, C., Gurkan, O. E., Fabre, A., Fischer, A., Diaz, M. G., Gonzalez-Lama, Y., Goulet, O., Guariso, G., Gurcan, N., Homan, M., Hugot, J. P., Jeziorski, E., Karanika, E., Lachaux, A., Lewindon, P., Lima, R., Magro, F., Major, J., Malamut, G., Mas, E., Mattyus, I., Mearin, L. M., Melek, J., Navas-Lopez, V. M., Paerregaard, A., Pelatan, C., Pigneur, B., Pais, I. P., Rebeuh, J., Romano, C., Siala, N., Strisciuglio, C., Tempia-Caliera, M., Tounian, P., Turner, D.,

Urbonas, V., Willot, S., Ruemmele, F. M., & Cerf-Bensussan, N. (2018). Diagnostic yield of next-generation sequencing in very early-onset inflammatory bowel diseases: A multicentre study. *Journal of Crohn's and Colitis*, *12*(9), 1104–1112. https://doi.org/10.1093/ecco-jcc/jjy068

Chassaing, B., Koren, O., Goodrich, J., Poole, A., Srinivasan, S., Ley, R. E., & Gewirtz, A. T. (2015). Colitis and Metabolic Syndrome. *Nature*, *519*(7541), 92–96. https://doi.org/10.1038/nature14232.Dietary

Cheng, L. E., Kanwar, B., Tcheurekdjian, H., Grenert, J. P., Muskat, M., Heyman, M. B., McCune, J. M., & Wara, D. W. (2009). Persistent systemic inflammation and atypical enterocolitis in patients with NEMO syndrome. *Clinical Immunology*, *132*(1), 124–131. https://doi.org/10.1016/j.clim.2009.03.514

Childs, K. S., & Goodbourn, S. (2003). Identification of novel co-repressor molecules for Interferon Regulatory Factor-2. *Nucleic Acids Research*, *31*(12), 3016–3026. https://doi.org/10.1093/nar/gkg431

Chou, J., Massaad, M. J., Wakim, R. H., Bainter, W., Dbaibo, G., & Geha, R. S. (2014). A novel mutation in FOXN1 resulting in SCID: A case report and literature review. *Clinical Immunology*, *155*(1), 30–32. https://doi.org/10.1016/j.clim.2014.08.005

Imielinski, M., Baldassano, R. N., Griffiths, A., Russell, R. K., Annese, V., Dubinsky, M., Kugathasan, S., Bradfield, J. P., Walters, T. D., Sleiman, P., Kim, C. E., Muise, A., Wang, K., Glessner, J. T., Saeed, S., Zhang, H., Frackelton, E. C., Hou, C., Flory, J. H., Otieno, G., Chiavacci, R. M., Grundmeier, R., Castro, M., Latiano, A., Dallapiccola, B., Stempak, J., Abrams, D. J., Taylor, K., McGovern, D., Heyman, M. B., Ferry, G. D., Kirschner, B., Lee, J., Essers, J., Grand, R., Stephens, M., Levine, A., Piccoli, D., Van Limbergen, J., Cucchiara, S., Monos, D. S., Guthery, S. L., Denson, L., Wilson, D. C., Grant, S. F. A., Daly, M., Silverberg, M. S., Satsangi, J., & Hakonarson, H. (2009). Common variants at five new loci associated with early-onset inflammatory bowel disease. *Nature Genetics*, *41*(12), 1335–1340. https://doi.org/10.1038/ng.489

Janecke, A. R., Heinz-Erian, P., Yin, J., Petersen, B. S., Franke, A., Lechner, S., Fuchs, I., Melancon, S., Uhlig, H. H., Travis, S., Marinier, E., Perisic, V., Ristic, N., Gerner, P., Booth, I. W., Wedenoja, S., Baumgartner, N., Vodopiutz, J., Frechette-Duval, M. C., De Lafollie, J., Persad, R., Warner, N., Tse, C. M., Sud, K., Zachos, N. C., Sarker, R., Zhu, X., Muise, A. M., Zimmer, K. P., Witt, H., Zoller, H., Donowitz, M., & Müller, T. (2015). Reduced sodium/proton exchanger NHE3 activity causes congenital sodium diarrhea. *Human Molecular Genetics*, *24*(23), 6614–6623. https://doi.org/10.1093/hmg/ddv367

Jarkovský, J., Benešová, K., Hejduk, K., Dušek, L., & Lukáš, M. (2017). Epidemiologie, hospitalizační léčba a migrace IBD pacientů za specializovanou péčí v České Republice. *Gastroenterologie a Hepatologie*, *71*(6), 501–509. https://doi.org/10.14735/amgh2017501

Jones, G. R., Bain, C. C., Fenton, T. M., Kelly, A., Brown, S. L., Ivens, A. C., Travis, M. A., Cook, P. C., & MacDonald, A. S. (2018). Dynamics of colon monocyte and macrophage activation during colitis. *Frontiers in Immunology*, *9*(NOV). https://doi.org/10.3389/fimmu.2018.02764

Joosse, M., Charbit-Henrion, F., Raatgeep, H., Lindenbergh-Kortleve, D. J., Costes, L. M., Nugteren, S., Malan, V., Nowak, J., Mearin, M. L., Escher, J., Cerf-Bensussan, N., & Samsom, J. N. (2018). Consequences of an IL2RA locus duplication in a very early onset inflammatory bowel disease patient. *Journal of Crohn's and Colitis*, *12*(supplement_1), S141–S142. https://doi.org/10.1093/ecco-jcc/jjx180.222

Jostins, L., Ripke, S., Weersma, R. K., Duerr, R. H., McGovern, D. P., Hui, K. Y., Lee, J. C., Philip Schumm, L., Sharma, Y., Anderson, C. A., Essers, J., Mitrovic, M., Ning, K., Cleynen, I., Theatre, E., Spain, S. L., Raychaudhuri, S., Goyette, P., Wei, Z., Abraham, C., Achkar, J. P., Ahmad, T., Amininejad, L., Ananthakrishnan, A N., Andersen, V., Andrews, J. M., Baidoo, L., Balschun, T., Bampton, P. A., Bitton, A., Boucher, G., Brand, S., Büning, C., Cohain, A., Cichon, S., D'Amato, M., De Jong, D., Devaney, K. L., Dubinsky, M., Edwards, C., Ellinghaus, D., Ferguson, L. R.,

Franchimont, D., Fransen, K., Gearry, R., Georges, M., Gieger, C., Glas, J., Haritunians, T., Hart, A., Hawkey, C., Hedl, M., Hu, X., Karlsen, T. H., Kupcinskas, L., Kugathasan, S., Latiano, A., Laukens, D., Lawrance, I. C., Lees, C. W., Louis, E., Mahy, G., Mansfield, J., Morgan, A. R., Mowat, C., Newman, W., Palmieri, O., Ponsioen, C. Y., Potocnik, U., Prescott, N. J., Regueiro, M., Rotter, J. I., Russell, R. K., Sanderson, J. D., Sans, M., Satsangi, J., Schreiber, S., Simms, L. A., Sventoraityte, J., Targan, S. R., Taylor, K. D., Tremelling, M., Verspaget, H. W., De Vos, M., Wijmenga, C., Wilson, D. C., Winkelmann, J., Xavier, R. J., Zeissig, S., Zhang, B., Zhang, C. K., Zhao, H., International IBD Genetics Consortium (IIBDGC), Silverberg. M. S-, Annese, V, Hakonarson. H., Brant, S. R., Radford-Smith, G., Mathew, C. G., Rioux, J. D., Schadt, E. E., Daly, M. J., Franke, A., Parkes, M., Vermeire, S., Barrett, J. C., & Cho, J. H. (2012). Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*, *491*(7422), 119–124. https://doi.org/10.1038/nature11582

Kammermeier, J., Drury, S., James, C. T., Dziubak, R., Ocaka, L., Elawad, M., Beales, P., Lench, N., Uhlig, H. H., Bacchelli, C., & Shah, N. (2014). Targeted gene panel sequencing in children with very early onset inflammatory bowel disease-evaluation and prospective analysis. *Journal of Medical Genetics*, *51*(11), 748–755. https://doi.org/10.1136/jmedgenet-2014-102624

Kanegane, H., Agematsu, K., Futatani, T., Sira, M. M., Suga, K., Sekiguchi, T., van Zelm, M. C., & Miyawaki, T. (2007). Novel mutations in a Japanese patient with CD19 deficiency. *Genes and Immunity*, *8*(8), 663–670. https://doi.org/10.1038/sj.gene.6364431

Kanneganti, T. D., Özören, N., Body-Malapel, M., Amer, A., Park, J. H., Franchi, L., Whitfield, J., Barchet, W., Colonna, M., Vandenabeele, P., Bertin, J., Coyle, A., Grant, E. P., Akira, S., & Núñez, G. (2006). Bacterial RNA and small antiviral compounds activate caspase-1 through cryopyrin/Nalp3. *Nature*, *440*(7081), 233–236. https://doi.org/10.1038/nature04517

Kaplan, G. G., Jackson, T., Sands, B. E., Frisch, M., Andersson, R. E., & Korzenik, J. (2008). The risk of developing Crohn's disease after an appendectomy: A meta-analysis. *American Journal of Gastroenterology*, *103*(11), 2925–2931. https://doi.org/10.1111/j.1572-0241.2008.02118.x

Kaplan, G. G., & Ng, S. C. (2017). Understanding and Preventing the Global Increase of Inflammatory Bowel Disease. *Gastroenterology*, *152*(2), 313-321.e2. https://doi.org/10.1053/j.gastro.2016.10.020

Karatzas, P. S., Mantzaris, G. J., Safioleas, M., & Gazouli, M. (2014). DNA methylation profile of genes involved in inflammation and autoimmunity in inflammatory bowel disease. *Medicine (United States)*, *93*(28), e309. https://doi.org/10.1097/MD.0000000000000309

Kaser, A., Zeissig, S., & Blumberg, R. S. (2010). Genes and environment: How will our concepts on the pathophysiology of IBD develop in the future? *Digestive Diseases*, *28*(3), 395–405. https://doi.org/10.1159/000320393

Keller, M. D., Pandey, R., Li, D., Glessner, J., Tian, L., Henrickson, S. E., Chinn, I. K., Monaco-Shawver, L., Heimall, J., Hou, C., Otieno, F. G., Jyonouchi, S., Calabrese, L., van Montfrans, J., Orange, J. S., & Hakonarson, H. (2016). Mutation in IRF2BP2 is responsible for a familial form of common variable immunodeficiency disorder. *Journal of Allergy and Clinical Immunology*, *138*(2), 544-550.e4. https://doi.org/10.1016/j.jaci.2016.01.018

Kelsen, J. R., Dawany, N., Martinez, A., Grochowski, C. M., Maurer, K., Rappaport, E., Piccoli, D. A., Baldassano, R. N., Mamula, P., Sullivan, K. E., & Devoto, M. (2015b). A de novo whole gene deletion of XIAP detected by exome sequencing analysis in very early onset inflammatory bowel disease: A case report. *BMC Gastroenterology*, *15*(1), 1–5. https://doi.org/10.1186/s12876-015-0394-z

Kelsen, J. R., Dawany, N., Moran, C. J., Petersen, B.-S., Sarmady, M., Sasson, A., Pauly-Hubbard, H., Martinez, A., Maurer, K., Soong, J., Rappaport, E., Franke, A., Keller, A., Winter, H. S., Mamula, P., Piccoli, D., Artis, D., Sonnenberg, G. F., Daly, M., Sullivan, K. E., Baldassano, R. N., & Devoto, M. (2015a). Exome Sequencing Analysis Reveals Variants in Primary Immunodeficiency

Genes in Patients With Very Early Onset Inflammatory Bowel Disease. *Gastroenterology*, *149*(6), 1415–1424. https://doi.org/10.1053/j.gastro.2015.07.006

Kircher, M., Witten, D. M., Jain, P., O'roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, *46*(3), 310–315. https://doi.org/10.1038/ng.2892

Kleinewietfeld, M., Manzel, A., Titze, J., Kvakan, H., Yosef, N., Linker, R. A., Muller, D. N., & Hafler, D. A. (2013). Sodium chloride drives autoimmune disease by the induction of pathogenic TH17 cells. *Nature*, *496*(7446), 518–522. https://doi.org/10.1038/nature11868

Klemann, C., Camacho-Ordonez, N., Yang, L., Eskandarian, Z., Rojas-Restrepo, J. L., Frede, N., Bulashevska, A., Heeg, M., Al-Ddafari, M. S., Premm, J., Seidl, M., Ammann, S., Sherkat, R., Radhakrishnan, N., Warnatz, K., Unger, S., Kobbe, R., Hüfner, A., Ronan Leahy, T., Ip, W., Burns, S. O., Fliegauf, M., & Grimbacher, B. (2019). Clinical and immunological phenotype of patients with primary immunodeficiency due to damaging mutations in NFKB2. *Frontiers in Immunology*, *10*(MAR), 1–17. https://doi.org/10.3389/fimmu.2019.00297

Kotlarz, D., Beier, R., Murugan, D., Diestelhorst, J., Jensen, O., Boztug, K., Pfeifer, D., Kreipe, H., Pfister, E. D., Baumann, U., Puchalka, J., Bohne, J., Egritas, O., Dalgic, B., Kolho, K. L., Sauerbrey, A., Buderus, S., Güngör, T., Enninger, A., Koda, Y. K. L., Guariso, G., Weiss, B., Corbacioglu, S., Socha, P., Uslu, N., Metin, A., Wahbeh, G. T., Husain, K., Ramadan, D., Al-Herz, W., Grimbacher, B., Sauer, M., Sykora, K. W., Koletzko, S., & Klein, C. (2012). Loss of interleukin-10 signaling and infantile inflammatory bowel disease: Implications for diagnosis and therapy. *Gastroenterology*, *143*(2), 347–355. https://doi.org/10.1053/j.gastro.2012.04.045

Kotlarz, D., Marquardt, B., Barøy, T., Lee, W. S., Konnikova, L., Hollizeck, S., Magg, T., Lehle, A. S., Walz, C., Borggraefe, I., Hauck, F., Bufler, P., Conca, R., Wall, S. M., Schumacher, E. M., Misceo, D., Frengen, E., Bentsen, B. S., Uhlig, H. H., Hopfner, K. P., Muise, A. M., Snapper, S. B., Strømme, P., & Klein, C. (2018). Human TGF-β1 deficiency causes severe inflammatory bowel disease and encephalopathy. *Nature Genetics*, *50*(3), 344–348. https://doi.org/10.1038/s41588-018-0063-6

Kuloglu, Z., Kansu, A., Serwas, N., Demir, A., Yaman, A., Ensari, A., & Boztug, K. (2015). Inflammatory bowel disease-like phenotype in a young girl with prolidase deficiency: a new spectrum of clinical manifestation. *Genetic Counseling*, *26*(2), 205–211.

Lanzi, G., Moratto, D., Vairo, D., Masneri, S., Delmonte, O., Paganini, T., Parolini, S., Tabellini, G., Mazza, C., Savoldi, G., Montin, D., Martino, S., Tovo, P., Pessach, I. M., Massaad, M. J., Ramesh, N., Porta, F., Plebani, A., Notarangelo, L. D., Geha, R.S., & Giliani, S. (2012). A novel primary human immunodeficiency due to deficiency in the WASP-interacting protein WIP. *Journal of Experimental Medicine*, *209*(1), 29–34. https://doi.org/10.1084/jem.20110896

Lee, D. W., Koo, J. S., Choe, J. W., Suh, S. J., Kim, S. Y., Hyun, J. J., Jung, S. W., Jung, Y. K., Yim, H. J., & Lee, S. W. (2017a). Diagnostic delay in inflammatory bowel disease increases the risk of intestinal surgery. *World Journal of Gastroenterology*, *23*(35), 6474–6481. https://doi.org/10.3748/wjg.v23.i35.6474

Lee, H., Deignan, J. L., Dorrani, N., Strom, S. P., Kantarci, S., Quintero-Rivera, F., Das, K., Toy, T., Harry, B., Yourshaw, M., Fox, M., Fogel, B. L., Martinez-Agosto, J. A., Wong, D. A., Chang, V. Y., Shieh, P. B., Palmer, C. G. S., Dipple, K. M., Grody, W. W., Vilain, E., & Nelson, S. F. (2014). Clinical Exome Sequencing for Genetic Identification of Rare Mendelian Disorders. *JAMA*, *312*(18), 1880. https://doi.org/10.1001/jama.2014.14604

Lee, C., Kim, B. G., Kim, J. H., Chun, J., Im, J. P., & Kim, J. S. (2017b). Sodium butyrate inhibits the NF-kappa B signaling pathway and histone deacetylation, and attenuates experimental colitis in an IL-10 independent manner. *International Immunopharmacology*, *51*(July), 47–56. https://doi.org/10.1016/j.intimp.2017.07.023

LeGuen, T., Jullien, L., Touzot, F., Schertzer, M., Gaillard, L., Perderiset, M., Carpentier, W., Nitschke, P., Picard, C., Couillault, G., Soulier, J., Fischer, A., Callebaut, I., Jabado, N., Londono-Vallejo, A., de Villartay, J. P., & Revy, P. (2013). Human RTEL1 deficiency causes hoyeraal-hreidarsson syndrome with short telomeres and genome instability. *Human Molecular Genetics*, *22*(16), 3239–3249. https://doi.org/10.1093/hmg/ddt178

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., O'Donnell-Luria, A. H., Ware, J. S., Hill, A. J., Cummings, B. B., Tukiainen, T., Birnbaum, D. P., Kosmicki, J. A., Duncan, L. E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., Cooper, D. N., Deflaux, N., DePristo, M., Do, R., Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Kurki, M. I., Moonshine, A. L., Natarajan, P., Orozco, L., Peloso, G. M., Poplin, R., Rivas, M. A., Ruano-Rubio, V., Rose, S. A., Ruderfer, D. M., Shakir, K., Stenson, P. D., Stevens, C., Thomas, B. P., Tiao, G., Tusie-Luna, M. T., Weisburd, B., Won, H. H., Yu, D., Altshuler, D. M., Ardissino, D., Boehnke, M., Danesh, J., Donnelly, S., Elosua, R., Florez, J. C., Gabriel, S. B., Getz, G., Glatt, S. J., Hultman, C. M., Kathiresan, S., Laakso, M., McCarroll, S., McCarthy, M. I., McGovern, D., McPherson, R., Neale, B. M., Palotie, A., Purcell, S. M., Saleheen, D., Scharf, J. M., Sklar, P., Sullivan, P. F., Tuomilehto, J., Tsuang, M. T., Watkins, H.C., Wilson, J. G., Daly, M. J., MacArthur, D. G., Abboud, H. E., Abecasis, G., Aguilar-Salinas, C. A., Arellano-Campos, O., Atzmon, G., Aukrust, I., Barr, C. L., Bell, G. I., Bergen, S., Bjørkhaug, L., Blangero, J., Bowden, D. W., Budman, C. L., Burtt, N. P., Centeno-Cruz, F., Chambers, J. C., Chambert, K., Clarke, R., Collins, R., Coppola, G., Córdova, E. J., Cortes, M. L., , Cox, N. J., Duggirala, R., Farrall, M., Fernandez-Lopez, J. C., Fontanillas, P., Frayling, T. M., Freimer, N. B., Fuchsberger, C., García-Ortiz, H., , Goel, A., Gómez-Vázquez, M. J., González-Villalpando, M. E., González-Villalpando, C., Grados, M. A., Groop, L., Haiman, C. A., Hanis, C. L., Hattersley, A. T., Henderson, B. E., Hopewell, J. C., Huerta-Chagoya, A., Islas-Andrade, S., Jacobs, S. B., Jalilzadeh, S., Jenkinson, C. P., Moran, J., Jiménez-Morale, S., Kähler, A., King, R. A., Kirov, G., Kooner, J. S., Kyriakou, T., Lee, J. Y., Lehman, D. M., Lyon, G., MacMahon, W., Magnusson, P. K., Mahajan, A., Marrugat, J., Martínez-Hernández, A., Mathews, C. A., McVean, G., Meigs, J. B., Meitinger, T., Mendoza-Caamal, E., Mercader, J. M., Mohlke, K. L., Moreno-Macías, H., Morris, A. P., Najmi, L. A., Njølstad, P. R., O'Donovan, M. C., Ordóñez-Sánchez, M. L., Owen, M. J., Park, T., Pauls, D. L., Posthuma, D., Revilla-Monsalve, C., Riba, L., Ripke, S., Rodríguez-Guillén, R., Rodríguez-Torres, M., Sandor, P., Seielstad, M., Sladek, R., Soberón, X., Spector, T. D., Tai,, S. E., Teslovich, T. M., Walford, G., Wilkens, L. R., & Williams, A. L. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, *536*(7616), 285–291. https://doi.org/10.1038/nature19057

Levine, A., Griffiths, A., Markowitz, J., Wilson, D. C., Turner, D., Russell, R. K., Fell, J., Ruemmele, F. M., Walters, T., Sherlock, M., Dubinsky, M., & Hyams, J. S. (2011). Pediatric modification of the Montreal classification for inflammatory bowel disease: The Paris classification. *Inflammatory Bowel Diseases*, *17*(6), 1314–1321. https://doi.org/10.1002/ibd.21493

Levine, A., Kugathasan, S., Annese, V., Biank, V., Leshinsky-Silver, E., Davidovich, O., Kimmel, G., Shamir, R., Orazio, P., Karban, A., Broeckel, U., & Cucchiara, S. (2007). Pediatric onset crohn's colitis is characterized by genotype-dependent age-related susceptibility. *Inflammatory Bowel Diseases*, *13*(12), 1509–1515. https://doi.org/10.1002/ibd.20244

Levy, J., Espanol-Boren, T., Thomas, C., Fischer, A., Tovo, P., Bordigoni, P., Resnick, I., Fasth, A., Baer, M., Gomez, L., Sanders, E. A. M., Tabone, M.-D., Plantaz, D., Etzioni, A., Monafo, V., Abinun, M., Hammarstrom, L., Abrahamsen, T., Jones, A., Finn,A., Klemola, T., DeVries, E., Sanal. O., Peitsch, M. C., & Notarangelo, L. D. (1997). Clinical spectrum of X-linked hyper-IgM syndrome. *The Journal of Pediatrics*, *131*(1), 47–54. https://doi.org/10.1016/S0022-3476(97)70123-9

Li, Q., Lee, C. H., Peters, L. A., Mastropaolo, L. A., Thoeni, C., Elkadri, A., Schwerd, T., Zhu, J., Zhang, B., Zhao, Y., Hao, K., Dinarzo, A., Hoffman, G., Kidd, B. A., Murchie, R., Al Adham, Z., Guo, C., Kotlarz, D., Cutz, E., Walters, T. D., Shouval, D. S., Curran, M., Dobrin, R., Brodmerkel, C., Snapper, S. B., Klein, C., Brumell, J. H., Hu, M., Nanan, R., Snanter-Nanan, B., Wong, M., Le

Deist, F., Haddad, E., Roifman, C. M., Deslandres, C., Griffiths, A. M., Gaskin, K. J., Uhlig, H. H., Schadt, E. E., & Muise, A. M. (2016). Variants in TRIM22 That Affect NOD2 Signaling Are Associated With Very-Early-Onset Inflammatory Bowel Disease. *Gastroenterology*, *150*(5), 1196–1207. https://doi.org/10.1053/j.gastro.2016.01.031

Lin, F., Spencer, D., Hatala, D. A., Levine, A. D., & Medof, M. E. (2004). Decay-Accelerating Factor Deficiency Increases Susceptibility to Dextran Sulfate Sodium-Induced Colitis: Role for Complement in Inflammatory Bowel Disease. *The Journal of Immunology*, *172*(6), 3836–3841. https://doi.org/10.4049/jimmunol.172.6.3836

Liu, J. Z., van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., Takahashi, A., Ripke, S., Lee, J. C., Jostins, L., Shah, T., Abedian, S., Cheon, J. H., Cho, J., Daryani, N. E., Franke, L., Fuyuno, Y., Hart, A., Juyal, R. C., Juyal, G., Kim, W. H., Morris, A. P., Poustchi, H., Newman, W. G., Midha, V., Orchard, T. R., Vahedi, H., Sood, A., Sung, J. J. Y., Malekzadeh, R., Westra, H. J., Yamazaki, K., Yang, S. K., Barrett, J. C., Franke, A., Alizadeh, B. Z., Parkes, M., B K, T., Daly, M. J., Kubo, M., Anderson, C. A., & Weersma, R. K. (2015). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature Genetics*, *47*(9), 979–986. https://doi.org/10.1038/ng.3359

Llewellyn, S. R., Britton, G. J., Contijoch, E. J., Vennaro, O. H., Mortha, A., Colombel, J.-F., Grinspan, A., Clemente, J. C., Merad, M., & Faith, J. J. (2018). Interactions Between Diet and the Intestinal Microbiota Alter Intestinal Permeability and Colitis Severity in Mice. *Gastroenterology*, *154*(4), 1037-1046.e2. https://doi.org/10.1053/j.gastro.2017.11.030

Lloyd-Price, J., Arze, C., Ananthakrishnan, A. N., Schirmer, M., Avila-Pacheco, J., Poon, T. W., Andrews, E., Ajami, N. J., Bonham, K. S., Brislawn, C. J., Casero, D., Courtney, H., Gonzalez, A., Graeber, T. G., Hall, A. B., Lake, K., Landers, C. J., Mallick, H., Plichta, D. R., Prasad, M., Rahnavard. G., Sauk, J., Shungin, D., Vázquez-Baeza, Y., White, R.,A. 3rd., Braun, J., Denson, L.A., Jansson, J. K., Knight, R., Kugathasan, S., McGovern, D. P. B., Petrosino, J. F., Stappenbeck, T.S., Winter, H. S., Clish, C. B., Franzosa, E. A., Vlamakis, H., Xavier. R,.J., & Huttenhower, C. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature*, *569*(7758), 655–662. https://doi.org/10.1038/s41586-019-1237-9

Lohr, N. J., Molleston, J. P., Strauss, K. A., Torres-Martinez, W., Sherman, E. A., Squires, R. H., Rider, N. L., Chikwava, K. R., Cummings, O. W., Morton, D. H., & Puffenberger, E. G. (2010). Human ITCH E3 Ubiquitin Ligase Deficiency Causes Syndromic Multisystem Autoimmune Disease. *American Journal of Human Genetics*, *86*(3), 447–453. https://doi.org/10.1016/j.ajhg.2010.01.028

Lozynska, L. Y., Plawski, A., Lozynska, M. R., Vytvytskyi, I., Lozynskyi, R. Y., Prokopchuk, N., & Tretiak, B. (2018). Variant of rare Hermansky - Pudlak syndrome associated with granulomatous colitis: Diagnostics, clinical course and treatment. *Experimental Oncology*, *40*(1), 73–78. https://doi.org/10.31768/2312-8852.2018.40(1):73-78

Lucas, C. L., Zhang, Y., Venida, A., Wang, Y., Hughes, J., McElwee, J., Butrick, M., Matthews, H., Price, S., Biancalana, M., Wang, X., Richards, M., Pozos, T., Barlan, I., Ozen, A., Rao, V. K., Su, H. C., & Lenardo, M. J. (2014). Heterozygous splice mutation in PIK3R1 causes human immunodeficiency with lymphoproliferation due to dominant activation of PI3K. *Journal of Experimental Medicine*, *211*(13), 2537–2547. https://doi.org/10.1084/jem.20141759

Luecken, M. D., & Theis, F. J. (2019). Current best practices in single-cell RNA-seq analysis: a tutorial. *Molecular Systems Biology*, *15*(6). https://doi.org/10.15252/msb.20188746

Lyons, J. J., Liu, Y., Ma, C. A., Yu, X., O'Connell, M. P., Lawrence, M. G., Zhang, Y., Karpe, K., Zhao, M., Siegel, A. M., Stone, K. D., Nelson, C., Jones, N., DiMaggio, T., Darnell, D. N., Mendoza-Caamal, E., Orozco, L., Hughes, J. D., McElwee, J., Holland, S. M., Freeman, A. F., & Milner, J. D. (2017). ERBIN deficiency links STAT3 and TGF-β pathway defects with atopy in humans. *The Journal of Experimental Medicine*, *214*(3), 669–680. https://doi.org/10.1084/jem.20161435

Lyons, J. J., Yu, X., Karpe, K. A., Treadwell, S. M., Ma, C. A., O'Connell, M. P., Sun, G., Hughes, J. D., Mehmet, H., McElwee, J., Holland, S. M., Freeman, A. F., & Milner, J. D. (2015). Convergence of Clinical and Cellular Phenotypes Among Patients with STAT3 and ERBB2IP Mutations. *Journal of Allergy and Clinical Immunology*, *135*(2), AB152. https://doi.org/10.1016/j.jaci.2014.12.1437

Ma, C. A., Xi, L., Cauff, B., DeZure, A., Freeman, A. F., Hambleton, S., Kleiner, G., Leahy, T. R., O'Sullivan, M., Makiya, M., O'Regan, G., Pittaluga, S., Niemela, J., Stoddard, J., Rosenzweig, S. D., Raffeld, M., Klion, A. D., & Milner, J. D. (2017). Somatic STAT5b gain-of-function mutations in early onset nonclonal eosinophilia, urticaria, dermatitis, and diarrhea. *Blood*, *129*(5), 650–653. https://doi.org/10.1182/blood-2016-09-737817

Ma, J. Y., Tong, J. L., & Ran, Z. H. (2016). Intestinal tuberculosis and Crohn's disease: Challenging differential diagnosis. *Journal of Digestive Diseases*, *17*(3), 155–161. https://doi.org/10.1111/1751-2980.12324

Maekawa, K., Yamada, M., Okura, Y., Sato, Y., Yamada, Y., Kawamura, N., & Ariga, T. (2010). X-linked agammaglobulinemia in a 10-year-old boy with a novel non-invariant splice-site mutation in Btk gene. *Blood Cells, Molecules, and Diseases*, *44*(4), 300–304. https://doi.org/10.1016/j.bcmd.2010.01.004

Mamula, P., Telega, G. W., Markowitz, J. E., Brown, K. A., Russo, P. A., Piccoli, D. A., & Baldassano, R. N. (2002). Inflammatory bowel disease in children 5 years of age and younger. *American Journal of Gastroenterology*, *97*(8), 2005–2010. https://doi.org/10.1016/S0002-9270(02)04272-7

Mao, L., Kitani, A., Similuk, M., Oler, A. J., Albenberg, L., Kelsen, J., Aktay, A., Quezado, M., Yao, M., Montgomery-Recht, K., Fuss, I. J., & Strober, W. (2018). Loss-of-function CARD8 mutation causes NLRP3 inflammasome activation and Crohn's disease. *Journal of Clinical Investigation*, *128*(5), 1793–1806. https://doi.org/10.1172/JCI98642

Mathieu, A.-L., Verronese, E., Rice, G. I., Fouyssac, F., Bertrand, Y., Picard, C., Chansel, M., Walter, J. E., Notarangelo, L. D., Butte, M. J., Nadeau, K. C., Csomos, K., Chen, D. J., Chen, K., Delgado, A., Rigal, C., Bardin, C., Schuetz, C., Moshous, D., Reumaux, H., Plenat, F., Phan, A., Zabot, M T.., Balme, B., Viel, S., Bienvenu, J., Cochat, P., van der Burg, M., Caux, C., Kemp, E. H., Rouvet, I., Malcus, C., Méritet, J. F., Lim, A., Crow, Y. J., Fabien, N., Ménétrier-Caux, C., De Villartay, J. P. Walzer, T., & Belot, A. (2015). PRKDC mutations associated with immunodeficiency, granuloma, and autoimmune regulator–dependent autoimmunity. *Journal of Allergy and Clinical Immunology*, *135*(6), 1578-1588.e5. https://doi.org/10.1016/j.jaci.2015.01.040

Matute, J. D., Arias, A. A., Wright, N. A. M., Wrobel, I., Waterhouse, C. C. M., Xing, J. L., Marchal, C. C., Stull, N. D., Lewis, D. B., Steele, M., Kellner, J. D., Yu, W., Meroueh, S. O., Nauseef, W. M., & Dinauer, M. C. (2009). A new genetic subgroup of chronic granulomatous disease with autosomal recessive mutations in p40 phox and selective defects in neutrophil NADPH oxidase activity. *Blood*, *114*(15), 3309–3315. https://doi.org/10.1182/blood-2009-07-231498

McCarl, C. A., Picard, C., Khalil, S., Kawasaki, T., Röther, J., Papolos, A., Kutok, J., Hivroz, C., LeDeist, F., Plogmann, K., Ehl, S., Notheis, G., Albert, M. H., Belohradsky, B. H., Kirschner, J., Rao, A., Fischer, A., & Feske, S. (2009). ORAI1 deficiency and lack of store-operated Ca2+ entry cause immunodeficiency, myopathy, and ectodermal dysplasia. *Journal of Allergy and Clinical Immunology*, *124*(6), 1311–1318. https://doi.org/10.1016/j.jaci.2009.10.007

McGhee, J., & Ginder, G. (1979). Mcghee.Pdf. In *Nature* (Vol. 280, Issue 5721, pp. 419–420). https://www.nature.com/articles/280419a0

McKinnon, M. L., Rozmus, J., Fung, S. Y., Hirschfeld, A. F., Del Bel, K. L., Thomas, L., Marr, N., Martin, S. D., Marwaha, A. K., Priatel, J. J., Tan, R., Senger, C., Tsang, A., Prendiville, J., Junker, A. K., Seear, M., Schultz, K. R., Sly, L. M., Holt, R. A., Patel, M. S., Friedman, J. M. & Turvey, S. E. (2014). Combined immunodeficiency associated with homozygous MALT1 mutations. *Journal of Allergy and Clinical Immunology*, *133*(5). https://doi.org/10.1016/j.jaci.2013.10.045

Meeths, M., Entesarian, M., Al-Herz, W., Chiang, S. C. C., Wood, S. M., Al-Ateeqi, W., Almazan, F., Boelens, J. J., Hasle, H., Ifversen, M., Lund, B., Van Den Berg, J. M., Gustafsson, B., Hjelmqvist, H., Nordenskjöld, M., Bryceson, Y. T., & Henter, J. I. (2010). Spectrum of clinical presentations in familial hemophagocytic lymphohistiocytosis type 5 patients with mutations in STXBP2. *Blood*, *116*(15), 2635–2643. https://doi.org/10.1182/blood-2010-05-282541

Meienberg, J., Bruggmann, R., Oexle, K., & Matyas, G. (2016). Clinical sequencing: is WGS the better WES? *Human Genetics*, *135*(3), 359–362. https://doi.org/10.1007/s00439-015-1631-9

Merico, D., Sharfe, N., Hu, P., Herbrick, J.-A., & Roifman, C. M. (2015). RelB deficiency causes combined immunodeficiency. *LymphoSign Journal*, *2*(3), 147–155. https://doi.org/10.14785/lpsn-2015-0005

Meyts, I., Bosch, B., Bolze, A., Boisson, B., Itan, Y., Belkadi, A., Pedergnana, V., Moens, L., Picard, C., Cobat, A., Bossuyt, X., Abel, L., & Casanova, J. L. (2016). Exome and genome sequencing for inborn errors of immunity. *Journal of Allergy and Clinical Immunology*, *138*(4), 957–969. https://doi.org/10.1016/j.jaci.2016.08.003

Minegishi, Y., Saito, M., Tsuchiya, S., Tsuge, I., Takada, H., Hara, T., Kawamura, N., Ariga, T., Pasic, S., Stojkovic, O., Metin, A., & Karasuyama, H. (2007). Dominant-negative mutations in the DNA-binding domain of STAT3 cause hyper-IgE syndrome. *Nature*, *448*(7157), 1058–1062. https://doi.org/10.1038/nature06096

Mistry, A., Scambler, T., Parry, D., Wood, M., Barcenas-Morales, G., Carter, C., Doffinger, R., & Savic, S. (2017). Glucose-6-Phosphatase Catalytic Subunit 3 (G6PC3) Deficiency Associated With Autoinflammatory Complications. *Frontiers in Immunology*, *8*(NOV), 1–7. https://doi.org/10.3389/fimmu.2017.01485

Mondot, S., Barreau, F., Al Nabhani, Z., Dussaillant, M., Le Roux, K., Doré, J., Leclerc, M., Hugot, J.-P., & Lepage, P. (2012). Altered gut microbiota composition in immune-impaired Nod2 −/− mice. *Gut*, *61*(4), 634–635. https://doi.org/10.1136/gutjnl-2011-300478

Mora, A. J., & Wolfsohn, D. M. (2011). The management of gastrointestinal disease in hermansky-pudlak syndrome. *Journal of Clinical Gastroenterology*, *45*(8), 700–702. https://doi.org/10.1097/MCG.0b013e3181fd2742

Mu, W., Lu, H. M., Chen, J., Li, S., & Elliott, A. M. (2016). Sanger Confirmation Is Required to Achieve Optimal Sensitivity and Specificity in Next-Generation Sequencing Panel Testing. *Journal of Molecular Diagnostics*, *18*(6), 923–932. https://doi.org/10.1016/j.jmoldx.2016.07.006

Muise, A. M., Snapper, S. B., & Kugathasan, S. (2012). The age of gene discovery in very early onset inflammatory bowel disease. *Gastroenterology*, *143*(2), 285–288. https://doi.org/10.1053/j.gastro.2012.06.025

Narumi, S., Amano, N., Ishii, T., Katsumata, N., Muroya, K., Adachi, M., Toyoshima, K., Tanaka, Y., Fukuzawa, R., Miyako, K., Kinjo, S., Ohga, S., Ihara, K., Inoue, H., Kinjo, T., Hara, T., Kohno, M., Yamada, S., Urano, H., Kitagawa, Y., Tsugawa, K., Higa, A., Miyawaki, M., Okutani, T.,Kizaki, Z., Hamada, H., Kihara, M., Shiga, K., Yamaguchi, T., Kenmochi, M., Kitajima, H., Fukami, M., Shimizu, A., Kudoh, J., Shibata, S., Okano, H., Miyake, N., Matsumoto, N., & Hasegawa, T. (2016). SAMD9 mutations cause a novel multisystem disorder, MIRAGE syndrome, and are associated with loss of chromosome 7. *Nature Genetics*, *48*(7), 792–797. https://doi.org/10.1038/ng.3569

Naviglio, S., Arrigo, S., Martelossi, S., Villanacci, V., Tommasini, A., Loganes, C., Fabretto, A., Vignola, S., Lonardi, S., & Ventura, A. (2014). Severe inflammatory bowel disease associated with congenital alteration of transforming growth factor beta signaling. *Journal of Crohn's and Colitis*, *8*(8), 770–774. https://doi.org/10.1016/j.crohns.2014.01.013

Nenci, A., Becker, C., Wullaert, A., Gareus, R., Van Loo, G., Danese, S., Huth, M., Nikolaev, A., Neufert, C., Madison, B., Gumucio, D., Neurath, M. F., & Pasparakis, M. (2007). Epithelial NEMO

links innate immunity to chronic intestinal inflammation. *Nature*, *446*(7135), 557–561. https://doi.org/10.1038/nature05698

Neves, J. F., Doffinger, R., Barcena-Morales, G., Martins, C., Papapietro, O., Plagnol, V., Curtis, J., Martins, M., Kumararatne, D., Cordeiro, A. I., Neves, C., Borrego, L. M., Katan, M., & Nejentsev, S. (2018). Novel PLCG2 Mutation in a Patient With APLAID and Cutis Laxa. *Frontiers in Immunology*, *9*(December), 2863. https://doi.org/10.3389/fimmu.2018.02863

Ng, P. C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Research*, *31*(13), 3812–3814. https://doi.org/10.1093/nar/gkg509

Ng, S. C., Shi, H. Y., Hamidi, N., Underwood, F. E., Tang, W., Benchimol, E. I., Panaccione, R., Ghosh, S., Wu, J. C. Y., Chan, F. K. L., Sung, J. J. Y., & Kaplan, G. G. (2017). Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *The Lancet*, *390*(10114), 2769–2778. https://doi.org/10.1016/S0140-6736(17)32448-0

Notarangelo, L. D., & Fleisher, T. A. (2017). Targeted strategies directed at the molecular defect: Toward precision medicine for select primary immunodeficiency disorders. *Journal of Allergy and Clinical Immunology*, *139*(3), 715–723. https://doi.org/10.1016/j.jaci.2017.01.004

Oh, S. H., Baek, J., Kim, K. M., Lee, E. J., Jung, Y., Lee, Y. J., Jin, H. S., Ye, B. D., Yang, S. K., Lee, J. K., Seo, E. J., Lim, H. T., Lee, I., & Song, K. (2015). Is whole exome sequencing clinically practical in the management of pediatric Crohn's disease? *Gut and Liver*, *9*(6), 767–775. https://doi.org/10.5009/gnl15176

Ohya, T., Yanagimachi, M., Iwasawa, K., Umetsu, S., Sogo, T., Inui, A., Fujisawa, T., & Ito, S. (2017). Childhood-onset inflammatory bowel diseases associated with mutation of Wiskott-Aldrich syndrome protein gene. *World Journal of Gastroenterology*, *23*(48), 8544–8552. https://doi.org/10.3748/wjg.v23.i48.5844

Okou, D. T., Mondal, K., Faubion, W. A., Kobrynski, L. J., Denson, L. A., Mulle, J. G., Ramachandran, D., Xiong, Y., Svingen, P., Patel, V., Bose, P., Waters, J. P., Prahalad, S., Cutler, D. J., Zwick, M. E., & Kugathasan, S. (2014). Exome Sequencing Identifies a Novel FOXP3 Mutation in a 2-Generation Family With Inflammatory Bowel Disease. *Journal of Pediatric Gastroenterology and Nutrition*, *58*(5), 561–568. https://doi.org/10.1097/MPG.0000000000000302

Oppmann, B., Lesley, R., Blom, B., Timans, J. C., Xu, Y., Hunte, B., Vega, F., Yu, N., Wang, J., Singh, K., Zonin, F., Vaisberg, E., Churakova, T., Liu, M. ru, Gorman, D., Wagner, J., Zurawski, S., Liu, Y. J., Abrams, J. S., Moore, K. W., Rennick, D., De Waal-Malefyt, R., Hannum, C., Bazan, J. F., & Kastelein, R. A. (2000). Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with biological activities similar as well as distinct from IL-12. *Immunity*, *13*(5), 715–725. https://doi.org/10.1016/S1074-7613(00)00070-4

Orholm, M., Iselius, L., Sorensen, T. I. A., Munkholm, P., Langholz, E., & Binder, V. (1993). Investigation of inheritance of chronic inflammatory bowel diseases by complex segregation analysis. *British Medical Journal*, *306*(6869), 20–24. https://doi.org/10.1136/bmj.306.6869.20

Ottaviano, G., Salvatore, S., Salvatoni, A., Martelossi, S., Ventura, A., & Naviglio, S. (2018). Ocular manifestations of paediatric inflammatory bowel disease: A systematic review and meta-analysis. In *Journal of Crohn's and Colitis* (Vol. 12, Issue 7). https://doi.org/10.1093/ecco-jcc/jjy029

Ozen, A., Comrie, W. A., Ardy, R. C., Conde, C. D., Dalgic, B., Beser, Ö. F., Morawski, A. R., Karakoc-Aydiner, E., Tutar, E., Baris, S., Ozcay, F., Serwas, N. K., Zhang, Y., Matthews, H. F., Pittaluga, S., Folio, L. R., Aksu, A. U., McElwee, J. J., Krolo, A., Kiykim, A., Baris, Z., Gulsan, M., Ogulur, I., Snapper, S. B., Houwen, R. H. J., Leavis, H. L., Ertem, D., Kain, R., Sari, S., Erkan, T., Su, H. C., Boztug, K. & Lenardo, M. J. (2017). CD55 deficiency, early-onset protein-losing enteropathy, and thrombosis. *New England Journal of Medicine*, *377*(1), 52–61. https://doi.org/10.1056/NEJMoa1615887

Pannicke, U., Baumann, B., Fuchs, S., Henneke, P., Rensing-Ehl, A., Rizzi, M., Janda, A., Hese, K., Schlesier, M., Holzmann, K., Borte, S., Laux, C., Rump, E. M., Rosenberg, A., Zelinski, T., Schrezenmeier, H., Wirth, T., Ehl, S., Schroeder, M. L., & Schwarz, K. (2013). Deficiency of innate and acquired immunity caused by an IKBKB mutation. *New England Journal of Medicine*, *369*(26), 2504–2514. https://doi.org/10.1056/NEJMoa1309199

Parlato, M., Charbit-Henrion, F., Pan, J., Romano, C., Duclaux-Loras, R., Le Du, M., Warner, N., Francalanci, P., Bruneau, J., Bras, M., Zarhrate, M., Bègue, B., Guegan, N., Rakotobe, S., Kapel, N., De Angelis, P., Griffiths, A. M., Fiedler, K., Crowley, E., Ruemmele, F., Muise, A. M., & Cerf-Bensussan, N. (2018). Human ALPI deficiency causes inflammatory bowel disease and highlights a key mechanism of gut homeostasis . *EMBO Molecular Medicine*, *10*(4), 1–12. https://doi.org/10.15252/emmm.201708483

Parry, D., Blumenthal, J., Tomar, R., Horowitz, S., Elder, M., & Gern, J. (1996). A 3-year-old boy with ZAP-70 deficiency, thrombocytopenia and ulcerative colitis. *Journal of Allergy and Clinical Immunology*, *97*(1), 390–390. https://doi.org/10.1016/S0091-6749(96)81048-7

Persson, P. G., Ahlbom, A., & Hellers, G. (1990). Inflammatory bowel disease and tobacco smoke - a case-control study. *Gut*, *31*(12), 1377–1381. https://doi.org/10.1136/gut.31.12.1377

Petersen, B.-S., August, D., Abt, R., Alddafari, M., Atarod, L., Baris, S., Bhavsar, H., Brinkert, F., Buchta, M., Bulashevska, A., Chee, R., Cordeiro, A. I., Dara, N., Dückers, G., Elmarsafy, A., Frede, N., Galal, N., Gerner, P., Glocker, E.-O., Goldacker, S., Hammermann, J., Hasselblatt, P., Havlicekova, Z., Hübscher, K., Jesenak, M., Karaca, N. E., Karakoc-Aydiner, E., Kharaghani, M. M., Kilic, S. S., Kiykim, A., Klein, C., Klemann, C., Kobbe, R., Kotlarz, D., Laass, M. W., Leahy, T. R., Mesdaghi, M., Mitton, S., Neves, J. F., Öztürk, B., Pereira, L. F., Rohr, J., Restrepo, J. L. R., Ruzaike, G., Saleh, N., Seneviratne, S., Senol, E., Speckmann, C., Tegtmeyer, D., Thankam, P., van der Werff Ten Bosch, J., von Bernuth, H., Zeissig, S., Zeissig, Y., Franke, A., & Grimbacher, B. (2017). Targeted Gene Panel Sequencing for Early-onset Inflammatory Bowel Disease and Chronic Diarrhea. *Inflammatory Bowel Diseases*, *0*(0), 1. https://doi.org/10.1097/MIB.0000000000001235

Picard, C., McCarl, C.-A., Papolos, A., Khalil, S., Lüthy, K., Hivroz, C., LeDeist, F., Rieux-Laucat, F., Rechavi, G., Rao, A., Fischer, A., & Feske, S. (2009). STIM1 Mutation Associated with a Syndrome of Immunodeficiency and Autoimmunity. *New England Journal of Medicine*, *360*(19), 1971–1980. https://doi.org/10.1056/NEJMoa0900082

Piovani, D., Danese, S., Peyrin-Biroulet, L., Nikolopoulos, G. K., Lytras, T., & Bonovas, S. (2019). Environmental Risk Factors for Inflammatory Bowel Diseases: An Umbrella Review of Meta-analyses. *Gastroenterology*, *157*(3), 647-659.e4. https://doi.org/10.1053/j.gastro.2019.04.016

Podolsky, D. K. (2002). Inflammatory Bowel Disease. *New England Journal of Medicine*, *347*(6), 417–429. https://doi.org/10.1056/NEJMra020831

Prenzel, F., & Uhlig, H. H. (2009). Frequency of indeterminate colitis in children and adults with IBD - a metaanalysis. *Journal of Crohn's and Colitis*, *3*(4), 277–281. https://doi.org/10.1016/j.crohns.2009.07.001

Puel, A., Ziegler, S. F., Buckley, R. H., & Leonard, W. J. (1998). Defective IL7R expression in T-B+NK+ severe combined immunodeficiency. *Nature Genetics*, *20*(4), 394–397. https://doi.org/10.1038/3877

Quartier, P., Bustamante, J., Sanal, O., Plebani, A., Debré, M., Deville, A., Litzman, J., Levy, J., Fermand, J. P., Lane, P., Horneff, G., Aksu, G., Yalçin, I., Davies, G., Tezcan, I., Ersoy, F., Catalan, N., Imai, K., Fischer, A., & Durandy, A. (2004). Clinical, immunologic and genetic analysis of 29 patients with autosomal recessive hyper-IgM syndrome due to Activation-Induced Cytidine Deaminase deficiency. *Clinical Immunology*, *110*(1), 22–29. https://doi.org/10.1016/j.clim.2003.10.007

Ramos, P. S., Criswell, L. A., Moser, K. L., Comeau, M. E., Williams, A. H., Pajewski, N. M., Chung, S. A., Graham, R. R., Zidovetzki, R., Kelly, J. A., Kaufman, K. M., Jacob, C. O., Vyse, T. J., Tsao, B. P., Kimberly, R. P., Gaffney, P. M., Alarcón-Riquelme, M. E., Harley, J. B., & Langefeld, C. D. (2011). A comprehensive analysis of shared loci between systemic lupus erythematosus (SLE) and sixteen autoimmune diseases reveals limited genetic overlap. *PLoS Genetics*, *7*(12). https://doi.org/10.1371/journal.pgen.1002406

Ranson, N., Veldhuis, M., Mitchell, B., Fanning, S., Cook, A. L., Kunde, D., & Eri, R. (2019). NLRP3-dependent and-independent processing of interleukin (IL)-1β in active ulcerative colitis. *International Journal of Molecular Sciences*, *20*(1). https://doi.org/10.3390/ijms20010057

Recio, M. J., Moreno-Pelayo, M. A., Kiliç, S. S., Guardo, A. C., Sanal, O., Allende, L. M., Pérez-Flores, V., Mencía, A., Modamio-Høybjør, S., Seoane, E., & Regueiro, J. R. (2007). Differential Biological Role of CD3 Chains Revealed by Human Immunodeficiencies. *The Journal of Immunology*, *178*(4), 2556–2564. https://doi.org/10.4049/jimmunol.178.4.2556

Requena, D., Maffucci, P., Bigio, B., Shang, L., Abhyankar, A., Boisson, B., Stenson, P. D., Cooper, D. N., Cunningham-Rundles, C., Casanova, J. L., Abel, L., & Itan, Y. (2018). CDG: An online server for detecting biologically closest disease-causing genes and its application to primary immunodeficiency. *Frontiers in Immunology*, *9*(JUN). https://doi.org/10.3389/fimmu.2018.01340

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, *17*(5), 405–424. https://doi.org/10.1038/gim.2015.30

Rizvi, S. A., Elder, M., & Beasley, G. (2019). A Novel Manifestation of Prolidase Deficiency in a Toddler Diagnosed with Very-early-onset Crohn Disease. *Journal of Pediatric Gastroenterology and Nutrition*, *69*(3), e89–e90. https://doi.org/10.1097/MPG.0000000000002402

Roldan, R., Ruiz, A. M., Miranda, M. D., & Collantes, E. (2008). Anakinra: New therapeutic approach in children with Familial Mediterranean Fever resistant to colchicine. *Joint Bone Spine*, *75*(4), 504–505. https://doi.org/10.1016/j.jbspin.2008.04.001

Roos, D., Kuhns, D. B., Maddalena, A., Bustamante, J., Kannengiesser, C., de Boer, M., van Leeuwen, K., Köker, M. Y., Wolach, B., Roesler, J., Malech, H. L., Holland, S. M., Gallin, J. I., & Stasia, M.-J. (2010). Hematologically important mutations: The autosomal recessive forms of chronic granulomatous disease (second update). *Blood Cells, Molecules, and Diseases*, *44*(4), 291–299. https://doi.org/10.1016/j.bcmd.2010.01.009

Roussel, L., Landekic, M., Golizeh, M., Gavino, C., Zhong, M. C., Chen, J., Faubert, D., Blanchet-Cohen, A., Dansereau, L., Parent, M. A., Marin, S., Luo, J., Le, C., Ford, B. R., Langelier, M., King, I. L., Divangahi, M., Foulkes, W. D., Veillette, A., & Vinh, D. C. (2018). Loss of human ICO SL results in combined immunodeficiency. *Journal of Experimental Medicine*, *215*(12), 3151–3164. https://doi.org/10.1084/jem.20180668

Rubin, S. J. S., Bai, L., Haileselassie, Y., Garay, G., Yun, C., Becker, L., Streett, S. E., Sinha, S. R., & Habtezion, A. (2019). Mass cytometry reveals systemic and local immune signatures that distinguish inflammatory bowel diseases. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-10387-7

Ruel, J., Ruane, D., Mehandru, S., Gower-Rousseau, C., & Colombel, J. F. (2014). IBD across the age spectrum - Is it the same disease? *Nature Reviews Gastroenterology and Hepatology*, *11*(2), 88–98. https://doi.org/10.1038/nrgastro.2013.240

Ruemmele, F. M., El Khoury, M. G., Talbotec, C., Maurage, C., Mougenot, J. F., Schmitz, J., & Goulet, O. (2006). Characteristics of inflammatory bowel disease with onset during the first year of life. *Journal of Pediatric Gastroenterology and Nutrition*, *43*(5), 603–609. https://doi.org/10.1097/01.mpg.0000237938.12674.e3

Saito, D., Hibi, N., Ozaki, R., Kikuchi, O., Sato, T., Tokunaga, S., Minowa, S., Ikezaki, O., Mitsui, T., Miura, M., Sakuraba, A., Hayashida, M., Miyoshi, J., Matsuura, M., Nakase, H., & Hisamatsu, T. (2019). MEFV Gene-Related Enterocolitis Account for Some Cases Diagnosed as Inflammatory Bowel Disease Unclassified. *Digestion*. https://doi.org/10.1159/000502640

Salzer, E., Kansu, A., Sic, H., Májek, P., Ikincioğullari, A., Dogu, F. E., Prengemann, N. K., Santos-Valente, E., Pickl, W. F., Bilic, I., Ban, S. A., Kuloğlu, Z., Demir, A. M., Ensari, A., Colinge, J., Rizzi, M., Eibel, H., & Boztug, K. (2014). Early-onset inflammatory bowel disease and common variable immunodeficiency-like disease caused by IL-21 deficiency. *Journal of Allergy and Clinical Immunology*, *133*(6). https://doi.org/10.1016/j.jaci.2014.02.034

Salzer, U., Bacchelli, C., Buckridge, S., Pan-Hammarström, Q., Jennings, S., Lougaris, V., Bergbreiter, A., Hagena, T., Birmelin, J., Plebani, A., Webster, A. D. B., Peter, H. H., Suez, D., Chapel, H., McLean-Tooke, A., Spickett, G. P., Anover-Sombke, S., Ochs, H. D., Urschel, S., Belohradsky, B. H., Ugrinovic, S., Kumararatne, D. S., Lawrence, T. C., Holm, A. M., Franco, J. L., Schulze, I., Schneider, P., Gertz, E. M., Schaffer, A. A., Hammarström, L., Thrasher, A. J., Gaspar, H. B. & Grimbacher, B. (2009). Relevance of biallelic versus monoallelic TNFRSF13B mutations in distinguishing disease-causing from risk-increasing TNFRSF13B variants in antibody deficiency syndromes. *Blood*, *113*(9), 1967–1976. https://doi.org/10.1182/blood-2008-02-141937

Samuels, M. E., Majewski, J., Alirezaie, N., Fernandez, I., Casals, F., Patey, N., Decaluwe, H., Gosselin, I., Haddad, E., Hodgkinson, A., Idaghdour, Y., Marchand, V., Michaud, J. L., Rodrigue, M. A., Desjardins, S., Dubois, S., Le Deist, F., Awadalla, P., Raymond, V., & Maranda, B. (2013). Exome sequencing identifies mutations in the gene TTC7A in French-Canadian cases with hereditary multiple intestinal atresia. *Journal of Medical Genetics*, *50*(5), 324–329. https://doi.org/10.1136/jmedgenet-2012-101483

Sanal, O., Jing, H., Ozgur, T., Ayvaz, D., Strauss-Albee, D. M., Ersoy-Evans, S., Tezcan, I., Turkkani, G., Matthews, H. F., Haliloglu, G., Yuce, A., Yalcin, B., Gokoz, O., Oguz, K. K., & Su, H. C. (2012). Additional diverse findings expand the clinical presentation of DOCK8 deficiency. *Journal of Clinical Immunology*, *32*(4), 698–708. https://doi.org/10.1007/s10875-012-9664-5

Saudi Mendeliome Group. (2015). Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. *Genome Biology*, *16*(1). https://doi.org/10.1186/s13059-015-0693-2

Segal, A. W., Cross, A. R., Garcia, R. C., Borregaard, N., Valerius, N. H., Soothill, J. F., & Jones, O. T. G. (1983). Absence of Cytochrome b -245 in Chronic Granulomatous Disease. *New England Journal of Medicine*, *308*(5), 245–251. https://doi.org/10.1056/NEJM198302033080503

Seibold, F., Weber, P., Schning, A., Mork, H., Goppel, S., & Scheurlen, M. (1996). Neutrophil antibodies (pANCA) in chronic liver disease and inflammatory bowel disease: do they react with different antigens? *European Journal of Gastroenterology & Hepatology*, *8*(11), 1095–1100. https://doi.org/10.1097/00042737-199611000-00012

Seumois, G., & Vijayanand, P. (2019). Single-cell analysis to understand the diversity of immune cell types that drive disease pathogenesis. *Journal of Allergy and Clinical Immunology*, *144*(5), 1150–1153. https://doi.org/10.1016/j.jaci.2019.09.014

Shanon, A., Levin, S., Holtzman, F., Brock-Sinai, F., & Abu-Said, A. (1988). Combined familial adenosine deaminase and purine nucleoside phosphorylase deficiencies. *Archives of Disease in Childhood*, *63*(8), 931–934. https://doi.org/10.1136/adc.63.8.931

Sherry, S. T., Ward, M., & Sirotkin, K. (1999). dbSNP - database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Research*, *9*(8), 677–679. https://doi.org/10.1101/gr.9.8.677

Shizuma, T. (2016). Concomitant thyroid disorders and inflammatory bowel disease: A literature review. *BioMed Research International*, *2016*. https://doi.org/10.1155/2016/5187061

Sharma, D., Malik, A., Guy, C. S., Karki, R., Vogel, P., & Kanneganti, T.-D. (2018). Pyrin Inflammasome Regulates Tight Junction Integrity to Restrict Colitis and Tumorigenesis. *Gastroenterology*, *154*(4), 948-964.e8. https://doi.org/10.1053/j.gastro.2017.11.276

Schepp, J., Atkinson, C., Qiao, F., Spearman, J., Gilkeson, G., & Tomlinson, S. (2009). Role of complement and the use of complement Inhibitors in the Acute DSS Model of IBD. *Inflammatory Bowel Diseases*, *15*(1), S31. https://doi.org/10.1097/00054725-200912002-00090

Schlapbach, L. J., Thiel, S., Kessler, U., Ammann, R. A., Aebi, C., & Jensenius, J. C. (2011). Congenital H-ficolin deficiency in premature infants with severe necrotising enterocolitis. *Gut*, *60*(10), 1438–1439. https://doi.org/10.1136/gut.2010.226027

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). Mutationtaster2: Mutation prediction for the deep-sequencing age. *Nature Methods*, *11*(4), 361–362. https://doi.org/10.1038/nmeth.2890

Schwerd, T., Bryant, R. V., Pandey, S., Capitani, M., Meran, L., Cazier, J. B., Jung, J., Mondal, K., Parkes, M., Mathew, C. G., Fiedler, K., McCarthy, D., WGS500 Consortium, Oxford IBD cohort study investigators, COLORS in IBD, UK IBD Genetics Consortium, Sullivan, P. B., Rodrigues, A., Travis, S. P. L., Moore, C., Sambrook, J., Ouwehand, W. H., Roberts, D. J., Danesh, J., INTERVAL Study, Russell, R. K., Wilson, D. C., Kelsen, J. R., Cornall, R., Denson, L. A., Kugathasan, S., Knaus, U. G., Serra, E. G., Anderson, C. A., Duerr, R. H., McGovern, D. P.B., Cho, J., Powrie, Fiona, Li, V. S.W., Muise, A. M., & Uhlig, H. H. (2018). NOX1 loss-of-function genetic variants in patients with inflammatory bowel disease. *Mucosal Immunology*, *11*(2), 562–574. https://doi.org/10.1038/mi.2017.74

Schwerd, T., Pandey, S., Yang, H. T., Bagola, K., Jameson, E., Jung, J., Lachmann, R. H., Shah, N., Patel, S. Y., Booth, C., Runz, H., Düker, G., Bettels, R., Rohrbach, M., Kugathasan, S., Chapel, H., Keshav, S., Elkadri, A., Platt, N., Muise, A. M., Koletzko, S., Xavier, R. J., Marquardt, T., Powrie, F., Wraith, J. E., Gyrd-Hansen, M., Platt, F. M., & Uhlig, H. H. (2017). Impaired antibacterial autophagy links granulomatous intestinal inflammation in Niemann-Pick disease type C1 and XIAP deficiency with NOD2 variants in Crohn's disease. *Gut*, *66*(6), 1060–1073. https://doi.org/10.1136/gutjnl-2015-310382

Sidiq, T., Yoshihama, S., Downs, I., & Kobayashi, K. S. (2016). Nod2: A critical regulator of ileal microbiota and Crohn's disease. *Frontiers in Immunology*, *7*(SEP), 18–20. https://doi.org/10.3389/fimmu.2016.00367

Silverberg, M. S., Satsangi, J., Ahmad, T., Arnott, I. D. R., Bernstein, C. N., Brant, S. R., Caprilli, R., Colombel, J. F., Gasche, C., Geboes, K., Jewell, D. P., Karban, A., Loftus, E. V., Peña, A. S., Riddell, R. H., Sachar, D. B., Schreiber, S., Steinhart, A. H., Targan, S. R., Vermeire, S., & Warren, B. F. (2005). Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Canadian Journal of Gastroenterology = Journal Canadien de Gastroenterologie*, *19 Suppl A*(September). https://doi.org/10.1155/2005/269076

Sina, C., Kemper, C., & Derer, S. (2018). The intestinal complement system in inflammatory bowel disease: Shaping intestinal barrier function. *Seminars in Immunology*, *37*(February), 66–73. https://doi.org/10.1016/j.smim.2018.02.008

Sokol, H., Seksik, P., Furet, J. P., Firmesse, O., Nion-Larmurier, I., Beaugerie, L., Cosnes, J., Corthier, G., Marteau, P., & Doraé, J. (2009). Low counts of faecalibacterium prausnitzii in colitis microbiota. *Inflammatory Bowel Diseases*, *15*(8), 1183–1189. https://doi.org/10.1002/ibd.20903

Speckmann, C., Lehmberg, K., Albert, M. H., Damgaard, R. B., Fritsch, M., Gyrd-Hansen, M., Rensing-Ehl, A., Vraetz, T., Grimbacher, B., Salzer, U., Fuchs, I., Ufheil, H., Belohradsky, B. H., Hassan, A., Cale, C. M., Elawad, M., Strahm, B., Schibli, S., Lauten, M., Kohl, M., Meerpohl, J. J., Rodeck, B., Kolb, R., Eberl, W., Soerensen, J., von Bernuth, H., Lorenz, M., Schwarz, K., zur Stadt, U., & Ehl, S. (2013). X-linked inhibitor of apoptosis (XIAP) deficiency: The spectrum of presenting

manifestations beyond hemophagocytic lymphohistiocytosis. *Clinical Immunology*, *149*(1), 133–141. https://doi.org/10.1016/j.clim.2013.07.004

Starokadomskyy, P., Gemelli, T., Rios, J. J., Xing, C., Wang, R. C., Li, H., Pokatayev, V., Dozmorov, I., Khan, S., Miyata, N., Fraile, G., Raj, P., Xu, Z., Xu, Z., Ma, L., Lin, Z., Wang, H., Yang, Y., Ben-Amitai, D., Orenstein, N., Mussaffi, H., Baselga, E., Tadini, G., Grunebaum, E., Sarajlija, A., Krzewski, K., Wakeland, E. K., Yan, N., de la Morena, M. T., Zinn, A.R., & Burstein, E. (2016). DNA polymerase-alpha regulates type I interferon activation through cytosolic RNA:DNA synthesis. *Nat Immunology*, *17(5)*(5), 495–504. https://doi.org/10.1038/ni.3409.DNA

Stengaard-Pedersen, K., Thiel, S., Gadjeva, M., Møller-Kristensen, M., Sørensen, R., Jensen, L. T., Sjöholm, A. G., Fugger, L., & Jensenius, J. C. (2003). Inherited deficiency of mannan-binding lectin-associated serine protease 2. *New England Journal of Medicine*, *349*(6), 554–560. https://doi.org/10.1056/NEJMoa022836

Sun, W., Han, X., Wu, S., & Yang, C. (2016). Tonsillectomy and the risk of inflammatory bowel disease: A systematic review and meta-analysis. *Journal of Gastroenterology and Hepatology (Australia)*, *31*(6), 1085–1094. https://doi.org/10.1111/jgh.13273

Sun, X., He, S., Lv, C., Sun, X., Wang, J., Zheng, W., & Wang, D. (2017). Analysis of murine and human Treg subsets in inflammatory bowel disease. *Molecular Medicine Reports*, *16*(3), 2893–2898. https://doi.org/10.3892/mmr.2017.6912

Sun, Y., Li, L., Xia, Y., Li, W., Wang, K., Wang, L., Miao, Y., & Ma, S. (2019). The gut microbiota heterogeneity and assembly changes associated with the IBD. *Scientific Reports*, *9*(1), 1–8. https://doi.org/10.1038/s41598-018-37143-z

Sura, S. P., Ahmed, A., Cheifetz, A. S., & Moss, A. C. (2014). Characteristics of Inflammatory Bowel Disease Serology in Patients With Indeterminate Colitis. *Journal of Clinical Gastroenterology*, *48*(4), 351–355. https://doi.org/10.1097/MCG.0000000000000083

Tangye, S. G., Al-Herz, W., Bousfiha, A., Chatila, T., Cunningham-Rundles, C., Etzioni, A., Franco, J. L., Holland, S. M., Klein, C., Morio, T., Ochs, H. D., Oksenhendler, E., Picard, C., Puck, J., Torgerson, T. R., Casanova, J. L., & Sullivan, K. E. (2020). Human Inborn Errors of Immunity: 2019 Update on the Classification from the International Union of Immunological Societies Expert Committee. *Journal of Clinical Immunology*, *40*(1), 24–64. https://doi.org/10.1007/s10875-019-00737-x

Teimourian, S., Zomorodian, E., Badalzadeh, M., Pouya, A., Kannengiesser, C., Mansouri, D., Cheraghi, T., & Parvaneh, N. (2008). Characterization of six novel mutations in CYBA: The gene causing autosomal recessive chronic granulomatous disease. *British Journal of Haematology*, *141*(6), 848–851. https://doi.org/10.1111/j.1365-2141.2008.07148.x

Telakis, E., & Tsironi, E. (2008). Indeterminate colitis - Definition, diagnosis, characteristics and management. *Annals of Gastroenterology*, *21*(3), 173–179.

The 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. https://doi.org/10.1038/nature15393

Theochari, N. A., Stefanopoulos, A., Mylonas, K. S., & Economopoulos, K. P. (2018). Antibiotics exposure and risk of inflammatory bowel disease: a systematic review. *Scandinavian Journal of Gastroenterology*, *53*(1), 1–7. https://doi.org/10.1080/00365521.2017.1386711

Thoeni, C., Hamilton, E., Elkadri, A., Murchie, R., Fiedler, K., Ovadia, A., Sharfe, N., Ngan, B. Y., Cutz, E., Nahum, A., Muise, A., & Roifman, C. M. (2016). The effects of STAT1 dysfunction on the gut. *LymphoSign Journal*, *3*, 19–33. https://doi.org/10.14785/lymphosign-2015-0012

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, *14*(2), 178–192. https://doi.org/10.1093/bib/bbs017

Uhlig, H. H., & Muise, A. M. (2017). Clinical Genomics in Inflammatory Bowel Disease. *Trends in Genetics*, *33*(9), 629–641. https://doi.org/10.1016/j.tig.2017.06.008

Uhlig, H. H., & Schwerd, T. (2016). From Genes to Mechanisms: The Expanding Spectrum of Monogenic Disorders Associated with Inflammatory Bowel Disease. *Inflammatory Bowel Diseases*, *22*(1), 202–212. https://doi.org/10.1097/MIB.0000000000000614

Uhlig, H. H., Schwerd, T., Koletzko, S., Shah, N., Kammermeier, J., Elkadri, A., Ouahed, J., Wilson, D. C., Travis, S. P., Turner, D., Klein, C., Snapper, S. B., & Muise, A. M. (2014). The diagnostic approach to monogenic very early onset inflammatory bowel disease. *Gastroenterology*, *147*(5), 990–1007. https://doi.org/10.1053/j.gastro.2014.07.023

Umeno, J., Hisamatsu, T., Esaki, M., Hirano, A., Kubokura, N., Asano, K., Kochi, S., Yanai, S., Fuyuno, Y., Shimamura, K., Hosoe, N., Ogata, H., Watanabe, T., Aoyagi, K., Ooi, H., Watanabe, K., Yasukawa, S., Hirai, F., Matsui, T., Iida, M., Yao, T., Hibi, T., Kosaki, K., Kanai, T., Kitazono, T., & Matsumoto, T. (2015). A Hereditary Enteropathy Caused by Mutations in the SLCO2A1 Gene, Encoding a Prostaglandin Transporter. *PLoS Genetics*, *11*(11), 1–15. https://doi.org/10.1371/journal.pgen.1005581

Uslu, N., Yüce, A., Demir, H., Saltik-Temizel, I. N., Usta, Y., Yilmaz, E., Beşbaş, N., Gürakan, F., Özen, H., & Özen, S. (2010). The association of inflammatory bowel disease and Mediterranean fever gene (MEFV) mutations in Turkish children. *Digestive Diseases and Sciences*, *55*(12), 3488–3494. https://doi.org/10.1007/s10620-010-1178-5

Uzel, G., Kleiner, D. E., Kuhns, D. B., & Holland, S. M. (2001). Dysfunctional LAD-1 neutrophils and colitis. *Gastroenterology*, *121*(4), 958–964. https://doi.org/10.1053/gast.2001.28022

Uzel, G., Sampaio, E. P., Lawrence, M. G., Hsu, A. P., Hackett, M., Dorsey, M. J., Noel, R. J., Verbsky, J. W., Freeman, A. F., Janssen, E., Bonilla, F. A., Pechacek, J., Chandrasekaran, P., Browne, S. K., Agharahimi, A., Gharib, A. M., Mannurita, S. C., Joon Yim, J., Gambineri, E., Torgerson, T., Tran, D. Q., Milner, J. D., & Holland, S. M. (2013). Dominant gain-of-function STAT1 mutations in FOXP3 WT IPEX- like Syndrome. *J Allergy Clin Immunol.*, *131*(6), 1611–1623. https://doi.org/10.1016/j.jaci.2012.11.054

Van De Vosse, E., Van Wengen, A., Van Geelen, J. A., De Boer, M., Roos, D., & Van Dissel, J. T. (2009). A novel mutation in NCF1 in an adult CGD patient with a liver abscess as first presentation. *Journal of Human Genetics*, *54*(6), 313–316. https://doi.org/10.1038/jhg.2009.24

Van Der Burg, M., Van Veelen, L. R., Verkaik, N. S., Wiegant, W. W., Hartwig, N. G., Barendregt, B. H., Brugmans, L., Raams, A., Jaspers, N. G. J., Zdzienicka, M. Z., Van Dongen, J. J. M., & Van Gent, D. C. (2006). A new type of radiosensitive T-B-NK+ severe combined immunodeficiency caused by a LIG4 mutation. *Journal of Clinical Investigation*, *116*(1), 137–145. https://doi.org/10.1172/JCI26121

Van Haaften-Visser, D. Y., Harakalova, M., Mocholi, E., Van Montfrans, J. M., Elkadri, A., Rieter, E., Fiedler, K., Van Hasselt, P. M., Triffaux, E. M. M., Van Haelst, M. M., Nijman, I. J., Kloosterman, W. P., Nieuwenhuis, E. E. S., Muise, A. M., Cuppen, E., Houwen, R. H. J., & Coffer, P. J. (2017). Ankyrin repeat and zinc-finger domain-containing 1 mutations are associated with infantile-onset inflammatory bowel disease. *Journal of Biological Chemistry*, *292*(19), 7904–7920. https://doi.org/10.1074/jbc.M116.772038

Van Limbergen, J., Russell, R. K., Nimmo, E. R., Drummond, H. E., Smith, L., Davies, G., Anderson, N. H., Gillett, P. M., McGrogan, P., Hassan, K., Weaver, L., Bisset, W. M., Mahdi, G., Wilson, D. C., & Satsangi, J. (2007). IL23R Arg381Gln is associated with childhood onset inflammatory bowel disease in Scotland. *Gut*, *56*(8), 1173–1174. https://doi.org/10.1136/gut.2007.122069

Van Rheenen, P. F., Van De Vijver, E., & Fidler, V. (2010). Faecal calprotectin for screening of patients with suspected inflammatory bowel disease: Diagnostic meta-analysis. *BMJ (Online)*, *341*(7765), 188. https://doi.org/10.1136/bmj.c3369

VanValkenburgh, J., Albu, D. I., Bapanpally, C., Casanova, S., Califano, D., Jones, D. M., Ignatowicz, L., Kawamoto, S., Fagarasan, S., Jenkins, N. A., Copeland, N. G., Liu, P., & Avram, D. (2011). Critical role of Bcl11b in suppressor function of T regulatory cells and prevention of inflammatory bowel disease. *Journal of Experimental Medicine*, *208*(10), 2069–2081. https://doi.org/10.1084/jem.20102683

Vavricka, S. R., Brun, L., Ballabeni, P., Pittet, V., Prinz Vavricka, B. M., Zeitz, J., Rogler, G., & Schoepfer, A. M. (2011). Frequency and risk factors for extraintestinal manifestations in the swiss inflammatory bowel disease cohort. *American Journal of Gastroenterology*, *106*(1), 110–119. https://doi.org/10.1038/ajg.2010.343

Verbsky, J. W., & Chatila, T. A. (2013). Immune dysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) and IPEX-related disorders: An evolving web of heritable autoimmune diseases. *Current Opinion in Pediatrics*, *25*(6), 708–714. https://doi.org/10.1097/MOP.0000000000000029

Vigone, M. C., Fugazzola, L., Zamproni, I., Passoni, A., Di Candia, S., Chiumello, G., Persani, L., & Weber, G. (2005). Persistent mild hypothyroidism associated with novel sequence variants of the DUOX2 gene in two siblings. *Human Mutation*, *26*(4), 395. https://doi.org/10.1002/humu.9372

Villani, A.-C., Lemire, M., Fortin, G., Louis, E., Silverberg, M. S., Collette, C., Baba, N., Libioulle, C., Belaiche, J., Bitton, A., Gaudet, D., Cohen, A., Langelier, D., Fortin, P. R., Wither, J. E., Sarfati, M., Rutgeerts, P., Rioux, J. D., Vermeire, S., Hodson, T. J., & Franchimont, D. (2009). *Common Variants in the NLRP3 Region Contribute to Crohn's Disease Susceptibility*. *41*(1), 71–76. https://doi.org/10.1038/ng285

Vivinus-Nébot, M., Frin-Mathy, G., Bzioueche, H., Dainese, R., Bernard, G., Anty, R., Filippi, J., Saint-Paul, M. C., Tulic, M. K., Verhasselt, V., Hébuterne, X., & Piche, T. (2014). Functional bowel symptoms in quiescent inflammatory bowel diseases: Role of epithelial barrier disruption and low-grade inflammation. *Gut*, *63*(5), 744–752. https://doi.org/10.1136/gutjnl-2012-304066

Volz, M. S., Nassir, M., Treese, C., von Winterfeld, M., Plöckinger, U., Epple, H. J., & Siegmund, B. (2015). Inflammatory bowel disease (IBD)-like disease in a case of a 33-year old man with glycogenosis 1b. *BMC Gastroenterology*, *15*(1), 1–4. https://doi.org/10.1186/s12876-015-0271-9

Waite, A. L., Schaner, P., Richards, N., Balci-Peynircioglu, B., Masters, S. L., Brydges, S. D., Fox, M., Hong, A., Yilmaz, E., Kastner, D. L., Reinherz, E. L., & Gumucio, D. L. (2009). Pyrin modulates the intracellular distribution of PSTPIP1. *PLoS ONE*, *4*(7), 1–15. https://doi.org/10.1371/journal.pone.0006147

Wang, L. L., Zhou, W., Zhao, W., Tian, Z. Q., Wang, W. F., Wang, X. F., & Chen, T. X. (2014). Clinical features and genetic analysis of 20 Chinese patients with X-linked hyper-IgM syndrome. *Journal of Immunology Research*, *2014*. https://doi.org/10.1155/2014/683160

Yang, X., Miyawaki, T., & Kanegane, H. (2012). SAP and XIAP deficiency in hemophagocytic lymphohistiocytosis. *Pediatrics International*, *54*(4), 447–454. https://doi.org/10.1111/j.1442-200X.2012.03683.x

Yang, Y., Muzny, D. M., Reid, J. G., Bainbridge, M. N., Willis, A., Ward, P. A., Braxton, A., Beuten, J., Xia, F., Niu, Z., Hardison, M., Person, R., Bekheirnia, M. R., Leduc, M. S., Kirby, A., Pham, P., Scull, J., Wang, M., Ding, Y., Plon, S. E., Lupski, J. R., Beaudet, A. L., Gibbs, R. A., & Eng, C. M. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *New England Journal of Medicine*, *369*(16), 1502–1511. https://doi.org/10.1056/NEJMoa1306555

Yeshi, K., Ruscher, R., Hunter, L., Daly, N. L., Loukas, A., & Wangchuk, P. (2020). Revisiting Inflammatory Bowel Disease: Pathology, Treatments, Challenges and Emerging Therapeutics Including Drug Leads from Natural Products. *Journal of Clinical Medicine*, 9(5), 1273. https://doi.org/10.3390/jcm9051273

Yu, Y. R., & Rodriguez, J. R. (2017). Clinical presentation of Crohn's, ulcerative colitis, and indeterminate colitis: Symptoms, extraintestinal manifestations, and disease phenotypes. *Seminars in Pediatric Surgery*, *26*(6), 349–355. https://doi.org/10.1053/j.sempedsurg.2017.10.003

Zeissig, S., Petersen, B.-S., Tomczak, M., Melum, E., Huc-Claustre, E., Dougan, S. K., Laerdahl, J. K., Stade, B., Forster, M., Schreiber, S., Weir, D., Leichtner, A. M., Franke, A., & Blumberg, R. S. (2015). Early-onset Crohn's disease and autoimmunity associated with a variant in CTLA-4. *Gut*, *64*(12), 1889–1897. https://doi.org/10.1136/gutjnl-2014-308541

Zhang, R. J., Sun, F., Chen, F., Fang, Y., Yan, C. Y., Zhang, C. R., Ying, Y. X., Wang, Z., Zhang, C. X., Wu, F. Y., Han, B., Liang, J., Zhao, S. X., & Song, H. D. (2020). The TPO mutation screening and genotype-phenotype analysis in 230 Chinese patients with congenital hypothyroidism. *Molecular and Cellular Endocrinology*, *506*, 110761. https://doi.org/10.1016/j.mce.2020.110761

Zhang, Z., Gothe, F., Pennamen, P., James, J. R., McDonald, D., Mata, C. P., Modis, Y., Alazami, A. M., Acres, M., Haller, W., Bowen, C., Döffinger, R., Sinclair, J., Brothers, S., Zhang, Y., Matthews, H. F., Naudion, S., Pelluard, F., Alajlan, H., Yamazaki, Y., Notarangelo, L. D., Thaventhiran, J. E., Engelhardt, K. R., Al-Mousa, H., Hambleton, S., Rooryck, C., Smith, K. G. C., & Lenardo, M. J. (2019). Human interleukin-2 receptor b mutations associated with defects in immunity and peripheral tolerance. *Journal of Experimental Medicine*, *216*(6), 1311–1327. https://doi.org/10.1084/jem.20182304

Zheng, C., Huang, Y., Ye, Z., Wang, Y., Tang, Z., Lu, J., Wu, J., Zhou, Y., Wang, L., Huang, Z., Yang, H., & Xue, A. (2018). Infantile onset intractable inflammatory bowel disease due to novel heterozygous mutations in TNFAIP3 (A20). *Inflammatory Bowel Diseases*, *24*(12), 2613–2620. https://doi.org/10.1093/IBD/IZY165

Zhou, Q., Lee, G. S., Brady, J., Datta, S., Katan, M., Sheikh, A., Martins, M. S., Bunney, T. D., Santich, B. H., Moir, S., Kuhns, D. B., Priel, D. A. L., Ombrello, A., Stone, D., Ombrello, M. J., Khan, J., Milner, J. D., Kastner, D. L., & Aksentijevich, I. (2012). A hypermorphic missense mutation in PLCG2, encoding phospholipase Cγ2, causes a dominantly inherited autoinflammatory disease with immunodeficiency. *American Journal of Human Genetics*, *91*(4), 713–720. https://doi.org/10.1016/j.ajhg.2012.08.006

Žuklys, S., Handel, A., Zhanybekova, S., Govani, F., Keller, M., Maio, S., E Mayer, C., Ying Teh, H., Hafen, K., Gallone, G., Barthlott, T., Ponting, C. P., & Holländer, G. A. (2016). Foxn1 regulates key target genes essential for T cell development in postnatal thymic epithelial cells. *Nature Immunology*, *17*(10), 1206–1215. https://doi.org/10.1038/ni.3537

Online sources

ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/, accessed January 21, 2020)

Genomics England - gene list for Infantile enterocolitis & monogenic inflammatory bowel disease (https://panelapp.genomicsengland.co.uk/panels/176/, accessed March 30, 2020)

IDT OligoAnalyzer (https://www.idtdna.com/pages/tools/oligoanalyzer, accessed May 10, 2020)

Integrated Genomics Viewer (https://igv.org/, accessed May 28, 2020)

NCBI Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/, accessed May 10, 2020)

NHLBI Exome Sequencing Project (https://evs.gs.washington.edu/EVS/, accessed January 21, 2020)

Online Mendelian Inheritance in Man (OMIM, http://omim.org, accessed January 21, 2020)

OMIM #147060: Hyper-IgE recurrent infection syndrome (https://www.omim.org/entry/147060, accessed June 1, 2020)

OMIM #191900: Muckle-Wells syndrome (https://www.omim.org/entry/191900, accessed June 1, 2020)

OMIM #301220: Pigmentary disorder, reticulate, with systemic manifestations, X-linked (https://www.omim.org/entry/301220, accessed June 1, 2020)

OMIM #304790: Immunodysregulation, polyendocrinopathy, and enteropathy, X-linked (https://www.omim.org/entry/304790, accessed June 1, 2020)

OMIM #604416: Pyogenic sterile arthritis, pyoderma gangrenosum, and acne (https://www.omim.org/entry/604416, accessed June 1, 2020)

OMIM #614878: Autoinflammation, antibody deficiency, and immune dysregulation syndrome (https://www.omim.org/entry/614878, accessed June 1, 2020)

OMIM #615952: Autoimmune disease, multisystem, infantile-onset, 1 (https://www.omim.org/entry/615952, accessed June 1, 2020)

OMIM #617237: Immunodeficiency 49 (https://www.omim.org/entry/617237, accessed June 1, 2020)

OMIM #617765: Immunodeficiency, common variable, 14 (https://www.omim.org/entry/617765, accessed June 1, 2020)

Pediatric Cancer Genomic Data Portal (https://pecan.stjude.cloud/, accessed January 21, 2020)

QIAGEN Ingenuity Variant Analysis (https://digitalinsights.qiagen.com/products/qiagen-ingenuity-variant-analysis/, accessed March 25, 2020)

The Genome Aggregation Database (https://gnomad.broadinstitute.org/, accessed January 21, 2020)

The Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/, accessed January 21, 2020)

UCSC Genome Browser (https://genome.ucsc.edu/, accessed May 10, 2020)

UCSC In-silico PCR (https://genome.ucsc.edu/cgi-bin/hgPcr, accessed May 10, 2020)

List of supplementary files

Supplementary file 1 Gene lists for variant filtering

Supplementary file 2 Parameters of identified variants in patients

Supplementary file 3 Sequences of customized primers

Supplementary file 1: Gene lists for variant filtering

CDG GENE LIST:

A2M	CACNA1S	CTLA4	FEN1	IL6	MPO	POLR1C	STX11
ABCB7	CALM1	CUL3	FGFR1	IL6R	MRPS22	POLR1D	STXBP1
ABL2	CAMK4	CUL4B	FGFR3	IL7R	MS4A2	POLR3A	SUPT16H
ACACA	CASR	CXCR4	FGG	INS	MSH2	POMC	SYN1
ACTA2	CAV1	CYBA	FLT1	INSR	MSH6	PPARA	TARDBP
ACTB	CAVIN1	CYBB	FLT4	IRAK3	MUC1	PRICKLE1	TBP
ACTG1	CBL	CYLD	FN1	IRF1	MUSK	PRKACA	TDGF1
ACTN2	CBLB	DCLRE1C	FOXP3	IRF4	MYBPC3	PRKAG2	TERT
ACTN4	CCM2	DDB2	FZD4	IRF5	MYC	PRKAR1A	TGFB1
ACVR1	CCND1	DKC1	GAA	IRF6	MYH2	PRKCA	TGFB2
ADAM10	CCR2	DLL1	GALT	ITGA2B	MYH7	PRKCG	TGFB3
ADAR	CCR5	DMD	GATA1	ITGA4	MYO1C	PRKDC	TGFBR1
ADCY6	CD14	DNMT3B	GATA3	ITGA6	NCF2	PRNP	TGFBR2
ADD1	CD151	DRD4	GATA4	ITGA9	NDN	PROS1	THPO
ADRA2B	CD19	EDNRA	GDNF	ITGAM	NFKB1	PSTPIP1	TIMP3
ADRA2C	CD209	EDNRB	GH1	ITGB2	NFKBIA	PTPN1	TLR2
AGT	CD247	EEF2	GHR	ITGB4	NGF	PTPN11	TLR4
AICDA	CD3D	EFNB1	GNA11	ITCH	NGFR	PTPRC	TMEM173
AKT1	CD3G	EGF	GNAI2	ITK	NHEJ1	RAC1	TNFRSF1A
ALB	CD4	EGFR	GNAI3	KALRN	NHLRC1	RAC2	TOPORS
ALDOA	CD40	EHMT1	GNAT2	KCNA1	NOS3	RAF1	TOR1A
ALOX12B	CD40LG	EIF4E	GP1BA	KCNA5	NOTCH1	RAI1	TP53
ANXA5	CD79A	EIF4G1	GP1BB	KDR	NTRK1	RAPSN	TREM2
APC	CD79B	EPHA2	GPI	KLK4	NTRK2	RASA1	TRIM2
APEX1	CD8A	EPM2A	HBEGF	KNG1	NTRK3	RB1	TRIM32
APOB	CDC6	EPOR	HGF	KRT1	OAS1	RELN	TRPC6
APP	CDK4	ERBB2	HIF1A	LAMA2	OCRL	ROR2	TYK2
APTX	CDK5RAP2	ERBB3	HLA-A	LAMB1	ODC1	RPL11	TYROBP
AR	CDK6	ERCC1	HLA-DRB1	LAMC1	OLR1	RPS14	UBB
ARHGAP26	CDKN1B	ERCC2	HMOX1	LCK	ORC1	RPS19	UBE2A
ASNS	CDKN2A	ERCC3	HPS1	LEFTY2	OSMR	RSPO1	UMPS
ATXN2	CDON	ERCC4	HRAS	LEP	PABPN1	RSPO4	USP15
AURKA	CENPJ	ERCC5	HRG	LIF	PAFAH1B1	RUNX2	USP9X
B2M	CFD	ERCC6	HSPB1	LIFR	PAK3	RYR1	VAMP1
BAG3	CFL2	ERCC8	CHMP2B	LIG1	PALB2	RYR2	VAPB
BARD1	CFTR	ESR1	CHMP4B	LRP6	PCNT	SCNN1A	VAV3
BAX	CIAO1	EZH2	CHN1	LRPPRC	PDGFB	SCNN1B	VCL
BCL2	CNBP	F12	CHUK	LRRK2	PDGFRA	SCNN1G	VCP
BCL3	COL10A1	F13A1	ICAM1	LY96	PDGFRB	SDC3	VEGFA
BCR	COL11A1	F2	IFNAR2	MAD2L1	PER1	SERPINA1	VHL
BDNF	COL17A1	F5	IFNG	MAG	PER2	SERPINB6	VWF
BLK	COL4A2	FADD	IFNGR1	MBL2	PEX2	SERPINE1	WRN
BLM	COL4A3	FANCA	IFNGR2	MCM6	PHKG2	SERPING1	XIAP
BRAF	COL4A4	FANCB	IGF1	MCM8	PIK3CD	SH3BP2	XPA
BRCA2	COL4A5	FANCC	IGF1R	MDM2	PIK3R1	SLC22A5	XPC
BRIP1	COL4A6	FANCD2	IGF2	MED12	PIK3R5	SMAD3	XRCC4
BTK	COL7A1	FANCE	IKBKG	MEF2A	PLA2G4A	SMAD4	ZAP70
BUB1B	CR1	FANCF	IL10	MIF	PLG	SMAD7	
C1QA	CRBN	FANCG	IL10RA	MLH1	PLXNB1	SMPD1	
C1QB	CREB1	FANCI	IL12B	MLH3	PMS2	SOD1	
C3	CSF2RA	FANCL	IL13	MMP1	PNKP	SOS1	
C4A	CSNK1D	FANCM	IL1B	MMP2	PNPLA6	SPARC	
C4B	CTDP1	FCGR2A	IL2RA	MMP9	PNPT1	SPTBN2	
C5	CTF1	FCGR3A	IL2RG	MPL	POLD1	SSTR5	

PID GENE LIST:

ACD	CD27	DNAJC21	IL10RA	MCM9	PLCG2	SAMD9	TICAN1
ACP5	CD3D	DNASE1L3	IL10RB	MEFV	PLEKHM1	SAMD9L	TIM1
ACTB	CD3E	DNASE2	IL12B	MKL1	PMS2	SAMHD1	TIM3
ADA	CD3G	DNMT3B	IL12RB1	MOGS	PNP	SBDS	TINF2
ADAM17	CD40	DOCK2	IL17F	MPO	POLA1	SEC61A1	TIRAP
ADAR	CD40LG	DOCK8	IL17RA	MRE11A	POLE	SEMA3E	TLR3
AHR	CD46	DUOX2	IL17RC	MS4A1	POLE2	SERPING1	TMC6
AICDA	CD55	EFL1	IL18	MSH6	POLR3A	SH2D1A	TMC8
AIRE	CD59	ELANE	IL1RN	MSN	POLR3B	SH3BP2	TMEM173
AK2	CD70	EP300	IL21	MTHFD1	POLR3E	SIRT1	TNFAIP3
AP1S3	CD79A	EPG5	IL21R	MVK	POLR3F	SLAMF8	TNFRSF11A
AP3B1	CD79B	ERBIN	IL2RA	MYD88	POMP	SLC11A1	TNFRSF13B
AP3D1	CD81	ERCC6L2	IL2RG	MYSM1	PRF1	SLC1A5	TNFRSF13C
APOL1	CD8A	EXTL3	IL36RN	NBAS	PRKCD	SLC29A3	TNFRSF1A
ARPC1	CDCA7	FAAP24	IL6ST	NBN	PRKDC	SLC35C1	TNFRSF4
ARPC1B	CEBPE	FADD	IL7R	NCF1	PSEN1	SLC37A4	TNFRSF6
ATM	CECR1	FAS	INO80	NCF2	PSENEN	SLC46A1	TNFSF11
ATP6AP1	CFB	FASLG	IRAK1	NCF4	PSMA3	SMARCAL1	TNFSF12
B2M	CFD	FAT4	IRAK4	NCSTN	PSMB4	SMARCD2	TNFSF6
BACH2	CFH	FCGR1	IRF2BP2	NFAT5	PSMB8	SMPD1	TOP2B
BCL10	CFHR1	FCGR3A	IRF3	NFC2	PSMB9	SMPD2	TPP2
BCL11B	CFHR2	FCN3	IRF7	NFKB1	PSTPIP1	SNX10	TRAC
BIM	CFHR3	FERMT3	IRF8	NFKB2	PTEN	SP110	TRAF3
BLM	CFHR4	FOXN1	ISG15	NFKBIA	PTPRC	SPINK5	TRAF3IP2
BLNK	CFHR5	FOXP3	ITGB2	NFKBIE	RAB27A	SRP54	TREX1
BLOC1S6p	CFI	FPR1	ITCH	NHEJ1	RAC2	STAT1	TRIM28
BTK	CFLAR	G6PC3	ITK	NHP2	RAD52	STAT2	TRNT1
C1QA	CFP	G6PD	JAGN1	NLRC4	RAG1	STAT3	TTC37
C1QB	CFTR	GATA2	JAK1	NLRP1	RAG2	STAT4	TTC7A
C1QC	CIITA	GFI1	JAK3	NLRP12	RANBP2	STAT5B	TYK2
C1R	CLCN7	GINS1	KDM6A	NLRP3	RASGRP1	STIM1	UNC119
C1S	CLEC7A	GZMB	KMT2D	NOD2	RBCK1	STK4	UNC13D
C2	CLPB	HAX1	KRAS	NOP10	RC3H1	STN1	UNC93B1
C3	COLEC11	HELLS	LACC1	NOX1	RECQL4	STX11	UNG
C4A	COPA	HMOX1	LAMTOR2	NPC1	RELA	STXBP2	USB1
C4B	CORO1A	HPS4	LAT	NPC2	RELB	TAOK2	USP18
C5	CR2	HYOU1	LCK	NRAS	RFX5	TAP1	VPS13B
C6	CREBBP	CHD7	LIG1	NSMCE3	RFXANK	TAP2	VPS45
C7	CSF2RA	ICOS	LIG4	NUR77	RFXAP	TAPBP	WAS
C8A	CSF2RB	IFIH1	LIPA	ORAI1	RHBDL4	TAZ	WDR1
C8B	CSF3R	IFNAR2	LPIN2	OSM	RHOH	TBK1	WIPF1
C8G	CTC1	IFNGR1	LRBA	OSTM1	RLTPR	TBX1	WRAP53
C9	CTLA4	IFNGR2	LRRC8A	OTULIN	RMRP	TCF3	XIAP
CARD11	CTPS1	IGHM	LYST	PARN	RNASEH2A	TCIRG1	XRCC4
CARD14	CTSC	IGKC	MAGT1	PCNA	RNASEH2B	TCN2	ZAP70
CARD9	CXCR4	IGKCD	MALT1	PD1	RNASEH2C	TERC	ZBTB24
CARMIL2	CYBA	IGLL1	MAP3K14	PDCD6	RNF168	TERT	
CASP10	CYBB	IKBKB	MASP1	PEPD	RNF31	TFRC	
CASP8	DCLRE1A	IKBKG	MASP2	PGM3	RNU4ATAC	TGFBR1	
CCBE1	DCLRE1B	IKZF1	MBL2	PIEZO1	RORC	TGFBR2	
CD19	DCLRE1C	IKZF3	MCM3	PIK3CD	RPSA	THBD	
CD247	DKC1	IL10	MCM4	PIK3R1	RTEL1	TICAM1	

IBD GENE LIST:

ADA	CYBA	G6PC3	IL10	LIG4	NCF4	SLC37A4	ZAP70
ADAM17	CYBB	GUCY2C	IL10RA	LRBA	PIK3R1	STAT1	
AICDA	DCLRE1C	HPS1	IL10RB	MASP2	PLCG2	STXBP2	
BTK	DKC1	HPS4	IL21	MEFV	RAG2	TTC37	
CD3G	DOCK8	HPS6	IL2RA	MVK	RTEL1	TTC7A	
CD40LG	FERMT1	ICOS	IL2RG	NCF1	SH2D1A	WAS	
COL7A1	FOXP3	IKBKG	ITGB2	NCF2	SKIV2L	XIAP	

IBD2 GENE LIST:

ADA	CTLA4	GUCY2C	IL2RG	NCF2	PIK3R1	RFXAP	TGFBR2
ADAM17	CYBA	HPS1	IL7R	NCF4	PLA2G4A	RTEL1	TNFAIP3
AICDA	CYBB	HPS4	IRF2BP2	NFAT5	PLCG2	SH2D1A	TNFRSF13B
ALPI	DCLRE1C	HPS6	ITGB2	NFKB1	PNP	SKIV2L	TRIM22
ANKZF1	DKC1	ICOS	ITCH	NFKB2	POLA1	SLC37A4	TRNT1
BTK	DOCK2	ICOSLG	JAK3	NFKBIA	PRKDC	SLC9A3	TTC37
CD19	DOCK8	IKBKB	LCK	NLRC4	PTEN	SLCO2A1	TTC7A
CD3D	DUOX2	IKBKG	LIG4	NLRC4	RAC2	STAT1	WAS
CD3E	EPCAM	IL10	LRBA	NOD2	RAG1	STAT3	WIPF1
CD3G	FCN3	IL10RA	MALT1	NOX1	RAG2	STAT5b	XIAP
CD40LG	FERMT1	IL10RB	MASP2	NPC1	RELA	STIM1	ZAP70
CD55	FOXN1	IL21	MEFV	ORAI1	RELB	STXBP2	ZBTB24
CIITA	FOXP3	IL2RA	MVK	PEPD	RFX5	TGFB1	
COL7A1	G6PC3	IL2RB	NCF1	PIK3CD	RFXANK	TGFBR1	

Supplementary file 2: Parameters of identified variants in patients

Patient #	Chromosome	Position	Reference Allele	Sample Allele	Variation Type	Gene Region	Gene Symbol	Protein Variant	Inferred Activity	Genotype	Sample Genotype Quality	Read Depth	Allele Fraction	Translation Impact	Classification	SIFT Function Prediction	PolyPhen-2 Function Prediction	CADD Score
2	5	65309431	T	A	SNV	Exonic	ERBIN	p.D147E	normal	Het	99	26	30.77	missense	Uncertain Significance	Damaging	Probably Damaging	24.400
3	16	3306389	C	T	SNV	Exonic	MEFV	p.V67M	loss	Het	99	39	46.15	missense	Uncertain Significance	Damaging	Probably Damaging	19 920
	1	234743467	T	G	SNV	Exonic; Intronic	IRF2BP2	p.T394P; p.T378P	normal	Het	114	64	48.44	missense	Uncertain Significance	Tolerated	Benign	17 360
	8	48872645	C	T	SNV	Promoter; Exonic	PRKDC	p.R14R	normal	Het	22	10	60	synonymous	Uncertain Significance			16 450
4	8	48801676	A	T	SNV	Exonic	PRKDC	p.A1393A; p.L1393H	loss	Het	65	29	58.62	missense; synonymous	Uncertain Significance			< 10
8	15	77328261	C	G	SNV	Exonic; ncRNA	PSTPIP1	p.Y359*; p.Y349*; p.Y365*; p.Y433*; p.Y368*	normal	Het	99	20	40	stop gain	Uncertain Significance			35 000
10	14	99640991	C	T	SNV	Exonic	BCL11B	p.G727S; p.G728S; p.G534S; p.G657S; p.G656S	normal	Het	99	79	56.96	missense	Uncertain Significance	Tolerated	Benign	23 200
11	1	247587343	G	A	SNV	Exonic	NLRP3	p.V198M; p.V200M	gain	Het	99	62	51.61	missense	Likely Benign	Tolerated	Benign	< 10
13	X	49108141	T	A	SNV	Exonic	FOXP3	p.H327L; p.H342L; p.H377L; p.H400L	loss	Hemi	99	45	100	missense	Uncertain Significance	Damaging	Possibly Damaging	25 900
16	16	50746086	C	T	SNV	Promoter; Exonic	NOD2	p.A755V; p.A728V	normal	Het	99	56	51.79	missense	Benign	Damaging	Probably Damaging	24 200
	15	45389859	G	A	SNV	Exonic	DUOX2	p.R1216W	loss	Het	99	37	45.95	missense	Uncertain Significance	Damaging	Possibly Damaging	29 500
17	15	45391884	C	T	SNV	Exonic	DUOX2	p.A1131T	loss	Het	99	55	40	missense	Likely Pathogenic	Damaging	Probably Damaging	29 100
19	16	81953129	G	A	SNV	Exonic	PLCG2	p.G699S	normal	Het	99	92	50	missense	Uncertain Significance	Tolerated	Probably Damaging	23 900

Patient #	Chromosome	Position	Reference Allele	Sample Allele	Variation Type	Gene Region	Gene Symbol	dbSNP ID	1000 Genomes Frequency	NHLBI ESP Frequency	NHLBI ESP European Frequency	ExAC Frequency	ExAC European Frequency	ExAC Homozygous Count	gnomAD Frequency	gnomAD European Frequency	gnomAD Homozygous Count	HGMD	CLINVAR ID
2	5	65309431	T	A	SNV	Exonic	ERBIN	757550675				0.002	0.003	0	0.003	0.005	0		
3	16	3306389	C	T	SNV	Exonic	MEFV	1422054832							0.002	0.001	0		
	1	234743467	T	G	SNV	Exonic; Intronic	IRF2BP2	138385624		0.015	0.023	0.010	0.016	0	0.006	0.012	0		
	8	48872645	C	T	SNV	Promoter; Exonic	PRKDC	564283387	0.040			0.148	0.107	0	0.046	0.028	0		RCV000433185.1
4	8	48801676	A	T	SNV	Exonic	PRKDC	774069970				0.007	0.011	0	0.006	0.009	0		
8	15	77328261	C	G	SNV	Exonic; ncRNA	PSTPIP1												
10	14	99640991	C	T	SNV	Exonic	BCL11B												
11	1	247587343	G	A	SNV	Exonic	NLRP3	121908147	0.399	0.684	0.953	0.826	1 132	5	0.832	1 156	17	CM013246 (DM?)	RCV000004619.5; RCV000248492.3; RCV000534651.2; RCV000312024.1; RCV000509555.2; RCV000224634.4
13	X	49108141	T	A	SNV	Exonic	FOXP3												
16	16	50746086	C	T	SNV	Promoter; Exonic	NOD2	61747625	0.020	0.323	0.465	0.232	0.354	0	0.252	0.411	1	CM015865 (DM?)	RCV000513725.1; RCV000528773.2; RCV000341511.1; RCV000284123.1
	15	45389859	G	A	SNV	Exonic	DUOX2	144886975		0.008	0.012	0.005	0.004	0	0.005	0.003	0		
17	15	45391884	C	T	SNV	Exonic	DUOX2	147540920		0.031	0.047	0.033	0.052	0	0.023	0.038	0		
19	16	81953129	G	A	SNV	Exonic	PLCG2	753618006				0.001	0.001	0	0.001	0.001	0		

Supplementary file 3: Sequences of customized primers

Patient #	Gene	Forward primer	Reverse primer	Comment
2	ERBIN	AGCCACTCATAACAGAAGCGAC	AAGTCAAAGCCACCAAACAAAGT	reference-specific
		AGCCACTCATAACAGAAGCGAC	AAGTCAAAGCCACCAAAGAAAGT	SNP-specific
3	MEFV	GGTCCCCTTTCCCACAAA	ATGTCTGCCAAGGCATGG	
4	IRF2BP2	CGGACTGTTGCTATTCCTCCT	CTTTGTGTTTCCAGTTGCAAGA	
8	PSTPIP1	CTCAGGATCAAAGACCCCG	ACCTCTTCCTCTGAGTGCCAT	
10	BCL11B	CGTGGAGAAGCGCAGG	CATCAAGGTGGAGAAGGACCT	
11	NLRP3	GGAGCTGCACCTTCCATTTC	TTGTGGATGGGTGGGTTTG	
13	FOXP3	AGACTCAGGTTGTGGCGGA	AGATGGCTCGGGGTAGGTC	
16	NOD2	CAACCTTCAGATCACAGCAGC	CCCACACTTAGCCTTGATGG	
17	DUOX2	GTTTCCCCCAACTAGACCCA	GAGGAAGAAGCCCCCAGATT	exon 25
		CCTTCCTGTCCCATCCTGA	CCATACCCTCAGCAGTCAGG	exon 28
19	PLCG2	ACCCACTTCTCTAAGGCTGGAT	ATGGTGAATACTCAGAGGTTTGCT	