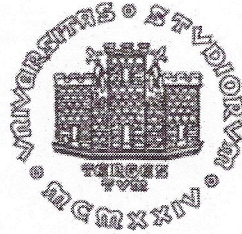


UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXII CICLO DEL DOTTORATO DI RICERCA IN

SCIENZE DELLA RIPRODUZIONE E DELLO SVILUPPO

VEO-IBD AS MODELS FOR PATHOGENIC STUDIES AND DEVELOPMENT OF PRECISION THERAPIES

Settore scientifico-disciplinare: MED/38 PEDIATRIA GENERALE E SPECIALISTICA

DOTTORANDA
ALESSIA PIN

COORDINATORE
PROF. PAOLO GASPARINI

SUPERVISORE DI TESI
DOTT. ALBERTO TOMMASINI

ANNO ACCADEMICO 2018/2019

ABSTRACT

Multiple monogenic disorders present as very early onset inflammatory bowel disease (VEOIBD) or as IBD with severe and atypical features. Establishing a genetic diagnosis may change patients' management and prognosis. In this study, we describe the diagnostic approach to suspected monogenic IBD in a real clinical setting, discussing genetic and phenotypic findings and therapeutic implications of molecular diagnosis. Monogenic VEOIBD diagnostic approach changed over time, especially after the advent of next generation sequencing (NGS) techniques. NGS should be preferred in patients with nonspecific phenotypes. Nevertheless, Sanger sequencing is still effective in patients with suggestive clinical and immunological findings. In a multicentric collaboration with Bambino Gesù Children's hospital, we developed a target gene panel sequencing (TGPS) including the most common monogenic diseases presenting with IBD symptoms, as first line of genetic approach for patient with non-specific phenotypes and negativity to this panel, we performed WES with an *in silico* analysis of 400 genes responsible for primary immunodeficiencies. 94 patients were included, and 13 (14%) reached a genetic diagnosis. Candidate sequencing was performed in 47 patients (50%), and NGS was performed in 85 patients (90%). Candidate sequencing had a good diagnostic performance only when guided by clinical features specific for known monogenic diseases, whereas NGS helped finding new causative genetic variants and would have anticipated one monogenic diagnosis (*XIAP*) and consequent bone marrow transplant (BMT). Genetic diagnosis impacted patient management in 11 patients (92%), 7 of whom underwent BMT. Although we identified 14% of monogenic disease in our cohort, the majority of cases remains without a genetic diagnosis. We hypothesized that transcriptome analysis by RNA sequencing (RNAseq) could help grouping multifactorial cases and correlating profiles with those found in distinct monogenic forms. We proposed a disease-similarity method for patients' stratification and the detection of possible biomarkers. 13 out of 94 patients, depending on RNA availability (4 monogenic and 9 nonmonogenic) described in the previous genetic workup, performed gene expression analysis by RNAseq of peripheral blood cells. We compared gene expression profile of the 4 monogenic IBD (XIAP, TTC37, DKC1, and LRBA) with nonmonogenic IBD and performed cluster analysis. The most evident impact on peripheral blood cells came

from XIAP and DKC1. TTC37 and LRBA did not show enriched pathway probably due to wrong sampling. Few nonmonogenic patients that presented extraintestinal manifestations (feature suggestive of monogenic defect) had a hybrid expression profile between monogenic and nonmonogenic IBD. Cluster and machine learning analyses might be applied to group patients by gene expression patterns in an unbiased manner. We performed an unsupervised analysis including our monogenic IBD, the nonmonogenic IBD from the cohort of Trieste and the first 13 genetically undefined VEOIBD and EOIBD enrolled within the collaborative project with University of Brescia, whose clinical collection data and genetic investigations are in progress. However, this data should be complemented by clinical reports and therapeutic management at the time of sampling to get more precise results and evaluate the obtained functional subgroups. Nevertheless, the characterization of more monogenic forms is a crucial point to expand this analysis and obtain more reliable results. The implementation of this knowledge may allow the use of monogenic disorders as prototypical diseases for the stratification and the therapeutic management of likely multifactorial cases towards a tailored therapy.

SUMMARY

1 INTRODUCTION......3

1.1 INFLAMMATORY BOWEL DISEASE: HISTORY, CLASSIFICATION, AND ETIOPATHOGENESIS......3

1.2 VERY EARLY ONSET INFLAMMATORY BOWEL DISEASES (VEOIBD) and MONOGENIC VEOIBD: WHEN SUSPECTING A MONOGENIC DEFECTS?......6

1.3 THE DIAGNOSTIC APPROACH OF VEOIBD IN THE ERA OF NEXT GENERATION SEQUENCING......13

1.3.1 GENOMICS IN VEOIBD......13

1.3.2 FUTURE PERSPECTIVE OF TRANSCRIPTOMICS and MULTI-OMICS APPROACH IN VEOIBD......14

1.4 TREATMENTS and PRECISION THERAPIES IN VEOIBD......16

2 OBJECTIVES......19

2.1 IDENTIFICATION OF MONOGENIC DISEASE WITHIN OUR EOIBD/VEOIBD GROUP and SET UP OF A GENETIC DIAGNOSTIC WORKFLOW......19

2.2 TRANSCRIPTOMICS AS GENETIC WORKUP COMPLEMENTATION AND AS PATIENTS' STRATIFICATION TOOL......19

3 MATERIALS AND METHODS......20

3.1 PATIENT POPULATION and STUDY......20

3.2 DIAGNOSTIC WORKFLOW......21

3.2.1 TARGETED GENE PANEL DESIGN......21

3.2.2 DNA LIBRARY PREPARATION AND RAW DATA ANALYSIS......24

3.2.3 WES AND DATA ANALYSIS......24

3.2.4 VARIANT SELECTION AND VALIDATION......24

3.2.5 STATISTICAL ANALYSES......25

3.3 RNAseq WORKFLOW......26

3.3.1 RNA EXTRACTION AND RAW DATA PROCESSING......26

3.3.2 DIFFERENTIAL GENE EXPRESSION ANALYSIS......26

3.3.3 CLUSTER ANALYSIS TO CLASSIFY PATIENTS INTO DIFFERENT SUBSETS......27

4 RESULTS and DISCUSSION......29

4.1 PATIENTS......29

4.2 GENETIC WORKUP AND DIAGNOSIS......29

4.3 RNAseq WORKUP......37

4.3.1 PATIENT DESCRIPTION......37

4.3.2 VEOIBD SUBGROUPS: THE WEIGHT OF THERAPY......39

4.3.3 VEOIBD SUBGROUPS: HEALTHY SUBJECTS COMPARISON (VEOIBD vs healthy subjects)......41

4.3.4 DIFFERENTIAL GENE EXPRESSION (monogenic vs nonmonogenic IBD)......44

4.3.5 PATHWAY ENRICHMENT ANALYSES of DEGs in MONOGENIC VEOIBD......45

4.3.6 SUBGROUPS IDENTIFICATION WITHIN VEOIBD......48

4.3.7 MACHINE LEARNING APPROACH TO IDENTIFY POSSIBLE CLUSTER-SPECIFIC GENES 50

4.3.8 UNSUPERVISED CLUSTERING ANALYSIS TO FIND SIMILAR DISEASES GROUP OF CLINICALLY AND GENETICALLY UNKNOWN VEOIBD PATIENTS......52

5 CONCLUSIONS......54

6 SUPPLEMENTARY MATERIAL..57
7 BIBLIOGRAPHY..60

1 INTRODUCTION

1.1 INFLAMMATORY BOWEL DISEASE: HISTORY, CLASSIFICATION, AND ETIOPATHOGENESIS

Inflammatory bowel disease (IBD) defines a group of complex chronic intestinal diseases, which includes ulcerative colitis (UC), Crohn's disease (CD), and IBD unclassified (IBD-U).

UC is characterized by intestinal mucosal inflammation, mainly limited to the colon, bloody diarrhea, abdominal pain and tenesmus [1].

Sir Samuel Wilks was the first physician who used the term "ulcerative colitis", referring to a condition similar to the one intended today [2].

The recognition of CD, as a different condition from UC, came in 1932 as the result of a publication by Crohn et al., after which the condition was named with the eponymous name of Crohn's disease [3, 4].

CD may involve any part of the gastrointestinal tract; the most common locations are the terminal ileum or the perianal region. The inflammatory lesions are transmural, involving all the layers of the intestine with a segmental and noncontinuous distribution characterized by the alternation of inflamed and non-inflamed regions [5]. The discontinuous inflammation is probably a consequence of a vicious circle arising from mucosal barrier defect, translocation of bacteria and inflammation. Moreover, a distinctive feature of CD is the presence of a chronic granulomatous inflammatory reaction.

Similarly to UC, the main symptoms are diarrhea, rectal bleeding, fatigue, and weight loss.

Even though CD clinical features are mostly limited to the gut, also extraintestinal manifestations are characteristic of this chronic inflammatory disorder, as uveitis, erythema nodosum, and arthritis.

The term indeterminate colitis, progressively replaced by IBD-U, was originally used referring to a pathological diagnosis related to colectomy biopsies, in which a specific diagnosis of UC or CD was difficult to reach [6]. More recently, IBD-U classification has been used when clinical, endoscopic and biopsy testing do not meet the diagnostic criteria of either UC or CD [7-9].

IBD can occur at any age with a higher prevalence in young adults, particularly between 30-40 years of age in the case of UC and between 20-30 years of age in the case of CD and can affect about one in 200 people in the developed countries [10, 11].
Although the pathogenesis of IBD is complex and not yet well established, recent studies indicated an involvement of various factors such as genetic susceptibility, external environment, and intestinal dysbiosis, contributing together to the development of the disease [12].
Epidemiological and molecular genetics studies of monozygotic twins highlighted the importance of genetics in the pathogenesis of IBD [13]. The models of inheritance are in most cases different from mendelian monogenic diseases but comparable to genetically complex disorders.
Genome-wide association studies (GWAS), based on linkage analyses, have identified more than 163 risk loci for IBD, encompassing about 300 potential candidate genes, such as locus IBD1 on chromosome 16, in which is located NOD2, the first gene discovered to correlate with IBD but especially with CD [14]. More recently, immunochip genotype data showed the implication of additional 38 loci in IBD risk, shared among different ethnic group [15].
Various genetic variants in NOD2 increase the probability to develop CD: from two to four times in people carriers for only one “risk allele”, heterozygous for a defined variant, and until forty times in individuals homozygous for a specific mutation.
Some loci contain risk variants associated both with both in UC and CD, as in the case of interleukin 23 receptor (IL23R)(Figure 1) [16].

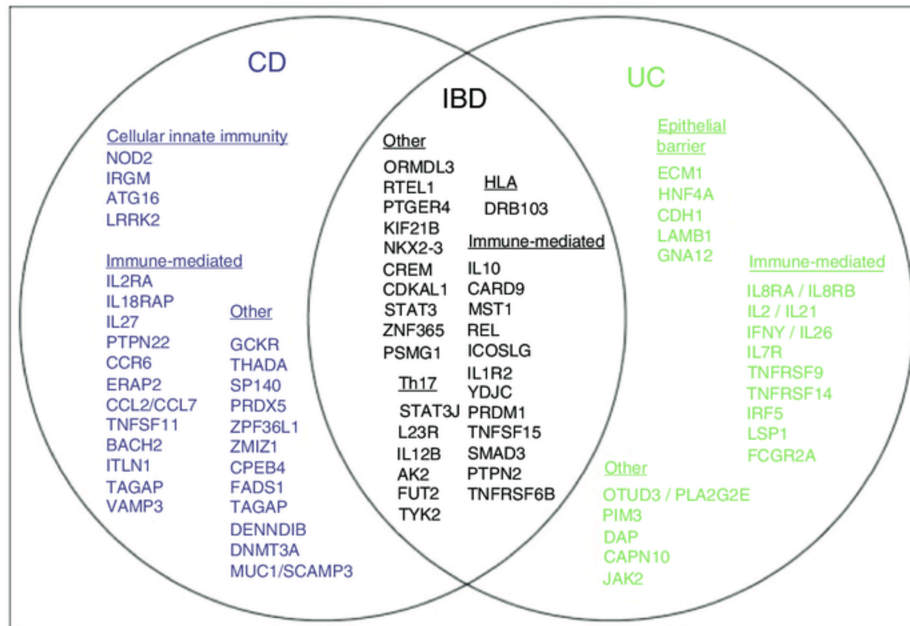


Figure 1.
Loci associated with IBD. Overlapping genes between CD and UC reported in black, gene associated to UC in green and genes associated with CD in blue [13].

Pathway enrichment analyses of candidate genes, within susceptibility *loci*, pointed out the involvement of three major biological processes, implicated in the activation of Th17 cells, in modulation of autophagy, and in immune-mediated responses.
About 70% of IBD susceptibility loci are shared with other complex autoimmune and autoinflammatory diseases, like psoriasis and ankylosing spondylitis, that often occur as extraintestinal manifestations of UC and CD [17, 18].
The identification of genes and risk *loci* associated with the development of IBD, is crucial to better understand the molecular pathways that underpin the disease and to help the therapeutic choice. However, only about 25% of IBD could be explained by genetic and molecular studies [19]. Other factors, like environment conditioning, have to be considered in the etiopathogenesis of the disease. Indeed, new epidemiological research, conducted in industrialized countries, suggested a key role of environmental factors (diet, nonsteroidal anti-inflammatory drug) in IBD development in genetically susceptible subjects [19, 20]. These aspects, synergistically with predisposing mutations, could lead to an alteration in the homeostasis between commensal bacteria and immune system, due to an imbalance of antimicrobial peptides (e.g. alpha-defensins by Paneth cells) [21], and dysregulated cytokine productions. In this context and thus in IBD pathogenesis, the transcription factor NF-κB seems to play a crucial role since it is involved in pro-inflammatory signaling.

Genetics might have a pivotal role in patients presenting an early onset IBD. The slogan goes: the earlier the onset, the higher is the probability to have a Mendelian disease (Figure 2).

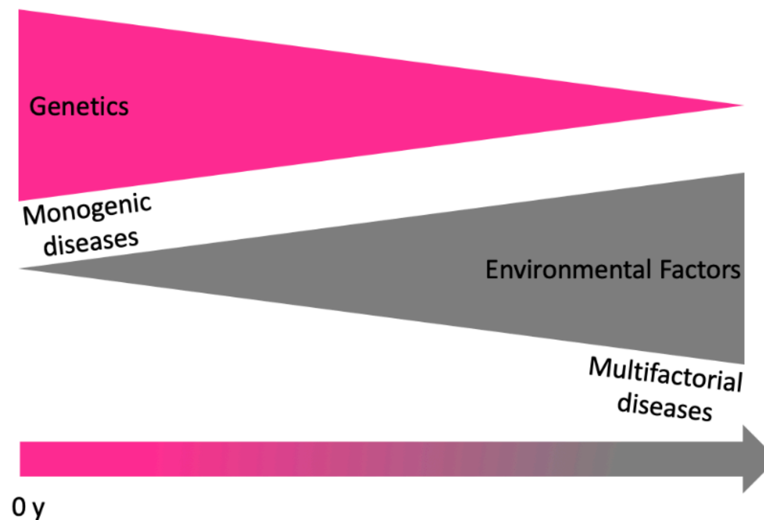


Figure 2.
Genetics and environmental weight from birth to adulthood.
Figure adapted from [22].

1.2 VERY EARLY ONSET INFLAMMATORY BOWEL DISEASES (VEOIBD) and MONOGENIC VEOIBD: WHEN SUSPECTING A MONOGENIC DEFECTS?

Even though most cases of IBD present as multifactorial disease, a broad spectrum of rare Mendelian disorders may occur with an IBD-like phenotype. The probability to develop a monogenic disease is higher when the disease manifests during the first years of life. A prompt diagnosis of such diseases is essential to determine the correct prognosis and the treatment strategy, which in some cases may include bone marrow transplantation (BMT) [23]. The age of onset is one of the first “red flag” to suspect a monogenic condition, alongside IBD location, progression, familial component, and response to therapies [24-31]. Monogenic defects are likely to be studied as prototypical diseases for a better understanding of IBD as they allow to analyze the effects of the dysregulation of single mechanisms [18, 32].

Pediatric IBD can be summarized mainly in five subgroups according to the age of onset (Table 1); pediatric IBD, early-onset IBD (EOIBD), very early onset IBD (VEOIBD), infantile IBD (IOIBD) and neonatal IBD.

The Montreal and the Pediatric Paris classifications defined, at first, two distinct subgroups, pediatric-onset IBD (children between 10 and 17 years of age, Paris A1b) and EOIBD (children younger than 10 years of age, Paris A1a) presenting a more severe disease compared to the one that occurs in adulthood [33].

Several literature records reported a relative enrichment of patients with a monogenic IBD in children with disease onset before 6 years of age, leading to propose VEOIBD as a new age group. However, more rarely monogenic defects, such as XIAP deficiency or other neutrophil defects, have been observed, also, in later onset IBD [34, 35] (Figure 3).

EOIBD and VEOIBD will be the focus of this thesis from here on out.

Group	Montreal/Paris classification	Age range of onset (y)	Phenotype notes
Pediatric onset IBD	Montreal A1	<17	-
EOIBD	Paris A1a	<10	More severe phenotypes than adolescence and adults
VEOIBD		<6	Enrichment of monogenic IBD, some case if IBD-U, resistance to conventional therapies
Infantile onset IBD		<2	Higher familial component rate with increased lethality, resistance to conventional therapies, risk to an underlined primary immunodeficiency.
Neonatal IBD		First 28 days of age	Extreme phenotype

Table 1. *Pediatric IBD subgroups. Table adapted from [31].*

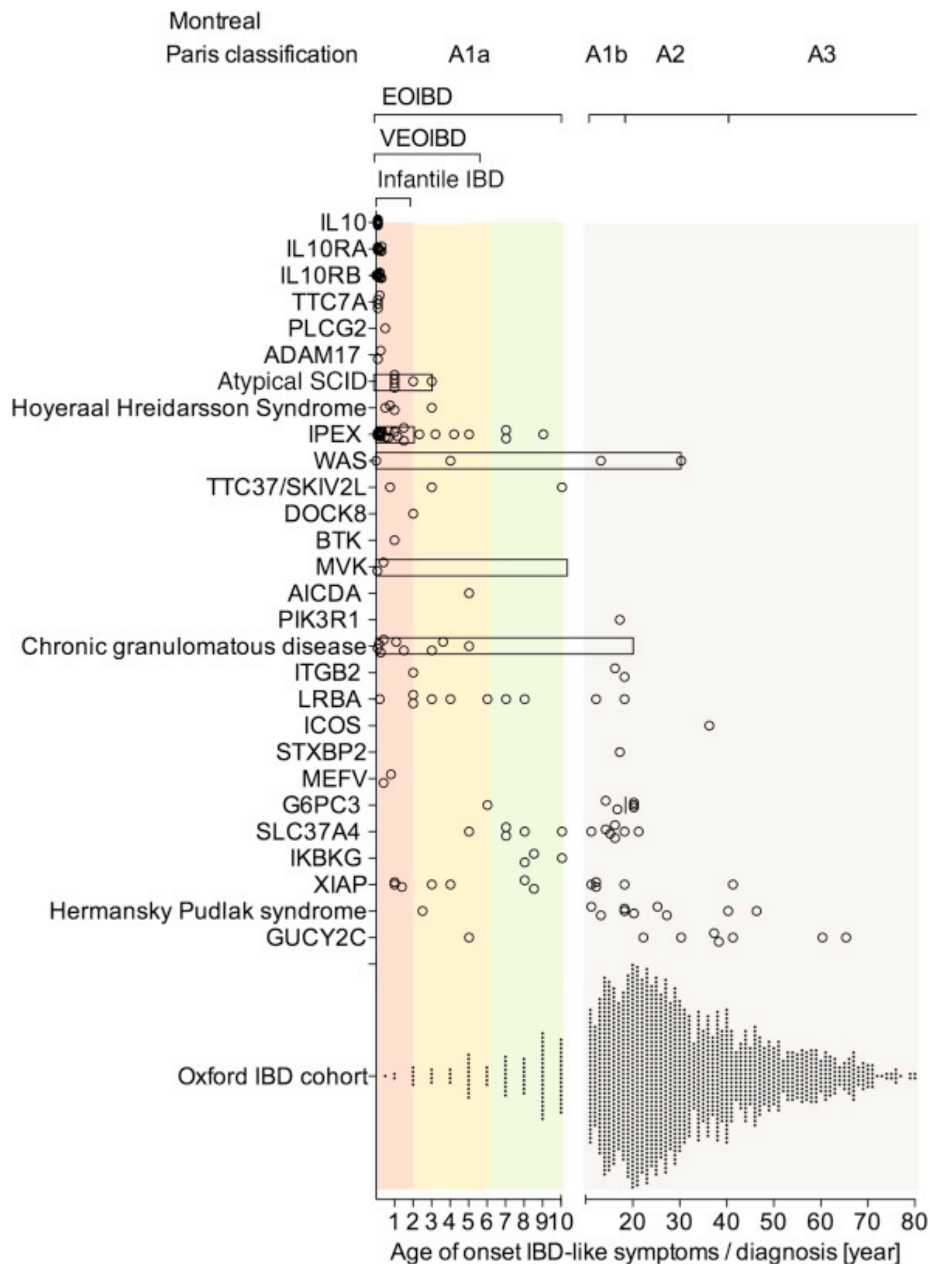


Figure 3.

Age of onset of IBD-like diseases in patients with a monogenic disease.
Genetic defects associated to IBD-like symptoms are summarized in this figure in comparison with an unselected IBD population (Oxford IBD cohort). Each dot represents one patient and bar represents the age of range of case series when individual data are not available. At the top of the figure, age range and Montreal/Paris classification are shown as reference [31].

These patients present a low response rate to conventional anti-inflammatory and immunomodulatory treatments. Some of them are labelled as IBD-U [36], supporting the difficulty to endoscopically categorize these “atypical” cases of IBD that could underline a more severe defect.

The IOIBD (younger than 2 years of age) and neonatal IBD (first month of age) groups, are likely characterized by higher presence of affected first-degree relatives with increased lethality [24-27, 31], as an indication of a stronger genetic component, even higher than VEOIBD, severe disease progression and resistance to immunosuppressive drugs with an increased risk to underneath a primary immunodeficiency.

Even though the spectrum of genotypes manifesting with an intestinal inflammation is increasing, only a small fraction of patients with VEOIBD (around 15%) may have a rare monogenic disorder.

The identification of genetic defects in IL10/IL10 receptor signaling as the cause of severe VEOIBD by family association studies and candidate sequencing, was a starting point for the identification of multiple monogenic disorders [37-39].

To date, disease causative variants have been identified in more than 50 genes (Table 2), that can be divided within different functional groups [23, 31]:

- Epithelial barrier dysregulation and epithelial response defects, causing Kindler syndrome, familial diarrhea;
- Neutropenia and defects in phagocyte bacterial killing, including chronic granulomatous disease (CGD);
- Hyperinflammatory and autoinflammatory disorders, such as mevalonate kinase deficiency;
- Defects in T and B lymphocytes development of regulation, as the cases of agammaglobulinemia and immune dysregulation, polyendocrinopathy, enteropathy, X linked (IPEX) syndrome and Wiskott Aldrich syndrome (WAS);
- Disorders affecting the immunoregulation, as IL10 signaling defects resulting in intestinal inflammation due to an impairment of the proinflammatory response;
- A group of disorders without a well-defined plausible functional mechanism, such as the trichohepatoenteric syndrome, which may involve a defect in epithelial cells causing intractable diarrhea.

Group	Disease	Gene	Inheritance	Reference
Epithelial barrier	Dystrophic bullosa	*COL7A1*	AR	[40]
	Kindler syndrome	*FERMT1*	AR	[41]
	X-linked ectodermal immunodeficiency	*IKBKG*	XL	[42, 43]
	TTC7A deficiency	*TTC7A*	AR	[44]
	ADAM17 deficiency	*ADAM17*	AR	[45]
	Familial diarrhea	*GUCY2C*	AD	[46]
Phagocyte defects	CGD	*CYBB*	XL	[47]
	CGD	*CYBA*	AR	[48]
	CGD	*NCF1*	AR	[47]
	CGD	*NCF2*	AR	[47]
	CGD	*NCF4*	AR	[49]
	Glycogen storage disease type 1b	*SLC37A4*	AR	[50]
	Congenital neutropenia	*G6PC3*	AR	[51]
	Leukocyte adhesion deficiency 1	*ITGB2*	AR	[52]
Hyperinflammatory disorders	Mevalonate kinase deficiency	*MVK*	AR	[53]
Autoinflammatory disorders	Phospholipase C-γ2 defects	*PLCG2*	AD	[54]
	Familial Mediterranean fever	*MEFV*	AR	[55, 56]
	Familial hemophagocytic lympgohistiocytosis type 5	*STXBP2*	AR	[57]
	X-linked lymphoproliferative syndrome 2 (XLP2)	*XIAP*	XL	[58]
	X-linked lymphoproliferative syndrome 1 (XLP1)	*SH2D1A*	XL	[59]
	Hermansky-Pudlak 1	*HPS1*	AR	[60]
	Hermansky-Pudlak 4	*HPS4*	AR	[61]
	Hermansky-Pudlak 6	*HPS6*	AR	[62]
T and B cells defects	CVID 1	*ICOS*	AR	[63]
	CVID 8	*LRBA*	AR	[64]
	IL-21 deficiency (CVID-like)	*IL21*	AR	[65]
	Agammaglobulinemia	*BTK*	XL	[66]
	Agammaglobulinemia	*PIK3R1*	AR	[67]
	Hyper IgM syndrome	*CD40LG*	XL	[68]
	Hyper IgM syndrome	*AICDA*	AR	[69]
	Wiskott Aldrich syndrome	*WAS*	XL	[70]
	Omenn syndrome	*DCLRE1C*	AR	[71]
	SCID	*ZAP70*	AR	[72]
	SCID/hyper IgM syndrome	*RAG2*	AR	[73]
	SCID	*IL2RG*	XL	[74]
	SCID	*LIG4*	AR	[75]
	SCID	*ADA*	AR	[75]

Group	Disease	Gene	Inheritance	Reference
T and B cells defects	SCID	*CD3γ*	AR	[76]
	Diskeratosis congenita	*DKC1*	XL	[77]
	Hoyeraal–Hreidarsson syndrome	*RTEL1*	AR	[78]
	Hyper IgE syndrome	*DOCK8*	AR	[79]
	IPEX	*FOXP3*	XL	[80]
Immunoregulation	IPEX-like	*IL2RA*	AR	[81]
	IPEX-like	*STAT1*	AD	[82]
	IL-10 signaling defects	*IL10RA*	AR	[83]
	IL-10 signaling defects	*IL10RB*	AR	[37]
	IL-10 signaling defects	*IL10*	AR	[84]
Others	MASP deficiency	*MASP2*	AR	[85]
	Trichohepatoenteric syndrome	*SKIV2L*	AR	[86]
	Trichohepatoenteric syndrome	*TTC37*	AR	[87]

Table 2. *Monogenic defects divided into functional subgroups.*
AR, autosomal recessive; XL, X-linked; AD, autosomal dominant.

Uhlig et al. [31] summarized in the powerful phenotypic aide-memoire "**YOUNG AGE MATTERS MOST**" the key points suggestive of a monogenic disease:

- **YOUNG AGE** onset;
- **M**ultiple family members and consanguinity;
- **A**utoimmunity;
- **T**hriving failure;
- **T**reatment with conventional medication fails;
- **E**ndocrine concerns;
- **R**ecurrent infections or unexplained fever;
- **S**evere perianal disease;
- **M**acrophage activation syndrome and hemophagocytic lymphohistiocytosis;
- **O**bstruction and atresia of intestine;
- **S**kin lesions and dental and hair abnormalities;
- **T**umors.

1.3 THE DIAGNOSTIC APPROACH OF VEOIBD IN THE ERA OF NEXT GENERATION SEQUENCING

1.3.1 GENOMICS IN VEOIBD

In these complex cases, interdisciplinary and multidisciplinary support from a team with pediatric gastroenterologists, immunologists, geneticists, bioinformaticians, and biotechnologists is crucial.

Classical approach to VEOIBD diagnosis is based on phenotypic and histological observations, excluding infections (e.g. CMV), followed by restricted specific-disease functional screening and genetic confirmation by candidate sequencing (Figure 4).

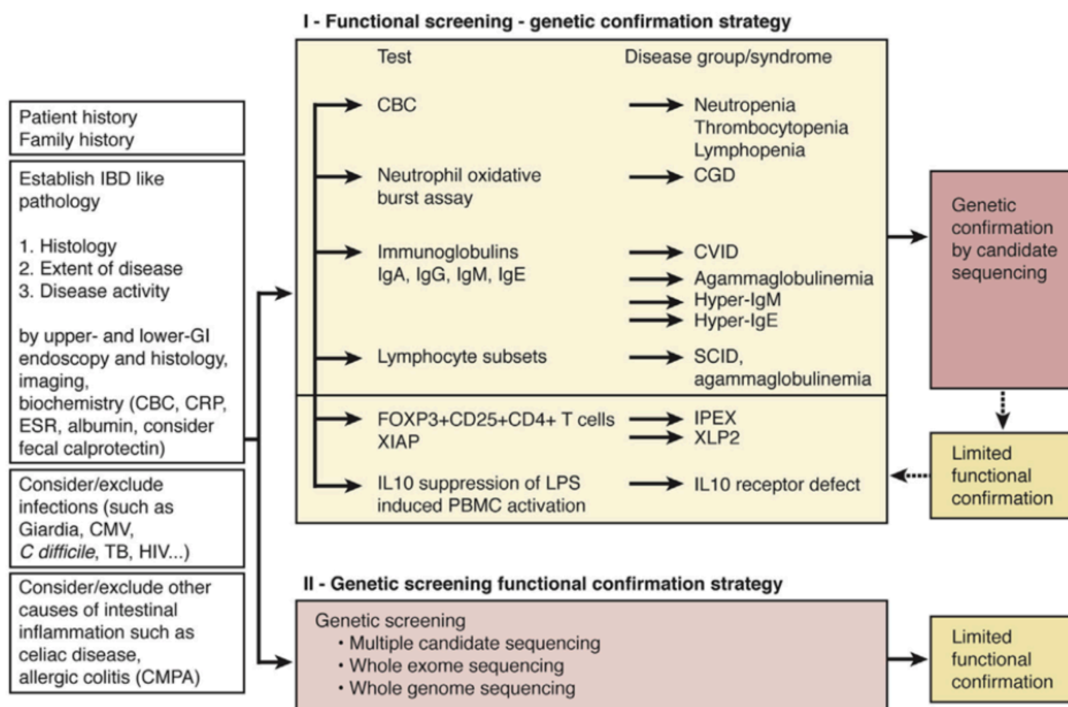


Figure 4.
VEOIBD classical and new diagnostic workflow [31].
CBC complete blood count, CRP C-reactive protein, ESR erythrocyte sedimentation rate, CMPA cow milk protein allergy.

In the last decade, with the advent of Next Generation Sequencing (NGS) technique, the diagnostic workflow of monogenic IBD changed the first line of diagnostics for VEOIBD: from deep phenotyping followed by candidate genes Sanger sequencing, towards multiple genes parallel sequencing using targeted gene panel sequencing (TGPS), and/or Whole Exome Sequencing (WES) (Figure 4).

In patients with nonspecific clinical phenotype, NGS has the advantage to allow analyzing simultaneously several genes, with shorter time to response and less expenses compared with sequential sequencing of single candidate genes [88].
A big edge of WES is to potentially highlight known and new causative variants both with a supervised approach, by the analysis of a selected gene set guided by a clinical suspicion, and with an unsupervised approach in the most atypical cases, sometimes leading to the detection of new candidate genes. To date Sanger sequencing is still required to confirm NGS output variants. However, when new variants of uncertain significance are found, further functional assays are essential to assess the role of newly detected damaging mutations, since computational prediction tools can be misleading and get false negative or false positive results.
The genomics of VEOIBD has been implemented also by Whole Genome Sequencing (WGS), useful in particular for some relevant IBD genes, especially for *IKBKG* and *NCF1* that are difficult to capture by exome sequencing techniques since result underrepresented [39, 88]. WGS offers high coverage sequencing across the genome and a much more uniform distribution of sequencing-quality parameters than WES, including promoters and enhancer binding regions, and allows the analysis of copy-number variants and the detection of inversions [89-92].
The perspective of genomics analyses implementation in VEOIBD diagnostic workup is the prompt detection of monogenic diseases, which in some patients can prevent unnecessary surgeries, severe infection, or tumors aiming to target therapy [39].

1.3.2 FUTURE PERSPECTIVE OF TRANSCRIPTOMICS and MULTI-OMICS APPROACH IN VEOIBD

The assessment of variant impacts that could affect RNA splicing, or gene expression levels, might be performed by RNA sequencing (RNAseq), as complementation of genomic technologies. A multi-omics approach could help the diagnosis of previously unsolved cases and especially RNAseq will improve the diagnostic yield by the detection of both the coding and non-coding variants that implicated splicing defects [39, 93].
Nevertheless, it is possible to use the measure of gene expression to stratify subjects on different clinical states, for instance, by the identification of prognostic biomarkers or by detection of cell-type-specific gene-signature, to clinically separate patients according to different disease courses and guide specific treatments [94].

However, the identification of biomarkers related to IBD is challenging given the high heterogeneity in disease progression and the variability in therapy response [95]. Indeed, biological specimens at the disease onset, when no therapy has been started, are available only for a minority of patients.
The majority of gene expression studies have been carried out on intestinal biopsies, which is the gold standard for assessing disease activity and severity. Intestinal biopsy sampling presents some drawbacks in terms of costs, discomfort for patients, procedure complications and in patients with an underneath immunodeficiency it may not be the most disease-representative tissue [96].
Serological and fecal biomarkers are widely used in clinical practice but present a low diagnostic accuracy to discriminate among various disease causes [96-98].
Microarray studies led to the development of a non-invasive test for IBD diagnosis, based on the expression profiles of peripheral leukocytes [99, 100]. Alsobrook et al. reported a six gene peripheral blood signature (*BLCAP, UBE2G1, GPX1, RAP1A, CALM3*, and *NONO*) able to distinguish patients with IBD from healthy controls with accuracy, sensitivity, and specificity rates of 84%, 89%, and 75%, respectively [100, 101]. In clinical practice there is no need for such a diagnostic tool, as common laboratory analysis, together with clinical and image data can easily guide the diagnostic process and since endoscopic and histologic examination is still required to study the response to treatments. Thus, the utility of expression profiling of blood sample is conflicting. J. Ostrowski et al., indeed, re-defined the utility of whole blood gene expression analysis in IBD diagnostic workup in pediatric and adult patients with IBD [100]. They identified 15 potential biomarkers (*S100A12, OPLAH, ATP9A, ANOS1, FCGR1A, ITGB4, UTS2R, MMP9, COX6B2, ANXA3, CACNA1E, GALNT14, IL18R1, KLRF1,* and *PFKFB3*) by RNAseq and subsequently validated by qPCR analyses in newly recruited IBD patients and controls. These findings confirmed a diagnostic potential of this whole blood-based test for pediatric patients with active IBD, but not for pediatric patients with inactive IBD, or adults with active or inactive disease [100].
To date, no whole blood signature has been demonstrated useful to sub-group subjects within the varied spectrum of VEOIBD, while it could be desirable to find gene expression profiles able to predict the response to distinct treatments.
The need to approach these highly heterogeneous diseases led, indeed, to the increasing generation and availability of digital data. The use of “big data” (intended

as large-volume and high-diversity biological, clinical, and lifestyle information) might be helpful in the approach to highly heterogeneous diseases, to define distinct subgroups of patients for the purpose of therapeutic stratification. Another method to analyze complex data and to stratify patients is a “machine learning” approach. Machine learning is a branch of statistics, well employed in finding specific patterns, making predictions and classifications or inferring new knowledge [102]. This method includes the unsupervised and supervised machine learning algorithms that aim to classify samples without or with *a priori* knowledge of their division into a specific category. Specifically, supervised algorithms are suited to solve classification problems where a known group is used as a training set to classify subsequent samples of unknown class.

The integration of unsupervised and supervised algorithms has been already used successfully in medicine and biology for example in cancer subtype identification and the discovery of novel drugs [103-107].

This approach has been applied to discriminate, for example, pediatric CD and UC patients based on endoscopic and histological disease location [107] but not in the diagnosis of VEOIBD.

This thesis project aims to stratify patients with VEOIBD, by RNAseq and machine learning approaches. In particular we will, if it is possible, separate cases with a higher genetic involvement from subjects with a multifactorial disease by exploiting gene expression profiles in peripheral blood from patients.

1.4 TREATMENTS and PRECISION THERAPIES IN VEOIBD

The most dramatic application of clinical genetics with the subsequent identification of pathogenic genetic variants is observed in IBD patients with prevalent immune defects that can be treated almost exclusively by BMT, as in the cases of *IL10RA* and *IL10RB* mutations [39]. Conversely, patients with epithelial defects (e.g *TTC7A* mutations) should not be treated with BMT since they might present a poor outcome [108].

Despite an increased detection of disease-causing gene variants associated with IBD-like symptoms, thanks to the advent of NGS technology (e.g. TGPS, WES, WGS), many subjects with a strong suspicion of a monogenic disorder, remained without a genetic diagnosis. In some cases, negative results may reflect low sensitivity of assays because of low sequencing coverage, epigenetic changes, and small noncoding

molecules impacting gene expression. In other cases, the disease can be a polygenic one presenting with severe symptoms that could resemble a mendelian disorder profile.

In these cases, a tailored therapy might be more difficult to develop and instead, symptomatic treatments are usually employed.

Historically, several small-molecule drugs (SMDs), including corticosteroids, immunomodulators (such as azathioprine, 6-mercaptopurine and methotrexate) and aminosalicylates have been used in the medical management of IBD [109, 110]. IBD treatment has been revolutionized by the introduction of biologic anti-tumour necrosis factor-α (TNF-α) drugs, like infliximab and adalimumab, within the first years of this century [110]. The link between TNF and IBD has been shown by several publications that reported patients with IBD presenting increased level of TNF in serum, stool or mucosal biopsies [111-113]. The remission rate by using anti-TNF-α drugs is about 60% [114], reaching mucosal healing of inflammation and restoration of intestinal epithelial integrity.

Due to its involvement in pro-inflammatory signaling, NF-κB became another attractive therapeutic target in IBD management. Corticosteroids, sulfasalazine, methotrexate and anti-TNF-α antibodies indirectly inhibit NF-κB-mediated inflammation [115-117]. Selective NF-κB inhibition can be obtained by drugs able to reduce the expression of this factor (e.g. specific antisense oligonucleotides) or modulator of NF-κB signaling by ubiquitylation and degradation of its regulator proteins.

Despite the potentiality of blocking NF-κB activities in IBD intervention, it is important to remember, also, its involvement in normal physiological cell functions. For this reason, to avoid severe side effects, it is crucial to attempt targeting specific cell types, as immune cells within the inflamed intestinal mucosa, or to block specific NF-κB subunits to minimize systemic toxic effects [117].

Another drug used for its immunomodulatory property in IBD, is thalidomide. Originally, thalidomide was used as an antiemetic agent in pregnancy but was withdrawn from the market due to its teratogenic effect. This drug has been re-introduced to treat various autoinflammatory diseases and it is shown to be effective and safe in treating refractory paediatric IBD. One of its mechanisms of action can exploit the inhibition of NF-κB activation. The molecular mechanism of thalidomide remained unclear for a long time. However, more recent research showed

that cereblon (CRBN) was found to be a direct target of thalidomide and to function as a substrate receptor of E3 ubiquitin ligase, creating, indeed, the complex thalidomide-CRBN-E3 ubiquitin ligase. These findings opened the doors to target new potential substrates leading to their ubiquitination to obtain the desirable anti-inflammatory effect [118].

This thesis project aims to describe a workflow to diagnose monogenic IBD in our cohort of patients and how establishing a genetic diagnosis may impact therapeutic management. Furthermore, we explored the possibility that monogenic diseases might be used as prototypical disease to understand similar deregulated mechanism in multifactorial forms.

2 OBJECTIVES

2.1 IDENTIFICATION OF MONOGENIC DISEASE WITHIN OUR EOIBD/VEOIBD GROUP and SET UP OF A GENETIC DIAGNOSTIC WORKFLOW

The early detection of monogenic diseases, manifested with IBD symptoms, is crucial to guide a precision therapy.

In this study we aim:

- To describe the diagnostic workup to suspected monogenic IBD in a real clinical multicentric setting during a 10-year period, focusing on the advantages and disadvantages of different diagnostic strategies;
- To provide the rate of monogenic diagnoses compared to previous literature studies conducted in other VEOIBD cohorts.

2.2 TRANSCRIPTOMICS AS GENETIC WORKUP COMPLEMENTATION AND AS PATIENTS' STRATIFICATION TOOL

Literature studies, about genetic VEOIBD analysis, reported that only 15/20% of patients with VEOIBD has a monogenic disease. The majority of cases remains without a genetic diagnosis, and thus likely multifactorial.

We hypothesized that transcriptome analysis by RNAseq:

- could help identifying VEOIBD subgroups guided by the comparison of gene expression profiles of prototypical monogenic diseases;
- could help distinguish nonmonogenic patients from distinct monogenic form;
- could be useful in the identification of functional profiles and enriched pathway characteristic of specific subgroups.

3 MATERIALS AND METHODS

3.1 PATIENT POPULATION and STUDY

1st analysis: Genomic study in VEOIBD

This was a multicenter observational cohort study. Patients diagnosed with VEOIBD and patients with early onset IBD with severe/atypical phenotypes (EOIBD s/a) managed at 2 main pediatric gastroenterology centers, Institute for Maternal and Child Health IRCCS Burlo Garofolo of Trieste and the Bambino Gesù Children's Hospital of Rome, in the last 10 years (2008 to 2017) and patients referred for a genetic workup from 9 external gastroenterology facilities were included.

The definition of severe/atypical phenotype was applied when at least one of the following clinical findings were present: severe perianal disease, recurrent/atypical infections, skin/annexes abnormalities, abnormal immune status, associated multiple/severe autoimmunity, history of macrophage activation syndrome or hemophagocytic lymphohistiocytosis, intestinal atresia, or early development of tumors. Demographic data and information on gastrointestinal disease, extraintestinal manifestations, and treatments were retrieved from medical records. In the first part of the study, information of interest was retrospectively collected from medical records and included in a dedicated database. Starting from 2015, newly diagnosed patients with VEOIBD and EOIBD s/a and patients without a previous definite genetic diagnosis were prospectively recruited for genetic workup. The work was conducted in accordance with the revised Declaration of Helsinki and was approved by Institute for Maternal and Child Health IRCCS Burlo Garofolo of Trieste and the Bambino Gesù Children's Hospital of Rome Ethics committee. Written informed parental consent was obtained for genetic analysis [119].

2nd analysis: Transcriptomic study in VEOIBD

Patients with VEOIBD and EOIBD s/a, either with a genetic diagnosis or prospectively recruited for transcriptomics analysis, in the last period (early 2019), coordinated by 2 main centers Institute for Institute for Maternal and Child Health IRCCS Burlo Garofolo of Trieste (RNAseq-Dataset 1, Trieste) and Asst degli Spedali Civili di Brescia/University of Brescia and patients referred for a transcriptomics

workup from 9 external gastroenterology (RNAseq-Dataset 2, Brescia) facilities were included.

The work was performed as complementation of the previous genetic workup of patients recruited at Institute for Institute for Institute for Maternal and Child Health IRCCS Burlo Garofolo of Trieste, if RNA were available at our biobank, and as new prospective study involving mainly patients from the cohort of Brescia.

A set of 4 young-age healthy individuals, 2 males and 2 females, was used as reference to compare patients' gene expression profiles.

The work was conducted in accordance with the revised Declaration of Helsinki and was approved by Burlo Garofolo and Asst degli Spedali Civili di Brescia.

Written informed parental consent was obtained for genetic and gene expression analyses.

The three cohort of patients and the group of controls are summarized in Table 1s (supplementary material).

3.2 DIAGNOSTIC WORKFLOW

In the prospective phase of the study, patients enrolled for genetic workup were screened using NGS technologies, with the exception of patients with well-defined phenotypes, suggestive of a specific monogenic disorder, for whom single gene sequencing was chosen. An IBD targeted gene panel sequencing (TGPS) analysis was performed in the majority of patients as the first line diagnostic tool. This IBD TGPS includes genes causative for the main monogenic diseases associated to IBD symptoms. Patients with a suspect of an IPEX-like performed also an extra TGPS panel containing *LRBA* and *CTLA4*. Beginning in 2017, WES replaced TGPS due to a significant decrease in WES costs. Whole exome sequencing analysis was initially restricted to a set of 400 genes plus the list of genes associated with primary immunodeficiency and related pathways as described by Kelsen et al [120]. Trio-WES was used in selected cases of patients with IOIBD and severe disease when parental DNA was available. Basic immunological workup included complete blood count, immunoglobulin levels, lymphocyte subsets, and neutrophil function studies [119].

3.2.1 TARGETED GENE PANEL DESIGN

Two custom-made panels for IBD TGPS were designed: the first panel, designed and performed by us, at Burlo Garofolo, included 30 genes (Panel A); the second panel,

designed and performed at Bambino Gesù Hospital, included 43 genes (Panel B). The full list of genes included in the panels and gene coverage is illustrated in Table 3. Gene selection for both panels was based on lists of genes suggested by Kammermeier et al, [88] Uhlig et al, [31] and Christodoulou et al. [121]. Genes associated with diseases presenting with well-defined phenotypes that had valid structured functional tests, such as Wiskott-Aldrich syndrome and Hyper IgM syndrome (HIGM), were not included in the panels. An extra small TGPS panel (data not shown) containing *LRBA* and *CTLA4* was designed, at a later time, for the diagnosis of IPEX-like subjects [119].

Functional Group	Condition	Gene	Chromosome	Inheritance	Coverage %	
					Panel A	Panel B
Epithelial barrier dysfunction	ADAM-17-deficiency	*ADAM17*	2p25	AR	98,78	-
	Familial diarrhoea	*GUCY2c*	12p13	AD	98,75	-
		EPCAM	2p21	AR	-	100
	Kindley syndrome	*FERMT1*	20p12	AR	98,29	-
	Neonatal inflammatory skin and bowel disease 2	*EGFR*	7p11	AR	97,13	-
	Gastric cancer, familial diffuse, with or without cleft lip and/or palate	*CDH1*	16q22	AD	-	100
Phagocyte defects	Leucocyte adhesion deficiency Type 1	*ITGB2*	21q22	AR	99,71	
	Severe Congenital neutropenia 4	*G6PC3*	17q21	AR	99,17	
Hyper/auto-inflammation	Phospholipase Cγ2- deficiency	*PLCG2*	16q23	AD	98,25	
	Familial haemophagocytic Lymphohistiocytosis Type 5	*STXBP2*	19p13	AR	98,02	
	X-linked lymphoproliferative syndrome Type 2	*XIAP*	Xq25	XLR	80,84	
	Hermansky-Pudlak syndrome	*HPS1*	10q23	AR	92,71	
	Hermansky–Pudlak syndrome	*HPS4*	22q12	AR	95,21	
	Hermansky–Pudlak syndrome	*HPS6*	10q24	AR	85,16	
	Autoinflammation with infantile enterocolitis	*NLRC4*	2p22	AD	100	
	Mevalonate kinase deficiency	*MVK*	12q24	AR	100	
Immune-regulation	X-linked immune dysregulation, Polyendocrinopathy, Enteropathy (IPEX) syndrome	*FOXP3*	Xp11	XLR	98,93	
	IL10 pathway defects	*IL10*	1q32	ND	79,99	
	IL10 pathway defects	*IL10RA*	11q23	AR	96,46	
	IL10 pathway defects	*IL10RB*	21q22	AR	85,85	
Disorder of apoptosis	Autoimmune disease, multisystem with facial dysmorphism	*ITCH*	20q11	AR	96,64	
	MASP2 deficiency	*MASP2*	1p36	AR	100	
Others	Trichohepato-enteric syndrome	*TTC37*	5q15	AR	97,11	
	Trichohepato-enteric syndrome	*SKIV2L*	6p21	AR	97,5	
	Multiple intestinal atresia	*TTC7A*	2p21	AR	97,5	
	Inflammatory bowel disease 1	*NOD2*	16q12	Mu	100	
	Inflammatory bowel disease 1	*TRIM22*	11p15	NA	99,14	
	Inflammatory bowel disease 10	*ATG16L1*	2q37	NA	100	
	Inflammatory bowel disease 17	*IL23R*	1p31	NA	90,43	
	Inflammatory bowel disease 12	*MST1*	3p21	NA	93,72	
	prolidase deficiency	*PEPD*	19q13	AR	89,28	
	Dyskeratosis congenita, X-linked	*DKC1*	Xq28	XLR	96,76	

Table 3. *Genes included in targeted gene panels and gene coverage.*

3.2.2 DNA LIBRARY PREPARATION AND RAW DATA ANALYSIS

DNA library were constructed using Ion Ampliseq Library Kit 2.0 and each sample was labelled with an Ion Xpress Barcode Adapters Kit (Life Technology, CA, USA) according to manufacturer's protocol. The sequencing step was performed on the Ion TorrentTM platform after the libraries' amplification on Ion Sphere Particles (ISP) using the Ion OneTouchTM 2 system (Life Technology, CA, USA).

The signal processing derived from the Ion sequencer was analyzed by the Torrent SuiteTM software v5.2 that also performed the base calling, the alignment of the trimmed reads to the human genome reference (GRch38/hg38) and the variant calling. The output file was further annotated using wANNOVAR software (http://wannovar.wglab.org/) [119, 122].

3.2.3 WES AND DATA ANALYSIS

For 6 patients, WES data were available besides the TGPS data and for 11 patients was performed directly the WES.

The library preparation and sequencing were carried out by the outsourced service from Macrogen Inc (Korea).

The exomes were enriched with SureSelect Human All Exon v4 Kits (Agilent Technologies, Santa Clara, CA, USA) and the sequencing of 2 X 150bp were made in Illumina HiSeq 2500 systems.

Data were analyzed using the Genome Analysis Toolkit (GATK), SAMtools, and Picard, according to documented best practices (https://software.broadinstitute.org/gatk/best-practices/). The annotation of VCF was performed using Annovar [119, 123]

3.2.4 VARIANT SELECTION AND VALIDATION

Data were filtered selecting nonsynonymous, nonsense, frameshift, splicing (about 10 nucleotides from the splice site), and variants, which were either absent or had a minor allele frequency (MAF) <0.02 (in case of recessive model) or MAF <0.001 (in case of dominant inheritance model and in case of de novo variants). Minor allele frequency selection was based on 1000 GenomesProject (1000genomes.org) database and ExAC browser (exac.broadinstitute.org). Moreover, all variants were interrogated by Genomic Evolutionary Rate Profiling (GERP) score [124] as a measure of the conservation of the genomic position. Genetic variants were classified according to

the American College of Medical Genetics (ACMG) guidelines [125] into “pathogenic,” “likely pathogenic,” or “variants of uncertain significance” using dedicated tools [126]. Nonsynonymous variants were further selected according to 5 different in silico prediction tools, namely CADD (score > 15) [127], Mutation Taster [128], Polyphen-2 [129], SIFT [130], and LRT [131]. Among the selected variants, those with a pathogenic prediction in at least 2 out of the 4 tools were retained. Human Splicing Finder v3.1 (umd.be/ HSF3) was used to predict the effect of splicing variants.

The clinical significance of variants, already described in public databases, and the association with specific phenotypes were investigated using OMIM (omim.org), ClinVar (ncbi. nlm.nih.gov/clinvar), and HGMD (Human Gene Mutation Database) professional. For novel mutations, pathogenicity was established with a functional assay, when available, or inferred from similar mutations with known clinical significance or based on the presence of highly specific clinical features [119].

Variants considered to be causative, according to the clinical phenotype and the mode of inheritance, were validated by Sanger Sequencing in patients and their parents, when available, after visualizing the read coverage of each mutation using the Integrative Genomics Viewer (IGV) (software.broadinstitute.org/software/igv) [132, 133]. Primers were designed using Primer Blast tool (ncbi.nlm.nih.gov/tools/primer-blast) and synthesized by Eurofins Genomics (eurofinsgenomics.eu). DNA regions were amplified by standard PCR protocols and sequenced in both directions. Sequences were evaluated using CodonCode Aligner 6.0.

3.2.5 STATISTICAL ANALYSES

Statistical analyses were made using GraphPad Prism version 8. Categorical variables were summarized as frequency and percentage and were compared across independent groups by the Fisher exact test. Numerical variables with asymmetrical distribution were summarized by median and interquartile range (IQR) and were compared by the Kruskal-Wallis test. A P value <0.05 was considered for significance.

3.3 RNAseq WORKFLOW

3.3.1 RNA EXTRACTION AND RAW DATA PROCESSING

Peripheral blood was collected in PAXgene Blood RNA Tubes (PreAnalytiX, Hombrechtikon, Switzerland) and, after two-hours incubation at room temperature, tubes were frozen at −20 °C until processing. Total RNA was extracted with PAXgene Blood RNA Kit (PreAnalytiX, Switzerland), following the manufacturer's instructions, and quantified with NanoDrop Spectrophotometer (Thermo Fisher, Waltham, MA, USA). RNA integrity was checked using an Agilent Technologies 2100 Bioanalyzer.

Transcriptome sequencing was performed using the TruSeq Stranded mRNA Sample Preparation kit (Illumina, San Diego, CA, USA) and sequenced on a NovaSeq 6000 platform (Illumina, San Diego, CA, USA), generating 2X100 bp paired-end reads (30 million reads per sample).

RNAseq raw data workflow was conducted as follows: quality control by FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), quality filtering by Trim Galore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), read alignment to hg38 using annotation from GENECODE v.31 (https://www.gencodegenes.org/) with STAR [134], reads counting into genes by featureCounts [135].

3.3.2 DIFFERENTIAL GENE EXPRESSION ANALYSIS

Differential gene expression analysis was performed in R by the Bioconductor package DESeq2 [136] both compared with healthy individual group and within different VEOIBD subgroups. Genes have been chosen by fold change greater than 2-fold increase/decrease and adjusted p-value < 0.05, according to the Benjamin-Hochberg method. Selected gene have been analyzed for pathway enrichment running the R package pathfindR (p-value 0.1) [137].

Principal component analysis (PCA), useful for data visualization, was conducted with DESeq2, to define the overall variability between samples. Before running PCA gender difference correction has been performed since the focus of the analysis was to highlight changes due to variability between individuals unrelated from gender. This step was performed by running the R function removeBatchEffect ("limma package").

Differential gene expression profile was conducted to make different comparisons that will be better described in Results and Discussion session:

- VEOIBD patients vs healthy individuals
- monogenic vs nonmonogenic VEOIBD
- XIAP_1TS vs 2 healthy individuals
- TTC37_4TS vs 2 healthy individuals
- DKC1_5TS vs 2 healthy individuals
- LRBA_6TS vs 2 healthy individuals

The DEGs output between monogenic and nonmonogenic VEOIBD was visualized by an heatmap in the Results and Discussion session. The heatmap plotted the 5000 most variable genes across all sample, calculated by the R function rowVars that estimate the sample variance within each row of a matrix.

3.3.3 CLUSTER ANALYSIS TO CLASSIFY PATIENTS INTO DIFFERENT SUBSETS

Unbiased patients subdivision, considering both genetically assessed and undiagnosed individuals, has been performed by an unsupervised non-hierarchical cluster (K-Means clustering) provided by R (http://www.R-project.org/). This analysis algorithm partitions patients into subgroups characterized by similar observations considering, in this case, the expression of coding protein genes. Clustering results were visualized employing the R functions fviz_cluster (“factoextra package”) that performs principal component analysis (PCA). Random sets were set on 40 and the number of centers have been choosen previously according to the "silhouette" method by fviz_nbclust (“factoextra package”).

3.3.4 MACHINE LEARNING APPROACH FOR SIGNATURE DISCOVERY

The identification of possible signatures, characterized different VEOIBD subgroups, has been performed by a machine learning approach that includes three binary classifiers: Partial Least Square Discriminant Analysis (PLS-DA), Random Forest (RF) and Support Vector Machines (SVM).

Gene expression data have been transformed by rlog function by DESeq2 package that converts the count data to the log2 scale in a way which minimizes differences

between samples for rows with small counts, and which normalizes with respect to library size.

This algorithm has been run in R by the package Biosigner [138] to detect possible genes or functional patterns, able to distinguish different VEOIBD clusters. Relevant features are found by this technique splitting into training and testing subsets (by bootstraping, controlling class proportion) [138]. In this analysis the number of bootstraps for resampling has been set to 50 and an adjusted p-value < 0.1 was considered. The output returns a level of each feature for the chosen classifiers, level S corresponds to the final signature that means to features which have been found significant in all the selection steps; features with level A have been found significant in all but the last selection, and so on for level B to D. Level E regroup the previous selection cycles.

Candidate genes have been analyzed calculating the z-score, to compare different population, from the rlog transformed values. Statistical analyses were made using GraphPad Prism version 8 ordinary one-way Anova for multiple comparisons or t-test when considering two groups.

A P value <0.05 was considered for significance.

4 RESULTS and DISCUSSION

4.1 PATIENTS

A total of 94 patients diagnosed with VEOIBD and EOIBD s/a were collected mainly by the 2 pediatric gastroenterology centers, Institute for Maternal and Child Health IRCCS Burlo Garofolo of Trieste and the Bambino Gesù Children's Hospital of Rome, and in lower portion from other 9 external gastroenterology facilities; of these, 55 patients (58.5%) had disease onset within the first 2 years of life, and 6 patients (6.4%) had disease onset between 2 and 6 years. 56 patients (59.6%) were males; 7 patients (7.4%) had a family history of IBD among first degree relatives; 2 patients (2.1%) had a sibling who had died in infancy or early childhood [119].

4.2 GENETIC WORKUP AND DIAGNOSIS

Genetic analyses have been performed in collaboration with Bambino Gesù Children's Hospital. We identified that about 15% of patients with VEOIBD has a rare monogenic disorder. The spectrum of monogenic diseases with intestinal inflammation could be caused by mutations in several genes mainly involved in immune system functions, inflammatory homeostasis, and intestinal epithelial barrier.

An integrative approach between molecular genetics and clinical workup is pivotal to diagnose monogenic IBD.

Over time, the use of NGS has become most common in clinical practice, particularly in patients with non-specific clinical and immune phenotypes, given the possibility to screen multiple genes simultaneously.

In our cohort, 47 patients (50%) underwent Sanger sequencing of 1 or multiple genes over time. In 8 patients, single gene sequencing was guided by the presence of supportive clinical and immunological features. Next generation sequencing was performed in 85 patients (90%) and consisted of TGPS in 70 of 85 patients (82%), "clinical WES", i.e. WES performed in the sole proband followed by in silico analysis of 400 target genes described by Kelsen et al [120], in 16 (19%), and trio-WES, with analysis of likely pathogenic variants in any gene, in 5 (6%). Of the patients who underwent NGS, 40 (45%) had been studied previously with a single gene approach with negative results. The proportion of patients who underwent NGS as the first

molecular analysis has increased over time. Among patients diagnosed with IBD before the year 2011, only 25% (7 of 29 patients) underwent NGS as the first genetic analysis; the proportion raised to 45% (16 of 35) between 2011 and 2014 and to 79% (23 of 29) after 2015.

NGS data interpretation required a variant selection process set-up. Figure 5 shows the variant selection workflow and reports the mean number of variants among all patients evaluated by NGS after each sequential step: total detected variants; variants filtered by allele frequency and/or in silico target analysis; variant filtered by consequence type; variant predicted pathogenic by bioinformatics tools; variants validated by concordance of clinical features, mode of inheritance and previous literature findings.

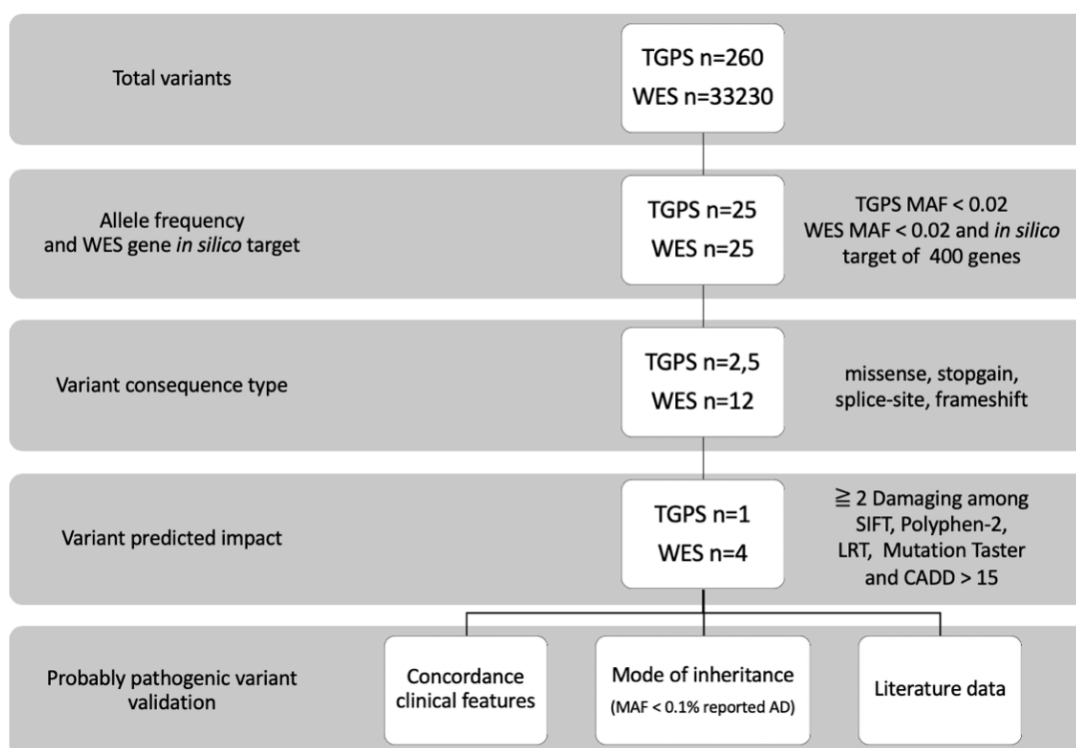


Figure 5.
Variant selection workflow. It shows the mean number of variants among all patients evaluated by NGS techniques.

Overall, genetic analysis revealed 13 cases (14%) of monogenic IBD. Table 1(a, Institute for Institute for Maternal and Child Health IRCCS Burlo Garofolo of Trieste (TS); b, Bambino Gesù Children’s Hospital [119].

Patient (Sex)	IBDonset (months)	Initial endoscopy	GI disease	Extraintestinal findings	Lab work-up	Treatment	Genetic Variants (zygosity)	Impact of genotype
1TS (M)	2	AI	Extensive colitis Apoptosis	Persistent fever, CMV infection, HLH	↑WBC	EN, steroids, AZA, Anti-TNF, tacrolimus, colectomy,	*XIAP* RefSeq NM_001167, c.1021_1022delAA:p.N341Yfs X7 (hem)	HSCT
2TS (M)	108	CD-like	Colitis, p.	Arthritis, vasculitis, PG, uveitis, nephritis,	↓PLT, ↑IgA,↓IgM,IgG	Steroids, anti-TNF, MTX cyclosporine, thalidomide, fistulotomy, colectomy	*WAS* Gene inversion (hem)	Anti IL-1, gene therapy
3TS (M)	0	EOS	Extensive colitis	CMV infection	↓PLT	Steroids	*WAS* RefSeq NM_000377c.257G>A :p.R86H (hem)	HSCT
4TS (F)	96	CD-like	Colitis, p	Trichorrhexis nodosa, syndromic facies, hepatopathy	↑Ig A	Anti-TNF	*TTC37* RefSeq NM_014639, c.4497-1G>A (hom)	Genetic counselling
5TS (M)	16	CD-like	Enterocolitis, apoptosis	Leukoplakia, nail dystrophy, skin reticulate	↓NK, B	Steroids, 5-ASA, anti-TNF, thalidomide, colectomy	*DKC1*;RefSeq NM_001363,c.146C>T:p.T49M rs121912304 (hem)	Danazol
6TS (M)	36	Gastric CD	Autoimmune gastritis	Autoimmunity	↓RTE ↑DNT	Prednisone tacrolimus, lansoprazole	*LRBA*; Refseq NM_006726: c.C6415T:p.R2139X c.C7315T:p.R2439X (c-het)	Abatacept
7TS (M)	20	CD-like	Enterocolitis, p, ileal fistulas	Recurrent respiratory infections	↓Treg,B ,↑IgM,↓IgA, IgG	EN, ileostomy	*CD40L*: RefSeq NM_000074,c.585dupA:p.L195fs (hem)	HSCT

Table 4a. *Monogenic IBD patients diagnosed at for Institute for Maternal and Child Health IRCCS Burlo Garofolo of Trieste.*

Patient (Sex)	IBDonset (months)	Initial endoscopy	GI disease	Extraintestinal findings	Lab work-up	Treatment	Defective gene	Impact of genotype
1RM (M)	48	IBD-U	Colitis	Sclerosing colangitis, cryptosporidium	↓B, ↓Ig, ↑Eos	EN, steroids,	*CD40L*: RefSeq NM_000074, c.410-2A>T (hem)	HSCT, liver transplant
2RM (M)	10	IBD-U	Enterocolitis	Liver ascess, eczema	DHR defective	EN, steroids, 5-ASA	*CYBA* RefSeq NM_000101 del ex6 (hom)	Prophylaxis
3RM (M)	30	CD-like	Colitis, p	Skin granulomas	DHR defective	EN, steroids, 5-ASA, AZA	*CYBB* NM_000397, c.252G> A 3' exon 3++(hem)	Prophylaxis
4RM (M)	70	CD-like	Enteropathy	Complicated EBV, HLH	↓Ig	EN, steroids, AZA, anti-TNF	*XIAP* RefSeq NM_001167, c.566T>C :p.L189P (hem)	HSCT
5RM (M)	1	CD-like	Enterocolitis	Candidiasis, psoriasis, opportunistic infections	↓PLT	EN, steroids	*FOXP3* RefSeq NM_014009:ex11,c.1078C>T :p.L360F (hem)	HSCT
6RM (M)	1	CD-like	Enterocolitis	Arthritis, infections, eczema	↓PLT, ↓WBC	EN, steroids, 5-ASA, cyclosporine.	*WAS*, na (hem)	HSCT

Table 4b. *Monogenic IBD patients diagnosed at Bambino Gesù Children's Hospital of Rome.*

Abbreviations: AI, autoimmune enteritis; AC, allergic colitis; EOS, eosinophilic enteropathy; p, perianal disease; PG, pyoderma gangrenosum, PLT, platelets, WBC, white blood cells; EN, enteral nutrition; AZA, azathioprine; MTX, methotrexate; BMT, bone marrow transplantation; MBC, Memory B cells; ++splice-site mutation; na, not available.

A subdivision of monogenic patients according to functional defect category is illustrated in Figure 6. A single gene approach was diagnostic in 8 out of 47 patients (2WAS, CYBA, CYBB, FOXP3, 2CD40L, XIAP). In 7 out of the 8 patients diagnosed by Sanger sequencing, the analysis was guided by the presence of disease specific features. One patient with XIAP deficiency (1TS) who had nonspecific presentation underwent sequential sequencing of multiple genes over a period of 15 months. During this time, the patient experienced recurrent bouts of HLH, failed several immunosuppressive therapies, became dependent on parenteral nutrition, and ultimately underwent a total colectomy. After the diagnosis of XIAP deficiency (1TS), he received BMT that led to a complete cure [139].

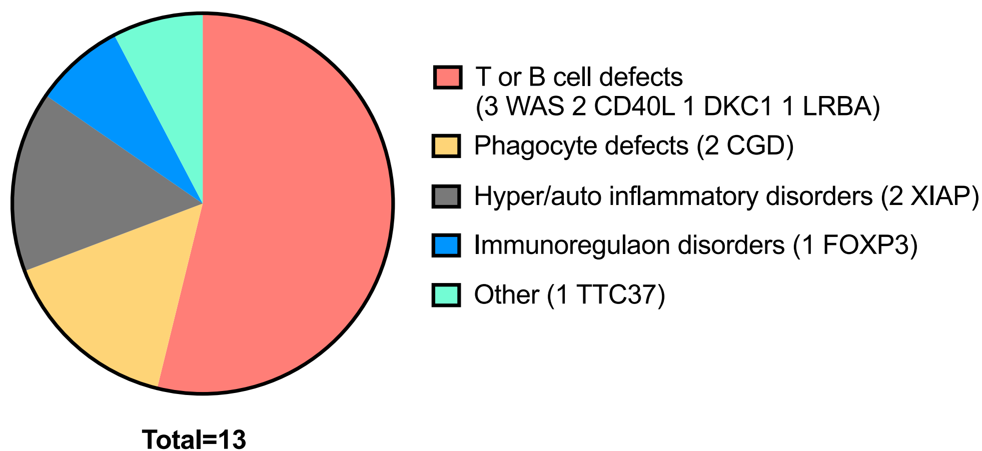


Figure 6.
Functional defect category of monogenic patients [119].

Next generation sequencing was performed as a first step in 46 patients and revealed causative genetic defects, all of them through TGPS, in 3 patients (6.5%) (i.e. TTC37_4TS, DKC1_5TS, XIAP_1TS defects). 40 patients underwent NGS as a second step. Among them, 2 patients with WAS (2TS, 3TS) and LRBA deficiency (6TS), in whom Sanger sequencing or IBD TGPS had not revealed mutations, were diagnosed by whole genome sequencing (elsewhere) that showed a large genomic inversion [140] and by a small TGPS including *LRBA* and *CTLA4* respectively. Additionally, with the use of WES, a rare homozygous variant on NOD2 nucleotide-binding domain was found in one male patient with IBD onset at the age of five months and associated arthritis. The role of such variant has been studied through bioinformatics and functional studies, which demonstrated that the consequence of the mutation was an auto-activation of NOD2-mediated NF-kB signaling, similar to that

described in patients with Blau Syndrome [141]. However, these findings were not enough to confirm a monogenic disease. The diagnostic steps and the rates of monogenic diagnosis with the different diagnostic approaches are summarized in Figure 7.

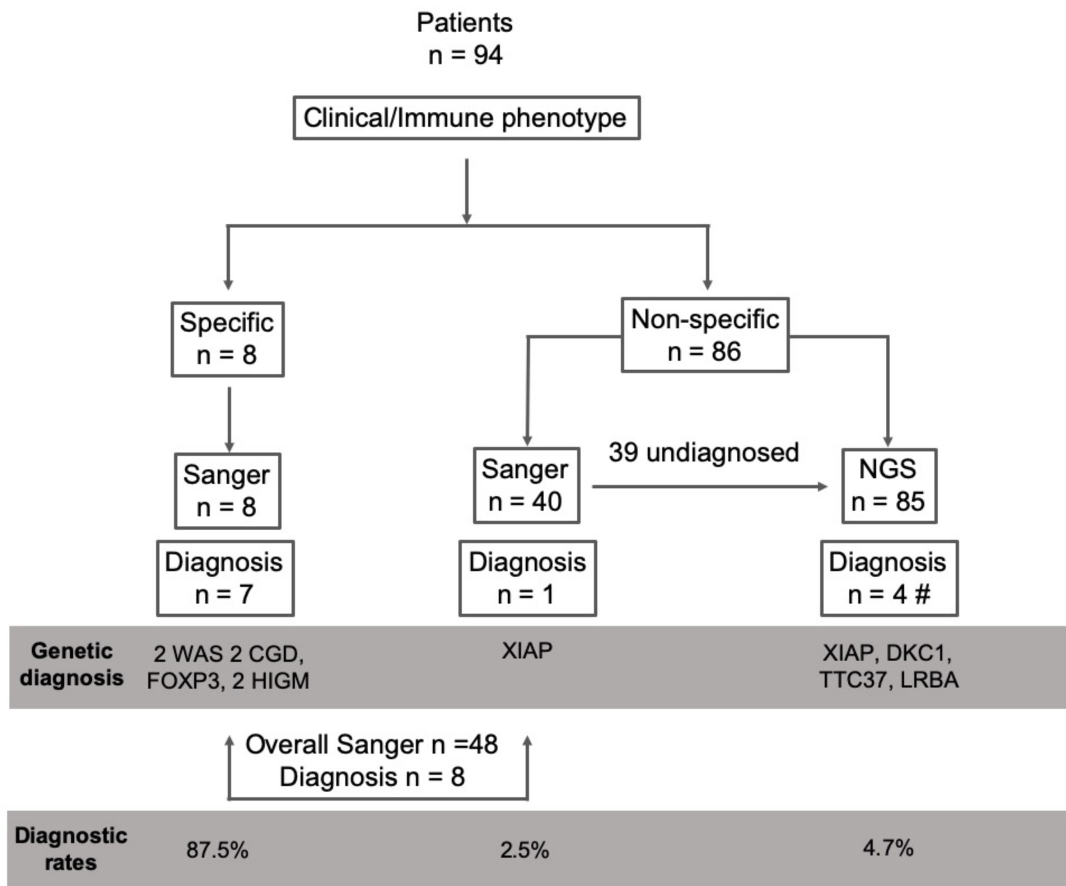


Figure 7.
Diagnostic steps and rates of monogenic diagnoses with the different diagnostic approach. (#plus 1 patient diagnosed elsewhere with WAS gene inversion through whole genome sequencing).

Genetic diagnosis impacted patient management in 12 patients (92.3%): 7 patients (2XIAP (1TS, 6RM), 2WAS (3TS, 4RM), 2CD40L (7TS, 1RM), FOXP3 (5RM)) underwent BMT; 1 patient with *WAS* (2TS) gene inversion introduced anti IL1 antagonist (anakinra), which led to the resolution of severe pyoderma gangrenosum and arthritis before undergoing gene therapy [140]; 2 patients (2RM, 3RM) with chronic granulomatous disease (CGD) introduced anti-infective prophylaxis; the patient (5TS) with dyskeratosis congenita (DKC1) introduced danazol as a telomere elongating therapy; the patient (6TS) with LPS-responsive beige-like anchor (LRBA) deficiency introduced cytotoxic T-lymphocyte–associated antigen 4-immunoglobulin (abatacept)

which led to an improvement of vomit episodes, evidence of lymphadenomegaly and fever [119].

The distribution of patients with a monogenic diagnosis according to age at IBD onset is illustrated in Figure 8.

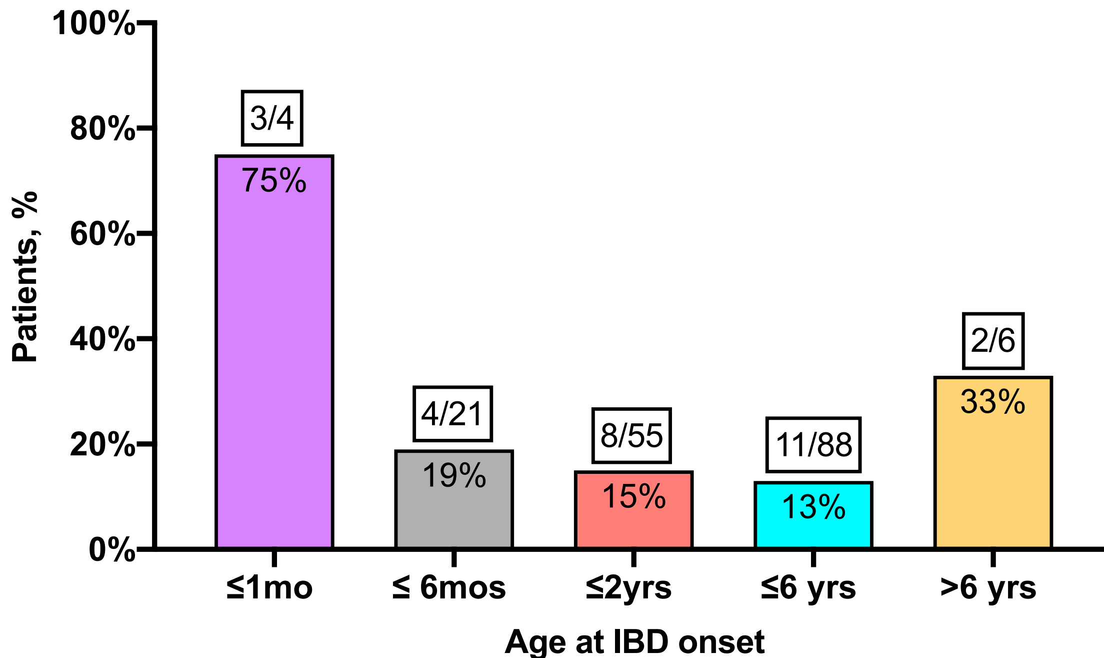


Figure 8.
Distribution of patients with monogenic IBD within different age groups [119].

In our cohort the diagnostic approach to suspected monogenic IBD has changed over time. Most of the patients with IBD onset before 2011 underwent a single gene approach. However, more recently, NGS is used as the first line diagnostic step in most of the patients. In our study the molecular diagnostic yield of NGS was 6.5% when performed as a first step (XIAP (1TS), DKC1 (5TS), TTC37 (6TS) defects) and 4.7 % overall. These rates are lower in comparison to previous VEOIBD work by Kammermeier et al. that reported a diagnostic yield of 16% performing a TGPS of 40 genes [88] and Charbit-Henrion et al who, using a TGPS with 66 genes, reported a variable diagnostic yield of 14% to up to 26.5% when TGPS was used either as a second line investigation or as a first screening, respectively [142]. This rate discrepancy could be due to a few factors. Firstly, in both cohorts there are some biases of patient selection: the majority of patients had a disease onset before the age of 2 years and the study by Charbit-Henrion et al [142] included only patients with a severe disease course and thus a likely higher probability for monogenic diseases.

Secondly, our TGPS do not include genes known to be associated with recognizable phenotypes or have valid functional tests for which we used a clinically and immunologically driven single gene approach. Including these genes in the TGPS panels might have increased the NGS diagnostic yield in our series. In our study, a single gene approach, indeed, lead to a good diagnostic performance only when guided by clinical and/or immunological features specific for known monogenic defects (detection of causative mutation in 7 out of 8 cases), such as CGD and WAS, and unsatisfactory results in patients with nonspecific phenotypes. In the latter subgroup, only 1 out of 40 patients (2.5%) could reach a molecular diagnosis of XIAP deficiency, and the diagnostic process in this case implied multiple single-gene-sequencing over 15 months. The use of NGS could have led to an earlier diagnosis and avoided the development of several complications in the period from the first symptoms manifestations and the molecular diagnosis.

Overall, in our cohort, a monogenic diagnosis has been reached in 14% of patients considering the combination of various genetic approaches. This result agrees with previous reports in which monogenic form were diagnosed in 12% of VEOIBD and 15% of IOIBD [88, 143]. However, we reported an increased rate of monogenic diagnosis among patients with disease onset before 6 months of life and particularly among patients with a disease onset during the first month of life; in these subgroups, a monogenic diagnosis could be established in 19% and 75% of patients, respectively. A molecular diagnosis was made also in two patients who manifested IBD symptoms after 6 years of age (WAS (2TS), TTC37 (4TS) defects), but probably they developed disease specific features earlier than IBD.

Interestingly, differently from other works, no patient with IL10 or IL10R defects were found in our cohort. Even though, the frequency *IL10* or *IL10R* mutations in the Italian population is unknown, it should be noted that most patients with IL10 pathways defect reported so far were of Arab, with a high frequency of history of consanguinity, or Asian descent. A possible explanation for our observation is that in our cohort, only 3 patients of non-white ethnicity (2 Arab and 1 Asian) were included, and parental consanguinity was not reported in any of the patients [38, 144].

From a clinical point of view, monogenic and nonmonogenic patients do not seem to differ in inflammatory bowel diseases severity. However, the presence of extraintestinal manifestation were present in all patients with a genetic molecular diagnosis.

Establishing a genetic diagnosis had an impact in the majority of cases. The most frequent consequence was BMT. BMT is a treatment of choice for example in XIAP defects that could lead to the intestinal inflammation remission. However, BMT is not suggested for epithelial barrier defect, since it fails to correct the epithelial disorder and patients had a poor outcome after BMT [145, 146]. In our cohort we identified 2 patients with a defect impacting the epithelial barrier (TTC37 (4TS), DKC1 (5TS)) that indeed were not treated with BMT [119]. Treatments targeted to correct epithelial dysfunction in these patients have not been developed so far. However, at least in the case of DKC1 defect, it could be reasonable to evaluate the potential of hormones stimulating gut epithelia (such a teduglutide, developed for short bowel syndrome).

4.3 RNAseq WORKUP

4.3.1 PATIENT DESCRIPTION

13 out of 94 patients who already conducted the previously described genetic workup, performed also gene expression analysis by RNAseq of peripheral blood cells. The choice of sample size was dependent on patients' RNA collection, which was available only for 13 subjects collected at Institute for Maternal and Child Health IRCCS Burlo Garofolo of Trieste (RNAseq-Dataset 1, Trieste). 4 had a diagnosis of monogenic disease (XIAP_1TS, TTC37_4TS), DKC1_5TS and LRBA_6TS defects). All of them, except for the patients with *TTC37* defect, had disease onset within the first 6 years of life. 8 were males.

Extraintestinal findings were reported in all monogenic IBD and in 4 out of 9 nonmonogenic IBD, as discussed above.

The laboratory workup showed an altered white blood cells count (WBC) in 2 patients, immunoglobulin impairment in 2 patients, c-reactive protein (CRP) increased in 4, fecal calprotectin (FCP) increased in 6 patients.

Table 5a reports patients' clinical, laboratory, and genetics findings.

An extra cohort of 13 patients with VEOIBD or EOIBD s/a were enrolled, in the last year, as the first part of a collaborative project with Spedali Civili di Brescia/University of Brescia (RNAseq-Dataset 2, Brescia, Table 5b). 2 patients had disease onset within the first 2 years of life and 4 had disease onset above 6 years. 6 patients were males. The genetic diagnostic workup is still in progress.

Patient (Sex)	IBDonset (months)	Initial endoscopy	Extraintestinal findings	Lab work-up	Treatment at RNA sampling time	Genetic results
1TS (M)	2	AI	Persistent fever, CMV infection, HLH	↑WBC	Anti-TNF	XIAP (monogenic)
4TS (F)	96	CD-like	Trichorrhexis nodosa, syndromic facies, hepatopathy	↑IgA	Anti-TNF	TTC37 (monogenic)
5TS (M)	16	CD-like	Leukoplakia, nail dystrophy, skin reticulate	↓NK, B	Danazol Thalidomide	DKC1 (monogenic)
6TS (M)	36	Gastric CD	Autoimmunity	↓RTE ↑DNT	Prednisone Tacrolimus, Lansoprazole	LRBA (monogenic)
8TS (F)	37	CD	Hyperferritinemia-cataract syndrome	↑CRP ↑FCP ↑PLT	Anti-TNF	Negative (nonmonogenic)
9TS (F)	13	IBD-U	None	↑CRP ↑FCP	Mesalazine	Negative (nonmonogenic)
10TS (M)	68	CD	MAS	↓WBC ↑IgA ↓IgM ↑FCP ↓PLT	No therapy	Negative (nonmonogenic)
11TS (F)	27	UC	NEC 30 weeks gestational age	↑CRP ↑FCP	Vitamin B12	Negative (nonmonogenic)
12TS (M)	3	UC	Axillary hydrosadenitis	Normal	Anti-TNF	Negative (nonmonogenic)
13TS (M)	57	UC	None	Normal	Mesalazine	Negative (nonmonogenic)
14TS (M)	3	EOS	None	↑FCP	No therapy	Negative (nonmonogenic)
15TS (M)	68	CD	None	↑FCP	Anti-TNF, Thalidomide, Salazopyrin	Negative (nonmonogenic)
16TS (F)	2	IBD-U	None	↓PLT	AZA	Negative (nonmonogenic)

***Table 5a**. RNAseq Dataset 1, Trieste*
Patients' clinical, laboratory, genetics finding. Each patient is labelled with a specific number. 1TS, 4TS, 5TS, 6TS are monogenic IBD previously described in Table 4a, the other patients 8TS-16S are nonmonogenic IBD firstly reported in this table.

Clinical and genetic characteristics of patients from RNAseq-Dataset 1, Trieste. Abbreviations: AI, autoimmune enteritis; CMV, cytomegalovirus; HLH, hemophagocytic lymphohistiocytosis; WBC, white blood cells; NK, natural killer cells, B, B cell; CRP, C-reactive protein; FCP, fecal calprotectin; PLT, platelets; MAS, macrophage activation syndrome; NEC, necrotizing enterocolitis; EOS, eosinophilic enteropathy; AZA, azathioprine; RTE, recent thymic emigrants; DNT, double negative T cell.

Patient n (Male/Female)	IBDonset (n/n tot)	Genetic analysis	Lab work-up
13 (6/7)	≤ 2 yrs (2/13) ≤ 6 yrs (7/13) > 6 yrs (4/13)	In progress	In progress

Table 5b. *RNAseq Dataset 2, Brescia.*

4.3.2 VEOIBD SUBGROUPS: THE WEIGHT OF THERAPY

The majority of patients with VEOIBD remains without a genetic diagnosis and thus probably likely multifactorial. The study of expression profiles by RNAseq may contribute to identify the defective mechanisms underlying nonmonogenic VEOIBD in comparison with monogenic IBD, allowing the definition of subgroups that present similar features.

We performed gene expression profile and clustering in monogenic IBD and nonmonogenic IBD from RNAseq-Dataset 1 compared to a group of four young-age healthy individuals.

Besides genetics, different factors may influence gene expression, such as gender and pharmacological treatments. The first step was to determine possible differences due to therapies.

Table 6 summarizes the therapeutic management. 4 of 13 patients took anti-TNF drug at the time of RNA sampling. 2/5 had monogenic IBD (TTC37_4TS and XIAP_1TS).

	n (%)
Anti-TNF	*4 (31)*
Danazol	*1 (8)*
Thalidomide	*1 (8)*
Prednisone	*1 (8)*
Tacrolimus	*1 (8)*
Lansoprazole	*1 (8)*
Mesalazine	*2 (15)*
Salazopyrin	*1 (8)*
AZA	*1 (8)*
Vitamin B12	*1 (8)*
No therapy	*2 (15)*

Table 6. *Summary of therapeutic approach for thirteen RNAseq-Dataset 1 patients. Abbreviations: AZA, azathioprine.*

PCA allowed to summarize and visualize the distribution of patients with VEOIBD considering protein-coding gene expression according to their pharmacological treatments. Legend of Figure 9 reported therapies shared from at least 2 subjects and the remaining are included in “other”.

Although patients that had “minimal therapies” (i.e. mesalazine), one patient with no therapy and healthy subjects clustered closely, no evident drug-specific groups are delineated, indicating a possible secondary role of therapies in influencing gene expression profiles in this cohort of patients. Gender differences were already corrected before running PCA.

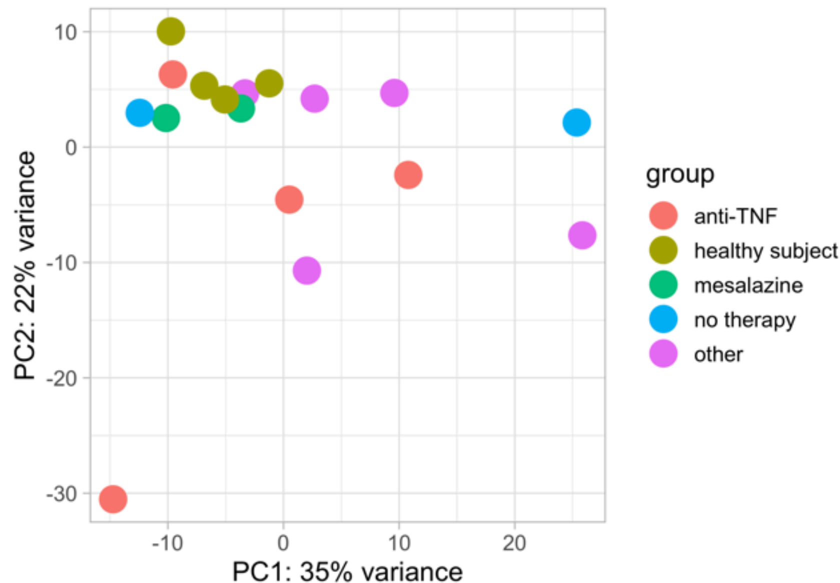


Figure 9.

PCA protein-coding gene expression focused on patients' distribution according to their pharmacological treatments. In red patients treated with anti-TNF, in light green healthy subjects, in dark green patients treated with mesalazine, in light blue patients without any treatments and in magenta single patients treated with other drugs (Danazol, Thalidomide, Prednisone, Tacrolimus, Lansoprazole, Salazopyrin, AZA and vitamin B12).
Each dot represents one patient.

4.3.3 VEOIBD SUBGROUPS: HEALTHY SUBJECTS COMPARISON (VEOIBD vs healthy subjects)

Given that the differences recorded among samples seemed only partially due to therapies, 2 unsupervised clustering algorithms (hierarchical and k-means) were ran to group patients with similar features. In this analysis were also included healthy individuals.

The choice of considering protein-coding genes, results from no variation detection in the overall distribution of patients after PCA (data not shown) when including all the polyA transcriptome or just the protein-coding subset.

The heatmap below (Figure 10a) plotted the 5000 most variable genes across all samples. Rows and columns are sorted by hierarchical clustering trees. Dendrograms, on the top of the figure, get close patients with similar features.

Before running the k-means algorithm, the optimal number of clusters was defined by the average silhouette method that measure the quality of clusters (Figure 10 b,c).

PCA analysis showed similar results (Figure 10d).

Figure 10.

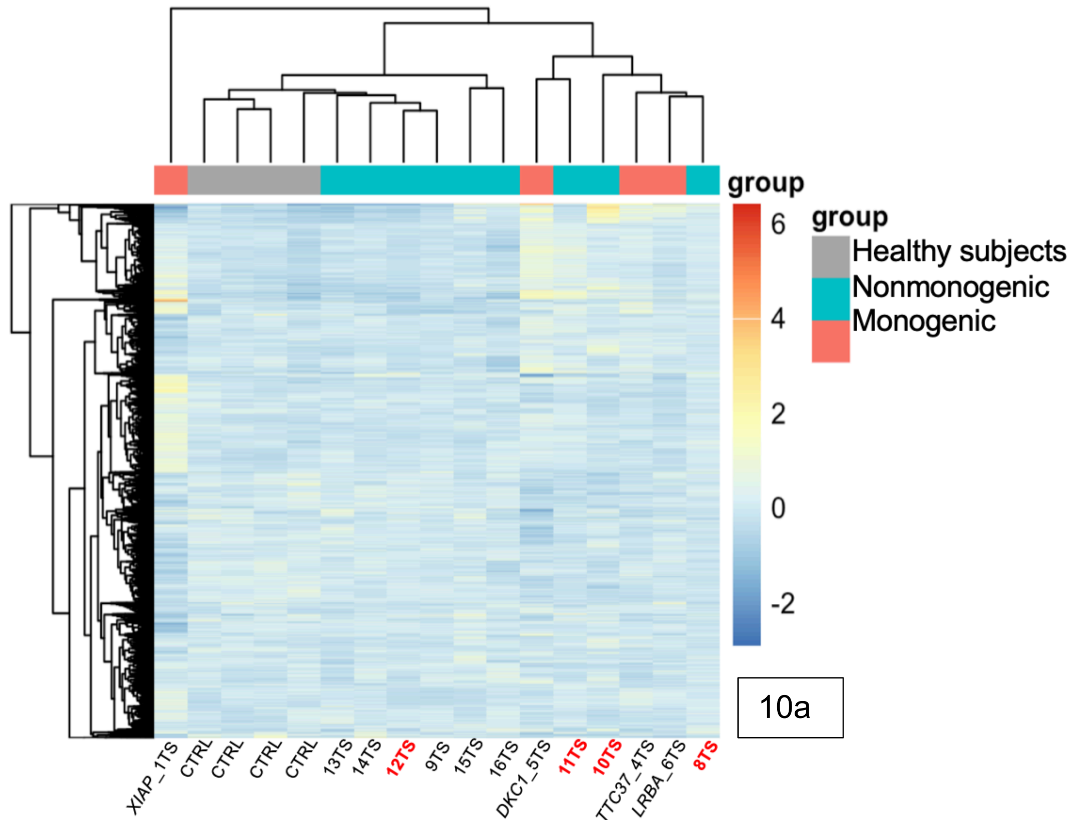


Figure 10a. Protein-coding genes unsupervised hierarchical clustering: healthy subjects in grey, monogenic patients in red, and nonmonogenic patients in turquoise.
Labelled in red (12TS, 11TS, 10TS, 8TS) patients manifested extraintestinal features and without a genetic diagnosis. The gradient from red to blue represents log2FoldChange gene expression values from the highest to the lowest values.

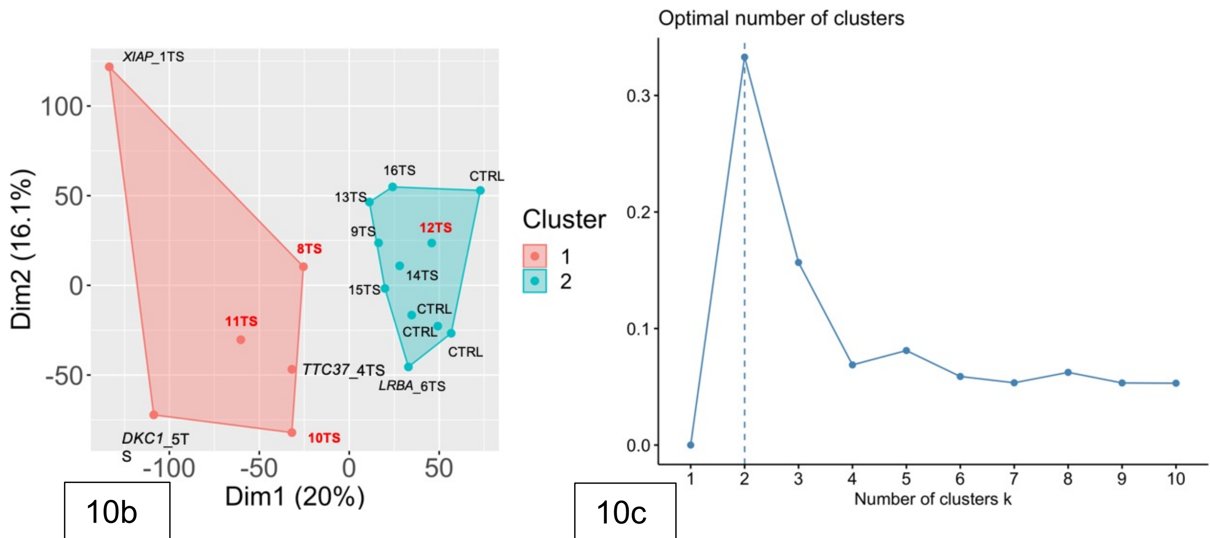


Figure 10b. Protein-coding genes clustering (K-means clustering). Patients' subgroups according to their similarities in protein-coding gene expression. This algorithm analyzed rlog transformed data from DESeq2 R package. Dim1 and Dim2 show the highest differences

between the main clusters; Cluster 1 in red: all monogenic patients except LRBA_6TS and all nonmonogenic patients manifested extraintestinal features except 12TS, Cluster 2 in turquoise: LRBA_6TS, nonmonogenic patients and healthy subjects.
Labelled in red patients without a genetic diagnosis who manifested extraintestinal features. Each dot represents one patient.
Figure 10c. The optimal number of clusters was determined by running silhouette method. Higher the average silhouette width, the better the number of clusters.

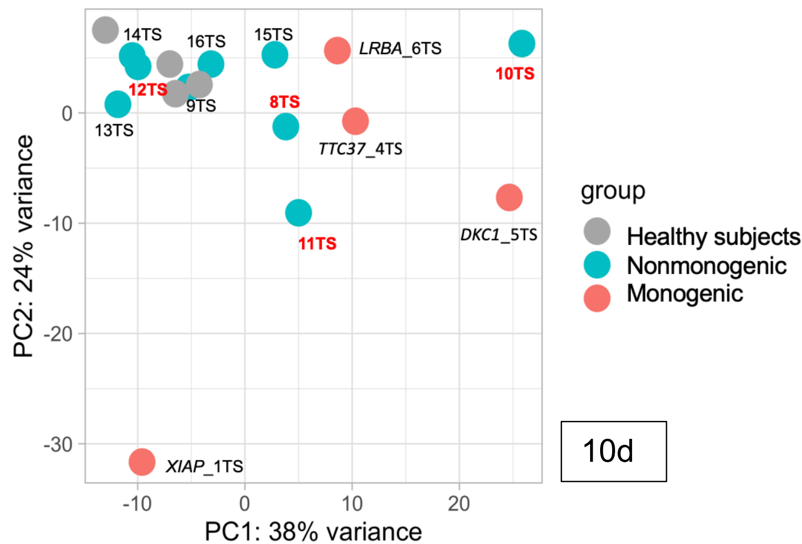


Figure 10d. Principal component analysis: in grey healthy subjects, in red monogenic patients and in turquoise nonmonogenic patients. PC1 and PC2 show the highest differences among all samples. Each dot represents a patient. Labelled in red patients without a genetic diagnosis who manifested extraintestinal features.

Clustering divided samples into 2 main groups. The first one is composed of 3 out of 4 monogenic IBD (XIAP_1TS, TTC37_4TS, and DKC1_5T) and, interestingly, 3 out of 4 nonmonogenic IBD that manifested extraintestinal features. However, as clearly highlighted by hierarchical clustering, the patient with XIAP deficiency (1TS) showed a diverse pattern compared to the other patients that formed the same cluster. The second cluster enclosed patient with LRBA deficiency (6TS), the remaining non-monogenic IBD, and all the four controls.

The presence of patients without any proven monogenic disease and with extraintestinal manifestation is highlighted in these analyses, since it is a feature highly supportive of a genetic defect. These patients, indeed, might underlie a disorder in which the genetic weight is higher than in other likely multifactorial IBD.

Another possible hypothesis is that subjects with extraintestinal features are more likely to have systemic features influencing peripheral blood expression profile.

Overall VEOIBD group presented distinct features than healthy individuals. However, differential gene expression and pathway enriched analyses didn't show any statistical significance.

4.3.4 DIFFERENTIAL GENE EXPRESSION ANALYSIS (monogenic vs nonmonogenic IBD)

The identification of distinct functional group within VEOIBD patients might be a useful way for patients' stratification and to support tailored therapies.

PCA was performed on VEOIBD samples to investigate possible changes in the overall variation of PC1 and PC2 (Figure 11), compared to the previous assessment in which the control group was included (Figure 10d). The results obtained were comparable.

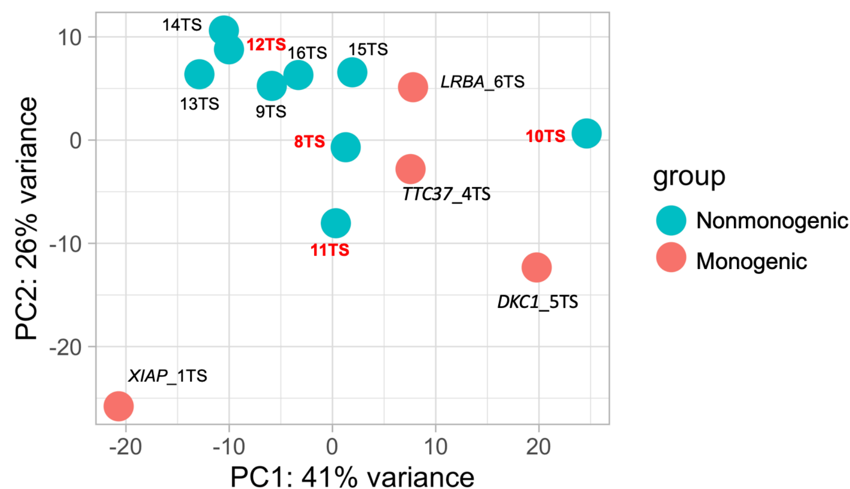


Figure 11.
Principal component analysis protein-coding genes: monogenic patients (red dots) and nonmonogenic patients (turquoise dots). PC1 and PC2 show the highest differences among all samples. Each dot represents a sample. Labelled in red patients without a genetic diagnosis who manifested extraintestinal features.

Monogenic and nonmonogenic IBD samples underwent to differential gene expression analysis. Figure 12 displays the significative (adjusted p-value < 0.05) differentially expressed genes (DEGs) between the two groups. The main pathway enriched is interleukin pathway, particularly IL10 signaling, according to Reactome database. The increased secretion of cytokines and chemokines may be the consequence of pathogens signaling due to an intestinal permeability alteration. A speculation might be that the inflammatory process, which arises in intestinal cell, is maintained also in peripheral blood as reflection of genetic alterations.

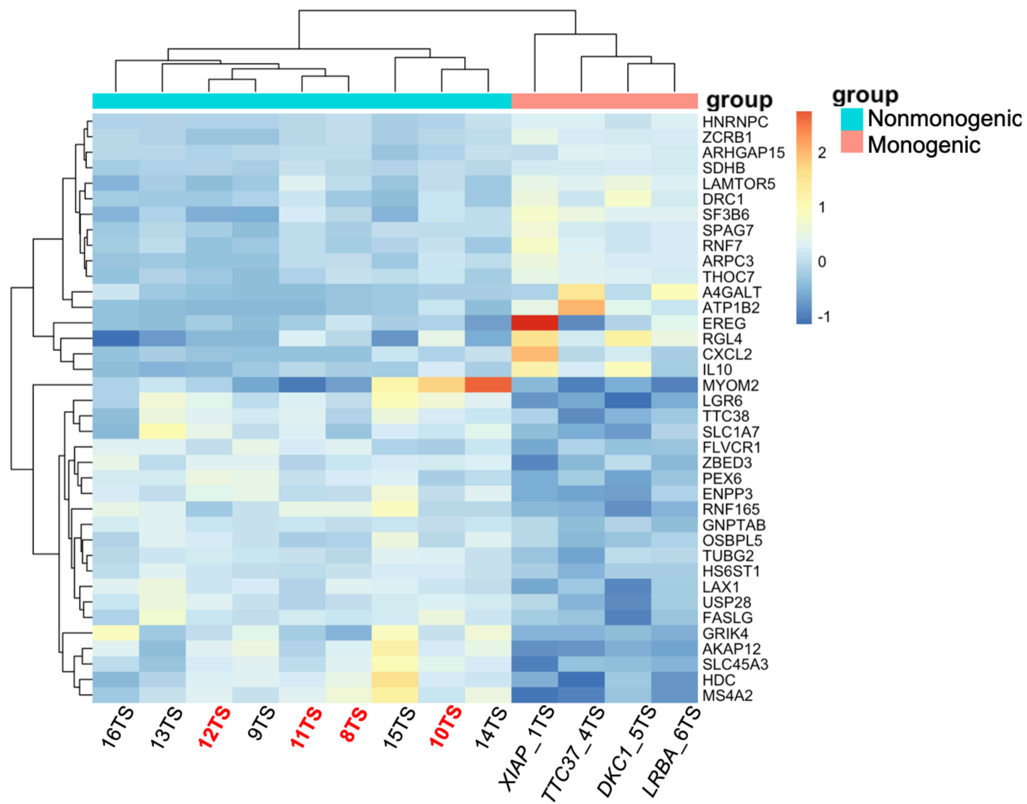


Figure 12.

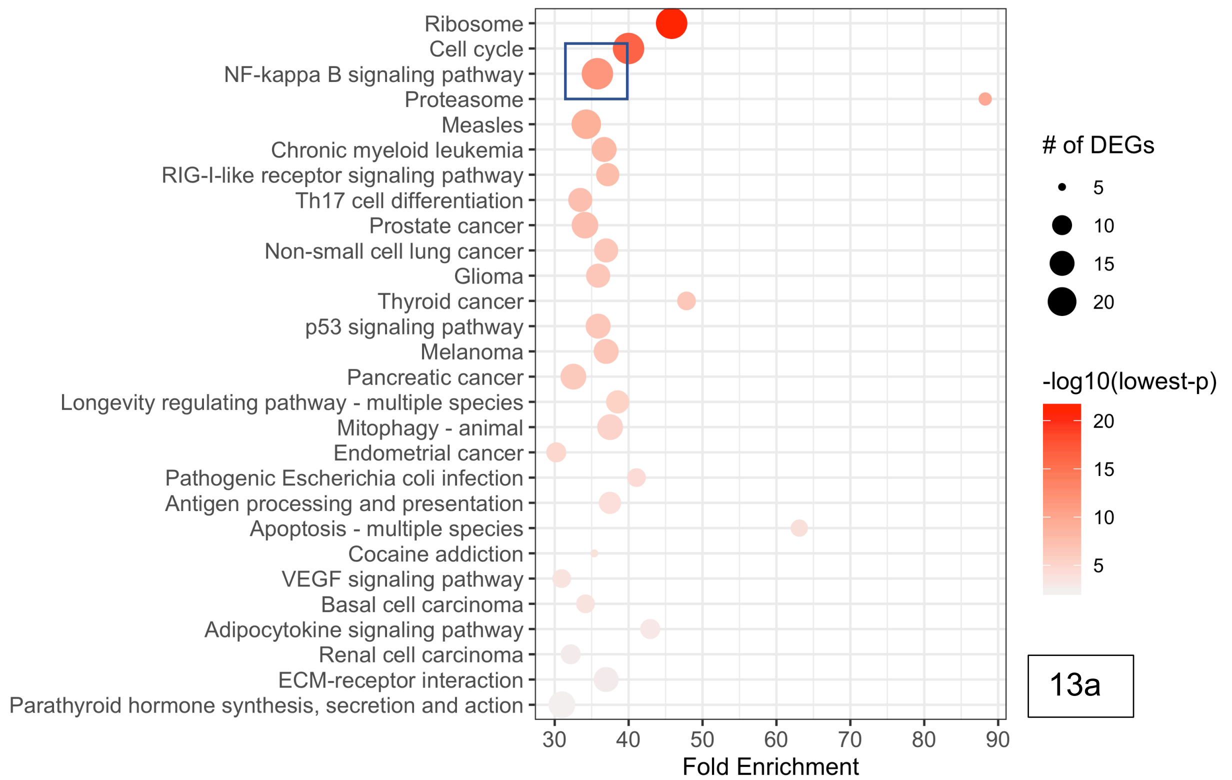
DEGs between monogenic and nonmonogenic patients. DEGs have been selected by fold change greater than 2-fold increase/decrease and adjusted p-value < 0.05. Labelled in red patients without a genetic diagnosis who manifested extraintestinal features. The gradient from red to blue represents log2FoldChange gene expression values from the highest to the lowest values (range: 3, -1).

4.3.5 PATHWAY ENRICHMENT ANALYSIS of DEGs in MONOGENIC VEOIBD

DEGs monogenic vs nonmonogenic VEOIBD output is swayed mostly by higher gene expression rates only in two monogenic patients (XIAP_1TS and DKC1_5TS) mediated with the other values. For this reason, a separate gene expression analysis of single monogenic IBD, in comparison to the control group (XIAP_1TS vs healthy subjects, TTC37_4TS vs healthy subjects, DKC1_5TS vs healthy subjects, LRBA_6TS vs healthy subjects) could give a more realistic picture (data not shown). A drawback of this analysis is the low statistical significance in comparing a single non-replicated sample with a group of control. This limitation might be unlikely to be solved when dealing with rare diseases. However, this method may be used to have a general overview about gene up or down-regulation under specific condition compared to healthy individuals. The exact fold change values should be considered with proper care.

Results and Discussion

The most noteworthy results, came from the parsing of XIAP (1TS) and DKC1 (5TS) defects both in terms of DEGs and enriched pathways (Figure 13 a,b).



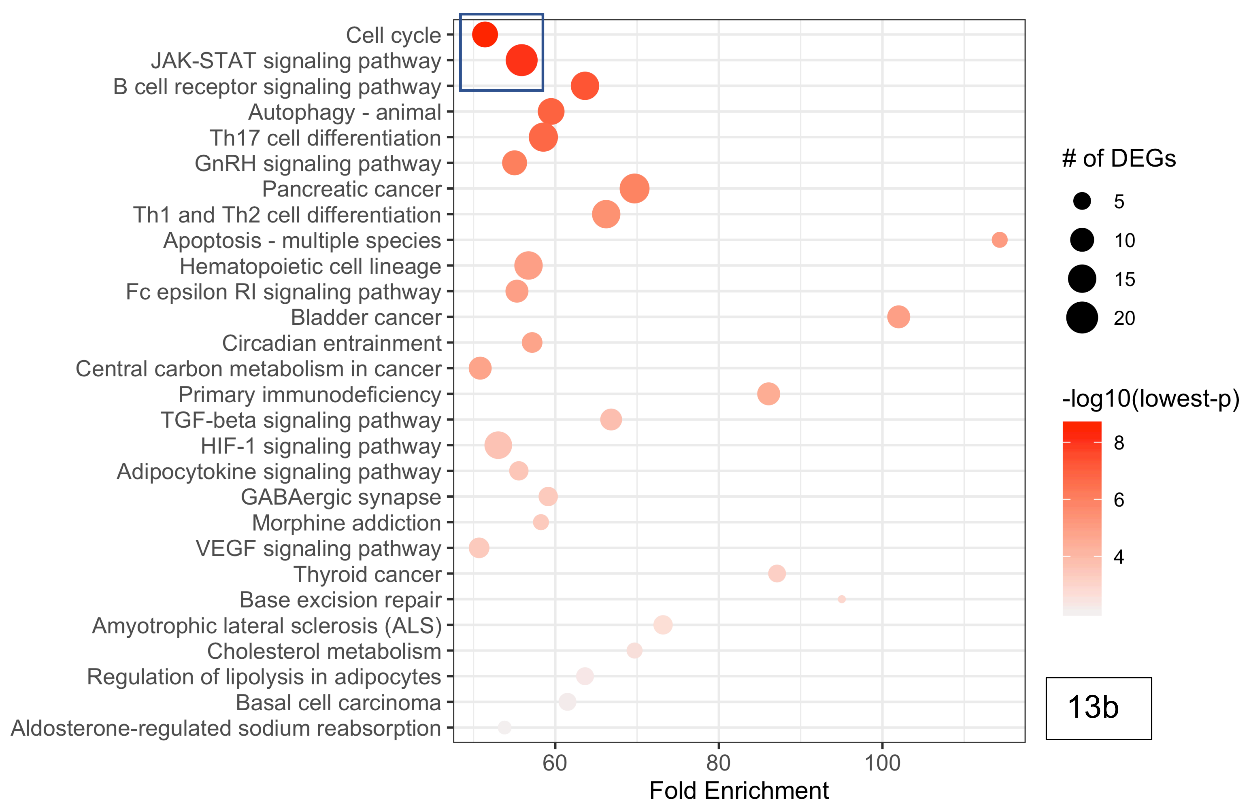


Figure 13.

Figure 13a. Bubble plot. Pathway enrichment analysis (KEGG database) of patient with XIAP defect (1TS) of DEGs detected in comparison with healthy subject group.

Figure 13b. Bubble plot Pathway enrichment analysis (KEGG database) of patient with DKC1 defect (5TS) of DEGs detected in comparison with healthy subject group.
The pathway represented are the most significative for p-value and Fold Enrichment. The size of the dot is directly proportional to the numbers of genes enriched; p-value gradient goes from the most to the less significative. The fold Enrichment is calculated considering the number of genes enriched in a specific pathway out the total number of genes present in that pathway. The box highlighted the pathways described in the main text.

XIAP deficiency (1TS) presented, indeed, a unique gene expression profile with an alteration of inflammatory cytokines. One of the most representative diseases-related enriched pathways was NF-κB. These results showed an up-regulation of both target genes (e.g. *IL6*) and negative regulators (*TNFAIP3*) of NF-κB signaling (Figure 13a).
DKC1 defect (5TS) presented an overrepresentation of the cell cycle pathway that may be attributable to the mutation effect on telomeres shortening and JAK-STAT pathway (Figure 13b) in agreement with a slightly positive Interferon Signature (data not shown).
TTC37 defect (4TS) and LRBA deficiency (6TS) didn't display any significant enriched pathway.

TTC37 mutations caused epithelial barrier defect and a translated systemic effect might be milder and more difficult to detect and resulted in a negative outcome of gene enrichment analysis in peripheral blood cells.

A lack of specific gene expression pattern of LRBA disorder could be, possibly explained focusing on LRBA protein function. It plays a role in the intracellular trafficking of CTLA4 avoiding it from lysosomal degradation and bring it back to the cell surface. CTLA4 is constitutively expressed on regulatory T cells. Accordingly, LRBA deficiency may have a higher impact on lymphocytes subset which is hidden by the overwhelming neutrophil population in peripheral blood.

The characterization of distinct profile of monogenic IBD can be, indeed, a useful tool to perform a disease-similarity analysis to group patient with common features.

4.3.6 SUBGROUPS IDENTIFICATION WITHIN VEOIBD

K-means clustering helped VEOIBD functional subgroups identifications. Hierarchical clustering reported similar features (data not shown).

Figure 14 displays the different clusters and, conversely from the previous analysis (Figure 10b) XIAP deficiency (1TS) standalone. The other two clusters remained almost unchanged. The within-group differences may be emphasized by the absence of the control group.

K-means clustering ran a second time within the selected 400 genes described by Kelsen et al [120], given the large involvement of monogenic IBD causative genes in primary immunodeficiency and its related pathways. The clustering output was the same, hypothesizing a possible role of the immune system in bringing group diversity.

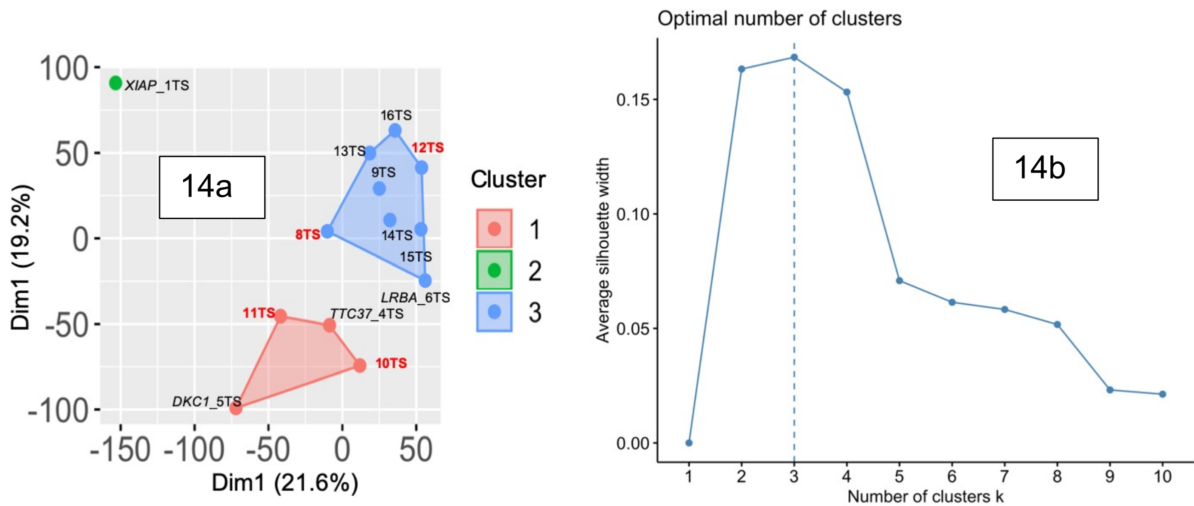


Figure 14.

Figure 14a. Cluster analysis results (K-means clustering) considering the expression profile of protein coding genes that divided into subgroups by gene expression similarities. This algorithm analyzed rlog transformed data from DESeq2 R package. Dim1 and Dim2 show the higher differences between the main clusters; Cluster 1 grouped 2/4 monogenic patients (DKC1_5TS and TTC37_4TS) and 2/4 nonmonogenic patients manifested extraintestinal features, Cluster 2 is composed only by the patient with XIAP deficiency (1TS) and Cluster 3 enclosed patient with LRBA deficiency (6TS), 2/4 nonmonogenic patients manifested extraintestinal features and nonmonogenic patients. Labelled in red patients manifested extraintestinal features and without a genetic diagnosis. Each dot represents one patient.
Figure 14b. Determination of optimal number of clusters to better group monogenic and nonmonogenic patients according to their protein-coding gene expression by running silhouette method.

Except from XIAP deficiency (1TS) that represented a separate case; which are effectively the main differences between cluster 1 and 3? (Figure 14a).
Firstly, a distinction is observed in the total number of significant DEGs (adjusted p-value < 0.05) calculated for each sample in comparison with healthy subjects.
Cluster 1 (10TS, 11TS, DKC1_5TS, TTC37_4TS) had a higher number of genes deviated from control conditions (Figure 15) than Cluster 2. Moreover, this feature seemed to be characteristic of monogenic disease, in general (Figure 15, Table 7). Patient from Cluster 1 slightly shared two common enriched pathways: Th1 and Th2 differentiation and FoxO signaling. This output is consistent with cytokines signaling revealed from DEGs pathway enrichment between monogenic and nonmonogenic IBD. This result might suggest a prevalence genetic component also in the two nonmonogenic patients with extraintestinal manifestations.

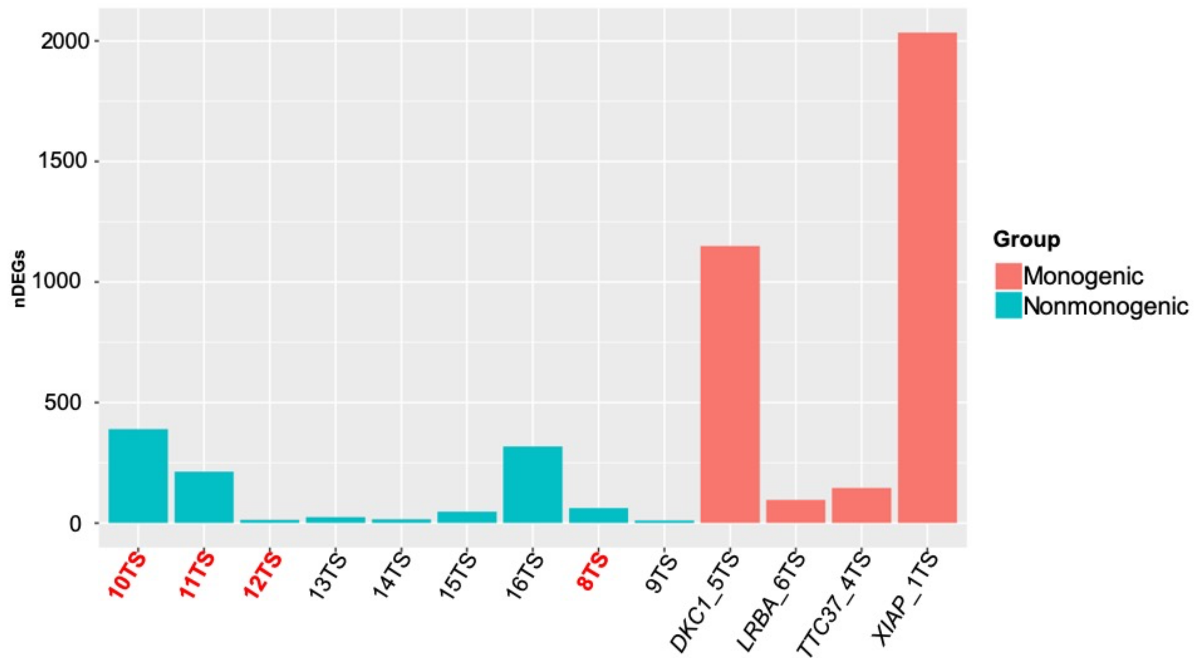


Figure 15.
Significant number of DEGs (adjusted p-value < 0.05 and fold change greater than 2-fold increase or decrease). nDEGs = number of DEGs. Labelled in red patients manifested extraintestinal features and without a genetic diagnosis.

Group	Total DEGs	Up-regulated DEGs (%)	Down-regulated DEGs (%)
Monogenic	856	527 (62)	329 (38)
Nonmonogenic	121	69 (57)	52 (43)

Table 7.
Summary DEGs number in monogenic and nonmonogenic IBD. The table reports the DEGs average of each group.

4.3.7 MACHINE LEARNING APPROACH TO IDENTIFY POSSIBLE CLUSTER-SPECIFIC GENES

The use of machine learning approach is widely applied to high-throughput studies that require effective computational and statistical methods for determining a minimal subset of biomarkers to discriminate two or more phenotypes, starting from larger number of candidate biomarkers.

Given the lack of strong features discriminating different clusters, we addressed the issue to the random forest algorithm by Biosigner (R package) as a possible method to identify cluster-specific genes (Cluster 1 and 3, Figure 14a).

Figure 16 reported the gene distinguishing the two clusters: TNF-related apoptosis inducing ligand, *TNFSF10* that was up-regulated in Cluster 1. This was one of the few genes classified as level S by the algorithm (i.e. corresponds to the final signature) and related to IBD and immune system. This gene is involved in NF-κB pathway and several findings referred its possible role in the pathogenesis of IBD contributing to the disruption of intestinal epithelium integrity by induction of epithelial cells apoptosis [147]. Its up-regulation is documented during intestinal inflammation. In our cohort, the real biological impact of *TNFSF10* up-regulation should be proven by functional assays, preceded by gene expression level confirmation by qPCR.

Unfortunately, our cohort was too small to obtain reliable and highly significant results. Machine learning approach, indeed, is more effective in larger studies.

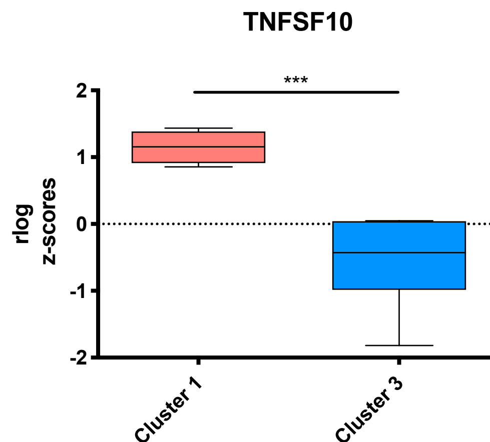


Figure 16.

*Random forest output in the cluster-specific gene identification. Random forest analyzed rlog transformed data from DESeq2 R package. Before statistical analysis, we performed z-score calculation of each value to compare the different population. Statistical analysis was performed by t-test by GraphPad Prism version 8. *** p-value = 0.0007.*

4.3.8 UNSUPERVISED CLUSTERING ANALYSIS TO FIND SIMILAR DISEASES GROUP OF CLINICALLY AND GENETICALLY UNKNOWN VEOIBD PATIENTS

Cluster and machine learning analyses might be applied to group patients by gene expression patterns in an unbiased manner. This method aims to detect undiagnosed patients' subgroups by a disease-similarity analysis in comparison to prototypical monogenic IBD.

RNAseq-Dataset 2 (Brescia) underwent to PCA and k-means clustering to determine gene expression variation compared to VEOIBD patients from RNAseq-Dataset 1 (Trieste) with established clinical and genetic data, and healthy individuals.

PCA plot (Figure 17) showed the presence of some outliers close to monogenic IBD XIAP_1TS and DKC1_5TS: 10TS, 11TS and one genetically undermined patient from RNAseq-Dataset 2.

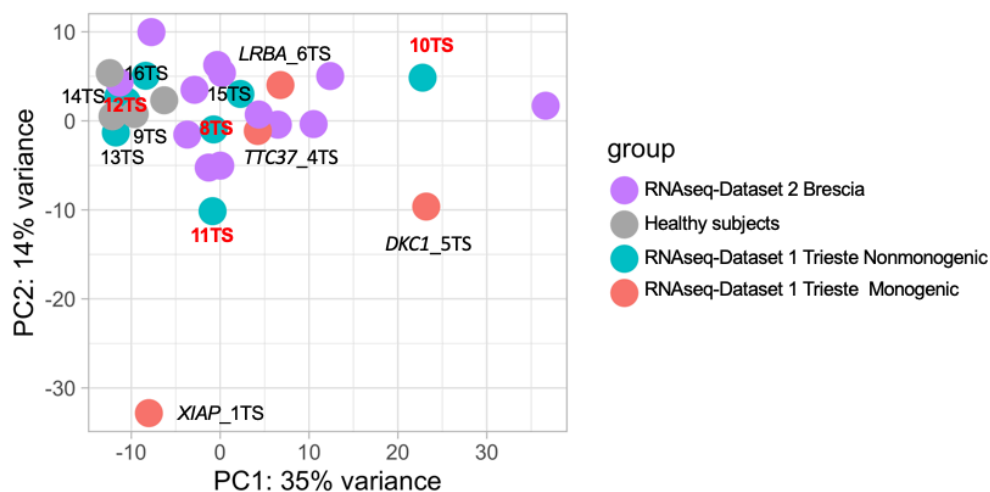


Figure 17.

PCA protein-coding genes: RNAseq-Dataset 1 monogenic patients (red dots), RNAseq-Dataset 1 nonmonogenic patients (turquoise dots), healthy subject (grey dots) and RNAseq-Dataset 2 (violet dots). PC1 and PC2 represent show the highest differences among all samples. Each dot represents a sample. Labelled in red patients manifested extraintestinal features and without a genetic diagnosis.

Clustering identified different subgroups compared to the analysis carried out only on RNAseq-Dataset 1 (Figure 18). Monogenic IBD are not closely distributed. Nevertheless, 3 out of 4 monogenic diseases (XIAP_1TS, DKC1_5TS, and LRBA_6TS defects) are along the edges of their cluster.

However, unsupervised and unbiased performance should be complemented by clinical reports and therapeutic management to get more precise results to the evaluation of the obtained functional subgroups.

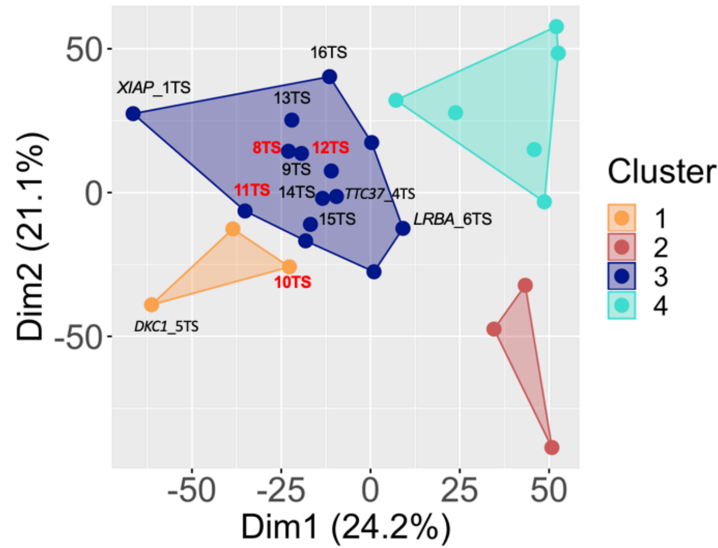


Figure 18.
Protein-coding genes clustering (K-means clustering). Patients' subgroups according to their similarities in protein-coding gene expression. This algorithm analyzed rlog transformed data from DESeq2 R package. Dim1 and Dim2 show the higher differences between the main clusters. Each dot represents one patient, no-labelled dots represent patient from RNAseq-Dataset 2 whose genetic and clinical data are not yet available. Labelled in red patients who manifested extraintestinal features and without a genetic diagnosis.

5 CONCLUSIONS

Our data showed that genetic diagnosis was established in almost 14% of patients with VEOIBD. This result is consistent with previous work that reported a rate of 16% [88, 148].

The “red flags” highly suggestive of a genetic defect are:

- Early age at disease onset
- Coexistence of extraintestinal manifestations
- Male sex

The gender should be considered, since a high number of disease causative genes presented a X-linked recessive inheritance, almost 10 out of 50 genes (20%) summarized by Uhlig et al [31].

Monogenic VEOIBD diagnostic approach changed over time, especially more recently after the advent of NGS techniques. NGS has the advantages to screen simultaneously multiple genes. For this reason, it should be preferred in patients with nonspecific phenotypes, especially in infants in whom the probability of a monogenic condition is higher and for whom timely diagnosis may have an impact on the patient’s management [119]. Nevertheless, Sanger sequencing is still effective in patients showing clinical and immunological findings pointing towards a specific diagnosis.

Considering the monogenic diagnostic rate (about 15%), the majority of patients with VEOIBD remained without causative mutations detection. The reason might be explained by a few factors, such as ineffective sequencing due to low coverage, alteration in the non-coding part of the genome and thus not detectable by WES analysis, or due to the involvement of few/several genes leading to likely multifactorial conditions.

We proposed a disease-similarity method for patients’ stratification and the detection of possible biomarkers. The idea behind is to compare gene expression profiles of undefined/likely multifactorial IBD to those found in distinct monogenic diseases.

To date, expression patterns in peripheral blood cells from monogenic IBD are not well defined, thus the delineation of distinct profiles could help the characterization of molecular mechanisms behind. The understanding of mendelian forms might shed

light on the pathogenesis of nonmonogenic IBD presenting similar pattern to one or more monogenic disorders.

For this type of analysis large cohorts are required, however dealing with rare disorders is not always feasible.

In this study, we obtained preliminary data that will be integrated in the future with an increased number of patients, in the context of the collaboration with University of Brescia.

We compared gene expression profile of 4 monogenic IBD (XIAP (1TS), TTC37 (4TS), DKC1 (5TS), and LRBA (6TS) defects) with nonmonogenic IBD and healthy subjects.

In general, VEOIBD and EOIBD s/a presented a different expression pattern compared to healthy individuals, even though differences weren't that stark.

A subgroup division was detectable also within VEOIBD.

The main differences were observable between monogenic and nonmonogenic IBD, mostly in cytokine production pathways, supporting a possible systemic effect related to genetic mutations. A distinctive feature of monogenic patients is the presence of a higher number of DEGs than the other subjects. However, nonmonogenic IBD that had extraintestinal manifestations, had characteristics more similar to the monogenic group. Obviously, monogenic forms are not all the same: XIAP and DKC1 defects had a higher weight on this examination and they might be caused the most effects observed.

XIAP deficiency presented a unique pattern compared to all the other samples. Studies conducted in mice proposed that xiap -/- neutrophils may contribute to hyper-inflammation and progression in certain pathologies seen in X-linked lymphoproliferative syndrome 2 patients [149]. This finding could suggest that in our patients the great inflammatory response might be directly related to neutrophils hyper-activation. The differences between the DKC1 defect is that the systemic effect observed might be a reflection of a major triggered intestinal inflammation rather than a direct effect on blood cells. We can hypothesize that the major consequence detected in peripheral blood cells was due to the impact of the mutation on telomerase shortening and senescence.

It does not mean that TTC37 defects and LRBA deficiency had normal gene expression profiles, but their effect could have been mitigated. In these cases, it might have been more appropriate to sample cells functionally affected by the mutations as intestinal cells for TTC37 and lymphocytes for LRBA.

In the context of high-throughput screening, machine learning may be a useful method to detect subgroups and characteristic signatures also in unbiased manner [150]. Our unsupervised analysis included our monogenic IBD, the nonmonogenic IBD from the cohort of Trieste and the first 13 genetically undefined VEOIBD and EOIBD enrolled within the collaborative project with University of Brescia whose clinical collection data and genetic investigations are in progress. This performance showed the presence of some outliers close to monogenic IBD XIAP and DKC1 including one patient from the cohort of Brescia, suggestive of a potential monogenic disease. However, this data should be complemented by clinical reports and therapeutic management at the time of sampling to get more precise results and evaluate the obtained functional subgroups.

The identification of candidate biomarkers is one of the possible applications of this analysis and might give the possibility to support the diagnosis and the treatment.

In the light of these considerations we can conclude that:

- Monogenic IBD present a distinct gene expression profiles compared to nonmonogenic IBD and healthy subjects;
- Monogenic diseases present a higher number of significant DEGs;
- Nonmonogenic IBD that had extraintestinal manifestations present hybrid genetic and phenotypic characteristics, between monogenic and likely multifactorial patients;
- The functional understanding of the underlying mechanism of distinct monogenic defects might be helpful in the selection of appropriate cell type to investigate for gene expression analyses.

Nevertheless, the characterization of more monogenic forms is a crucial point to expand this analysis. The implementation of this knowledge may allow the use of monogenic disorders as prototypical diseases for the stratification and the therapeutic management of likely multifactorial cases towards a tailored therapy.

6 SUPPLEMENTARY MATERIAL

Patient	Cohort	Genetic result	Sanger	TGPS	WES or WGS	RNAseq
1TS	Trieste	monogenic	•			•
2TS	Trieste	monogenic	•		•	
3TS	Trieste	monogenic	•			
4TS	Trieste	monogenic		•		•
5TS	Trieste	monogenic		•		•
6TS	Trieste	monogenic	•	•		•
7TS	Trieste	monogenic	•			
8TS	Trieste	nonmonogenic	•	•		•
9TS	Trieste	nonmonogenic	•	•		•
10TS	Trieste	nonmonogenic	•	•	•	•
11TS	Trieste	nonmonogenic	•	•		•
12TS	Trieste	nonmonogenic	•	•		•
13TS	Trieste	nonmonogenic	•	•		•
14TS	Trieste	nonmonogenic	•	•		•
15TS	Trieste	nonmonogenic	•	•	•	•
16TS	Trieste	nonmonogenic	•		•	•
17TS	Trieste	nonmonogenic	•		•	
18TS	Trieste	nonmonogenic	•	•		
19TS	Trieste	nonmonogenic	•	•		
20TS	Trieste	nonmonogenic	•	•		
21TS	Trieste	nonmonogenic	•		•	
22TS	Trieste	nonmonogenic	•	•		
23TS	Trieste	nonmonogenic	•	•		
24TS	Trieste	nonmonogenic	•	•		
25TS	Trieste	nonmonogenic	•	•		
26TS	Trieste	nonmonogenic	•		•	
27TS	Trieste	nonmonogenic	•	•		
28TS	Trieste	nonmonogenic	•	•		
29TS	Trieste	nonmonogenic	•	•		
30TS	Trieste	nonmonogenic	•		•	
31TS	Trieste	nonmonogenic	•	•		
32TS	Trieste	nonmonogenic	•		•	
33TS	Trieste	nonmonogenic	•		•	
34TS	Trieste	nonmonogenic	•	•		
35TS	Trieste	nonmonogenic	•	•		
36TS	Trieste	nonmonogenic	•	•		
37TS	Trieste	nonmonogenic	•	•		

Patient	Cohort	Genetic result	Sanger	TGPS	WES or WGS	RNAseq
38TS	Trieste	nonmonogenic	•	•		
39TS	Trieste	nonmonogenic	•	•		
40TS	Trieste	nonmonogenic	•	•		
41TS	Trieste	nonmonogenic	•	•		
42TS	Trieste	nonmonogenic	•		•	
43TS	Trieste	nonmonogenic	•	•	•	
44TS	Trieste	nonmonogenic	•	•	•	
45TS	Trieste	nonmonogenic		•		
46TS	Trieste	nonmonogenic		•		
47TS	Trieste	nonmonogenic		•		
48TS	Trieste	nonmonogenic		•		
49TS	Trieste	nonmonogenic		•		
50TS	Trieste	nonmonogenic		•		
51TS	Trieste	nonmonogenic		•		
52TS	Trieste	nonmonogenic		•		
53TS	Trieste	nonmonogenic		•		
54TS	Trieste	nonmonogenic			•	
55TS	Trieste	nonmonogenic		•	•	
56TS	Trieste	nonmonogenic		•	•	
57TS	Trieste	nonmonogenic		•		
58TS	Trieste	nonmonogenic		•	•	
59TS	Trieste	nonmonogenic			•	
60TS	Trieste	nonmonogenic				
61TS	Trieste	nonmonogenic			•	
62TS	Trieste	nonmonogenic		•		
63TS	Trieste	nonmonogenic			•	
64TS	Trieste	nonmonogenic			•	
65TS	Trieste	nonmonogenic		•		
66TS	Trieste	nonmonogenic	•	•		
67TS	Trieste	nonmonogenic	•	•		
68TS	Trieste	nonmonogenic	•	•		
69TS	Trieste	nonmonogenic		•		
70TS	Trieste	nonmonogenic		•		
71TS	Trieste	nonmonogenic		•		
72TS	Trieste	nonmonogenic		•		
1RM	Rome	monogenic	•			
2RM	Rome	monogenic	•			
3RM	Rome	monogenic		•		
4RM	Rome	monogenic	•			

Patient	Cohort	Genetic result	Sanger	TGPS	WES or WGS	RNAseq
5RM	Rome	monogenic	•			
6RM	Rome	monogenic	•			
7RM	Rome	nonmonogenic		•		
8RM	Rome	nonmonogenic		•		
9RM	Rome	nonmonogenic		•		
10RM	Rome	nonmonogenic		•		
11RM	Rome	nonmonogenic		•		
12RM	Rome	nonmonogenic		•		
13RM	Rome	nonmonogenic		•		
14RM	Rome	nonmonogenic		•		
15RM	Rome	nonmonogenic		•		
16RM	Rome	nonmonogenic		•		
17RM	Rome	nonmonogenic		•		
18RM	Rome	nonmonogenic		•		
19RM	Rome	nonmonogenic		•		
20RM	Rome	nonmonogenic		•		
21RM	Rome	nonmonogenic		•		
22RM	Rome	nonmonogenic		•		
-	Brescia	In progress				•
-	Brescia	In progress				•
-	Brescia	In progress				•
-	Brescia	In progress				•
-	Brescia	In progress				•
-	Brescia	In progress				•
-	Brescia	In progress				•
-	Brescia	In progress				•
-	Brescia	In progress				•
-	Brescia	In progress				•
-	Brescia	In progress				•
-	Brescia	In progress				•
-	Brescia	In progress				•
CTRL	-	Healthy subject				•
CTRL	-	Healthy subject				•
CTRL	-	Healthy subject				•
CTRL	-	Healthy subject				•

Table 1s.
Summary of patients and controls analyzed in this study (cohort, genetic results, genetic tests)
Abbreviations: TGPS, target gene panel sequencing; WES, whole exome sequencing; WGS, whole genome sequencing; RNAseq, RNA sequenging.

7 BIBLIOGRAPHY

[1] A. Kornbluth, D.B. Sachar, P.P.C.o.t.A.C.o. Gastroenterology, Ulcerative colitis practice guidelines in adults: American College Of Gastroenterology, Practice Parameters Committee, Am J Gastroenterol, 105 (2010) 501-523; quiz 524.

[2] W. S., The morbid appearance of the
intestine of Miss Banks, Medical Times and Gazzette, 1859.

[3] B.B. Crohn, L. Ginzburg, G.D. Oppenheimer, Regional ileitis: a pathologic and clinical entity. 1932, Mt Sinai J Med, 67 (2000) 263-268.

[4] D.J. Mulder, A.J. Noble, C.J. Justinich, J.M. Duffin, A tale of two diseases: the history of inflammatory bowel disease, J Crohns Colitis, 8 (2014) 341-348.

[5] C. Abraham, J.H. Cho, Inflammatory bowel disease, N Engl J Med, 361 (2009) 2066-2078.

[6] A.B. Price, Overlap in the spectrum of non-specific inflammatory bowel disease--'colitis indeterminate', J Clin Pathol, 31 (1978) 567-577.

[7] A. Pera, P. Bellando, D. Caldera, V. Ponti, M. Astegiano, C. Barletti, E. David, A. Arrigoni, G. Rocca, G. Verme, Colonoscopy in inflammatory bowel disease. Diagnostic accuracy and proposal of an endoscopic score, Gastroenterology, 92 (1987) 181-185.

[8] V. Maunoury, G. Savoye, A. Bourreille, Y. Bouhnik, M. Jarry, S. Sacher-Huvelin, E. Ben Soussan, E. Lerebours, J.P. Galmiche, J.F. Colombel, Value of wireless capsule endoscopy in patients with indeterminate colitis (inflammatory bowel disease type unclassified), Inflamm Bowel Dis, 13 (2007) 152-155.

[9] W.J. Tremaine, Is indeterminate colitis determinable?, Curr Gastroenterol Rep, 14 (2012) 162-165.

[10] T.M. Connelly, A.S. Berg, L. Harris, D. Brinton, S. Deiling, W.A. Koltun, Genetic determinants associated with early age of diagnosis of IBD, Dis Colon Rectum, 58 (2015) 321-327.

[11] I. Cleynen, G. Boucher, L. Jostins, L.P. Schumm, S. Zeissig, T. Ahmad, V. Andersen, J.M. Andrews, V. Annese, S. Brand, S.R. Brant, J.H. Cho, M.J. Daly, M. Dubinsky, R.H. Duerr, L.R. Ferguson, A. Franke, R.B. Gearry, P. Goyette, H. Hakonarson, J. Halfvarson, J.R. Hov, H. Huang, N.A. Kennedy, L. Kupcinskas, I.C. Lawrance, J.C. Lee, J. Satsangi, S. Schreiber, E. Théâtre, A.E. van der Meulen-de Jong, R.K. Weersma, D.C. Wilson, M. Parkes, S. Vermeire, J.D. Rioux, J. Mansfield, M.S. Silverberg, G. Radford-Smith, D.P. McGovern, J.C. Barrett, C.W. Lees, I.I.B.D.G.

Consortium, Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study, Lancet, 387 (2016) 156-167.

[12] D. Knights, K.G. Lassen, R.J. Xavier, Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome, Gut, 62 (2013) 1505-1510.

[13] W.E. Ek, M. D'Amato, J. Halfvarson, The history of genetics in inflammatory bowel disease, Ann Gastroenterol, 27 (2014) 294-303.

[14] T. Ahmad, J. Satsangi, D. McGovern, M. Bunce, D.P. Jewell, Review article: the genetics of inflammatory bowel disease, Aliment Pharmacol Ther, 15 (2001) 731-748.

[15] J.Z. Liu, S. van Sommeren, H. Huang, S.C. Ng, R. Alberts, A. Takahashi, S. Ripke, J.C. Lee, L. Jostins, T. Shah, S. Abedian, J.H. Cheon, J. Cho, N.E. Dayani, L. Franke, Y. Fuyuno, A. Hart, R.C. Juyal, G. Juyal, W.H. Kim, A.P. Morris, H. Poustchi, W.G. Newman, V. Midha, T.R. Orchard, H. Vahedi, A. Sood, J.Y. Sung, R. Malekzadeh, H.J. Westra, K. Yamazaki, S.K. Yang, J.C. Barrett, B.Z. Alizadeh, M. Parkes, T. Bk, M.J. Daly, M. Kubo, C.A. Anderson, R.K. Weersma, I.M.S.G. Consortium, I.I.G. Consortium, Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations, Nat Genet, 47 (2015) 979-986.

[16] R.H. Duerr, K.D. Taylor, S.R. Brant, J.D. Rioux, M.S. Silverberg, M.J. Daly, A.H. Steinhart, C. Abraham, M. Regueiro, A. Griffiths, T. Dassopoulos, A. Bitton, H. Yang, S. Targan, L.W. Datta, E.O. Kistner, L.P. Schumm, A.T. Lee, P.K. Gregersen, M.M. Barmada, J.I. Rotter, D.L. Nicolae, J.H. Cho, A genome-wide association study identifies IL23R as an inflammatory bowel disease gene, Science, 314 (2006) 1461-1463.

[17] L. Jostins, S. Ripke, R.K. Weersma, R.H. Duerr, D.P. McGovern, K.Y. Hui, J.C. Lee, L.P. Schumm, Y. Sharma, C.A. Anderson, J. Essers, M. Mitrovic, K. Ning, I. Cleynen, E. Theatre, S.L. Spain, S. Raychaudhuri, P. Goyette, Z. Wei, C. Abraham, J.P. Achkar, T. Ahmad, L. Amininejad, A.N. Ananthakrishnan, V. Andersen, J.M. Andrews, L. Baidoo, T. Balschun, P.A. Bampton, A. Bitton, G. Boucher, S. Brand, C. Büning, A. Cohain, S. Cichon, M. D'Amato, D. De Jong, K.L. Devaney, M. Dubinsky, C. Edwards, D. Ellinghaus, L.R. Ferguson, D. Franchimont, K. Fransen, R. Gearry, M. Georges, C. Gieger, J. Glas, T. Haritunians, A. Hart, C. Hawkey, M. Hedl, X. Hu, T.H. Karlsen, L. Kupcinskas, S. Kugathasan, A. Latiano, D. Laukens, I.C. Lawrance, C.W. Lees, E. Louis, G. Mahy, J. Mansfield, A.R. Morgan, C. Mowat, W. Newman, O. Palmieri, C.Y. Ponsioen, U. Potocnik, N.J. Prescott, M. Regueiro, J.I. Rotter, R.K.

Russell, J.D. Sanderson, M. Sans, J. Satsangi, S. Schreiber, L.A. Simms, J. Sventoraityte, S.R. Targan, K.D. Taylor, M. Tremelling, H.W. Verspaget, M. De Vos, C. Wijmenga, D.C. Wilson, J. Winkelmann, R.J. Xavier, S. Zeissig, B. Zhang, C.K. Zhang, H. Zhao, M.S. Silverberg, V. Annese, H. Hakonarson, S.R. Brant, G. Radford-Smith, C.G. Mathew, J.D. Rioux, E.E. Schadt, M.J. Daly, A. Franke, M. Parkes, S. Vermeire, J.C. Barrett, J.H. Cho, I.I.G.C. (IIBDGC), Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease, Nature, 491 (2012) 119-124.

[18] G.P. Ramos, K.A. Papadakis, Mechanisms of Disease: Inflammatory Bowel Diseases, Mayo Clin Proc, 94 (2019) 155-165.

[19] D.S. Shouval, P.A. Rufo, The Role of Environmental Factors in the Pathogenesis of Inflammatory Bowel Diseases: A Review, JAMA Pediatr, 171 (2017) 999-1005.

[20] S.C. Ng, H.Y. Shi, N. Hamidi, F.E. Underwood, W. Tang, E.I. Benchimol, R. Panaccione, S. Ghosh, J.C.Y. Wu, F.K.L. Chan, J.J.Y. Sung, G.G. Kaplan, Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies, Lancet, 390 (2018) 2769-2778.

[21] J. Wehkamp, J. Harder, M. Weichenthal, M. Schwab, E. Schäffeler, M. Schlee, K.R. Herrlinger, A. Stallmach, F. Noack, P. Fritz, J.M. Schröder, C.L. Bevins, K. Fellermann, E.F. Stange, NOD2 (CARD15) mutations in Crohn's disease are associated with diminished mucosal alpha-defensin expression, Gut, 53 (2004) 1658-1664.

[22] D.T. Okou, S. Kugathasan, Role of genetics in pediatric inflammatory bowel disease, Inflamm Bowel Dis, 20 (2014) 1878-1884.

[23] H.H. Uhlig, Monogenic diseases associated with intestinal inflammation: implications for the understanding of inflammatory bowel disease, Gut, 62 (2013) 1795-1805.

[24] N. Thapar, N. Shah, A.D. Ramsay, K.J. Lindley, P.J. Milla, Long-term outcome of intractable ulcerating enterocolitis of infancy, J Pediatr Gastroenterol Nutr, 40 (2005) 582-588.

[25] F.M. Ruemmele, M.G. El Khoury, C. Talbotec, C. Maurage, J.F. Mougenot, J. Schmitz, O. Goulet, Characteristics of inflammatory bowel disease with onset during the first year of life, J Pediatr Gastroenterol Nutr, 43 (2006) 603-609.

[26] M.B. Heyman, B.S. Kirschner, B.D. Gold, G. Ferry, R. Baldassano, S.A. Cohen, H.S. Winter, P. Fain, C. King, T. Smith, H.B. El-Serag, Children with early-onset

inflammatory bowel disease (IBD): analysis of a pediatric IBD consortium registry, J Pediatr, 146 (2005) 35-40.
[27] T. Paul, A. Birnbaum, D.K. Pal, N. Pittman, C. Ceballos, N.S. LeLeiko, K. Benkov, Distinct phenotype of early childhood inflammatory bowel disease, J Clin Gastroenterol, 40 (2006) 583-586.
[28] Z. Cannioto, I. Berti, S. Martelossi, I. Bruno, N. Giurici, S. Crovella, A. Ventura, IBD and IBD mimicking enterocolitis in children younger than 2 years of age, Eur J Pediatr, 168 (2009) 149-155.
[29] J. Ruel, D. Ruane, S. Mehandru, C. Gower-Rousseau, J.F. Colombel, IBD across the age spectrum: is it the same disease?, Nat Rev Gastroenterol Hepatol, 11 (2014) 88-98.
[30] J. Van Limbergen, R.K. Russell, H.E. Drummond, M.C. Aldhous, N.K. Round, E.R. Nimmo, L. Smith, P.M. Gillett, P. McGrogan, L.T. Weaver, W.M. Bisset, G. Mahdi, I.D. Arnott, J. Satsangi, D.C. Wilson, Definition of phenotypic characteristics of childhood-onset inflammatory bowel disease, Gastroenterology, 135 (2008) 1114-1122.
[31] H.H. Uhlig, T. Schwerd, S. Koletzko, N. Shah, J. Kammermeier, A. Elkadri, J. Ouahed, D.C. Wilson, S.P. Travis, D. Turner, C. Klein, S.B. Snapper, A.M. Muise, C.i.I.S.G.a. NEOPICS, The diagnostic approach to monogenic very early onset inflammatory bowel disease, Gastroenterology, 147 (2014) 990-1007.e1003.
[32] D.P. McGovern, S. Kugathasan, J.H. Cho, Genetics of Inflammatory Bowel Diseases, Gastroenterology, 149 (2015) 1163-1176.e1162.
[33] A. Levine, A. Griffiths, J. Markowitz, D.C. Wilson, D. Turner, R.K. Russell, J. Fell, F.M. Ruemmele, T. Walters, M. Sherlock, M. Dubinsky, J.S. Hyams, Pediatric modification of the Montreal classification for inflammatory bowel disease: the Paris classification, Inflamm Bowel Dis, 17 (2011) 1314-1321.
[34] S.K. Khangura, N. Kamal, N. Ho, M. Quezado, X. Zhao, B. Marciano, J. Simpson, C. Zerbe, G. Uzel, M.D. Yao, S.S. DeRavin, C. Hadigan, D.B. Kuhns, J.I. Gallin, H.L. Malech, S.M. Holland, T. Heller, Gastrointestinal Features of Chronic Granulomatous Disease Found During Endoscopy, Clin Gastroenterol Hepatol, 14 (2016) 395-402.e395.
[35] M. Quaranta, R. Wilson, E. Gonçalves Serra, S. Pandey, T. Schwerd, K. Gilmour, P. Klenerman, F. Powrie, S. Keshav, S.P.L. Travis, C.A. Anderson, H.H. Uhlig, Consequences of Identifying XIAP Deficiency in an Adult Patient With Inflammatory Bowel Disease, Gastroenterology, 155 (2018) 231-234.

[36] F. Prenzel, H.H. Uhlig, Frequency of indeterminate colitis in children and adults with IBD - a metaanalysis, J Crohns Colitis, 3 (2009) 277-281.

[37] E.O. Glocker, N. Frede, M. Perro, N. Sebire, M. Elawad, N. Shah, B. Grimbacher, Infant colitis--it's in the genes, Lancet, 376 (2010) 1272.

[38] E.O. Glocker, D. Kotlarz, K. Boztug, E.M. Gertz, A.A. Schäffer, F. Noyan, M. Perro, J. Diestelhorst, A. Allroth, D. Murugan, N. Hätscher, D. Pfeifer, K.W. Sykora, M. Sauer, H. Kreipe, M. Lacher, R. Nustede, C. Woellner, U. Baumann, U. Salzer, S. Koletzko, N. Shah, A.W. Segal, A. Sauerbrey, S. Buderus, S.B. Snapper, B. Grimbacher, C. Klein, Inflammatory bowel disease and mutations affecting the interleukin-10 receptor, N Engl J Med, 361 (2009) 2033-2045.

[39] H.H. Uhlig, A.M. Muise, Clinical Genomics in Inflammatory Bowel Disease, Trends Genet, 33 (2017) 629-641.

[40] E.B. Freeman, J. Köglmeier, A.E. Martinez, J.E. Mellerio, L. Haynes, N.J. Sebire, K.J. Lindley, N. Shah, Gastrointestinal complications of epidermolysis bullosa in children, Br J Dermatol, 158 (2008) 1308-1314.

[41] J.S. Kern, C. Herz, E. Haan, D. Moore, S. Nottelmann, T. von Lilien, P. Greiner, A. Schmitt-Graeff, O.G. Opitz, L. Bruckner-Tuderman, C. Has, Chronic colitis due to an epithelial barrier defect: the role of kindlin-1 isoforms, J Pathol, 213 (2007) 462-470.

[42] L.E. Cheng, B. Kanwar, H. Tcheurekdjian, J.P. Grenert, M. Muskat, M.B. Heyman, J.M. McCune, D.W. Wara, Persistent systemic inflammation and atypical enterocolitis in patients with NEMO syndrome, Clin Immunol, 132 (2009) 124-131.

[43] T. Mizukami, M. Obara, R. Nishikomori, T. Kawai, Y. Tahara, N. Sameshima, K. Marutsuka, H. Nakase, N. Kimura, T. Heike, H. Nunoi, Successful treatment with infliximab for inflammatory colitis in a patient with X-linked anhidrotic ectodermal dysplasia with immunodeficiency, J Clin Immunol, 32 (2012) 39-49.

[44] Y. Avitzur, C. Guo, L.A. Mastropaolo, E. Bahrami, H. Chen, Z. Zhao, A. Elkadri, S. Dhillon, R. Murchie, R. Fattouh, H. Huynh, J.L. Walker, P.W. Wales, E. Cutz, Y. Kakuta, J. Dudley, J. Kammermeier, F. Powrie, N. Shah, C. Walz, M. Nathrath, D. Kotlarz, J. Puchaka, J.R. Krieger, T. Racek, T. Kirchner, T.D. Walters, J.H. Brumell, A.M. Griffiths, N. Rezaei, P. Rashtian, M. Najafi, M. Monajemzadeh, S. Pelsue, D.P. McGovern, H.H. Uhlig, E. Schadt, C. Klein, S.B. Snapper, A.M. Muise, Mutations in tetratricopeptide repeat domain 7A result in a severe form of very early onset inflammatory bowel disease, Gastroenterology, 146 (2014) 1028-1039.

[45] D.C. Blaydon, P. Biancheri, W.L. Di, V. Plagnol, R.M. Cabral, M.A. Brooke, D.A. van Heel, F. Ruschendorf, M. Toynbee, A. Walne, E.A. O'Toole, J.E. Martin, K. Lindley, T. Vulliamy, D.J. Abrams, T.T. MacDonald, J.I. Harper, D.P. Kelsell, Inflammatory skin and bowel disease linked to ADAM17 deletion, N Engl J Med, 365 (2011) 1502-1508.

[46] T. Fiskerstrand, N. Arshad, B.I. Haukanes, R.R. Tronstad, K.D. Pham, S. Johansson, B. Håvik, S.L. Tønder, S.E. Levy, D. Brackman, H. Boman, K.H. Biswas, J. Apold, N. Hovdenak, S.S. Visweswariah, P.M. Knappskog, Familial diarrhea syndrome caused by an activating GUCY2C mutation, N Engl J Med, 366 (2012) 1586-1595.

[47] M.G. Schäppi, V.V. Smith, D. Goldblatt, K.J. Lindley, P.J. Milla, Colitis in chronic granulomatous disease, Arch Dis Child, 84 (2001) 147-151.

[48] A. Al-Bousafy, A. Al-Tubuly, E. Dawi, S. Zaroog, I. Schulze, Libyan Boy with Autosomal Recessive Trait (P22-phox Defect) of Chronic Granulomatous Disease, Libyan J Med, 1 (2006) 162-171.

[49] J.D. Matute, A.A. Arias, N.A. Wright, I. Wrobel, C.C. Waterhouse, X.J. Li, C.C. Marchal, N.D. Stull, D.B. Lewis, M. Steele, J.D. Kellner, W. Yu, S.O. Meroueh, W.M. Nauseef, M.C. Dinauer, 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 (2009) 3309-3315.

[50] G. Visser, J.P. Rake, J. Fernandes, P. Labrune, J.V. Leonard, S. Moses, K. Ullrich, G.P. Smit, Neutropenia, neutrophil dysfunction, and inflammatory bowel disease in glycogen storage disease type Ib: results of the European Study on Glycogen Storage Disease type I, J Pediatr, 137 (2000) 187-191.

[51] P. Bégin, N. Patey, P. Mueller, A. Rasquin, A. Sirard, C. Klein, E. Haddad, É. Drouin, F. Le Deist, Inflammatory bowel disease and T cell lymphopenia in G6PC3 deficiency, J Clin Immunol, 33 (2013) 520-525.

[52] I.D. D'Agata, K. Paradis, Z. Chad, Y. Bonny, E. Seidman, Leucocyte adhesion deficiency presenting as a chronic ileocolitis, Gut, 39 (1996) 605-608.

[53] B. Bader-Meunier, B. Florkin, J. Sibilia, C. Acquaviva, E. Hachulla, G. Grateau, O. Richer, C.M. Farber, M. Fischbach, V. Hentgen, P. Jego, C. Laroche, B. Neven, T. Lequerré, A. Mathian, I. Pellier, I. Touitou, D. Rabier, A.M. Prieur, L. Cuisset, P. Quartier, S.S.F.p.I.R.e.I.M.I.e. Pédiatrie), C.C.R.e. Inflammations), Mevalonate kinase deficiency: a survey of 50 patients, Pediatrics, 128 (2011) e152-159.

[54] Q. Zhou, G.S. Lee, J. Brady, S. Datta, M. Katan, A. Sheikh, M.S. Martins, T.D. Bunney, B.H. Santich, S. Moir, D.B. Kuhns, D.A. Long Priel, A. Ombrello, D. Stone, M.J. Ombrello, J. Khan, J.D. Milner, D.L. Kastner, I. Aksentijevich, A hypermorphic missense mutation in PLCG2, encoding phospholipase Cγ2, causes a dominantly inherited autoinflammatory disease with immunodeficiency, Am J Hum Genet, 91 (2012) 713-720.

[55] O. Egritas, B. Dalgic, Infantile colitis as a novel presentation of familial Mediterranean fever responding to colchicine therapy, J Pediatr Gastroenterol Nutr, 53 (2011) 102-105.

[56] C.J. Cardinale, J.R. Kelsen, R.N. Baldassano, H. Hakonarson, Impact of exome sequencing in inflammatory bowel disease, World J Gastroenterol, 19 (2013) 6721-6729.

[57] M. Meeths, M. Entesarian, W. Al-Herz, S.C. Chiang, S.M. Wood, W. Al-Ateeqi, F. Almazan, J.J. Boelens, H. Hasle, M. Ifversen, B. Lund, J.M. van den Berg, B. Gustafsson, H. Hjelmqvist, M. Nordenskjöld, Y.T. Bryceson, J.I. Henter, Spectrum of clinical presentations in familial hemophagocytic lymphohistiocytosis type 5 patients with mutations in STXBP2, Blood, 116 (2010) 2635-2643.

[58] Y. Zeissig, B.S. Petersen, S. Milutinovic, E. Bosse, G. Mayr, K. Peuker, J. Hartwig, A. Keller, M. Kohl, M.W. Laass, S. Billmann-Born, H. Brandau, A.C. Feller, C. Röcken, M. Schrappe, P. Rosenstiel, J.C. Reed, S. Schreiber, A. Franke, S. Zeissig, XIAP variants in male Crohn's disease, Gut, 64 (2015) 66-76.

[59] C. Booth, K.C. Gilmour, P. Veys, A.R. Gennery, M.A. Slatter, H. Chapel, P.T. Heath, C.G. Steward, O. Smith, A. O'Meara, H. Kerrigan, N. Mahlaoui, M. Cavazzana-Calvo, A. Fischer, D. Moshous, S. Blanche, J. Pachlopnik Schmid, J. Pachlopnick-Schmid, S. Latour, G. de Saint-Basile, M. Albert, G. Notheis, N. Rieber, B. Strahm, H. Ritterbusch, A. Lankester, N.G. Hartwig, I. Meyts, A. Plebani, A. Soresina, A. Finocchi, C. Pignata, E. Cirillo, S. Bonanomi, C. Peters, K. Kalwak, S. Pasic, P. Sedlacek, J. Jazbec, H. Kanegane, K.E. Nichols, I.C. Hanson, N. Kapoor, E. Haddad, M. Cowan, S. Choo, J. Smart, P.D. Arkwright, H.B. Gaspar, X-linked lymphoproliferative disease due to SAP/SH2D1A deficiency: a multicenter study on the manifestations, management and outcome of the disease, Blood, 117 (2011) 53-62.

[60] D. Hazzan, S. Seward, H. Stock, S. Zisman, K. Gabriel, N. Harpaz, J.J. Bauer, Crohn's-like colitis, enterocolitis and perianal disease in Hermansky-Pudlak syndrome, Colorectal Dis, 8 (2006) 539-543.

[61] P.D. Anderson, M. Huizing, D.A. Claassen, J. White, W.A. Gahl, Hermansky-Pudlak syndrome type 4 (HPS-4): clinical and molecular characteristics, Hum Genet, 113 (2003) 10-17.

[62] A.J. Mora, D.M. Wolfsohn, The management of gastrointestinal disease in Hermansky-Pudlak syndrome, J Clin Gastroenterol, 45 (2011) 700-702.

[63] N. Takahashi, K. Matsumoto, H. Saito, T. Nanki, N. Miyasaka, T. Kobata, M. Azuma, S.K. Lee, S. Mizutani, T. Morio, Impaired CD4 and CD8 effector function and decreased memory T cell populations in ICOS-deficient patients, J Immunol, 182 (2009) 5515-5527.

[64] S.O. Burns, H.L. Zenner, V. Plagnol, J. Curtis, K. Mok, M. Eisenhut, D. Kumararatne, R. Doffinger, A.J. Thrasher, S. Nejentsev, LRBA gene deletion in a patient presenting with autoimmunity without hypogammaglobulinemia, J Allergy Clin Immunol, 130 (2012) 1428-1432.

[65] E. Salzer, A. Kansu, H. Sic, P. Májek, A. Ikincioğullari, F.E. Dogu, N.K. Prengemann, E. Santos-Valente, W.F. Pickl, I. Bilic, S.A. Ban, Z. Kuloğlu, A.M. Demir, A. Ensari, J. Colinge, M. Rizzi, H. Eibel, K. Boztug, Early-onset inflammatory bowel disease and common variable immunodeficiency-like disease caused by IL-21 deficiency, J Allergy Clin Immunol, 133 (2014) 1651-1659.e1612.

[66] S. Agarwal, L. Mayer, Pathogenesis and treatment of gastrointestinal disease in antibody deficiency syndromes, J Allergy Clin Immunol, 124 (2009) 658-664.

[67] M.E. Conley, A.K. Dobbs, A.M. Quintana, A. Bosompem, Y.D. Wang, E. Coustan-Smith, A.M. Smith, E.E. Perez, P.J. Murray, Agammaglobulinemia and absent B lineage cells in a patient lacking the p85α subunit of PI3K, J Exp Med, 209 (2012) 463-470.

[68] J. Levy, T. Espanol-Boren, C. Thomas, A. Fischer, P. Tovo, P. Bordigoni, I. Resnick, A. Fasth, M. Baer, L. Gomez, E.A. Sanders, M.D. Tabone, D. Plantaz, A. Etzioni, V. Monafo, M. Abinun, L. Hammarstrom, T. Abrahamsen, A. Jones, A. Finn, T. Klemola, E. DeVries, O. Sanal, M.C. Peitsch, L.D. Notarangelo, Clinical spectrum of X-linked hyper-IgM syndrome, J Pediatr, 131 (1997) 47-54.

[69] P. Quartier, J. Bustamante, O. Sanal, A. Plebani, M. Debré, A. Deville, J. Litzman, J. Levy, J.P. Fermand, P. Lane, G. Horneff, G. Aksu, I. Yalçin, G. Davies, I. Tezcan, F. Ersoy, N. Catalan, K. Imai, A. Fischer, A. Durandy, Clinical, immunologic and genetic analysis of 29 patients with autosomal recessive hyper-IgM syndrome due to Activation-Induced Cytidine Deaminase deficiency, Clin Immunol, 110 (2004) 22-29.

[70] M. Catucci, M.C. Castiello, F. Pala, M. Bosticardo, A. Villa, Autoimmunity in wiskott-Aldrich syndrome: an unsolved enigma, Front Immunol, 3 (2012) 209.

[71] J. Rohr, U. Pannicke, M. Döring, A. Schmitt-Graeff, E. Wiech, A. Busch, C. Speckmann, I. Müller, P. Lang, R. Handgretinger, P. Fisch, K. Schwarz, S. Ehl, Chronic inflammatory bowel disease as key manifestation of atypical ARTEMIS deficiency, J Clin Immunol, 30 (2010) 314-320.

[72] Parry DE, et al., A 3-year-old boy with ZAP-70 deficiency, thrombocytopenia and ulcerative colitis, Journal of Allergy and Clinical Immunology, 1996, pp. 390.

[73] J. Chou, R. Hanna-Wakim, I. Tirosh, J. Kane, D. Fraulino, Y.N. Lee, S. Ghanem, I. Mahfouz, A. Mégarbané, G. Lefranc, A. Inati, G. Dbaibo, S. Giliani, L.D. Notarangelo, R.S. Geha, M.J. Massaad, A novel homozygous mutation in recombination activating gene 2 in 2 relatives with different clinical phenotypes: Omenn syndrome and hyper-IgM syndrome, J Allergy Clin Immunol, 130 (2012) 1414-1416.

[74] G. de Saint-Basile, F. Le Deist, M. Caniglia, Y. Lebranchu, C. Griscelli, A. Fischer, Genetic study of a new X-linked recessive immunodeficiency syndrome, J Clin Invest, 89 (1992) 861-866.

[75] K. Felgentreff, R. Perez-Becker, C. Speckmann, K. Schwarz, K. Kalwak, G. Markelj, T. Avcin, W. Qasim, E.G. Davies, T. Niehues, S. Ehl, Clinical and immunological manifestations of patients with atypical severe combined immunodeficiency, Clin Immunol, 141 (2011) 73-82.

[76] T.T. Ozgür, G.T. Asal, D. Cetinkaya, D. Orhan, S.S. Kiliç, Y. Usta, H. Ozen, I. Tezcan, Hematopoietic stem cell transplantation in a CD3 gamma-deficient infant with inflammatory bowel disease, Pediatr Transplant, 12 (2008) 910-913.

[77] I. Borggraefe, S. Koletzko, T. Arenz, M. Fuehrer, F. Hoffmann, I. Dokal, T. Vulliamy, V. Weiler, M. Griese, B.H. Belohradsky, T. Lang, Severe variant of x-linked dyskeratosis congenita (Hoyeraal-Hreidarsson Syndrome) causes significant enterocolitis in early infancy, J Pediatr Gastroenterol Nutr, 49 (2009) 359-363.

[78] B.J. Ballew, V. Joseph, S. De, G. Sarek, J.B. Vannier, T. Stracker, K.A. Schrader, T.N. Small, R. O'Reilly, C. Manschreck, M.M. Harlan Fleischut, L. Zhang, J. Sullivan, K. Stratton, M. Yeager, K. Jacobs, N. Giri, B.P. Alter, J. Boland, L. Burdett, K. Offit, S.J. Boulton, S.A. Savage, J.H. Petrini, A recessive founder mutation in regulator of telomere elongation helicase 1, RTEL1, underlies severe immunodeficiency and features of Hoyeraal Hreidarsson syndrome, PLoS Genet, 9 (2013) e1003695.

[79] O. Sanal, H. Jing, T. Ozgur, D. Ayvaz, D.M. Strauss-Albee, S. Ersoy-Evans, I. Tezcan, G. Turkkani, H.F. Matthews, G. Haliloglu, A. Yuce, B. Yalcin, O. Gokoz, K.K. Oguz, H.C. Su, Additional diverse findings expand the clinical presentation of DOCK8 deficiency, J Clin Immunol, 32 (2012) 698-708.

[80] F. Barzaghi, L. Passerini, R. Bacchetta, Immune dysregulation, polyendocrinopathy, enteropathy, x-linked syndrome: a paradigm of immunodeficiency with autoimmunity, Front Immunol, 3 (2012) 211.

[81] A.A. Caudy, S.T. Reddy, T. Chatila, J.P. Atkinson, J.W. Verbsky, CD25 deficiency causes an immune dysregulation, polyendocrinopathy, enteropathy, X-linked-like syndrome, and defective IL-10 expression from CD4 lymphocytes, J Allergy Clin Immunol, 119 (2007) 482-487.

[82] G. Uzel, E.P. Sampaio, M.G. Lawrence, A.P. Hsu, M. Hackett, M.J. Dorsey, R.J. Noel, J.W. Verbsky, A.F. Freeman, E. Janssen, F.A. Bonilla, J. Pechacek, P. Chandrasekaran, S.K. Browne, A. Agharahimi, A.M. Gharib, S.C. Mannurita, J.J. Yim, E. Gambineri, T. Torgerson, D.Q. Tran, J.D. Milner, S.M. Holland, Dominant gain-of-function STAT1 mutations in FOXP3 wild-type immune dysregulation-polyendocrinopathy-enteropathy-X-linked-like syndrome, J Allergy Clin Immunol, 131 (2013) 1611-1623.

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

[84] B. Begue, J. Verdier, F. Rieux-Laucat, O. Goulet, A. Morali, D. Canioni, J.P. Hugot, C. Daussy, V. Verkarre, B. Pigneur, A. Fischer, C. Klein, N. Cerf-Bensussan, F.M. Ruemmele, Defective IL10 signaling defining a subgroup of patients with inflammatory bowel disease, Am J Gastroenterol, 106 (2011) 1544-1555.

[85] K. Stengaard-Pedersen, S. Thiel, M. Gadjeva, M. Møller-Kristensen, R. Sørensen, L.T. Jensen, A.G. Sjøholm, L. Fugger, J.C. Jensenius, Inherited deficiency of mannan-binding lectin-associated serine protease 2, N Engl J Med, 349 (2003) 554-560.

[86] O. Egritas, B. Dalgic, M. Onder, Tricho-hepato-enteric syndrome presenting with mild colitis, Eur J Pediatr, 168 (2009) 933-935.

[87] A. Fabre, B. Charroux, C. Martinez-Vinson, B. Roquelaure, E. Odul, E. Sayar, H. Smith, V. Colomb, N. Andre, J.P. Hugot, O. Goulet, C. Lacoste, J. Sarles, J. Royet, N. Levy, C. Badens, SKIV2L mutations cause syndromic diarrhea, or trichohepatoenteric syndrome, Am J Hum Genet, 90 (2012) 689-692.

[88] J. Kammermeier, S. Drury, C.T. James, R. Dziubak, L. Ocaka, M. Elawad, P. Beales, N. Lench, H.H. Uhlig, C. Bacchelli, N. Shah, Targeted gene panel sequencing in children with very early onset inflammatory bowel disease--evaluation and prospective analysis, J Med Genet, 51 (2014) 748-755.

[89] B.S. Petersen, B. Fredrich, M.P. Hoeppner, D. Ellinghaus, A. Franke, Opportunities and challenges of whole-genome and -exome sequencing, BMC Genet, 18 (2017) 14.

[90] J.C. Taylor, H.C. Martin, S. Lise, J. Broxholme, J.B. Cazier, A. Rimmer, A. Kanapin, G. Lunter, S. Fiddy, C. Allan, A.R. Aricescu, M. Attar, C. Babbs, J. Becq, D. Beeson, C. Bento, P. Bignell, E. Blair, V.J. Buckle, K. Bull, O. Cais, H. Cario, H. Chapel, R.R. Copley, R. Cornall, J. Craft, K. Dahan, E.E. Davenport, C. Dendrou, O. Devuyst, A.L. Fenwick, J. Flint, L. Fugger, R.D. Gilbert, A. Goriely, A. Green, I.H. Greger, R. Grocock, A.V. Gruszczyk, R. Hastings, E. Hatton, D. Higgs, A. Hill, C. Holmes, M. Howard, L. Hughes, P. Humburg, D. Johnson, F. Karpe, Z. Kingsbury, U. Kini, J.C. Knight, J. Krohn, S. Lamble, C. Langman, L. Lonie, J. Luck, D. McCarthy, S.J. McGowan, M.F. McMullin, K.A. Miller, L. Murray, A.H. Németh, M.A. Nesbit, D. Nutt, E. Ormondroyd, A.B. Oturai, A. Pagnamenta, S.Y. Patel, M. Percy, N. Petousi, P. Piazza, S.E. Piret, G. Polanco-Echeverry, N. Popitsch, F. Powrie, C. Pugh, L. Quek, P.A. Robbins, K. Robson, A. Russo, N. Sahgal, P.A. van Schouwenburg, A. Schuh, E. Silverman, A. Simmons, P.S. Sørensen, E. Sweeney, J. Taylor, R.V. Thakker, I. Tomlinson, A. Trebes, S.R. Twigg, H.H. Uhlig, P. Vyas, T. Vyse, S.A. Wall, H. Watkins, M.P. Whyte, L. Witty, B. Wright, C. Yau, D. Buck, S. Humphray, P.J. Ratcliffe, J.I. Bell, A.O. Wilkie, D. Bentley, P. Donnelly, G. McVean, Factors influencing success of clinical genome sequencing across a broad spectrum of disorders, Nat Genet, 47 (2015) 717-726.

[91] A. Belkadi, A. Bolze, Y. Itan, A. Cobat, Q.B. Vincent, A. Antipenko, L. Shang, B. Boisson, J.L. Casanova, L. Abel, Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants, Proc Natl Acad Sci U S A, 112 (2015) 5473-5478.

[92] M. Hegde, A. Santani, R. Mao, A. Ferreira-Gonzalez, K.E. Weck, K.V. Voelkerding, Development and Validation of Clinical Whole-Exome and Whole-Genome Sequencing for Detection of Germline Variants in Inherited Disease, Arch Pathol Lab Med, 141 (2017) 798-805.

[93] B.B. Cummings, J.L. Marshall, T. Tukiainen, M. Lek, S. Donkervoort, A.R. Foley, V. Bolduc, L.B. Waddell, S.A. Sandaradura, G.L. O'Grady, E. Estrella, H.M. Reddy, F. Zhao, B. Weisburd, K.J. Karczewski, A.H. O'Donnell-Luria, D. Birnbaum, A. Sarkozy, Y. Hu, H. Gonorazky, K. Claeys, H. Joshi, A. Bournazos, E.C. Oates, R. Ghaoui, M.R. Davis, N.G. Laing, A. Topf, P.B. Kang, A.H. Beggs, K.N. North, V. Straub, J.J. Dowling, F. Muntoni, N.F. Clarke, S.T. Cooper, C.G. Bönnemann, D.G. MacArthur, G.-T.E. Consortium, Improving genetic diagnosis in Mendelian disease with transcriptome sequencing, Sci Transl Med, 9 (2017).

[94] J.C. Lee, P.A. Lyons, E.F. McKinney, J.M. Sowerby, E.J. Carr, F. Bredin, H.M. Rickman, H. Ratlamwala, A. Hatton, T.F. Rayner, M. Parkes, K.G. Smith, Gene expression profiling of CD8+ T cells predicts prognosis in patients with Crohn disease and ulcerative colitis, J Clin Invest, 121 (2011) 4170-4179.

[95] D. Biasci, J.C. Lee, N.M. Noor, D.R. Pombal, M. Hou, N. Lewis, T. Ahmad, A. Hart, M. Parkes, E.F. McKinney, P.A. Lyons, K.G.C. Smith, A blood-based prognostic biomarker in IBD, Gut, 68 (2019) 1386-1395.

[96] N. Planell, M.C. Masamunt, R.F. Leal, L. Rodríguez, M. Esteller, J.J. Lozano, A. Ramírez, M.L.S. Ayrizono, C.S.R. Coy, I. Alfaro, I. Ordás, S. Visvanathan, E. Ricart, J. Guardiola, J. Panés, A. Salas, Usefulness of Transcriptional Blood Biomarkers as a Non-invasive Surrogate Marker of Mucosal Healing and Endoscopic Response in Ulcerative Colitis, J Crohns Colitis, 11 (2017) 1335-1346.

[97] G. D'Haens, M. Ferrante, S. Vermeire, F. Baert, M. Noman, L. Moortgat, P. Geens, D. Iwens, I. Aerden, G. Van Assche, G. Van Olmen, P. Rutgeerts, Fecal calprotectin is a surrogate marker for endoscopic lesions in inflammatory bowel disease, Inflamm Bowel Dis, 18 (2012) 2218-2224.

[98] R.W. Stidham, P.D. Higgins, Value of mucosal assessment and biomarkers in inflammatory bowel disease, Expert Rev Gastroenterol Hepatol, 4 (2010) 285-291.

[99] M.E. Burczynski, R.L. Peterson, N.C. Twine, K.A. Zuberek, B.J. Brodeur, L. Casciotti, V. Maganti, P.S. Reddy, A. Strahs, F. Immermann, W. Spinelli, U. Schwertschlag, A.M. Slager, M.M. Cotreau, A.J. Dorner, Molecular classification of

Crohn's disease and ulcerative colitis patients using transcriptional profiles in peripheral blood mononuclear cells, J Mol Diagn, 8 (2006) 51-61.
[100] J. Ostrowski, M. Dabrowska, I. Lazowska, A. Paziewska, A. Balabas, A. Kluska, M. Kulecka, J. Karczmarski, F. Ambrozkiewicz, M. Piatkowska, K. Goryca, N. Zeber-Lubecka, J. Kierkus, P. Socha, M. Lodyga, M. Klopocka, B. Iwanczak, K. Bak-Drabik, J. Walkowiak, P. Radwan, U. Grzybowska-Chlebowczyk, B. Korczowski, T. Starzynska, M. Mikula, Redefining the Practical Utility of Blood Transcriptome Biomarkers in Inflammatory Bowel Diseases, J Crohns Colitis, 13 (2019) 626-633.
[101] A. J., Novel genomic biomarkers that differentiate between inflammatory bowel disease and normal patients using peripheral blood specimens, Am J Gastroenterol, 2008.
[102] P. Olivera, S. Danese, N. Jay, G. Natoli, L. Peyrin-Biroulet, Big data in IBD: a look into the future, Nat Rev Gastroenterol Hepatol, 16 (2019) 312-321.
[103] R. Upstill-Goddard, D. Eccles, S. Ennis, S. Rafiq, W. Tapper, J. Fliege, A. Collins, Support Vector Machine classifier for estrogen receptor positive and negative early-onset breast cancer, PLoS One, 8 (2013) e68606.
[104] E. Capriotti, R.B. Altman, A new disease-specific machine learning approach for the prediction of cancer-causing missense variants, Genomics, 98 (2011) 310-317.
[105] A.N. Lima, E.A. Philot, G.H. Trossini, L.P. Scott, V.G. Maltarollo, K.M. Honorio, Use of machine learning approaches for novel drug discovery, Expert Opin Drug Discov, 11 (2016) 225-239.
[106] C. Mathé, M.F. Sagot, T. Schiex, P. Rouzé, Current methods of gene prediction, their strengths and weaknesses, Nucleic Acids Res, 30 (2002) 4103-4117.
[107] E. Mossotto, J.J. Ashton, T. Coelho, R.M. Beattie, B.D. MacArthur, S. Ennis, Classification of Paediatric Inflammatory Bowel Disease using Machine Learning, Sci Rep, 7 (2017) 2427.
[108] J. Kammermeier, G. Lucchini, S.Y. Pai, A. Worth, D. Rampling, P. Amrolia, J. Silva, R. Chiesa, K. Rao, G. Noble-Jamieson, M. Gasparetto, D. Ellershaw, H. Uhlig, N. Sebire, M. Elawad, L. Notarangelo, N. Shah, P. Veys, Stem cell transplantation for tetratricopeptide repeat domain 7A deficiency: long-term follow-up, Blood, 128 (2016) 1306-1308.
[109] W.A. Faubion, E.V. Loftus, W.S. Harmsen, A.R. Zinsmeister, W.J. Sandborn, The natural history of corticosteroid therapy for inflammatory bowel disease: a population-based study, Gastroenterology, 121 (2001) 255-260.

[110] P. Olivera, S. Danese, L. Peyrin-Biroulet, Next generation of small molecules in inflammatory bowel disease, Gut, 66 (2017) 199-209.

[111] S.H. Murch, V.A. Lamkin, M.O. Savage, J.A. Walker-Smith, T.T. MacDonald, Serum concentrations of tumour necrosis factor alpha in childhood chronic inflammatory bowel disease, Gut, 32 (1991) 913-917.

[112] C.P. Braegger, S. Nicholls, S.H. Murch, S. Stephens, T.T. MacDonald, Tumour necrosis factor alpha in stool as a marker of intestinal inflammation, Lancet, 339 (1992) 89-91.

[113] E.J. Breese, C.A. Michie, S.W. Nicholls, S.H. Murch, C.B. Williams, P. Domizio, J.A. Walker-Smith, T.T. MacDonald, Tumor necrosis factor alpha-producing cells in the intestinal mucosa of children with inflammatory bowel disease, Gastroenterology, 106 (1994) 1455-1466.

[114] L. Peyrin-Biroulet, M. Lémann, Review article: remission rates achievable by current therapies for inflammatory bowel disease, Aliment Pharmacol Ther, 33 (2011) 870-879.

[115] N. Auphan, J.A. DiDonato, C. Rosette, A. Helmberg, M. Karin, Immunosuppression by glucocorticoids: inhibition of NF-kappa B activity through induction of I kappa B synthesis, Science, 270 (1995) 286-290.

[116] S. Majumdar, B.B. Aggarwal, Methotrexate suppresses NF-kappaB activation through inhibition of IkappaBalpha phosphorylation and degradation, J Immunol, 167 (2001) 2911-2920.

[117] I. Atreya, R. Atreya, M.F. Neurath, NF-kappaB in inflammatory bowel disease, J Intern Med, 263 (2008) 591-596.

[118] T. Ito, H. Handa, Cereblon and its downstream substrates as molecular targets of immunomodulatory drugs, Int J Hematol, 104 (2016) 293-299.

[119] S. Lega, A. Pin, S. Arrigo, C. Cifaldi, M. Girardelli, A.M. Bianco, M. Malamisura, G. Angelino, S. Faraci, F. Rea, E.F. Romeo, M. Aloi, C. Romano, A. Barabino, S. Martelossi, A. Tommasini, G. Di Matteo, C. Cancrini, P. De Angelis, A. Finocchi, M. Bramuzzo, Diagnostic Approach to Monogenic Inflammatory Bowel Disease in Clinical Practice: A Ten-Year Multicentric Experience, Inflamm Bowel Dis, (2019).

[120] J.R. Kelsen, N. Dawany, C.J. Moran, B.S. Petersen, M. Sarmady, A. Sasson, H. Pauly-Hubbard, A. Martinez, K. Maurer, J. Soong, E. Rappaport, A. Franke, A. Keller, H.S. Winter, P. Mamula, D. Piccoli, D. Artis, G.F. Sonnenberg, M. Daly, K.E. Sullivan, R.N. Baldassano, M. Devoto, Exome sequencing analysis reveals variants in primary

immunodeficiency genes in patients with very early onset inflammatory bowel disease, Gastroenterology, 149 (2015) 1415-1424.

[121] K. Christodoulou, A.E. Wiskin, J. Gibson, W. Tapper, C. Willis, N.A. Afzal, R. Upstill-Goddard, J.W. Holloway, M.A. Simpson, R.M. Beattie, A. Collins, S. Ennis, Next generation exome sequencing of paediatric inflammatory bowel disease patients identifies rare and novel variants in candidate genes, Gut, 62 (2013) 977-984.

[122] X. Chang, K. Wang, wANNOVAR: annotating genetic variants for personal genomes via the web, J Med Genet, 49 (2012) 433-436.

[123] K. Wang, M. Li, H. Hakonarson, ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data, Nucleic Acids Res, 38 (2010) e164.

[124] G.M. Cooper, D.L. Goode, S.B. Ng, A. Sidow, M.J. Bamshad, J. Shendure, D.A. Nickerson, Single-nucleotide evolutionary constraint scores highlight disease-causing mutations, Nat Methods, 7 (2010) 250-251.

[125] S. Richards, N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, W.W. Grody, M. Hegde, E. Lyon, E. Spector, K. Voelkerding, H.L. Rehm, A.L.Q.A. Committee, 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, Genet Med, 17 (2015) 405-424.

[126] Q. Li, K. Wang, InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines, Am J Hum Genet, 100 (2017) 267-280.

[127] M. Kircher, D.M. Witten, P. Jain, B.J. O'Roak, G.M. Cooper, J. Shendure, A general framework for estimating the relative pathogenicity of human genetic variants, Nat Genet, 46 (2014) 310-315.

[128] J.M. Schwarz, C. Rödelsperger, M. Schuelke, D. Seelow, MutationTaster evaluates disease-causing potential of sequence alterations, Nat Methods, 7 (2010) 575-576.

[129] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, S.R. Sunyaev, A method and server for predicting damaging missense mutations, Nat Methods, 7 (2010) 248-249.

[130] P. Kumar, S. Henikoff, P.C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm, Nat Protoc, 4 (2009) 1073-1081.

[131] S. Chun, J.C. Fay, Identification of deleterious mutations within three human genomes, Genome Res, 19 (2009) 1553-1561.

[132] J.T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E.S. Lander, G. Getz, J.P. Mesirov, Integrative genomics viewer, Nat Biotechnol, 29 (2011) 24-26.
[133] H. Thorvaldsdóttir, J.T. Robinson, J.P. Mesirov, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration, Brief Bioinform, 14 (2013) 178-192.
[134] A. Dobin, C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T.R. Gingeras, STAR: ultrafast universal RNA-seq aligner, Bioinformatics, 29 (2013) 15-21.
[135] Y. Liao, G.K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features, Bioinformatics, 30 (2014) 923-930.
[136] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol, 15 (2014) 550.
[137] E. Ulgen, O. Ozisik, O.U. Sezerman, pathfindR: An R Package for Comprehensive Identification of Enriched Pathways in Omics Data Through Active Subnetworks, Front Genet, 10 (2019) 858.
[138] P. Rinaudo, S. Boudah, C. Junot, E.A. Thévenot, biosigner: A New Method for the Discovery of Significant Molecular Signatures from Omics Data, Front Mol Biosci, 3 (2016) 26.
[139] M. Girardelli, S. Arrigo, A. Barabino, C. Loganes, G. Morreale, S. Crovella, A. Tommasini, A.M. Bianco, The diagnostic challenge of very early-onset enterocolitis in an infant with XIAP deficiency, BMC Pediatr, 15 (2015) 208.
[140] I. Brigida, S. Scaramuzza, D. Lazarevic, D. Cittaro, F. Ferrua, L. Leonardelli, M. Alessio, O. Forma, C. Lanzani, G. Viarengo, F. Ciceri, M. Jankovic, F. Pesce, A. Aiuti, M.P. Cicalese, A novel genomic inversion in Wiskott-Aldrich-associated autoinflammation, J Allergy Clin Immunol, 138 (2016) 619-622.e617.
[141] M. Girardelli, C. Loganes, A. Pin, E. Stacul, E. Decleva, D. Vozzi, G. Baj, C. De Giacomo, A. Tommasini, A.M. Bianco, Novel NOD2 Mutation in Early-Onset Inflammatory Bowel Phenotype, Inflamm Bowel Dis, 24 (2018) 1204-1212.
[142] F. Charbit-Henrion, M. Parlato, S. Hanein, R. Duclaux-Loras, J. Nowak, B. Begue, S. Rakotobe, J. Bruneau, C. Fourrage, O. Alibeu, F. Rieux-Laucat, E. Lévy, M.C. Stolzenberg, F. Mazerolles, S. Latour, C. Lenoir, A. Fischer, C. Picard, M. Aloi, J. Amil Dias, M. Ben Hariz, A. Bourrier, C. Breuer, A. Breton, J. Bronski, S. Buderus, M. Cananzi, S. Coopman, C. Crémilleux, A. Dabadie, C. Dumant-Forest, O. Egritas

Gurkan, A. Fabre, M. German Diaz, Y. Gonzalez-Lama, O. Goulet, G. Guariso, N. Gurcan, M. Homan, J.P. Hugot, E. Jeziorski, E. Karanika, A. Lachaux, P. Lewindon, R. Lima, F. Magro, J. Major, G. Malamut, E. Mas, I. Mattyus, L.M. Mearin, J. Melek, V.M. Navas-Lopez, A. Paerregaard, C. Pelatan, B. Pigneur, I. Pinto Pais, J. Rebeuh, C. Romano, N. Siala, C. Strisciuglio, M. Tempia-Caliera, P. Tounian, D. Turner, V. Urbonas, S. Willot, F.M. Ruemmele, N. Cerf-Bensussan, Diagnostic Yield of Next-Generation Sequencing in Very Early-Onset Inflammatory Bowel Diseases: A Multicenter Study, J Crohns Colitis, (2018).

[143] J. Kammermeier, R. Dziubak, M. Pescarin, S. Drury, H. Godwin, K. Reeve, S. Chadokufa, B. Huggett, S. Sider, C. James, N. Acton, E. Cernat, M. Gasparetto, G. Noble-Jamieson, F. Kiparissi, M. Elawad, P.L. Beales, N.J. Sebire, K. Gilmour, H.H. Uhlig, C. Bacchelli, N. Shah, Phenotypic and Genotypic Characterisation of Inflammatory Bowel Disease Presenting Before the Age of 2 years, J Crohns Colitis, 11 (2017) 60-69.

[144] C. Zheng, Y. Huang, W. Hu, J. Shi, Z. Ye, X. Qian, Z. Huang, A. Xue, Y. Wang, J. Lu, Z. Tang, J. Wu, L. Wang, K. Peng, Y. Zhou, S. Miao, H. Sun, Phenotypic Characterization of Very Early-Onset Inflammatory Bowel Disease with Interleukin-10 Signaling Deficiency: Based on a Large Cohort Study, Inflamm Bowel Dis, 25 (2019) 756-766.

[145] K.R. Engelhardt, N. Shah, I. Faizura-Yeop, D.F. Kocacik Uygun, N. Frede, A.M. Muise, E. Shteyer, S. Filiz, R. Chee, M. Elawad, B. Hartmann, P.D. Arkwright, C. Dvorak, C. Klein, J.M. Puck, B. Grimbacher, E.O. Glocker, Clinical outcome in IL-10- and IL-10 receptor-deficient patients with or without hematopoietic stem cell transplantation, J Allergy Clin Immunol, 131 (2013) 825-830.

[146] O.H. Nielsen, E.C. LaCasse, How genetic testing can lead to targeted management of XIAP deficiency-related inflammatory bowel disease, Genet Med, 19 (2017) 133-143.

[147] B. Begue, H. Wajant, J.C. Bambou, L. Dubuquoy, D. Siegmund, J.F. Beaulieu, D. Canioni, D. Berrebi, N. Brousse, P. Desreumaux, J. Schmitz, M.J. Lentze, O. Goulet, N. Cerf-Bensussan, F.M. Ruemmele, Implication of TNF-related apoptosis-inducing ligand in inflammatory intestinal epithelial lesions, Gastroenterology, 130 (2006) 1962-1974.

[148] Y.H. Fang, Y.Y. Luo, J.D. Yu, J.G. Lou, J. Chen, Phenotypic and genotypic characterization of inflammatory bowel disease in children under six years of age in China, World J Gastroenterol, 24 (2018) 1035-1045.

[149] S. Wicki, U. Gurzeler, W. Wei-Lynn Wong, P.J. Jost, D. Bachmann, T. Kaufmann, Loss of XIAP facilitates switch to TNFα-induced necroptosis in mouse neutrophils, Cell Death Dis, 7 (2016) e2422.

[150] G. I., An Introduction to Variable and Feature Selection, Journal of Machine Learning Research, 2003, pp. 1157-1182.