UNIVERSITY OF IOANNINA
SCHOOL OF HEALTH SCIENCES
FACULTY OF MEDICINE

SECTOR OF INTERNAL MEDICINE
DEPARTMENT OF GASTROENTEROLOGY

Study of the genetic polymorphisms of IBD in NW Greece

VASILEIOS E. TSIANOS
MEDICAL DOCTOR - BIOLOGIST
(MD, BSc, MSc)

PhD THESIS

IOANNINA 2019

UNIVERSITY OF IOANNINA
SCHOOL OF HEALTH SCIENCES
FACULTY OF MEDICINE

SECTOR OF INTERNAL MEDICINE
DEPARTMENT OF GASTROENTEROLOGY

Study of the genetic polymorphisms of IBD in NW Greece

VASILEIOS E. TSIANOS
MEDICAL DOCTOR - BIOLOGIST
(MD, BSc, MSc)

PhD THESIS

IOANNINA 2019

«Η έγκριση της διδακτορικής διατριβής από την Ιατρική Σχολή του Πανεπιστημίου Ιωαννίνων δεν υποδηλώνει αποδοχή των γνωμών του συγγραφέα Ν. 5343/32, άρθρο 202, παράγραφος 2 (νομική κατοχύρωση του Ιατρικού Τμήματος)».

Ημερομηνία αίτησης του κ. Τσιάνου Βασιλείου: 20-6-2012

Ημερομηνία ορισμού Τριμελούς Συμβουλευτικής Επιτροπής: 740α/3-7-2012

Μέλη Τριμελούς Συμβουλευτικής Επιτροπής:
Επιβλέπων
Τσιάνος Επαμεινώνδας, Καθηγητής Παθολογίας του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων
Μέλη
Γεωργίου Ιωάννης, Καθηγητής Ιατρικής Γενετικής και Υποβοηθούμενης Αναπαραγωγής της Μαιευτικής-Γυναικολογίας του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων
Φριλίγγος Ευστάθιος, Αναπληρωτής Καθηγητής Βιολογικής Χημείας του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων

Ημερομηνία ορισμού θέματος: 4-9-2012
«Συστηματική μελέτη των γονιδιακών πολυμορφισμών της ΙΦΠΕ στη ΒΔ Ελλάδα»

Ημερομηνία ανασυγκρότησης της Τριμελούς Συμβουλευτικής Επιτροπής: 901α/17-7-2019
Επιβλέπων
Κατσάνος Κωνσταντίνος, Αναπληρωτής Καθηγητής Γαστρεντερολογίας του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων
Μέλη
Γεωργίου Ιωάννης, Καθηγητής Ιατρικής Γενετικής και Υποβοηθούμενης Αναπαραγωγής της Μαιευτικής-Γυναικολογίας του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων
Φριλίγγος Ευστάθιος, Καθηγητής Βιολογικής Χημείας του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων

ΟΡΙΣΜΟΣ ΕΠΤΑΜΕΛΟΥΣ ΕΞΕΤΑΣΤΙΚΗΣ ΕΠΙΤΡΟΠΗΣ 904α/1-10-2019

Γεωργίου Ιωάννης	Καθηγητής Ιατρικής Γενετικής και Υποβοηθούμενης Αναπαραγωγής της Μαιευτικής-Γυναικολογίας, του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων
Χρήστου Λεωνίδας	Καθηγητής Παθολογίας του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων
Χριστοδούλου Δημήτριος	Καθηγητής Γαστρεντερολογίας του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων
Φριλίγγος Ευστάθιος	Καθηγητής Βιολογικής Χημείας του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων
Κατσάνος Κωνσταντίνος	Αναπληρωτής Καθηγητής Γαστρεντερολογίας του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων
Τσιάρα Σταυρούλα	Αναπληρώτρια Καθηγήτρια Παθολογίας του Τμήματος Ιατρικής του Πανεπιστημίου Ιωαννίνων
Τζάλλας Αλέξανδρος	Επίκουρος Καθηγητής, Βιοϊατρικής Τεχνολογίας με έμφαση στην Επεξεργασία και Ανάλυση Βιοϊατρικών Δεδομένων του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του Πανεπιστημίου Ιωαννίνων

Έγκριση Διδακτορικής Διατριβής με βαθμό **«ΑΡΙΣΤΑ»** στις **29-10-2019**

Ιωάννινα 1-11-2019

ΠΡΟΕΔΡΟΣ ΤΟΥ ΤΜΗΜΑΤΟΣ ΙΑΤΡΙΚΗΣ

Άννα Μπατιστάτου

Καθηγήτρια Παθολογικής Ανατομίας

Η Γραμματέας του Τμήματος

ΜΑΡΙΑ ΚΑΠΙΤΟΠΟΥΛΟΥ

ACKNOWLEDGMENTS

I would like to express my great appreciation to Associate Professor Konstantinos Katsanos who introduced me to the multidimensional side of the IBD genetic studies and encouraged me to study this topic, Professor Ioannis Georgiou and Professor Stathis Frillingos who introduced me to the concepts and techniques of genetic studies and basic research.

I am particularly grateful to Alexandros Tzallas and Evangelos Karvounis for their help and for pushing me to delve deeper into all aspects of scientific research, to Charis Kostoulas, Gioula Georgitsi, Kleopatra Gkaralea and Ioanna Mpoumpa for their support in the lab and Spyridoula Koutsouki for her multifaceted help and support.

Also, I would like to thank my mother Eleni, my sisters Zoi, Tina and Christina as well as my Dimitra for their full encouragement, patience and support all these years.

Last, but not least, I would like to thank my father, Professor Emeritus of Internal Medicine Epameinondas Tsianos who has initiated me to medicine, scientific research and thinking and is always there to provide proper guidance both at medical/scientific level and at personal level as well.

V.T.

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω βαθύτατα τον Αναπληρωτή Καθηγητή κ. Κωνσταντίνο Κατσάνο ο οποίος με εισήγαγε στην πολυδιάστατη πλευρά των γονιδιακών μελετών στην ΙΦΠΕ και με ενθάρρυνε να ασχοληθώ με το θέμα αυτό, τον Καθηγητή κ. Ιωάννη Γεωργίου και τον Καθηγητή κ. Ευστάθιο Φριλίγγο οι οποίοι με εισήγαγαν στις έννοιες και τεχνικές των γονιδιακών μελετών και της βασικής έρευνας.

Είμαι ιδιαίτερα ευγνώμων στους Αλέξανδρο Τζάλλα και Ευάγγελο Καρβούνη για την βοήθεια που μου έδωσαν και με ώθησαν να εμβαθύνω σε όλες τις διαστάσεις της επιστημονικής έρευνας, στον Χάρη Κωστούλα και στις Γιούλα Γεωργίτση, Κλεοπάτρα Γκαραλέα και Ιωάννα Μπούμπα για την βοήθειά τους στο εργαστήριο, καθώς και την Σπυριδούλα Κουτσούκη για την πολύπλευρη βοήθεια και στήριξη.

Θα ήθελα επίσης να ευχαριστήσω τη μητέρα μου Ελένη, τις αδερφές μου Ζωή, Τίνα και Χριστίνα καθώς και την Δήμητρά μου για την αμέριστη συμπαράσταση, υπομονή και στήριξη όλα αυτά τα χρόνια.

Τέλος, πάνω από όλους, θα ήθελα να ευχαριστήσω τον πατέρα μου, Ομότιμο Καθηγητή Παθολογίας, κ. Επαμεινώνδα Τσιάνο ο οποίος με μύησε στην ιατρική, την επιστημονική έρευνα και σκέψη και πάντοτε βρίσκεται εκεί για να προσφέρει σωστή καθοδήγηση τόσο σε ιατρικό/επιστημονικό επίπεδο, αλλά και προσωπικό.

V.T.

δύο γὰρ, ἐπιστήμη τε καὶ δόξα,
ὧν τὸ μὲν ἐπίστασθαι ποιέει, τὸ δὲ ἀγνοεῖν

Ἱπποκράτης, Νόμος

Contents

List of figures

List of tables

List of abbreviation

5-ASA	5-aminosalicylic acid
6-MP	6-mercaptopurine
ATG16L1	autophagy related 16 like 1
AZA	azathioprine
CARD15	caspase recruitment domain-containing protein 15
CD	Crohn's Disease
CI	Confidence Interval
DSS	Decision Support System
ECM1	extracellular matrix protein 1
EIMs	extra-intestinal manifestations
GWAS	Genome Wide Association Studies
IBD	Inflammatory Bowel Disease
IBS	Irritable Bowel Syndrome
IFN	interferon
IL	interleukin
MTX	methotrexate
NOD2	Nucleotide Binding Oligomerization Domain Containing 2
NSAIDs	Nonsteroidal Anti-inflammatory Drugs
OR	Odds Ratio
PCR	polymerase chain reaction
RT-PCR	real-time polymerase chain reaction
SNP	Single-Nucleotide Polymorphism
TNF	tumor necrosis factor
UC	Ulcerative Colitis
ΕΚ	ελκώδης κολίτιδα
ΙΦΠΕ	Ιδιοπαθής Φλεγμονώδης Πάθηση του Εντέρου
NC	νόσος Crohn
ΠΓΝΙ	Περιφερειακό Γενικό Νοσοκομείο Ιωαννίνων

A. Introduction and aim of the study

A1. Epidemiology of Inflammatory Bowel Disease

Inflammatory Bowel Disease (IBD) comprises two idiopathic relapsing and remitting disorders of the gastrointestinal tract, Ulcerative Colitis (UC) and Crohn's Disease (CD). Over recent decades, the incidence of IBD has been steadily increasing, in some locations more than others varying by geographic location. Combined, UC and CD, affect more than 5 million people worldwide, with approximately 1.4 million in USA and 3 million in Europe. Both adult and pediatric-onset IBD new cases are reported worldwide especially in previous low-incident regions (Figure 1) such as Africa, Asia, South America and South-eastern Europe[1]. This increasing prevalence of IBD can be partially attributed to longer life expectancy.

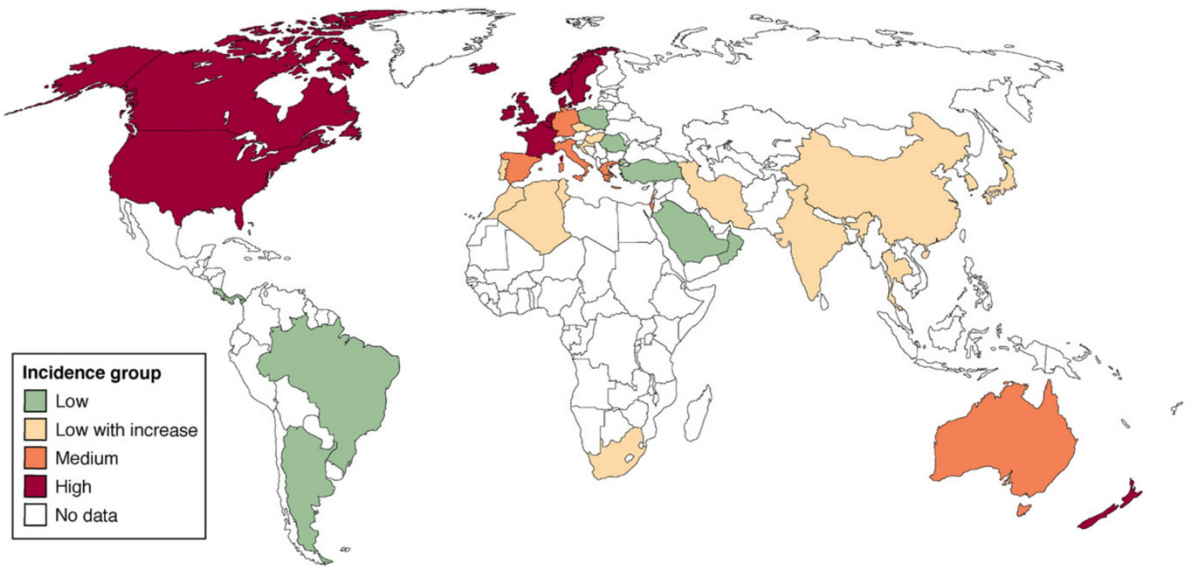
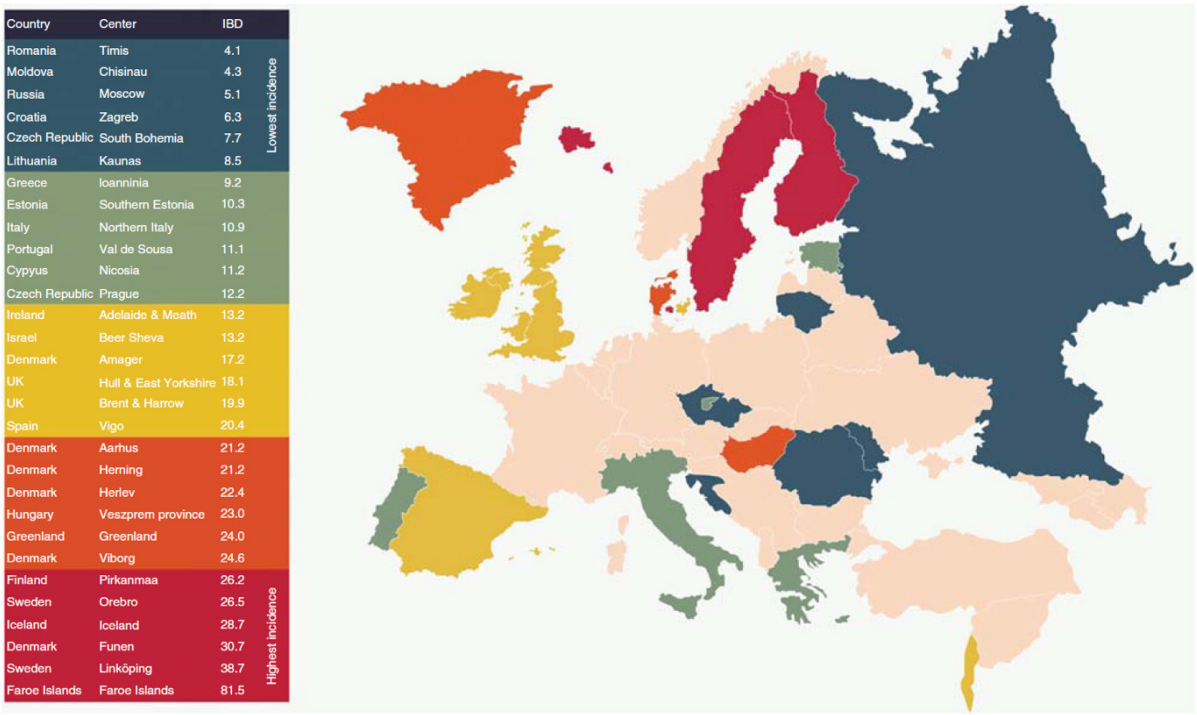


Figure 1 Worldwide incidence of IBD[2]

The highest incidences of IBD have been reported in North America, the United Kingdom and Northern Europe, with the highest incidence in the world being reported in Faroe Islands[2]. Such high incidences may indicate common etiologic factors, in these regions at least. There seems to be a north-to-south gradient with higher incidence rates of both CD and UC in northern locations compared with southern ones. Also, the incidence of UC is greater than that of CD, except in Canada and several areas of Europe, including Northwest Greece[3,4], although this has been changing over the past decades.

In Europe in particular, there is an incidence grade from North to South and East to West, with the highest incidence rate reported in North-West Europe, especially in Faroe Islands as mentioned above[5] (Figure 2).



Country	Center	IBD	
Romania	Timis	4.1	Lowest incidence
Moldova	Chisinau	4.3	
Russia	Moscow	5.1	
Croatia	Zagreb	6.3	
Czech Republic	South Bohemia	7.7	
Lithuania	Kaunas	8.5	
Greece	Ioanninia	9.2	
Estonia	Southern Estonia	10.3	
Italy	Northern Italy	10.9	
Portugal	Val de Sousa	11.1	
Cypyus	Nicosia	11.2	
Czech Republic	Prague	12.2	
Ireland	Adelaide & Meath	13.2	
Israel	Beer Sheva	13.2	
Denmark	Amager	17.2	
UK	Hull & East Yorkshire	18.1	
UK	Brent & Harrow	19.9	
Spain	Vigo	20.4	
Denmark	Aarhus	21.2	
Denmark	Herning	21.2	
Denmark	Herlev	22.4	
Hungary	Veszprem province	23.0	
Greenland	Greenland	24.0	
Denmark	Viborg	24.6	
Finland	Pirkanmaa	26.2	Highest incidence
Sweden	Orebro	26.5	
Iceland	Iceland	28.7	
Denmark	Funen	30.7	
Sweden	Linköping	38.7	
Faroe Islands	Faroe Islands	81.5	

Figure 2 Incidence of IBD in Europe[5]

Moreover, IBD have some distinct demographic features (age of onset, sex, race and ethnicity) attributing to different incidence rates among these subgroups. Although IBD can occur at any age, the age of onset of IBD patients is between 15 and 30 years, with UC patients being diagnosed in their 30s and 40s and CD patients in general between 20 and 30 years of age. In addition, there is a possible second peak between 50 and 80 years of age. It is not clear whether this second peak relates to greater susceptibility to disease with older age, the late expression of an earlier environmental exposure, or higher rates of health care utilization in older persons[1]. Regarding sex, there are small differences among IBD patients, with a slightly higher prevalence in female patients, especially in adult-onset CD, and male predominance in UC. Concerning, race and ethnicity, there are substantial differences in incidence rates, at some subgroups, even among high incidence populations. For example, while New Zealand has one of the highest IBD incidences, native New Zealanders (Aboriginals) have very low. Another well described paradigm is that of

Ashkenazi Jews with a very high incidence, compared to the Sephardic ones. Also, adult Hispanic Americans and Asian Americans seem to have a lower prevalence of CD than non-Hispanic whites. However, these ethnic and racial differences may be related to environmental and lifestyle factors as well as due to underlying genetic differences; yet it is remarkable the fact that when individuals migrate to other areas, they tend to adjust to the local incidence rates[6].

Finally, our area of study, North-West (NW) Greece, has remarkable characteristics compared to other Greek regions as has been previously reported[3]. It used to be a secluded area where population migration was limited which might have played a crucial role in shaping the genetic pool composition. Exceptional in this area is the continuous low incidence of CD compared to UC, despite the phenomenal rise of CD's incidence in the past few years[4].

A2. Clinical presentation of Inflammatory Bowel Disease

Inflammatory Bowel Disease are diseases of the gastrointestinal tract. Despite a substantial overlap, the two disorders, CD and UC, have different pathologic and clinical characteristics (Figure 3).

	Ulcerative colitis	Crohn's disease
Clinical features		
Haematochezia	Common	Rare
Passage of mucus or pus	Common	Rare
Small-bowel disease	No (except backwash ileitis)	Yes
Can affect upper-gastrointestinal tract	No	Yes
Abdominal mass	Rare	Sometimes in right lower quadrant
Extraintestinal manifestations	Common	Common
Small-bowel obstruction	Rarely	Common
Colonic obstruction	Rarely	Common
Fistulas and perianal disease	No	Common
Biochemical features		
Anti-neutrophil cytoplasmic antibodies	Common	Rarely
Anti-saccharomyces cerevisiae antibodies	Rarely	Common
Pathological features		
Transmural mucosal inflammation	No	Yes
Distorted crypt architecture	Yes	Uncommon
Cryptitis and crypt abscesses	Yes	Yes
Granulomas	No	Yes, but rarely in mucosal biopsies
Fissures and skip lesions	Rarely	Common

Figure 3 Differential diagnosis of UC and CD[7]

Ulcerative colitis: Ulcerative colitis is a chronic inflammatory condition characterized by relapsing and remitting episodes of inflammation limited to the mucosal layer of the colon. It usually begins in the rectum, and either remains there or spreads proximally in a continuous fashion.

To describe the degree of large bowel involvement, different terms have been employed. Montreal classification has classified UC depending on the anatomic extent of involvement; thus, patients can be classified as having proctitis, left-sided colitis (involving the sigmoid colon with or without involvement of the descending colon), or

pancolitis. Occasionally, in severe pancolitis, the distal ileum is involved developing ileal inflammation ("backwash ileitis"), which may complicate the differentiation from CD ileocolitis[8].

In detail, Montreal classification categorizes UC patients in three subgroups:

- E1 - Ulcerative proctitis: Involvement limited to the rectum (that is, proximal extent of inflammation is distal to the rectosigmoid junction)
- E2 - Left sided UC (distal UC): Involvement limited to a proportion of the colorectum distal to the splenic flexure
- E3 - Extensive UC (pancolitis): Involvement extends proximal to the splenic flexure

The severity of UC is generally classified endoscopically as mild, moderate, or severe disease, according to Mayo Endoscopic Scoring of Ulcerative Colitis[9] (Figure 4).

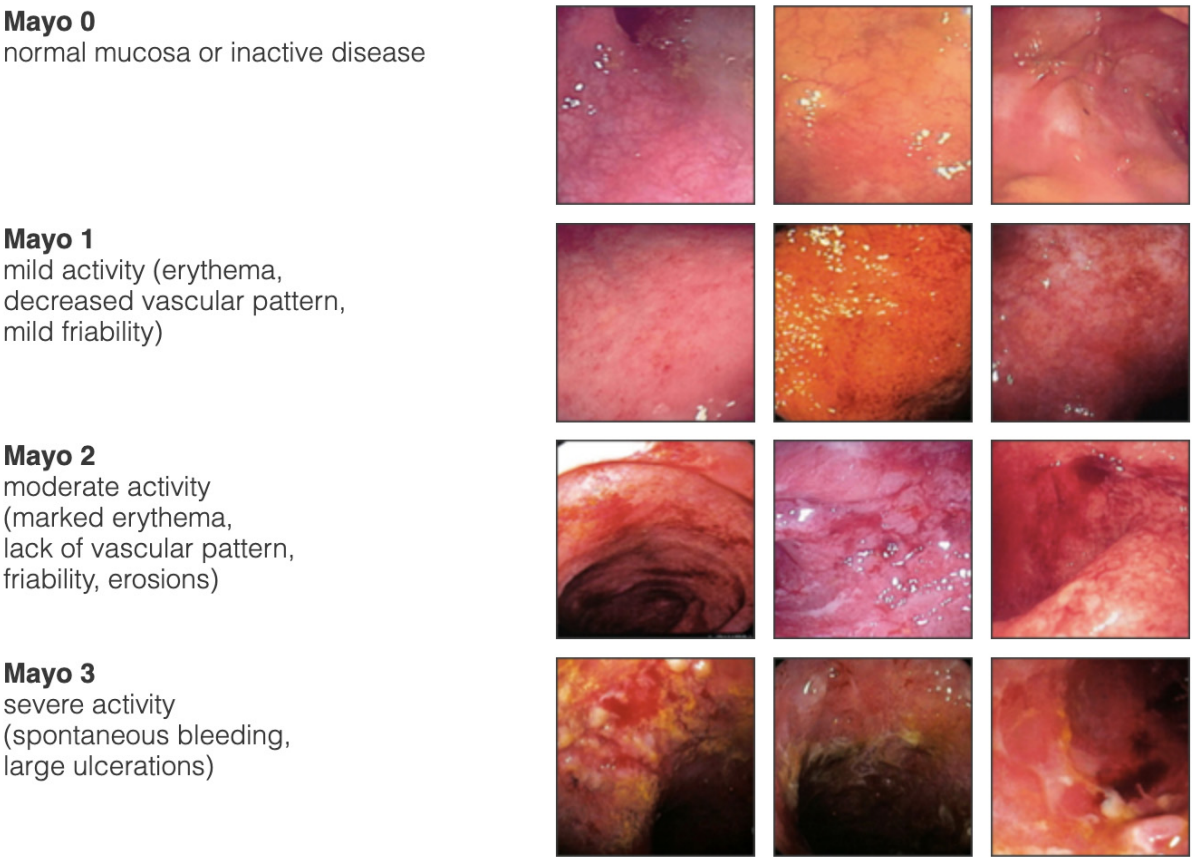


Figure 4 Mayo Endoscopic Scoring of Ulcerative Colitis[9]

The most common symptoms of ulcerative colitis are abdominal pain and cramping and frequent diarrhea, often with blood, pus, or mucus in the stool. Other signs and symptoms include nausea, loss of appetite, fatigue, and fevers. Chronic bleeding from the inflamed and ulcerated intestinal tissue can cause anaemia[10].

Crohn's disease: Crohn's disease can affect any part of the gastrointestinal tract. Is characterized by transmural inflammation and by segments of normal-appearing bowel interrupted by areas of disease (skip lesions). The transmural inflammatory nature of CD may lead to fibrosis and strictures and to obstructive clinical presentations that are not typically seen in patients with UC. Transmural inflammation may also result in sinus tracts, giving rise to micro-perforations and fistula formation. Crohn's disease most commonly involves the ileum and proximal colon; however, any part of the gastrointestinal tract may be affected[10].
Montreal classification has classified CD depending on three different aspects; age at diagnosis, location and behaviour[8]:

- Age at diagnosis:
  - A1 below 16 y
  - A2 between 17 and 40 y
  - A3 above 40 y
- Location:
  - L1 ileal
  - L2 colonic
  - L3 ileocolonic
  - L4 isolated upper disease (is a modifier that can be added to L1–L3 when concomitant upper gastrointestinal disease is present)
- Behavior:
  - B1 non-stricturing, non-penetrating
  - B2 stricturing
  - B3 penetrating
  - p perianal disease modifier (is added to B1–B3 when concomitant perianal disease is present)

The clinical presentation of CD patients may be subtle and varies considerably. Is largely dependent on the location of the disease, the intensity of the inflammation, and presence of specific intestinal and extraintestinal complications and includes diarrhea, abdominal pain, fever, clinical signs of bowel obstruction, as well as passage of blood or mucus or both[7].

Moreover, both UC and CD can present with or develop during disease course many systematic complications, called extra-intestinal manifestations (EIMs), involving the skin, joints, eyes, kidneys or liver/biliary tree[11] (Figure 5 and 6). Their occurrence can differ among areas; considering NW Greece, EIMs are not rare especially in CD patients[12].

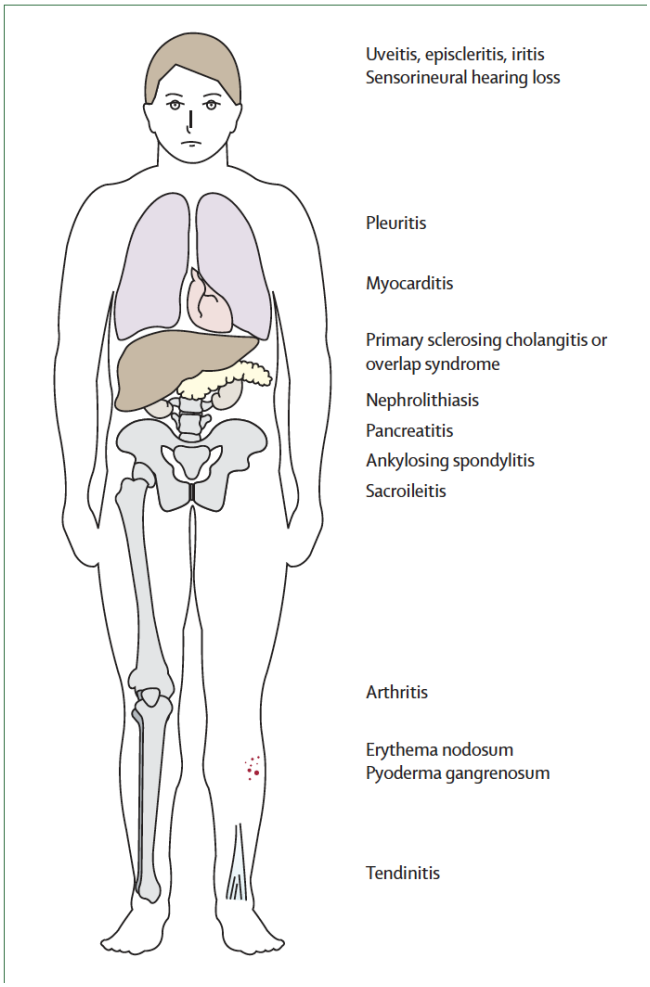


Figure 5 Extra-intestinal manifestations (EIMs) of IBD[7]

Sites	Extraintestinal manifestations
Musculoskeletal system	• Arthritis: colitic type, ankylosing spondylitis, isolated joint involvement • Hypertrophic osteoarthropathy: clubbing, periostitis • Miscellaneous manifestations: osteoporosis, aseptic necrosis, polymyositis
Dermatologic and oral systems	• Reactive lesions: erythema nodosum, pyoderma gangrenosum, aphthous ulcers, necrotizingvasculitis • Specific lesions: fissures, fistulas, oral Crohn's disease, drug rashes • Nutritional deficiencies: acrodermatitis enteropathica, purpura, glossitis, hair loss, brittle nails • Associated diseases: vitiligo, psoriasis, amyloidosis
Hepatopancreatobiliary system	• Primary sclerosing cholangitis, bile-duct carcinoma • Associated inflammation: autoimmune chronic active hepatitis, pericholangitis, portal fibrosis, cirrhosis, granulomatous disease • Metabolic manifestations: fatty liver, gallstones associated with ileal Crohn's disease
Ocular system	• Uveitis/iritis, episcleritis, scleromalacia, corneal ulcers, retinal vascular disease
Metabolic system	• Growth retardation in children and adolescents, delayed sexual maturation
Renal system	• Calcium oxalate stones

Figure 6 Extra intestinal manifestations of IBD[11]

A3. Etiopathogenesis of Inflammatory Bowel Disease

Although the exact cause of IBD remains unknown, increasing epidemiological and laboratory data suggest that it results from a delicate correlation among four overlapping factors (Figure 7): genetic susceptibility, environmental factors (i.e. diet, smoking, infectious disease), immune response (innate and adaptive) and gut microbiota[13,14]. While both Ulcerative Colitis and Crohn's Disease are classified under the IBD flag, genetic and other predisposing factors may differ significantly between the two.

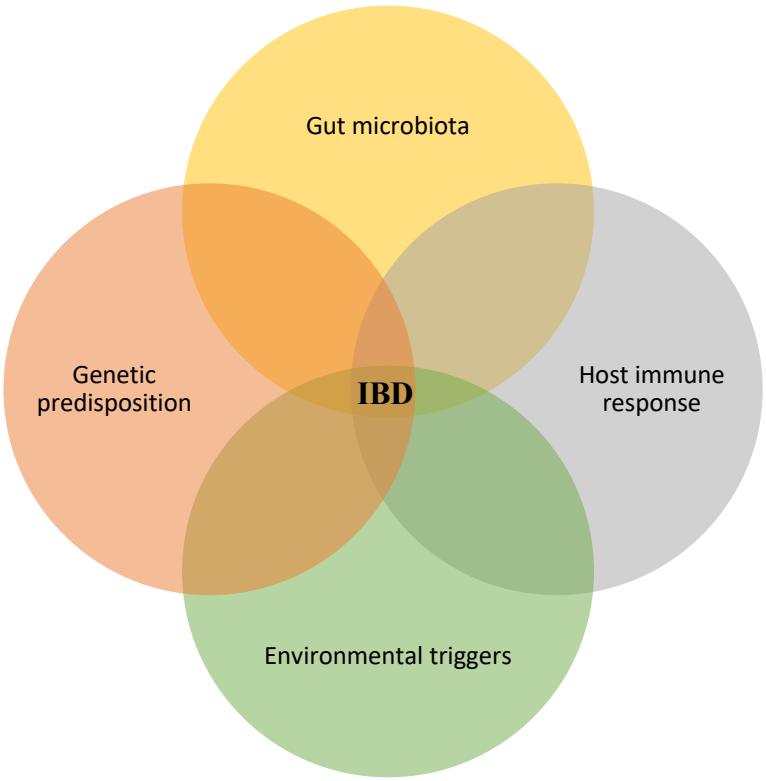


Figure 7 Inflammatory Bowel Disease Pathogenesis

A3.1 Genetics

Scientists' focus is placed upon genetics not only for the understanding of the etiopathogenesis of IBD, but also for improving its treatment and course. Numerous studies have demonstrated the vital role of genetic susceptibility in the development of IBD[15–17], however, due to the genetic complexity of the disease, a single gene trait alone cannot elucidate the pathogenesis of it. Thoughts are, that CD and UC, are likely to be related, heterogenous, polygenic disorders sharing some but not all susceptibility loci and

there is no single Mendelian pattern of inheritance for neither of the two IBD forms[15–18]. It is established knowledge that the variant phenotype of IBD is a result of delicate gene-gene, or even between allelic variants, and environment-gene interactions[19,20].

Ethnic and familial factors: Incidence of IBD among family members is consistently observed allowing one to think that there must be a strong genetic context relating to the pathogenesis of the disease[21]. Studies have shown that the most distinct risk factor of IBD occurrence is having an affected relative[22]. In first-degree relatives, the frequency of IBD can be more than a third, especially among twins[23], and also tends to be higher in CD rather than UC, thus genetic factors appear to play a more important role in CD than UC[13]. Studies have shown that around 15% of CD patients have an affected family member[24]. What is more, there is a high rate of concordance for IBD in monozygotic twins, again particularly in CD[25]. Twin studies have demonstrated 50% concordance to monozygotic twins compared to less than 10% in dizygotic[26]. Moreover, evidence such as the higher prevalence of IBD in Ashkenazi than Sephardic Jews[21] and in Northern Europeans than in Southern[27], further reinforce the genetic substrate of the disease[28], suggesting that, in some subgroups, genetic factors may play a more crucial role. Newer evidence have attributed those differences in the presence of higher IBD risk alleles and more in absolute number among the affected subgroups[29].

Genetic studies: To date, more than 200 risk loci have been identified, in a genome wide scale, representing various pathways in IBD development[30,31]; while results from genetic studies suggest that the two diseases are distinct, sharing some but not all susceptibility genes/loci[28]. What is more, not only which risk loci are present to the infected individual's genome, but also the number of those loci, seems to play a major role on the phenotype of the disease[32]. For instance, mutations of the *CARD15* gene (R702W, G908R and 1007fsinsC) at the IBD 1 locus on chromosome 16 increase the risk of CD by fortyfold, if all three mutations exist concurrently[33,34]. However, because *CARD15* mutations account for only about 20-30% of cases worldwide (i.e. not linked with CD in Japanese patients), this genetic risk factor is neither necessary not sufficient for the development of CD[35]. Thus, a more complex interaction/correlation among several gene mutations that affect immune response, defective colonic mucus, gut permeability and also how all these interact with the gut microbiota, seems to play, as new studies suggest, a key role in the development, phenotype and response to treatment of IBD[36]. A wide range of scientific studies have been deployed over the years in order to clarify the genetic role in the IBD entity. All the genes, loci and polymorphisms that have been identified, despite their effect on the immune system,

they can be divided into those who influence innate immunity, adaptive immunity, autophagy, integrity of epithelial barrier, oxidative stress responsiveness and reaction with microbes[37]. Linkage studies have firstly recognized and associated genetic loci with a chromosome 16 locus[38], Genome-Wide Association Studies (GWAS) have helped to identify numerous IBD susceptibility loci[39,40], exome and whole exome sequencing have elucidate the association of IBD and other diseases such as colorectal cancer[41] and also have helped in the diagnosis of rare IBD cases[42].

A3.2 Environmental factors

Epidemiological and other evidence have identified a number of environmental factors that may play a role in the etiopathogenesis of IBD (Figure 8); such as cigarette smoking, dietary habits, specific infections, certain drugs, stress and appendicectomy as an independent factor[43,44].

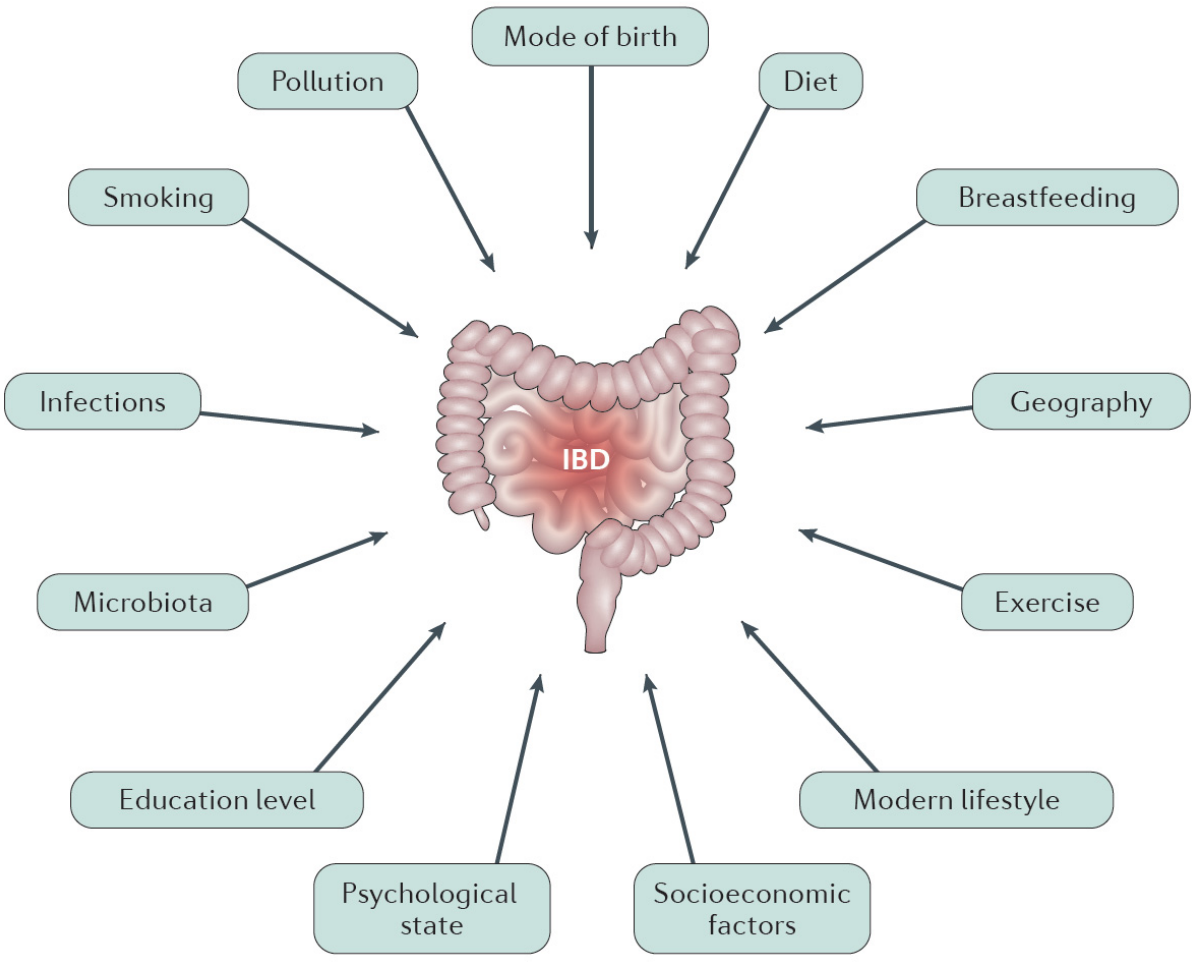


Figure 8 Environmental risk factors of IBD[44]

Smoking: Smoking is perhaps the most important of the environmental factors, being crucial for both UC and CD development and severity. Not only nicotine itself but other substances as well, such as free radicals and CO, are involved in a complex mechanism of effects aiming various targets, namely mucus layer, cytokine production, macrophage function and microcirculation[45]. A history of recent cessation of smoking is common in patients presenting with UC for the first time, and nicotine patches seems to have a modest therapeutic benefit. On the contrary, in CD, smoking increases the risk of relapse and of surgery, while cessation improves the natural course of the disease[46]. Nicotine and other constitutes of tobacco smoke, and vaping as well, seem to have a variety of effects on the inflammatory response and are under investigation of why are beneficial in patients with UC yet harmful in those with CD[47].

Diet: It is logical to anticipate a connection between diet and IBD as the latter affects the site of nutrient absorption. Gut microbiota research demonstrated the key role of proper nutrition in the preservation of a healthy microbiome and how deviations from that can have catastrophic effects in gut health causing various disorders, with IBD being among them[48]. Although, a strict cause-effect relation has not been proven yet, given the fact that gut microbiota is unique in every individual, studies have shown that a "Westernized" type of diet cause a prevalence of *Bacteroides* compared to an agrarian type of diet were the *Prevotella* genus predominates[49].Moreover, specific diets have been used as treatments, for example patients with active CD improve when their ordinary food is replaced by a liquid formula diet, and they may deteriorate thereafter on the introduction of specific foods, such as high-residue food that may cause bloating[50].

Specific infections: The arguing of whether IBD is caused by an infectious pathogen or not, is long lasting. Studies have failed to provide evidence of a single pathogen causing inflammation. What is now widely accepted is that multiple infectious agents, along with other etiological factors, contribute to IBD development. For example, despite its resemblance to, and occasional onset after, infective diarrhea, there is no evidence that UC is due to single infective agent. The possible roles of pathogenic E. coli and sulphate-reducing bacteria are under investigation. Epidemiological, molecular biological and serological research has suggested initiating roles for Mycobacterium paratuberculosis, the measles virus and vaccination, and Listeria monocytogenes in the pathogenesis of CD, but available data are controversial and require further evaluation.[51,52]

Drugs: Aspirin and nonsteroidal anti-inflammatory drugs (NSAIDs) have a well-recognized effect in the gastrointestinal track. Though, an IBD onset or relapse triggering effect has not been effectively validated, a dose-dependent, prolonged and frequent use of NSAIDs, but not aspirin, has been demonstrated to increase the risk of UC and CD development[53]. NSAIDs may precipitate a relapse of IBD perhaps as a result of inhibition of the synthesis of cytoprotective prostaglandins. Antibiotics, secondary to changes in enteric flora, increase the risk of IBD, with that being clearly demonstrated in child-onset disease where antibiotic use within the first year of life is common among this patients[54]. Moreover, the oral contraceptive pill has been associated epidemiologically in particular with CD and when is stopped the risk of IBD development drops to that of the non-exposed population[55].

Stress: Psychological stress is common in patients with IBD, particularly those with CD, due to the unpleasant, chronic and intractable nature of their illness. It is possible, however, that in some patients, stress may itself trigger pathogenesis or relapse of IBD. Individuals who have low stress levels are shown to have a reduced risk of IBD onset[56]. Depression, anxiety and other situations of perceived stress, possibly play a key role in disease deterioration. Although the use of antidepressants has been found to reduce the number of IBD relapses, a Cochrane review demonstrated no benefit of psychological interventions in IBD[57].

Appendectomy: Previous appendectomy is a protective factor against developing UC. It has been suggested that T lymphocytes in an inflamed appendix could trigger inflammation in the more distal large bowel in genetically predisposed individuals.[58]

A3.3 Immune response

The human immune system is of high complexity and still not completely understood. Contributing to tissue damage in IBD, intestinal immune system incorporates all mucosal immune cells as well as nonimmune cells, such as epithelial, mesenchymal and endothelial cells. Both branches of immune response, innate and adaptive immunity (Figure 9), contribute equally to the pathogenesis of IBD, having as effect a nonspecific inflammation caused by massively produced proinflammatory cytokines and chemokines, such as IL-1, IL-6, IL-8, TNF-α, TL1A and many others.[59]

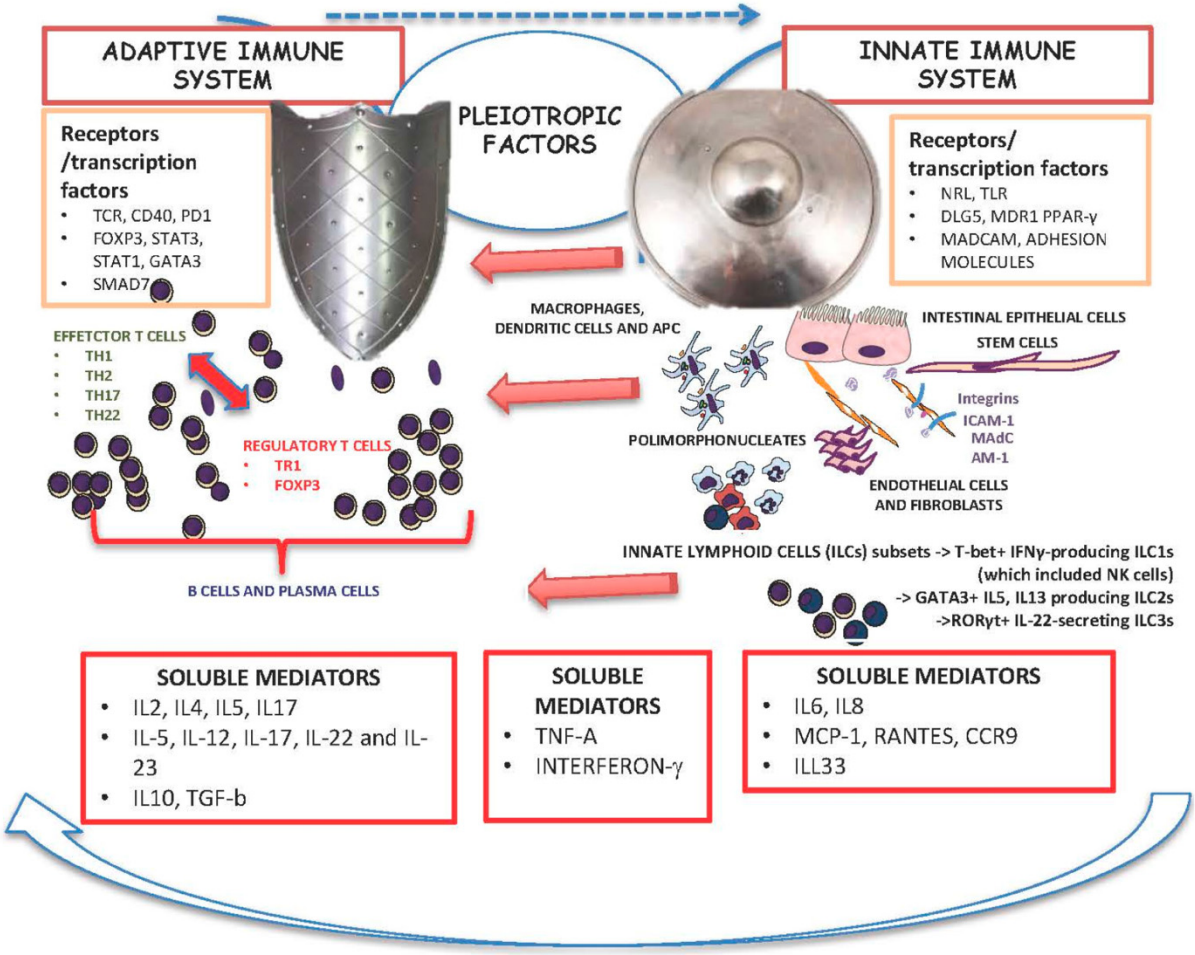


Figure 9 Innate and adaptive immune system in IBD[60]

Adaptive immunity: Adaptive immunity has been long known to play a leading role in the pathogenesis of intestinal inflammation in IBD (Figure 10). T-helper cells (Th) are of major significance in mediation of adaptive immune system. Specifically, Th1 cells eliminate intracellular pathogens, Th2 cells protect against parasites and are mediating allergic reactions and Th17 play a part in extracellular bacteria and fungi clearance[61]. When adaptive immune system is activated there is an increased production of mucosal antibodies such as IgG_1, especially in UC patients, and a markedly increased work from $CD4^+$ Th cells with raised levels of IL-12, IFN-γ, TNF-α and IL-18 from Th1 and IL-5 and IL-13 from Th2 in CD and UC patients respectively. Also, a subset of T-helper cells, Th17, whose differentiation is promoted by IL-23, are common in the mucosa of CD patients and genetic variances of *IL23R* gene are associated with prevalence of the disease in this patients[62].

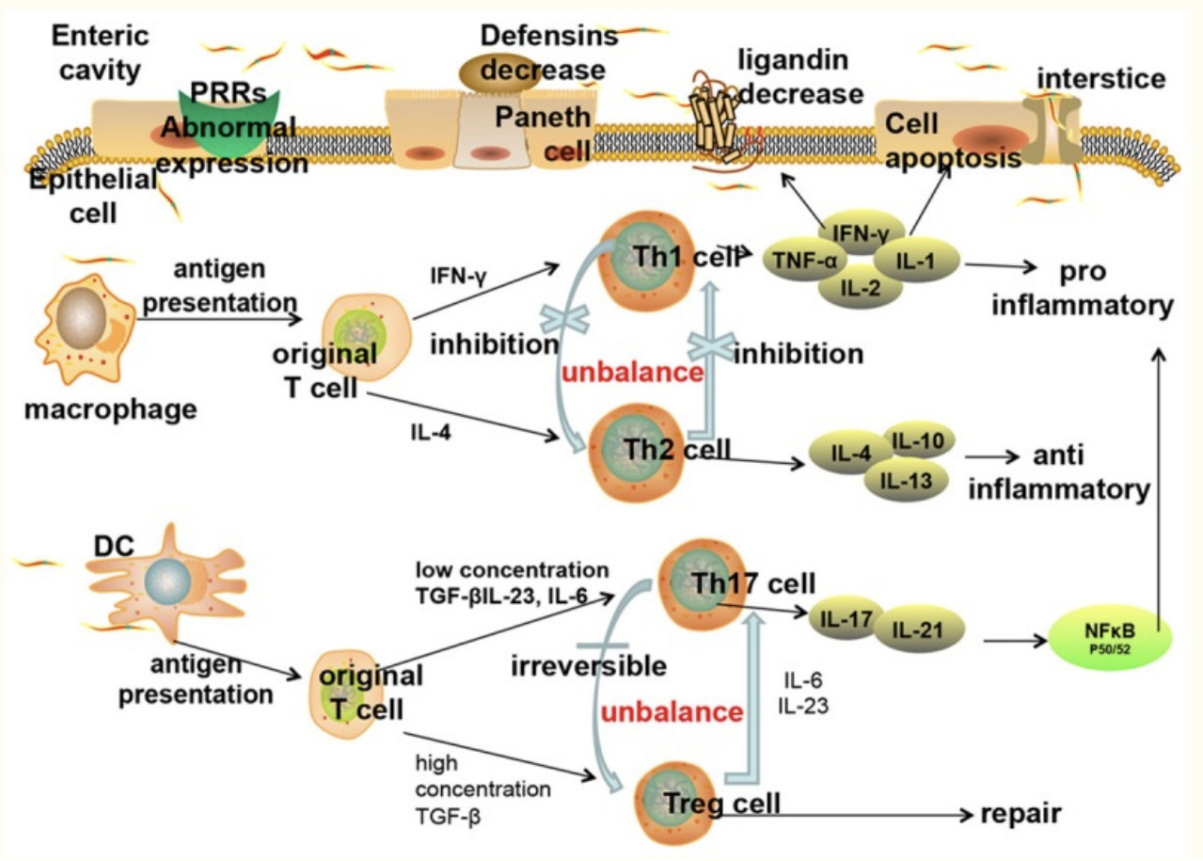


Figure 10 Adaptive immune system in IBD[63]

Innate immunity: Numerous recent genetic association studies have identified the role of various innate immune response genes in IBD pathogenesis, mainly in CD and to a lesser extent in UC. The first line of defense against any hostile stimulus is provided by innate immune responses which are mediated by various immune cells such as macrophages, neutrophils, monocytes and dendritic cells, and also nonimmune cells such as epithelial, endothelial and mesenchymal cells. One of the first signs of intestinal inflammation is the infiltration of gut mucosa and epithelium by neutrophils through impairment of epithelial barrier function, oxidative stress and tissue damage and a continuation of the inflammation due to multiple inflammatory mediators' release[64]. Impaired function of epithelial barrier is more prominent in UC patients, while autophagy, antimicrobial peptides production and innate microbial sensing are of foremost importance in CD pathogenesis. In UC patients carrying SNPs of epithelial barrier regulation genes, such as *HNF4A*, *CDH1*, *GNA12*, *LAMB1* and *ECM1*, the epithelial barrier is constituted defective with increased intestinal permeability. In particular, polymorphisms of the *ECM1* gene (extracellular matrix protein-1, a protein that is involved in epithelial barrier formation) can lead to tissue injury resulting in

intestinal ulcers and scaring in UC patients[65]. Furthermore, *NOD2* protein/*CARD15* gene variants were the first to be recognized to play a major role in CD pathogenesis. When bacteria enter the intestine, wild-type *NOD2* protein activation leads to cytokine production and clearance of bacteria[33]. On the other hand, autophagy is a process that mediates resistance to intracellular pathogens and defects in that process have been associated with CD pathogenesis. Mutations in *ATG16L1* and *IRGM* genes are strongly associated with autophagy dysregulation. Specifically, carrying of *ATG16L1* SNPs is associated with changes in Paneth cells and goblet cells, a decreased ability to clear bacteria, and an increased secretion of cytokines[66]. A typical case of innate immunity defectiveness, that contributes to intestinal inflammation, is the reduced production of IL-10 and IL-12 (immunoregulatory cytokines) from dendritic cells and excessive production of IL-1β and IL-6 (proinflammatory cytokines) in CD patients carrying *NOD2* variants and *ATG16L1* SNPs respectively.

Moreover, there is a close correlation between the adaptive and innate immune systems which is attributed to dendritic cells who are responsible for T-cell activation which are consequently activate the adaptive immune system.

A3.4 Gut microbiota

Gut microbiota is instituted at birth, but changes rapidly during the first year of life and then usually remains fairly stable. Fluctuations may occur mainly due to environmental factors or in disease. Such changes in the intestinal microbiota composition can affect homeostasis through various signaling pathways, thus affecting the interactions between bacteria and the host organism. Production and proper function of the gut's antimicrobial proteins, as well as epithelial, NK-T, Th17 and macrophage cells, are depended on the gut microbiota and the ability of the organism to recognize and respond to this microbiota[67] (Figure 11). Case-control studies that have investigated the intestinal microbiota in both CD and UC patients at inflamed and noninflamed segments, have shown that flora biodiversity is significantly reduced in fecal microbiome in IBD patients compared to healthy individuals[68]. What is more, the microbiota in IBD patients is unstable than that in healthy individuals. In the healthy intestinal flora, dominant phyla are the Firmicutes and Bacteroidetes which contribute to the production of epithelial metabolic substrates. On the other hand, CD patients' microbiota is characterized by a relative lack of these phyla, and an over-representation of enterobacteria, while in UC patients Escherichia coli predominates against Clostridium spp. In other words, in the inflamed intestinal mucosa, anti-inflammatory bacteria are over-run by proinflammatory ones[48].

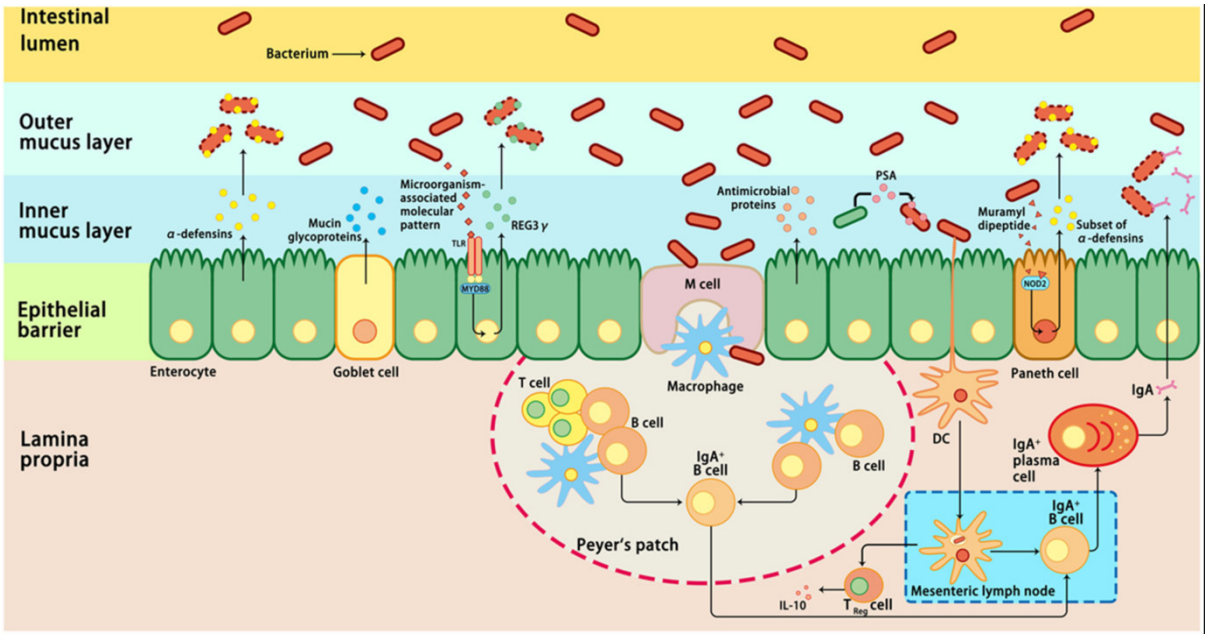


Figure 11 Host immune response to bacteria in IBD pathogenesis[67]

A3.5 Lessons from animal models

Animal models have provided valuable insights into the essential mechanisms responsible for maintaining a well-balanced intestinal immune system and the underlying defects in the gut-associated lymphoid tissue (GALT), all involved in the etiopathogenesis of IBD. Animal "knock-out" models of genes that affect immune system function have revealed that these immune defects can cause the development of mucosal inflammation. Results driven from such studies have shown that the absence or impaired function of proteins or cells involved in regulating the innate or adaptive immune system can cause mucosal inflammation and also that there is a continuous interaction between host and intestinal microbes, contributing to the protection or the inflammation of the intestinal mucosa[69]. These animal models provide valuable information and can help in developing and testing new therapeutic strategies.

A4. Therapy and treatment of Inflammatory Bowel Disease

Therapy and treatment of IBD aims to induce and maintain disease remission as well as ameliorate secondary effects, rather than aiming on modification or reversal of the underlying etiopathogenesis. There are many drugs available for the treatment and maintenance of remission of IBD, used as a standalone therapy or in combination, all depending on the severity, extent and the extraintestinal involvement of the disease. Drugs that are routinely used are corticosteroids, aminosalicylates (5-ASA), immunosuppressive agents, such as azathioprine and 6-mercaptopurine and immunodulators (biologic agents) such as infliximab and adalimumab. In some cases, other drugs that can be proven helpful are metronidazole and broad-spectrum antibiotics, methotrexate, bismuth and arsenical salts and cholestyramine. Although therapy protocols may exist, each patient's treatment is individual and determined by whether it's CD or UC, the severity and extent as well as the presence of EIMs.[70,71]

Ulcerative colitis: Treatment of UC is dependent upon severity and location of the disease. It is common practice to start the treatment on a step-up manner; aminosalicylates as first-line therapy and biologic agents as last.

For new cases of mild to moderate proctitis or left-sided UC and of pancolitis (Figure 12), induction to therapy usually starts with 5-aminosalicylic (5-ASA); usually with oral or rectal mesalazine and oral sulfasalazine[72]. Poor responders in 5-ASA are then treated with oral corticosteroids (i.e. prednisone) for 4 weeks plus azathioprine/mercaptopurine in order to induce remission. If again no response is achieved, hospitalization and intravenous corticosteroids are the next step prior to entering into biologic treatment with anti-TNFα agents (i.e. infliximab and adalimumab) and immunosuppressants like cyclosporine and tacrolimus[70].

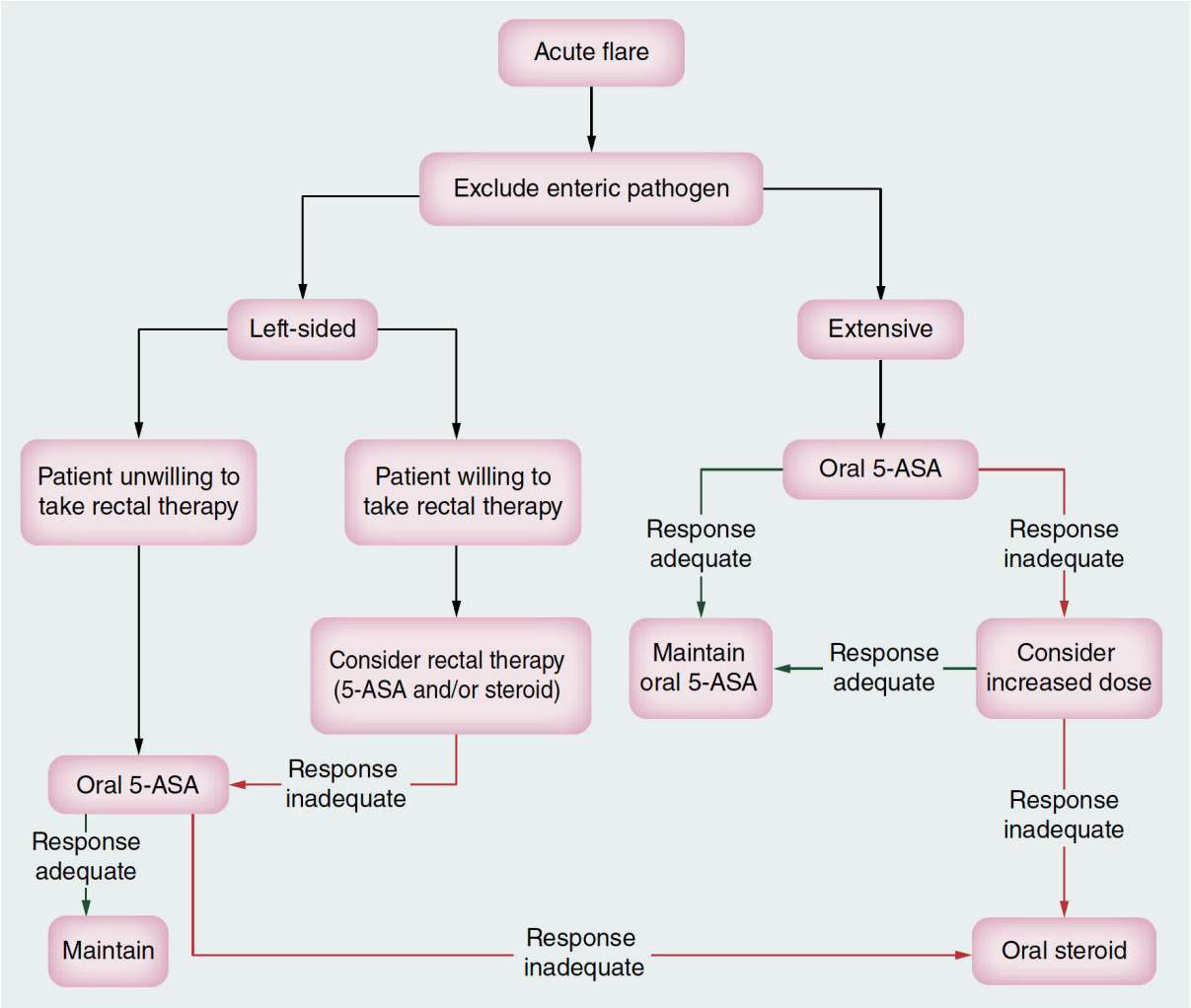


Figure 12 Mild to moderate UC therapeutic management algorithm[73]

Moderate to severe or fulminant UC cases (Figure 13), along with 5-ASA, corticosteroids and azathioprine/mercaptopurine treatment can also be treated with infliximab or adalimumab[70,72].

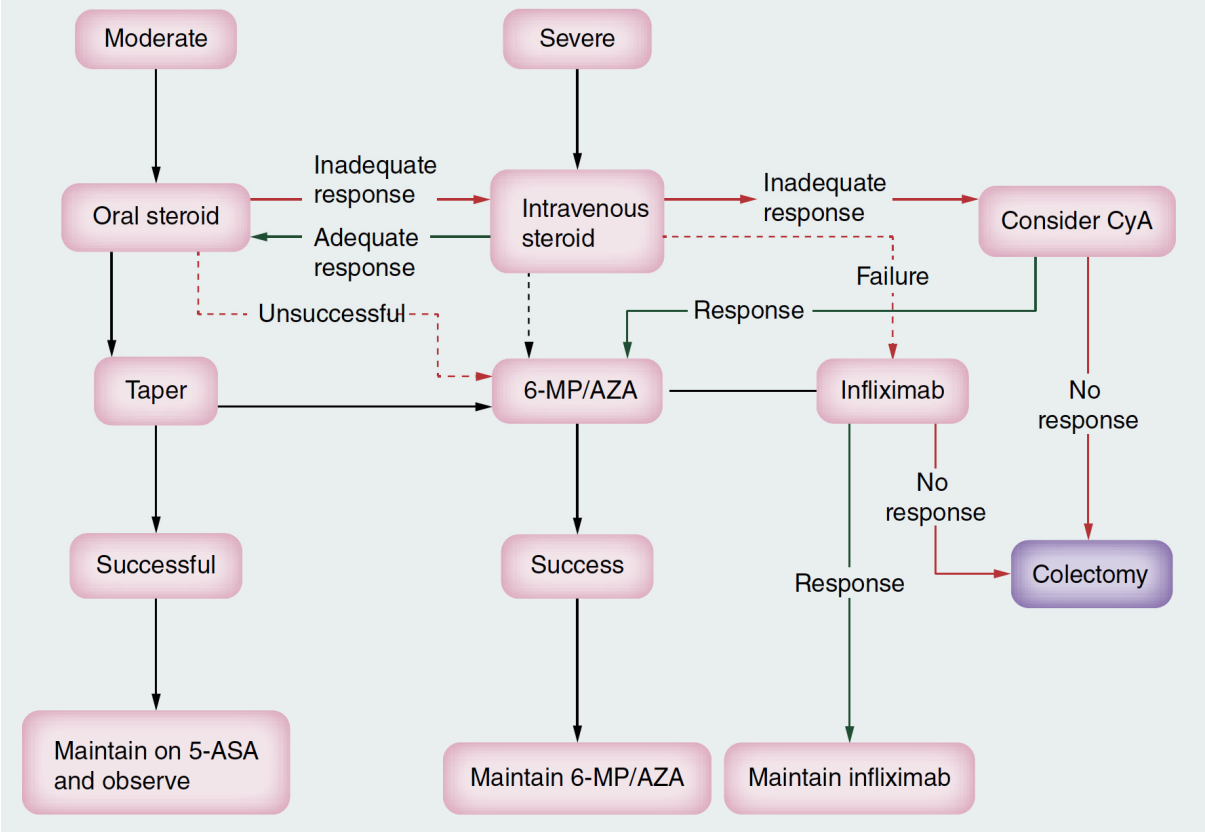


Figure 13 Moderate to severe UC management algorithm[73]

When remission is achieved, maintenance of remission is managed with oral 5-ASA or rectal 5-ASA for proctitis or left-sided cases. If, while on 5-ASA, relapse occurs then the patients who are steroid-dependent and those with severe UC, azathioprine/mercaptopurine are the drug of choice. Moreover, steroid-dependent patients can maintain remission with infliximab which is steroid-sparing[70].

Finally, when pharmaceutical treatments fail or an episode of fulminant UC occurs, surgery can act as a last resort rescue therapy. It is reserved for severe and difficult to treat cases and is indicated in life threatening emergencies as a definite solution of UC complications (perforation, refractory rectal bleeding and toxic megacolon). Total colectomy with j-pouch formation is the surgical procedure of choice[74].

Crohn's disease: Similar to UC, medical treatment for CD is approached depending on the location of the disease (Figure 14).

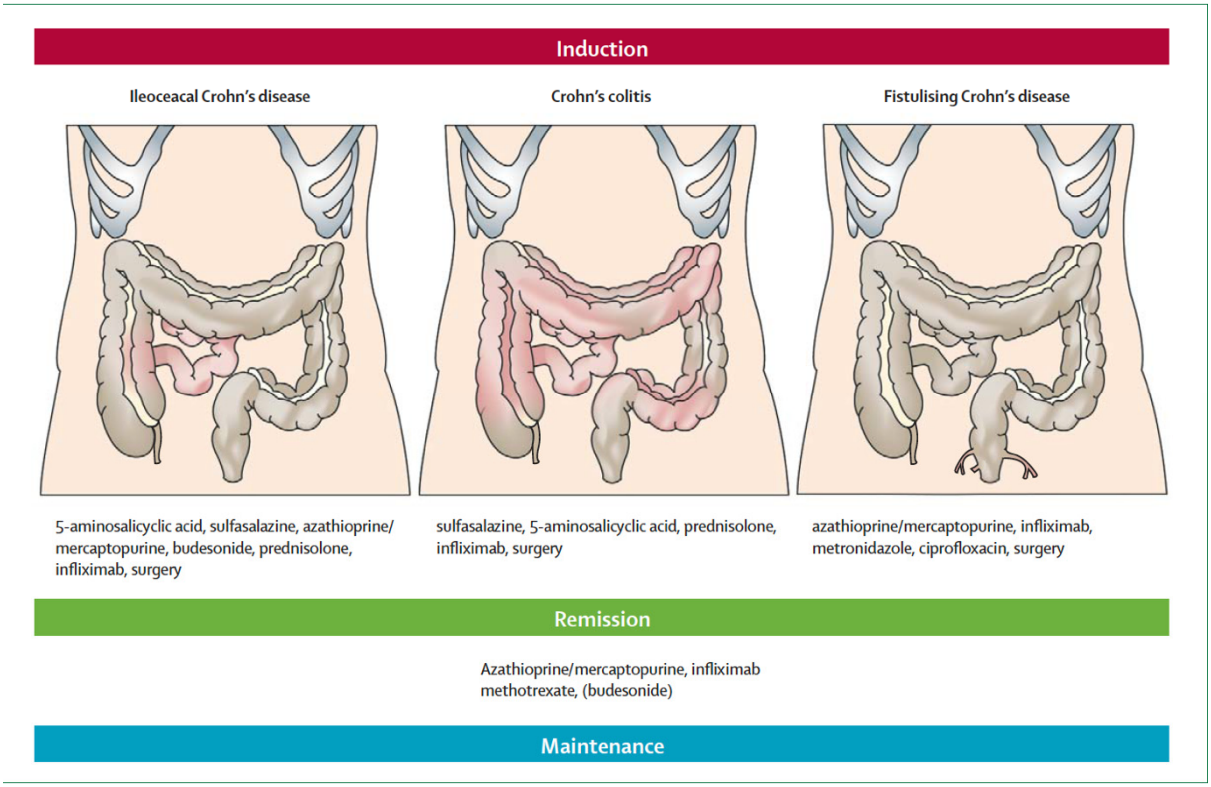


Figure 14 Therapeutic management of CD[7]

Mild to moderate Crohn's disease is traditionally treated with 5-ASA. Sulfasalazine is an effective agent for induction of remission in active disease, especially to those with colonic involvement; though some patients may develop tolerance. Interestingly, mesalazine has not been proven effective. Patients with ileal or right colonic involvement or both are treated with budesonide, which markedly reduces side-effects of systemic corticosteroids and has similar efficacy to prednisolone for the induction of remission in active CD. Additionally, antibiotics fail to induce remission in active CD. In a top-up approach, patients with mild to moderate CD who do not respond to sulfasalazine or budesonide, and patients with moderate to severe disease are treated with oral prednisone. If again this doesn't prove effective, next step is azathioprine/mercaptopurine and methotrexate. Patients with moderate to severe disease who fail to remit with sulfasalazine, budesonide, conventional corticosteroids, and azathioprine/mercaptopurine or methotrexate can be treated with infliximab or adalimumab. A newer approach in severe cases of CD, is a 'top-down' strategy[75] (Figure 15).

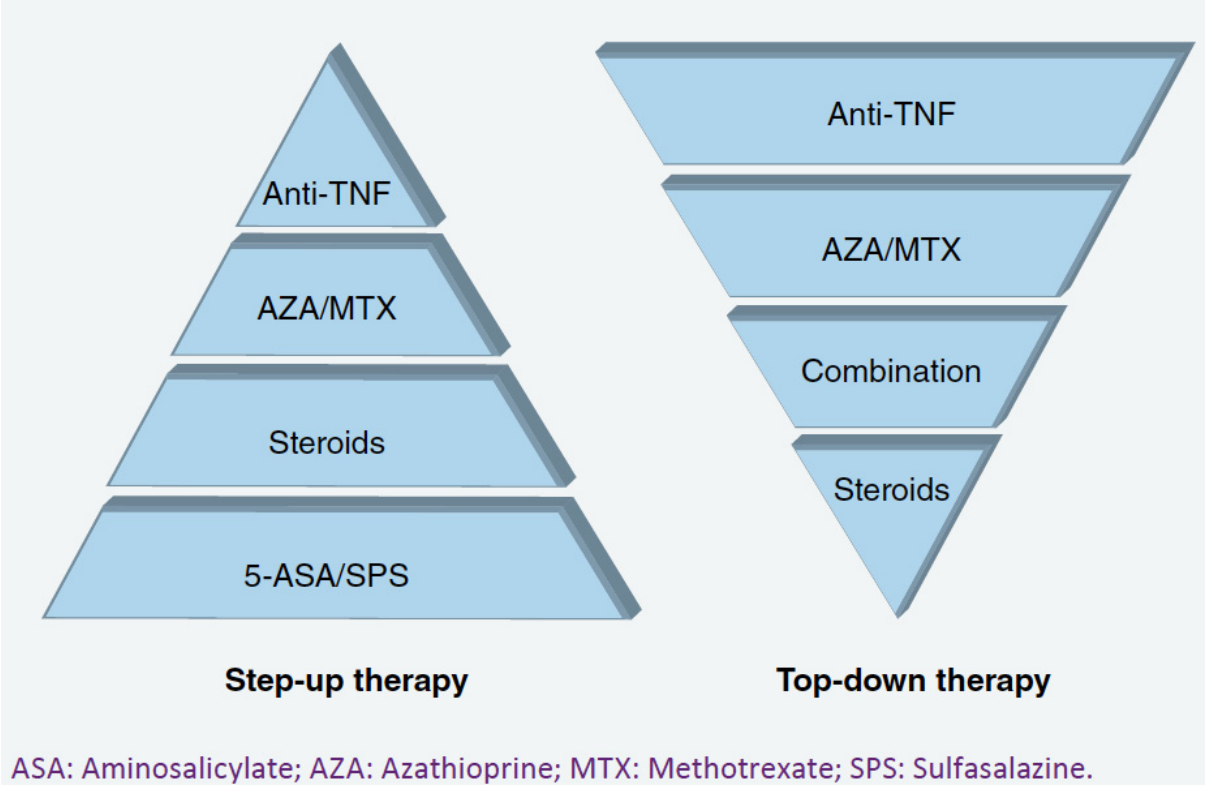


Figure 15 Step-up and top-down therapy for CD patients[76]

As shown in Figure 14, maintenance of remission in CD patients is achieved with oral 5-ASA and budesonide. Steroid dependent patients and those with moderate to severe disease can be remain in remission with azathioprine/mercaptopurine, or methotrexate. Infliximab and adalimumab are effective for maintenance of remission, steroid-sparing, and mucosal healing in patients who are unable to maintain remission or who remain steroid dependent despite treatment with azathioprine/mercaptopurine, or methotrexate[77].

Fistulising Crohn's disease needs a diversified therapeutic approach (Figure 14). As first-line treatment of fistulising CD is the use of antibiotics such as ciprofloxacin or metronidazole, with second-line treatment the use of azathioprine/mercaptopurine. Patients with refractory disease with no improvement with the abovementioned therapies are then treated with infliximab or adalimumab[78].

Finally, unlike ulcerative colitis, surgical treatment won't cure Crohn's disease; though it might be necessary in order to achieve remission in refractory cases. Indications for colonic CD cases are as for UC. Other indications are part or complete bowel obstruction due to fibro-stenosis, abdominal abscesses and fistulas of various locations[79]. Post-operative maintenance of remission is partially achieved with azathioprine/mercaptopurine and metronidazole[71].

A5. Genetics of Inflammatory Bowel Disease

Numerous observations to humans and animal models indicate significant genetic susceptibility contributing to IBD development. To date, more than 200 susceptible loci have been identified, in a genome wide scale, representing various pathways in IBD development[30,80]. The majority of this loci is associated with both CD and UC and over 70% of them are also associated with other immunomediated diseases (i.e. ankylosing spondylitis, psoriasis and celiac disease)[81,82].

Family and twin studies have played the most compelling role in providing clinical evidence of IBD's heritability risk, suggesting also a more prominent role of genetic factors in CD rather than UC[23]; IBD doesn't appear to follow a Mendelian pattern of inheritance. First-degree relatives of IBD patients are more likely to develop the disease compared to the general population. Also, children whom their parents suffer from IBD have a greater risk of developing the disease in earlier age[21]. Furthermore, the heritable pattern of IBD occurrence is also supported by the clinical features of the disease. Concordance of the disease location and behavior among family members is described[83].

On the other hand, animal studies, mainly on rodents, have shown that even one alteration in any of the susceptible genes can lead to disease development (IBD-like syndromes), while in humans there is a potential aggregation effect of several loci contributing to IBD phenotype[84,85].

Analysis of the various genes involved in IBD pathogenesis as well as of the functional properties of the proteins encoded by these genes, has assisted in identifying the pathways involved in the pathogenesis of the disease (Figure 16).

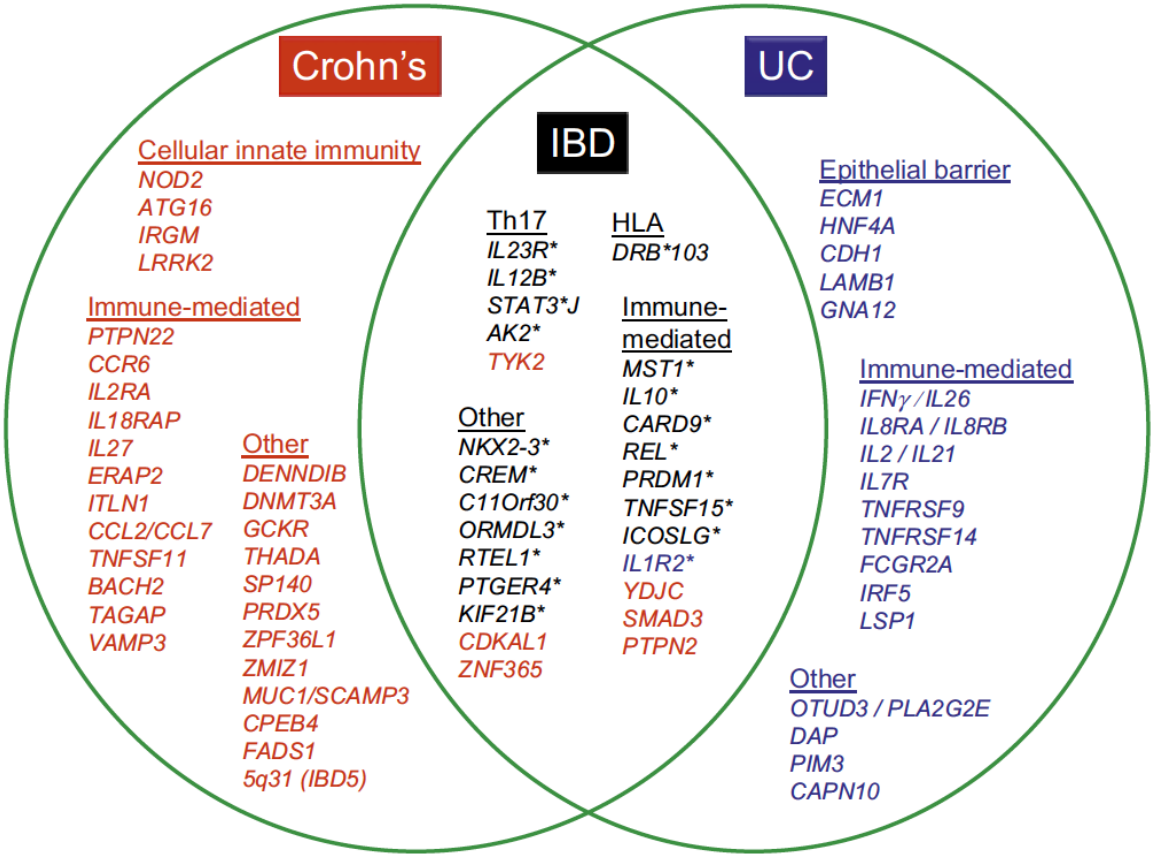


Figure 16 Inflammatory bowel disease susceptibility loci[82]

One of the first genes recognized to contribute in the development of CD is the *CARD15* gene of the IBD1 locus on chromosome 16 which encodes for the *NOD2* protein. NOD2 plays a crucial role in pathways involved in innate immunity responsible for recognizing microbial products. There are many, up to 30, polymorphisms of the *CARD15* gene that have been associated with CD and not UC, but only three (Arg702Trp, Gly908Arg and Leuc1007insC) are the more common[86]. Interestingly, IBD related *CARD15* polymorphisms are reported to impact CD occurrence only in European Caucasians populations, and are completely absent in Chinese, Japanese and African-American populations implicating the importance of ethnic variations in IBD development[87].

Another crucial pathway in the development of inflammation is the autophagy pathway. Its role has been implicated in several immune-related processes that all potentially can influence the pathogenesis of IBD including the elimination of intracellular microorganisms (xenophagy), recycling of the intracellular organelles (mitophagy), antigen presentation, secretion and vesicular trafficking and cytokine-based regulation of the inflammation[88,89].

Numerous autophagy-related genes have been identified to play a role in CD such as *ATG16L1, DAP1, SMURF1, IRGM, LRRK2* and *NOD2* as well[90]. One of the most related risk alleles of the *ATG16L1* gene that is associated with CD development is the T300A (rs2241880), which results in impairment of autophagy and antigen presentation. Both these processes are enhanced by MDP (muramyl dipeptide) which is a bacterial ligament for CARD15/*NOD2* receptor[91]. That overlapping of these two genes suggests that these pathways are integrated and defective in some patients with CD.

Additionally, adaptive immune genes are shown to have active involvement in IBD pathogenesis. Genes involved in regulating the innate immunity pathway include *IL23R, IL12B, STAT3, JAK2, TNFSF15, TYK2,* genes that regulate both IL-17 and IL-23 receptors which have been implicated for increased IBD risk, as well as genes that regulate IL-10 immune pathway (*IL-10, STAT3, JAK2,* and *TYK2*) which is also implicated for having an independently link to both CD and UC[92].

Furthermore, epithelial barrier has crucial importance in the regulation of intestinal homeostasis. A dysfunction of the barrier can result in a pronounced inflammatory response resulting from the increased antigen translocation across the epithelium. It is established knowledge that IBD patients have an increased intestinal permeability. Genes involved in the regulation of epithelial barrier function have been associated with IBD risk, mainly with UC and to a lesser extent with CD. Genes included in this group are *OCTN2, ECM1, CDH1, HNF4A, LAMB1, GNA12* and *DLG5* (mainly in CD)[81]. In addition, genes (i.e. *XBP1*) that control Paneth cells function and ER stress have been identified in CD patients[93].

Added to the above, genetic variants seems to influence disease location, behavior, prognosis and the presence of EIMs. For instance, T300A variant of *ATG16L1* gene is linked to predisposing to ileal CD[94]. *CARD15/NOD2* is linked with upper GI Crohn's disease, fibrostenosis and the need of surgery[95]. Genes related with EIMs include *FcRL3, HLADRB*103, HLAB*27, HLA-B*44, HLA-B*35*[32,96].

Finally, it is clearly understood that no worldwide pattern exists concerning IBD genes and associated polymorphisms. Studies across the globe have yielded controversial results. Polymorphisms in genes, especially of those who regulate autophagy pathways, such as *ATG16L1* and *CARD15/NOD2* that play crucial part in IBD pathogenesis in some populations (i.e. those of Northern European ancestry), at others do not (i.e. Japanese)[97]. On the other hand, shared genes across populations do exist (i.e. *TNFSF15, FCGR2A, HLA* alleles)[98].

A5.1 *ATG16L1* T300A (rs2241880)

Autophagy related 16 like 1 (*ATG16L1*) gene, which encodes for the ATG16L1 protein, is located on chromosome 2 in position q37.1 (Figure 17).

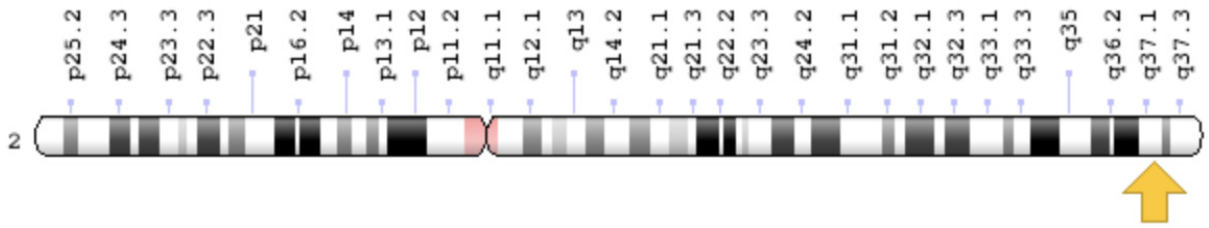


Figure 17 ATG16L1 gene position on Ch2q37.1[99]

Two independent GWASs identified *ATG16L1* as a susceptibility locus implicating the role of autophagy in CD development but not UC[100,101]. *ATG16L1* is a core autophagy protein playing multiple roles in the immune system, including xenophagy, antigen presentation, production and secretion of IL-1β. Dysregulation of the protein, due to genetic variants, leads to CD-related processes[102].

The best and most studied *ATG16L1* genetics variant is rs2241880. It entails a threonine to alanine substitution (T300A) and it is strongly associated with Crohn's disease development and especially with ileal involvement[94]. Many studies of different populations confirm these findings[103–105], though controversies in ethnic variations do exist[106–108]. Of note, is the confirmed association of the presence of rs2241880 along with two other SNPs, rs6596075 of *IBD5* gene and rs17221417 of *CARD15*. Coexistence of this three variants in an individual increases the risk of CD development by 20fold[94]. What is more, T300A has been used in prognostic models to diagnose CD, to calculate CD risk, or differentiate CD from UC as well IBD from IBS[109].

A5.2 *ECM1* T130M (rs3737240) and G290S (rs13294)

Extracellular Matrix Protein 1 (*ECM1*) gene, which encodes for ECM1 protein, is located in chromosome 1 in position 1q21.2 (Figure 18).

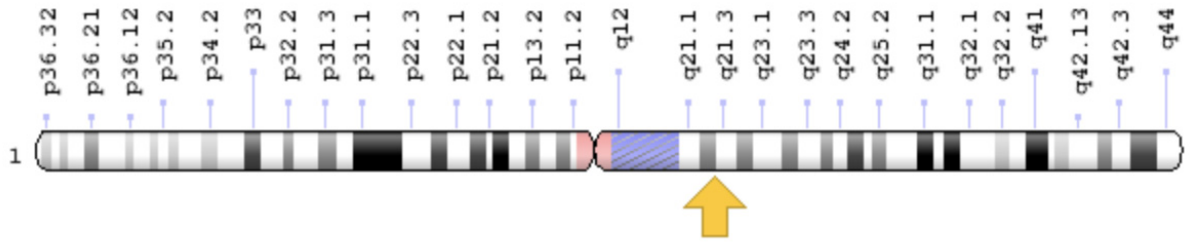


Figure 18 ECM1 gene position on Ch1q21.2[110]

ECM1 locus, unknown since 2008, was identified as a UC susceptibility gene but not for CD[65]. *ECM1* protein is a secreted glycoprotein playing significant role in endothelial cell growth and differentiation, angiogenesis, as well as intestinal epithelial homeostasis, securing its cohesion. Dysregulation of the protein, due to genetic variants, leads to UC-related processes, such as ulceration and scarring[111].

Two variants of *ECM1* gene are significantly associated with high risk of UC development; rs3737240, a threonine to methionine substitution (T130M), and rs13294, a glycine to serine substitution (G290S)[65], though controversial reports exist here as well[112]. Again, as with T300M (*ATG16L1)*, these two UC-related variants have been employed in prognostic models to diagnose UC, to calculate UC risk, or differentiate UC from CD as well as IBD from IBS[109].

A5.3 Future challenges of IBD genetics

Over the past decades, a tremendous number of genetic studies have brought to light many strong evidences of the genetic susceptibility in IBD development. However, the more answer they yielded the more questions are posed. Although many of these identified IBD loci have been correlated significantly in a disease-association manner, their importance in sub-phenotypes and disease behavior are yet to be characterized. Fine mapping of the approximately 230 known IBD susceptibility loci will help to determine how all these contribute to disease risk. What is more, while up until now most of the studies were on

population of European ancestry, studying the differences and the associations in different ethnic groups will yield perplexing yet elucidating results. In addition, more accurate genotype-phenotype studies through prognostic models, will enlighten clinician's role in predicting disease outcomes and response to treatment.

Nowadays studies are focusing on networking or clustering all these genetic variances emphasizing on elucidating the functional complexities underlined. Although having now gone beyond genetic association and linkage studies, with whole genome and whole exome sequencing replacing them in the majority of the studies, scientist's focus is upon the impact of epigenetic alterations on disease occurrence.

Although epigenetics is an emerging field, its result adds extra complexity to the IBD genetic substrate. It is anticipated that, similar to genetic studies, won't address all issues, but it will offer a great value of information for larger bioanalytical models that will encompass all new and emerging IBD-related research fields (i.e. transcriptomics, metagenomics and metabolomics).[39,113]

A6. Aim of the study

With this study, focusing mainly on the population of Northwest Greece, we aimed at understanding the natural course of the disease, study the predisposing factors and related genes in order to reveal underlying genetic associations and determine early clinical, genetic and immunological predictors of outcome and response to treatment.

Our primary focus on achieving that was with evaluating if in our genotypic study findings, a predictive and/or prognostic association with disease development or a specific clinical phenotype exist and if these findings correlate with existing data from other regions worldwide.

In particular, we studied 3 single nucleotide polymorphisms (SNPs) from 2 genes; rs2241880 (T300A) of *ATG16L1* gene and rs3737240 (T130M) and rs13294 (G290S) of *ECM1* gene. We genotyped all our study subjects (223 healthy volunteers and 205 IBD patients) for all 3 aforementioned SNPs using the RT-PCR method. The obtained genotypic results were then investigated for association with disease susceptibility as well as for genotype-phenotype correlations.

B. Materials and methods

B1. Study cohort

We recruited a total of 428 individuals. Of them, 205 were unrelated IBD patients (108 CD patients and 97 UC patients) and 223 unrelated healthy blood donors (control group). All IBD patients participating in the study were followed up at the Outpatient Clinic and the Gastroenterology Department of the University Hospital of Ioannina, Greece and all the healthy blood donors were attendees of the University Hospital of Ioannina's Blood Bank. All study subjects were originated from the NW Greece region and were of Caucasian ethnicity (Figure 19).

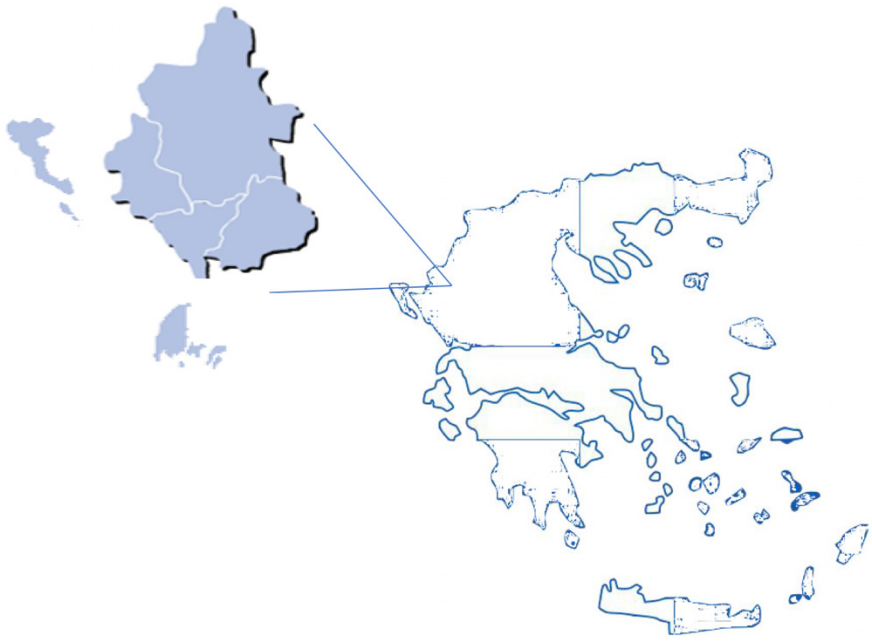


Figure 19 North-western Greece (study region)

For all patients, along with whole blood samples (for DNA extraction), we collected demographic and clinical data from the hospital's registry. The data that were collected included: current age and age at onset, type of diagnosis (CD or UC) and clinical details of the disease (extent, severity, behaviour), presence and type of extra intestinal manifestations (EIMs) or other autoimmune disease, history of appendicectomy, cholecystectomy and tonsillectomy, and if they needed surgical operation for their disease.

Diagnosis of either CD or UC was based on standard clinical, endoscopic, radiological, and histological criteria[114,115]. Both CD phenotype (age at onset, disease location and behavior) and UC phenotype (extent and severity) were determined according to the Montreal Classification[8]. Presence of EIMs and/or other autoimmune disease was established by relevant specialists.

For all healthy volunteers, along with whole blood samples (for DNA extraction), we collected demographic and health-related data in a form of an interview. Recruitment of the control group subjects was based on not having any gut or liver related disease.

All subjects were informed of the nature of the study and signed the informed consent form. Demographic and clinical data of the study subjects are presented in Table 1 and Table 2.

Table 1 Demographic data of the study population

	CD	**UC**	**CONTROL**
Total number	108	97	223
Sex (male/female)	66/42	56/41	173/50
Age at data collection (mean, ± SD, range)	41.7 ± 14.7 (15 – 78)	49.6 ± 16.6 (18 – 86)	38.3 ± 10.3 (20 – 71)
Age at diagnosis (mean, ± SD, range)	32.7 ± 13.4 (13 – 64)	39.3 ± 14.7 (15 – 83)	
≤16 17 - 40 >40	6 (5.6%) 76 (70.4%) 26 (24.1%)	2 (2.1%) 50 (51.5%) 45 (46.4%)	
≤16 >17	6 (5.6%) 102 (94.4%)	2 (2.1%) 95 (97.9%)	
Smoking	50 (46.3%)	41 (42.3%)	

Table 2 Clinical data of the study population

	CD	UC
Disease Location		
L1 - Ileal	33 (30.6%)	
L2 - Colonic	32 (29.6%)	
L3 - Ileocolitis	43 (39.8%)	
L4 - Upper gastrointestinal	6 (5.6%)	
Disease behavior		
B1 - Nonstricturing, nonpenetrating	54 (50.0%)	
B2 – Stricturing	22 (20.4%)	
B3 – Penetrating	21 (19.4%)	
B2 + B3	11 (10.2%)	
p - Perianal disease	26 (24.1%)	
UC extent		
E1 – Ulcerative proctitis		10 (10.3%)
E2 – Left sided		61 (62.9%)
E3 - Pancolitis		26 (26.8%)
UC severity		
Mild/Moderate		73 (75.3%)
Severe		24 (24.7%)
EIMs	67 (62%)	46 (47.4%)
1 EIM	37 (34.3%)	21 (21.6%)
>1 EIMs	30 (27.8%)	25 (25.8%)
Joint	45 (41.7%)	32 (33.0%)
Osteoporosis	18 (16.7%)	7 (7.2%)
Skin/Oral	26 (24.1%)	18 (18.6%)
Ocular	10 (9.3%)	10 (10.3%)
PSC	1 (0.9%)	2 (2.1%)
Vascular	2 (1.9%)	3 (3.1%)
Nephrolithiasis	5 (4.6%)	5 (5.2%)
Other autoimmune disease	6 (5.6%)	1 (1.0%)
Operated	16 (14.8%)	3 (3.1%)
anti-TNFα	63 (65.6%)	33 (34.4%)
Cholecystectomy	7 (6.5%)	6 (6.2%)
Appendicectomy	20 (18.5%)	11 (11.3%)
Tonsillectomy	16 (14.8%)	16 (16.5%)

B2. Study protocol

The presented study was carried out at the Research Laboratory of Hepato-Gastroenterology, Division of Gastroenterology, Faculty of Medicine School of Health Sciences, University of Ioannina, Ioannina, Greece.

B2.1 Whole blood collection and DNA Extraction

Ten mL (10 mL) of whole blood from a peripheral vein was collected in standard EDTA tubes from all IBD patients and healthy volunteers.

Genomic DNA was extracted from the whole blood samples with Nucleospin Blood XL kit (Macherey-Nagel, Germany) according to the manufacturer's protocol (available at https://www.mnnet.com/Portals/8/attachments/Redakteure_Bio/Protocols/Genomic%20DNA/UM_gDNABlod.pdf). No modifications of the protocol were necessary. Figure 20 demonstrates the protocol at a glance.

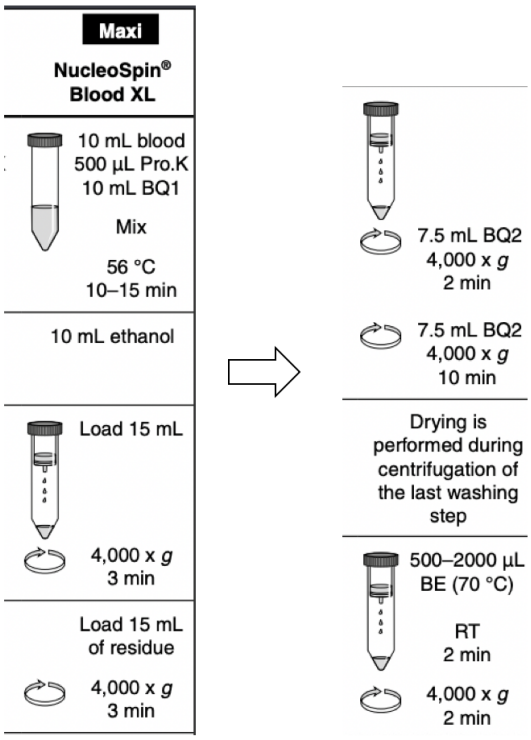


Figure 20 Nucleospin Blood XL protocol at a glance[116]

Extracted DNA yield as well as purification were calculated based on spectrophotometry measured with *NanoDrop™ 1000*. Mean yield was 500µg with a mean concentration of 350-400ng/µL and a mean $A_{260/280}$ ratio of 1.7-1.8 ($A_{260/280}$: ratio of sample absorbance at 260 and

280 nm. The ratio of absorbance at 260 and 280 nm is used to assess the purity of DNA. A ratio of ~1.8 is generally accepted as "pure" for DNA). Afterwards, all DNA samples were labelled accordingly (CD for Crohn's Disease, UC for ulcerative colitis and CTRL for the control group) and given a serial number. Then, samples were stored at -80^{o} C for later use.

B2.2 Genotyping

Three SNPs of two genes were investigated in this study. Specifically, rs2241880 (T300A)[117] of *ATG16L1* gene and rs3737240 (T130M)[118] and rs13294 (G290S)[119] of *ECM1* gene. The selection of these genes and their respective polymorphisms was based on current bibliography. Both genes and their aforementioned polymorphism are strongly associated with IBD (*ATG16L1* with CD and *ECM1* with UC)[120,121] and also, they have been both used in screening and prognostic algorithms[109].

All DNA samples were prepared for genotyping using the KAPA PROBE FAST qPCR Kit Master Mix (© Roche Sequencing). The preparation of the reaction specimen was based on manufacturer's protocol (available at https://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Roche/Datasheet/1/pfukbdat.pdf):

- The reaction components were: PCR-grade water (© Jena Biosciences), Mastermix (Kapa Probe Fast qPCR kit, © Roche Sequencing), forward (Fw) primer, reverse primer (Rv), SNP-specific probe of each allele (total of 2) and DNA sample
- The total reaction volume was 20μL consisting of: 1μL genomic DNA (~65ng), 10μL Mastermix, 7.8μL PCR-grade water, 0.4μL forward primer, 0.4μL reverse primer and 0.2μL of each of the two SNP-specific probes.

Oligonucleotide primers (forward and reverse) and SNP-specific probes were synthesized by VBC-Biotech Services GmbH (Vienna, Austria).

The following oligonucleotide primers and SNP-specific probes were used:

- ATG16L1-T300A-Fw: 5'-TGA AGC ATA CTT ACG AAG ACA CAC-3'
- ATG16L1-T300A-Rv: 5'-TGT CTC TTC CTT CCC AGT CC-3'
- ATG16L1-T300A-T: 5'-CCA GAA CCA GGA TGA GTA TCC ACA T-3'

- ATG16L1-T300A-C: 5'-CAG AAC CAG GAT GAG CAT CCA CAT-3'
- ECM1-T130M-Fw: 5'-CCC CAG ATT CTT TCA ATC CTC-3'
- ECM1-T130M-Rv: 5'-AGG ACT CAG GTT CTG GAT GG-3'
- ECM1-T130M-C: 5'-TTT CCC CAT TCC AGG AAC GCC AGC TCC ATT-3'
- ECM1-T130M-T: 5'-TTT CCC CAT TCC AGG AAT GCC AGC TCC ATT-3'
- ECM1-G290S-Fw: 5'-CCC AAC TAT GAC CGG GAC-3'
- ECM1-G290S-Rv: 5'-GCA ACT TAC TGC TTG GTG AG-3'
- ECM1-G290S-G: 5'-CTT GAC CAT TGA CAT CGG TCG AG-3'
- ECM1-G290S-A: 5'-CTT GAC CAT TGA CAT CAG TCG AGT C-3'

Genotyping was carried out using the RotorGene 3000 RealTime-PCR system (Corbett Research, Australia) and allelic discrimination was based on the RotorGene 3000 software (http://www.lth.se/fileadmin/sciblu/QC/RG3000.pdf).

The cycling protocol was determined as follows:

- enzyme activation at 95^{o} C for 3 minutes
- followed by 40 two-step cycles of denaturation at 95^{o} C for 3 seconds
- annealing-elongation at 60^{o} C for 20 seconds

The protocol was applied in all samples for all three SNPs.

Following an initial run, three samples from each SNP (a total of 9), one wildtype, one mutant and one heterozygote, that have given the strongest signal of genotype confirmation (Figure 21, RotorGene 3000 software) were selected for subsequent verification by dsDNA sequencing. Sequencing services were performed by VBC-Biotech Services GmbH (Vienna, Austria). DNA samples were prepared for sequencing according to VBC-Biotech protocol (available at https://www.nucleics.com/DNA_sequencing_support/sequencing-service/vbc-genomics.html).

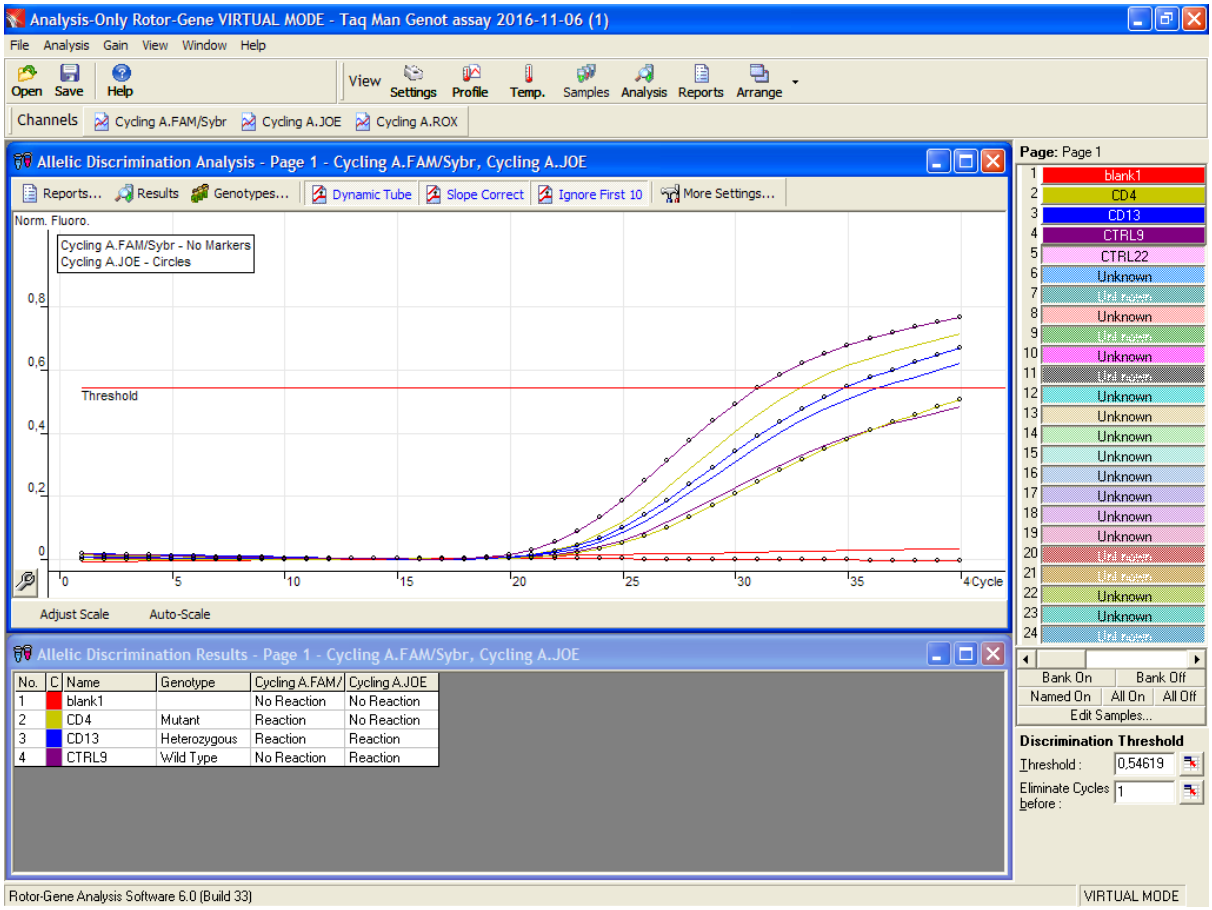


Figure 21 RT-PCR genotyping (RotorGene 3000)

Next, after the verification, these 9 samples (which were correctly chosen and subsequently verified by dsDNA sequencing) were used as control samples in each and every run. Each triplet (wildtype, heterozygote and mutant) for its specific SNP.
Obtained results were then used for statistical analyses.

B2.3 Decision Support System (DSS)

As a side project of this thesis was the building of an innovative decision support system (DSS) that would function as a prognostic model helping the clinician to better evaluate and treat patients with IBD. In order to achieve that, the DSS needs to incorporate two modules, a data repository module and a knowledge extraction and statistics module.
The Data Repository module is a centralized data repository for annotation data. Clinical and demographic data and data from genotypic and serological markers' studies/screening are incorporated into the database. For the clinical and demographic data, hard-copy

medical records have been digitized and registered to the database. Genotypic study results (main aim of the presented study) are incorporated as well and have been registered to the database. Serological marker's study results (not part of the presented study) have been registered to the database.

The DSS, in the front end, will provide tracking, data query, report generation, process management functions, data handling as well as statistics, data mining and knowledge extraction capability (Knowledge Discovery and Statistics module incorporated in the back-end of the system). Moreover, the module will contain a Data Representation module that will handle the presentation of the extracted knowledge from the patients' data. Figure 22 demonstrates the DSS workflow.

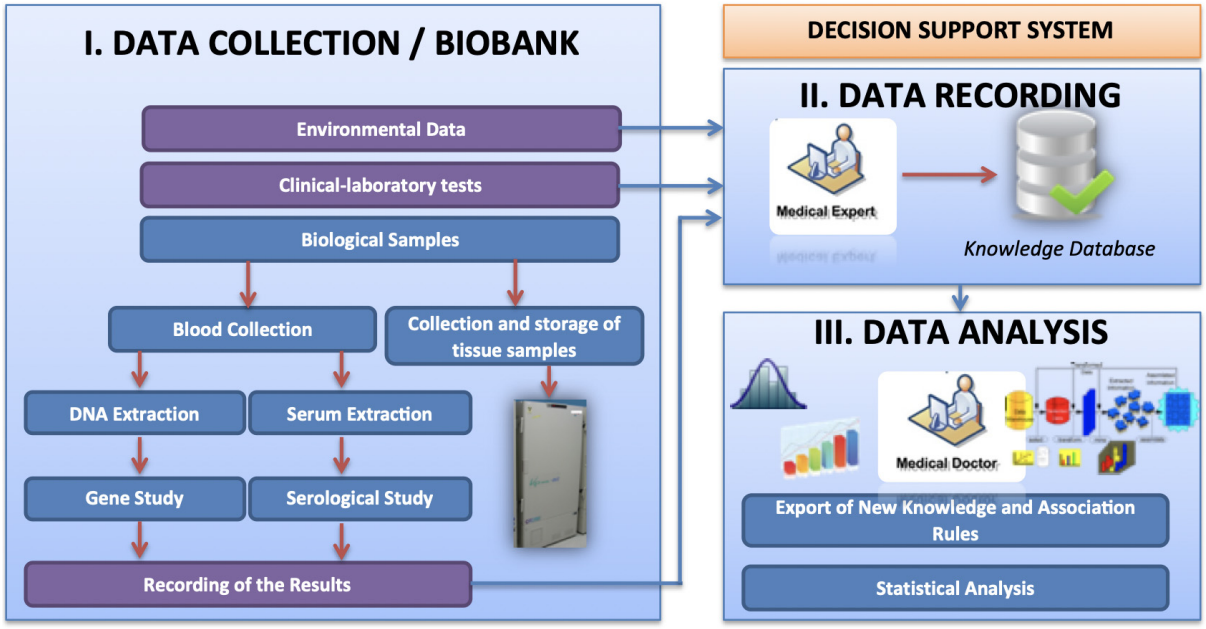


Figure 22 Decision Support System (DSS) workflow

Data Repository module: Centralized data repository for annotation data (clinical, demographic and experimental data), sample source and handling information, processing and quality assurance information, as well as inventory and process flow data.

Will provide tracking, data query, report generation, process management functions, data handling as well as statistics, data mining and knowledge extraction capability.

This module will incorporate two types of patients; existing and new.

Existing patients:

- Collection of old Patients' Medical Records
- Digitization of medical records into electronic files
- Design and implementation of Data Base System
- Copying data to the System Database

New patients:

- Collection and reporting of clinical-laboratory data from patients with IBD
- Blood sampling after signed informed consent of each patient
- Bio-samples are processed in the lab (serum/DNA extraction)
- Encoding based on the disease (Crohn's Disease (CD), Ulcerative Colitis (UC) or Indeterminate Colitis (IC)
- Storage in -80˚C freezer for further laboratory analysis (gene/serological study).
- Genotyping for various IBD susceptible genes/polymorphisms

Knowledge Discovery/Statistics module: Provide tracking, data query, report generation, process management functions, data handling as well as statistics, data mining and knowledge extraction capability.

Moreover, the module will contain a Data Representation module that will handle the presentation of the extracted knowledge from the patients' data.

Existing and extracted knowledge aims to support the physician's decision to make specific adjustments regarding the treatment plan and the understanding of clinical-laboratory interdependencies in the patient data.

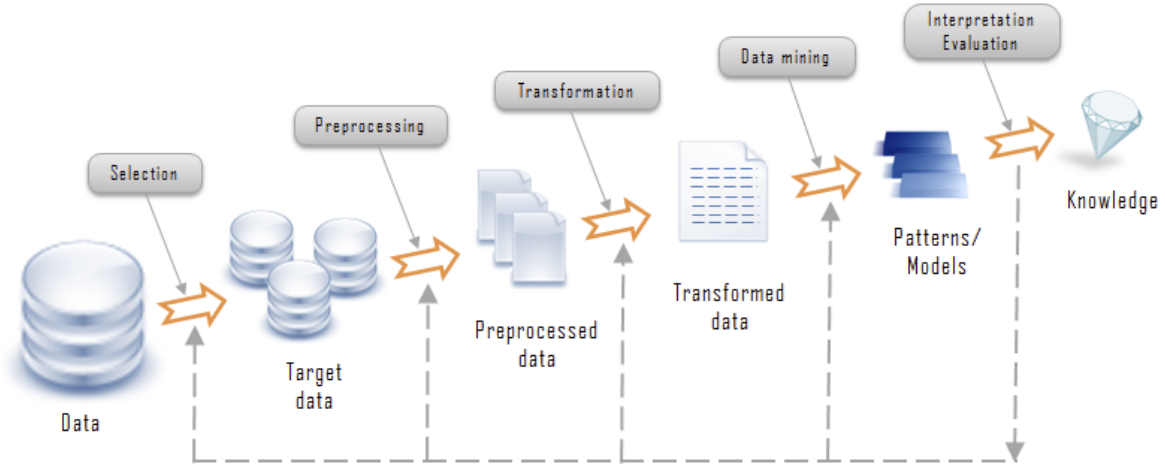


Figure 23 Knowledge discovery techniques

B3. Statistical analysis

The control group was investigated for conformity with Hardy-Weinberg equilibrium ($p > 0.05$) in all three SNPs. Hardy-Weinberg equilibrium test was performed on OEGE calculator (available at http://www.oege.org/software/hwe-mr-calc.shtml).

Allele and genotype frequencies among groups were calculated using Chi-Squared test (χ^2) or Fisher's exact test.

Association assessment of clinical, demographic and genotypic data was implemented by using regression analysis (binary logistic or linear where appropriate) and the results were expressed as odds ratio (OR) with a confidence interval of 95% (95% CI).

Phi-coefficient (Φ) was calculated for genotype-phenotype correlation tests.

For all statistical analyses, a two-tailed p-value of <0.05 was considered as statistically significant.

Statistical analysis was performed using the Jamovi software (The jamovi project (2019). Jamovi (Version 1.0) [Computer Software], retrieved from https://www.jamovi.org).

Additionally, linkage disequilibrium between rs3737240 and rs13294 of *ECM1* gene was tested using the SNPstats software (http://bioinfo.iconcologia.net).

C. Results

C1. Genotyping results

The total number of study subjects was 428, 108 CD patients, 97 UC patients and 223 healthy individuals (control group). All subjects were genotyped in order to examine possible associations of the three single nucleotide polymorphisms with IBD patients in NW Greece. The healthy control group was in Hardy-Weinberg equilibrium for all 3 SNPs (p_{value}>0.05).

Allele and genotype frequencies for the T300A *ATG16L1* polymorphism (rs2241880) are presented in Table 3; AA represents wildtype, AG heterozygotes and GG mutants.

Table 3 ATG16L1 rs2241880 allele and genotype frequencies in CD, UC and control group (Fisher's exact test, odds ratio and confidence intervals were estimated using allele frequencies in 2 × 2 contingency tables)

	Alleles				**Genotypes**				
ATG16L1 (rs2241880)	A	G	G allele freq. (%)	p [OR (95% CI)]	AA	AG	GG	GG genotype freq. (%)	p [OR (95% CI)]
CD	77	139	64.4	0.029* [1.45 (1.04-2.03)]	11	55	42	38.9	0.134 [1.48 (0.91-2.40)]
UC	80	114	58.8	0.436 [1.15 (0.82-1.61)]	14	52	31	32.0	0.733 [1.09 (0.65-1.83)]
IBD	157	253	61.7	0.061 [1.30 (0.99-1.71)]	25	107	73	35.6	0.257 [1.29 (0.86-1.93)]
Control group	199	247	55.4		43	113	67	30.0	

*p_{value} <0.05

The frequency of G allele of T300A polymorphism was 64.4%, 58.8% and 55.4% in CD, UC and healthy individuals respectively. When compared to the control group, the frequency of G allele in Crohn's disease patients was significantly higher (p = 0.029; OR = 1.45, 95% CI 1.04-2.03), while it showed no significant association with UC patients.

Correspondingly, T130M and G290S *ECM1* polymorphisms (rs3737240, rs13294) are shown in Tables 4 and 5. For T130M, CC represents wildtype, CT heterozygotes and TT mutants, while for G290S, GG represents wildtype, GA heterozygotes and AA mutants.

Table 4 ECM1 rs3737240 allele and genotype frequencies in CD, UC and control group (Fisher's exact test, odds ratio and confidence intervals were estimated using allele frequencies in 2 × 2 contingency tables)

			Alleles					**Genotypes**	
ECM1 (rs3737240)	C	T	T allele freq. (%)	p [OR (95% CI)]	CC	CT	TT	TT genotype freq. (%)	p [OR (95% CI)]
CD	116	100	46.3	0.617 [1.09 (0.79-1.51)]	26	64	18	16.7	0.550 [0.79 (0.43-1.45)]
UC	117	77	39.7	0.298 [0.83 (0.59-1.17)]	34	49	14	14.4	0.273 [0.67 (0.35-1.28)]
IBD	233	177	43.2	0.78 [0.96 (0.73-1.26)]	60	113	32	15.6	0.257 [0.73 (0.44-1.21)]
Control group	249	197	44.2		71	107	45	20.2	

Table 5 ECM1 rs13294 allele and genotype frequencies in CD, UC and control group (Fisher's exact test, odds ratio and confidence intervals were estimated using allele frequencies in 2 × 2 contingency tables)

			Alleles					**Genotypes**	
ECM1 (rs13294)	G	A	A allele freq. (%)	p [OR (95% CI)]	GG	GA	AA	AA genotype freq. (%)	p [OR (95% CI)]
CD	116	100	46.3	0.505 [1.13 (0.81-1.57)]	26	64	18	16.7	0.550 [0.79 (0.43-1.45)]
UC	120	74	38.2	0.257 [0.81 (0.57-1.14)]	36	48	13	13.4	0.159 [0.61 (0.31-1.20)]
IBD	236	174	42.4	0.836 [0.97 (0.74-1.27)]	62	112	31	15.1	0.205 [0.71 (0.43-1.17)]
Control group	253	193	43.3		75	103	45	20.2	

The frequency of T allele of T130M was 46.3%, 39.7% and 44.2% for CD, UC and control group respectively, while the frequency of A allele of G290S mutation was 46.3%, 38.2% and 43.3% respectively. No strong associations between either of the two SNPs of *ECM1* gene and our study group were found (p_{value} >0.05).

Furthermore, a potential additive effect of the studied alleles of the three SNPs was tested. Investigating the additive effect of G allele of the T300A SNP (rs2241880, *ATG16L1*) in CD patients, we found that carriers of two G alleles (mutant group) compared to those carrying only one G allele (heterozygotes), were 1.3 times more susceptible to CD (Table 6), which was statistically significant (GG: p = 0.022; OR: 2.450; 95% CI: 1.14-5.27, AG: p = 0.087; OR: 1.903; 95% CI: 0.91-3.97).

Table 6 Additive effect of G allele, ATG16L1; T300A (Binary logistic regression analyses)

	p_{value}	Odds Ratio	95% CI
Crohn's disease			
AG	0.087	1.90	0.91-3.97
GG	0.022*	2.45	1.14-5.27

*p_{value} <0.05

No additive effect of either of the two *ECM1* SNPs was found.

Furthermore, comparison of the two IBD groups (CD and UC) with each other, from a genotypic point of view, showed no significant difference for any of the 3 SNPs (p = 0.290; OR: 1.27; 95% CI: 0.85-1.89 for T300A, p = 0.195; OR: 1.31; 95% CI: 0.88-1.94 for T130M and p = 0.109; OR: 1.40; 95% CI: 0.94-2.07 for G290S).

In addition, we tested for potential associations between genotype and disease phenotype or certain clinical features.

In CD patients, the presence of one or two G alleles (AG+GG genotypes) of the T300A polymorphism (rs2241880, *ATG16L1*), indicated a possible protective effect against developing a penetrating phenotype (B3 behavior according to Montreal classification[8]) with p = 0.015; OR: 0.20, 95% CI: 0.05-0.74, while in UC patients, presence of one or two G alleles (AG+GG genotypes) of the T300A polymorphism (rs2241880, *ATG16L1*), indicated a possible protective effect against developing joint-involving EIMs, with p = 0.038; OR: 0.31, 95% CI: 0.10-0.97 (Table 5). However, when measure analyses of these associations performed, by using phi-coefficient (Φ) test, we found that these findings are of mild association (Φ = 0.251 and Φ = 0.211 respectively). Furthermore, in CD patients carrying T300A SNP (AG+GG genotype), we found an indication of a possible protective effect against the need of cholecystectomy (p = 0.022; OR: 0.12, 95% CI: 0.02–0.60), with a mild to moderate association (Φ = 0.284) of this finding though. However, the number of CD patients who underwent cholecystectomy (post-diagnosis) was small (n=7). Genotype-phenotype association results are shown in Table 7.

Table 7 ATG16L1 T300A Genotype and phenotype associations (Fisher's exact test and correlation test)

ATG16L1 T300A (rs2241880)	Crohn's Disease			Ulcerative Colitis		
	p_{value}	OR (95% CI)	Φ	p_{value}	OR (95% CI)	Φ
B3 - Penetrating behavior	0.015	0.20 (0.05-0.74)	0.251	-	-	-
Cholecystectomy	0.022	0.12 (0.02 – 0.60)	0.284	-	-	-
Joint-involving EIMs	-	-	-	0.038	0.31 (0.10-0.97)	0.211

p_{value}: AG+GG vs AA, Φ: Phi-coefficient

No association was found between the age at onset, CD location, UC extent and severity, presence of EIMs or other immune disease, need of operation, anti-TNFα treatment, appendicectomy, tonsillectomy and any of the three SNPs for either CD or UC patients (Table 8).

Table 8 Genotype - phenotype associations

	Crohn's Disease						Ulcerative Colitis					
	ATG16L1 (rs2241880)		ECM1 (rs3737240)		ECM1 (rs13294)		ATG16L1 (rs2241880)		ECM1 (rs3737240)		ECM1 (rs13294)	
	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)
Age at onset	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
CD Location	NS	NS	NS	NS	NS	NS	-	-	-	-	-	-
Ileal involvement	NS	NS	NS	NS	NS	NS	-	-	-	-	-	-
B3 – Penetrating behaviour	0.015	0.20 (0.05-0.74)	NS	NS	NS	NS	-	-	-	-	-	-
UC Extent	-	-	-	-	-	-	NS	NS	NS	NS	NS	NS
UC Severity	-	-	-	-	-	-	NS	NS	NS	NS	NS	NS
Joint-involving EIMs	NS	NS	NS	NS	NS	NS	0.038	0.31 (0.10-0.97)	NS	NS	NS	NS
EIMs	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Other Autoimmune Disease	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Cholecystectomy	0.022	0.12 (0.02 – 0.60)	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Tonsillectomy	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Appendectomy	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Operated	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
anti-TNFα	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

NS: not significant

Finally, the two polymorphisms of *ECM1* gene were not found to be in linkage disequilibrium (Table 9).

Table 9 ECM1 gene linkage disequilibrium study results

Linkage disequilibrium analysis

D statistic

	rs3737240	rs13294
rs3737240	.	0.2413
rs13294	.	.

D' statistic

	rs3737240	rs13294
rs3737240	.	0.9997
rs13294	.	.

r statistic

	rs3737240	rs13294
rs3737240	.	0.9832
rs13294	.	.

P-values

	rs3737240	rs13294
rs3737240	.	0
rs13294	.	.

Haplotype analysis

Haplotype frequencies estimation (n=428)

	rs3737240	rs13294	Total	group.0	group.1	Cumulative frequency
1	C	G	0.5631	0.5583	0.5683	0.5631
2	T	A	0.4287	0.4327	0.4244	0.9918
3	T	G	0.0082	0.009	0.0073	1

Haplotype association with response (n=428, crude analysis)

	rs3737240	rs13294	Freq	OR (95% CI)	P-value
1	C	G	0.5631	1.00	---
2	T	A	0.4287	0.96 (0.73 - 1.27)	0.77
rare	*	*	0.0082	0.79 (0.17 - 3.62)	0.76

Global haplotype association p-value: 0.93

C2. Decision support system (DSS) results

Expected results of the DSS are defined as interactions between the user (i.e. physician) and the system.

Possible interactions between the user and the system:

- Create a predictive bioinformatic model for the disease's diagnosis and course with clinical and laboratory data.
- Refine reasons (i.e. genetic, serologic, histologic, diet) for differences in IBD incidence and severity of course in our area compared to other European areas.
- Referral clinical centre and Biobank of excellence in the Western Balkans

The DSS will represent a unique multidisciplinary combined database, which will include genetic, serological, histological, clinical and environmental data of a homogeneous population in a computerized bioinformatics manner. It will be possible to study in combination of clinical, environmental and laboratory data. The center will be possible to collaborate and contribute valuable data to many international leading IBD projects. The computerized Bio-Database will serve as a model and a basis for research in other chronic diseases in our well-defined population. This project will enrich our experience and will contribute towards a better IBD education and training in our medical, nursing and laboratory personnel. This will clearly impact the quality of care and quality of life of our IBD patients. Finally, as IBD is regarded to be a multifactorial disease we hope to better define some factors that clearly predispose to certain IBD phenotypes and IBD disease course.

Possible scenarios of the user-system interaction:

- Shows the probability of a surgery need
- Displays the most appropriate treatment for each individual patient
- Recommends modifications/ changes of the treatment plan for an individual patient
- Shows probability of exacerbations in the disease process
- Show probability of recession of the disease
- Displays the possible extent of the disease
- Shows the probability of occurrence of extra-intestinal manifestations (EIMs)

D. Discussion – Conclusion

Inflammatory Bowel Disease is affecting millions worldwide. Even areas where the disease used to be of low occurrence, now are having a spurt and the pattern is changing dynamically; possibly due to a more "westernized" type of living. The etiopathology of IBD still remains unknown. Scientists across the globe unanimously suggest that a fine interaction among genetic susceptibility, environmental factors, immune response and gut microbiota may hold the answer. However, studies have shown that a globally pattern of genetic influence, heritability, environmental triggers and gut fauna does not exist. In order to investigate that though, all available data would offer a great impact on elucidating the true face of IBD; thus, studies from various areas and ethnicities are needed.

This is the first study of *ATG16L1* and *ECM1* genes and their polymorphisms in the Northwest Greece region, a previously well described sheltered area[3]. In a previous genotypic study from this area focusing on the *NOD2/CARD15* gene it was shown that no association with the studied polymorphisms and CD susceptibility[122] exists, a divergent result compared to most studies from other areas[123,124]. Furthermore, the role of T300A (rs2241880, *ATG16L1*) and T130M (rs3737240, *ECM1*) and G290S (rs13294, *ECM1*) polymorphisms in the development of CD and UC respectively, is well established in the bibliography[65,101], despite some ethnic variabilities[106,112]. Our study replicates that the T300A (rs2241880, *ATG16L1*) polymorphism predisposes to CD in our cohort with also an additive effect of G allele in CD patients (individuals carrying two copies of G allele are 1.3 times more susceptible to CD compared to those carrying only one copy), but failed to demonstrate any association of *ECM1* gene's polymorphisms with UC susceptibility.

On the other hand, despite the strong occurrence of the G allele in the CD group compared to the control group, a clear distinction among IBD patients (CD vs UC analysis) could not be established. Similarly, as mentioned earlier, *ECM1* gene's SNP investigation in our population failed to replicate existing data[65,121] and, again, no distinction among IBD patients was found. Hence, in our study group, it is not possible to differentiate the underling disease (CD or UC) based on genotypic-phenotypic associations, probably due to ethnic variations, but more patients and more widely associated susceptibility genes are needed to drive to a definite conclusion.

On the contrary, we found some interesting protective associations of certain phenotypes and the IBD patients. In the CD group, patients who are carriers (AG and GG patients) of G allele of T300A (rs2241880, *ATG16L1*) polymorphism were found to associate with a possible protective effect against penetrating behavior (B3 phenotype according to the Montreal classification[8]), a finding that disagrees with another study's results where the majority of the patients were found to have a penetrating behavior[94], suggesting that a potential environmental or ethnicity trigger may be present in our cohort playing a role in developing such a phenotype. Again, in the CD group, carriers of the G allele (AG and GG patients) of the T300A (rs2241880, *ATG16L1*) polymorphism were found to have a mild association with a protective effect against the need of a cholecystectomy. In the bibliography, gallbladder disease is well described in IBD patients and is mainly associated with Crohn's disease[125], though a recent metanalyses of Zhang et al[126] concluded that despite the apparent association of CD and gallbladder disease, other factors such as CD location, number of relapses and ileal surgery were identified as independent variables for developing cholelithiasis, but more studies are required for a definite answer. Moreover, a protective effect of G allele (T300A polymorphism) against joint-involving EIMs in UC patient was found. As have been described in the past, by *Christodoulou et al*[12], and this study confirms, EIMs are not rare in our IBD cohort (62% of CD patients and 47.4% of UC patients, in this study) and data from other studies as well suggest that a close genetic correlation between IBD and EIMs does exist[127]. However, such associations between IBD susceptibility genes and EIMs occurrence could not be demonstrated in the presented study.

When other clinical data (age at onset, CD location, CD behavior, UC extent, UC severity, need of therapeutic operation, anti-TNFα therapy) were analyzed for any possible linkage with the aforementioned SNPs, no significant associations were drawn while in the bibliography such associations exist[94,105,128–130].

Regarding the decision support system, we aimed at developing a prognostic tool that, by incorporating clinical, demographic and research data, would help understand the natural course of the disease, study the predisposing factors and related genes and determine early clinical, genetic and immunological predictors of outcome and response to treatment, as well as help clinicians to better evaluate and treat IBD patients. Building such a system will contribute even more to IBD knowledge and research and hopefully lead to a more personalized type of medicine, bearing in mind that there are no diseases, but only patients. Such efforts are currently an ongoing trend on a worldwide level[109,131].

To conclude, as was shown by *Tsianos et al* in 2003[4], CD is less frequent than UC in our study area, the area of North-western Greece. Thus, the findings of our current study, concerning the significant association of T300A polymorphism with CD susceptibility, point to a strong genetic background which plays a crucial role in CD occurrence to our population, and an additive effect of T300A G allele, though further investigation including more patients and more susceptibility genes will provide a better understanding.

E. Abstract

Crohn's Disease (CD) and Ulcerative Colitis (UC), are well described disease entities with unknown etiopathogenesis affecting millions worldwide. Environmental, genetic, gut microbiota and host immune response correlations have been implicated. Genetic susceptibility across different geographic areas and ethnicities varies significantly. Northwestern Greece is a well-defined geographic area with a very high homogeneity of the population, thus a strong genetic background is implicated.

The role of susceptibility gene polymorphisms, such as *ATG16L1* T300A (rs2241880) and *ECM1* T130M (rs3737240) and G290S (rs13294), is well described, although controversial findings have been reported.

In our study, two hundred and five unrelated IBD patients (108 CD patients and 97 UC patients), and 223 healthy unrelated blood donors (control group) from the Northwest Greece area, were genotyped for rs2241880 (T300A), rs3737240 (T130M) and rs13294 (G290S) single nucleotide polymorphisms. Genotyping was performed with Real-Time PCR.

Our results suggest that the frequency of G allele (of the T300A polymorphism) in CD patients, compared to the control group, was significantly higher ($p = 0.029$; OR = 1.45, 95% CI 1.04-2.03). Carriers of two G alleles (T300A), compared to those carrying only one, were 1.3 times more susceptible to CD ($p = 0.022$; OR: 2.450; 95% CI: 1.14-5.27), implying an additive effect of G allele. In CD patients, presence of the T300A polymorphism, showed a protective effect against developing a penetrating phenotype ($p = 0.015$; OR: 0.20, 95% CI: 0.05-0.74) or needing cholecystectomy ($p = 0.022$; OR: 0.12, 95% CI: 0.02–0.60). In UC patients, presence of the T300A polymorphism, was protective against developing joint-involving EIMs ($p = 0.038$; OR: 0.31, 95% CI: 0.10-0.97). No association of the SNPs of the *ECM1* gene and UC patients was found in our study.

To conclude, our study, concerning the significant association of T300A polymorphism with CD susceptibility, imposes a strong genetic background in CD occurrence to our population, and also an additive effect of T300A G allele.

F. Εκτεταμένη Περίληψη

Η Ιδιοπαθής Φλεγμονώδης Πάθηση των Εντέρων (ΙΦΠΕ: Νόσος Crohn και Ελκώδης Κολίτιδα) σαν παράδειγμα χρόνιου νοσήματος είναι ομάδα παθήσεων αδιευκρίνιστης αιτιολογίας η οποία έχει περιγραφεί από τους Ιπποκρατικούς χρόνους. Οι πάσχοντες παγκοσμίως ξεπερνούν τα 10,000,000 και στην Ελλάδα τους 10,000 με μεγάλο κόστος στις υπηρεσίες υγείας και φανερή αντανάκλαση σε παραμέτρους ποιότητας ζωής, καθώς η ΙΦΠΕ συνοδεύεται από σημαντικά σωματικά συμπτώματα (κοιλιακό άλγος, διάρροιες, αίσθημα ακράτειας, αποβολή αίματος από το ορθό κ.ά.), αλλά και ψυχικά συμπτώματα με κοινωνικές επιπτώσεις.

Η πάθηση διαδράμει με εξάρσεις και υφέσεις και συχνά οδηγεί σε παρατεταμένες νοσηλείες. Προσβάλλει κυρίως το λεπτό και το παχύ έντερο σε ασθενείς οποιασδήποτε ηλικίας, κυρίως όμως άτομα νεαρής παραγωγικής ηλικίας με σημαντική νοσηρότητα και, σε ορισμένες περιπτώσεις, θνητότητα. Σύμφωνα με πρόσφατες επιδημιολογικές μελέτες από την περιοχή της Ηπείρου, ο αριθμός των ασθενών αυτών στη ΒΔ Ελλάδα συνεχώς αυξάνεται υποδηλώνοντας μια σύνθετη αλληλεπίδραση περιβαλλοντικών και γενετικών παραγόντων.

Τα τελευταία χρόνια, η έρευνα όσον αφορά την αιτιοπαθογένεια της ΙΦΠΕ έχει αυξηθεί κατακόρυφα καθώς το πρόβλημα είναι συνεχώς αυξανόμενο και θεραπευτικά περίπλοκο. Ιδιαίτερα ενδιαφέρον είναι ότι υπάρχει ποικιλότητα της επίδρασης γενετικών παραγόντων καθώς κάθε φυλή, αλλά και κάθε καλά προσδιορισμένη γεωγραφικά πληθυσμιακή ομάδα, φαίνεται να έχει ιδιαίτερα γονιδιακά χαρακτηριστικά που σχετίζονται με την πάθηση.

Η Ήπειρος, και η ΒΔ Ελλάδα γενικότερα, αποτελούν μία τέτοια καλά προσδιορισμένη περιοχή, όπου ο αριθμός των ασθενών είναι αρκετά μεγάλος και παρουσιάζει αυξητική τάση, αν και η συχνότητα των ασθενών με νόσο Crohn (NC) παραμένει πολύ μικρή. Η παρατήρηση αυτή θα μπορούσε να αποδοθεί είτε σε κάποιο περιβαλλοντικό παράγοντα που πιθανόν ελλείπει από την περιοχή είτε σε κάποιο ειδικότερο, γενετικό παράγοντα που περιορίζει τη συχνότητα της NC.

Η μελέτη γενετικών τόπων που θα μπορούσαν να ενέχονται στην αιτιοπαθογένεια της ΙΦΠΕ έχει εντατικοποιηθεί χωρίς όμως προς το παρόν να έχει δειχθεί ο γενετικός τόπος ή τόποι που ενέχονται άμεσα. Παρ' όλα αυτά ορισμένες μεταλλάξεις (πολυμορφισμοί) γονιδίων δείχνουν ότι διαδραματίζουν σημαντικό ρόλο στην τελική έκφραση και βαρύτητα της πάθησης αλλά και στην πρόβλεψη ανταπόκρισης στη θεραπεία.

Πρόσφατες μελέτες ανάλυσης γονιδιώματος (Genome Wide Association Studies, GWAS) ανέδειξαν πάνω από 200 γενετικούς τόπους (*loci*) που σχετίζονται με την αιτιοπαθογένεια της ΙΦΠΕ, με τα 30 περίπου να είναι κοινά και στις δύο νόσους (NC, ΕΚ), ενώ ορισμένα συνδέονται ειδικά με την μία εκ των δύο (όπως το *ATG16L1* με την NC και το *ECM1* με την ΕΚ) και θα μπορούσαν να αξιοποιηθούν για την διαφοροδιάγνωση. Εν γένει, τα γονίδια αυτά θα μπορούσαν να προσφέρουν τη δυνατότητα έγκαιρης διάγνωσης ή/και πρόγνωσης της ΙΦΠΕ, υπό την προϋπόθεση ότι θα έχει προηγηθεί μια συστηματική ανάλυση του γενετικού υποβάθρου των ασθενών της αντίστοιχης γεωγραφικής περιοχής.

Στόχος της παρούσας διδακτορικής διατριβής ήταν να μελετηθεί ο πληθυσμός της ΒΔ Ελλάδας που πάσχει από ΙΦΠΕ, ως προς το γενετικό προφίλ της νόσου, χρησιμοποιώντας αναλυτικές τεχνικές προσδιορισμού γονιδιακών πολυμορφισμών. Συγκεκριμένα, έγινε ανάλυση, με RT-PCR, των γονιδιακών πολυμορφισμών rs2241880 (T300A) του γονιδίου *ATG16L1* και των rs3737240 (T130M) και rs13294 (G290S) του γονιδίου *ECM1* σε ασθενείς με ΙΦΠΕ (ΕΚ, NC) και σύγκρισή τους με δείγματα υγιών μαρτύρων (controls).

Συγκεκριμένα, μελετήθηκαν 205 ασθενείς με ΙΦΠΕ (108 με NC και 97 με ΕΚ) και 223 υγιείς μάρτυρες από την περιοχή της ΒΔ Ελλάδος. Τα αποτελέσματα μας υποδεικνύουν ότι η συχνότητα του G αλληλίου (του T300A πολυμορφισμού) στους ασθενείς με NC, συγκρινόμενη με αυτή των controls, ήταν σημαντικά υψηλότερη ($p = 0.029$; OR = 1.45, 95% CI 1.04-2.03). Επίσης, οι ασθενείς που ήταν φορείς δύο G αλληλίων (του T300A πολυμορφισμού), συγκρινόμενοι με αυτούς που ήταν μόνο ενός, ήταν 1.3 φορές ποιο επιδεκτικοί στο να αναπτύξουν NC ($p = 0.022$; OR: 2.450; 95% CI: 1.14-5.27), υποδηλώνοντας μία προσθετική επίδραση του G αλληλίου. Ακόμα, στους ασθενείς με NC, η παρουσία του T300A πολυμορφισμού, φάνηκε να επιφέρει μία προστατευτική επίδραση όσον αφορά την ανάπτυξη διεισδυτικού φαινοτύπου της νόσου ($p = 0.015$; OR: 0.20, 95% CI: 0.05-0.74), καθώς και στην ανάγκη για χολοκυστεκτομή. Στους ασθενείς με ΕΚ, η παρουσία του T300A πολυμορφισμού, φάνηκε να έχει προστατευτική επίδραση όσον αφορά την ανάπτυξη εξωεντερικών εκδηλώσεων από τις αρθρώσεις ($p = 0.038$; OR: 0.31, 95% CI: 0.10-0.97). Δεν βρέθηκε καμία συσχέτιση στον πληθυσμό μας μεταξύ των πολυμορφισμών του *ECM1* γονιδίου και των ασθενών με ΕΚ.

Κλείνοντας, η παρούσα μελέτη, λαμβάνοντας υπόψιν το σημαντικό συσχετισμό του T300A πολυμορφισμού με τη NC, υπαγορεύει ένα ισχυρό γενετικό υπόβαθρο στην εμφάνιση της NC στον πληθυσμό μελέτης μας καθώς και μία προσθετική επίδραση του G αλληλίου του T300A πολυμορφισμού.

G. References

1 Burisch J, Jess T, Martinato M, Lakatos PL. The burden of inflammatory bowel disease in Europe. *Journal of Crohn's and Colitis* 2013; **7**: 322–337.

2 Cosnes J, Gower–Rousseau C, Seksik P, Cortot A. Epidemiology and Natural History of Inflammatory Bowel Diseases. *Gastroenterology* 2011; **140**: 1785-1794.e4.

3 Tsianos EV, Masalas CN, Merkouropoulos M, Dalekos GN, Logan RF. Incidence of inflammatory bowel disease in north west Greece: rarity of Crohn's disease in an area where ulcerative colitis is common. *Gut* 1994; **35**: 369–372.

4 Tsianos EV, Katsanos KH, Christodoulou D, Dimoliatis I, Kogevinas A, Logan RF. Continuing low incidence of Crohn's disease in Northwest Greece. *Digestive and Liver Disease* 2003; **35**: 99–103.

5 Burisch J, Munkholm P. Inflammatory bowel disease epidemiology: *Current Opinion in Gastroenterology* 2013; **29**: 357–362.

6 Gearry RB, Richardson AK, Frampton CM, Dodgshun AJ, Barclay ML. Population-based cases control study of inflammatory bowel disease risk factors. *Journal of Gastroenterology and Hepatology* 2010; **25**: 325–333.

7 Baumgart DC, Sandborn WJ. Inflammatory bowel disease: clinical aspects and established and evolving therapies. *The Lancet* 2007; **369**: 1641–1657.

8 Satsangi J. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut* 2006; **55**: 749–753.

9 Schroeder KW, Tremaine WJ, Ilstrup DM. Coated Oral 5-Aminosalicylic Acid Therapy for Mildly to Moderately Active Ulcerative Colitis. *New England Journal of Medicine* 1987; **317**: 1625–1629.

10 Fakhoury M, Al-Salami H, Negrulj R, Mooranian A. Inflammatory bowel disease: clinical aspects and treatments. *Journal of Inflammation Research* 2014; : 113.

11 Levine JS, Burakoff R. Extraintestinal manifestations of inflammatory bowel disease. *Gastroenterol Hepatol (N Y)* 2011; **7**: 235–241.

12 Christodoulou DK, Katsanos KH, Kitsanou M, Stergiopoulou C, Hatzis J, Tsianos EV. Frequency of extraintestinal manifestations in patients with inflammatory bowel disease in Northwest Greece and review of the literature. *Digestive and Liver Disease* 2002; **34**: 781–786.

13 Fiocchi C. Inflammatory Bowel Disease: Etiology and Pathogenesis. 1998; **115**: 24.

14 Loddo I, Romano C. Inflammatory Bowel Disease: Genetics, Epigenetics, and Pathogenesis. *Frontiers in Immunology* 2015; **6**. doi:10.3389/fimmu.2015.00551.

15 Boyapati R, Satsangi J, Tzer Ho G. Pathogenesis of Crohn's disease. *F1000Prime Reports* 2015; **7**. doi:10.12703/P7-44.

16 Gardet A, Xavier RJ. Common alleles that influence autophagy and the risk for inflammatory bowel disease. *Current Opinion in Immunology* 2012; **24**: 522–529.

17 Hanauer SB. Update on the etiology, pathogenesis and diagnosis of ulcerative colitis. *Nature Clinical Practice Gastroenterology & Hepatology* 2004; **1**: 26–31.

18 Hanauer SB. Inflammatory bowel disease: Epidemiology, pathogenesis, and therapeutic opportunities: *Inflammatory Bowel Diseases* 2006; **12**: S3–S9.

19 Zhang Y-Z. Inflammatory bowel disease: Pathogenesis. *World Journal of Gastroenterology* 2014; **20**: 91.

20 Renz H, von Mutius E, Brandtzaeg P, Cookson WO, Autenrieth IB, Haller D. Gene-environment interactions in chronic inflammatory disease. *Nature Immunology* 2011; **12**: 273–277.

21 Russell RK, Satsangi J. IBD: a family affair. *Best Practice & Research Clinical Gastroenterology* 2004; **18**: 525–539.

22 Peeters M, Cortot A, Vermeire S, Colombel J-F. Familial and sporadic inflammatory bowel disease: Different entities? *Inflammatory Bowel Diseases* 2000; **6**: 314–320.

23 Halme L. Family and twin studies in inflammatory bowel disease. *World Journal of Gastroenterology* 2006; **12**: 3668.

24 Ananthakrishnan AN. Epidemiology and risk factors for IBD. *Nature Reviews Gastroenterology & Hepatology* 2015; **12**: 205–217.

25 M. Orholm, V. Binder, T. I. A. Søre. Concordance of Inflammatory Bowel Disease among Danish Twins: Results of a Nationwide Study. *Scandinavian Journal of Gastroenterology* 2000; **35**: 1075–1081.

26 Cho JH, Brant SR. Recent Insights Into the Genetics of Inflammatory Bowel Disease. *Gastroenterology* 2011; **140**: 1704-1712.e2.

27 Burisch J, Munkholm P. The epidemiology of inflammatory bowel disease. *Scandinavian Journal of Gastroenterology* 2015; **50**: 942–951.

28 Bonen D. The genetics of inflammatory bowel disease. *Gastroenterology* 2003; **124**: 521–536.

29 Rivas MA, Avila BE, Koskela J, Huang H, Stevens C, Pirinen M *et al.* Insights into the genetic epidemiology of Crohn's and rare diseases in the Ashkenazi Jewish population. *PLOS Genetics* 2018; **14**: e1007329.

30 de Lange KM, Moutsianas L, Lee JC, Lamb CA, Luo Y, Kennedy NA *et al.* Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nature Genetics* 2017; **49**: 256–261.

31 Luo Y, de Lange KM, Jostins L, Moutsianas L, Randall J, Kennedy NA *et al.* Exploring the genetic architecture of inflammatory bowel disease by whole-genome sequencing identifies association at ADCY7. *Nature Genetics* 2017; **49**: 186–192.

32 UK IBD Genetics Consortium, Lee JC, Biasci D, Roberts R, Gearry RB, Mansfield JC *et al.* Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn's disease. *Nature Genetics* 2017; **49**: 262–268.

33 Hugot J-P, Chamaillard M, Zouali H, Lesage S, Cézard J-P, Belaiche J *et al.* Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001; **411**: 599–603.

34 Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R *et al.* A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 2001; **411**: 603–606.

35 Hugot J-P. CARD15/NOD2 Mutations in Crohn's Disease. *Annals of the New York Academy of Sciences* 2006; **1072**: 9–18.

36 de Souza HSP, Fiocchi C, Iliopoulos D. The IBD interactome: an integrated view of aetiology, pathogenesis and therapy. *Nature Reviews Gastroenterology & Hepatology* 2017; **14**: 739–749.

37 Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. *Nature* 2011; **474**: 307–317.

38 Hugot J-P, Laurent-Puig P, Gower-Rousseau C, Olson JM, Lee JC, Beaugerie L *et al.* Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature* 1996; **379**: 821–823.

39 Van Limbergen J, Radford-Smith G, Satsangi J. Advances in IBD genetics. *Nature Reviews Gastroenterology & Hepatology* 2014; **11**: 372–385.

40 Shih DQ, Targan SR. Insights into IBD pathogenesis. *Current Gastroenterology Reports* 2009; **11**: 473.

41 Robles AI, Traverso G, Zhang M, Roberts NJ, Khan MA, Joseph C *et al.* Whole-Exome Sequencing Analyses of Inflammatory Bowel Disease−Associated Colorectal Cancers. *Gastroenterology* 2016; **150**: 931–943.

42 Worthey EA, Mayer AN, Syverson GD, Helbling D, Bonacci BB, Decker B *et al.* Making a definitive diagnosis: Successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genetics in Medicine* 2011; **13**: 255–262.

43 Molodecky NA, Kaplan GG. Environmental Risk Factors for Inflammatory Bowel Disease. ; : 8.

44 Ananthakrishnan AN, Bernstein CN, Iliopoulos D, Macpherson A, Neurath MF, Ali RAR *et al.* Environmental triggers in IBD: a review of progress and evidence. *Nature Reviews Gastroenterology & Hepatology* 2018; **15**: 39–49.

45 Cosnes J. Tobacco and IBD: relevance in the understanding of disease mechanisms and clinical practice. *Best Practice & Research Clinical Gastroenterology* 2004; **18**: 481–496.

46 Cosnes J, Beaugerie L, Carbonnel F, Gendre J. Smoking cessation and the course of Crohn's disease: An intervention study. *Gastroenterology* 2001; **120**: 1093–1099.

47 Savin Z, Kivity S, Yonath H, Yehuda S. Smoking and the intestinal microbiome. *Archives of Microbiology* 2018; **200**: 677–684.

48 Wu GD, Bushmanc FD, Lewis JD. Diet, the human gut microbiota, and IBD. *Anaerobe* 2013; **24**: 117–120.

49 Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA *et al.* Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. *Science* 2011; **334**: 105–108.

50 Griffiths AM, Ohlsson A, Sherman PM, Sutherland LR. Meta-analysis of enteral nutrition as a primary treatment of active crohn's disease. *Gastroenterology* 1995; **108**: 1056–1067.

51 Ohkusa T, Nomura T, Sato N. The Role of Bacterial Infection in the Pathogenesis of Inflammatory Bowel Disease. *Internal Medicine* 2004; **43**: 534–539.

52 Irving PM, Gibson PR. Infections and IBD. *Nature Clinical Practice Gastroenterology & Hepatology* 2008; **5**: 18–27.

53 Ananthakrishnan AN, Higuchi LM, Huang ES, Khalili H, Richter JM, Fuchs CS *et al.* Aspirin, nonsteroidal anti-inflammatory drug use, and risk for Crohn disease and ulcerative colitis: a cohort study. *Ann Intern Med* 2012; **156**: 350–359.

54 Shaw SY, Blanchard JF, Bernstein CN. Association Between the Use of Antibiotics in the First Year of Life and Pediatric Inflammatory Bowel Disease: *American Journal of Gastroenterology* 2010; **105**: 2687–2692.

55 Cornish JA, Tan E, Simillis C, Clark SK, Teare J, Tekkis PP. The Risk of Oral Contraceptives in the Etiology of Inflammatory Bowel Disease: A Meta-Analysis. *The American Journal of Gastroenterology* 2008; **103**: 2394–2400.

56 Goodhand J, Rampton D. Psychological stress and coping in IBD. *Gut* 2008; **57**: 1345–1347.

57 Timmer A, Preiss JC, Motschall E, Rücker G, Jantschek G, Moser G. Psychological interventions for treatment of inflammatory bowel disease. *Cochrane Database of Systematic Reviews* 2011. doi: 10.1002/14651858.CD006913.pub2.

58 Rutgeerts P, D'Haens G, Hiele M, Geboes K, Vantrappen G. Appendectomy protects against ulcerative colitis. *Gastroenterology* 1994; **106**: 1251–1253.

59 de Souza HSP, Fiocchi C. Immunopathogenesis of IBD: current state of the art. *Nature Reviews Gastroenterology & Hepatology* 2016; **13**: 13–27.

60 Holleran G, Lopetuso L, Petito V, Graziani C, Ianiro G, McNamara D *et al.* The Innate and Adaptive Immune System as Targets for Biologic Therapies in Inflammatory Bowel Disease. *International Journal of Molecular Sciences* 2017; **18**: 2020.

61 Romagnani S. Lymphokine Production by Human T Cells in Disease States. *Annual Review of Immunology* 1994; **12**: 227–257.

62 Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ *et al.* A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 2006; **314**: 1461–1463.

63 Huang Y, Chen Z. Inflammatory bowel disease related innate immunity and adaptive immunity. *Am J Transl Res* 2016; **8**: 2490–2497.

64 Geremia A, Biancheri P, Allan P, Corazza GR, Di Sabatino A. Innate and adaptive immunity in inflammatory bowel disease. *Autoimmunity Reviews* 2014; **13**: 3–10.

65 the Wellcome Trust Case Control Consortium, Fisher SA, Tremelling M, Anderson CA, Gwilliam R, Bumpstead S *et al.* Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. *Nature Genetics* 2008; **40**: 710–712.

66 Fritz T, Niederreiter L, Adolph T, Blumberg RS, Kaser A. Crohn's disease: NOD2, autophagy and ER stress converge. *Gut* 2011; **60**: 1580–1588.

67 Zhang M, Sun K, Wu Y, Yang Y, Tso P, Wu Z. Interactions between Intestinal Microbiota and Host Immune Response in Inflammatory Bowel Disease. *Frontiers in Immunology* 2017; **8**. doi:10.3389/fimmu.2017.00942.

68 Turpin W, Goethel A, Bedrani L, Croitoru, MDCM K. Determinants of IBD Heritability: Genes, Bugs, and More. *Inflammatory Bowel Diseases* 2018; **24**: 1133–1148.

69 Bhan AK, Mizoguthi E, Smith RN, Mizoguchi A. Colitis in transgenic and knockout animals as models of human inflammatory bowel disease. *Immunological Reviews* 1999; **169**: 195–207.

70 Dignass A, Lindsay JO, Sturm A, Windsor A, Colombel J-F, Allez M *et al.* Second European evidence-based consensus on the diagnosis and management of ulcerative colitis Part 2: Current management. *Journal of Crohn's and Colitis* 2012; **6**: 991–1030.

71 Dignass A, Van Assche G, Lindsay JO, Lémann M, Söderholm J, Colombel JF *et al.* The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Current management. *Journal of Crohn's and Colitis* 2010; **4**: 28–62.

72 Harbord M, Eliakim R, Bettenworth D, Karmiris K, Katsanos K, Kopylov U *et al.* Third European Evidence-based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 2: Current Management. *Journal of Crohn's and Colitis* 2017; **11**: 769–784.

73 Koutroubakis IE, Future Medicine Ltd. *Advances in the management of inflammatory bowel disease*. Future Medicine Ltd (Unitec House, 2 Albert Place, London N3 1QB, UK): London, England, 2013 http://www.futuremedicine.com/doi/pdf/10.2217/9781780842196 (accessed 12 Sep2019).

74 Berg DF, Bahadursingh AM, Kaminski DL, Longo WE. Acute surgical emergencies in inflammatory bowel disease. *The American Journal of Surgery* 2002; **184**: 45–51.

75 D'Haens G, Baert F, van Assche G, Caenepeel P, Vergauwe P, Tuynman H *et al.* Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. *The Lancet* 2008; **371**: 660–667.

76 Lin MV, Blonski W, Lichtenstein GR. What is the optimal therapy for Crohn's disease: step-up or top-down? *Expert Review of Gastroenterology & Hepatology* 2010; **4**: 167–180.

77 Hanauer SB, Feagan BG, Lichtenstein GR, Mayer LF, Schreiber S, Colombel JF *et al.* Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial. *The Lancet* 2002; **359**: 1541–1549.

78 Gecse KB, Bemelman W, Kamm MA, Stoker J, Khanna R, Ng SC *et al.* A global consensus on the classification, diagnosis and multidisciplinary treatment of perianal fistulising Crohn's disease. *Gut* 2014; **63**: 1381–1392.

79 Larson DW, Pemberton JH. Current concepts and controversies in surgery for IBD. *Gastroenterology* 2004; **126**: 1611–1619.

80 Satsangi J, Jewell DP, Bell JI. The genetics of inflammatory bowel disease. *Gut* 1997; **40**: 572–574.

81 International Multiple Sclerosis Genetics Consortium, International IBD Genetics Consortium, Liu JZ, van Sommeren S, Huang H, Ng SC *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature Genetics* 2015; **47**: 979–986.

82 Lees CW, Barrett JC, Parkes M, Satsangi J. New IBD genetics: common pathways with other diseases. *Gut* 2011; **60**: 1739–1753.

83 Bayless TM, Tokayer AZ, Polito JM, Quaskey SA, Mellits ED, Harris ML. Crohn's disease: concordance for site and clinical type in affected family members--potential hereditary influences. *Gastroenterology* 1996; **111**: 573–579.

84 Kaser A, Zeissig S, Blumberg RS. Inflammatory bowel disease. *Annu Rev Immunol* 2010; **28**: 573–621.

85 Ye BD, McGovern DPB. Genetic variation in IBD: progress, clues to pathogenesis and possible clinical utility. *Expert Rev Clin Immunol* 2016; **12**: 1091–1107.

86 Cho JH. Significant role of genetics in IBD: the NOD2 gene. *Rev Gastroenterol Disord* 2003; **3 Suppl 1**: S18-22.

87 Ng SC, Tsoi KKF, Kamm MA, Xia B, Wu J, Chan FKL *et al.* Genetics of inflammatory bowel disease in Asia: Systematic review and meta-analysis: *Inflammatory Bowel Diseases* 2012; **18**: 1164–1176.

88 Parkes M. Evidence from Genetics for a Role of Autophagy and Innate Immunity in IBD Pathogenesis. *Digestive Diseases* 2012; **30**: 330–333.

89 Jones SA, Mills KHG, Harris J. Autophagy and inflammatory diseases. *Immunology and Cell Biology* 2013; **91**: 250–258.

90 Franke A, McGovern DPB, Barrett JC, Wang K, Radford-Smith GL, Ahmad T *et al.* Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nature Genetics* 2010; **42**: 1118–1125.

91 Travassos LH, Carneiro LAM, Ramjeet M, Hussey S, Kim Y-G, Magalhães JG *et al.* Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nature Immunology* 2010; **11**: 55–62.

92 Kaser A, Blumberg RS. Adaptive immunity in inflammatory bowel disease: state of the art: *Current Opinion in Gastroenterology* 2008; **24**: 455–461.

93 Kaser A, Lee A-H, Franke A, Glickman JN, Zeissig S, Tilg H *et al.* XBP1 links ER stress to intestinal inflammation and confers genetic risk for human inflammatory bowel disease. *Cell* 2008; **134**: 743–756.

94 Prescott NJ, Fisher SA, Franke A, Hampe J, Onnie CM, Soars D *et al.* A Nonsynonymous SNP in ATG16L1 Predisposes to Ileal Crohn's Disease and Is Independent of CARD15 and IBD5. *Gastroenterology* 2007; **132**: 1665–1671.

95 Abreu MT, Taylor KD, Lin Y-C, Hang T, Gaiennie J, Landers CJ *et al.* Mutations in NOD2 are associated with fibrostenosing disease in patients with Crohn's disease. *Gastroenterology* 2002; **123**: 679–688.

96 Cleynen I, Boucher G, Jostins L, Schumm LP, Zeissig S, Ahmad T *et al.* Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *Lancet* 2016; **387**: 156–167.

97 Inoue N, Tamura K, Kinouchi Y, Fukuda Y, Takahashi S, Ogura Y *et al.* Lack of common NOD2 variants in Japanese patients with Crohn's disease. *Gastroenterology* 2002; **123**: 86–91.

98 The International IBD Genetics Consortium (IIBDGC), Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP *et al.* Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012; **491**: 119–124.

99 ATG16L1 autophagy related 16 like 1 [Homo sapiens (human)] - Gene - NCBI. https://www.ncbi.nlm.nih.gov/gene/55054 (accessed 13 Sep2019).

100 Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A *et al.* Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nature Genetics* 2007; **39**: 596–604.

101 Hampe J, Franke A, Rosenstiel P, Till A, Teuber M, Huse K *et al.* A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nature Genetics* 2007; **39**: 207–211.

102 Cummings FJR, Cooney R, Pathan S, Anderson CA, Barrett JC, Beckly J *et al.* Confirmation of the role of ATG16l1 as a Crohn's disease susceptibility gene: *Inflammatory Bowel Diseases* 2007; **13**: 941–946.

103 Lacher M, Schroepf S, Ballauff A, Lohse P, von Schweinitz D, Kappler R *et al.* Autophagy 16-like 1 rs2241880 G allele is associated with Crohn's disease in German children. *Acta Paediatrica* 2009; **98**: 1835–1840.

104 Roberts RL, Gearry RB, Hollis-Moffatt JE, Miller AL, Reid J, Abkevich V *et al.* IL23R R381Q and ATG16L1 T300A are strongly associated with Crohn's disease in a study of New Zealand Caucasians with inflammatory bowel disease. *Am J Gastroenterol* 2007; **102**: 2754–2761.

105 Gazouli M. *NOD2* / *CARD15*, *ATG16L1* and *IL23R* gene polymorphisms and childhood-onset of Crohn's disease. *World Journal of Gastroenterology* 2010; **16**: 1753.

106 Walker DG, Bancil AS, Rai PS, Williams HR, Chambers J, Kooner JS *et al.* Ethnic variation in the frequency of IBD related polymorphisms in IRGM, ATG16L1 and IL23R. *Gut* 2011; **60**: A223–A223.

107 Yang S-K, Park M, Lim J, Park SH, Ye BD, Lee I *et al.* Contribution of IL23R but not ATG16L1 to Crohn's disease susceptibility in Koreans: *Inflammatory Bowel Diseases* 2009; **15**: 1385–1390.

108 Pugazhendhi S, Baskaran K, Santhanam S, Ramakrishna BS. Association of ATG16L1 gene haplotype with inflammatory bowel disease in Indians. *PLOS ONE* 2017; **12**: e0178291.

109 Plevy S, Silverberg MS, Lockton S, Stockfisch T, Croner L, Stachelski J *et al.* Combined Serological, Genetic, and Inflammatory Markers Differentiate Non-IBD, Crohn's Disease, and Ulcerative Colitis Patients: *Inflammatory Bowel Diseases* 2013; **19**: 1139–1148.

110 ECM1 extracellular matrix protein 1 [Homo sapiens (human)] - Gene - NCBI. https://www.ncbi.nlm.nih.gov/gene/1893 (accessed 13 Sep2019).

111 McCole DF. IBD candidate genes and intestinal barrier regulation. *Inflamm Bowel Dis* 2014; **20**: 1829–1849.

112 Shi J, Zhou L, Zhernakova A, Qian J, Zhu F, Sun G *et al.* Haplotype-based analysis of ulcerative colitis risk loci identifies both IL2 and IL21 as susceptibility genes in Han Chinese: *Inflammatory Bowel Diseases* 2011; **17**: 2472–2479.

113 Ventham NT, Kennedy NA, Nimmo ER, Satsangi J. Beyond Gene Discovery in Inflammatory Bowel Disease: The Emerging Role of Epigenetics. *Gastroenterology* 2013; **145**: 293–308.

114 Podolsky DK. Inflammatory Bowel Disease. *The New England Journal of Medicine* 2002; **13**.

115 Lennard-Jones JE. Classification of Inflammatory Bowel Disease. *Scandinavian Journal of Gastroenterology* 1989; **24**: 2–6.

116 UM_gDNABlood.pdf. https://www.mn-net.com/Portals/8/attachments/Redakteure_Bio/Protocols/Genomic%20DNA/UM_gDNABlood.pdf (accessed 15 Sep2019).

117 rs2241880 RefSNP Report - dbSNP - NCBI. https://www.ncbi.nlm.nih.gov/snp/rs2241880 (accessed 15 Sep2019).

118 rs3737240 RefSNP Report - dbSNP - NCBI. https://www.ncbi.nlm.nih.gov/snp/rs3737240 (accessed 15 Sep2019).

119 rs13294 RefSNP Report - dbSNP - NCBI. https://www.ncbi.nlm.nih.gov/snp/rs13294 (accessed 15 Sep2019).

120 Glas J, Konrad A, Schmechel S, Dambacher J, Seiderer J, Schroff F *et al.* The *ATG16L1* Gene Variants rs2241879 and rs2241880 (T300A) Are Strongly Associated with Susceptibility to Crohn's Disease in the German Population. *The American Journal of Gastroenterology* 2008; **103**: 682–691.

121 Adali G, Ersoy Tunali N, Yorulmaz E, Tiryakioglu NO, Guray Mungan S, Ulasoglu C *et al.* Extracellular matrix protein 1 gene rs3737240 single nucleotide polymorphism is associated with ulcerative colitis in Turkish patients. *The Turkish Journal of Gastroenterology* 2017; **28**: 254–259.

122 Economou M. Crohn's disease incidence evolution in North-western Greece is not associated with alteration of NOD2/CARD15 variants. *World Journal of Gastroenterology* 2007; **13**: 5116.

123 Lesage S, Zouali H, Cézard J-P, Colombel J-F, Belaiche J, Almer S *et al.* CARD15/NOD2 Mutational Analysis and Genotype-Phenotype Correlation in 612 Patients with Inflammatory Bowel Disease. *The American Journal of Human Genetics* 2002; **70**: 845–857.

124 Eckmann L, Karin M. NOD2 and Crohn's Disease: Loss or Gain of Function? *Immunity* 2005; **22**: 661–667.

125 Parente F, Pastore L, Bargiggia S, Cucino C, Greco S, Molteni M *et al.* Incidence and risk factors for gallstones in patients with inflammatory bowel disease: A large case-control study. *Hepatology* 2007; **45**: 1267–1274.

126 Zhang FM, Xu CF, Shan GD, Chen HT, Xu GQ. Is gallstone disease associated with inflammatory bowel diseases? A meta-analysis: Gallstone disease and IBD. *Journal of Digestive Diseases* 2015; **16**: 634–641.

127 van Sommeren S, Janse M, Karjalainen J, Fehrmann R, Franke L, Fu J *et al.* Extraintestinal Manifestations and Complications in Inflammatory Bowel Disease: From Shared Genetics to Shared Biological Pathways. *Inflammatory Bowel Diseases* 2014; : 1.

128 Lakatos PL, Szamosi T, Szilvasi A, Molnar E, Lakatos L, Kovacs A *et al.* ATG16L1 and IL23 receptor (IL23R) genes are associated with disease susceptibility in Hungarian CD patients. *Digestive and Liver Disease* 2008; **40**: 867–873.

129 Aida I, Meddour Y, Kadiri H, Smara M, Bousseloub A, Kecili L *et al.* T300A variant of AT16L1 gene in a cohort of Algerian Crohn disease patients. *Current Research in Translational Medicine* 2018; **66**: 9–14.

130 Nuij VJAA, Peppelenbosch MP, van der Woude CJ, Fuhler GM. Genetic polymorphism in ATG16L1 gene is associated with adalimumab use in inflammatory bowel disease. *Journal of Translational Medicine* 2017; **15**. doi:10.1186/s12967-017-1355-9.

131 Gerich ME, McGovern DPB. Towards personalized care in IBD. *Nature Reviews Gastroenterology & Hepatology* 2014; **11**: 287–299.